DEVELOPMENT OF NOVEL SPARSE-AWARE ALGORITHMS FOR ADAPTIVE SYSTEM IDENTIFICATION

Submitted in partial fulfilment of the requirements
for the award of the degree of

Doctor of Philosophy

by
P.Rakesh
(Roll No. 701350)

Under the Guidance of

Prof. T. Kishore Kumar
Professor

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY
WARANGAL – 506004, T.S, INDIA

JULY-2019

APPROVAL SHEET

This thesis entitled "**Development of Novel Sparse-Aware Algorithms for Adaptive System Identification**" by Mr. **P. Rakesh** is approved for the degree of **Doctor of Philosophy**.

Examiners

____

____

Supervisor

____

Prof. T.Kishore Kumar

Professor, Electronics and Communication Engineering Department,

NIT WARANGAL

Chairman

____

Prof. N. Bheema Rao

Head, Electronics and Communication Engineering Department,

NIT WARANGAL

Date:

Place:

DECLARATION

I hereby declare that the matter embodied in this thesis entitled "**Development of Novel Sparse-Aware Algorithms for Adaptive System Identification"** is based entirely on the results of the investigations and research work carried out by me under the supervision of **Prof. T. Kishore Kumar**, Department of Electronics and Communication Engineering, National Institute of Technology Warangal. I declare that this work is original and has not been submitted in part or full, for any degree or diploma to this or any other University.

I declare that this written submission represents my ideas in my own words and where others ideas or words have been included. I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

P. Rakesh

Roll No: 701350

Date:

Place: Warangal

Department of Electronics and Communication Engineering
National Institute of Technology
Warangal – 506 004, Telangana, India

CERTIFICATE

This is to certify that the dissertation work entitled **"Development of Novel Sparse-Aware Algorithms for Adaptive System Identification"**, which is being submitted by Mr. P. Rakesh (Roll No. 701350), is a bonafide work submitted to National Institute of Technology Warangal in partial fulfilment of the requirement for the award of the degree of *Doctor of Philosophy in Electronics and Communication Engineering*.

To the best of our knowledge, the work incorporated in this thesis has not been submitted elsewhere for the award of any degree.

Prof. T. Kishore Kumar
Supervisor
Department of ECE
National Institute of Technology
Warangal – 506004

Dedicated to my beloved

Parents, Wife & Daughter

CONTENTS

ACKNOWLEDGEMENTS

It gives me immense pleasure to express my deep sense of gratitude and thanks to my supervisor **Prof. T. Kishore Kumar** (National Institute of Technology Warangal, NITW), for his invaluable guidance, support and suggestions. His knowledge, suggestions, and discussions helped me to become a capable researcher. He has shown me the interesting side of this wonderful multidisciplinary area. His encouragement also helped me to overcome the difficulties encountered in my research as well as in my personal life also.

I am thankful to the current HOD **Prof. N. Bheema Rao**, Dept. of ECE and the former HODs, Prof. T.Kishore Kumar and Prof. N.V.S.N. Sarma, for giving me an opportunity to carryout my research work in the department.

I take this privilege to thank all my Doctoral Scrutiny Committee members, **Prof. Y.N. Reddy**, Department of Mathematics, **Dr. C. B. Rama Rao**, Professor, Department of Electronics and Communication Engineering, **Dr. L.Anjaneyulu**, Professor, Department of Electronics and Communication Engineering, for their detailed review, constructive suggestions and excellent advice during the progress of this research work.

I want to mention my sincere and special thanks to **Dr. Felix Albu**, Professor, Department of Electronics, Valahia University of Targoviste, Targoviste, Romania for giving me his precious critique and proofreading which has profoundly improved the composition of my publications. Also, I would like to thank **Dr. Tõnu Trump**, Professor, Tallinn University of Technology, Tallinn, Estonia for accepting me as a visting Ph.D scholar and for his guidance and continuous support that have helped me so much during this period.

I also appreciate the encouragement from teaching, non-teaching members and fraternity of Dept. of E.C.E. in N.I.T. Warangal. They have always been encouraging and supportive.

I take this opportunity to convey my regards to Dr. P Venkata Subba Reddy and Dr. K V Kadambari, Assistant Professor, Department of Computer Science and Engineering for their motivation and support throughout my work. I wish to thank all my friends for their encouragement and good wishes. Thanks to Sivaprasad Nandyala, Sunnydayal V, Sudeep Surendran, Prasad Nizampatnam, Surekha Reddy, Siva Priyanka, Ravi B, Kiran D and Sunil Kumar K, Department of Electronics and Communication Engineering for being always next to me.

I would like to thank the MHRD, Govt. of India for the financial support of this work and INTERWEAVE, Erasmus Mundus Programme.

I acknowledge my gratitude to all my teachers and colleagues at various places for supporting and cooperating me to complete this work.

Finally, I would like to express my deepest gratitude to my parents P. Komaraiah and P. Padma, my wife for their relentless support and boundless love. I render my respect to my in-laws, my sister and brother-in-law for giving me mental support and constant encouragement. They have motivated and helped me to complete my thesis work successfully.

P.Rakesh

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVATIONS

AEC	: acoustic echo cancellation
ANC	: adaptive noise cancellation
APA	: affine projection algorithm
AP-NCA	: affine projection normalized correlation algorithm
AR process	: auto regressive process
ASI	: adaptive system identification
BG	: Bernoulli-Gaussian
CIM	: correntropy induced metric
CIR	: channel impulse response
EMSE	: excess mean square error
ERLE	: echo return loss enhancement
GMM	: Gaussain mixture model
HOEP	: high-order error power
IPNLMS	: improved proportionate NLMS
LA	: Lyapunov theory-based adaptive filtering
LASSO	: least absolute shrinkage and selection operator
LMAT	: least mean absolute third
LMMN	: least-mean mixed- norm algorithm
LMS	: least mean square
MMSE	: minimum mean square error
MSD	: mean square deviation
MSE	: mean square error
NCA	: normalized correlation algorithm

NLMS	:	normalized least mean square
NLMAT	:	normalized least mean absolute third
NRMN	:	normalized robust mixed norm
NWM	:	normalized weight misalignment
PNLMS	:	proportionate NLMS
RLS	:	recursive least squares
RZA-LMS	:	reweighted zero-attracting LMS
SLMMN	:	sigmoid least-mean mixed- norm algorithm
SOS	:	second-order statistics
WGN	:	white Gaussian noise
ZA-LMS	:	zero-attracting LMS

ABSTRACT

Adaptive filtering algorithms play a key role in adaptive signal processing, especially for applications where real-time estimation of some unknown parameters is required.

In this thesis, the significance of adaptive filters in system identification configuration is considered. Adaptive echo cancellation and channel estimation are the two prominent communications applications in system identification configuration.

In order to determine the transfer function estimate for an unknown digital or analog system, one can use the adaptive system identification (ASI). System identification describes the task of identifying an existing unknown system and adaptive filters are widely used for this application. In many scenarios of system identification, the impulse response of underlying system is presumed to be sparse which means most of its coefficients are zeros (inactive) and have few non-zero values (active). The basic methodology behind the sparse system identification is to make use of the prior sparse information to improve its filtering/estimation performance.

The traditional system identification algorithms are generally sparsity agnostic in nature viz. they are unaware of the underlying system sparsity which makes their application impractical for system identification. In order to exploit system sparsity, sparse adaptive filters are extensively used. Sparse-aware adaptive filtering algorithms offer improved performance. Hence, in this thesis, we consider the development of novel sparse adaptive algorithms for system identification. This thesis comprises of four parts:

1. Combinational approaches of adaptive filters for sparse system identification.
2. Sparse adaptive algorithms based on Lyapunov Stability for system identification.

3. Robust sparse system identification algorithms for impulsive noise environments.
4. Complex domain sparse adaptive algorithms for system identification.

In the first part, utilizing the benefits of combining two adaptive filters through a mixing parameter, we propose an affine combination of two Improved Proportionate Normalized LMS (IPNLMS) filters. Further, we also propose an affine combination of Reweighted Zero Attracting-NLMS (RZA-NLMS) and NLMS algorithm for system identification with variable sparsity. The combination approach provides the robust solution to alleviate the convergence speed vs steady-state error tradeoff, as well as to efficiently increase the filter robustness to time varying sparsity of the system.

In the second part, we consider the Lyapunov theory-based adaptive filter (LA) which offers to improve the convergence and stability, and also overcome the problems faced by gradient descent-based adaptive filtering techniques. In order to address system sparsity, the Zero-Attracting Lyapunov Adaptation algorithm (ZA-LA), the Reweighted Zero-Attracting Lyapunov Adaptation algorithm (RZA-LA) and an affine combination of the LA and ZA-LA adaptive filters are proposed. We show that the proposed sparse Lyapunov algorithms outperform the Least Mean Square (LMS) algorithm and its sparse counterpart (ZA-LMS and RZA-LMS) for both white input and colored input cases in terms of Mean Square Deviation (MSD) and Mean Square Error (MSE). The proposed affine combination filter is also robust in identifying the system with variable sparsity.

In the third part, we investigate the estimation performance of adaptive algorithms under the impulsive noise conditions. The novel sparse algorithms based on high-order error power (HOEP) criterion i.e., Normalized Least Mean Absolute Third (NLMAT) are proposed to mitigate the adverse effects of impulsive noise and to utilize the sparsity phenomenon effectively. Modified Least-Mean Mixed-Norm algorithm which is based on sigmoid function (SLMMN) is also developed to achieve robust performance against impulsive noise and the corresponding sparse SLMMN algorithms are proposed in the sparse system identification context.

In the fourth part, we discuss the complex domain sparse adaptive filter algorithms by incorporating different sparse penalty terms into the affine projection normalized correlation algorithm (APNCA). The proposed algorithms address the system sparsity as well as robustness against impulsive noise and achieves faster convergence for a correlated input.

All the proposed algorithms in this work are tested by extensive computer simulations. The results demonstrate significant performance improvement in terms of convergence rate, robustness against impulsive noise and steady-state error.

Chapter 1

Introduction

CHAPTER 1

1.1 Introduction

The area of digital signal processing is steadily developing over the past decades. The rapid growth in this field has evolved into various specialized topics. One of the examples of digital signal processing system is the use of adaptive filters [1-3]. An adaptive filter is a self-operating filter that relies on an adaptive algorithm to perform satisfactorily in an environment where its relevant characteristics are not available. The main advantage of using adaptive filters is that they work excellently in an unknown environment resulting in better performance of the filter. Owing to the various configurations of digital signal processors along with adaptive algorithms, they are widely used in many areas like radar, telecommunications, echo cancellation, noise reduction, channel estimation, etc. [4], [5]. Fig.1.1 illustrates the general configuration for an adaptive filter.

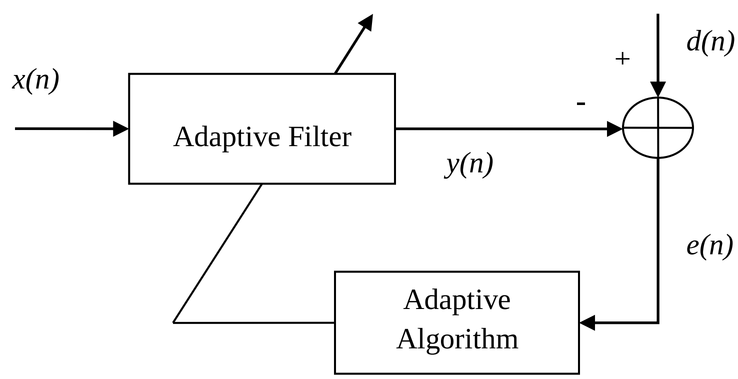


Fig. 1.1: General Adaptive Filter Configuration

where,

$x(n)$: the input signal to the adaptive filter,

$y(n)$: the corresponding output signal,

$d(n)$: the desired response signal, and

$e(n)$: the error signal.

An adaptive filtering algorithm uses an iterative method to adjust its coefficients to minimize the error signal (cost function). The most popular criterion used for updating the adaptive algorithm is the minimum mean square error (MMSE) and the least square error. Several families of adaptive algorithms such as least mean square (LMS) and its variants [6-14] recursive least squares (RLS) [2], [15], [16] are developed. They differ in their filter weight updating scheme, computational complexity and convergence speed.

Adaptive filters are generally evaluated by their convergence rate, steady-state mean squared error, computational complexity, and numerical stability. Widrow's LMS adaptive filter [7] is the first in chronology. It is computationally simple and numerically stable but slow in convergence, particularly when the filter input data is correlated. Subsequently, over the past few decades, numerous algorithms have been developed for improving the performance rate. However, such techniques are normally more computationally complex. Thus, one needs to balance the tradeoff between complexity and performance measure.

By the manner in which the desired signal is extracted, the adaptive filtering applications may be classified as four classes: identification, inverse modeling, prediction and interference cancellation [2].

The advantages of using adaptive filters for solving the real world problems can be summarized into four points:

- The adaptive filters do not introduce any significant delay in the filter output.
- Adaptive filters are capable of tracking variations of signal statistics or time-varying systems.
- The adaptive filter updates its parameters each time the new signal sample is available, thereby saving memory.
- The adaptive filter, in general, is much simpler to code in software or to implement in hardware.

However, adaptive filtering is governed by certain limitations. The convergence speed, being one of the main characteristic performance measures of an adaptive algorithm, measures how fast the filter adapts its coefficients to the desired state. Real time applications such as adaptive noise cancellation (ANC) require usually fast convergence. Adaptive filtering is an iterative method which implies the existence of a step size that controls the adaptation of the filter's parameters. As a result, the choice of the step size has a direct impact on the convergence speed of the algorithm. In addition, the length of the adaptive filter is another factor that affects the convergence of the iterative process. For applications where the impulse response of a room needs to be modeled several hundreds of filter taps are required. The longer the filter is, the slower the convergence of the filter coefficients. The tradeoff between the initial convergence speed and the mean-square error in steady state is controlled by the step size of the adaptive filtering algorithm. Large step size leads to a fast initial convergence but the algorithm also exhibits a large mean-square error in the steady state and in contrary, small step size slows down the convergence but results in a small steady state error. Recently, there has been an interest in the combination scheme that is able to optimize the trade-off between convergence speed and steady state error. The scheme consists of two adaptive filters which are updated independently and the outputs of these filters are combined through a mixing parameter λ.

In the recent years, sparsity property has been very popular among researchers in the area of signal processing using adaptive filtering [17], [18], [19], [20] image processing and statistical estimation [21], [22], [23], [24]. If most of the entries of a vector are zeros but only a few ones have significant values, the vector is said to be sparse. Sparsity is used in adaptive filtering in different manners and offers us many advantages. Actually, in adaptive filtering, many systems are generally assumed to be linear. But in some cases, like in digital TV transmissions channels [25] and echo paths [26], [27], a few components of the impulse response have significant value while the rest are zeros or near to zeros. For example, a network echo path has an active region only in a narrow interval with significant values and the rest of the impulse response coefficients is zero or negligible. An acoustic echo path also has similar characteristics as that of the network echo with a little more complicated structure depending on the movement and distance between the microphone and loudspeaker. Eventually, such systems are said to be sparse systems. By utilizing such sparse prior information we can improve the filtering/estimation performance. However, standard adaptive filters do not exploit such information. An approach to promote sparsity is based on assigning proportional step sizes to different taps according to their magnitudes, such as the proportionate normalized least-mean square (PNLMS). Recently, based on recent progress in compressive sensing, the spare information can be utilized by inducing sparse penalties into the cost function of the adaptive algorithm.

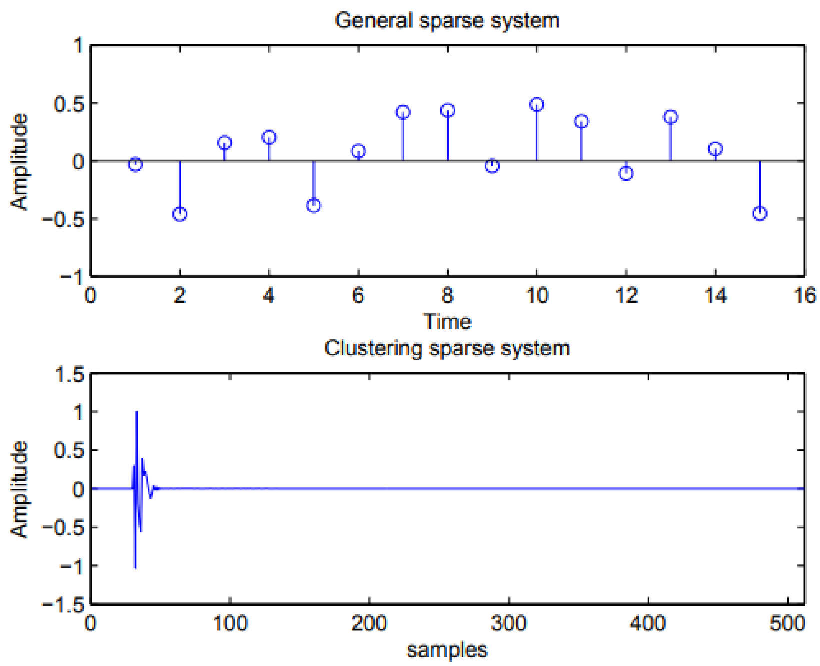


Fig. 1.2: Typical sparse systems

A general sparse system having scattered non-zero distribution throughout the system response is shown in Fig. 1.2 (top) and a clustering sparse system consisting of one or more clusters of non-zero coefficients along the entire system response (for example: echo path) is provided in Fig. 1.2 (bottom).

1.2 System identification

The system identification approach is used to model an unknown system. Adaptive System Identification (ASI) describes the task of identifying an existing unknown system and using adaptive filters we can estimate a model of the system [28]. The adaptive filtering applications such as channel identification [29], network and acoustic echo cancellation [30], [31], [32], adaptive noise cancellation [33] performs the system identification task. The block diagram of ASI is shown in Fig. 1.3. In this configuration the same input signal $x(n)$ is fed to both the unknown system and the adaptive filter. The output of the adaptive filter $y(n)$ is subtracted from the desired signal $d(n)$ to get the estimation error, $e(n) = d(n) - y(n)$. The error signal $e(n)$ will be minimized with the adaptive filter's adaptation process so that the adaptive model approximates the unknown system from the view point of input/output. The mean square error (MSE) becomes zero when the output of unknown system $d(n)$ is free from the observation noise.

In practicality the input signal is combined with the noise, thus leading to a high value in the error signal. There are various types of noise such as Gaussian noise and Impulsive (colored) noise. By using the adaptive filters, we reduce this error signal to a minimum value. In other words, minimal error signal means purity of the input signal and correctness in the output signal thereby improving the performance accuracy.

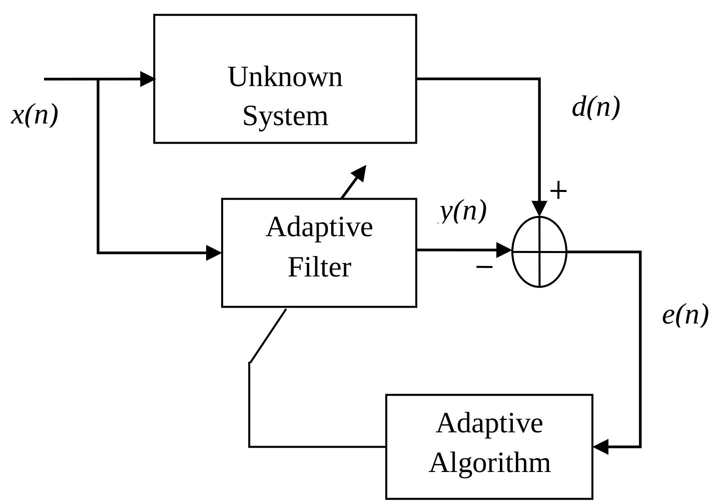


Fig. 1.3: Adaptive filter for system identification

Based on some statistical criterion the adaptive filter's objective is to minimize the cost function of the error signal. The Gaussian distribution has been by far the most popular statistical distribution in signal processing. The minimum mean-square error (MMSE) criterion based on L_2 norm is well known and is adequate to provide effective estimation under the Gaussian assumption. Algorithms that are based on MMSE rule work better when the noise is Gaussian in nature. However, these techniques suffer from serious performance degradation in the case of impulsive noise environments. Thus, we propose new algorithms for addressing these drawbacks.

Impulsive, non-Gaussian noise often occurs in practical applications and is characterized by large-magnitude outliers with heavy-tailed distributions [34], [35], [36]. Some examples of non-Gaussian noise are underwater acoustic noise, communication over power lines, and noise on telephone lines [37]-[41]. In acoustic echo cancellation (AEC), doubletalk situations can also be viewed as an impulsive interference source.

1.3 Problem Statement

Most of the real-life applications are sparse in nature. A sparse system is one in which a large percentage of the energy is concentrated to only a few coefficients. System sparsity can be observed in applications like wireless communication, network and acoustic echo cancellers, radar imaging, etc. However, these algorithms are sparse agnostic i.e., they are unaware of underlying sparse impulse response.

Hence, there is a need to develop novel sparse adaptive algorithms to exploit the system sparsity in the system identification context and also to exhibit robust performance in the presence of impulsive noise environments.

1.4 Motivation

The LMS and the normalized LMS algorithms are the most widely used adaptive filtering algorithms in system identification application because of their low computational complexity and simplicity [6], [8]. Still, they display performance trade-offs that can hinder their use in practice, such as the compromise between convergence rate and steady-state error. To address this tradeoff problem, combination of adaptive filters has been proposed. In many situations, the impulse response of unknown system is assumed to be sparse, which means most of its coefficients are zeros (inactive) while only a few coefficients are large (active). The basic idea of sparse system identification is to try to incorporate these prior sparse information to improve the filtering/estimation performance.

Sparse-aware adaptive filter algorithms offer improved performance when the system impulse response is highly sparse. In order to exploit system sparsity, new sparse algorithms has to be developed by including various sparse penalties into the cost function of the conventional algorithms. Hence, in the first part of the thesis, sparse-aware adaptive algorithms are developed for system identification to handle variable sparsity as well as to achieve faster convergence rate and lower steady-state error.

An important issue in system identification is the effect of measurement noise on the results. The measurement noise is often assumed to be a random process with finite second-order statistics (SOS), making the MSE an appropriate metric for estimation error. However, in real-world environments, the noise encountered is more impulsive in nature than that predicted by a Gaussian distribution. The noises that exhibit impulsive behaviour frequently produce large amplitude outliers and the traditional algorithms fail to converge in such noise cases. Hence, in the latter part of the thesis, robust sparse adaptive filtering algorithms are developed for impulsive noise environments.

1.5 Objectives

The objectives of the work are:

- To propose novel algorithms using affine combination of adaptive filters for echo cancellation and system identification with variable sparsity.
- Implementation of novel sparse techniques based on Lyapunov Stability for system identification.
- Development of robust sparse algorithms for adaptive system identification under impulsive noise environments.
- To develop complex domain sparse adaptive algorithms for system identification.

1.6 Contributions of the Thesis

- The first approach focuses on adaptive system identification using affine combination method. To improve the convergence rate and steady-state error tradeoff and to increase the robustness to system with different degrees of sparsity, an affine combination of two Improved Proportionate NLMS (IPNLMS) filters [1**] is developed. Also, another scheme of affine combination of NLMS and Reweighted Zero Attracting-NLMS (RZA-NLMS) which is achieved by including log-sum penalty into the cost function of the NLMS algorithm is developed for identification of unknown system with variable sparsity [1*].

- The second approach is based on adaptive filtering technique called Lyapunov Theory-based Adaptive Filtering (LA) which is used to overcome the limitations of LMS algorithm. To promote sparsity, novel sparse algorithms using different sparse constraints into the cost function of the LA algorithm are developed. Also, an affine combination based filter namely, Affine Combined Lyapunov Adaptation (ACLA) filter is proposed to identify the system with variable sparsity [2**].

- The third approach is based on adaptive sparse system identification under impulsive noise environments. Normalized Least Mean Absolute Third (NLMAT) algorithm which is based on high-order error power (HOEP) condition and is independent of the power of the input signal and has good immunity to impulsive noise. In order to exploit system sparsity and to combat impulsive noise effect, novel sparse algorithms [3**] are developed by integrating different sparse penalties into the cost function of NLMAT algorithm. To mitigate the adverse causes of impulsive noise effectively, another algorithm based on sigmoid function i.e., sigmoid least-mean mixed- norm algorithm (SLMMN) is proposed and novel sparse algorithms are developed by integrating different sparse constraints into the cost function of SLMMN algorithm to exploit system sparsity [3*].

- Fourth approach is based on Affine Projection Normalized Correlation Algorithm (AP-NCA) which is developed in the complex domain to achieve faster convergence for a correlated input and also robust against impulsive noises. For sparse system identification novel sparse AP-NCA algorithms are developed [2*].

The proposed algorithms are implemented in MATLAB.

1.7 Organization of the Thesis

The thesis is organized into seven chapters:

Chapter 1 gives an introduction of adaptive filtering and system identification application, motivation and scope, objectives and contributions of the thesis.

In **Chapter 2**, literature review of sparse-aware adaptive algorithms is presented and overview of adaptive algorithms under impulsive noise environments is discussed.

Chapter 3 details the proposed adaptive combination approaches for sparse system identification.

Chapter 4 discusses the novel sparse algorithms based on Lyapunov Stability for adaptive system identification.

Chapter 5 deals with the novel sparse algorithms under impulsive noise environments.

Chapter 6 discusses the complex domain adaptive system identification using sparse affine projection normalized correlation algorithms under impulsive noises.

Finally, in **Chapter 7**, the conclusion of the thesis is summarized and future scope of the work is given.

Chapter 2

Literature Survey

CHAPTER 2

2.1 Introduction

Adaptive system identification includes many applications such as echo interference cancellation, sparse channel estimation, and adaptive beamforming. In many practical scenarios of system identification, the unknown system impulse response can be assumed to be sparse with varying degree of sparsity. A sparse system has only a few active coefficients while most of its coefficients are inactive. Some of the examples where sparse systems are encountered are network and acoustic echo cancellers [27], [42], wireless multipath channels [23], underwater acoustic communications [43]. Least Mean Square (LMS) and Normalized LMS (NLMS) are the most widely used adaptive filters for system identification [6], e.g., channel estimation, due to their simplicity. In practice, the LMS is highly sensitive to the characteristics of the input signals and has slow convergence [44]. The drawback of RLS algorithm is its high computational complexity. In addition, most of the adaptive filtering schemes suffer from so-called local minima problem, i.e., the optimization search may stop at a local minimum of the cost function in the parameter space if the initial values are arbitrarily chosen. To overcome these problems, a new adaptive filtering technique called Lyapunov Theory-based Adaptive Filtering (LA) is proposed [45]. In [46], a Lyapunov function of the error signal is defined and by properly choosing the parameter update rule, the weights of the filter are adaptively adjusted based on Lyapunov stability theory so that the error can asymptotically converge to zero. The design of Lyapunov adaptive filters is independent of the stochastic properties of the random input perturbations. Therefore, the local minima problem occurred in the gradient search-based adaptive filters can be avoided. New adaptive filter designs based on Lyapunov stability theory are proposed in [47-50].

However, they fail to exploit the system sparse structure and their performance is reduced when estimating the sparse channels, which makes their application unpractical for sparse system identification. The filtering/estimation performance of the unknown system can be further improved by utilizing the existing sparse information.

The echo paths (for both network and acoustic echo cancellation scenarios) have a specific property of system sparsity, which can be used in order to help the adaptation process. The Fig. 2.1 shows the sparse system identification based on adaptive filtering algorithm.

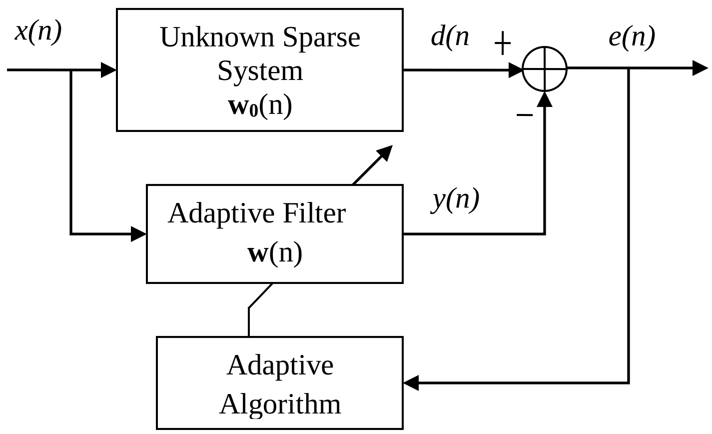


Fig. 2.1: Adaptive algorithm based sparse system identification

2.2 Sparse Adaptive Filters: A review

Exploiting the sparseness property of the echo paths, sparse adaptive filters are extensively used in AEC applications. Several sparse-aware adaptive filtering algorithms exist on proportionate-type methods [51]-[54] and compressed sensing techniques [21], [22], [55], [56]. One of the first sparse adaptive filtering algorithms for acoustic echo cancellation is Proportionate Normalized LMS (PNLMS) [51]. The underlying principle of PNLMS is to adapt each coefficient with an adaptation gain proportional to its own magnitude. This makes the PNLMS algorithm to achieve faster initial convergence rate when the impulse response is considered as sparse. However, after few iterations the convergence rate decreases remarkably, thus making the PNLMS performance much slower than NLMS. An Improved PNLMS [52] algorithm was proposed to exploit the 'proportionate' idea by introducing a controlled mixture of proportionate (PNLMS) and non-proportionate (NLMS) adaptation. Much research has been done during the last decades in order to develop proportionate-type algorithms for promoting sparsity in system identification [19], [57]-[84].

In a separate development, an alternative approach to identify sparse systems has been proposed by introducing different norm penalty terms to the existing algorithms [85]-[107]. The $\ell 1$-norm penalty was incorporated into the cost function of conventional LMS algorithm, which resulted in the Zero-Attracting LMS (ZA-LMS) algorithm. Reweighted zero-attracting LMS (RZA-LMS) algorithm was obtained by inducing log-sum penalty term into the standard LMS [85]. Moreover, it is worth mentioning that some algorithms, such as those from [108]-[116], have been developed combining both the proportionate-type and norm-based strategies.

However, there exists a tradeoff between convergence and steady state error for these sparse algorithms and also tend to show poor performance for systems with time varying sparsity. Hence, to handle the situation with variable system sparsity and to alleviate the tradeoff behaviour, adaptive filtering using combinational approaches are proposed [117]-[136].

In the convex combination adaptive filtering [120], the two component filters are updated according to their own weight adaptation rule and their outputs are combined through a mixing parameter, λ. Instead of directly updating λ(n), we define it as the output of a sigmoidal activation function,

$$\lambda(n) = sgm[a(n)] = \frac{1}{1+e^{-a(n)}}, \tag{2.1}$$

and update $a(n)$ using

$$a(n+1) = a(n) + \mu_a e(n)[e_2(n) - e_1(n)]\lambda(n)[1-\lambda(n)] \tag{2.2}$$

In [122] the adaptation of the mixing parameter is done using the normalized version which is very robust when the SNR is not known a priori or when it is time-varying.

The power-normalized least mean square algorithm is used to minimize the square of the error $e^2(n) = [d(n) - y(n)]^2$ in Eq. 2.2,

$$a(n+1) = a(n) + \frac{\mu_a}{p(n)} e(n)[e_2(n) - e_1(n)] \frac{d\lambda(n)}{da(n)}, \quad (2.3)$$

where μ_a is a step-size for the adaptation of $a(n)$, $p(n)$ is a rough estimate of the power of $[e_2(n) - e_1(n)]$, and

$$\frac{d\lambda(n)}{da(n)} = \frac{dsgm[a(n)]}{da(n)} = \lambda(n)[1 - \lambda(n)] \quad (2.4)$$

we update $p(n)$ using

$$p(n) = \beta p(n-1) + (1-\beta)[e_2(n) - e_1(n)]^2 \quad (2.5)$$

where 'β' is the forgetting factor. For instance, choosing β = 0.9 provides a good approximation for faster adaptation of $p(n)$.

2.3 Overview of Adaptive Algorithms under Impulsive Noise

Most of the algorithms discussed in section 2.2 are based on the renowned MSE criterion, assuming that the background noise is Gaussian. Since the MSE criterion is based on the second-order statistics consideration, it makes sense in the signal processing with Gaussian assumption. Whereas, the noisy conditions which are generated physically or man-made, deviate from the assumption of Gaussian distribution because of the impulsive behaviour. Some of the examples, multiple access interference in radio channels [137], double-talk in acoustic echo-cancellation [138], and other scenarios [139]–[140]. These impulsive distribution problems, i.e., the non-Gaussian heavy-tailed distribution problems, cannot be satisfactorily solved by the MSE criterion based algorithms thus requiring the use of robust adaptive filters.

Symmetric alpha-Stable (SαS) distribution:

Symmetric alpha-Stable (SαS) distribution is a classic non-Gaussian distribution which can model many impulsive noise processes in communications channels, and, in fact, includes the Gaussian density as a special case [36].

Generally, a SαS random distribution can be described by its characteristic function

$$\emptyset(\boldsymbol{t}) = exp(j\mu t - \gamma|t|^{\alpha}) \quad (2.6)$$

The characteristic exponent $\alpha \in (0,2]$ gives the impulsiveness of the noise (if α is small it leads to more number of impulses), μ is the location parameter and the dispersion γ ($\gamma > 0$) controls the extent of the distribution around μ and similar to the variance of Gaussian random variable. For $\alpha = 2$, the SαS probability density function (pdf) is comparable to the Gaussian pdf & γ becomes half of the variance.

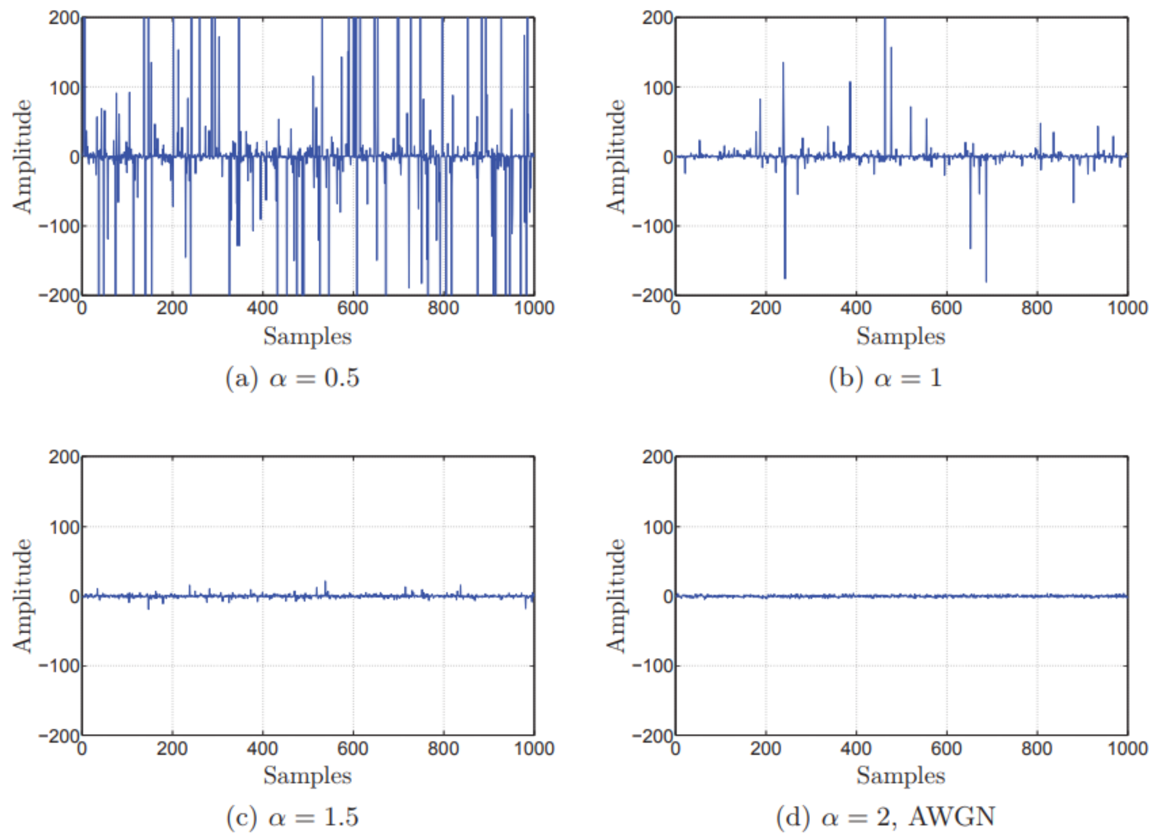


Fig. 2.2: Impulsive noise generated using SαS distribution for different values of α

Bernoulli-Gaussian (BG) distribution:

The impulsive noise modeled by a Bernoulli-Gaussian (BG) distribution is represented by [141]

$$q(t) = \alpha(t)I(t) \quad (2.7)$$

$I(t) \sim \mathcal{N}(0, \sigma_I^2)$ and $\alpha(t)$ is a binary process described by the probability $p(\alpha(t) = 1) = P,\ p(\alpha(t) = 0) = 1 - P,$ where P represents the probability of occurrence of the impulsive interference $I(t)$. Note that the variance of $q(t)$ is given by $\sigma_q^2 = P\sigma_I^2$.

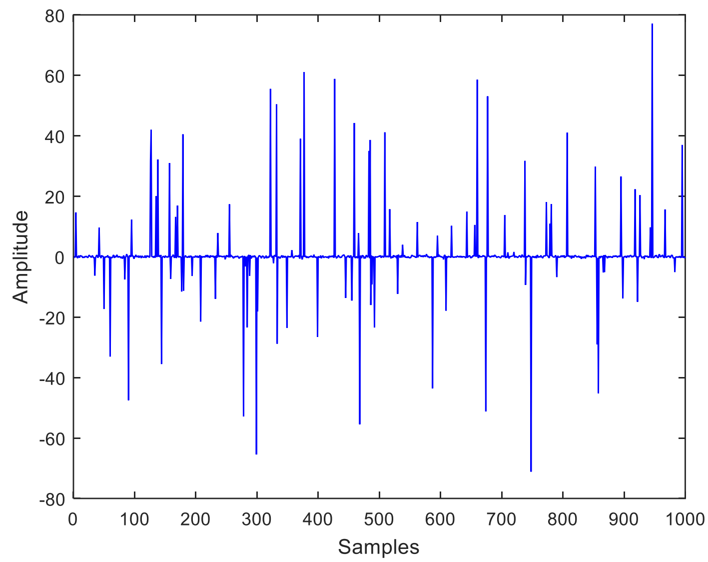


Fig. 2.3: Impulsive noise generated using Bernoulli-Gaussian method

In the recent years, adaptive filtering algorithms that are based on high-order error power (HOEP) conditions are proposed [142]-[145] to improve the convergence performance and mitigate the noise interference effectively. Least mean absolute third (LMAT) algorithm which is an adaptive HOEP algorithm is based on the minimization of the mean of the absolute error value to the third power [144], [145]. The LMAT algorithm outperforms the LMS algorithm for many noise distributions [146] and often exhibit faster convergence than the LMS algorithm. But, its convergence performance depends strongly on the input signal power. To alleviate these hostile effects, a normalized form of LMAT (NLMAT) algorithm is proposed in [147]. NLMAT algorithm is independent of the input signal power and achieves a strong capability of tolerance towards impulsive noise and good stability.

To overcome the sensitivity issues of LMS [6] and LMF [142], the least-mean mixed-norm (LMMN) algorithm which is a convex, linear combination of LMS and LMF is proposed in [149]-[150]. However, the performance of LMMN algorithm degrades seriously due to impulsive interferences which exist in practical environments [151]-[153].

To improve the filter performance for colored signals, an Affine Projection Algorithm (APA) has been proposed [11], [154]. For a larger projection order, the APA algorithm has faster convergence, but the steady-state error is higher resulting in a convergence vs steady-state error tradeoff. However, these methods may be unreliable in estimating the systems under non-Gaussian impulsive noise environments. Normalized correlation algorithm (NCA) for complex-domain adaptive filtering is proposed [155] that is highly robust against severe impulsive noise environments. In order to utilize the benefits of APA and NCA, a complex domain Affine Projection Normalized Correlation Algorithm (AP-NCA) is proposed [156]. AP-NCA algorithm achieves faster convergence for a correlated input and is also robust against impulsive noises.

Regrettably, the algorithms mentioned in this section have limited performance under sparse system identification. Considering the compressive sensing framework [21], [22], [157] and the least absolute shrinkage and selection operator (LASSO) [158], a great deal of attention has received recently in developing numerous adaptive algorithms which incorporate the sparsity of a system and robustness against impulsive noise [159]-[181].

2.4 Summary

This chapter summarizes the review of sparse adaptive algorithms for system identification is presented. A wide variety of algorithms based on proportionate ideas, norm constraints and combinational approaches are detailed. Disadvantage of conventional algorithms is discussed in this chapter. The major difference between Gaussian and non-Gaussian (impulsive) noise is highlighted and adaptive algorithms under impulsive noise environments are discussed.

Chapter 3

Adaptive Combination Approaches for Sparse System Identification

CHAPTER 3

3.1 Motivation

Conventional adaptive filtering algorithms like LMS and Normalized LMS are widely used in system identification applications. However, the performance of these algorithms is degraded when the system's echo path is sparse in nature as in network and acoustic echo cancellation (AEC) scenarios. To address the sparseness property of the echo paths, sparse adaptive filters are extensively used in AEC applications. Several sparsity aware adaptive filter algorithms exist on proportionate-type methods and compressed sensing techniques. Proportionate-type filters are considered as suitable candidate to achieve better performance for sparse echo paths, but they fail to exploit the time varying system sparsity. Considering the ongoing progress in compressive sensing and inspired by LASSO method, a different approach to identify sparse systems based on introducing norm-penalties into the cost function of the existing algorithms has been developed. However, they tend to show poor performance for systems with time varying sparsity. Moreover, there exists some compromise between their steady-state error and convergence speed. To overcome these limitations, the combination approaches of two adaptive filters that combines the output of individual filters through a mixing parameter are proposed.

3.2 Introduction

In system identification applications viz., echo cancellation and channel estimation, the concept of adaptive filter theory is widely used. In general, we often encounter systems with sparse impulse response like network and acoustic echo cancellation systems. A sparse system impulse response consists of negligible number of active coefficients while most of them are inactive [19]. Low complexity adaptive algorithms such as LMS and NLMS that are often used in echo cancellation tend to show slow performance as they apply uniform step size across all filter coefficients. Further, irrespective of the adaptive algorithm used, a tradeoff between MSE and convergence speed always exists [2]. These conventional system identification algorithms are also incapable of utilizing the existing system sparse structure and their performance is reduced when estimating the sparse channels. Taking into account the time varying sparsity of echo path, it is essential to obtain a robust solution that is acceptable to perform effectively with different echo path channels [117].

In an echo cancellation set-up as shown in Fig. 3.1, the weight vector $\mathbf{w}(n)$ of the adaptive filter estimates the echo path impulse response $\mathbf{w_0}(n)$, and produces the output y(n) which is subtracted from the microphone signal, d(n) [2]. The goal of an echo canceller is to eliminate the undesired echo signal by replicating the echo signal and subtract the echo from the echo corrupted signal.

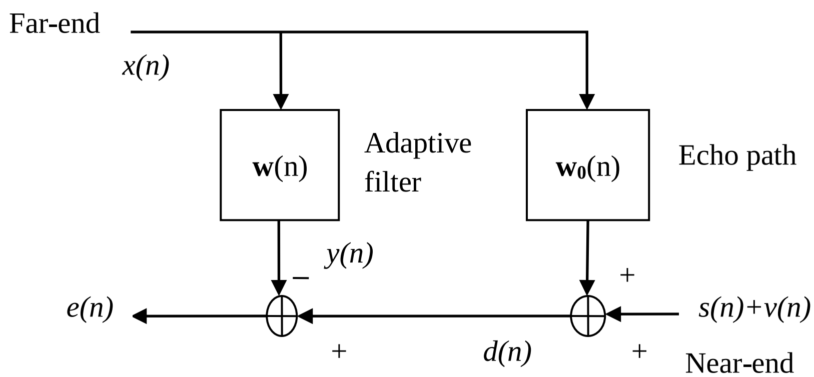


Fig. 3.1: General echo canceller configuration

An echo canceller used to model a sparse echo channel usually requires long adaptive filters; hence the conventional adaptive algorithms suffer from slow convergence [26]. When the sparsity level increases, the traditional methods such as LMS and NLMS algorithms fail to exploit the system sparsity, whereas they perform well for non-sparse systems [3], [123]. To avoid these problems, proportionate adaptive algorithms are developed [17], [53], in which every coefficient is assigned different step size parameter. The convergence speed of Proportionate NLMS (PNLMS) filter [51] is faster than NLMS for sparse echo paths, but NLMS exhibits better performance when the system is not so sparse. Improved Proportionate Normalized LMS (IPNLMS) algorithm was proposed in [52]. Motivated from the ideas of compressed Sensing (CS) [21], [55], Zero-Attracting

LMS (ZA-LMS) has been presented in [85]. In ZA-LMS algorithm ℓ1-norm penalty term is added into the standard LMS error function which makes its implementation simple. Based on reweighted ℓ1-minimization sparse recovery algorithm [20], an improved version of ZA-LMS namely, Reweighted ZA-LMS (RZA-LMS) which enforces the zero attracting term for different taps using reweighted step sizes is also developed in [85]. In [182], several sparsity induced NLMS algorithms for system identification is proposed.

3.3 Proportionate update adaptive filters

3.3.1 Proportionate Normalized LMS (PNLMS)

The update equation of the LMS algorithm is derived as

$$\overline{w}(n+1) = \overline{w}(n) + \mu e(n)\overline{x}(n) \tag{3.1}$$

The update equation of the 'proportionate LMS' (PLMS) algorithm is given as

$$\overline{w}(n+1) = \overline{w}(n) + \mu \overline{Q}(n)e(n)\overline{x}(n) \tag{3.2}$$

where $\overline{Q}(n)$ is the diagonal 'tap selection matrix' that is equipped into the standard LMS. From the convergence analysis of PLMS in both the mean and the mean square [62], the bound on the step-size, μ, is given by

$$0 < \mu < \frac{2}{\overline{x}^T(n)\overline{Q}(n)\overline{x}(n)} \tag{3.3}$$

The a posteriori error is defined as $\tilde{e}(n) = d(n) - \overline{x}^T(n)\overline{w}(n+1)$

We derive the PNLMS by estimating the range of values of μ for which $|\tilde{e}(n)| < |e(n)|$.

By performing a first-order Taylor series expansion of $|\tilde{e}(n)|$ around $|e(n)|$, [63] that is,

$$|\tilde{e}(n)|^2 = |e(n)|^2 + \sum_{i=1}^{M} \frac{\partial |e(n)|^2}{\partial w_i(n)} \Delta w_i(n) \tag{3.4}$$

The partial derivative in (3.4) can be obtained from (3.1) as

$$\frac{\partial |e(n)|^2}{\partial w_i(n)} = -2e(n)\overline{x}(n-i+1) = -2e(n)x_i(n), \quad i = 1,2,...,M \tag{3.5}$$

Whilst the update $\Delta w_i(n)$ in (3.2) is given by

$$\Delta w_i(n) = \mu e(n) q_i(n) x_i(n). \qquad i = 1,2,...,M \tag{3.6}$$

A substitution of (3.5)–(3.6) into (3.4) yields

$$\begin{aligned} |\tilde{e}(n)|^2 &= |e(n)|^2 - 2\mu\left[e(n)\sum_{i=1}^{M} x_i(n)\right]\left[e(n)\sum_{i=1}^{M} g_i(n)x_i(n)\right] \\ &= |e(n)|^2 - 2\mu|e(n)|^2 \bar{x}^T(n)\bar{Q}(n)\bar{x}(n) \\ &= |e(n)|^2\left[1 - 2\mu\bar{x}^T(n)\bar{Q}(n)\bar{x}(n)\right] \end{aligned} \tag{3.7}$$

For the output error to vanish towards zero as $n \to \infty$, we require

$$|\tilde{e}(n)|^2 \le |e(n)|^2\left[1 - 2\mu\bar{x}^T(n)\bar{Q}(n)\bar{x}(n)\right] \tag{3.8}$$

As the squared error terms are non-negative, this will occur if and only if

$$\left[1 - 2\mu\bar{x}^T(n)\bar{Q}(n)\bar{x}(n)\right] < 1 \tag{3.9}$$

resulting in the following bounds on the range of the step-size

$$0 < \mu \le \frac{1}{\bar{x}^T(n)\bar{Q}(n)\bar{x}(n)} \tag{3.10}$$

From (3.10), to minimise the a posteriori error $\tilde{e}(n)$ and equip the PLMS with an optimal learning rate, we have

$$\bar{w}(n+1) = \bar{w}(n) + \mu\frac{\bar{Q}(n)e(n)\bar{x}(n)}{\bar{x}^T(n)e(n)\bar{x}(n)} \tag{3.11}$$

The Eq. (3.11) is the weight update equation of the 'Proportionate Normalized LMS' (PNLMS) algorithm.

To prevent numerical instabilities, a small positive constant δ_{PNLMS} is added in the denominator of (3.11),

$$\bar{w}(n+1) = \bar{w}(n) + \frac{\mu\bar{Q}(n)\bar{x}(n)e(n)}{\bar{x}^T(n)\bar{Q}(n)\bar{x}(n) + \delta_{PNLMS}} \tag{3.12}$$

$where\ \delta_{PNLMS} = \sigma_x^2/M$

M is the length of the adaptive filter.

$\bar{Q}(n) = diag\{q_0(n), \ldots, q_{M-1}(n)\}$ controls the step size. The elements of the control matrix is as shown below

$$q_l(n) = \frac{k_l(n)}{(1/M)\sum_{i=0}^{M-1} k_i(n)}, l = 0,1,\dots,M-1 \quad (3.13)$$

$$k_l(n) = \max(\rho * max[\gamma, |w_0(n)|, \dots, |w_{M-1}(n)|\,], |w_l(n)|) \quad (3.14)$$

The typical value of ρ is $5/M$ and γ is 0.01, respectively. The parameter ρ avoids the filter coefficients from halting if they are negligible compared to the largest coefficient, likewise the parameter γ avoids $w(n)$ from terminating in the initialization step.

The PNLMS algorithm gives faster initial convergence and then slows down substantially. In addition, the PNLMS algorithm is susceptible to the level of sparseness of the system, i.e., for non sparse echo path its convergence rate is decreased. NLMS outperforms PNLMS in case of less sparse or dense impulse response.

3.3.2 Improved Proportionate Normalized LMS (IPNLMS)

The benefits of PNLMS increase with the sparseness of the system and reduce as the unknown system becomes more dispersive. In the search for an adaptation rule which gives performance always better than NLMS and PNLMS, regardless of whether the unknown systems is sparse or dispersive, improved PNLMS (IPNLMS) is proposed [52].

In IPNLMS, the update equation is defined as

$$\overline{w}(n+1) = \overline{w}(n) + \frac{\mu \overline{Q}(n)\overline{x}(n)e(n)}{\overline{x}^T(n)\overline{Q}(n)\overline{x}(n) + \delta_{IPNLMS}} \quad (3.15)$$

Where Q is a diagonal matrix, $\overline{Q}(n) = diag\{q_0(n), q_1(n), \dots q_{M-1}(n)\}$

$$q_l(n) = \frac{q_l(n)}{\|\overline{q}(n)\|_1} = \frac{1-k}{2M} + (1+k)\frac{|w_l(n)|}{2\|\overline{w}(n)\|_1 + \varepsilon},\ l = 0,1,\dots,M-1 \quad (3.16)$$

The IPNLMS parameters μ and k must be properly chosen because,

1) IPNLMS algorithm exhibits faster convergence with larger step size and slower residual misadjustment with smaller step size μ

2) Selection of PNLMS filter with $k = 1$ achieves faster convergence for strongly sparse channels, but has degraded performance for dispersive (non-sparse) channels.

3.4 Proposed Adaptive Affine Combination of IPNLMS Filters

3.4.1 Introduction

Recently, the combination approach of adaptive filters has gained much importance for system identification applications [120], [183], as it provides robustness against systems with varying sparsity and also achieves better performance than each of the combining filters separately. Obtaining the mixing parameter λ(n), through which the combination of the outputs of component filters is performed is crucial in this approach. In [120], the convex combination approach is used where λ(n) in defined by a sigmoid function, i.e., λ(n) is restricted to lie in the range [0, 1]. An approach based on the affine combination of two NLMS adaptive filters is proposed in [126], [127], [128], where λ(n) is calculated from the two component filter output signals. The affine combination as a generalization of convex combination is studied in [184] and two new schemes for estimating the optimal unrealizable affine combiner were proposed and in [125], affine combination analysis was extended for colored inputs and nonstationary environments. In [185], it is shown that the affine combination of two LMS adaptive filters is best suited for multipath mitigation in GPS applications.

3.4.2 Affine Combination approach

Let us consider two adaptive filters ($\mathbf{w}_1$ and $\mathbf{w}_2$) combined in a manner as shown in Fig. 3.2. The input signal is denoted as $\bar{x}(n)$ which is fed to the adaptive filters. The two filters are adapted using their own set of rules and the outputs are combined as

$$y(n) = \lambda(n) y_1(n) + [1 - \lambda(n)] y_2(n), \tag{3.17}$$

to achieve a better performance for the combined filter. $\lambda(n)$ is the mixing parameter. $y_1(n)$ and $y_2(n)$ denote the output of individual filters i.e., $y_i(n) = w_i^T(n-1)\bar{x}(n),\ i=1,2.$

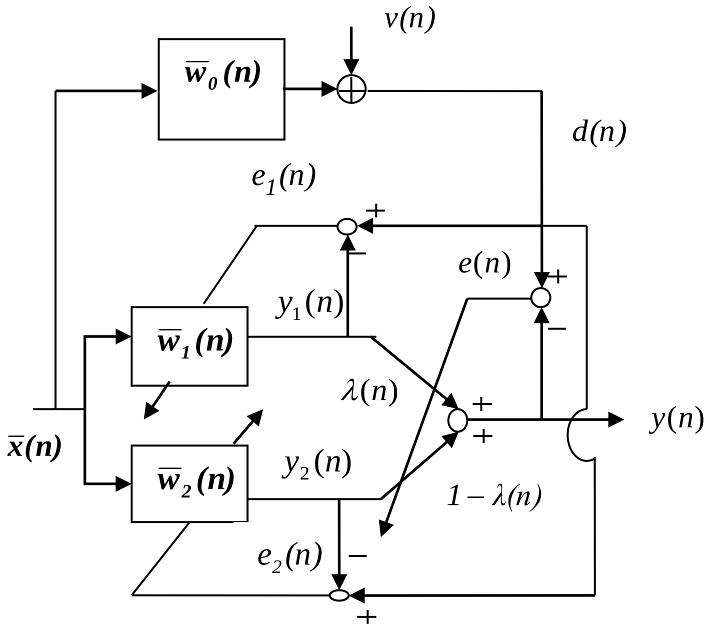


Fig. 3.2: Affine combination of two adaptive filters

As we employ the IPNLMS as component filters, we consider the update equation (3.15) for their adaptation. The parameters $\{\mu_1, k_1\}$ and $\{\mu_2, k_2\}$ are selected in one of the following ways:

a) $\mu_1 > \mu_2$ and $k_1 = k_2$: With this configuration, the combined filter achieves faster convergence with filter having step size μ_1 and minimum steady-state error with filter having step size μ_2, simultaneously.

b) $\mu_1 = \mu_2$, $k_1 < 0$, $k_2 \approx 1$: With this setting, the robustness against systems with varying sparsity and better convergence performance can be achieved.

3.4.3 Derivation of mixing parameter, λ

In our affine combination approach, the mixing parameter '$\lambda(n)$' is any real number, whereas in convex combination approach [120], $\lambda(n)$ lies between 0 and 1.

$$e_i(n) = d(n) - w_i^H(n-1)\bar{x}(n), \tag{3.18}$$

$$d(n) = \bar{w}_0^T \bar{x}(n) + v(n) \tag{3.19}$$

where, $d(n)$ is the desired signal and it is assumed that $v(n)$ is Gaussian noise signal with zero mean and independent of all other signals statistically.

The *a priori* system error signal, $e_a(n)$ is defined as

$$e_a(n) = y_0(n) - y(n) = y_0(n) - \lambda(n)y_1(n) - [1-\lambda(n)]y_2(n) \tag{3.20}$$

where $y_0(n)$ is the output signal of the true system i.e., $y_0(n) = \bar{w}_0^T \bar{x}(n) = d(n) - v(n),$ and $y(n)$ is the output of the adaptive filter.

$\lambda(n)$ is obtained by minimizing the mean square of the *a priori* system error. The derivative of $E[e_a^2(n)]$ with respect to $\lambda(n)$ is given by

$$\begin{aligned}\frac{\partial E\left[e_a^2(n)\right]}{\partial \lambda(n)} &= 2E\left[(y_0(n) - \lambda(n)y_1(n) - (1-\lambda(n))y_2(n))(-y_1(n) + y_2(n))\right] \\ &= 2E\left[(y_0(n) - y_2(n) - \lambda(n)y_1(n) + \lambda(n)y_2(n))(y_2(n) - y_1(n))\right] \\ &= 2E\left[(y_0(n) - y_2(n))(y_2(n) - y_1(n)) + \lambda(n)(y_2(n) - y_1(n))^2\right]\end{aligned} \tag{3.21}$$

Equating the derivative in (3.21) to zero yields

$$\lambda(n) = \frac{E[(d(n) - y_2(n))(y_1(n) - y_2(n))]}{E[(y_1(n) - y_2(n))^2]} \tag{3.22}$$

where the true system output signal, $y_0(n)$ is replaced by its observable noisy version, $d(n)$. So as to achieve a feasible algorithm, we replace the E[.] operators in (3.22) with exponential smoothing of the type

$$p_u(n) = (1-\gamma)p_u(n-1) + \gamma u^2(n), \tag{3.23}$$

where $u(n)$ is the signal to be averaged, $p_u(n)$ is the averaged quantity, and $\gamma = 0.01$. These averaged quantities were then used in (3.22) to obtain λ.

3.4.4 Simulation Results

In this section, we evaluate the performance of the proposed affine combination of two IPNLMS filters by considering the two parameter settings as discussed earlier. The input signal $x(n)$ of 100000 samples is generated and is considered to be white Gaussian noise (WGN). The noise signal $v(n)$, with variance σ_0^2 is added to the reference signal to get an SNR of 70dB initially, and it is changed to SNR= 30dB at n= 80000. The echo path impulse response is generated synthetically using the method given in [59]. The adaptive filter coefficients were initialized to zero vectors. The length of the two component filters is set to M=512. A change in the echo path by circular shift operation is observed at n=50000 sample index to study the filter's reconvergence ability.

The following performance metrics are used to evaluate the proposed filters.

Normalized Weight Misalignment (NWM) evaluates the convergence of the adaptive algorithm. It is defined by

$$\text{NWM(n)=20log10}\frac{\left\|\boldsymbol{w_0}-\boldsymbol{w}(n)\right\|_2^2}{\left\|\boldsymbol{w_0}\right\|_2^2}(dB) \tag{3.24}$$

Echo Return Loss Enhancement (ERLE) measures the attenuation of the echo path. A higher ERLE leads to better reduction in echo.

$$\text{ERLE(n)=}\frac{E\left[d(n)-v(n)\right]^2}{E\left[e(n)-v(n)\right]^2} \tag{3.25}$$

a) With μ_1>μ_2 and k_1=k_2:

The step size values for each component filter are selected as μ_1= 1 and μ_2= 0.2 and the constant k parameter is fixed to $k_1 = k_2 = -0.5$. By using this setting in our combination approach we try to lessen the convergence speed vs steady-state error tradeoff. The channel echo path with sparse impulse response is shown in Fig. 3.3.

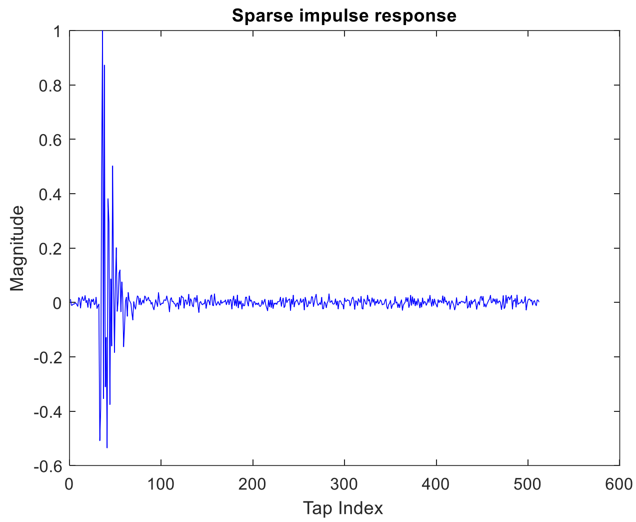


Fig. 3.3: Echo path with sparse impulse response

Fig. 3.4 illustrates the convergence performance of our proposed combination approach to which we will refer hereafter as CIPNLMS. The misalignment curves for the component filters are also displayed for comparison. From the figure, it is observed that the IPNLMS filter with step size μ_1 achieves faster convergence and the filter with step size μ_2 achieves smaller steady-state error. Thus, our CIPNLMS filter keeps the best property of each of the component filter at each time instant i.e., faster convergence and lower steady-state error. This fact is clear even when the echo path is changed (at n=50000). Further, the CIPNLMS filter follows μ_2 - filter steady state value when SNR decreases to 30dB (at n=80000) showing its robust performance.

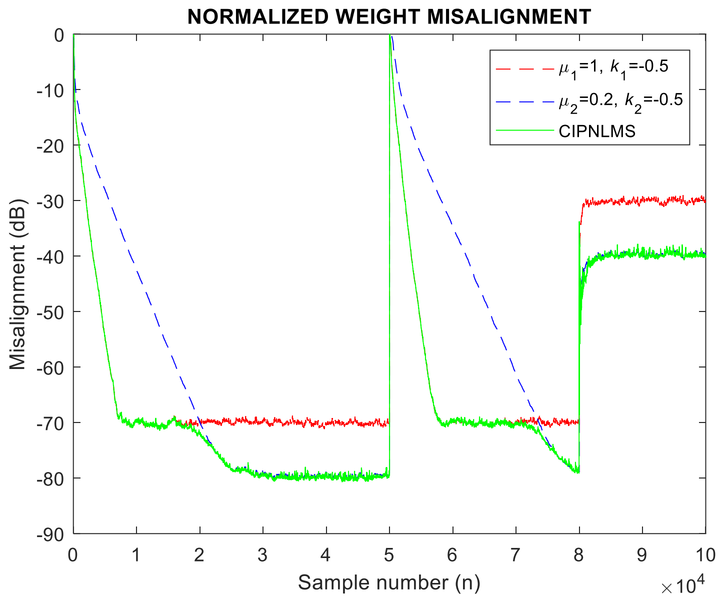


Fig. 3.4: Misalignment performance evaluation of the proposed CIPNLMS filter.

From Fig. 3.5, it is evident that the ERLE value of filter 1 is high during the start of the iteration and after a sudden change in the echo path. But filter 2 dominates filter 1 after a period of time and even at low SNR conditions (n=80000). The ERLE of the CIPNLMS filter always attains the highest value. Thus, our proposed filter achieves higher reduction in echo at every time instant.

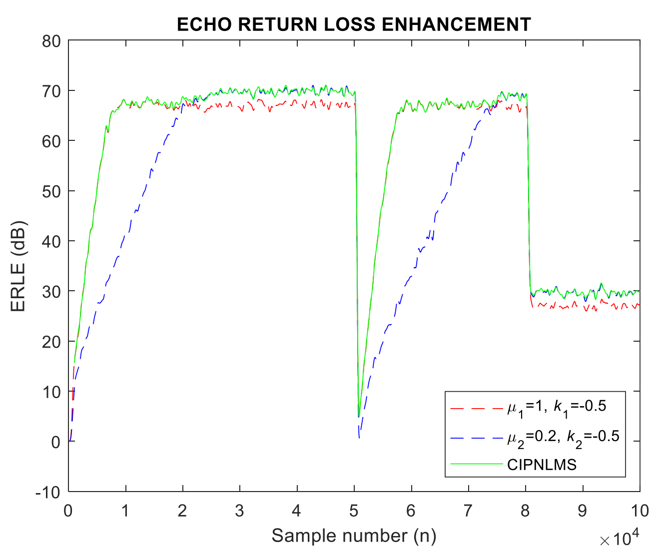


Fig. 3.5: ERLE performance evaluation of the proposed CIPNLMS filter.

Fig. 3.6 depicts the evolution of mixing parameter, λ(n) for the proposed affine combination filter. It can be concluded that the λ(n) is not restricted to lie between [0,1] as in the convex combination approach [120] and it can take any real number.

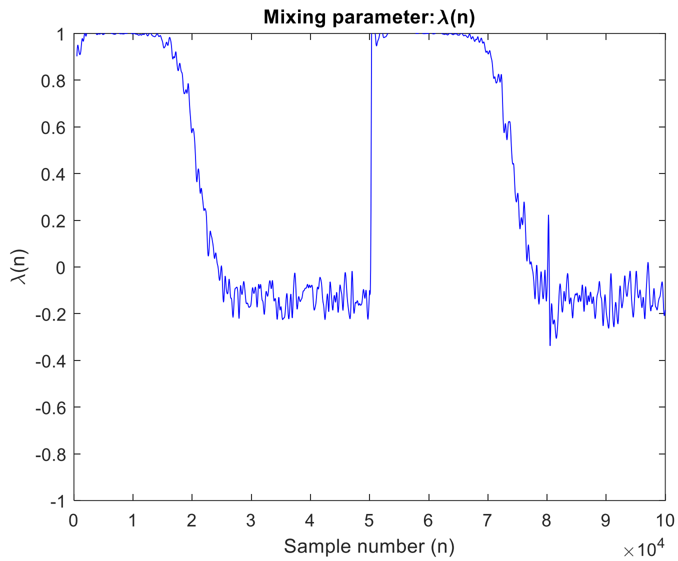


Fig. 3.6: Mixing parameter, $\lambda(n)$ for the proposed affine combination.

b) With $k_1 < 0$, $k_2 \approx 1$ and $\mu_1 = \mu_2$:

We consider selecting the step size value of individual filters as $\mu_1 = \mu_2 = 0.5$; and for parameter k, $k_1 = -0.5$ and $k_2 = 0.9$. With this setting we evaluate the robustness of the proposed combination filter to systems with variable sparsity. The impulse response of the non sparse and sparse echo channels is represented in Fig. 3.7. We carried our simulations assuming that the system echo path is non-sparse initially and later it has changed to a completely sparse system at n= 50000. The WGN noise $v(n)$,with SNR=20dB is added to the reference signal.

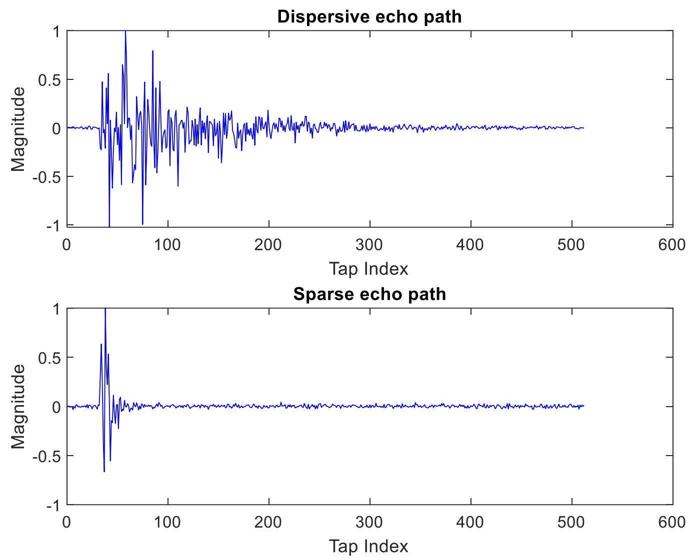


Fig. 3.7: Echo path impulse responses: Dispersive (top), Sparse echo path (bottom).

For non sparse echo path, as shown in Fig. 3.8, the filter with k_1 = -0.5 guarantees fast convergence and for sparse systems the filter with k_2 = 0.9 provides good convergence properties and the CIPNLMS filter always inherits the best component filter performance at each time instant. Thus, we figure out that the proposed approach is robust to system with different degrees of sparsity.

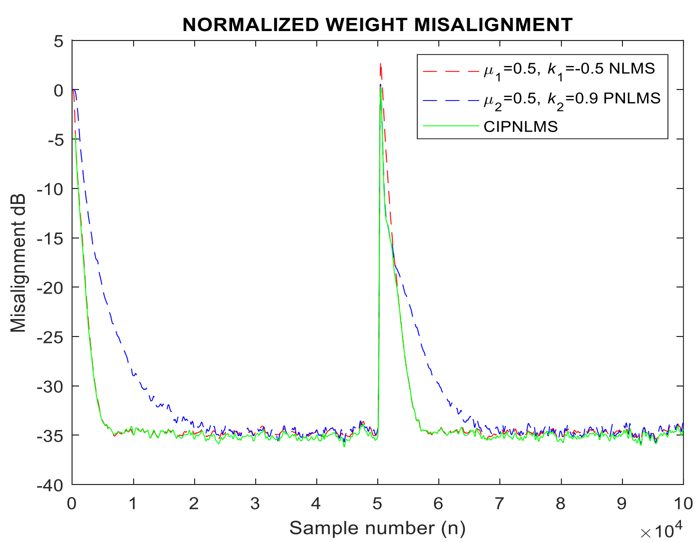


Fig. 3.8: Misalignment performance evaluation of the proposed CIPNLMS filter.

From Fig. 3.9, the CIPNLMS filter maintains the high ERLE value at each iteration by choosing the best of the two component filters depending on the sparsity of the system.

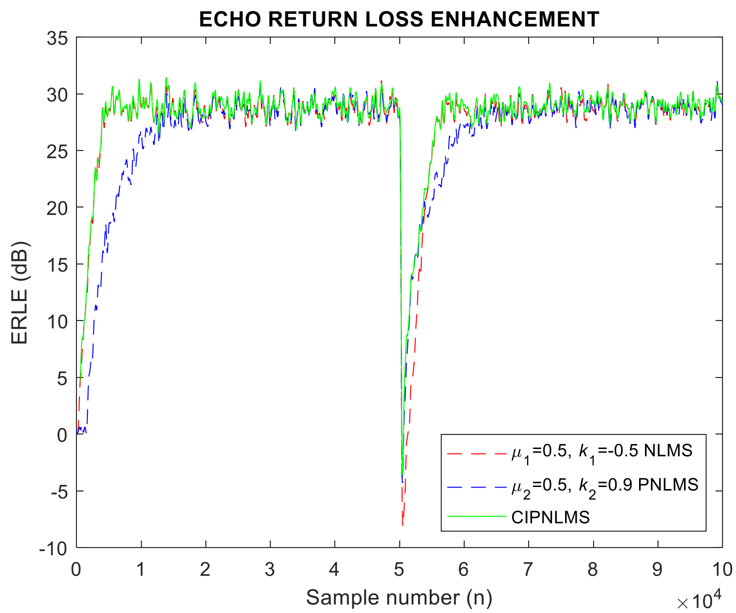


Fig. 3.9: ERLE performance evaluation of the proposed CIPNLMS filter.

3.5 Normalized LMS and Reweighted Zero-Attracting NLMS (RZA-NLMS) Algorithms

Standard NLMS:

The NLMS algorithm is obtained by modifying the step size of LMS given by

$$\mu(n) = \frac{\alpha}{\bar{x}^T(n)\bar{x}(n)} \tag{3.26}$$

where α ($0<\alpha<2$) is the normalized step size.

The normalized LMS (NLMS) filter weight updation is governed by

$$\overline{w}(n+1)=\overline{w}(n)+\frac{\alpha}{\|\overline{x}(n)\|^2}e(n)\overline{x}(n) \tag{3.27}$$

Reweighted ZA-NLMS (RZA-NLMS):

The error function, $L_1(n)$ of ZA-LMS [85] is represented as

$$L_1(n)=\underbrace{\frac{1}{2}e^2(n)}_{\substack{LMS error\\ function}}+\gamma_{ZA}\underbrace{\|\overline{w}(n)\|_1}_{l_1\text{-}norm} \tag{3.28}$$

The ZA-LMS filter update is given by

$$\overline{w}(n+1)=\overline{w}(n)+\mu e(n)\overline{x}(n)-\rho_{ZA}sgn(\overline{w}(n)) \tag{3.29}$$

where $\rho_{ZA}=\mu\gamma_{ZA}$ and the sign function $sgn(\cdot)$ is given by

$$\begin{aligned} sgn(x) &= x/|x|, x\neq 0 \\ &= 0, \quad x=0 \end{aligned} \tag{3.30}$$

γ_{ZA} is a regularization parameter of ZA-LMS.

The efficiency of ZA-LMS algorithm deteriorates as it forces all taps to zero uniformly and is not suitable for non sparse systems.

The RZA-LMS error function is represented by

$$L_2(n)=\frac{1}{2}e^2(n)+\gamma_{RZA}\sum_{i=1}^{N}log(1+\varepsilon|w_i(n)|) \tag{3.31}$$

The RZA-LMS filter update is defined as

$$\overline{w}(n+1)=\overline{w}(n)+\mu e(n)\overline{x}(n)-\rho_{RZA}\frac{sgn(\overline{w}(n))}{1+\varepsilon_{RZA}|\overline{w}(n)|} \tag{3.32}$$

where $\rho_{RZA} = \mu\gamma_{RZA}\varepsilon_{RZA}$ with regularization parameter $\gamma_{RZA} > 0$, and threshold $\varepsilon_{RZA} > 0$.

As the time invariant step size parameter μ of LMS filters cannot guarantee stability, we modify the ZA-LMS and RZA-LMS filters by applying the normalization term in their update equations.

For ZA-NLMS, the weight update is

$$\overline{w}(n+1) = \overline{w}(n) \quad \rho_{ZAN} sgn(\overline{w}(n)) + \frac{\alpha}{\|\overline{x}(n)\|^2} e(n)\overline{x}(n) \tag{3.33}$$

where $\rho_{ZAN} = \alpha\gamma_{ZAN}$ and γ_{ZAN} is a regularization parameter of ZA-NLMS algorithm.

For RZA-NLMS, the weight update is

$$\overline{w}(n+1) = \overline{w}(n) \quad \rho_{RZAN} \frac{sgn(\overline{w}(n))}{1 + \varepsilon_{RZAN}|\overline{w}(n)|} + \frac{\alpha}{\|\overline{x}(n)\|^2} e(n)\overline{x}(n) \tag{3.34}$$

where $\rho_{RZAN} = \alpha\gamma_{RZAN}\varepsilon_{RZAN}$, $\varepsilon_{RZAN} = 20$ and γ_{RZAN} is a regularization parameter of RZA-NLMS algorithm.

3.6 Proposed Affine Combination of NLMS and RZA-NLMS Filters

3.6.1 Description of the combination approach

Let us consider two adaptive filters ($\mathbf{w}_1$ and $\mathbf{w}_2$) combined using our proposed affine combination approach as shown in Fig. 3.2. Filter 1 is adapted using RZA-NLMS algorithm (3.34) and Filter 2 uses the standard NLMS algorithm (3.27). The same input signal $\overline{x}(n)$ is given to both the filters. The output of the proposed filter is given by

$$y(n) = \lambda(n)y_1(n) + [1 - \lambda(n)]y_2(n) \tag{3.35}$$

$\lambda(n)$ is called the mixing parameter, which is not restricted to lie between 0 and 1 as in the convex combination approach [120], $y_1(n)$ and $y_2(n)$ denote the outputs of the combining filters $\mathbf{w}_1$(n) and $\mathbf{w}_2$(n) respectively, i.e., $y_i(n) = w_i^T(n-1)\bar{x}(n),\ i=1,2$.

The desired signal $d(n)$ is defined by

$$d(n) = \bar{w}_0^T \bar{x}(n) + v(n) \tag{3.36}$$

where the noise, $v(n)$ is zero mean Gaussian signal, and do not depend on the other signals statistically. The vector $\bar{\boldsymbol{w}}_0$ is the true weight vector we try to estimate with the proposed approach.

The a priori system error signal is written as

$$e_a(n) = y_0(n) - y(n) = y_0(n) - \lambda(n) y_1(n) - [1 - \lambda(n)] y_2(n) \tag{3.37}$$

and $$y_0(n) = \bar{w}_0^H \bar{x}(n) = d(n) - v(n), \tag{3.38}$$

The derivative of $E[e_a^2(n)]$ with respect to $\lambda(n)$ is given by

$$\begin{aligned}\frac{\partial E\left[e_a^2(n)\right]}{\partial \lambda(n)} &= 2E\left[(y_0(n) - \lambda(n)y_1(n) - (1-\lambda(n))y_2(n))(-y_1(n) + y_2(n))\right] \\ &= 2E\left[(y_0(n) - y_2(n) - \lambda(n)y_1(n) + \lambda(n)y_2(n))(y_2(n) - y_1(n))\right] \\ &= 2E\left[(y_0(n) - y_2(n))(y_2(n) - y_1(n)) + \lambda(n)(y_2(n) - y_1(n))^2\right] \end{aligned} \tag{3.39}$$

The derivative in (3.39) is equated to zero to obtain λ(n),

$$\lambda(n) = \frac{E[(d(n) - y_2(n))(y_1(n) - y_2(n))]}{E[(y_1(n) - y_2(n))^2]} \tag{3.40}$$

where the true system output, $y_0(n)$ is replaced by $d(n)$ which is a valid assumption.

As solving for λ(n) from (3.40) involves hard expectation operations in both the numerator and denominator and to avoid division by zero, it is necessary to find a convenient algorithm. Therefore we use the exponential averaging of the form

$$p_u(n) = (1-\gamma)p_u(n-1) + \gamma u^2(n), \tag{3.41}$$

where $p_u(n)$ is the averaged quantity, $u(n)$ is the signal to be averaged, and $\gamma = 0.01$. These results obtained for both the numerator and denominator are substituted in (3.40) to obtain λ.

3.6.2 Simulation Results

We assess the performance of the proposed affine combination of RZA-NLMS and NLMS filters for adaptive system identification with variable sparsity. We consider simulating two scenarios wherein, Scenario 1 uses general sparse system in which the active coefficients may be randomly located and Scenario 2 consists of clustered sparse system, where a cluster is a gathering of active coefficients.

For Scenario 1, let us assume that the system to be identified has 16 tap coefficients. The input signal, x(n) of 35000 samples is generated. Initially, the system has only two active coefficients and all other coefficients are set to zero i.e., a highly sparse system. After 10000 samples, the semi sparse system with six elements is obtained. Finally, after 20000 samples, a non-sparse system is created with all active tap coefficients. Fig. 3.10 shows the impulse response of a general sparse system generated with different sparse levels. x(n) is considered to be a white Gaussian noise (WGN) with zero mean and unit variance. The noise, $v(n)$ with variance σ_0^2 is added to the reference signal to get an SNR=30dB.

The following performance measures are used as the evaluation metric.

$$\text{Excess MSE}_i\text{(n)} = E[e_i(n) - v(n)]^2, i = 1,2 \tag{3.42}$$

Normalized Weight Misalignment (NWM) (3.24) and Echo Return Loss Enhancement (ERLE) as defined in (3.25)

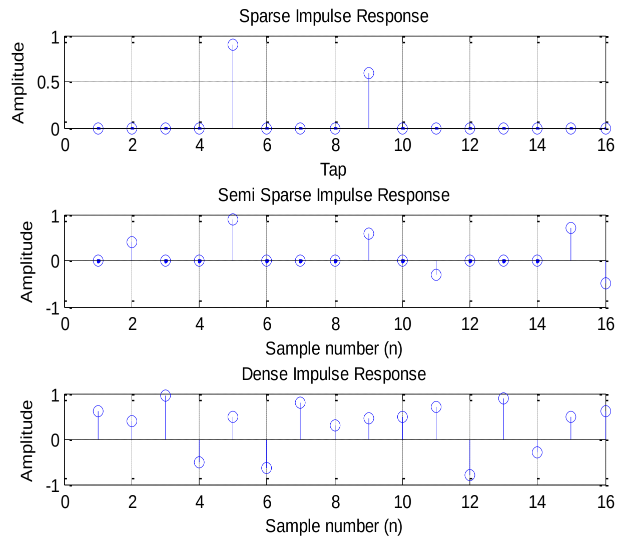


Fig. 3.10: The impulse responses of a general sparse system (Scenario 1)

RZA-NLMS and NLMS adaptive filter coefficients were initialized to zero values with length equivalent to the unknown system. The parameter set for RZA-NLMS filter are: α =0.3, ρ_{RZAN} =0.0008 and ε_{RZAN} =20. We use the same normalized step-size value α for both the filters. The Excess MSE and misalignment performance of the proposed method are compared in Fig. 3.11 and Fig. 3.12 respectively. The average of 100 trials is used in evaluating the results.

For the first scenario, it is verified from Fig. 3.11 that the proposed affine combination filter always achieves the steady state value of the individual filter which has the lower value. For the comparison, we have also simulated the combination approach of two NLMS filters in [126] to which we will refer hereafter as combined NLMS (CNLMS) filter. One of the NLMS filter in CNLMS has the same step size as α and for the other NLMS filter we choose the step size value 0.1. When the system is considered as sparse and semi sparse, our proposed filter attains the EMSE value of that of the RZA-NLMS filter (red line) and for the non-sparse system it attains to that of the NLMS based filter (dotted blue line). From the figure it is clear that the CNLMS filter fails to identify the sparse system.

The misalignment value of our proposed combination is much better than the CNLMS filter for sparse and semi sparse systems and achieves close to that of the RZA-NLMS filter. This is illustrated in Fig. 3.12. In the case of non-sparse system, NWM of our proposed filter follows the NLMS component filter.

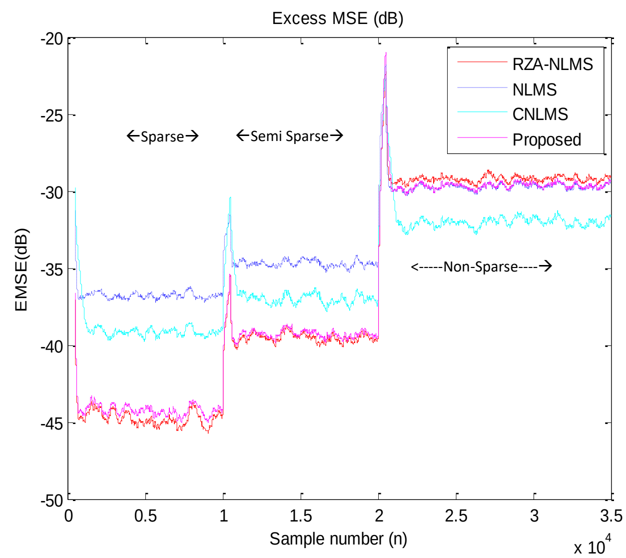


Fig. 3.11: Excess MSE plot comparison with the proposed method.

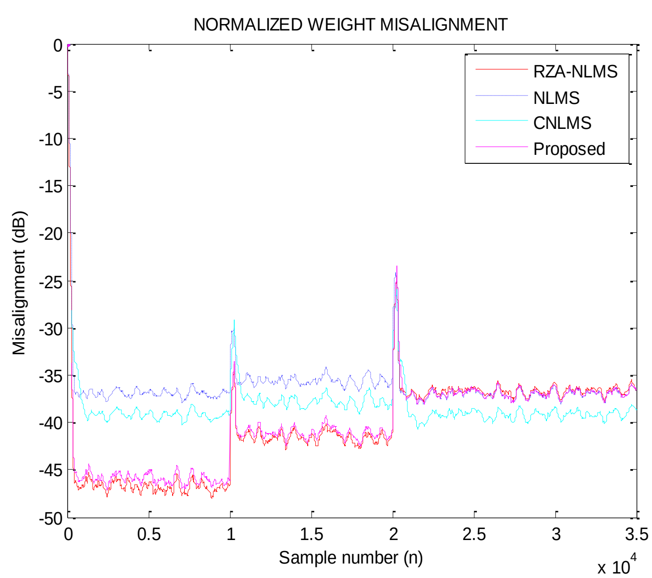


Fig. 3.12: NWM plot comparison with the proposed method.

For Scenario 2, the impulse response of a clustered sparse system is shown in Fig. 3.13. The system has 800 tap elements. Initially, the system has two clusters with active taps at [405; 429] and [569; 590], respectively and is considered to be highly sparse in nature. A change in the system echo path from sparse to semi sparse is observed at 10000 sample index with moderate active coefficients, and then the system is changed to non-sparse at sample index 20000 with all non-zero taps. The input signal, x(n) is considered to be same as in Scenario 1 and the noise variance σ_0^{2} is set to 0.1. The parameters for RZA-NLMS are taken as α=0.8, ρ_{RZAN}=0.00008 and ε_{RZAN}=20.

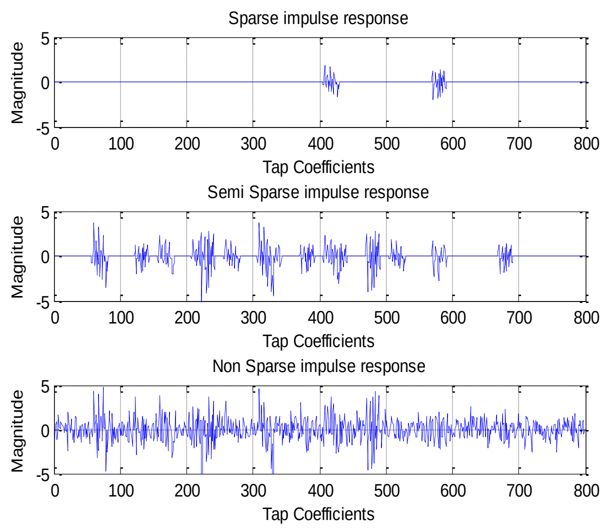


Fig. 3.13: The impulse responses of a clustered sparse system (Scenario 2)

The Excess MSE performance of our proposed combination filter for the second scenario is analyzed from Fig. 3.14. For sparse and semi sparse systems, the proposed filter achieves steady state value of that of the RZA-NLMS filter and for non-sparse system, it achieves the EMSE value of the NLMS filter which is smaller than RZA-NLMS filter. For the system with sparse and semi sparse nature the CNLMS filter has reduced convergence speed, and for non sparse system it has the same convergence speed as our proposed combination filter. Hence, our proposed combination approach guarantees good convergence properties under different sparsity conditions.

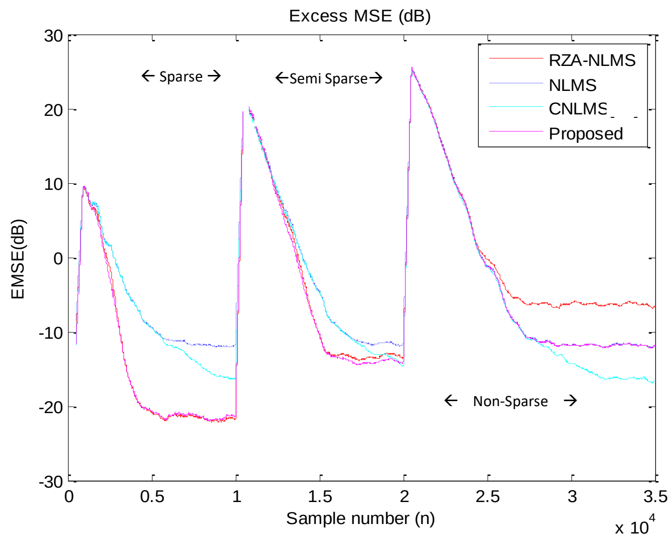


Fig. 3.14: Excess MSE plot comparison with the proposed method

Fig. 3.15 shows the misalignment (dB) of our proposed filter which is same as the RZA-NLMS filter for sparse system and outperforms the two individual filters in the system with semi sparse nature. For the non- sparse system, the RZA-NLMS filter has reduced performance and the misalignment of our proposed method follows the NLMS filter. Thus, the proposed affine combination of two adaptive filters shows better performance than each of the combining filters separately. Hence the affine combination of RZA-NLMS and NLMS filter is capable of providing robust performance in identifying the system with variable sparsity.

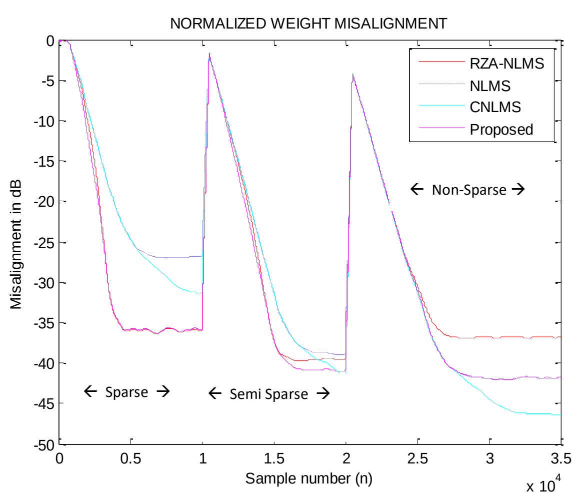


Fig. 3.15: NWM plot comparison with the proposed method.

Fig. 3.16 shows the evolution of mixing parameter, λ(n) for the proposed affine combination filter. It can be concluded that the mixing parameter λ(n) is not restricted to lie in the range [0, 1] as in convex combination approach [120]. It can take any real number.

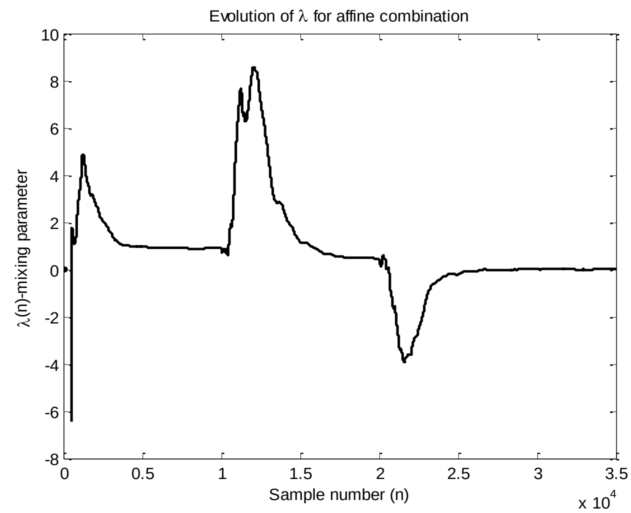


Fig. 3.16: Evolution of mixing parameter, λ(n) for the proposed affine combination.

We will now illustrate the performance of our proposed combination filter considering the input speech signal as a segment of 3.5sec sampled at 8000 Hz. We shall consider the sparse and non-sparse echo paths of Scenario 1 shown in Fig. 3.10 and evaluate the robust behaviour of the proposed combination for this variable sparsity system. The system echo path is switched from the sparse to the non sparse echo path at time t=2sec. The ERLE is considered as the performance metric and simulations are averaged over 100 trials.

Fig. 3.17 represents the ERLE evolution of our proposed filter. When the system echo path is sparse, the RZA-NLMS filter achieves higher ERLE and the combination filter behaves as RZA-NLMS. At t=2sec, when the system echo path is switched to non-sparse, the combination filter attains the value of the other component filter i.e., NLMS. Hence depending on the degree of system sparsity, the proposed combination filter achieves higher ERLE at every iteration and behaves as the most effective component filter.

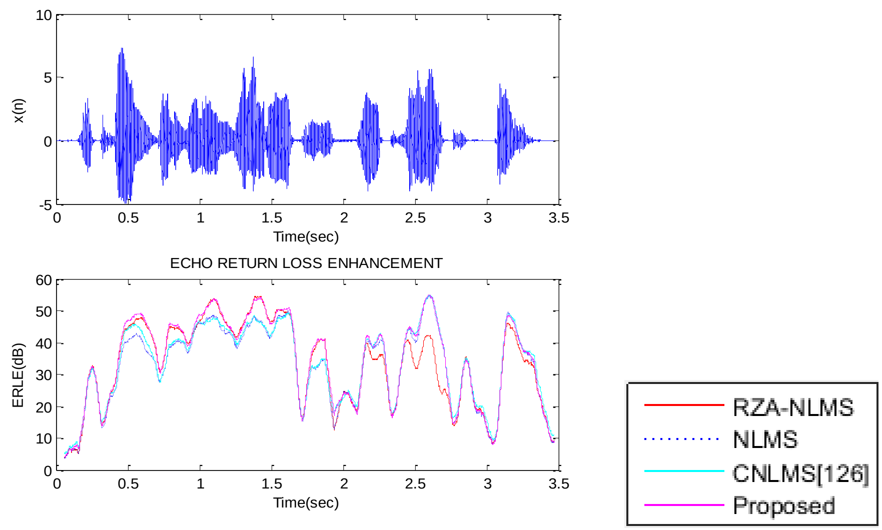


Fig. 3.17: ERLE performance of the proposed filter. The input is a speech signal (top). At t = 2s the echo path changes from sparse to non-sparse. The ERLE performance comparison with the proposed filter (bottom).

3.7 Summary

An adaptive affine combination of two IPNLMS filters that achieves better steady state performance at each iteration and attains robustness to systems with varying degrees of sparsity is proposed. Performance measures Normalized Weight Misalignment (NWM), Echo Return Loss Enhancement (ERLE) are utilized to validate the proposed method. The simulation results exemplifies that the proposed combination approach with different parameter settings has improved the robustness of proportionate filters to systems with varying degrees of sparsity and also alleviated the convergence speed vs steady state error tradeoff of adaptive filters.

Standard NLMS algorithm does not have prior sparse information of the system and fail to exploit the system sparsity. Sparsity aware RZA-NLMS algorithm provides improved steady-state behaviour for highly sparse system but fails to exhibit better performance for non-sparse systems. Hence, an approach that uses an adaptive affine combination of standard NLMS and Reweighted ZA-NLMS (RZA-NLMS) algorithm is proposed that achieves better steady state performance at each iteration. Performance measures Excess

Mean Square Error (EMSE), Normalized Weight Misalignment (NWM), Echo Return Lossless Enhancement (ERLE) are utilized to validate the proposed method. The results illustrates that the proposed combination filter attains improved robustness to systems with varying degrees of sparsity.

Chapter 4

Adaptive System Identification Using Sparse Algorithms based on Lyapunov Stability Theory

CHAPTER 4

4.1Motivation

The idea of adaptive filters is most widely used in the unknown system identification. In contrast to the existing gradient-search related adaptive filtering methods, the Lyapunov Theory-based adaptive filter provides enhanced convergence rate and stability. For the sparse system model, the performance of Lyapunov Adaptive (LA) filter is worsened as it fails to utilize the system sparsity. To cope up with this situation, two algorithms namely, Zero-Attracting Lyapunov Adaptation algorithm (ZA-LA) which relies on $\ell 1$-norm relaxation and Reweighted Zero-Attracting Lyapunov Adaptation algorithm (RZA-LA) by applying log-sum penalty are introduced to improve the convergence performance for the sparse system identification. Further to exploit the systems with variable sparsity, an affine combination scheme of the LA and proposed ZA-LA filters is also developed.

4.2 Introduction

Adaptive filtering algorithms play a significant role in system identification applications e.g., channel estimation and echo cancellation [2]. Ideally the adaptive filter with high convergence rate, stability, good tracking capability and robustness to random noise is desirable for many applications. The widely used optimization technique for optimal filter design is the gradient descent method viz., Least Mean Square (LMS) algorithm [6]. Lyapunov Adaptive Filtering (LA) algorithms in the sense of the Lyapunov stability theory were proposed in [45], [46], [47], [186], [187] to overcome the problems faced by gradient descent-based techniques such as slow rate of convergence, sensitivity to variations in the eigenvalue spread and local minima problem. Moreover, the LA algorithm is independent of the stochastic properties of the input signal and additive noise. However, the LA algorithm suffers from poor convergence performance when the underlying system is identified as sparse such as network and acoustic echo path [27], digital TV transmission channel [25], and underwater channel [43], [188]. In general, the sparse FIR system is characterized by its impulse response which consists of very few active coefficients among many inactive ones [189–191].

Conventional adaptive algorithms neglect the sparse information which is present in the system that leads to degrade their performance when estimating the sparse channels. Recent studies on system identification specify that by utilizing the a priori knowledge about the system sparsity, the estimation performance can be improved substantially. This motivated the researchers towards developing sparse adaptive filtering algorithms in the last few years. In [85], sparse LMS algorithms for system identification are developed by incorporating $\ell 1$-norm penalty and log-sum penalty into the cost function of LMS algorithm. ZA-LMS algorithm is easy to implement and performs well for the highly sparse system, whereas it fails for the system with less sparsity. Reweighted ZA-LMS (RZA-LMS) algorithm performs better than ZA-LMS in less sparse conditions but, at the cost of increased complexity. Following these ideas, we propose two sparse Lyapunov Adaptation algorithms namely, ZA-LA and RZA-LA with application to adaptive sparse system identification.

Over the past decade or so, a combination of adaptive filters has proven to be an efficient way to handle systems with variable sparsity. In [120], [192], an adaptive convex combination of two LMS filters with different parameter setting is proposed to alleviate the speed of convergence vs the residual error trade-off. In [119], a mixture approach of adaptively combining LMS and ZALMS algorithm using a convex combination has been proposed to achieve robustness against time-varying system sparsity. This approach is extended [193] to colored input signal with a convex combination of the Affine Projection Algorithm (APA) and Zero Attracting APA (ZA-APA), and in [194] steady state mean square analysis of convex combination in the context of acoustic echo cancellation is performed. In [117], a convex combination approach using another variant of sparse adaptive filtering i.e., Improved Proportionate Normalized LMS (IPNLMS) algorithms is proposed. In all these works, the authors have used the convex combination to combine effectively the outputs of the individual adaptive filters. The affine combination as a generalization of the convex combination is studied in [184] and in [125] affine combination analysis was extended for colored inputs and nonstationary environments. In [128], transient analysis for the affine combination of two NLMS adaptive filters is studied. In [195], [196], it is demonstrated that affine combination results in faster convergence than the convex combination of two adaptive filters. So, in this work we use affine combination scheme to combine LA and ZA-LA filters to handle systems with variable sparsity.

4.3 Review of LA Algorithms

The basic idea of sparse system identification is to improve the filtering/estimation performance by utilizing the inherent sparse structure information. The block diagram of the sparse system identification is shown in Fig. 4.1. We consider an N length FIR filter coefficients vector $\boldsymbol{w_0} = \left[w_0, w_1, ... w_{N-1}\right]^T$ and the input signal vector $\boldsymbol{\overline{x}(n)} = [x(n), x(n-1), ... x(n-N+1)]^T$ are considered. The input $\boldsymbol{\overline{x}(n)}$ is applied to both the adaptive filter and the unknown sparse system. The desired signal $d(n)$ that is generally corrupted by the observation noise $v(n)$ is, $d(n) = \boldsymbol{w_0^T} \boldsymbol{\overline{x}(n)} + v(n)$. The output estimate $y(n)$ of the adaptive filter $\boldsymbol{\overline{w}(n)}$ is subtracted from the reference signal $d(n)$ to produce an error signal $e(n)$ [197]. The error signal $e(n)$ is then used by the adaptive algorithm in an iterative manner to manipulate the filter coefficients such that the error is minimized.

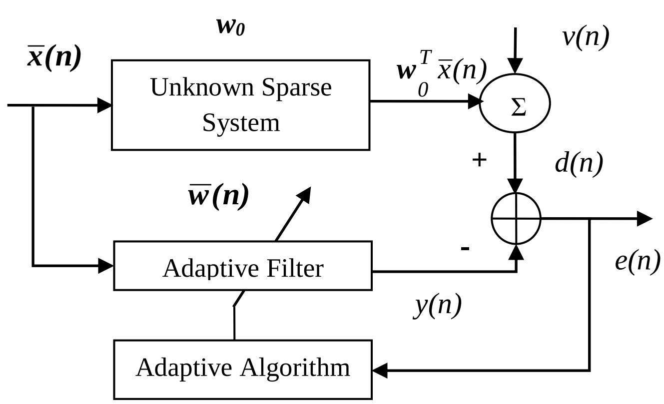


Fig. 4.1: Sparse system identification

The cost function $V(n)$ of the LA filter is defined as the square of error between the desired reference input and filter output which is considered as the Lyapunov function [45], [186],

$$V(n) = e^2(n) \tag{4.1}$$

where,
$$e(n) = d(n) - y(n) \tag{4.2}$$

The Lyapunov adaptive filter weight update law is properly chosen such that $\Delta V(k) = V(k) - V(k-1)$ is negative definite at each iteration. According to Lyapunov stability theory [198], when $\Delta V(k) < 0$, the output of the adaptive filter can asymptotically converge to the desired signal which means that the error $e(n) \rightarrow 0$.

For the given desired signal '$d(n)$' and filter output, $y(n) = \boldsymbol{\bar{w}^T(n)\bar{x}(n)}$, the LA weight update rule is as follows:

$$\boldsymbol{\bar{w}(n)} = \boldsymbol{\bar{w}(n-1)} + g(n)\alpha(n) \tag{4.3}$$

and

$$g(n) = \frac{x(n)}{\lambda_1 + \|x(n)\|^2}\left[1 - \kappa\frac{|e(n-1)|}{\lambda_2 + |\alpha(n)|}\right] \tag{4.4}$$

where $g(n)$ is the adaptation gain and $\alpha(n)$ is the a priori estimation error defined as

$$\alpha(n) = d(n) - \boldsymbol{\bar{w}^T(n-1)\bar{x}(n)} \tag{4.5}$$

$0 \le \kappa < 1$ and λ_1, λ_2 are small positive constants to prevent the singularities of the adaptation gain. The parameter κ controls the convergence rate of the algorithm.

To improve the LA filter tracking performance and robustness, the Lyapunov function is redefined as $V(n) = \beta^n e^2(n)$ and the adaptation gain $g(n)$ of the Lyapunov adaptive algorithm is modified [47], [187] as

$$g(n) = \frac{\bar{x}(n)}{\lambda_1 + \|\bar{x}(n)\|^2}\left[1 - \frac{|e(n-1)|}{\lambda_2 + |\alpha(n)|\beta^{n/2}}\right] \tag{4.6}$$

where $\beta > 1$ is a constant parameter.

To achieve faster convergence rate and lower steady state error performance in case of noisy environments, a step size parameter μ is included in the adaptation gain $g(n)$[48]:

$$g(n) = \frac{\mu \bar{x}(n)}{\lambda_1 + \|\bar{x}(n)\|^2} \left[1 - \kappa \frac{|e(n-1)|}{\lambda_2 + |\alpha(n)|} \right] \tag{4.7}$$

where μ : step size parameter.

The adaptation gain $g(n)$ for the Lyapunov adaptive algorithm can also be defined as

$$g(n) = \frac{\bar{x}(n)}{\lambda_1 + \|\bar{x}(n)\|^2} \left[1 - \frac{|e(n-1)|}{\lambda_2 + \beta^{\frac{n}{2}}(n-1)\, |\alpha(n)|} \right] \tag{4.8}$$

where the adaptive adaptation gain rate is adjustable in order to improve the tracking capability of the algorithm [49], [50].

$$\beta(n) = 1 + \frac{e^2(n-1)}{e^2(n)} \tag{4.9}$$

4.4 Proposed Sparse LA Algorithms

4.4.1 Zero-Attracting Lyapunov Adaptation Algorithm (ZA-LA)

The cost function $V_1(n)$ of ZA-LA is defined by inducing ℓ1-norm penalty in the Lyapunov function of LA filter as

$$V_1(n) = \underbrace{\beta^n e^2(n)}_{V(n)} + \gamma_{ZA} \underbrace{\left\| \bar{\boldsymbol{w}}(\boldsymbol{n}) \right\|_1}_{l_1 - norm} \tag{4.10}$$

where $\gamma_{ZA} > 0$ denotes a regularization parameter which balances the error term and system sparsity.

Then, we have

$$\Delta V_1(n) = V_1(n) - V_1(n-1)$$

$$= \beta^n e^2(n) - \beta^{n-1} e^2(n-1) + \gamma_{ZA}(\|\bar{\boldsymbol{w}}\boldsymbol{(n)}\|_1 - \|\bar{\boldsymbol{w}}\boldsymbol{(n-1)}\|_1)$$

$$= \beta^n [d(n) - \bar{\boldsymbol{w}}^{\boldsymbol{T}}\boldsymbol{(n)}\bar{\boldsymbol{x}}\boldsymbol{(n)}]^2 - \beta^{n-1} e^2(n-1) + \gamma_{ZA}(\|\bar{\boldsymbol{w}}\boldsymbol{(n)}\|_1 - \|\bar{\boldsymbol{w}}\boldsymbol{(n-1)}\|_1)$$

$$= \beta^n [d(n) - (\bar{\boldsymbol{w}}^{\boldsymbol{T}}\boldsymbol{(n-1)} + g^T(n)\alpha(n))\bar{\boldsymbol{x}}\boldsymbol{(n)}]^2 - \beta^{n-1} e^2(n-1) + \gamma_{ZA}(\|\bar{\boldsymbol{w}}\boldsymbol{(n)}\|_1 - \|\bar{\boldsymbol{w}}\boldsymbol{(n-1)}\|_1)$$

$$= \beta^n [\alpha(n) - \alpha(n) g^T(n)\bar{\boldsymbol{x}}\boldsymbol{(n)}]^2 - \beta^{n-1} e^2(n-1) + \gamma_{ZA}(\|\bar{\boldsymbol{w}}\boldsymbol{(n)}\|_1 - \|\bar{\boldsymbol{w}}\boldsymbol{(n-1)}\|_1) \tag{4.11}$$

Substituting the adaptation gain $g(n)$ given in (4.6) into (4.11), we obtain

$$\Delta V_1(n) = -(\beta^{n-1} - 1)e^2(n-1) + \gamma_{ZA}(\|\bar{\boldsymbol{w}}\boldsymbol{(n)}\|_1 - \|\bar{\boldsymbol{w}}\boldsymbol{(n-1)}\|_1) \tag{4.12}$$

The negative definiteness of the Lyapunov function, $V_1(n)$, must be satisfied subjecting to the inequality constraint $\Delta V_1(n) = V_1(n) - V_1(n-1) < 0$ which contributes to the asymptotical stability of the filtering algorithm in the sense of Lyapunov. The first part on the right hand side of (4.12) achieves negative definiteness for all $n > 1$ and $\beta > 1$ [47], [49]. By replacing γ_{ZA} with $-\gamma_{ZA}$, the negative definiteness of $V_1(n)$, is satisfied $\Delta V_1(n) < 0$.

The ZA-LA filter update rule is given by

$$\bar{\boldsymbol{w}}\boldsymbol{(n)} = \bar{\boldsymbol{w}}\boldsymbol{(n-1)} + g(n)\alpha(n) - \gamma_{ZA} sgn(\bar{\boldsymbol{w}}\boldsymbol{(n)}) \tag{4.13}$$

where $sgn(\cdot)$ is the sign function. The ZA-LA algorithm complexity is slightly higher than that of LA algorithm due to the third term of (4.13).

4.4.2 Reweighted Zero-Attracting Lyapunov Adaptation Algorithm (RZA-LA)

The RZA-LA cost function is represented by

$$V_2(n) = \beta^n e^2(n) + \gamma_{RZA} \sum_{i=1}^{N} log(1+\varepsilon\left|w_i(n)\right|) \quad (4.14)$$

where $\gamma_{RZA} > 0$ is a regularization parameter which balances the estimation error term and sparsity of $\boldsymbol{\overline{w}(n)}$.

According to the log–det heuristic approach used in [199], [200] the zero attractor term in (4.14) yields to a convex optimization problem.

To establish this connection, consider the problem

$$\begin{aligned} &\text{minimize } \sum_i log(1+\varepsilon\left|w_i\right|) \\ &\text{subject to } \overline{w} \in R^n \end{aligned} \quad (4.15)$$

Iterative linearization of this objective function gives

$$\boldsymbol{\overline{w}(n+1)} = argmin \sum_i \frac{\left|w_i\right|}{1+\varepsilon\left|w_i(n)\right|} \quad (4.16)$$

If $w_i(n)$ is small, its weighting factor in the next minimization step, $(1+\varepsilon\left|w_i(n)\right|)^{-1}$, is large. So the small entries in $\boldsymbol{\overline{w}(n)}$ are pushed towards zero. Thus, the log-sum penalty function has the potential to be much more sparsity encouraging than the ℓ1-norm.

The RZA-LA filter update is defined as

$$\boldsymbol{\overline{w}(n)} = \boldsymbol{\overline{w}(n-1)} + g(n)\alpha(n) - \gamma_{RZA} \frac{sgn(\boldsymbol{\overline{w}(n)})}{(1+\varepsilon_{RZA}\left|\boldsymbol{\overline{w}(n)}\right|)} \quad (4.17)$$

The RZA-LA algorithm complexity is slightly higher than that of LA algorithm due to the third term of (4.17).

4.5 Computational Complexity

The numerical complexity in terms of additions, multiplications, and divisions of the proposed sparse adaptive algorithms and those of the LMS algorithm and its sparse variants are shown in Table 4.1.

Table 4.1: Comparison of computational complexity of the investigated algorithms

Algorithms	Addition	Multiplication	Division
LMS	2N	2N+1	-
ZA-LMS	3N	3N+1	-
RZA-LMS	3N+1	3N+2	N
LA	3N+5	3N+3	N+1
ZA-LA	4N+5	4N+3	N+1
RZA-LA	4N+6	4N+4	2N+1

It can be seen from Table 4.1, that the proposed sparse algorithms have a moderate computational complexity increase when compared with the original algorithms.

4.6 Proposed Affine Combination of LA and ZA-LA Algorithms (ACLA)

In order to handle the system with varying level of sparseness, we have proposed to combine LA and ZA-LA algorithm using an affine combination approach. The configuration of the proposed affine combination scheme is shown in Fig. 4.2 in which Filter 1 is updated using LA algorithm (4.3), and Filter 2 is updated using ZA-LA algorithm as given in (4.13) respectively. We will hereafter call this filter as the Affine Combined Lyapunov Adaptation (ACLA) filter.

The output signal of ACLA filter is given by

$$y(n) = \lambda(n) y_1(n) + [1 - \lambda(n)] y_2(n) \tag{4.18}$$

where $\lambda(n)$ is the mixing parameter and can be any real number, and $y_1(n)$ and $y_2(n)$ denotes the output of the individual filters, i.e. $y_i(n) = \bar{W}_i^T(n-1)\bar{x}(n), i = 1, 2.$ The a priori error signal $e_a(n)$ is obtained by subtracting the output signal of ACLA filter from the output signal of the unknown system,

$$e_a(n) = y_0(n) - y(n) = y_0(n) - \lambda(n)y_1(n) - [1-\lambda(n)]y_2(n) \tag{4.19}$$

where,

$$y_0(n) = \overline{w}_0^{\,T}\,\overline{x}(n) = d(n) - v(n) \tag{4.20}$$

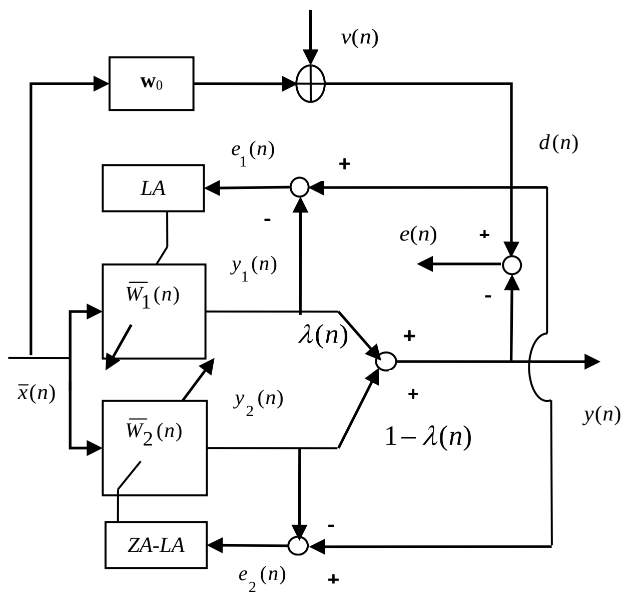


Fig. 4.2: Affine combination of two adaptive filters with mixing parameter λ(n)

To find the mixing parameter λ(n), the derivative of mean square of the a priori error with respect to λ(n) is evaluated and equated to zero.

$$\begin{aligned}\frac{\partial E\left[e_a^2(n)\right]}{\partial \lambda(n)} &= 2E\left[(y_0(n) - \lambda(n)y_1(n) - (1-\lambda(n))y_2(n))(-y_1(n) + y_2(n))\right] \\ &= 2E\left[(y_0(n) - y_2(n) - \lambda(n)y_1(n) + \lambda(n)y_2(n))(y_2(n) - y_1(n))\right] \\ &= 2E\left[(y_0(n) - y_2(n))(y_2(n) - y_1(n)) + \lambda(n)(y_2(n) - y_1(n))^2\right]\end{aligned} \tag{4.21}$$

Setting, $\frac{\partial E\left[e_a^2(n)\right]}{\partial \lambda(n)} = 0$

$$\lambda(n) = \frac{E[(d(n) - y_2(n))(y_1(n) - y_2(n))]}{E[(y_1(n) - y_2(n))^2]} \tag{4.22}$$

where the unknown system output, $y_0(n)$, is replaced by *d(n)*, which is a valid assumption.

The mathematical expectation in the numerator and the denominator of (4.22) are replaced by exponential averaging

$$p_u(n) = (1-\gamma)\,p_u(n-1) + \gamma u^2(n) \tag{4.23}$$

where $p_u(n)$ is the averaged quantity, *u(n)* is the signal to be averaged, and $\gamma = 0.01$. These results obtained for both the numerator and denominator are substituted in (4.22) to obtain λ(n).

4.7 Stability Analysis of ACLA filter

Let us first provide the convergence analysis of LA filter.

When the adaptive filter coefficient vector $\overline{W}(n)$ is updated by (4.3) and (4.6), the Lyapunov function is chosen as

$$V(n) = \beta^n e^2(n) \tag{4.24}$$

According to Lyapunov stability theory [198], the tracking error e(n) will asymptotically converge to zero.

$$e(n) = d(n) - W^T(n)\bar{x}(n) \tag{4.25}$$

$$= d(n) - (\overline{W}^T(n-1) + g^T(n)\alpha(n))\bar{x}(n)$$

$$= d(n) - \overline{W}^T(n-1)\bar{x}(n) - g^T(n)\alpha(n)\bar{x}(n)$$

$$= \alpha(n) - g^T(n)\alpha(n)\bar{x}(n)$$

$$= \alpha(n) - \frac{\bar{x}^T(n)}{\|\bar{x}(n)\|^2}\left[1 - \frac{|e(n-1)|}{|\alpha(n)|\beta^{n/2}}\right]\alpha(n)\bar{x}(n)$$

$$= \alpha(n) - \alpha(n)\left[1 - \frac{|e(n-1)|}{|\alpha(n)|\beta^{n/2}}\right]$$

$$= \frac{\alpha(n)}{|\alpha(n)|}\frac{|e(n-1)|}{\beta^{n/2}}$$

$$= |e(n-1)|\beta^{-n/2}\,\mathrm{sgn}(\alpha(n))$$

$$\therefore |e(n)| = |e(n-1)|\beta^{-n/2} \tag{4.26}$$

$$|e(1)| = |e(0)|\beta^{-1/2}$$

$$|e(2)| = |e(1)|\beta^{-2/2} = |e(0)|\beta^{-(2+1)/2}$$

$$\vdots$$

$$|e(n)| = |e(n-1)|\beta^{-n/2} = \cdots = |e(0)|\beta^{-(n+1)n/4}$$

Therefore, the tracking error e(n) converges to zero exponentially according to

$$|e(n)| = |e(0)|\beta^{-(n+1)n/4} \tag{4.27}$$

Now, consider the output of the individual filters of ACLA filter which is expressed as

$$y_i(n) = \overline{W}_i^T(n-1)\bar{x}(n), i = 1,2 \tag{4.28}$$

and the overall system error is given by

$$e(n) = d(n) - y(n) \tag{4.29}$$

where,

$$d(n) = w_0^T \bar{x}(n) + v(n) \tag{4.30}$$

and $y(n)$ is the ACLA filter output as given in (4.18).

Equation (4.18) can be written as

$$y(n) = \lambda(n)\overline{W_1}^T(n)\bar{x}(n) + [1-\lambda(n)]\overline{W_2}^T(n)\bar{x}(n)$$

$$= \left\{\lambda(n)[\overline{W_1}(n) - \overline{W_2}(n)] + \overline{W_2}(n)\right\}^T \bar{x}(n)$$

$$= \left\{\lambda(n)\overline{W_{12}}(n) + \overline{W_2}(n)\right\}^T \bar{x}(n) \tag{4.31}$$

where,

$$\overline{W_{12}}(n) = \overline{W_1}(n) - \overline{W_2}(n) \tag{4.32}$$

The equivalent weight vector $\overline{W_c}$ of the combined filter can be expressed as

$$\overline{W_c}(n) = \lambda(n)\overline{W_1}(n) + [1-\lambda(n)]\overline{W_2}(n)$$

$$= \lambda(n)[\overline{W_1}(n) - W_2(n)] + \overline{W_2}(n)$$

$$= \lambda(n)\overline{W_{12}}(n) + \overline{W_2}(n) \tag{4.33}$$

Using (4.30) and (4.31) in (4.29), we get

$$e(n) = w_0^T \bar{x}(n) + v(n) - \left\{\lambda(n)\overline{W_{12}}(n) + \overline{W_2}(n)\right\}^T \bar{x}(n)$$

$$= v(n) + \left[\overline{W_{02}}(n) - \lambda(n)\overline{W_{12}}(n)\right]^T \bar{x}(n) \tag{4.34}$$

where,

$$\overline{W}_{02}(n) = w_0 - \overline{W}_2(n) \tag{4.35}$$

The Mean Square Deviation (MSD) of the ACLA filter at time n is

$$MSD_c(n) = E\left\{[w_0 - \overline{W}_c(n)]^T[w_0 - \overline{W}_c(n)]\right\}$$

$$= E\left\{\left[w_0 - \lambda(n)\overline{W}_{12}(n) - W_2(n)\right]^T\left[w_0 - \lambda(n)\overline{W}_{12}(n) - W_2(n)\right]\right\}$$

$$= E\left\{\left[\overline{W}_{02}(n) - \lambda(n)\overline{W}_{12}(n)\right]^T\left[\overline{W}_{02}(n) - \lambda(n)\overline{W}_{12}(n)\right]\right\}$$

$$= E\left[\overline{W}_{02}^T(n)\overline{W}_{02}(n)\right] - 2E\left[\lambda(n)\overline{W}_{02}^T(n)\overline{W}_{12}(n)\right] + E\left[\lambda^2(n)\overline{W}_{12}^T(n)\overline{W}_{12}(n)\right] \tag{4.36}$$

Substituting (4.22) in (4.36) yields

$$MSD_c(n) = E\left[\overline{W}_{02}^T(n)\overline{W}_{02}(n)\right] - E\left\{\frac{[\overline{W}_{02}^T(n)\overline{W}_{12}(n)]^2}{\overline{W}_{12}^T(n)\overline{W}_{12}(n)}\right\} \tag{4.37}$$

The first term of (4.37) corresponds to the MSD of the second adaptive filter, $MSD_2(n)$ and since the MSD is a positive quantity, it indicates that $MSD_c(n)$ is smaller than $MSD_2(n)$.

Equation (4.33) can also be expressed as

$$\overline{W}_c(n) = \overline{W}_1(n) - [1 - \lambda(n)]\overline{W}_{12}(n) \tag{4.38}$$

Now by inserting (4.38) in the first line of (4.36),

$$MSD_c(n) = E\left[\overline{W}_{01}^T(n)\overline{W}_{01}(n)\right] - E\left\{\frac{[\overline{W}_{01}^T(n)\overline{W}_{12}(n)]^2}{\overline{W}_{12}^T(n)\overline{W}_{12}(n)}\right\} \tag{4.39}$$

where,

$$\overline{W}_{01}(n) = w_0 - \overline{W}_1(n) \tag{4.40}$$

The first term of (4.39) corresponds to the MSD of the first adaptive filter, $MSD_1(n)$ and since the MSD is a positive quantity, it indicates that $MSD_C(n)$ is smaller than $MSD_1(n)$. Thus, the combined filter performs at least as well as the best component filter or better than any of them, for every n and the stability is guaranteed.

4.8 Simulation Results

This section shows the simulations that are carried out to evaluate the performance of our proposed sparse algorithms. The length of unknown system $\mathbf{w}_0$ is set to N = 16, and its impulse response consists of only one non-zero value at random position index and zeroes elsewhere making the system highly sparse as shown in Fig. 4.3.

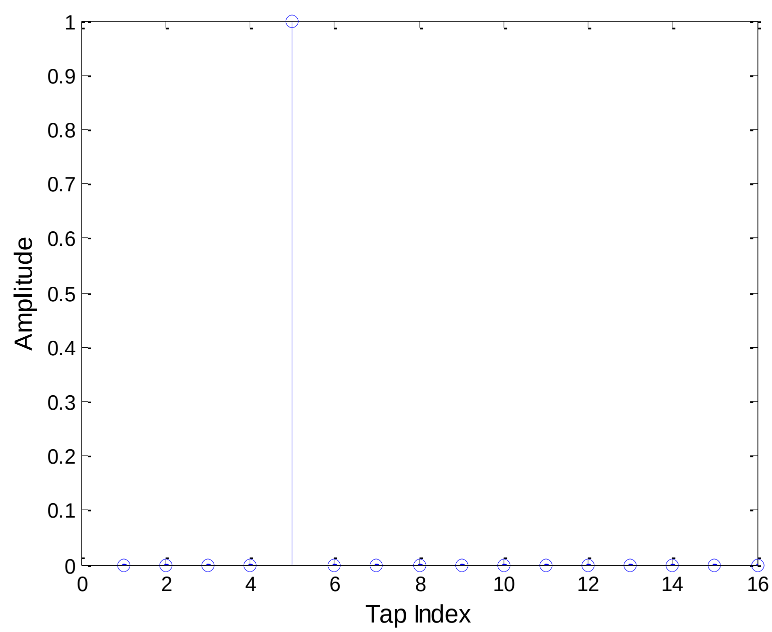


Fig. 4.3: Impulse response of a sparse system

The input signal $\bar{x}(n)$ is considered in two ways:

Case 1: Gaussian random signal with zero mean and unit variance, $N(0,1)$ and

Case 2: Correlated/Colored signal generated by passing a white Gaussian, u(n) through a first-order autoregressive process, AR(1) with a pole 0.8 that is represented as $\bar{x}(n) = 0.8\bar{x}(n-1) + \bar{u}(n)$.

The output of the system is corrupted by an independent white Gaussian noise with variance 0.001.

The performance metrics used to evaluate the proposed algorithms are Mean Square Deviation (MSD), which is defined as

$$MSD\{\overline{W}(n)\} = E\left\{\left\|\overline{w}_0 - \overline{W}(n)\right\|_2^2\right\} \tag{4.41}$$

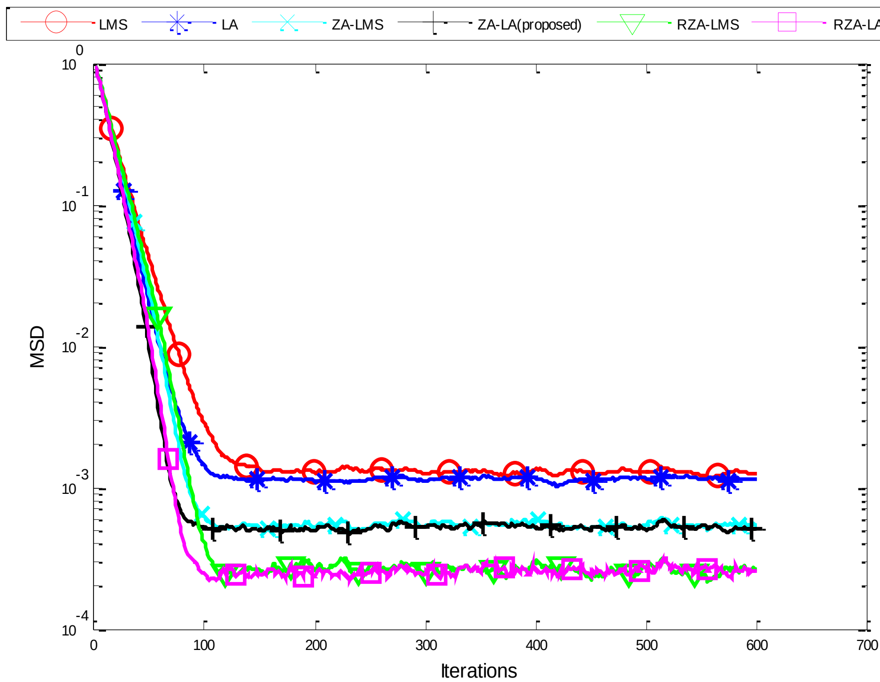
where $E\{\bullet\}$ denotes expectation operator, and $\overline{w}_0$ and $\overline{W}(n)$ are the true FIR filter vector and its adaptive estimator, respectively.

and the MSE is given as

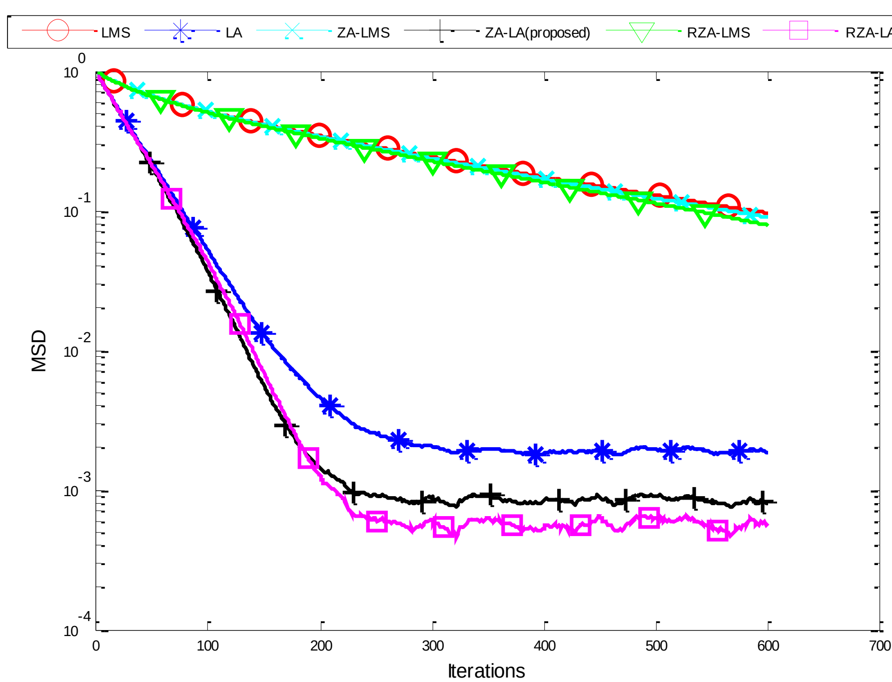
$$MSE(n) = E\{e^2(n)\} \tag{4.42}$$

The average of 200 trials is used in evaluating the results.

From the simulation results shown in Fig. 4.4 (a), for the Gaussian input and when the system is highly sparse, it is observed that our proposed Sparse LA filters (ZA-LA & RZA-LA) converge faster than the existing LMS and LA algorithms which cannot exploit the sparseness information present in the system. In Fig. 4.4 (b), for the case of colored input, it is observed that the proposed LA algorithms have converged while the LMS algorithms fail to converge.



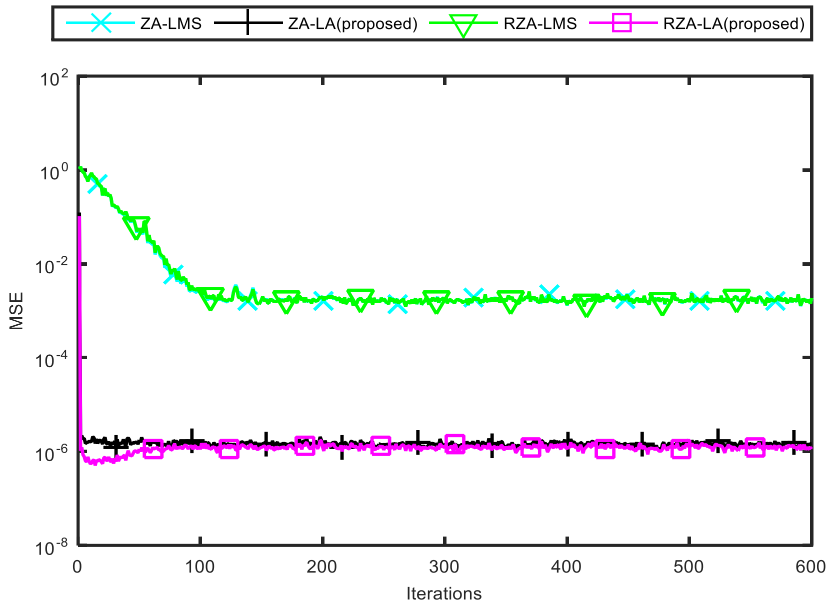
(a)



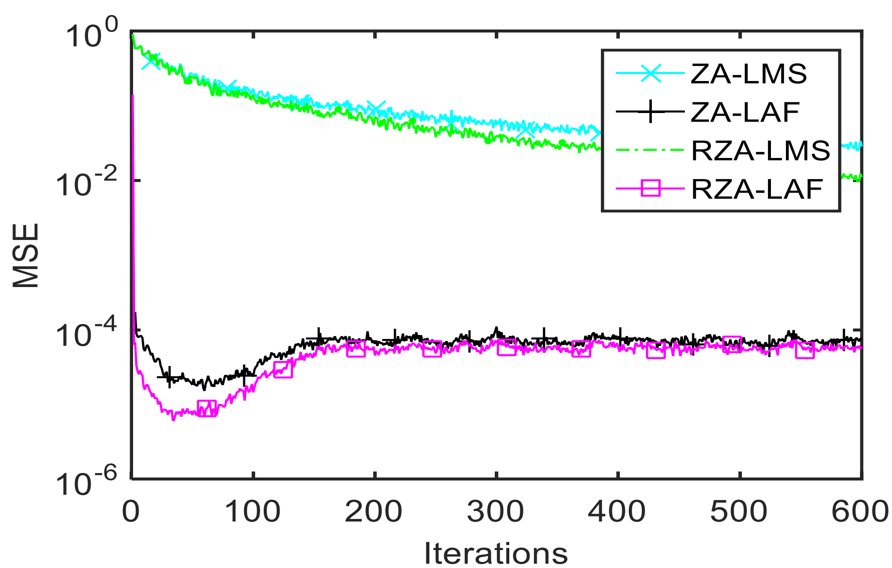
(b)

Fig. 4.4: MSD comparison of the proposed sparse LA algorithms with existing adaptive algorithms for highly sparse system with (a) white input, (b) colored input

The MSE performance of the proposed LA algorithms is depicted in Fig. 4.5. It is observed that the MSE of proposed LA algorithms is lower than the LMS algorithm and their sparse counterpart.



(a)



(b)

Fig. 4.5: MSE comparison of the proposed sparse LA algorithms with existing adaptive algorithms for highly sparse system with (a) white input, (b) colored input

Next, we have considered that the unknown system has all its tap coefficients set to non-zero values, i.e. a non-sparse system zero values, i.e. a non-sparse system. Fig. 4.6 shows the impulse response of the used non-sparse system.

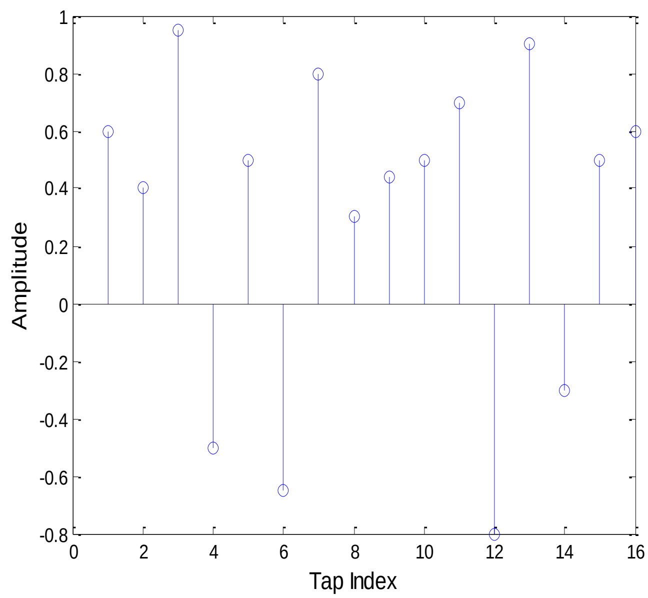
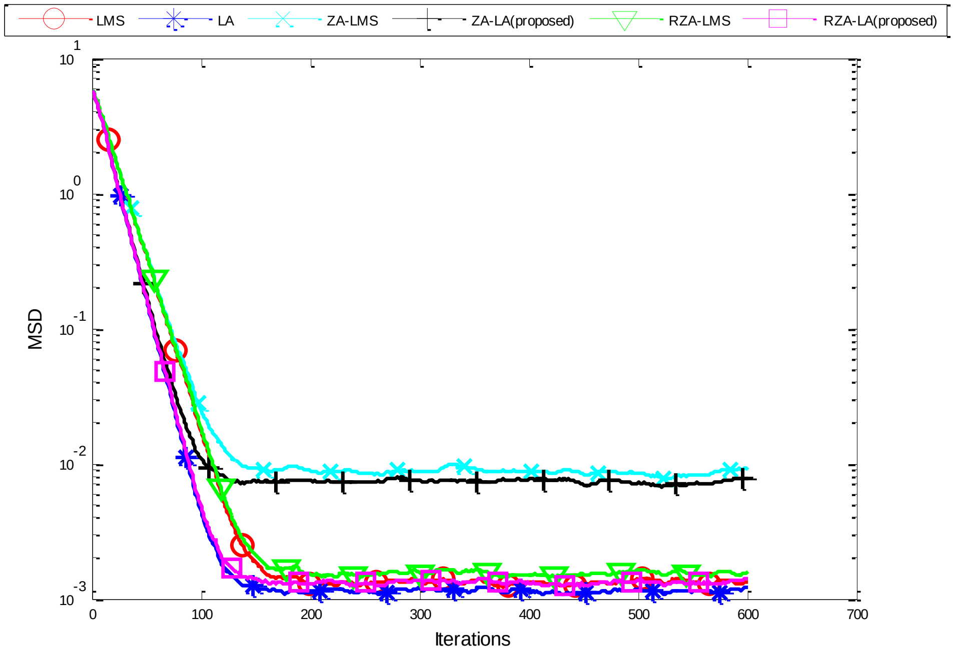
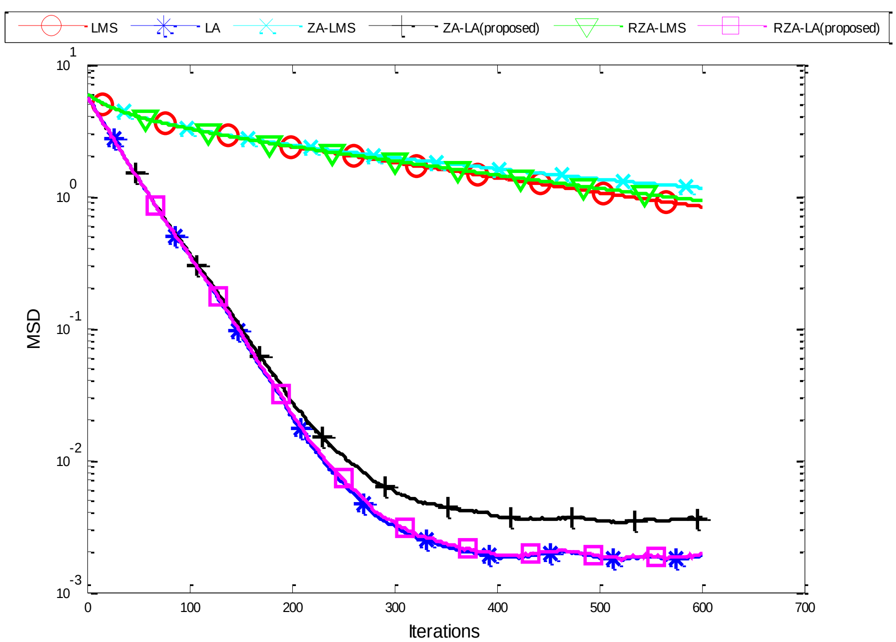


Fig. 4.6: Impulse response of a non-sparse system

With the Gaussian input, the proposed ZA-LA algorithm converges faster than ZA-LMS algorithm, but it exhibits high steady state error when the system is non-sparse as shown in Fig. 4.7 (a). The convergence of RZA-LA algorithm is the same as that of the LA algorithm and much better than the LMS algorithm. In the case of colored input, we observe from Fig. 4.7 (b) that the performance of our proposed algorithms is superior to that of LMS algorithm and their sparse counterpart.



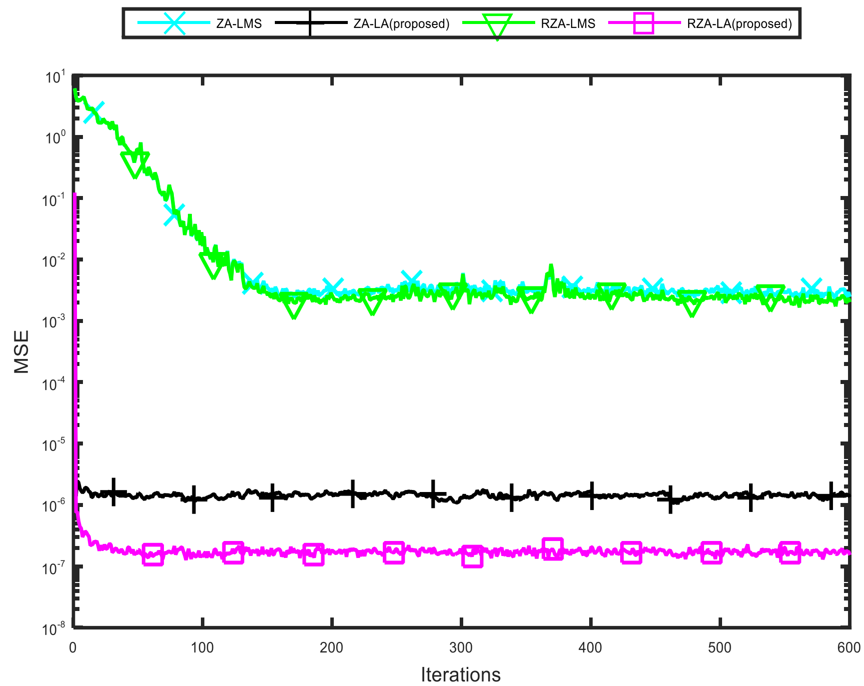
(a)



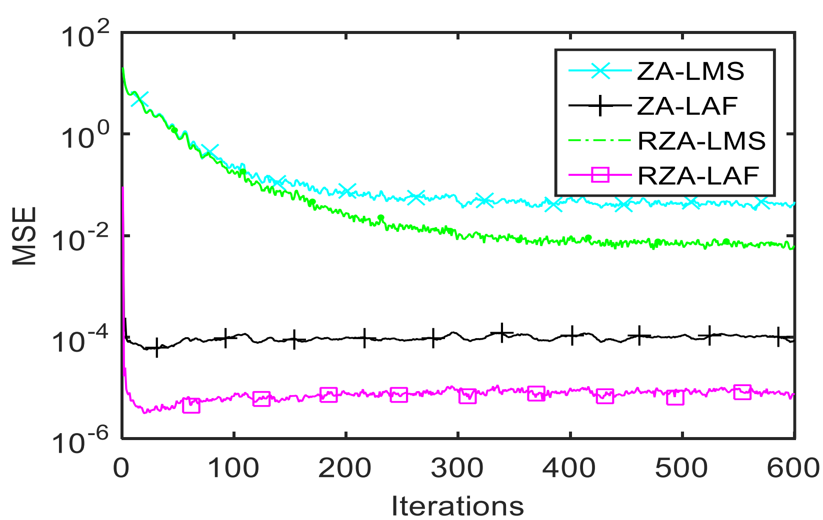
(b)

Fig. 4.7: MSD comparison of the proposed sparse LA algorithms with existing adaptive algorithms for non-sparse system with (a) white input, (b) colored input

The mean square error (MSE) of proposed LA algorithms is shown in Fig. 4.8. The RZA-LA algorithm achieves lower MSE value than that of the ZA-LA algorithm which is lower than that of the LMS algorithm and its sparse counterpart.



(a)



(b)

Fig. 4.8: MSE comparison of the proposed sparse LA algorithms with existing adaptive algorithms for non-sparse system with (a) white input, (b) colored input

The performance of the ACLA filter is analyzed for identifying the system of length N=16 with variable sparsity. Initially, the system is assumed to be a highly sparse system with impulse response as shown in Fig. 4.3. At 600th sample, the system is abruptly converted to a non sparse system with impulse response as shown in Fig. 4.6. The learning curves of ACLA filter are shown in Fig. 4.9 and Fig. 4.10 for white and colored input cases respectively.

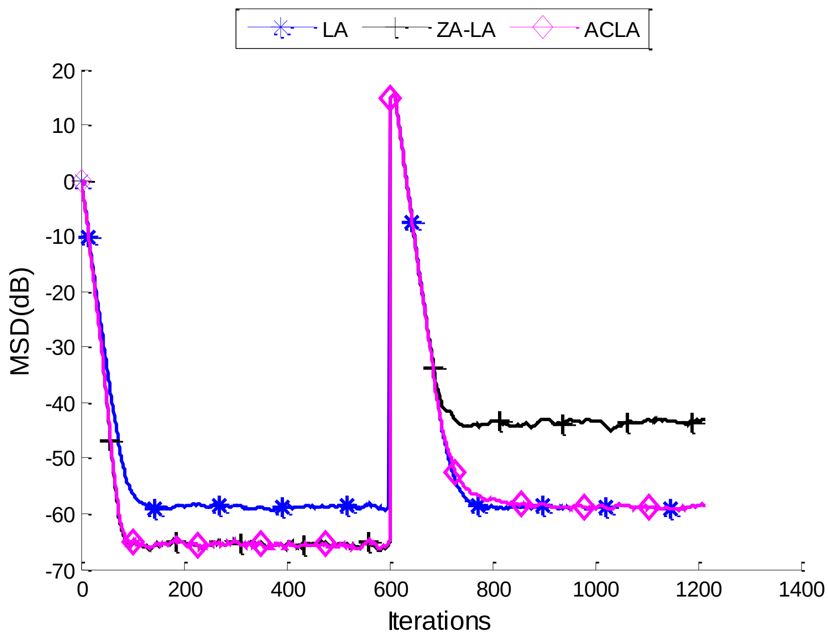


Fig. 4.9: Tracking and steady-state performance of ACLA filter for white input

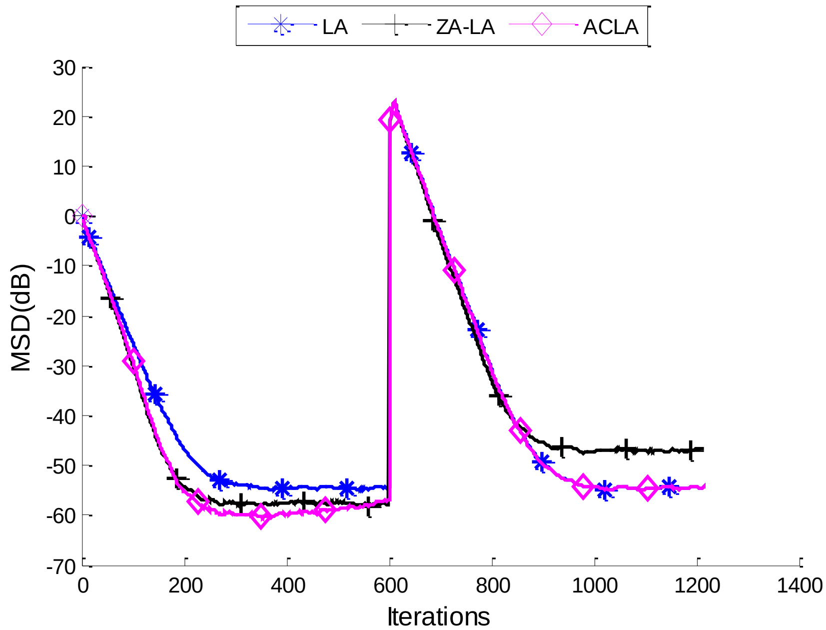


Fig. 4.10: Tracking and steady-state performance of ACLA filter for colored input

From Fig. 4.9, it can be seen clearly that for white input case, the ACLA filter achieves faster convergence and better steady state behaviour. When the system is highly sparse (before 600th sample), the ACLA filter attains the lower steady state value of ZA-LA filter, and when the system is changed to a non-sparse system (at 600th sample), it achieves the steady state value of LA filter. When the input is considered as colored input, the steady state value of the ACLA filter is slightly better than that of the independent filters for a highly sparse system and when the system is converted to a non-sparse system, the ACLA filter behaves like LA filter that achieves the lower steady state value as shown in Fig. 4.10. Thus, the proposed ACLA filter is robust in identifying the systems with variable sparsity. Fig. 4.11 shows the evolution of mixing parameter, $\lambda(n)$ for the ACLA filter.

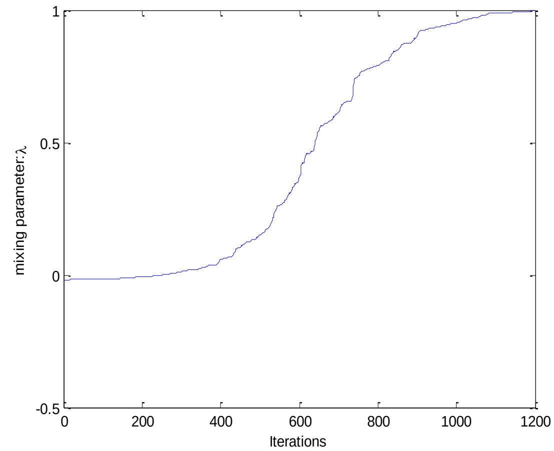


Fig. 4.11: Evolution of mixing parameter, $\lambda(n)$ for the proposed affine combination.

The behaviour of the proposed algorithms with input as a real speech signal sampled at 8 kHz is evaluated in the next simulation. The system is considered to have a varying degree of sparsity. Initially, the system is assumed to be sparse with impulse response shown in Fig. 4.3, and at time t=1.75 sec, it is changed to the non-sparse system as shown in Fig. 4.6. The simulations are averaged over 100 trials. The Echo Return Loss Enhancement (ERLE) is used as the performance metric and is defined as [125],

$$ERLE(dB) = 10\log\frac{E[(d(n)-v(n)]^2}{E[(e(n)-v(n))]^2} \tag{4.43}$$

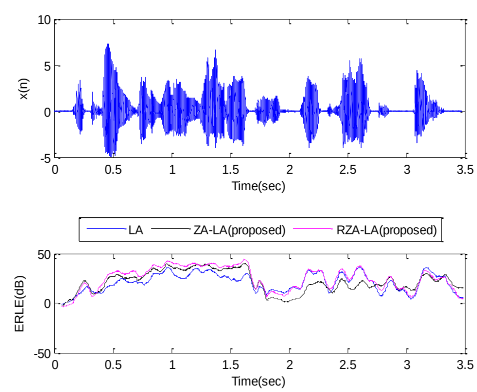


Fig. 4.12: ERLE performance plot. The input signal, x(n) is a real speech signal sampled at 8 kHz (top). At t = 1.75 sec the system impulse response changes from sparse to non-sparse. The ERLE comparison of the proposed filters (bottom).

It is observed from Fig. 4.12 that the proposed RZA-LA filter achieves higher ERLE values and the ERLE of ZA-LA filter is better than that of the LA filter when the system is sparse. At t=1.75 sec, the system's impulse response is switched to non sparse and it is observed that the ERLE performance of ZA-LA is reduced while the RZA-LA filter still attains the higher ERLE value closer to the LA filter which performs well under non sparse conditions. Hence, the robustness of the proposed algorithms is verified.

4.9 Summary

The standard Lyapunov Adaptation algorithm does not exploit sparsity present in an unknown sparse system. In this chapter, we proposed two novel algorithms, ZA-LA and RZA-LA to improve adaptive sparse system identification performance. From the simulation results, the effectiveness of the proposed algorithms is verified for both white input and colored input case in terms of MSD and MSE.

Also, an Affine Combined Lyapunov Adaptation (ACLA) filter is presented to identify the systems with variable sparsity. The proposed combination filter exhibits robustness and achieves lower steady state value irrespective of the level of sparseness for the unknown system. The added complexity of the proposed algorithms is worth considering due to the increased performance.

Chapter 5

Novel Sparse Algorithms under Impulsive Noise Environments

CHAPTER 5

5.1 Motivation

The conventional adaptive filtering algorithms based on minimum mean square error (MMSE) criterion are best suited for the Gaussian noise assumption. However, in practical applications exhibiting the non-Gaussian noise characteristics the performance of these algorithms is unsatisfactory. The Normalized Least Mean Absolute Third (NLMAT) algorithm using the high-order error power (HOEP) criterion is best suited for the impulsive noise environments, but has reduced estimation performance in case of sparse systems.

The performance of the least-mean mixed-norm (LMMN) algorithm degrades seriously due to impulsive interferences. The LMMN adaptive algorithm is modified utilizing the sigmoid cost function to combat the effect of impulsive noise interference. This new algorithm is called as sigmoid LMMN (SLMMN) algorithm. Unfortunately, the proposed SLMMN algorithm cannot utilize the *a priori* sparse structure of the system.

In this chapter, several sparse algorithms are proposed by inducing sparse-penalty functions into the standard NLMAT and the sigmoid LMMN algorithms in order to exploit the system with different levels of sparsity under impulsive noise.

5.2 Introduction

Adaptive filtering algorithms have received much attention over the past decades and are widely used for diverse applications such as system identification, adaptive beamforming, channel estimation, and interference cancellation [1], [2], [3], [201]. The LMS and NLMS are the most popular adaptive algorithms widely used for system identification due to their simplicity and low computational complexity. But, they suffer from reduction in the performance in the presence of colored noise/ impulsive interferences and slow convergence [153]. Therefore, a normalized robust mixed-norm RMN (NRMN) algorithm [202], [204] was presented in order to overcome these drawbacks. In [203], a generalisation of the mixed norm stochastic gradient descent algorithms based on least exponential (LE) algorithms has been presented. NRMN uses the variable step size rather than using the fixed step-size of the RMN algorithm [205]. But, it depends on the variance of the white noise and impulsive noise which is unknown in practice. In the recent years, adaptive filtering algorithms that were based on high-order error power (HOEP) conditions were proposed [142], [143], [144] which can improve the convergence rate performance and mitigate the noise interference effectively. The least mean absolute third (LMAT) algorithm depends on minimizing the mean of the absolute error value to the third power [144,145]. The problem of local minimum is overcome in LMAT algorithm since the error function is a perfect convex function with respect to the filter coefficients. The LMAT algorithm often converges faster than the LMS algorithm and is suitable for various noise conditions [146]. To alleviate the dependence of the input signal power effect, a normalized form of LMAT (NLMAT) algorithm is proposed in [147]. The NLMAT algorithm exhibits good stability and can mitigate non-Gaussian impulsive noise.

In many physical scenarios [25], [26], [148], the unknown system to be identified exhibit sparse representation and it is worthy to note that if the system sparsity is properly used, then the identifying performance can be improved. However, all of the above-mentioned algorithms do not utilize such sparse prior information present in the system and may lose some estimation performance.

Recently, many sparse adaptive filter algorithms that exploit system sparsity have been proposed, the well-known ones are the proportionate normalized LMS (PNLMS) algorithm [51] and its variants [52-54], [57], [59]. On the other hand, motivated by the LASSO [158] and recent advances in compressive sensing [21], [22], [55], a different way for sparse system identification has been proposed in [85]. The approach applies $\ell 1$ relaxation, to improve the performance of the LMS algorithm and the convergence analysis is performed in [206]. To achieve further performance improvement in sparse system identification, reweighted $\ell 1$-norm penalty LMS (RL1-LMS) [96], [207] and Non-uniform norm constraint LMS (NNC-LMS) [88] algorithms were also proposed. Following this idea, sparse-aware LMF algorithms are proposed [208], [209], [210].

Recently, a novel $\ell 0$-norm approximate method based on the correntropy induced metric (CIM) [211] is widely used in sparse channel estimation [179], [212], [213], [214]. However, these methods may be unreliable in estimating the system under non-Gaussian impulsive noise environments. In [215] the impulsive noise is modelled as a sparse vector in the time domain and proved useful for a powerline communication application. Fractional adaptive identification algorithms have been applied for parameter estimation in channel equalization, linear and nonlinear control autoregressive moving average model [216]-[219]. It is observed that fractional based identification algorithms outperform standard estimation methods in terms of accuracy, convergence, stability and robustness.

The Normalized LMAT algorithm has been successfully validated for system identification under impulsive noise environments [147]. In this chapter, the sparse NLMAT algorithms based on different sparsity penalty terms are proposed to deal with sparse system identification under impulsive noise environment and various noise distributions. The following algorithms that integrate similar approaches presented above are proposed: the Zero-Attracting NLMAT (ZA-NLMAT), Reweighted Zero-Attracting NLMAT (RZA-NLMAT), Reweighted $\ell 1$-norm NLMAT (RL1-NLMAT), Non-uniform Norm Constraint NLMAT (NNC-NLMAT) and Correntropy-Induced Metric NLMAT (CIM-NLMAT).

The least-mean mixed-norm (LMMN) algorithm to overcome the sensitivity issues of LMS and LMF is proposed in [149]. Also, the reweighted least-mean mixed-norm algorithm for sparse channel estimation [102] is recently proposed. However, the performance of LMMN algorithm degrades seriously due to impulsive interferences which exist in practical environments [151], [152], [153]. A recent study focuses on the nonlinear sigmoid function which can be used in the traditional cost function of the adaptive filtering algorithms to improve the robustness to impulsive noise [220], [221], [222].

In this chapter, we also propose a modified LMMN algorithm based on the sigmoid cost function. This new algorithm is called as Sigmoid LMMN (SLMMN) which can improve the estimation performance. Unfortunately, the SLMMN algorithm cannot utilize the a priori sparse structure of the system. Hence, two sparsity promoting algorithms, Zero attracting (ZA)-SLMMN and Reweighted Zero attracting (RZA)-SLMMN are proposed to address the system sparsity.

5.3 Review of LMAT and Normalized LMAT Algorithms

The general block diagram of sparse system identification using an adaptive filter is shown in Fig. 5.1.

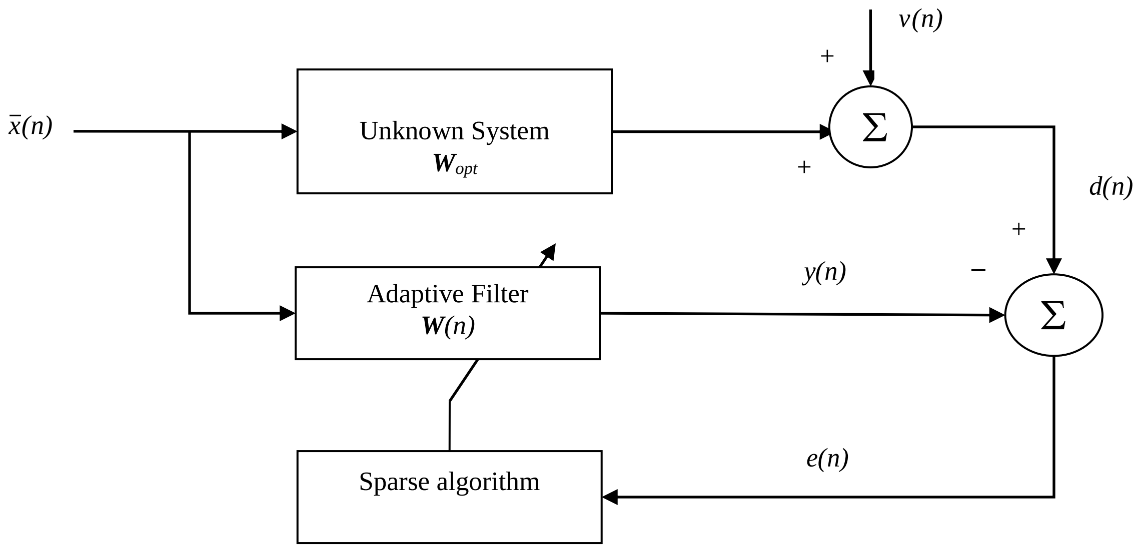


Fig. 5.1: Schematic diagram of sparse system identification

where,

the desired response: $d(n) = \boldsymbol{W}_{opt}{}^{T}\bar{x}(n) + v(n)$,

weight vector: $\boldsymbol{w}_{opt}$ of length L,

input vector: $\bar{x}(n) = [x(n), x(n \quad 1), \ldots. x(n - L + 1)]^{T}$ and

background noise: $v(n)$, with various noise distributions (Gaussian, uniform, Rayleigh and exponential) and impulsive noise.

5.3.1 LMAT Algorithm

The objective function of LMAT algorithm is

$$J_{LMAT}(n) = |e(n)|^3$$

$$= |d(n) - y(n)|^3 \quad (5.1)$$

where $y(n) = \overline{W}^T(n)\overline{x}(n)$ is the output of the adaptive filter, $e(n) = d(n) - y(n)$ denotes the error signal, and $\overline{W}(n) = [w_1(n), w_2(n), ...w_L(n)]^T$ is the weight vector of the adaptive filter.

The gradient descent method is used to minimize $J_{LMAT}(n)$:

$$\overline{W}(n+1) = \overline{W}(n) - \frac{\mu}{3}\frac{\partial J_{LMAT}(n)}{\partial \overline{W}(n)} \quad (5.2)$$

By substituting Eq. (5.1) in the above equation, the LMAT algorithm's update equation is given by

$$\overline{W}(n+1) = \overline{W}(n) + \mu e^2(n) sgn[e(n)]\overline{x}(n) \quad (5.3)$$

where the positive constant μ is the step size parameter.

$sgn(x)$ denotes the sign function of x

The drawback of the LMAT algorithm is that its convergence performance is highly dependent on the power of the input signal.

5.3.2 Normalized LMAT Algorithm

To avoid the limitation of the LMAT algorithm, NLMAT algorithm [147] is derived by considering the following minimization problem [12]:

$$\min_{\overline{W}(n+1)} \left\{ \frac{1}{3} \left| d(n) - \overline{W}^T(n+1)\bar{x}(n) \right|^3 + \frac{1}{2} \|\bar{x}(n)\|^2 \left\| \overline{W}(n+1) - \overline{W}(n) \right\|^2 \right\} \tag{5.4}$$

where $\|\bullet\|$ is the Euclidean norm of a vector.

Derivating Eq. (5.4) with respect to $\overline{W}(n+1)$ and equating to zero yields

$$\overline{W}(n+1) = \overline{W}(n) + \frac{e^2(n) sgn[e(n)] \bar{x}(n)}{\bar{x}^T(n)\bar{x}(n)} \tag{5.5}$$

The weight update equation for the NLMAT algorithm is given by

$$\overline{W}(n+1) = \overline{W}(n) + \mu \frac{e^2(n) sgn[e(n)] \bar{x}(n)}{\bar{x}^T(n)\bar{x}(n) + \delta} \tag{5.6}$$

where μ is a step-size parameter, and δ is a small positive constant to prevent division by zero when $\bar{x}^T(n)\bar{x}(n)$ vanishes.

In the presence of impulsive noises, the squared error term $e^2(n)$ in Eq. (5.6) might degrade the performance of NLMAT algorithm and hence we assign an upper-bound e_{up} to $e^2(n)$.

Thus, the NLMAT algorithm is modified as

$$\overline{W}(n+1) = \overline{W}(n) + \mu \frac{sgn[e(n)] \bar{x}(n)}{\bar{x}^T(n)\bar{x}(n) + \delta} \min\left\{ e^2(n), e_{up} \right\} \tag{5.7}$$

where e_{up} is the upper-bound assigned to $e^2(n)$ in Eq. (5.6) and is expressed as

$$e_{up} = \frac{\sqrt{2\pi}\sigma_e(n)}{\mu} \tag{5.8}$$

The estimate of standard deviation $\sigma_e(n)$ is given by the method in [202]-[205]

$$\sigma_e(n) = \sqrt{\frac{O^T(n) T_w O(n)}{N_w - K}} \tag{5.9}$$

where T_w is the diagonal matrix defined as, $T_w = diag[1,......,1,0,........0]$, that sets the last K elements of $O(n)$ to zero, obtaining an unbiased estimate $\sigma_e(n)$ with the remaining (N_w-K) elements.

$O(n) = sort\left(\left[|e(n)|,.....|e(n-N_w+1)|\right]^T\right)$ contains the N_w most latest values of $e(n)$ arranged in the increasing order of the absolute value.

In general, N_w and K is chosen as $N_w = L$ and $K \geq 1 + \lfloor LPr \rfloor$ where $\lfloor \bullet \rfloor$ is the floor operator and Pr is the probability of the impulsive noise occurrence.

5.4 Proposed Sparse NLMAT Algorithms

To exploit the system sparsity and robustness against impulsive noise, several sparse NLMAT algorithms are proposed by inducing effective sparsity constraints into the standard NLMAT namely, Zero-Attracting NLMAT, Reweighted Zero-Attracting NLMAT, Reweighted ℓ1-norm (RL1)-NLMAT, Non-uniform Norm Constraint (NNC)-NLMAT and Correntropy-Induced Metric (CIM)-NLMAT.

The update equation of LMAT sparse algorithm can be generalized as

$$\overline{W}(n+1) = \underbrace{\underbrace{\overline{W}(n) + Adaptive\ error\ update}_{LMAT} + Sparse\ penalty}_{sparse\ LMAT} \tag{5.10}$$

5.4.1 Zero-Attracting NLMAT (ZA-NLMAT)

The cost function of ZA-LMAT filter with ℓ1-norm penalty is:

$$J_{ZA}(n) = \frac{1}{3}|e(n)|^3 + \lambda_{ZA}\left\|\overline{W}(n)\right\|_1 \tag{5.11}$$

The updating equation of ZA-LMAT filter can be written as

$$\overline{W}(n+1) = \overline{W}(n) - \mu \frac{\partial J_{ZA}(n)}{\partial \overline{W}(n)} \tag{5.12}$$

$$\overline{W}(n+1) = \overline{W}(n) + \mu e^2(n) sgn[e(n)] \overline{x}(n) - \rho_{ZA} sgn(\overline{W}(n) \tag{5.13}$$

where $\rho_{ZA} = \mu\lambda_{ZA}$.

Based on the update Eq. (5.6), the NLMAT filter update equation is generalized as

$$\overline{W}(n+1) = \underbrace{\underbrace{\overline{W}(n) + \begin{matrix} Normalized\ \ Adaptive \\ error\ update \end{matrix}}_{NLMAT} + Sparse\ penalty}_{sparse\ NLMAT} \tag{5.14}$$

In order to avoid the stability issues of Eq. (5.13), the modified form is represented as

$$\overline{W}(n+1) = \overline{W}(n) + \mu \frac{e^2(n) sgn[e(n)] \overline{x}(n)}{\overline{x}^T(n)\overline{x}(n) + \delta} - \rho_{ZA} sgn(\overline{W}(n)) \tag{5.15}$$

Eq. (5.15) corresponds to the updated equation of sparse NLMAT filter.

The update rule of the modified sparse NLMAT algorithm is:

$$\overline{W}(n+1) = \overline{W}(n) + \mu \frac{sgn[e(n)] \overline{x}(n)}{\overline{x}^T(n)\overline{x}(n) + \delta} \min\{e^2(n), e_{up}\} - \rho_{ZA} sgn(\overline{W}(n)) \tag{5.16}$$

which is termed as the Zero-Attracting NLMAT (ZA-NLMAT).

The ZA-NLMAT algorithm based on $\ell 1$-norm penalty is easy to implement and performs well for the system with high sparsity, whereas struggles for the system with less sparsity.

5.4.2 Reweighted Zero-Attracting NLMAT (RZA-NLMAT)

The cost function of the Reweighted ZA-LMAT algorithm is derived by introducing the log-sum penalty

$$J_{RZA}(n) = \frac{1}{3}|e(n)|^3 + \lambda_{RZA} \sum_{i=1}^{L} \log(1 + \varepsilon_{RZA}|w_i(n)|) \tag{5.17}$$

The *i*th filter coefficient is then updated as

$$w_i(n+1) = w_i(n) - \mu \frac{\partial J_{RZA}(n)}{\partial w_i(n)}$$

$$= w_i(n) + \mu e^2(n) sgn[e(n)] x_i(n) - \rho_{RZA} \frac{sgn(w_i(n))}{1+\varepsilon_{RZA}|w_i(n)|} \tag{5.18}$$

The RZA-LMAT update equation can be expressed in vector form:

$$\overline{W}(n+1) = \overline{W}(n) + \mu e^2(n) sgn[e(n)] \overline{x}(n) - \rho_{RZA} \frac{sgn(\overline{W}(n))}{1+\varepsilon_{RZA}|\overline{W}(n)|} \tag{5.19}$$

By using $\lambda_{RZA} \sum_{i=1}^{L} log(1+\varepsilon_{RZA}|w_i(n)|)$ as a sparse penalty in Eq. (5.14), the RZA-NLMAT update equation is denoted by

$$\overline{W}(n+1) = \overline{W}(n) + \mu \frac{sgn[e(n)] \overline{x}(n)}{\overline{x}^T(n)\overline{x}(n)+\delta} min\{e^2(n), e_{up}\} - \rho_{RZA} \frac{sgn(\overline{W}(n))}{1+\varepsilon_{RZA}|\overline{W}(n)|} \tag{5.20}$$

where $\rho_{RZA} = \mu \lambda_{RZA} \varepsilon_{RZA}$ and $\lambda_{RZA} > 0$ is the regularization parameter for RZA-NLMAT.

A logarithmic penalty which is a close measure of ℓ0-norm is considered in RZA-NLMAT. This makes RZA-NLMAT to exhibit a better performance than the ZA-NLMAT. However, the cost function Eq. (5.17) is not convex and the convergence analysis is problematic for Eq. (5.20).

5.4.3 Reweighted ℓ1-norm NLMAT (RL1-NLMAT)

Since the complexity of using the ℓ0-norm penalty is high, a term more similar to the ℓ0-norm i.e., the reweighted ℓ1-norm penalty is used in the proposed RL1-NLMAT algorithm.

The cost function of the reweighted ℓ1-norm LMAT algorithm is given by

$$J_{RL1}(n) = \frac{1}{3}|e(n)|^3 + \lambda_{RL1} \| \bar{f}(n) \overline{W}(n) \|_1 \tag{5.21}$$

where λ_{RL1} is the parameter related with the penalty parameter and the elements of $\bar{f}(n)$ are set to

$$\left[\bar{f}(n)\right]_i = \frac{1}{\delta_{RL1} + \left|\left[\overline{W}(n-1)\right]_i\right|}, \quad i = 0,1,.....L-1 \tag{5.22}$$

with δ_{RL1} is a positive number and hence $\left[\bar{f}(n)\right]_i > 0$ for $i = 0,1,.....L-1$

Differentiating Eq. (5.21) with respect to $\overline{W}(n)$, the RL1-LMAT update rule is

$$\overline{W}(n+1) = \overline{W}(n) - \mu\frac{\partial J_{RL1}(n)}{\partial \overline{W}(n)}$$

$$= \overline{W}(n) + \mu e^2(n)sgn[e(n)]\bar{x}(n) - \rho_{RL1}\frac{sgn\left(\overline{W}(n)\right)}{\delta_{RL1} + \left|\overline{W}(n-1)\right|} \tag{5.23}$$

According to the NLMAT in Eq. (5.7), the update equation of RL1-NLMAT can be written as

$$\overline{W}(n+1) = \overline{W}(n) + \mu\frac{sgn[e(n)]\bar{x}(n)}{\bar{x}^T(n)\bar{x}(n)+\delta}min\left\{e^2(n),e_{up}\right\} - \rho_{RL1}\frac{sgn\left(\overline{W}(n)\right)}{\delta_{RL1} + \left|\overline{W}(n-1)\right|} \tag{5.24}$$

where $\rho_{RL1} = \mu\lambda_{RL1}$.

The cost function Eq. (5.21) is convex unlike the cost function for the RZA-NLMAT. Therefore, the algorithm is guaranteed to converge to the global minimum under some conditions.

5.4.4 Non-uniform Norm Constraint NLMAT (NNC-NLMAT)

In all the above algorithms, there is no adjustable factor that can efficiently adapt the norm penalty itself to the unknown sparse finite impulse response of the system. To further improve the performance of sparse system identification, the non-uniform *p*-norm like constraint is incorporated into NLMAT algorithm.

Let us consider the cost function of sparse NLMAT with *p*-norm like constraint as

$$J(n) = \frac{1}{3}|e(n)|^3 + \lambda\left\|\overline{W}(n)\right\|_p^p \tag{5.25}$$

where $\left\|\overline{W}(n)\right\|_p^p = \sum_{i=1}^{L}\left|w_i(n)\right|^p$ is called L_p^p-*norm* or *p*-norm like, $0 \le p \le 1$

The gradient of the cost function $J(n)$ with respect to $\overline{W}(n)$ is

$$\nabla J(n) = \frac{\partial(\frac{1}{3}|e(n)|^3)}{\partial \overline{W}(n)} + \lambda \frac{\partial \|\overline{W}(n)\|_p^p}{\partial \overline{W}(n)} \tag{5.26}$$

Thus the gradient descent recursion of the filter coefficient vector is

$$w_i(n+1) = w_i(n) - \mu \nabla J(n)$$

$$= w_i(n) + \mu e^2(n) sgn[e(n)] x(n-i) - \rho \frac{p\, sgn(w_i(n))}{|w_i(n)|^{1-p}}, \forall 0 \le i < L \tag{5.27}$$

The zero attractor term in the Eq. (5.27) is produced by the $p-$norm-like constraint which will cause an estimation error for the desired sparsity exploitation. To solve this problem, the non-uniform p-norm like definition which uses a different value of p for each of the L entries in $\overline{W}(n)$ is provided,

$$\|\overline{W}(n)\|_{p,L}^p = \sum_{i=1}^{L} |w_i(n)|^{p_i}, 0 \le p_i \le 1 \tag{5.28}$$

The new cost function using the non-uniform p-norm-penalty is given as

$$J_{NNC}(n) = \frac{1}{3}|e(n)|^3 + \lambda_{NNC} \|\overline{W}(n)\|_{p,L}^p \tag{5.29}$$

The corresponding gradient descent recursion equation is

$$w_i(n+1) = w_i(n) + \mu e^2(n) sgn[e(n)] x(n-i) - \rho_{NNC} \frac{p_i\, sgn(w_i(n))}{|w_i(n)|^{1-p_i}}, \forall\ 0 \le i < L \tag{5.30}$$

$$g(n) = E[|w_i(n)|], \forall\ 0 \le i < L \tag{5.31}$$

$$w_i(n+1) = w_i(n) + \mu e^2(n) sgn[e(n)] x(n-i) - \rho_{NNC} f_i sgn(w_i(n)), \forall 0 \le i < L \tag{5.32}$$

where

$$f_i = \frac{sgn\left[g(n) - |w_i(n)|\right] + 1}{2}, \forall 0 \le i < L \tag{5.33}$$

and $\rho_{NNC} = \mu \lambda_{NNC}$.

The reweighted zero attraction which is used to reduce the bias is introduced to Eq. (5.33).

The weight update equation of NNC-LMAT algorithm is:

$$w_i(n+1) = w_i(n) + \mu e^2(n)sgn[e(n)]x(n-i) - \rho_{NNC} \frac{f_i sgn(w_i(n))}{1+\varepsilon_{NNC}|w_i(n)|}, \forall 0 \le i < L \tag{5.34}$$

where $\varepsilon_{NNC} > 0$

The weight update equation of NNC-NLMAT algorithm can be written in vector form as

$$\overline{W}(n+1) = \overline{W}(n) + \mu \frac{sgn[e(n)]\overline{x}(n)}{\overline{x}^T(n)\overline{x}(n)+\delta} min\{e^2(n), e_{up}\} - \frac{\rho_{NNC} F sgn(\overline{W}(n))}{1+\varepsilon_{NNC}|\overline{W}(n)|} \tag{5.35}$$

where $\boldsymbol{F}$ is defined as

$$\boldsymbol{F} = \begin{bmatrix} f_0 & 0\cdots 0 & \\ 0 & f_1 \cdots 0 & \\ \vdots & 0 \ddots .0 & \\ 0 & 0 \cdots f_{L-1} & \end{bmatrix}_{LxL} \tag{5.36}$$

5.4.5 Correntropy Induced Metric NLMAT (CIM-NLMAT)

Due to the superiority of the correntropy induced metric (CIM) to approximate the ℓ0-norm, CIM is used as the penalty term to impart sparsity in the NLMAT algorithm.

The similarity between two random vectors $\boldsymbol{p} = \{p_1, p_2, ...p_L\}$ and $\boldsymbol{q} = \{q_1, q_2, ...q_L\}$ in kernel space can be measured using CIM which is described as

$$CIM(\boldsymbol{p},\boldsymbol{q}) = \left(k(0) - \hat{V}(\boldsymbol{p},\boldsymbol{q})\right)^{1/2} \tag{5.37}$$

where

$$k(0) = \frac{1}{\sigma\sqrt{2\pi}}, \tag{5.38}$$

$$\hat{V}(\boldsymbol{p},\boldsymbol{q}) = \frac{1}{L}\sum_{i=1}^{L} k(p_i, q_i) \tag{5.39}$$

For the Gaussian kernel, $k(p,q) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{e^2}{2\sigma^2}\right)$ (5.40)

here $e = p - q$ and σ is the kernel width.

The CIM provides a good approximation for the ℓ0-norm that can be represented as

$$\|\boldsymbol{p}\|_0 \sim CIM^2(\boldsymbol{p},0) = \frac{k(0)}{L}\sum_{i=1}^{L}\left(1 - exp\left(-\frac{(p_i)^2}{2\sigma^2}\right)\right) \quad (5.41)$$

The Gaussian kernel-based CIM is integrated with cost function of LMAT algorithm i.e.,

$$J_{CIM}(n) = \frac{1}{3}|e(n)|^3 + \lambda_{CIM} CIM^2(\overline{W}(n),0)$$

$$= \frac{1}{3}|e(n)|^3 + \lambda_{CIM}\frac{k(0)}{L}\sum_{i=1}^{L}\left(1 - exp\left(-\frac{(w_i(n))^2}{2\sigma^2}\right)\right) \quad (5.42)$$

The gradient of the cost function $J_{CIM}(n)$ with respect to $\overline{W}(n)$ is

$$\nabla J_{CIM}(n) = \frac{\partial J_{CIM}(n)}{\partial W(n)}$$

$$= -e^2(n) sgn[e(n)] x(n-i) + \lambda_{CIM}\frac{1}{L\sigma^3\sqrt{2\pi}} w_i(n) exp\left(-\frac{(w_i(n))^2}{2\sigma^2}\right) \quad (5.43)$$

The weight update equation of CIM-LMAT is denoted by

$$w_i(n+1) = w_i(n) - \mu\nabla J_{CIM}(n)$$

$$= w_i(n) + \mu e^2(n) sgn[e(n)] x(n-i) - \rho_{CIM}\frac{1}{L\sigma^3\sqrt{2\pi}} w_i(n) exp\left(-\frac{(w_i(n))^2}{2\sigma^2}\right) \quad (5.44)$$

where $\rho_{CIM} = \mu\lambda_{CIM} > 0$ is a regularization term which balances the estimation error and sparsity penalty.

Equation (5.44) can be rewritten in matrix form as

$$\overline{W}(n+1)=\overline{W}(n)+\mu e^2(n)sgn[e(n)]\overline{x}(n)-\rho_{CIM}\frac{1}{L\sigma^3\sqrt{2\pi}}\overline{W}(n)exp\left(-\frac{\left\|\overline{W}(n)\right\|^2}{2\sigma^2}\right) \quad (5.45)$$

By using $\lambda_{CIM}\frac{k(0)}{L}\sum_{i=1}^{L}\left(1-exp\left(-\frac{(w_i(n))^2}{2\sigma^2}\right)\right)$ as a sparse penalty in Eq. (5.14), the CIM-NLMAT update equation is given by

$$w_i(n+1)=w_i(n)+\mu\frac{sgn[e(n)]\overline{x}(n)}{\sum_{i=0}^{L-1}(x(n-i))^2+\delta}min\{e^2(n),e_{up}\}x(n-i)-\rho_{CIM}\frac{1}{L\sigma^3\sqrt{2\pi}}w_i(n)exp\left(-\frac{(w_i(n))^2}{2\sigma^2}\right), \quad (5.46)$$

The matrix form of CIM-NLMAT algorithm is expressed as

$$\overline{W}(n+1)=\overline{W}(n)+\mu\frac{sgn[e(n)]\overline{x}(n)}{\overline{x}^T(n)\overline{x}(n)+\delta}min\{e^2(n),e_{up}\}-\rho_{CIM}\frac{1}{L\sigma^3\sqrt{2\pi}}\overline{W}(n)exp\left(-\frac{(\overline{W}(n))^2}{2\sigma^2}\right) \quad (5.47)$$

5.4.6. Computational Complexity

The numerical complexity of the proposed sparse algorithms in terms of additions, multiplications, divisions, square-roots, and comparisons per iteration is shown in Table 5.1.

Table 5.1: Comparison of computational complexity of the investigated algorithms

Algorithms	Additions	Multiplications	Divisions	Square-roots	Comparisons	Exponents
NLMAT [147]	4*L*+2	4*L*+1	2	1	*L*ln*L*+2	-
ZA-NLMAT	5*L*+2	5*L*+1	2	1	*L*ln*L*+2	-
RZA-NLMAT	5*L*+3	5*L*+2	*L*+2	1	*L*ln*L*+2	-
RL1-NLMAT	5*L*+2	5*L*+1	*L*+2	1	*L*ln*L*+2	-
NNC-NLMAT	5*L*+3	6*L*+2	*L*+2	1	*L*ln*L*+2	-
CIM-NLMAT	4*L*+2	7*L*+1	*L*+2	1	*L*ln*L*+2	L

5.4.7. Simulation Results

In this section, the performance of the proposed sparse algorithms is evaluated in the context of system identification using various noise distributions and impulsive noise environment. The unknown system $\boldsymbol{W}_{opt}$ is of length L=16 and its channel impulse response (CIR) is assumed to be sparse in the time domain. The adaptive filter is also assumed to be of the same length. The proposed algorithms are compared under different sparsity levels S=1 & S=4. The active coefficients are uniformly distributed in the interval (-1, 1) and the position of the nonzero taps in the CIR is randomly chosen. The Gaussian white noise with variance ${\sigma_x}^2 = 1$ is considered as the input signal $\bar{x}(n)$. The correlated signal $\bar{z}(n)$ is obtained using a first-order autoregressive process, AR(1), with a pole 0.5 and is given by $\bar{z}(n) = 0.5\bar{z}(n-1) + \bar{x}(n)$. The system background noise comprises of

Case 1: only White Gaussian noise with $N(0,1)$,

Case 2: impulsive noise + White Gaussian noise with $N(0,1)$,

Case 3: impulsive noise + uniformly distributed noise within the range (-1, 1),

Case 4: impulsive noise + Rayleigh distributed noise with 1 and

Case 5: impulsive noise + Exponential distributed noise with 2.

The impulsive noise generated by the Bernoulli-Gaussian (BG) process is given as $\xi(n) = a(n)I(n)$, where $a(n)$ is a white Gaussian signal with $N(0,\sigma_a^2)$ and $I(n)$ is a Bernoulli process expressed by the probability $p\{I(n)=1\} = Pr$, $p\{I(n)=0\} = 1 - Pr,$ where Pr represents the probability of the impulsive noise occurrence. We choose Pr =0.01 and ${\sigma_a}^2 = 10^4/12$.

The performance metrics Mean Square Deviation (MSD) and Excess Mean Square Error (EMSE) are used to evaluate the performance of the proposed algorithms which are expressed as

$$MSD(dB)=10\log_{10}\left\|\mathbf{W}_{opt}-\overline{W}(n)\right\|_2^2 \tag{5.48}$$

and

$$EMSE(dB)=10\log_{10}[\varepsilon(n)]^2,\text{ respectively.} \tag{5.49}$$

$\varepsilon(n)=\theta^T(n)\overline{x}(n)$, where $\theta(n)=\mathbf{W}_{opt}-\overline{W}(n)$

The average of 100 independent trials with SNR=20 dB is used in evaluating the results.

In order to show the effectiveness of the proposed sparse NLMAT algorithms, a comparison with the NRMN algorithms is performed. In Fig. 5.2, the simulation results for the proposed algorithms are shown for the white Gaussian input and Case 1 as the background noise for the system with sparsity S=1. The simulation results shown in Fig. 5.3 are carried out for the white Gaussian input with Case 2 background noise with sparsity level S=1. It can be seen from Figs. 5.2 and 5.3 that the proposed sparse NLMAT algorithms exhibit better performance than NLMAT and NRMN algorithms in terms of MSD for the very sparse system. Moreover, the proposed CIM-NLMAT algorithm achieves lower steady state error.

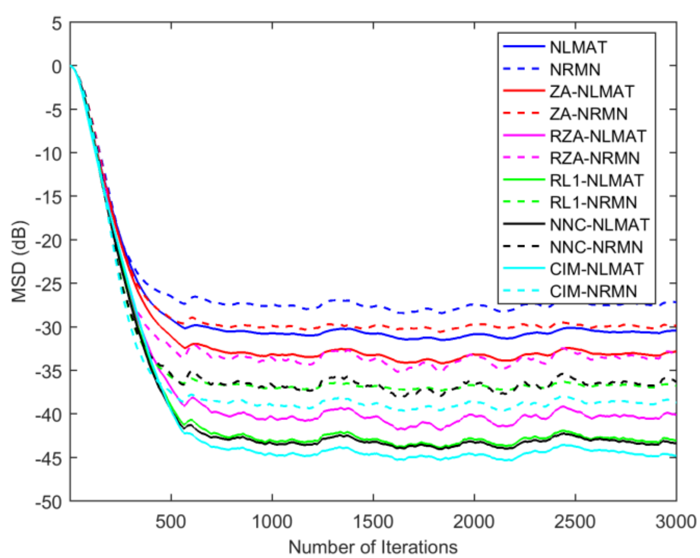


Fig. 5.2: MSD performance for the proposed algorithms with Case 1 as the background noise and the Gaussian white input signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 5\times10^{-5}$, $\rho_{RZA} = 3\times10^{-4}$, $\rho_{RL1} = 1\times10^{-5}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 1\times10^{-3}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 2\times10^{-3}$, $\sigma = 0.05$

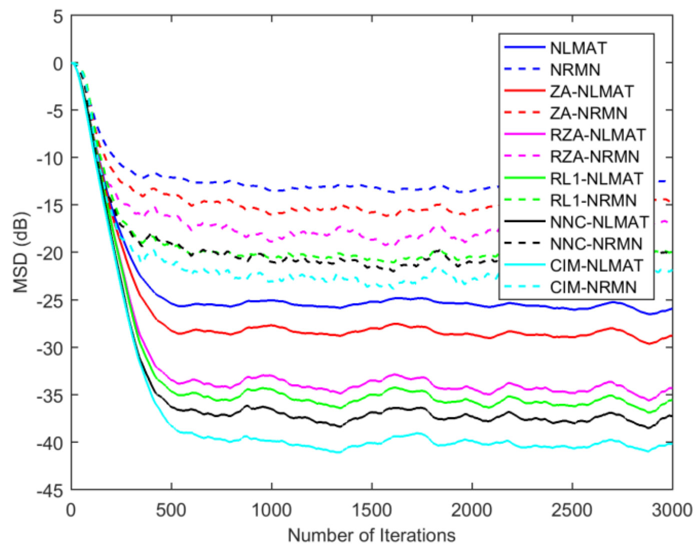


Fig. 5.3: MSD performance for the proposed algorithms with Case 2 as the background noise and the white Gaussian input signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 4\times10^{-4}$, $\varepsilon_{RZA} = 20$, $\rho_{RL1} = 1\times10^{-5}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 1\times10^{-3}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 2\times10^{-3}$, $\sigma = 0.05$

In Fig. 5.4, the simulation results for the proposed algorithms are shown for the white Gaussian input and Case 3 as the background noise for the system with sparsity S=1. In Fig. 5.5, the input is white Gaussian with Case 4 background noise for the system with sparsity S=1. In Fig. 5.6, the input is white Gaussian signal and the background noise as Case 5 for the system with sparsity S=1.

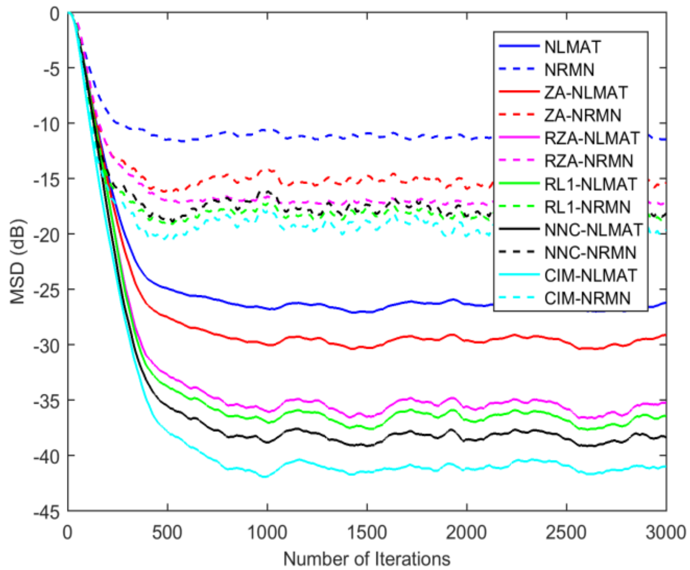


Fig. 5.4: MSD performance for the proposed algorithms with Case 3 background noise and the white Gaussian input signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}, \rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 4\times10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 1\times10^{-3}, \varepsilon_{NNC} = 20$, $\rho_{CIM} = 2\times10^{-3}, \sigma = 0.05$

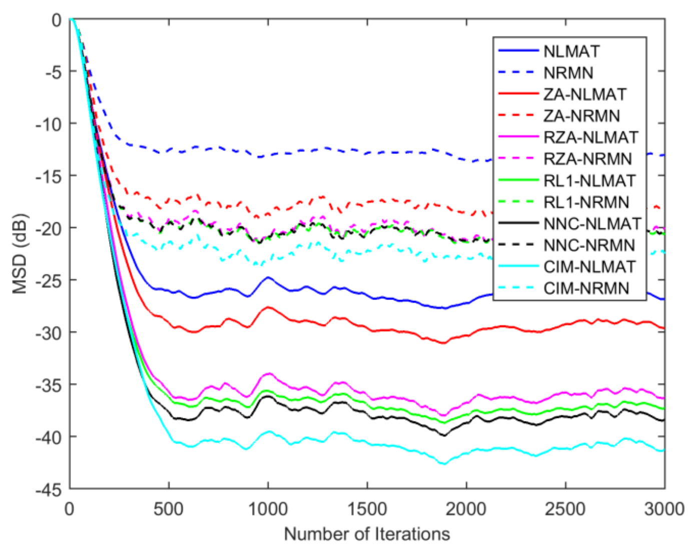


Fig. 5.5: MSD performance for the proposed algorithms with Case 4 background noise and the input is white Gaussian signal for the system with sparsity S=1. The simulation

parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times 10^{-3}$, $\rho_{ZA} = 1\times 10^{-4}, \rho_{RZA} = 5\times 10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 3\times 10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 1\times 10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 2\times 10^{-3}, \sigma = 0.05$

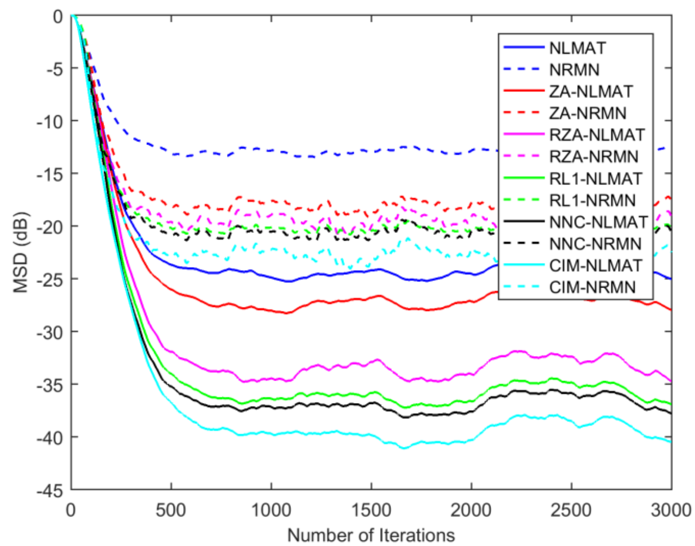


Fig. 5.6: MSD performance for the proposed algorithms with Case 5 as the background noise and the white Gaussian input signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times 10^{-3}$, $\rho_{ZA} = 1\times 10^{-4}, \rho_{RZA} = 5\times 10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 3\times 10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 2\times 10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 2\times 10^{-3}, \sigma = 0.05$

It can be easily seen from Figs. 5.4, 5.5 and 5.6 that the proposed sparse NLMAT algorithms provide better performance than NLMAT and NRMN algorithms in terms of MSD for the very sparse system. As shown above, the proposed CIM-NLMAT algorithm achieves lower steady state error too.

The EMSE values of the proposed algorithms obtained for different noise cases with uncorrelated input and system sparsity S=1 are given in Table 5.2. It is confirmed that the proposed sparse algorithms outperform the NLMAT algorithm in identifying a sparse system.

Table 5.2: Comparison of EMSE values for different NLMAT algorithms with Uncorrelated input signal and different noise cases with system sparsity S=1

EMSE in dB						
Uncorrelated input signal with	**NLMAT**	**ZA-NLMAT**	**RZA-NLMAT**	**RL1-NLMAT**	**NNC-NLMAT**	**CIM-NLMAT**
Gaussian noise	-30.5626	-32.2503	-35.4460	-36.0610	-36.1562	-36.3304
Gaussian + impulsive noise	-24.2397	-25.7307	-27.6974	-28.0753	-29.1522	-30.8891
Uniformly + impulsive noise	-24.4995	-25.9092	-27.6775	-28.0074	-29.1906	-31.4352
Rayleigh + impulsive noise	-24.7076	-26.1345	-28.3539	-29.8174	-29.4793	-31.5281
Exponential + impulsive noise	-23.2402	-24.6765	- 26.8894	-28.1389	-28.3890	-29.1493

In the simulations shown in Figs. 5.7 – 5.11 the input signal is the correlated/colored input and the system background noise from Case 1 to Case 5 are considered accordingly with system sparsity S=1. It is observed from Figs. 5.7–5.11 that the proposed sparse NLMAT algorithms exhibit better performance than NLMAT and NRMN algorithms in terms of MSD for the very sparse system and colored input. Moreover, like for the previous simulations, the proposed CIM-NLMAT algorithm achieves the lowest steady state error.

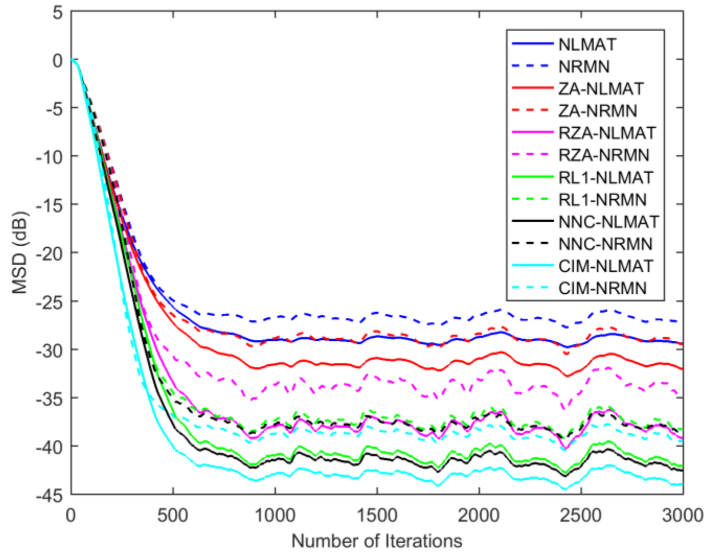


Fig. 5.7: MSD performance for the proposed algorithms with Case 1 as the background noise and the input is the correlated signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 5\times10^{-5}, \rho_{RZA} = 3\times10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 1\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 2\times10^{-3}, \sigma = 0.05$

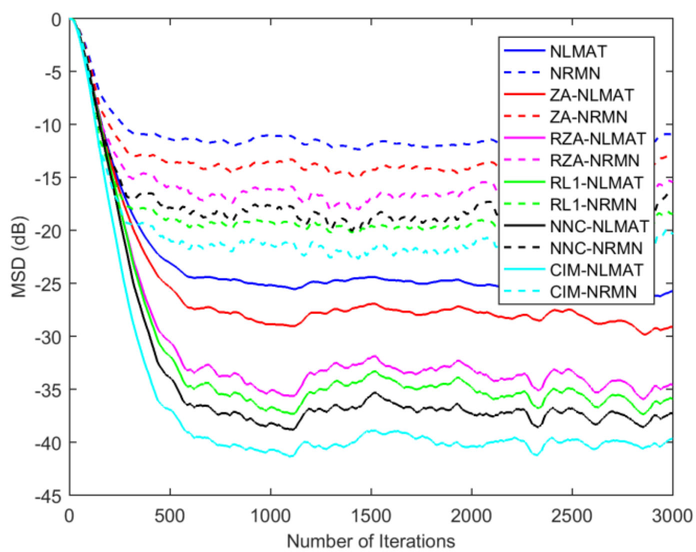


Fig. 5.8: MSD performance for the proposed algorithms with Case 2 background noise and the input is the correlated signal for the system with sparsity S=1. The simulation

parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}, \rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 4\times10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 1\times10^{-3}, \varepsilon_{NNC} = 20$, $\rho_{CIM} = 2\times10^{-3}, \sigma = 0.05$

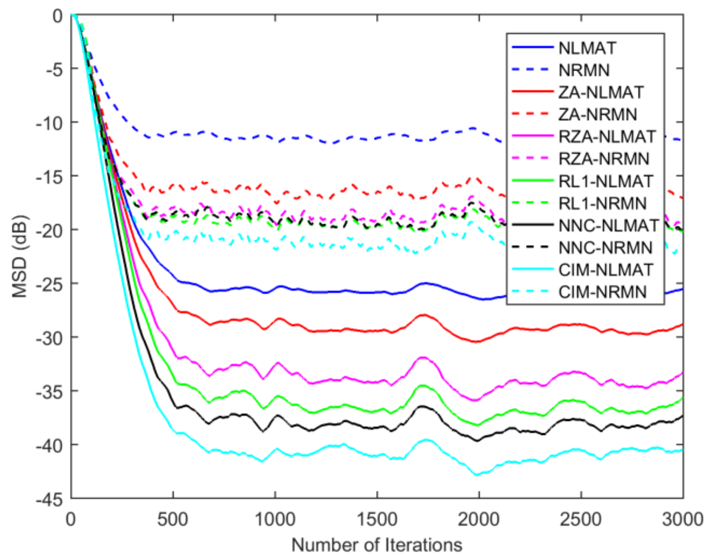


Fig. 5.9: MSD performance of the proposed algorithms with Case 3 background noise and the correlated input signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}, \rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 3\times10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 1\times10^{-3}, \varepsilon_{NNC} = 20$, $\rho_{CIM} = 2\times10^{-3}, \sigma = 0.05$

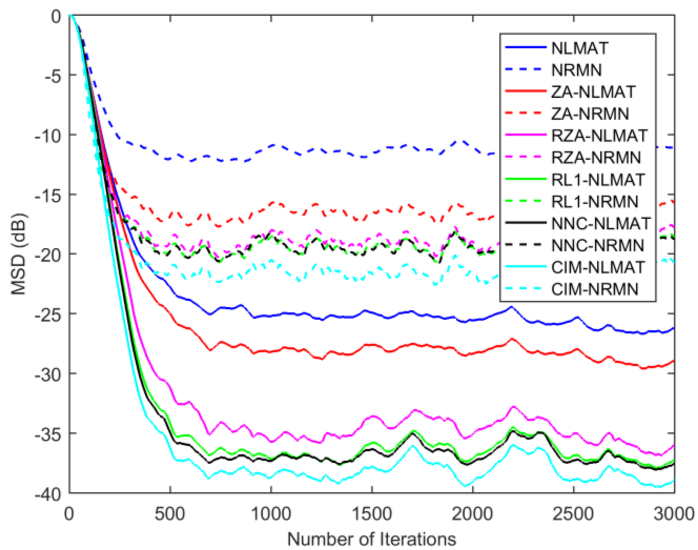


Fig. 5.10: MSD performance for the proposed algorithms with Case 4 background noise and the input is the correlated signal for the system with sparsity S=1. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8$, $\delta = 1\times10^{-3}$, $\rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 5\times10^{-4}$, $\varepsilon_{RZA} = 20$, $\rho_{RL1} = 3\times10^{-5}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 3\times10^{-3}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 1\times10^{-3}$, $\sigma = 0.05$

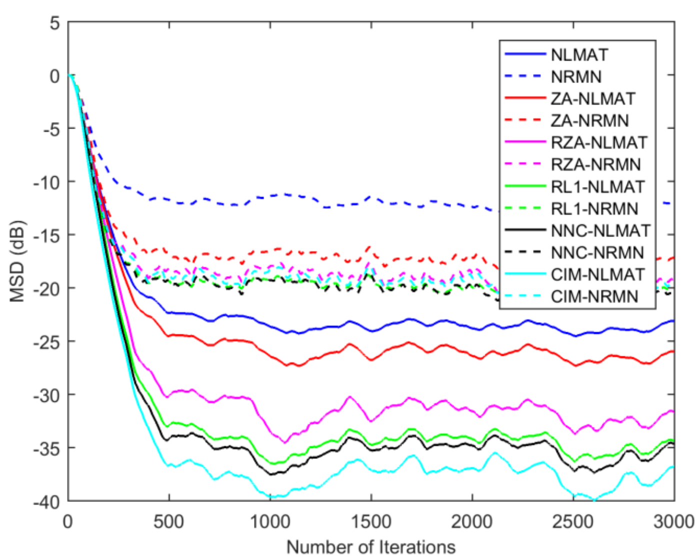


Fig. 5.11: MSD performance for the proposed algorithms with Case 5 background noise and the input is the correlated signal for the system with sparsity S=1. The simulation

parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 1\times10^{-4}, \rho_{RZA} = 5\times10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 3\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 2\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 2\times10^{-3}, \sigma = 0.05$

The EMSE values of the proposed algorithms obtained for different noise cases with correlated/colored input and system sparsity S=1 are given in Tables 5.3 respectively. It clearly shows that the proposed sparse algorithms outperform the NLMAT algorithm in identifying a sparse system.

Table 5.3: Comparison of EMSE values for different NLMAT algorithms with Correlated/Colored input signal and different noise cases with system sparsity S=1

EMSE in dB						
Correlated input signal with	**NLMAT**	**ZA-NLMAT**	**RZA-NLMAT**	**RL1-NLMAT**	**NNC-NLMAT**	**CIM-NLMAT**
Gaussian noise	-27.7545	-28.9331	-31.5727	-32.5171	-32.6784	-32.9729
Gaussian + impulsive noise	-23.7591	-25.1377	-27.0692	-27.5438	-28.5263	-30.3892
Uniformly + impulsive noise	-22.9686	-24.2891	-25.3021	-25.9263	-26.7201	-28.4806
Rayleigh + impulsive noise	-22.1036	-23.3135	-25.2280	-26.6034	-27.1472	-27.4192
Exponential + impulsive noise	-21.6826	-22.8867	-25.1323	-26.7080	-26.9945	-27.8838

In Figs. 5.12–5.21, the performance of the proposed algorithms when the system sparsity is changed to S=4 is shown.

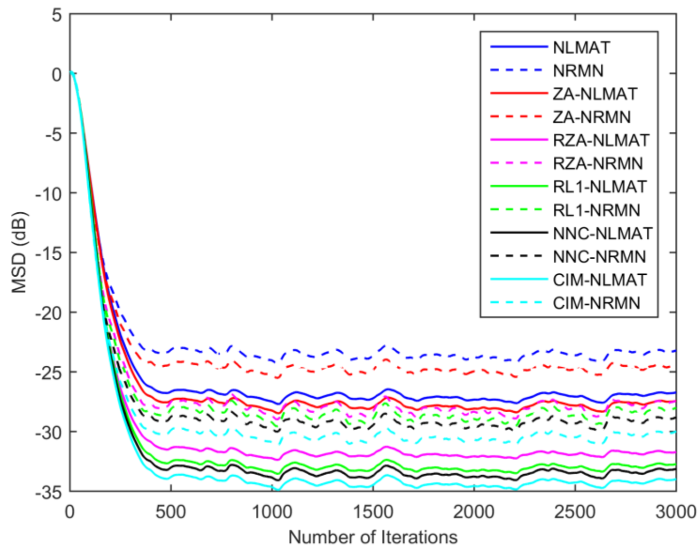


Fig. 5.12: MSD performance for the proposed algorithms with Case 1 as the background noise and the Gaussian white input signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 5\times10^{-3}, \rho_{RZA} = 5\times10^{-4}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-4}, \delta_{RL1} = 0.01, \rho_{NNC} = 5\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 5\times10^{-3}, \sigma = 0.05$

In Fig. 5.12, the simulation results for the proposed algorithms are shown for the white Gaussian input with Case 1 as the background noise for the system with sparsity S=4. The simulation results shown in Fig. 5.13 are carried out for the white Gaussian input with Case 2 background noise for system sparsity level S=4. It can be seen from Figs. 5.12 and 5.13 that the proposed sparse NLMAT algorithms exhibit better performance than NLMAT and NRMN algorithms in terms of MSD even after changing the system sparsity to S=4.

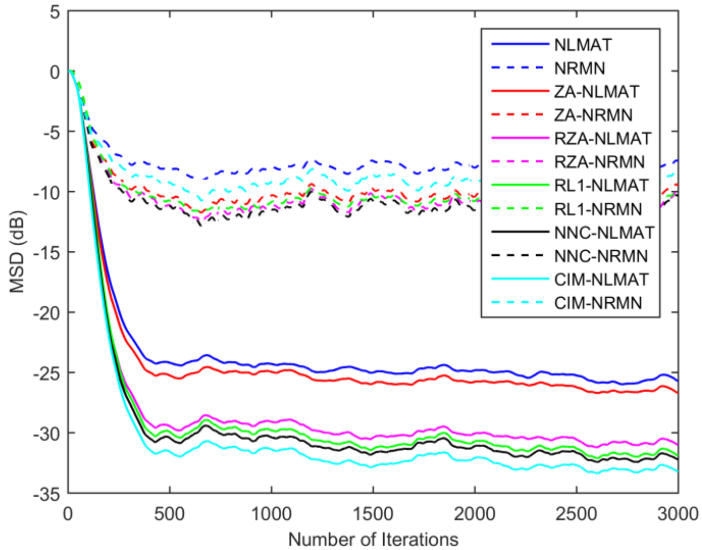


Fig. 5.13: MSD performance for the proposed algorithms with Case 2 as the background noise and the white Gaussian input signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 5\times10^{-4}$, $\rho_{RZA} = 5\times10^{-3}$, $\varepsilon_{RZA} = 20$, $\rho_{RL1} = 8\times10^{-5}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 5\times10^{-3}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 5\times10^{-3}$, $\sigma = 0.05$

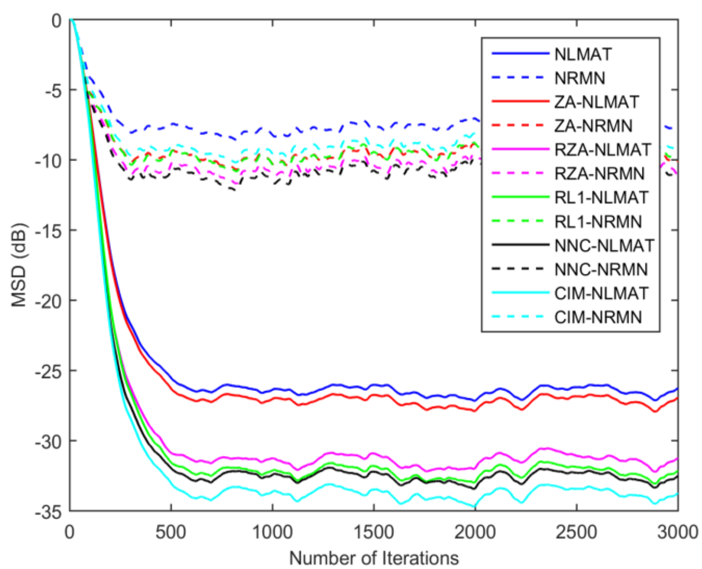


Fig. 5.14: MSD performance for the proposed algorithms with Case 3 background noise and the white Gaussian input signal for the system with sparsity S=4. The simulation

parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 5\times10^{-3}$, $\varepsilon_{RZA} = 20$, $\rho_{RL1} = 1\times10^{-4}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 8\times10^{-3}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 5\times10^{-3}$, $\sigma = 0.05$

In Fig. 5.14, the simulation results for the proposed algorithms are shown for the white Gaussian input while the background noise is Case 3 for the system with sparsity S=4. In Fig. 5.15, the input is white Gaussian with background noise as Case 4 for the system with sparsity S=4. In Fig. 5.16, the input is white Gaussian signal and the background noise is Case 5 for the system with sparsity S=4.

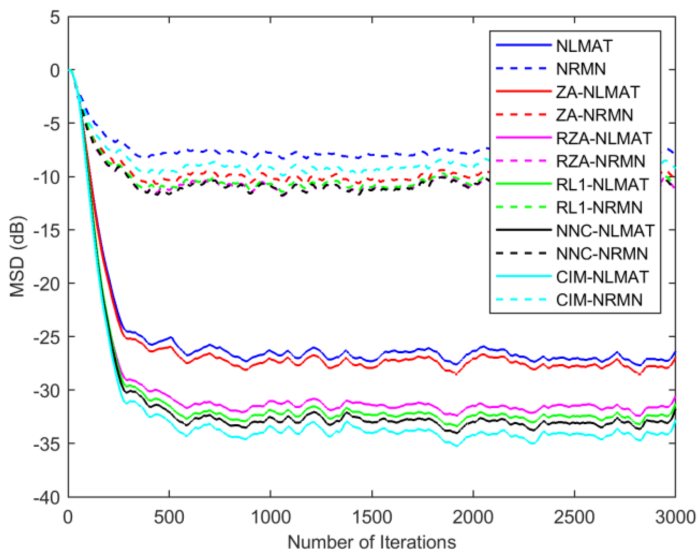


Fig. 5.15: MSD performance for the proposed algorithms with Case 4 as the background noise and the input is white Gaussian signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 1\times10^{-4}$, $\rho_{RZA} = 5\times10^{-3}$, $\varepsilon_{RZA} = 20$, $\rho_{RL1} = 1\times10^{-4}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 5\times10^{-3}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 5\times10^{-3}$, $\sigma = 0.05$

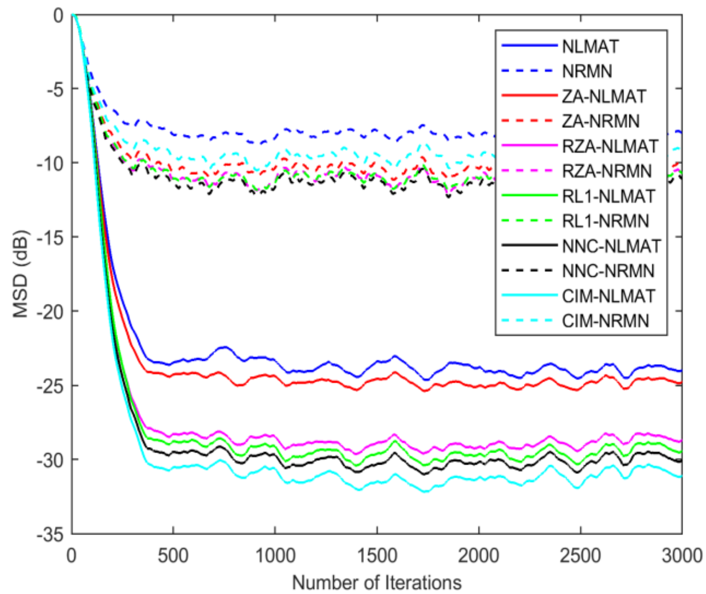


Fig. 5.16: MSD performance for the proposed algorithms with background noise as Case 5 and the white Gaussian input signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 5\times10^{-4}, \rho_{RZA} = 5\times10^{-3}, \varepsilon_{RZA} = 20, \rho_{RL1} = 8\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 6\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 5\times10^{-3}, \sigma = 0.05$

It can be easily seen from Figs. 5.14, 5.15 and 5.16 that the proposed sparse NLMAT algorithms provide better performance than NLMAT and NRMN algorithms in terms of MSD even after changing the system sparsity to S=4. As shown above, the proposed CIM-NLMAT algorithm achieves lower steady state error too.

The EMSE values of the proposed algorithms obtained under different noise cases with uncorrelated input and system sparsity S=4 are given in Table 5.4. It is confirmed that the proposed sparse algorithms outperform the NLMAT algorithm in identifying a sparse system.

Table 5.4: Comparison of EMSE values for different NLMAT algorithms with Uncorrelated input signal and different noise cases with system sparsity S=4

EMSE in dB						
Uncorrelated input signal with	**NLMAT**	**ZA-NLMAT**	**RZA-NLMAT**	**RL1-NLMAT**	**NNC-NLMAT**	**CIM-NLMAT**
Gaussian noise	-20.9392	-21.3551	-23.0535	-23.3768	-23.6006	-23.8147
Gaussian + impulsive noise	-17.9256	-18.4067	-20.1009	-20.2364	-20.4554	-20.7783
Uniformly + impulsive noise	-20.1611	-20.5581	-22.6783	-22.8724	-23.0897	-23.6132
Rayleigh + impulsive noise	-20.0824	-20.3172	-22.1620	-22.4268	-22.8624	-23.0831
Exponential + impulsive noise	-17.3953	-17.8863	- 19.5835	-19.7195	-19.9825	-20.3212

In the simulations shown in Figs. 5.17 – 5.21 the input signal is the correlated/colored input and the system background noise from Case 1 to Case 5 are considered accordingly with system sparsity changed to S=4. It is observed from Figs. 5.17–5.21 that the proposed sparse NLMAT algorithms exhibit better performance than NLMAT and NRMN algorithms in terms of MSD even after changing the system sparsity to S=4. Moreover, like for the previous simulations, the proposed CIM-NLMAT algorithm achieves the lowest steady state error.

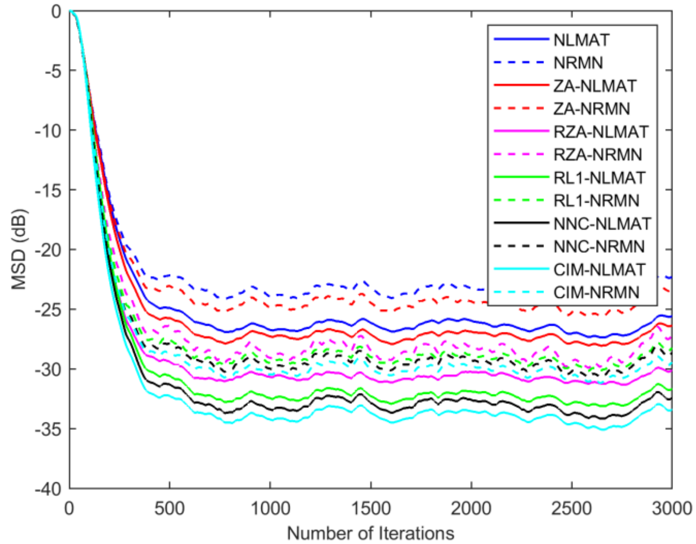


Fig. 5.17: MSD performance for the proposed algorithms with Case 1 as the background noise and the input is the correlated signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 1\times10^{-4}, \rho_{RZA} = 5\times10^{-3}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-4}, \delta_{RL1} = 0.01, \rho_{NNC} = 5\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 5\times10^{-3}, \sigma = 0.05$

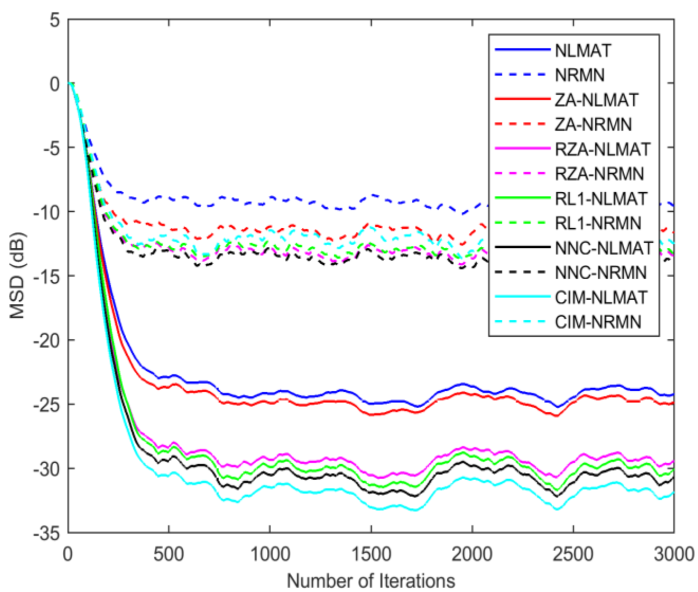


Fig. 5.18: MSD performance for the proposed algorithms with Case 2 as the background noise and the input is the correlated signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 4\times10^{-4}, \rho_{RZA} = 4\times10^{-3}, \varepsilon_{RZA} = 20, \rho_{RL1} = 8\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 5\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 5\times10^{-3}, \sigma = 0.05$

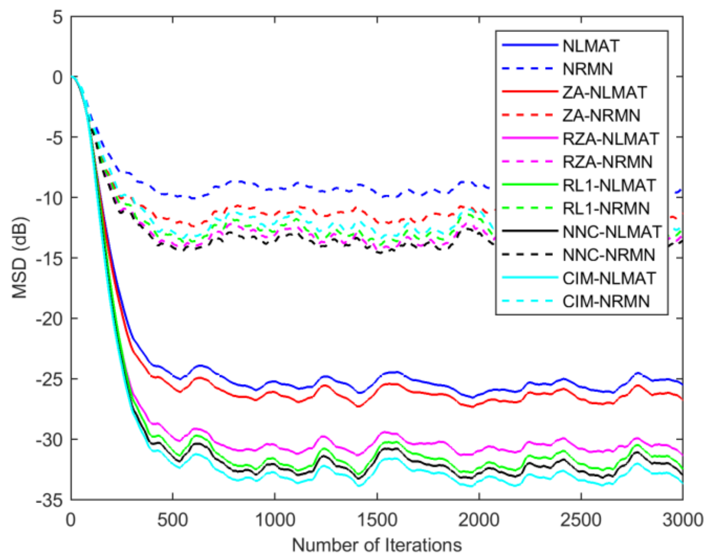


Fig. 5.19: MSD performance for the proposed algorithms with Case 3 background noise and the correlated input signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 3\times10^{-4}, \rho_{RZA} = 5\times10^{-3}, \varepsilon_{RZA} = 20, \rho_{RL1} = 7\times10^{-5}, \delta_{RL1} = 0.01, \rho_{NNC} = 5\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 8\times10^{-3}, \sigma = 0.05$

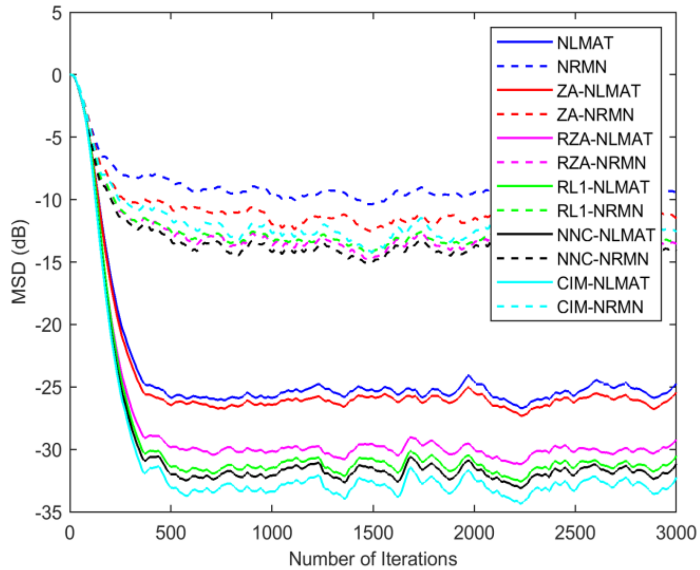


Fig. 5.20: MSD performance for the proposed algorithms with Case 4 background noise and the input is the correlated signal for the system with sparsity S=4. The simulation parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 2\times10^{-4}, \rho_{RZA} = 5\times10^{-3}, \varepsilon_{RZA} = 20, \rho_{RL1} = 1\times10^{-4}, \delta_{RL1} = 0.01, \rho_{NNC} = 8\times10^{-3}$, $\varepsilon_{NNC} = 20, \rho_{CIM} = 5\times10^{-3}, \sigma = 0.05$

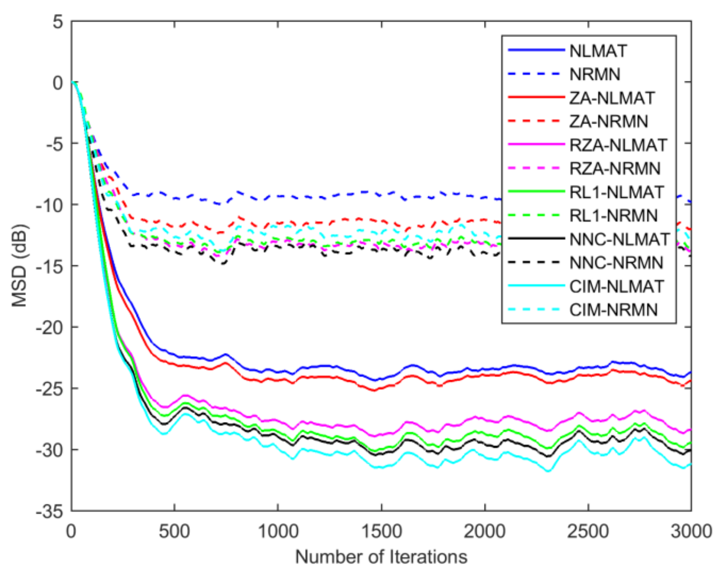


Fig. 5.21: MSD performance for the proposed algorithms with Case 5 background noise and the input is the correlated signal for the system with sparsity S=4. The simulation

parameters for sparse NLMAT algorithms are given as $\mu = 0.8, \delta = 1\times10^{-3}$, $\rho_{ZA} = 2\times10^{-4}$, $\rho_{RZA} = 5\times10^{-3}$, $\varepsilon_{RZA} = 20$, $\rho_{RL1} = 2\times10^{-4}$, $\delta_{RL1} = 0.01$, $\rho_{NNC} = 1\times10^{-2}$, $\varepsilon_{NNC} = 20$, $\rho_{CIM} = 5\times10^{-3}$, $\sigma = 0.05$

The EMSE values of the proposed algorithms obtained for different noise cases with correlated/colored input and system sparsity S=4 are given in Table 5.5. It can be shown that the proposed sparse algorithms outperform the NLMAT algorithm in identifying a sparse system.

Table 5.5: Comparison of EMSE values for different NLMAT algorithms with Correlated/Colored input signal and different noise cases with system sparsity S=4

	EMSE in dB					
Correlated input signal with	**NLMAT**	**ZA-NLMAT**	**RZA-NLMAT**	**RL1-NLMAT**	**NNC-NLMAT**	**CIM-NLMAT**
Gaussian noise	-19.1247	-19.5721	-21.0284	-21.3306	-21.4909	-21.7910
Gaussian + impulsive noise	-16.8195	-17.2241	-18.8023	-18.8600	-19.0803	-19.5058
Uniformly + impulsive noise	-19.1671	-19.6554	-21.2712	-21.5192	-21.7299	-21.7390
Rayleigh + impulsive noise	-17.3378	-17.6872	-19.3570	-19.5731	-19.7939	-20.3204
Exponential + impulsive noise	-15.3710	-15.7622	-17.2763	-17.4946	-17.7603	-18.2423

Let us now consider a network echo cancellation (NEC) system with the echo path impulse response of length L=512 as shown in Fig. 5.22. This is a sparse impulse response.

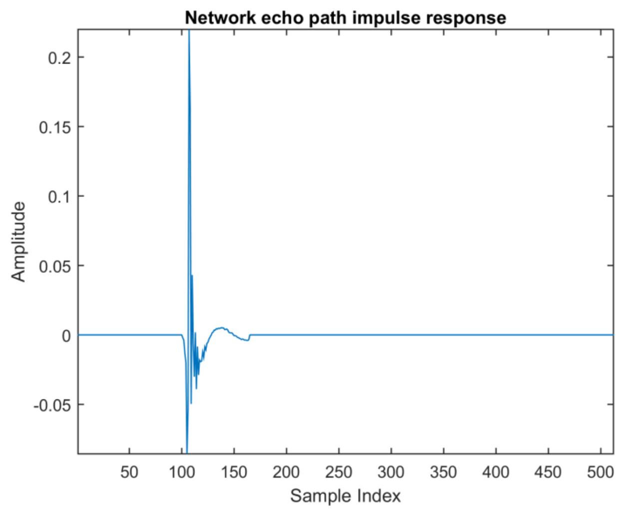


Fig. 5.22: Network echo path impulse response

In Fig. 5.23, the simulation results for the proposed algorithms are shown for the white Gaussian input and when the background noise consists of both white Gaussian noise with SNR of 20dB, and impulsive noise. From Fig. 5.23, it clearly depicts that the proposed sparse NLMAT algorithms exhibit better performance than the NLMAT algorithm for long echo paths with sparse impulse response.

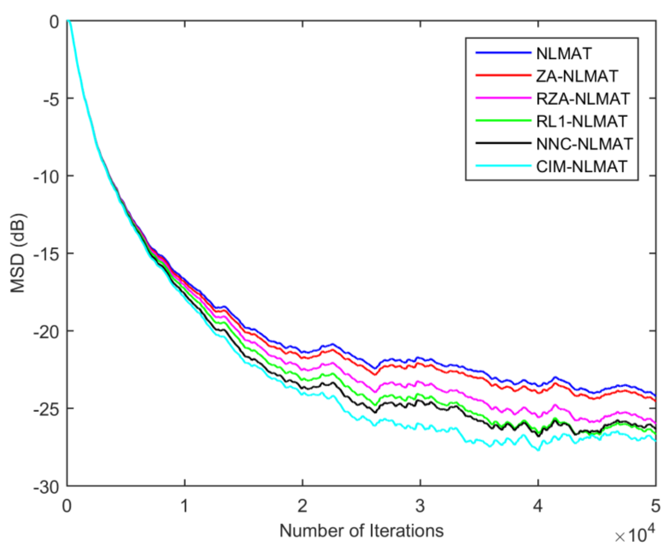


Fig. 5.23: MSD performance of the proposed algorithms in a NEC sparse system with white Gaussian noise and impulsive noise as the background noise and the input is white Gaussian signal.

In Fig. 5.24, the input signal is the correlated/colored input and the system noise comprises of both white Gaussian noise and impulsive noise. It is observed that the proposed sparse NLMAT algorithms exhibit better performance than NLMAT algorithm. Moreover, like for the previous simulations, the proposed CIM-NLMAT algorithm achieves the lowest steady state error for long echo paths with sparse impulse response.

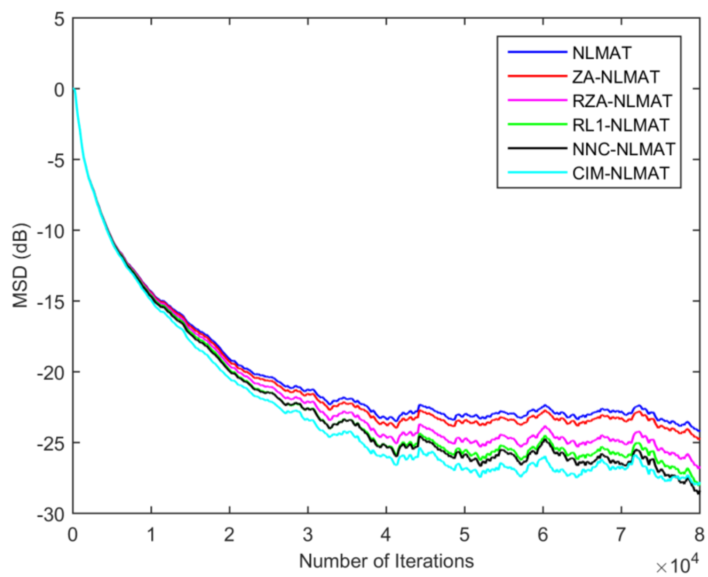


Fig. 5.24: MSD performance of the proposed algorithms in a NEC sparse system with white Gaussian noise and impulsive noise as the background noise and the input is the AR(1) correlated signal.

5.5 Least Mean Mixed Norm (LMMN) Algorithm

The output of the unknown system as described in Fig. 5.1 is given as

$$d(n) = \mathbf{x}^T(n)\mathbf{w}_{opt} + v(n), \tag{5.50}$$

where $\boldsymbol{x}(n) = [x(n), x(n-1) \ldots . x(n-L+1)]^T$ denotes an $Lx1$ input signal vector and $\boldsymbol{W}_{opt} = [w_0, w_1 \ldots . w_{L-1}]^T$ is an unknown weight vector. The measurement noise $v(n)$ consists of white Gaussian noise and impulsive noise. The system error signal is defined as

$$e(n) = d(n) - y(n) = d(n) - \boldsymbol{x}^T(n)\boldsymbol{W}(n-1), \tag{5.51}$$

where, $y(n)$ is the output of the adaptive filter $\boldsymbol{W}(n)$.

The cost function of the LMMN algorithm can be expressed as

$$J_{LMMN}(n) = \frac{\lambda}{2} E\{e^2(n)\} + \frac{1-\lambda}{4} E\{e^4(n)\}, \tag{5.52}$$

which is a combination of LMS and LMF algorithm cost functions and λ is the mixing parameter, $0 \le \lambda \le 1$.

The LMMN weight update equation is derived using the gradient minimization technique

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \mu \hat{\nabla}_{\boldsymbol{W}(n)} J_{LMMN}(n)$$

$$= \boldsymbol{W}(n) + \mu e(n)\{\lambda + (1-\lambda)e^2(n)\}\boldsymbol{x}(n), \tag{5.53}$$

where μ is the step-size of LMMN algorithm.

5.6 Proposed Modified LMMN Algorithm Based on Sigmoid Function (SLMMN)

We firstly define the sigmoid function as [223], [224]

$$S(n) = sgm[\alpha J_{LMMN}(n)] = \frac{1}{1 + e^{-\alpha J_{LMMN}(n)}}, \tag{5.54}$$

where α is the steepness parameter of the sigmoid function.

The modified cost function of the LMMN algorithm based on (5.54) can be expressed as

$$J_{SLMMN}(n) = \frac{1}{\alpha} S(n) = \frac{1}{\alpha} \frac{1}{1+e^{-\alpha J_{LMMN}(n)}}. \tag{5.55}$$

On differentiating the sigmoid LMMN (SLMMN) cost function (5.55) with respect to $\boldsymbol{W}(n)$ yields

$$\hat{\nabla}_{\boldsymbol{W}(n)} J_{SLMMN}(n) = \frac{\partial J_{SLMMN}(n)}{\partial \boldsymbol{W}(n)}$$

$$= \frac{1}{\alpha} \frac{\partial S(n)}{\partial J_{LMMN}(n)} \frac{\partial J_{LMMN}(n)}{\partial W(n)}$$

$$= S(n)[1-S(n)] \hat{\nabla}_{\boldsymbol{W}(n)} J_{LMMN}(n) \tag{5.56}$$

The weight update rule of the proposed SLMMN algorithm is given by

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \mu \hat{\nabla}_{\boldsymbol{W}(n)} J_{SLMMN}(n). \tag{5.57}$$

Substituting (5.56) into (5.57), we get

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \mu S(n)[1-S(n)] \hat{\nabla}_{\boldsymbol{W}(n)} J_{LMMN}(n)$$

$$= \boldsymbol{W}(n) + \mu S(n)[1-S(n)] e(n) \left\{ \lambda + (1-\lambda) e^2(n) \right\} \boldsymbol{x}(n), \tag{5.58}$$

where, $S(n) = sgm\left[\alpha \left\{ \frac{\lambda}{2}\left(e^2(n)\right) + \frac{1-\lambda}{4}\left(e^4(n)\right) \right\} \right]$

$$= \frac{1}{1+e^{-\alpha\left\{ \frac{\lambda}{2}\left(e^2(n)\right) + \frac{1-\lambda}{4}\left(e^4(n)\right) \right\}}}. \tag{5.59}$$

5.7 Proposed Sparse SLMMN Algorithms

To exploit the system sparsity, two sparse algorithms are proposed by introducing ℓ1-norm and log-sum penalties into the SLMMN namely, Zero Attracting SLMMN (ZA-SLMMN) and Reweighted Zero Attracting SLMMN (RZA-SLMMN) respectively.

5.7.1. Zero Attracting SLMMN (ZA-SLMMN) Algorithm

Let the cost function of ZA-SLMMN algorithm denoted by

$$J_{ZA-SLMMN}(n) = J_{SLMMN}(n) + \gamma_{ZA}\|\boldsymbol{W}(n)\|_1 \tag{5.60}$$

where γ_{ZA} is the regularization parameter which balances the estimation error and $\|\boldsymbol{W}(n)\|_1$

Using the gradient descent rule, the ZA-SLMMN algorithm update is defined as

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \mu\hat{\nabla}_{\boldsymbol{W}(n)} J_{ZA\text{-}SLMMN}(n) \tag{5.61}$$

where,

$$\hat{\nabla}_{\boldsymbol{W}(n)} J_{ZA\text{-}SLMMN}(n) = \frac{\partial J_{ZA\text{-}SLMMN}(n)}{\partial \boldsymbol{W}(n)}$$

$$= \frac{\partial J_{SLMMN}(n)}{\partial \boldsymbol{W}(n)} + \gamma_{ZA} sgn(\boldsymbol{W}(n)). \tag{5.62}$$

Using (5.56) in (5.62) and substituting into (5.61), we get

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \mu S(n)[1 - S(n)]e(n)\{\lambda + (1-\lambda)e^2(n)\}\boldsymbol{x}(n) - \rho_{ZA} sgn(\boldsymbol{W}(n)) \tag{5.63}$$

where, $sgn(.)$ is the signum function, $\rho_{ZA} = \mu\gamma_{ZA}$, and

$$S(n) = sgm\left[\alpha\left\{\frac{\lambda}{2}\left(e^2(n)\right) + \frac{1-\lambda}{4}\left(e^4(n)\right) + \gamma_{ZA}\|\boldsymbol{W}(n)\|_1\right\}\right]. \tag{5.64}$$

Equation (5.63) corresponds to the weight updating of the proposed ZA-SLMMN algorithm.

5.7.2. Reweighted Zero Attracting SLMMN (RZA-SLMMN) Algorithm

The cost function of RZA-SLMMN algorithm is obtained by introducing the log-sum penalty [20] into the SLMMN cost function as

$$J_{RZA-SLMMN}(n) = J_{SLMMN}(n) + \gamma_{RZA}\sum_{i=0}^{L-1} log\left(1 + \varepsilon_{RZA}|w_i(n)|\right) \tag{5.65}$$

where γ_{RZA} is the regularization parameter.

The weight update rule of RZA-SLMMN algorithm is derived as

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \mu \hat{\nabla}_{\boldsymbol{W}(n)} J_{RZA\text{-}SLMMN}(n). \tag{5.66}$$

On differentiating the second term in (5.66) with respect to $\boldsymbol{W}(n)$, yields the following equation that corresponds to the weight updating of the proposed RZA-SLMMN algorithm.

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \mu S(n)[1 - S(n)]e(n)\left\{\lambda + (1-\lambda)e^2(n)\right\}\boldsymbol{x}(n) - \rho_{RZA} \frac{sgn(\boldsymbol{W}(n))}{1 + \varepsilon_{RZA}\left|\boldsymbol{W}(n)\right|}, \tag{5.67}$$

where, $\rho_{RZA} = \mu \gamma_{RZA} \varepsilon_{RZA}$,

$$S(n) = sgm\left[\alpha\left\{\frac{\lambda}{2}\left(e^2(n)\right) + \frac{1-\lambda}{4}\left(e^4(n)\right) + \sum_{i=0}^{L-1} log\left(1 + \varepsilon_{RZA}\left|w_i(n)\right|\right)\right\}\right]. \tag{5.68}$$

5.7.3. Simulation Results

The performance of the proposed SLMMN algorithms in the system identification scenario is evaluated in this section. The length of the unknown system to be identified is set as L = 16 with system sparsity of K = {1, 4, 8} and the adaptive filter is also assumed to have the same length. The correlated (colored) input signal is generated by filtering a Gaussian white noise with variance $\sigma_x^2 = 1$ (0 dB) through a first-order autoregressive system, AR(1), with a pole at 0.8. The system noise $v(n)$ contains white Gaussian noise with SNR = 20dB and Bernoulli-Gaussian (B-G) distributed impulsive noise. B-G noise is generated as $q(n) = b(n)v_i(n)$. $b(n)$ is a binary process, illustrated by the probability $p(b(n)=1) = P$, $p(b(n)=0) = 1 - P$, where P represents the probability of occurrence of the impulsive noise. $v_i(n)$ is assumed to be a zero-mean white Gaussian noise with variance $\sigma_{v_i}^2$. The normalized mean-square deviation (NMSD) defined as

$NMSD(n) = 10\log_{10} \frac{\left\| \boldsymbol{W}_{opt} - \boldsymbol{W}(n) \right\|_2^2}{\left\| \boldsymbol{W}_{opt} \right\|_2^2} (dB)$ is used to estimate the performance of the proposed algorithms. The average of 100 trials is used in evaluating the results. The simulation parameters setting are as follows: $\mu = 0.04$, $P = 0.01$, $\sigma_{v_i}^2 = 10^4/12$, $\lambda = 0.5$, $\alpha = 0.6$, $\rho_{ZA} = 5x10^{-5}$, $\rho_{RZA} = 1x10^{-4}$, and $\varepsilon_{RZA} = 20$.

From Figs. 5.25, 5.26 and 5.27, it can be seen that the SLMMN, ZA-SLMMN and RZA-SLMMN yield better steady state performances than the sparse LMP algorithms (ZA-LMP and RZA-LMP), while the LMMN algorithm does not converge in the presence of impulsive noise. Hence, the proposed algorithms are robust against impulsive noise and are capable of handling the system with different sparsity levels, K={1, 4, 8}. The RZA-SLMMN algorithm exhibits superior performance and achieves the lowest steady-state error in all the cases.

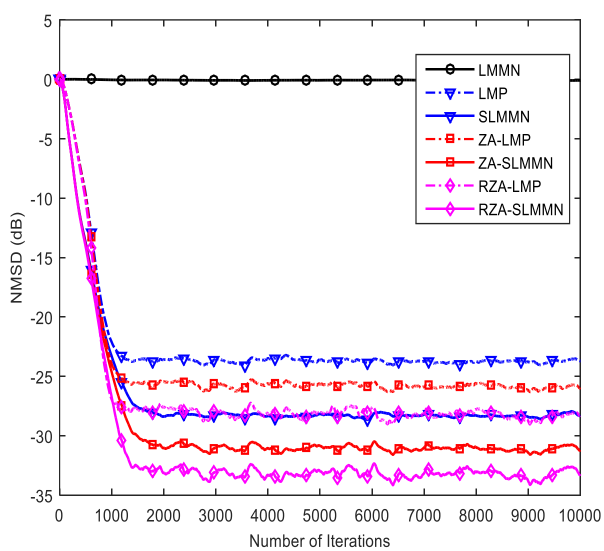


Fig. 5.25: NMSD comparison of the proposed SLMMN algorithms for the system with sparsity K =1 and in the presence of impulsive noise

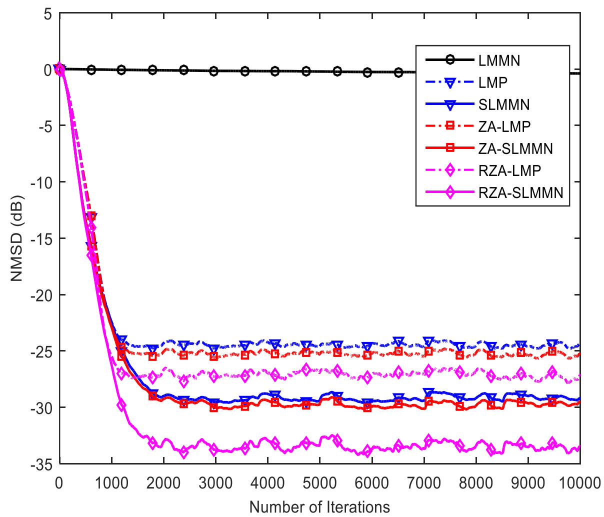


Fig. 5.26: NMSD comparison of the proposed SLMMN algorithms for the system with sparsity K =4 and in the presence of impulsive noise

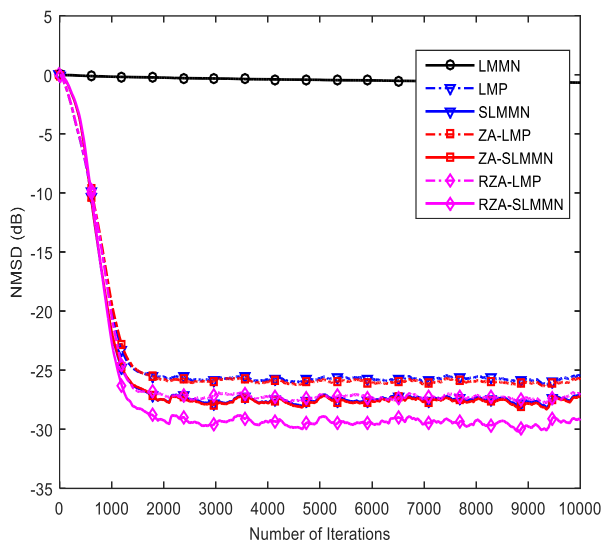


Fig. 5.27: NMSD comparison of the proposed SLMMN algorithms for the system with sparsity K =8 and in the presence of impulsive noise

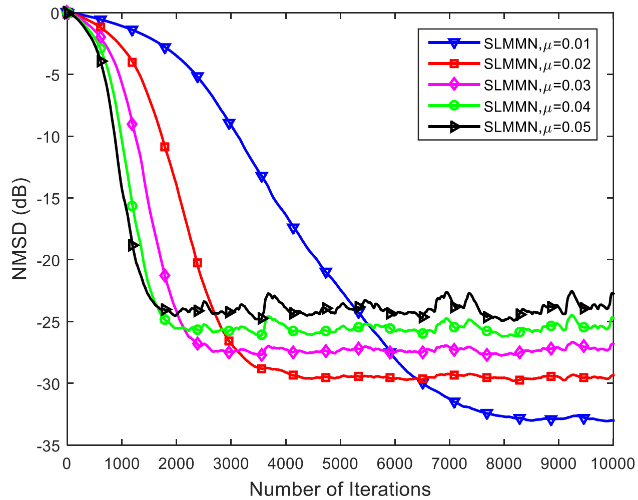


Fig. 5.28: NMSD of the proposed SLMMN algorithm with different step-size parameter μ

It can be noticed from Fig. 5.28 that increasing the step-size value leads to an increased convergence rate of the proposed SLMMN algorithm, but also results in high steady-state error.

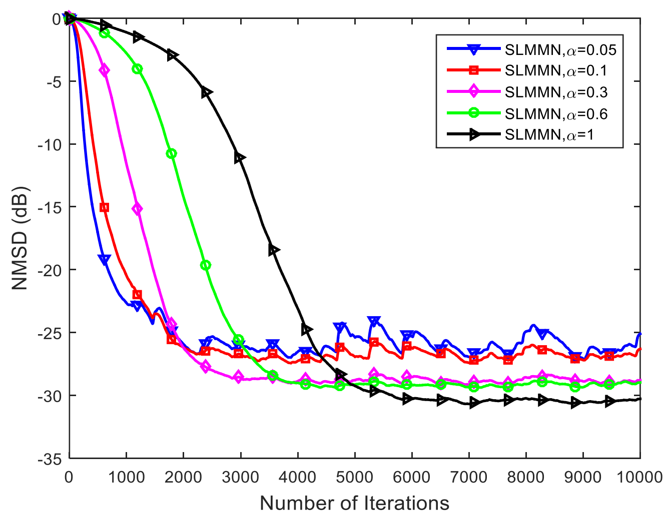


Fig. 5.29: NMSD of the proposed SLMMN algorithm with different α

As can be seen from Fig. 5.29, the greater the constant α, the slower will be the convergence, but the lower will be the steady-state misadjustment. Depending on the particular practical application the proper choice of the parameters of the SLMMN algorithm can highly reduce the number of needed iterations if a certain NMSD performance is needed.

5.8 Summary

The Normalized LMAT algorithm based on high-order error power (HOEP) criterion achieves improved performance and mitigates the noise interference effectively, but it does not promote sparsity. Hence, in this chapter, we have proposed different sparse Normalized LMAT algorithms in the sparse system identification context. From the simulation results, it is verified that our proposed sparse algorithms are capable of exploiting the system sparsity as well as providing robustness to impulsive noise. Moreover, the proposed CIM-NLMAT algorithm exhibit superior performance in the presence of different types of noise.

The LMMN algorithm which is reported to overcome the sensitivity and to improve the misadjustment performance, fails to converge in the presence of non-Gaussian impulsive intereferences. So, a modified LMMN algorithm using the sigmoid function, i.e., SLMMN algorithm is proposed which is robust against impulsive noise. By incorporating different sparsity penalties into the SLMMN filter, ZA-SLMMN and RZA-SLMMN algorithms are derived. The proposed sparse algorithms are capable of estimating effectively the system with different sparsity levels and achieves lowest steady-state misadjustment compared to the SLMMN algorithm.

Chapter 6

Sparse Constrained Algorithms for Complex Domain Adaptive System Identification

CHAPTER 6

6.1 Motivation

Sparse adaptive filters are used extensively for enhancing the filter performance in a sparse system. The affine projection algorithm (APA) is effective in improving the convergence speed for strongly correlated input signals, but it is very sensitive to impulsive noise. Normalized Correlation Algorithm (NCA) is robust in impulsive noise environments. The affine projection normalized correlation algorithm (AP-NCA) used in complex-domain adaptive filters, combines the benefits of APA and NCA and it does not take into account the underlying sparsity information of the system. In this chapter, we develop sparse AP-NCA algorithms to exploit system sparsity as well as to mitigate impulsive noise with correlated complex-valued input. Simulation results show that the proposed algorithms exhibit better performance than the AP-NCA for a sparse system.

6.2 Introduction

The complex valued signals are of fundamental interest and arise frequently in applications such as communications, optics, and acoustics. A complex-valued random variable is considered circular if it has a rotation invariant distribution, and is otherwise known as noncircular [234] [235].

A vector is called circular if its probability distribution is rotationally invariant, i.e., z and $z' = e^{j\alpha}z$ have the same probability distribution for any given real α.

The covariance matrix of z' is $C_{z'z'} = E\{z'z'^H\} = E\{e^{j\alpha}zz^H e^{-j\alpha}\} = C_{zz}$ (6.1)

On the other hand,

The pseudo covariance matrix of z' is $\tilde{C}_{z'z'} = E\{z'z'^T\} = E\{e^{j\alpha}zz^T e^{j\alpha}\} = e^{j2\alpha}\tilde{C}_{zz}$ (6.2)

Equation (6.2) is true for arbitrary α if and only if $\tilde{C}_{zz} = 0$. z is called proper, where the complementary covariance matrix vanishes, $\tilde{C}_{zz} = 0$, otherwise improper. Because the

Gaussian distribution is completely determined by second-order statistics, a complex Gaussian random vector z is circular if and only if it is zero-mean and proper.

The degree of noncircularity can be quantified by the circularity measure r, defined as the magnitude of the circularity quotient $\rho(z) = re^{j\theta} = {\tau_z}^2 / {\sigma_z}^2$, where

$$r = |\rho(z)| = \frac{|\tau_z|^2}{{\sigma_z}^2}, r \in [0,1] \quad (6.3)$$

measures the degree of noncircularity in the complex signal, with the circularity angle $\theta = \arg(\rho(z))$ indicating orientation of the distribution. Note that, $r = 0$ corresponds to a purely circular signal, with θ not providing additional information about the distribution, while $r = 1$, corresponding to a highly noncircular signal [236].

Adaptive filtering algorithms have received much attention over the past decades and are widely used for diverse applications such as system identification, interference cancellation, and channel estimation. In recent years, sparse adaptive filters have been developed to exploit the system sparse information and the performance can be greatly improved when compared with the conventional algorithms such as Least Mean Square (LMS) and Affine Projection Algorithm (APA) [2], [6]. Based on the assumption of the Gaussian noise model, sparse algorithms are derived by applying the $\ell 1$-norm relaxation into the LMS cost function [85], [89], sparsity-aware ℓp-norm penalized and reweighted $\ell 1$-norm penalized LMS algorithms are derived in [96], [215], and sparsity-aware affine projection adaptive algorithms for system identification are proposed in [90], [95]. However, these methods may be unreliable in estimating the systems under non-Gaussian impulsive noise environments. For example, the least mean square (LMS) [85] algorithm performance is affected by strong impulsive noise [153]. Several sign algorithms (SA) have been proposed in [225], [226], [227] to suppress impulsive noise under the assumption of the dense impulse response. In [228], the standard sign least mean square (SLMS) algorithm was proposed in order to achieve the robustness against impulsive noise. For adaptive filters defined in the complex-domain, the Normalized correlation algorithm (NCA) was proposed [155] for robust filtering in severe impulsive noise environments. In [178], [180], [229], considering the sparse information in a wireless channel, several sparse SLMS algorithms were proposed to exploit system sparsity and to mitigate non-Gaussian impulsive noise. In [230], [231] a flexible zero attractor constraint is utilized in sparse channel estimation under the mixed Gaussian noise environment. However, when the input signal is strongly correlated the performance of sparse SLMS algorithms deteriorates.

When the input to the adaptive filter is assumed to be colored (correlated) input, the standard LMS filter may converge slowly. To improve the filter performance for colored signals, the Affine Projection Algorithm has been proposed [11]. For a large projection order, the APA algorithm has faster convergence, but the steady-state error is higher resulting in a convergence vs steady-state error tradeoff.

In order to utilize the benefits of APA and NCA, the Affine Projection Normalized Correlation Algorithm (AP-NCA) was proposed [156]. The AP-NCA achieves faster convergence for a correlated input and is also robust against impulsive noises. To fully take advantage of the sparse structure present in the system, in this chapter, we propose sparse AP-NCA algorithms with different sparse norm constraint functions.

6.3 Impulse Noise Models

The stochastic models used to generate impulse noise are presented in this section. We observe that the two types of impulse noise entering adaptive filtering systems can be the observation noise and another at the input of adaptive filter.

A) Gaussian mixture model (GMM)

A model often used for impulsive observation noise is the Gaussian mixture model (GMM) [232]. GMM is a combination of two independent noise sources $v^{(1)}(n)$ and $v^{(2)}(n)$. The noise source $v^{(1)}(n)$ has a variance $\sigma^2_{v_1}$ with probability of occurrence $(1-\varphi)$, and the noise source $v^{(2)}(n)$ has $\sigma^2_{v_2}$ with the probability of occurrence φ. Usually, $\sigma^2_{v_2} >> \sigma^2_{v_1}$. The GMM distribution is given as

$$p(v(n)) = (1-\varphi)N(0,\sigma_{v_1}^2) + \varphi N(0,\sigma_{v_2}^2) \tag{6.4}$$

The variance of GMM is given by $\sigma_v^2 = E\{v^2(n)\} = (1-\varphi)\sigma_{v_1}^2 + \varphi\sigma_{v_2}^2$. Note that $v(n)$ will reduce to Gaussian noise model if $\varphi = 0$.

B) Bernoulli-Gaussian Model (B-G)

When impulse noise enters the reference input $\bar{x}(n)$, the filter input $b(n)$ is written as $b(n) = x(n) + q(n)$. $q(n)$ is the impulse noise modeled by a Bernoulli-Gaussian (BG) process [233], given as $q(n) = \alpha(n)v_a(n)$, with $v_a(n)$ assumed to be a White Gaussian process, and its variance is $\sigma_{v_a}^2$. $\alpha(n)$ is a binary process, described by the probability $p(\alpha(n)=1) = P,\ p(\alpha(n)=0) = 1-P$, where P represents the probability of occurrence of the impulsive noise, $v_a(n)$.

6.4 Affine Projection Normalized Correlation Algorithm (APNCA)

The system identification problem is shown in Fig. 6.1. Let $\bar{x}(n) = [x(n), x(n-1), ...x(n-L+1)]^T \in C^{Lx1}$ be the filter input vector of length L. $\bar{x}(n)$ is the complex-valued regressor process. The output signal from an unknown system with tap coefficient vector h is given by $u(n) = \boldsymbol{h}^T\bar{x}(n)$. $\bar{W}(n) = [w(n), w(n-1), ...w(n-L+1)]^T$ is an estimate of h at iteration n and L is the length of the adaptive filter. The update equation for the APA is given by

$$\bar{W}(n+1) = \bar{W}(n) + \mu\bar{X}(n)\left[\bar{X}^H(n)\bar{X}(n) + \varepsilon I_M\right]^{-1}\bar{e}^*(n), \tag{6.5}$$

where $\bar{X}(n) = [\bar{x}(n), \bar{x}(n-1), ...\bar{x}(n-M+1)] \in C^{LxM}$, is the input signal matrix and M is the projection order. μ is the step-size of APA filter, ε is the regularization term, I_M is the MxM identity matrix, $(.)^H$ is the conjugate transpose, $(.)^*$ is the complex conjugate, $(.)^T$ is the transpose of a matrix or a vector.

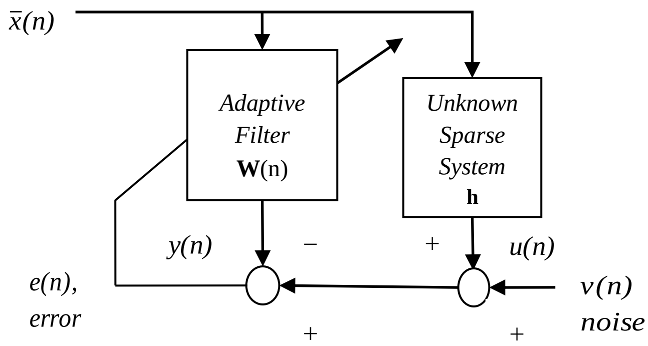


Fig. 6.1: Block diagram of adaptive sparse system identification

$$\bar{e}^{*}(n) = \bar{X}^{H}(n)\boldsymbol{\theta}(n) + \mathbf{v}^{*}(n), \tag{6.5}$$

where $\boldsymbol{\theta}(n) = \boldsymbol{h} - \bar{W}(n),$ is the misalignment vector and,

$\mathbf{v}(n) = [v(n), v(n-1).....v(n-M+1)]^{T} \in C^{M\text{x}1}$ is the noise vector. If M=1, the APA algorithm simplifies to NLMS algorithm.

The update equation for Normalized Correlation Algorithm (NCA) is given by

$$\bar{W}(n+1) = \bar{W}(n) + \mu\bar{z}(n) / \|\bar{z}(n)\|_{1}, \tag{6.7}$$

where, $\bar{z}(n) = \; z_{0}(n), z_{1}(n),....z_{L-1}(n)^{\;T} = e(n)\bar{x}(n)$ is the correlation vector and $\|\bar{z}(n)\|_{1}$ is the Euclidean norm of the correlation vector. Since $\|\bar{z}(n)\|_{1} = |e(n)|.\|\bar{x}(n)\|,$ the update equation can be rewritten as

$$\bar{W}(n+1) = \bar{W}(n) + \mu\varphi_{e}^{*}(n)\bar{x}(n) / \|\bar{x}(n)\|_{l}, \tag{6.8}$$

with $\varphi_{e}(n) = e(n)/|e(n)|$.

The Affine Projection Normalized Correlation Algorithm (AP-NCA) is updated as follows

$$\bar{W}(n+1) = \bar{W}(n) + \mu\bar{X}(n)\left[\bar{X}^{H}(n)\bar{X}(n) + \varepsilon I_{M}\right]^{-1/2}\bar{\varphi}_{e}^{\,*}(n) \tag{6.9}$$

where, $\bar{\varphi}_e(n) = \left[\varphi_e(n), \varphi_e(n-1), ...\varphi_e(n-M+1)\right]^T \in C^{M\text{x}1}$. If M=1, the AP-NCA algorithm behaves as NCA algorithm.

6.5 Proposed Sparse APNCA Algorithms

To exploit the system sparsity and robustness against impulsive noises, four sparse algorithms are proposed by introducing effective sparsity constraints into the standard AP-NCA namely, Zero Attracting AP-NCA (ZA-APNCA), Reweighted Zero Attracting AP-NCA (RZA-APNCA), Reweighted L1-norm AP-NCA (RL1-APNCA) and Flexible Zero Attracting AP-NCA (FZA-APNCA).

6.5.1. The Zero Attracting AP-NCA (ZA-APNCA) Algorithm

Let the cost function of ZA-APNCA algorithm denoted by

$$J_{ZA}\left(\bar{W}(n)\right) = J\left(\bar{W}(n)\right) + \lambda_{ZA}\left\|\bar{W}(n)\right\|_1 \tag{6.10}$$

Where $J\left(\bar{W}(n)\right)$ is the cost function related to AP-NCA algorithm without sparsity constraint and λ_{ZA} is the regularization parameter which balances the estimation error and $\left\|\bar{W}(n)\right\|_1$.

The weight update equation of ZA-APNCA algorithm is derived as

$$\bar{W}(n+1) = \bar{W}(n) + \mu\bar{X}(n)\left[\bar{X}^H(n)\bar{X}(n) + \varepsilon I_M\right]^{-1/2}\bar{\varphi}_e^*(n) - \rho_{ZA}sgn(\bar{W}(n)) \tag{6.11}$$

where $\rho_{ZA} = \mu\lambda_{ZA}$ and $\text{sgn}(.)$ denotes the well-known sign function.

6.5.2 The Reweighted Zero Attracting AP-NCA (RZA-APNCA) Algorithm

Let the cost function of RZA-APNCA algorithm be

$$J_{RZA}\left(\bar{W}(n)\right) = J\left(\bar{W}(n)\right) + \lambda_{RZA}\sum_{i=0}^{L-1}\log(1+\varepsilon_{RZA}\left|w_i(n)\right|), \tag{6.12}$$

where λ_{RZA} is the regularization parameter which balances the estimation error and $\sum_{i=0}^{L-1}\log(1+\varepsilon_{RZA}|w_i(n)|)$.

The weight update equation of RZA-APNCA algorithm is derived as

$$\overline{W}(n+1)=\overline{W}(n)+\mu\overline{X}(n)\left[\overline{X}^H(n)\overline{X}(n)+\varepsilon I_M\right]^{-1/2}\overline{\varphi}_e^*(n)-\frac{\rho_{RZA}sgn(\overline{W}(n))}{1+\varepsilon_{RZA}|\overline{W}(n)|} \tag{6.13}$$

where $\rho_{RZA}=\mu\lambda_{RZA}\varepsilon_{RZA}$

6.5.3 The Reweighted ℓ1-norm AP-NCA (RL1-APNCA) Algorithm

Let the cost function of RL1-APNCA algorithm be

$$J_{RL1}\left(\overline{W}(n)\right)=J\left(\overline{W}(n)\right)+\lambda_{RL1}\left\|\overline{f}(n)\overline{W}(n)\right\|_1 \tag{6.14}$$

where λ_{RL1} is the weight associated with the penalty term and

$$\left[\overline{f}(n)\right]_i=\frac{1}{\delta_{RL1}+\left|[\overline{W}(n-1)]_i\right|},\ i=0,1,.....L-1 \tag{6.15}$$

$\delta_{RL1}>0$ and hence $\left[\bar{f}(n)\right]_i>0$ for $i=0,1,.....L-1$.

The weight update equation of RL1-APNCA algorithm is derived as

$$\overline{W}(n+1)=\overline{W}(n)+\mu\overline{X}(n)\left[\overline{X}^H(n)\overline{X}(n)+\varepsilon I_M\right]^{-1/2}\overline{\varphi}_e^*(n)-\frac{\rho_{RL1}sgn(\overline{W}(n))}{\delta_{RL1}+|\overline{W}(n-1)|} \tag{6.16}$$

where $\rho_{RL1}=\mu\lambda_{RL1}$.

6.5.4 The Flexible Zero Attracting AP-NCA (FZA-APNCA) Algorithm

The flexible zero attractor is realized using the approximation parameter adjustment function defined as

$$S_\alpha(\bar{W}(n)) = (1+\alpha^{-1})\left(1-e^{-\alpha|\bar{W}(n)|}\right), \tag{6.17}$$

where α is a small positive constant.

The modified cost function obtained by incorporating $S_\alpha(\bar{W}(n))$ function into the AP-NCA cost function is the following

$$J_{FZA}\left(\bar{W}(n)\right) = J\left(\bar{W}(n)\right) + \rho_{FZA} S_\alpha(\bar{W}(n)). \tag{6.18}$$

The weight update equation of FZA-APNCA algorithm is derived as

$$\bar{W}(n+1) = \bar{W}(n) + \mu\bar{X}(n)\left[\bar{X}^H(n)\bar{X}(n) + \varepsilon I_M\right]^{-1/2} \bar{\varphi}_e^*(n) - \rho_{FZA} S'_\alpha(\bar{W}(n)) \tag{6.19}$$

where,

$$S'_\alpha(\bar{W}(n+1)) = (\alpha+1)e^{(-\alpha|\bar{W}(n+1)|)}\operatorname{sgn}(\bar{W}(n+1)) \tag{6.20}$$

6.6 Simulation Results

In this section, we evaluate the performance of the proposed sparse adaptive algorithms in the context of system identification. The length of the unknown system is set as L = 16 with system sparsity of K = {1, 4, 8} and the adaptive filter is also assumed to have the same length. The correlated (colored) input signal is generated by using a Gaussian white noise with variance $\sigma_x^2 = 1$ (0 dB) through a first-order autoregressive process, AR(1), with a 0.5 pole. The system noise $v(n)$ contains white Gaussian noise with SNR = 20 dB and impulse noise. The algorithms are compared based on the performance of the Mean Square Error (MSE) between the actual and estimated CIR. The average of 100 trials is used in evaluating the results.

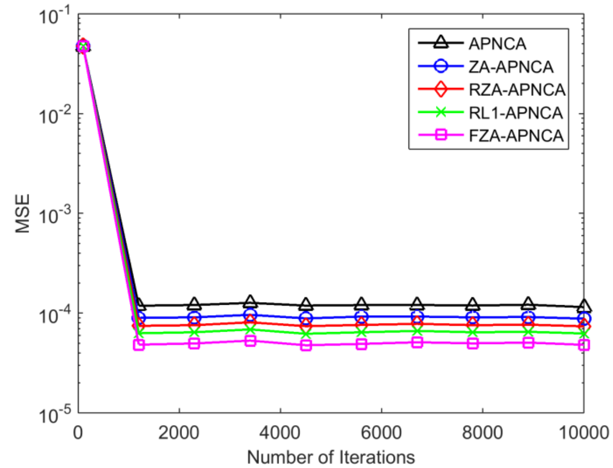
Detailed parameters for computer simulation are listed in Table 6.1.

Table 6.1: Simulation Parameters

Parameters	**Values**
Input Signal	Correlated/Colored Input: AR(1) Gaussian process with pole 0.5; $x(n)$=0.5$x(n-1)$ +$u(n)$
Unknown System Length	L=16
No. of nonzero coefficients	System sparsity, K={1, 4, 8}
Distribution of nonzero coefficients	Random Gaussian distribution $N(0,1)$
Projection order	M=4
SNR	20 dB
Noise types	Case 1: "white" Gaussian noise, $\sigma^2{}_v$ = 0.01 (–20 dB) Case 2: Observation noise: Gaussian Mixture Model (GMM) φ = 0.1, $\sigma^2_{v_1}$ = 0.01 (-20 dB), $\sigma^2_{v_2}$ = 10 (10 dB). Case 3: Impulse noise at filter input: Bernoulli-Gaussian (B-G) model $p_{v_a} = 0.1$, $\sigma^2_{v_a}$ =1000 (30 dB) Case 4: GMM & impulse noise at filter input

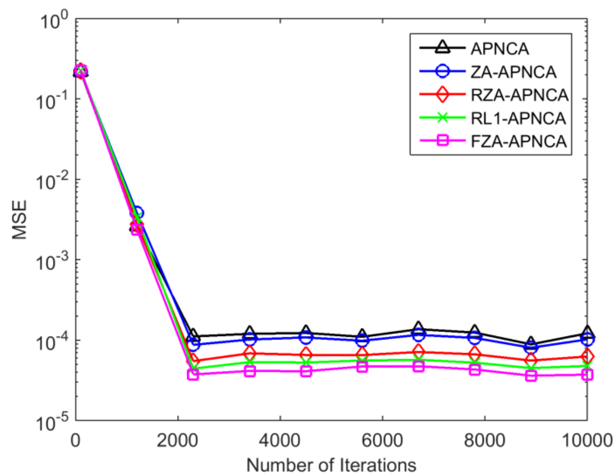
Comparison of the proposed sparse AP-NCA algorithms under noise case 1

The performance of the proposed sparse algorithms under the assumption of "white" Gaussian noise is shown in Fig. 6.2. It can be noticed that the proposed sparse APNCA algorithms exhibit better performance in terms of MSE when the system is highly sparse and it reduces as the system sparsity increases. The FZA-APNCA algorithm achieves minimum steady state error value.



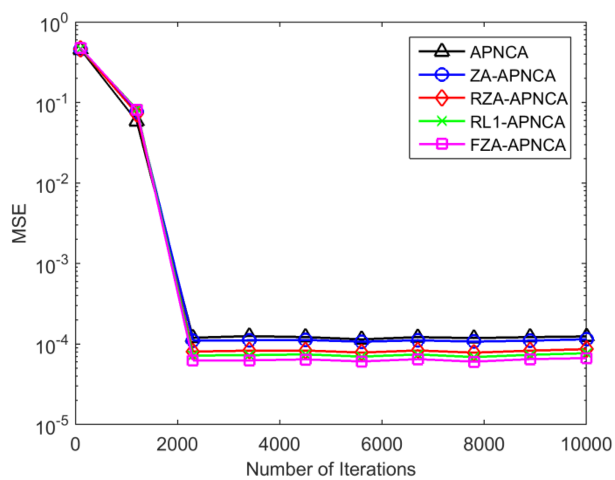
APNCA
ZA-APNCA
RZA-APNCA
RL1-APNCA
FZA-APNCA
MSE
Number of Iterations

(a)



APNCA
ZA-APNCA
RZA-APNCA
RL1-APNCA
FZA-APNCA
MSE
Number of Iterations

(b)



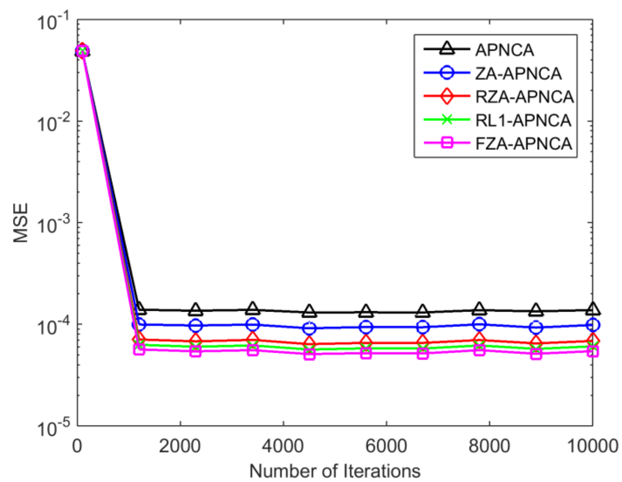
APNCA
ZA-APNCA
RZA-APNCA
RL1-APNCA
FZA-APNCA
MSE
Number of Iterations

(c)

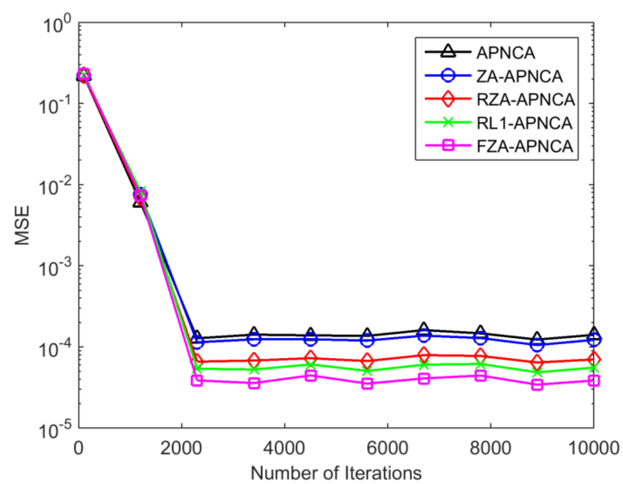
Fig. 6.2: MSEs of the proposed sparse AP-NCA algorithms for noise case 1 ("white" Gaussian) with the projection order, M=4 and different system sparsity of, (a) K=1, (b) K=4 and (c) K=8.

Comparison of the proposed sparse AP-NCA algorithms under noise case 2

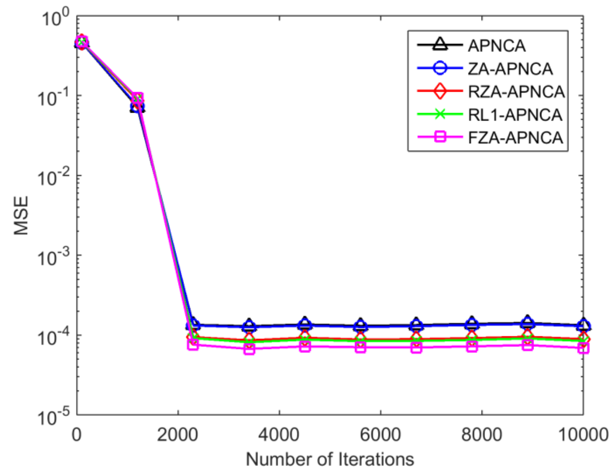
The performance of the proposed sparse algorithms under the assumption of GMM modeled impulsive observation noise is shown in Fig. 6.3. It can be noticed that the proposed sparse APNCA algorithms exhibit better performance in terms of MSE when the system is highly sparse and it reduces as the system sparsity increases. The FZA-APNCA achieves minimum steady state error value.



(a)



(b)

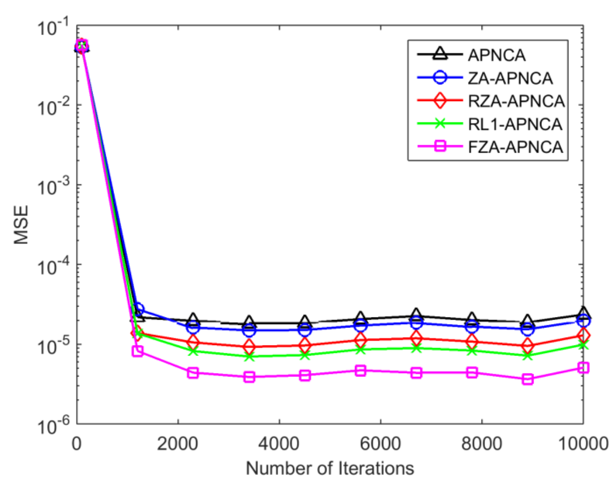


(c)

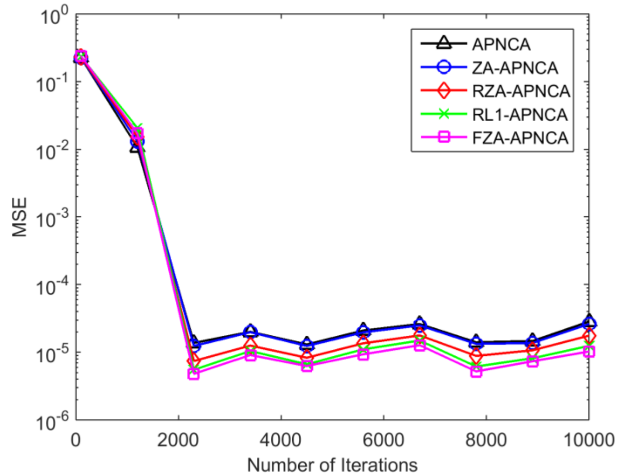
Fig. 6.3: MSEs of the proposed sparse AP-NCA algorithms for noise case 2 (impulsive observation noise: GMM) with the projection order, M=4 and different system sparsity of, (a) K=1, (b) K=4 and (c) K=8.

Comparison of the proposed sparse AP-NCA algorithms under noise case 3

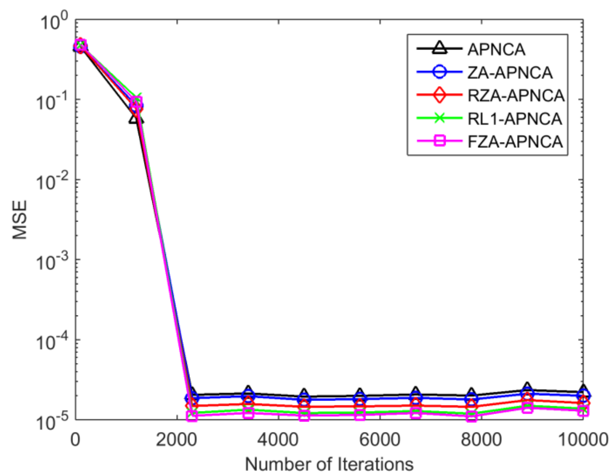
The performance of the proposed sparse algorithms under the assumption of impulse noise at filter input is shown in Fig. 6.4. It can be noticed that the proposed sparse APNCA algorithms exhibit better performance in terms of MSE when the system is highly sparse and it reduces as the system sparsity increases.



(a)



(b)



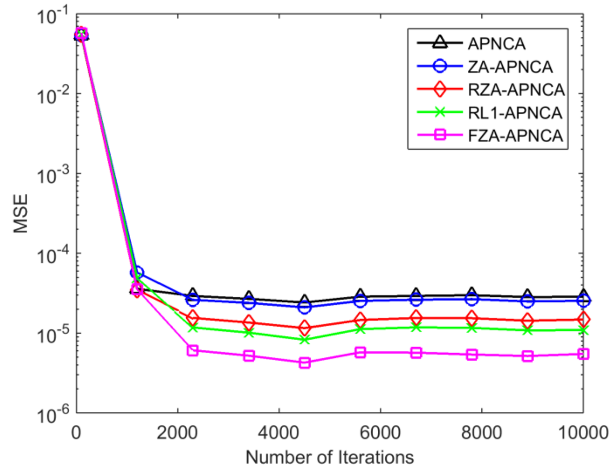
(c)

Fig. 6.4: MSEs of the proposed sparse AP-NCA algorithms for noise case 3 (impulse noise at filter input:B-G) with the projection order, M=4 and different system sparsity of, (a) K=1, (b) K=4 and (c) K=8.

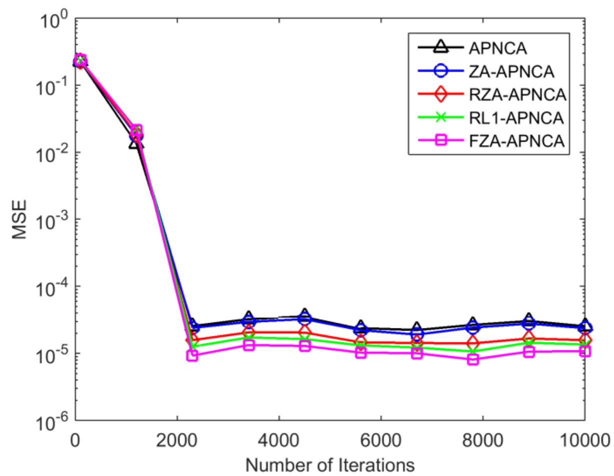
Comparison of the proposed sparse AP-NCA algorithms under noise case 4

The performance of the proposed sparse algorithms under the assumption of GMM observation noise & impulse noise at filter input is shown in Fig. 6.5. It can be noticed that the proposed sparse APNCA algorithms exhibit better performance in terms of MSE when the system is highly sparse and it reduces as the system sparsity increases. The

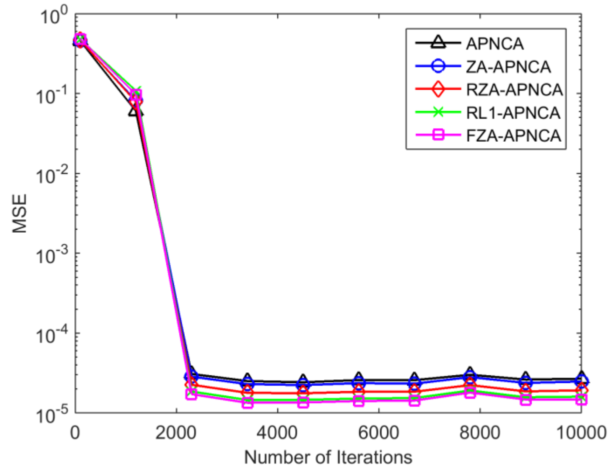
FZA-APNCA achieves minimum steady state error value. Similar results were obtained for higher L values in all previous cases.



(a)



(b)

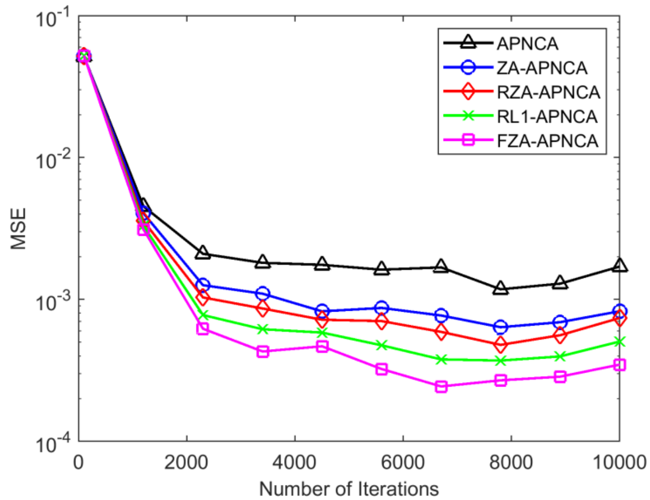


(c)

Fig. 6.5: MSEs of the proposed sparse AP-NCA algorithms for noise case 4 (GMM noise & impulse noise at filter input) with the projection order, M=4 and different system sparsity of, (a) K=1, (b) K=4 and (c) K=8.

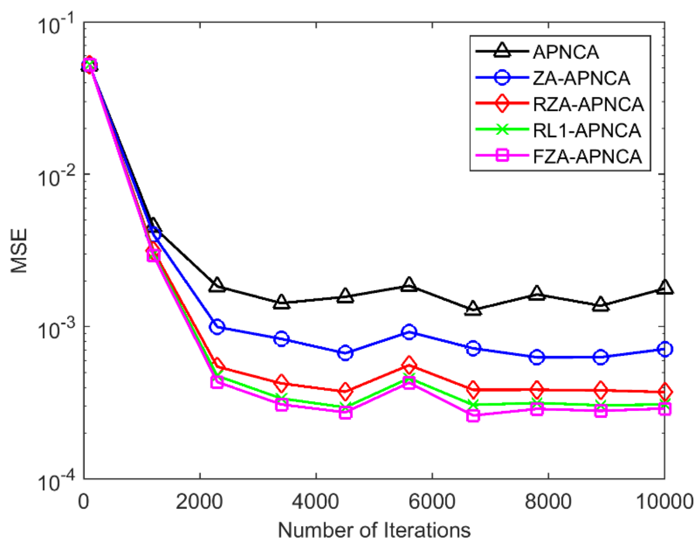
Now, let us evaluate the performance of the proposed AP-NCA algorithms for signals with complex noncircularity properties. The input is considered to be 16-QAM noncircular complex signal and the system is assumed to be of length L=16 with sparsity K={1, 4, 8}.

From Fig. 6.6, it is observed that the proposed algorithms exhibit better performance in the MSE sense for different noise cases shown in Table 6.1. The system sparsity is assumed to be K=1. The FZA-APNCA algorithm achieves minimum steady state error value.



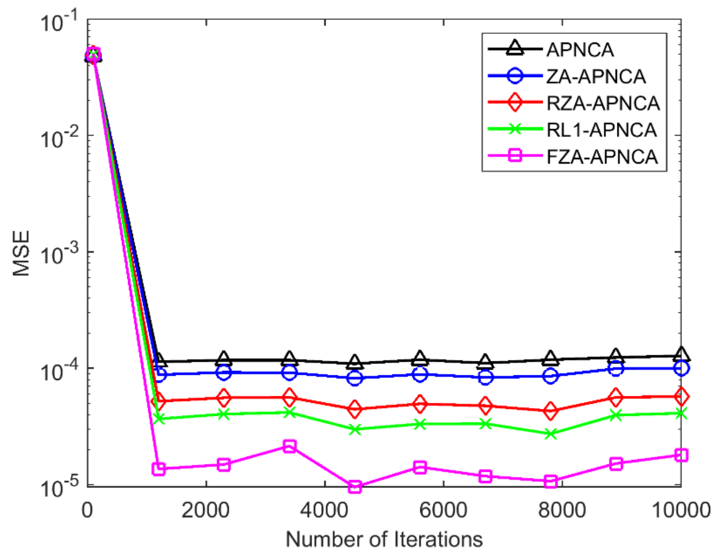
APNCA
ZA-APNCA
RZA-APNCA
RL1-APNCA
FZA-APNCA
MSE
Number of Iterations

(a)

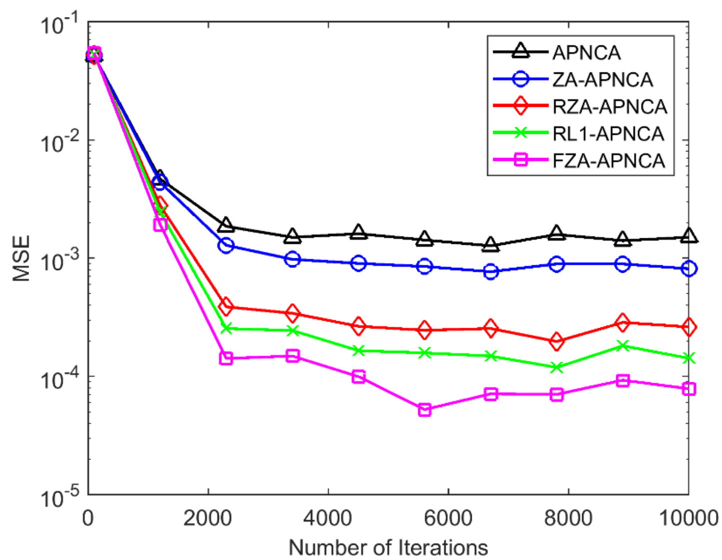


APNCA
ZA-APNCA
RZA-APNCA
RL1-APNCA
FZA-APNCA
MSE
Number of Iterations

(b)



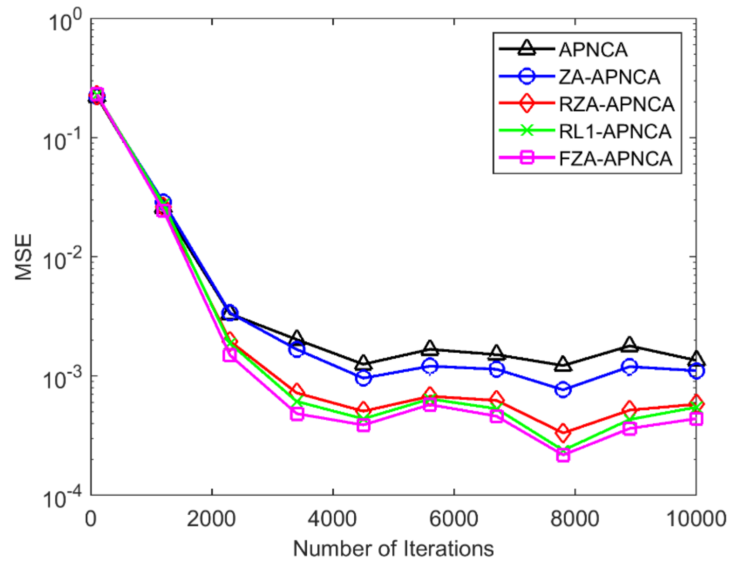
(c)



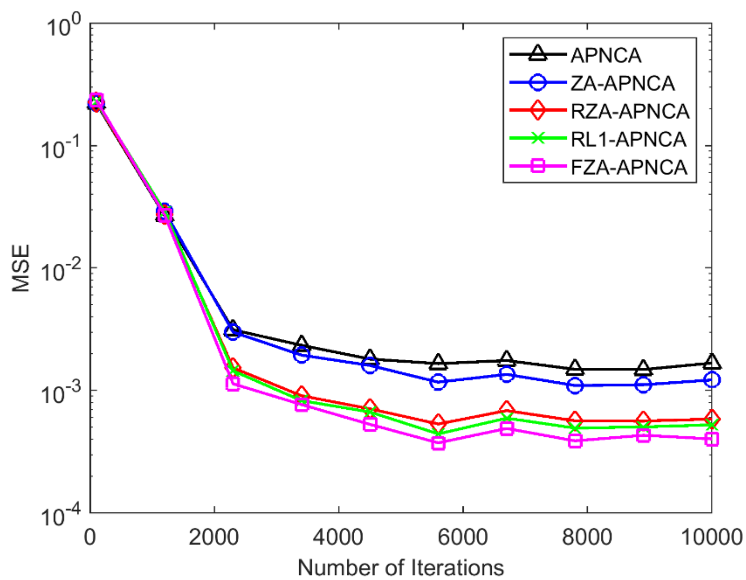
(d)

Fig. 6.6: MSEs of the proposed sparse AP-NCA algorithms for different noise cases with the projection order, M=4 and system sparsity of, K=1. (a) "white" Gaussian, (b) impulsive observation noise: GMM, (c) impulse noise at filter input:B-G, and (d) GMM noise & impulse noise at filter input.

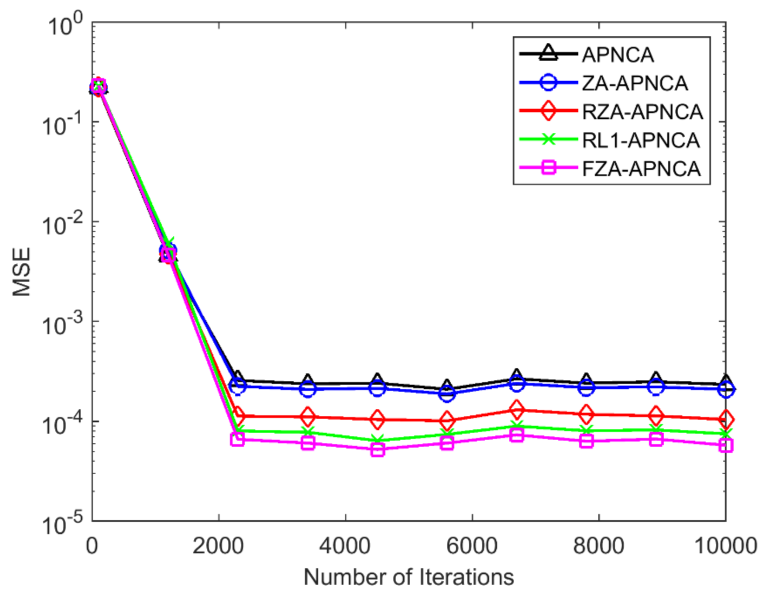
From Fig. 6.7, it is observed that the proposed algorithms exhibit better performance in the MSE sense for different noise cases shown in Table 6.1. The system sparsity is assumed to be K=4. The FZA-APNCA algorithm achieves minimum steady state error value.



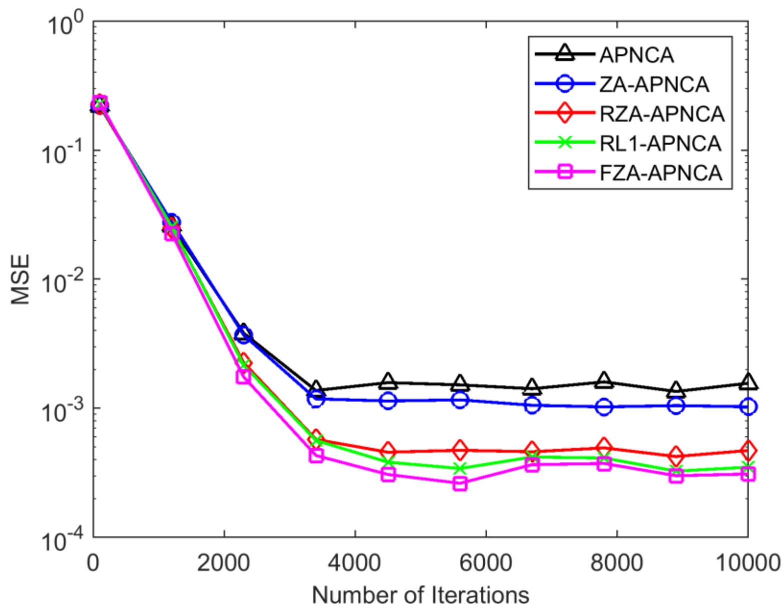
(a)



(b)



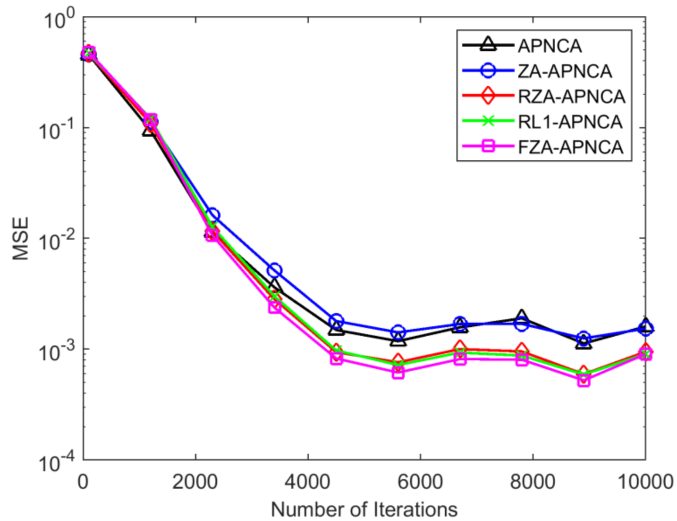
(c)



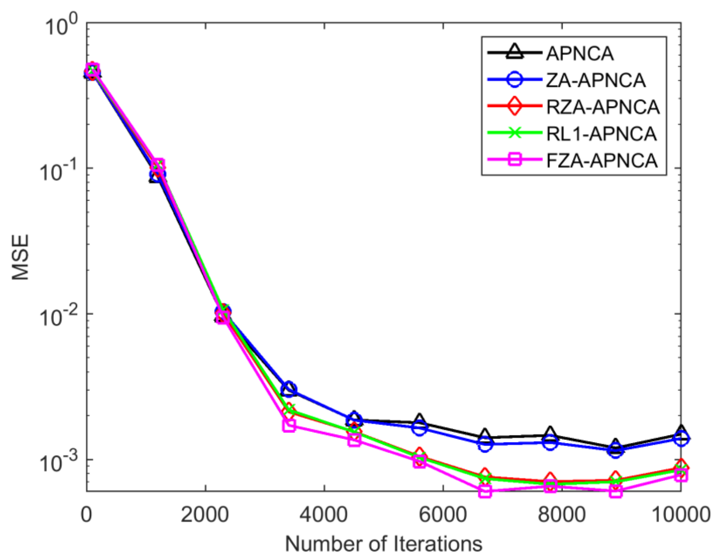
(d)

Fig. 6.7: MSEs of the proposed sparse AP-NCA algorithms for different noise cases with the projection order, M=4 and system sparsity of, K=4. (a) “white” Gaussian, (b) impulsive observation noise: GMM, (c) impulse noise at filter input:B-G, and (d) GMM noise & impulse noise at filter input.

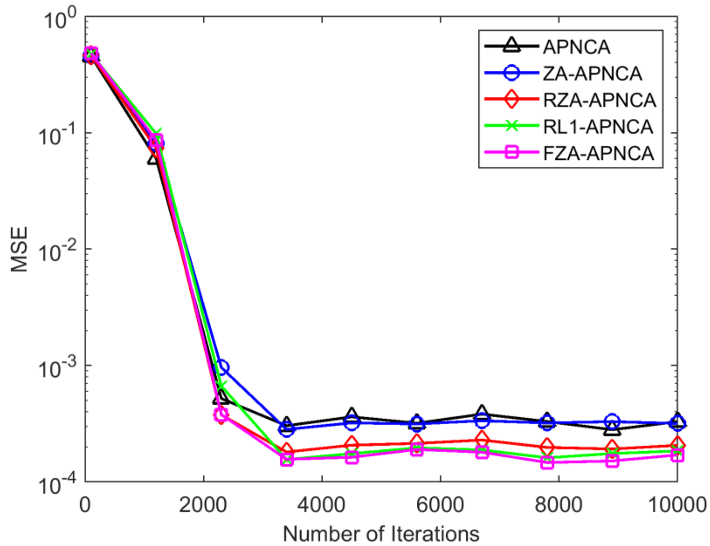
From Fig. 6.8, it is observed that the proposed algorithms exhibit better performance in the MSE sense for different noise cases shown in Table 6.1. The system sparsity is assumed to be K=8. The FZA-APNCA algorithm achieves minimum steady state error value.



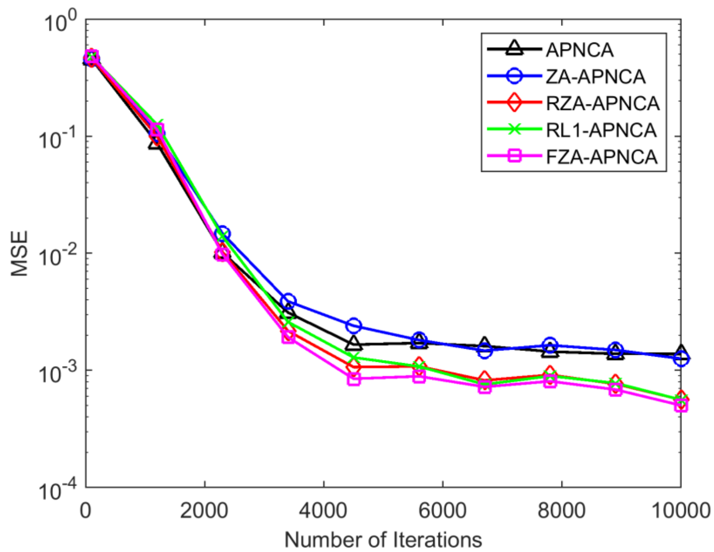
(a)



(b)



(c)



(d)

Fig. 6.8: MSEs of the proposed sparse AP-NCA algorithms for different noise cases with the projection order, M=4 and system sparsity of, K=8. (a) "white" Gaussian, (b) impulsive observation noise: GMM, (c) impulse noise at filter input:B-G, and (d) GMM noise & impulse noise at filter input.

6.7 Summary

The AP-NCA algorithm developed for adaptive filters in the complex domain has faster convergence for correlated inputs and at the same time highly robust in the presence of impulsive noise, but it does not promote sparsity. Hence, in this chapter, we have proposed four sparse APNCA algorithms in the sparse system identification context. Simulation results validate our proposed sparse algorithms in exploiting the system sparsity as well as robust to impulsive observation noise and impulsive filter input in the complex domain. Moreover, the proposed FZA-APNCA algorithm exhibit superior performance in Gaussian and non-Gaussian noise environments.

Chapter 7

Conclusions and Future scope

CHAPTER 7

7.1 Conclusions

This thesis deals with the adaptive filtering algorithms to system identification configuration. In applications such as network and acoustic echo cancellation, the impulse response of the echo channel is usually considered to be sparse with only a small number of non zero taps in the presence of large number of inactive taps. This thesis considers the problem of identifying the unknown system with time varying system sparsity. The problem of sparse system identification was formulated under the assumption of system background noise described as Gaussian and non-Gaussian impulsive noise. Given the above formulation of the problem, several novel contributions in the field of sparse system identification are proposed.

The research that has been carried out in this thesis can be divided in four parts. In the first part (Chapter 3) combinational approaches of two adaptive filters for system identification was proposed and studied. In the second part (Chapter 4) sparse adaptive algorithms based on Lyapunov stability theory to exploit system sparsity was proposed. The third part (Chapter 5) is concerned with development of sparse algorithms for system identification under impulsive noise environments. Finally, in the fourth part (Chapter 6) complex domain adaptive system identification using sparse APNCA algorithms under impulsive noises was proposed.

In Chapter 3, an adaptive affine combination of two IPNLMS filters is proposed. In this combination approach, the two adaptive filters are adapted independently and the output of individual filters is combined through a mixing parameter. The proposed approach tends to alleviate the convergence speed vs steady-state error tradeoff, as well as efficiently increase the IPNLMS filter robustness to time varying system sparsity. In a separate arrangement, Reweighted Zero Attracting-NLMS (RZA-NLMS) algorithm is developed to exploit system sparsity by introducing log-sum penalty into the cost function of the NLMS algorithm. In order to identify the system with varying sparseness, an affine combination of RZA-NLMS and NLMS algorithm is also proposed. The performance metrics Misalignment, Excess MSE and ERLE are used to validate the effectiveness of the proposed approaches.

In Chapter 4, new sparse adaptive algorithms namely, the Zero-Attracting Lyapunov Adaptation algorithm (ZA-LA), the Reweighted Zero-Attracting Lyapunov Adaptation algorithm (RZA-LA) and an affine combination of the LA and ZA-LA algorithms are proposed for sparse system identification. Adaptive algorithms based on Lyapunov stability theory offers improved convergence and stability, and overcome the problems faced by gradient descent-based adaptive filtering techniques. Performance measures MSD and MSE shows that the proposed algorithms performs better than the LMS algorithm and its sparse counterpart (ZA-LMS and RZA-LMS) for both white input and colored input cases.

In Chapter 5, novel sparse algorithms were developed under impulsive noise environments. Five different algorithms namely, ZA-NLMAT, RZA-NLMAT, RL1-LMAT, NNC-NLMAT and CIM-NLMAT are proposed by incorporating different sparsity constraints into Normalized Least Mean Absolute Third (NLMAT) which is based on high-order error power (HOEP) condition. Further, Modified Least-Mean Mixed-Norm algorithm which is based on sigmoid function (SLMMN) is also proposed to mitigate the adverse effects of impulsive noise and ZA-SLMMN, RZA-SLMMN algorithms are derived in sparse system identification context. The proposed algorithms outperforms the existing algorithms in terms of MSD and EMSE values and achieve robust performance against impulsive noise and are capable of exploiting the system with different levels of sparsity.

In Chapter 6, four different sparse algorithms in the complex domain namely, ZA-APNCA, RZA-APNCA, RL1-APNCA, and FZA-APNCA are developed using affine projection normalized correlation algorithm (APNCA). APNCA achieves faster convergence for a correlated input and is also robust against impulsive noise. Simulation results have shown that the proposed sparse algorithms outperform the existing APNCA in non-Gaussian environments in terms of MSE.

7.2 Future Scope

In chapter 5, the proposed SLMMN algorithms depending on the fixed step size μ and steepness parameter α did not guarantee the fast convergence and lower steady state error simultaneously. Hence, variable step-size (VSS) methods can be incorporated in the proposed algorithms to address this issue.

Actually in this work, the proposed algorithms were tested using MATLAB. Both input and noise signals were artificially generated in the same manner as in other related works. So, it is recommended that, the proposed algorithms be tested in real-time applications with real speech and noise signals as a future investigation.

Another possible future work could be the extension to multi-variable (MIMO) systems, since the work in this thesis was formulated to SISO systems.

References

[1] P.S.R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Second Edition, Kluwer Academic Publishers, Boston, 2004.
[2] S. Haykin, *Adaptive Filter Theory*, Fourth edition, Prentice Hall, New Jersey, 2002.
[3] A.H. Sayed, Fundamentals of Adaptive Filtering, John Wiley & Sons, New York, 2003.
[4] J. G. Avalos, J. C. Sanchez and J. Velazquez., "*Applications of Adaptive Filtering*," in Adaptive Filtering Applications, Lino Garcia (Ed.), InTech, pp. 1–20, 2011.
[5] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, John Wiley & Sons, New York, 1998.
[6] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Inc. Englewood Cliffs, N.J. 1985.
[7] B. Widrow and M.E. Hoff, Jr., "Adaptive Switching Circuits," *IRE WESCON Convention Record*, part 4, pp. 96-104, August 1960.
[8] S. Haykin, B. Widrow (Editor), *Least-Mean-Square Adaptive Filters*, Wiley, 2003.
[9] S.C. Douglas, "A family of normalized LMS algorithms," *IEEE Signal Processing Letters*, Vol. 1, No. 3, pp. 49-51, March 1994.
[10] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP34, pp. 793–806, Aug. 1986.
[11] K. Ozeki and T. Umeda, "An Adaptive Filtering Algorithm Using an Orthogonal Projection to an Affine Subspace and Its Properties," *Electronics and Communications in Japan*, Vol.67-A, No. 5, pp.19-27, 1984.
[12] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Processing*, Vol. 41, pp. 2811–2825, Sept. 1993.
[13] S. L. Gay and S. Tavathia, "The fast affine projection algorithm," *in Proc. ICASSP-95*, Detroit, USA, Vol.5, pp. 3023–3026, May 1995.
[14] S. G. Sankaran and A. A. Louis Beex, "Normalized LMS algorithm with orthogonal correction factors," *in Proc. 31st Asilomar Conference*, pp.1670–1673, Nov. 1997.
[15] M.L. Honig and D.G. Messerschmitt, *Adaptive Filters - Structures, Algorithms and Applications*, Kluwer Academic Publishers, 1984.
[16] A.H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, 11(3), pp. 18–60, July 1994.
[17] R. K. Martin, W. A. Sethares, R. C. Williamson, and C. R. J. Jr., "Exploiting sparsity in adaptive filters," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1883-1894, Aug. 2002.
[18] B. D. Rao and B. Song, "Adaptive filtering algorithms for promoting sparsity," *ICASSP*, pp. 361-364, Jun. 2003.
[19] C. Paleologu, J. Benesty, and S. Ciochina, *Sparse Adaptive Filters for Echo Cancellation*, San Rafael, CA: Morgan and Claypool, 2010.
[20] E. J. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted $\ell 1$ minimization," *Journal of Fourier Analysis and Applications*, 14(5), pp. 877–905, 2008.

[21] E. J. Candès, "Compressive sampling," *in Proc. Int. Congress of Mathematicians*, vol. 3, pp. 1433–1452, 2006.
[22] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[23] U. W. Bajwa, J. Haupt, G. Raz and R. Nowak, "Compressed Channel Sensing," *42nd Annual Conference on Information Sciences and Systems, CISS'08*, Princeton, pp. 5-10, March 2008.
[24] C. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Commun. Mag*., vol. 48, no. 11, pp. 164 –174, Nov. 2010.
[25] W. F. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions", *in Proc. IEEE*, vol. 83, no. 6, pp. 958-981, 1995.
[26] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*, Berlin, Germany: Springer-Verlag, 2001.
[27] J. Radecki, Z. Zilic, and K. Radecka, "Echo cancellation in IP networks", *in Proc. of the 45th Midwest Symposium on Circuits and Systems*, vol. 2, pp. 219-222, August, 2002.
[28] N. Kalouptsidis and S. Theodoridis, *Adaptive System Identification and Signal Processing Algorithms*, Prentice-Hall, Englewood Cliffs, NJ, 1993.
[29] J. G. Proakis, and M. Salehi, *Communication Systems Engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1994.
[30] M. M. Sondhi, "An adaptive echo canceller," *Bell Sys. Tech. J*., 46, pp. 497–511, March 1967.
[31] D. G. Messerschmitt, "Echo cancellation in speech and data transmission," *IEEE J. Sel. Areas Commun.*, SAC-2(2), pp. 283–297, March 1984.
[32] K. Murano, S. Unagami, and F. Amano, "Echo cancellation and applications," *IEEE Commun. Mag*., 28(1), pp. 49–55, Jan. 1990.
[33] B. Widrow, J. R. Glover, Jr., McCool, et al., "Adaptive noise cancelling: principles and applications," *Proc. IEEE*, 63(12), pp. 1692–1716, Dec. 1975.
[34] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*, New York:Springer, 1988.
[35] E. J. Wegmcin, S. C. Schwartz, and J. B. Thomas, Eds., *Topics in Non-Gaussian Signal Processing*, New York: 1989
[36] C.L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, John Wiley & Sons, New York, 1995.
[37] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band power line communications," *Electromagnetic Compatibility, IEEE Transactions on*, vol. 44, no. 1, pp. 249 –258, Feb 2002.
[38] M. Bouvet and S.C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acous. Speech and Signal Proc*., Vol. 37, pp. 621-626, 1989.
[39] D. Zha, T. Qiu, "Underwater sources location in non-Gaussian impulsive noise environments," *Digit. Signal Process*., 16 (2), pp. 149–163, 2006.
[40] B.W. Stuck and B. Kleiner, "A Statistical Analysis of the Telephone Noise," *Bell Syst. Tech. J.*, Vol. 53, no. 7, pp. 1263-1320, 1974.

[41] D. Dragomir, R. Dragomir, B. Pantelimon, S. Puscoci, “Modeling and Measuring the Pulse Noise in Telecommunication Cables,” *in Proc. of the 7th International Conference of Electromechanical and Power Systems, SIELMEN 2009*, Iasi, Romania, Vol. I, pp. 210-212, Oct. 2009.
[42] V. V. Krishna, J. Rayala and B. Slade, “Algorithmic and Implementation Aspects of Echo Cancellation in Packet Voice Networks”, *in Proc. 36th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, pp. 1252-1257, 2002.
[43] M. Kocic, D. Brady and M. Stojanovic, “Sparse Equalization for Real-Time Digital Underwater Acoustic Communications,” *in Proc. IEEE OCEANS*, vol. 3, pp. 1417-1422, 1995.
[44] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, John Wiley & Sons, 2008.
[45] Z. Man, H. R. Wu, W. Lai and T. Nguyen, “Design of Adaptive Filters Using Lyapunov Stability Theory,” *Proc. of the 6th IEEE International Workshop on Intelligent Signal Processing and Communication Systems*, pp. 304-308, Nov. 1998.
[46] S. K. Phooi, Z. Man and H. R. Wu, “Nonlinear adaptive RBF Neural Filter with Lyapunov Adaptation Algorithm and its application to nonlinear channel equalization,” *Proc. of the 5th International Symposium on Signal Processing and its Applications*, pp. 151-154, August 1999.
[47] H. Zhao and J. Zhang, “Filtered-s Lyapunov algorithm for active control of nonlinear noise processes,” *In Proceedings of the 9th International Conference on Signal Processing*, Beijing, China, pp. 311–314, 2008.
[48] E. C. Mengüc and N. Acir, “Lyapunov stability theory based adaptive filter algorithm for noisy measurements,” *2013 UKSim 15th International Conference on Computer Modelling and Simulation*, Cambridge, UK, pp. 451–454, 2013.
[49] E. C. Mengüc and N. Acir, “A novel adaptive filter design using Lyapunov stability theory,” *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 23, no. 6, pp. 719–728, 2015.
[50] E. C. Mengüc and N. Acir, “Real-Time Implementation of Lyapunov Stability Theory-Based Adaptive Filter on FPGA,” *IEICE Transactions on Electronics*, vol. E99.C, no. 1, pp. 129–137, 2016.
[51] D. L. Duttweiler, “Proportionate normalized least-mean- squares adaptation in echo cancellers,” *IEEE Trans.Speech Audio Process*., vol.8, no.5, pp. 508–518, July 2000.
[52] J. Benesty and S. L. Gay, “An improved PNLMS algorithm,” *in Proc. IEEE ICASSP-02*, Orlando USA , vol. 2, pp. 1881–1884 , Nov. 2002.
[53] Z. Chen, S. Haykin and S. L. Gay, “*Proportionate Adaptation:New Paradigms in Adaptive Filters*,” Ch. 8, Wiley, New York, 2005,
[54] H. Deng and M. Doroslova˘cki, “Improving Convergence of the PNLMS Algorithm for Sparse Impulse Response Identification,” *IEEE Signal Processing Letters*, vol. 12, no.3, pp. 181–184 , May 2005.
[55] R. Baraniuk, “Compressive sensing,” *in IEEE Signal Process. Mag*., vol. 25, pp. 21-30, March 2007.
[56] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Process. Mag*., vol. 21, no. 3, pp. 21–30, Mar. 2008.

[57] H. Deng and M. I. Doroslovacki, “Proportionate adaptive algorithms for network echo cancellation,” *IEEE Trans. Signal Process*., vol. 54, no. 5, pp. 1794–1803, May 2006.
[58] C. Paleologu, J. Benesty and S. Ciochina, “An improved proportionate NLMS algorithm based on the ℓ_0-norm,” *in Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, TX, pp. 309–312, Mar. 2010.
[59] A. W. H. Khong and P. A. Naylor, “Efficient use of sparse adaptive filters,” *in Proc. 40th Asilomar Conf. Signals, Syst., Comput*., Pacific Grove, CA, pp. 1375–1379, Oct. 2006.
[60] P.A. Naylor, J. Cui and M. Brookes, “Adaptive algorithms for sparse echo cancellation,” Signal Process., vol. 86, no. 6, pp. 1182-1192, 2006.
[61] B. Jelfs, D. P. Mandic and A. Cichocki, “A unifying approach to the derivation of the class of PNLMS algorithms,” *in Proc. 15th Int. Conf. Digit. Signal Process*., Cardiff, U.K., pp. 35-38, Jul. 2007.
[62] B. Jelfs and D. P. Mandic, “A unifying framework for the analysis of proportionate NLMS algorithms,” International Journal of Adaptive Control and Signal Processing, vol. 29, no. 9, pp. 1073-1085, 2015.
[63] E. Soria-Olivas, J. Calpe-Maravilla, J. Guerrero-Martinez, M. Martinez-Sober, J. Espi-Lopez, “An easy demonstration of the optimum value of the adaptation constant in the LMS algorithm,” IEEE Transactions on Education, 41(1):81, 1998.
[64] K. T. Wagner and M. I. Doroslovacki, “Gain allocation in proportionate-type NLMS algorithms for fast decay of output error at all times,” *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.(ICASSP)*, Taipei, Taiwan, pp. 3117-3120, Apr. 2009.
[65] F. C. de Souza, O. J. Tobias, R. Seara and D. R. Morgan, “Alternative approach for computing the activation factor of the PNLMS algorithm,” *in Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Glasgow, U.K., pp. 2633-2637, Aug. 2009.
[66] J. Cui, P. A. Naylor and D. T. Brown, “An improved IPNLMS algorithm for echo cancellation in packet switched networks,” *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Montreal, QC, Canada, vol. 4, pp. 141–144, May 2004.
[67] P. Loganathan, A.W Khong, P. Naylor, A class of sparseness-controlled algorithms for echo cancellation, *Audio Speech Lang. Process. IEEE Trans.*, 17(8), 1591–1601, 2009.
[68] F. C. de Souza, O. J. Tobias, R. Seara and D. R. Morgan, “A PNLMS algorithm with individual activation factors,” *IEEE Trans. Signal Process*., vol. 58, no. 4, pp. 2036–2047, Apr. 2010.
[69] A. W. H. Khong, P. A. Naylor and J. Benesty, “A low delay and fast converging improved proportionate algorithm for sparse system identification,” *in EURASIP J. Audio, Speech, Music Process.*, vol. 2007, pp. 1–8, 2007.
[70] F. C. de Souza, R. Seara and D. R. Morgan, An enhanced IAF-PNLMS adaptive algorithm for sparse impulse response identification, *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3301–3307, Jun. 2012.
[71] F. L. Perez, E. V. Kuhn, F. C. de Souza and R. Seara, “A novel gain distribution policy based on individual-coefficient convergence for PNLMS-type algorithms,” *Signal Processing*, vol. 138, pp. 294, 2017.

[72] Y Zhang, C. Wang and N. Li, “A fast converging algorithm for systems with highly sparse impulse response,” *IEEE Fifth International Joint Conference on Computational Sciences and Optimization*, vol. 12, no. 3, pp. 831–834, 2012.
[73] L. Liu, M. Fukumoto and S. Saiki, “An improved mu-law proportionate NLMS algorithm,” *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP),* Las Vegas, NE, USA, Apr. 2008, pp. 3797–3800.
[74] K. Sakhnov, “An improved proportionate affine projection algorithm for network echo cancellation,” *in Proc. 15th Int. Conf. Syst., Signals Image Process*., Bratislava, Slovakia, pp. 125–128, Jun. 2008.
[75] C. Paleologu, S. Ciochina and J. Benesty, “An efficient proportionate affine projection algorithm for echo cancellation,” *IEEE Signal Process. Lett*., vol. 17, no. 2, pp. 165–168, Feb. 2010.
[76] R. Arablouei, K. Dogançay and S. Perreau, “Proportionate affine projection algorithm with selective projections for sparse system identification,” *in Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf*., pp. 362–366, 2010.
[77] C. Stanciu, C. Anghel, C. Paleologu, J. Benesty, F. Albu and S. Ciochina, “A proportionate affine projection algorithm using dichotomous coordinate descent iterations,” *in Proc. 10th Int. Symp. Signals, Circuits Syst. (ISSCS)*, pp. 1–4, 2011.
[78] C. Paleologu and J. Benesty, “Proportionate affine projection algorithms from a basis pursuit perspective,” *in Proc. IEEE Int. Symp. Circuits Syst*., Seoul, South Korea, pp. 2757–2760, May 2012.
[79] H. Zhao, Y. Yu, X. Zeng and Z. He, “Memory proportionate APA with individual activation factors for acoustic echo cancellation,” *IEEE Trans. Audio, Speech, Language Process*., vol. 22, no. 6, pp. 1047–1055, Jun. 2014.
[80] P. Loganathan, E. A. P. Habets and P. A. Naylor, “A partitioned block proportionate adaptive algorithm for acoustic echo cancellation,” *in Proc. APSIPA Annu. Summit Conf*., Biopolis, Singapore, Dec. 2010.
[81] J. Liu and S. L. Grant, “Proportionate adaptive filtering for block-sparse system identification,” *IEEE/ACM Trans. Audio, Speech, Language Process*., vol. 24, no. 4, pp. 623–630, Apr. 2016.
[82] J. Liu and S. L. Grant, “Proportionate affine projection algorithms for block-sparse system identification,” *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP),* Shanghai, China, pp. 529–533, Jan. 2016.
[83] Z. Jin, Y. Li and J. Liu, “An Improved Set-Membership Proportionate Adaptive Algorithm for a Block-Sparse System,” *Symmetry*, 10 (3), 75, pp. 1-15, 2018.
[84] Z. Wu, S. Peng, B. Chen, H. Zhao and J. C. Principe, “Proportionate minimum error entropy algorithm for sparse system identification,” *Entropy,* 17, pp. 5995–6006, 2015.
[85] Y. Chen, Y. Gu and A. O. Hero, III, “Sparse LMS for system identification,” *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09),* Taipei, Taiwan, pp. 3125–3128, April 2009.
[86] Y. Gu, J. Jin and S. Mei, “ℓ_0-norm constraint LMS algorithm for sparse system identification,” *IEEE Signal Process. Lett*., 16 (9), pp. 774-777, 2009.
[87] G. Su, J. Jin, Y. Gu and J. Wang, “Performance analysis of ℓ_0-norm constraint least mean square algorithm,” *IEEE Trans. Signal Process*., vol. 60, no. 5, pp. 2223–2235, May 2012.

[88] Y. Li and M. Hamamura, “Zero-attracting variable-step-size least mean square algorithms for adaptive sparse channel estimation,” *Int. J. Adapt. Control Signal Process.*, vol. 29, no. 9, pp. 1189-1206, 2015.
[89] J. Jin, Q. Qu and Y. Gu, “Robust zero-point attraction least mean square algorithm on near sparse system identification,” *IET Signal Processing*, 7(3), pp. 210-218, 2013.
[90] R. Meng, R. C. de Lamare, and V.H. Nascimento, “Sparsity-aware affine projection adaptive algorithms for system identification,” *in Proc. Sensor Signal Process. Defence (SSPD)*, London, U.K., pp. 1–5, Sep. 2011.
[91] F. Y. Wu and F. Tong, “Non-uniform norm constraint LMS algorithm for sparse system identification,” *IEEE Commun. Lett.*, vol. 17, no. 2, pp. 385–388, Jan. 2013.
[92] F. Y. Wu and F. Tong, “Gradient optimization p-norm-like constraint LMS algorithm for sparse system estimation,” *Signal Process.*, vol. 93, no. 4, pp. 961–967, Apr. 2013.
[93] C. Wang, Y. Zhang, Y. Wei and N. Li, “A new ℓ_0-LMS algorithm with adaptive zero attractor,” *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2150–2153, Dec. 2015.
[94] Y. Li and M. Hamamura, “Smooth approximation ℓ_0-norm constrained affine projection algorithm and its applications in sparse channel estimation,” *The Scientific World Journal*, vol. 2014, pp. 1-14, 2014.
[95] M. V. S. Lima, W. A. Martins and P. S. Z. Diniz, “Affine projection algorithms for sparse system identification,” *Proceedings of the IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP '13)*, pp. 5666-5670, May 2013.
[96] O. Taheri and S. A. Vorobyov, “Sparse channel estimation with Lp-norm and reweighted L1-norm penalized least mean squares,” In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2864-2867, May 2011.
[97] Y. Li, C. Zhang and S. Wang, “Low complexity non-uniform penalized affine projection algorithm for sparse system identification,” *Circuits Systems and Signal Processing*, vol. 35, no. 5, pp. 1611-1624, 2016.
[98] Y. Li, W. Li, W. Yu, J. Wan and Z. Li, “Sparse adaptive channel estimation based on lp-norm-penalized affine projection algorithm,” *International Journal of Antennas and Propagation*, vol.2014, pp.1-8, 2014.
[99] G. Gui, A. Mehbodniya and F. Adachi, “Least mean square/fourth algorithm for adaptive sparse channel estimation,” *in Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, London, U.K., pp. 296–300, Sept. 2013.
[100] Y. Li, Y. Wang and T. Jiang, “Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation,” *Signal Processing*, vol. 128, pp. 243-251, 2016.
[101] Y. Li, Y. Wang and T. Jiang, “Low complexity norm-adaption least mean square/fourth algorithm and its applications for sparse channel estimation,” *IEEE* Wireless Communications and Networking (WCNC 2016), pp. 3-6, April 2016.
[102] Y. Li, Y. Wang and F. Albu, “Sparse Channel Estimation Based on a Reweighted Least-Mean Mixed-Norm Adaptive Filter Algorithm,” in Proc. of EUSIPCO, Budapest, Hungary, pp. 2380-2384, 2016.

[103] B. Babadi, N. Kalouptsidis and V. Tarokh, "SPARLS: the sparse RLS algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4013-4025, Aug. 2010.

[104] D. Angelosante, J. A. Bazerque and G. B. Giannakis, "Online adaptive estimation of sparse signals: where RLS meets the ℓ_1-norm," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3436-3447, Jul. 2010.

[105] X. Hong, J. Gao and S. Chen, "Zero attracting recursive least squares algorithms," *IEEE Transactions on Vehicular Technology*, Vol. 66, no. 1, pp. 213- 221, 2016.

[106] E. M. Eksioglu, "Group sparse RLS algorithms," *Int. J. Adapt. Contr. Signal Process.*, vol. 28, no. 12, pp. 1398–1412, 2014.

[107] S. Jiang and Y. Gu, "Block-sparsity-induced adaptive filter for multi clustering system identification," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5318–5330, Oct. 15, 2015.

[108] Y. Li and M. Hamamura, "An improved proportionate normalized least-mean-square algorithm for broadband multipath channel estimation," *The Scientific World Journal*, vol. 2014, pp. 1–9, 2014.

[109] R.L. Das and M. Chakraborty, "A zero attracting proportionate normalized least mean square algorithm," *2012 Asia-Pacific Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1-4, 2012.

[110] R. L. Das and M. Chakraborty, "Improving the performance of the PNLMS algorithm using ℓ_1-norm regularization," *IEEE Trans. Audio, Speech, Language Process.*, vol. 24, no. 7, pp. 1280–1290, Jul. 2016.

[111] H. Zhao and Y. Yu, "Improved efficient proportionate affine projection algorithm based on l0-norm for sparse system identification," *The Journal of Engineering*, Vol. 2014, Issue 1, pp. 20 – 23, 2014.

[112] J. Liu, S. L. Grant and J. Benesty, "A low complexity reweighted proportionate affine projection algorithm with memory and row action projection," *EURASIP Journal on Advances in Signal Processing*, 2015(1), pp. 1–12, 2015.

[113] Z. Qu and Z. Zheng, "An efficient ℓ_0-norm constraint memory improved proportionate affine projection algorithm," *Circuits, Syst., Signal Process.*, 36 (8), pp. 3448-3456, 2017.

[114] Y. Li, Y. Wang, F. Albu and J. Jiang, "A general zero attraction proportionate normalized maximum correntropy criterion algorithm for sparse system identification," *Symmetry*, vol. 9, no. 10, 229, pp. 1-13, 2017.

[115] Y. Li, Y. Wang and L. Sun, "A Proportionate Normalized Maximum Correntropy Criterion Algorithm with Correntropy Induced Metric Constraint for Identifying Sparse Systems," *Symmetry*, Vol. 10, no. 12, 683, 2018.

[116] F. Albu, Y. Li and Y. Wang, "The *lp*-norm Proportionate Normalized Least Mean Square Algorithm for Active Noise Control," *in Proc. of ICSTCC 2017*, Sinaia, Romania, pp. 401-405, 2017.

[117] J. Arenas-Garcia and A. R. Figueiras-Vidal, "Adaptive combination of proportionate filters for sparse echo cancellation," *in IEEE Trans. Audio, Speech Language Process.*, vol. 17, no. 6, pp. 1087-1098, August, 2009.

[118] J. Arenas-Garcia, M. de Diego, L. A. Azpicueta-Ruiz, M. Ferrer and A. Gonzalez, "Combinations of proportionate adaptive filters in acoustics: an application to ANC," *in Proc. EUSIPCO*, pp.1269-1273, 2011.

[119] B. K. Das and M. Chakraborty, “Sparse Adaptive Filtering by an Adaptive Convex Combination of the LMS and the ZA LMS Algorithms,” *in IEEE Trans. Circuits and Systems I*, vol. 61, no. 5, pp. 1499-1507, May, 2014.
[120] M. Martinez-Ramon, J. Arenas-Garcia, A. NaviaVazquez and A. R. Figueiras-Vidal, “An Adaptive Combination of Adaptive Filters for Plant Identification,” *In Proc. 14th International Conference on Digital Signal Processing*, Santorini, Greece, pp. 1195-1198, 2002.
[121] Y. Zhang and J. A. Chambers, “Convex combination of adaptive filters for a variable tap-length LMS algorithm,” *IEEE Signal Process Lett.*, vol. 13, no. 10, pp. 628-631, 2006.
[122] L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal and J.Arenas-Garcia, “A Normalized Adaptation Scheme for the Convex Combination of Two Adaptive Filters,” *In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, pp.3301-3304, 2008.
[123] A. Gonzalo-Ayuso, M. T. M. Silva, V. H. Nascimento and J. Arenas-Garc´ıa, “Improving sparse echo cancellation via convex combination of two NLMS filters with different lengths,” *in 2012 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, Sept. 2012.
[124] M. Ferrer, M. Diego and A. Gonzalez, G. Pinero, “Convex combination of affine projection algorithms,” *17th European Signal Processing Conference (EUSIPCO)*, pp. 431-435, 2009.
[125] R. Candido, M. T. M. Silva and V. H. Nascimento, “Affine combinations of adaptive filters,” *in Proc. 42nd Asilomar Conf. Signal, Syst. Comput*., pp. 236-240, Oct. 2008.
[126] T. Trump, “An output signal based combination of two NLMS adaptive algorithms,” *In Proc. 16th International Conference on Digital Signal Processing*, Santorini, Greece. pp. 1-6, 2009.
[127] T. Trump, “Steady state analysis of an output signal based combination of two NLMS adaptive filters,” *In Proc. 17th European Signal Processing Conference (EUSIPCO*, Glasgow, Scotland, pp. 1720–1724, 2009.
[128] T. Trump, “Transient analysis of an output signal based combination of two NLMS adaptive filters,” *IEEE international workshop on machine learning for signal processing*, pp. 244-249, 2010.
[129] J. Ni and F. Li, “Adaptive combination of subband adaptive filters for acoustic echo cancellation,” *IEEE Trans Consum Electron.*, 56 (3), pp. 1549-1555, 2010.
[130] L. Lu, H. Zhao, Z. He and B. Chen, “A novel sign adaptation scheme for convex combination of two adaptive filters,” *International Journal of Electronics and Communications*, 69(11), pp. 1590–1598, 2015.
[131] K. H. Kim, S. E. Kim, Y. S. Choi and W. J. Song, “A low-complexity complementary pair affine projection adaptive filter,” *IEICE Transactions on Fundamentals of Electronics,Communications and Computer Sciences*, vol. 97, no. 10, pp. 2074–2078, 2014.
[132] V. H. Nascimento, M. T. M. Silva and J. Arenas-Garc´ıa, “A low-cost implementation strategy for combinations of adaptive filters,” *in Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Process*., Vancouver, Canada, pp. 5671–5675, 2013.

[133] A. T. Erdogan, S. S. Kozat and A. C. Singer, "Comparison of convex combination and affine combination of adaptive filters," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, Taipei, Taiwan, pp. 3089–3092, 2009.

[134] S. Peng, Z. Wu, Y. Zhou and B. Chen, "Adaptive convex combination filter under minimum error entropy criterion," *in Proc. 2016 IEEE Int. Conf. Dig. Signal Process*., pp. 44–48, 2016.

[135] F. Huang, J. Zhang and Y. Pang, "A novel combination scheme of proportionate filter," *Sig. Process*. 143, pp. 222–231, 2018.

[136] Y. Wang, Y. Li, J. C. M. Bermudez, X. Han, "An adaptive combination constrained proportionate normalized maximum correntropy criterion algorithm for sparse channel estimations," *EURASIP J. Adv. Signal Process*., 2018: 58, pp. 1-13, 2018.

[137] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Electromagn. Compat*., vol. 44, no. 1, pp. 249–258, Feb. 2002.

[138] T. G¨ansler et al., "Double-talk robust fast converging algorithms for network echo cancelation," *IEEE Trans. Speech, Audio Process*., vol. 8, no. 6, pp. 656–663, 2000.

[139] P. Kismode, "Alpha-stable distributions in signal processing of audio signals," *in Proc. 41st Conf. Simulation Modelling*, pp. 87–94, Sep. 2000.

[140] F. Wen, "Diffusion least-mean P-power algorithms for distributed estimation in alpha-stable noise environments," *Electron. Lett*., vol. 49, no. 21, pp. 1355–1356, Oct. 2013.

[141] S. Koike, "Analysis of normalized least mean modulus algorithm for adaptive filters in impulsive noise environments," *in Proc. IEEE ISAPCS*, Chengdu, China, Dec. 2010.

[142] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Trans. Inf. Theory*, 30, pp. 275–283, 1984.

[143] S. Pei and C. Tseng, "Least mean p-power error criterion for adaptive FIR filters," *IEEE J. Sel. Areas Commun*., 12, pp. 1540–1547, 1994.

[144] S. H. Cho, S. D. Kim, H. P. Moon and J. Y. Na, "Least mean absolute third (LMAT) adaptive algorithm: Mean and mean-squared convergence properties," *Proc. Sixth Western Pacific Reg. Acoust. Conf.*, Hong Kong, November 1997.

[145] S. Cho and D.K. Sang, "Adaptive filters based on the high order error statistics," *IEEE Asia Pacific Conference on Circuits and Systems*, Seoul, South Korea, pp. 109-112, November 1996.

[146] Y. H. Lee, J. D. Mok, S. D. Kim and S. H. Cho, "Performance of least mean absolute third (LMAT) adaptive algorithm in various noise environments," *Electron. Lett.*, vol. 34, no. 3, pp. 241–243, Feb. 1998.

[147] H. Zhao, Y. Yu, S. Gao and Z. He, "A new normalized LMAT algorithm and its performance analysis," *Signal Process*., 105(12), pp. 399-409, 2014.

[148] F. Adachi and E. Kudoh, "New direction of broadband wireless technology," *Wirel. Commun. Mob. Comput.*, 7(8), pp. 969-983, 2007.

[149] J. A. Chambers, O. Tanrikulu and S. A. G. Constantinides, "Least mean mixed-norm adaptive filtering," *Electronics Letters*, vol.30, no.19, pp.1574-1575, 1994.

[150] O. Tanrikulu and J. A. Chambers, "Convergence and steady-state properties of the least-mean mixed-norm (LMMN) adaptive algorithm," *IEEE Proc. Vision, Image and Signal Processing*, Vol. 143, No. 3, 1996.

[151] I. Pitas and A. N. Venetsanopoulos, "Adaptive filters based on order statistics," *IEEE Trans. Signal Process*., vol. 39, no. 2, pp. 518-522, Feb. 1991.

[152] P. Salemhier, "Adaptive rank order based filters," *Signal Process*., vol. 27, no. 1, pp. 1-25, Apr. 1992.

[153] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proc. IEEE*, vol. 81, no. 7, pp. 986–1010, Jul. 1993.

[154] S. G. Sankaran and A. A. Beex, "Convergence behavior of affine projection algorithms," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1086-1096, Apr. 2000.

[155] S. Koike, "Analysis of normalized correlation algorithm for adaptive filters in impulsive noise environments," *19th European Signal Processing Conference (EUSIPCO)*, pp. 141-145, 2011.

[156] S. Koike, "Analysis of Affine Projection Normalized Correlation Algorithm," *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2016.

[157] E. J. Candes, J. Romberg and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[158] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society B: Methodological*, vol. 58, no. 1, pp. 267–288, 1996.

[159] Z. Yang, Y. R. Zheng and S. L. Grant, "Proportionate affine projection sign algorithms for network echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 8, pp. 2273–2284, Nov. 2011.

[160] F. Albu and H.K. Kwan, "Memory improved proportionate affine projection sign algorithm," *Electron. Lett.*, vol. 48, no. 20, pp. 1279–1281, 2012.

[161] Y. Yu and H. Zhao, "An efficient memory-improved proportionate affine projection sign algorithm based on l0-norm for sparse system identification," in *Foundations of Intelligent Systems, Advances in Intelligent Systems and Computing*, vol. 277, Springer, Berlin, Heidelberg, pp. 509–518, 2014.

[162] J. Liu and S.L. Grant, "Block sparse memory improved proportionate affine projection sign algorithm," *Electronics Letters*, vol. 51, no. 24, pp. 2001–2003, 2015.

[163] Y. Yu, H. Zhao and B. Chen, "Sparseness-controlled proportionate affine projection sign algorithms for acoustic echo cancellation," *Circuits Syst. Signal Process.*, vol. 34, no. 12, pp. 3933-3948, 2015.

[164] G. Gui, L. Xu, and N. Shimoi, "ROSA: Robust sparse adaptive channel estimation in the presence of impulsive noises," *CoRR, abs/1503.01105*, pp. 1-5, 2015.

[165] Y. Choi, "A new subband adaptive filtering algorithm for sparse system identification with impulsive noise," *J. Appl. Math*., PP. 1–7, 2014.

[166] M. O. Sayin, Y. Yilmaz, A. Demir, et al., "The Krylov-proportionate normalized least mean fourth approach: formulation and performance analysis," *Signal Process*., 109, pp. 1–13, 2015.

[167] Y. Wang, Y. Li and R. Yang, "A sparsity-aware proportionate normalized maximum correntropy criterion algorithm for sparse system identification in non-gaussian environment," *In Proceedings of the 25th European Signal Processing Conference (EUSIPCO)*, Kos Island, Greece, pp. 246–250, 2017.
[168] R.L. Das and M. Narwaria, "Lorentzian based adaptive filters for impulsive noise environments," *IEEE Trans. Circuits Syst. I, Reg. Pap*., vol. 64, no. 6, pp. 1529–1539, 2017.
[169] W. Ma, H. Qu, J. Zhao, B. Chen and J. C. Principe, "Sparsity aware minimum error entropy algorithms," *IEEE Int. Conf. on Acoustics, Speech, and Signal Process*., Brisbane, Queensland, Australia, pp. 19–24, 2015.
[170] W. Ma, H. Qu, G. Gui, J. Zhao, L. Xu and B. Chen, "Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments," *J. Frankl. Inst.-Eng. Appl. Math*., vol. 352, no. 7, pp. 2708–2727, 2015.
[171] W. Ma, J. Duan, B. Chen, G. Gui and W. Man, "Recursive generalized maximum correntropy criterion algorithm with sparse penalty constraints for system identification," *Asian J. Control*, vol. 19, no. 3, pp. 1164–1172, 2017.
[172] W. Ma, D. Zheng, Z. Zhang, J. Duan and B. Chen, "Robust proportionate adaptive filter based on maximum correntropy criterion for sparse system identification in impulsive noise environments," *Signal Image Video Process*., vol. 12, no. 1, pp. 117–124, 2018.
[173] T. Zhang, J. Dan and G. Gui, "IMAC: Impulsive-mitigation adaptive sparse channel estimation based on Gaussian-mixture model," *CoRR*, abs/1503.00800, 2014.
[174] G. Gui, H. Wei, N. Shimoi and L. Xu, "Regularization parameter selection method for sign LMS with reweighted L1-norm constriant algorithm," *CoRR*, abs/1503.03608, 2015.
[175] Md. Z. Ali Bhotto, M. O. Ahmad and M. N. S. Swamy, "Robust shrinkage affine-projection sign adaptive-filtering algorithms for impulsive noise environments," *IEEE Transactions on Signal Processing*, vol. 62, no. 13, pp. 3349-3359, 2014.
[176] S. Zhang, J. Zhang and H. Han, "Robust shrinkage normalized sign algorithm in an impulsive noise environment," *IEEE Trans. Circuits Syst. II Exp. Br*., vol. 64, no. 1, pp. 91-95, 2017.
[177] W. Ma, B. Chen, H. Qu and J. Zhao, "Sparse least mean p-power algorithms for channel estimation in the presence of impulsive noise," *Signal, Image and Video Process*., vol. 10, no. 3, pp. 503-510, 2016.
[178] G. Gui, L. Xu, W. Ma and B. Chen, "Robust adaptive sparse channel estimation in the presence of impulsive noises," *in IEEE International Conference on Digital Signal Processing*, pp. 628–632, 2015.
[179] Y. Li, Z. Jin, Y. Wang and R. Yang, "A robust sparse adaptive filtering algorithm with a correntropy induced metric constraint for broadband multi-path channel estimation," *Entropy*, vol. 18, no. 10, pp. 1–14, 2016.
[180] T. Zhang and G. Gui, "Sign Function Based Sparse Adaptive Filtering Algorithms for Robust Channel Estimation under Non-Gaussian Noise Environments," *Algorithms*, 9, 54, pp. 1-18, 2016.

[181] Z. Wu, S. Peng, W. Ma, B. Chen and J. C. Principe, "Minimum Error Entropy Algorithm with Sparsity Penalty Constraints," *Entropy*, 17, pp. 3419–3437, 2015.

[182] G. Gui and F. Adachi, "Improved least mean square algorithm with application to adaptive sparse channel estimation," *EURASIP Journal on Wireless Communications and Networking*, 2013:204, pp. 1-18, 2013.

[183] J. Arenas-García, A. R. Figueiras-Vidal and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Trans. Signal Process.*, vol. 54, pp. 1078–1090, 2006.

[184] N. J. Bershad, J. C. M. Bermudez and J. Y. Tourneret, "An Affine Combination of Two LMS Adaptive Filters—Transient Mean-Square Analysis," *IEEE Trans. on Signal Processing*, vol. 56, no..5, pp. 1853–1864, 2008.

[185] Y. Kamatham, B. Kinnara and M. K. Kartan, "Mitigation of GPS multipath using affine combination of two LMS adaptive filters," *Proc. of International Conf. on Signal Processing, Informatics, Communication and Energy Systems*, Feb. 2015.

[186] K. P. Seng, Z. Man and H. R. Wu, "Lyapunov-theory-based radial basis function networks for adaptive filtering," *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 49, no. 8, pp. 1215–1220, 2002.

[187] Z. Man, H. R. Wu, S. Liu and et al., "A new adaptive backpropagation algorithm based on Lyapunov stability theory for neural networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1580–1581, 2006.

[188] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE Journal of Oceanic Engineering*, vol. 32, no. 4, pp. 927–939, 2007.

[189] S. Y. Shih and K. C. Chen, "Compressed sensing construction of spectrum map for routing in cognitive radio networks," *Wireless Communication and Mobile Computing*, vol. 12, no. 18, pp. 1592–1607, 2012.

[190] L. Dai, Z. Wang and Z. Yang, "Compressive sensing based time domain synchronous OFDM transmission for vehicular communications," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 460–469, 2013.

[191] G. Gui, A. Mehbodniya and F. Adachi, "Bayesian sparse channel estimation and data detection for OFDM communication systems," *In Proceedings 2013 IEEE 78th Vehicular Technology Conference (VTC-Fall)*, LasVegas (USA), pp. 1–5, 2013.

[192] J. Arenas-García, M. Martínez-Ramón, V. Gómez-Verdejo, et al., "Multiple plant identifier via adaptive LMS convex combination," *In Proceedings IEEE International Symposium on Intelligent Signal Processing*, Budapest (Hungary), pp. 137–142, 2003.

[193] V. .C. Gogineni, M. Chakraborty, "Adaptive convex combination of APA and ZA-APA algorithms for sparse system identification," 19 pages. [Online] Cited 2015-09-10. Available at: https://arxiv.org/pdf/1509.03203.pdf.

[194] S. Radhika and S. Arumugam, "Steady state mean square analysis of convex combination of ZA-APA and APA for acoustic echo cancellation," *In Proceedings of International Symposium on Intelligent Systems Technologies and Applications*, pp. 437–446, 2015.

[195] M. Butsenko and T. Trump, "An affine combination of adaptive filters for sparse impulse response identification," *In 23rd Telecommunications Forum (TELFOR)*, pp. 396–399, 2015.

[196] M. Butsenko and T. Trump, “An affine combination of adaptive filters for channels with different sparsity levels,” *Telfor Journal*, vol. 8, no. 1, pp. 32–37, 2016.

[197] B. Mulgrew and C. F. N. Cowan, *Adaptive Filters and Equalizers*, MA (USA): Kluwer Academic Publishers, 1988.

[198] J. J. E. Slotine and W. Li, *Applied Nonlinear Control*, NJ (USA): Prentice Hall, 1991.

[199] M. Fazel, Matrix rank minimization with applications. Ph.D. thesis, Electrical Engineering Department, Stanford University, vol. 54, pp. 1-130, 2002.

[200] M. Fazel, H. Hindi and S.P. Boyd, “Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices,” *Proceedings of the American Control Conference*, Denver, Colorado (USA), pp. 2156–2162, 2003.

[201] M. M. Sondhi, “The history of echo cancellation,” *IEEE Signal Process. Mag*., vol. 23, no. 5, pp. 95-102, 2006.

[202] D. P. Mandic, E. V. Papoulis and C. G. Boukis, “A normalized mixed-norm adaptive filtering algorithm robust under impulsive noise interference,” *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’03)*, pp. 333-336, April 2003.

[203] C. Boukis, D. P. Mandic and A. G. Constantinides, “A class of stochastic gradient algorithms with exponentiated error cost functions,” Digital Signal Processing, vol. 19, no. 2, pp. 201-212, 2009.

[204] E. V. Papoulis and T. Stathaki, “A normalized robust mixed-Norm adaptive algorithm for system identification,” *IEEE Signal Proc. Lett*., vol. 11, no. 1, pp. 56-59, 2004.

[205] J. Chambers and A. Avlonitis, “A robust mixed-norm adaptive filter algorithm,” *IEEE Signal Proc. Lett.*, vol. 4, no.2, pp. 46-48, 1997.

[206] K. Shi and P. Shi, “Convergence analysis of sparse LMS algorithms with l1-norm penalty,” *Signal Process*., vol. 90, no. 12, pp. 3289-3293, 2010.

[207] O. Taheri and S. A. Vorobyov, “Reweighted ℓ_1-norm penalized LMS for sparse channel estimation and its analysis,” *Signal Process*., vol. 104, pp. 70–79, Nov. 2014.

[208] G. Gui, W. Peng and F. Adachi, “Sparse least mean fourth algorithm for adaptive channel estimation in low signal-to-noise ratio region,” *Int. J. Commun. Syst*. vol. 27, no. 11, pp. 3147-3157, 2014.

[209] G. Gui, L. Xu and F. Adachi, “RZA-NLMF algorithm-based adaptive sparse sensing for realizing compressive sensing,” *EURASIP Journal on Advances in Signal Process*., 2014:125, pp. 1-10, 2014.

[210] Y. Chen, G. Gui, L. Xu and N. Shimoi, “Improved Adaptive Sparse Channel Estimation Using Re-Weighted L1-norm Normalized Least Mean Fourth Algorithm,” *54th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pp. 689-694, July 2015.

[211] S. Seth and J. C. Principe, “Compressed signal reconstruction using the correntropy induced metric,” *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3845–3848, 2008.

[212] G. Gui, L. Dai, B. Zheng, L. Xu and F. Adachi, “Correntropy Induced Metric Penalized Sparse RLS Algorithm to Improve Adaptive System Identification,” *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Nanjing, China, May 2016.

[213] Y. Wang, Y. Li, F. Albu and R. Yang, “Sparse Channel Estimation Using Correntropy Induced Metric Criterion Based SM-NLMS Algorithm,” *In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, San Francisco, USA, pp. 1–6, March 2017.

[214] Z. Jin, Y. Li and Y. Wang, “An Enhanced Set-Membership PNLMS Algorithm with a Correntropy Induced Metric Constraint for Acoustic Channel Estimation,” Entropy, 19(6), pp. 1-14, 2017.

[215] J. Lin, M. Nassar and B. L. Evans, “Impulsive noise mitigation in powerline communications using sparse Bayesian learning,” *IEEE J. Sel. Areas in Commun*., vol. 31, no. 7, pp. 1172-1183, 2013.

[216] M. A. Z. Raja and N. I. Chaudhary, “Two-stage fractional least mean square identification algorithm for parameter estimation of CARMA systems,” *Signal Processing*, 107, pp. 327-339, 2015.

[217] M. S. Aslam, N. I. Chaudhary and M. A. Z. Raja, “A sliding-window approximation-based fractional adaptive strategy for Hammerstein nonlinear ARMAX systems,” *Nonlinear Dynamics*, vol. 87, no. 1, pp. 519-533, 2017.

[218] N. I. Chaudhary, S. Zubair, M. A. Z. Raja, “A new computing approach for power signal modeling using fractional adaptive algorithms,” *ISA transactions*, 68, pp. 189-202, 2017.

[219] N. I. Chaudhary, M. Ahmed, et al., “Design of normalized fractional adaptive algorithms for parameter estimation of control autoregressive autoregressive systems,” *Applied Mathematical Modelling*, 55, pp. 698-715, 2018.

[220] F. Huang, J. Zhang and S. Zhang, “A family of robust adaptive filtering algorithms based on sigmoid cost,” *Signal Processing*, 2018.

[221] Y. Chen, J. Tian and Y. Liu, “Variable step size LMS algorithm based on modified sigmoid function,” *2014 International Conference on Audio, Language and Image Processing*, Shanghai, pp. 627-630, 2014.

[222] S. C. Douglas and T. H. Y. Meng, “An adaptive edge detection method using a modified sigmoid-LMS algorithm,” *Twenty-Third Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, California, USA, pp.252-256, 1989.

[223] J. C. Principe, N. R. Euliano and W. C. Lefebvre, *Neural and Adaptive Systems: Fundamentals through Simulations*. New York: John Wiley & Sons, 2000.

[224] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ: Pearson Education, 2009.

[225] T. Shao, Y. Zheng and J. Benesty, “An affine projection sign algorithm robust against impulsive interferences,” *IEEE Signal Processing Letters*, vol. 17, no. 4, pp. 327–330, 2010.

[226] J. Shin, J. Yoo and P. Park, “Variable step-size affine projection sign algorithm,” *Electronic Letters*, vol. 48, no. 9, pp. 483–485, 2012.

[227] J. Yoo, J. Shin and P. Park, “Variable step-size affine projection sign algorithm,” *IEEE Transactions on Circuits and Systems—II: Express Briefs*, vol. 61, no. 4, pp. 274–278, 2014.

[228] A. H. Sayed, *Adaptive Filters*, John Wiley & Sons: Hoboken, New Jersey, USA, 2008.

[229] T. Zhang and G. Gui, "Reweighted Factor Selection for SLMS-RL1 Algorithm under Gaussian Mixture Noise Environments," *Algorithms*, vol. 8, no. 4, pp. 799-809, 2015.

[230] Y. Li, Y. Wang and R. Yang, "An Improved Maximum Correntropy Criterion Algorithm with Flexible Zero Attractor for Sparse Channel Estimation in Mixed Gaussian Noise Environment," *2017 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, pp. 316-319, 2017.

[231] Y. Li, Y. Wang, R. Yang and F. Albu, "A Soft Parameter Function Penalized Normalized Maximum Correntropy Criterion Algorithm for Sparse System Identification," *Entropy*, vol. 19, pp. 1-16, 2017.

[232] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class A and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1129–1149, 1999.

[233] L. R. Vega, H. Rey, J. Benesty and S. Tressens, "A new robust variable step-size NLMS algorithm," *IEEE Trans. Signal Process*., vol. 56, no.5, pp. 1878–1893, 2008.

[234] B. Picinbono, "On circularity," IEEE Transactions on Signal Processing, 42(12), pp. 3473–3482, 1994.

[235] T. Adali, P. J. Schreier and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," IEEE Trans. Signal Process., vol. 59, no. 11, pp. 5101–5125, 2011.

[236] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, Wiley, 2009.

LIST OF PUBLICATIONS

International Journals

1**. Pogula Rakesh, T. Kishore Kumar, “Adaptive Affine Combination of Two IPNLMS Filters for Robust Sparse Echo Cancellation”, **Journal of Telecommunication, Electronic and Computer Engineering (JTEC)**, PRINT ISSN: 2180-1843, eISSN: 2289-8131, Vol. 9 No. 1-3, pp.35-39, 2017. **[SCOPUS Journal]**

2**. Rakesh Pogula, T. Kishore Kumar, Felix Albu, “Novel Sparse Algorithms based on Lyapunov Stability for Adaptive System Identification”, **Radioengineering**, Vol. 27(1), pp. 270-280, Apr. 2018. doi: 10.13164/re.2018.0270 **[SCI Journal, Impact factor: 1.048]**

3**. Rakesh Pogula, T. Kishore Kumar, Felix Albu, “Robust Sparse Normalized LMAT Algorithms for Adaptive System Identification Under Impulsive Noise Environments”, **Circuits, Systems, and Signal Processing (CSSP)**, pp. 1-32, 2019. doi: https://doi.org/10.1007/s00034-019-01111-3 **[SCI Journal, Springer, Impact factor: 1.998]**

International Conferences

1*. Pogula Rakesh, Dr. T. Kishore Kumar, “An Affine Combination of Two Adaptive Filters for System Identification with Variable Sparsity”, **2016 5th International Conference on Advances in Computing, Communications and Informatics (ICACCI)**, Jaipur, India, pp. 281-286, 21-14 Sept., 2016. doi: 10.1109/ICACCI.2016.7732060 **[IEEE Xplore]**

2*. Pogula Rakesh, T. Kishore Kumar, Felix Albu, “Complex Domain Adaptive System Identification Using Sparse Affine Projection Normalized Correlation Algorithms Under Impulsive Noises”, in Proc. of the **12th International Conference on Communications** (**COMM 2018**), Bucharest, Romania, pp. 61-66, 14-16 June, 2018. doi: 10.1109/ICComm.2018.8484791 **[IEEE Xplore]**

3*. Rakesh Pogula, T. Kishore Kumar, Felix Albu, “Modified Least-Mean Mixed-Norm Algorithms For Adaptive Sparse System Identification Under Impulsive Noise Environment”, **2019 42nd International Conference on Telecommunications and Signal Processing (TSP)**, Budapest, Hungary, pp. 557-561, 1-3 July, 2019. **[IEEE]**