University of Mississippi

eGrove

Electronic Theses and Dissertations Graduate School

1-1-2023

Detection of Rayleigh Faded Signal using Two Distributed Sensors

Sayantan Saha

Follow this and additional works at: https://egrove.olemiss.edu/etd

Recommended Citation

Saha, Sayantan, "Detection of Rayleigh Faded Signal using Two Distributed Sensors" (2023). *Electronic Theses and Dissertations*. 2572.
https://egrove.olemiss.edu/etd/2572

This Thesis is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

DETECTION OF RAYLEIGH FADED SIGNALS USING TWO DISTRIBUTED SENSORS

A Thesis
presented in partial fulfillment of requirements
for the degree of Master of Science
in the Electrical Engineering
The University of Mississippi

by

Sayantan Saha

May, 2023

Copyright Sayantan Saha 2023
ALL RIGHTS RESERVED

ABSTRACT

Distributed detection(DD) is an area of interest in the Cognitive Radio(CR) Network domain. With the help of geographically dispersed sensors and a Fusion Center(FC) (i.e. the network element which generates the global network decision), a DD system can determine the presence or absence of a primary user signal. This prevents any harmful interference by a secondary user who is trying to gain access to the unoccupied spectrum opportunistically. DD system gained its popularity in spectrum analysis due to the limited bandwidth of the reporting channels between the sensors and their fusion center.

But this process can be hindered by fading, a significant issue in wireless communication. Due to fading, the received signal can be varied significantly over amplitude and phase. This phenomenon increases the difficulty to detect a signal.

In this thesis, we studied detecting a faded signal in a Gaussian noise environment. Fading is often modeled as a random process. Here the faded signal has been assumed to be Rayleigh distributed. In the first part of the thesis, the detection of the signal has been converted into a binary hypothesis testing problem. We use a DD system with two parallel sensors for the detection process. By calculating the probability of detection error (P_e), we examine how accurately this particular system can detect the faded signal. In the second part, the Genetic Algorithm (GA), which is a well-known optimization algorithm, has been used to get the sub-optimal local decision rule for the sensors. And the thesis has concluded with the results which establish the effectiveness and validity of the proposed framework in a practical sensor system.

DEDICATION

This thesis work is dedicated to my parents and all the lovely people in my life who believed in me, supported me, and encouraged me persistently through my problems and difficulties. It is their prayers and love which make me able to get such honor and success.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Prof. Dr. Lei Cao, who has the attitude and the substance of genius. He continually and convincingly conveyed a spirit of adventure in regard to research and scholarship and excitement in regard to teaching. I would also like to mention Prof Dr. Ramanarayan Viswanathan for his time-to-time support in the process. Without their guidance and persistent help, this thesis would not have been possible. Special thanks to Dr. John Daigle for his advice and for serving on this thesis committee.

I also gratefully acknowledge the assistance of my friends and all the other people who helped me.

University of Mississippi　　　　Sayantan Saha

May 2023

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Distributed Detection and Fading Channels

In the area of Cognitive Radio (CR) networks, Spectrum Sensing (SS) plays a vital role to identify spectrum availability for the secondary Radio user (SU). It is a practical and effective way to prevent interference and collision with the primary user (PU) by finding spectrum holes and spectrum opportunities for CR transmission. SS can be done using either one sensor or cooperatively (Cooperative SS). It has been studied over the years that cooperative sensing performs better than the one-sensor SS due to its capability to exploit the advantage of spatial diversity [3]. There are mainly two ways to do the detection in the cooperative SS system: centralized detection(CD) [4]-[5] and decentralized detection(DD) [6].

A DD system can find a vacant licensed band for secondary transmission without any harmful interference to the primary user with the help of geographically displaced sensors and a fusion center (FC). In the case of DD, each sensor sends processed data to an FC whereas, in CD, all the sensors provide raw data to FC [7]. In this thesis work, we mainly focused on decentralized detection (DD). The DD system is becoming more appealing for wireless communication over CD due to its bandwidth efficiency and other advantages like the relatively low cost of sensors, the availability of high-speed communication networks, and low computational complexity have encouraged research in this topic [8].

Bandwidth limitations and high data rate costs forced system designers to quantize the data at each sensor before it is transmitted to the FC. The sensors of a DD network transfer the sensed information to the FC in either of two quantized (compressed) forms: single-bit local decision (hard-decision) [3] and multi-bit local decision (soft-decision) [8]. The FC then makes the final decision regarding the presence or absence of a primary user using these local decisions.

The focus of our thesis is to analyze a dual-sensor DD system capable of detecting signals in the presence of fading channels.. In wireless communication, the communication between a transmitter and a receiver is not ideal and the transmitted signal may go through various kinds of attenuation including path loss, multi-path attenuation, etc. This attenuation of the received signal amplitude and phase variation over time or space is known as Fading. Fading depends on various factors like time, radio frequency, and the path or position of the transmitter/receiver. In the fixed scenario, fading depends on atmospheric conditions such as rainfall, lightning, etc. In a mobile scenario, fading depends on obstacles over the path which are varying over time. These obstacles create complex transmission effects on the transmitted signal. Different fading types can be implemented through Rayleigh, Rician, Nakagami, Weibull, and other distributions.

Among the above-mentioned models, Rayleigh is the most commonly used one in wireless communication. In the Rayleigh fading model, it is assumed that the variation of a transmitted signal through a communication channel occurs according to a Rayleigh Distribution, i.e. that the received signal power can be treated as the sum of two signal components that are uncorrelated, squared Gaussian random variables. In this thesis work, it has been assumed that the received data at the two sensors can be correlated. Rayleigh fading is very useful and applicable in situations where the transmitter and the receiver are not propagating within the line-of-sight (LOS) region. Fig. 1.1 depicts the scenario.

Therefore, this fading model is a reasonable choice for radio signal propagation in heavily built-up urban areas, as well as for tropospheric and ionospheric signal propagation [9]. For this research work, we have chosen this particular model due to its applicability in such scenarios.

With this added complexity of fading and correlation, the performance of the DD system

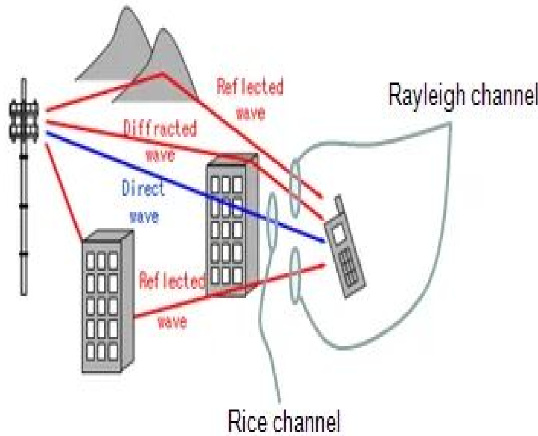

Figure 1.1. Rayleigh Fading Chanel

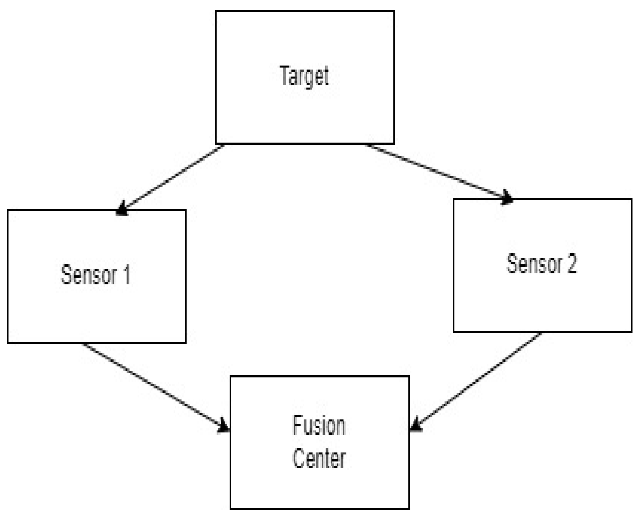

Figure 1.2. Two Sensor Distributed Detection System

has been assessed. The DD system under consideration consists of two sensors and an FC. The diagram of the system has been shown in Fig. 1.2. This DD system works as an on-off keying system to correctly detect the status of the presence or absence of the primary user.

The effectiveness of the system has been determined by calculating the probability of error (P_e) value. P_e value is the summation of probability of false alarm(P_f) weighted by the prior probability of H_0 and probability of miss(P_m) weighted by the prior probability of H_1. P_f defines the probability when the system falsely identifies the presence of PU in the spectrum. And P_m defines the probability when the PU is present but the system cannot detect it properly. With the combination of these two values, The probability of the system making an error can be decided.

In the end, Genetic Algorithm (GA) has been used to optimize the process. As implemented in [8], it helped to find the sub-optimal local decision rules for the sensors. In the result section of the thesis it has been shown that the system performance closely matches the centralized likelihood ratio test(CLRT) values with the increment of quantization levels.

1.2 Organization of Thesis

The rest of the thesis work is organized in the following manner. Chapter II describes the network model at first. It has explained how the sensors and FC communicate with each other and the working procedure of a receiver for a Rayleigh faded signal. In this chapter, it has also been shown that the problem at hand has been converted into a binary hypothesis testing problem. Chapter III addresses the different Bi-variate exponential distribution models and the corresponding calculation of P_{e}. In this section, we have also generated Bi-variate exponential distribution using an FGM (Ferlie-Gumbel-Morgenstern) copula. Results generated using the copula model showed that this model also performs in a similar manner to the other two established models. Chapter IV discusses the Genetic algorithm and its complexity. Finally, the thesis concludes with the presentation of simulation results in Chapter V and the overall conclusion in Chapter VI.

CHAPTER 2

SYSTEM MODEL AND PROBLEM STATEMENT

The system under consideration is an on-off keying system. It is the simplest form of ASK (amplitude shift keying) modulation scheme. Here the digital data represents the presence or absence of a carrier wave. It contains a binary communication channel where the source transmits over a T-second interval. In this process, a digital '1' transfers using a sinusoidal waveform, and a digital '0' stands for no transmission. Fig. 2.1 depicts the on-off keying modulation scheme.

The receiver observes the transmitted signal in the presence of additive zero-mean white Gaussian noise $v(t)$ with variance $\sigma^2 = N_0/2$, where N_0 is the two-sided power spectral density. In the case of transmission over fading channels, the attenuation in the received signal can be modeled as a random variable. A random change in phase also occurs during reception as well. Hence the problem of deciding the presence of a signal can be depicted as a binary hypothesis testing problem, which can be modeled in the following manner:

$$H_1 : Z(t) = \sqrt{\frac{2}{T}} Acos(\omega_c t + \Theta) + v(t) \quad 0 \le t \le T \tag{2.1}$$

$$H_0 : Z(t) = v(t) \quad 0 \le t \le T \tag{2.2}$$

where A and Θ are random variables with known prior distributions. Here it has been assumed that a Rayleigh Fading Channel is used for communication purposes. That is, A is Rayleigh distributed. Θ is uniformly distributed in $(0, 2\pi)$. A and Θ are independent of each other.

We can see that the observations under H_0 do not depend on A and θ. So the conditional likelihood function $\lambda(Z|A, \Theta)$ has been obtained first. Then it was averaged over the joint

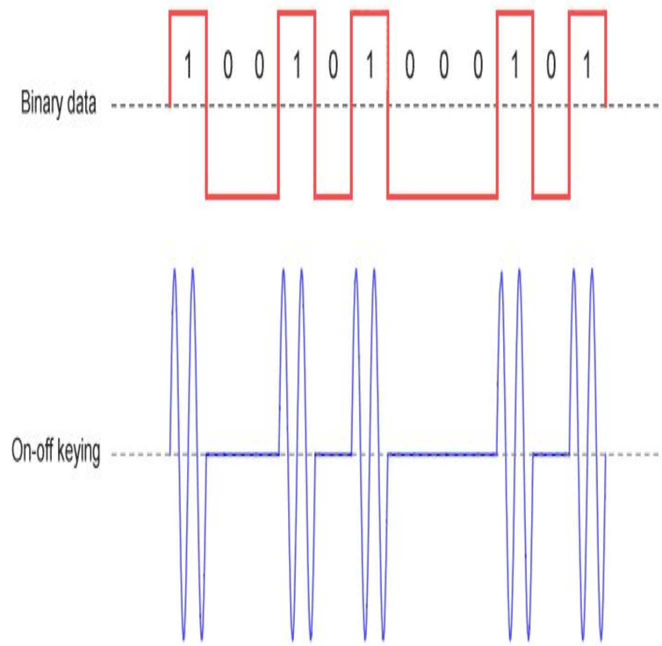

Figure 2.1. On-Off keying modulation

density function of A and θ to find the likelihood ratio $\lambda(Z)$:

$$\lambda(Z) == \int_{R_A}\int_{R_\Theta} \lambda(Z|a,\theta)P(a,\theta)dad\theta \tag{2.3}$$

Where a and θ are samples of A and Θ. And R_A and R_Θ are the ranges of A and Θ, respectively. Now the received signal $Z(t)$ can be expanded into an ortho-normal basis function to derive the conditional likelihood ratio. The first two chosen ortho-normal basis functions are:

$$g_1(t) = \sqrt{\frac{2}{T}}cos(\omega_c t), \qquad g_2(t) = \sqrt{\frac{2}{T}}sin(\omega_c t).$$

The rest of the ortho-normal functions can be chosen arbitrarily. With the help of these two basis functions, the received signal can be represented into two independent and

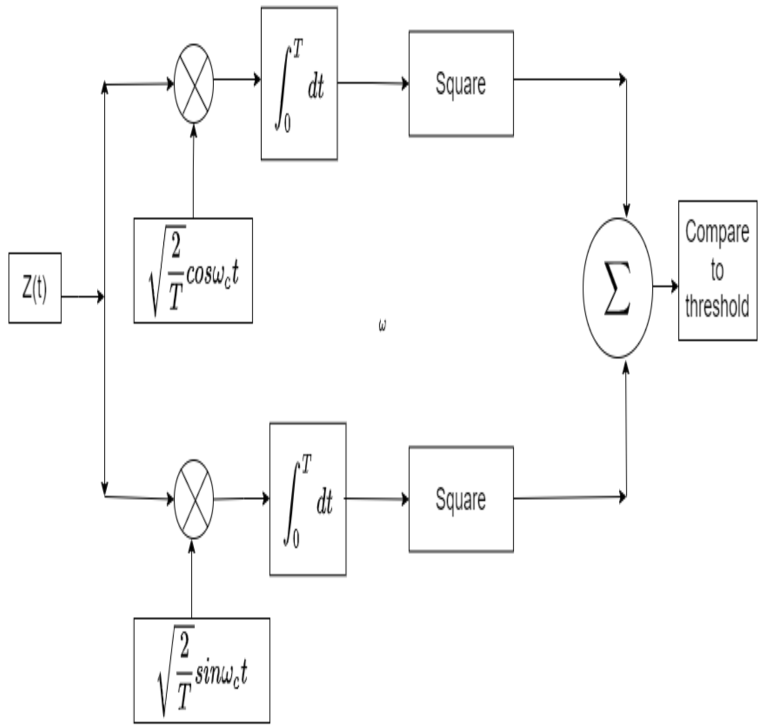

Figure 2.2. Receiver Structure of the system

orthogonal components Z_1 and Z_2 [10]. Only these two signal components will depend on the fact which hypothesis is true.

The process is shown in Fig. 2.2. Let

$$Z_1 = \int_0^T Z(t)g_1(t)dt \qquad\qquad Z_2 = \int_0^T Z(t)g_2(t)dt.$$

From the above discussion, it is very much comprehensible that Z_1, Z_2 are independent and identically distributed Gaussian variables with zero means. The variance for these two signal components are the same which is σ^2 under the H_0 , $(\sigma^2 + \sigma_s^2)$ under H_1. Now, the conditional likelihood ratio can be written as:

$$\lambda(Z|a,\theta) = \frac{P(Z_1, Z_2|a, \theta, H_1)}{P(Z_1, Z_2|a, \theta, H_0)} \tag{2.4}$$

Using these two signal components the probability distribution for the received signal under the H_1 hypothesis will be:

$$\begin{aligned} P(Z_1, Z_2|H_1) &= \int_{-\pi}^{\pi}\int_{0}^{\infty} P(Z_1 Z_2|a, \theta, H_1) da d\theta \\ &= \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_s^2)}} exp\left(\frac{-Z_1^2}{2(\sigma^2+\sigma_s^2)}\right) \\ &\quad \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_s^2)}} exp\left(\frac{-Z_2^2}{2(\sigma^2+\sigma_s^2)}\right) \end{aligned} \tag{2.5}$$

Similarly, the probability distribution for the received signal under H_0 hypothesis will be:

$$P(Z_1, Z_2|H_0) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(\frac{-Z_1^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}\sigma} exp\left(\frac{-Z_2^2}{2\sigma^2}\right) \tag{2.6}$$

The decision rule for the system is:

$$\begin{aligned} \lambda(Z) &= \frac{P(Z_1, Z_2|H_1)}{P(Z_1, Z_2|H_0)} \lessgtr \eta \\ \text{where the decision} &= \begin{cases} \mathrm{H}_1 & \text{if } \lambda > \eta \\ \mathrm{H}_0 & \text{if } \lambda < \eta \end{cases} \end{aligned} \tag{2.7}$$

here $\eta = \frac{P(H_0)}{P(H_1)}$, $P(H_0)$ and $P(H_1)$ are the priori probabilities of the hypotheses. After putting $P(Z_1, Z_2|H_1)$, $P(Z_1, Z_2|H_0)$ values and simplifying, the decision rule becomes:

$$\begin{aligned} &Z_1^2 + Z_2^2 \lessgtr t \\ &\text{where } t = \frac{\ln(\eta\frac{\sigma^2+\sigma_s^2}{\sigma^2})}{\frac{1}{2\sigma^2} - \frac{1}{2(\sigma^2+\sigma_s^2)}} \end{aligned} \tag{2.8}$$

Here Z_1 and Z_2 are two Gaussian distributions. So the summation of squares of Z_1, Z_2 will be an exponential distribution. Equivalently, for each sensor, the detection becomes to test hypotheses with the following distributions.

$$H_1 : X \backsim Exp(2(\sigma^2 + \sigma_s^2)) \tag{2.9}$$

$$H_0 : X \backsim Exp(2\sigma^2) \tag{2.10}$$

Where, $X = Z_1^2 + Z_2^2$

In our work, we assume the use of two sensors in a parallel network as shown in Fig. 1.2.

It is also assumed that observations in both sensors could be correlated. As a result, the problem becomes testing different Bivariate exponential distributions that are based on different hypotheses.

CHAPTER 3

DISCUSSION ON DIFFERENT BIVARIATE EXPONENTIAL DISTRIBUTION

After establishing the Binary Hypothesis Problem in the last chapter, different Bi-variate Exponential distributions have been used to find the solution for the problem. For this purpose, we have mainly used two joint Cumulative Distribution Functions (CDF) which are presented in [1]. From these CDF formulas, the PDFs have been generated in the context of the problem. These PDFs are used to calculate the P_e values. To extend the correlation range of observations, we have also used another PDF based on the Farlie-Gumbel Morgenstern (FGM) copula [11]. Through simulation results, it has been introduced that the former PDFs are performing similarly to the copula model. It has also been shown that the curves for other PDFs and the Copula model come very close to each other with the increasing value of quantization.

In the rest of the section, a detailed discussion of the different Bi-variate formulas has been presented. Firstly, we have discussed the CDFs and PDFs formulas presented in the paper. Also, some of the characteristics and traits have been presented. In the case of the copula, an introductory discussion regarding the copula has been presented which follows the characteristics of the FGM copula. We finish our discussion of the section by describing the formula which has been used to calculate the probability of error (P_e) using these different PDFs.

3.1 *First* Bi-variate Distribution formula

The first considered Bi-variate function given in [1] is the following:

$$
\begin{aligned}
F_{X_1,X_2}(x_1,x_2) = 1 - exp(-x_1) - exp(-x_2) \\
+exp(-x_1 - x_2 - \delta x_1 x_2)
\end{aligned}
\tag{3.1}
$$

for two different random variables X_1 and X_2 which are exponentially distributed. The boundary values for Eq.(3.1) are:

$$
\begin{aligned}
F(x_1, 0) &= 0; \\
F(x_2, 0) &= 0; \\
F(0, 0) &= 0; \\
F(\infty, \infty) &= 1;
\end{aligned}
\tag{3.2}
$$

In this CDF equation, we can see a parameter δ which can lead Eq.(3.1) to independence if the value of δ becomes 0. The range of this delta parameter lies between 0 to 1, i.e.,:

$$
0 \leq \delta \leq 1 \tag{3.3}
$$

This above inequality can be proven in the following manner:
At first, if we derive the PDF function from Eq. (3.1) then the equation should look like the following:

$$
f(x_1, y_1) = e^{-x_1(1+\delta x_2)-x_2}[(1 + \delta x_1)(1 + \delta x_2) - \delta] \tag{3.4}
$$

where,

$$
\begin{aligned}
f(x_1, \infty) &= 0; \\
f(\infty, x_2) &= 0; \\
f(0, 0) &= 1 - \delta;
\end{aligned}
\tag{3.5}
$$

Using the last equation and the non-negativity of a density function, it can be shown that the $\delta \leq 1$.

Now, according to the inequality which is true for any Bi-variate distribution is:

$$F(x_1, x_2) \leq F(x_1) \tag{3.6}$$

Using this above condition, we can simplify Eq.(3.1) in the following way:

$$-x_1(1 + \delta x_2) \leq 0 \tag{3.7}$$

for all x_1 and x_2. Due to the non-negativity clause of the distribution function, the δ parameter cannot be negative. Therefore $\delta \geq 0$.

The attributes of the first Bi-variate formula which has been described by the authors are:

1. Marginal probabilities are exponential:

$$\begin{aligned} f_{X_1}(x_1) &= \frac{\partial F(x_1, \infty)}{\partial x_1} = \frac{\partial(1 - exp(-x_1))}{\partial x_1} \\ &= \begin{cases} exp(-x_1) & \text{when } x_1 \geq 0 \\ 0 & \text{when } x_1 < 0 \end{cases} \text{with mean value 1} \end{aligned} \tag{3.8}$$

   Similarly, X_2 also has an exponential distribution with a mean value of 1.

2. It has been mentioned earlier that $\delta = 0$ leads to independence which means $F(x_1, x_2) = F(x_1)F(x_2)$. But δ is not the correlation co-efficient ρ. The relationship between δ and ρ is as follows [1]

$$\rho = -\frac{exp(\frac{1}{\delta})}{\delta} E_i(-\delta^{-1}) \tag{3.9}$$

   where, E_i is the Exponential Integral function. We have already proved that

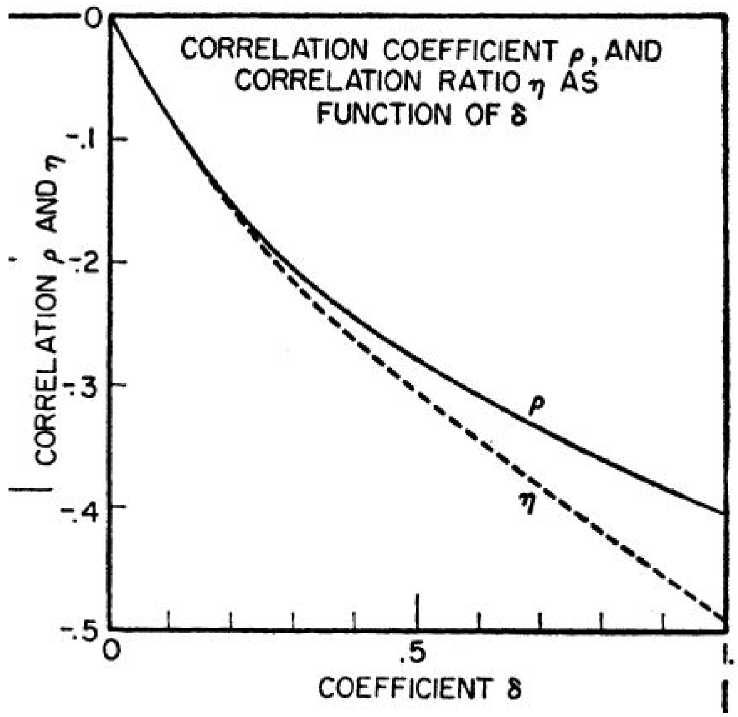

Figure 3.1. Relationship between δ and ρ from [1]

the range of parameter δ lies within 0 to 1, i.e. $0 \leq \delta \leq 1$. So, the range of ρ is $-0.40365 \leq \rho \leq 0$, which means that the correlation coefficient always stays non-positive. The relationship between δ and correlation coefficient ρ has been shown graphically in Fig. 3.1.

3. The other parameter η in the graph is the correlation ratio. The authors of [1] describes the relationship between η and δ as:

$$\eta^2(x_1|x_2) = -\frac{\delta}{3} - \frac{1}{6} - \frac{\rho}{6\delta} \tag{3.10}$$

For increasing value of δ, the correlation ratio decreases from 0 to -0.4837, as shown in the graph.

Under the context of the problem in hand, the random variables under consideration (i.e. observation of each sensor) should be:

1. Under H_1: Exponential distribution with mean $\theta^1 = \frac{1}{2(\sigma^2+\sigma_S{}^2)}$

2. Under H_0: Exponential distribution with mean $\theta^0 = \frac{1}{2\sigma^2}$

Now, let us assume two exponential variables X_1', X_2', which are represented as $X_1' = \theta_1 X_1$ and $X_2' = \theta_2 X_2$.

The joint CDF for X_1', X_2' can be written as

$$\begin{aligned} F_{X_1',X_2'}(x_1, x_2) &= P(X_1' < x_1, X_2' < x_2) \\ &= P(\theta X_1 < x_1, \theta X_2 < x_2) \\ &= P\left(X_1 \leq \frac{x_1}{\theta_1}, X_2 \leq \frac{x_2}{\theta_2}\right). \end{aligned} \tag{3.11}$$

Using these modified random variables, the joint CDF will look like

$$\begin{aligned} F_{X_1',X_2'}(x_1, x_2) = 1 - exp\left(-\frac{x_1}{\theta_1}\right) - exp\left(-\frac{x_2}{\theta_2}\right) \\ +exp\left(-\frac{x_1}{\theta_1} - \frac{x_2}{\theta_2} - \frac{\delta x_1 x_2}{\theta_1\theta_2}\right). \end{aligned} \tag{3.12}$$

And the joint PDF will be

$$f_{X_1',X_2'}(x_1, x_2) = \frac{1}{\theta_1\theta_2} exp\left(-\frac{x_1}{\theta_1} - \frac{x_2}{\theta_2}\right) exp\left(-\delta\left(\frac{x_1}{\theta_1}\right)\left(\frac{x_2}{\theta_2}\right)\right)\left[\left(1+\delta\left(\frac{x_1}{\theta_1}\right)\right)\left(1+\delta\left(\frac{x_2}{\theta_2}\right)-\delta\right)\right]. \tag{3.13}$$

Calculation of Eq.(3.13) from Eq.(3.12) is presented in Appendix A. The above equation is a general formula for the joint PDF under both hypotheses H_0 and H_1. Replacing θ_1 and θ_2 will give the PDF equations under H_1 and H_0, respectively. In other words, under H_1 we have $\theta_1 = \theta_2 = \theta^1$, and under H_0 we have $\theta_1 = \theta_2 = \theta^0$.

3.2 *Second* Bi-variate Distribution formula

In the last section, we have seen that the value of the correlation coefficient can only be negative. Now in this second formula, the correlation coefficient value ranges from negative to positive.

The discussion for the second formula can be started from the fact that it has been presented in [1]. If there are two given probability functions like $F(X_1)$ and $F(X_2)$ then a Bi-variate can be constructed in the means of the equation like :

$$F(x_1, x_2) = F(x_1)F(x_2)[1 + \alpha\{1 - F(x_1)\}\{1 - F(x_2)\}] \tag{3.14}$$

where,

$$-1 \leq \alpha \leq 1 \tag{3.15}$$

And the corresponding density function is given by:

$$f(x_1, x_2) = f(x_1)f(x_2)[1 + \alpha\{2F(x_1) - 1\}\{2F(x_2) - 1\}] \tag{3.16}$$

A Bi-variate probability function $F(x_1, x_2)$ with monotonically increasing marginal distribution should preserve these following conditions [1]:

1. F(- ∞, x_2) = F(x_1, - ∞) = 0

   F(x_1, ∞) = F(x_1); F(∞, x_2) = F(x_2); F(∞, ∞) = 1

2. The Probability content for every rectangle is non-negative, that is, for every $x_{11} \leq x_{12}, x_{21} \leq x_{22}$,

   Prob$\{x_{11} < X \leq x_{12},\ x_{21} < X \leq x_{22}\}$

   = F(x_{12}, x_{22}) - F(x_{12}, x_{21}) - F(x_{11}, x_{22}) + F(x_{11}, x_{21})

   which can also be written in the following manner:

$$\frac{\partial^2 F}{\partial x_1 \partial x_2} = f(x_1, x_2) \geq 0$$

Now, Eq.(3.14) and Eq.(3.16) holds conditions 1 and 2, and $F(X_1)$ and $F(X_2)$ are exponential functions.

So the second Bi-variate probability function which has been derived and presented in the paper[1] is:

$$\begin{aligned} F(x_1, x_2) &= (1 - e^{-x_1})(1 - e^{-x_2})[1 + \alpha e^{-x_1 - x_2}]; \\ f(x_1, x_2) &= e^{-x_1 - x_2}[1 + \alpha(2e^{-x_1} - 1)(2e^{-x_2} - 1)]; \end{aligned} \tag{3.17}$$

The relationship between parameter α and correlation coefficient ρ is:

$$\rho = \frac{\alpha}{4} \tag{3.18}$$

From Eq.(3.15) and Eq.(3.18), it can be easily derived that the range of correlation coefficient (ρ) is:

$$-0.25 \leq \rho \leq 0.25 \tag{3.19}$$

In contrast to the previous case, ρ can also be positive over here.

The relationship between the correlation ratio and the α parameter is:

$$\eta(x_1|x_2) = -\frac{\alpha}{2\sqrt{3}} \tag{3.20}$$

Hence, the relationship between $\eta(x_1|x_2)$ and ρ is:

$$\eta(x_1|x_2) = \frac{2\rho}{\sqrt{3}} \tag{3.21}$$

and the correlation value varies in the interval of -0.28867 to +0.28867.

Now we need to generate the PDF from this *second* formula in the context of the problem at hand. The equation which has been derived from the formula is

$$f_{X_1',X_2'}(x_1,x_2) = \frac{1}{\theta_1\theta_2} exp\left(-\frac{x_1}{\theta_1} - \frac{x_2}{\theta_2}\right) \left(1+\alpha\left(1-2exp\left(-\frac{x_1}{\theta_1}\right)\right) \left(1-2exp\left(-\frac{x_2}{\theta_2}\right)\right)\right) \tag{3.22}$$

From this general equation, the PDFs under H_1 and H_0 can be generated simply by replacing θ_1, θ_2 with θ^1 (for H_1 hypothesis) and θ^0 (for H_0 hypothesis) similar to the process followed for the *formula* 1.

After generating these PDF equations, we can use these formulas to calculate the probability of false alarm (P_{f}) and the probability of miss-detection (P_{m}). In the earlier section of the thesis, we have already mentioned that P_f is the probability of making a wrong detection when the primary user is not present, and P_m is the miss detection of the presence of the primary user. These two components together give the value of P_e which measures what is the overall probability of the system making a mistake.

The formula for the Probability of False Alarm (P_f) is

$$P_f = \iint\limits_R f_{X_1',X_2'}(x_1,x_2|H_0)\,dx_1\,dx_2 \tag{3.23}$$

The next thing we need to calculate is Probability of Detection(P_d) is

$$P_d = \iint\limits_R f_{X_1',X_2'}(x_1,x_2|H_1)\,dx_1\,dx_2 \tag{3.24}$$

The R represents the region within the observation space where the decision is 1 if (x_1, x_2) falls over it. P_d defines the probability of the system making the correct decision about the occupation of the spectrum by the primary user. The probability of miss (P_m) can be calculated using P_d by simply $P_m = 1 - P_d$. For the calculation purpose, the observation regions can be divided into multiple non-overlap small areas using thresholds of t_1 and t_2, with $t_1 \in \{t_{11}, t_{12}, ..., t_{1n}\}$, $t_2 \in \{t_{21}, t_{22}, ..., t_{2n}\}$, where t_1 and t_2 represent the set of thresholds for the observations in sensor 1 and sensor 2. After getting these two values, we can calculate the probability of error as $P_e = P_f P(H_0) + P_m P(H_1)$ which is the metric to judge the performance of detection.

3.3 *Copula* Model:

To consider a bi-variate model which is more general than the one presented in [1], we also use a model based on the FGM (Farlie-Gumble-Morgenstern) copula [12] - [13].

Copula is a computationally convenient way to describe the dependency between two random variables. We know that correlation also defines the dependency between two variables. But that is a more linear dependency. Using copula, we mostly describe the nonlinear dependency. Also, a copula is a non-redundant way to find out the joint distribution. It is because a two-dimensional CDF also contains information about one-dimensional CDFs. It means that when we are calculating the 2D values we also need to calculate the 1D values as well. But if we assume that the 1D CDF values are already known then soliciting them will be unnecessary. Therefore it is more desirable to describe the dependence in a more non-redundant way. In these situations the copula become very useful.

A general formula for the copula is:

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)) \tag{3.25}$$

Where X and Y are two random variables with the cumulative function $F_X(x) = Prob(X \leq x)$ and $F_Y(y) = Prob(Y \leq y)$ respectively.

For a two-dimensional distribution function to be a copula, it must satisfy some properties which can be easily derived from the definition of the copula:

1. C(0,v) = C(u, 0) = 0;

2. C(1, v) = v;

3. C(u,1) = u;

There are many different families of copula existed. But among them, the following Farlie-Gumble-Morgenstern (FGM) copula turns out to be very successful.

The general formula for FGM copula is:

$$C(u, v) = u \cdot v + \alpha \cdot u \cdot (1 - u) \cdot v \cdot (1 - v) \tag{3.26}$$

where u, v are uniformly distributed in $[0, 1]$, and α is a dependency parameter in the range of [-1,1]. If α reaches zero then the FGM copula corresponds to independence. u, v have positive dependence for $\alpha > 0$ and negative dependence for $\alpha < 0$.

The class of FGM copulas is a popular choice when working in two dimensions because of its simplicity in computation and the exact calculus of polynomial functions. According to [11], the fastest-to-compute copulas are polynomial copulas. And among them, the ones with the smallest possible degrees are better. The FGM copula fits in that category perfectly as mentioned in [14], [15]. It is also mentioned that when fuzzy logic is used to understand the dependency between random variables, the FGM copulas are naturally suitable.

Now from equation 3.26 if we derive the density function then it should be:

$$c(u, v) = 1 + \alpha(2u - 1)(2v - 1). \tag{3.27}$$

Converting the joint CDF equation in the context of the problem in hand should be:

$$\begin{aligned} F_{X_1', X_2'}(x_1, x_2) = &\left(1 - exp\left(-\frac{x_1}{\theta_1}\right)\right)\left(1 - exp\left(-\frac{x_2}{\theta_2}\right)\right) \\ &\left(1 + \alpha\left(exp\left(-\frac{x_1}{\theta_1}\right)\right)\left(exp\left(-\frac{x_2}{\theta_2}\right)\right)\right). \end{aligned} \tag{3.28}$$

The corresponding PDF is

$$
\begin{aligned}
f_{X_1',X_2'}(x_1,x_2) = \frac{1}{\theta_1\theta_2} exp\left(\frac{-x_1}{\theta_1}\right) exp\left(\frac{-x_2}{\theta_2}\right) \\
\left[1+\alpha\left(2exp\left(\frac{-x_1}{\theta_1}\right)-1\right)\left(2exp\left(\frac{-x_2}{\theta_2}\right)-1\right)\right].
\end{aligned}
\tag{3.29}
$$

The above PDF is used as the hypothesis distribution to check the probability of detection error, which is similar to the previously used bi-variate exponential formulas. In the results, it has been shown that all three of these distribution formulas are performed similarly.

CHAPTER 4

GENETIC ALGORITHM TO FIND THE MINIMUM P_e

This chapter discusses how the Genetic Algorithm is being used to produce a better-optimized system.

4.1 Overview of the genetic algorithm (GA)

A type of optimized algorithm is the Genetic algorithm [16], which is used to resolve issues in a variety of fields, including engineering, medicine, and finance. The main idea of this algorithm is based on natural selection, a process that drives biological evolution. A flowchart of the Genetic algorithm is presented in Fig.4.1. The algorithm starts with the initial population. And with every iteration, it creates a new generation by the selected parents of the old one. The selection for the parents of the new generation is mainly done in three ways:

- *Selection Rule* selects a set of parents for the next generation depending on their score.
- *Crossover* procedure takes two parents and combines them to make two new offspring.
- *Mutation* applies random changes to individual offspring.

With each generation, the population "evolves" towards an optimal solution.

The Genetic Algorithm has several advantages, which is the reason behind its widespread application and utilization, like:

- It can solve issues that are challenging to solve using conventional techniques.
- Problems that are stochastic can be solved using Genetic Algorithm.

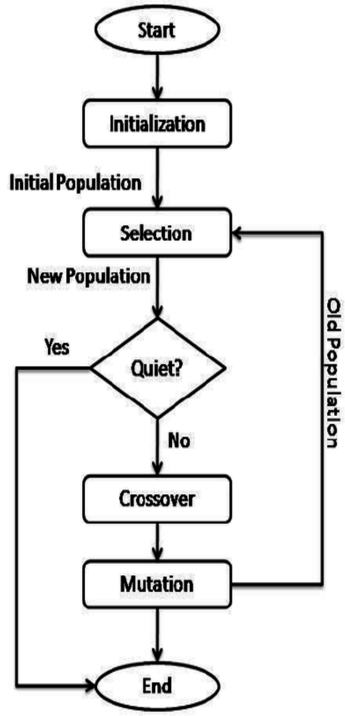

Figure 4.1. Flowchart of Genetic Algorithm [2]

- It can resolve issues with numerous objectives and restrictions.
- It is very useful when multiple local optima exist for a specific problem.

Due to these advantages, the genetic Algorithm has been employed in this specific problem to find the quantizers used in sensors that produce the minimum probability of error in this decentralized detection environment [8]. The pseudo-code of this algorithm is given in Algorithm 1. The algorithm first divides the observation region of the sensors into small areas, then assigns those small areas with suitable codewords (CW). The main purpose of the algorithm is to find the codeword assignment to the areas from a pool of codewords which generates the minimum probability of detection error. The likelihood Ratio test (LRT) is used at the FC for the final decision.

4.2 *GA complexity*

The time efficiency of the algorithm is judged on the following two aspects:

Algorithm 1. Genetic Algorithm pseudo code

1: Initialize Detection parameters:

   - π= Prior Prob
   - mh_0 = mean underH_0
   - mh_1 = mean under H_1
   - parameter values for different formulas

2: Divide the total observation range of both sensors into small intervals using thresholds.
3: Measure and store the probability mass of each small area under both hypotheses
4: Initialize Genetic Algorithm(GA) parameters:

   - nP= number of population
   - nI = number of Iteration
   - Cr = crossover rate
   - Mr = mutation rate

5: Start GA
6: Generate initial population of codewords(CW) of size nP
   (current CW <nP)
7: Calculate P_{e}:

   - assign the small intervals with CWs
   - Randomly assign CWs for each interval of both sensors.
   - calculate total probability mass for each set of combinations under both hypotheses using the stored value
   - make decision of H_1 and H_0 using LRT Fusion rule
   - summing the total probability mass under hypothesis H_0 when the decision is 1 to calculate False Alarm (P_{f})
   - summing the total probability mass under hypothesis H_1 when the decision is 0 to calculate Miss Detection (P_{m})
   - Calculate P_e: $P(\mathrm{H}_0)\mathrm{P}_{\mathrm{f}}+\mathrm{P}(H_1)P_{\mathrm{m}}$ where $P(\mathrm{H}_0)$ and $P(\mathrm{H}_1)$ are prior probabilities
   - P_{e} value

8: Store the min P_{e} (current iteration <nI)
9: Select two parent CW using Roullette Wheel selection
10: Perform crossover to produce the offspring CW (random probability< Mr)
11: Mutate the offspring CW
12: Calculate P_{e} for each off-spring CW and update best solution
13: Add offspring to population and remove least-fit solution
14: population=new-population
15: Return(min P_{e} value, best CW)
16: End GA

1. *Counting the number of Fitness function calculation*: The amount of time the fitness function can be calculated using the following step:

$$nP \times nI = (\text{number of population}) \times (\text{number of iteration})$$

2. *Time complexity to calculate the fitness function for each set of population*: As we have already stored the value of probability mass for each small area under both hypotheses, the time required for calculating the prob. mass for one set of CW is a constant. We have defined this time with PM_{CW}. For two different hypotheses, the total time required for each set of the population is

$$(PM_{\mathrm{CW(H_0)}} + PM_{\mathrm{CW(H_1)}}) \times nP$$
$$\text{where, } PM_{CW(H_j)}, j = 0, 1$$

CHAPTER 5

PERFORMANCE EVALUATION

In this section, the simulation studies have been provided. The detection performance is evaluated for different Bi-variate exponential distributions. P_{e} has been taken as the performance metric. The GA parameters were the same for all cases, which are *No. of Iterations* = 200, *number of Population* = 500, *Crossover rate* = 0.8, *Mutation rate*= 0.05. Also, the curves are compared among three types of quantizers i.e. 1-bit, 2-bit, and 3-bit.

1. *Scenario 1*: The performance of P_{e} has been tested against the delta (δ) parameter in Fig.5.1. The joint pdf, which has been used to generate this figure, is from the first model in [1]. As mentioned earlier, the δ parameter is inversely related to the correlation coefficient (ρ) which is always non-positive in value in this scenario. This relation has been depicted in the figure where it can be seen that P_{e} value decreases with the increasing value of δ. The prior probabilities are equal and the mean values under H_0 and H_1 are set as [4,4] and [2,2] respectively.

2. *Scenario 2*: In Fig. 5.2, the P_{e} vs P(H_1) plot has been presented. The mean values under both hypotheses stay the same as the previous scenario, and the δ value was fixed at 0.5. This figure has used the same probability of distribution function. As we can see from the figure all the quantization curves give the highest P_{e} at $P(H_1) = 0.6$. It is because the highest uncertainty happens when the prior probability value stays in this range.

3. *Scenario 3*: Fig. 5.3 plots P_{e} vs α. This α parameter comes from the second formula of the Bivariate exponential distribution from Gumble's paper [1]. The range of α lies between -1 to 1. And it bears a relationship with the correlation coefficient, i.e., $\rho = \frac{\alpha}{4}$.

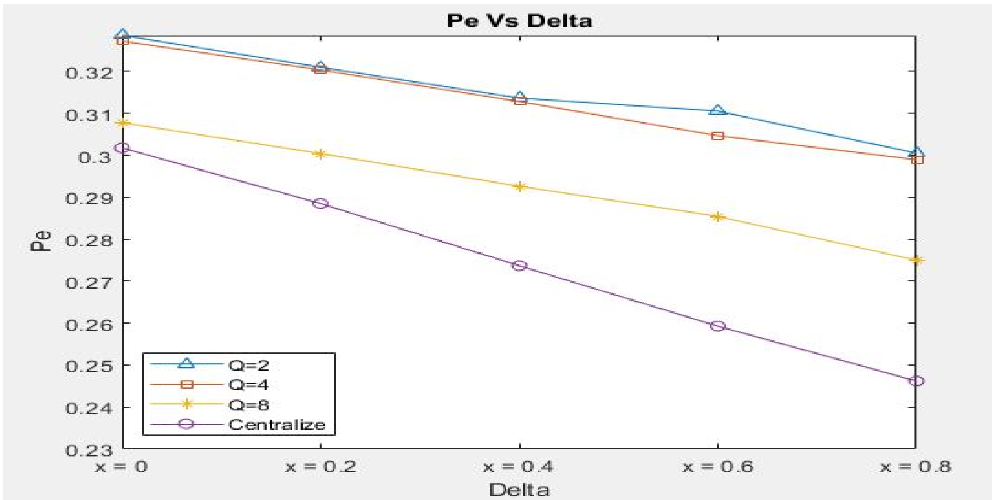

Figure 5.1. P_e Vs δ

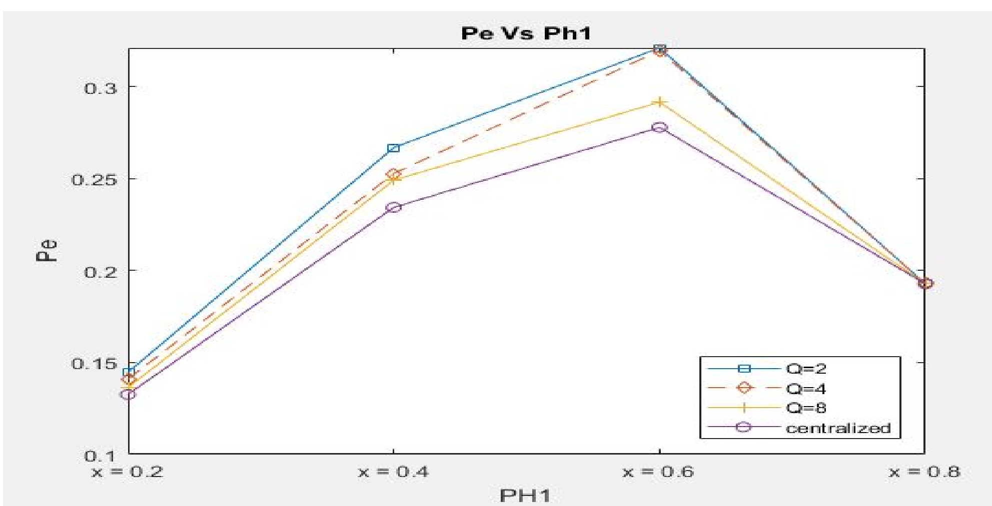

Figure 5.2. P_e Vs Prior Probability(using the first formula)

Using this second formula of Bivariate exponential distribution, the nature of the dependence of P_e over the positive value ρ can be studied. In this figure, the value of α increases (which means ρ is also increasing), and the probability of error (P_e) increases as well. Also, the P_e curves are getting closer to the centralized one, i.e., without quantization, with the increasing value of three quantization level.

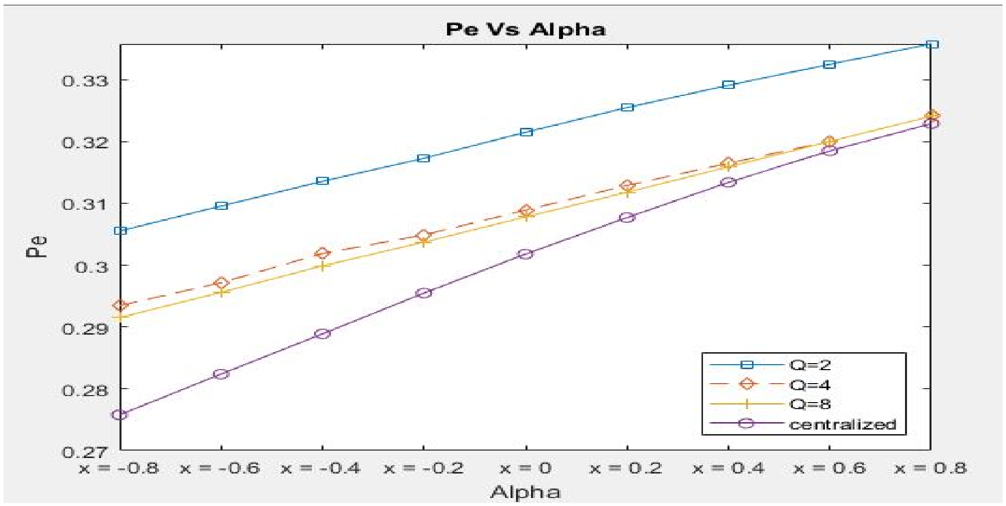

Figure 5.3. P_e Vs α

4. *Scenario 4*: Fig. 5.4 again studied the P_e vs $P(H_1)$ to examine whether any changes take place due to the change of Bi-variate exponential distribution formula. But in the figure, we can notice that the behavior of the curve is similar to the earlier P_e vs $P(H_1)$ curve which has been generated using the *first* formula of [1] of Bivariate exponential distribution. The value of α, which has been used in this figure is 0.25. So the correlation coefficient (ρ) value is 0.0625.

   Establishing a relationship between the α and the δ parameter (from the first formula) is difficult. So, a detailed comparison between the two curves has not been presented here.

5. *Scenario 5*: In the Fig. 5.5, the performance of the copula formula has been shown. The α parameter(dependency parameter value) value is 0.5. As we can see in the figure that the P_e values for the different prior probabilities are similar to the values which

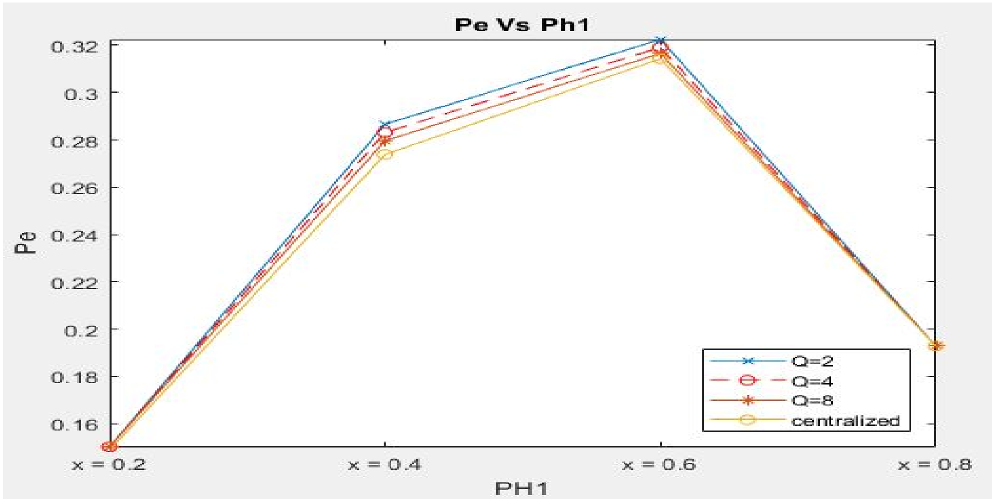

Figure 5.4. Pe Vs Prior Probability(using the second formula)

we got using the previous formulas. The curve reaches closer to the centralized value with the change in the quantization levels.

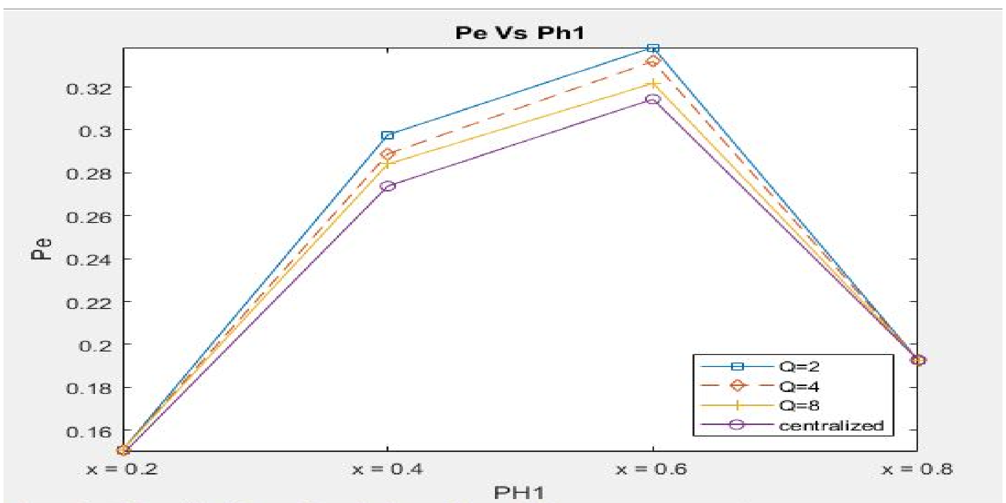

Figure 5.5. P_e Vs $P(H_1)$ (Using the FGM copula)

CHAPTER 6

CONCLUSION

In this thesis work, it has been shown how distributed detection performs over a fading channel with possible correlation in a two-sensors network. It has been mentioned earlier how fading is very common in today's wireless communication environment. As wireless communication becomes part of our everyday life, a system that can properly detect despite the difficulty of fading is very important.

Furthermore, we investigated the impact of correlation on the system's performance by examining both negative and positive correlations. We also presented the system's performance using a copula model that we generated. Simulation results show that our generated model performs similarly to two established models taken from [1].

It had been shown how the problem has been transformed into a binary hypothesis testing problem where the observations can be represented in the form of a Bi-variate exponential distribution. And last but not least, a Genetic Algorithm is proposed in search of the Sub-optimal solution. Using simulation it has been shown how the performance of the system gets closer to the centralized likelihood ratio solution with increasing quantization levels.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] E. J. Gumbel, "Bivariate exponential distributions," *Journal of the American Statistical Association*, vol. 55, no. 292, pp. 698–707, 1960.

[2] A. Dastanpour, S. Ibrahim, R. Mashinchi, and A. Selamat, "Using genetic algorithm to support artificial neural network for intrusion detection system," *Journal of Communication and Computer*, vol. 11, pp. 143–147, 2014.

[3] H. Kasasbeh, L. Cao, and R. Viswanathan, "Hard decision based distributed detection in multi-sensor system over noise correlated sensing channels," in *2016 Annual Conference on Information Science and Systems (CISS)*. IEEE, 2016, pp. 280–285.

[4] H. L. Van Trees, *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons, 2004.

[5] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 1998.

[6] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: distributed detection of a known signal in dependent gaussian noise," *IEEE Transactions on signal processing*, vol. 48, no. 12, pp. 3266–3279, 2000.

[7] L. Khalid and A. Anpalagan, "Cooperative sensing with correlated local decisions in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 2, pp. 843–849, 2011.

[8] H. Kasasbeh, L. Cao, and R. Viswanathan, "Soft-decision based distributed detection over correlated sensing channels," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2017, pp. 1–6.

[9] T. S. Rappaport, "Wireless communications–principles and practice, (the book end)," *Microwave Journal*, vol. 45, no. 12, pp. 128–129, 2002.

[10] M. D. Srinath, P. Rajasekaran, and R. Viswanathan, *Introduction to statistical signal processing with applications*. Prentice-Hall, Inc., 1995.

[11] S. Sriboonchitta and V. Kreinovich, "Why are FGM copulas successful? a simple explanation," *Advances in Fuzzy Systems*, vol. 2018, 2018.

[12] D. J. Farlie, "The performance of some correlation coefficients for a general bivariate distribution," *Biometrika*, vol. 47, no. 3/4, pp. 307–323, 1960.

[13] H. Bekrizadeh, G. A. Parham, and M. R. Zadkarmi, “The new generalization of farlie–gumbel–morgenstern copulas,” *Applied Mathematical Sciences*, vol. 6, no. 71, pp. 3527–3533, 2012.

[14] V. Kreinovich, S. Sriboonchitta, and V.-N. Huynh, *Robustness in Econometrics*. Springer, 2017.

[15] Z. Wei, D. Kim, T. Wang, and T. Teetranont, “A multivariate generalized fgm copulas and its application to multiple regression,” *Robustness in Econometrics*, pp. 363–378, 2017.

[16] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.* MIT press, 1992.

APPENDICES

APPENDIX A

DERIVATION OF THE PDF EQUATION FROM THE CDF

In this appendix, we show the derivations of the joint PDF equation from the joint CDF equation from the first formula of [1]. This derivation is done in the context of the problem at hand.

The joint CDF equation is:

$$\begin{aligned} F_{X_1',X_2'}(x_1,x_2) = 1 - exp\left(-\frac{x_1}{\theta_1}\right) - exp\left(-\frac{x_2}{\theta_2}\right) \\ +exp\left(-\frac{x_1}{\theta_1} - \frac{x_2}{\theta_2} - \frac{\delta x_1 x_2}{\theta_1 \theta_2}\right). \end{aligned} \tag{A.1}$$

Now,

$$f_{X_1',X_2'}(x_1,x_2) = \frac{\partial^2 F_{X_1',X_2'}(x_1,x_2)}{\partial x_1 \partial x_2}$$

$$= \frac{\partial}{\partial x_2}\left[\frac{\partial}{\partial x_1}\left(1 - exp\left(\frac{-x_1}{\theta}\right) - exp\left(\frac{-x_2}{\theta}\right) + exp\left(-\frac{x_1}{\theta} - \frac{x_2}{\theta} - \frac{\delta x_1 x_2}{\theta^2}\right)\right)\right]$$

$$= \frac{\partial}{\partial x_2}\left[\frac{1}{\theta}exp\left(\frac{-x_1}{\theta}\right) - \frac{1}{\theta}exp\left(-\frac{x_1}{\theta} - \frac{x_2}{\theta} - \frac{\delta x_1 x_2}{\theta}\right) - \frac{\delta x_2}{\theta^2}exp\left(-\frac{x_1}{\theta} - \frac{x_2}{\theta} - \frac{\delta x_1 x_2}{\theta^2}\right)\right]$$

$$f_{X_1',X_2'}(x_1,x_2) = \frac{1}{\theta^2}exp\left[-\frac{x_1}{\theta} - \frac{x_2}{\theta}\right] exp\left[\left(-\delta\left(\frac{x_1}{\theta}\right)\left(\frac{x_2}{\theta}\right)\right)\left[1 + \frac{\delta x_1}{\theta} + \frac{\delta x_2}{\theta} + \frac{\delta^2 x_1 x_2}{\theta^2} - \delta\right]\right]$$

After simplifying and rearranging the last equation we will get

$$f_{X_1',X_2'}(x_1,x_2) = \frac{1}{\theta_1\theta_2} exp\left(-\frac{x_1}{\theta_1} - \frac{x_2}{\theta_2}\right) exp\left(-\delta\left(\frac{x_1}{\theta_1}\right)\left(\frac{x_2}{\theta_2}\right)\right)\left[\left(1+\delta\left(\frac{x_1}{\theta_1}\right)\right)\left(1+\delta\left(\frac{x_2}{\theta_2}\right)-\delta\right)\right]. \tag{A.2}$$

VITA

Sayantan Saha was born in Kolkata, India, in 1990. He received his Bachelor of Technology degree in Electronics and Communication Engineering from Maulana Abul Kalam Azad University of Technology, West Bengal, in 2013. He then went on to obtain a Master of Technology degree in Telecommunication Engineering from the same university in 2015. In August 2017, he started pursuing an M.S. degree in electrical engineering at The University of Mississippi, during which time he served as a research assistant under Dr. Lei Cao. Additionally, he worked as an instructor for Electric Circuit Lab under the guidance of Dr. Ramanarayan Viswanathan.