DIVERGENCE AND BAYES ERROR BASED SOFT DECISION FOR DECENTRALIZED SIGNAL DETECTION OF CORRELATED SENSOR DATA

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

by

Roopa Shree Rajanna

December 2015

ABSTRACT

Spectrum sensing in cognitive radio involves the determination of presence or absence of a primary user signal so that secondary users may opportunistically gain access when the spectrum is unoccupied. In decentralized sensing scheme, two or more secondary users (SUs) sense the spectrum, process individual observation and then pass quantized data to a fusion center, where the decision on signal present hypothesis or signal absent hypothesis is made. The reporting channel between a SU and the fusion center (FC) is typically bandlimited and error prone.

In this thesis we consider the problem of design of quantizers at the SUs assuming i) the secondary users' observations have known joint distributions, conditioned on the signal present and absent hypothesis and ii) the reporting channel is a binary symmetric channel with known error probability. The quantizer design is to obtain, at the minimum a locally optimal design based on optimizing one of three possible criteria: i) KL divergence metric ii) Chernoff metric iii) direct metric, that is Bayes error, with all metrics computed for the data received at the FC. A successful quantizer design must consider three integrated-issues simultaneously, i.e., quantizer thresholds, binary codeword assignment, and error resilience.

Numerical simulation results are obtained assuming that two SUs' observations are conditionally distributed as bivariate Gaussian, with identical marginal distributions and higher mean value under the signal present hypothesis. The results indicate that all of the quantizer designs possess the inherent property of resilience to channel errors, that is, given the limit of D bits transmission through the reporting channel, the quantizer would represent the data with less than the maximum allowable 2^D quantization intervals, with appropriate binary codewords, thus using the unused codewords for combating channel error. The higher the channel error, the coarser will be the quantization of user observation. The results also

indicate that the quantizer designed with the knowledge of known joint distributions of SUs' data outperform, in the sense of KL divergence, Chernoff, or Bayes error, as the case may be, when compared to the case of quantizer designed as though the users' data were statistically independent. Hence, for obtaining better performance at the fusion center, it is not possible to ignore correlation, especially when the correlation is significant.

DEDICATION

In dedication to my parents for making me be who I am, and my loving husband for supporting me all the way!

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my advisors, Prof Dr. Ramanarayanan Viswanathan and Prof Dr. Lei Cao, who have the attitude and the substance of a genius: They continually and convincingly conveyed a spirit of adventure in regard to research and scholarship, and an excitement in regard to teaching. 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 the thesis committee.

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

University, Mississippi Roopa Shree Rajanna

August 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AWGN additive white gaussian noise

BEP bit error probability

BER bit error rate

BPSK binary phase shift keying

BSC binary symmetric channel

CR cognitive radio

FC fusion center

KL Kullback-Leibler

LLR log likelihood ratio

LRT likelihood ratio test

PMF probability mass function

PSD power spectral density

PU primary user

RF radio frequency

ROC receiver operating characteristic

SNR signal to noise ratio

SU secondary user

CHAPTER 1

INTRODUCTION

The problem addressed in the thesis is the design of quantizers for digital representation of sensing data, received at SUs' in a cooperative spectrum sensing scheme for cognitive radio (CR) systems. This is an extension of previous work where the SUs data were assumed statistically independent. Before we introduce the specific problem, we provide a brief introduction to cognitive radio systems and cooperative spectrum sensing.

1.1 COGNITIVE RADIO SYSTEMS

Cognitive radio is one of the new long term developments taking place in radio communications technology. The fundamental task of each CR user in CR networks, in the most primitive sense, is to detect the licensed users, also known as primary users (PUs), if they are present and identify the available spectrum if they are absent. This is usually achieved by sensing the radio frequency (RF) environment, a process called spectrum sensing $[1]-[2]$. In $[3]-[4]$, Mitola stated that CRs are capable of sensing their environment, learning about their radio resources and user/application requirements, and adapting behavior by optimizing their own performance in response to user requests.

1.2 COOPERATIVE SPECTRUM SENSING

As mentioned, spectrum sensing is a key function of CR to prevent the harmful interference with licensed users and identify the available spectrum for improving the spectrum's utilization. Spectrum sensing with different detection techniques are reviewed in [5]. This spectrum sensing is challenging in low signal-to-noise ratio (SNR) regions, which can be caused by severe fading, shadowing or blocking in the CR sensing channel. One possible

approach to increase the spectral estimation reliability and decrease the probability of interference of CRs to existing radio systems is by using distributed spectrum sensing. In such a distributed approach, the spectrum occupancy is determined by the joint work of cognitive radios, as opposed to being determined individually by each cognitive radio. However, when cooperating CR users are geographically proximate, they may experience similar environmental shadowing effects and thus have correlated shadowing [6]. Hence, correlated shadowing could cause the sensors' observations to be correlated. Also, in general, the complexity of the cooperative spectrum sensing approach is higher; it needs a control channel and increases the traffic overhead. Cooperative sensing can be implemented in two fashions: Centralized and Decentralized. In centralized scheme, each sensor transmits all of its observation to a FC. The FC will aggregate the local information and make a final decision. In decentralized detection scheme, also called distributed detection, each sensor sends a summary of its own observations (soft or hard decisions) to the fusion center. The fusion center makes a decision on the basis of the summary it has received. We digress a little to briefly review distributed detection literature

Distributed detection has attracted substantial attention for its application to the cooperative spectrum sensing in cognitive radios [7]. A review of distributed signal detection procedures was discussed in [8]. In application scenarios involving geographically distributed radios, such as a wireless communication system, distributed spectrum sensing approaches are worth considering due to the variability of the radio signal, as suggested in [9]. Such methods may significantly increase the reliability of the spectrum estimation process, at the expense of computational complexity and power/bandwidth usage for the transmission of spectrum sensing information. The detection performance can be primarily determined on the basis of two metrics: probability of false alarm, which denotes the probability of a CR user declaring that a PU is present when the spectrum is actually free, and probability of detection, which denotes the probability of a CR user declaring that a PU is present when the spectrum is indeed occupied by the PU. Since a miss in the detection will cause interference

with the PU and a false alarm will reduce the spectral efficiency, it is usual to assume that the probability of false alarm is minimized subject to the constraint of the probability of miss (Neyman-Pearson criterion). The performance of a distributed detection system with the given local decision rules and correlated local decisions was studied in [10], and the optimum decision fusion rule in Neyman-Pearson sense was derived and analyzed.

The performance of cooperative spectrum schemes based on hard decision and soft decision has been compared in [11] − [13] for the case in which decisions are sent over error-free reporting channels. Spectrum sensing considering soft decision and hard decision in the presence of channel error is studied in [14]. The authors have highlighted the advantages of using soft decisions, instead of hard decisions, to improve the performance of cooperative spectrum sensing, specially in the presence of reporting channel errors. Channel coding is a technique used for controlling errors in data transmission over unreliable or noisy communication channels. It does this by introducing redundant data, called error correcting code, prior to data transmission or storage. This operation involves increased data packet size. So, rather than using forward error correction coding, a quantizer can be designed at the SU with error resilient feature. Some time, the performance loss due to quantization in signal detection and estimation is investigated in terms of Ali-and-Silvey divergence loss [15]. A fundamental issue related to quantization and error resilience design is the binary codeword assignment. The quantizer design must consider the quantization thresholds, binary codeword assignment and error resilience simultaneously.

1.3 THESIS WORK

An earlier work looked at the quantizer design for independent and identical SU's data [16]. The contribution of this thesis is to present a joint quantizer design that considers local quantization, codeword assignments and channel error resilience from KL and Chernoff metric divergences perspective, for correlated secondary users' data. Further goal is to examine the design based on Bayes error, instead of divergent measures. The quantizer

design procedure is based on a two stage algorithm and is developed using Flynn and Gray algorithm [17], which is executed iteratively for guaranteed convergence. The secondary user is allowed to transmit the quantized data only with a limited rate. Problem is to maximize divergence computed for the data at the FC. The design is based on person-by-person optimization and therefore provides a local optimality. Also, investigation of error probability of a likelihood ratio test at the FC, for different types of local quantizers, is considered. We observe that by applying the proposed quantizer design, improvement in KL and Chernoff metric divergences are achieved over the quantizer designed as though the SU observations were statistically independent.

1.4 ORGANIZATION OF THESIS

The rest of the thesis is organized as follows. In chapter 2, we present system model and concepts which are essential in understanding quantizer design. We first design the quantizers based on distributional divergence and Bayes error criterion based on our proposed algorithms. Numerical results are then analyzed in chapter 3. In chapter 4, we summarize the content of thesis, and discuss the contributions of our work. Recommendations for further research in this topic are also included in this chapter.

CHAPTER 2

JOINT QUANTIZER DESIGN FOR SPECTRUM SENSING

In this chapter, we formulate quantizer design algorithms with the assumption that sensors' observations (i.e., SUs' data) are distributed as bivariate Gaussian. We consider CR network with two SUs, which can detect the presence of PU. Let H_0 be the null hypothesis (i.e, PU is absent) and H_1 be the alternative hypothesis (i.e., PU is present) with priors of π_0 and $1-\pi_0$, respectively as shown in Figure 2.1. Each SU receiver senses the spectrum signal and makes a local log-likelihood ratio (LLR) statistic, denoted L_n $(n = 1, 2)$. Assume that the SUs' data are correlated with correlation coefficient $\rho\,(-1 \leq \rho \leq 1)$. Information L_n is quantized into a D-bits binary codeword, L_n^{su}, which is then transmitted through a binary symmetric channel with bit error probability (BEP) P_b. At the fusion center, L_n^{su} is received as L_n^{fc}. The best decision rule, i.e, LLR detection rule is used at the fusion center to arrive at a final decision. Let

$$\Lambda = \log \frac{P(L_1^{\text{fc}}, L_2^{\text{fc}}|H_1)}{P(L_1^{\text{fc}}, L_2^{\text{fc}}|H_0)}. \tag{2.1}$$

Using a threshold η, the FC implements the test

$$\Lambda \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{2.2}$$

The false alarm probability and miss detection probability are defined as

$$\alpha = P_f = \int_{\eta}^{+\infty} P(\Lambda|H_0)d\Lambda. \tag{2.3}$$

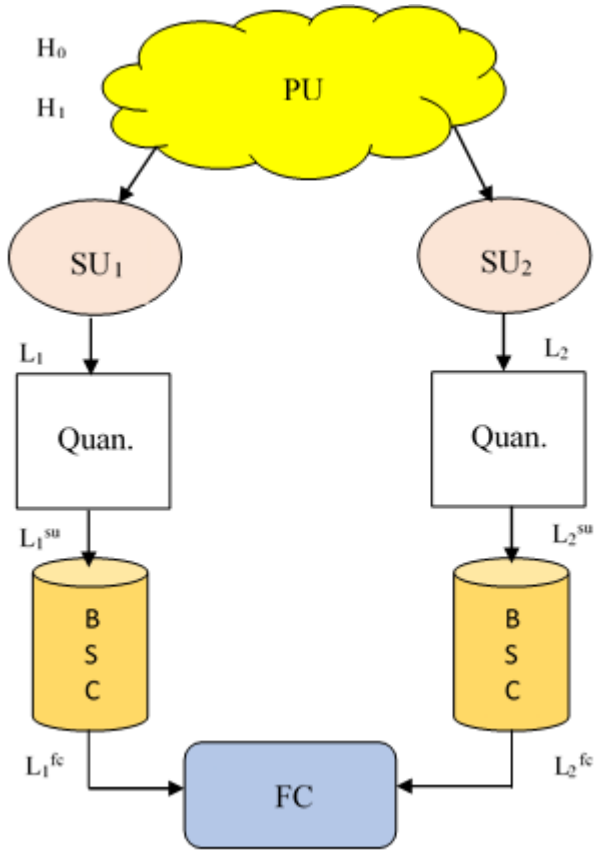

Figure 2.1. Sensors and data fusion model

$$\beta = P_m = 1 - P_d = 1 - \int_{\eta}^{+\infty} P(\Lambda|H_1)d\Lambda. \tag{2.4}$$

2.1 QUANTIZATION INTERVALS AND CODEWORD ASSIGNMENT

The quantizers and reporting channels between SUs and FC are shown in Figure. 2.1. Let W ($W = 2^D$) be the number of levels of the quantizer as shown in Figure. 2.2.

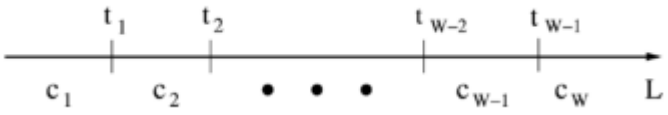

Figure 2.2. Quantizer threshold

Let t_m be the quantization thresholds, where $m = 0, \ldots, W$, $t_0 = -\infty$ and $t_W = +\infty$. Since

both the SUs' data are correlated, calculation of probability mass of an interval considers both SU's LLRs. Hence, combining quantizer thresholds of both SU leads to a two-dimensional cell structure with LLR of SU1 on x-axis and LLR of SU2 on y-axis as shown in Figure 2.3. We denote $P_j^{\text{su}}(i,k)$, $i,k = 1,\ldots,W$ as the probability that the quantized symbol is in the cell (i,k) under hypothesis $j, j = 0,1$.

Then probability mass of each cell is given by

$$P_j^{\text{su}}(i,k) = \int_{t_{i-1}}^{t_i}\int_{t_{k-1}}^{t_k} f(L_1,L_2|H_j)dL_1dL_2. \tag{2.5}$$

where

$$f(L_1,L_2) = \frac{1}{2\pi\sigma_0\sigma_1\sqrt{1-\rho^2}}\exp\left[-\frac{z}{2(1-\rho^2)}\right], \tag{2.6}$$

$$z \equiv \frac{(L_1-\mu_0)^2}{\sigma_0^2} - \frac{2\rho(L_1-\mu_0)(L_2-\mu_1)}{\sigma_0\sigma_1} + \frac{(L_2-\mu_1)^2}{\sigma_1^2}.$$

The codewords assignments for these W quantization intervals are (i,c_i) for SU1 and $(k,c\text{`}_k)$ for SU2, $i,k = 1,\ldots,W$, where $c_i$ and $c\text{`}_k$ are D-bits binary sequence. Concatenation of $c_i, c\text{`}_k$ binary sequences represents the $(i,k)^{\text{th}}$ cell's codeword. This concatenated codeword $c_q$, $q = 1,\ldots,W*W$ is in the vector space, $V(2^{2D})$, consisting of all $2^{2D}$ bits binary sequences, i.e., $c_i \in V(2^D)$ and $c\text{`}_k \in V(2^D)$. Let $P_j^{\text{su}} = [P_j^{\text{su}}(1),\ldots,0,\ldots,0,\ldots,P_j^{\text{su}}(W*W)]$ be the probability mass of the quantized data under hypothesis, j for all possible codeword combinations. Then probability mass of each codeword combination is calculated as the sum of the probability mass of the cells with the respective codeword, c_q. Existence of the zero elements means that the number of quantization levels is less than 2^D for each SU and hence, some D-bits binary sequences are not used for the codeword assignment. This quantized data passes through a binary symmetric channel (BSC) with bit error probability, P_b. A unique probability transition matrix, $P_T = \{p_{ij}\}, i,j = 1,\ldots,2^{2D}$, can be obtained

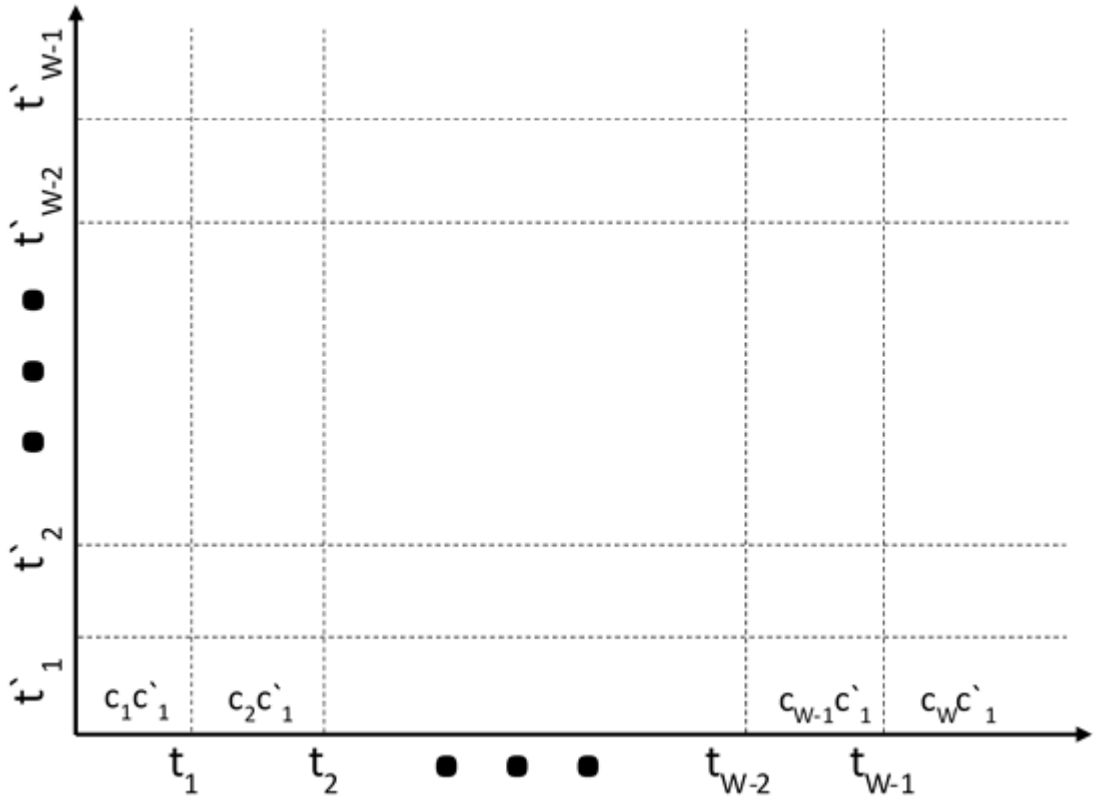

Figure 2.3. 2-D cell structure

among all possible vectors in $V(2^{2D})$. Let $v_i, v_j \in V(2^{2D})$, then

$$p_{i,j} = {P_b}^{H(v_i,v_j)}(1-P_b)^{D-H(v_i,v_j)}. \tag{2.7}$$

where H (v_i, v_j) is the hamming distance between v_i and v_j. The quantized data L^{su}, which is a D-bits binary codeword, is transmitted through BSC. This codeword may be changed by the channel and received at the fusion center as L^{fc}. The probability mass of the L^{fc} for all codeword combinations under hypothesis j can be calculated once the indices of the quantized codewords are aligned with that of P_T and can be computed as

$$P_j^{\text{fc}} = P_j^{\text{su}} * P_T. \tag{2.8}$$

Final performance is determined in terms of distributional divergence or Bayes error criterion.

Therefore, the main objective is to find the quantization thresholds $t_i, t_k, i, k = 1, \ldots, W-1$, and the codewords $c_i, c`_k, i, k = 1, \ldots, W$, that maximizes a divergence measure between P_0^{fc} and P_1^{fc} for the quantizer designed for optimizing a divergence metric or that minimizes the probability of error for the quantizer designed for Bayes error criterion.

2.2 QUANTIZATION SCHEMES, JOINT QUANTIZATION AND ERROR RESILIENCE

A quantization scheme is designed based on a distributional divergence (KL or Chernoff) or Bayes error criterion. A two-stage iterative algorithm is discussed below with guaranteed convergence to a local optimum scheme.

2.2.1 DISTRIBUTIONAL DIVERGENCE

Divergence metrics are used to measure the difference between two distributions and can be applied to detection problems [8]. Consider two distributions, $P_0(x)$ and $P_1(x)$. The divergence can be measured by the expectation of a function of the likelihood ratio, i.e., $E[h(\phi(x))]$, where $\phi(x) = P_1(x)/P_0(x)$. It is shown in [18] that there is no loss of generality in considering the expectation relative to P_0, compared to the expectation relative to a mixture distribution consisting of P_0 and P_1. Considering probability mass function (PMF), the KL divergence between P_0 and P_1 is given by

$$D(P_0||P_1) = \sum_i P_0(i) \log \frac{P_0(i)}{P_1(i)} = E_0 \left[-\log \frac{P_1}{P_0} \right]. \tag{2.9}$$

Steins lemma [19] shows that for a large number of SUs, the asymptotic miss probability β, for any false alarm probability α, is

$$\lim_{n \to \infty} \beta_n = e^{-nD(P_0 \,||\, P_1)}. \tag{2.10}$$

In [15], considering the tradeoff between the asymptotic values, α and β, it is shown that

$$\alpha \approx e^{-nD(P_\lambda \,||\, P_0)}$$

$$\beta \approx e^{-nD(P_\lambda \,||\, P_1)}. \tag{2.11}$$

where

$$P_\lambda(i) = \frac{P_0(i)^{1-\lambda} P_1(i)^\lambda}{\sum_j P_0(j)^{1-\lambda} P_1(j)^\lambda},\ 0 \le \lambda \le 1. \tag{2.12}$$

When λ satisfies $D(P_\lambda \,||\, P_0) = D(P_\lambda \,||\, P_1)$, the smallest detection error is achieved. It can be proved [20] that Bayes probability error $P_e = \pi_0\alpha + \pi_1\beta$ satisfies,

$$\frac{1}{N}\log P_e \le \log \sum_i P_0(i)^{1-\lambda} P_1(i)^\lambda = \log E_0\left[\left(\frac{P_1}{P_0}\right)^\lambda\right]. \tag{2.13}$$

We define $E_0\left[-\left(\frac{P_1}{P_0}\right)^\lambda\right]$ as the Chernoff metric in the sequel.
We consider a joint quantizer design based on one of these two divergence measurements. For the two PMF distributions received at the FC under different hypothesis, one optimal design is to maximize the KL divergence in (2.9) or minimize the exponent in (2.13).

2.2.2 MINIMUM ERROR PROBABILITY

Consider the binary hypothesis test $H_0 : x \sim f_0(x), H_1 : x \sim f_1(x)$. Let π_i be the prior probability of hypothesis H_i. In making a decision in any binary hypothesis testing problem, we have four possibilities to consider:

(a) H_0 is the true hypothesis, the test decides D_0;

(b) H_1 is the true hypothesis, the test decides D_1;

(c) H_0 is the true hypothesis, the test decides D_1; and

(d) H_1 is the true hypothesis, the test decides D_0.

The first two correspond to correct choices. The last two correspond to errors. In the statistical literature (c) is called a type I error and (d) is a type II error. We assume prior probabilities π_0 and π_1 are known. The probability of making an error, denoted P_e, is

$$\begin{aligned} P_e &= \Pr[H_1] \Pr[\text{Decide } H_0 | H_1] + \Pr[H_0] \Pr[\text{Decide } H_1 | H_0] \\ &= \pi_1 \Pr\left[D_0 | H_1\right] + \pi_0 \Pr\left[D_1 | H_0\right]. \end{aligned} \tag{2.14}$$

In a practical communication system, the consequences of each type of error are not equally important. The consequence of saying target is present when in fact there is none, is quite different from saying that no target is present when in fact there is. To reflect these differences, costs may be assigned to each type of error. Define C_{ij} as the cost associated with choosing hypothesis H_i when actually hypothesis H_j is true. In terms of log likelihood ratio function, the optimum Bayes rule that minimizes Bayes cost is given by [21],

$$\Lambda \underset{H_0}{\overset{H_1}{\gtrless}} \ln\left(\frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})}\right). \tag{2.15}$$

where Λ is same as equation (2.1). In communication system, it is usual to minimize the average error probability. Zero cost is associated with a correct decision, and the errors of each kind are assigned equal cost. Therefore, assume that $C_{00} = C_{11} = 0$ and $C_{10} = C_{01} = 1$ and the average cost is the average error probability P_e. For minimizing the average error probability, the decision rule becomes

$$\Lambda \underset{H_0}{\overset{H_1}{\gtrless}} \ln\left(\frac{P(H_0)}{P(H_1)}\right). \tag{2.16}$$

In the numerical computation, for minimum error probability criterion, we assume equal prior probabilities i.e., $P(H_0) = P(H_1)$. Hence, right hand side of above inequality will be equal to 0. In the next section we propose a suboptimal quantization algorithm (algorithm

3) for minimizing error probability.

2.2.3 ALGORITHMS

In our scenario, we denote P_0 and P_1 as P_0^{fc} and P_1^{fc}. Altering the local quantization thresholds and the codewords assignments, changes these two PMFs. To initiate the algorithm, we start by representing the range of LLR of the local SUs' observations before transmission as the union of a large number of non-overlapped small cells, each with a given probability mass under the mixture of two distributions, i.e., $\pi_0 f(L|H_0) + (1-\pi_0) f(L|H_1)$. Each cell will be assigned an initial codeword. Since the number of cells generally far exceeds 2^D, each of the possible codewords would be assigned to more than one cell. The algorithms for different divergence metrics and minimum error probability are summarized as follows.

Algorithm 1: Joint Quantizer and Error Resilience Design Based on Maximizing KL-Divergence

1. Initialization

   Set the rate limit, D

   Divide the range of local LLR of each SU into small intervals with a number much larger than 2^D,

   Assign initial codeword $v \in V(2^D)$ to each interval using regular scalar quantization,

   Considering correlated SU's data, combine the two local LLR value's intervals to form 2-dimension cells.

   Calculate $P_0^{\text{su}}(i,k)$ and $P_1^{\text{su}}(i,k)$ of each cell based on the codeword assignment and the probability mass using (2.5),

   $P_j^{\text{fc}} = P_j^{\text{su}} P_T \ , j = 0, 1$

   Div-old $= 0$,

   Div $= E_0 \left[-log \frac{P_1^{\text{fc}}}{P_0^{\text{fc}}} \right]$

2. Iteration Loop (person-by-person)

   While Div $-$ Divold $>$ err do

Div-old = Div,

(a) For each cell,

Update its codeword assignment if the resulting Div from (2.9) is increased (update first SU's codeword by keeping second SU's codeword constant),

End-For

(b) With the new codeword assignment:

Find P_j^{su}, and $P_j^{\text{fc}} = P_j^{\text{su}} * P_T,\ j = 0, 1,$

Calculate Div using (2.9).

Repeat (a) and (b) by keeping newly assigned first SU's codeword constant and update second SU's codeword.

Calculate Div using (2.9)

End-While

3. Post-processing

Combine adjacent cells with the same codeword assignment and output the quantizers thresholds and codewords.

Algorithm 2: Joint Quantizer and Error Resilience Design Based on Maximizing Chernoff Metric

Same as Algorithm 1 except replace the divergence with Chernoff metric.

Algorithm 3: Joint Quantizer Design Based on Minimum Error Probability.

1. Initialization

Set the rate limit, D

Divide the range of local LLR of each SU into small intervals with a number much larger than 2^D,

Assign initial codeword $v \in V(2^D)$ to each interval using regular scalar quantization,

Considering correlated SU's data, combine the two local LLR value's intervals to form

2-dimension cells.

Calculate $P_0^{\text{su}}(i,k)$ and $P_1^{\text{su}}(i,k)$ of each cell based on the codeword assignment and the probability mass using (2.5),

$P_j^{\text{fc}} = P_j^{\text{su}} * P_T \ , j = 0, 1$

Likelihood ratio Λ is formed using the PMFs at the FC using (2.16).

$P(H_0) = P(H_1)$

α and β are calculated using (2.3) and (2.4)

$P_e \,\text{old} = 0,$

P_e is calculated from (2.14).

2. Iteration Loop (Person-by-Person)

   While $|P_e - P_e \,\text{old}| > err$ do

   $P_e \,\text{old} = P_e,$

   (a) For each cell,

   Update its codeword assignment if the resulting error probability P_e using (2.14) is decreased (update first SU's codeword by keeping second SU's codeword constant),

   End-For

   (b) With the new codeword assignment:

   Find P_j^{su}, and $P_j^{\text{fc}} = P_j^{\text{su}} P_T, j = 0, 1,$

   Calculate P_e.

   Repeat (a) and (b) by keeping newly assigned first SU's codeword constant and update second SU's codeword.

   Calculate P_e

   End-While

3. Post-processing

   Combine adjacent cells with the same codeword assignment and output the quantizers

thresholds and codewords.

There are some techniques in practice that could improve these algorithms. For example, if two adjacent cells have the same codeword during several iterations, they can be combined into one cell. Since each update of codeword assignment is for one cell, while keeping the codewords of other cells fixed, the above algorithms present linear computational complexity.

2.2.4 JOINT QUANTIZATION AND ERROR RESILIENCE

As stated in algorithms, the codeword of each small cell is updated independently. Hence, it is possible that at the end, some binary sequences have never been chosen for the codeword assignment. In other words, the number of quantization intervals could be less than 2^D, although each interval is always represented by a D-bits sequence. In [16] it is shown that quantizer with less quantization intervals i.e., less than 2^D works much better than the quantizer that uses all D-bits in quantization, when channel error exists. This indicates that optimal quantizer actually uses some portion of the total D-bits ($< D$) in quantization for the purpose of data representation, while leaving the rest of the bits for the combating channel errors. LLR at the SU is quantized with a number of quantization levels less than 2^D, when channel errors are present, depending on the error probability. However, it should be noted that the number of different codewords at the FC is always 2^D. Every possible D-bits binary sequences can be obtained at FC due to channel errors. Finally, at the FC, we have two probability mass function P_0^{fc} and P_1^{fc}, with respect to all possible binary sequences of received codewords.

2.3 CENTRALIZED DETECTION WITH GAUSSIAN CHANNELS

The purpose of this analysis is to provide a performance benchmark for comparison of decentralized quantized schemes with a central scheme. Assume in centralized and un-quantized detection scheme, the FC knows everything about the practical scenario: observation in its analog domain, Gaussian noise variance in the sensing process, and the

spectral density of additive white Gaussian noise (AWGN) channel for transmission of the observation.

Suppose that $H_0 : X \sim N(-\mu_0, \sigma^2)$ and $H_1 : X \sim N(\mu_1, \sigma^2)$, with the prior probability of π_0 and $\pi_1 = 1 - \pi_0$, respectively. In the decentralized case, X is quantized into a D-bits string, which is transmitted through a BSC channel with bit error rate (BER), P_b, and further received by the FC. While in the centralized version, the un-quantized X is directly transmitted through an equivalent AWGN channel $N(0, \sigma_n^2)$. Now, the objective is to find σ_n^2 so that the FC can appropriately account for the reporting channel errors. Consider only one SU for simplicity.

For un-quantized case: The energy/symbol of X is

$$E_s = \overline{X^2} = (\mu_0^2 + \sigma^2)\pi_0 + (\mu_1^2 + \sigma^2)\pi_1. \tag{2.17}$$

For quantized case: X is quantized into D-bits, then the energy/bit is

$$E_b = \frac{E_s}{D} = \frac{\overline{X^2}}{D}. \tag{2.18}$$

With binary phase shift keying (BPSK) transmission

$$P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right). \tag{2.19}$$

where N_0 is the one-sided power spectral density (PSD), i.e, the noise power per Hz. Theoretically, two symbols can be transmitted per Hz. Then, channel noise power (i.e., variance σ_n^2) per symbol is $\sigma_n^2 = \frac{N_0}{2}$.

Combining all the above, we obtain

$$\sigma_n^2 = \frac{\overline{X^2}}{[Q^{-1}(P_b)]^2 D}. \tag{2.20}$$

To summarize, σ_n^2 is related to μ_0, μ_1, σ, priors and D. The key point is to keep the overall energy of transmitting one D-bit string in the quantized case the same as the energy of transmitting the original X in the un-quantized case.

There is one glitch in the above when using BPSK. It's known that transmitting one symbol uses channel once. In the BPSK case, each bit is a symbol. Therefore, the D-bit string uses the channel D times in total. Given the time period T_s to transmit the X and the D-bit string being the same (so as to keep the same power), BPSK expands channel bandwidth by D times, although we always have the same N_0 for each channel use. However, this issue can be easily solved by using a D-ary modulation scheme so that the D-bit string is grouped into one symbol for transmission and uses the channel only once per T_s second. In this case, (2.19) needs to be changed based on the specific modulation scheme and subsequently (2.20) needs change as well. In our work, without changing anything in simulation, we can simply update σ_n^2 to recalculate the performance of the centralized case and then we can safely say that we have transmitted the quantized data with the corresponding modulation scheme. Lets assume (X_1', X_2') be the data received at the FC instead of (X_1, X_2) after transmitting through an equivalent AWGN channel $N(0, \sigma_n^2)$ for the centralized scheme.
Likelihood ratio test (LRT) for the data received at FC for centralized scheme is,

$$\frac{(\mu_1 - \mu_0)(X_1' + X_2')}{\sigma'^2(1+\rho')} \underset{H_0}{\overset{H_1}{\gtrless}} \ln \eta. \tag{2.21}$$

where $\rho' = \dfrac{\rho\sigma^2}{\sigma^2 + \sigma_n^2}$ and $\sigma'^2 = \sigma^2 + \sigma_n^2$
Proof: See Appendix A shown below.

APPENDIX A

The proof consists of LRT derivation for bivariate Gaussian distribution. Consider the test of $H_0 : X_1, X_2 \sim N(\mu_0, \mu_0, \sigma, \sigma, \rho)$ vs. $H_1 : X_1, X_2 \sim N(\mu_1, \mu_1, \sigma, \sigma, \rho)$, the LRT is

given by,

$$\Lambda = \frac{f_1(X_1, X_2)}{f_0(X_1, X_2)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{2.22}$$

$$\Lambda = \frac{\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left[-\frac{\frac{(X_1-\mu_1)^2}{\sigma^2} - \frac{2\rho(X_1-\mu_1)(X_2-\mu_1)}{\sigma^2} + \frac{(X_2-\mu_1)^2}{\sigma^2}}{2(1-\rho^2)}\right]}{\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left[-\frac{\frac{(X_1-\mu_0)^2}{\sigma^2} - \frac{2\rho(X_1-\mu_0)(X_2-\mu_0)}{\sigma^2} + \frac{(X_2-\mu_0)^2}{\sigma^2}}{2(1-\rho^2)}\right]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

By taking natural log on both the sides for above equation, it simplifies to

$$\ln\Lambda = \left[-\frac{\frac{(X_1-\mu_1)^2}{\sigma^2} - \frac{2\rho(X_1-\mu_1)(X_2-\mu_1)}{\sigma^2} + \frac{(X_2-\mu_1)^2}{\sigma^2}}{2(1-\rho^2)}\right] + \left[-\frac{\frac{(X_1-\mu_0)^2}{\sigma^2} - \frac{2\rho(X_1-\mu_0)(X_2-\mu_0)}{\sigma^2} + \frac{(X_2-\mu_0)^2}{\sigma^2}}{2(1-\rho^2)}\right] \underset{H_0}{\overset{H_1}{\gtrless}} \ln\eta$$

After simplification, we have the equivalent test,

$$\frac{(\mu_1-\mu_0)(X_1+X_2)}{\sigma^2(1+\rho)} \underset{H_0}{\overset{H_1}{\gtrless}} ln\,\eta. \tag{2.23}$$

Hence, we consider LRT for the data received at FC for centralized scheme, assuming X_1', X_2' to be the data received at the FC after transmission through an equivalent AWGN channel $N(0, \sigma_n^2)$.

The equivalent LRT of the received data at FC reduces to

$$\frac{(\mu_1-\mu_0)(X_1'+X_2')}{\sigma'^2(1+\rho')} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\,\eta. \tag{2.24}$$

We assume covariance of the SU's data are same, before and after transmission through the

AWGN channel. Hence, $\mathrm{Cov}(X_1, X_2) = \mathrm{Cov}(X_1', X_2')$

i.e., $\rho\sqrt{\mathrm{Var}(X_1)\mathrm{Var}(X_2)} = \rho'\sqrt{\mathrm{Var}(X_1')\mathrm{Var}(X_2')}$ which is equivalent to

$\rho\sigma^2 = \rho'\sigma'^2$

where $\sigma'^2 = (\sigma^2 + \sigma_n^2)$. Therefore, $\rho' = \dfrac{\rho\sigma^2}{\sigma^2 + \sigma_n^2}$

Probability of false alarm and probability of detection can be found in terms of Q-function [21] for the LRT specified in (2.24).

$$P_f = Q\left(\frac{-2\mu_0}{\sqrt{2\,\sigma'\,(1+\rho')}}\right). \tag{2.25}$$

$$P_d = Q\left(\frac{-2\mu_1}{\sqrt{2\,\sigma'\,(1+\rho')}}\right). \tag{2.26}$$

CHAPTER 3

NUMERICAL ANALYSIS AND RESULTS

In this chapter, we present some detailed numerical results of our proposed quantization algorithms. We compare the quantizers designed for dependent SUs observation with the quantizers designed with independence assumption. Also, we analyze the FC LRT performance using ROC curve and probability of error.

3.1 KL DIVERGENCE-BASED QUANTIZER

We consider testing of mean of Gaussian as a simple model to represent the signal present or absent hypothesis. Let $H_0 : X_1, X_2 \sim N(-1, -1, \sigma^2, \sigma^2, \rho)$ and $H_1 : X_1, X_2 \sim N(1, 1, \sigma^2, \sigma^2, \rho)$, and prior probabilities $\pi_0 = \pi_1 = 0.5$, where X_1, X_2 are the observations at SU1 and SU2, respectively. This is equivalent to the case that one hypothesis has zero mean and the other has a mean 2, when the observation is added to 1. It can be seen that the normalized log-likelihood ratio L at each SU is equivalent to the observation, x. To start the algorithm, the initial codeword assignment is set as regular scalar quantization with equal intervals over $(-s, s)$, with $(-\infty, -s)$ and (s, ∞) forming the end intervals. For designing this quantizer, we considered specific cases where $\sigma = 0.5$, channel error probability $P_b = [0, 0.05]$, and different correlation coefficients ρ, which ranges from -1 to 1. For numerical computation, different values of $\rho\,[-0.9, -0.75, -0.5, 0, 0.5, 0.75, 0.9]$ and $D = [1, 2]$, are considered. In Table 3.1, we have summarized the quantizer thresholds and codeword assignments for the quantizer design for algorithm 1, when $D = 1$ and no channel error, $P_b = 0$. Table 3.2 shows the KL divergence values for the quantizer design for algorithm 1 for the same case, $D = 1$ and $P_b = 0$.

Table 3.1. Quantizer thresholds and codeword assignment for algorithm 1, when channel error probability, $P_b = 0$ and $D = 1$

$\rho = -0.9$				$\rho = -0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.5]$ $[-0.5, \infty)$	00 01	00 01	$(-\infty, -0.5]$ $[-0.5, \infty)$	$(-\infty, -0.5]$ $[-0.5, \infty)$	00 01	00 01	$(-\infty, -0.5]$ $[-0.5, \infty)$
$\rho = -0.5$				$\rho = 0$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.5]$ $[-0.5, \infty)$	00 01	00 01	$(-\infty, -0.5]$ $[-0.5, \infty)$	$(-\infty, -0.6]$ $[-0.6, \infty)$	00 01	00 01	$(-\infty, -0.6]$ $[-0.6, \infty)$
$\rho = 0.5$				$\rho = 0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.7]$ $[-0.7, \infty)$	00 01	00 01	$(-\infty, -0.7]$ $[-0.7, \infty)$	$(-\infty, -0.7]$ $[-0.7, \infty)$	00 01	00 01	$(-\infty, -0.7]$ $[-0.7, \infty)$
$\rho = 0.9$							
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.				
$(-\infty, -1.2]$ $[-1.2, 0]$ $[0, \infty)$	00 01 00	00 01	$(-\infty, -0.7]$ $[-0.7, \infty)$				

Table 3.2. KL divergence values for the quantizer designed using algorithm 1 with channel error probability, $P_b = 0$ and $D = 1$

ρ	-0.9	-0.75	-0.5	0	0.5	0.75	0.9
KL-DIV	68.037	30.227	17.226	10.447	7.730	6.675	6.972

In Table 3.1, we observe that as the ρ tends towards total positive correlation (+1), quantization intervals tend to be non-contiguous [22] i.e., quantizer intervals with the same codeword assignment are not adjacent to each other. In the same table (Table 3.1), quantizer thresholds for SU1 and $\rho = 0.9$, codeword 00 is assigned for two non-adjacent intervals i.e., $(-\infty, -1.2]$ and $[0, \infty)$. Table 3.2 shows that KL divergence value for the quantizer designed for negative correlation ($\rho = -0.9$) is maximum whereas that of the quantizer designed for positive correlation ($\rho = 0.9$) is minimum. Hence, negative correlation gives better

performance, although physically it is not conceivable to expect negative correlation. In Tables 3.3 and 3.4, we have summarized the quantizer thresholds with codewords assignment and KL divergence values for the quantizer design based on algorithm 1, when $D = 1$ and channel error probability, $P_b = 0.05$.

Table 3.3. Quantizer thresholds and codeword assignment for algorithm 1, when channel error probability, $P_b = 0.05$ and $D = 1$

$\rho = -0.9$				$\rho = -0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.1]$ $[-0.1, \infty)$	00 01	00 01	$(-\infty, -0.1]$ $[-0.1, \infty)$	$(-\infty, -0.1]$ $[-0.1, \infty)$	00 01	00 01	$(-\infty, -0.1]$ $[-0.1, \infty)$
$\rho = -0.5$				$\rho = 0$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.1]$ $[-0.1, \infty)$	00 01	00 01	$(-\infty, -0.1]$ $[-0.1, \infty)$	$(-\infty, -0.1]$ $[-0.1, \infty)$	00 01	00 01	$(-\infty, -0.1]$ $[-0.1, \infty)$
$\rho = 0.5$				$\rho = 0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.2]$ $[-0.2, \infty)$	00 01	00 01	$(-\infty, -0.2]$ $[-0.2, \infty)$	$(-\infty, -0.3]$ $[-0.3, \infty)$	00 01	00 01	$(-\infty, -0.3]$ $[-0.3, \infty)$
$\rho = 0.9$							
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.				
$(-\infty, -0.4]$ $[-0.4, \infty)$	00 01	00 01	$(-\infty, -0.2]$ $[-0.2, \infty)$				

Table 3.4. KL divergence values for the quantizer designed using algorithm 1 for channel error probability, $P_b = 0.05$ and $D = 1$

ρ	-0.9	-0.75	-0.5	0	0.5	0.75	0.9
KL-DIV	4.804	4.804	4.804	4.768	4.581	4.394	4.247

We tabulated the quantizer thresholds and KL divergence values for the quantizer designed by algorithm 1 for $P_b = 0$ and $D = 2$ in Tables 3.5 and 3.6, respectively. With a larger D value ($D = 2$), observation similar to $D = 1$ case is observed, that is when ρ tends

Table 3.5. Quantizer thresholds and codeword assignment for algorithm 1, when channel error probability, $P_b = 0$ and $D = 2$

$\rho = -0.9$				$\rho = -0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -1.1]$	00	00	$(-\infty, -1.1]$	$(-\infty, -1.1]$	00	00	$(-\infty, -1.1]$
$[-1.1, -0.6]$	01	01	$[-1.1, -0.6]$	$[-1.1, -0.6]$	01	01	$[-1.1, -0.6]$
$[-0.6, 0]$	10	10	$[-0.6, 0]$	$[-0.6, 0]$	10	10	$[-0.6, 0]$
$[0, \infty)$	11	11	$[0, \infty)$	$[0, \infty)$	11	11	$[0, \infty)$
$\rho = -0.5$				$\rho = 0$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -1.2]$	00	00	$(-\infty, -1.1]$	$(-\infty, -1.2]$	00	00	$(-\infty, -1.2]$
$[-1.2, -0.7]$	01	01	$[-1.1, -0.6]$	$[-1.2, -0.7]$	01	01	$[-1.2, -0.7]$
$[-0.7, -0.1]$	10	10	$[-0.6, 0]$	$[-0.7, -0.2]$	10	10	$[-0.7, -0.2]$
$[-0.1, \infty)$	11	11	$[0, \infty)$	$[-0.2, \infty)$	11	11	$[-0.2, \infty)$
$\rho = 0.5$				$\rho = 0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -1.1]$	00	00	$(-\infty, -1.1]$	$(-\infty, -1.1]$	00	00	$(-\infty, -1.1]$
$[-1.1, -0.6]$	01	01	$[-1.1, -0.6]$	$[-1.1, -0.6]$	01	01	$[-1.1, -0.6]$
$[-0.6, 0]$	10	10	$[-0.6, 0]$	$[-0.6, 0]$	10	10	$[-0.6, 0]$
$[0, \infty)$	11	11	$[0, \infty)$	$[0, \infty)$	11	11	$[0, \infty)$

$\rho = 0.9$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -1.7]$	00	00	$(-\infty, -1.1]$
$[-1.7, -1.4]$	10	01	$[-1.1, -0.7]$
$[-1.4, -0.9]$	11	11	$[-0.7, -0.3]$
$[-0.9, -0.4]$	01	10	$[-0.3, \infty)$
$[-0.4, 0]$	10		
$[0, 0.5]$	00		
$[0.5, \infty)$	11		

Table 3.6. KL divergence for the quantizer designed using algorithm 1 for channel error probability, $P_b = 0$ and $D = 2$

ρ	-0.9	-0.75	-0.5	0	0.5	0.75	0.9
KL-DIV	116.557	49.2706	26.239	14.139	9.7882	8.454	7.957

towards total positive correlation (+1), quantization intervals tend to be non-contiguous i.e., it results in multiple non-adjacent intervals with the same codeword assignment. Quantizer

thresholds for SU1 and $\rho = 0.9$ in Table 3.5 shows that, codeword 00 is assigned for two non-adjacent intervals, that is, $(-\infty, -1.7]$ and $[0, 0.5]$. As before, Table 3.6 shows that KL divergence of the quantizer designed for negative correlation is maximum.

Table 3.7. Quantizer thresholds and codeword assignment for algorithm 1, when channel error probability, $P_b = 0.05$ and $D = 2$

$\rho = -0.9$				$\rho = -0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.4]$	01	01	$(-\infty, -0.4]$	$(-\infty, -0.4]$	01	01	$(-\infty, -0.4]$
$[-0.4, 0]$	00	00	$[-0.4, 0]$	$[-0.4, 0]$	00	00	$[-0.4, 0]$
$[0, \infty)$	10	10	$[0, \infty)$	$[0, \infty)$	10	10	$[0, \infty)$
$\rho = -0.5$				$\rho = 0$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.4]$	01	01	$(-\infty, -0.4]$	$(-\infty, -0.5]$	01	01	$(-\infty, -0.5]$
$[-0.4, 0]$	00	00	$[-0.4, 0]$	$[-0.5, -0.1]$	00	00	$[-0.5, -0.1]$
$[0, \infty)$	10	10	$[0, \infty)$	$[-0.1, \infty)$	10	10	$[-0.1, \infty)$
$\rho = 0.5$				$\rho = 0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.6]$	01	01	$(-\infty, -0.6]$	$(-\infty, -0.7]$	01	01	$(-\infty, -0.7]$
$[-0.6, -0.2]$	00	00	$[-0.6, -0.2]$	$[-0.7, -0.2]$	00	00	$[-0.7, -0.2]$
$[-0.2, \infty)$	10	10	$[-0.2, \infty)$	$[-0.2, \infty)$	10	10	$[-0.2, \infty)$
$\rho = 0.9$							
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.				
$(-\infty, -0.8]$	01	01	$(-\infty, -0.7]$				
$[-0.8, -0.3]$	00	00	$[-0.7, -0.1]$				
$[-0.3, \infty)$	10	10	$[-0.1, \infty)$				

Table 3.8. KL divergence for the quantizer designed using algorithm 1 for channel error probability, $P_b = 0.05$ and $D = 2$

ρ	-0.9	-0.75	-0.5	0	0.5	0.75	0.9
KL-DIV	9.0848	9.0847	9.0836	8.8328	7.6769	7.0017	6.6248

In Table 3.7 and 3.8, we have outlined the quantizer thresholds with codeword assignments and KL divergence values for the quantizer design based on algorithm 1, when $D = 2$ and

channel error probability, $P_b = 0.05$. Table 3.8 shows the KL divergence values for the quantizer design based on algorithm 1 when $D = 2$ and $P_b = 0.05$. We can observe in Table 3.7, that when channel errors exist, the quantization scheme provides for error resilience. It can be seen that the codeword 11 is not used for data representation, but is left unassigned for combating channel errors.

The curve of KL divergence as a function of channel error probability with correlation coefficient as a parameter is shown in Figure 3.1. Since BEP in practical systems is generally less than 0.1, P_b range of $[0, 0.1]$ is used in Figure 3.1. We observe that with the increase of channel error probability, the KL divergence value decreases. This is to be expected. In Figure 3.2, for $D = 2$ and $\sigma = 0.5$, we have plotted KL divergence for the quantizers designed, when $P_b = [0, 0.05, 0.1]$ and for different correlation coefficients. Both Figures 3.1 and 3.2 show that negative correlation is beneficial, i.e., increases KL divergence.

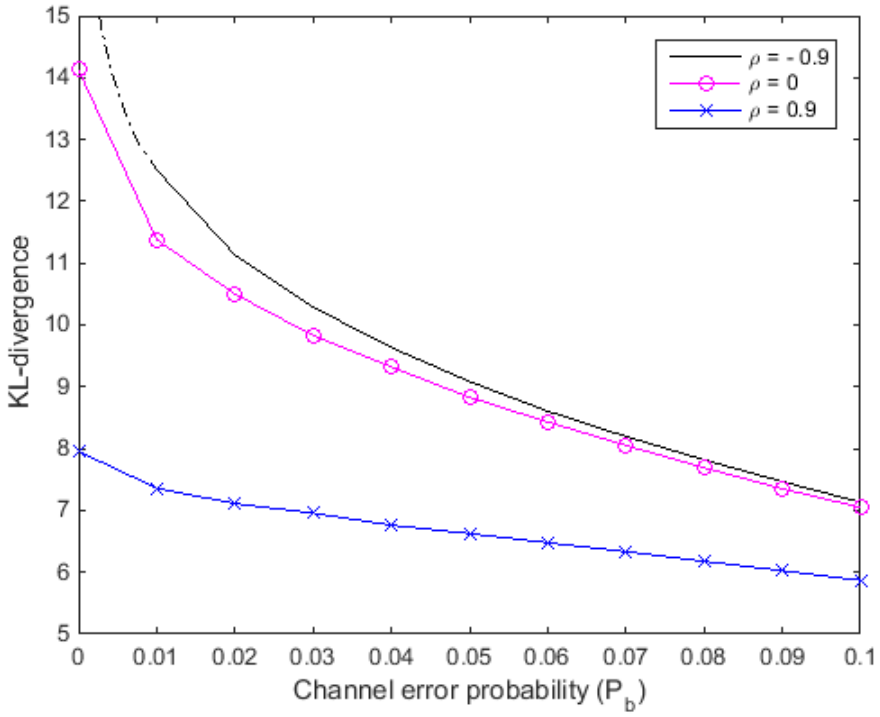

Figure 3.1. The variation of KL divergence as a function of channel error probability (P_b) with correlation coefficient (ρ) as a parameter, when $D = 2$

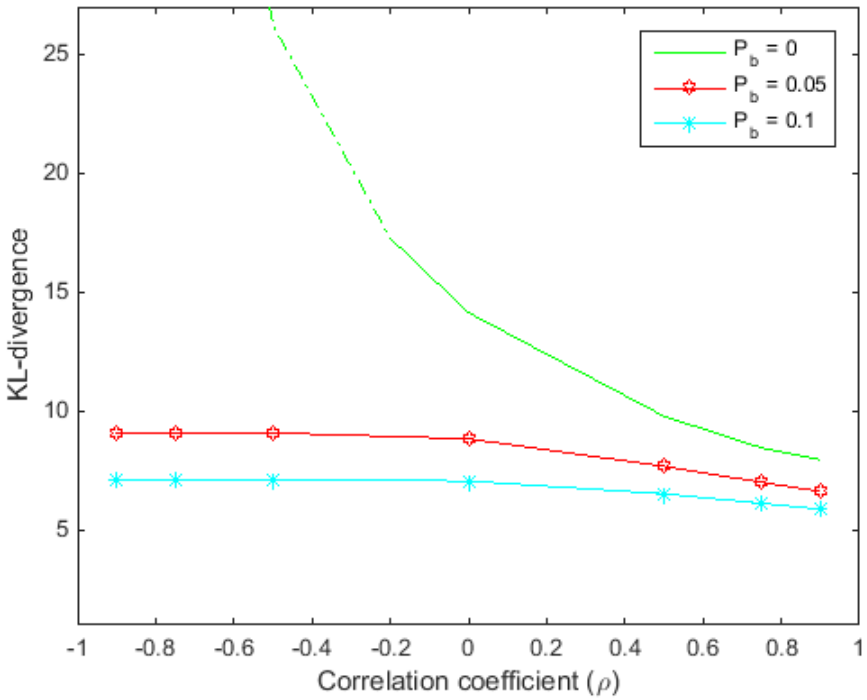

Figure 3.2. The variation of KL divergence as a function of correlation coefficient (ρ) for different channel error probability (P_b) when $D = 2$.

3.2 CHERNOFF METRIC-BASED QUANTIZER

With the same parameter setup used in the design of KL divergence quantizer, we now consider the quantizers based on algorithm 2. We have summarized the quantizer thresholds and codeword assignments for algorithm 2, when $D = 2$ and no channel error, $P_b = 0$ in Table 3.9. Table 3.10 shows the negative Chernoff metric values for the quantizer design using algorithm 2, when $D = 2$ and $P_b = 0$.

We observe that the results are similar to that of the quantizer design based on algorithm 1. As ρ tends towards total positive correlation (+1), quantization intervals tend to be non-contiguous i.e., quantizer intervals with same codeword assignment are not adjacent to each other. In Table 3.9, quantizer thresholds for SU1 and $\rho = 0.9$ shows that codeword 10 is assigned for two non-adjacent intervals i.e., $(-\infty, -1.1]$ and $[0.8, 1.4]$. Table 3.10 shows that Chernoff metric value of the quantizer designed for negative correlation is larger than

that of the quantizer designed for independent SUs (i.e. $\rho = 0$). Figure 3.3 shows the

Table 3.9. Quantizer thresholds and codeword assignment for algorithm 2, when channel error probability, $P_b = 0$ and $D = 2$

$\rho = -0.9$				$\rho = -0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.9]$	00	00	$(-\infty, -0.9]$	$(-\infty, -0.7]$	00	00	$(-\infty, -0.7]$
$[-0.9, 0.4]$	01	01	$[-0.9, 0.4]$	$[-0.7, 0.2]$	01	01	$[-0.7, 0.2]$
$[0.4, 1.2]$	10	10	$[0.4, 1.2]$	$[0.2, 1.2]$	10	10	$[0.2, 1.2]$
$[1.2, \infty)$	11	11	$[1.2, \infty)$	$[1.2, \infty)$	11	11	$[1.2, \infty)$
$\rho = -0.5$				$\rho = 0$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.6]$	00	00	$(-\infty, -0.6]$	$(-\infty, -0.5]$	00	00	$(-\infty, -0.5]$
$[-0.6, 0.1]$	01	01	$[-0.6, 0.1]$	$[-0.5, 0]$	01	01	$[-0.5, 0]$
$[0.1, 0.8]$	10	10	$[0.1, 0.8]$	$[0, 0.5]$	10	10	$[0, 0.5]$
$[0.8, \infty)$	11	11	$[0.8, \infty)$	$[0.5, \infty)$	11	11	$[0.5, \infty)$
$\rho = 0.5$				$\rho = 0.75$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.	SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -0.5]$	00	00	$(-\infty, -0.5]$	$(-\infty, -0.4]$	00	00	$(-\infty, -0.8]$
$[-0.5, 0]$	01	01	$[-0.5, 0]$	$[-0.4, 0.1]$	01	11	$[-0.8, -0.2]$
$[0, 0.5]$	10	10	$[0, 0.5]$	$[0.1, 0.6]$	10	01	$[-0.2, 0.3]$
$[0.5, \infty)$	11	11	$[0.5, \infty)$	$[0.6, \infty)$	11	10	$[0.3, 1.2)$
						00	$[1.2, \infty)$

$\rho = 0.9$			
SU1 thresh.	SU1 label	SU2 label	SU2 thresh.
$(-\infty, -1.1]$	10	00	$(-\infty, -0.3]$
$[-1.1, -0.8]$	11	01	$[-0.3, 0]$
$[-0.8, -0.5]$	00	10	$[0, 0.8]$
$[-0.5, 0.2]$	01	11	$[0.8, \infty)$
$[0.2, 0.5]$	00		
$[0.5, 0.8]$	11		
$[0.8, 1.4]$	10		
$[1.4, \infty)$	01		

curve of negative Chernoff metric as a function of channel error probability, with correlation coefficient, ρ, as a parameter, when $D = 2$. As expected, the performance corresponding to lower channel error probability is better that of the case of higher error. Figure 3.4 shows

Table 3.10. Negative Chernoff metric values for the quantizer designed using algorithm 2 for channel error probability, $P_b = 0$ and $D = 2$

ρ	-0.9	-0.75	-0.5	0	0.5	0.75	0.9
Chernoff metric	-8.59E-06	-2.48E-04	-2.92E-03	-3.07E-02	-8.83E-02	-1.22E-01	-1.39E-01

Chernoff metric values with respect to correlation coefficients ρ. Both Figures 3.3 and 3.4 show that Chernoff values decrease with increasing correlation coefficient.

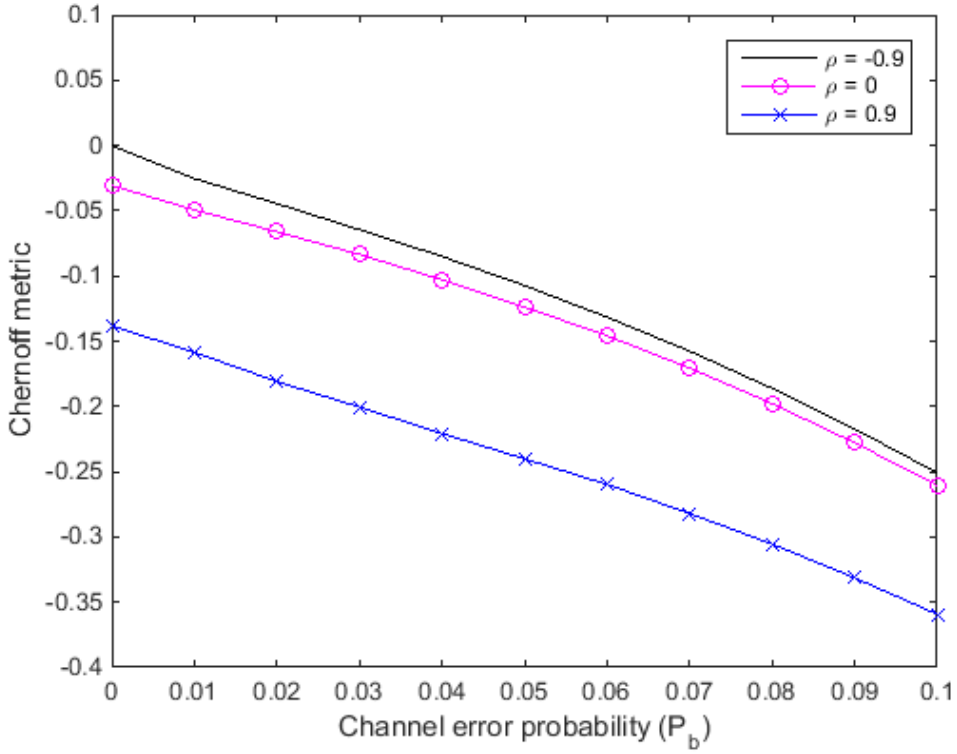

Figure 3.3. The variation of Chernoff metric as a function of channel error probability (P_b) with correlation coefficient (ρ) as a parameter when $D = 2$

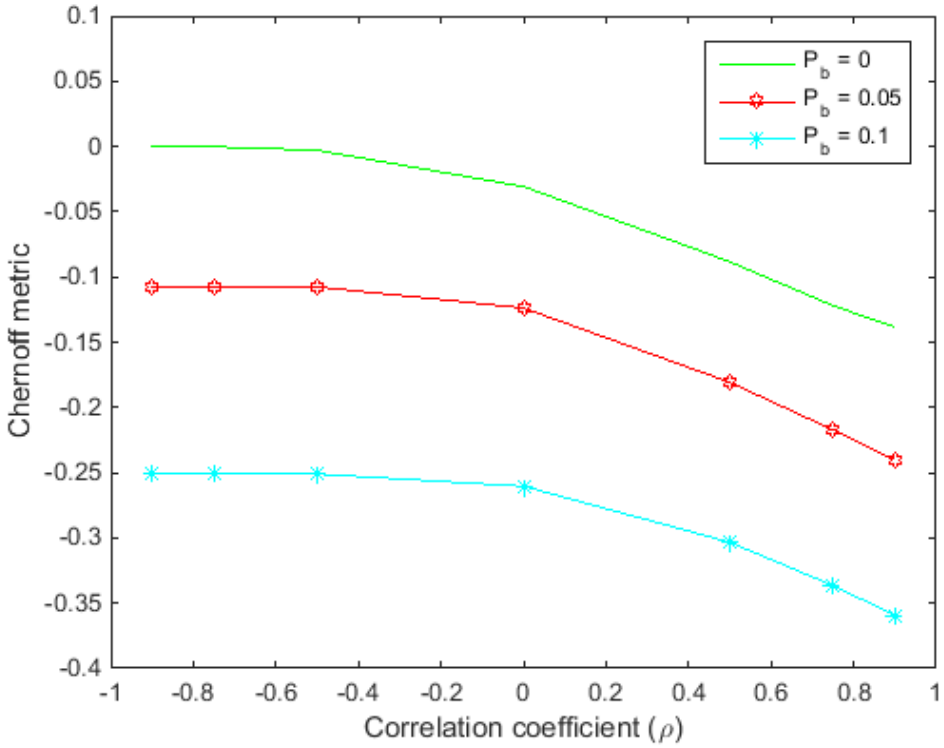

Figure 3.4. The variation of Chernoff metric as a function of correlation coefficient (ρ) for different channel error probability (P_b) when $D = 2$

3.3 DETECTOR PERFORMANCE ANALYSIS USING ROC CURVE

A detector's performance is measured by its ability to achieve a certain probability of detection and probability of false alarm for a given sensor observation signal-to-noise ratio (SNR). Typically, examination of a detector's receiver operating characteristic (ROC) curves provides insight into its performance.

We plotted the ROC curve for the quantizer based on algorithm 1 designed with the assumption of independent SUs data and ROC curve of the quantizer based on algorithm 1 with the actual dependent SUs data. In the independent assumption case also, the probability mass functions for the sensors' observations were calculated using the true bivariate joint density, although the quantizers were designed based on independence assumption. Probability of false alarm and probability of detection are calculated using (2.3) and (2.4). Figures

$3.5 - 3.10$ show the plots of ROC for the cases mentioned below. In order to get more clarity on behavior of the system for lower probability of false alarm, we plotted logarithmic value of probability of false alarm on x-axis and probability of detection on y-axis.

We perform numerical computation for each sensing SNR, with distinct mean and standard deviation values, correlation coefficients, $\rho = [-0.75, -0.5, 0, 0.5, 0.75]$ and channel error probabilities, BER $(P_b) = [0, 0.05, 0.1]$. Numerical values considered are listed below. However, we show the graph of one case for each scenario, as the behavior happens to be similar.

High sensing SNR

(i) $H_0 \sim N(-1, -1, 0.5, 0.5, \rho), H_1 \sim N(1, 1, 0.5, 0.5, \rho)$ (Figure 3.5 and 3.6).

(ii) $H_0 \sim N(-2, -2, 1, 1, \rho), H_1 \sim N(2, 2, 1, 1, \rho)$.

Average sensing SNR

(i) $H_0 \sim N(-0.5, -0.5, 0.5, 0.5, \rho), H_1 \sim N(0.5, 0.5, 0.5, 0.5, \rho)$ (Figure 3.7 and 3.8).

(ii) $H_0 \sim N(-1, -1, 1, 1, \rho), H_1 \sim N(1, 1, 1, 1, \rho)$.

Low sensing SNR

(i) $H_0 \sim N(-1, -1, 2, 2, \rho), H_1 \sim N(1, 1, 2, 2, \rho)$.

(ii) $H_0 \sim N(-0.5, -0.5, 1, 1, \rho), H_1 \sim N(0.5, 0.5, 1, 1, \rho)$ (Figure 3.9 and 3.10)

In the following figures of ROC plot, the terms Quantizer-KL-1 and Quantizer-KL-2 are defined, follows:

Quantizer-KL-1: Quantizer based on algorithm 1 with correlated SU information.

Quantizer-KL-2: Quantizer based on algorithm 1 with independent SU information.

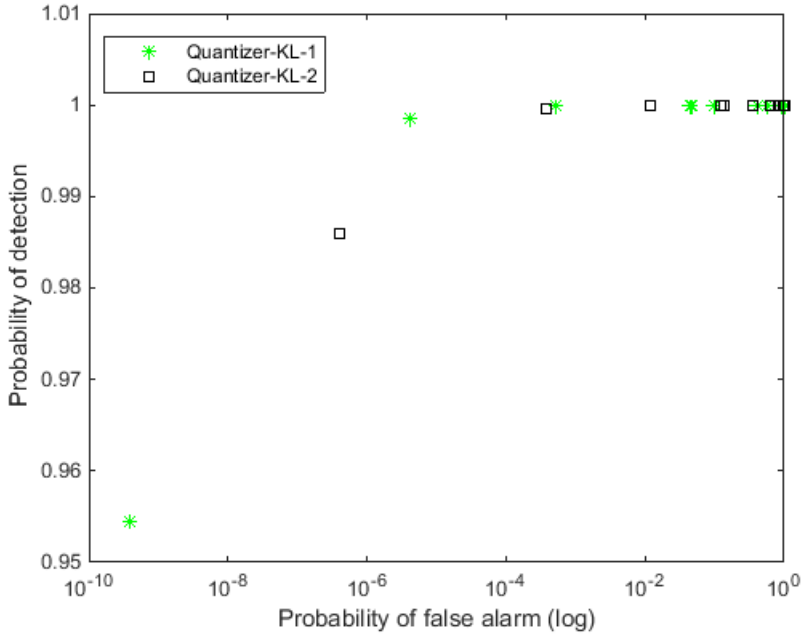

Figure 3.5. ROC comparison between different KL divergence quantizers for channel error probability, $P_b = 0$ and $D = 2$. $H_0 \sim N(-1, -1, 0.5, 0.5, -0.75), H_1 \sim N(1, 1, 0.5, 0.5, -0.75)$

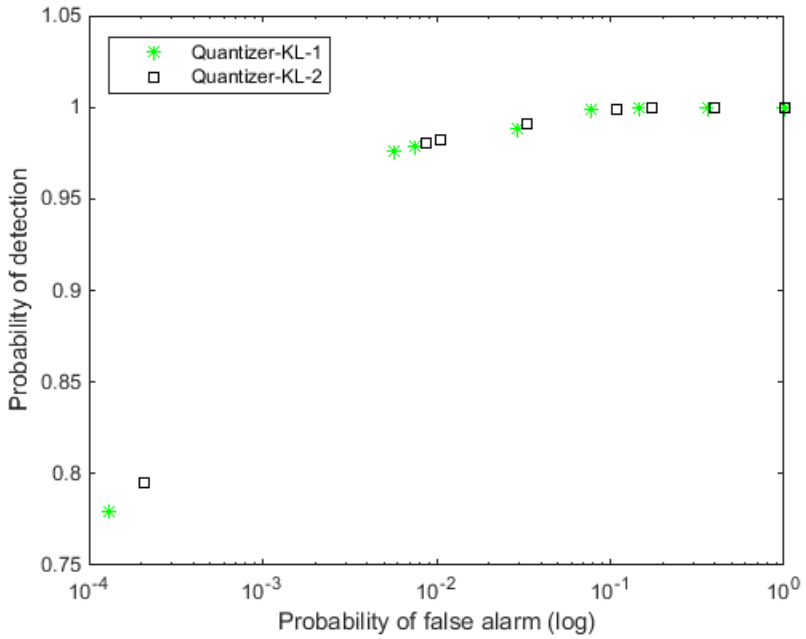

Figure 3.6. ROC comparison between different KL divergence quantizers for channel error probability, $P_b = 0.05$ and $D = 2$. $H_0 \sim N(-1, -1, 0.5, 0.5, -0.75), H_1 \sim N(1, 1, 0.5, 0.5, -0.75)$

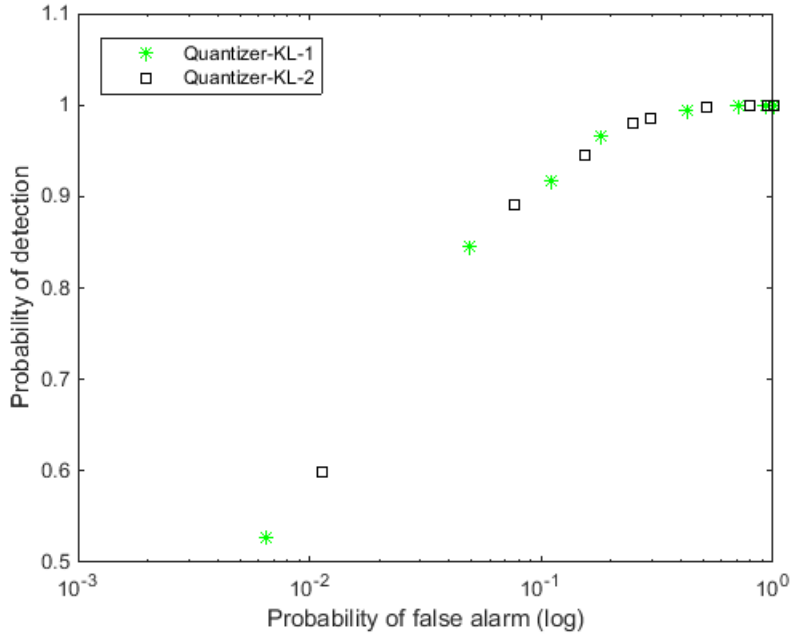

Figure 3.7. ROC comparison between different KL divergence quantizers for channel error probability ($P_b = 0$) and $D = 2$. $H_0 \sim N(-0.5, -0.5, 0.5, 0.5, 0), H_1 \sim N(0.5, 0.5, 0.5, 0.5, 0)$

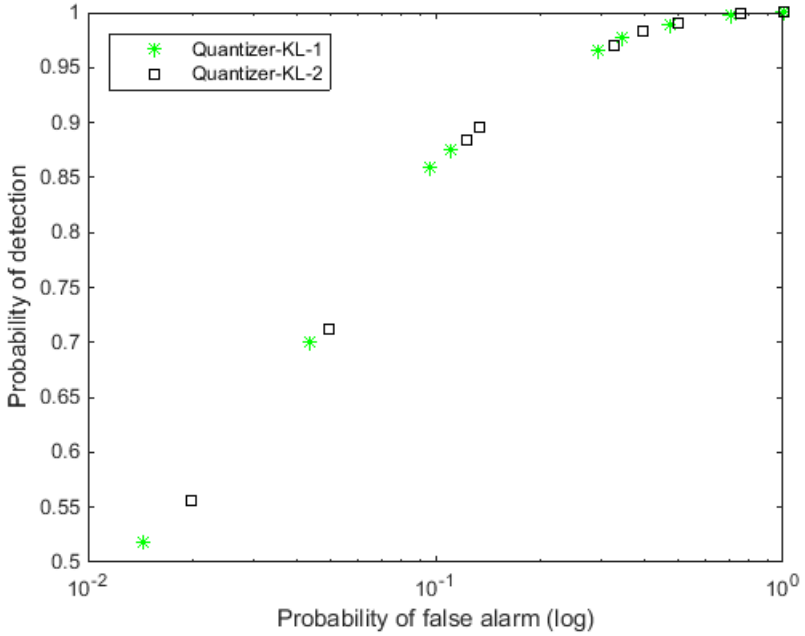

Figure 3.8. ROC comparison between different KL divergence quantizers for channel error probability, $P_b = 0.05$ and $D = 2$. $H_0 \sim N(-0.5, -0.5, 0.5, 0.5, 0), H_1 \sim N(0.5, 0.5, 0.5, 0.5, 0)$

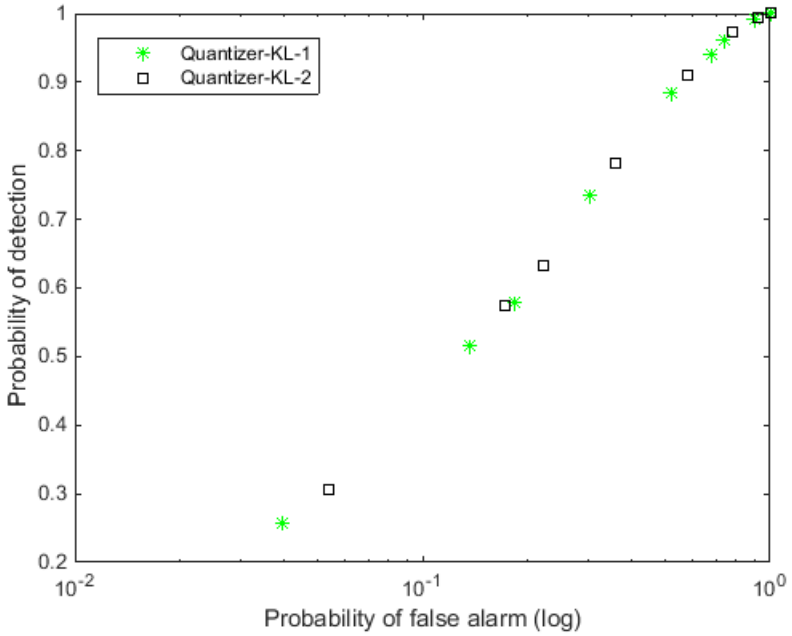

Figure 3.9. ROC comparison between different KL divergence quantizers for channel error probability, $P_b = 0$ and $D = 2$. $H_0 \sim N(-0.5, -0.5, 1, 1, 0.5), H_1 \sim N(0.5, 0.5, 1, 1, 0.5)$

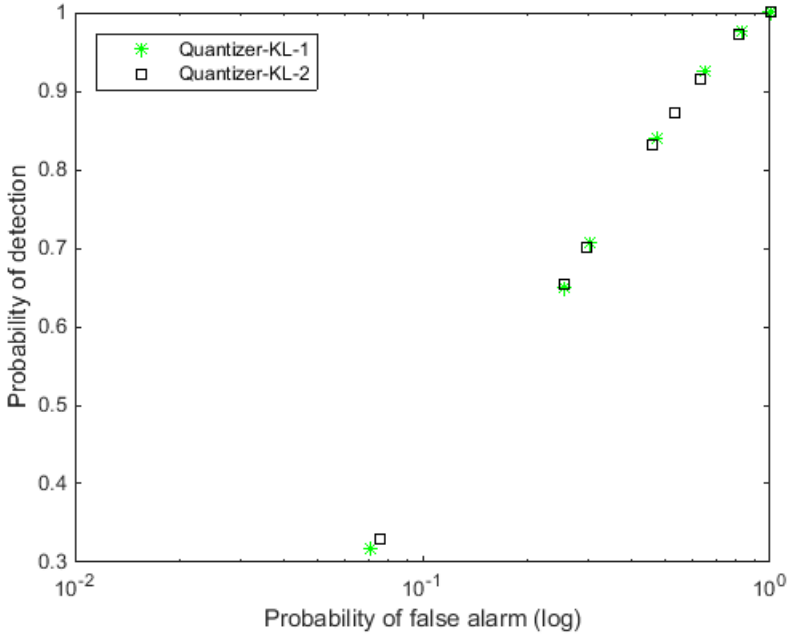

Figure 3.10. ROC comparison between different KL divergence quantizers for channel error probability, $P_b = 0.05$ and $D = 2$. $H_0 \sim N(-0.5, -0.5, 1, 1, 0.5), H_1 \sim N(0.5, 0.5, 1, 1, 0.5)$

As the SU's observations are discrete in nature, comparing the performances of the quantizers based on algorithm 1, with correlated and independence assumptions at a particular α value was not obtainable, unless one uses randomization. Using randomization, although considering correlation of SU's data improved divergence compared to that of independence assumption, it turns out that there was not much to gain in ROC behavior. The non-symmetric nature of KL divergence could have also contributed to this behavior.

3.4 BAYES ERROR BASED QUANTIZER AND PROBABILITY OF ERROR

In this section we compute probability of error for the quantizer based on algorithm 3 and compare it with the probability of error obtained with Chernoff based quantizer and the centralized test. Probability of error was calculated using (2.14), α, and β from (2.3) and (2.4), respectively. We present computation results of probability of error for different quantizers by considering all the cases mentioned in the ROC curve analysis section. Because of the space constraint, we tabulated the result for the scenario $H_0 \sim N(-1,-1,0.5,0.5,\rho), H_1 \sim N(1,1,0.5,0.5,\rho)$, $\rho = [-09,-0.5,0,0.5,0.9]$ and BER $(P_b) = [0, 0.05]$ only. In the quantizer column of Table 3.11, Corr. Bayes represents the quantizer designed using algorithm 3 for dependent SU's data, Corr. Chernoff represents the quantizer designed using algorithm 2 for dependent SU's data and Ind. Chernoff represents the quantizer designed using algorithm 2 assuming as though SU's observations were independent.

As to be expected, probability of error is minimum for the centralized test. The error rate for centralized test is several orders less than the error rate of others, especially for negative correlation coefficients. For positive correlation, the performance improvement for centralized test over others is marginal. Excluding the centralized test, Table 3.11 shows that P_e is minimum for the quantizer based on Bayes error. In the case of Chernoff metric based quantizer, the probability of error for the quantizer design with the true correlated sensor observation model is less than that of the quantizer designed as though the SUs data

Table 3.11. Comparison of probability of error for different quantizers, $H_0 \sim N(-1,-1,0.5,0.5,\rho)$ and $H_1 \sim N(1,1,0.5,0.5,\rho)$. BER $(P_b) = [0, 0.05]$ and $D = 2$.

ρ	Quantizer	BER $(P_b) = 0$	CT	BER $(P_b) = 0.05$	CT
-0.9	Corr. Bayes Corr. Chernoff Ind. Chernoff	2.569E-09 4.201E-07 9.676E-04	9.360E-20	0.019896 0.019970 0.020085	0.000081
-0.5	Corr. Bayes Corr. Chernoff Ind. Chernoff	0.000144 0.000299 0.001150	3.167E-05	0.019923 0.019990 0.020101	0.001965
0	Corr. Bayes Corr. Chernoff Ind. Chernoff	0.003091 0.004929 0.004957	0.002339	0.021158 0.021330 0.020085	0.009662
0.5	Corr. Bayes Corr. Chernoff Ind. Chernoff	0.011221 0.014156 0.014562	0.010461	0.024445 0.024553 0.024643	0.021730
0.9	Corr. Bayes Corr. Chernoff Ind. Chernoff	0.020285 0.022750 0.022436	0.020087	0.032803 0.038171 0.039504	0.032857

were independent. In fact, the difference in error probabilities is significant in some data points. Hence, for obtaining better performance at the fusion center, it is not possible to ignore correlation, especially when the correlation is significant.

CHAPTER 4

CONCLUSION AND FUTURE RESEARCH

In this thesis, we have studied decentralized signal detection with correlated observations with a focus on how correlation affects the soft decision fusion with channel errors. The problem studied here is put in the context of cooperative spectrum sensing by secondary users in a cognitive radio system. The work shows that when channel errors exist, there is a trade off in using the available transmission resource, i.e., trade the limited number of bits, between the sensor data representation and error resilience.

The design of quantizer for SU observation was based on three different criteria, namely, the KL divergence, Chernoff metric and Bayes error. Since every Ali-and-Silvey divergence metric must have a convex function kernel, design under other divergence metrics may also follow the results presented here. However, we must point out that using divergence metrics is reasonable with very large number of SUs (asymptotic assumption). Also, we assumed the same statistical model in sensing and transmission for all SUs. Study of the fusion performance for a limited number of SUs with heterogeneous sensing and transmission statistics is interesting and may require different optimization rules. In addition, the current quantizer design is still a sub-optimal process. Investigation of the local optimal regions for the codeword assignment could be an interesting work in the future which might require a study of the subspace spanned by the resultant quantizer codewords.

One more interesting issue is to look at the relationship between the number of bits (D) and the detection performance, given a modulation scheme and a quantization scheme. Let $\hat{X}$ be the quantized value of X. Since a bit error can vastly change the value of $\hat{X}$ due to its error position, the detection performance based on the received D-bits string might

not be able to approach the centralized performance, which only depends on the value of $X + N(0, \sigma_n^2)$. This asymptotic performance gap from the centralized case needs further investigation.

Another extension is to consider a practical scenario of known primary user location with two secondary users placed at random points within a specified geometrical region. Assuming a path loss model and energy detectors at the SUs, appropriate joint probability distributions under the signal and no-signal hypothesis can be formulated. The specific locations of the SUs may dictate non-identical marginal probability distributions at the two sensors, thus providing a generalization to the identical marginal case studied in the thesis.

BIBLIOGRAPHY

[1] I.F. Akyildiz, W.Y. Lee, K.R. Chowdhury, "*CRAHNs: Cognitive radio ad hoc networks*", Ad Hoc Networks 7 (5) (2009) 810 − 836.

[2] J. Ma, G. Li, B.H. Juang, "Signal processing in cognitive radio", *Proceedings of the IEEE*, vol. 97 (5) (2009) 805 − 823.

[3] J. Mitola III and G. Q. Maguire, Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications Magazine*, vol. 6, no. 4, pp. 1318, Aug. 1999.

[4] J. Mitola III, "*Cognitive radio: An integrated agent architecture for software defined radio*," Doctor of Technology dissertation, Royal Inst. Technol. (KTH), Stockholm, Sweden, 2000.

[5] B. Ahsant and R. Viswanathan, "*A review of cooperative spectrum sensing in cognitive radios*," vol. 1, 2013, pp 69 − 80, Advancement in Sensing Technology Smart Sensors, Measurement and Instrumentation, Springer-Verlag.

[6] J Unnikrishnan and V.V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio", *IEEE Journal of selected topics in signal processing*, vol. 2, No. 1, Feb. 2008.

[7] S.Mishra, A.Sahai, and R. Brodersen, "Cooperative sensing among cognitive radios," presented at *IEEE Int. Conf. Commun.*, Istanbul, Turkey, June 2006.

[8] R. Viswanathan and P.K. Varshney, "Distributed detection with multiple sensors: Part I - fundamentals", *Proceedings of the IEEE*, Vol. 85, no. 1, Jan 1997.

[9] G. Ganesan and Y. Li, "*Cooperative spectrum sensing in cognitive radio networks*," in Proc. Dynamic Spectrum Access Nets., Baltimore, MD, 2005, pp. 137 − 143.

[10] E. Drakopoulos and C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593 − 606, Jul. 1991.

[11] C. Lee and J. Chao, "Optimum local decision space partitioning for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 536 − 544, Jul. 1989.

[12] E. Visotsky, S. Kuffner, and R. Peterson, "*On collaborative detection of TV transmissions in support of dynamic spectrum sharing*," presented at the Int. Conf. Dynam. Spectrum Access Net. Baltimore, MD, Nov. 2005.

[13] M. Mustonen, M. Matinmikko, and A. Mammela, "*Cooperative spectrum sensing using quantized soft decision combining*," presented at the 4$^{\text{th}}$ Ann. Conf. Cognitive Radio Oriented Wireless Netw. Commun., Hannover, Germany, Jun. 2009.

[14] S. Chaudari, J. Lunden, V. Koivunen and H. Vincent Poor, "Cooperative sensing with imperfect reporting channels: hard decisions or soft decisions?," *IEEE Trans. Signal Processing*, vol. 60, no. 2, Jan. 2012.

[15] H. V. Poor, "Fine quantization in signal detection and estimation," *IEEE Trans. On Information Theory*, vol. 34, no. 5, pp. 960 – 972, Sept. 1988.

[16] L. Cao and R. Viswanathan, "Divergence-based soft decision for error resilient decentralized signal detection", *IEEE Trans. on Signal processing*, vol. 62, no. 19, Oct. 2014.

[17] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 773 – 787, Nov. 1987.

[18] S. M. Ali and S. D. Silvey, "*A general class of coefficients of divergence of one distribution from another*", J. Roy. Statistic. Soc. B (Methodology.), vol. 28, no. 1, pp. 131 – 142, 1966.

[19] A. Dembo and O. Zeitouni, "*Large deviations techniques and applications*", 2nd ed. New York, NY, USA: Springer, 1998.

[20] T. M. Cover and J. A. Thomas, "*Elements of information theory*," 2nd ed. New York, NY, USA: Wiley, 2006.

[21] M.D. Srinath, P.K. Rajasekaran and R. Viswanathan, "*Introduction to Statistical signal processing with applications*", Prentice-Hall, Inc. Upper Saddle River, NJ, USA 1995

[22] P. Willett, P. Swaszek and R. Blum, "The good, bad, and ugly: distributed detection of a known signal in dependent Gaussian noise", *IEEE Trans. on signal processing*, vol. 48, no. 12, December 2000

VITA

Born in Bangalore, India in 1987, Roopashree Rajanna received the B.S. degree in Electronics and communication engineering from Visvesvaraya Technological University, India in June, 2009. Starting August 2013, she is working towards the M.S. degree in electrical engineering at The University of Mississippi, University, MS. Earlier she worked as a NSS Engineer for Nokia Siemens Networks (Currently Nokia Solutions Network) in Bangalore, India. During her M.S. program, she worked first as a graduate assistant for the school of engineering Dean's office. Then worked as a teaching assistant and a research assistant for Dr. Ramanarayanan Viswanathan and Dr. Lei Cao. She was awarded first prize for her poster presentation at UM NSF-BWAC planning grant workshop.