ROBUST ADAPTIVE ALGORITHMS FOR UNDERWATER ACOUSTIC CHANNEL ESTIMATION AND THEIR PERFORMANCE ANALYSIS

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

By
Iman Marivani
September 2017

Robust Adaptive Algorithms for Underwater Acoustic Channel Estimation and Their Performance Analysis
By Iman Marivani
September 2017

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

Süleyman Serdar Kozat(Advisor)

Tolga Mete Duman

Çağatay Candan

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

ROBUST ADAPTIVE ALGORITHMS FOR UNDERWATER ACOUSTIC CHANNEL ESTIMATION AND THEIR PERFORMANCE ANALYSIS

Iman Marivani
M.Sc. in Electrical and Electronics Engineering
Advisor: Süleyman Serdar Kozat
September 2017

We introduce a novel family of adaptive robust channel estimators for highly challenging underwater acoustic (UWA) channels. Since the underwater environment is highly non-stationary and subjected to impulsive noise, we use adaptive filtering techniques based on minimization of a logarithmic cost function, which results in a better trade-off between the convergence rate and the steady state performance of the algorithm. To improve the convergence performance of the conventional first and second order linear estimation methods while mitigating the stability issues related to impulsive noise, we intrinsically combine different norms of the error in the cost function using a logarithmic term. Hence, we achieve a comparable convergence rate to the faster algorithms, while significantly enhancing the stability against impulsive noise in such an adverse communication medium. Furthermore, we provide a thorough analysis for the tracking and steady-state performances of our proposed methods in the presence of impulsive noise. In our analysis, we not only consider the impulsive noise, but also take into account the frequency and phase offsets commonly experienced in real life experiments. We demonstrate the performance of our algorithms through highly realistic experiments performed on accurately simulated underwater acoustic channels.

Keywords: Underwater communications, robust channel estimation, logarithmic cost function, impulsive noise, performance analysis, carrier offsets.

ÖZET

SUALTİ AKUSTIK KANAL KESTIRIMINDE SAĞLAM ADAPTIF ALGORITMALAR VE PERFORMANS ANALIZI

Iman Marivani
Elektrik ve Elektronik Mühendisliği, yüksek lisans
Tez Danışmanı: Süleyman Serdar Kozat
Eylül 2017

Son derece zorlu sualtı akustik kanalları için yeni bir adaptif sağlam kanal tahmincileri ailesini sunuyoruz. Sualtı ortamı oldukça hareketli olduğundan ve dürtüsel gürültüye maruz kaldığından, logaritmik bir maliyet fonksiyonunun küçültülmesine dayanan adaptif tekrarlamalı teknikler kullanırız ki, yakınsama oranı ile algoritmanın kararlı durum performansı arasında daha iyi bir denge oluşur. Dürtüsel gürültüye ilişkin istikrar sorunlarını hafifletirken geleneksel birinci ve ikinci dereceden doğrusal tahmin yöntemlerinin yakınsama performansını arttırmak için, logaritmik bir terim kullanarak maliyet fonksiyonundaki hatanın farklı normlarını özünde birleştirdik. Bu nedenle, olumsuz bir iletişim ortamında dürtüsel gürültüye karşı istikrarı olumlu bir şekilde artırırken, daha hızlı algoritmalara kıyaslanabilir yakınsama oranı elde ediyoruz. Ayrıca, dürtüsel gürültü varlığında önerilen yöntemlerin izlenmesi ve sabit durum performansları için kapsamlı bir analiz yapıyoruz. Analizlerimizde, sadece dürtüsel gürültüyü deüil, aynı zamanda gerçek yaşam deneylerinde sıkça karşılaılan frekans ve faz uzantılarını da dikkate alıyoruz. Doğru şekilde simüle edilmiş sualtı akustik kanalları üzerinde gerçekleştirilen oldukça gerçekçi deneyler yoluyla algoritmalarımızın performansını sergiliyoruz.

Anahtar sözcükler: Sualtı iletişimi, sağlam kanal tahmini, logaritmik maliyet fonksiyonu, dürtüsel ses, performans analizi, taşıyıcı uzantılar..

Acknowledgement

I would like to thank Assoc. Prof. Suleyman Serdar Kozat for his great supervision, encouragement and huge support. I have learned so many things from him during my M.S. studies. I would like to state my deep gratitude to Prof. Tolga Mete Duman and Prof. Çağatay Candan for allocating their time to investigate my work and providing me with invaluable comments to make this thesis stronger. Also, I would like to thank all of my mentors in Bilkent University, especially, Assoc. Prof. Sinan Gezici, and Prof. Ömer Morgül, for their invaluable guidance and support during my master studies. Last but not least, I would like to dedicate this thesis to the enormous love and support of my family, my mother, father, and sisters, who had to stand my rare visits. They were always there for me and I would like to express my love and appreciation for them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Underwater acoustic (UWA) communication has attracted much attention in recent years due to proliferation of new and exciting applications such as marine environmental monitoring and sea bottom resource exploitation [2, 3, 4]. However, due to the constant movement of waves, multi-path propagation, large delay spreads, Doppler effects, and frequency dependent propagation loss [4, 5], the underwater acoustic channel is considered as one of the most detrimental communication mediums in use today [6, 5]. In these mediums, channel equalization [7, 8] plays a key role in providing reliable high data rate communication [4]. One reason of signal filtering is that we need to restrict the bandwidth of the signal, therefore, we can attain an efficient division of existing resources. Moreover, a high number of practical channels that we are dealing with in real life are bandpass, meaning that they usually have different responses to inputs with various frequencies, and UWA is an example of real life channels which we discuss here. The impulse response of a simple bandlimited channel is an ideal lowpass filter which cuases the sequential symbols intersect with neighboring symbols. Hence, there will be an intersymbol interference (ISI) decreasing the performance of the system. Equalization is a method that filters the received signal in order to eliminate the effect of ISI caused by channel impulse response (CIR) [9]. Note that, in order to combat the effects of long and time varying channel impulse response,

orthogonal frequency division multiplexing (OFDM) seems to be an elegant solution [10]. However, one needs an accurate estimate of the time varying channel to be used for OFDM as well as equalization. Furthermore, due to rapidly changing and unpredictable nature of underwater environment, such processing should be adaptive [4]. Nevertheless, the impulsive nature of the ambient noise in UWA channels introduces stability issues for adaptive channel estimation [11, 12]. To this end, we propose and analyze the performance of a family of new adaptive linear channel estimators, which provides a relatively fast convergence rate as well as strong stability against the ambient noise in the UWA channels.

Although the additive white Gaussian noise (AWGN) model is widely used in digital and wireless communication contexts, this model is insufficient to appropriately address the ambient noise in UWA channels [13, 14, 15]. For example, in warm shallow waters, the high frequency noise component is dominated by the "impulsive noise" [16, 17, 18, 19] resulted from numerous noise sources such as marine life, shipping traffic, underwater explosives, and offshore oil exploration-production [20]. The impulsive noise consists of relatively short duration, infrequent, high amplitude noise pulses. Here, in order to rectify the undesirable effects of UWA channels, especially to mitigate the effects of the impulsive noise, we introduce a radical approach to adaptive channel estimation.

In [21], the authors propose a low-complexity decision feedback equalizer, which employs a sphere detection-Viterbi algorithm (SDVA) with two radii in a decision feedback equalization (DFE). [21] provides a simple operating condition for the SDVA, under which the algorithm only searches a small fraction of the lattice points lying between two radii, hence, provides a lower complexity. However, the impulsive noise in underwater channels degenerates the performance of the SDVA-algorithm. Thus, some preprocessing steps can reduce the impulsive noise effect and improve the performance. Also, in [22], for localization of multiple acoustic sources in an impulsive-noise environment, the authors employ a data-adaptive zero-memory nonlinear preprocessor to process each sample of the contaminated data to obtain less-noisy data. Specifically, if the sample's value is less than a threshold, they assume it is not affected by the impulsive noise, hence, its value does not change during the process. However, samples above

the threshold are assumed to be corrupted by a large noise impulse and are suppressed. Note that the threshold is determined based on the noise statistics, hence, one has to estimate the noise statistics.

Linear adaptive channel estimation methods (e.g., sign algorithm (SA), least mean squares (LMS) or least mean fourth (LMF) algorithms [23, 24]) are the simplest as well as low complexity methods. However, the conventional linear estimators either have a slow convergence speed (e.g., sign algorithm (SA)) or suffer from stability issues (e.g., LMF) in impulsive noise environments, hence, cannot fully address the problem [25]. These methods are commonly based on minimization of a cost function of the form $C(e_m) \triangleq E\left[|e_m|^k\right]$ (where $E[.]$ indicates the expectation), using the stochastic gradient descent method [26, 27]. However, there is always a trade-off between the "robustness" of such algorithms against impulsive noise and their convergence speed [25]. In this sense, the algorithms that use lower powers of the error as the cost function (e.g., SA [27]) provide a better robustness, while, the algorithms based on higher powers of the error, usually exhibit faster convergence [25].

The mixed norm algorithms, combining different norms of the error in the cost function, are used to achieve a better trade-off between robustness and convergence speed [28, 29]. Nevertheless, optimization of the combination parameters in such algorithms needs "a priori" knowledge of the noise and input signal statistics [25]. On the contrary, the mixture of experts algorithms [30, 31, 32], adaptively learn the best combination parameters. However, such algorithms are infeasible in UWA scenarios due to the high computational complexity resulted from running several different algorithms in parallel [25, 33].

Recently, in [34] and [35], the authors proposed recursive least squares (RLS)-type robust adaptive estimation and equalization methods, which leverage the sparsity of the underwater channels by adding an l_0-norm to the cost function. However, the methods in [34] and [35] are not completely adaptive in the sense that they need a few threshold values to be determined in advance, and the values of these thresholds are highly dependent on the transmitted data statistics. In [36], a hyperbolic function, e.g., $C(e_m) = \cosh(e_m)$, is used as the cost function

to inherently combine different powers of the error. Moreover, [37] uses a sparse least mean p-power (LMP) algorithm to adaptively provide a robust estimate of the underwater channel.

In this thesis, in order to obtain a comparable performance to the robust algorithms, while retaining the fast convergence of conventional least square methods, we use a logarithmic function as a regularization term in the cost function of the well-known adaptive methods. In this sense, we choose a conventional method that uses a power of the error as the cost, e.g, least mean squares (LMS), and improve that method through adding a logarithmic term to its cost function. Due to the characteristics of the logarithmic functions, when the error is high, e.g., when there is an impulse, the cost function resembles the cost function of the original method, while for the small errors a correction term is added. The correction term includes the higher powers of the error, which yields a faster convergence. Hence, we intrinsically mitigate the effect of the impulsive noise pulses and provide an improved robustness, while increase the convergence speed when there is no impulsive noise.

Specifically, we present first and second order channel estimation algorithms using the logarithmic cost. The first order methods, logarithmic cost least mean absolute (LCLMA) and logarithmic cost least mean squares (LCLMS), are based on the first and second powers of the error as in SA and LMS. Similarly, we use the first and second powers of the error to introduce logarithmic cost recursive least squares (LCRLS), and logarithmic cost recursive least absolutes (LCRLA), which are second order (RLS-type) algorithms. An alternative design could use a conventional equalizer, aided by a preprocessing step that performs impulsive noise reduction. For instance, if OFDM is used, the probability density function of the noise-free signal is approximately Gaussian, since an OFDM signal consists of a sum of many independent and identically distributed signals and due to the Central Limit Theorem, results in a nearly Gaussian distributed signal [38]. Therefore, a Bayesian minimum mean squared error (MMSE) approach, applied to the signal-plus-impulsive-noise, would reduce the impulsive noise [39, 40]. Note that this approach would maximize the SNR at the output of the preprocessing step [40]. After reducing the impulsive noise in the preprocessing step, a

conventional equalizer (designed for Gaussian noise) could be employed.

Moreover, we provide a thorough analysis of the tracking and steady state performance of our algorithms. In these analyses, in order to be consistent with the real world UWA channels, we assume that there are frequency and phase offsets as well as impulsive noise in the channel. For the first order algorithms, we avoid mentioning all of the intermediate steps of the analyses, instead we discuss the important steps, use the results of [41], and establish the final results based on them. However, for the second order analyses, we completely explain the steps since it is a nontrivial extension of the existing analysis methods. We show the improved performance of our algorithms and the correctness of our analyses through highly realistic simulations.

The thesis is organized as follows: In chapter 2, we introduce the notation and describe the problem mathematically. Then, in chapter 3, we provide a family of stable and fast converging channel estimators based on logarithmic cost functions. We provide the performance analysis for our methods in chapter 4. We demonstrate the performance of the presented methods through highly realistic simulations in chapter 5 and conclude the thesis with several remarks.

Chapter 2

Problem description

2.1 Notations

All vectors are column vectors and are denoted by boldface lower case letters and all matrices are denoted by boldface upper case letters. For a vector $\boldsymbol{x}$, $\boldsymbol{x}^H$ is the Hermitian transpose. We denote the all zero vector of size $L \times 1$ by $\mathbf{0}_L$. In addition, $E[x]$ denotes the expectation of the random variable x, and $\mathrm{Tr}(\boldsymbol{A})$ denotes the trace of the matrix $\boldsymbol{A}$. Also, the discrete time variables are shown as subscripts, while the continuous time variable t is parenthesized. Furthermore, for a vector $\boldsymbol{w}$ the squared l_2-norm is defined as $\|\boldsymbol{w}\|^2 \triangleq \boldsymbol{w}^H\boldsymbol{w}$, and the weighted squared l_2-norm is defined as $\|\boldsymbol{w}\|^2_{\boldsymbol{P}} \triangleq \boldsymbol{w}^T\boldsymbol{P}\boldsymbol{w}$, where $\boldsymbol{P}$ is a positive definite weighting matrix.

2.2 Setup

As depicted in Fig. 2.1, we denote the received signal by $r(t)$, $r(t) \in \mathbb{R}$, and our aim is to determine the transmitted symbols $\{s_m\}_{m\geq 1}$. To transmit the symbols $\{s_m\}_{m\geq 1}$, we use the raised cosine pulse shaping filter $g(t)$, which generates the

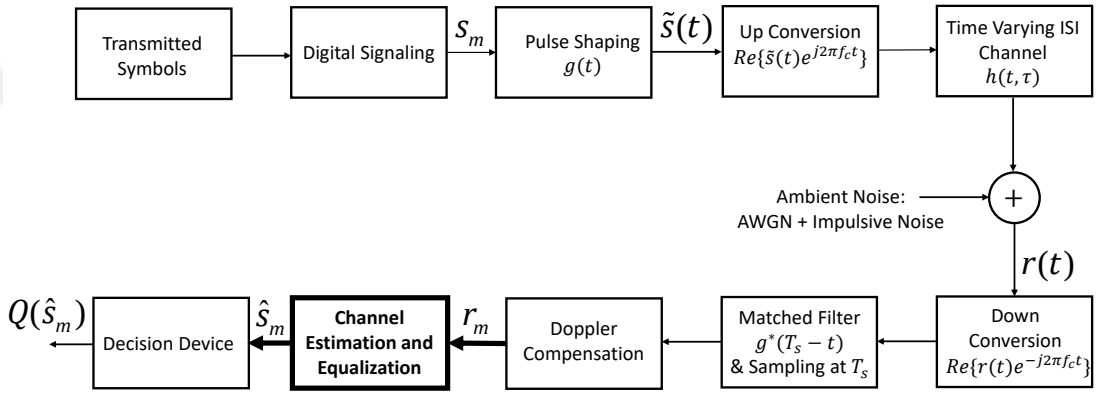

Figure 2.1: The block diagram of the model we use for the transmitted and received signals. The transmitted data $\{s_m\}_{m\geq 1}$ are modulated, and after pulse shaping with a raised cosine filter $g(t)$ and up-conversion to the carrier frequency f_c, pass through a time varying intersymbol interference (ISI) channel $h(t,\tau)$. The received signal is the output of the ISI channel contaminated with the ambient noise $n(t)$. The equalizer is fed with r_m and generates the soft output $\hat{s}_m$. $Q(\hat{s}_m)$ is the hard estmation for the m^{th} transmitted symbol. We use the hard estimates $Q(\hat{s}_m)$ for adapting the channel estimator, and use the channel estimation to reduce the ISI.

signal $\tilde{s}(t)$, and then up-convert the signal to the carrier frequency f_c, and send it through the channel. Using the linear time varying convolution between $\tilde{s}(t)$ and $h(t,\tau)$ [42], the received signal at time t is

$$r(t) = \int_0^{\tau_{max}} \tilde{s}(t-\tau)h(t,\tau)d\tau + n(t), \tag{2.1}$$

where $\tilde{s}(t)$ is the transmitted signal after pulse shaping at time t. $h(t,\tau)$ indicates the CIR at time t corresponding to the impulse sent at time $t-\tau$ and τ_{max} shows the maximum delay spread of the channel. Also, $n(t)$ is the ambient noise of the channel, which is represented as

$$n(t) = n_g(t) + n_i(t), \tag{2.2}$$

where $n_g(t)$ indicates the white Gaussian part of the noise and $n_i(t)$ indicates the impulsive part. Note that the effects of time delay and phase deviations are usually addressed at the front-end of the receiver, hence, in this thesis, we do not deal with these problems explicitly. We then sample the received signal every T_s seconds (T_s is the symbol duration) such that the discrete time channel model is represented as [42]

$$r_m = \sum_{k=0}^{L-1} s_{m-k}h_{m,k} + n_m, \tag{2.3}$$

and

$$n_m = n_{g,m} + n_{i,m}, \tag{2.4}$$

where $L = \lfloor \tau_{max}/T_s \rfloor$ indicates the length of the CIR. In the following chapters, we use the discrete time model of the channel to address the estimation and equalization problems. Note that (2.3) is a causal representation of the channel, i.e., the channel output at time m depends only on the transmitted symbols before m.

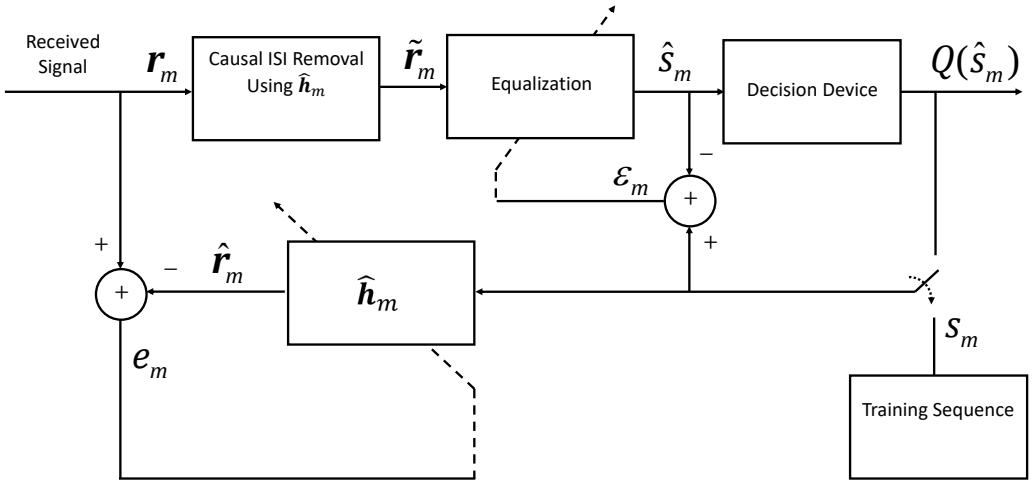

Figure 2.2: The block diagram of our adaptive channel estimator and equalizer. We present new algorithms for the channel estimation block (which determines how should $\hat{\boldsymbol{h}}_m$ be updated based on the error amount e_m), which yields improved performance over the conventional methods in the impulsive noise environments.

2.3 Channel Estimation and Causal ISI Removal

Our aim is to estimate the communication channel, $\boldsymbol{h}_m$, and based on that, estimate the transmitted symbols $\{s_m\}_{m\geq 1}$ according to the channel outputs $\{r_m\}_{m\geq 1}$. We introduce a new adaptation method to obtain an accurate estimate of the causal part of the channel, denoted by $\hat{\boldsymbol{h}}_m$. Mathematically, the output of the estimated channel is $\hat{r}_m = \hat{\boldsymbol{h}}_m^T \boldsymbol{s}_m$, where $\boldsymbol{s}_m \triangleq [s_m, \ldots, s_{m-L+1}]^T$ is the vector of the current and past $L-1$ transmitted symbols. We use the error $e_m = r_m - \hat{r}_m$ to adapt the estimated channel after processing each sample, as shown in Fig. 2.2. As depicted in Fig. 2.2, $\boldsymbol{s}_m$ consists of the hard estimates of the transmitted symbols, i.e., in the training mode we use s_m, however, in the decision directed mode we use the quantized estimates $Q(\hat{s}_m)$, as the input to the channel estimator.

As depicted in Fig. 2.2, we first remove the inter-symbol interference (ISI) effect generated by the causal part of the channel, i.e., $\boldsymbol{h}_m$, to obtain "cleaned" symbols $\tilde{r}_m$. We then use the past N_c cleaned symbols at each time, in the causal part of the equalizer, to further reduce the ISI effect. Since the channel has a length of L, in order to obtain cleaned symbols $\tilde{r}_m$, we must remove the effect of the past $L-1$ symbols from each received symbol r_m, based on the estimated channel at time m. Therefore, inserting in a matrix form, we have

$$\tilde{\boldsymbol{r}}_{m+1} = \boldsymbol{r}_{m+1} - \hat{\boldsymbol{H}}_m[s_{m-N_c-L+2}, ..., s_m]^T, \tag{2.5}$$

where $\hat{\boldsymbol{H}}_m$ is the channel convolutional matrix defined as

$$\hat{\boldsymbol{H}}_m \triangleq \begin{pmatrix} \hat{\boldsymbol{h}}_{m-N_c+1}^T & & \boldsymbol{0}_{N_c-1}^T \\ 0 & \hat{\boldsymbol{h}}_{m-N_c+2}^T & \boldsymbol{0}_{N_c-2}^T \\ \vdots & \ddots & \vdots \\ \boldsymbol{0}_{N_c-1}^T & & \hat{\boldsymbol{h}}_m^T \end{pmatrix}_{N_c \times (N_c+L-1)}.$$

In addition, $\boldsymbol{r}_m = [r_{m-N_c+1}, ..., r_m]$ is the $N_c \times 1$ received signal vector, and $\tilde{\boldsymbol{r}}_m = [\tilde{r}_{m-N_c+1}, ..., \tilde{r}_m]$ is the cleaned version of the received signal vector ($\boldsymbol{r}_m$) to be used as the input of the causal part of the equalizer.

In order to update the channel estimate, i.e., $\hat{\boldsymbol{h}}_m$, a cost function $C(e_m)$ is defined, e.g., in LMS method the cost is defined as $C(e_m) = E[|e_m|^2]$. Then, we derive an algorithm for updating $\hat{\boldsymbol{h}}_m$ based on minimization of this cost function using the stochastic gradient descent method [27]. Therefore, we update the tap weights vector as follows.

$$\hat{\boldsymbol{h}}_{m+1} = \hat{\boldsymbol{h}}_m - \frac{1}{2}\mu \nabla_{\hat{\boldsymbol{h}}_m} \tilde{C}(e_m),$$

where $\nabla_{\hat{\boldsymbol{h}}_m} \tilde{C}(e_m)$ denotes an instantaneous approximation of the gradient of the cost function $C(e_m)$ with respect to $\hat{\boldsymbol{h}}_m$, i.e., the gradient obtained by removing the expectation and taking the gradient of the term inside [27]. As an example, in LMS method $C(e_m) = E[|e_m|^2]$, hence we update the tap weights vector of an LMS estimator as

$$\hat{\boldsymbol{h}}_{m+1} = \hat{\boldsymbol{h}}_m + \mu e_m \boldsymbol{s}_m, \tag{2.6}$$

where $\mu > 0$ is the learning rate. Note that according to [27], for the LMS algorithm the learning rate (step size) should satisfy the following inequality:

$$\mu \leq \frac{2}{\lambda_{max}}$$

where λ_{max} shows the maximum eigenvalue of the signal covariance matrix.

The updating expression in (2.6) is the well-known LMS update. However, we seek to provide a more robust updating algorithm through minimization of a different cost function. Note that the cost functions of the form $C(e_m) = E[|e_m|^k]$, which consist of only the k^{th} power of the error, either have a slow convergence, or do not perform well, from the stability viewpoint, in an impulsive noise environment [25]. Therefore, in chapter 3, we use a logarithmic term in the cost function, which intrinsically introduces different powers of the error into the cost function. As a result, when an impulsive error occurs, the algorithm mitigates the effect of that sample in updating the equalizer coefficients by simply using the lower order norms of the error, whereas in impulse-free environments the algorithm accelerates the convergence using higher order norms of the error.

2.4 Channel Equalization

To obtain the estimate $\hat{s}_m$, we use a linear channel equalizer, which is mathematically represented as $\hat{s}_m = \boldsymbol{w}_m^T \check{\boldsymbol{r}}_m$, where $\check{\boldsymbol{r}}_m \triangleq [\tilde{r}_{m-N_c+1}, \ldots, \tilde{r}_m, r_{m+1}, \ldots, r_{m+N_a}]^T$, N_a and N_c represent the lengths of anti-causal and causal equalizers, respectively, and $\boldsymbol{w}_m \triangleq [w_{-N_c+1,m}, \ldots, w_{N_a,m}]^T$ is the tap weights vector of the linear equalizer at time m. We define the error in estimating the transmitted symbol s_m as $\epsilon_m = s_m - \hat{s}_m$. We then use the conventional normalized LMS (NLMS) method to update the equalizer. In the rest of the thesis, we focus on the channel estimation part of the system.

Chapter 3

Logarithmic Cost Functions

In this chapter, we explain our method mathematically and introduce two channel estimators based on the logarithmic costs [25]. Thus, we define the new cost as

$$C(e_m) \triangleq \Phi(e_m) - \frac{1}{a}\ln(1 + a\Phi(e_m)), \tag{3.1}$$

where $a > 0$ is a design parameter. From the properties of the logarithmic function [25], we deduce that when $a\Phi(e_m) \leq 1$

$$C(e_m) \to \Phi(e_m) - \frac{1}{a}(\ a\Phi(e_m) - \frac{a^2}{2}\Phi^2(e_m) + \frac{a^3}{3}\Phi^3(e_m) + \dots)$$

$$C(e_m) \to \frac{a}{2}\Phi^2(e_m) - \frac{a^2}{3}\Phi^3(e_m) + \dots$$

and for the big errors, e.g, the presence of impulsive noise, if $e_m \to \infty$, Then, $\Phi(e_m) \to \infty$ and note that the linear function grows faster than the logarithmic function. Then we have,

$$C(e_m) \to \Phi(e_m), \quad \text{as} \ \ e_m \to \infty.$$

It is straightforward to show that the new cost function $C(e_m)$ is a convex

function of e_m, if the cost function $\Phi(e_m)$ is convex. Therefore, in the new algorithm, we seek to minimize a cost function that is mainly consisted of the first and second powers of the primary cost function $\Phi(e_m)$, based on the error amount.

3.1 First Order Methods

In this section, we use first order methods to adaptively adjust the channel estimate $\hat{\boldsymbol{h}}_m$. For this purpose, we use the stochastic gradient method [27] to derive a recursive expression for updating $\hat{\boldsymbol{h}}_m$. This yields,

$$\begin{aligned}\hat{\boldsymbol{h}}_{m+1} &= \hat{\boldsymbol{h}}_m - \frac{1}{2}\mu\nabla_{\hat{\boldsymbol{h}}_m} C(e_m) \\ &= \hat{\boldsymbol{h}}_m - \mu\, \frac{a\Phi(e_m)}{1+a\Phi(e_m)} \left[\nabla_{\hat{\boldsymbol{h}}_m} \Phi(e_m)\right].\end{aligned}$$

Particularly, in this framework, we adopt a well-known cost function, e.g., $\Phi(e_m) = E[e_m^2]$, as the primary cost function $\Phi(e_m)$. Suppose that $\Phi(e_m)$ is the expectation of another function $\varphi(e_m)$, i.e., $\Phi(e_m) = E[\varphi(e_m)]$. Since we do not know the exact function, we use instantaneous approximation which is the use of the realizations of the random function. Then, using the instantaneous approximation for $\Phi(e_m)$ [27], the general stochastic gradient update is given by

$$\hat{\boldsymbol{h}}_{m+1} = \hat{\boldsymbol{h}}_m + \mu\, \frac{a\varphi(e_m)}{1+a\varphi(e_m)} \left[\nabla_{e_m}\varphi(e_m)\right] \boldsymbol{s}_m. \tag{3.2}$$

We present two channel estimation methods based on the introduced approach. We then show the superior performance of these methods through highly realistic experiments in chapter 5.

1. *The logarithmic cost least mean squares (LCLMS)*:
   Here, we adopt $\varphi(e_m) = e_m^2$, which (according to (3.2)) yields the following

update on the tap weights vector

$$\begin{aligned}\hat{\boldsymbol{h}}_{m+1} &= \hat{\boldsymbol{h}}_m + \mu \, \frac{a \, e_m^2}{1 + a \, e_m^2} \, [2e_m] \, \boldsymbol{s}_m \\ &= \hat{\boldsymbol{h}}_m + \mu' \, \frac{e_m^3}{1 + a \, e_m^2} \boldsymbol{s}_m.\end{aligned}$$

2. *The logarithmic cost least mean absolutes (LCLMA)*:
   In this case, we adopt $\varphi(e_m) = |e_m|$, and according to (3.2), obtain the following update on the tap weights vector,

$$\hat{\boldsymbol{h}}_{m+1} = \hat{\boldsymbol{h}}_m + \mu \, \frac{a \, |e_m|}{1 + a \, |e_m|} \, [\text{sign}(e_m)] \, \boldsymbol{s}_m$$

As shown in the simulations, the LCLMS algorithm results in an improved convergence speed over the conventional LMS algorithm, while achieving the comparable stability to LMS method. Similarly, we achieve an improved convergence speed performance over the conventional SA algorithm by using LCLMA algorithm, while preserving the robustness against impulsive noise. Hence, these are elegant alternatives to the conventional methods for UWA channel estimation.

3.2 Second Order Methods

By defining the cost function as $J_m \triangleq \sum_{i=0}^{m} \lambda^{m-i} C(e_i)$, we seek to update $\hat{\boldsymbol{h}}_m$ in order to minimize J_m. Therefore, by solving $\nabla_{\hat{\boldsymbol{h}}} J_m\big|_{\hat{\boldsymbol{h}}=\hat{\boldsymbol{h}}_{m+1}} = 0$, we have

$$\begin{aligned}
\nabla_{\hat{\boldsymbol{h}}} J_m &= \sum_{i=0}^{m} \lambda^{m-i} \nabla_{\hat{\boldsymbol{h}}} C(e_i) \\
&= \sum_{i=0}^{m} \lambda^{m-i} \frac{\partial C(e_i)}{\partial \varphi(e_i)} \frac{\partial \varphi(e_i)}{\partial e_i} \nabla_{\hat{\boldsymbol{h}}} e_i \\
&= \sum_{i=0}^{m} \lambda^{m-i} \frac{a\varphi(e_i)}{1+a\varphi(e_i)} \frac{\partial \varphi(e_i)}{\partial e_i} (-\boldsymbol{s}_i) \\
&= \sum_{i=0}^{m} \lambda^{m-i} \frac{a\frac{\varphi(e_i)}{e_i}}{1+a\varphi(e_i)} \frac{\partial \varphi(e_i)}{\partial e_i} (-\boldsymbol{s}_i) e_i \\
&= \sum_{i=0}^{m} \lambda^{m-i} w(e_i)(-\boldsymbol{s}_i)(r_i - \boldsymbol{s}_i^H \hat{\boldsymbol{h}}),
\end{aligned} \tag{3.3}$$

where $w(e_i) = \dfrac{a\frac{\varphi(e_i)}{e_i}}{1+a\varphi(e_i)} \dfrac{\partial \varphi(e_i)}{\partial e_i}$ is considered as the weight of the i^{th} sample. Finally, the equation $\nabla_{\hat{\boldsymbol{h}}} J_m\big|_{\hat{\boldsymbol{h}}=\hat{\boldsymbol{h}}_{m+1}} = 0$ yields the following equation for $\hat{\boldsymbol{h}}_{m+1}$

$$\sum_{i=0}^{m} \lambda^{m-i} w(e_i) \boldsymbol{s}_i \boldsymbol{s}_i^H \hat{\boldsymbol{h}}_{m+1} = \sum_{i=0}^{m} \lambda^{m-i} w(e_i) \boldsymbol{s}_i r_i,$$

Hence [26]

$$\hat{\boldsymbol{h}}_{m+1} = \boldsymbol{\Omega}_m^{-1} \boldsymbol{\psi}_m,$$

where $\boldsymbol{\Omega}_m = \sum_{i=0}^{m} \lambda^{m-i} w(e_i) \boldsymbol{s}_i \boldsymbol{s}_i^T$ and $\boldsymbol{\psi}_m = \sum_{i=0}^{m} \lambda^{m-i} w(e_i) \boldsymbol{s}_i r_i$. We also observe that

$$\boldsymbol{\Omega}_m = \lambda \boldsymbol{\Omega}_{m-1} + w(e_m) \boldsymbol{s}_m \boldsymbol{s}_m^H, \tag{3.4}$$

and

$$\boldsymbol{\psi}_m = \lambda \boldsymbol{\psi}_{m-1} + w(e_m) \boldsymbol{s}_m r_m. \tag{3.5}$$

Thus, by using the matrix inversion lemma [27], $\boldsymbol{\Omega}_m^{-1}$ can be calculated as follows.

$$\boldsymbol{\Omega}_m^{-1} = \lambda^{-1}\left(\boldsymbol{\Omega}_{m-1}^{-1} - \frac{1}{\frac{\lambda}{w(e_m)} + \boldsymbol{s}_m^H\boldsymbol{\Omega}_{m-1}^{-1}\boldsymbol{s}_m}\,\boldsymbol{\Omega}_{m-1}^{-1}\boldsymbol{s}_m\boldsymbol{s}_m^H\boldsymbol{\Omega}_{m-1}^{-1}\right).$$

Now by defining $\boldsymbol{P}_m \triangleq \boldsymbol{\Omega}_m^{-1}$, and $\boldsymbol{g}_m \triangleq \frac{w(e_m)\boldsymbol{P}_{m-1}\boldsymbol{s}_m}{\lambda + w(e_m)\boldsymbol{s}_m^H\boldsymbol{P}_{m-1}\boldsymbol{s}_m}$, we have

$$\boldsymbol{P}_m = \lambda^{-1}(\boldsymbol{P}_{m-1} - \boldsymbol{g}_m\boldsymbol{s}_m^H\boldsymbol{P}_{m-1}), \tag{3.6}$$

and rearranging the terms in the definition of $\boldsymbol{g}_m$ yields

$$\begin{aligned}\boldsymbol{g}_m &= \frac{w(e_m)}{\lambda}\left[\boldsymbol{P}_{m-1} - \boldsymbol{g}_m\boldsymbol{s}_m^H\boldsymbol{P}_{m-1}\right]\boldsymbol{s}_m \\ &= w(e_m)\boldsymbol{P}_m\boldsymbol{s}_m.\end{aligned} \tag{3.7}$$

Substituting (3.5) and (3.6) in the expression $\hat{\boldsymbol{h}}_{m+1} = \boldsymbol{P}_m\boldsymbol{\psi}_m$ and expanding it results in

$$\begin{aligned}\hat{\boldsymbol{h}}_{m+1} &= \lambda\boldsymbol{P}_m\boldsymbol{\psi}_{m-1} + w(e_m)\boldsymbol{P}_m\boldsymbol{s}_m r_m \\ &= (\boldsymbol{P}_{m-1} - \boldsymbol{g}_m\boldsymbol{s}_m^H\boldsymbol{P}_{m-1})\boldsymbol{\psi}_{m-1} + \boldsymbol{g}_m r_m \\ &= \hat{\boldsymbol{h}}_m + \boldsymbol{g}_m(r_m - \boldsymbol{s}_m^H\hat{\boldsymbol{h}}_m).\end{aligned} \tag{3.8}$$

Hence, the final second order updating algorithm is

$$\begin{aligned}e_m &= r_m - \boldsymbol{s}_m^H\hat{\boldsymbol{h}}_m, \\ \boldsymbol{g}_m &= \frac{w(e_m)\boldsymbol{P}_{m-1}\boldsymbol{s}_m}{\lambda + w(e_m)\boldsymbol{s}_m^H\boldsymbol{P}_{m-1}\boldsymbol{s}_m}, \\ \hat{\boldsymbol{h}}_{m+1} &= \hat{\boldsymbol{h}}_m + e_m\boldsymbol{g}_m, \\ \boldsymbol{P}_m &= \lambda^{-1}\left(\boldsymbol{P}_{m-1} - \boldsymbol{g}_m\boldsymbol{s}_m^H\boldsymbol{P}_{m-1}\right),\end{aligned} \tag{3.9}$$

where $\boldsymbol{P}_0 = v^{-1}\boldsymbol{I}$, and $0 < v \ll 1$. We point out that for the $\varphi(e_m) = |e_m|^2$, we have $w(e_m) = \frac{2a\ |e_m|^2}{1+a\ |e_m|^2}$. However, according to (3.3) multiplying the $w(e_m)$ by a constant term does not affect the algorithm. Therefore, we use $w(e_m) = \frac{|e_m|^2}{1+a\ |e_m|^2}$ in (3.9) to obtain the logarithmic cost recursive least squares (LCRLS) algorithm. Moreover, by using $\varphi(e_m) = |e_m|$, we achieve $w(e_m) = \frac{1}{1+a\ |e_m|}$, which results in

the logarithmic cost recursive least absolutes (LCRLA) algorithm.

Chapter 4

Performance Analysis

In this chapter we provide a thorough analysis for the mean square error (MSE) performance of the proposed methods. Since the channel is highly time-varying, we use the notion of excess MSE (EMSE) as in [27]. Furthermore, in order to be more realistic, we assume carrier frequency offset as well as impulsive ambient noise. We first present the impulsive noise model used in the analyses, then we establish our analyses based on the widely used assumptions in the literature. Note that although for the sake of notational simplicity, we assume $a = 1$ in all algorithms, the results can be readily extended to other values of a.

In order to analyze the tracking performance of the introduced algorithms, we assume a random walk model [27] for the discrete channel vector $\boldsymbol{h}_m$ that yields the minimum mean squared error such that

$$\begin{aligned} \boldsymbol{h}_m &= \boldsymbol{h} + \boldsymbol{\theta}_m, \\ \boldsymbol{\theta}_{m+1} &= \alpha \boldsymbol{\theta}_m + \boldsymbol{q}_m. \end{aligned} \tag{4.1}$$

Moreover, due to the carrier offsets, we consider the following model for the received data

$$r_m = \boldsymbol{s}_m^H \boldsymbol{h}_m e^{j\Omega m} + n_m. \tag{4.2}$$

We define $\tilde{\boldsymbol{h}}_m \triangleq \boldsymbol{h}_m e^{j\Omega m} - \hat{\boldsymbol{h}}_m$ and $\boldsymbol{q}_m \in \mathbb{R}^L$ is a zero-mean vector process with

covariance matrix $E[\boldsymbol{q}_m \boldsymbol{q}_m^H] = \boldsymbol{Q}$.

$$\text{MSE} = \sigma_n^2 + \lim_{m\to\infty} E|\boldsymbol{s}_m^H \tilde{\boldsymbol{h}}_m|^2. \tag{4.3}$$

We define the EMSE as

$$\zeta \triangleq \lim_{m\to\infty} E|\boldsymbol{s}_m^H \tilde{\boldsymbol{h}}_m|^2 = \lim_{m\to\infty} E|e_{a,m}|^2. \tag{4.4}$$

We next introduce the impulsive noise model we use in our analyses.

Impulsive noise model: Since the received signal is subjected to impulsive noise, we model the estimation noise as a summation of two independent random terms [43, 44] as

$$n_m = n_{g,m} + z_m \, n_{i,m},$$

where $n_{g,m}$ is the ordinary AWGN noise signal that is zero-mean Gaussian with variance σ_g^2 and $n_{i,m}$ is the impulsive part of the noise that is also a zero-mean Gaussian random process with a significantly large variance σ_i^2. Here, z_m is generated through a Bernoulli random process and determines the occurrence of the impulses in the noise signal with $p_Z(z_m = 1) = \nu_i$ and $p_Z(z_m = 0) = 1 - \nu_i$, where ν_i is the frequency of the impulses. The overall probability density function of the noise signal n_m is given by

$$p_n(n_m) = \frac{1-\nu_i}{\sqrt{2\pi}\sigma_g} \exp\left(-\frac{n_m^2}{2\sigma_g^2}\right) + \frac{\nu_i}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{n_m^2}{2\sigma_n^2}\right),$$

where $\sigma_n^2 = \sigma_g^2 + \sigma_i^2$.

4.1 MSE Analysis of the First Order Methods

Assuming that $\boldsymbol{q}_m$ is independent from the received and noise signals, at the steady state, we have the following general variance relation for an adaptive filter

with the error nonlinearity $g(e_m)$ [41, 27],

$$\begin{aligned}2\mathrm{Re}\left\{E\left[e_{a,m}\ g(e_m)\right]\right\} = \mu E\left[\|\boldsymbol{s}_m\|^2|g(e_m)|^2\right] + \mu^{-1}\mathrm{Tr}(\boldsymbol{Q}) \\ + \mu^{-1}|1-e^{j\Omega}|^2\|\boldsymbol{h}\|^2 + \mu^{-1}|1-\alpha e^{j\Omega}|^2\mathrm{Tr}(\boldsymbol{\Theta}) \\ - 2\mu^{-1}\mathrm{Re}\left\{(1-e^{-j\Omega})\boldsymbol{h}E\left[(\tilde{\boldsymbol{h}}_{m-1}-\mu\boldsymbol{s}_m g(e_m))e^{-j\Omega(m-1)}\right]\right\} \\ - 2\mu^{-1}\mathrm{Re}\left\{(1-\alpha^* e^{-j\Omega})\boldsymbol{h}E\left[\boldsymbol{\theta}_{m-1}(\tilde{\boldsymbol{h}}_{m-1}-\mu\boldsymbol{s}_m g(e_m))e^{-j\Omega(m-1)}\right]\right\},\end{aligned} \tag{4.5}$$

where

$$\boldsymbol{\Theta} \triangleq \lim_{m\to\infty} E\left[\boldsymbol{\theta}_m\boldsymbol{\theta}_m^*\right] = \frac{1}{1-|\alpha|^2}\boldsymbol{Q},$$

and $g(e_m)$ is the nonlinear error function [45]. Moreover, $e_m = e_{a,m}+n_m$, in which $e_{a,m} = \boldsymbol{s}_m\tilde{\boldsymbol{h}}_m$ is a priori estimation error and n_m is the estimation noise, i.e., the error resulted from the optimal linear estimator. For the proposed algorithms, $g(e_m)$ is defined as

$$g(e_m) \triangleq \frac{\partial\varphi(e_m)}{\partial e_m}\frac{\varphi(e_m)}{1+\varphi(e_m)}. \tag{4.6}$$

We note that in an impulse-free environment, i.e., when $e_m \ll 1$, we have

$$\frac{a\varphi(e_m)}{1+a\varphi(e_m)} \approx a\varphi(e_m). \tag{4.7}$$

On the other hand, in an impulsive noise environment we have

$$\frac{a\varphi(e_m)}{1+a\varphi(e_m)} \approx 1. \tag{4.8}$$

Therefore, we decompose the left hand side of (4.5) as

$$\begin{aligned}E\left[|e_{a,m}|^2\right] &= E\left[|e_{a,m}|^2 \;\middle|\; z_m=1\right]p_Z(z_m=1) + E\left[|e_{a,m}|^2 \;\middle|\; z_m=0\right]p_Z(z_m=0) \\ &= E\left[|e_{a,m}|^2 \;\middle|\; g(e_m)=\frac{\partial\varphi(e_m)}{\partial e_m}\right]\nu_i + E\left[|e_{a,m}|^2 \;\middle|\; g(e_m)=\frac{\partial\varphi(e_m)}{\partial e_m}\varphi(e_m)\right](1-\nu_i),\end{aligned} \tag{4.9}$$

and we also note that

$$E\left[e_{a,m}\ g(e_m)\ \middle|\ z_m=1\right]=E\left[e_{a,m}\frac{\partial\varphi(e(t))}{\partial e(t)}\right],$$
$$E\left[e_{a,m}\ g(e_m)\ \middle|\ z_m=0\right]=E\left[e_{a,m}\ \frac{\partial\varphi(e(t))}{\partial e(t)}\varphi(e(t))\right]. \tag{4.10}$$

We use (4.10) to calculate the terms on the right hand side of (4.9). We also note that for the LCLMS and LCLMA methods, $\varphi(e_m)=|e_m|^2$ and $\varphi(e_m)=|e_m|$, respectively. In our analysis, we use the following assumptions:

Assumption 1:

The noise signal n_m is a zero-mean, independently, and identically distributed (i.i.d.), Gaussian random variable, and independent from $\boldsymbol{s}_m$. The transmitted signal $\boldsymbol{s}_m$ is also a zero-mean i.i.d. Gaussian random variable with the auto-correlation matrix $\boldsymbol{R_S}\triangleq E\left[\boldsymbol{s}_m\boldsymbol{s}_m^H\right]$.

Assumption 2:

The a priori estimation error $e_{a,m}$ has Gaussian distribution. This assumption is reasonable by the Assumption 1, whenever L_c is sufficiently large and the learning rate μ is sufficiently small [45].

Assumption 3:

The random processes $\|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m}$ and $g^2(e_m)$ are uncorrelated, which results in the following separation

$$E\left[\|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m}g^2(e_m)\right]=E\left[\|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m}\right]E\left[g^2(e_m)\right].$$

Based on (4.5), Assumptions 1-3, and the Price's Theorem [27] for $E\left[e_{a,m}\ \mathrm{sign}(e_m)\right]$, it can be straightforwardly shown that [27, 41]

$$E\left[|e_{a,m}|^2\ \middle|\ g(e_m)=\mathrm{csgn}(e_m)\right]=\zeta^{\mathrm{SA}}$$
$$E\left[|e_{a,m}|^2\ \middle|\ g(e_m)=e_m\right]=\zeta^{\mathrm{LMS}}$$
$$E\left[|e_{a,m}|^2\ \middle|\ g(e_m)=e_m^3\right]=\zeta^{\mathrm{LMF}}. \tag{4.11}$$

However, in [41], the excess MSE for SA, LMS, and LMF methods are calculated as follows,

$$\begin{aligned}\zeta^{\text{LCLMA}} &= \nu_i \zeta^{\text{SA}} + (1-\nu_i)\zeta^{\text{LMS}} \\ &= \nu_i \frac{2\mu \ \text{Tr}(\boldsymbol{R_s}) + \mu^{-1}\beta(\eta)}{2\eta} + (1-\nu_i)\frac{\mu\sigma_g^2 \ \text{Tr}(\boldsymbol{R_s}) + \mu^{-1}\beta(1)}{2},\end{aligned} \tag{4.12}$$

where

$$\begin{aligned}\eta &= \frac{2}{\sqrt{\pi(\zeta^{SA} + \sigma_n^2)}}, \\ \beta(\gamma) &= |1 - e^{j\Omega}|^2 \ \text{Re}\left\{\text{Tr}\left((\boldsymbol{I} - 2(\boldsymbol{X}_\gamma - \mu\boldsymbol{R_s}))\boldsymbol{h}\boldsymbol{h}^H\right)\right\} \\ &\quad + |1 - \alpha e^{j\Omega}|^2 \ \text{Re}\left\{\text{Tr}\left((\boldsymbol{I} - 2(\alpha^*\boldsymbol{X}_{\gamma,\alpha} - \mu\boldsymbol{R_s}))\boldsymbol{\Theta}\right)\right\} \\ &\quad + \text{Re}\left\{\text{Tr}\left((\boldsymbol{I} + 2(e^{j\Omega} - \alpha^*)\boldsymbol{X}_\alpha)\boldsymbol{Q}\right)\right\},\end{aligned} \tag{4.13}$$

and for $\forall\gamma$

$$\begin{aligned}\boldsymbol{X}_\gamma &= (\boldsymbol{I} - \mu\gamma\boldsymbol{R_s})\left[\boldsymbol{I} - e^{j\Omega}(\boldsymbol{I} - \mu\gamma\boldsymbol{R_s})^{-1}\right]^{-1}, \\ \boldsymbol{X}_{\gamma,\alpha} &= (\boldsymbol{I} - \mu\gamma\boldsymbol{R_s})\left[\alpha^*\boldsymbol{I} - e^{j\Omega}(\boldsymbol{I} - \mu\gamma\boldsymbol{R_s})^{-1}\right]^{-1}.\end{aligned} \tag{4.14}$$

Similarly, we obtain the following results for the excess MSE of the LCLMS method.

$$\begin{aligned}\zeta^{\text{LCLMS}} &= \nu_i \zeta^{\text{LMS}} + (1-\nu_i)\zeta^{\text{LMF}} \\ &= \nu_i \frac{\mu\sigma_n^2 \ \text{Tr}(\boldsymbol{R_s}) + \mu^{-1}\beta(1)}{2} + (1-\nu_i)\frac{\mu\xi^6 \ \text{Tr}(\boldsymbol{R_s}) + \mu^{-1}\beta(3\sigma_g^2)}{6\sigma_g^2},\end{aligned} \tag{4.15}$$

where

$$\xi^6 \triangleq E\left[|n_{g,m}|^6\right]. \tag{4.16}$$

4.2 MSE Analysis of the Second Order Methods

Here, we use the same assumptions as the first order analyses. According to (3.7), and defining $g(e_m) \triangleq e_m w(e_m)$, we have the recursion $\hat{\boldsymbol{h}}_{m+1} = \hat{\boldsymbol{h}}_m + g(e_m)\boldsymbol{P}_m \boldsymbol{s}_m$. Thus,

$$\tilde{\boldsymbol{h}}_{m+1} = \tilde{\boldsymbol{h}}_m - \boldsymbol{P}_m \boldsymbol{s}_m g(e_m) + \boldsymbol{c}_m e^{j\Omega m}, \tag{4.17}$$

which yields the following relation between the a priori and a posteriori error estimations

$$\begin{aligned} e_{p,m} &= \boldsymbol{s}_m^H(\tilde{\boldsymbol{h}}_{m+1} - \boldsymbol{c}_m e^{j\Omega m}) \\ &= e_{a,m} - \|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m}\, g(e_m). \end{aligned} \tag{4.18}$$

Equivalently, we observe that

$$\tilde{\boldsymbol{h}}_{m+1} - \boldsymbol{c}_m e^{j\Omega m} = \tilde{\boldsymbol{h}}_m - \boldsymbol{P}_m \boldsymbol{s}_m \frac{e_{a,m} - e_{p,m}}{\|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m}}. \tag{4.19}$$

Therefore, we have

$$\|\tilde{\boldsymbol{h}}_{m+1} - \boldsymbol{c}_m e^{j\Omega m}\|^2_{\boldsymbol{P}_m^{-1}} = \|\tilde{\boldsymbol{h}}_m\|^2_{\boldsymbol{P}_m^{-1}} + \frac{|e_{p,m}|^2 - |e_{a,m}|^2}{\|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m}}. \tag{4.20}$$

Assuming that $\boldsymbol{P}_m^{-1}$ converges to its mean value, $E[\boldsymbol{P}_m^{-1}] = \boldsymbol{P}^{-1}$, in the steady state, we have $E[\|\tilde{\boldsymbol{h}}_{m+1}\|^2_{\boldsymbol{P}^{-1}_{m+1}}] = E[\|\tilde{\boldsymbol{h}}_m\|^2_{\boldsymbol{P}_m^{-1}}]$ [41], which results in

$$\begin{aligned} E\left[e_{a,m} g(e_m)\right] = E\left[\|\boldsymbol{s}_m\|^2_{\boldsymbol{P}_m} |g(e_m)|^2\right] + \|\boldsymbol{c}_m\|^2_{\boldsymbol{P}_m^{-1}} \\ + 2\mathrm{Re}\left\{E\left[e^{-j\Omega m}\boldsymbol{c}_m^H \boldsymbol{P}_m^{-1}\left(\tilde{\boldsymbol{h}}_m - \boldsymbol{P}_m \boldsymbol{s}_m g(e_m)\right)\right]\right\}. \end{aligned} \tag{4.21}$$

We next investigate three cases : $g(e_m) = |e_m|$, $g(e_m) = |e_m|^3$ and $g(e_m) = \mathrm{sign}(e_m)$. We define $\boldsymbol{v}_m \triangleq E[\boldsymbol{P}_m^{-1}\tilde{\boldsymbol{h}}_m]$ and $\boldsymbol{\Sigma} \triangleq E[\boldsymbol{P}_m^{-1}\tilde{\boldsymbol{h}}_m \boldsymbol{\theta}_m^H]$. In addition, we assume that in the steady state, $\boldsymbol{P}_m$ is independent of $\tilde{\boldsymbol{h}}_m$ and $\boldsymbol{c}_m$.

1. For $g(e_m) = |e_m|$,

$$\boldsymbol{v}_{m+1} = (\boldsymbol{I} - \boldsymbol{R_S})\boldsymbol{v}_m + \boldsymbol{P}^{-1}\boldsymbol{h}(e^{j\Omega} - 1)e^{j\Omega m},$$
$$\boldsymbol{\Sigma}_{m+1} = \alpha(\boldsymbol{I} - \boldsymbol{R_S})\boldsymbol{\Sigma}_m - \boldsymbol{P}^{-1}\boldsymbol{C}e^{j\Omega m}. \tag{4.22}$$

With a method similar to [41], we obtain

$$E[\boldsymbol{P}_m^{-1}\tilde{\boldsymbol{h}}_m] = \boldsymbol{v}e^{j\Omega m},$$
$$E[\boldsymbol{P}_m^{-1}\tilde{\boldsymbol{h}}_m\boldsymbol{\theta}_m^T] = \boldsymbol{\Sigma}e^{j\Omega m}, \tag{4.23}$$

where

$$\boldsymbol{v} = [\boldsymbol{I} - \boldsymbol{R_S} - e^{j\Omega}\boldsymbol{I}]^{-1}\boldsymbol{P}^{-1}\boldsymbol{h}(1 - e^{j\Omega}),$$
$$\boldsymbol{\Sigma} = [\alpha(\boldsymbol{I} - \boldsymbol{R_S}) - e^{j\Omega}\boldsymbol{I}]^{-1}\boldsymbol{P}^{-1}\boldsymbol{C}, \tag{4.24}$$

and

$$\boldsymbol{C} = \alpha^*[1 - \alpha e^{j\Omega}]\boldsymbol{\Theta} - e^{j\Omega}\boldsymbol{Q}, \tag{4.25}$$

which yields the following expression for the excess MSE

$$[1 - \mathrm{Tr}(\boldsymbol{R_S}\boldsymbol{P})]\ \zeta_{|e_m|} = \mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{Q}) + |1 - e^{j\Omega}|^2\|\boldsymbol{h}\|^2 + |1 - \alpha e^{j\Omega}|^2\mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{\Theta})$$
$$- 2\mathrm{Re}\left\{(1 - e^{-j\Omega})\boldsymbol{h}^T(\boldsymbol{I} - \boldsymbol{R_S})\boldsymbol{v} + \mathrm{Tr}\left[(1 - \alpha^* e^{-j\Omega})(\boldsymbol{I} - \boldsymbol{R_S})\boldsymbol{\Sigma}\right]\right\} \tag{4.26}$$

2. For $g(e_m) = |e_m|^3$,
   Since in our final analysis, this case appears when the noise is Gaussian (non-impulsive), we assume that the noise power is σ_g^2. Furthermore, we use the assumption $|e_{a,m}|^2 \ll \sigma_g^2$ [41], hence by substituting $g(e_m)$ in (4.21) and simplifying the result, we obtain the following expression for EMSE:

$$[3\sigma_g^2 - 15\xi^4\mathrm{Tr}(\boldsymbol{R_S}\boldsymbol{P})]\ \zeta_{|e_m|^3} = \xi^6\mathrm{Tr}(\boldsymbol{R_S}\boldsymbol{P}) + \mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{Q})$$
$$+ |1 - e^{j\Omega}|^2\|\boldsymbol{h}\|^2 + |1 - \alpha e^{j\Omega}|^2\mathrm{Tr}(\boldsymbol{P}^{-1}\boldsymbol{\Theta}). \tag{4.27}$$

3. For $g(e_m) = \text{sign}(e_m)$,
   By using the price theorem, we have

$$\eta\ \zeta_{\text{sign}(e_m)} = \text{Tr}(\boldsymbol{R_S P}) + \text{Tr}(\boldsymbol{P}^{-1}\boldsymbol{Q}) + |1 - e^{j\Omega}|^2\|\boldsymbol{h}\|^2 + |1 - \alpha e^{j\Omega}|^2\text{Tr}(\boldsymbol{P}^{-1}\boldsymbol{\Theta}) - 2\text{Re}\left\{(1 - e^{-j\Omega})\boldsymbol{h}^T(\boldsymbol{I} - \eta\boldsymbol{R_S})\boldsymbol{v} + \text{Tr}\left[(1 - \alpha^* e^{-j\Omega})(\boldsymbol{I} - \eta\boldsymbol{R_S})\boldsymbol{\Sigma}\right]\right\}, \tag{4.28}$$

   where $\eta = \sqrt{\frac{2}{\pi\ (\zeta+\sigma_n^2)}}$. Solving this equation for ζ, we obtain the EMSE.

Finally, we achieve the following expressions for the EMSE of the second order methods.

$$\zeta^{\text{LCRLS}} = \nu_i\zeta_{|e_m|} + (1 - \nu_i)\zeta_{|e_m|^3}, \tag{4.29}$$

and

$$\zeta^{\text{LCRLA}} = \nu_i\zeta_{\text{sign}(e_m)} + (1 - \nu_i)\zeta_{e_m}. \tag{4.30}$$

4.3 Verification of the analytical results

We use a random 10-tap channel to transmit 10000 bits generated by a Turyn sequence [46] (without pule shaping), and we consider that the channel follows the model of (4.1), where $\sigma_q = 10^{-4}$, $\Omega = 10^{-5}$ and $\alpha = 0.9$. The Gaussian part of the noise has a variance of 10^{-6}, while that of the impulsive part is 10^{-2}. We average the results over 30 iterations. According to the results in Figs. 4.1 and 4.2 for LCLMS and LCRLS algorithms, we see that there is a good match between the theoretical and simulation results which is similar for LCLMA and LCRLA methods.

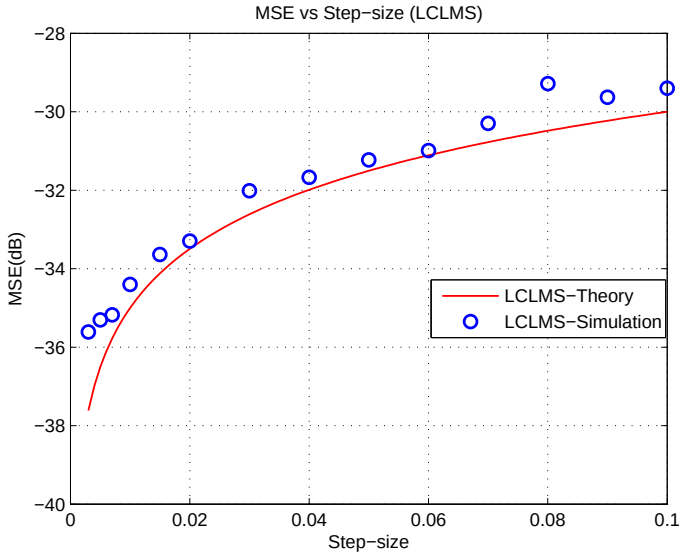

Figure 4.1: The comparison between the simulation results and theoretical results for the MSE of LCLMS algorithm.

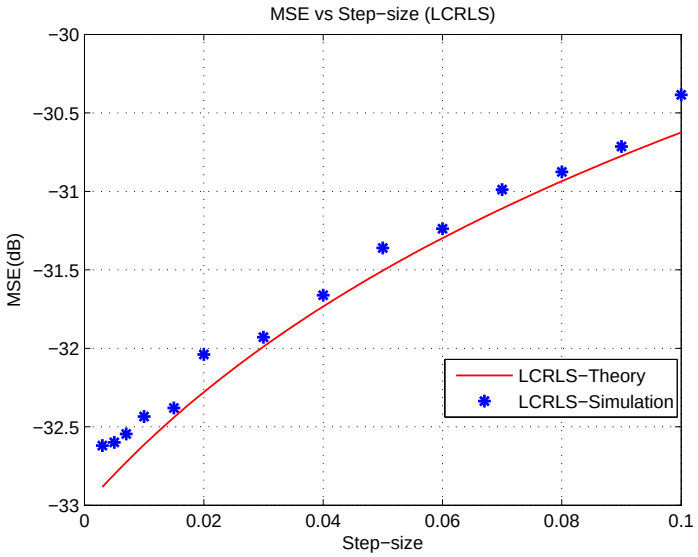

Figure 4.2: The comparison between the simulation results and theoretical results for the MSE of LCRLS algorithm.

Chapter 5

Simulations and Conclusion

5.1 Setup

In this section, we examine the performance of our algorithm under a highly realistic underwater acoustic channel equalization scenario through a highly accurate modeling of the underwater channels introduced in [1]. We add impulsive noise to the channel output to compare the robustness of the algorithms against impulsive noise. The simulation configurations and the parameters used for simulating the channel are presented in the Table 5.1. We send 60000 bits generated by repeating a Turyn sequence [46] (with a length of 28 bits), over the simulated UWA channel shown in Fig. 5.1. In addition, the system setup is the same as the one described at Section 2.2. Also, we calculate the SNR after down converting and matched filtering (i.e., from the baseband signal). The step sizes are set to $\mu = 0.01$ for all algorithms. In all of the algorithms we use a 10-tap NLMS equalizer, with the causal and anti-causal parts each of length 5. The length of the channel estimator is set to 140, since according to Fig. 5.1, $\tau_{max} \approx 35$ ms and $T_s = 1/4$ ms (because the symbol rate is 4 kHz).

We use a raised cosine pulse shaped signal with a roll-off factor of 0.25

and transmit the data at a rate of $4\text{k}\frac{\text{symbol}}{\text{second}}$, i.e., each symbol has a duration of 0.25 ms. Since we use the BPSK modulation, the signal bandwidth is $BW = (1+\beta)R = (1+0.25) \times 4\text{kHz} = 5\text{kHz}$, where β and R are roll-off factor and the transmitting rate, respectively. We use 4 samples to represent each symbol, hence, the time resolution of the signal (the time distance between two signal samples) is $T_s = 1/16$ ms.

In order to simulate the channel, we should use a time resolution $dt = T_s = 1/16$ ms, which corresponds to using a sample rate of 16 kHz. Moreover, in practice, the maximum delay of an underwater channel is usually less than 50 ms [1], hence, we will observe $T_{obs} = 62.5$ ms of the channel. One may note that the choice of T_{obs} implies the frequency resolution we use for simulating the channel, i.e., $df = 1/T_{obs} = 16$ Hz. Note that since the maximum delay of the channel is around 35 ms, the effective channel length is almost $35/0.25 = 140$ symbol durations. Also, we use a coherence time of $T_{SS} = 400$ ms, since according to [6], for a general purpose design one should consider a coherence time of several hundred milliseconds. Furthermore, the sound speed in the water for this experiment is $c = 1500$ m/s [47] and the spreading factor is set to $k = 1.7$ [1].

The results are averaged over 30 repetitions, and show the superior performance of our robust algorithm over other methods. We compare the bit error rate (BER) as well as the normalized time mean squared errors (MSE) of different algorithms to show the efficacy of our methods. In particular, to precisely evaluate the tracking performance of the algorithms, we compute the MSE exactly after the causal ISI removal block, i.e., the MSE at time step n is defined as $\text{MSE} \triangleq \frac{1}{n}\sum_{m=1}^{n}(s_m - \tilde{r}_m)^2$.

Parameters	Values
Transmitter (Tx) depth	80 m
Receiver (Rx) depth	50 m
Distance between Tx and Rx	1 km
Carrier frequency (f_c)	15 kHz
Signal bandwidth	5 kHz
Sample rate	16 kHz
Frequency resolution (df)	16 Hz
Time resolution (dt)	1/16 ms
Symbol rate	4 kHz
Maximum multipath delay (τ_{max})	35 ms
Coherence time of the small scale variants (T_{SS})	0.4 s
Total duration of simulated channel (T_{tot})	16 s

Table 5.1: The simulated channel configurations.

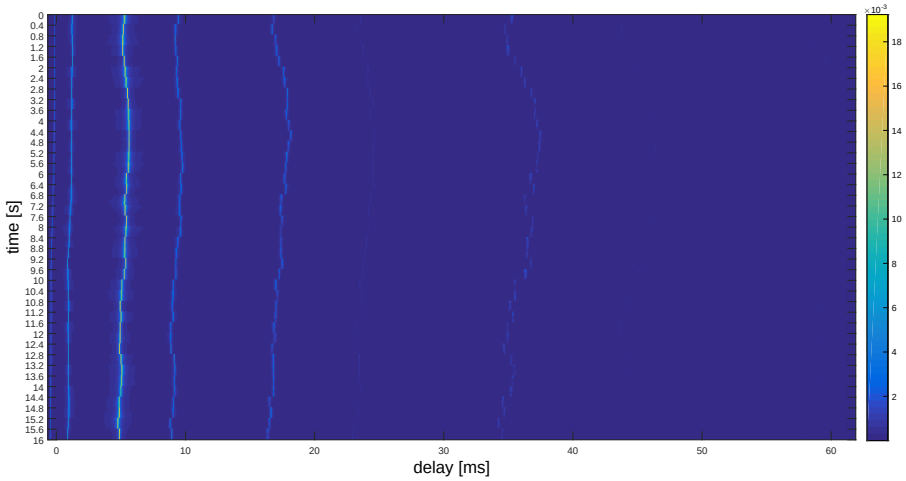

Figure 5.1: Time evolution of the magnitude baseband impulse response of the generated channel [1]

5.2 Results and Discussion

We perform two experiments to evaluate the performance of the algorithms in (1) different SNR values (2) and different impulse probabilities. In the first experiment, in order to indicate the BER vs. SNR performance of our algorithms, i.e, LCLMA, LCLMS, LCRLS and LCRLA, we use a mixture of Gaussian and a 5% impulsive noise model, i.e., $\nu_i = 0.05$, with the variance (power) of the impulsive noise being set to 10^4 times that of the Gaussian noise. In these experiments, we compare the BER and MSE of our methods with those of the conventional DFE and the state-of-the-art algorithms in the literature including: reweighted zero-attracting least mean p-power (RZALMP) [37], improved least sum of exponentials (ILSE) [36], l_0-RLS [34] (indicated as LZRLS in the figures), as well as the conventional SA, RLS and LMS algorithms. We emphasize that all of these algorithms are designed to combat the impulsive noise through minimization of different cost functions summarized in Table 5.2. For the sake of clarity, we demonstrate the results of each experiment in two different plots, one for the first order algorithms, i.e., LCLMA, LCLMS, LMS, SA, ILSE, RZALMP and another for the second order algorithms, i.e., LCRLA, LCRLS, RLS and l_0-RLS. We have also provided the complexity comparisons of different methods in Table 5.3.

As shown in the Figs. 5.2 and 5.3, our algorithms outperform all other algorithms in all SNR values. Also, observe that the conventional DFE cannot perform well in lower SNRs, while in high SNRs, it delivers a comparable performance to our methods. Moreover, the Figs. 5.4 and 5.5 depict the MSE results for the first order and second order algorithms over time at SNR = 20dB, respectively (One can observe the MSE comparison at first time steps in Figs. 5.6 and 5.7). Again, the MSE results in Figs. 5.4 and 5.5, show how our algorithms can successfully track the channel variations in this highly non-stationary and impulsive noise environment, yielding a superior performance related to the other competitors.

Algorithm	**Cost function**
SA	$\|e_m\|$
LMS	$\|e_m\|^2$
LCLMS	$\|e_m\|^2 - \frac{1}{a}\ln(1+a\|e_m\|^2)$
LCLMA	$\|e_m\| - \frac{1}{a}\ln(1+a\|e_m\|)$
ILSE	$\frac{1}{\lambda}\cosh(\lambda e_m)$
RZALMP	$\|e_m\|^p + \lambda\sum_{i=1}^{L}\log(1+\frac{\|\hat{h}_i\|}{\delta})$
RLS	$\sum_{i=1}^{m}\lambda^{m-i}\|e_i\|^2$
LCRLS	$\sum_{i=1}^{m}\lambda^{m-i}\|e_i\|^2 - \frac{1}{a}\ln(1+a\sum_{i=1}^{m}\lambda^{m-i}\|e_i\|^2)$
LCRLA	$\sum_{i=1}^{m}\lambda^{m-i}\|e_i\| - \frac{1}{a}\ln(1+a\sum_{i=1}^{m}\lambda^{m-i}\|e_i\|)$
l_0-RLS	$\sum_{i=0}^{m}\lambda^{m-i}\|e_i\|^2 + \zeta\|\hat{\mathbf{h}}_n\|_0$, where $\|\hat{\mathbf{h}}_n\|_0 \approx \sum_{k=0}^{K-1}(1-\exp(-\eta\|\hat{h}_k\|))$

Table 5.2: Cost functions of the competing algorithms.

Algorithm	$\times$	$+$	$/$	sign
SA	$2L$	$2L$		1
LMS	$2L+1$	$2L$		
LCLMS	$2L+5$	$2L+1$	1	
LCLMA	$2L+3$	$2L+1$	1	1
ILSE	$L+M+2$	$L+N+1$	1	
RZALMP	$3L+M+1$	$3L+N$	L	$L+1$
RLS	$4L^2+3L$	$3L^2+L$	$L+1$	
LCRLS	$4L^2+4L+3$	$3L^2+L+1$	$L+2$	
LCRLA	$4L^2+4L+2$	$3L^2+L+1$	$L+2$	1
l_0-RLS	$7L^2+6L+LM+2$	$6L^2+LN+1$	$L+3$	L

Table 5.3: Computational complexity comparison of different algorithms. Number of each operation (on real numbers) needed by different algorithms per one sample processing is provided in the table. We have assumed that the exponentiating to a non-integer number needs M multiplication and N additions. As we see in this table without any significant increase in the computational complexity, compared to the traditional methods we obtain a superior performance.

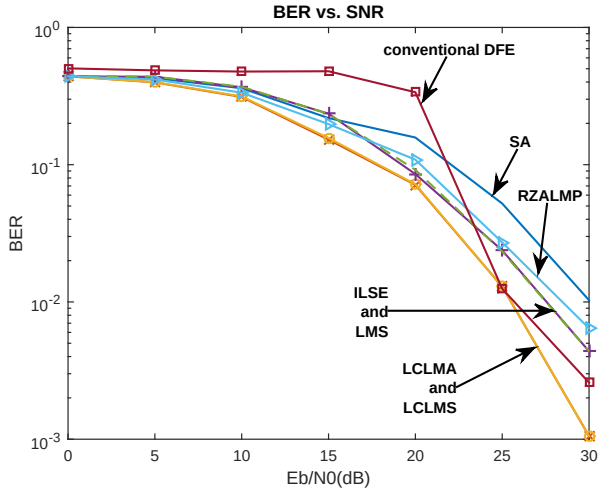

Figure 5.2: BER vs SNR of the first order algorithms, in a 5% impulsive noise environment. This figure shows the superior performance of LCLMS and LCLMA algorithms.

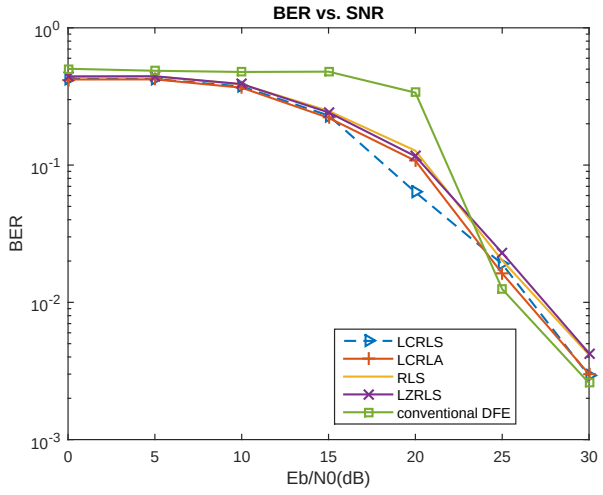

Figure 5.3: BER vs SNR of the second order algorithms, in a 5% impulsive noise environment. This figure shows the superior performance of LCRLS algorithm.

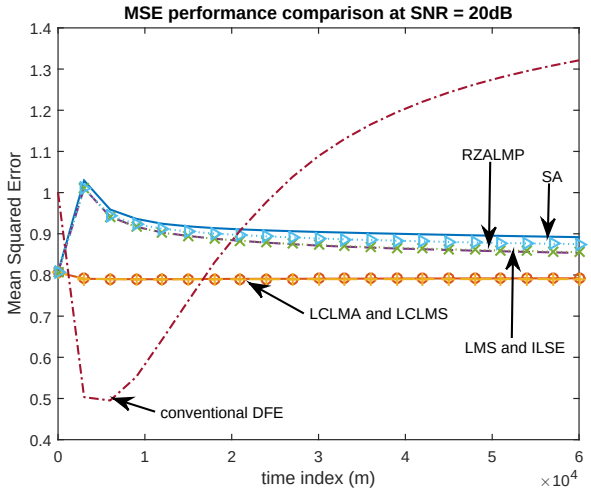

Figure 5.4: MSE of the first order algorithms, in 5% impulsive noise environment. This figure shows the superior convergence performance of the LCLMS and LCLMA methods at SNR = 20 dB.

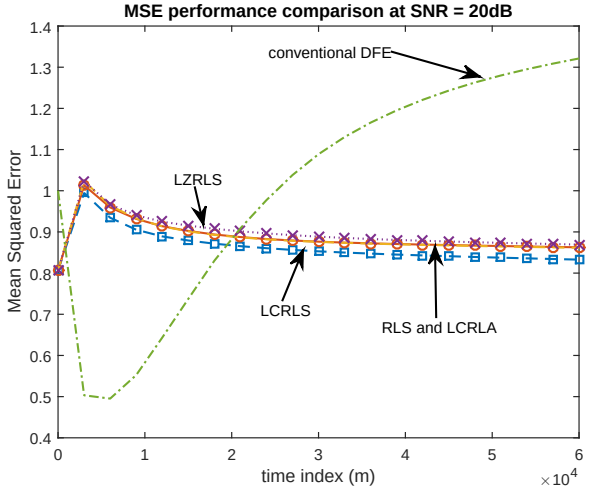

Figure 5.5: MSE of the second order algorithms, in 5% impulsive noise environment. This figure shows the superior convergence performance of the LCRLS method at SNR = 20 dB.

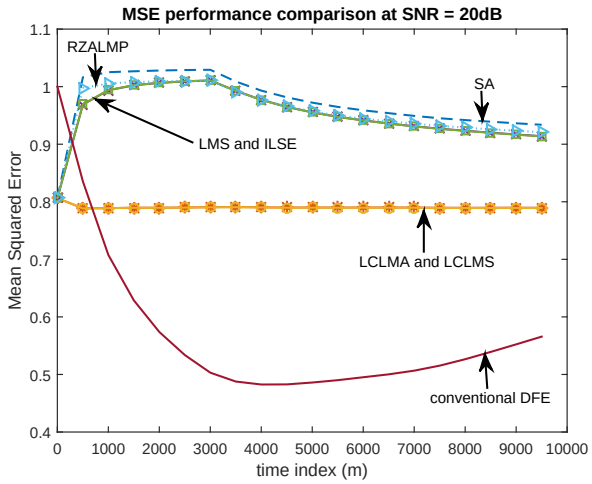

Figure 5.6: MSE comparison for different first order methods.

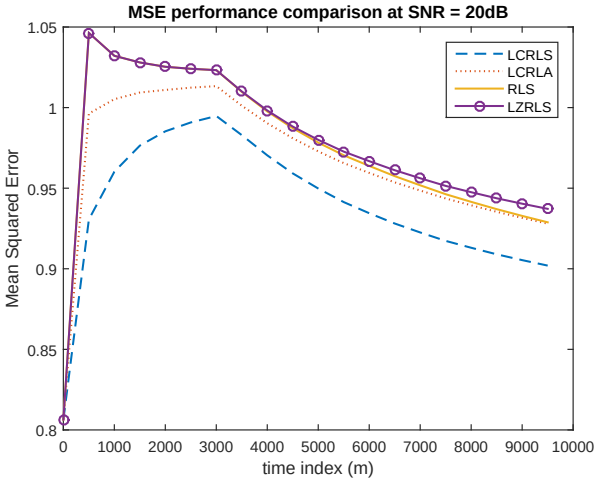

Figure 5.7: MSE comparison for different second order methods.

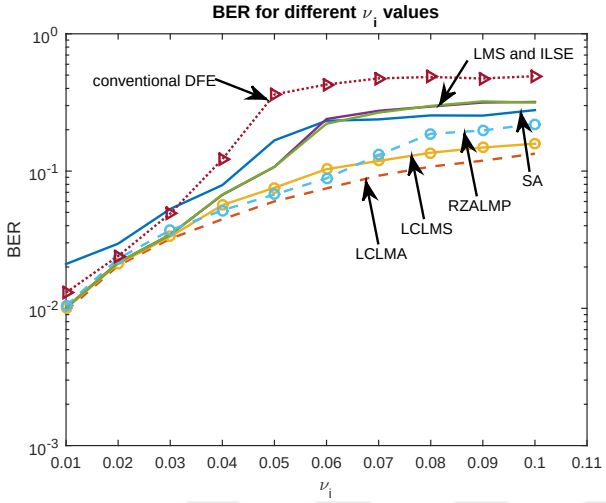

Figure 5.8: BER at different impulse probabilities for the first order algorithms. The experiments are done at SNR= 20 dB.

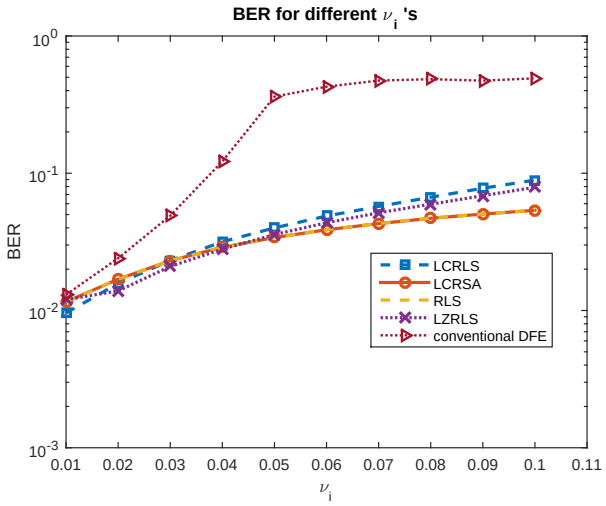

Figure 5.9: BER at different impulse probabilities for the second order algorithms. The experiments are done at SNR= 20 dB.

We also investigate the effect of the impulse probability, ν_i, on the performance of the proposed algorithms. To this end, we set the SNR to 20 dB and obtain the BER at different impulse probabilities, as shown in Figs. 5.8 and 5.9. Based on these simulations, when the impulse probability increases, our algorithms significantly outperform other methods. This superior performance

makes our algorithms suitable candidates for the highly impulsive noise real life channels.

5.3 Conclusion

In this thesis, we presented a novel family of linear channel estimation algorithms based on the logarithmic cost functions. Specifically, we introduce two first order methods, i.e., LCLMA and LCLMS, as well as two second order methods, i.e., LCRLA and LCRLS, as robust adaptive estimators for underwater acoustic channels. We implement these algorithms in a decision feedback equalization framework and remove the ISI in two consecutive stages, i.e., a channel estimation based equalizer followed by a blind equalizer. Our methods achieve a comparable convergence rate to the algorithms seeking to minimize higher order norms of the error, while maintaining the same stability of the lower order norm methods, with similar computational complexity to the conventional methods. We provide the tracking and steady-state performance analysis of the proposed algorithms both in the impulse-free and impulsive noise environments, in the presence of frequency and phase offsets, which are among the common impairments in underwater acoustic communications. Finally, we show the enhanced performance of the new algorithms in a highly realistic UWA communication scenario.

Bibliography

[1] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, 2013.

[2] M. K. Banavar, J. J. Zhang, B. Chakraborty, H. Kwon, Y. Li, H. Jiang, A. Spanias, C. Tepedelenlioglu, C. Chakrabarti, and A. Papandreou-Suppappola, "An overview of recent advances on distributed and agile sensing algorithms and implementation," *Digital Signal Processing*, vol. 39, pp. 1–14, 2015.

[3] A. Gunes and M. B. Guldogan, "Joint underwater target detection and tracking with the bernoulli filter using an acoustic vector sensor," *Digital Signal Processing*, vol. 48, pp. 246–258, 2016.

[4] A. C. Singer, J. K. Nelson, and S. S. Kozat, "Signal processing for underwater acoustic communications," *IEEE Communications Magazine*, vol. 47, pp. 90–96, January 2009.

[5] D. B. Kilfoyle and A. B. Baggeroer, "The state of the art in underwater acoustic telemetry," *Oceanic Engineering, IEEE Journal of*, vol. 25, no. 1, pp. 4–27, 2000.

[6] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, pp. 84–89, January 2009.

[7] D. Peng, Y. Xiang, H. Trinh, and Z. Man, "Adaptive blind equalization of time-varying simo systems driven by qpsk inputs," *Digit. Signal Process.*, vol. 23, pp. 268–274, Jan. 2013.

[8] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[9] T. Wong and T. Lok, "Theory of Digital Communications," 2017.

[10] A. Song and M. Badiey, "Generalized equalization for underwater acoustic communications," *Proceedings of OCEANS 2005 MTS/IEEE*, pp. 1522 – 1527, 2005.

[11] Y. R. Zheng and V. H. Nascimento, "Two variable step-size adaptive algorithms for non-gaussian interference environment using fractionally lower-order moment minimization," *Digital Signal Processing*, vol. 23, no. 3, pp. 831–844, 2013.

[12] X. Xu, S. Zhou, H. Sun, A. K. Morozov, and Y. Zhang, "Impulsive noise suppression in per-survivor processing based dsss systems," *Oceans - St. John's, 2014*, pp. 1–5, September 2014.

[13] S. Banerjee and M. Agrawal, "Underwater acoustic communication in the presence of heavy-tailed impulsive noise with bi-parameter cauchy-gaussian mixture model," *Ocean Electronics (SYMPOL), 2013*, pp. 1–7, October 2013.

[14] S. Banerjee and M. Agrawal, "Underwater acoustic noise with generalized gaussian statistics: Effects on error performance," *OCEANS - Bergen, 2013 MTS/IEEE*, pp. 1–8, June 2013.

[15] S. Banerjee and M. Agrawal, "On the performance of underwater communication system in noise with gaussian mixture statistics," *2014 Twentieth National Conference on Communications (NCC)*, pp. 1–6, March 2014.

[16] T. Y. Min and M. Chitre, "Localization of impulsive sources in the ocean using the method of images," *Oceans - St. John's, 2014*, pp. 1–6, September 2014.

[17] D. Zha and T. Qiu, "Underwater sources location in non-gaussian impulsive noise environments," *Digital Signal Processing*, vol. 16, no. 2, pp. 149–163, 2006.

[18] J. He and Z. Liu, "Underwater acoustic azimuth and elevation angle estimation using spatial invariance of two identically oriented vector hydrophones at unknown locations in impulsive noise," *Digit. Signal Process.*, vol. 19, pp. 452–462, May 2009.

[19] J. Zhang and Y. Pang, "Pipelined robust m-estimate adaptive second-order volterra filter against impulsive noise," *Digital Signal Processing*, vol. 26, pp. 71–80, 2014.

[20] J. A. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean," *Inter-Research, Marine Ecology Progress Series (MEPS)*, vol. 395, pp. 5–20, December 2009.

[21] H.-Y. Cheng and A.-Y. A. Wu, "Unified low-complexity decision feedback equalizer with adjustable double radius constraint," *Digit. Signal Process.*, vol. 51, pp. 82–91, Apr. 2016.

[22] Z. Madadi, G. V. Anand, and A. B. Premkumar, "Three-dimensional localization of multiple acoustic sources in shallow ocean with non-gaussian noise," *Digit. Signal Process.*, vol. 32, pp. 85–99, Sept. 2014.

[23] M. U. Otaru, A. Zerguine, and L. Cheded, "Channel equalization using simplified least mean-fourth algorithm," *Digital Signal Processing*, vol. 21, no. 3, pp. 447–465, 2011.

[24] A. Zerguine, "Convergence and steady-state analysis of the normalized least mean fourth algorithm," *Digit. Signal Process.*, vol. 17, pp. 17–31, Jan. 2007.

[25] M. O. Sayin, N. D. Vanli, and S. S. Kozat, "A novel family of adaptive filtering algorithms based on the logarithmic cost," *Signal Processing, IEEE Transactions on*, vol. 62, pp. 4411 – 4424, September 2014.

[26] S. Haykin, *Adaptive filter theory, 4th edition.* Prentice Hall information and system sciences series, Upper Saddle River, N.J. Prentice Hall, 2002.

[27] A. H. Sayed, *Fundamentals of Adaptive Filtering.* NJ: John Wiley & Sons, 2003.

[28] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides, "Least mean mixed-norm adaptive filtering," *Electron. Lett.*, vol. 30, no. 19, pp. 1574–1575, 1994.

[29] J. Chambers and A. Avlonitis, "A robust mixed-norm adaptive filter algorithm," *IEEE Signal Process. Lett.*, vol. 4, no. 2, pp. 46–48, 1997.

[30] J. Arenas-Garcia, V. Gomez-Verdejo, and A. R. Figueiras-Vidal, "New algorithms for improved adaptive convex combination of LMS transversal filters," *IEEE Trans. Instrumentation and Measurement*, vol. 54, no. 6, pp. 2239–2249, 2005.

[31] J. Arenas-Garcia, A. R. Figueiras-Vidal, and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1078–1090, 2006.

[32] M. T. M. Silva and V. Nascimento, "Improving the tracking capability of adaptive filters via convex combination," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3137–3149, 2008.

[33] S. S. Kozat, A. T. Erdogan, A. C. Singer, and A. H. Sayed, "Steady-state MSE performance analysis of mixture approaches to adaptive filtering," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4050–4063, 2010.

[34] K. Pelekanakis and M. Chitre, "Adaptive sparse channel estimation under symmetric alpha-stable noise," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 3183–3195, June 2014.

[35] K. Pelekanakis and M. Chitre, "Robust equalization of mobile underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 40, pp. 775–784, Oct 2015.

[36] S. Wang, Y. Zheng, S. Duan, L. Wang, and C. K. Tse, "A class of improved least sum of exponentials algorithms," *Signal Processing*, vol. 128, pp. 340 – 349, 2016.

[37] W. Ma, B. Chen, H. Qu, and J. Zhao, "Sparse least mean p-power algorithms for channel estimation in the presence of impulsive noise," *Signal, Image and Video Processing*, vol. 10, no. 3, pp. 503–510, 2016.

[38] C. van den Bos, M. H. L. Ksuwenhoven, and W. A. Serdijn, "Effect of smooth nonlinear distortion on ofdm symbol error rate," *IEEE Transactions on Communications*, vol. 49, pp. 1510–1514, Sep 2001.

[39] P. Banelli, "Bayesian estimation of a gaussian source in middleton's class-a impulsive noise," *IEEE Signal Processing Letters*, vol. 20, pp. 956–959, Oct 2013.

[40] L. Rugini and P. Banelli, "On the equivalence of maximum snr and mmse estimation: Applications to additive non-gaussian channels and quantized observations," *Trans. Sig. Proc.*, vol. 64, pp. 6190–6199, Dec. 2016.

[41] N. R. Yousef and A. H. Sayed, "Ability of adaptive filters to track carrier offsets and channel nonstationarities," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1533–1544, Jul 2002.

[42] F. Qu and L. Yang, "Basis expansion model for underwater acoustic channels?," in *OCEANS 2008*, pp. 1–7, Sept 2008.

[43] X. Wang and H. V. Poor, "Joint channel estimation and symbol detection in Rayleigh flat-fading channels with impulsive noise," *IEEE Comm. Lett.*, vol. 1, no. 1, pp. 19–21, 1997.

[44] S.-C. Chan and Y.-X. Zou, "A recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise: fast algorithm and convergence performance analysis," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 975–991, 2004.

[45] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters with error nonlinearities," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 653–663, 2003.

[46] R. J. Turyn, "Sequences with small correlation," *In Error Correcting Codes: Proceedings of a Symposium*, pp. 195–228, 1968.

[47] P. C. Etter, *Underwater Acoustic Modeling and Simulation*. CRC Press, fourth ed., 2013.