

AMERICAN UNIVERSITY OF BEIRUT

DECISION THEORY UNDER “QUANTIZED
FILTERING”

by

SARA AHMAD JALALEDDINE

A thesis

submitted in partial fulfillment of the requirements
for the degree of Master of Engineering
to the Department of Electrical and Computer Engineering
of the Faculty of Engineering and Architecture
at the American University of Beirut

Beirut, Lebanon
September 2016

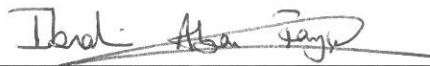
AMERICAN UNIVERSITY OF BEIRUT

DECISION THEORY UNDER “QUANTIZED FILTERING”

by

SARA AHMAD JALALEDDINE

Approved by:



Dr. Ibrahim Abou Faycal, Associate Professor

Advisor

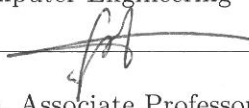
Electrical and Computer Engineering



Dr. Mohammad Mansour, Professor

Member of Committee

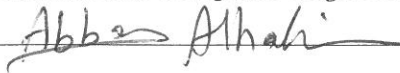
Electrical and Computer Engineering



Dr. Fadi Karamah, Associate Professor

Member of Committee

Electrical and Computer Engineering



Dr. Abbas Alhakim, Associate Professor

Member of Committee

Mathematics

Date of thesis defense: September 9, 2016

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Jalaleddine Sara Ahmad
Last First Middle

Master's Thesis Master's Project Doctoral Dissertation

I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project**, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

Sara
Signature

9/21/2016
Date

Acknowledgements

First I would like to express my deepest appreciation to my supervisor, Professor Ibrahim Abou Faycal, for all his time and effort in guiding me in my thesis. His persistence in helping me in many challenging problems, eased my way to complete this dissertation. Not only that, I would like to thank Prof. Ibrahim Abou Faycal for all his support and continuous encouragement throughout our work together.

Also, I would like to give my greatest gratitude to Professor Mohammad Mansour, Professor Fadi Karamah and Professor Abbas Alhakim on their valuable comments. This aided me to improve my thesis. Additionally, I would like to thank Ms. Rabab Abi Shakra and the Electrical and Computer Engineering Department's staff, in general, for their cooperation.

Sincere thanks to my beloved parents and siblings Nour, Reem and Khalil for their support during the last two years and their encouragement throughout the past years. I thank God for having them in my life.

Furthermore, I would like to thank my friends, especially the ECE ladies I met during my stay at AUB. Their presence motivated me to keep working with high enthusiasm. Last but not least, I would like to express my gratefulness to every person that helped me in completing this work.

An Abstract of the Thesis of

Sara Ahmad Jalaleddine for Master of Engineering
Major: Electrical and Computer Engineering

Title: Decision Theory Under “Quantized Filtering”

Convolution is an important operation in signal processing and analysis. It is a mathematical operation on two input functions which results in a third function that is typically viewed as a transformed version of one of the original functions. Convolution is used in several contexts that include probability, statistics, computer vision, image and signal processing, electrical engineering, and differential equations. In the context of communication system design, one of the initial tasks of a receiver is to detect the presence of a packet through filtering and hence convolution: The convolution is between the received signal and typically a matched filter, used to detect the presence of a training sequence or synchronizing sequence. This process is computationally intensive and usually runs for a long duration. For this reason, several methods have been proposed and implemented to alleviate this computational burden.

In our work, we aim to decrease the required number of multiplications through quantizing the matched filter, without increasing the amount of input to output latency. This of course comes at the expense of “performance”. With a view toward detection application, the selection of the quantized filter is done through applying decision theory techniques and corresponding quality measures. Cases of several stochastic noise models are studied and analysed. The performance of the proposed scheme was measured through plotting the Operating Characteristic curve, and also through the rates of exponential decay using large deviation theory. It is found that in all studied cases, the design of the suboptimal structure was immune to Signal-to-Noise Ratio values and also typically to the various quality measures and operating points that were considered.

Contents

Acknowledgements	v
Abstract	vi
List of Figures	x
List of Tables	xiii
1 Introduction	1
1.1 Packet Detection	2
2 Literature Review	4
2.1 Computing Convolution using FFT	4
2.2 Block Convolution using the Overlap Add Method	5
2.3 Fast Wavelet Transform	5
3 Our Proposed Approach	6
3.1 Quantization Process	6
3.2 The Receiver's Model and the Detection Process	7
3.3 The Models to be Studied	8
4 Used Techniques	10
4.1 Quantizer Optimization Using Kmeans Approach	11
4.2 Quantizer Optimization Using Neyman-Pearson Approach	12
4.3 Optimization Using the Large Deviation Approach	15
4.3.1 The Gärtner-Ellis theorem	15
5 Model 1: The Received Signal is Subject to Additive Gaussian Noise	19
5.1 Model Description	19
5.2 The Likelihood Ratio Test	21
5.2.1 Real Case	21
5.2.2 Complex Case	21
5.3 Neyman-Pearson Approach	22

5.3.1	Real Case	22
5.3.2	Complex Case	25
5.4	Maximum Likelihood Rule	30
5.5	Minimising Error Rate using Large Deviations	31
5.5.1	For Real Signal Cases	32
5.5.2	For complex Signal Cases	36
5.6	Unknown Location	41
5.6.1	Numerical Results	43
6	Model 2: Signal with Rayleigh Fading	48
6.1	Model Description	48
6.2	Likelihood Ratio Test	49
6.3	Minimising Error Rate using Large Deviations	50
6.3.1	The Probability of a Miss (P_M)	50
6.3.2	The Probability of False Alarm (P_F)	52
6.4	Minimising Error Rate through Decision Theory	53
6.4.1	The Probability of a miss (P_M)	54
6.4.2	The Probability of False Alarm (P_F)	54
6.4.3	Numerical Results	55
7	Model 3: Signal Subject to Noise having a Laplace Distribution	62
7.1	Model Description	62
7.2	Likelihood Ratio Test	63
7.3	Minimising Error Rate using Large Deviations	63
7.3.1	For Real Signal Cases	63
8	Model 4: Signal Subject to Noise having an Exponential Distribution	68
8.1	Model Description	68
8.2	Likelihood Ratio Test	69
8.3	Minimising Error Rate using Large Deviations	69
8.3.1	For Real Signal Cases	69
9	Conclusion	71
	Appendices	72
A	Different Stochastic Distributions	72
A.1	Gaussian Distribution	72
A.2	Complex Gaussian Distribution	72
A.3	Laplace Distribution	72
A.4	Exponential Distribution	73
B	Abbreviations	74

List of Figures

3.1	A Quantized Filter used to detect Barker convolved with a Sinc Sequence	7
3.2	Model at the Receiver	8
5.1	Model 1 at the Receiver	19
5.2	PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 4 dB. The performance with no quantization is shown using solid line.	23
5.3	PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, SNR= 9.2898dB. The performance with no quantization is shown using solid line.	24
5.4	PD vs P_F for five quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR= 4 dB.	25
5.5	PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zaddoff Chu sequence detection, SNR= 4.9136dB. The performance with no quantization is shown using solid line of 16 quantization points.	27
5.6	PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zaddoff Chu sequence detection, SNR= 0.1424dB. The performance with no quantization is shown using solid line of 16 quantization points.	28
5.7	PD vs P_F for five quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Zadoff Chu, for SNR=4.9136 dB.	30
5.8	plot the results of $\Lambda^*(\tau)$, $\Lambda(\theta)$ and $\theta\tau$ for $\mu = 0.5244$	34
5.9	plot the results of $\Lambda^*(\tau)$, $\Lambda(\theta)$ and $\theta\tau$ for $\mu = 1.0489$	34

5.10	$\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 9.2898 dB. The performance with no quantization is shown using ‘*’ line.	36
5.11	$\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 17.0713 dB. The performance with no quantization is shown using ‘*’ line.	37
5.12	$\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 4.9136 dB. The performance with no quantization is shown using solid line.	41
5.13	p_Y for a bivariate normal distribution with mean of $[-1 \ 1]$ and variance 0.5	44
5.14	PD vs P_F for optimized theoretical quantization points, starting for selected number of quantization points of Barker with Sinc sequence detection, SNR= 4dB. The performance with no quantization is shown using dotted line. Here maximum peak was studied in a packet having a total length $L = 535$ and packet is located at index $I=2$	45
5.15	PD vs P_F for two quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR=4 dB . Here maximum peak was studied in a packet having a total length $L = 535$ and packet is located at index $I=2$	45
5.16	PD vs P_F for No quantization, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR=4 dB. Here maximum peak was studied in a packet having a total length $L = 535$ and the packet is located at $I=2$	46
5.17	PD vs P_F for no quantization, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR=4 dB. Here maximum peak was studied in a packet having a total length $L = 542$ and the packet is located at $I=2$	46
5.18	PD vs P_F for No quantization, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR= $SNR = 17.0713$ dB. Here maximum peak was studied in a packet having a total length $L = 557$ and the packet is located at $I=10$	47

6.1	Model 2 at the Receiver	48
6.2	P_D vs P_F for $\frac{\sigma_0^2}{\sigma_1^2} = 1$ and $\frac{\sigma_0^2}{\sigma_1^2} = 0.1$	55
6.3	P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 4 dB. The performance with no quantization is shown using * line.	56
6.4	P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, SNR= 9.2898dB. The performance with no quantization is shown using * line.	59
6.5	P_D vs P_F for two quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR= 4 dB.	60
6.6	P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zadoff Chu sequence detection, SNR= 0.1424dB. The performance with no quantization is shown using solid line of 16 quantization points.	60
6.7	P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zadoff Chu sequence detection, SNR= 4.9136dB. The performance with no quantization is shown using solid line of 16 quantization points.	61
6.8	P_D vs P_F for two quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Zadoff Chu sequence, for SNR=0.1424 dB.	61
7.1	Model 3 at the Receiver	62
7.2	$\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for $b = 2$. The performance with no quantization is shown using '*' line.	66
7.3	$\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for $b = 10$. The performance with no quantization is shown using '*' line.	67
8.1	Model 4 at the Receiver	68

List of Tables

5.1	The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=4 dB.	26
5.2	The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=-2.9287 dB.	26
5.3	The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu, for SNR=4.9136 dB.	29
5.4	The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu, for SNR=0.1424 dB.	29
6.1	The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=4 dB.	57
6.2	The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=-2.9287 dB.	57
6.3	The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu for SNR=4.9136 dB.	58
6.4	The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu for SNR=0.1424 dB.	58
7.1	The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for $b = 2$	66
7.2	The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for $b = 10$	67

Chapter 1

Introduction

Convolution is an important operation in signal processing and analysis. Through convolution, one can construct the output of a system using the original input signal and the impulse response. In other words, convolution is a mathematical operation on two input functions (such as f and Y), that produces a third function that is typically viewed as a transformed version of one of the original functions. Convolution is used in several contexts that include probability, statistics, computer vision, image and signal processing, electrical engineering, and differential equations.

In discrete systems [1], the convolution of two functions is given by,

$$(f * Y)[k] = \sum_{i=-\infty}^{\infty} f[i]Y[k - i]. \quad (1.1)$$

For two finite discrete sequences of length N each, the linear or aperiodic convolution will result in an output of length $2N - 1$. However, calculating the convolution of two functions is computationally intensive. Direct application of equation (1) using a tapped delay line architecture requires an order of $\mathcal{O}(N^2)$ multiplications. For example, in a typical audio signal, a three second impulse response sampled at 44.1 kHz will require of the order of $(3 \times 44,100)^2$ or 35 billion operations to convolve with another input of the same length.

For this reason, several methods have been proposed and implemented to overcome this computational problem. One approach is to find the convolution of two signals by multiplying the Fourier transform of the input signals, followed by computing inverse Fourier transform of their product,

$$F^{-1}[F[f]F[Y]] = f * Y. \quad (1.2)$$

However, the approach of using the frequency domain introduces significant input to output latency. The input signals should be initially buffered and then

transformed into the frequency domain. After that, the output will be inversely transformed into the time domain. All of this buffering will result in a minimum of $2 \times N$ samples latency.

Other convolution methods were proposed, which aim to decrease the required number of computations. Some of these techniques include block convolution.

A similar operation to convolution is correlation, where it uses equation (1) but with no time reversal. So equation (1) becomes,

$$(f * Y)[k] = \sum_{i=-\infty}^{\infty} f[i]Y[i - k].$$

The above equation is also known as a sliding dot product or sliding inner-product. We use correlation to measure the similarity between two series. Thus, it is widely used in packet detection.

1.1 Packet Detection

In the context of communication system design, one of the initial tasks of a receiver is to detect the presence of a packet, especially in multi-user system. Classical Decision theory results dictate the use of a correlator or a "Matched Filter" to detect the presence of a training sequence/ synchronising sequence.

In our work, we aim to decrease the required number of multiplications, without increasing the amount of input to output latency. This of course comes at the expense of "performance". Our quality measures naturally related to those of the detection problem:

- i) If the maximum amplitude is greater than a certain threshold.
- ii) If the location of the maximum amplitude is within a certain range.

The aim is to prove that up to a certain number of quantization levels of the filter, the filter was so close in performance to the original filter and thus was able to detect the desired sequence up to close percentage to the original one.

With a view toward detection application, the choice of the quantization points will be based on applying Decision Theory concepts, section 4.2 explains it in more details. Moreover, Probability Theory tools such as large deviation theory, which deals with the behaviour at which probabilities of events' tails of certain distributions decay asymptotically, are used in our study to find the optimal points.

More precisely, the Gärtner-Ellis theorem will be used, followed by applying the Fenchel-Legendre transform to find an optimized result through increasing the efficiency of the quantizer used. This idea will be elaborated in more details in Section 4.3. In addition to both techniques, Kmeans quantization method is also applied in our study to find the optimal points.

Cases of several stochastic noise models are studied and analysed. Different types of signal packets with one or more noises of diverse nature having different Signal-to-Noise (SNR) are considered, in order to produce numerous outputs when convolved/ correlated with various quantized version of the original filter. Each scenario will be studied and analysed in a different chapter. The performance of the proposed scheme was measured through plotting the Operating Characteristic curve, and also through the rates of exponential decay using large deviation theory. It is found that in all studied cases, the design of the suboptimal structure was immune to Signal-to-Noise Ratio values and also typically to the various quality measures and operating points that were considered.

The thesis is organized as follows. In chapter 2, we review the previous work which is deemed relevant, that aims to decrease the computations required to calculate convolution. In chapter 3, the proposed design is introduced. Chapter 4 explains in details the techniques used to optimize the filter. Chapter 5 through chapter 8, study models with different stochastic noise. Conclusion is discussed in chapter 9.

Chapter 2

Literature Review

As stated in chapter 1, several approaches were adopted to overcome the amount of input to output latency without changing the expected output when computing correlation/ convolution. Section 2.1 explains the FFT technique, while section 2.2 talks about the block convolution method.

2.1 Computing Convolution using FFT

One of the approaches used in discrete cases, is multiplication in the frequency domain or also known as “Fast Convolution [2], which translates to circular convolution in the time domain (Rabiner and Gold, 1975). The length of the output of this convolution is equal to the length of the longest sequence. This is given by,

$$R[k] = IDFT(R[w]) = IDFT(Y[w].f[w]), \quad (2.1)$$

where w represents the discrete frequency variable, Y the signal and f the filter.

The result of the above equation will yield the required linear convolution result only if $Y[n]$ and $f[n]$ are padded with zeros prior to the DFT (Discrete Fourier Transform). By this, their respective lengths will become $2N - 1$. This has been known as “fast” convolution because the DFT is computed with the Fast Fourier Transform (FFT) and its inverse (IFFT) (Rabiner and Gold, 1975). However, this method excludes the implementation of a real-time input without special consideration for the lengths of the signals.

Comparing this method to the traditional time domain convolution, it is much more efficient. This is because, the number of operations required for the given inputs is around $2(2N - 1) \times \log_2(2N - 1)$ operations for the FFTs, $2 \times (2N - 1)$ operations for the complex multiplication of $Y[k]$ and $f[k]$, and another $(2N - 1) \times \log_2(2N - 1)$ operations for the IDFT. This means this approach results in

a notable decrease in the number of calculation required. Therefore the total number for multiplication using FFT is $\mathcal{O}(N \times \log_2 N)$

However, this method is limited to cases where the input signal's length is finite and completely defined. Alternative methods which allow the convolution to be performed in consecutive sections, allowing very long or indeterminate signals to be convolved with the desired impulse response are known as block convolutions.

2.2 Block Convolution using the Overlap Add Method

The overlap add method is an efficient way used to find the discrete convolution with a very long signal with a finite impulse response (FIR) filter. It is based on the concept of dividing the signal into multiple convolutions of $f[k]$ with short “ j ” segments of $Y[k]$ of length L , where L is the segment length and j is the number of broken segments. Therefore, $R[k]$ can be represented as a sum of short convolutions. In other words, to find the result of $R[k]$, first the original signal is partitioned into non-overlapping “ j ” sequences, followed by calculating the discrete Fourier transforms of the sequences through multiplying the FFT of $Y[i]$ with the FFT of $f[k]$. After that, we recover $R[k]$ by calculating inverse FFT of each $R_j[w]$, then the resulting output signal is reconstructed by overlapping and adding the $R_j[k]$. The overlap arises from the fact that a linear convolution is always longer than the original sequences.

2.3 Fast Wavelet Transform

For a stationary signal, Fourier transform is sufficient to analyse the signal. However, in many applications, a signal may have transitory or nonstationary aspects. Fourier analysis is unable to detect such events and therefore not suitable to describe them. To overcome this limitation, that is to gain information in both time and frequency domain, another type of transform called wavelet transform [3] can be used. This technique is used in engineering and computer science for data compression and signal processing. Similar to FFT, there exists a technique that efficiently computes the wavelet transform much faster, which is called the Fast Wavelet Transform (FWT). The number of computations in the convolution process depends on the wavelet basis function being used.

Chapter 3

Our Proposed Approach

In this thesis, we propose a new idea to detect a sequence in the packet received \mathbf{Y} , which is based on quantizing the amplitude of the filter before the convolution process between the packet and the filter. The classical optimal detection procedure was to use a filter that is equal to the sequence and perform convolution, this is known as the matched filter. This is equivalent to use a filter \mathbf{f} that is conjugated time-reversed version of the sequence, which performs convolution with the packet. By quantizing the filter \mathbf{f} , we will be able to decrease the required number of multiplications required for the convolution process. For example, let $\hat{\mathbf{f}}$ denotes the quantized version of the filter, then:

- 1) if $\hat{\mathbf{f}}$ is constant for all i

$$\langle \mathbf{Y}, \hat{\mathbf{f}} \rangle = \sum_{i=0}^{N-1} \hat{f}[i]Y[k-i] \rightarrow \hat{f} \sum Y \rightarrow 1 \text{ Multiplication}$$

- 2) if $\hat{\mathbf{f}}$ consists of 2 quantization points

$$\langle \mathbf{Y}, \hat{\mathbf{f}} \rangle = \sum_{i=0}^{N-1} \hat{f}[i]Y[k-i] \rightarrow 2 \text{ Multiplications}$$

This chapter will be as follows: Section 3.1 discusses how the quantization process will be applied. The Receiver's Model and detection process will be introduced in section 3.2. Finally, section 3.3 introduces all models to be studied.

3.1 Quantization Process

As it is known, the main function of a quantizer is to change a signal into a "quantized signal" [4], which is then communicated through a digital communication system. In other words, it is a process that transforms a possible analog

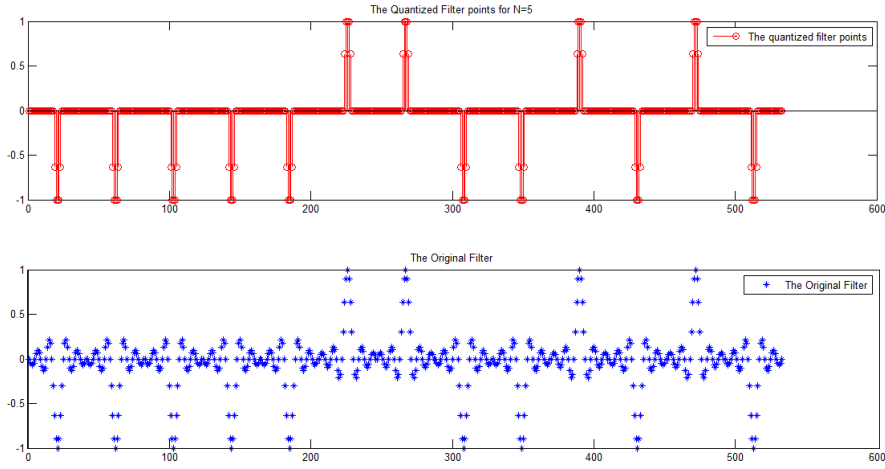


Figure 3.1: A Quantized Filter used to detect Barker convolved with a Sinc Sequence

sequence into a sequence of discrete finite values. However, performing quantization to a signal results in a loss in the amplitude information. In our research we adopted signal quantization and applied it to quantize the original detection filter. Figure 3.1 is an example of quantizing a filter that is used to detect a Barker convolved with a Sinc Sequence. The blue discrete points are the original filter points. While the red ones are the quantized version of the filter, where five points were used to design it.

Moreover, to find the “Best Quantizer”, we will perform several simulations using Matlab through testing different techniques proposed and find which one results in minimal error. Chapter 4 discusses the several ways used to find the quantization points and the comparison process used.

3.2 The Receiver’s Model and the Detection Process

Figure 3.2 represents the model of the receiver. \mathbf{Y} is the sequence being received of length L , which consists of either:

- 1) The desired signal \mathbf{f} , located at location I and of length N , to be detected in the presence of additive stochastic noise.
- 2) Noise only

$\hat{\mathbf{f}}$ is the filter used to detect the presence of sequence \mathbf{f} . It is the quantized version of the desired conjugated time-reversed version of the signal \mathbf{f} . At the receiver,

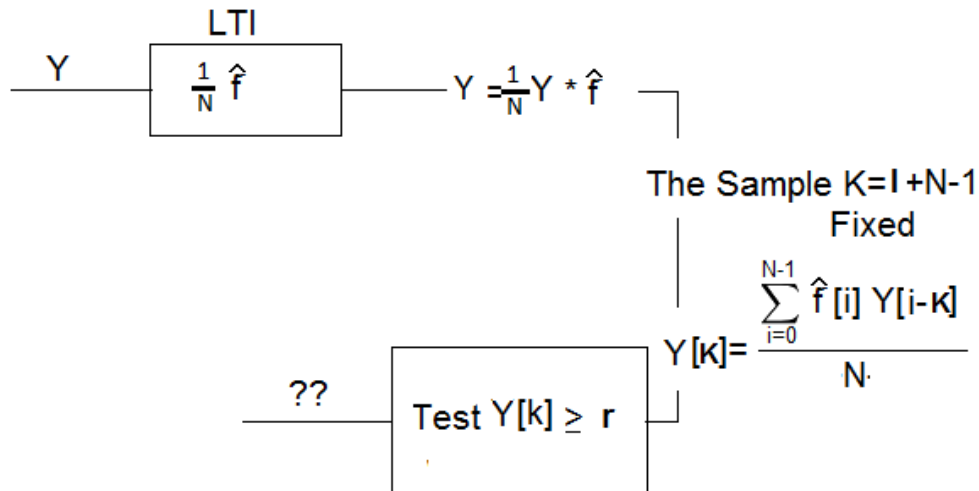


Figure 3.2: Model at the Receiver

the discrete signal \mathbf{Y} convolves with $\hat{\mathbf{f}}$. The result of this convolution \mathbf{Y} will have a peak. Two different scenarios will be considered:

- 1) We assume this peak should be located at K , where $K = I + N - 1$. $Y[K]$ will be normalized and checked if it is greater than the threshold \mathbf{r} . If this was the case, the signal \mathbf{f} is present in the packet, else \mathbf{f} is not present.
- 2) The location K is unknown to the detector. In this case, maximum peak resulting from the convolution process will be checked if it is greater than \mathbf{r} . Since we know originally where the packet is located, our detection analysis will be based on checking if:
 - It is at location K .
 - It is greater than a certain threshold.

That is,

$$P(\max = K \quad \& \quad Y[K] \geq \mathbf{r})$$

3.3 The Models to be Studied

As stated in the beginning of this chapter, detection is based on performing correlation/convolution between the filter and the received packet, which in our study “quantized filter” and the received packet. In addition, in section 3.2, we stated that received packet \mathbf{Y} is subjected to additive stochastic noise. The different stochastic noise cases that we will study are:

- Additive Gaussian Noise

- Rayleigh Fading
- Laplacian Noise
- Exponentially Distributed Noise

Each case will be studied in a different chapter, where all applicable optimizing Filter techniques will be analysed and compared. Chapter 5 studies the Additive Gaussian Noise case. Chapter 6 has the Rayleigh Fading Case. Laplacian Noise is considered in chapter 7. Finally, Exponentially Distributed Noise is in chapter 8.

Chapter 4

Used Techniques

Several methods and techniques will be used in order to develop an efficient quantizer with minimal error. In general, to measure the overall performance of a quantizer one can calculate the distortion error, so we may say a quantizer is a “good scalar quantizer” if the distortion’s error is of a small value. However, in our case, our main goal is to increase the probability of detection P_D (that is minimising the probability of a miss P_M) and minimize the probability of false alarm P_F .

In general, to design a quantizer, two things should be found (a) the Code words or Representation points” and (b) the Cells or Quantization Regions. To do that, it is assumed that the amplitude distribution of the quantizer of the input is known. There are several methods used to find (a) and (b) which we will look at and test. One of those techniques is by adapting the idea of “Kmeans” clustering for every model designed. Another is to optimize our quantizer through applying the “Gärtner-Ellis theorem. These and other techniques will be elaborated more in the coming subsections.

For all adopted criteria, we find the best quantizing points using Matlab to solve constrained nonlinear multivariable optimization problem (Fmincon) [5,6], which uses sequential quadratic programming (SQP). In each of this method’s iterations, the function solves a quadratic programming (QP) sub problem. Later, after each iteration, it updates the estimate of the Hessian, which is the second derivative of the Lagrangian, using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. Moreover, this technique depends on the initial values given. For this reason, we have tested the techniques using two different initial quantization points:

- 1) Way 1: The quantized points chosen are equally spaced between the “minimum” and “maximum points” of the original filter points.
For example if the filter f has the points: 1 2 5 6 7 and 3 points are required. The initial points selected will be: 1 4 7

- 2) Way 2: The quantized points are chosen based on the idea of selecting, from the original filter points, the points that are farthest from the rest. For example if the filter f has the points: 1 2 5 6 7 and 3 points are required. The initial points selected will be: 1 5 7, since those three points in the sequence are the points that are farthest from each other.

Based on the initial results we got, the more appropriate initial values were obtained using the Fmincon approach using Way 2.

The comparison process between all techniques will be done through producing an Operating Characteristic (OC) curve through simulations using Matlab. This curve is a plot of Probability of Detection and Probability of False Alarm (P_D vs P_F), where those values are produced through changing the threshold in the LRT. During the simulation process in Matlab, a specific random seed was set to all simulations to make the comparison more reliable.

4.1 Quantizer Optimization Using Kmeans Approach

The method of kmeans clustering, discussed in [7–10], can be used to optimize the quantizer, for the purpose of minimising the Mean Squared Error (MSE).

Kmeans is a method of vector quantization that uses cluster analysis. It aims to partition a given data into k clusters, where each observation belongs to the cluster with the nearest approximate value. By this the data space is partitioned into Voronoi cells. In our search, we created two different codes using Matlab that uses the idea of Kmeans clustering. One of them splits the clusters with the maximum distortion, while the other splits clusters with the greatest number of data. It was found that the latter case resulted in a better performance. Despite that difference, both are created to comply with any required distortion, not just MSE.

The different distortion measures that were studied are:

1. MSE: $|f_i - \hat{f}_c|^2$, which is the distortion originally used by kmeans
2. - $|f_i - \hat{f}_c|$
3. $|f_i - \hat{f}_c|^4$
4. Maximum Absolute Error (MAE): $\max |f_i - \hat{f}_c|$
5. $|f_i - \hat{f}_c|^{1.8}$

6. $|f_i - \hat{f}_c|^{2.2}$

where f_i denotes the different points of the original filter and \hat{f}_c are the quantization points used to represent the filter.

The reason behind using different distortions is that each will result in different quantization points and it is not clear which one yields better results in a decision theory setting. In other words, finding when $\sum_{x \in \delta_j} d(x, \alpha)$ is minimum depends on

the function $d(x, \alpha)$, which is the distortion measure. For example:

1- Mean square $d(x, \alpha) = (x - \alpha)^2$ will result in $\alpha^* = (1/N) \sum x$, which is the mean.

2- However, if we took the case when $d(x, \alpha) = |x - \alpha|$, then it will result in α^* being the median.

After comparing initial results from the above listed distortions, it was found that the OC curves generated for several models of different signals were really close in performance. As a result, the Kmeans with its original idea was adopted to be studied and compared in all models to all other techniques. In other words, the MSE distortion will be used to cluster data by splitting the ones with the greatest amount.

4.2 Quantizer Optimization Using Neyman-Pearson Approach

Decision theory tools are widely used in designing communication systems, biomedical tools, radar and sonar implementations, etc. Since our design deals with signal detection, decision theory tools seems a suitable approach to be used in the designing process.

In general, one may have several hypothesis, which may be represented as: H_0, H_1, \dots, H_N this is called hypothesis testing problem. In other words, let Y be the result of a random experiment which belongs to set \mathcal{Y} . There exists a device that decides based on the value of Y which will be the “best” out of those N hypothesis. However, in this thesis, a binary problem is considered that is only two hypothesis will be studied: The first is H_0 which only contains noise, while the second hypothesis will be H_1 which consists of a combination of the desired signal and noise. This binary problem is called Binary Hypothesis Testing.

Bayes' Risk Formulation

For an optimal design three basic things should be considered. First, it is assumed that each hypothesis has an apriori, denoted by:

$$Pr[H = H_0] = P_0$$

$$Pr[H = H_1] = P_1$$

where H denotes the true hypothesis.

Second, a measurement model should be adopted. To get the measurement model, the relationship between the hypothesis and observations is defined by a probabilistic model through two likelihood functions: $p_{Y|H}(y|H_0)$ and $p_{Y|H}(y|H_1)$.

Third, cost c_{ij} of deciding on H_i when H_j is true should be studied in order to minimize the average cost of the design.

Thus for an optimal design that minimizes the average cost, all above listed requirements are combined to form:

$$\frac{p_{Y|H}(Y|H_1)}{p_{Y|H}(Y|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \frac{c_{10} - c_{00} P_0}{c_{01} - c_{11} P_1} = \eta, \quad (4.1)$$

where $c_{10} > c_{00}$ and $c_{01} > c_{11}$, i.e the cost of making the incorrect decision is greater than the cost of making the correct one.

The ratio of the left-hand side of the above equation is the Likelihood Ratio Test (LRT), which we will represent by:

$$\mathcal{L}(Y) = \frac{p_{Y|H}(Y|H_1)}{p_{Y|H}(Y|H_0)}$$

For performance specifications, Probability of Detection (P_D) and Probability of False Alarm (P_F) are found through

$$P_D = Pr(\hat{H} = H_1 | H = H_1) = \int_{\mathcal{Y}_1} p_{Y|H}(y|H_1) dy, \quad (4.2)$$

and

$$P_F = Pr(\hat{H} = H_1 | H = H_0) = \int_{\mathcal{Y}_1} p_{Y|H}(y|H_0) dy, \quad (4.3)$$

where \mathcal{Y}_1 is the decision region of H_1 :

$$\mathcal{Y}_1 = \{y \in \mathcal{Y}, \hat{H}(y) = H_1\},$$

and \hat{H} is the decided decision, which is either H_0 and H_1 . Thus the region of H_0 which is \mathcal{Y}_0 will be

$$\mathcal{Y}_0 = \mathcal{Y} \setminus \mathcal{Y}_1 = \{y \in \mathcal{Y}, \hat{H}(y) = H_0\}$$

The Probability of a Miss (P_M) can be defined as:

$$P_M = 1 - P_D = Pr(\hat{H} = H_0 | H = H_1) = \int_{\mathcal{Y}_0} p_{Y|H}(y|H_1) dy \quad (4.4)$$

A special case from the Binary Hypothesis Testing will be considered, which is the Maximum Likelihood (ML) rule.

Maximum Likelihood Rule

In the ML rule we set the cost assignments to be as follows:

$$c_{ij} = 1 - \delta_{ij} = \begin{cases} 1 & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}, \quad (4.5)$$

and assume the apriories are equal. That is

$$P_0 = P_1 = \frac{1}{2}.$$

Therefore, in equation (4.1), using ML rule, η equals to 1 :

$$\frac{p_{Y|H_1}}{p_{Y|H_0}} \underset{H_0}{\overset{H_1}{\geq}} 1. \quad (4.6)$$

This rule is useful in cases where no costs and apriori are assigned, which is useful in our design process. By this we mean that no hypothesis is favored over another.

Neyman-Pearson

The Neyman-Pearson Approach Found to be suitable when the costs and apriori are unknown and setting $\eta = 1$, as in the ML rule, is unreasonable. A quality criterion which depends on the probabilities of detection and that of false alarm seems more appropriate in such scenario. Since, for any detection process, there are always a probability of detection and a probability of false alarm associated with it. Below is the objective from using of the Neyman-Pearson Approach:

$$\max_{P_F \leq \alpha} P_D.$$

This means that we need to maximize the probability of detection while the probability of false alarm is bounded.

The solution for this maximization will be formulated through

$$\max_{P_F \leq \alpha} P_D = \min_{P_F \leq \alpha} (1 - P_D) = \min_{P_F \leq \alpha} 1 - \int_{\mathcal{Y}_1} p_{Y|H}(y|H_1) dy = \min_{P_F \leq \alpha} \int_{\mathcal{Y}_0} p_{Y|H}(y|H_1) dy,$$

and the constraint is

$$P_F \leq \alpha \iff \int_{\mathcal{Y}_1} p_{Y|H}(y|H_0) dy \leq \alpha.$$

In general, the LRT solution for this approach is

$$\frac{p_{Y|H}(Y|H_1)}{p_{Y|H}(Y|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \zeta, \quad (4.7)$$

where ζ is selected such that $P_F(\zeta) = \alpha$.

4.3 Optimization Using the Large Deviation Approach

Large deviations deal with the behaviour at which tails of probabilities of events of certain distributions decay asymptotically. It is widely used in topics, such as probability, statistics, queueing theory, communication theory and statistical mechanics. Since our quality measures are probability quantities, large deviation concepts seem adequate whenever the problem in the question decrease asymptotically.

4.3.1 The Gärtner-Ellis theorem

Since arbitrary random sequences $\{\mathbf{Y}[\mathbf{n}]\}$ are studied in our work, the Gärtner-Ellis theorem [11] will be possibly useful along with the Frenchel-Legendre transform. The Gärtner-Ellis theorem is an (nonconvex) extension formula produced by twisting the conventional large deviation rate function around a continuous functional. The OC of the LRT's is a plot of how P_F and P_D vary as a function of a threshold \mathbf{r} . Therefore, the Gärtner-Ellis theorem will be applied for minimizing probability of miss (P_M) and probability of false alarm (P_F). We state the theorem below:

Theorem Let $\mathcal{X} = \mathbb{R}^d$ be a complete separable space for each $n \in \mathbb{N}$, having a norming constant $\{a_n, n \in \mathbb{N}\}$ which are a sequence of positive numbers tending to ∞ , $\{(\Omega_n, \mathcal{F}_n, P_n), n \in \mathbb{N}\}$ is the probability space and a random vector \mathbf{W}_n which maps Ω_n into \mathcal{X} . Assuming the limit $\Lambda(\boldsymbol{\theta})$ exists and is finite for $\boldsymbol{\theta} \in \mathbb{R}_d$ defined to be

$$\Lambda(\boldsymbol{\theta}) = \lim_{n \rightarrow \infty} \frac{1}{a_n} \Lambda_n(a_n \boldsymbol{\theta}), \quad (4.8)$$

define

$$\Lambda_n(\boldsymbol{\theta}) = \ln \Phi_{W_n}(\boldsymbol{\theta}), \quad (4.9)$$

where $\Phi_{W_n}(\boldsymbol{\theta})$, the moment generating function of \mathbf{W}_n is given by

$$\Phi_{W_n}(\boldsymbol{\theta}) = E[e^{\langle \boldsymbol{\theta}, \mathbf{W}_n \rangle}]. \quad (4.10)$$

The Fenchel-Legendre transform of $\Lambda(\boldsymbol{\theta})$ is:

$$\Lambda^*(\mathbf{x}) = \sup_{\boldsymbol{\theta} \in \mathbb{R}_d} (\langle \boldsymbol{\theta}, \mathbf{x} \rangle - \Lambda(\boldsymbol{\theta})). \quad (4.11)$$

Then

i) For every closed subset F :

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_n\{\mathbf{W}_n \in F\} \leq -\Lambda^*(F) = -\inf_{\mathbf{x} \in F} \Lambda^*(\mathbf{x}) \quad (4.12)$$

ii) For every open subset G :

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_n\{\mathbf{W}_n \in G\} \geq -\Lambda^*(G) = -\inf_{\mathbf{x} \in G} \Lambda^*(\mathbf{x}). \quad (4.13)$$

Here the norming constant $a_n = n$ which is the filter's length, and the sets are typically $G = \{\mathbf{Y}[\mathbf{n}] < \mathbf{r}\}$ and $F = \bar{G}$.

Special Case:

In the scalar case, let G be the set $(-\infty, r)$

$$\begin{aligned} \Lambda^*(G) &= \inf_{x \in (-\infty, r)} \left\{ \sup_{\theta} (\theta x - \Lambda(\theta)) \right\} \\ &= \inf_{x \in (-\infty, 0)} \left\{ \sup_{\theta} (\theta x + \theta r - \Lambda(\theta)) \right\} \end{aligned} \quad (4.14)$$

Let x^* be an optimal solution and $\theta^*(x^*)$ be the corresponding optimal θ .

1) If $x^* < 0$ and $\theta^*(x^*) > 0$,

we claim that there exists a solution x_0 and $\theta(x_0)$ where x_0 is smaller.

Take $x = x^* - \epsilon$; for $\theta > 0$,

$$\sup_{\theta > 0} (\theta(x^* - \epsilon) + \theta r - \Lambda(\theta)) < \sup_{\theta > 0} (\theta x^* + \theta r - \Lambda(\theta)) \leq (\theta^* x^* + \theta^* r - \Lambda(\theta^*)).$$

Assuming continuity of θ^* , $\theta^*(x^* - \epsilon)$ will yield a supremum value that is smaller or equal. Since x^* is the solution, the values of sup will be equal but $x_0 = x^* - \epsilon < x^*$, which is a contradiction.

2) If $x^* < 0$ and $\theta^*(x^*) < 0$,

we claim that there exists a solution x_0 and $\theta(x_0)$ where x_0 is larger.

Take $x = x^* + \epsilon$; for $\theta < 0$,

$$\sup_{\theta < 0} (\theta(x^* + \epsilon) + \theta r - \Lambda(\theta)) < \sup_{\theta < 0} (\theta x^* + \theta r - \Lambda(\theta)) \leq (\theta^* x^* + \theta^* r - \Lambda(\theta^*)).$$

Assuming continuity of θ^* , $\theta^*(x^* + \epsilon)$ will yield a supremum value that is smaller or equal. Since x^* is the solution, the values of sup will be equal but $x_0 = x^* + \epsilon > x^*$, which is a contradiction.

3) if $\theta^*(x^*) = 0$,

$$\Lambda^*(G) = -\Lambda(0),$$

so we get the value of $-\Lambda(0)$.

4) $x^* = 0$,

$$\Lambda^*(G) = \sup_{\theta} (\theta r - \Lambda(\theta))$$

i) $\theta^*(x^*) > \delta$,

Take $x = -\epsilon$; for $\theta > \delta$,

$$\sup_{\theta > \delta} (\theta(-\epsilon) + \theta r - \Lambda(\theta)) < \sup_{\theta > \delta} (\theta r - \Lambda(\theta)) \leq \theta^* r - \Lambda(\theta^*).$$

Assuming continuity of θ^* , $\theta^*(-\epsilon)$ will yield a supremum value that is smaller or equal. Since x^* is the solution, the values of sup will be equal but $x_0 = -\epsilon < x^*$, which is a contradiction.

In conclusion $x^* = 0$ and $\theta^*(x^*) \leq 0$ is the solution.

The Gärtner-Ellis theorem becomes for the scalar case:

$$P(Y[n] < r) \cong e^{-n\Lambda^*(r)} \quad (4.15)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P(Y[n] < r) = -\Lambda^*(r)$$

where the rate $\Lambda^*(r)$ is the Fenchel-Legendre transform of $\Lambda(\theta)$, i.e.

$$\Lambda^*(r) = \sup_{\theta < 0} (\langle \theta, r \rangle - \Lambda(\theta)) \quad (4.16)$$

where $\Lambda(\theta)$ is the cumulant generating function $Y[n]$

$$\Lambda(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_n(n\theta), \quad (4.17)$$

and

$$\Lambda_n(\theta) = \ln \Phi_{Y_n}(\theta), \quad (4.18)$$

and

$$\Phi_{Y_n}(\theta) = E[e^{\langle \theta, Y[n] \rangle}]. \quad (4.19)$$

Here we introduced the Gärtner-Ellis theorem which states in layman terms that when the scaled cumulant generating function $\Lambda(\boldsymbol{\theta})$ of $\mathbf{Y}[\mathbf{n}]$ is differentiable, then $\mathbf{Y}[\mathbf{n}]$ obeys a large deviation principle with a rate function $\Lambda^*(\mathbf{r})$ which is the Legendre-Fenchel transform of $\Lambda(\boldsymbol{\theta})$.

In what follows, for each model we study, the Gärtner-Ellis theorem will be used in order to find the best quantization points of the filter to increase its probability of detection (P_D), i.e decrease its probability of miss (P_M) and minimize probability of false alarm (P_F).

Chapter 5

Model 1: The Received Signal is Subject to Additive Gaussian Noise

We assume the received signal is subject to Additive Gaussian Noise. The model description is explained in section 5.1. Ways to optimize the receiver's device are applied using Neyman Pearson in section 5.3, Maximum Likelihood in section 5.4 and Large Deviations in section 5.5. Likelihood Ratio Test is discussed section 5.2.

5.1 Model Description

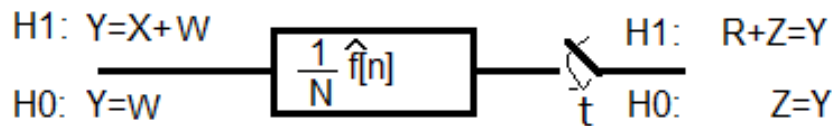


Figure 5.1: Model 1 at the Receiver

Lets consider the above model, signal $X[k]$ is subject to a White Gaussian Additive Noise $W[k]$ with $W[k] \sim \mathcal{N}(0, \sigma^2)$. The received signal $Y[k]$ is convolved by filter \hat{f} . After sampling at $t = K$, this will produce

1) H_0 : $Y = Z$ where,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[K-i] \quad (5.1)$$

and $Z \sim \mathcal{N}(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2})$.

therefore,

$$Y \sim \mathcal{N}\left(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2}\right).$$

which we will represent as

$$Y \sim \mathcal{N}(0, \sigma_1^2).$$

2) H_1 : $Y = R + Z$ where,

$$R = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]x[K-i] \quad (5.2)$$

and

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[K-i]$$

therefore,

$$Y \sim \mathcal{N}\left(\frac{\sum \hat{f}[i]x[K-i]}{N}, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2}\right).$$

which we will represent it as

$$Y \sim \mathcal{N}(\mu_K, \sigma_1^2).$$

In section 3.2, we stated two different scenarios, which we will study in this chapter. Section 5.3 to section 5.5 will use scenario 1 to optimize the filter's performance. Scenario 2 will be analysed in section 5.6.

5.2 The Likelihood Ratio Test

The Likelihood Ratio Test (LRT) parametrized by ζ is

$$\frac{p_{Y|H_1}}{p_{Y|H_0}} \underset{H_0}{\overset{H_1}{\geq}} \zeta, \quad \zeta > 0.$$

We will discuss two cases, the real case signal in section 5.2.1, and the complex case signal in section 5.2.2. In both cases we set $\frac{1}{N} \sum_{i=0}^{N-1} |\hat{f}[i]|^2 = 1$, i.e. $\hat{f}[i]$ has a unit power, and $\sigma_1^2 = \frac{\sigma^2}{N}$.

5.2.1 Real Case

Assuming real quantities, for $\mu \geq 0$ the LRT is,

$$\frac{\frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y-\mu_K)^2}{2\sigma_1^2}}}{\frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{Y^2}{2\sigma_1^2}}} \underset{H_0}{\overset{H_1}{\geq}} \zeta, \quad (5.3)$$

or equivalently,

$$Y^2 - (Y - \mu_K)^2 \underset{H_0}{\overset{H_1}{\geq}} 2\sigma_1^2 \ln \zeta, \quad (5.4)$$

$$Y \underset{H_0}{\overset{H_1}{\geq}} \frac{2\sigma_1^2 \ln \zeta + \mu_K^2}{2\mu_K} \hat{=} \eta. \quad (5.5)$$

5.2.2 Complex Case

Assuming complex quantities,

$$\frac{\frac{1}{2\pi\sigma_1^2} e^{-\frac{|Y-\mu_K|^2}{2\sigma_1^2}}}{\frac{1}{2\pi\sigma_1^2} e^{-\frac{|Y|^2}{2\sigma_1^2}}} \underset{H_0}{\overset{H_1}{\geq}} \zeta, \quad (5.6)$$

equivalent to

$$|Y|^2 - |Y - \mu_K|^2 \underset{H_0}{\overset{H_1}{\geq}} 2\sigma_1^2 \ln \zeta, \quad (5.7)$$

$$\begin{array}{c} H_1 \\ \Re(Y e^{-j\theta_{\mu_K}}) \geq \frac{2\sigma_1^2 \ln \zeta + |\mu_K|^2}{2|\mu_K|} \hat{=} \eta. \\ < \\ H_0 \end{array} \quad (5.8)$$

5.3 Neyman-Pearson Approach

Neyman-Pearson will be studied for real signal case in section 5.3.1 and complex signal case 5.3.2.

5.3.1 Real Case

For $\mu > 0$, using η defined in equation (5.5),

The Probability of a Miss (P_M):

$$P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(Y-\mu_K)^2}{2\sigma_1^2}} dY. \quad (5.9)$$

Setting $U = \frac{Y-\mu_K}{\sigma_1}$, so equation (5.9) becomes

$$P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_{-\infty}^{\frac{\eta-\mu_K}{\sigma_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{U^2}{2}} dU. \quad (5.10)$$

If one were to minimize P_M then one should minimize $\frac{\eta-\mu}{\sigma_1}$.

Now we will consider the case of a probability of a false alarm (P_F):

$$P_F = P(\hat{H} = H_1 | H_0) = \int_{\eta}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{Y^2}{2\sigma_1^2}} dY. \quad (5.11)$$

Let $U = \frac{Y}{\sigma_1}$, so equation (5.11) will change to

$$P_F = P(\hat{H} = H_1 | H_0) = \int_{\frac{\eta}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{U^2}{2}} dU \quad (5.12)$$

and one should maximise $\frac{\eta}{\sigma_1}$ to minimize the probability of false Alarm.

Using the Neyman-Pearson approach, we upper bound P_F by α and maximize P_D .

$$\text{” arg max}_{P_F \leq \alpha} P_D \text{”} \Leftrightarrow \text{arg max } P_D + \lambda(P_F - \alpha) \quad (5.13)$$

where λ is the Lagrange multiplier of the constraint.

Note that for $\mu_K < 0$, P_M will be $1 - P_M$ and P_F will be $1 - P_F$ found above, that is maximize $\frac{\eta - \mu}{\sigma_1}$ is equivalent to minimize $\frac{-\eta + \mu}{\sigma_1}$, and the problems are equivalent.

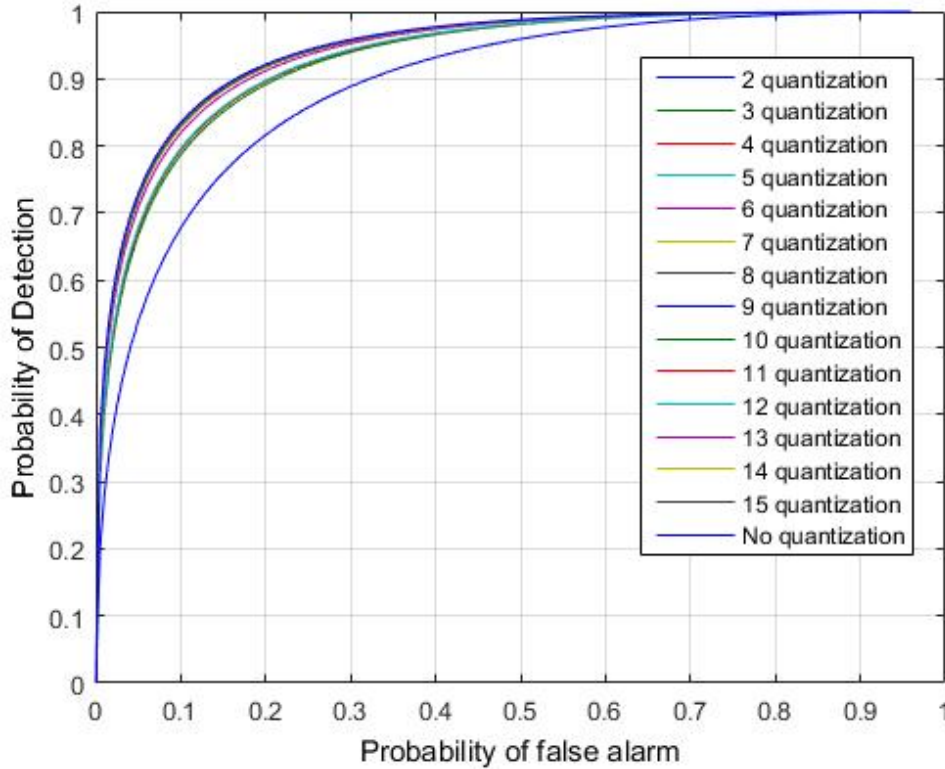


Figure 5.2: PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 4 dB. The performance with no quantization is shown using solid line.

Numerical Results

For the real case, best \hat{f} was searched in Matlab using `fmincon`. For each number of quantization points, P_F was set to be at most to α , and the value of α ranged from 0.001 to 0.99. While changing the value of α , it was found that the quantization points were equal, see table 5.1. We conjecture that optimal \hat{f} is not dependant on α (for a given SNR). In addition, the value of the noise variance was varied from 10 to 10000. The difference between \hat{f} values was found to be also equal, refer to table 5.1 and table 5.2 to compare the quantization points for SNR=4 db and SNR=-2.9287dB. Based on the above, for a given number of quantization points, we propose using one quantized filter \hat{f} for all operating

points and SNRs. Figure 5.2 is one sample from the produced results of P_D vs $P_F = \alpha$ for SNR= 4 dB during the search process of finding the quantization points. Figure 5.3 shows a similar results for SNR= 9.2898 dB. We conclude from both graphs, starting at three quantization points the filter’s performance is “close” to the original one. Figure 5.4 is a comparison between the theoretical plot given five quantization points and the simulation of 100000 packets using the same filter \hat{f} produced from the fmincon search process for SNR=4dB, which validates our theoretical analysis.

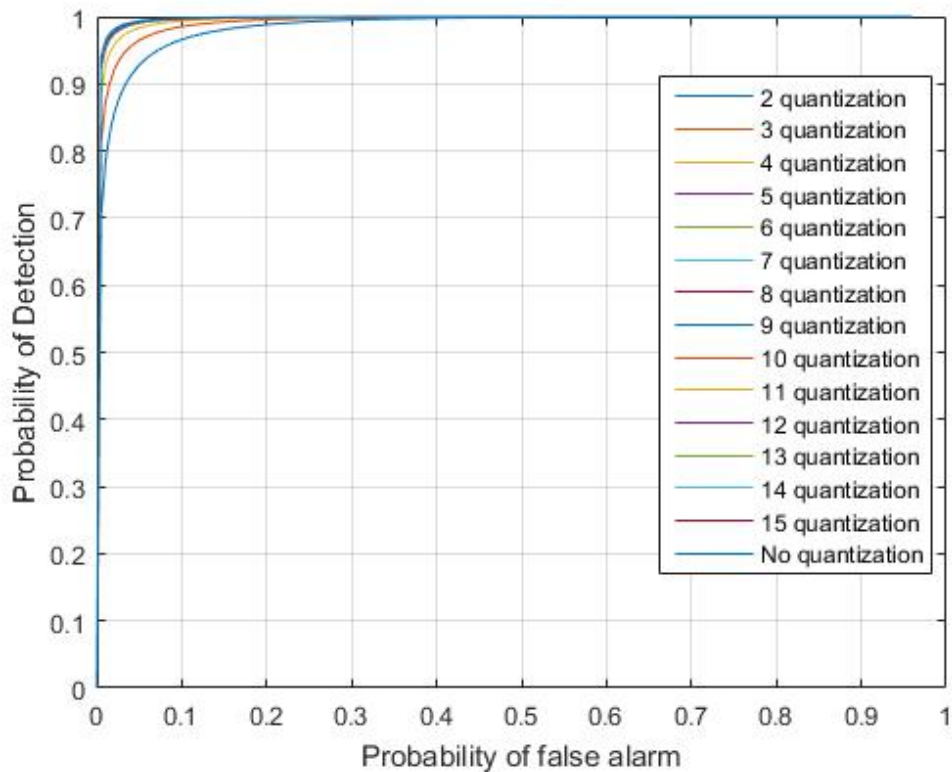


Figure 5.3: P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, SNR= 9.2898dB. The performance with no quantization is shown using solid line.

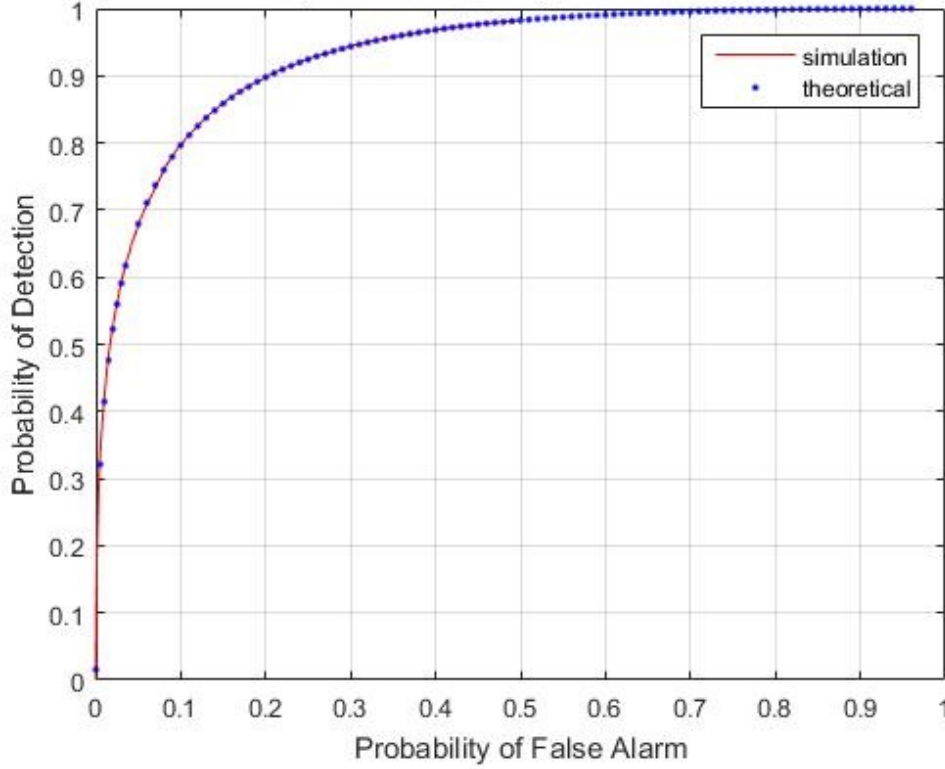


Figure 5.4: P_D vs P_F for five quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR= 4 dB.

5.3.2 Complex Case

Using η defined in equation (5.8), the Probability of a Miss (P_M) is:

$$P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_{\Re(Ye^{-j\theta_{\mu_K}}) < \eta} \frac{1}{2\pi\sigma_1^2} e^{-\frac{|Y-\mu_K|^2}{2\sigma_1^2}} dY, \quad (5.14)$$

Let $V = (Ye^{-j\theta_{\mu_K}})$,

$$P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_{\Re(V) \leq \eta} \frac{1}{2\pi\sigma_1^2} e^{-\frac{|Ve^{j\theta_{\mu_K}} - \mu_K|^2}{2\sigma_1^2}} dV = \int_{\Re(V) \leq \eta} \frac{1}{2\pi\sigma_1^2} e^{-\frac{|V-\mu_1 e^{-j\theta_{\mu_K}}|^2}{2\sigma_1^2}} dV, \quad (5.15)$$

Set $U = \frac{V - |\mu_K|}{\sigma_1}$, so equation (5.15) becomes

$$P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_{\Re(U) \leq \frac{\eta - |\mu_K|}{\sigma_1}} \frac{1}{2\pi} e^{-\frac{|U|^2}{2}} dU. \quad (5.16)$$

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.039983292683474	-0.999682864612971
0.05	0.039983292683474	-0.999682864612971
0.15	0.039983292683474	- 0.999682864612971
0.2	0.039983292683474	-0.999682864612971
0.4	0.039983292683474	- 0.999682864612971
0.5	0.039983292683474	- 0.999682864612971
0.6	0.039983292683474	- 0.999682864612971
0.7	0.039983292683474	-0.999682864612971
0.8	0.039983292683474	- 0.999682864612971
0.9	0.039983292683474	-0.999682864612971
0.99	0.039983292683474	- 0.999682864612971

Table 5.1: The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=4 dB.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.039983292683474	-0.999682864612971
0.05	0.039983292683474	-0.999682864612971
0.15	0.039983292683474	-0.999682864612971
0.2	0.039983292683474	-0.999682864612971
0.4	0.039983292683474	-0.999682864612971
0.5	0.039983292683474	-0.999682864612971
0.6	0.039983292683474	-0.999682864612971
0.7	0.039983292683474	- 0.999682864612971
0.8	0.039983292683474	-0.999682864612971
0.9	0.039983292683474	- 0.999682864612971
0.99	0.039983292683474	-0.999682864612971

Table 5.2: The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=-2.9287 dB.

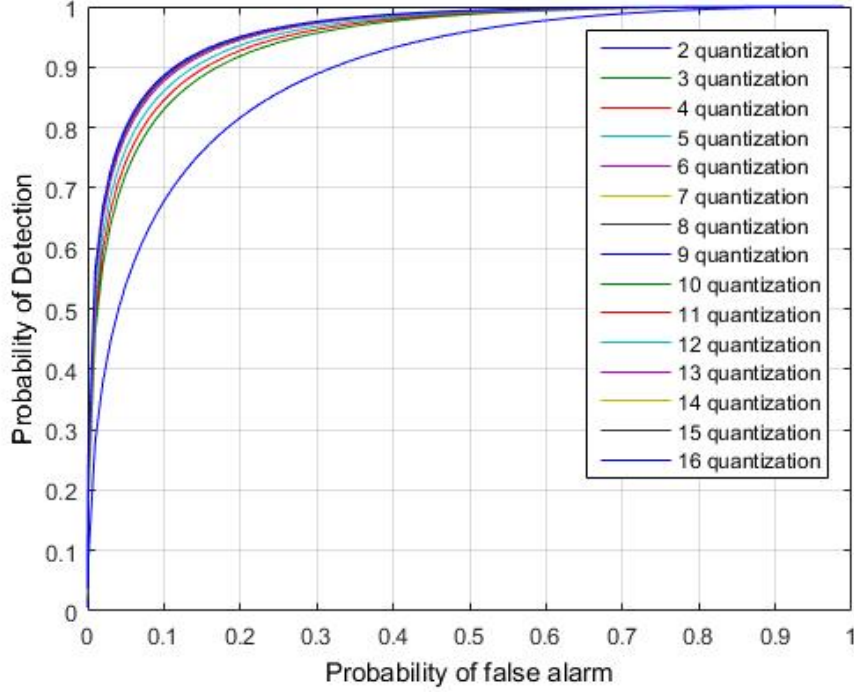


Figure 5.5: PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zaddoff Chu sequence detection, SNR= 4.9136dB. The performance with no quantization is shown using solid line of 16 quantization points.

Thus one should minimize $\frac{\eta - |\mu_K|}{\sqrt{\sigma_1}}$ if one were to minimize (P_M).

Now we will consider the case of a probability of a false alarm (P_F):

$$P_F = P(\hat{H} = H_1 | H_0) = \int_{\Re(Y e^{-j\theta\mu_K}) \geq \eta} \frac{1}{2\pi\sigma_1^2} e^{-\frac{|Y|^2}{2\sigma_1^2}} dY \quad (5.17)$$

Let $V = (Y e^{-j\theta\mu_K})$,

$$P_F = P(\hat{H} = H_1 | H_0) = \int_{\Re(V) \geq \eta} \frac{1}{2\pi\sigma_1^2} e^{-\frac{|V|^2}{2\sigma_1^2}} dV = \int_{\Re(U) \geq \frac{\eta}{\sigma_1}} \frac{1}{2\pi} e^{-\frac{|U|^2}{2}} dU \quad (5.18)$$

Thus we should maximise $\frac{\eta}{\sigma_1}$ to minimize the probability of false Alarm. The Neyman-Pearson probability is hence

$$” \arg \max_{P_F \leq \alpha} P_D ” \Leftrightarrow \arg \max P_D + \lambda(P_F - \alpha) \quad (5.19)$$

Numerical Results

Similar procedure was applied as in real case and the results were found to be identical to the real case results, refer to table 5.3 and table 5.4. Thus same conclusion is made and one quantized filter \hat{f} maybe used for all operating points and SNRs. The small difference in the values of quantization points when SNR is changed is due to numerical precision. Figure 5.5 is a sample case from the produced results of P_D vs $P_F = \alpha$ for SNR=4.9136 dB obtained by the search process of finding the quantization points. Figure 5.6 shows a similar results for SNR= 0.14248 dB. We conclude from both figures, starting from three quantization points the performance of the filter is close to the original one. Figure 5.7 is a comparison between the theoretical plot given four quantization points and the simulation of 100000 packets using the same filter \hat{f} produced from the fmincon search process for SNR=4.9136dB, which validates our theoretical analysis.

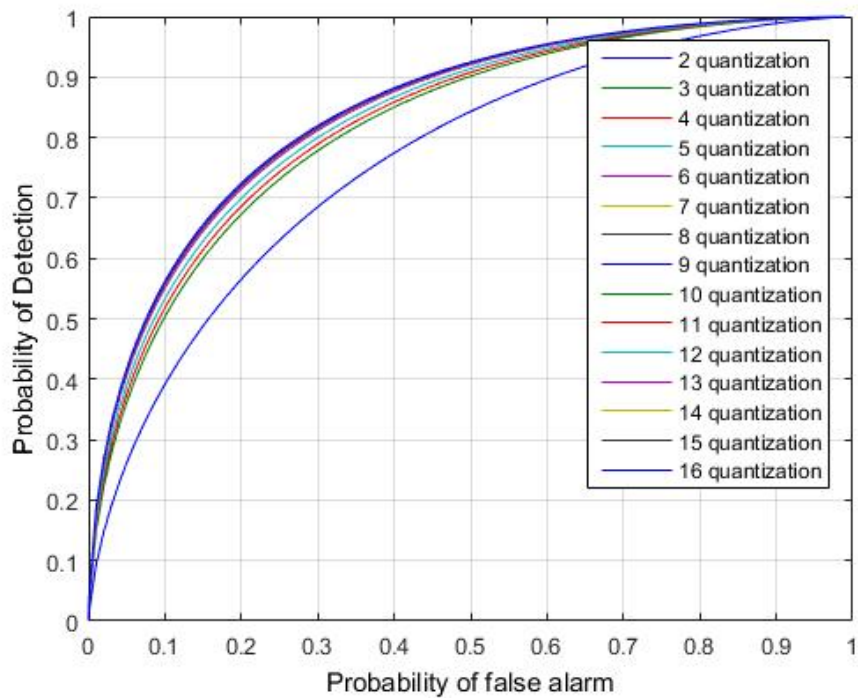


Figure 5.6: P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zadoff Chu sequence detection, SNR= 0.1424dB. The performance with no quantization is shown using solid line of 16 quantization points.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.99390989982605 + 0.46477139217882i	-0.91599532894826 - 0.14714914075528i
0.05	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.15	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.2	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.4	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.5	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.6	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.7	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.8	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.9	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i
0.99	0.94823359967661 + 0.46828407660848i	-0.94357658477569 - 0.15532369871884i

Table 5.3: The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu, for SNR=4.9136 dB.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.99649426948682 + 0.46388104997698i	-0.90860769404907 - 0.15516409372289i
0.05	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.15	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.2	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.4	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.5	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.6	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.7	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.8	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.9	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i
0.99	0.95363673856140 + 0.45914657952403i	-0.94458800741223 - 0.14462309677832i

Table 5.4: The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu, for SNR=0.1424 dB.

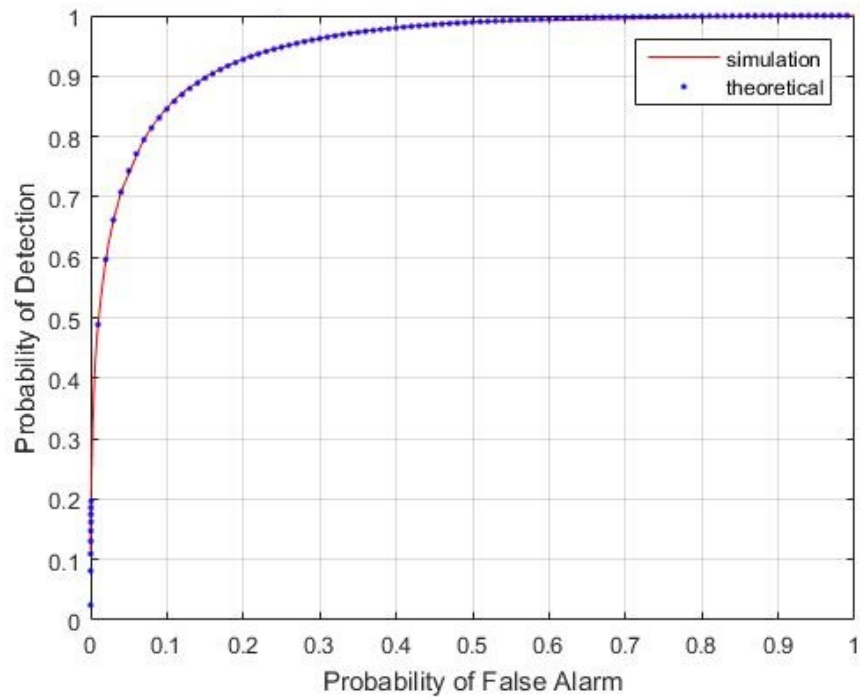


Figure 5.7: PD vs P_F for five quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Zadoff Chu, for SNR=4.9136 dB.

5.4 Maximum Likelihood Rule

The ML device is a LRT parameter (1) and is a min distance device,

$$\frac{p_{Y|H_1}}{p_{Y|H_0}} \underset{H_0}{\overset{H_1}{\geq}} 1$$

For Real quantities, the decision rule for $\mu \geq 0$

$$Y \underset{H_0}{\overset{H_1}{\geq}} \frac{\mu_K}{2} \tag{5.20}$$

and the probability of error $Q(\frac{\mu_K}{2\sigma_1})$. Hence one should we should maximise $\frac{\mu_K}{\sigma_1}$, i.e.

$$\arg \max_f \left(\frac{\mu_K}{\sigma_1} \right) = \arg \max_{\hat{f}: \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]=1} \frac{\sum_{i=-\infty}^{\infty} \hat{f}[i] x[K-i]}{N\sigma_1} = \arg \max_{\hat{f}: \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]=1} = \sum_{i=-\infty}^{\infty} \hat{f}[i] x[K-i] \quad (5.21)$$

From equation (5.21), it is concluded that we should maximise $\sum_{i=-\infty}^{\infty} \hat{f}[i] x[K-i]$

Now if $\mu_K < 0$ the opposite case will be required that is minimize $\frac{\mu_K}{\sigma_1}$ to minimize error.

Note we need to maximise $|\mu_K|$ in both cases.

For Complex Case, the ML device used will be

$$\frac{\frac{1}{2\pi\sigma_1^2} e^{-\frac{|Y-\mu_K|^2}{2\sigma_1^2}}}{\frac{1}{2\pi\sigma_1^2} e^{-\frac{|Y|^2}{2\sigma_1^2}}} \begin{matrix} H_1 \\ \geq 1 \\ < \\ H_0 \end{matrix} \quad (5.22)$$

equivalent to,

$$\Re(Y\mu_K^*) \begin{matrix} H_1 \\ \geq \frac{\mu_K\mu_K^*}{2} \\ < \\ H_0 \end{matrix} \quad (5.23)$$

with probability of error is $Q(\frac{|\mu_K|}{2\sigma_1})$ and we should maximise $\frac{|\mu_K|}{\sigma_1}$ which is as in the real case,

$$\arg \max_f \left(\frac{|\mu_K|}{\sigma_1} \right) = \arg \max_{\hat{f}: \frac{1}{N} \sum_{i=0}^{N-1} \|\hat{f}[i]\|^2=1} \frac{\sum_{i=-\infty}^{\infty} |\hat{f}[i] x^*[K-i]|}{N\sigma_1} = \arg \max_{\hat{f}: \frac{1}{N} \sum_{i=0}^{N-1} \|\hat{f}[i]\|^2=1} \sum_{i=-\infty}^{\infty} |\hat{f}[i] x^*[K-i]| \quad (5.24)$$

5.5 Minimising Error Rate using Large Deviations

As stated in section 3.2 the amplitude at location K of the signal will first be checked. If it is less than a threshold \mathbf{r} it will be an error. Using the Gärtner-Ellis theorem introduced in section 4.3, we minimize both the probability of a miss (P_M) and probability of false alarm (P_F) in the experiments. Each case will be studied for both real (section 5.5.1) and complex (section 5.5.2) signals.

5.5.1 For Real Signal Cases

The Probability of a Miss (P_M)

Given hypothesis H_1 , the output signal $Y = R + Z$, as discussed in section 5.1, Y is distributed as

$$Y \sim \mathcal{N} \left(\frac{\sum_{i=0}^{N-1} \hat{f}[i] X[K-i]}{N}, \frac{\sigma^2 \sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N^2} \right)$$

Then, starting from equation (4.19),

$$\Phi_{Y_N}(\theta) = e^{\theta \frac{\sum_{i=0}^{N-1} \hat{f}[i] X[K-i]}{N} + \frac{\theta^2}{2} \frac{\sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N^2}} \quad (5.25)$$

then

$$\Lambda_N(N\theta) = \theta \sum_{i=0}^{N-1} \hat{f}[i] X[K-i] + \sigma^2 \frac{\theta^2}{2} \sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i] \quad (5.26)$$

Now take

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = \lim_{N \rightarrow \infty} \theta \frac{\sum_{i=0}^{N-1} \hat{f}[i] X[K-i]}{N} + \sigma^2 \frac{\theta^2}{2} \lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N} \quad (5.27)$$

denoting $\mu = \lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} f[i] X[K-i]}{N}$, equation (5.27) becomes,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = \theta \mu + \sigma^2 \frac{\theta^2}{2} \lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N} \quad (5.28)$$

If we set $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N} = 1$, i.e \hat{f} has a “unit power”,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = \theta \mu + \sigma^2 \frac{\theta^2}{2} \quad (5.29)$$

Therefore, equation (4.17), which represents the Fenchel-Legendre transform, can be written as,

$$\Lambda^*(\tau) = \sup_{\theta < 0} \left(\theta(\tau - \mu) - \frac{\theta^2}{2} \sigma^2 \right) \quad (5.30)$$

Calculate the derivative and set it to zero,

$$\frac{\partial \left(\theta(\tau - \mu) - \frac{\theta^2}{2} \sigma^2 \right)}{\partial \theta} = 0 \quad (5.31)$$

and hence, the optimal θ is

$$\theta = \begin{cases} \frac{(\tau - \mu)}{\sigma^2} & \text{Acceptable if } \tau - \mu \leq 0 \\ & \text{i.e } \tau \leq \mu \\ 0 & \text{o.w} \end{cases} \quad (5.32)$$

substitute θ in $\Lambda^*(\tau)$ in a given \hat{f} :

$$\Lambda^*(\tau) = \begin{cases} \frac{(\tau - \mu)^2}{2\sigma^2} & \text{if } \tau \leq \mu \\ 0 & \text{o.w} \end{cases}$$

Next we find the maximum of $\Lambda^*(\tau)$ over $\hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i] = 1$

$$\arg \max_{\substack{\hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i] = 1 \\ \tau \leq \mu}} \left(\frac{(\tau - \mu)^2}{2\sigma^2} \right) = \arg \max_{\substack{\hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i] = 1 \\ \tau \leq \mu}} (\mu - \tau)^2 = \arg \max_{\substack{\hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i] = 1 \\ \tau \leq \mu}} \mu$$

Therefore, after optimizing over \hat{f}

$$\Lambda^*(\tau) = \begin{cases} \frac{1}{2\sigma^2} \max_{\hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i] = 1} (\mu - \tau)^2 & \tau \leq \mu \\ 0 & \text{o.w.} \end{cases} \quad (5.33)$$

Figures 5.8 and 5.9 plot the term $\Lambda^*(\tau)$ (green line), $\Lambda(\theta)$ (red line) and $\theta\tau$ for $\mu = 0.5244$ in Figure 5.8 and $\mu = 1.0489$ in Figure 5.9. The parameter τ was equal to 0.2622 in both figures. Figure 5.9 resulted in a higher peak of $\Lambda^*(\tau)$ since its mean μ is higher which confirms our analysis.

The Probability of False Alarm (P_F)

Consider hypothesis H_0 , the output signal $Y = Z$, as discussed in Section 5.1, Y is distributed as

$$Y \sim \mathcal{N} \left(0, \frac{\sigma^2 \sum_{i=0}^{L-1} \hat{f}[i]\hat{f}[i]}{N^2} \right)$$

Then, starting from equation (4.19),

$$\Phi_{R_N}(\theta) = e^{\frac{\theta^2}{2} \frac{\sigma^2 \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i]}{N^2}} \quad (5.34)$$

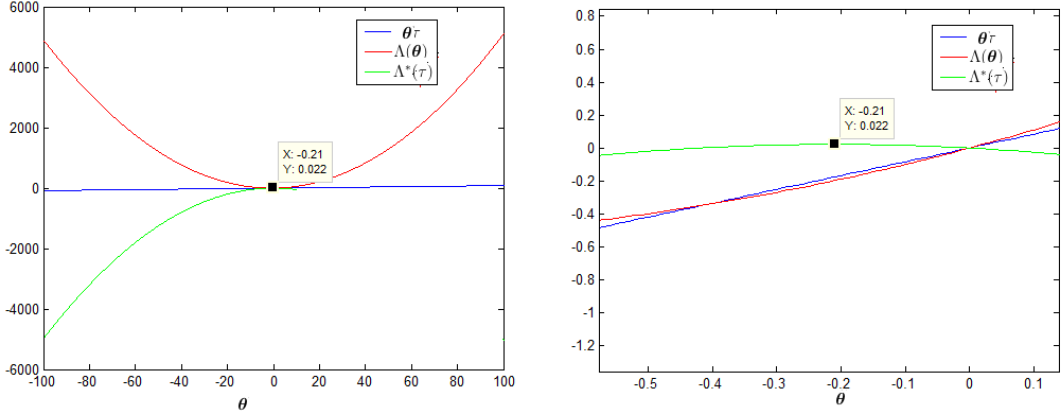


Figure 5.8: plot the results of $\Lambda^*(\tau)$, $\Lambda(\theta)$ and $\theta\tau$ for $\mu = 0.5244$

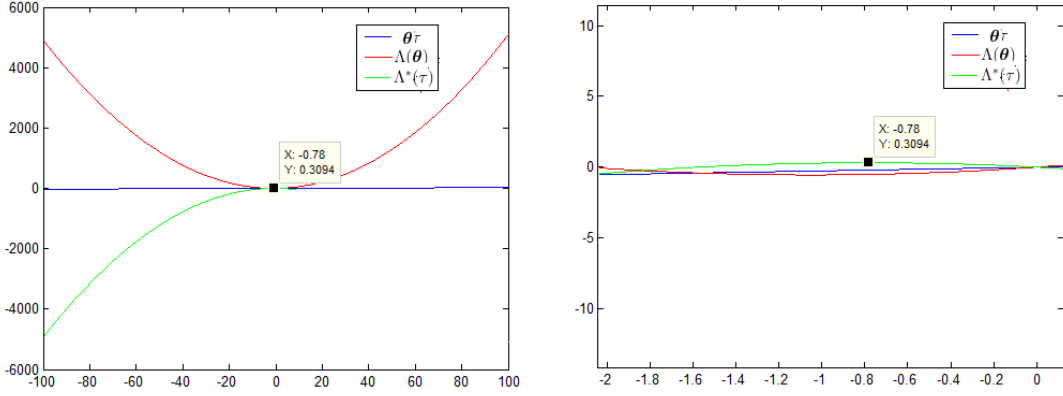


Figure 5.9: plot the results of $\Lambda^*(\tau)$, $\Lambda(\theta)$ and $\theta\tau$ for $\mu = 1.0489$

then

$$\Lambda_N(N\theta) = \sigma^2 \frac{\theta^2}{2} \sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i] \quad (5.35)$$

Now take

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = \sigma^2 \frac{\theta^2}{2} \lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N} \quad (5.36)$$

setting $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}[i] \hat{f}[i]}{N} = 1$, i.e \hat{f} is a unit power,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = \sigma^2 \frac{\theta^2}{2} \quad (5.37)$$

Therefore, equation (4.17), which represents Fenchel-Legendre transform, can be written as,

$$\Lambda^*(\tau) = \sup_{\theta > 0} \left(\theta\tau - \frac{\theta^2}{2}\sigma^2 \right) \quad (5.38)$$

Calculate the derivative and set it to zero,

$$\frac{\partial \left(\theta\tau - \frac{\theta^2}{2}\sigma^2 \right)}{\partial \theta} = 0 \quad (5.39)$$

since $\tau \geq 0$, we will get,

$$\theta = \left(\frac{\tau}{\sigma^2} \right) \quad (5.40)$$

substitute θ in equation (5.38)

$$\Lambda^*(\tau) = \left(\frac{\tau^2}{2\sigma^2} \right).$$

Now maximizing over \hat{f}

$$\Lambda^*(\tau) = \max_{\hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \hat{f}[i]\hat{f}[i] = 1} \left(\frac{\tau^2}{2\sigma^2} \right) = \frac{\tau^2}{2\sigma^2}$$

Numerical Results

For the real case, best \hat{f} was searched in Matlab using `fmincon`, starting with the points obtained using Neyman Pearson Approach. It was found that the optimal \hat{f} points were equal to the points from Neyman Pearson Approach. Based on the above, same conclusions are made as in section 5.3.1. After that, $\Lambda^*(\tau)$ for P_M and P_F were found and plotted. Figure 5.10 is one sample from the produced results of $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for SNR= 9.2898 dB during the search process of finding the quantization points. It is clear that there is better performance when the number of quantization points increases, which is logical. Figure 5.11 is another plot of $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for SNR= 17.0713 dB during the search process of finding the quantization points. Since it has a higher SNR than figure 5.10, figure 5.11 shows better performance.

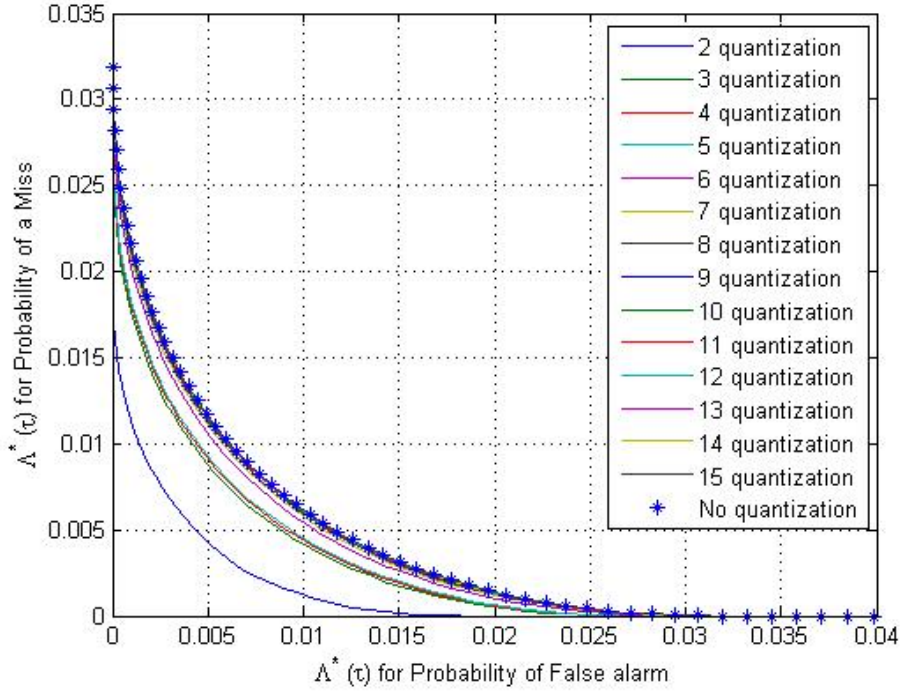


Figure 5.10: $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 9.2898 dB. The performance with no quantization is shown using ‘*’ line.

5.5.2 For complex Signal Cases

The Probability of a Miss (P_M)

To minimize P_M , consider hypothesis H_1 having the output signal $Y = R + Z$ as discussed in the Model Description Section. It follows

$$Y \sim \mathcal{N} \left(\frac{\sum_{i=0}^{N-1} \hat{f}[i]X[K-i]}{N}, \frac{\sigma^2 \sum_{i=0}^{L-1} \hat{f}[i]\hat{f}[i]}{N^2} \right)$$

However, for complex signal analysis using the Gärtner-Ellis theorem, it is more practical to deal with it as a 2D variable. That is, Y 's mean \mathbf{m}_Y and $\mathbf{\Lambda}_Y$ covariance matrix are represented as a vector quantity consisting of its real and

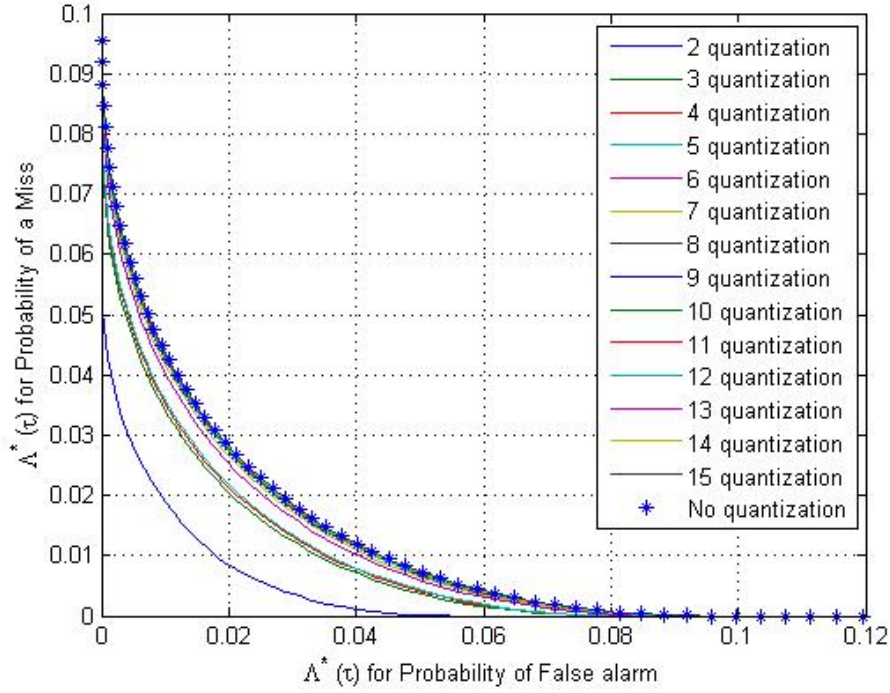


Figure 5.11: $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 17.0713 dB. The performance with no quantization is shown using ‘*’ line.

imaginary parts such that

$$\mathbf{m}_Y = \begin{bmatrix} \frac{\sum_{i=0}^{N-1} \hat{f}_R[i]X_R[K-i] - \hat{f}_I[i]X_I[K-i]}{N} \\ \frac{\sum_{i=0}^{N-1} \hat{f}_R[i]X_I[K-i] + \hat{f}_I[i]X_R[K-i]}{N} \end{bmatrix} = \begin{bmatrix} m_R \\ m_I \end{bmatrix} \quad (5.41)$$

$$\mathbf{\Lambda}_Y = \begin{bmatrix} \frac{\sigma^2 \sum_{i=0}^{L-1} \hat{f}_R[i]\hat{f}_R[i] + \hat{f}_I[i]\hat{f}_I[i]}{N^2} & 0 \\ 0 & \frac{\sigma^2 \sum_{i=0}^{L-1} \hat{f}_R[i]\hat{f}_R[i] + \hat{f}_I[i]\hat{f}_I[i]}{N^2} \end{bmatrix} = \sigma^2 \frac{\|\hat{f}\|^2}{N^2} \mathbf{I} \quad (5.42)$$

Thus,

$$Y \sim \mathcal{N}(\mathbf{m}_Y, \mathbf{\Lambda}_Y)$$

Here $\boldsymbol{\theta}$ and \mathbf{r} are two dimensional, denoted as

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad \mathbf{r} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}$$

Thus, starting from equation (4.19) ,

$$\Phi_{R_N}(\boldsymbol{\theta}) = e^{\theta_1 m_1 + \theta_2 m_2 + \frac{1}{2} \boldsymbol{\theta}^T \boldsymbol{\Lambda}_Y \boldsymbol{\theta}}, \quad (5.43)$$

and

$$\begin{aligned} \Lambda_N(N\boldsymbol{\theta}) &= N\theta_1 m_1 + N\theta_2 m_2 + \frac{N^2}{2} \boldsymbol{\theta}^T \boldsymbol{\Lambda}_Y \boldsymbol{\theta} \\ &= \theta_1 \sum_{i=0}^{N-1} (\hat{f}_R[i] X_R[K-i] - \hat{f}_I[i] X_I[K-i]) + \theta_2 \sum_{i=0}^{N-1} (\hat{f}_R[i] X_I[K-i] + \hat{f}_I[i] X_R[K-i]) \\ &\quad + \frac{1}{2} \sigma^2 \|\hat{f}\|^2 \boldsymbol{\theta}^T \mathbf{I} \boldsymbol{\theta} \end{aligned} \quad (5.44)$$

Now take

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \theta_1 \frac{\sum_{i=0}^{N-1} \hat{f}_R[i] X_R[K-i]}{N} + \lim_{N \rightarrow \infty} \theta_2 \frac{\sum_{i=0}^{N-1} \hat{f}_I[i] X_I[K-i]}{N} + \sigma^2 \frac{1}{2} \lim_{N \rightarrow \infty} \frac{\|\hat{f}\|^2 \boldsymbol{\theta}^T \mathbf{I} \boldsymbol{\theta}}{N} \quad (5.45)$$

let $\boldsymbol{\mu} = [\mu_K \ \mu_2]$ where $\mu = \lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} f[i] X[K-i]}{N}$, so equation (5.45) becomes,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \theta_1 \mu_K + \theta_2 \mu_2 + \sigma^2 \frac{1}{2} \lim_{N \rightarrow \infty} \frac{\|\hat{f}\|^2 \boldsymbol{\theta}^T \mathbf{I} \boldsymbol{\theta}}{N} \quad (5.46)$$

Setting $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}_R[i] \hat{f}_R[i] + \hat{f}_I[i] \hat{f}_I[i]}{N} = 1$, i.e $\hat{f}[i]$ is unit-power. Thus $\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta})$ will equal to,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \theta_1 \mu_K + \theta_2 \mu_2 + \frac{\sigma^2}{2} (\theta_1^2 + \theta_2^2) \quad (5.47)$$

Therefore, equation (4.17), which represents Fenchel-Legendre transform, can be written as,

$$\Lambda^*(\mathbf{r}) = \sup_{\boldsymbol{\theta} \in \mathbb{R}^2} \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - \langle \boldsymbol{\theta}, \boldsymbol{\mu} \rangle - \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{2} \sigma^2 \right) \quad (5.48)$$

This means find the maximum of $\Lambda^*(\boldsymbol{\tau})$

$$\text{over } \hat{f} : \frac{1}{N} \sum_{i=0}^{N-1} \|\hat{f}[i]\|^2 = 1$$

Calculate the derivatives and set them to zero,

$$\frac{\partial \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - \langle \boldsymbol{\theta}, \boldsymbol{\mu} \rangle - \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{2} \sigma^2 \right)}{\partial \theta_i} = 0 \quad (5.49)$$

for $i = 1, 2$,

$$\forall i \quad \theta_i = \begin{cases} \frac{(\tau_i - \mu_i)}{\sigma^2} & \text{if } \tau_i - \mu_i \leq 0 \\ & \text{i.e } \tau_i \leq \mu_i \\ 0 & \text{o.w} \end{cases} \quad (5.50)$$

Now substitute $\boldsymbol{\theta}$ in $\Lambda^*(\mathbf{r})$ for a given \hat{f} :

$$\Lambda^*(\mathbf{r}) = \begin{cases} \frac{(\tau_i - \mu_i)^2}{2\sigma^2} & \text{if } \tau_i \leq \mu_i \\ & i = 1, 2 \\ 0 & \text{o.w} \end{cases}$$

Since we are interested in minimizing the probability of a miss, based on equation (5.8), this means we need to compute

$$\frac{1}{2\sigma^2} \inf_{\mathbf{r} \in \mathcal{R}} \|\mathbf{r} - \boldsymbol{\mu}\|^2, \quad \text{where } \mathcal{R} : \{\Re(\mathbf{r}e^{-j\theta_{\mu_K}}) \leq \eta\}$$

which is equal

$$\frac{1}{2\sigma^2} \inf_{\mathbf{r} \in \mathcal{R}} \|\mathbf{r}e^{-j\theta_{\mu_K}} - \|\boldsymbol{\mu}\|\|^2 = \frac{1}{2\sigma^2} \inf_{\mathbf{r} \in \mathcal{R}' = \{\Re(\mathbf{r}) \leq \eta\}} \|\mathbf{r} - \|\boldsymbol{\mu}\|\|^2 \quad (5.51)$$

$$\Lambda^*(\mathbf{r}) = \begin{cases} \frac{(\|\boldsymbol{\mu}\| - \eta)^2}{2\sigma^2} & \text{if } \|\boldsymbol{\mu}\| > \eta \\ 0 & \text{if } \|\boldsymbol{\mu}\| \leq \eta \end{cases}$$

Next we optimize over \hat{f} ,

$$\Lambda^*(\mathbf{r}) = \begin{cases} \frac{1}{2\sigma^2} \max_{\hat{f}: \frac{1}{N} \sum_{i=0}^{N-1} \|\hat{f}\|^2 = 1} (\|\boldsymbol{\mu}\| - \eta)^2 & \|\boldsymbol{\mu}\| > \eta \\ 0 & \|\boldsymbol{\mu}\| \leq \eta \end{cases} \quad (5.52)$$

The Probability of False Alarm (P_F)

Given hypothesis H_0 , the output signal Y as discussed in Section 5.1 and is distributed as

$$Y \sim \mathcal{N}\left(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2}\right)$$

Using 2D representation, covariance matrix $\boldsymbol{\Lambda}_Y$

$$\boldsymbol{\Lambda}_Y = \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \mathbf{I} \quad (5.53)$$

Thus,

$$Y \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_Y).$$

Now,

$$\Phi_{Y_N}(\boldsymbol{\theta}) = e^{\frac{1}{2} \boldsymbol{\theta}^T \boldsymbol{\Lambda}_Y \boldsymbol{\theta}}, \quad (5.54)$$

and

$$\begin{aligned} \Lambda_N(N\boldsymbol{\theta}) &= \frac{N^2}{2} \boldsymbol{\theta}^T \boldsymbol{\Lambda}_Y \boldsymbol{\theta} \\ &= \frac{N^2}{2} \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \boldsymbol{\theta}^T \mathbf{I} \boldsymbol{\theta} \end{aligned} \quad (5.55)$$

Now take

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{2} \lim_{N \rightarrow \infty} \frac{\sigma^2 \sum |\hat{f}[i]|^2 \|\boldsymbol{\theta}\|^2}{N} \quad (5.56)$$

and setting $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}_R[i] \hat{f}_R[i] + \hat{f}_I[i] \hat{f}_I[i]}{N} = 1$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{2} \lim_{N \rightarrow \infty} \sigma^2 (\theta_1^2 + \theta_2^2) \quad (5.57)$$

Therefore, equation (4.17), which represents Fenchel-Legendre transform, can be written as,

$$\Lambda^*(\mathbf{r}) = \sup_{\boldsymbol{\theta} \in \mathbb{R}^2} \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - \sigma^2 \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{2} \right) \quad (5.58)$$

Calculate the derivative of each θ_i , for $i = 1, 2$, and set it to zero,

$$\frac{\partial \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - \sigma^2 \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{2} \right)}{\partial \theta_i} = 0 \quad (5.59)$$

we will get

$$\forall i \quad \theta_i = \left(\frac{\tau_i}{\sigma^2} \right) \quad (5.60)$$

substitute $\boldsymbol{\theta}$ in

$$\Lambda^*(\mathbf{r}) = \sum_{i=1}^2 \left(\frac{\tau_i^2}{2\sigma^2} \right) = \inf_{\mathbf{r} \notin \mathcal{R}} \frac{\|\mathbf{r} e^{-j\theta_{\mu_K}}\|^2}{2\sigma^2} = \frac{1}{2\sigma^2} \inf_{\mathbf{r} \notin \mathcal{R}'} \|\mathbf{r}\|^2$$

since this is independent of \hat{f} , the optimal rate of decay for P_F is

$$\Lambda^*(\mathbf{r}) = \begin{cases} \frac{\eta^2}{2\sigma^2} & \text{if } \eta > 0 \\ 0 & \text{if } \eta \leq 0 \end{cases} \quad (5.61)$$

Numerical Results

Similar procedure was applied as in real case and the results were found to be identical to the real case results. Figure 5.12 is one sample from the produced results of $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for SNR= 4.9136 dB during the search process of finding the quantization points. Therefore, same conclusions are made as in the real case.

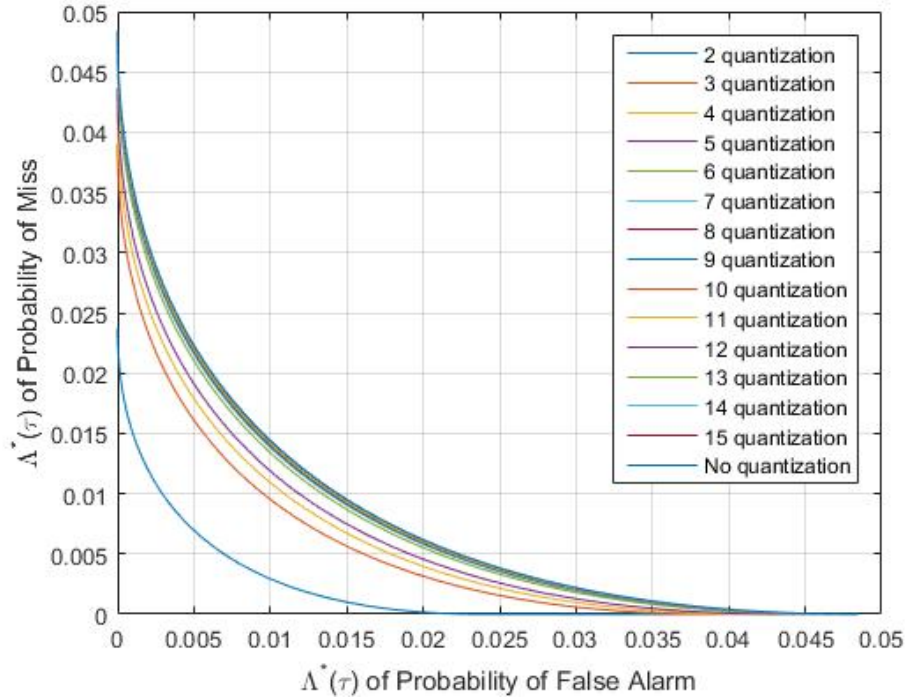


Figure 5.12: $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 4.9136 dB. The performance with no quantization is shown using solid line.

5.6 Unknown Location

In this section, the second scenario stated in section 3.2 will be studied. Thus, the Probability of Detection (P_D) will be,

$$P_D = P(\max = K \quad \& \quad Y[K] \geq r)$$

The resulted sequence after convolution is distributed as a Multivariate Normal Distribution. The amplitude of each point, starting from point $t = N$ and ending at point $t = L$, will be considered. Thus, the total length of the Multivariate Normal Distribution is $T = N - L + 1$. The points in the T -dimensional random vector have the following distribution, refer to figure 5.1,

1) H_1 at each point t :

$$Z[t] = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[t-i]$$

therefore,

$$Y[t] \sim \mathcal{N} \left(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \right).$$

2) H_1 at each point t :

$$R[t] = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]x[t-i] \quad (5.62)$$

and

$$Z[t] = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[t-i]$$

therefore,

$$Y[t] \sim \mathcal{N} \left(\frac{\sum \hat{f}[i]x[t-i]}{N}, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \right).$$

In which we will represent it as

$$Y[t] \sim \mathcal{N} (\mu_t, \sigma_1^2)$$

where μ_t is the the mean at each point t and the variance σ_1^2 is constant at all points. Finally, the multivariate normal distribution of a T -dimensional random vector $\mathbf{Y} = [Y_N, \dots, Y_L]$ can be written using a vector notation:

$$1) H_0 : \frac{1}{N} \mathbf{A} \mathbf{W} = \mathbf{Y}.$$

$$2) H_1 : \frac{1}{N} \mathbf{A} \mathbf{x} + \frac{1}{N} \mathbf{A} \mathbf{W} = \mathbf{Y}$$

$$\mathbf{Y} \sim \mathcal{N} \left(\frac{1}{N} \mathbf{A} \mathbf{x}, \frac{1}{N^2} \mathbf{A} \Lambda_{\mathbf{W}} \mathbf{A}^H \right),$$

denoted by,

$$\mathbf{Y} \sim \mathcal{N} (\boldsymbol{\mu}, \boldsymbol{\Sigma}). \quad (5.63)$$

with T -dimensional mean vector

$$\boldsymbol{\mu} = [\mu_N, \dots, \mu_L]$$

and $T \times T$ covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{N,N+1}^2 & \sigma_{N,N+2} & \dots & \sigma_{N,L} \\ \sigma_{N+1,N} & \sigma_1^2 & \sigma_{N+1,N+2} & \dots & \sigma_{N+1,L} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{L,N} & \sigma_{L,N+1} & \sigma_{L,N+2} & \dots & \sigma_1^2 \end{bmatrix},$$

$$\text{where } \sigma_{t,j} = \frac{\sigma^2 \sum_{i=1}^N \hat{f}[i] \hat{f}[i+(j-t)]}{N^2}.$$

The probability density function of it is represented by

$$p_y(Y_N, \dots, Y_L) = \frac{1}{\sqrt{(2\pi)^T |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})} \quad (5.64)$$

To find the probability that the maximum point is located at $t = K$, first we will consider a Bivariate case, that is $Y = [Y_1 \ Y_2]$. It has a mean vector of

$$\boldsymbol{\mu} = [\mu_1 \ \mu_2]$$

and a 2×2 variance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}.$$

The probability that Y_2 is maximum

$$\begin{aligned} \int_{\mathbf{r}}^{\infty} \int_{-\infty}^{Y_2} p_Y(Y_1, Y_2) dY_1 dY_2 &= \int_{\mathbf{r}}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_2-\mu_2)^2}{2\sigma_1^2}} \int_{-\infty}^{Y_2} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_1-\mu_1)^2}{2\sigma_1^2}} dY_1 dY_2 \\ &= \int_{\mathbf{r}}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_2-\mu_2)^2}{2\sigma_1^2}} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_1-\mu_1)^2}{2\sigma_1^2}} \right. \\ &\quad \left. - \int_{Y_2}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_1-\mu_1)^2}{2\sigma_1^2}} \right) dY_1 dY_2 \\ &= \int_{\mathbf{r}}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_2-\mu_2)^2}{2\sigma_1^2}} \left(1 - \int_{Y_2}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_1-\mu_1)^2}{2\sigma_1^2}} \right) dY_1 dY_2 \\ &= Q\left(\frac{\mathbf{r}-\mu_2}{\sigma_1}\right) - \int_{\mathbf{r}}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(Y_2-\mu_2)^2}{2\sigma_1^2}} Q\left(\frac{Y_2-\mu_2}{\sigma_1}\right) dY_2. \end{aligned} \quad (5.65)$$

The right integral of equation (5.65) will be express through approximation. Figure 5.13 is an example of the pdf of a bivariate case. The red arrows are pointing to the region when Y_2 is maximum. Using equation (5.65), it was found that $P_D = P(\max = K \ \& \ Y[K] \geq \mathbf{r}) = 0.977$.

In case of T -dimensional random vector $\mathbf{Y} = [Y_N, \dots, Y_K, \dots, Y_L]$ equation (5.65) becomes

$$P_D = P(\max = K \ \& \ Y[K] \geq \mathbf{r}) = \int_{-\infty}^{\infty} \int_{\eta_N}^{Y_K} \dots \int_{-\infty}^{Y_K} p_y(Y_N, \dots, Y_L) dY_L \dots dY_L. \quad (5.66)$$

5.6.1 Numerical Results

Figure 5.14 is the produced theoretical result for several quantization points using the analysis in this section for size $T = 3$. Figure 5.15 and figure 5.15 are a

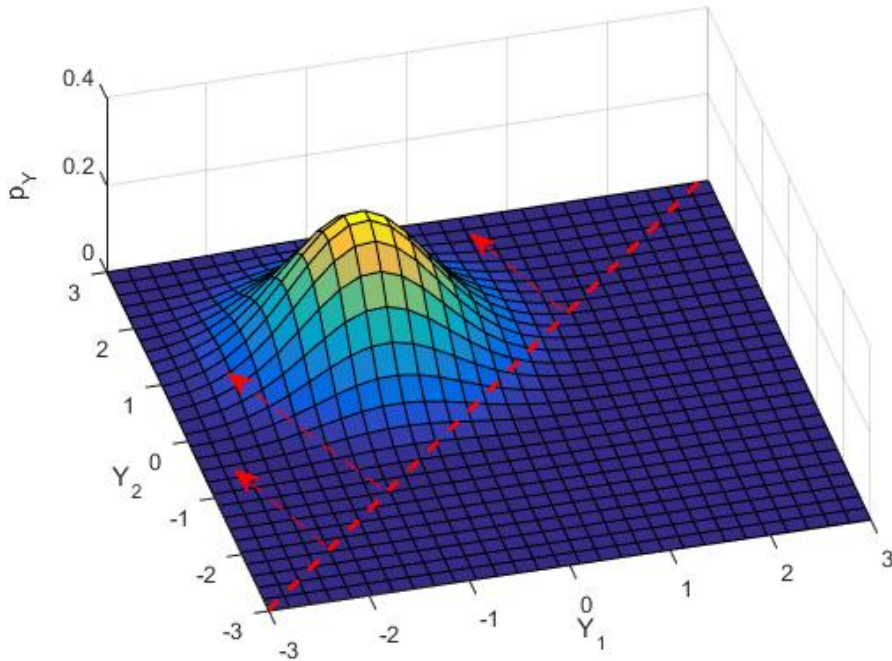


Figure 5.13: p_Y for a bivariate normal distribution with mean of $[-1 \ 1]$ and variance 0.5

comparison between the simulation results and theoretical of $P_D = P(\max_{K \leq Y[K]} \geq \mathbf{r})$ for SNR= 4 for size $T = 3$, which proves equation (5.65). The slight difference in the plots in figure 5.15 is due numerical precision of the tolerance used to find the Cumulative Distribution Function (CDF) using Matlab. Figure 5.17 is a comparison between the simulation results and theoretical of a packet length of 542 and SNR=4 dB. This figure shows that as the packet length increases, the probability of detection decrease. Figure 5.18 is a comparison between the simulation results and theoretical of a packet length of 542 and SNR=31.05076 dB. **Note** : Matlab could not handle more than a multivariate variable of size 25 to find its Cumulative Distribution Function (CDF), since it was the function used for equation (5.65).

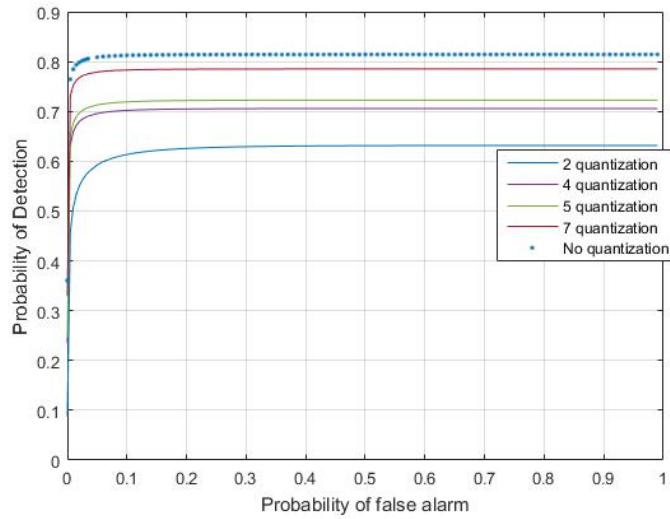


Figure 5.14: PD vs P_F for optimized theoretical quantization points, starting for selected number of quantization points of Barker with Sinc sequence detection, SNR= 4dB. The performance with no quantization is shown using dotted line. Here maximum peak was studied in a packet having a total length $L = 535$ and packet is located at index $I=2$.

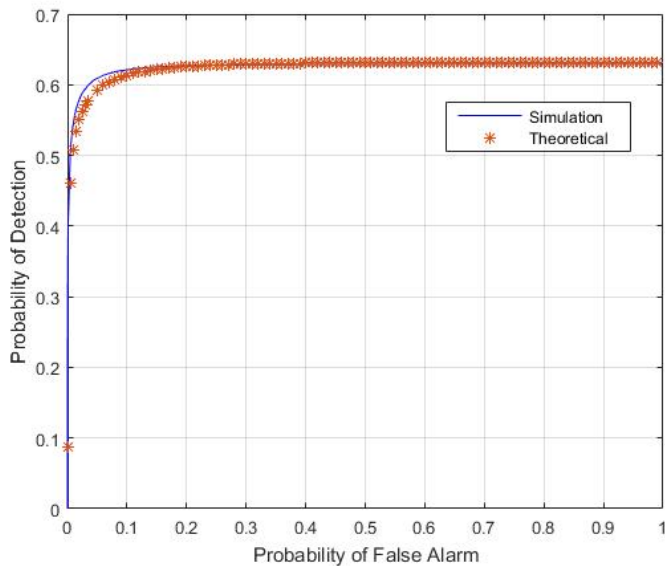


Figure 5.15: PD vs P_F for two quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR=4 dB . Here maximum peak was studied in a packet having a total length $L = 535$ and packet is located at index $I=2$.

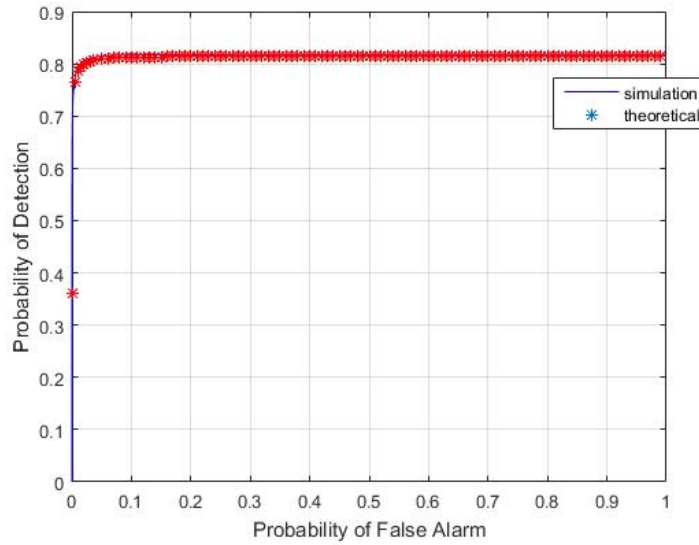


Figure 5.16: PD vs P_F for No quantization, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR=4 dB. Here maximum peak was studied in a packet having a total length $L = 535$ and the packet is located at $I=2$

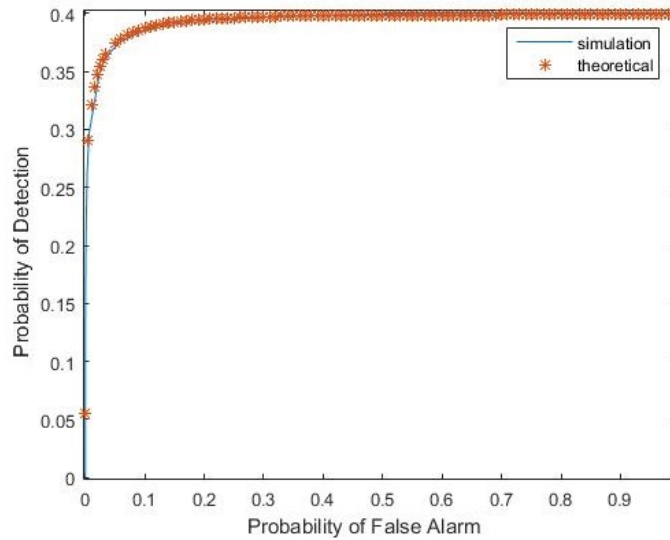


Figure 5.17: PD vs P_F for no quantization, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR=4 dB. Here maximum peak was studied in a packet having a total length $L = 542$ and the packet is located at $I=2$

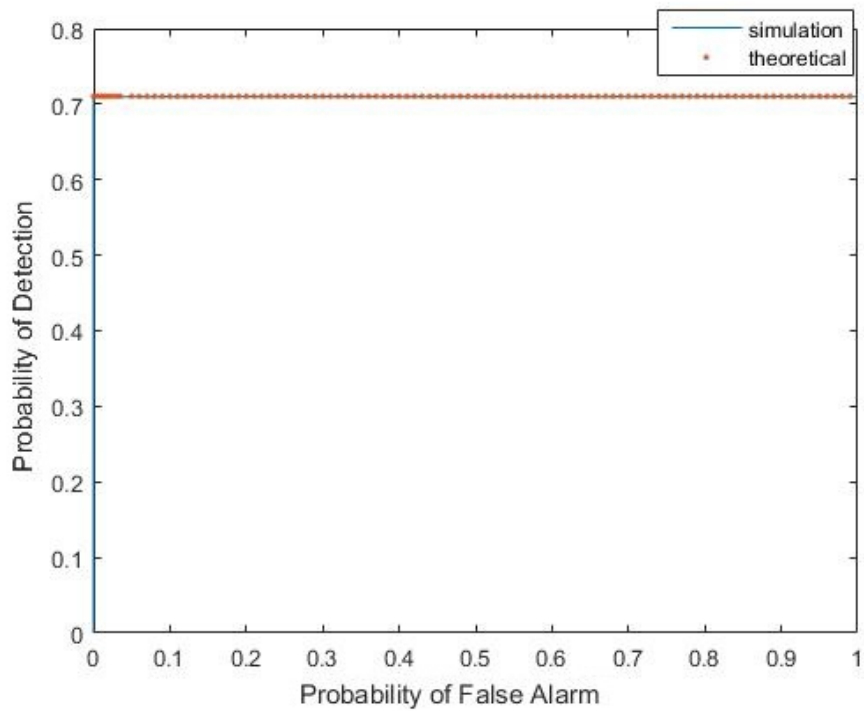


Figure 5.18: PD vs P_F for No quantization, a comparison between the optimized theoretical quantization points results with the simulation results of 50000 packets using the same produced filter \hat{f} of Barker with Sinc, for $SNR=SNR = 17.0713$ dB. Here maximum peak was studied in a packet having a total length $L = 557$ and the packet is located at $I=10$

Chapter 6

Model 2: Signal with Rayleigh Fading

We assume that the received signal was subject to Rayleigh scattering and additive noise. The model description is in section 6.1. Likelihood Ratio Test is derived in section 6.2. Ways to optimize the filter using Large Deviation is studied in section 6.3, and using Decision Theory Tools in section 6.4.

6.1 Model Description

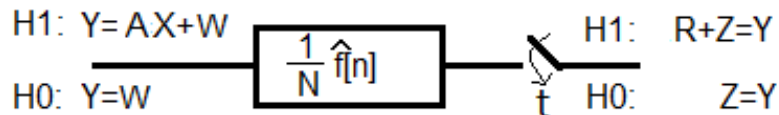


Figure 6.1: Model 2 at the Receiver

Lets consider the above model, $X[k]$ is subject to with $A[k] \sim \mathcal{N}(0, \sigma_R^2)$ and $W[k]$ is Gaussian IID $W[k] \sim \mathcal{N}(0, \sigma^2)$. The received signal $Y[k]$ is convolved by filter \hat{f} , where \hat{f} is the quantized version of the signal f required to be detected. After sampling at $t=K$, this will produce

1) $H_0: Y = Z$ where,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i] W[K - i] \quad (6.1)$$

and $Z \sim \mathcal{N}(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2})$.

2) $H_1 : Y = R + Z$ where,

$$R = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]x[K-i]A[K-i] \quad (6.2)$$

and,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[K-i]$$

where $R \sim \mathcal{N}(0, \frac{\sigma_R^2 \sum |\hat{f}[i]x[K-i]|^2}{N^2})$ and $Z \sim \mathcal{N}(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2})$

For each case, H_1 and H_0 ,

1) For H_1 case:

$$H_1 : Y \sim \mathcal{N} \left(0, \frac{\sigma_R^2 \sum |\hat{f}[i]x[k-i]|^2 + \sigma^2 \sum |\hat{f}[i]|^2}{N^2} \right).$$

which we will represent it as

$$H_1 : Y \sim \mathcal{N}(0, 2\sigma_1^2).$$

and

2) For H_0 case:

$$H_0 : Y \sim \mathcal{N} \left(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \right).$$

which we will represent it as

$$H_0 : Y \sim \mathcal{N}(0, 2\sigma_0^2).$$

Initially, we assume the location K of the output is known, and an error is made if at this location a test fails.

6.2 Likelihood Ratio Test

The LRT device used to detect this signal will be based on the following equation:

$$\frac{1}{\sigma_1^2} e^{\frac{-|Y|^2}{2\sigma_1^2}} \underset{H_0}{\overset{H_1}{\geq}} \frac{1}{\sigma_0^2} e^{\frac{-|Y|^2}{2\sigma_0^2}} \zeta \quad (6.3)$$

After simplifications, equation (26) becomes,

$$\frac{|Y|^2}{2\sigma_0^2} - \frac{|Y|^2}{2\sigma_1^2} + 2 \ln\left(\frac{\sigma_0}{\sigma_1}\right) \begin{array}{l} \geq \\ < \end{array} \begin{array}{l} H_1 \\ H_0 \end{array} \ln \zeta \quad (6.4)$$

which is equal to

$$|Y|^2 \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \begin{array}{l} \geq \\ < \end{array} \begin{array}{l} H_1 \\ H_0 \end{array} 2 \ln \zeta + 4 \ln\left(\frac{\sigma_1}{\sigma_0}\right) = 2 \ln\left(\zeta \frac{\sigma_1^2}{\sigma_0^2}\right) \quad (6.5)$$

For $\zeta \leq \frac{\sigma_1^2}{\sigma_0^2}$, decide always on H_1 . Otherwise, the decision will be base on,

$$|Y| \begin{array}{l} \geq \\ < \end{array} \begin{array}{l} H_1 \\ H_0 \end{array} \sqrt{\frac{2 \ln\left(\zeta \frac{\sigma_1^2}{\sigma_0^2}\right) \sigma_1^2 \sigma_0^2}{\sigma_1^2 - \sigma_0^2}} = \eta \quad (6.6)$$

Therefore,

$$\left. \begin{array}{l} \mathcal{Y}_1 = \{\mathbf{Y}, \text{ such that } |\mathbf{Y}| \geq \eta\} \\ \mathcal{Y}_0 = \{\mathbf{Y}, \text{ such that } |\mathbf{Y}| < \eta\} \end{array} \right\} \begin{array}{l} \text{either } \mathbf{Y} \in \mathbb{R}^2 \\ \text{or } \mathbf{Y} \in \mathbb{C} \end{array}$$

6.3 Minimising Error Rate using Large Deviations

Rayleigh is a random complex number whose real and imaginary components are independently and identically distributed Gaussian. For this reason, in complex signal analysis using the Gärtner-Ellis theorem, it is more practical to deal with it as a 2D variable.

6.3.1 The Probability of a Miss (P_M)

Given hypothesis H_1 , the output signal $Y = R + Z$ as discussed in Section 6.1 and is distributed as

$$Y \sim \mathcal{N}\left(0, \frac{\sigma_R^2 \sum \hat{f}[i] x[k-i]^2 + \sigma^2 \sum |\hat{f}[i]|^2}{N^2}\right)$$

Note that Y is a complex variable, equivalently a 2D vector

$$Y \sim \mathcal{N}(\mathbf{m}_Y, \Lambda_Y).$$

with a mean

$$\mathbf{m}_Y = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} m_R \\ m_I \end{bmatrix} \quad (6.7)$$

and covariance matrix Λ_Y

$$\Lambda_Y = \frac{\sigma_R^2 \sum \hat{f}[i]x[K-i]^2 + \sigma^2 \sum |\hat{f}[i]|^2}{2N^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{\sigma_R^2 \sum \hat{f}[i]x[K-i]^2 + \sigma^2 \sum |\hat{f}[i]|^2}{2N^2} \mathbf{I} \quad (6.8)$$

Let $\boldsymbol{\theta}$ and \mathbf{r} be

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad \mathbf{r} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}.$$

Then, starting from equation (4.19),

$$\Phi_{Y_N}(\boldsymbol{\theta}) = e^{\frac{1}{2}\boldsymbol{\theta}^T \Lambda_Y \boldsymbol{\theta}} \quad (6.9)$$

then

$$\begin{aligned} \Lambda_N(N\boldsymbol{\theta}) &= \frac{N^2}{2} \boldsymbol{\theta}^T \Lambda_Y \boldsymbol{\theta} \\ &= \frac{N^2}{2} \frac{\sigma_R^2 \sum \hat{f}[i]x[K-i]^2 + \sigma^2 \sum |\hat{f}[i]|^2}{2N^2} \boldsymbol{\theta}^T \mathbf{I} \boldsymbol{\theta} \end{aligned} \quad (6.10)$$

Now take

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{4} \lim_{N \rightarrow \infty} \frac{(\sigma_R^2 \sum \hat{f}[i]x[K-i]^2 + \sigma^2 \sum |\hat{f}[i]|^2) \|\boldsymbol{\theta}\|^2}{N} \quad (6.11)$$

If we set $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}_R[i] \hat{f}_R[i] + \hat{f}_I[i] \hat{f}_I[i]}{N} = 1$, i.e \hat{f} is a ‘‘unit power’’

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{4} \left(\lim_{N \rightarrow \infty} \frac{\sigma_R^2 \sum_{i=0}^{N-1} |\hat{f}[i]x[K-i]|^2}{N} + \sigma^2 \right) (\theta_1^2 + \theta_2^2) \quad (6.12)$$

denoting $\mu = \lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} |\hat{f}[i]x[K-i]|^2}{N}$ then equation (6.23) becomes,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{4} (\sigma_R^2 \mu + \sigma^2) (\theta_1^2 + \theta_2^2) \quad (6.13)$$

Therefore, equation (4.17), which represents Fenchel-Legendre transform, can be written as,

$$\Lambda^*(\mathbf{r}) = \sup_{\boldsymbol{\theta} \in \mathbb{R}^2} \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - (\sigma_R^2 \mu + \sigma^2) \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{4} \right) \quad (6.14)$$

Calculate the derivatives of each θ_i , for $i = 1, 2$, and set it to zero,

$$\frac{\partial \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - (\sigma_R^2 \mu + \sigma^2) \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{4} \right)}{\partial \theta_i} = 0 \quad (6.15)$$

we will get

$$\forall i \quad \theta_i = \left(\frac{2\tau_i}{\sigma_R^2 \mu + \sigma^2} \right) \quad (6.16)$$

substitute $\boldsymbol{\theta}$ in

$$\Lambda^*(\mathbf{r}) = \sum_{i=1}^2 \left(\frac{\tau_i^2}{(\sigma_R^2 \mu + \sigma^2)} \right) = \frac{\|\mathbf{r}\|^2}{(\sigma_R^2 \mu + \sigma^2)}$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log P_M(\mathbf{Y}[\mathbf{n}] \in \mathcal{Y}_0) \leq - \inf_{\mathbf{r} \in \mathcal{Y}_0} \frac{\|\mathbf{r}\|^2}{(\sigma_R^2 \mu + \sigma^2)} = 0 \quad (6.17)$$

Which indicates that P_M does not decay exponentially to zero with N .

6.3.2 The Probability of False Alarm (P_F)

Given hypothesis H_0 , the output signal Y as discussed in Section 6.1 and is distributed as

$$Y \sim \mathcal{N} \left(0, \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \right)$$

equivalently,

$$Y \sim \mathcal{N}(\mathbf{m}_Y, \boldsymbol{\Lambda}_Y).$$

with a mean

$$\mathbf{m}_Y = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} m_R \\ m_I \end{bmatrix} \quad (6.18)$$

and covariance matrix $\boldsymbol{\Lambda}_Y$

$$\boldsymbol{\Lambda}_Y = \frac{\sigma^2 \sum |\hat{f}[i]|^2}{2N^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{\sigma^2 \sum |\hat{f}[i]|^2}{2N^2} \mathbf{I}. \quad (6.19)$$

Let $\boldsymbol{\theta}$ and \mathbf{r} be

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad \mathbf{r} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}.$$

Then, starting from equation (4.19),

$$\Phi_{Y_N}(\boldsymbol{\theta}) = e^{\frac{1}{2} \boldsymbol{\theta}^T \boldsymbol{\Lambda}_Y \boldsymbol{\theta}} \quad (6.20)$$

then

$$\begin{aligned} \Lambda_N(N\boldsymbol{\theta}) &= \frac{N^2}{2} \boldsymbol{\theta}^T \boldsymbol{\Lambda}_Y \boldsymbol{\theta} \\ &= \frac{N^2}{4} \frac{\sigma^2 \sum |\hat{f}[i]|^2}{N^2} \boldsymbol{\theta}^T \mathbf{I} \boldsymbol{\theta} \end{aligned} \quad (6.21)$$

Now take

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{4} \lim_{N \rightarrow \infty} \frac{\sigma^2 \sum |\hat{f}[i]|^2 \|\boldsymbol{\theta}\|^2}{N} \quad (6.22)$$

If we set $\lim_{N \rightarrow \infty} \frac{\sum_{i=0}^{N-1} \hat{f}_R[i] \hat{f}_R[i] + \hat{f}_I[i] \hat{f}_I[i]}{N} = 1$, i.e \hat{f} is a “unit power”

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\boldsymbol{\theta}) = \frac{1}{4} \lim_{N \rightarrow \infty} \sigma^2 (\theta_1^2 + \theta_2^2) \quad (6.23)$$

Using equation (4.17),

$$\Lambda^*(\mathbf{r}) = \sup_{\boldsymbol{\theta} \in \mathbb{R}^2} \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - \sigma^2 \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{4} \right) \quad (6.24)$$

Calculate the derivative of each θ_i , for $i = 1, 2$, and set it to zero,

$$\frac{\partial \left(\langle \boldsymbol{\theta}, \mathbf{r} \rangle - \sigma^2 \frac{\boldsymbol{\theta}^T \boldsymbol{\theta}}{4} \right)}{\partial \theta_i} = 0 \quad (6.25)$$

we will get

$$\forall i \quad \theta_i = \left(\frac{2\tau_i}{\sigma^2} \right) \quad (6.26)$$

substitute $\boldsymbol{\theta}$ in

$$\Lambda^*(\mathbf{r}) = \sum_{i=1}^2 \left(\frac{\tau_i^2}{\sigma^2} \right) = \frac{\|\mathbf{r}\|^2}{\sigma^2}$$

Finally

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log P_F(\mathbf{Y}[\mathbf{n}] \in \mathcal{Y}_1) \geq - \inf_{\mathbf{r} \in \mathcal{Y}_1} \frac{\|\mathbf{r}\|^2}{\sigma^2} = \frac{\eta^2}{\sigma^2} \quad (6.27)$$

6.4 Minimising Error Rate through Decision Theory

LRT device for Neyman-Pearson was found to be in equation (6.6)

$$\begin{array}{l} H_1 \\ |Y| \geq \sqrt{\frac{2 \ln(\zeta \frac{\sigma_1^2}{\sigma_0^2}) \sigma_1^2 \sigma_0^2}{\sigma_1^2 - \sigma_0^2}} \\ < \\ H_0 \end{array}$$

If ζ is set to 1, this becomes a special case which is the ML Device. That is, the ML device used to detect this signal will be based on the following equation:

$$\begin{array}{l} H_1 \\ \frac{1}{\sigma_1^2} e^{-\frac{|Y|^2}{2\sigma_1^2}} \geq \frac{1}{\sigma_0^2} e^{-\frac{|Y|^2}{2\sigma_0^2}} \\ < \\ H_0 \end{array} \quad (6.28)$$

and the final decision will be based on

$$|Y| \begin{matrix} \geq \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \sqrt{\frac{4 \ln(\frac{\sigma_1}{\sigma_0}) \sigma_1^2 \sigma_0^2}{\sigma_1^2 - \sigma_0^2}}. \quad (6.29)$$

Based on the choice of ζ the following will be applied to minimize P_M and P_F :

6.4.1 The Probability of a miss (P_M)

$$P(\text{Error}) : P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_0^\eta \frac{Y}{\sigma_1^2} e^{-\frac{Y^2}{2\sigma_1^2}} dY \quad (6.30)$$

Let $U = \frac{Y}{\sigma_1}$ and $dU = \frac{dY}{\sigma_1}$, so equation (6.30) will change to

$$P(\text{Error}) : P_M = 1 - P_D = P(\hat{H} = H_0 | H_1) = \int_0^{\frac{\eta}{\sigma_1}} U e^{-\frac{U^2}{2}} dU = 1 - e^{-\frac{\eta^2}{2\sigma_1^2}} \quad (6.31)$$

Thus we should minimize $\frac{\eta}{\sqrt{\sigma_1}}$ to minimize (P_M).

6.4.2 The Probability of False Alarm (P_F)

$$P_F = P(\hat{H} = H_1 | H_0) = \int_\eta^\infty \frac{Y}{\sigma_0^2} e^{-\frac{Y^2}{2\sigma_0^2}} dY \quad (6.32)$$

Let $U = \frac{Y}{\sigma_0}$ and $dU = \frac{dY}{\sigma_0}$, so equation (6.32) will change to

$$P_F = P(\hat{H} = H_1 | H_0) = \int_{\frac{\eta}{\sigma_0}}^\infty U e^{-\frac{U^2}{2}} dU = e^{-\frac{\eta^2}{2\sigma_0^2}} \quad (6.33)$$

Thus we should maximise $\frac{\eta}{\sigma_0}$ to minimize the probability of false Alarm. As concluded from before we should minimize $\frac{\eta}{\sigma_1}$ and maximise $\frac{\eta}{\sigma_0}$. This can be proved as follows:

The relation between P_D and P_F was found to be

$$\left. \begin{matrix} P_D = e^{-\frac{\eta^2}{2\sigma_1^2}} \\ P_F = e^{-\frac{\eta^2}{2\sigma_0^2}} \end{matrix} \right\} P_D = P_F^{\frac{\sigma_0^2}{\sigma_1^2}} \quad (6.34)$$

Therefore maximize P_D , $\frac{\sigma_0^2}{\sigma_1^2}$ should be minimized, since $0 < \frac{\sigma_0^2}{\sigma_1^2} < 1$, so

$$\begin{aligned} \min \frac{\sigma_0^2}{\sigma_1^2} &= \arg \max \frac{\sigma_1^2}{\sigma_0^2} = \arg \max_f \frac{\sigma_R^2 |\sum f[i]x[k-i]|^2 + \sigma^2 \sum |\hat{f}[i]|^2}{\sigma^2 \sum |\hat{f}[i]|^2} = \\ & \arg \max_f \frac{\sigma_R^2 |\sum f[i]x[k-i]|^2}{\sigma^2 \sum |\hat{f}[i]|^2} + 1 = \arg \max_f \frac{\sigma_R^2 |\sum f[i]x[k-i]|^2}{\sigma^2 \sum |\hat{f}[i]|^2} \end{aligned} \quad (6.35)$$

Figure 6.2 shows the result from minimizing the ratio $\frac{\sigma_0^2}{\sigma_1^2}$. It is clear that the slope increased when $\frac{\sigma_0^2}{\sigma_1^2}$ was minimized.

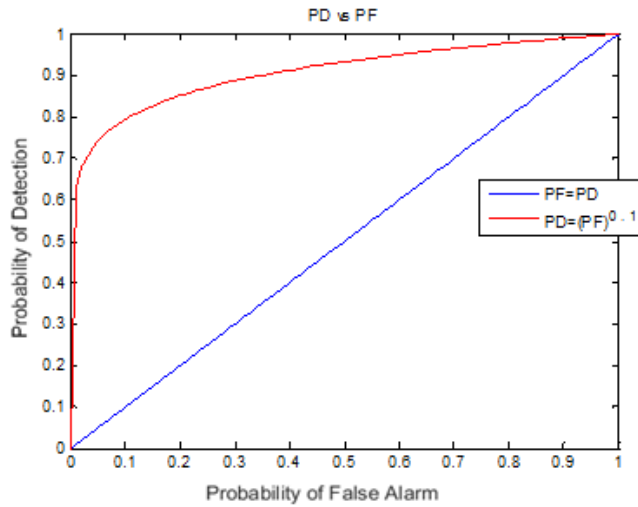


Figure 6.2: P_D vs P_F for $\frac{\sigma_0^2}{\sigma_1^2} = 1$ and $\frac{\sigma_0^2}{\sigma_1^2} = 0.1$

6.4.3 Numerical Results

For the real case signal that was subject to Rayleigh scattering and additive noise, best \hat{f} was searched in Matlab using `fmincon`. For each number of quantization points, P_F was set to be at most to α , and the value of α ranged from 0.001 to 0.99. Similar to Chapter 5 results, as the value of α varied, the produced quantization points were equal, see table 6.1. In addition, the value of the noise variance was varied to produce several SNRs. The difference between \hat{f} was found to be also small and less than 0.01 difference, refer to table 6.1 and table 6.2 to compare the quantization points for SNR=4 dB and SNR=-2.9287dB. Based on the above, same conclusion is made as in Chapter 5 and one quantized filter \hat{f} maybe used for all operating points and SNRs. Figure 6.3 is one sample from the produced results of P_D vs $P_F = \alpha$ for SNR= 4 dB during the search process of finding the quantization points. Figure 6.4 shows a similar results for SNR= 9.2898 dB. From both figures, we conclude starting from three quantization points the

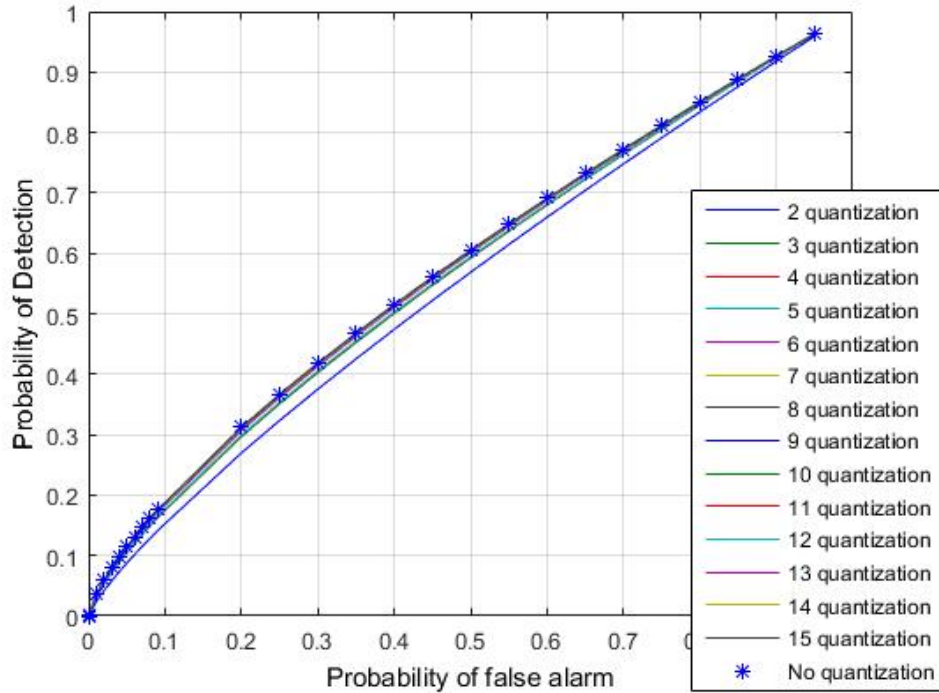


Figure 6.3: P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for SNR= 4 dB. The performance with no quantization is shown using * line.

performance of the filter is so close to the performance to the original one. Figure 6.5 is a comparison between the theoretical plot given two quantization points and the simulation of 100000 packets using the same filter \hat{f} produced from the `fmincon` search process for SNR=4dB, which validates our theoretical analysis. Also a signal with complex values was tested. Similar procedure was applied as in real case and the results were found to be identical to the real case results, refer to table 6.3 and table 6.4. Thus same conclusion is made and one quantized filter \hat{f} maybe used for all operating points and SNRs. Figure 6.6 is a sample case from the produced results of P_D vs $P_F = \alpha$ for SNR=0.1424 dB obtained by the search process of finding the quantization points. Figure 6.7 shows a similar results for SNR= 4.9136 dB. From both figures, we conclude starting from three quantization points the performance of the filter is so close in performance to the original one. Figure 6.8 is a comparison between the theoretical plot given two quantization points and the simulation of 100000 packets using the same filter \hat{f} produced from the `fmincon` search process for SNR=0.1424dB, which validates our theoretical analysis.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.0399263801533052	-0.9982560548344744
0.05	0.0399263801533052	-0.9982560548344744
0.15	0.0399263801533052	-0.9982560548344744
0.2	0.0399263801533052	-0.9982560548344744
0.4	0.0399263801533052	-0.9982560548344744
0.5	0.0399263801533052	-0.9982560548344744
0.6	0.0399263801533052	-0.9982560548344744
0.7	0.0399263801533052	-0.9982560548344744
0.8	0.0399263801533052	-0.9982560548344744
0.9	0.0399263801533052	-0.9982560548344744
0.99	0.0399263801533052	-0.9982560548344744

Table 6.1: The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=4 dB.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.0399926053219320	-0.999991460896565
0.05	0.0399926053219320	-0.999991460896565
0.15	0.0399926053219320	-0.999991460896565
0.2	0.0399926053219320	-0.999991460896565
0.4	0.0399926053219320	-0.999991460896565
0.5	0.0399926053219320	-0.999991460896565
0.6	0.0399926053219320	-0.999991460896565
0.7	0.0399926053219320	-0.999991460896565
0.8	0.0399926053219320	-0.999991460896565
0.9	0.0399926053219320	-0.999991460896565
0.99	0.0399926053219320	-0.999991460896565

Table 6.2: The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for SNR=-2.9287 dB.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	1.00000000+ 0.258163735993003i	-0.75093096409019 + 0.0002636497994105i
0.05	1.00000000 + 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.15	1.00000000+ 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.2	1.00000000+ 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.4	1.00000000+ 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.5	1.00000000+ 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.6	1.00000000 + 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.7	1.00000000 + 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.8	1.00000000 + 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.9	1.00000000+ 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i
0.99	1.00000000 + 0.2581637359930i	-0.75093096409019 + 0.00026364979941051i

Table 6.3: The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu for SNR=4.9136 dB.

PF= α	First Quantization Point	Second Quantization Point
$1.0 \times e^{-05}$	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.05	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.15	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.2	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.4	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.5	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.6	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.7	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.8	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.9	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i
0.99	0.9998796634043 + 0.2673225535602i	-0.7525033118705 - 0.0061791377514439i

Table 6.4: The produced two quantization points to create the optimal filter \hat{f} of Zadoff Chu for SNR=0.1424 dB.

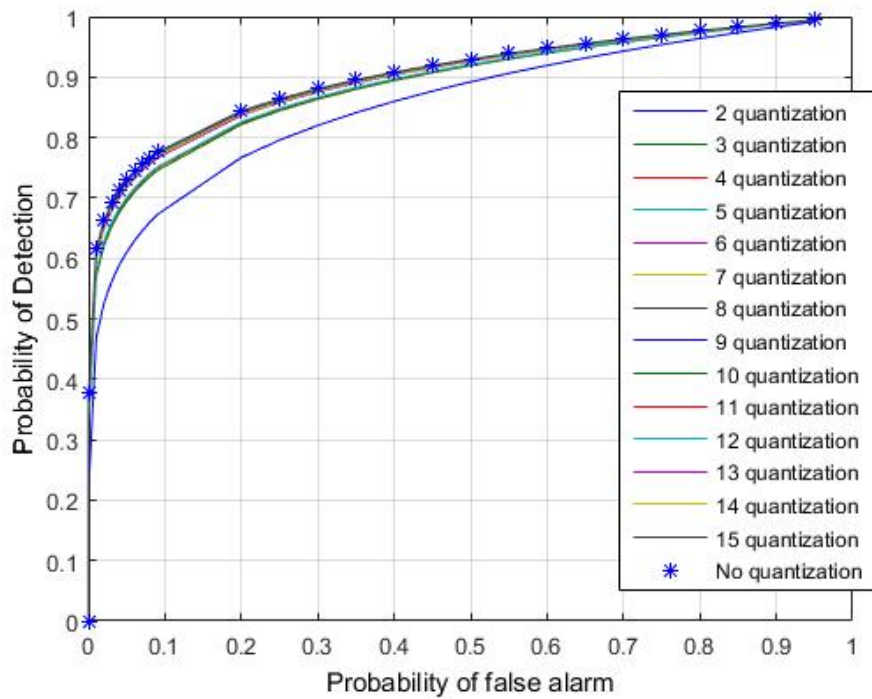


Figure 6.4: PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, SNR= 9.2898dB. The performance with no quantization is shown using * line.

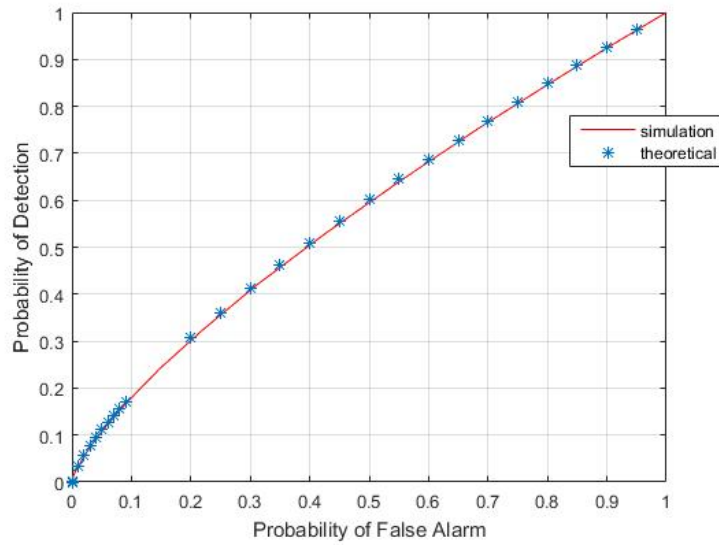


Figure 6.5: PD vs P_F for two quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Barker with Sinc, for SNR= 4 dB.

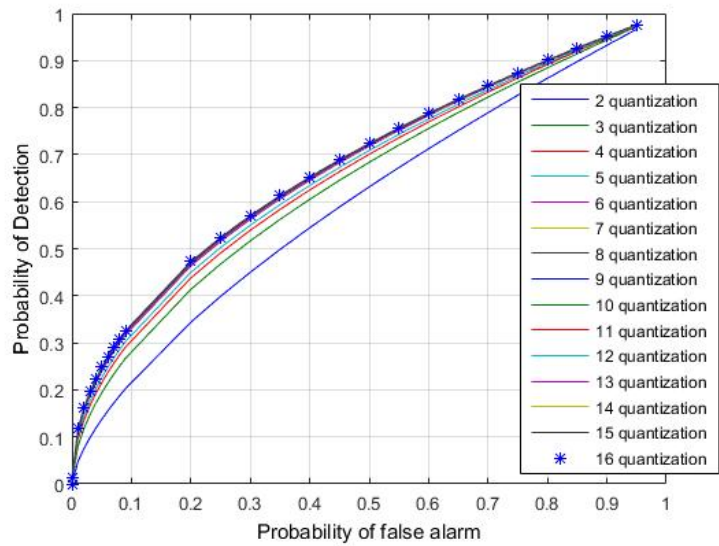


Figure 6.6: PD vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zadoff Chu sequence detection, SNR= 0.1424dB. The performance with no quantization is shown using solid line of 16 quantization points.

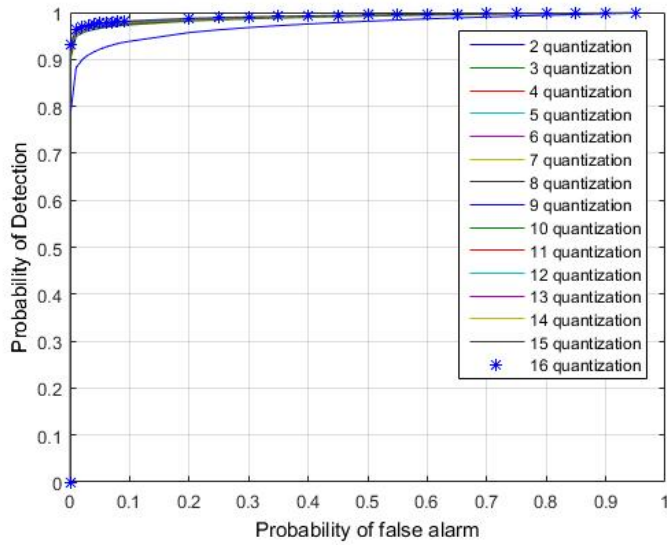


Figure 6.7: P_D vs P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Zadoff Chu sequence detection, SNR= 4.9136dB. The performance with no quantization is shown using solid line of 16 quantization points.

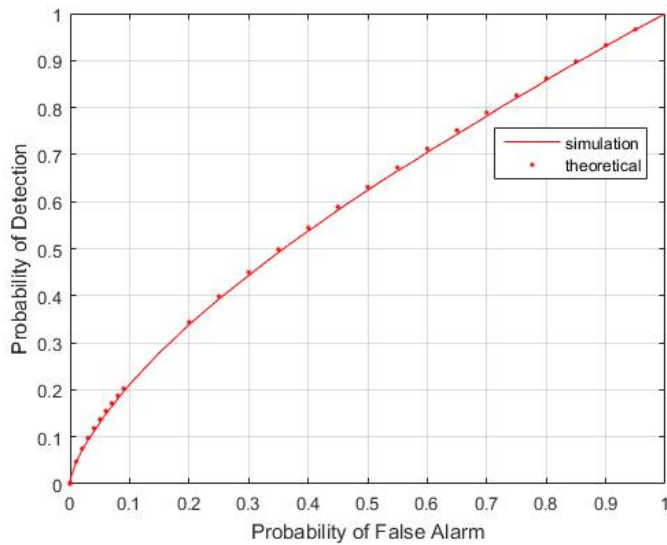


Figure 6.8: P_D vs P_F for two quantization points, a comparison between the optimized theoretical quantization points results with the simulation results of 100000 packets using the same produced filter \hat{f} of Zadoff Chu sequence, for SNR=0.1424 dB.

Chapter 7

Model 3: Signal Subject to Noise having a Laplace Distribution

In this section we consider the signal being received is exposed to noise having a Laplace Distribution. We optimize the filter through Large Deviations.

7.1 Model Description

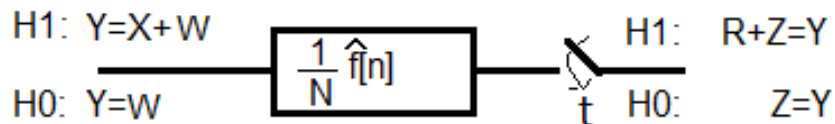


Figure 7.1: Model 3 at the Receiver

Signal $X[k]$ is subject to a Laplacian Noise $W[k]$ with $W[k] \sim Laplace(0, b)$. After filtering and sampling at $t=K$, it was found

1) H_0 : $Y = Z$ where,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i] W[K - i] \quad (7.1)$$

with mean 0 and variance $\frac{2b^2 \sum |\hat{f}[i]|^2}{N^2}$

2) $H_1 : Y = R + Z$ where,

$$R = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]x[K-i] \quad (7.2)$$

and,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[K-i]$$

with mean $\frac{\sum \hat{f}[i]x[K-i]}{N}$ and variance $\frac{2b^2 \sum |\hat{f}[i]|^2}{N^2}$

Initially, we assume the location K of the output is known, and an error is made if at this location a test fails.

7.2 Likelihood Ratio Test

The distribution of Y under H_1 or H_0 is difficult to work with [12]. We instead work with the characteristic function below.

7.3 Minimising Error Rate using Large Deviations

As stated in section 3.2 the amplitude at location K of the signal will first be checked. If it is less than a threshold \mathbf{r} it will be an error. Using this concept, the technique of Gärtner-Ellis theorem introduced in section 4.3 will be applied on cases such that we minimize both the probability of a miss (P_M) and probability of false alarm (P_F). Only real signals will be studied in this chapter.

7.3.1 For Real Signal Cases

The Probability of a Miss (P_M)

Since finding the distribution of \mathbf{Y} is complex, it is wise to work with it through its moment generating function. Then, starting from equation (4.19),

$$\Phi_{Y_N}(\theta) = \mathbb{E}[e^{\theta Y}] = \prod_{i=1}^N \frac{e^{\frac{\hat{f}[i]x[K-i]\theta}{N}}}{1 - \frac{b^2|\hat{f}[i]|^2}{N^2}\theta^2} = \frac{e^{\sum_{i=1}^N \frac{\hat{f}[i]x[K-i]\theta}{N}}}{\prod_{i=1}^N \left(1 - \frac{b^2|\hat{f}[i]|^2}{N^2}\theta^2\right)}, \quad \text{where } |\theta| < \frac{N}{b \max |\hat{f}[i]|} \quad (7.3)$$

then

$$\begin{aligned} \Lambda_N(N\theta) &= \sum_{i=1}^N \left(\hat{f}[i]x[K-i]\theta \right) - \ln \prod_{i=1}^N \left(1 - b^2|\hat{f}[i]|^2\theta^2 \right) \\ &= \sum_{i=1}^N \left(\hat{f}[i]x[K-i]\theta \right) - \sum_{i=1}^N \ln \left(1 - b^2|\hat{f}[i]|^2\theta^2 \right) \end{aligned} \quad (7.4)$$

Now

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) &= \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N (|\hat{f}[i]x[K-i]|)}{N} \theta - \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \ln(1 - b^2 |\hat{f}[i]|^2 \theta^2)}{N} \\ &= \theta \mu - \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln \left(1 - b^2 |\hat{f}[i]|^2 \theta^2 \right),\end{aligned}\tag{7.5}$$

where $\mu = \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N (|\hat{f}[i]x[K-i]|)}{N}$.

The Fenchel-Legendre transform in equation (4.17), can be written as,

$$\begin{aligned}\Lambda^*(\tau) &= \sup_{\frac{-1}{b \max |\hat{f}[i]]} < \theta < 0} \left(\theta \tau - \theta \mu + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln \left(1 - b^2 |\hat{f}[i]|^2 \theta^2 \right) \right) \\ &= \sup_{\frac{-1}{\max |\hat{f}[i]]} < \theta < 0} \left(\theta \frac{\tau}{b} - \theta \frac{\mu}{b} + g(\theta) \right),\end{aligned}\tag{7.6}$$

where $g(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln \left(1 - |\hat{f}[i]|^2 x^2 \right)$, defined on $|x| < \frac{1}{\max |\hat{f}[i]|}$

The Probability of False Alarm (P_F)

Starting from equation (4.19),

$$\Phi_{Y_N}(\theta) = \mathbb{E}[e^{\theta Y}] = \prod_{i=1}^N \frac{1}{1 - \frac{b^2 |\hat{f}[i]|^2}{N^2} \theta^2} = \frac{1}{\prod_{i=1}^N \left(1 - \frac{b^2 |\hat{f}[i]|^2}{N^2} \theta^2 \right)}, \quad \text{where } |\theta| < \frac{N}{b \max |\hat{f}[i]|}\tag{7.7}$$

then

$$\begin{aligned}\Lambda_N(N\theta) &= -\ln \prod_{i=1}^N \left(1 - b^2 |\hat{f}[i]|^2 \theta^2 \right), \quad |\theta| < \frac{N}{b \max |\hat{f}[i]|} \\ &= -\sum_{i=1}^N \ln \left(1 - b^2 |\hat{f}[i]|^2 \theta^2 \right)\end{aligned}\tag{7.8}$$

Now

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = -\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \ln \left(1 - b^2 |\hat{f}[i]|^2 \theta^2 \right)}{N}\tag{7.9}$$

Therefore, equation 4.17, which represents Fenchel-Legendre transform, can be written as,

$$\Lambda^*(\tau) = \sup_{0 < \theta < \frac{1}{b \max |\hat{f}[i]}} \left(\theta \tau + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln \left(1 - b^2 |\hat{f}[i]|^2 \theta^2 \right) \right)\tag{7.10}$$

Numerical Results

We compute for every \hat{f} the function $g(x)$ (where N is large enough) by evaluation of a fine grid between $\frac{1}{\max \hat{f}}$ and 0 and use linear approximation in between. For each number of quantization points, ranges of τ was studied, where τ ranged from 0 to μ . While changing the value of τ , it was found that the quantization points were so close in value, see table 7.1 and difference was less than 0.01 which may be considered small, and maybe due numerical precision. We conjecture that optimal \hat{f} is not dependant on τ (for a given b). In addition, several values of the b was considered. The difference between \hat{f} , for different values of b , was found to be also small and less than 0.01, refer to table 7.1 and table 7.2 to compare the quantization points for $b = 2$ and $b = 10$. Based on the above, for a given number of quantization points, we propose using one quantized filter \hat{f} for all operating points and SNRs. $\Lambda^*(\tau)$ for P_M and P_F were found and plotted. Figure 7.2 is one sample from the produced results of $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for $b = 2$ during the search process of finding the quantization points. It is clear that there is better performance when the number of quantization points increases, which is logical. Figure 7.3 is another result when b set to 10. Comparing both figures, figure 7.2 resulted in a better performance which is reasonable. We conclude from both figures, starting from three quantization points the performance of the filter “approximately” is so close to the performance to the original one

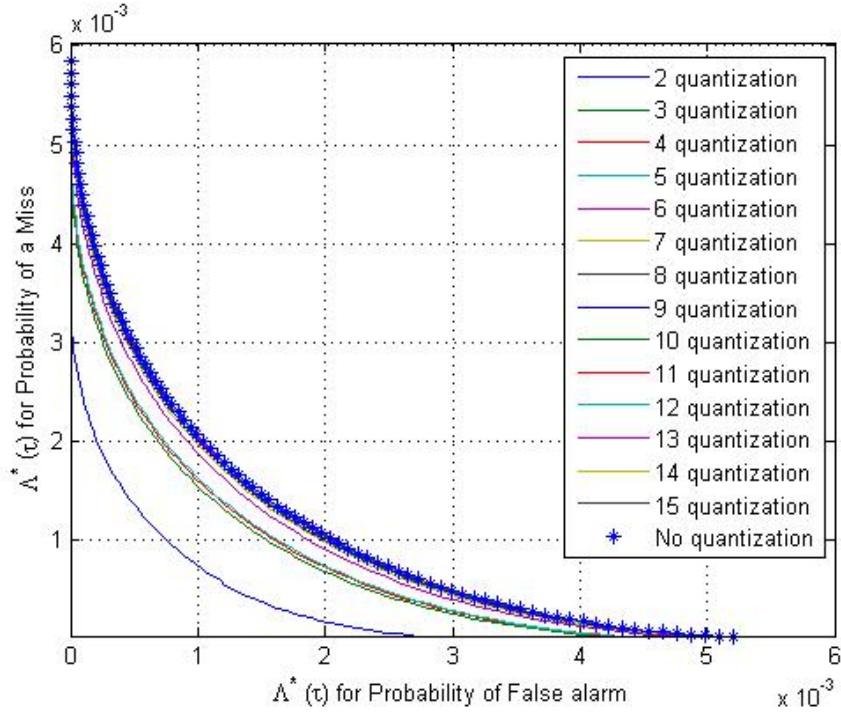


Figure 7.2: $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for $b = 2$. The performance with no quantization is shown using '*' line.

τ	First Quantization Point	Second Quantization Point
0.001	0.0415646212712	-0.999682864612971
0.003	0.041964242618645	-1
0.005	0.04211558721833	-1
0.025	0.04586362683160	-1
0.04	0.04586362683160	-1
0.05	0.04586362683160	-1
0.06	0.04586362683160	-1
0.07	0.04586362683160	-1
0.073	0.04586362683160	-1

Table 7.1: The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for $b = 2$.

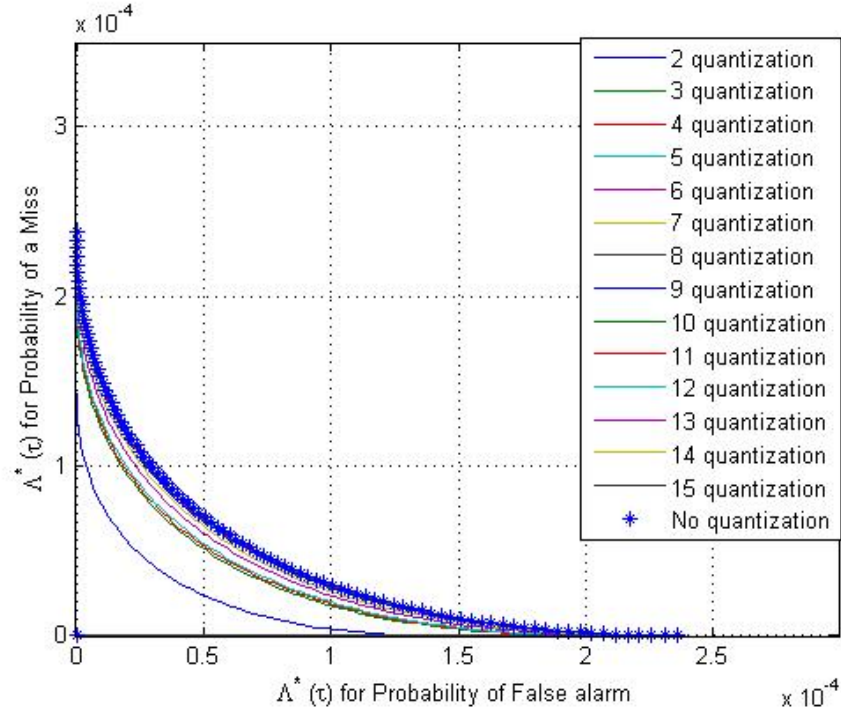


Figure 7.3: $\Lambda^*(\tau)$ for P_M vs $\Lambda^*(\tau)$ for P_F for optimized theoretical quantization points, starting from 2 quantization points to 15 quantization points of Barker with Sinc sequence detection, for $b = 10$. The performance with no quantization is shown using ‘*’ line.

τ	First Quantization Point	Second Quantization Point
0.001	0.04093252341851	-0.999682864612971
0.003	0.04093238551205	-1
0.005	0.04093234105777	-1
0.025	0.04095146135136	-1
0.04	0.04095146135136	-1
0.05	0.04095146135136	-1
0.06	0.04095146135136	-1
0.07	0.04095146135136	-1
0.073	0.04095146135136	-1

Table 7.2: The produced two quantization points to create the optimal filter \hat{f} of Barker with Sinc for $b = 10$.

Chapter 8

Model 4: Signal Subject to Noise having an Exponential Distribution

In this section we consider the signal being received to be exposed to noise having an Exponential Distribution. We take the Large Deviations in what follows.

8.1 Model Description

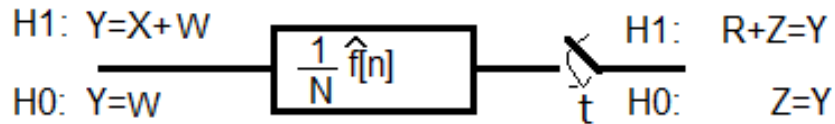


Figure 8.1: Model 4 at the Receiver

Signal $X[k]$ is subject to a Noise $W[k]$ with $W[k] \sim \text{EXP}(\lambda)$. After filtering and sampling at $t=K$,

1) $H_0: Y = Z$ where,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[K - i] \tag{8.1}$$

with mean $\frac{\sum \hat{f}[i]}{N\lambda}$ and variance $\frac{\sum \hat{f}[i]^2}{\lambda^2 N^2}$

2) $H_1 : Y = R + Z$ where,

$$R = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]x[K-i] \quad (8.2)$$

and,

$$Z = \frac{1}{N} \sum_{i=-\infty}^{\infty} \hat{f}[i]W[K-i]$$

with mean $\frac{\sum \hat{f}[i]x[K-i]}{N} + \frac{\sum \hat{f}[i]}{\lambda N}$ and variance $\frac{\sum |\hat{f}[i]|^2}{\lambda^2 N^2}$

Initially, we assume the location K of the output is known, and an error is made if at this location a test fails.

8.2 Likelihood Ratio Test

The distribution of Y under H_1 or H_0 is difficult to work with [12]. We instead work with the characteristic function below.

8.3 Minimising Error Rate using Large Deviations

As stated in section 3.2 the amplitude at location K of the signal will first be checked. If it is less than a threshold \mathbf{r} it will be an error. The technique of Gärtner-Ellis theorem introduced in section 4.3 will be applied on cases such that we minimize both the probability of a miss (P_M) and probability of false alarm (P_F). Only real signals will be studied in this chapter.

8.3.1 For Real Signal Cases

The Probability of a Miss (P_M)

Starting from equation (4.19),

$$\Phi_{Y_N}(\theta) = \mathbb{E}[e^{\theta Y}] = e^{\theta R} \mathbb{E}[e^{\theta Z}] = e^{\theta \frac{1}{N} \sum_{i=1}^N \hat{f}[i]x[K-i]} \prod_{i=1}^N \frac{\lambda}{\lambda - \frac{\hat{f}[i]}{N}\theta}, \quad \text{where } \begin{cases} \theta < \frac{N\lambda}{\max \hat{f}[i]}, & \text{if } \max \hat{f}[i] > 0 \\ \theta > \frac{N\lambda}{\min \hat{f}[i]}, & \text{if } \min \hat{f}[i] < 0 \end{cases}$$

then

$$\begin{aligned} \Lambda_N(N\theta) &= \sum_{i=1}^N \left(\hat{f}[i]x[K-i]\theta \right) + \ln \prod_{i=1}^N \left(\frac{\lambda}{\lambda - \hat{f}[i]\theta} \right) \\ &= \sum_{i=1}^N \left(\hat{f}[i]x[K-i]\theta \right) + N \ln \lambda - \sum_{i=1}^N \ln \left(\lambda - \hat{f}[i]\theta \right). \end{aligned}$$

Now

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) &= \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N (\hat{f}[i]x[K-i])}{N} \theta + \ln \lambda - \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \ln(\lambda - \hat{f}[i]\theta)}{N} \\ &= \theta\mu + \ln \lambda - \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln(\lambda - \hat{f}[i]\theta).\end{aligned}$$

where $\mu = \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N (|\hat{f}[i]x[K-i]|)}{N}$

Therefore, equation 4.17, which represents Fenchel-Legendre transform, can be written as,

$$\begin{aligned}\Lambda^*(\tau) &= \sup_{\frac{-\lambda}{\min \hat{f}[i]} < \theta < 0} \left(\theta\tau - \theta\mu - \ln \lambda + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln(\lambda - \hat{f}[i]\theta) \right) \\ &= \sup_{\frac{-\lambda}{\min \hat{f}[i]} < \theta < 0} (\theta\tau - \theta\mu - \ln \lambda + g(\theta)),\end{aligned}$$

where $g(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln(\lambda - \hat{f}[i]x)$, defined on $\frac{\lambda}{\min \hat{f}[i]} < x < \frac{\lambda}{\max \hat{f}[i]}$

The Probability of False Alarm (P_F)

Starting from equation (4.19),

$$\Phi_{Y_N}(\theta) = \mathbb{E}[e^{\theta Y}] = \prod_{i=1}^N \frac{\lambda}{\lambda - \frac{\hat{f}[i]}{N}\theta}, \quad \text{where} \quad \begin{cases} \theta < \frac{N\lambda}{\max \hat{f}[i]}, & \text{if } \max \hat{f}[i] > 0 \\ \theta > \frac{N\lambda}{\min \hat{f}[i]}, & \text{if } \min \hat{f}[i] < 0 \end{cases}$$

then

$$\begin{aligned}\Lambda_N(N\theta) &= \ln \prod_{i=1}^N \left(\frac{\lambda}{\lambda - \frac{\hat{f}[i]}{N}\theta} \right) \\ &= N \ln \lambda - \sum_{i=1}^N \ln(\lambda - \hat{f}[i]\theta).\end{aligned}$$

Now

$$\lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(N\theta) = \ln \lambda - \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N \ln(\lambda - \hat{f}[i]\theta)}{N}$$

Therefore, equation (4.17), which represents Fenchel-Legendre transform, can be written as,

$$\begin{aligned}\Lambda^*(\tau) &= \sup_{0 < \theta < \frac{\lambda}{\max \hat{f}[i]}} \left(\theta\tau - \ln \lambda + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln(\lambda - \hat{f}[i]\theta) \right) \\ &= \sup_{0 < \theta < \frac{\lambda}{\max \hat{f}[i]}} (\theta\tau - \ln \lambda + g(\theta)),\end{aligned}$$

Chapter 9

Conclusion

In this work, we reduce the computational complexity of the correlation/convolution operation through quantizing the amplitude of one of the functions; this could be the impulse response of a receive filter. The number of multiplications has been hence noticeably decreased. With a view toward detection application, the selection of the quantized points is done through applying decision theory techniques and corresponding quality measures. Different models of stochastic noises are studied in different chapters, and theoretical analysis is validated through Matlab simulations. The performance of the proposed schemes is measured through plotting the Operating Characteristic curve, as well as through the rates of exponential decay using large deviation theory. Examining the performance in various scenarios of quantized versions of the filter, it is clear from our analysis that the total number of required quantization points, in order for the quantized filter to have close performance to the original one, is less than \log_2 of the number of the total points forming the sequence. As a conclusion, we are able to create a filter that is “close” in performance to the original one but with reduced complexity in the number of multiplications during correlation/convolution process.

Appendix A

Different Stochastic Distributions

This section includes the distributions of different stochastic noise used in this thesis.

A.1 Gaussian Distribution

The probability density function of the gaussian distribution:

$$p(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (\text{A.1})$$

where μ is the mean and σ^2 is the variance.

Its moment Generating function:

$$M(t) = e^{(\mu t + \frac{1}{2}\sigma^2 t^2)}. \quad (\text{A.2})$$

A.2 Complex Gaussian Distribution

The probability density function of a i.i.d complex gaussian distribution

$$p(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{|x-\mu|^2}{2\sigma^2}} \quad (\text{A.3})$$

where μ is the mean and σ^2 is the variance.

Its moment Generating function:

$$M(t) = e^{\frac{1}{2}\sigma^2 |t|^2}. \quad (\text{A.4})$$

A.3 Laplace Distribution

The probability density function of the Laplace distribution:

$$\frac{1}{2b} e^{-\frac{|x-\mu|}{b}} \quad (\text{A.5})$$

where μ is the mean and $2b^2$ is the variance.
Its moment Generating function:

$$\frac{e^{\mu t}}{1 - b^2 t^2} \tag{A.6}$$

A.4 Exponential Distribution

The probability density function of the exponential distribution:

$$\lambda e^{-\lambda x}, \tag{A.7}$$

where the mean is λ^{-1} and variance λ^{-2} Its moment Generating function:

$$\frac{\lambda}{\lambda - t}, \text{ for } t < \lambda \tag{A.8}$$

Appendix B

Abbreviations

AWGN	Additive White Gaussian Noise
DFT	Discrete Fourier Transform
CDF	Cumulative Distribution Function
FFT	Fast Fourier Transform
FWT	Fast Wavelet Transform
LRT	Likelihood Ratio Test
MGF	Moment Generating Function
OP	Operating Characteristic
P_D	Probability of Detection
P_F	Probability of False Alarm
P_M	Probability of a Miss
QP	Quadratic Programming
SNR	Signal-to-Noise Ratio
SQP	Sequential Quadratic Programming

Bibliography

- [1] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*, ch. Discrete-Time signals and systems, pp. 8–70. Upper Saddle River: Pearson Higher education, Inc, 3 ed., 2010.
- [2] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*, ch. Computation of the Discrete Fourier Transform, pp. 629–699. Upper Saddle River: Pearson Higher education, Inc, 3 ed., 2010.
- [3] G. Strang, “Wavelet transforms versus fourier transforms,” *Bulletin of the American Mathematical Society*, vol. 28, no. 2, pp. 288–305, 2003.
- [4] I. C. Abou-Faycal, *Communication Systems*, ch. Quantization, pp. 50–55. 2012.
- [5] M. J. D. Powell, “A fast algorithm for nonlinearly constrained optimization calculations,” in *Lecture Notes in Mathematics* (G. Watson, ed.), vol. 630, Springer Verlag.
- [6] M. J. D. Powell, “The convergence of variable metric methods for nonlinearly constrained optimization calculations,” in *Nonlinear Programming 3* (R. R. M. Olvi L. Mangasarian and S. M. Robinson, eds.), Academic Press.
- [7] H. J. egou, M. Douze, and C. Schmid, “Product quantization for nearest neighbor search,” *PAMI*, vol. 33, no. 1.
- [8] K. H. T. Ge, Q. Ke, and J. Sun, “Optimized product quantization for approximate nearest neighbor search,” *In CVPR*.
- [9] K. He, F. Wen, and J. Sun, “K-means hashing: An affinitypreserving quantization method for learning binary compact codes,” *In CVPR*.
- [10] M. Norouzi and D. Fleet, “Cartesian k-means,” *In CVPR*.
- [11] R. S. Ellis, “An overview of the theory of large deviations and applications to statistical mechanics,” *Scandinavian Actuarial Journal*, vol. 1995, no. 1, pp. 97–142, 1995.

- [12] M. Nassralla, M. Mansour, and L. Jalloul, “A low-complexity detection algorithm for the primary synchronization signal in lte,” 2015.
- [13] M. Akkouchi, “On the convolution of exponential distributions,” *J. Chungcheong Math. Soc.*, vol. 21, no. 4, pp. 501–510, 2008.