ROBUST ADAPTIVE FILTERING ALGORITHMS FOR IMPULSIVE NOISE ENVIRONMENTS

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND
ELECTRONICS ENGINEERING
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Gül Aydın
July 1996

ROBUST ADAPTIVE FILTERING ALGORITHMS FOR IMPULSIVE NOISE ENVIRONMENTS

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND

ELECTRONICS ENGINEERING

AND THE INSTITUTE OF ENGINEERING AND SCIENCES

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

By

Gül Aydın

July 1996

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. A. Enis Çetin(Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Orhan Arıkan

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Billur Barshan

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

ROBUST ADAPTIVE FILTERING ALGORITHMS FOR IMPULSIVE NOISE ENVIRONMENTS

Gül Aydın
M.S. in Electrical and Electronics Engineering
Supervisor: Assoc. Prof. Dr. A. Enis Çetin
July 1996

In this thesis, robust adaptive filtering algorithms are introduced for impulsive noise environments which can be modeled as α-stable distributions and/or ε-contaminated Gaussian distributions. The algorithms are developed using the Fractional Lower Order Statistics concept. Robust performance is obtained.

Keywords : Adaptive Filtering, α-stable distributions, Fractional Lower Order Moments, ε-contaminated Gaussian distributions

ÖZET

PATLAMALI GÜRÜLTÜLÜ ORTAMLAR İÇİN DAYANIKLI SÜZGEÇLEME ALGORİTMALARI

Gül Aydın
Elektrik ve Elektronik Mühendisliği Bölümü Yüksek Lisans
Tez yöneticisi: Assoc. Prof. Dr. A. Enis Çetin
Temmuz 1996

Bu tezde, patlamalı gürültülerin bulunduğu ortamlarda çalışabilecek dayanıklı süzgeçleme algoritmaları geliştirilmiştir. Bu algoritmalar geliştirilirken patlamalı gürültüler kararlı dağılımlarla ve/veya ε 'la kirletilmiş Gauss dağılımı ile modellenmiştir. Tanıtılan algoritmalarda Kesirsel Düşük Derece İstatistiği (KDDİ) kullanılmıştır. Sonuçta dayanıklı performance elde edilmiştir.

Anahtar Kelimeler : Süzgeçleme Algoritmaları, Gauss olmayan dağılımlar, Kararlı dağılımlar, ε'la kirletilmiş Gauss dağılımı, Kesirsel Düşük Derece İstatistiği.

ACKNOWLEDGMENTS

I would like to thank Dr. A. Enis Çetin, Dr. Orhan Arıkan and Dr. Billur Barshan for their supervision, guidance, suggestions and encouragement throughout the development of this thesis.

Many thanks also to *Birsel and Lütfiye* for their valuable ideas and help.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

The Gaussian distribution has been by far the most popular statistical distribution in signal processing. This distribution has been widely used to model the noise in the analysis of signals corrupted during transmission or measurement. Many theorems in the fields of communications, speech analysis, estimation and detection have used Gaussian noise assumption. Having no a priori information about the noise process, this is a well justified assumption due to the "Central Limit Theorem" [1]. The second reason that makes a Gaussian noise assumption attractive is its analytical tractability. The final reason is that Gaussian processes are easy to describe, i.e., only the first two moments are sufficient to characterize a Gaussian process.

Unfortunately, there are also many phenomena in signal processing which are decidedly non-Gaussian [2]-[4]. Some examples of non-Gaussian noise are atmospheric noise [5]-[7], underwater acoustic noise [8], noise on telephone lines [9, 10], low frequency electromagnetic interference [11] and degradations on aged audio recordings [12]. The common property of most of these type of noises is that they show impulsive behaviour, that is they produce large-amplitude outliers much more frequently than Gaussian noise. These high amplitude samples correspond to very low probability parts of the Gaussian probability density function, therefore they are not likely to be the samples from a Gaussian probability density function. This indicates that impulsive noises can only be represented by heavier tailed probability density functions than Gaussian distribution.

Several attempts were made in developing a statistical model for impulsive noise [4, 5], [12]-[16]. In this thesis, we concentrate on impulsive type noise using the statistical model of α-stable distributions developed by Nikias and Shao [17] and of ε-contaminated Gaussian distributions [3]. We develop some adaptive filtering algorithms.

An adaptive filter is a learning and self designing system which adjusts its parameters by a recursive algorithm to optimize some performance criteria. In this thesis, robust adaptive filtering algorithms are developed for impulsive noise environments.

A brief outline of the chapters in this thesis is given below :

In Chapter 2, α-stable distributions and some of their properties will be presented. We also give information about ε-contaminated Gaussian distributions.

In Chapter 3, a family of the NLMS algorithms [18] are investigated. The performance of these algorithms are examined in impulsive noise environments for system identification both theoretically and experimentally. We also study the performance under additive observation noise. We observe that they have a very poor performance in impulsive noise environments.

In Chapter 4, we introduce a new family of adaptive filtering algorithms based on the Fractional Lower Order Statistics (FLOS) concept. Some properties of the proposed FLOS based algorithms are investigated. We also present other existing FLOS based adaptive filtering algorithms for impulsive noise environments and compare their performance.

In Chapter 5, we modify the Least Mean Mixed Norm algorithms using the FLOS. We investigate the properties of the LMMN algorithms in impulsive noise environments. Then the proposed family of the algorithms is investigated. Simulation studies will be presented.

In Chapter 6, proposed algorithms are compared and directions for the future work are addressed.

Chapter 2

SOME HEAVY TAILED DISTRIBUTIONS

2.1 Introduction

In this chapter, we introduce α-stable distributions and ε-contaminated Gaussian distributions. We review their characteristics and statistical properties. The proofs of most of the results are omitted but can be found in probability and statistics literature such as [1],[20]-[26].

2.2 α-Stable Distributions

The characteristic function of α-stable distributions are given by :

$$\phi(t) = \exp\{iat - \gamma|t|^{\alpha}[1 + i\beta \text{sign}(t) w(t, \alpha)]\} \tag{2.1}$$

where

$$w(t,\alpha) = \begin{cases} \tan\left(\frac{\alpha\pi}{2}\right) & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\log|t| & \text{if } \alpha = 1. \end{cases} \tag{2.2}$$

and

$$-\infty < a < \infty, \ \ \gamma > 0, \ \ 0 < \alpha \leq 2, \ \ -1 \leq \beta \leq 1. \tag{2.3}$$

In the above expressions, a is the location parameter, γ is the dispersion, β is the index of skewness and α is the characteristic exponent. When $\beta = 0$, the distribution is called symmetric α-stable (SαS). The characteristic exponent, α, controls the tails of the distribution. For $0 < \alpha < 2$, the distributions have algebraic tails which are significantly heavier than the exponential tail of the Gaussian distribution. The smaller the value of the α, the heavier the tails of the distributions. When $\alpha = 2$, the relevant stable distribution is Gaussian. Cauchy distribution is also an α-stable distribution with $\alpha = 1$ and $\beta = 0$.

The α-stable distributions are called *standard* when $a = 0$ and $\gamma = 1$. Then, one can see that if X is an α-stable random variable with parameters α, β, γ and a then $(X - a)/\gamma^{\frac{1}{\alpha}}$ is standard with characteristic exponent α and skewness β.

By taking the inverse Fourier transform of the characteristic function, it is easy to show that the standard stable density function is given by

$$f(x; \alpha, \beta) = \frac{1}{\pi} \int_0^\infty \exp(-t^\alpha) \cos[xt + \beta t^\alpha w(t, \alpha)] dt. \tag{2.4}$$

Note that $f(x; \alpha, \beta) = f(-x; \alpha, -\beta)$. It can also be shown that the probability density functions of the α-stable distributions are bounded and have derivatives of arbitrary orders [21]. Unfortunately, no closed-form expressions exist for the general α-stable density and distribution functions, except for the Gaussian ($\alpha = 2$), Cauchy ($\alpha = 1, \beta = 0$), and Pearson ($\alpha = \frac{1}{2}, \beta = -1$) distributions [26]. But power series expansions of their density functions are available. The standard α-stable density function can be expanded into convergent power series as follows [1], [21, 22], [26, 27]. For $x > 0$,

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) \left(\frac{x}{r}\right)^{-\alpha k} \sin\left[\frac{k\pi}{2}(\alpha + \xi)\right] & \text{for } 0 < \alpha < 1 \\ \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\frac{k}{\alpha} + 1) \left(\frac{x}{r}\right)^{k} \sin\left[\frac{k\pi}{2\alpha}(\alpha + \xi)\right] & \text{for } 1 < \alpha \leq 2 \end{cases}$$

where

$$\eta = \beta \tan(\pi\alpha/2), \tag{2.5}$$

$$r = (1 + \eta^2)^{-1/(2\alpha)} \tag{2.6}$$

$$\xi = -(2/\pi) \arctan(\eta). \tag{2.7}$$

The Γ is the usual gamma function defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt. \tag{2.8}$$

In particular, the standard SαS density function is given by

$$f_\alpha(x) = \begin{cases} \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1)(x)^{-\alpha k} \sin\left[\frac{k\pi\alpha}{2}\right] & \text{for } 0 < \alpha < 1 \\ \frac{1}{\pi(x^2+1)} & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k!} \Gamma(\frac{2k+1}{\alpha}) x^{2k} & \text{for } 1 < \alpha < 2 \\ \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}} & \text{for } \alpha = 2. \end{cases}$$

The standard SαS density functions for a few values of the characteristic exponent α are shown in Figure 2.1. Observe that SαS densities have many features of the Gaussian density. They are smooth, unimodal, symmetric with respect to the median and bell-shaped. A detailed comparison between the standard normal and SαS density functions shows that non-Gaussian stable density functions depart from the corresponding Gaussian density in the following ways. For small absolute values of x, the SαS densities are more peaked than the normal. For some intermediate range of $|x|$, the SαS distributions have lower densities than the normal. Most importantly, unlike the Gaussian density which has exponential tails, the α-stable densities have algebraic tails [23]. Thus, the SαS densities have heavier tails than the Gaussian density [see Figure 2.1 (b)].

2.2.1 Basic Properties of α-Stable Distributions

Two of the most important properties of the α-stable distributions are the *Stability Property* and the *Generalized Central Limit Theorem.* They are responsible for much of the appeal of the stable distribution as a statistical model of uncertainty.

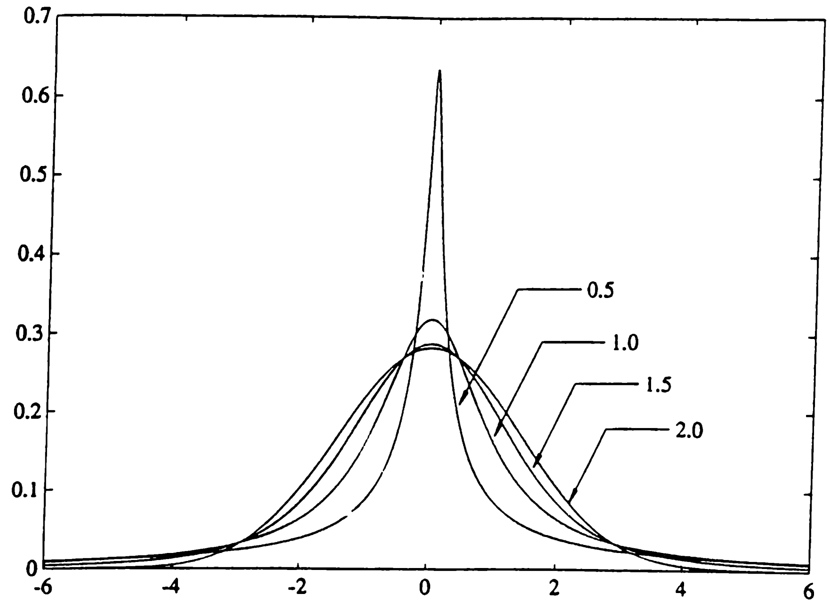
The stability property is actually a defining characteristic of α-stable distributions.

Theorem 1 (Stability Property) : A random variable X is α-stable if and only if for any independent random variables X_1, X_2 with the same distribution as X, and for arbitrary constants a_1, a_2, there exist constants a and b such that

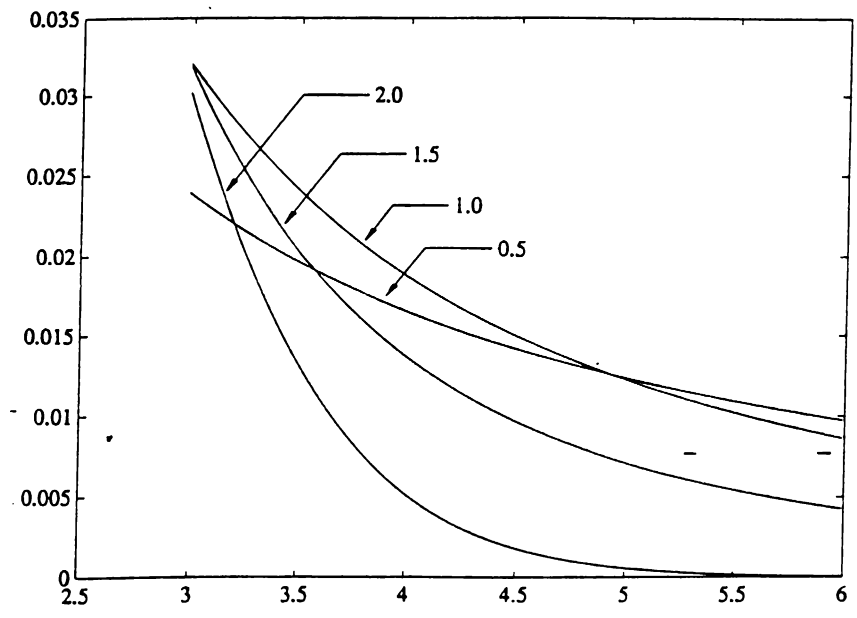
$$a_1X_1 + a_2X_2 \stackrel{d}{=} aX + b \tag{2.9}$$

where the notation $X \stackrel{d}{=} Y$ means that X and Y have the same distributions. By using the characteristic function of the α-stable distributions, one can easily show a more general statement : if $X_1, X_2, ..., X_n$ are independent and have α-stable distribution functions with the same (α, β), then all the linear combinations of the form $\sum_{j=1}^{n} a_jX_j$ are α-stable with the same parameters α and β.

As a consequence of the stability property, it can be shown that α-stable distributions are the only possible limit distributions for sums of independent identically distributed (i.i.d.) random variables. This is known as the Generalized Central Limit Theorem and is formally stated as follows [20]:



(a)



(b)

Figure 2.1: Density functions of SαS distributions for different values of characteristic exponent α :(a) the overall densities and (b) the tails of the densities.

Theorem 2 (Generalized Central Limit Theorem) : X is the limit in distribution of normalized sums of the form

$$S_n = (X_1 + ... + X_1)/a_n - b_n \tag{2.10}$$

where $X_1, X_2, ...,$ are i.i.d. and $a_n \to \infty$, if and only if X is stable.

In particular, if X_i's are i.i.d. and have finite variances then the limiting distribution is Gaussian. This is of course the result of the ordinary Central Limit Theorem.

The main cause of different behaviours of the Gaussian and α-stable distributions is their tails. It can be shown [28], [29] that for α-stable random variable X with zero location parameter and dispersion γ,

$$\lim_{t\to\infty} t^{\alpha} P(|X| > t) = \gamma C(\alpha) \tag{2.11}$$

where $C(\alpha)$ is a positive constant depending on α. The α-stable distributions have inverse power (i.e. :algebraic) tails while Gaussian distribution has exponential tails. This fact shows that the tails of α-stable distributions are much heavier than the tails of the Gaussian distributions.

Equation (2.11) has an important consequence that the second-order moment of α-stable distributions, except for the limiting case $\alpha = 2$, does not exist. This can be written as in the following proposition [17]:

Proposition : Let X be an α-stable random variable. If $0 < \alpha < 2$ then

$$\mathbf{E}[|X|^p] = \infty, \quad \text{if } \; p \geq \alpha \tag{2.12}$$

and

$$\mathbf{E}[|X|^p] < \infty, \quad \text{if } \; 0 \leq p < \alpha. \tag{2.13}$$

If $\alpha = 2$, then

$$\mathbf{E}[|X|^p] < \infty, \quad \text{for all } \; p \geq 0. \tag{2.14}$$

Therefore, α-stable distributions have no finite first or higher-order moments for $0 < \alpha \leq 1$; they have finite first-order moments and all the fractional moments of order p for $1 < \alpha < 2$ where $p < \alpha$; and all the moments exist for $\alpha = 2$. Note also that α-stable distributions have infinite variances.

2.3 Symmetric α-Stable Random Variables and Processes

A real random variable (r.v.) X is SαS, if its characteristics function is of the form :

$$\phi(t) = \exp\{iat - \gamma|t|^\alpha\} \tag{2.15}$$

where $0 < \alpha \leq 2$ is the characteristic exponent, $\gamma > 0$ is the dispersion, and $-\infty < a < \infty$ is the location parameter. When $\alpha = 2$, X is Gaussian and when $\alpha = 1$, X is Cauchy.

2.3.1 Fractional Lower-Order Moments

Although the second-order moment of a SαS random variable with $0 < \alpha < 2$ does not exist, all the moments of order less than α do exist and are called the *fractional lower-order moments* or FLOM's. The FLOM's of a SαS random variable can be easily found from its dispersion and characteristic exponent as follow.

Theorem 3 : Let X be a SαS r.v. with zero location parameter and dispersion γ. Then,

$$\mathbf{E}[|X|^p] = C(p,\alpha)\gamma^{p/\alpha} \tag{2.16}$$

for $0 < p < \alpha$, where

$$C(p,\alpha) = \frac{2^{p+1}\Gamma(\frac{p+1}{2})\Gamma(-p/\alpha)}{\alpha\pi\Gamma(-p/2)} \tag{2.17}$$

depends only on α and p, not on X. In this expression Γ is defined in (2.8).

This important result was first proved by Zolotarev using the Mellin-Stieljes transform [30]. Cambanis and Miller rediscovered [31] it by using a property of the characteristic function derived in [32]. An elementary proof of the theorem using basic properties of the gamma function is given in [17].

A fundamental difficulty in stable signal processing with lower-order moments is that the tools of the Hilbert space theory are no longer applicable. Although the linear space of a Gaussian process is a Hilbert space, the linear space of the α-stable distributions is a Banach space for $1 \leq \alpha < 2$ and only a metric space for $0 < \alpha < 1$ [17], [33].

2.4 ε-contaminated Gaussian Distributions

The ε-contaminated Gaussian mixture density has the probability density function of $(1-\varepsilon)N(0,\sigma^2)+\varepsilon N(0,\lambda\sigma^2)$. This family of the distribution is characterized by the mixing parameter ε which is the fraction of the contamination. The outliers of the distribution are also Gaussian but have λ times the variance of the dominant distribution, resulting in an impulsive behaviour.

A typical realization of ε-contaminated Gaussian distribution is given in the following figure. In this figure we take $\varepsilon = 0.1$, $\lambda = 10$ and $\sigma^2 = 1$.

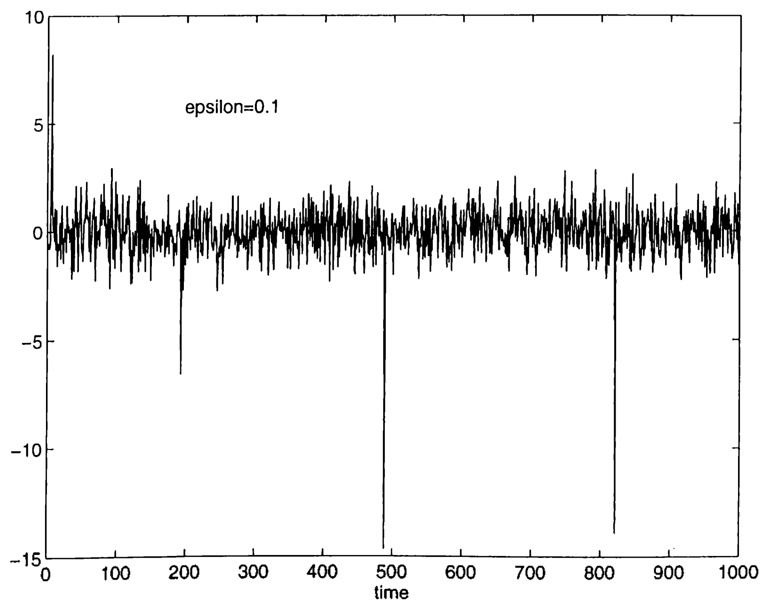


Figure 2.2: A typical realization of ε-contaminated Gaussian mixture.

2.5 Conclusion

In this chapter, we briefly review some heavy tailed distributions, namely α-stable distributions and ε-contaminated Gaussian distributions. Through the thesis we use these two classes of distributions for modeling impulsive noise environments.

Chapter 3

A FAMILY OF NORMALIZED LEAST MEAN SQUARE ALGORITHM

3.1 Introduction

Normalized Least Mean Square (NLMS) algorithm is first developed by Nagumo and Noda [36] and Albert and Gardner [37] independently. This algorithm is also called the projection algorithm [35]. In [36], a modified version of the NLMS [36] is introduced. Properties of the NLMS algorithm are studied in [34] and in [18] a generalized family of the NLMS algorithms are presented.

In this chapter, we first review the above algorithms with some of their properties. Then using the adaptive filtering configuration of Figure 3.1 we investigate the performance of the NLMS type algorithms in impulsive noise environments.

3.2 NLMS Algorithm

The NLMS algorithm of [36] and [37] has the following update equation :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k}{\sum_{m=0}^{M-1} x_{k-m}^2} \mathbf{X}_k \tag{3.1}$$

where $\mathbf{W}_k = [w_{0,k} \dots w_{M-1,k}]^T$ are the tap weights of the adaptive filter at time k, $\mathbf{X}_k = [x_k \dots x_{k-M+1}]^T$ are the M samples of the input data in filter memory at time k, $e_k = d_k - \mathbf{W}_k{}^T \mathbf{X}_k$ is the error between the adaptive filter output and the desired signal d_k, and μ is the step size which should be appropriately determined.

A variety of the theoretical results for NLMS algorithm such as conditions for convergence, rates of convergence and the effects of errors due to digital implementation of the algorithm are given in [38, 39].

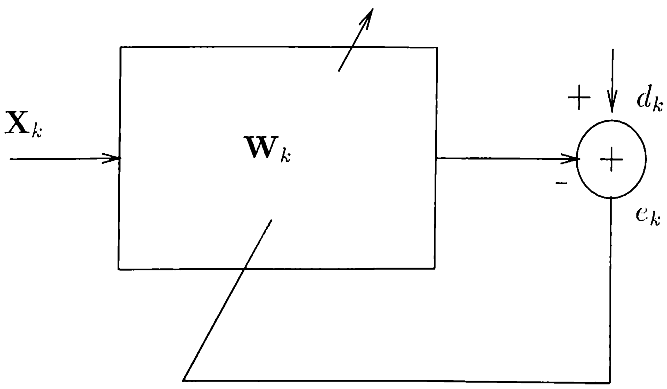


Figure 3.1: Adaptive filtering block diagram

3.2.1 Performance of NLMS Algorithm in Impulsive Noise Environments

Consider an AR(L) α-stable process, defined as follows,

$$x_k = \sum_{i=1}^{L} a_i x_{k-i} + u_k \tag{3.2}$$

where u_k is a symmetric α-stable (SαS) sequence of random variables. The random variable x_k is also symmetric α-stable (SαS) with the same parameters as u_k [17], [40] if $\{a_i\}$ is an absolutely summable sequence.

The NLMS algorithm of Equation (3.1) is used to identify an AR system driven by an i.i.d. SαS random process, u_k, with three different values of the α parameter as well as ε-contaminated Gaussian noise. AR(2) process has the parameters $a_1 = 0.99$ and $a_2 = -0.1$. The tap weight values versus time plots are given in Figure 3.2. We used $\alpha = 1.1$, 1.2 and 1.5 values for the $S\alpha S$ process and standard Gaussian random process $N(0,1)$ contaminated by another Gaussian process $N(0,10)$ with contamination rate $\varepsilon = 0.1$. As it can be seen from Figure 3.2, the performance of the algorithm is far from satisfactory. We repeat the experiment under additive observation noise. For various α values and ε-contaminated Gaussian random processes, the performance is seen in Figure 3.3. Again, we conclude that the performance is far from satisfactory.

Note that in this thesis all the simulations are obtained by averaging 100 independent trials of the experiment and for each trial, a different computer realization of the process u_k is used.

3.3 A Modified NLMS Algorithm

The NLMS algorithms requires a minimum number of one additional multiplication, division and addition over the usual LMS algorithm [41], which has the following update equation :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_k \mathbf{X}_k, \tag{3.3}$$

to implement for shift-input data. Even so, the multipliers required for the algorithm update may still be prohibitive in certain high-data-rate applications. In these situations, it is useful to determine modified versions of the NLMS algorithm of Equation (3.1) while reducing the computation per iteration. One such modified algorithm, first suggested by Nagumo and Noda [36], is :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k}{\sum_{m=0}^{M-1} |x_{k-m}|} \operatorname{sign}(\mathbf{X}_k). \tag{3.4}$$

This update is similar to that of Equation (3.1) but allows the nonlinear transformation of the input data vector elements.

3.3.1 Performance of the Modified NLMS Algorithm in Impulsive Environments

The same experiment of Figure 3.2 is performed for the modified NLMS algorithm of Equation (3.4). The plot is in Figure 3.4. The concluding remark is again that this algorithm is unsatisfactory in impulsive noise environments.

3.4 A Family of NLMS Algorithms

In [18] a family of NLMS algorithms is derived with the motivation of the NLMS and the modified NLMS algorithms mentioned above. The update equation is given by :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_k F_q(\mathbf{X}_k) \tag{3.5}$$

$$[F_q(\mathbf{X}_k)]_i = \begin{cases} \frac{|x_{k-i}|^{q-1}\mathrm{sign}(x_{k-i})}{\sum_{m=0}^{M-1}|x_{k-m}|^q} & \text{if } 1 \leq q < \infty \\ \\ \frac{1}{x_{k-n}}\delta_{i-n} & \text{if } q = \infty \end{cases} \tag{3.6}$$

where $[F_q(.)]_i$ denotes the i^{th} element of the vector-valued function $F_q(.)$, n is any one of the integers $0, 1, ..., M-1$ such that $|x_{k-n}| = \max_{0\leq j\leq M-1}|x_{k-j}|$, and δ_j is the Kronecker delta function. For $q = 2$, this algorithm reduces to the NLMS algorithm of Equation (3.1) and for $q = 1$, this update reduces to the modified NLMS algorithm of Equation (3.4). For $\mu = 1$, this algorithm is shown to be the solution of the following optimization problem [18] :

$$\text{minimize} \quad ||\mathbf{W}_{k+1} - \mathbf{W}_k||_p \tag{3.7}$$

$$\text{subject to} \quad d_k - \mathbf{W}_{k+1}^T\mathbf{X}_k = 0 \tag{3.8}$$

where $||.||_p$ denotes the L_p norm, and p satisfies the Hölder inequality $1/p + 1/q = 1$. Therefore, the adaptation algorithm of Equations (3.5) and (3.6) provides the minimum change in an L_p-norm sense of the tap weights to exactly satisfy the filtering relationship between the input data and the desired response at time k, similar to a projection in the L_2-norm case. Investigating the algorithm of Equations (3.5) and (3.6) for $q = \infty$, that is for L_1-norm case, the update equation takes the following form:

$$w_{i,k+1} = \begin{cases} w_{i,k} + \mu\frac{e_k}{x_{k-i}} & \text{if } |x_{k-i}| = \max_{0\leq j\leq M-1}|x_{k-j}| \\ \\ w_{i,k} & \text{otherwise.} \end{cases} \tag{3.9}$$

In this update equation, the maximum absolute value of the input data vector $|x_{k-i}| = \max_{0\leq j\leq M-1}|x_{k-j}|$ is supposed to be unique. In the case of being not unique, a single filter tap weight of the set $\{w_{i,k} : |x_{k-i}| = \max_{0\leq j\leq M-1}|x_{k-j}|\}$ is chosen randomly for updating. Therefore the only filter tap weight changed at time k is a tap weight associated with an input sample that has the largest absolute value of all input data samples currently in the filter memory.

3.4.1 Derivation

In this section we review the derivation of [18] in which it is shown that the algorithms of Equations (3.5) and (3.6) solve the optimization problem in Equations (3.7) and (3.8). This derivation follows a similar derivation of [42] for the modified NLMS algorithm and uses the following theorem :

Theorem : Let **A** be a nonzero vector contained in the vector space R^M, and b be a scalar quantity. Then, the minimum L_p-norm solution vector Z to a consistent linear equation $\mathbf{A}^T\mathbf{Z} = b$ is given by

$$\mathbf{Z} = bF_q(\mathbf{A}) \tag{3.10}$$

where the vector function $F_q(.)$ is given by Equation (3.6).

Proof : Let a_i and z_i denote the i^{th} elements of the vectors **A** and **Z**, respectively. Then

$$|b| = |\sum_{i=0}^{M-1} a_i z_i| \leq ||\mathbf{Z}||_p ||\mathbf{A}||_q \tag{3.11}$$

where the inequality follows from the Hölder inequality with $1/p + 1/q = 1$. Thus, for the nonzero vector **A**, we have

$$||\mathbf{Z}||_p \geq \frac{|b|}{||\mathbf{A}||_q}. \tag{3.12}$$

Consequently, the following inequality holds :

$$\min_{\mathbf{A}^T\mathbf{Z}=b} ||\mathbf{Z}||_p \geq \frac{|b|}{||\mathbf{A}||_q}. \tag{3.13}$$

Let $\tilde{\mathbf{Z}}$ be a solution vector to the equation $\mathbf{A}^T\mathbf{Z} = b$. Note that $\tilde{\mathbf{Z}}$ is not unique but that it satisfies

$$||\tilde{\mathbf{Z}}||_p \geq \min_{\mathbf{A}^T\mathbf{Z}=b} ||\mathbf{Z}||_p \tag{3.14}$$

for all $||\mathbf{Z}||$ in R^M. Now, let

$$\tilde{\mathbf{Z}} = bF_q(\mathbf{A}). \tag{3.15}$$

It can be seen that for $1 \leq q < \infty$

$$||\tilde{\mathbf{Z}}||_p = |b| \left(\frac{\sum_{i=0}^{M-1} |a_i|^{p(q-1)}}{||\mathbf{A}||_q^{pq}} \right)^{1/p} = \frac{|b|}{||\mathbf{A}||_q} \left(\frac{\sum_{i=0}^{M-1} |a_i|^{p(q-1)}}{||\mathbf{A}||_q^{p(q-1)}} \right)^{1/p} \tag{3.16}$$

Using the relationship $p = q/(q-1)$, the term inside the parenthesis of (3.16) can be shown to be equal to one. Thus, from (3.13) and (3.16), we have

$$||\tilde{\mathbf{Z}}||_p = \min_{\mathbf{A}^T\mathbf{Z}=b}||\mathbf{Z}||_p. \tag{3.17}$$

Considering the case ($p = 1, q = \infty$), it is found from (3.13) and (3.16) that

$$||\tilde{\mathbf{Z}}||_1 = \frac{|b|}{||\mathbf{A}||_\infty} = \min_{\mathbf{A}^T\mathbf{Z}=b}||\mathbf{Z}||_1. \tag{3.18}$$

Therefore, Equation (3.10) follows.

To see how the theorem enables the solution to the problem posed in (3.7) and (3.8), assign $\mathbf{Z} = \mathbf{W}_{k+1} - \mathbf{W}_k$, $\mathbf{A} = \mathbf{X}_k$, and $b = e_k$. Then, from the definition of the error e_k we have

$$e_k = d_k - \mathbf{X}_k^T\mathbf{W}_k = (d_k - \mathbf{X}_k^T\mathbf{W}_{k+1}) + \mathbf{X}_k^T(\mathbf{W}_{k+1} - \mathbf{W}_k). \tag{3.19}$$

If the constraint in Equation (3.8) is satisfied, then from the assignments of $\mathbf{Z}$, $\mathbf{A}$, and b

$$\mathbf{X}_k^T(\mathbf{W}_{k+1} - \mathbf{W}_k) = e_k \rightarrow \mathbf{A}^T\mathbf{Z} = b \tag{3.20}$$

and thus, the optimization problem in Equation (3.7) and (3.8) is the same as the minimization of $||\mathbf{Z}||_p$ subject to $\mathbf{A}^T\mathbf{Z} = b$. Therefore, from Equation (3.17), the optimum update for $\mathbf{W}_k$ is given by (3.5) and (3.6).

3.4.2 Variance Analysis for Impulsive Environments

In this section we will show that :

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \infty \tag{3.21}$$

for the family of the NLMS algorithms presented above, for both finite and infinite q.

Finite q case : Equation (3.5) and (3.6) can be written for each sample of $\mathbf{W}_k$ as follows :

$$w_{i,k+1} = w_{i,k} + \mu e_k \frac{|x_{k-i}|^{q-1}\text{sign}(x_{k-i})}{\sum_{m=0}^{M-1}|x_{k-m}|^q} \tag{3.22}$$

and

$$\mathbf{E}[(w_{i,k+1} - w_{i,k})^2] = \mu^2\mathbf{E}\left[\left(\frac{e_k|x_{k-i}|^{q-1}}{\sum_{m=0}^{M-1}|x_{k-m}|^q}\right)^2\right]. \tag{3.23}$$

If the algorithm converges at steady state, it can be assumed that the error, e_k, and the samples of the random process $\mathbf{X}_k$ are uncorrelated. Thus the right hand side of the last equation can be written as :

$$\mu^2 \mathbf{E}\left[\left(\frac{e_k |x_{k-i}|^{q-1}}{\sum_{m=0}^{M-1} |x_{k-m}|^q}\right)^2\right] = \mu^2 \mathbf{E}[{e_k}^2] \mathbf{E}\left[\left(\frac{|x_{k-i}|^{q-1}}{\sum_{m=0}^{M-1} |x_{k-m}|^q}\right)^2\right]. \tag{3.24}$$

In the following, we will show that $\mathbf{E}[{e_k}^2]$ in Equation (3.24) is infinite and the last expectation is strictly positive for at least one value of i, $0 \leq i \leq M-1$. This way, we will be able to conclude that the left side of Equation (3.24) is infinite for at least one value of i, hence the claim in Equation (3.21) is true. For this purpose, let us investigate the last expectation for the index j, which is chosen such that $|x_{k-j}| = \max_{0 \leq m \leq M-1} |x_{k-m}|$. Then, we have

$$\frac{|x_{k-j}|^{q-1}}{\sum_{m=0}^{M-1} |x_{k-m}|^q} \geq \frac{|x_{k-j}|^{q-1}}{M |x_{k-j}|^q} = \frac{1}{M |x_{k-j}|} \tag{3.25}$$

implying,

$$\mathbf{E}\left[\left(\frac{|x_{k-j}|^{q-1}}{\sum_{m=0}^{M-1} |x_{k-m}|^q}\right)^2\right] \geq \frac{1}{M^2} \mathbf{E}\left[\frac{1}{|x_{k-j}|^2}\right]. \tag{3.26}$$

Using Jensen's inequality for the last term we obtain :

$$\frac{1}{M^2} \mathbf{E}\left[\frac{1}{|x_{k-j}|^2}\right] > \frac{1}{M^2} \frac{1}{\mathbf{E}[|x_{k-j}|^2]} \tag{3.27}$$

where the right hand side is zero since $\mathbf{E}[|x_{k-j}|^2] = \infty$. So, we can say that the second expectation term in the right hand side of Equation (3.24) is strictly greater than zero.

As for the first expectation term, $\mathbf{E}[{e_k}^2]$, it includes some linear combinations of $\mathbf{E}[{x_{k-i}}^2]$. Knowing that the samples of $\mathbf{X}_k$ are SαS random variables, it is clear that $\mathbf{E}[{e_k}^2]$ has an infinite value. Therefore :

$$\mathbf{E}[(w_{j,k+1} - w_{j,k})^2] = \infty \tag{3.28}$$

implying

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \infty. \tag{3.29}$$

Infinite q case : In this case, Equation (3.9) can be rewritten as :

$$w_{i,k+1} = \begin{cases} w_{i,k} + \mu \frac{e_k}{x_{k-i}} & \text{if } |x_{k-i}| = \max_{0 \leq m \leq M-1} |x_{k-m}| \\ w_{i,k} & \text{otherwise} \end{cases} \tag{3.30}$$

and

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \mu^2 \mathbf{E}\left[\frac{e_k^2}{x_{k-i}^2}\right]. \tag{3.31}$$

Again assuming that the error and the samples of the input vector are uncorrelated and the system converges, this last equation can be expressed as

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \mu^2 \mathbf{E}[{e_k}^2]\mathbf{E}\left[\frac{1}{{x_{k-i}}^2}\right]. \tag{3.32}$$

Using the Jensen's inequality, [44], the second expectation term in the right hand side of Equation (3.32) may be written as :

$$\mathbf{E}\left[\frac{1}{{x_{k-i}}^2}\right] > \frac{1}{\mathbf{E}[{x_{k-i}}^2]}. \tag{3.33}$$

Since $\mathbf{E}[{x_{k-i}}^2] = \infty$, we can say that the left hand side of Equation (3.33) is strictly greater than zero.

Similar to the finite q case, $\mathbf{E}[{e_k}^2]$ is infinite since it includes some linear combinations of $\mathbf{E}[{x_{k-i}}^2]$. Therefore, we can conclude that

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \infty. \tag{3.34}$$

It can be also shown that in the case of Gaussian excitation and α-stable observation noise, i.e., x_k is a Gaussian AR sequence and $d_k = \mathbf{W}_k^T\mathbf{X}_k + n_k$ where n_k is α-stable, the variance of the update term of Equation (3.22) and Equation (3.30) is not finite, either.

3.4.3 Performance Analysis of the Family of NLMS Algorithms in Impulsive Environments

We repeat the same experiment for the family of the NLMS algorithm of Equations (3.5) and (3.6) as in Section 3.3.1 for a particular value of q and see from the Figure 3.5 that this algorithm is also far from satisfactory when the α-stable distributions are used.

3.5 The Family of NLMS Algorithms with Variable Step Size

The performance of the family of NLMS algorithms is improved for the Gaussian environments using a variable step size, μ_k [43] instead of μ in Equation

(3.5). The variable step size μ_k is :

$$\mu_k = \mu_{k-1} + \rho e_k e_{k-1} F(\mathbf{X}_{k-1})^T \mathbf{X}_k \tag{3.35}$$

where ρ is a convergence parameter for the step size.

3.5.1 Derivation

In this subsection, we review the derivation for μ_k, [43]. In [43] the update equation is considered in a more general form, i.e.,

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu_k f(e_k) F(\mathbf{X}_k). \tag{3.36}$$

Following the stochastic gradient-descent procedure as in [45] and [46], it can be written

$$\mu_k = \mu_{k-1} - \rho \frac{\partial \phi(e_k)}{\partial \mu_{k-1}}. \tag{3.37}$$

The function $\phi(.)$ denotes the relevant cost function to be minimized. Also

$$\frac{\partial \phi(e_k)}{\partial \mu_{k-1}} = \frac{\partial \phi(e_k)}{\partial e_k} \frac{\partial e_k}{\partial \mu_{k-1}}. \tag{3.38}$$

Following the notation in [47], $f(e_k) = \partial \phi(e_k)/\partial e_k$ and putting

$$\mathbf{W}_k = \mathbf{W}_{k-1} + \mu_{k-1} f(e_{k-1}) F(\mathbf{X}_{k-1}) \tag{3.39}$$

into the expression of e_k ,

$$e_k = d_k - \mathbf{W}_{k-1}\mathbf{X}_k - \mu_{k-1} f(e_{k-1}) F(\mathbf{X}_{k-1})^T \mathbf{X}_k \tag{3.40}$$

is obtained. Thus, we get

$$\frac{\partial e_k}{\partial \mu_{k-1}} = -f(e_{k-1}) F(\mathbf{X}_{k-1})^T \mathbf{X}_k. \tag{3.41}$$

Combining (3.37), (3.38) and (3.41) yields the step size update as

$$\mu_k = \mu_{k-1} + \rho f(e_k) f(e_{k-1}) F(\mathbf{X}_{k-1})^T \mathbf{X}_k. \tag{3.42}$$

Since in our case $f(e_k) = e_k$, the last form of the step size is as given in Equation (3.35).

3.5.2 Performance of the Family of the NLMS Algorithms with Variable Step Size in Impulsive Noise

The same experiment of the Section 3.4.3 is performed again. As can be seen from Figure 3.6 the performance of the algorithm is unsatisfactory in the case of α-stable distributions. This result is also obvious from the fact that the expected value of the variable step size of Equation (3.35) is infinite in the case of α-stable distributions.

3.6 Conclusion

In this chapter we review a family of NLMS algorithms. The performance of the algorithms are assessed in impulsive noise environments. The obtained plots are far from satisfactory. The degradation is also shown theoretically by finding the variance of the update term as infinite by using α-stable random processes. After all, it is clear that we need to develop some robust algorithms for impulsive noise environments.

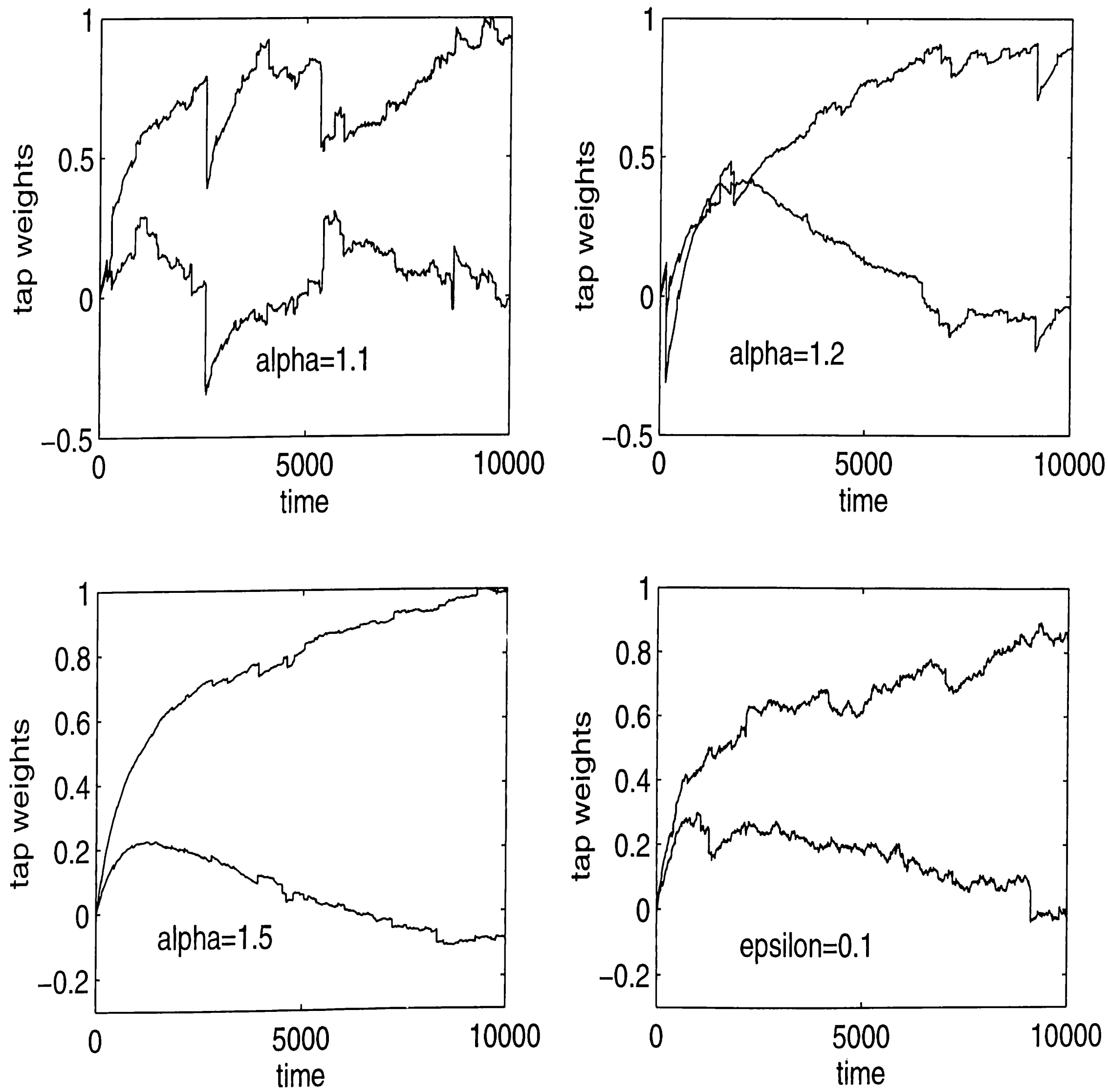


Figure 3.2: Transient behavior of tap weight adaptations for the NLMS algorithm of Equation (3.1) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

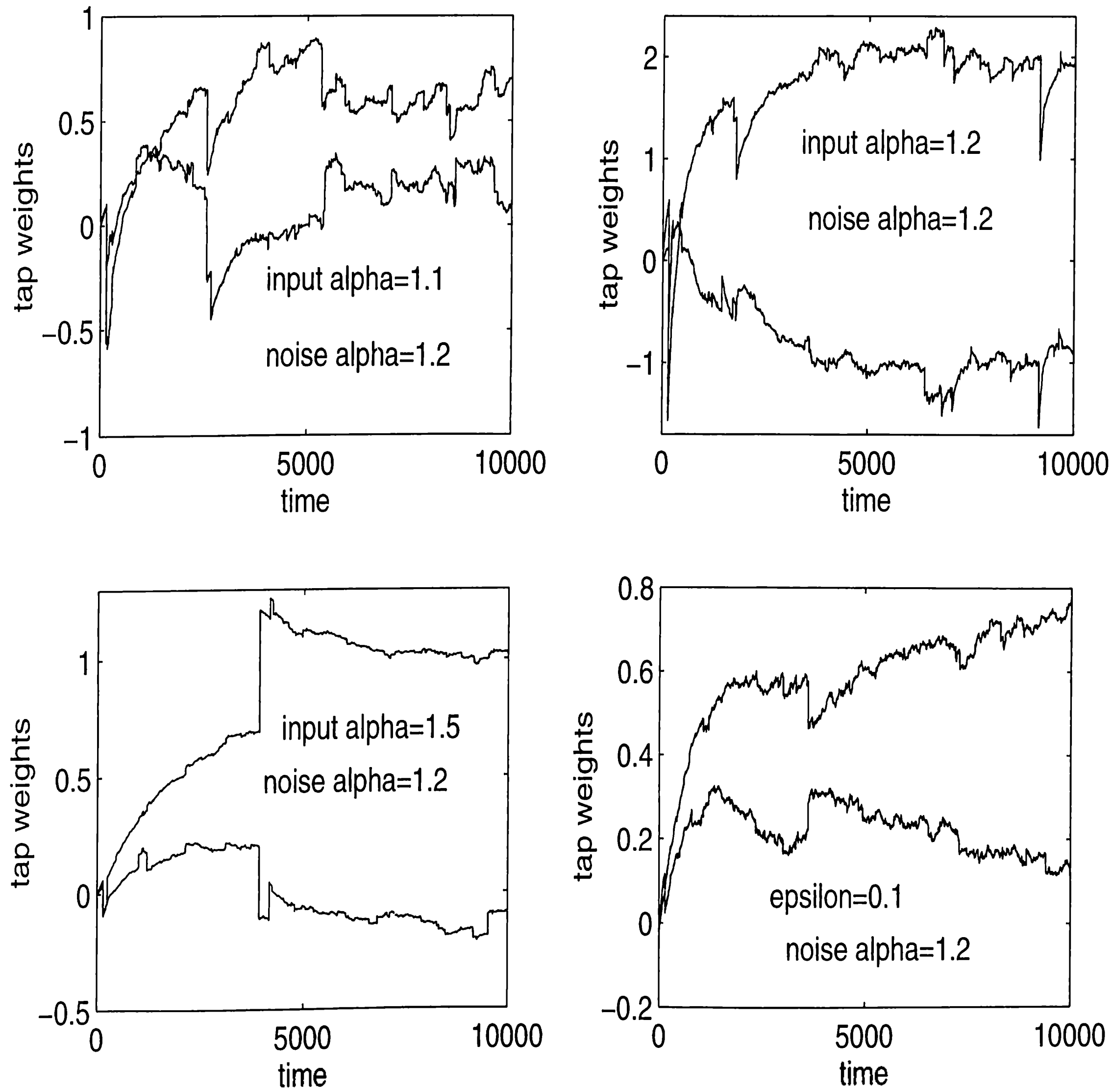


Figure 3.3: Transient behavior of tap weight adaptations for the NLMS algorithm of Equation (3.1) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$ under additive observation noise whose $\alpha = 1.2$. The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

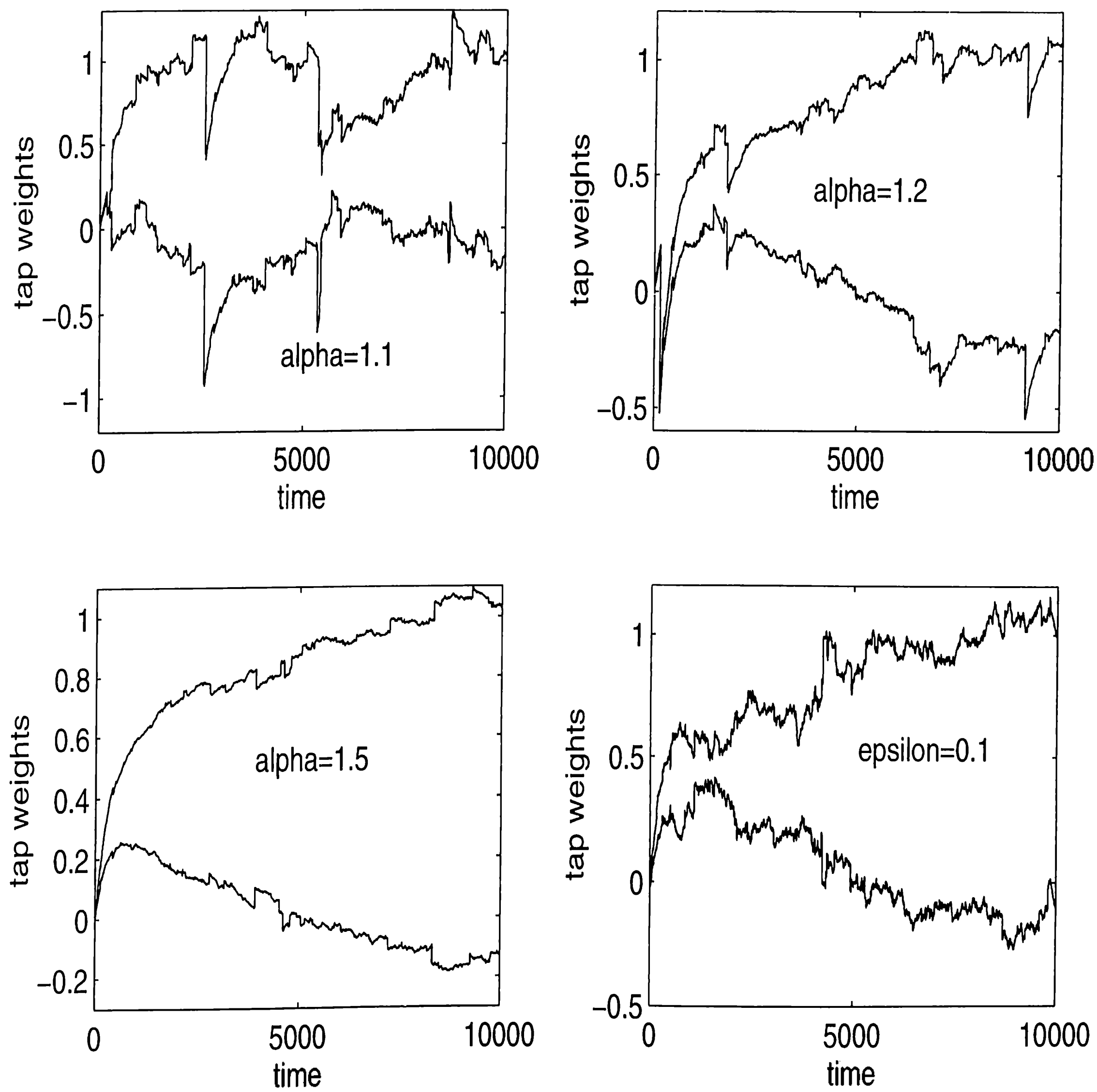


Figure 3.4: Transient behavior of tap weight adaptations for the modified NLMS algorithm of Equation (3.4) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

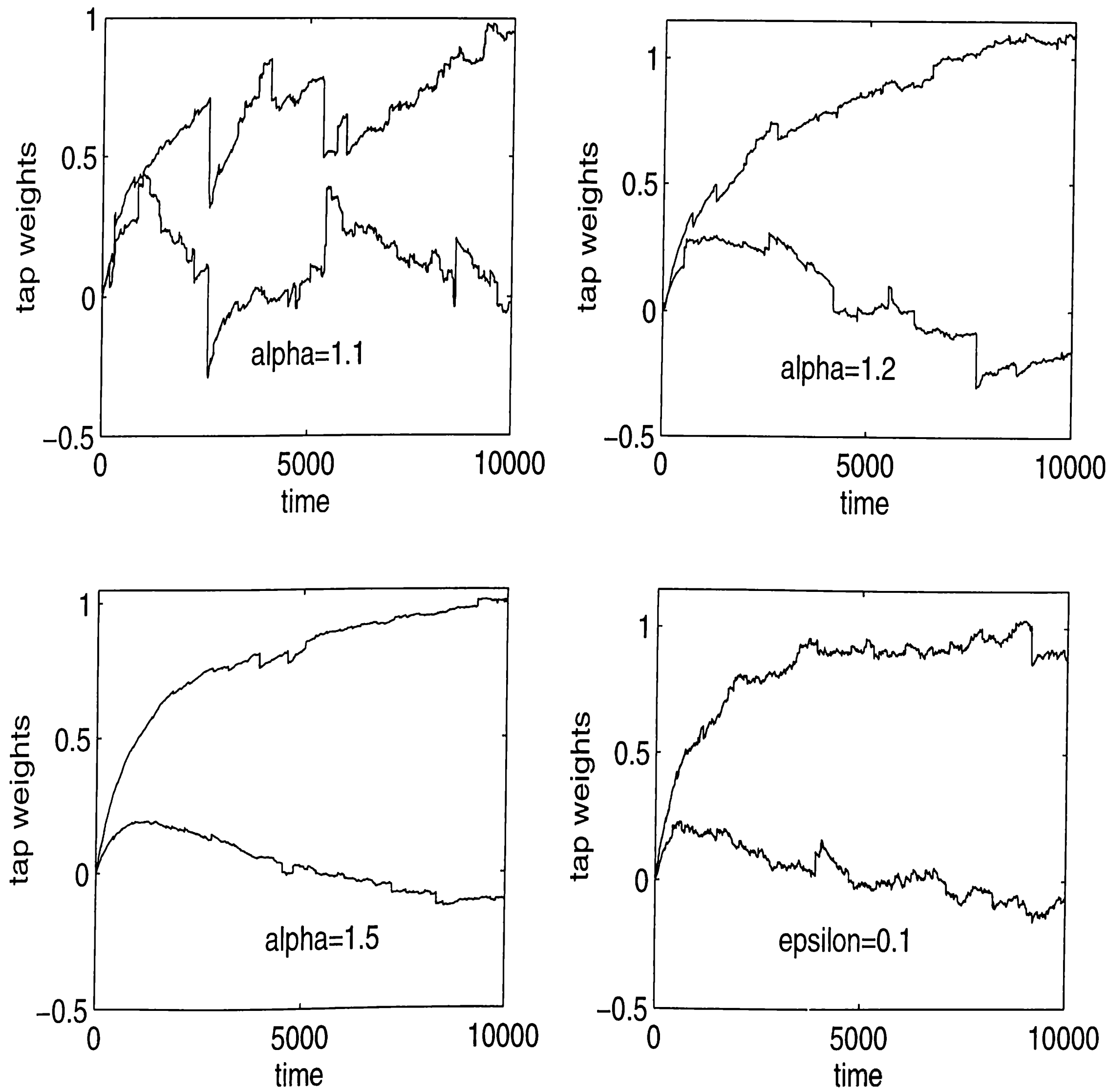


Figure 3.5: Transient behavior of tap weight adaptations for the family of the NLMS algorithm of Equations (3.5) and (3.6) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

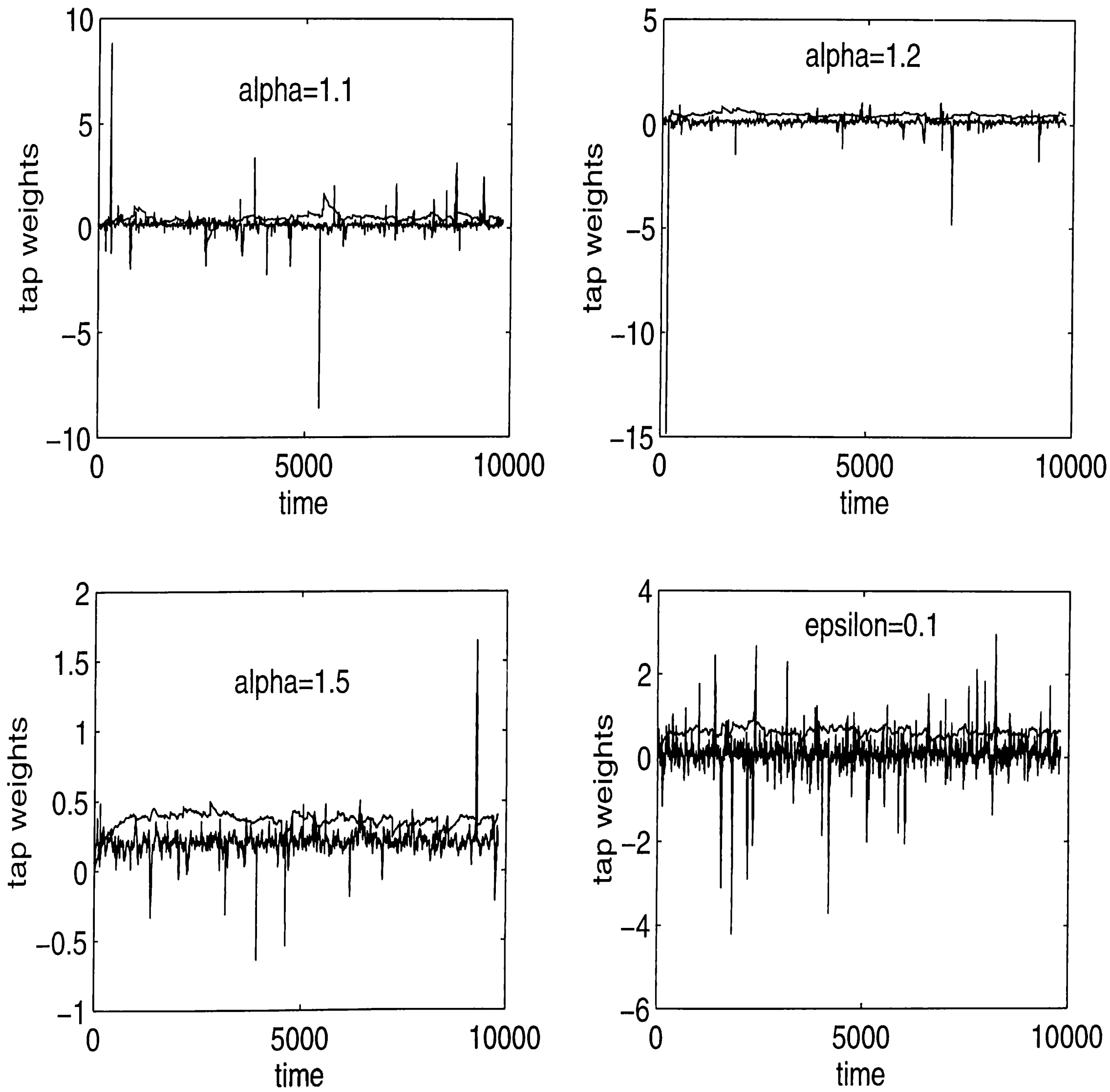


Figure 3.6: Transient behavior of tap weight adaptations for the family of the NLMS algorithm of Equations (3.5) and (3.6) with variable step size of Equation (3.35) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

Chapter 4

FRACTIONAL LOWER-ORDER STATISTICS BASED ALGORITHMS

4.1 Introduction

The well-known NLMS algorithm is based on a second-order cost function. Therefore it exhibits poor performance under α-stable noise. We robustify this algorithm using the Fractional Lower Order Statistics (FLOS). In this chapter, proposed adaptive algorithms utilize fractional moments and correlations.

In the following section, the proposed FLOS based family of algorithms and some of their properties are investigated. In Section 4.3 and 4.4 some existing FLOS based adaptive filtering algorithms for α-stable distributions are reviewed. Following these, simulation results are presented. The performance of these algorithms are investigated for systems with unknown orders in Section 4.5. In Section 4.6 "Momentum" FLOS based adaptive algorithms are proposed with relevant simulations. In Section 4.7, "Median" FLOS based adaptive algorithms are introduced and in Section 4.8, we modify the FLOS based algorithms using a prenonlinearity at the input and the desired signal. "Delayed" FLOS based algorithms are presented in Section 4.9. Extensive simulation studies are also presented.

Let us introduce the following notation [17]. For any two real numbers z

and $y \geq 0$ as :

$$z^{<y>} \triangleq |z|^y \text{sign}(z)$$

where sign(.) is the *signum* function.

4.2 Proposed FLOS Based Algorithm

As it is discussed in Section 3.4.2, the variance of the update term of the algorithms presented in Chapter 3 is not finite in impulsive environments. In order to achieve a finite variance, i.e.,

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] < \infty \tag{4.1}$$

we modify the algorithms of Chapter 3 using Fractional Lower Order Statistics (FLOS) concept.

The fractional lower-order moment (FLOM) of the error, $\mathbf{E}[|e_k|^b]$, is finite for $0 < b < \alpha$. Based on this observation we define the following update equation [51], [52], [53] :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_k{}^{<a>} F_q(\mathbf{X}_k) \tag{4.2}$$

$$[F_q(\mathbf{X}_k)]_i = \begin{cases} \frac{|x_{k-i}|^{(q-1)a}\text{sign}(x_{k-i})}{\sum_{m=0}^{M-1}|x_{k-m}|^{qa}} \text{ if } 2 < q < \infty \\ \\ \frac{1}{x_{k-n}^{<a>}}\delta_{i-n} \text{ if } q = \infty \end{cases} \tag{4.3}$$

where $[F_q(.)]_i$ denotes the i^{th} element of the vector-valued function $F_q(.)$, q satisfies the relation $1/\alpha + 1/q < 1$. The FLOS parameter $a > 0$ obeys the following inequality :

$$a < 1/2. \tag{4.4}$$

Also, it is easy to check that actual weights form a stationary point of the iterations. Hence, it is expected that the above class of algorithms which we call FLOS based will have better performance than those in Chapter 3, in impulsive environments.

In the following we will prove Equation (4.4) both for finite and infinite q case.

Finite q case : The update equation of the FLOS based algorithms is given by :

$$w_{i,k+1} = w_{i,k} + \mu e_k^{<a>} \frac{|x_{k-i}|^{(q-1)a}\mathrm{sign}(x_{k-i})}{\sum_{m=0}^{M-1} |x_{k-m}|^{qa}} \tag{4.5}$$

for each sample of the weight vector $\mathbf{W}_k$. Equation (4.5) implies that :

$$\mathbf{E}[(w_{i,k+1} - w_{i,k})^2] = \mu^2 \mathbf{E}\left[\left(\frac{|e_k|^a |x_{k-i}|^{(q-1)a}}{\sum_{m=0}^{M-1} |x_{k-m}|^{qa}}\right)^2\right] \tag{4.6}$$

which can be written as :

$$\mu^2 \mathbf{E}\left[\left(\frac{|e_k|^a |x_{k-i}|^{(q-1)a}}{\sum_{m=0}^{M-1} |x_{k-m}|^{qa}}\right)^2\right] = \mu^2 \mathbf{E}[|e_k|^{2a}] \mathbf{E}\left[\left(\frac{|x_{k-i}|^{(q-1)a}}{\sum_{m=0}^{M-1} |x_{k-m}|^{qa}}\right)^2\right], \tag{4.7}$$

by assuming that the error is uncorrelated with the past samples, x_{k-i} , of the input at steady state.

To have a finite value of the first expectation term in the right hand side, we should have the following inequality:

$$a < \alpha/2. \tag{4.8}$$

For the second expectation term of the right hand side of Equation (4.7), let $|x_{k-j}| = \max_{0 \leq m \leq M-1} |x_{k-m}|$, then the term inside the expectation can be written as :

$$\frac{|x_{k-i}|^{(q-1)a}}{\sum_{m=0}^{M-1} |x_{k-m}|^{qa}} \leq \frac{|x_{k-j}|^{(q-1)a}}{\sum_{m=0}^{M-1} |x_{k-m}|^{qa}} \leq \frac{|x_{k-j}|^{(q-1)a}}{|x_{k-j}|^{qa}} = \frac{1}{|x_{k-j}|^{a}}. \tag{4.9}$$

From here, we have to find the value of a satisfying the following:

$$\mathbf{E}\left[\frac{1}{|x_{k-j}|^{2a}}\right] < \infty. \tag{4.10}$$

It is shown in [61] that the property given in Equation (2.13) is also valid for the interval $-1 < p < \alpha$. Therefore the condition in Equation (4.1) is satisfied when

$$a < 1/2. \tag{4.11}$$

Since the value of α is in the interval [1,2), in the applications of adaptive filtering for impulsive environments, taking the simultaneous solution of

Equation (4.8) and (4.11) we get the condition for a as in Equation (4.11), i.e., $a < 1/2$ having :

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] < \infty. \tag{4.12}$$

Infinite q case : The update equation of the FLOS based algorithm can be rewritten as :

$$w_{i,k+1} = \begin{cases} w_{i,k} + \mu \frac{e_k^{<a>}}{x_{k-i}^{<a>}} & \text{if } |x_{k-i}| = \max_{0 \le m \le M-1} |x_{k-m}| \\ w_{i,k} & \text{otherwise.} \end{cases} \tag{4.13}$$

So we consider for the analysis only the value of i for which $|x_{k-i}| = \max_{0 \le m \le M-1} |x_{k-m}|$. Then, we may write

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \mu^2 \mathbf{E}\left[\frac{|e_k|^{2a}}{|x_{k-i}|^{2a}}\right]. \tag{4.14}$$

Again assuming that the error and samples of the input vector are uncorrelated and the system converges, this last equation can be expressed as :

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \mu^2 \mathbf{E}[|e_k|^{2a}] \mathbf{E}\left[\frac{1}{|x_{k-i}|^{2a}}\right]. \tag{4.15}$$

To have a finite value for the $\mathbf{E}[|e_k|^{2a}]$ we must have the condition of Equation (4.8) for a. Following the same arguments as in finite q case, for the second expectation term we have for a again the condition of Equation (4.11). Taking the intersection region, we obtain $a < 1/2$ for having

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] < \infty. \tag{4.16}$$

4.3 Least Mean-p Norm (LMP) Algorithm

The first of the FLOS based adaptive filtering algorithms for impulsive noise environments is called Least Mean-p Norm (LMP) [17] algorithm with the following update equation :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_k^{<p-1>} \mathbf{X}_k \tag{4.17}$$

The cost function of this algorithm is given by,

$$J_k = \mathbf{E}[|e_k|^p] = \mathbf{E}[|d_k - \mathbf{W}_k^T \mathbf{X}_k|^p] \tag{4.18}$$

where $1 \le p < \alpha$.

There is no closed-form solution for the set of the coefficients minimizing Equation (4.18). But knowing that J_k is convex, a stochastic gradient method to solve the coefficients as in Equation (4.17) can be used. The algorithm in Equation (4.17) is called Least Mean Absolute Deviation (LMAD) when $p = 1$. The LMAD is actually the familiar signed LMS algorithm, although it is derived in a different context.

4.4 Normalized Least Mean-p Norm (NLMP) Algorithm

With the motivation of the NLMS algorithm, recently LMP algorithm is normalized giving the Normalized Least Mean p-Norm (NLMP) [33] algorithm by the following update equation :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k^{<p-1>}}{||\mathbf{X}_k||_p^p + \lambda} \mathbf{X}_k \quad \text{for} \quad 1 \leq p < \alpha \tag{4.19}$$

where $\mu, \lambda > 0$ are appropriately chosen update parameters. When $p = 1$ the algorithm in Equation (4.19) is called Normalized Least Mean Absolute Deviation (NLMAD) [33], having the following update equation:

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{\text{sign}(e_k)}{||\mathbf{X}_k||_1 + \lambda} \mathbf{X}_k \tag{4.20}$$

The NLMP algorithm is shown to outperform the other existing algorithms when α-stable distributions are used, [33]. Therefore, during the simulation studies we will only consider the NLMP algorithm and the algorithm that we propose in Section 4.2.

4.4.1 Simulation Studies

In Figure 4.1, the system identification problem of Chapter 3 is considered. A comparison study is performed for the proposed FLOS based algorithm of Equations (4.2) and (4.3) and the NLMP algorithm of Equation (4.19). The same comparison study is made on Figure 4.2 for an AR(3) process with coefficients $a_1 = 0.99$, $a_2 = -0.152$, $a_3 = -0.097$.

In Figure 4.3 the system mismatch, [50], $||\mathbf{W}_k - \mathbf{W}_*||_2^2$, where $\mathbf{W}_k$ and $\mathbf{W}_*$ are the current tap weight and the optimal solution vectors, respectively, of the AR(3) process defined above versus time plots are given for both algorithms.

In Figure 4.4 the proposed FLOS based algorithm is investigated under additive impulsive observation noise with the same AR(2) process as above. The degradation of the algorithm is seen from the plot clearly.

In Figure 4.5 the degradation of the NLMP algorithm under additive impulsive observation noise is investigated with the same AR(2) process as in Figure 4.1.

In Figure 4.6 a comparison study under additive impulsive noise is plotted for the proposed FLOS based algorithm and the NLMP algorithm. From this plot it is seen that they have comparable performance.

All the plots are obtained by 100 independent trials and to get a fair comparison between the algorithms, the step size of the algorithms are adjusted so that the steady state variances of the tap weights are equal.

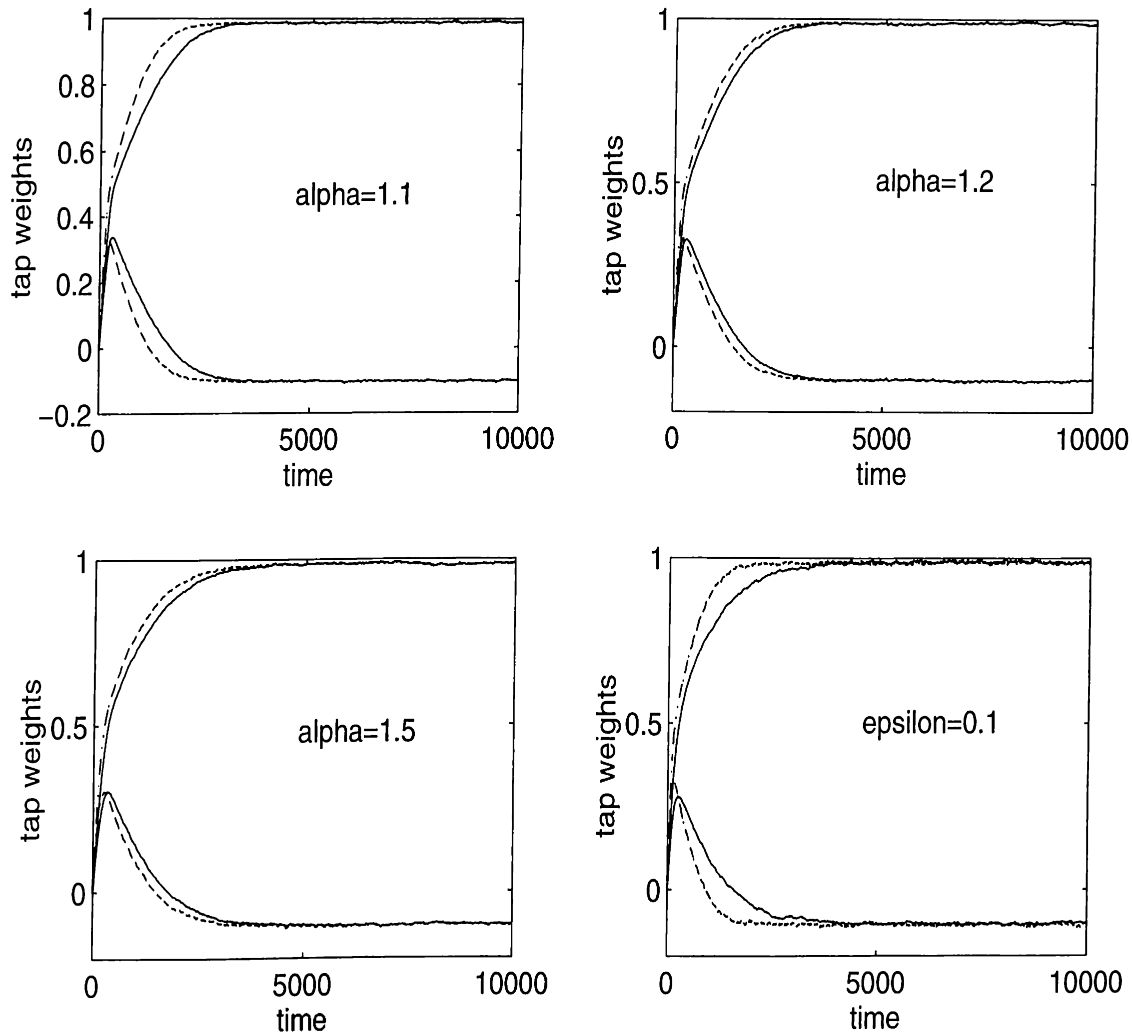


Figure 4.1: Transient behavior of tap weight adaptations for the proposed FLOS based algorithm (dashed line) of Equations (4.2) and (4.3) and the NLMP algorithm (solid line) of Equation (4.19) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

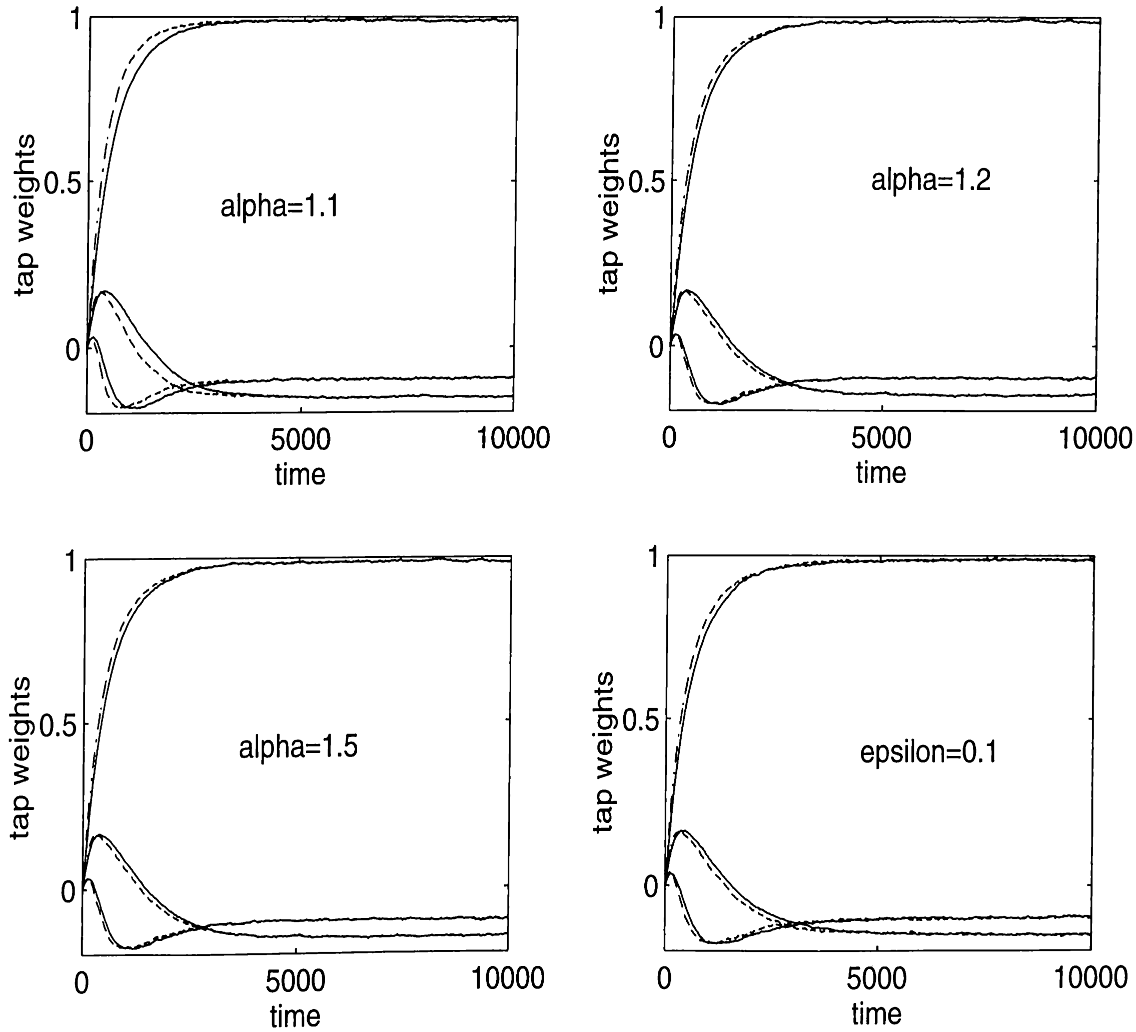


Figure 4.2: Transient behavior of tap weight adaptations for the proposed FLOS based algorithm (dashed line) of Equations (4.2) and (4.3) and the NLMP algorithm (solid line) of Equation (4.19) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(3) process parameters are $a_1 = 0.99$, $a_2 = -0.152$ and $a_3 = -0.097$.

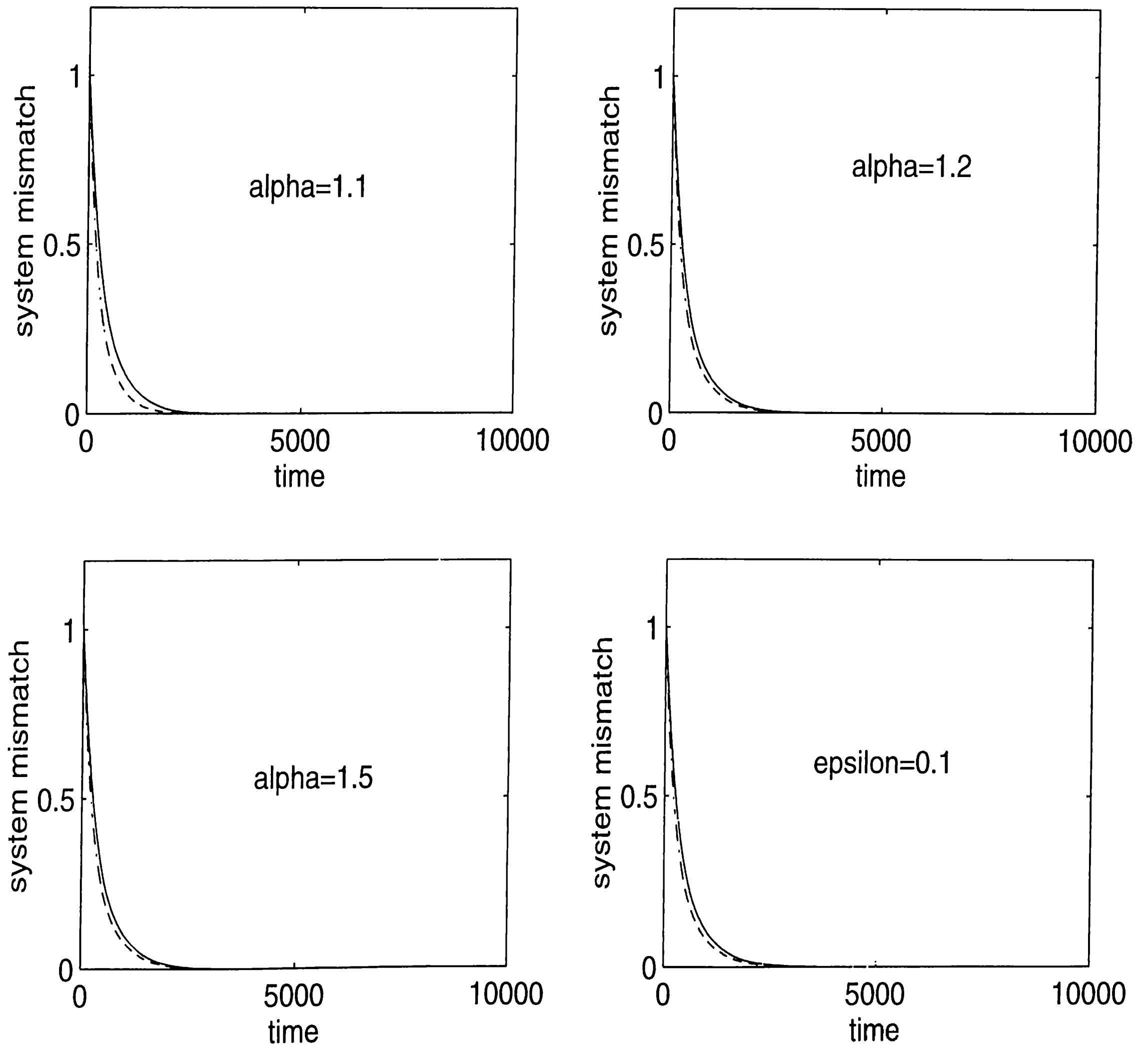


Figure 4.3: The system mismatch, $||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2$, versus time is plotted for the proposed FLOS based algorithm (dashed line) of Equations (4.2) and (4.3) and the NLMP algorithm (solid line) of (4.19) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$ and $\varepsilon = 0.1$. The AR(3) parameters are $a_1 = 0.99$, $a_2 = -0.152$ and $a_3 = -0.097$.

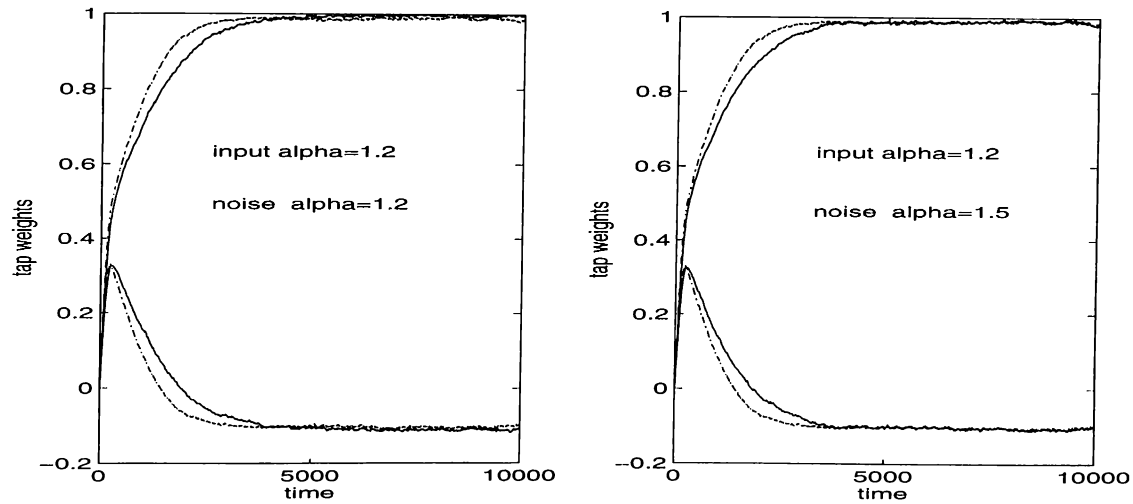


Figure 4.4: Transient behavior of the tap weight adaptations for the proposed FLOS based algorithm of Equations (4.2) and (4.3) for $\alpha = 1.2$ under additive impulsive observation noise (solid line) when the noise distribution has α values as 1.2 and 1.5, respectively. For comparison the performance under no additive observation noise is also plotted (dashed line). The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

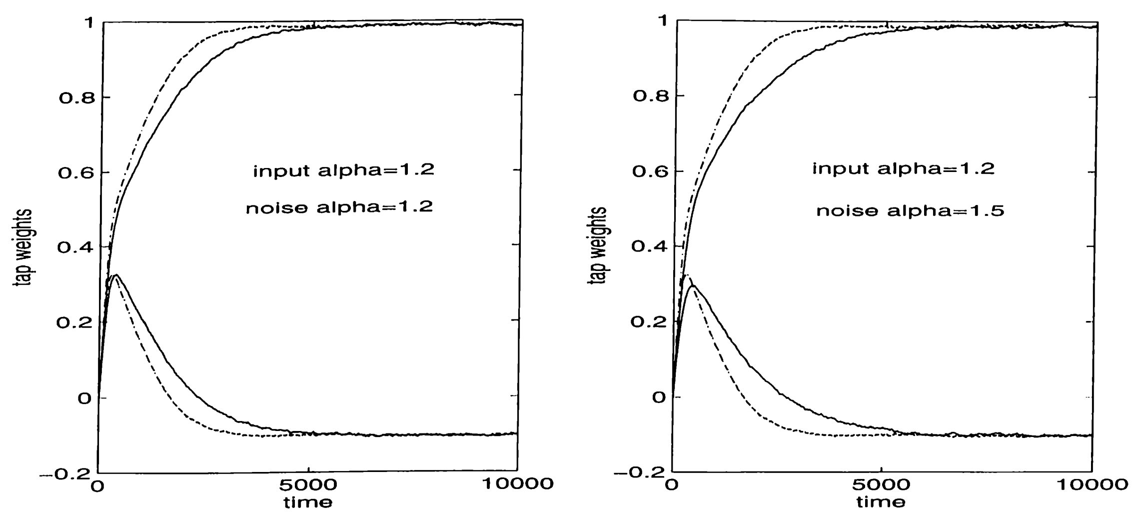


Figure 4.5: Transient behavior of the tap weight adaptations for the NLMP algorithm of Equations (4.19) for $\alpha = 1.2$ under additive impulsive observation noise (solid line) when the noise distribution has α values 1.2 and 1.5, respectively. For comparison the performance under no additive observation noise is also plotted (dashed line). The AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

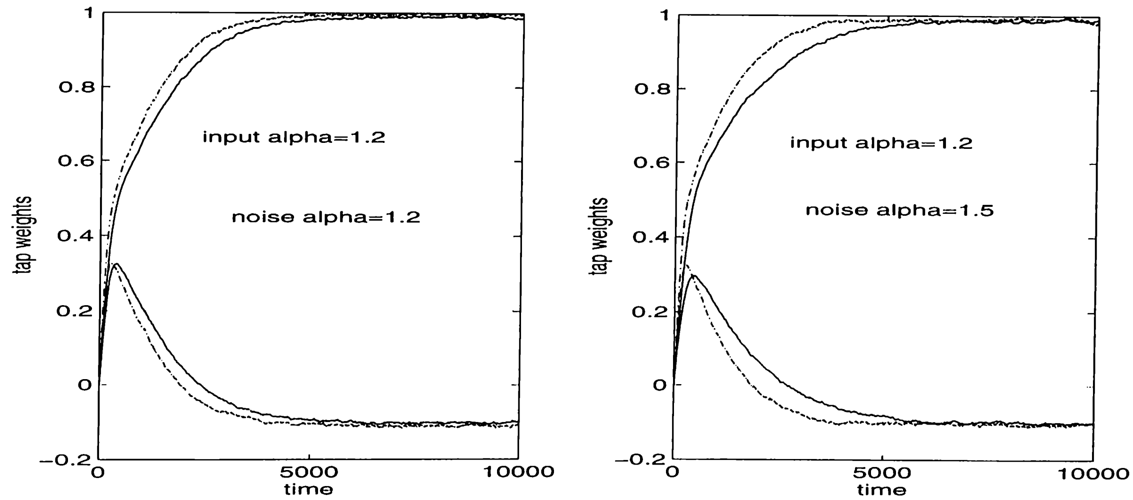


Figure 4.6: Transient behavior of the tap weight adaptations for the proposed FLOS based algorithm of Equations (4.2) and (4.3) (dashed line), and the NLMP algorithm (solid line) of Equations (4.19), for $\alpha = 1.2$ under additive impulsive observation noise when the noise distribution has α values 1.2 and 1.5, respectively.

4.5 Performance of the FLOS Based Algorithm for Systems with Unknown Orders

In practice, there may be some systems with unknown order. In this section, we try to see the performance of the proposed algorithm of Equations (4.2) and (4.3) and the NLMP algorithm of Equation (4.19) when the order of the system is unknown.

For this purpose, we generate the AR(2) SαS process with $\alpha = 1.2$ and $a_1 = 0.99$ and $a_2 = -0.1$ that we used before. Then we try to find the 5 tap weight FIR filter for this system. In Figure 4.7, we plot the transient behaviors of these 5 tap weights for both of the algorithms and see their steady state values as follows : $a_1 = 0.9858$, $a_2 = -0.0956$, $a_3 = -0.0047$, $a_4 = 0.0014$ and $a_5 = -0.0030$. As it is seen from these results a_3, a_4 and a_5 are so small that they can be neglected, i.e., the system can be thought as a second-order system.

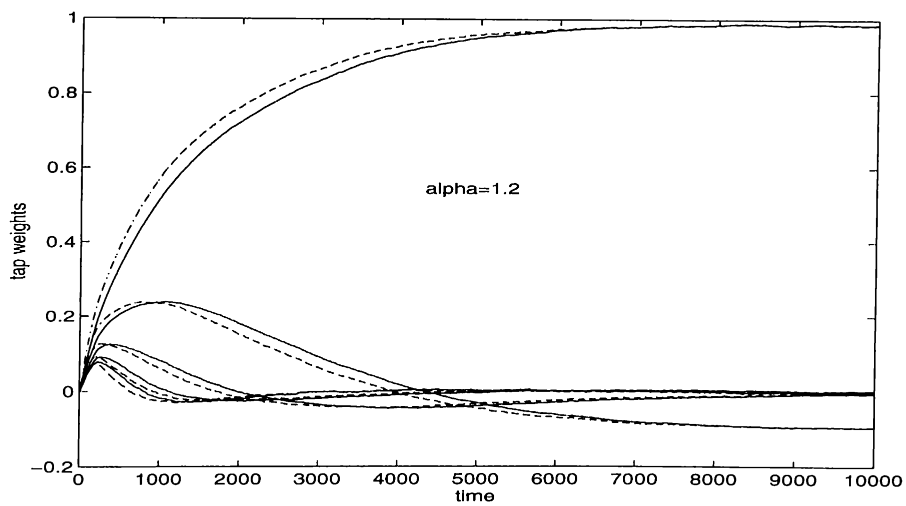


Figure 4.7: Transient behavior of the tap weight adaptations for the proposed FLOS based algorithm of Equations (4.2) and (4.3) (dashed line), and the NLMP algorithm (solid line) of Equations (4.19), for $\alpha = 1.2$. The AR(2) system, with $a_1 = 0.99$ and $a_2 = -0.1$, is modeled by an AR(5) system.

4.6 "Momentum" FLOS Based Adaptive Algorithms

The algorithms presented in this chapter until now are all based on the instantaneous value of the gradient vector. When the signals are Gaussian and stationary, ignoring the past values and considering just the values of the gradient vector at time k for evaluating the coefficients at the k^{th} step, is a reasonable approximation [49]. In impulsive noise environments, the current observation may be an outlier and the corresponding update term may be useless. Therefore the use of past values may provide robustness and improve convergence speed.

In this section, with the motivation of "Momentum" LMS algorithm [54], in addition to instantaneous value of the gradient vector, we deal also with some of its past values. By doing so, we expect to accelerate the algorithms.

4.6.1 "Momentum" FLOS Based Algorithm

The cost function of the algorithms in Equation (4.2) and (4.3) will be taken as

$$J_k = \sum_{n=k-j}^{k} \mathbf{E}[||\mathbf{W}_{n+1} - \mathbf{W}_n||_p^p] \quad \text{for} \quad 0 < j < k \quad \text{and} \quad 1 \leq p < \alpha \quad (4.21)$$

The corresponding update equation of the algorithm will have the following form :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \sum_{n=k-j}^{k} e_n^{<a>} F_q(\mathbf{X}_n) \quad (4.22)$$

where

$$[F_q(\mathbf{X}_n)]_i = \begin{cases} \frac{|x_{n-i}|^{(q-1)a}\text{sign}(x_{n-i})}{\sum_{m=0}^{M-1} |x_{n-m}|^{qa}} & \text{if } 2 < q < \infty \\ \\ \frac{1}{x_{n-m}^{<a>}} \delta_{n-m} & \text{if } q = \infty \end{cases} \quad (4.23)$$

for the i^{th} entry.

4.6.2 "Momentum" NLMP Algorithm

With the help of the motivation given above, we take the cost function of the NLMP algorithm of Equation (4.19) as

$$J_k = \sum_{n=k-j}^{k} \mathbf{E}[|e_n|^p] \quad \text{for} \quad 0 < j < k \quad \text{and} \quad 1 \leq p < \alpha \quad (4.24)$$

Using this new cost function we modify the NLMP algorithm by the following update equation :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \sum_{n=k-j}^{k} \frac{e_n^{<p-1>}}{||\mathbf{X}_n||_p^p + \lambda} \mathbf{X}_n. \quad (4.25)$$

The additional terms introduce a "Momentum" effect during convergence [54, 55], and they serve as the estimate of the previous gradients. Also by intuition, if the previous weight charges is large, then adding a fraction of this amount to the current update will "accelerate" the descent procedure to the global minimum. As a result, smoother and faster convergence can be expected for the weight vector. This improvement is achieved at the expense of the extra storage requirement of the past weight vectors and an additional scalar vector multiplication.

4.6.3 Simulation Studies

In Figure 4.8 the proposed FLOS based algorithm of Equations (4.2) and (4.3) is plotted with the "Momentum" FLOS based algorithm of Equations (4.22) and (4.23) using the AR(2) process with coefficients $a_1 = 0.99$ and $a_2 = -0.1$. In Figure 4.9 the NLMP algorithm of Equation (4.19) and the "Momentum" NLMP algorithm of Equation (4.25) is plotted for the same system identification problem. In Figure 4.10 we plot the system mismatch of both of the "Momentum" FLOS based algorithms mentioned above by generating an AR(5) process. The process parameters are $a_1 = 0.89$, $a_2 = -0.152$, $a_3 = 0.1$, $a_4 = -0.197$ and $a_5 = 0.097$.

We also compare the "Momentum" FLOS based algorithms with additive observation noise within themselves and each other. In Figure 4.11 "Momentum" FLOS based algorithm of Equations (4.22) and (4.23) with and without additive observation noise is plotted with the same AR(2) process, as above. In Figure 4.12 the "Momentum" NLMP algorithm of Equation (4.25) with and without additive observation noise is plotted again for the same AR(2) process. Lastly, in Figure 4.13 both of the "Momentum" algorithms under additive observation noise is plotted for the same system identification problem. We take $j = 1$ from Figure 4.8 to Figure 4.13.

In Figure 4.14 with the AR(5) α-stable process above we investigate the effect of the added last j terms to the update term in both of the algorithms of Equations (4.22) and (4.23) and Equation (4.25). We plot the system mismatch for various $j = 1, 3$ and 5 values. For $j = 1$ the proposed FLOS based algorithm of Equations (4.22) and (4.23) converges around 2500 time steps, whereas for $j = 5$, it converges around 1000 time steps. Similarly, the algorithm of Equation (4.25) converges in 3000 time steps for $j = 1$ and 1000 time steps for $j = 5$. However, increasing j means also decreasing the space in the memory. From the plots it is observed that, there is a great improvement in the results when j is increased from 1 to 3. However, there is not much difference when j is increased from 3 to 5. We reach a point of diminishing returns at $j = 3$ and the use of too many past values of the gradient vector does not improve the convergence further.

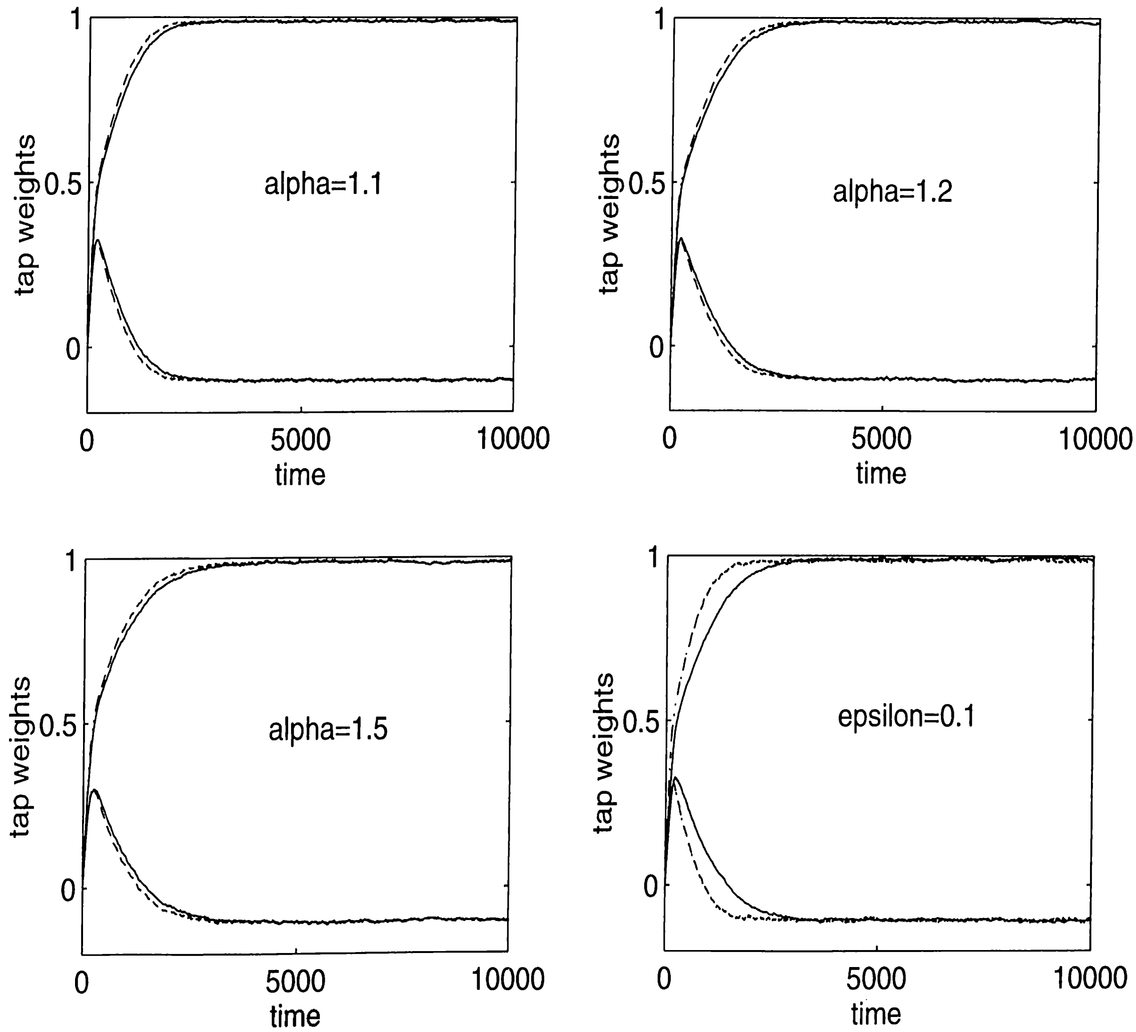


Figure 4.8: Transient behavior of tap weight adaptations for the "Momentum" FLOS based algorithm (dashed line) of Equations (4.22) and (4.23) and the proposed FLOS based algorithm (solid line) of Equations (4.2) and (4.3) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$, and $\varepsilon = 0.1$. The AR(2) parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

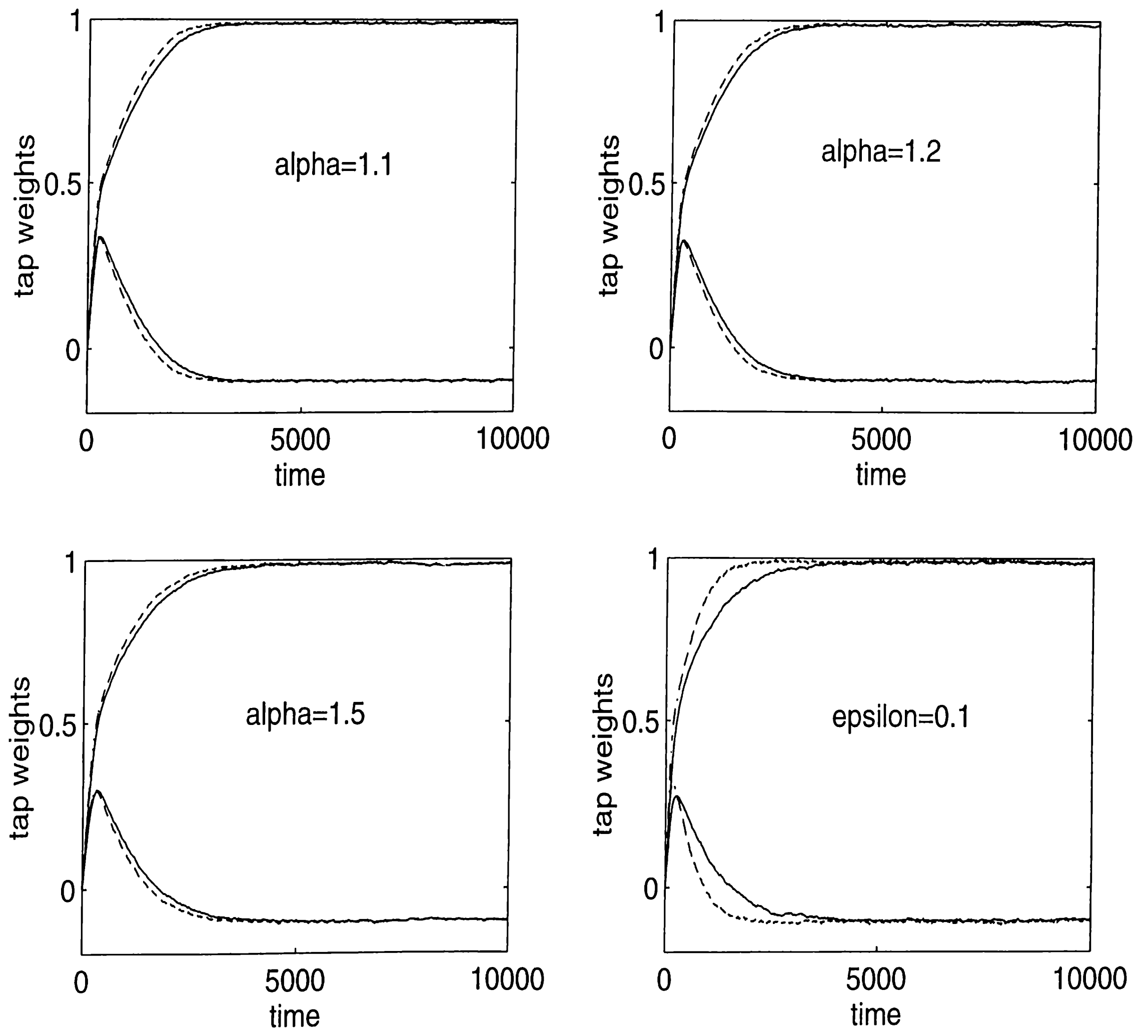


Figure 4.9: Transient behavior of tap weight adaptations for the "Momentum" NLMP algorithm (dashed line) of Equation (4.25) and the NLMP algorithm (solid line) of Equations (4.19) for $\alpha = 1.1$, $\alpha = 1.2$, $\alpha = 1.5$, and $\varepsilon = 0.1$. The AR(2) parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

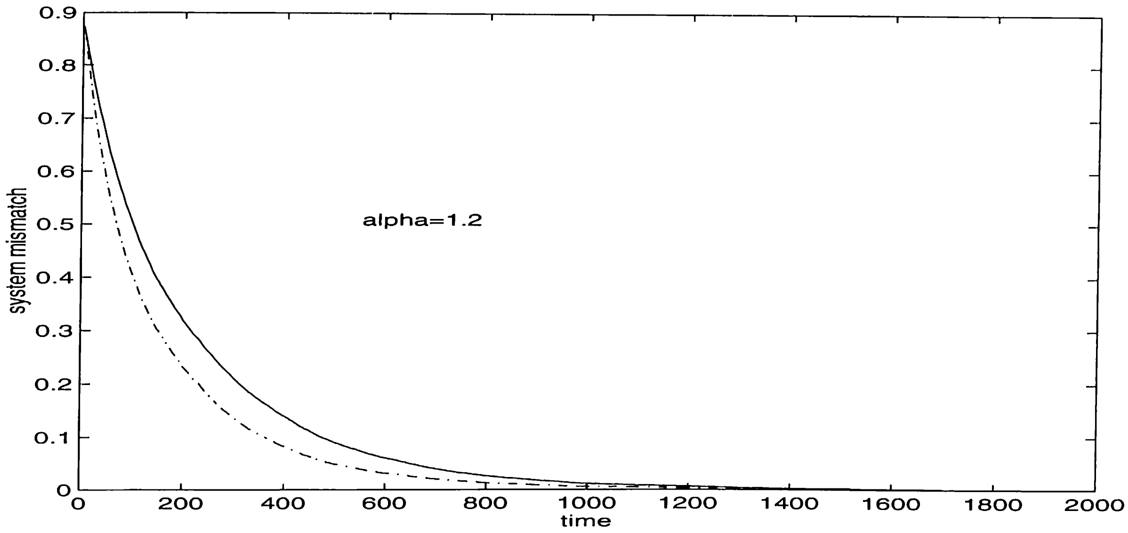


Figure 4.10: The system mismatch for the "Momentum" FLOS based algorithm (dashed line) of Equations (4.22) and (4.23) and the "Momentum" NLMP algorithm (solid line) of Equations (4.25) for $\alpha = 1.2$. The AR(5) parameters are $a_1 = 0.89$, $a_2 = -0.152$, $a_3 = 0.1$, $a_4 = -0.197$ and $a_5 = 0.097$.

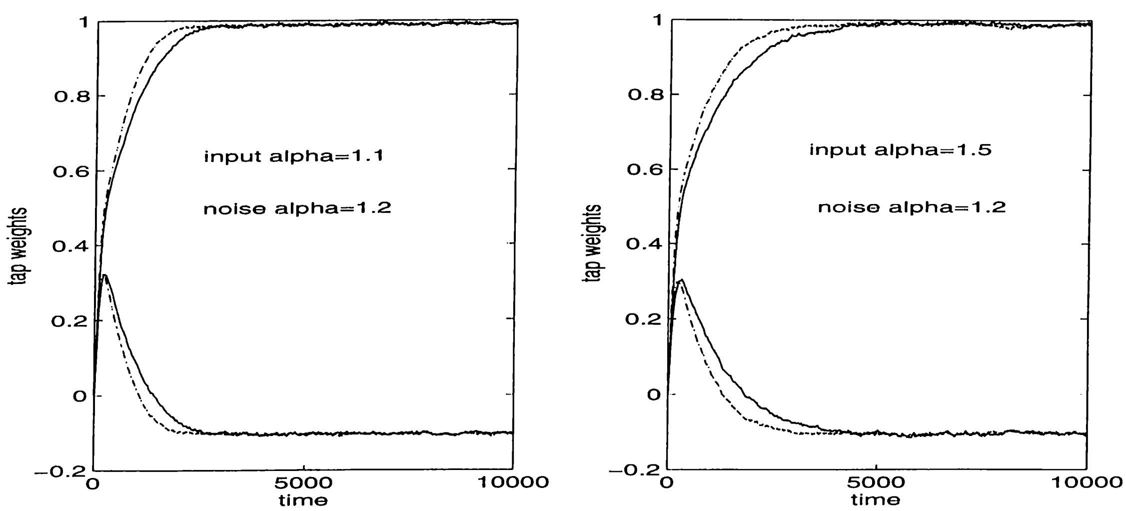


Figure 4.11: Transient behavior of tap weight adaptations for the "Momentum" FLOS based algorithm of Equation (4.22) and (4.23) for $\alpha = 1.1$ and $\alpha = 1.5$ under additive impulsive observation noise with $\alpha = 1.2$ (solid line). For comparison the performance under no additive observation noise is also plotted (dashed line). The AR(2) parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

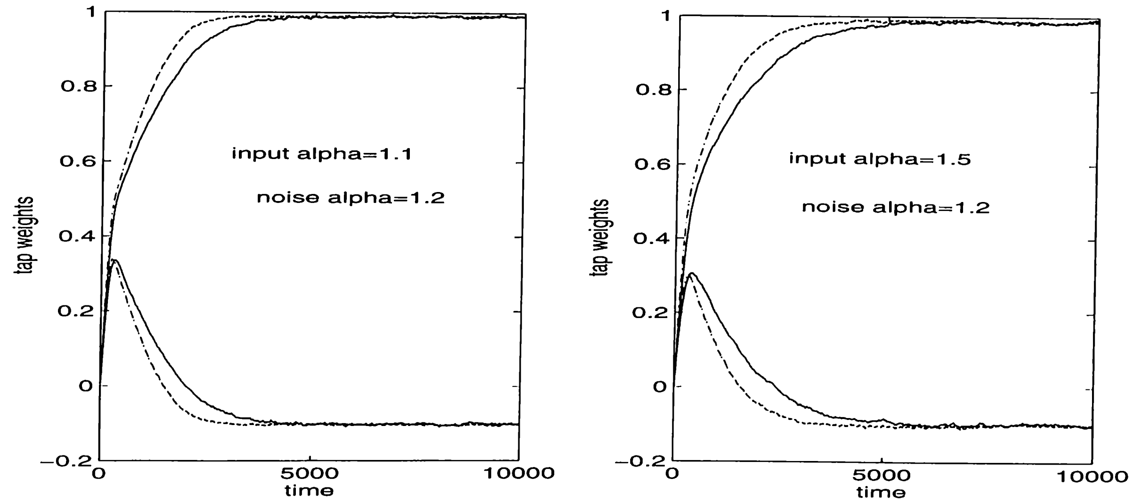


Figure 4.12: Transient behavior of tap weight adaptations for the "Momentum" NLMP algorithm of Equation (4.25) for $\alpha = 1.1$ and $\alpha = 1.5$ under additive impulsive observation noise with $\alpha = 1.2$ (solid line). For comparison the performance under no additive observation noise is also plotted (dashed line). The AR(2) parameters are $a_1 = 0.99$ and $a_2 = -0.1$

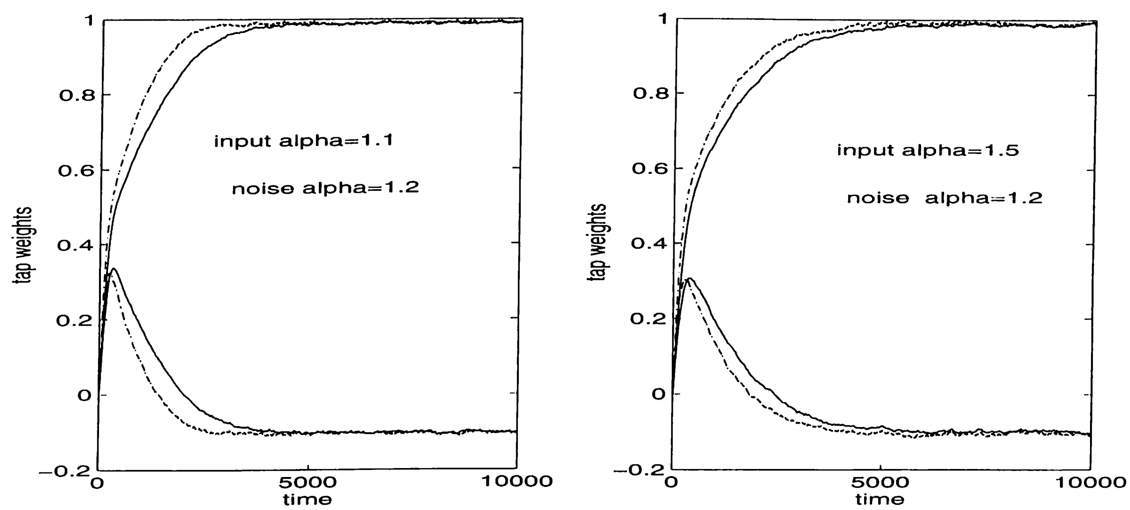
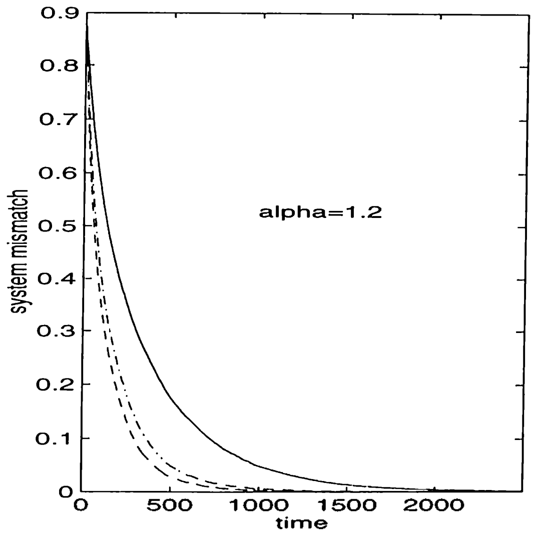


Figure 4.13: Transient behavior of tap weight adaptations for the "Momentum" FLOS based algorithm (dashed line) of Equations (4.22) and (4.23) and the "Momentum" NLMP algorithm (solid line) of Equation (4.25) for $\alpha = 1.1$ and $\alpha = 1.5$ under additive impulsive observation noise with $\alpha = 1.2$. The AR(2) parameters are $a_1 = 0.99$ and $a_2 = -0.1$.



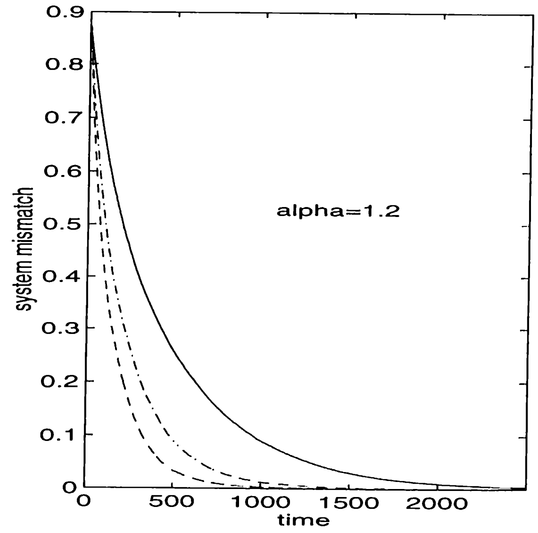


Figure 4.14: System mismatch for the proposed "Momentum" FLOS based algorithm of Equations (4.22) and (4.23) (left) and for the "Momentum" NLMP algorithm of Equation (4.25) (right). Solid line for $j = 1$, dashdot line for $j = 3$ and dashed line for $j = 5$. The AR(5) parameters are $a_1 = 0.89$, $a_2 = -0.152$, $a_3 = 0.1$, $a_4 = -0.197$ and $a_5 = 0.097$.

4.7 "Median" FLOS Based Adaptive Algorithms

Order statistics filtering is very effective in impulsive noise environments [56]. For some practical purposes the added last j terms of the update term of Equations (4.22) and (4.23) and Equation (4.25) may be median filtered, i.e., we may reorganize the update equations for the algorithm of Equations (4.22) and (4.23) as :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \text{median}(\mathbf{U}) \tag{4.26}$$

where

$$\mathbf{U} = [e_{k-j}{}^{<a>} F_q(\mathbf{X}_{k-j}), ..., e_k{}^{<a>} F_q(\mathbf{X}_k)]^T \tag{4.27}$$

with the same $F_q(.)$ used before and for the algorithm of Equation (4.25) as :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \text{median}(\mathbf{V}) \tag{4.28}$$

where

$$\mathbf{V} = \left[\frac{e_{k-j}^{<p-1>}}{||\mathbf{X}_{k-j}||_p^p + \lambda} \mathbf{X}_{k-j}, ..., \frac{e_k^{<p-1>}}{||\mathbf{X}_k||_p^p + \lambda} \mathbf{X}_k \right]^T \tag{4.29}$$

The major drawback of median filtering is that the estimators are biased [57] if the distribution of the input is asymmetric or if its mean and median have different signs. This point is shown in the following section.

4.7.1 Simulation Studies

In this subsection, for the AR(5) process used before, we plot in Figure 4.15 the transient behavior of the first tap weight for both of the algorithms of Equation (4.26) and (4.28) and see that although we choose SαS distribution as the input, we have a biased value for $a_1 = 0.89$ in both of the algorithms. We also show the algorithms of Equations (4.2) and (4.3) and the algorithm of Equation (4.19) on the same plot. We see that although the transient behaviour of the tap weights improves the steady state, tap weights are biased.

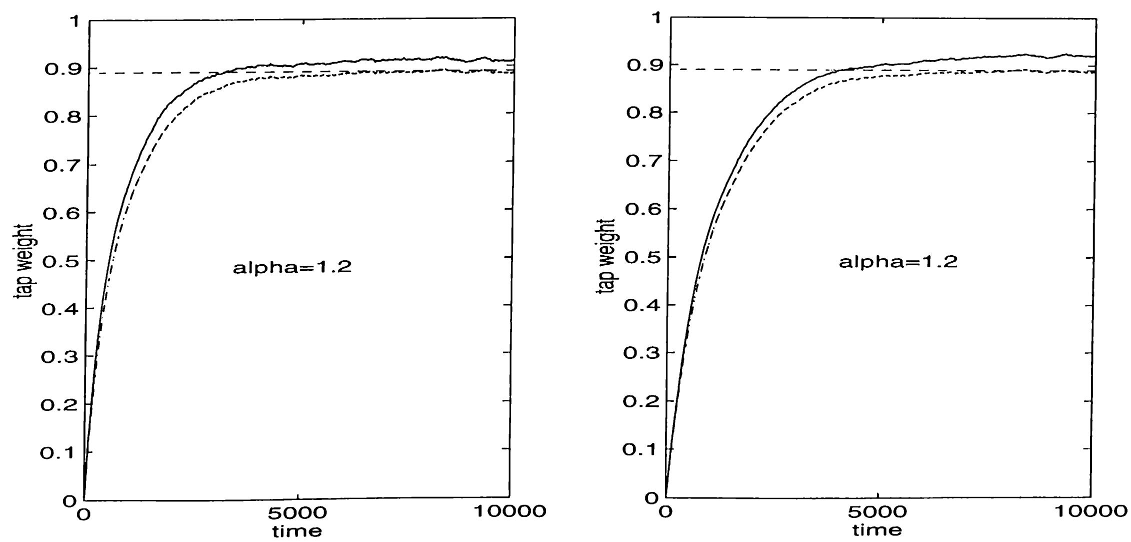


Figure 4.15: Transient behaviour of the first tap weight for "Median" FLOS based algorithm of Equations (4.26) (left) and for the "Median" NLMP algorithm of Equation (4.28) (right). Also the algorithms of Equation (4.2) and (4.3) (left) and Equation (4.19) (right) with dashed line. True value, $a_1 = 0.89$, is plotted by the dashed line.

4.7.2 Variation of the Parameter α with Median Filtering

In this section, we investigate how the length of the median filter affects the parameter α of any α-stable sequence. For this purpose we take two α-stable sequences with $\alpha = 1.1$ and $\alpha = 1.2$. We median filtered these sequences with the filter length N. Then, we estimate the value of α by using the *Koutrovelos Method* and *linear regression* [17] for this median filtered sequence. The results are shown in Table 4.1. In the first column of the table, there is the length of the median filter, in the second column, the estimated α values when the true value of the $\alpha = 1.1$ and in the third column, those estimated values for the true value of $\alpha = 1.2$.

N	$\hat{\alpha}(\alpha_{true} = 1.1)$	$\hat{\alpha}(\alpha_{true} = 1.2)$
3	1.6375	1.7593
5	1.8600	1.9421
7	1.9593	1.9763
9	1.9764	1.9851
11	1.9822	1.9881

Table 4.1: Table of computation results of α for different values of N

From this table, it is seen that as the length of the median filter increases, the sequence tends to behave as if it is Gaussian. Being aware of this fact, we tested all the filter lengths above. The transient performance gets better, but again there is a bias at steady state.

4.8 Use of the Prenonlinearity in FLOS Based Adaptive Algorithms

In this section, we investigate both of the algorithms using a soft limiter [58], [59]. By doing so, we first passed the input and the desired signal through the nonlinearity of Figure 4.16. In [58], [59] it is shown that for the NLMP algorithm of Equation (4.19), using this kind of nonlinearity improves the performance, giving a small bias.

We also investigate the performance of the "Momentum" FLOS based algorithms under this prenonlinearity.

4.8.1 Simulation Studies

To see the performance of both of the algorithms of Equations (4.2) and (4.3) and Equation (4.19) using nonlinearity, we plot the tap weight adaptations of

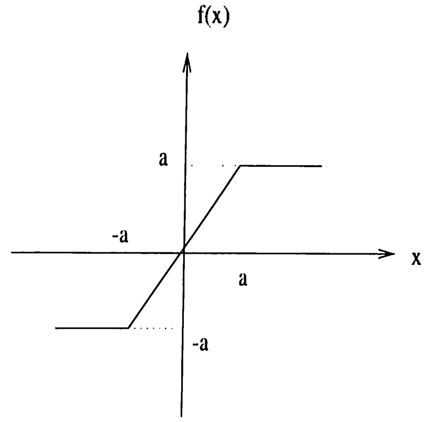


Figure 4.16: The nonlinearity used for the input and the desired signal.

the AR(2) process above in Figure 4.17. As it is seen from the plot there is a small bias at the output.

In Figure 4.18, we plot both of the "Momentum" FLOS based algorithms of Equations (4.22) and (4.23) and Equation (4.25) for the j values of $0, 1, 3$ and 5 using the nonlinearity of Figure 4.16 for the input and desired signal considering the AR(5) process used before. It can be deduced that as j increases the performance of both of the algorithms improve.

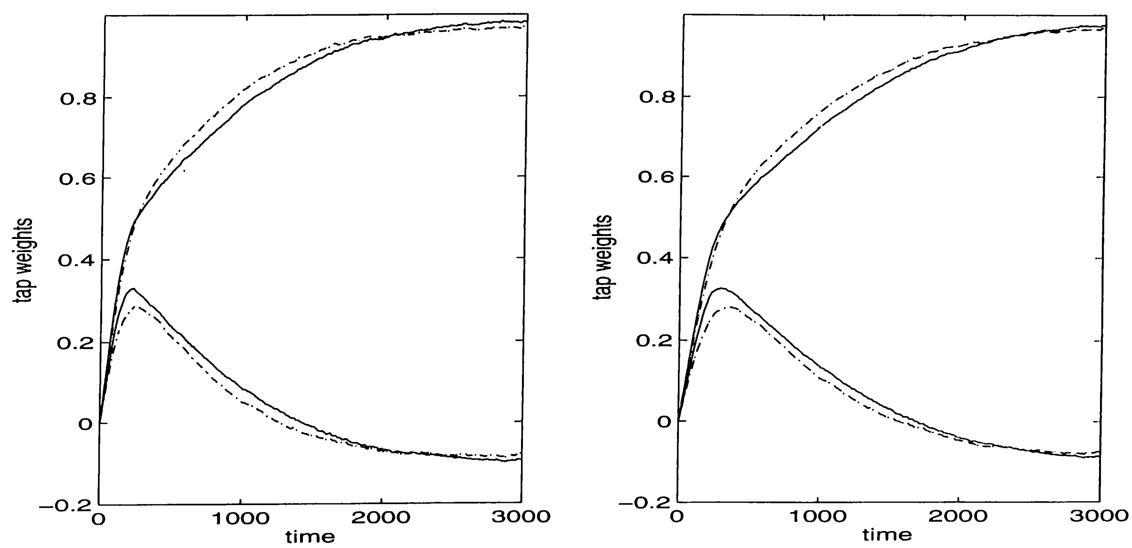


Figure 4.17: FLOS based algorithm of Equations (4.2) and (4.3) (left) and NLMP algorithm of Equation (4.19) (right) with (dashed) and without (solid) prenonlinearity, respectively.

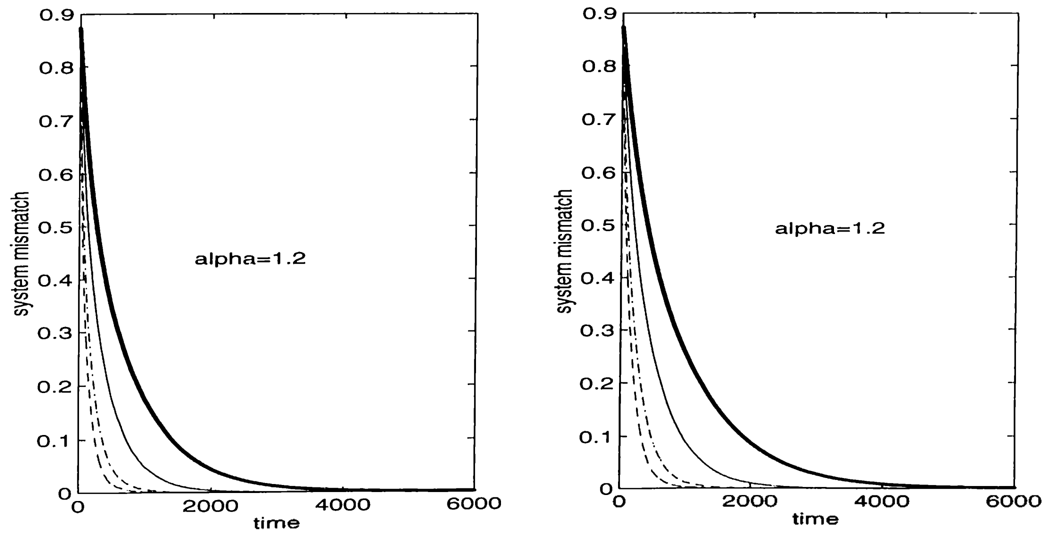


Figure 4.18: "Momentum" FLOS based algorithm of Equations (4.22) and (4.23) (left) and "Momentum" NLMP algorithm of Equation (4.25) (right) by using nonlinearity, for $j = 0, 1, 3$ and 5 for the heavy solid, solid, dashdot and dashed line, respectively.

4.9 "Delayed" FLOS Based Adaptive Algorithms

In some applications of adaptive filtering, the adaptation algorithm can be implemented only with a delay, h, in the coefficient update [60].

In this section, with the motivation of "Delayed" LMS algorithm, [58], we investigate the "Delayed" version of the FLOS based algorithms and compare their performance. The update equation for the proposed FLOS based algorithm of Equations (4.2) and (4.3) will be

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_{k-h}^{<a>} F_q(\mathbf{X}_{k-h}) \tag{4.30}$$

and the update equation of the NLMP algorithm of Equation (4.19) will be

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_{k-h}^{<p-1>} \mathbf{X}_{k-h}}{||\mathbf{X}_{k-h}||_p^p + \lambda} \tag{4.31}$$

where $e_{k-h} = d_{k-h} - \mathbf{W}_{k-h}{}^T \mathbf{X}_{k-h}$ and h is the delay. The term $F_q(\mathbf{X}_{k-h})$ is the term $F_q(\mathbf{X}_k)$ which is given before, with k is replaced by $k - h$.

4.9.1 Simulation Studies

In this subsection, we investigate the performance of the "Delayed" FLOS based algorithms under the system identification problem with the AR(2) process that we encountered before.

In Figure 4.19, we plot the system mismatch for the proposed algorithms of above. The simulations are again the average of 100 independent trials.

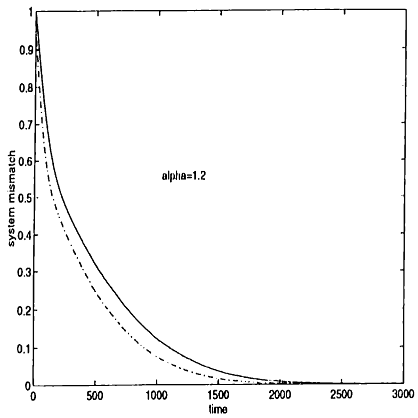


Figure 4.19: The system mismatch, $||\mathbf{W}_k - \mathbf{W}_*||_2^2$, for the "delayed" FLOS based algorithm of Equations (4.30) (dashed line), and the "delayed" NLMP algorithm of Equation (4.31) (solid) line. The AR(2) process parameters are $a_1 = 0.99$, $a_2 = -0.1$.

4.10 Computational Complexity of the FLOS Based Algorithms

The proposed FLOS based algorithm of Equation (4.2) and (4.3) can also be written as :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k^{<a>}}{||\mathbf{X}_k||_{qa}^{qa}} \tilde{\mathbf{X}}_k \tag{4.32}$$

where

$$\tilde{\mathbf{X}}_k = [x_k^{<(q-1)a>} \ldots x_{k-M+1}^{<(q-1)a>}]^T$$

and the NLMP algorithm which can be rewritten as :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k^{<p-1>}}{||\mathbf{X}_k||_p^p + \lambda} \mathbf{X}_k \tag{4.33}$$

for $p < \alpha$. If we compare the computational complexity of these two algorithms, it can be seen that the only difference is the nonlinear transformation of the input vector in the algorithm of Equation (4.32). Since the vector $\tilde{\mathbf{X}}_k$ can be rewritten as :

$$\tilde{\mathbf{X}}_k = \left[\frac{x_k{}^{<qa>}}{x_k} \cdots \frac{x_{k-M+1}{}^{<qa>}}{x_{k-M+1}}\right]^T = \left[\frac{x_k{}^{<qa>}}{x_k} \quad \tilde{\mathbf{X}}_{k-1}(1:M-1)\right]^T, \qquad (4.34)$$

only the term $\frac{x_k{}^{<qa>}}{x_k}$ is needed for recursive evaluation of $\tilde{\mathbf{X}}_k$, at each time step. The term $\frac{x_k{}^{<qa>}}{x_k}$ can be computed by power series expansion and it can be closely approximated by using a few multiplications independent of the filter length M. Therefore, the complexity of both NLMP and FLOS based algorithms are the same.

4.11 Conclusion

In this chapter we present new adaptive filtering algorithms in the presence of α-stable distributions. The proposed algorithms are developed with the motivation of FLOS concept and the family of NLMS algorithms of Chapter 3. The performance of the proposed algorithms are investigated under various additive observation noise. We also accelerate the algorithms that we propose and the NLMP algorithm. The performance of both of the accelerated algorithms under additive observation noise is also investigated. The median filtered versions and the prenonlinear filtered version of both of the algorithms are also presented. For some practical purposes, the delayed version of the algorithms are also explained. At the end, we see that the accelerated version of the proposed FLOS based algorithms gives the best results among all of the different versions of the algorithms. Ending with this result, we also deal with the computational complexity of the proposed FLOS based algorithm and see that it is of the same order with the NLMP algorithm.

Chapter 5

ROBUST LEAST MEAN MIXED NORM ADAPTIVE FILTERING

5.1 Introduction

In this chapter, Least Mean Mixed Norm (LMMN) algorithm [19] is robustified using the Fractional Lower Order Statistics (FLOS). In the next section, we briefly review the LMMN algorithm and some of its properties in impulsive noise environments. In Section 5.3, we present the Robust Least Mean Mixed Norm (RLMMN) algorithm and some of its properties.

5.2 The LMMN Algorithm

The LMMN algorithm is based on the Least Mean Square (LMS)[41] and Least Mean Fourth (LMF) [62] algorithms. The cost function and the update equation is given by :

$$J_k = \lambda \mathbf{E}[e_k^2] + (1-\lambda)\mathbf{E}[e_k^4] \tag{5.1}$$

and

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2\mu e_k[\lambda + 2(1-\lambda)e_k^2]\mathbf{X}_k, \tag{5.2}$$

respectively. The scalar $0 \leq \lambda \leq 1$ is the mixing parameter.

5.2.1 Performance of the LMMN Algorithm in Impulsive Environments

The algorithm of Equation (5.2) is investigated for α-stable distributions and it is shown that the variance of the update term is not finite, i.e. ,

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \infty. \tag{5.3}$$

Proof : Let

$$e_k = d_k - \mathbf{W}_k{}^T\mathbf{X}_k = \mathbf{V}_k{}^T\mathbf{X}_k \tag{5.4}$$

where

$$\mathbf{V}_k \triangleq \mathbf{W}^* - \mathbf{W}_k \tag{5.5}$$

is the *weight error vector*. Since $\mathbf{V}_k - \mathbf{V}_{k+1} = \mathbf{W}_{k+1} - \mathbf{W}_k$, the weight update equation of the LMMN algorithm can be written in terms of $\mathbf{V}_k$ as :

$$\mathbf{V}_{k+1} = \mathbf{V}_k - 2\mu e_k[\lambda + 2(1-\lambda)e_k^2]\mathbf{X}_k \tag{5.6}$$

and

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = 4\mu^2\mathbf{E}[(\mathbf{V}_k^T\mathbf{X}_k)^2[\lambda + 2(1-\lambda)(\mathbf{V}_k^T\mathbf{X}_k)^2]^2||\mathbf{X}_k||_2^2]. \tag{5.7}$$

We just want to show the unboundedness of the last equation. Considering the first term after expanding the parenthesis is sufficient. In other words, let us consider the term $\mathbf{E}[(\mathbf{V}_k^T\mathbf{X}_k)^2||\mathbf{X}_k||_2^2]$. We may use the "independence theory assumption" of [48] and assume that $\mathbf{V}_k$ and $\mathbf{X}_k$ are independent. Then, we have

$$\mathbf{E}[(\mathbf{V}_k^T\mathbf{X}_k)^2||\mathbf{X}_k||_2^2] = \sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{E}[v_{k-i}v_{k-j}]\mathbf{E}[x_{k-i}x_{k-j}||\mathbf{X}_k||_2^2]$$

$$= \sum_{i=1}^{M}\mathbf{E}[v_{k-i}^2]\mathbf{E}[x_{k-i}^2||\mathbf{X}_k||_2^2] + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M}\mathbf{E}[v_{k-i}v_{k-j}]\mathbf{E}[x_{k-i}x_{k-j}||\mathbf{X}_k||_2^2] \tag{5.8}$$

It is easy to see that $\mathbf{E}[x_{k-i}^2||\mathbf{X}_k||_2^2] = \infty$, since it has the terms like $\mathbf{E}[x_{k-i}^4]$ within this expectation, so the concluding result is

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \infty. \tag{5.9}$$

Assuming the error and the input are uncorrelated at steady state we will give a different proof :

Equation (5.2) can be written as :

$$||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2 = 4\mu^2 e_k^2[\lambda + 2(1-\lambda)e_k^2]^2||\mathbf{X}_k||_2^2. \tag{5.10}$$

At steady state if the algorithm converges, it can be assumed that the error, e_k, and the samples of the random process $\mathbf{X}_k$ are uncorrelated. Thus, taking the expected value of the last equation, we may write the right hand side as :

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = 4\mu^2 \mathbf{E}[e_k^2[\lambda + 2(1-\lambda)e_k^2]^2]\mathbf{E}[||\mathbf{X}_k||_2^2]. \tag{5.11}$$

In this expression, we immediately see that $\mathbf{E}[||\mathbf{X}_k||_2^2]$ is infinity in the case of SαS random processes. Therefore without any further investigation we can say that

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = \infty. \tag{5.12}$$

5.2.2 Simulation Studies

In this subsection, we consider the same system identification problem discussed in Chapter 4. We generate a realization of the AR(2) process with parameters $a_1 = 0.99$ and $a_2 = -0.1$. We plot a realization of the transient behaviour of the tap weights in Figure 5.1. The performance seen in this plot is unacceptable. The time axis cannot be extended beyond the value in the plot. After those time values, the system blows up and the tap weight values go to nonsense numbers like 10^{180}. In this plot, we take the value of the $\alpha = 1.2$, and the mixing parameter $\lambda = 0.1$, 0.3, 0.6 and 0.9.

5.3 Robust LMMN (RLMMN) Adaptive Filtering

In this section, a new family of descent type algorithms are presented. As discussed in Section 5.2.1 the update term of the LMMN algorithm does not have finite variance. In order to achieve finite variance, i.e.,

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] < \infty \tag{5.13}$$

we modify the algorithm of Section 5.2 using Fractional Lower Order Statistics (FLOS) concept.

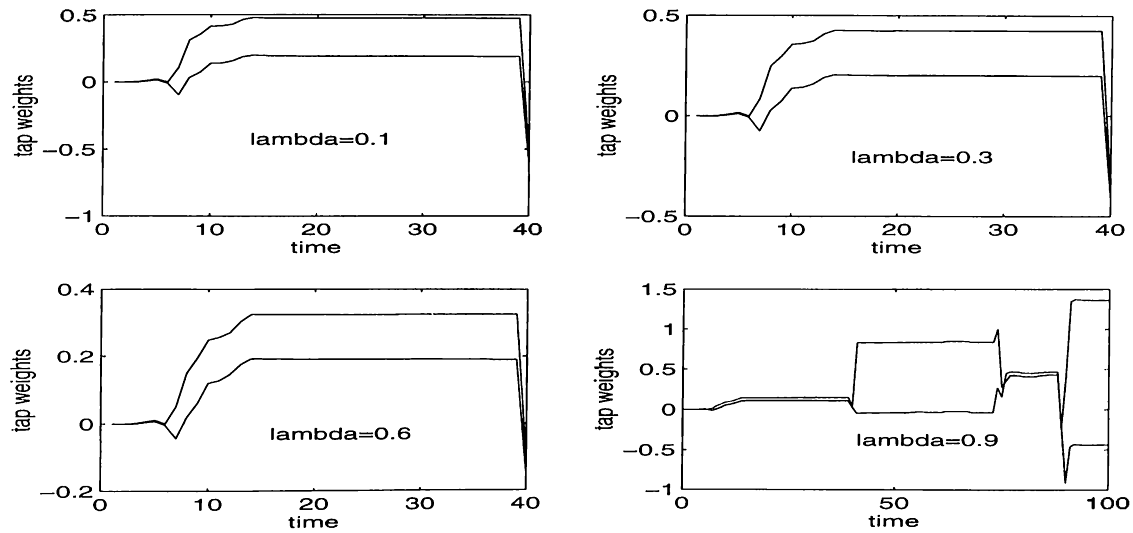


Figure 5.1: Transient behaviour of the tap weights for the LMMN algorithm for $\alpha = 1.2$.

In impulsive noise, the variance of the error, e_k, is not finite in LMMN algorithms of Section 5.2. However, the fractional lower order error,

$$e_k{}^{<b>} \triangleq |e_k|^b \text{sign}(e_k) \quad \text{for} \quad 0 < b < \alpha, \tag{5.14}$$

has a finite variance. Based on this observation we define the Robust LMMN (RLMMN) algorithm with the following update equation [63] :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2\mu e_k^{<a>}(\lambda + 2(1-\lambda)\{e_k^{<a>}\}^2)\tilde{\mathbf{X}}_k \tag{5.15}$$

where $\tilde{\mathbf{X}}_k = [x_k{}^{<a>}...x_{k-M+1}{}^{<a>}]^T$ and M is the order of the filter. It can be shown that the FLOS parameter $a > 0$ satisfies the relation :

$$a < \alpha/8 \tag{5.16}$$

to have the condition of Equation (5.13).

Proof : The weight update equation of the RLMMN algorithm can be written in terms of $\mathbf{V}_k$ as :

$$\mathbf{V}_{k+1} = \mathbf{V}_k - 2\mu(\mathbf{V}_k{}^T\mathbf{X}_k)^{<a>}\left[\lambda + 2(1-\lambda)\{(\mathbf{V}_k{}^T\mathbf{X}_k)^{<a>}\}^2\right]\tilde{\mathbf{X}}_k. \tag{5.17}$$

From here we have :

$$||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2 = 4\mu^2\{(\mathbf{V}_k{}^T\mathbf{X}_k)^{<a>}\}^2\left[\lambda + 2(1-\lambda)\{(\mathbf{V}_k{}^T\mathbf{X}_k)^{<a>}\}^2\right]^2||\tilde{\mathbf{X}}_k||_2^2.$$

For $|\mathbf{V}_k{}^T\mathbf{X}_k|$ using the Hölder inequality we may write

$$|\mathbf{V}_k{}^T\mathbf{X}_k| \leq ||\mathbf{V}_k||_1||\mathbf{X}_k||_\infty. \tag{5.18}$$

Note that $\mathbf{E}[||\mathbf{V}_k||_1] < \infty$ and $\mathbf{E}[||\mathbf{X}_k||_\infty] < \infty$. Therefore

$$||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2 \leq 4\mu^2 ||\mathbf{V}_k||_1^{2a} ||\mathbf{X}_k||_\infty^{2a}$$
$$\left[\lambda^2 + 2\lambda(1-\lambda)||\mathbf{V}_k||_1^{2a}||\mathbf{X}_k||_\infty^{2a} + 4(1-\lambda)^2||\mathbf{V}_k||_1^{4a}||\mathbf{X}_k||_\infty^{4a}\right] ||\tilde{\mathbf{X}}_k||_2^2 \quad (5.19)$$

Since we assume that $\mathbf{V}_k$ and $\mathbf{X}_k$ are independent we may write

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] \leq 4\mu^2[\lambda^2 \mathbf{E}[||\mathbf{V}_k||_1^{2a}]\mathbf{E}[||\mathbf{X}_k||_\infty^{2a}||\tilde{\mathbf{X}}_k||_2^2]$$
$$+2\lambda(1-\lambda)\mathbf{E}[||\mathbf{V}_k||_1^{4a}||\mathbf{X}_k||_\infty^{4a}||\tilde{\mathbf{X}}_k||_2^2]$$
$$+4(1-\lambda)^2\mathbf{E}[||\mathbf{V}_k||_1^{6a}]\mathbf{E}[||\mathbf{X}_k||_\infty^{6a}||\tilde{\mathbf{X}}_k||_2^2]]. \quad (5.20)$$

It is obvious but nevertheless, we also have

$$\mathbf{E}[||\mathbf{V}_k||_1^{4a}||\mathbf{X}_k||_\infty^{4a}||\tilde{\mathbf{X}}_k||_2^2] \leq \mathbf{E}[||\mathbf{V}_k||_1^{4a}]\mathbf{E}[||\mathbf{X}_k||_\infty^{4a}||\tilde{\mathbf{X}}_k||_2^2]$$

where the last equation is written using independence assumption. Now, let us consider as $x_{k-j} \triangleq ||\mathbf{X}_k||_\infty$ and use the last two equation we clearly get

$$a < \alpha/8 \quad (5.21)$$

as a sufficient condition for the boundedness of $\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2]$. In this analysis we also use the terms $\mathbf{E}[||\mathbf{V}_k||_1^{2a}]$, $\mathbf{E}[||\mathbf{V}_k||_1^{4a}]$ and $\mathbf{E}[||\mathbf{V}_k||_1^{6a}]$. Even if $\mathbf{V}_k$ is a SαS vector process, these terms are also finite with $a < \alpha/8$.

Assuming the error and the input are uncorrelated at steady state, we can obtain another bound for a:

The update equation of the RLMMN algorithm is given by

$$||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2 = 4\mu^2|e_k|^{2a}(\lambda + 2(1-\lambda)\{e_k^{<a>}\}^2)^2||\tilde{\mathbf{X}}_k||_2^2. \quad (5.22)$$

and

$$\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2] = 4\mu^2\mathbf{E}[|e_k|^{2a}(\lambda + 2(1-\lambda)\{e_k^{<a>}\}^2)^2]\mathbf{E}[||\tilde{\mathbf{X}}_k||_2^2] \quad (5.23)$$

right hand side can be reorganized as

$$4\mu^2\mathbf{E}[|e_k|^{2a}(\lambda^2 + 4\lambda(1-\lambda)|e_k|^{2a} + 4(1-\lambda)^2|e_k|^{4a})]\mathbf{E}[||\tilde{\mathbf{X}}_k||_2^2] \quad (5.24)$$

From here we see that

$$a < \alpha/6 \quad (5.25)$$

as a sufficient condition for the boundedness of $\mathbf{E}[||\mathbf{W}_{k+1} - \mathbf{W}_k||_2^2]$. Note that the first bound is tighter than the second bound. We also experimentally tested the performance of the algorithm depending on the value of the FLOS parameter a. For $a < \alpha/8$ the algorithm converges. For $\alpha/8 < a < \alpha/6$ the algorithm again converges but it requires a smaller step size μ. Finally, we also tested $a > \alpha/6$ case and observed the diverging performance of the algorithm. We expect this behaviour from the theoretical results, too.

5.3.1 Simulation Studies

In this subsection, the FLOS based RLMMN algorithm is compared to the Normalized Least Mean-p Norm (NLMP) algorithm of Chapter 4. The update equation of the NLMP algorithm is given by, [33] :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k^{<p-1>}}{||\mathbf{X}_k||_p^p + \varepsilon} \mathbf{X}_k \tag{5.26}$$

where $1 \leq p < \alpha$ and $\mu, \varepsilon > 0$ are appropriately chosen update parameters.

The algorithm of Equation (5.2) is also compared to the other proposed FLOS based algorithm of Chapter 4 whose update equation is given by, [51],[52],[53] :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k{}^{<a>}}{||\mathbf{X}_k||_{qa}^{qa}} \tilde{\mathbf{X}}_k \tag{5.27}$$

where

$$\tilde{\mathbf{X}}_k = [x_k{}^{<(q-1)a>} \ldots x_{k-M+1}{}^{<(q-1)a>}]^T,$$

the FLOS parameter $a < 1/2$ and q satisfies the relation $1/\alpha + 1/q < 1$. It is shown in Chapter 4 that the algorithms of Equation (5.26) and (5.27) have comparable performances.

In Figure 5.2 and Figure 5.3, we deal with system identification problem for $\alpha = 1.2$. We generate the AR(5) process with parameters $a_1 = 0.89$, $a_2 = -0.152$, $a_3 = 0.1$, $a_4 = -0.197$ and $a_5 = 0.097$ which is driven by an i.i.d. SαS random process, u_k, of Chapter 4. The system mismatch, $||\mathbf{W}_k - \mathbf{W}_*||_2^2$, where $\mathbf{W}_k$ and $\mathbf{W}_*$ are the current tap weight and the optimal solution vectors, respectively, is plotted. It is shown in these two figures that the performance of the RLMMN algorithm is comparable with the other two algorithms of Equation (5.26) and (5.27) depending on the value of the mixing parameter

λ. In Figure 5.2 we take $\lambda = 0.9$ and in Figure 5.3 we take $\lambda = 0.1$. For the RLMMN algorithm the FLOS parameter a is taken as 0.1 for these two figures.

With the motivation of the RLMMN algorithm, consider the mixed cost function $J_k = \lambda \mathbf{E}[|e_k|] + (1-\lambda)\mathbf{E}[|e_k|^3]$. The corresponding update equation can be as follows :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \text{sign}(e_k)[\lambda + 3(1-\lambda)\{e_k{}^{<a>}\}^2]\tilde{\mathbf{X}}_k \tag{5.28}$$

where $\tilde{\mathbf{X}}_k = [x_k{}^{<a>} ... x_{k-M+1}{}^{<a>}]^T$. With a similar theoretical analysis of Section 5.3, we derive the condition on a as :

$$a < \alpha/6. \tag{5.29}$$

For the experimental analysis, we deal with the AR(2) system identification problem of Chapter 4. We plot the transient behaviour of the tap weights for the RLMMN algorithm and the algorithm of Equation (5.28) in Figure 5.4. As it can be seen from the plot the performances of the algorithms are comparable.

All the plots are obtained by averaging 100 independent trials and to get a fair comparison between the algorithms, the step size of the algorithms are adjusted so that the steady state variances of the tap weights are equal.

We also tested the performance of the RLMMN algorithm as a function of the parameter λ for two different values of the α. In Table 5.1 we tabulated the results. In the first colomn the value of the mixing parameter, in the second colomn the convergence time for $\alpha = 1.2$, and in the third colomn for $\alpha = 1.5$ are given. For this purpose we just consider the first tap weight, i.e., $a_1 = 0.89$.

λ	*Time step*, $\alpha = 1.2$	*Time step*, $\alpha = 1.5$
0.1	6342	6201
0.3	5936	5270
0.5	5824	4931
0.7	5054	4424
0.9	4232	4145

Table 5.1: λ versus the convergence speed.

The computational complexity of the proposed algorithm of Equation (5.15) is slightly lower than those of algorithms of Equation (5.26) and (5.27) since it does not require the normalization terms of those algorithms. However, a major problem arises for this algorithm due to the following fact. When the same linear filtering operation is done on the input and the desired signal, the step size of the proposed algorithm should be rearranged whereas the algorithms of

Equation (5.26) and Equation (5.27) do not require such a rearrangement. For this purpose, a time varying step size may be used which is another research subject for the time being for us. We did not investigate further properties of this algorithm such as "Momentum" RLMMN, "Median" RLMMN and "Delayed" RLMMN algorithms before we find a time varying step size.

5.4 Conclusion

In this chapter, new adaptive filtering algorithms for impulsive noise environments are introduced. These algorithms are developed using Fractional Lower Order Statistics (FLOS) concept. The performance of the algorithms are compared to the FLOS based algorithms of Chapter 4. It is observed that the new algorithms have a comparable performance depending on the value of the mixing parameter and have a robust performance in impulsive noise environments.

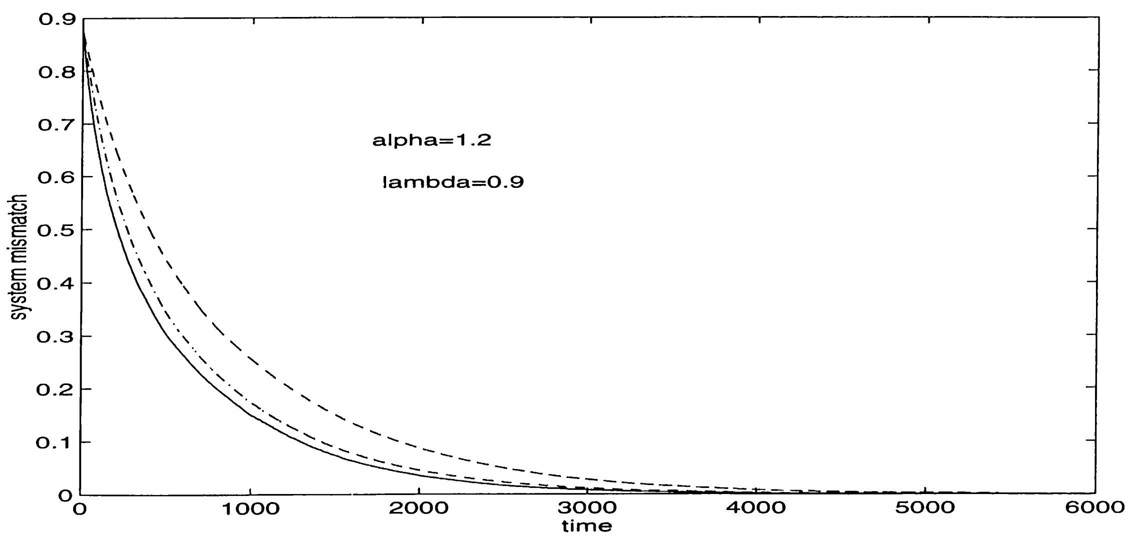


Figure 5.2: The system mismatch for RLMMN algorithm (solid), the algorithm of Equation (5.27) (dashdotted) and Equation (5.26) (dashed) for $\alpha = 1.2$ and $\lambda = 0.9$.

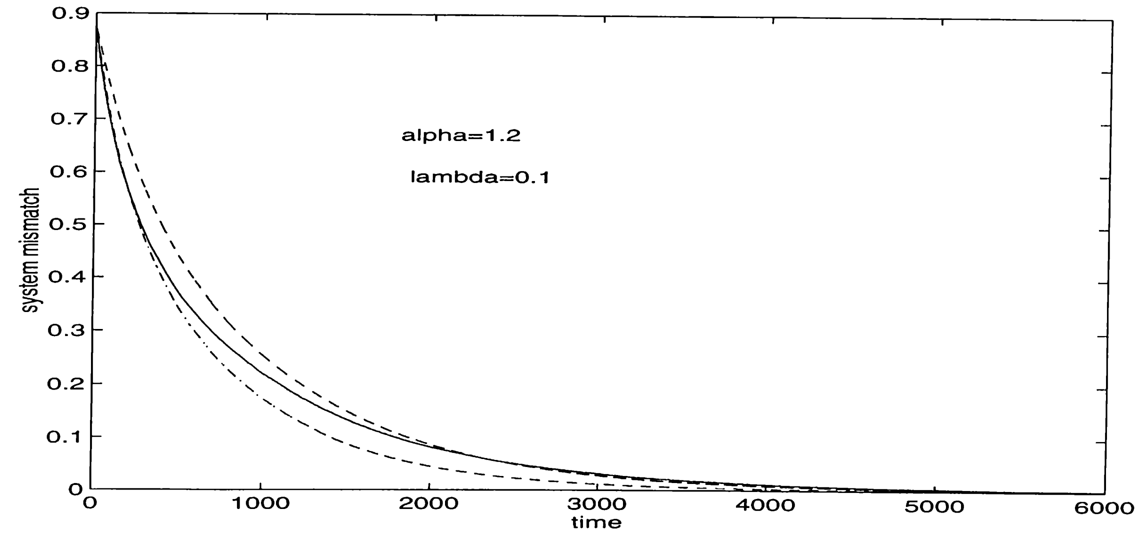


Figure 5.3: The system mismatch for RLMMN algorithm (solid), the algorithm of Equation (5.27) (dashdotted) and Equation (5.26) (dashed) for $\alpha = 1.2$ and $\lambda = 0.1$.

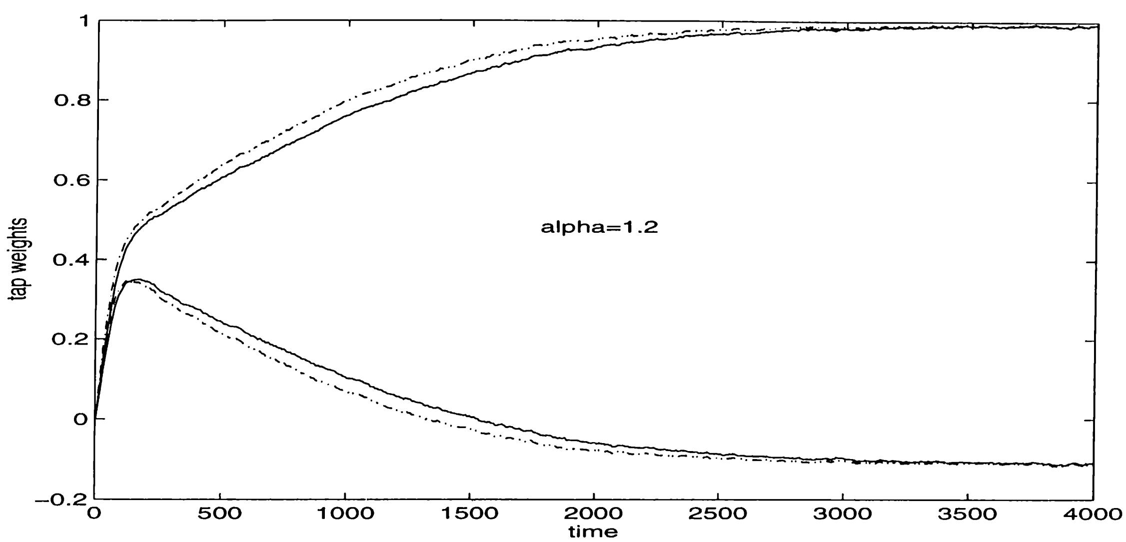


Figure 5.4: Transient behaviour of the tap weights for RLMMN algorithm (dashed) and the algorithm of Equation (5.28) (solid) for $\alpha = 1.2$ and $\lambda = 0.1$. AR(2) process parameters are $a_1 = 0.99$ and $a_2 = -0.1$.

Chapter 6

CONCLUSION

The main objective of this thesis is to develop robust adaptive filtering algorithms for impulsive noise environments. The new adaptive algorithms are based on the Fractional Lower Order Statistics (FLOS) concept.

The novel algorithms can be classified into two categories. In the first category, the generalized family of Normalized Least Mean Square algorithm [18], is robustified using FLOS. The performance of the algorithm is also compared to the Normalized Least Mean p-Norm (NLMP) algorithm. The FLOS based algorithm and the NLMP algorithm are also accelerated using the "Momentum" technique and in this way a faster convergence behaviour is achieved. Nonlinear methods are incorporated into the adaptive algorithms and their convergence behaviour is studied. It is observed that the performance of the algorithms increases by using nonlinear methods. However, there is a small bias for the estimators. It is experimentally seen that "Momentum" version of the proposed FLOS based algorithm provides the fastest convergence performance.

In the second category, the family of Least Mean Mixed Norm (LMMN) algorithm is [19] robustifed again using the FLOS concept. The performance of the proposed Robust Least Mean Mixed Norm (RLMMN) algorithm is also compared with the FLOS based algorithm of the first category and the NLMP algorithm for impulsive noise environments. Depending on the value of the mixture parameter, the RLMMN algorithm has a comparable performance with the FLOS based algorithm of the first category and the Normalized Least Mean p-Norm (NLMP) algorithm. Based on the set of the simulations, it can be deduced that the proposed algorithms have good convergence behaviour in impulsive noise environments.

The computational complexity of the proposed algorithms is not high and shown to be comparable to the NLMP algorithm.

For the future work, we would like to improve the performance of these algorithms using time varying step size. Also, we will investigate the model order selection in an extensive manner.

REFERENCES

[1] W.Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, John Wiley&Sons, 1966.

[2] E.J.Wegman and J.G.Smith, Eds., *Statistical Signal Processing.* New York: Marcel Dekker, 1984.

[3] S. A. Kassam, *Signal Detection in Non-Gaussian Noise.* New York: Springer, 1988.

[4] E.J.Wegman, S.C.Schwartz, and J.B.Thomas, Eds.,*Topics in Non-Gaussian Signal Processing.* New York: 1989.

[5] E.C. Field, Jr. and M. Lewinstein, "Amplitude-probability distribution model for VLF/ELF atmospheric noise," *IEEE Trans.Communications,* Vol. COM-26, no. 1, pp.83-87,1978.

[6] A.Giordano and F.Haber," Modeling of atmospheric noise," *Radio Science,* Vol. 7, pp.1101-1123,1972.

[7] M.P.Shinde and S.N.Gupta, "Signal detection in the presence of atmospheric noise in tropics ," *IEEE Trans.Communications,* Vol. COM-22, pp. 1055-1063, Aug. 1974.

[8] 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.

[9] J. Berger and B.Mandelbrot, " A new model for error clustering in telephone circuits," *IBM J. Research and Development,* Vol.7, pp. 224-236, 1963.

[10] 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.

[11] D.Middleton," Statistical physical models of electromagnetic interference, " *IEEE Transactions on Electromagnetic Compat.*, Vol. EMC-19, pp. 106-127, 1977.

[12] S.J. Godsill, "The Restoration of Degraded Audio Signals," Ph.D. Thesis, Department of Engineering, University of Cambridge, UK, 1993.

[13] P.Mertz, " Model of impulsive noise for data transmision," *IRE Trans. Communication Systems,* Vol. 9, no.6, pp. 130-137, 1961.

[14] H. Hall, " A New Model for Impulsive Phenomena : Application to Atmospheric-noise Communication Channel," *Technical Report* 3412-8, 66-052, Stanford University, 1966.

[15] L.A. Berry, " Understanding Middleton's canonical formulafor class a noise," *IEEE Transactions on Electromagnetic Compat.*, Vol. 23, pp.337-344, 1981.

[16] K. Furutsu and T. Ishida, " On the theory of amplitude distribution of impulsive random noise," *Journal of Applied Physics,* Vol. 32, no. 7, 1961.

[17] M.Shao and C.L.Nikias, "Signal processing with fractional lower order moments : Stable processes and their applications, " *Proc. IEEE*, Vol. 81, pp. 986-1009, 1993.

[18] S.C.Douglas, "A family of normalized LMS algorithms." *IEEE Signal Processing Letters,* Vol. 1, No. 3, March 1994.

[19] J.A. Chambers, O. Tanrıkulu and A.G. Constantinides, " Least mean mixed-norm adaptive filtering," *Electronics Letters*, Vol. 30, No. 19, pp. 1574-1575, Sep. 1994.

[20] L.Breiman, *Probability* Reading, MA: Addision-Wesley, 1968.

[21] V.Zolotarev, *One-Dimensional Stable Distributions.* Providence, RI :American Mathematical Society, 1986.

[22] E. Lukacs, *Characteristic Functions.* London : Griffin, 1960.

[23] B.V. Gnedenko and A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables.* Reading, MA : Addision-Wesley, 1968.

[24] G. Miller, "Properties of certain symmetric stable distributions", *Journal of Multivariate Anal.,* Vol. 8, pp. 346-360, 1978.

[25] V.J.Paulauskas, " Some remarks on multivariate stable distributions," *Journal of Multivariate Anal.,* Vol. 6, pp. 356-368, 1978.

[26] D.R. Holt and E.L. Crow, "Tables and graphs of stable probabilty density functions," *Journal of Research Nat. Bureau of Standard-B. Mathematical Sciences,* Vol. 77B, no.3&4,pp. 143-198, 1978.

[27] H. Bergstrom, "On some expansions of stable distribution functions," *Ark. Math.,* Vol. 2, pp. 375-378, 1952.

[28] Paul Lévy, *Calcul des Probabilités*, Paris : Gauthier-Villars.

[29] A. Weron, " Stable Processes and Measures : A survey," in *Probability Theory on Vector Spaces III,* S.Cambains, Ed. Berlin: Springer, 1983, pp. 306-364.

[30] V.M. Zolotarev, "Mellin-stieltjes transforms in probability theory," *Theory Prob. Appl.*, vol. 2, no. 4, pp. 433-460, 1957.

[31] S. Cambanis and G. Miller, " Linear problems in p^{th} order and stable processes," *SIAM J. Appl. Math.,* Vol. 41, pp. 43-69, Aug. 1981.

[32] S.J.Wolfe, "On the local behaviour of characteristic functions," *Ann. Prob.,* Vol. 1, no.5, pp. 862-866,1973.

[33] O. Arıkan, A. E. Çetin, and E. Erzin, " Adaptive filtering for non-gaussian stable processes," *IEEE Transaction on Signal Processing Letters*, Vol.1, pp.1-3, Nov.1994.

[34] R.R. Bitmead and B.D. Anderson, " Performance of adaptive estimation algorithms in dependent random environments," *IEEE Trans. Auto. Control,* Vol. AC-25, pp. 788-794, Aug. 1980.

[35] G.C.Goodwin and K.S.Sin, *Adaptive Filtering, Prediction, and Control.* Englewood Cliffs, NJ:Prentice-Hall,1984.

[36] J.Nagumo and A.Noda, "A learning method for system identification."*IEEE Trans. Automat. Contr.,* Vol. AC-12, no.3, pp. 282-287, June 1967.

[37] A.E. Albert and L.S. Gardner, Jr., " Stochastic Approximation and Nonlinear Regression," (MIT Press, 1967).

[38] A. Weiss and D. Mitra ," Some mathematical results on the effects on digital adaptive filters of implementation errors and noise,"*Proceedings*

IEEE International Conference on Acous.Speech and Signal Processing, (Tulsa, Oklohoma, April 1978), pp113-117.

[39] A. Weiss and D. Mitra ," Digital adaptive filters : Conditions for convergence, rates of convergence, effects of noise and errors arising from the implementation," *IEEE Trans. Information Theory,* Vol. IT-25, pp. 637-652, Nov. 1979.

[40] Y.Hosoya, "Discrete time stable processes and their certain properties," *Ann. Prob.*, Vol.6, No 1, pp.94-105, 1978.

[41] B.Widrow and M.E. Hoff, Jr., "Adaptive Switching Circuits," Tech. Rep. 1553-1, Stanford Electron Labs., Stanford, CA, June 1960.

[42] S.H. Cho, Y.S. Kim, and J.A.Cadzow, " Adaptive FIR filtering based on minimum L_∞-norm," in *Proc. IEEE Pacific Rim Conf. Commun. Comput. Signal Processing,* (Victoria, Canada), May 1991, pp. 643-646, Vol. 2.

[43] S.C.Douglas, "Generalized gradient adaptive step sizes for stochastic gradient adaptive filters." *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing*, Vol. 2, pp. 1396-1399, 1995

[44] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.

[45] Y.K.Shin and J.G.Lee, " A study on the fast convergence algorithm for the LMS adaptive filter design," Proc. KIEE, Vol. 19, No. 5, pp. 12-19, Oct.1985.

[46] V.J. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive filter size," *IEEE Trans. Signal Processing,* Vol. SP-41, no. 6, pp. 2075-87, June 1993.

[47] S.C. Douglas and T.H.-Y. Meng, " Stochastic gradient adaptation under general error criteria,"*IEEE Trans. Signal Processing,* Vol. SP-42, no. 6, pp. 1335-51, June 1994.

[48] S.Haykin *Adaptive Filter Theory*, Prentice Hall, NJ, 1986

[49] Treichler, Johnson, Larimore, *Theory and Design of Adaptive Filters*, New York, NY : John Wiley and Sons, 1986.

[50] O. Tanrıkulu and A.G. Constantinides, " The LMK algorithm with time-varying forgetting factor for adaptive system identification in additive output noise," *Proceedings IEEE International Conference on Acous.Speech and Signal Processing* , 1996.

[51] G. Çiydem and A.E.Çetin, "Robust adaptive filtering algorithms for α-stable distributions," *Proceedings CISS Conference on Information Science and Systems*, 1996.

[52] G. Çiydem and A.E.Çetin, "Gauss dağılımına sahip olmayan rastgele süreçler için dayanıklı uyarlamalı süzgeçleme algoritmaları," *4. Sinyal İşleme ve Uygulamaları Kurultayı.* Vol. 1, pp. 263, 1996.

[53] G. Aydın, A. E. Çetin, O. Arıkan, " Robust adaptive filtering algorithms for α-stable random processes,"*IEEE Transactions on Circuits and Systems-II*, Submitted for publication.

[54] S. Roy and J.J. Shynk, " Analysis of momentum LMS algorithm," *IEEE Trans. Acous. Speech and Signal Proc.,* Vol. 38, No. 12, pp. 2088-2098, Dec. 1990.

[55] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propogation," in *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*, D.E. Rumalhart and J.L.McClelland, Eds. Cambridge, MA: M.I.T. Press, 1986, pp. 318-362.

[56] S. A. Kassam and H. V. Poor, "Robust tecniques for signal processing: a survey, " *Proc. of IEEE,* Vol. 73, pp. 433-481, March 1985.

[57] W.A. Sethares and J.A. Bucklew " Local stability of the median LMS filter, " *IEEE Transactions on Signal Processing,* Vol . 42, pp. 2901-2905, Nov. 1994.

[58] O. Arıkan, M. Belge, A. Enis Çetin, and E. Erzin, "Adaptive filtering approaches for non-gaussian stable processes,"*Proc. IEEE International Conference on Acoustic, Speech and Signal Processing* 1995, Vol. 2, pp. 1400-1403.

[59] Murat Belge, " Investigations on Efficient Adaptation Algorithms," Master Thesis, Bilkent University, Sept. 1995.

[60] G. Long, F. Ling, and J. G. Proakis, " The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. Acous. Speech and Signal Proc.,* Vol. 37, no.9 pp. 1397-1405, Sept, 1989.

[61] X. Ma and C.L. Nikias, "Parameter estimation and blind channel identification for impulsive signal environments," *IEEE Trans. Signal Processing*, Vol. 43, pp. 2884-2897, Dec. 1995.

[62] Walach, E., and Widrow, B. : " The least mean fourth (LMF) adaptive algorithm and its family,"*IEEE Trans.,* 1984, IT-30, pp.275-283.

[63] G. Aydın, O. Tanrıkulu, A. E. Çetin, "Robust least mean mixed norm adaptive filtering for α-stable random processes," *IEEE International Symposium on Circuits and Systems, ISCAS'97*, Submitted for publication.