DISTRIBUTED SPECTRUM SENSING VIA ENERGY DETECTION IN COGNITIVE RADIOS

by

Can Altay

B.S., Electronics and Communication Engineering, Yıldız Technical University, 2010

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University
2013

DISTRIBUTED SPECTRUM SENSING VIA ENERGY DETECTION IN COGNITIVE RADIOS

APPROVED BY:

Prof. Hakan Deliç (Thesis Supervisor)	....
Prof. Ayşın Baytan Ertüzün	....
Assoc. Prof. Tuna Tuğcu	....

DATE OF APPROVAL: 17.01.2013

ACKNOWLEDGEMENTS

First of all, I would like to thank to Prof. Hakan Deliç for his patience when he listens my ideas and his remarkable advices which always helped me to widen my perspective.

I thank my jury comitee members Prof. Ayşın Baytan Ertüzün and Assoc. Prof. Tuna Tuğcu for their constructive advices at my thesis defence.

I also want to thank to Assoc. Prof. Tuna Tuğcu and Dr. Birkan Yılmaz and other people from SATLAB for their allowance in their research group before this thesis. Their research group have been always welcoming and collaborative.

I thank to all members of WCL for their warm and friendly environment. We have spent great amount of time at laboratory together. Also, I thank to Artun Oyman for his sincere conversation whenever I break to free my mind.

Lastly, I have my special thanks to my parents Tayfun and Nigar and my brother Gökhan for their confidence on my decisions and their endless love and support.

ABSTRACT

DISTRIBUTED SPECTRUM SENSING VIA ENERGY DETECTION IN COGNITIVE RADIOS

Energy Detection is widely studied spectrum sensing system to detect signal energy due to its simplicity and lack of requirement signal characteristics. This thesis focuses on two main problems on distributed detection of energy detectors. The first problem is threshold finding in the nodes due to achieve constant false alarm probability (P_F) for binary decisions. The second problem is to derive optimum quantization of node observations due to minimum distortion and to give a expression on number of bits should nodes quantize their observations. For adaptive threshold finding, two algorithms are proposed. The first algorithm depends on minimizing the distance between instantenous P_F and desired P_F by Netwon's method. This algorithm is implemented for both mathematically calculated and estimated P_F values. The second algorithm depends on generating a discrete random variable and comparing it with the previous decision in every iteration. By this algorithm, system adjusts its threshold to any changes in the noise variance and works totally independent from $P_F(\lambda)$. In the second focus of the thesis, minimum distortion quantization of two hypothesis together is derived for Lloyd-Max algorithm and its performance is observed. Number of quantized bits of observations is also in the scope of this thesis. A lower bound on rate-distortion function is derived for the signal which consists of two hypotheses. Another approach on number of bits is handled by Chernoff bound which allows to compare performance of different bit quantizatons.

ÖZET

BİLİŞSEL RADYOLARDA ENERJİ ALGILAMA İLE DAĞINIK SPEKTRUM SEZME

Uygulama kolaylığı ve sinyal karakteristiklerine ihtiyaç duyamaması sebebiyle enerii algılama, spectrum sezme metodu olarak literatürde genişçe çalışılmış bir yöntemdir. Bu tez kapsamında, iki ana enerji algılama ile dağınık spektrum sezme problemine odaklanılmıştır. İlk problem tek bit karar veren algılama cihazlarında yanlış alarm olasılığını sabit düzeyde tutacak eşik değeri bulan algoritma geliştirmektir. İkinci problem ise gözlem kuantalanmasında en az bozulmayı sağlayacak kuantalama sevıyelerinin ve değerlerinin elde edilmesi ve algılama cihazlarının gözlemlerini kuantalayacakları bit sayısını elde etmek. Adaptif eşik değeri bulmak için iki algoritma önerilmiştir. İlk algoritma Newton yöntemini kullanarak anlık P_F ve istenen P_F arasındaki uzaklığı azaltmak üzerine kuruludur. Bu algoritma matematiksel olarak hesaplanmış P_F için de, kestirilmiş P_F değeri için de uygulanmıştır. İkinci algoritma her bir iterasonda yeni bir ayrık raslantı değişkeni oluşturup, bu raslantı değişkenini bir önceki algılama kararı ile karşılaştırarak çalışmaktadır. Bu algoritma, sistem eşik değerini gürültü varyasındaki her türlü değişime göre ayarlayabilmekte ve matematiksel $P_F(\lambda)$ fonksiyonundan tamamen bağımsız çalışmaktadır. İkinci problem için, Lloyd-Max algoritması ile, iki hipotezin birlikte en az bozulma ile kuantalaması bulunmuş ve performansı gözlemlenmiştir. Kuantalanan bit sayısı da ayrıca bu problem altında incelenmiştir. Iki hipotezin birleşimi olan sinyal üzerinden oran-bozulma fonksiyona bir alt sınır elde edilmiştir. Ayrıca farklı bir yaklaşım olarak, kuantalanmış gözlemler için Chernoff sınırı elde edilmiş ve bu sınırla farklı bit kuantalamaların performansları karşılaştırılmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS

D	Maximum allowable excepted distortion
f_j	PDF of jth Hypothesis
f_c	Central frequency of sensing
$I(a,b)$	Mutual Information
$I_m(.)$	Modified bessel function of first kind
$h(a)$	Entropy
$K(\lambda)$	Kullback-Liebler distance between instantenous P_F and α
m	Number of quantization levels
M	Number of samples
N	Number of sensing nodes
$\mathcal{N}(a,b)$	Gaussian distribution with a mean and b variance
N_0	One-sided PSD of noise
N_{02}	Two-sided PSD of noise
N_{win}	Number of summed decisions for estimating P_F
O_j	Output of Energy Detection for jth Hypothesis
$p_j(q_i)$	$\Pr(\lambda_{i+1} > z \geq \lambda_i)$ for H_j Hypothesis
P_D	Probability of detection of node
$P_{D,C}$	Probability of detection of fusion center
P_F	Probability of false alarm of node
$P_{F,C}$	Probability of false alarm of fusion center
P_M	Probability of miss of node
$P_{M,C}$	Probability of miss of fusion center
q_i	Quantization level
$Q(a)$	Q function
$Q_c(a,b)$	Marcum Q function
$R(D)$	Rate-distortion function
u_i	Test output of ith node
u_{des}	Discrete random variable ($u_{des} \in \{0,1\}$) with distribution of desired P_F

u_{fc}	Decision of fusion center
T	Sensing time interval
W	Sensing bandwidth
α	Desired P_F level
β	Threshold update coefficient
γ	Received SNR
$\bar{\gamma}$	Mean of received SNR
γ_T	Sum of node SNR values
$\gamma_u(.,.)$	Regularized upper incomplete gamma function
$\gamma_l(.,.)$	Regularized lower incomplete gamma function
$\Gamma(.)$	Gamma function
$\Gamma_u(.,.)$	Upper incomplete gamma function
$\Gamma_l(.,.)$	Lower incomplete gamma function
θ	Threshold addition update coefficient
λ	Threshold of node
λ_{Add}	Threshold addition
λ_{Add}^{Bound}	Lower bound of threshold addition
λ_i	Quantization threshold
λ_C	Threshold of fusion center
μ_j	Mean of sensing observation of H_j Hypothesis
$\mu_{q,j}$	Mean of quantized sensing observations of H_j Hypothesis
π_j	A priori probability of jth Hypothesis
σ_j^2	Variance of sensing observation of H_j Hypothesis
$\sigma_{q,j}^2$	Variance of quantized sensing observation of H_j Hypothesis
σ_n^2	Noise Variance
$\psi(a)$	Digamma function
χ_a^2	Chi-square distribution with a degree of freedom
$\chi_a^2(b)$	Noncentral chi-square distribution with a degree of freedom and b noncentrality parameter

LIST OF ACRONYMS/ABBREVIATIONS

AWGN	Additive White Gaussian Noise
BEP	Bit Error Probability
BPF	Band-pass Filter
CLT	Central Limit Theorem
CR	Cognitive Radio
ED	Energy Detection
EGC	Equal Gain Combining
I/Q	Inphase and Quadrature
KKT	Karush-Kuhn-Tucker
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
PDF	Probability Density Function
PSD	Power Spectral Density
SNR	Signal to Noise Ratio

1. INTRODUCTION

Cognitive Radio (CR) is a tranceiver based system to find and communicate through unused spectrum holes. The motivation of CR is the current under-utilization of spectrum bands. According to the spectrum occupancy studies in several different cities around the world [1–3], the spectrum is used inefficiently. Even though the spectrum bands are assigned to users for long terms and high prices, many communication bands are underutilized because of their technology or number of users. As a result of long-term occupancies, it is an expensive and difficult task to change the technology and use the bands more efficiently.

To solve this problem, the idea of a more intelligently working radio is proposed in [4] and [5]. CR is thought to be a secondary user of the frequency band, using the empty spectrum channels for its communication without affecting the primary user of the band. By this technology, the primary owner of the band can still guarantee quality of communication service to its users, while the secondary users' service quality depends on the usage of the band by the primary users.

CR has four main functions to form its intelligent structure [6]. The spectrum sensing function senses the channels of the band for the primary user existence. When a primary user is detected at a channel, CR transmission should vacate the channel in a pre-defined time interval and continue its transmission from another available channel if it exists. This functionality is called spectrum mobility. Before implementing the spectrum mobility, the best available channels have been chosen by spectrum decision function. Besides, the spectrum mobility also depends on spectrum sharing functionality which is to choose spectrum hole by taking capacity requirements of other secondary users into account.

1.1. Spectrum Sensing

The spectrum sensing is the key feature of CR due to unknown properties of the primary user transmission and avoidance of disturbance to primary users. Although the latest developments in CR regulation favor the database approach [7] instead, spectrum sensing can provide reliable opportunistic spectrum access in dynamic and unpredictable environments. Furthermore, spectrum sensing in CR is not taken as obsolete and studies on spectrum sensing goes on literally.

There are several proposed methods for signal detection [8]. Most of the methods depend on taking samples to detect analog signal Wave-form based sensing is a method which senses by taking correlation between received signal samples and known signal pattern. This method performs well with help of the knowledge of signal pattern [9]. The matched filter is a proposed method for the signal detection which is thought to be optimum sensing model [10]. It is also the only sensing model which can be implemented without sampling among others. But it requires perfect knowledge of the signal and has high complexity. Another proposed method is cyclostationarity-based method which differentiates the noise from the signal by noise's wide sense stationarity and redundancy of signal periodicities [11]. It can be implemented with or without knowledge of cyclic frequencies of the signal. Radio identification based sensing method depends on monitoring and identifying the transmission technology of the primary user [12]. Such an identification provides higher accuracy and higher dimensional knowledge. There are other proposed sensing methods such as multitaper spectrum estimate and wavelet based detection but our focus for this thesis will be on energy detection (ED).

1.2. Energy Detection

As introduced in the previous section, ED depends on comparing the sum of squares of the signal samples with a threshold. According to well known study of Urkowitz [13], energy of time and band limited signal can be approximated well by sum of finite number of squared samples. As the number of samples increases, the approximation of signal energy becomes better and detection performance increases. This model proposes that when the signal exists, it is assumed to be Gaussian with non-zero mean. For a good comparison on other ED signal models, one can read [14]. In our signal model, the output of ED is chi-square distribution in the case of absence of signal. When the signal exists, the output is noncentral chi-square distribution, which has a noncentrality parameter that depends on Signal-to-noise ratio (SNR) of the signal.

According to the analysis of detection with high noise uncertainity and low SNR, when the SNR is below a certain value, the signal is not detectable regardless of how many samples are taken. This certain SNR level is called SNR wall, and this concept is introduced in [15, 16].

For the case when the signal exists, the signal faces fading effect due to wireless medium. Small scale flat fading channel models such as Rayleigh, Nakagami, Rician are widely studied by Digham in [17, 18]. As large scale fading, log-normal shadowing channel is evaluated early by Ghasemi [19] and more widely studied by Atapattu in [20, 21] by mixing small scale fading channels. The fading channels are mostly handled by assuming SNR value as a random variable. It is assumed to be constant in additive white gaussian noise (AWGN) channel and has different random characteristics for different channel models. Also, the multi-antenna energy detection is studied under antenna correlation in [22].

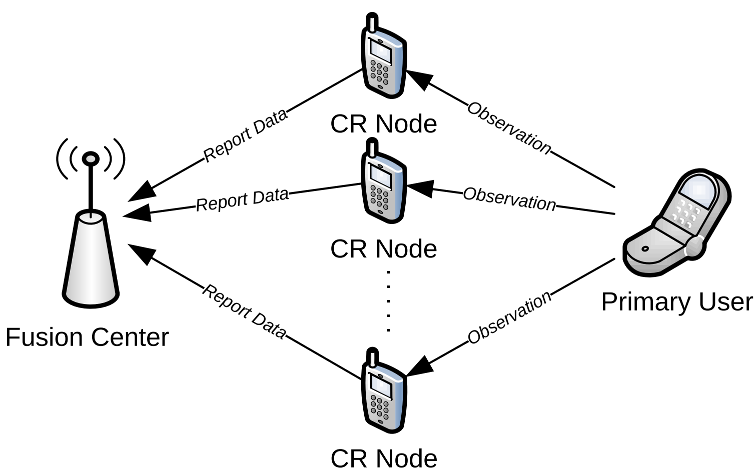


Figure 1.1. Distributed Spectrum Sensing in Cognitive Radios.

Since the fading and low SNR cases affect the sensing performance, distributed sensing is studied widely in the literature. In this thesis, a central unit which is called fusion center exists to gather node sensing data and make final, more reliable decision on the signal existence in the certain channels (Figure 1.1). There are also ad hoc mode studies in the literature [23] which are out of the scope of this thesis. Distributed spectrum sensing studies differ in their data fusion methods and transmitted data of each sensing node. Soft (output of ED) and hard (detection results of each node) data can be transmitted from the nodes. For equal number of sensing users, soft data deliver better performance. According to [24], performance of soft data can be outperformed by hard data as number of sensing nodes gets larger. As data fusion method in this thesis, we adopt equal gain combining which consists of summing the node sensing data and comparing it with a threshold. Other combining methods are selective combining, switch and stay combining [17].

Hard data can be called single bit decision of sensing node. Since it depends on sending only one bit to the fusion center, the performance is vulnerable to bit error probability (BEP) in the reporting channel between sensing nodes and fusion center [25, 26]. Moreover, the performance improvement by feedback from fusion center to sensing nodes is studied for Bayesian [27, 28] and Neyman-Pearson hypothesis testing [29]. Robust decision making in the presence of malicious nodes is studied in [30].

Inevitably, there are many studies for optimization of data fusion with different perspectives. For data fusion by weighted sum of soft data, optimization of weights is solved with simplified problem as maximum ratio combining in [31] and without simplification in [32]. Moreover, optimal thresholds of both fusion center and nodes have been studied for hard data in [33, 34]. Optimization studies have continued not only on detection of the signal in one band but several bands. Optimum multi-band sensing studies are handled in [35, 36].

1.3. Adaptive Methods

Adjusting threshold adaptively is another challenging problem due to difficult mathematical calculation of desired threshold and changing variance. In the literature, there are lots of studies to adjust threshold due to estimation on noise power [37], although computational difficulty of threshold still exists. A good study on the performance of the estimated variance case with an original SNR wall phenomenon can be found in [38].

Another perspective is to estimate instantenous probability values to adjust threshold. For Bayesian hypothesis testing, a threshold updating approach which minimizes the Bayesian risk is proposed in [39] with estimation of probability and similar approach is proposed for Neyman-Pearson hypothesis testing in [40].

As another aspect, there are simpler proposed algorithms which obey neither Bayesian nor Neyman-Pearson detection rules to adjust threshold. An algorithm is proposed in [41] to adjust thresholds with a technique from image binarization literature. Alternatively, thresholds can be updated with pre-defined constant additions when the signal exists [42].

1.4. Quantization of Observations

Even though adaptive algorithms for threshold finding are studied for hard data case, quantzed observations are generally derived mathematically in the literature. There are several studies which derive the formulation of performance [43, 44] in the literature of quantization of observations in CR.

As we include quantization of observations in CR and before CR, many studies exists to find optimum quantization levels and thresholds. An older study [45] derives optimum quantizer for minimum distortion, asymptotically. For derivation of optimum quantizers, deflection coefficient [46] is used as the optimization criterion. According to the deflection coefficient, the optimum quantization levels for quantization the thresholds are derived in [47,48], analytically. Also, for more simplified version of deflection coefficient (which is the difference of means of quantized observations), optimum quantization levels and thresholds are derived in [49].

Another aspect for the quantization of observations is the number of bits, which observations will quantize to keep desired performance. The question of how many bits is already studied as rate-distortion theory in quantization literature for minimum distortion [50]. This question for quantization of observations is also studied by Chernoff bound on error for detection literature [46]. This bound gives a performance measure for detection and is well studied for sensing observations in [51].

1.5. Outline

The rest of the thesis is organized as follows. In Chapter 2, signal model and probability derivations with or without distributed sensing are presented. In chapter 3, two adaptive threshold finding methods are proposed. The first algorthm implements the threshold finding by minimizing the distance between desired and instantenous probability of false alarm (P_F) values with Newton's method and the second algorithm finds the threshold by generating a discrete random variable and comparing it to previous decisions, iteratively. Chapter 4 presents optimum minimum distortion quantization by Lloyd-Max algorithm with its performance. Chapter 5 presents mathematical derivations on the question of "How many bits should nodes quantize their observations?". To achieve a mathematical expression on this question, rate-distortion theory and Chernoff bound are considered, separately. Finally, Chapter 6 gives the concluding remarks and directions for future works.

2. SIGNAL MODEL AND ANALYTICAL DERIVATIONS

2.1. Signal Model for Energy Detection

The received signal for two hypothesis represented as

$$\begin{aligned} y(t) &= n(t) & H_0\ (\textit{white space}), \\ y(t) &= h_j \cdot s(t) + n(t) & H_1\ (\textit{occupied}) \end{aligned} \tag{2.1}$$

where $y(t)$ is the signal received by the CR node, $s(t)$ is the transmitted signal of the primary user, $n(t)$ is AWGN with zero mean, and h_j is the gain of the channel of the jth sensing node, respectively. H_1 represents the case that transmission exists, and H_0 represents the case when there is no transmission by the primary user.

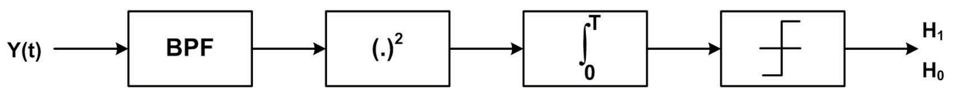


Figure 2.1. Block Diagram of Energy Detector.

The block diagram of energy detector for a desired band is shown in Figure 2.1. The received signal first passes through bandpass filter with desired bandwidth. Then, the output of the filter is squared and integrated over the desired sensing time interval to get a measure of energy of the received signal. The output is compared with a threshold for detection purpose.

There are two cases for deterministic signal detection: known or unknown noise power spectral densities (PSD). For the known noise PSD, transfer function of bandpass filter is

$$H_{bandpass}(f) = \begin{cases} \sqrt{\frac{2}{N_0}}, & |f - f_c| \leq W \\ 0, & |f - f_c| > W \end{cases} \tag{2.2}$$

where f_c is central frequency, W is bandwidth of observation, and N_0 is one-sided PSD. Two-sided PSD equals half the value of one-sided PSD ($N_{02} = \frac{N_0}{2}$). With this transfer function, bandpass filter normalizes the variance of the noise.

According to the sampling theorem [52], noise can be expressed as

$$y(t) = n(t) = \sum_{i=-\infty}^{\infty} n_i sinc(2Wt - i) \tag{2.3}$$

where $n_i = n\left(\frac{i}{2W}\right) \sim \mathcal{N}(0, N_0 W)$. $\mathcal{N}(a, b)$ represents Gaussian distribution with mean a and variance b. By using

$$\int_{-\infty}^{\infty} \text{sinc}(2Wt - i)\text{sinc}(2Wt - j)dt = \begin{cases} \frac{1}{2W}, & i = j \\ 0, & i \neq j, \end{cases} \tag{2.4}$$

integral of the squared signal at Equation 2.3 can be written as

$$\int_{-\infty}^{\infty} n^2(t)dt = \frac{1}{2W} \sum_{i=-\infty}^{\infty} n_i^2. \tag{2.5}$$

The observation which is the output of energy detector (Figure 2.1) is

$$O_{H_0} = \frac{1}{N_{02}} \left(\int_0^T n^2(t)dt \right) = \frac{2}{N_0} \left(\int_0^T n^2(t)dt \right) \simeq \frac{2}{N_0} \left(\frac{1}{2W} \sum_{i=1}^{M} n_i^2 \right) = \sum_{i=1}^{M} \acute{n}_i^2 \tag{2.6}$$

where T is the observation interval and $\acute{n}_i = \frac{n_i}{\sqrt{N_0 W}} \sim \mathcal{N}(0, 1)$. In this formula, M is number of the samples taken to approximate the energy of noise. M should be chosen sufficiently large for a good approximation. Sufficient number of samples is $M = 2TW$ due to [13]. This result is obtained by certain fall off at eigenvalues of band and time limited functions, when more than $2TW$ samples are taken [53]. In other words, first $2TW$ sampled band and time limited eigenfunctions has obvious higher eigenvalues than the eigenfunctions with higher samples. The interval between each two samples is $\frac{T}{M}$.

According to sampling theorem, when the signal exists, the received signal can be expressed as

$$y(t) = h_j \cdot s(t) + n(t) = \sum_{i=-\infty}^{\infty} (h \cdot s_i + n_i) sinc(2Wt - i) \tag{2.7}$$

where $s_i = s\left(\frac{i}{2W}\right)$. The output of the ED can be expressed similar to Equation 2.6.

$$O_{H1} = \frac{2}{N_0}\left(\int_0^T (h_j \cdot s(t) + n(t))^2 dt\right) \simeq \frac{2}{N_0}\left(\frac{1}{2W}\sum_{i=1}^{M}(h_j \cdot s_i + n_i)^2\right) = \sum_{i=1}^{M} \acute{s}_i^2 \tag{2.8}$$

where $\acute{s}_i = \frac{(h \cdot s_i + n_i)}{\sqrt{N_0 W}} \sim \mathcal{N}\left(\frac{h \cdot s_i}{\sqrt{N_0 W}}, 1\right)$. Furthermore, the two resulting distributions under H_0 and H_1 are

$$\begin{aligned} O_{H_0}(j) &\sim \chi_M^2 \\ O_{H_1}(j) &\sim \chi_M^2\left(M\frac{h_j^2 \cdot s_i^2}{N_0 W}\right) = \chi_M^2(M\gamma_j) \end{aligned} \tag{2.9}$$

where γ_j is SNR value for jth node, χ_M^2 is the central chi-square distribution, and $\chi_M^2(a)$ denotes the noncentral chi-square distribution with a noncentrality parameter [54].

If the PSD of the noise is not known, the transfer function of the bandpass filter can be

$$H_{bandpass}(f) = \begin{cases} 1, & |f - f_c| \leq W \\ 0, & |f - f_c| > W. \end{cases} \tag{2.10}$$

without normalizing the PSD of noise. As a result of this transfer function, every sample is square of a gaussian random variable with variance $\sigma_n^2 = N_{02}$. Resulting distributions of the output of energy detector can be shown as,

$$\begin{aligned} \frac{O_{H_0}(j)}{\sigma_n^2} &\sim \chi_M^2 \\ \frac{O_{H_1}(j)}{\sigma_n^2} &\sim \chi_M^2(M\gamma_j). \end{aligned} \tag{2.11}$$

Due to the band-pass filters, in-phase and quadrature (I/Q) components of the signal may need to be considered. But according to derivations in [13], taking the I/Q components into account does not change the result of the formulations.

2.1.1. Energy Detection with Central Limit Theorem

If the number of samples is asymptotically large, the output of the energy detector can be assumed to be Gaussian due to Central Limit Theorem (CLT). But there is still a small difference between exact Chi-square distributions and Gaussian approximation regardless of how many samples due to nonnegativity of squared sample random variables. For the energy detector which normalizes the noise variance with filter in Equation 2.2, the output distributions are [32]

$$\begin{aligned} O_{H_0}(j) &\sim \mathcal{N}\big(M, 2M\big) \\ O_{H_1}(j) &\sim \mathcal{N}\big(M(1+\gamma_j), 2M(1+2\gamma_j)\big). \end{aligned} \tag{2.12}$$

If the variance of the noise is unknown and a filter with transfer function (2.10) is used, the output distributions become

$$\begin{aligned} O_{H_0}(j) &\sim \mathcal{N}(M\sigma_n^2, 2M\sigma_n^4) \\ O_{H_1}(j) &\sim \mathcal{N}(M\sigma_n^2(1+\gamma_j), 2M\sigma_n^4(1+2\gamma_j)). \end{aligned} \tag{2.13}$$

2.2. Probability Derivations

Derivations of P_F and probability of detection (P_D) for the ED are represented in this section. P_F is the probability of deciding H_1 when H_0 is true and P_D is probability of detecting H_1 when H_1 is true.

2.2.1. AWGN Channel

Probability density functions (PDF) of the output of the ED in Equation 2.9 are chi-square and noncentral chi-square distributions [54],

$$H_1 : f_1(x) = \frac{1}{2} exp\left(- \frac{x + M\gamma}{2} \right) \left[\frac{x}{M\gamma} \right]^{\frac{M}{4} - \frac{1}{2}} I_{\frac{M}{2}-1}(\sqrt{M\gamma x}) \tag{2.14}$$

$$H_0 : f_0(x) = \frac{x^{\frac{M}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{M}{2}} \Gamma(\frac{M}{2})} \tag{2.15}$$

where M is number of samples, γ is received SNR, $\Gamma(.)$ is gamma function, and $I_m(.)$ is mth order modified Bessel function of first the kind [54]. Modified Bessel function of first kind has a series expansion as

$$I_m(t) = \sum_{k=0}^{\infty} \frac{(\frac{t}{2})^{m+2k}}{k!\Gamma(m+k+1)}. \tag{2.16}$$

For finding P_F, integration of $f_0(x)$ yields regularized upper incomplete gamma function which is examined at Appendix A. Index of this function (u) represents its property of being upper.

$$P_F(\lambda) = \int_{\lambda}^{\infty} f_0(z) dz = \gamma_u \left(\frac{M}{2}, \frac{\lambda}{2} \right) \tag{2.17}$$

This equation can be calculated by the finite summation of Eq. A.4 if the number of samples (M) is even.

P_D can be found by integrating $f_1(x)$ which can be expressed as Marcum Q function [55],

$$P_D(\lambda) = \int_{\lambda}^{\infty} f_1(z) dz = Q_{\frac{M}{2}}(\sqrt{M\gamma}, \sqrt{\lambda}). \tag{2.18}$$

where

$$Q_c(a,b) = \int_b^{\infty} z\left(\frac{z}{a}\right)^{c-1} e^{-\frac{z^2+a^2}{2}} I_{c-1}(az)dz, \quad c \geq 1 \tag{2.19}$$

In the AWGN channel, γ value assumed to be constant and same at each node. In other channels, derivations become more complicated due to random γ values.

2.2.2. Rayleigh Channel

Under the Rayleigh fading channel, the signal amplitude follows the Rayleigh distribution. Therefore, γ follows the exponential distribution. PDF of γ in the Rayleigh channel SNR value is

$$f(\gamma) = \frac{1}{\overline{\gamma}} \exp(-\frac{\gamma}{\overline{\gamma}}) \tag{2.20}$$

where $\overline{\gamma}$ is the mean received SNR value of CR nodes. Furthermore, P_D depends on Eq. 2.20 for determining γ and equation. 2.14 for determining O_i of the current node in the Rayleigh fading channel. P_D is calculated as [17]

$$P_D(x) = e^{-x/2} \sum_{i=0}^{\frac{M}{2}-2} \frac{1}{i!}\left(\frac{x}{2}\right)^i + \left(\frac{1+\overline{\gamma}}{\overline{\gamma}}\right)^{\frac{M}{2}-1} \left[e^{-\frac{x}{2(1+\overline{\gamma})}} - e^{-\frac{x}{2}} \sum_{i=0}^{\frac{M}{2}-2} \frac{1}{i!}\left(\frac{x\overline{\gamma}}{2(1+\overline{\gamma})}\right)^i \right]. \tag{2.21}$$

2.3. Distributed Spectrum Sensing

Distributed sensing is achieved by a fusion center which gathers test outputs (u_i) of each node (Figure 2.2). These test outputs may be binary decisions, several bit quantization of observations(O_i) or soft data of O_i itself. In the next subsections, formulation of P_F and P_D is represented for different type of test outputs.

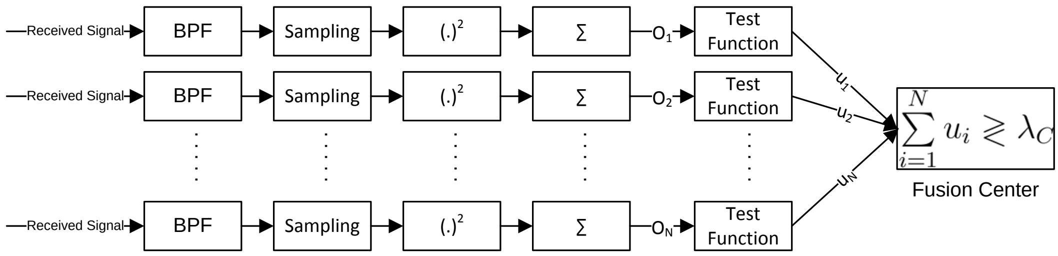


Figure 2.2. Block diagram of Distributed Spectrum Sensing.

2.3.1. Sum of Binary Decisions

In this case, every CR node sends 1 bit decisions about the signal existence to fusion center. Because of only one bit, performance is free from the values corresponding to each bit. In this study, 1 bit decision is taken zero for H_0 case and one for H_1 case. In the formulation of sum of binary decisions, fusion center threshold (λ_C) is assumed to be an integer. A non-integer λ_C does not change the performance.

Sum of the binary decisions can be taken as binomial distribution and P_F and P_D can be calculated as [56]

$$P_{D,C}(\lambda_C) = \sum_{k=\lambda_C}^{N} \binom{N}{k} (1 - P_D)^{N-k} (P_D)^k \tag{2.22}$$

$$P_{F,C}(\lambda_C) = \sum_{k=\lambda_C}^{N} \binom{N}{k} (1 - P_F)^{N-k} (P_F)^k \tag{2.23}$$

where N is number of sensing nodes, $P_{D,C}(\lambda_C)$ and $P_{F,C}(\lambda_C)$ are probability values at fusion center and P_D and P_F are local probability values at nodes.

As Neyman-Pearson detection rule requires to keep P_F at a desired level (α), threshold should be chosen due to inverse of Equation 2.23. As a reason of discrete behaviour of sum of binary decisions, Equation 2.23 may not be able to guarantee a λ_C value which achieves $P_F(\lambda_C) = \alpha$ for any α. Furthermore, a probability value (r) defined as randomization parameter is included [46]. When λ_C equals to sum of binary decisions, fusion center decides H_1 by randomization of r probability to maintain

$P_F(\lambda_C) = \alpha$. By randomization, $P_{F,C}$ formula becomes

$$P_{F,C}(\lambda_C, r) = \sum_{k=\lambda_C+1}^{N} \binom{N}{k}(1-P_F)^{N-k}(P_F)^k + r\binom{N}{\lambda_C}(1-P_F)^{N-\lambda_C}(P_F)^{\lambda_C} \quad (2.24)$$

where

$$r = \frac{\alpha - P_{F,C}(\lambda_C + 1)}{P_{F,C}(\lambda_C) - P_{F,C}(\lambda_C + 1)} \quad (2.25)$$

and α is desired P_F. It can be noticed that r is calculated from α and $P_F(\lambda_C)$ without r.

By the same method, P_D can be calculated by

$$P_{D,C}(\lambda_C, r) = \sum_{k=\lambda_C+1}^{N} \binom{N}{k}(1-P_D)^{N-k}P_D^k + r\binom{N}{\lambda_C}(1-P_D)^{N-\lambda_C}P_D^{\lambda_C}. \quad (2.26)$$

2.3.2. Sum of Soft Decisions

In this case, we assume that the output of the ED is directly sent to the fusion center as u_i (Figure 2.2) and all gathered soft u_i values summed at the fusion center. This sum is compared with λ_C by fusion center to detect. This method is called equal gain combining. The sum of ED observations are again chi-square or noncentral chi-square distributions with multiplied degree of freedom with number of users (N) [17].

$$O_{H_0}(Nusers) \sim \chi^2_{NM} \quad (2.27)$$

$$O_{H_1}(Nusers) \sim \chi^2_{NM}(M\gamma_T) \quad (2.28)$$

where γ_T is sum of SNR values of nodes,

$$\gamma_T = \sum_{i=1}^{N} \gamma_i. \quad (2.29)$$

Furthermore, $P_{F,C}(\lambda_C)$ and $P_{D,C}(\lambda_C)$ can be calculated with a similar way with the case without cooperation.

$$P_{F,C}(\lambda_C) = \gamma_u\left(\frac{MN}{2}, \frac{\lambda_C}{2}\right) \tag{2.30}$$

$$P_{D,C}(\lambda_C) = Q_{\frac{MN}{2}}(\sqrt{M\gamma_T}, \sqrt{\lambda_C}) \tag{2.31}$$

For AWGN channels, γ_T can be calculated by multiplying the SNR of the received signal at the nodes by N [17], if each node is assumed to have same received signal SNR.

$$\gamma_T = N\gamma \tag{2.32}$$

Since SNR has exponential distribution in Rayleigh channel,

$$f_{\gamma_i}(x) = \frac{1}{\overline{\gamma}}e^{-\frac{x}{\overline{\gamma}}}, \tag{2.33}$$

γ_T can be calculated by

$$f_{\gamma_T}(x) = \frac{1}{\overline{\gamma}^N}\frac{x^{N-1}}{(N-1)!}e^{-\frac{x}{\overline{\gamma}}} \tag{2.34}$$

which is the distribution of the sum of exponential random variables [56]. Since the sum of two independent random variables is the convolution of their PDFs, this distribution can be derived by taking convolution of exponential distributions several times.

2.3.3. Sum of Quantized Observations

The sensing observation of a node can be sent to the fusion center in multiple bits which has performance between the previous two cases. Since the number of bits that should be sent is a design criterion, the formulation of the performance of the multiple bit observations is presented in this section.

Even though there are probability derivations of fusion of several bit observations by binomial distribution in the literature [44], a more generalized probability formulation is derived in this subsection. To evaluate the formulation, discrete PDF of quantized observations is used as if it is continuous by

$$f_{q,j}(x) = \sum_{i=1}^{m} p_j(q_i) v(q_i - x) \tag{2.35}$$

where

$$p_j(q_i) = \Pr(\lambda_{i+1} > z \geq \lambda_i) = \int_{\lambda_i}^{\lambda_{i+1}} f_j(z) dz \tag{2.36}$$

and

$$v(x) = \begin{cases} \infty, & x = 0, \\ 0, & otherwise, \end{cases}, \quad \int_{-\infty}^{\infty} v(x) dx = 1 \tag{2.37}$$

where $f_{q,j}(x)$ is continuous PDF of quantized observation of H_j Hypothesis, q_i values are the quantization levels, λ_i values are the quantization thresholds, and $f_j(x)$ is the output distributon of ED of H_j Hypothesis.

Calculation of the sum of discrete quantized data of sensing nodes can be calculated by convolution [57, 58]. Convolution of two quantized random variables is

$$\begin{aligned} f_{2users,j}(x) &= \int_{-\infty}^{\infty} f_{local,j}(z) f_{local,j}(x-z) dz \\ &= \int_{-\infty}^{\infty} \left[\sum_{a=1}^{m} p_j(q_a) v(q_a - z) \right] \left[\sum_{b=1}^{m} p_j(q_b) v(q_b - x + z) \right] dz \\ &= \sum_{a=1}^{m} \sum_{b=1}^{m} p_j(q_a) p_j(q_b) \int_{-\infty}^{\infty} v(q_a - z) v(q_b - x + z) dz \\ &= \sum_{a=1}^{m} \sum_{b=1}^{m} p_j(q_a) p_j(q_b) v(q_a + q_b - x), \end{aligned} \tag{2.38}$$

which is the summation of m^2 elements. By using convolution for each users data, this can be generalized on

$$f_{Nusers,j}(x) = \sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m}\dots p_j(q_a)p_j(q_b)p_j(q_c)\dots v(q_a + q_b + q_c + \cdots - x) \quad (2.39)$$

which is the summation of m^N elements. There are N summation symbols where each symbol is for quantization levels of a sensing node.

The good thing about this summation is that the x variable is only in $v(.)$. This eases the calculation of $P_{F,C}$. By taking integral from λ_C to ∞, the probability of false alarm becomes

$$P_{F,C}(\lambda_C) = \sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m}\dots p_0(q_a)p_0(q_b)p_0(q_c)\dots s(q_a + q_b + q_c + \cdots - \lambda_C) \quad (2.40)$$

where

$$s(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (2.41)$$

Since Equation 2.40 depends on a discrete random variable, it still requires a randomization probability r to achieve the desired $P_{F,C}$ value (α). This randomization can be added by an update at $s(x)$ function.

$$P_{F,C}(\lambda_C, r) = \sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m}\dots p_0(q_a)p_0(q_b)p_0(q_c)\dots s'(q_a + q_b + q_c + \cdots - \lambda_C) \quad (2.42)$$

where

$$s'(x) = \begin{cases} 1, & x \geq 0, \\ r, & x < 0, \end{cases} \quad (2.43)$$

and

$$r = \frac{\alpha - P_{F,C}(\lambda_C)}{1 - P_{F,C}(\lambda_C)}. \tag{2.44}$$

In the fusion center, this randomization is done when the sum of data is below the threshold. Fusion center decides that the signal exists when the sum is above the threshold. When the sum is below the threshold, fusion center does randomization and again decides the signal exists with probability r to maintain $P_{F,C} = \alpha$.

For Neyman-Pearson hypothesis testing, λ_C should be calculated from inverse of $P_{F,C}$ function. Since Equation 2.42 is not invertible, its inverse can be taken by numerical methods such as bisection method. In the analysis of this thesis, the inverse of the Equation 2.40 is calculated by bisection method to find λ_C and r is calculated using 2.44.

Similarly, $P_{D,C}$ can be calculated on

$$P_{D,C}(\lambda_C) = \sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m} \dots p_1(q_a)p_1(q_b)p_1(q_c)\dots s(q_a + q_b + q_c + \cdots - \lambda_C) \tag{2.45}$$

and

$$P_{D,C}(\lambda_C, r) = \sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m} \dots p_1(q_a)p_1(q_b)p_1(q_c)\dots s'(q_a + q_b + q_c + \cdots - \lambda_C). \tag{2.46}$$

2.3.4. Sum of Observations with Central Limit Theorem

The formulation of summed decisions for all cases of data is already given in the previous sections. According to the CLT [56], when the number of users is large enough, it is possible to approximate probability values by assuming that the sum has as Gaussian distribution. This makes formulation easier.

Probability values of the sum is

$$\begin{aligned} P_{D,C}(\lambda_C) &= Q\left(\frac{\lambda_C - N\mu_1}{\sqrt{N}\sigma_1}\right) &(2.47)\\ P_{F,C}(\lambda_C) &= Q\left(\frac{\lambda_C - N\mu_0}{\sqrt{N}\sigma_0}\right) &(2.48)\end{aligned}$$

where μ_1, μ_0 are mean values and σ_1^2, σ_0^2 are variance values of H_1 and H_0 hypotheses, respectively. These values can be found in Section 2.1.1.

Furthermore, for desired P_F (α), λ_C can be calculated by

$$\lambda_C = Q^{-1}(\alpha)\sqrt{N}\sigma_0 + N\mu_0. \tag{2.49}$$

which requires calculation of inverse of Q function. By putting this to $P_{D,C}$ formula, we can get [32]

$$\begin{aligned} P_{D,C}(\alpha) &= Q\left(\frac{Q^{-1}(\alpha)\sqrt{N}\sigma_0 + N\mu_0 - N\mu_1}{\sqrt{N}\sigma_1}\right) \\ &= Q\left(Q^{-1}(\alpha)\frac{\sigma_0}{\sigma_1} - \sqrt{N}\frac{\mu_1 - \mu_0}{\sigma_1}\right) &(2.50)\end{aligned}$$

which gives directly performance formulation. Since taking inverse of $P_{F,C}(\lambda_C)$ was a hard task for the formulation of the previous subsections, the $P_{D,C}(P_{F,C}^{-1}(\alpha))$ formula is easier to implement by the help of CLT. The disadvantage of this formulation is the requirement of large number of users.

For quantized observations, the same formulation is valid [56]

$$\begin{aligned} P_{D,C}(\lambda_C) &= Q\left(\frac{\lambda_C - N\mu_{q,1}}{\sqrt{N}\sigma_{q,1}}\right) &(2.51)\\ P_{F,C}(\lambda_C) &= Q\left(\frac{\lambda_C - N\mu_{q,0}}{\sqrt{N}\sigma_{q,0}}\right) &(2.52)\end{aligned}$$

where

$$\mu_{q,j} = \sum_{i=0}^{m} q_i \int_{\lambda_i}^{\lambda_{i+1}} f_j(x)dx \tag{2.53}$$

$$\sigma_{q,j}^2 = \sum_{i=0}^{m} (q_i - \mu_{q,j})^2 \int_{\lambda_i}^{\lambda_{i+1}} f_j(x)dx \tag{2.54}$$

are means and variances of quantized sensing observations for H_j Hypothesis. But in this case, not only the number of users (N), but also the number of quantization levels (m) should be large. This formulation cannot give exact desired probability values because of lack of randomization, r.

Again, the performance formulation of $P_{D,C}(P_{F,C}^{-1}(\alpha))$ can be derived for quantized observations (large number of users and quantization levels) as,

$$\lambda = Q^{-1}(\alpha)\sqrt{N}\sigma_{q,0} + N\mu_{q,0} \tag{2.55}$$

$$P_{D,C}(\alpha) = Q\left(Q^{-1}(\alpha)\frac{\sigma_{q,0}}{\sigma_{q,1}} - \sqrt{N}\frac{\mu_{q,1} - \mu_{q,0}}{\sigma_{q,1}} \right). \tag{2.56}$$

3. ADAPTIVE THRESHOLD FINDING

Adaptive threshold finding methods, which are presented at Section 1.3, on spectrum sensing mainly depend on variance estimation or threshold finding methods without detection perspective. Spectrum sensing with noise power estimation requires large number of samples for good performance [37, 38]. The methods which do not depend on estimation generally do not account for conventional detection rules (Bayesian, Neyman-Pearson, Minimax detection rules).

According to Neyman-Pearson detection rule, design criteria is to keep P_F at a constant level and maximize P_D. To keep P_F at a desired level (α), sensing nodes should be able to take inverse of $P_F(\lambda)$ (Section 2.2) and should know channel characteristic which is the noise variance. The focus of this section is to design algorithms which find necessary detection threshold iteratively for $P_F = \alpha$ without knowing exact noise variance and using the inverse of $P_F(\lambda)$.

3.1. Threshold Finding with Newton's Method

To find the threshold, a distance between desired P_F and instantenous P_F have to be defined and should be minimized iteratively. For this study, Kullback-Leibler distance (also known as Relative Entropy) [50] will be implemented due to its sensitivity as a distance between probability values. Due to convexity of Kullback-Leibler distance, an iterative convex optimization algorithm have to be included. Newton's method for optimization [59] is a well known method for finding minimum of a convex function, iteratively. Even though Newton's Method requires first and second derivatives of convex function, next subsection points out that computation of derivatives of $P_F(\lambda)$ is easier than compuation of $P_F(\lambda)$ itself. The perspective of this algorithm is similar to one which is proposed at [29].

3.1.1. Newton's Method for Analytical $P_F(\lambda)$

Firstly, threshold finding algorithm of Newton's method is handled by only analytical functions. For every iteration, the threshold update can be given as [59]

$$\lambda_{n+1} = \lambda_n + \lambda_{n,Add} = \lambda_n - \frac{\frac{\partial K(\lambda_n)}{\partial \lambda_n}}{\frac{\partial^2 K(\lambda_n)}{\partial \lambda_n^2}} \tag{3.1}$$

where $K(\lambda_n)$ is Kullback distance between α and $P_F(\lambda_n)$. This distance and its derivatives are [50]

$$\begin{aligned} K(\lambda) &= \alpha \, log \frac{\alpha}{P_F(\lambda)} + (1-\alpha) log \frac{1-\alpha}{1-P_F(\lambda)} && (3.2) \\ K'(\lambda) &= P_F'(\lambda) \left[\frac{1-\alpha}{1-P_F(\lambda)} - \frac{\alpha}{P_F(\lambda)} \right] && (3.3) \\ K''(\lambda) &= P_F''(\lambda) \left[\frac{1-\alpha}{1-P_F(\lambda)} - \frac{\alpha}{P_F(\lambda)} \right] + P_F'(\lambda)^2 \left[\frac{1-\alpha}{[1-P_F(\lambda)]^2} - \frac{\alpha}{P_F(\lambda)^2} \right] && (3.4) \end{aligned}$$

where $K'(\lambda)$ and $K''(\lambda)$ are first and second derivatives of Kullback-Leibler distance. According to Equation 2.11, for each noisy observation

$$O_{H0} \gtreqless \lambda = \sigma_n^2 \lambda^* \tag{3.5}$$

where λ^* is the threshold when $\sigma_n^2 = 1$ and λ is the actual threshold for decision making. Furthermore, calculation of P_F and its derivatives are

$$\begin{aligned} P_F(\lambda^*) &= \frac{1}{\Gamma(\frac{M}{2})} \int_{\frac{\lambda}{2}}^{\infty} t^{\frac{M}{2}-1} e^{-t} dt && (3.6) \\ P_F'(\lambda^*) &= \frac{-1}{2\Gamma(\frac{M}{2})} \left(\frac{\lambda^*}{2} \right)^{\frac{M}{2}-1} e^{-\frac{\lambda^*}{2}} && (3.7) \\ P_F''(\lambda^*) &= P_F'(\lambda^*) \left[\frac{1}{\lambda^*} \left(\frac{M}{2} - 1 \right) - \frac{1}{2} \right] && (3.8) \end{aligned}$$

where $P_F'(\lambda^*)$ and $P_F''(\lambda^*)$ are first and second derivatives of $P_F(\lambda^*)$. Then, by putting all derivatives together, threshold update in Equation 3.1 becomes

$$\begin{aligned}\frac{\frac{\partial K(\lambda^*)}{\partial \lambda^*}}{\frac{\partial^2 K(\lambda^*)}{\partial \lambda^{*2}}} &= \left[\frac{P_F''(\lambda^*)}{P_F'(\lambda^*)} + P_F'(\lambda^*)\frac{\left(P_F(\lambda^*)-\alpha\right)^2+\alpha\left(1-\alpha\right)}{P_F(\lambda^*)\left(1-P_F(\lambda^*)\right)\left(P_F(\lambda^*)-\alpha\right)}\right]^{-1} \\ &= \left[\frac{M-2}{2\lambda^*} - \frac{1}{2} - \frac{\lambda^{*\frac{M}{2}-1}e^{-\frac{\lambda^*}{2}}}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})}\frac{\left(P_F(\lambda^*)-\alpha\right)^2+\alpha\left(1-\alpha\right)}{P_F(\lambda^*)\left(1-P_F(\lambda^*)\right)\left(P_F(\lambda^*)-\alpha\right)}\right]^{-1} \quad (3.9)\end{aligned}$$

which is the analytical solution for iterative threshold finding by Newton's method.

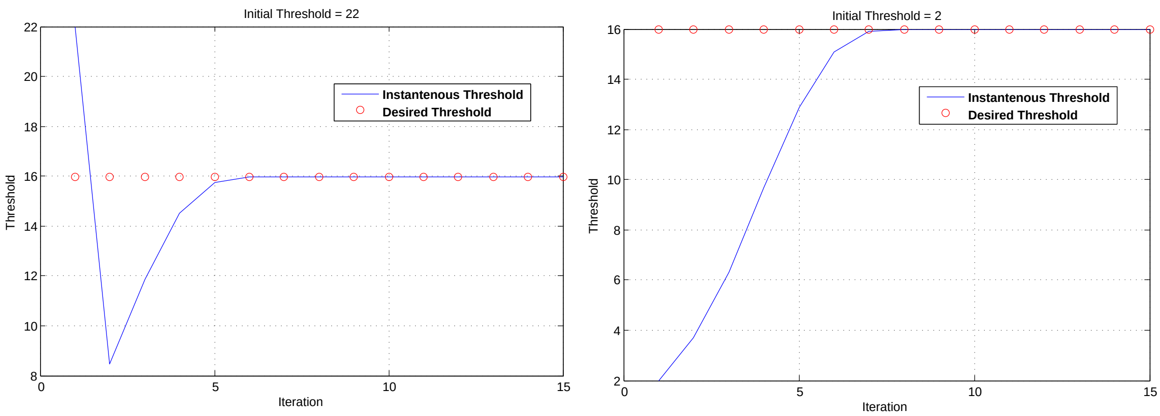


Figure 3.1. Plots of Newton's Method results for analytical $P_F(\lambda)$ function.

In figure 3.1, the first figure shows the algorithm for higher initial threshold. One can observe that by only one iteration, threshold drops from 22 to around 9 in value. This sudden quick move is a disadvantage for this method. When initial threshold is higher, threshold may even be negative which is an unwanted situation. $P_F(\lambda)$ and its derivatives are defined for non-negative λ values. Furthermore, initial threshold has to be given low and there has to be lower bound which prevents threshold from taking negative values.

3.1.2. Newton's Method for Estimated P_F

As analytical results works perfectly, the second step on the algorithm is to replace $P_F(\lambda^*)$ function with its estimate. Furthermore, sensing device does not necessarily calculate $P_F(\lambda)$ and becomes robust to small changes at the noise variance. Estimation of probability is executed by [39]

$$\hat{P}_F = \frac{1}{N_{win}} \sum_{i=1}^{N_{win}} u_i \tag{3.10}$$

where u_i is previous decisions ($u_i \in \{0, 1\}$) and N_{win} is number of decisions which are summed. The mean and variance of estimate is

$$E[\hat{P}_F] = P_F, \quad VAR(\hat{P}_F) = \tfrac{P_F(1-P_F)}{N_{win}} \tag{3.11}$$

which shows that there is no bias and the variance of the estimate decreases with the decrease of P_F and with the increase of N_{win}. Furthermore, adaptive system makes N_{win} detections before each threshold update. Threshold update is done by

$$\frac{\frac{\partial K(\lambda^*)}{\partial \lambda^*}}{\frac{\partial^2 K(\lambda^*)}{\partial \lambda^{*\,2}}} = \left[\frac{1}{\lambda^*}\left(\frac{M}{2} - 1\right) - \frac{1}{2} - \frac{\lambda^{*\,\frac{M}{2}-1} e^{-\frac{\lambda^*}{2}}}{2^{\frac{M}{2}} \Gamma(\frac{M}{2})} \frac{\left(\hat{P}_F - \alpha\right)^2 + \alpha(1-\alpha)}{\hat{P}_F(1-\hat{P}_F)(\hat{P}_F - \alpha))} \right]^{-1} \tag{3.12}$$

which is the same with Equation 3.9 except $P_F(\lambda)$ functions are replaced by its estimates.

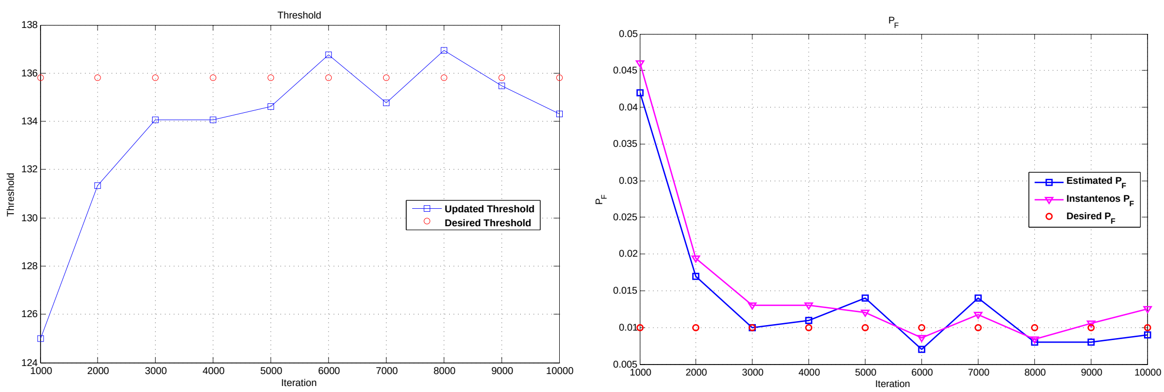


Figure 3.2. Plots of results of Newton's method for estimated $\hat{P}_F$ under constant $\sigma_n^2 = 1$ and $N_{win} = 1000$.

Figure 3.2, shows changes of threshold and P_F values by iteration. Since N_{win} is equal to 1000, threshold is updated at every $1000th$ iteration. Each iteration corresponds to a detection which results a new u_i. Figure 3.2b also denotes how good is the estimate of P_F. By increasing the quality of estimation, a more accurate threshold can be obtained.

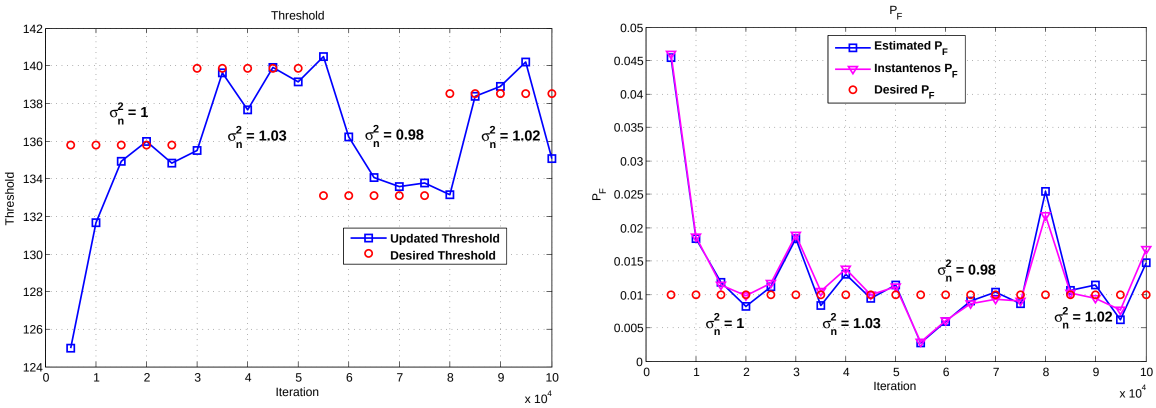


Figure 3.3. Plots of results of Newton's method for estimated $\hat{P}_F$ under varying variance around $\sigma_n^2 = 1$ and $N_{win} = 5000$.

Figure 3.3 illustrates how the algorithm behaves for small changes of variance. This time, P_F estimation is more accurate with higher N_{win}. Noise variance changes in every 2.5×10^4 iterations and the exact variance values are written on figures. Even though the desired P_F does not change, corresponding desired thresholds changes due to variance and the algorithm again trys to follow desired threshold values. In Figure 3.3b, there are bigger deviations after 2.5×10^4th, 5×10^4th, and 7.5×10^4th iterations because of sudden variance changes. But, the algorithm recovers the threshold at next threshold update.

3.2. Threshold Finding with Randomization

Threshold finding algotihms generally require estimation as previous algorithm based on Newton's method also does. Inevitably, after variance estimation they require to implement inverse of $P_F(\lambda)$ or at least functions which depend on $P_F(\lambda)$. In this section, an algorithm which does not require either any estimation or any mathematical function depends on $P_F(\lambda)$ is proposed. Without these requirements, this algorithm generates a new random variable and compares its previous decision to update the threshold and keep P_F probability around α. Although this algorithm is so simple to execute, it gives better performance than previous threshold finding algorithm with

Newton's method.

The proposed algorithm works at only H_0 hypothesis. Furthermore, this algorithm should be deployed in a time interval which is already known as H_0 or for the iterations when fusion center decided H_0 (by assuming fusion center has perfectly low false alarm probability). Threshold update is done for each iteration by

$$\lambda_{i+1} = \lambda_i + \lambda_{Add} \times [u_i - u_{des}] = \lambda_n + \lambda_{Add} \times d_i \tag{3.13}$$

where λ_{Add} is threshold addition, $u_{des} \in \{0, 1\}$ is newly generated random variable with

$$\Pr(u_{des} = 1) = \alpha \quad , \quad \Pr(u_{des} = 0) = 1 - \alpha, \tag{3.14}$$

and $d_i \in \{-1, 0, 1\}$ is random variable of direction with

$$\Pr(d_i = 1) = (1 - \alpha)P_{F,i} \tag{3.15}$$

$$\Pr(d_i = 0) = \alpha P_{F,i} + (1 - \alpha)(1 - P_{F,i}) \tag{3.16}$$

$$\Pr(d_i = -1) = \alpha(1 - P_{F,i}) \tag{3.17}$$

distribution. For constant $\alpha = 0.1$, $\Pr(d = 1)$ and $\Pr(d = -1)$ values are shown in Figure 3.4. When P_F is greater than α, $\Pr(d = 1)$ is greater than $\Pr(d = -1)$. Furthermore, P_F is more likely to decrease due to the increasing threshold. In the

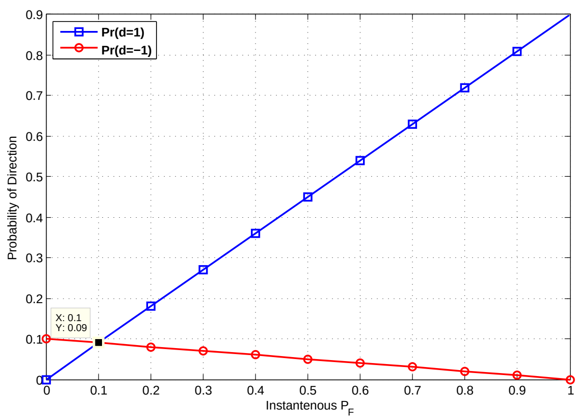


Figure 3.4. $\Pr(d{=}1)$ and $\Pr(d{=}-1)$ values according to instantenous P_F for $\alpha = 0.1$.

same way, when P_F is less than α, in the next detection P_F is more likely to be greater. Slopes of lines also result that initial threshold should be low for faster approach to α.

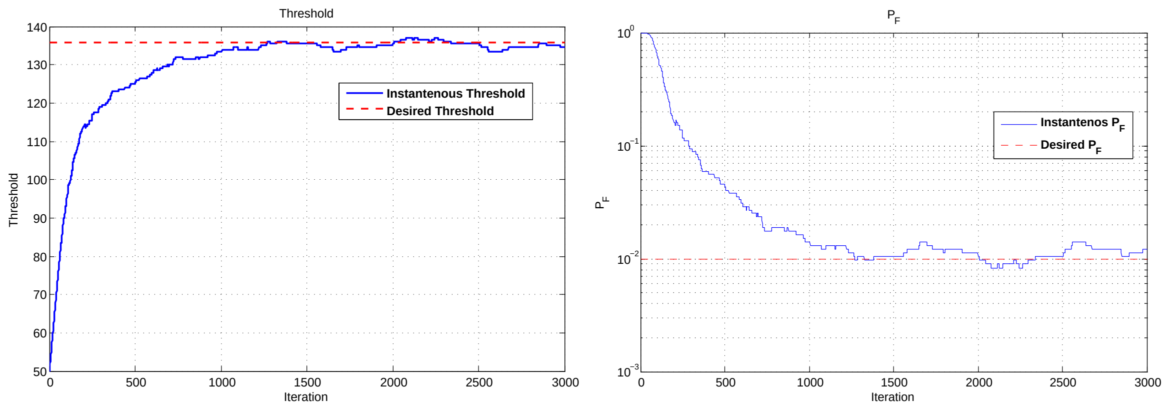


Figure 3.5. Plots of threshold and P_F of threshold finding with randomization for constant $\lambda_{Add} = 0.5$, and $\alpha = 0.01$.

Figure 3.5 presents threshold and P_F curves iteratively. The system starts from a low threshold and quickly converges to the desired level. Then, curves move around desired level and produce a variance. This variance is the result of randomness and the value of λ_{Add}. For high values of λ_{Add}, the algorithm converges fast but has higher variance of output. In same way, for low values of λ_{Add}, the algorithm behaves slower but has lower variance at convergence. To provide fast and less varying output, the update rule for λ_{Add} will be represented in the next subsection.

3.2.1. Updating λ_{Add}

Updating λ_{Add} is an important design parameter which may lead the algorithm unexpected circumstances because of randomization. First of all, update criterion should be stringent due to prevent updating rapidly. In the proposed algorithm, the update criterion is to check if the sum of previous $\frac{\theta}{\alpha}$ decisions equals θ. Probability of threshold update is

$$\Pr(\lambda_{Add} \Leftarrow \beta \times \lambda_{Add}) = \Pr\left(\sum_{i=1}^{\frac{\theta}{\alpha}} u_i = \theta\right) \tag{3.18}$$

where $0.5 < \beta < 1$ is update coefficient and θ is a positive integer which makes $\frac{\theta}{\alpha}$ integer. By assuming each $P_{F,i}$ equals to each other, Equation 3.18 becomes

$$\Pr\left(\sum_{i=1}^{\frac{\theta}{\alpha}} u_i = \theta\right) = \binom{\frac{\theta}{\alpha}}{\theta} P_F^{\theta}(1 - P_F)^{\frac{\theta}{\alpha}-\theta}. \tag{3.19}$$

By this update criterion, the maximum probability of decreasing threshold is at α. Equation 3.19 is drawn for different θ values Figure 3.6. As θ increases, the curve becomes narrower and pile more on α. However, $\Pr\left(\sum_{i=1}^{\frac{\theta}{\alpha}} u_i = \theta\right)$ decreases with increase in θ which makes the algorithm converge slowly. Furthermore, θ should be chosen well. For low, θ algorithm may decrease λ_{Add} too quick to zero and P_F may converge to an unwanted probability value randomly. Against undesired convergences, we propose a lower bound for λ_{Add}. By this lower bound (λ_{Add}^{Bound}), P_F has still ran-

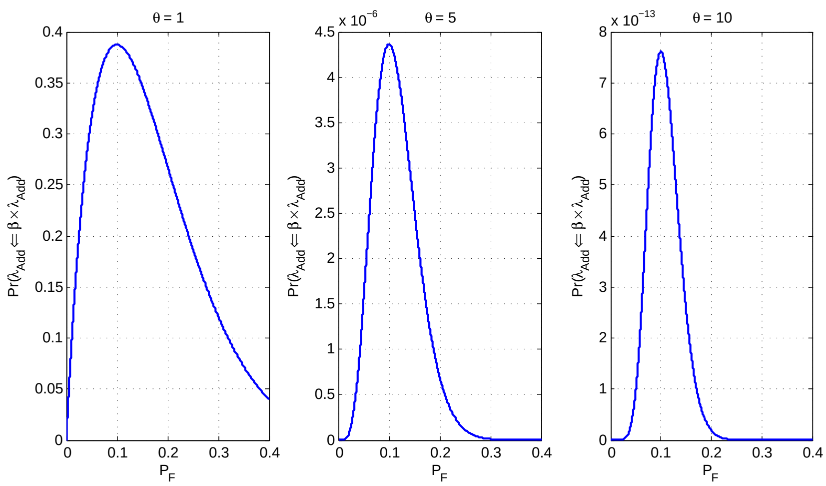


Figure 3.6. Plot of $\Pr(\lambda_{Add} \Leftarrow \beta \times \lambda_{Add})$ with respect to instantenous P_F for $\alpha = 0.1$.

domness when it converges. Then, P_F can still move to α from undesired converged level. Figure 3.7 presents the overall algorithm. This figure obviously shows how stringent the λ_{Add} rule is. k index prevents the algorithm to update λ_{Add} adjacently. As this index exists, after each λ_{Add} update, the algorithm waits $\frac{\theta}{\alpha}$ iterations to check for λ_{Add} update again. Another parameter is the previous decision of fusion center (u_{fc}).

This algorithm works when the channel known as H_0 hypothesis. Furthermore, we assume the probability of miss at fusion center ($P_{F,C}$) is so low and it is broadcasted as feedback from the fusion center to sensing nodes. This feedback should be sent before every threshold update, because the algorithm cannot work without it.

Initiate Select a initial value for λ_{Add};
$k = 0$;
while $u_{fc} = 0$ **do**
  Generate a discrete random variable u_{des} by Equation 3.14;
  Update threshold by Equation 3.13;
  if $k > \frac{\theta}{\alpha}$ **then**
    if $\lambda_{Add} \geq \lambda_{Add}^{Bound}$ **then**
      if $\sum_{i=1}^{\frac{\theta}{\alpha}} u_i = 1$ **then**
        $\lambda_{Add} \Leftarrow \beta \times \lambda_{Add}$;
        $k \Leftarrow 0$;
      end if
    end if
  end if
  $k \Leftarrow k + 1$;
end while

Figure 3.7. Threshold Finding Algorithm with Randomization.

Figure 3.8 shows the performance of algorithm with λ_{Add} update. As the algorithm approachs to α, it starts to decrease its λ_{Add}. Finally, it converges to the desired level with less variance than the algorithm in Figure 3.5. When the algorithm converges, it still has randomness around α due to λ_{Add}^{Bound}. If λ_{Add}^{Bound} would not exist, λ_{Add} would converge to zero and algorithm would converge as a smooth line. But, level of convergence may not be same with α. Both Figures 3.8 and 3.9 has extra curve of threshold finder by noise variance estimation to compare. In every iteration,

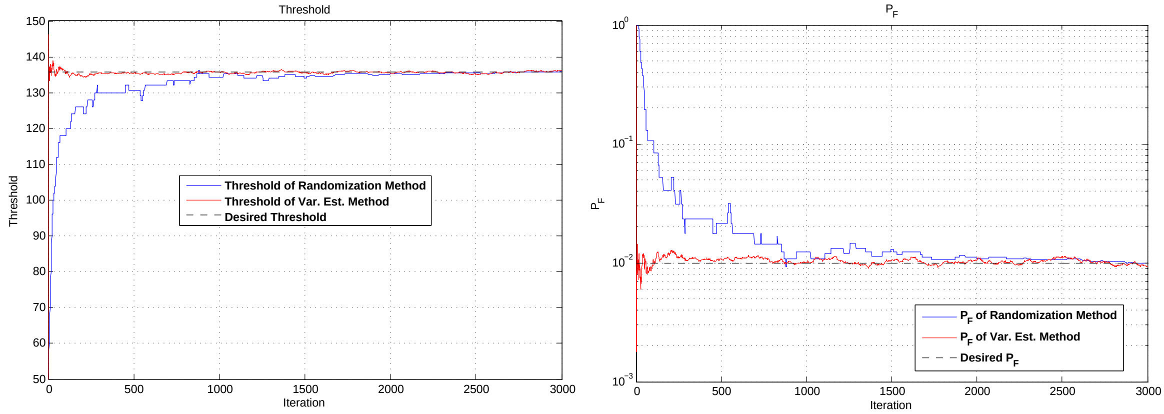


Figure 3.8. Plots of threshold and P_F of threshold finding with randomization for initial $\lambda_{Add} = 2$, $\beta = 0.7$, $\theta = 2$, $\lambda_{Add}^{Bound} = 0.1$, and $\alpha = 0.01$.

this threshold finder estimates the noise variance by summing the squares of previous samples and inserts this estimate and α into inverse of $P_F(\lambda)$ function to find threshold. Noise variance estimation is implemented by maximum 60000 previous samples. The reason of this much sample is to adapt the algorithm to our manual variance changes in Figure 3.9.

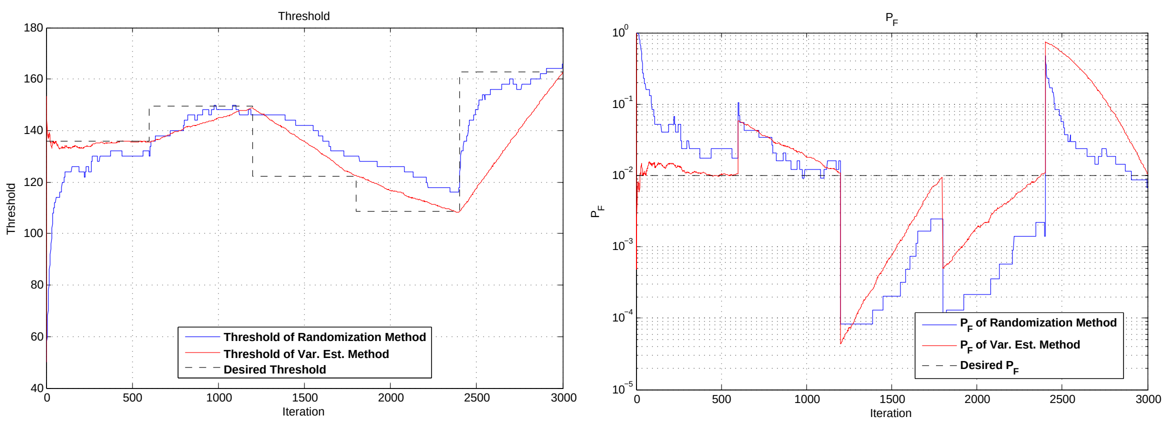


Figure 3.9. Plots of threshold and P_F of threshold finding with randomization for different variance values, constant $\lambda_{Add} = 2$, and $\alpha = 0.01$.

Figure 3.9 represents the algorithm under changing noise variance. The algorithm in this figures has constant λ_{Add}. This design is chosen to provide agility against changing noise variance. If the algorithm would lower λ_{Add}, agility of the algorithm to variance changes would be worse. Furthermore, when noise variance changes suddenly, algorithm instantly converges to a new desired threshold level. Those sudden changes in instantenous P_F is the result of sudden manually changes in the noise variance. Moreover, when desired the threshold becomes lower due to low variance, the algorithm moves slower (at regions of $\sigma_n^2 = 0.9$ and $\sigma_n^2 = 0.8$). The reason is again the slopes in Figure 3.4.

4. OPTIMUM QUANTIZATION

Quantization levels and thresholds are represented as,

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix} \quad , \quad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{m+1} \end{bmatrix} \tag{4.1}$$

where q_i values are quantization levels, λ_i values are quantization thresholds (except λ_1 and λ_{m+1}) and m is the number of quantization levels. λ_1 and λ_{m+1} are the lower and the upper boundaries of probability distributions. For ED, lower and upper boundaries of signal are $\lambda_1 = 0$ and $\lambda_{m+1} = \infty$

The optimization of quantization levels has its mathematical solution by maximization of deflection coefficient [47]. But the optimization of quantization thresholds is a harder problem and generally solved by numerical optimization algorithms. In the approach of this thesis, optimum quantization for distortion is derived by Lloyd-Max algorithm [60] and its performance on detection problem is observed.

According to [61], overall distribution of the output of the ED can be written as a mixture of two hypotheses.

$$\begin{aligned} f(x) &= \sum_{i=0}^{1} f(x|H_i)\,\Pr(H_i) \\ &= \pi_1 f_1(x) + \pi_0 f_0(x) \end{aligned} \tag{4.2}$$

where π_1 and π_0 are the *a priori* probabilities of H_1 and H_0 hypotheses, respectively, and $\pi_1 + \pi_0$ equals 1. The approach of optimization is derived by taking the overall PDF of the sum of PDFs of two hypotheses. By minimizing the mean square error (MSE) of distribution $f(x)$, quantizer chooses quantization levels which represents the exact signal (includes both hypotheses) by minimizing the MSE.

MSE is

$$MSE = E[(x-q_i)^2] = \sum_{i=1}^{m} \int_{\lambda_i}^{\lambda_{i+1}} (x-q_i)^2 f(x)dx. \tag{4.3}$$

The approach of optimization is derived by taking the overall PDF of sum of two hypotheses. Optimization of MSE is derived by taking the derivatives

$$\begin{aligned}
\frac{\partial MSE}{\partial \lambda_k} &= \frac{\partial}{\partial \lambda_k}\Big[\int_{\lambda_k}^{\lambda_{k+1}} (x-q_k)^2 f(x)dx + \int_{\lambda_{k-1}}^{\lambda_k} (x-q_{k-1})^2 f(x)dx\Big] \\
&= \frac{\partial}{\partial \lambda_k}\Big[\Big(\frac{\partial^{-1}}{\partial x^{-1}}(x-q_k)^2 f(x)\Big)\Big|_{\lambda_k}^{\lambda_{k+1}} + \Big(\frac{\partial^{-1}}{\partial x^{-1}}(x-q_{k-1})^2 f(x)\Big)\Big|_{\lambda_{k-1}}^{\lambda_k}\Big] \\
&= \frac{\partial}{\partial \lambda_k}\Big[\Big(\frac{\partial^{-1}}{\partial x^{-1}}(x-q_k)^2 f(x)\Big)\Big|_{\lambda_k} - \Big(\frac{\partial^{-1}}{\partial x^{-1}}(x-q_{k-1})^2 f(x)\Big)\Big|_{\lambda_k}\Big] \\
&= \Big[(\lambda_k - q_k)^2 - (\lambda_k - q_{k-1})^2\Big] f(\lambda_k) \\
0 &= 2\lambda_k(q_k + q_{k-1}) + q_k^2 - q_{k-1}^2 \\
\lambda_k &= \frac{q_k + q_{k-1}}{2}
\end{aligned} \tag{4.4}$$

and

$$\begin{aligned}
\frac{\partial MSE}{\partial q_k} &= \frac{\partial}{\partial q_k}\Big[\int_{\lambda_k}^{\lambda_{k+1}} (x-q_k)^2 f(x)dx\Big] \\
&= \frac{\partial}{\partial q_k}\Big[-2q_k \int_{\lambda_k}^{\lambda_{k+1}} x f(x)dx + q_k^2 \int_{\lambda_k}^{\lambda_{k+1}} f(x)dx\Big] \\
0 &= -2\int_{\lambda_k}^{\lambda_{k+1}} x f(x)dx + 2q_k \int_{\lambda_k}^{\lambda_{k+1}} f(x)dx \\
q_k &= \frac{\int_{\lambda_k}^{\lambda_{k+1}} x f(x)dx}{\int_{\lambda_k}^{\lambda_{k+1}} f(x)dx}.
\end{aligned} \tag{4.5}$$

These derivatives show that optimum thresholds for minimum MSE ($MMSE$) are arithmetic average of succesive quantization levels which are calculated by integrating the distribution over integrals defined by succesive quantization thresholds. These two results are identical to [60].

Initial Values: Select or estimate initial λ_i values;
Calculate q_i values by (4.5);
while $(new\, MSE) \neq (previous\, MSE)$ **do**
  Calculate λ_i values by (4.4);
  Calculate q_i values by (4.5);
  Calculate MSE by (4.3) for checking converged MSE;
end while

Figure 4.1. Lloyd-Max Algorithm.

Figure 4.1 shows Lloyd-Max algorithm. By implementing equations, MSE converges to its minimum value. The stopping rule is chosen by checking the convergence of MSE. Because the PDF consists of two distributions of hypotheses, Lloyd-Max algorithm may converge to locally optimum values of MSE. Furthermore, initial values should be chosen around mean values of the two hypotheses.

For two hypotheses, Equation 4.5 becomes

$$q_k = \frac{\pi_1 \int_{\lambda_k}^{\lambda_{k+1}} x f_1(x)dx + \pi_0 \int_{\lambda_k}^{\lambda_{k+1}} x f_0(x)dx}{\pi_1 \int_{\lambda_k}^{\lambda_{k+1}} f_1(x)dx + \pi_0 \int_{\lambda_k}^{\lambda_{k+1}} f_0(x)dx}. \tag{4.6}$$

In the following sections, derivation of this formula under Gaussian and chi-square distributions are introduced.

4.1. Lloyd-Max Algortihm for Gaussian Hypotheses

Solution of the optimum quantization depends on the evaluation of Equation 4.5. In this section, derivation for two Gaussian hypotheses is introduced. The integrals in the denominator of Equation 4.6 are calculated by Q function. The numerator integrals of Equation 4.6 can be calculated on

$$\begin{aligned}
\int_b^a xg(x)dx &= \int_b^a \frac{x}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x-\mu)^2}{2\sigma^2})dx \\
&= \int_b^a \left[\frac{x-\mu}{\sqrt{2\pi\sigma^2}} exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big) + \frac{\mu}{\sqrt{2\pi\sigma^2}} exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)\right] dx \\
&= \left[-\frac{\sigma}{\sqrt{2\pi}} exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)\right]_b^a + \mu\left[Q\Big(\frac{b-\mu}{\sigma}\Big) - Q\Big(\frac{a-\mu}{\sigma}\Big)\right] \\
&= \sigma^2\big[g(b) - g(a)\big] + \mu\big[\Pr(a > x \geq b)\big] \qquad (4.7)
\end{aligned}$$

where $g(x)$ is the pdf of Gaussian and $\Pr(a > x \geq b)$ is the probability of that x is between a and b.

By placing solutions of integrals to Equation 4.6,

$$q_i = \frac{\pi_1\sigma_1^2[f_1(\lambda_i) - f_1(\lambda_{i+1})] + \pi_1\mu_1 p_1(q_i) + \pi_0\sigma_0^2[f_0(\lambda_i) - f_0(\lambda_{i+1})] + \pi_0\mu_0 p_0(q_i)}{\pi_1 p_1(q_i) + \pi_0 p_0(q_i)} \qquad (4.8)$$

where $p_j(q_i) = \Pr(\lambda_{i+1} > x \geq \lambda_i)$ for hypothesis $H_j$and $f_1(x)$ and $f_0(x)$ are Gaussian PDFs of H_1 and H_0 hypotheses, respectively.

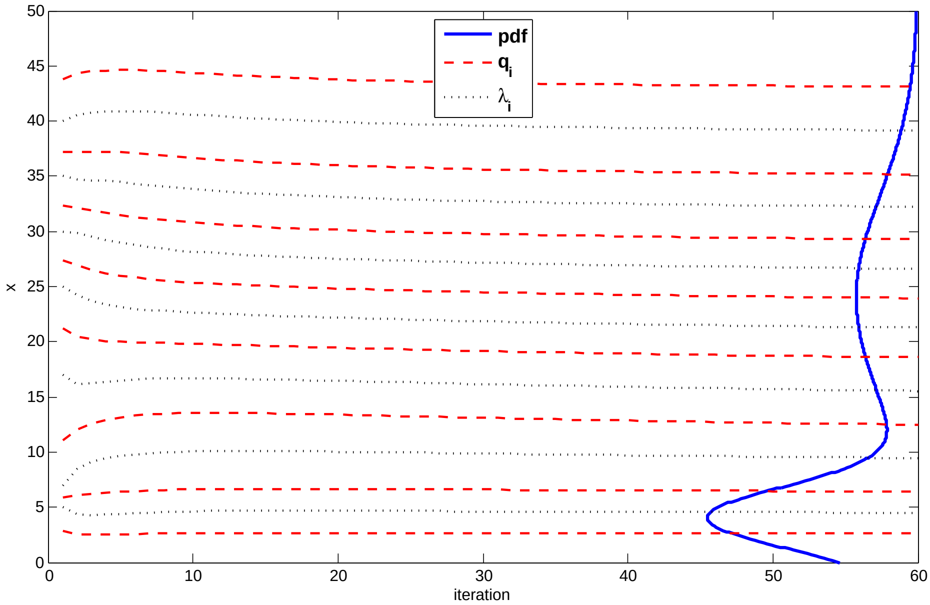


Figure 4.2. Implementation of Lloyd-Max algorithm for two Gaussian hypotheses, $H_0 \sim \mathcal{N}(4, 8)$, $H_1 \sim \mathcal{N}(24, 88)$, and $\pi_1 = \pi_0 = 0.5$.

In the Figure 4.2, iterative implementation of algorithm in Figure 4.1 is presented for two Gaussian hypotheses. Plots of the representational PDF (Equation 4.2) is included to show where the algorithm converges.

4.1.1. 1 Bit Quantization for Gaussian

The one bit analytical solution ($\lambda_1 = -\infty > q_1 > \lambda_2 > q_2 > \lambda_3 = \infty$) is considered in this section. By using Equations 4.5 and 4.7, the quantization level, q_i, for a single Gaussian distribution is

$$\begin{aligned} q_i &= \frac{\int_{\lambda_k}^{\lambda_{k+1}} x f(x) dx}{\int_{\lambda_k}^{\lambda_{k+1}} f(x) dx} \\ &= \frac{\sigma^2[f(\lambda_i) - f(\lambda_{i+1})] + \mu[\Pr(\lambda_{i+1} > x \geq \lambda_i)]}{\Pr(\lambda_{i+1} > x \geq \lambda_i)} \\ &= \sigma^2 \frac{f(\lambda_i) - f(\lambda_{i+1})}{\Pr(\lambda_{i+1} > x \geq \lambda_i)} + \mu. \end{aligned} \tag{4.9}$$

Then, by inserting this formula into equation 4.4,

$$\begin{aligned} \lambda_2 &= \frac{q_1 + q_2}{2} \\ &= \frac{\sigma^2}{2}\left[\frac{f(-\infty) - f(\lambda_2)}{Q(-\infty) - Q(\frac{\lambda_2 - \mu}{\sigma})} - \frac{f(\lambda_2) - f(\infty)}{Q(\frac{\lambda_2 - \mu}{\sigma}) - Q(\infty)}\right] + \mu \\ &= \frac{\sigma^2 f(\lambda_2)}{2}\left[\frac{1 - 2Q(\frac{\lambda_2 - \mu}{\sigma})}{[1 - Q(\frac{\lambda_2 - \mu}{\sigma})]Q(\frac{\lambda_2 - \mu}{\sigma})}\right] + \mu \\ &= \mu \end{aligned} \tag{4.10}$$

which gives the optimum threshold for a single Gaussian distribution. Then, quantization levels can be calculated by

$$q_1 = \sigma^2\left[\frac{f(-\infty) - f(\lambda_2)}{Q(-\infty) - Q(\frac{\lambda_2 - \mu}{\sigma})}\right] + \mu = \sigma\frac{1}{1/2}\frac{-1}{\sqrt{2\pi\sigma^2}} + \mu = \mu - \sqrt{\frac{2\sigma^2}{\pi}} \tag{4.11}$$

$$q_2 = \mu + \sqrt{\frac{2\sigma^2}{\pi}}. \tag{4.12}$$

which is also a result from [50]. For two Gaussian hypotheses by using Equation 4.4,

$$
\begin{aligned}
\lambda_2 &= \frac{1}{2}\frac{\pi_1\sigma_1^2[f_1(-\infty)-f_1(\lambda_2)]+\pi_1\mu_1 p_1(q_1)+\pi_0\sigma_0^2[f_0(-\infty)-f_0(\lambda_2)]+\pi_0\mu_0 p_0(q_1)}{\pi_1 p_1(q_1)+\pi_0 p_0(q_1)} \\
&+\frac{1}{2}\frac{\pi_1\sigma_1^2[f_1(\lambda_2)-f_1(\infty)]+\pi_1\mu_1 p_1(q_2)+\pi_0\sigma_0^2[f_0(\lambda_2)-f_0(\infty)]+\pi_0\mu_0 p_0(q_2)}{\pi_1 p_1(q_2)+\pi_0 p_0(q_2)} \\
&= \frac{2k-1}{2k(1-k)}[\pi_1\sigma_1^2 f_1(\lambda_2)+\pi_0\sigma_0^2 f_0(\lambda_2)+\pi_1\mu_1 p_1(q_2)+\pi_0\mu_0 p_0(q_2)] \\
&\qquad +\frac{\pi_1\mu_1+\pi_0\mu_0}{2k}
\end{aligned}
\tag{4.13}
$$

where $k=\pi_1 p_1(q_1)+\pi_0 p_0(q_1)$ and $(1-k)=\pi_1 p_1(q_2)+\pi_0 p_0(q_2)$. If $\sigma_1^2=\sigma_0^2$ and $\pi_1=\pi_0$, solution of this equation is $\lambda_2=\frac{\mu_1+\mu_0}{2}$. Otherwise, the solution can be acquired by numerical methods on root finding. q_1 and q_2 can be found by putting λ_2 and other parameters to Equation 4.8.

4.2. Lloyd-Max Algorithm for Chi and Noncentral Chi Square Hypotheses

As the PDFs of chi and noncentral chi square distributions are represented in Equations 2.14 and 2.15, derivation of integrals at formula of q_i (Equation 4.6) is introduced in this section.

Now, we will calculate four integrals which are required to calculate Equation 4.6. The first integral at the denominator of Equation 4.6 is

$$
\begin{aligned}
\int_{\lambda_i}^{\lambda_{i+1}} f_0(x)dx &= \int_{\lambda_i}^{\infty} f_0(x)dx - \int_{\lambda_{i+1}}^{\infty} f_0(x)dx \\
&= \gamma_u\left(\frac{M}{2},\frac{\lambda_i}{2}\right) - \gamma_u\left(\frac{M}{2},\frac{\lambda_{i+1}}{2}\right)
\end{aligned}
\tag{4.14}
$$

where $\gamma_u(a,b)$is the regularized upper incomplete gamma function (Appendix A). The second integral is

$$
\int_{\lambda_i}^{\lambda_{i+1}} f_1(x)dx = Q_{\frac{M}{2}}(\sqrt{M\gamma},\sqrt{\lambda_i}) - Q_{\frac{M}{2}}(\sqrt{M\gamma},\sqrt{\lambda_{i+1}}) \tag{4.15}
$$

where $Q_m(a,b)$ is Marcum Q Function [55]. This equation is widely used to describe the cumulative distribution function (CDF) of a noncentral chi-square distribution. Here, we will derive another expression which depends on the incomplete gamma function.

$$
\begin{aligned}
\int_{\lambda_i}^{\lambda_{i+1}} f_1(x)dx &= \int_{\lambda_i}^{\infty} f_1(x)dx - \int_{\lambda_{i+1}}^{\infty} f_1(x)dx \\
\int_{\lambda_i}^{\infty} f_1(x)dx &= \frac{1}{2}\exp\Big(-\frac{M\gamma}{2}\Big)(M\gamma)^{\frac{1}{2}-\frac{M}{4}} \int_{\lambda_i}^{\infty} e^{-\frac{x}{2}} x^{\frac{M}{4}-\frac{1}{2}} I_{\frac{M}{2}-1}(\sqrt{M\gamma x})dx \\
&= \frac{1}{2}\exp\Big(-\frac{M\gamma}{2}\Big)(M\gamma)^{\frac{1}{2}-\frac{M}{4}} \sum_{k=0}^{\infty} \frac{(\frac{M\gamma}{4})^{\frac{M}{4}-\frac{1}{2}+k}\int_{\lambda_i}^{\infty} e^{-\frac{x}{2}} x^{\frac{M}{2}-1+k}dx}{k!\Gamma(\frac{M}{2}+k)} \\
&= \frac{1}{2}\exp\Big(-\frac{M\gamma}{2}\Big)(M\gamma)^{\frac{1}{2}-\frac{M}{4}} \sum_{k=0}^{\infty} \frac{(\frac{M\gamma}{4})^{\frac{M}{4}-\frac{1}{2}+k} 2^{\frac{M}{2}+k}\Gamma_u(\frac{M}{2}+k,\frac{\lambda_i}{2})}{k!\Gamma(\frac{M}{2}+k)} \\
&= \exp\Big(-\frac{M\gamma}{2}\Big)\sum_{k=0}^{\infty}\Big(\frac{M\gamma}{2}\Big)^k \frac{1}{k!}\gamma_u\Big(\frac{M}{2}+k,\frac{\lambda_i}{2}\Big) \\
\int_{\lambda_i}^{\lambda_{i+1}} f_1(x)dx &= \exp\Big(-\frac{M\gamma}{2}\Big)\sum_{k=0}^{\infty}\Big(\frac{M\gamma}{2}\Big)^k \frac{1}{k!} \\
&\qquad \times\Big[\gamma_u\Big(\frac{M}{2}+k,\frac{\lambda_i}{2}\Big) - \gamma_u\Big(\frac{M}{2}+k,\frac{\lambda_{i+1}}{2}\Big)\Big]
\end{aligned} \tag{4.16}
$$

The first integral of the numerator of q_i is

$$
\begin{aligned}
\int_{\lambda_i}^{\lambda_{i+1}} xf_0(x)dx &= \int_{\lambda_i}^{\infty} xf_0(x)dx - \int_{\lambda_{i+1}}^{\infty} xf_0(x)dx \\
&= \frac{1}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})}\Big[\int_{\lambda_i}^{\infty} x^{\frac{M}{2}} e^{-\frac{x}{2}}dx - \int_{\lambda_{i+1}}^{\infty} x^{\frac{M}{2}} e^{-\frac{x}{2}}dx\Big] \\
&= M\Big[\gamma_u\Big(\frac{M}{2}+1,\frac{\lambda_i}{2}\Big) - \gamma_u\Big(\frac{M}{2}+1,\frac{\lambda_{i+1}}{2}\Big)\Big]
\end{aligned} \tag{4.17}
$$

which has a solution similar to Equation 4.14. The calculation of second integral of the numerator is similar to Equation 4.16.

$$\int_{\lambda_i}^{\lambda_{i+1}} x f_1(x)dx = 2\exp\left(-\frac{M\gamma}{2}\right)\sum_{k=0}^{\infty}\left(\frac{M\gamma}{2}\right)^k \frac{\left(\frac{M}{2}+k\right)}{k!} \left[\gamma_u\left(\frac{M}{2}+k+1,\frac{\lambda_i}{2}\right)-\gamma_u\left(\frac{M}{2}+k+1,\frac{\lambda_{i+1}}{2}\right)\right]. \tag{4.18}$$

Then, calculation of q_i can be done by putting Equations 4.14, 4.16, 4.17, and 4.18 into Equation 4.6. Infinite sums in equations can be approximated closely by finite summation consisting of 60-70 elements, which is decided by trial and error. One may notice that PDF in Figure 4.2 shows the Gaussian approximation of PDF in Figure 4.3.

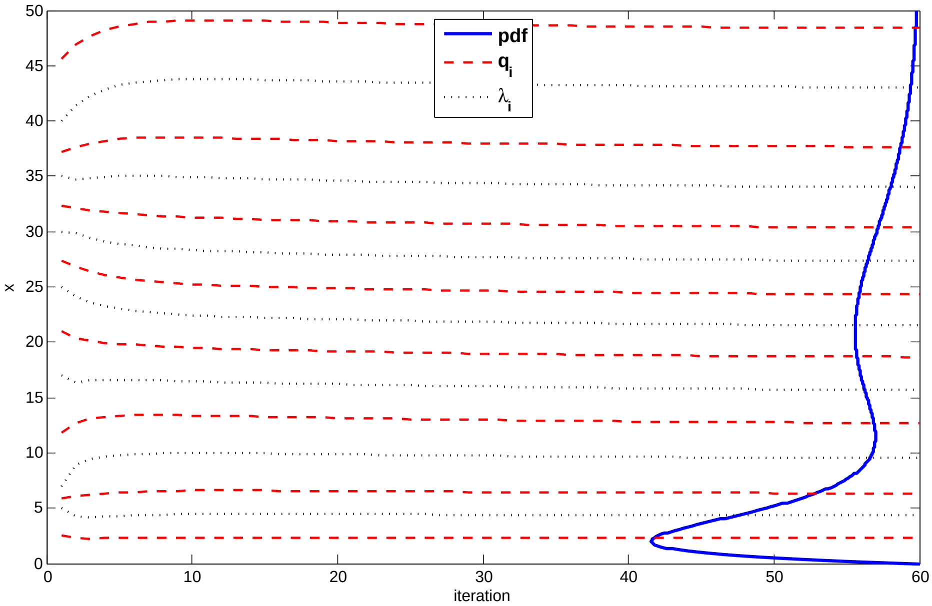


Figure 4.3. Implementation of Lloyd-Max algorithm on chi and noncentral chi square hypotheses, $M = 4$, $\gamma = 5$ and $\pi_1 = \pi_0 = 0.5$.

4.3. Quantization According to Noise Power

In Sections 2.1 and 2.1.1, the formulation of the output of the ED may be without normalization of noise power. In this case, noise power should affect the quantization levels and the thresholds. This problem can be evaluated by transformation $y = \sigma_n^2 x$. As a result of this transformation, new quantization values are $q_{i,new} = \sigma_n^2 q_i$ and $\lambda_{i,new} = \sigma_n^2 \lambda_i$ which can be adjusted in Equations 2.39, 4.4, and 4.5.

4.4. Simulation Results

The P_D results, which are presented in this section, are acquired by simulations. All simulations in this section realize a system consisting of one fusion center and 10 CR nodes ($N = 10$). The P_D value of every point in the figures is calculated by 10,000,000 runs. For each point, inverse of $P_{F,C}(\lambda_C)$ is calculated by numerical method to get desired threshold for desired $P_{F,C}$ value according to Neyman-Pearson detection rule.

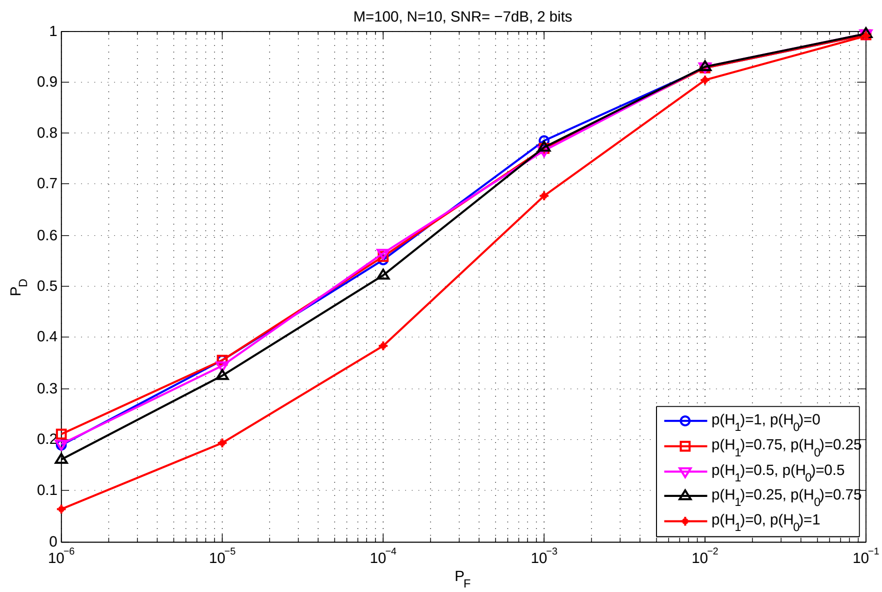


Figure 4.4. ROC curves for 2 bits MMSE quantization with different a priori probabilities, $M = 100$, $SNR = -7$dB, and $N = 10$.

In Figure 4.4, MMSE quantization is performed for different *a priori* probability values. One may notice that the algorithms which take that π_0 is higher than 0.5 perform worse. Quantization of those algorithms dominates on the distribution of H_0. On the other hand, when π_1 is higher than 0.5, performance is not affected vitally. This shows the quantization of the distribution of H_1 has more importance and the *a priori*

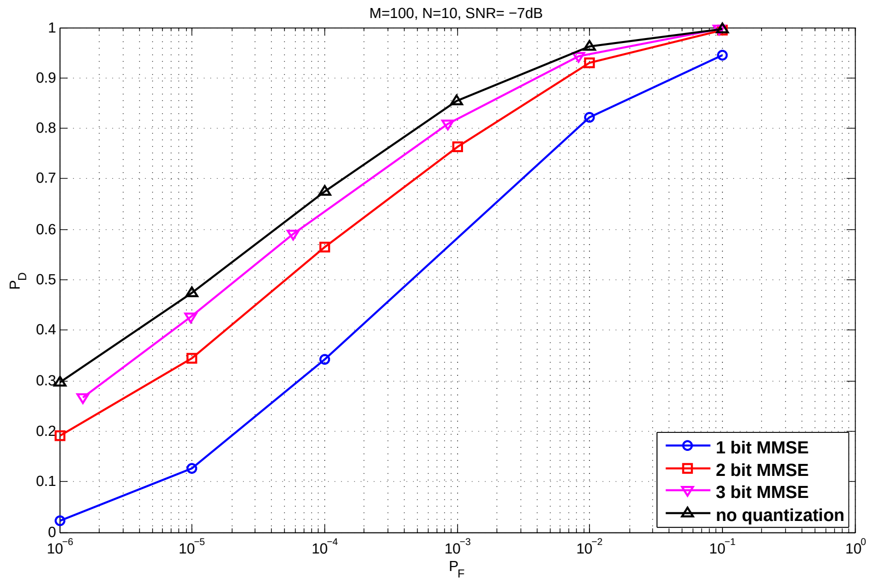


Figure 4.5. ROC curves for MMSE quantization with different bits, $M = 100$, $SNR = -7$dB, $N = 10$, and $\pi_1 = \pi_0 = 0.5$.

probabilities should be chosen as $1 - \pi_0 = \pi_1 \geq 0.5$. The reason of this is the higher variance of H_1 distribution. The Figure 4.5 shows performance improvement of the MMSE by increase number of bits. The 1 bit case has obviously worse performance. What makes that curve apart from the others is parameters of simulation environment (M, N, SNR).

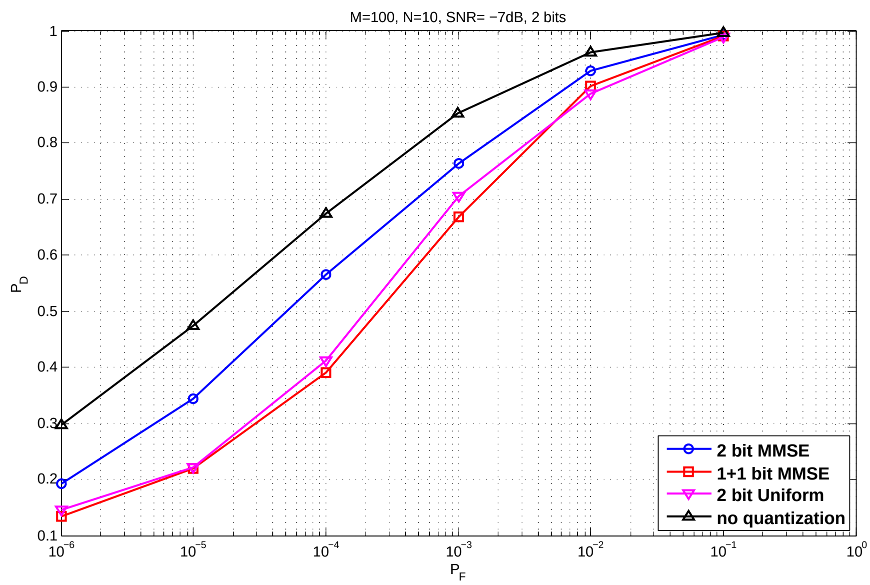


Figure 4.6. ROC curves for different quantization models, $M = 100$, $SNR = -7$dB, $N = 10$, and $\pi_1 = \pi_0 = 0.5$.

Figures 4.6 and 4.7 introduce comparison of MMSE with some other quantization models. The uniform quantization curves are implemented by same size quantization intervals between thresholds and quantization levels are in the middle of those intervals. λ_2 and λ_4 are assumed to be mean values of two distributions for positioning. In $1+1$ bit MMSE quantization, Lloyd-Max algorithm is used to find quantization for each hypothesis separately and a threshold which is at the average of the means of

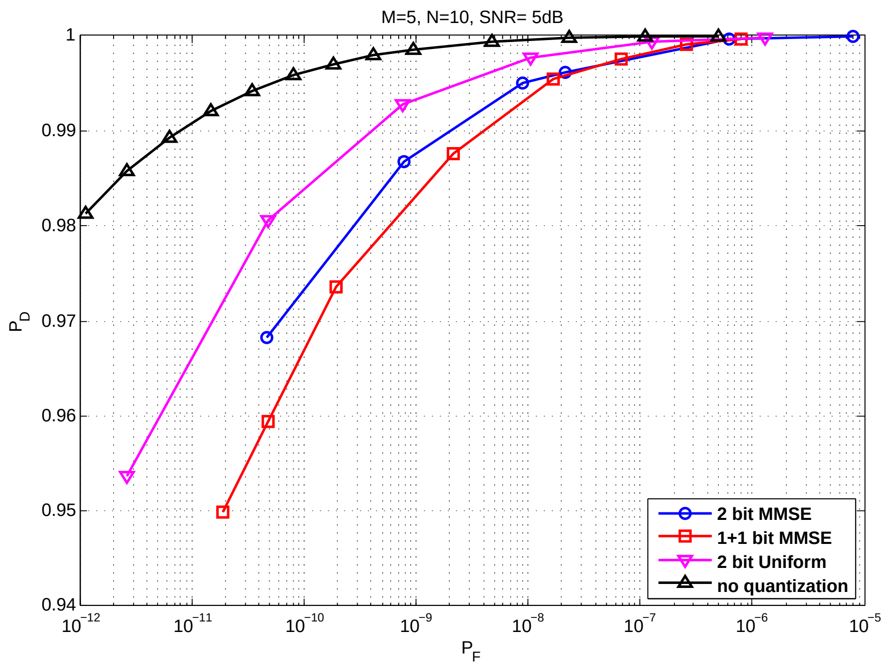


Figure 4.7. ROC curves for different quantization models, $M = 5$, $SNR = 5$dB, $N = 10$, and $\pi_1 = \pi_0 = 0.5$.

distributions is used for extra 1 bit. No quantization curves show the performance of the soft data which has higher performance as expected. Figure 4.6 shows how good the performance of 2 bit quantization with squared error distortion is at low SNR case. In Figure 4.7, the performance of MMSE quantizer at high SNR is evaluated. As SNR and the distance between hypotheses increase, performance of 2 bit quantization with squared error distortion becomes worse than low SNR case. For $M = 5$, distributions may not be a good approximation for signal energy, but this result is included to show the effect of high SNR. Since multi-bit quantization requirement is only needed at low SNR case, this result does not affect overall performance deeply. The randomization r parameter is not calculated in this figure due to its comptational difficulty at extremely low P_F values.

5. NUMBER OF BITS OF QUANTIZED OBSERVATIONS

This chapter focuses on proposing an analytical expression on how many bits should sensing nodes quantize their observations to.

5.1. Information Theoretical Approach

The question of how many bits should we quantize is widely studied as rate-distortion theory in quantization literature. Rate-distortion function [50] is given as,

$$R(D) = \min_{\sum p(x)p(q|x)d(x,q) \leq D} I(X;q) \tag{5.1}$$

where D is maximum allowable expected distortion and $I(.;.)$ is the mutual information. Rate-distortion function monotonically increases while distortion decreases. For the quantization of a continuous signal, MSE is generally prefered as the distortion measure. In this thesis, primarily MSE optimization will be evaluated for ED.

5.1.1. Gaussian Hypotheses

Rate distortion function of a single Gaussian distribution [50] for $MSE \leq D$ is given as

$$\begin{aligned} R(D) &= h(X) - h(X|q) \\ &\geq \frac{1}{2}\log\left(\frac{\sigma^2}{D}\right) \end{aligned} \tag{5.2}$$

which only depends on maximum expected distortion level and the variance of the Gaussian distribution. In this equation, $h(.)$ is differential entropy and $h(.|.)$ is conditional differential entropy.

In our approach, we will insert mixed distribution of

$$f(x) = \pi_1 f_1(x) + \pi_0 f_0(x) \tag{5.3}$$

into Equation 5.2. Since calculation of the entropy of mixed distribution is a hard task, we will derive a lower bound on their entropy by using concavity of entropy on probability distribution.

$$\begin{aligned}
I(X;q) &\geq h(X) - h(\mathcal{N}(0, E(X-q)^2)) \\
&\geq \pi_1 h_1(x) + \pi_0 h_0(x) - h(N(0, E(X-q)^2)) \\
&= \frac{\pi_1}{2}\log(2\pi e\sigma_1^2) + \frac{\pi_0}{2}\log(2\pi e\sigma_0^2) - \frac{1}{2}\log(2\pi e D) \\
&= \frac{1}{2}\log\left(\frac{(\sigma_1^2)^{\pi_1}(\sigma_0^2)^{\pi_0}}{D}\right)
\end{aligned} \tag{5.4}$$

where $h_1(x)$ and $h_0(x)$ are entropies of $f_1(x)$ and $f_0(x)$ distributions.

5.1.2. For Chi and Noncentral Chi-Square Hypotheses

After deriving lower bound for rate distortion theory, we will extend our derivations for entropy of chi-square and noncentral chi-square distributions. Entropy of chi-square distribution is

$$\begin{aligned}
h_0(x) &= -\int_0^\infty \frac{x^{\frac{M}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})} log\left(\frac{x^{\frac{M}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})}\right)dx \\
&= log\Big(2^{\frac{M}{2}}\Gamma\Big(\frac{M}{2}\Big)\Big) - \int_0^\infty \frac{x^{\frac{M}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})} log(x)^{\frac{M}{2}-1}dx + \int_0^\infty \frac{x^{\frac{M}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})}\frac{x}{2}dx \\
&= log\Big(2^{\frac{M}{2}}\Gamma\Big(\frac{M}{2}\Big)\Big) + \frac{M}{2} + \Big(1-\frac{M}{2}\Big)\int_0^\infty \frac{x^{\frac{M}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}}\Gamma(\frac{M}{2})} log(x)dx \\
&= log\Big(2^{\frac{M}{2}}\Gamma\Big(\frac{M}{2}\Big)\Big) + \frac{M}{2} + \Big(1-\frac{M}{2}\Big)\Big[\psi\Big(\frac{M}{2}\Big) + log2\Big] && (5.5) \\
&= \frac{M}{2} + log\Big(2\Gamma\Big(\frac{M}{2}\Big)\Big) + \Big(1-\frac{M}{2}\Big)\psi\Big(\frac{M}{2}\Big) && (5.6)
\end{aligned}$$

where $\psi(.)$ is digamma (psi) function which is $\psi(x) = \frac{d}{dx}\ln\Gamma(x)$. The integral solution at Equation 5.5 is from 4.352.1st Equation of [62] on page 573.

Since we cannot find exact entropy of noncentral chi-square distribution, we derive a lower bound for it by using the concavity of entropy. The Equation 2.14 can be written as

$$f_1(x) = \sum_{k=0}^{\infty} \theta_k \frac{1}{\theta_k} \frac{exp\Big(-\frac{x+M\gamma}{2}\Big)}{2^{\frac{M}{2}+k}k!\Gamma(\frac{M}{2}+k)} \Big(\frac{M\gamma}{2}\Big)^k x^{\frac{M}{2}+k-1} \tag{5.7}$$

where

$$f_1(x) = \sum_{k=0}^{\infty} \theta_k = 1. \tag{5.8}$$

By using concavity of entropy on the probability distribution, we have

$$h_1(x) \geq \sum_{k=0}^{\infty} \theta_k \zeta(k) \tag{5.9}$$

where

$$\begin{aligned}
\zeta(k) &= -\frac{1}{\theta_k}\int_0^{\infty} \frac{exp\big(-\frac{x+M\gamma}{2}\big)\big(\frac{M\gamma}{2}\big)^k x^{\frac{M}{2}+k-1}}{2^{\frac{M}{2}+k}k!\Gamma(\frac{M}{2}+k)} log\Bigg(\frac{1}{\theta_k}\frac{exp\big(-\frac{x+M\gamma}{2}\big)\big(\frac{M\gamma}{2}\big)^k x^{\frac{M}{2}+k-1}}{2^{\frac{M}{2}+k}k!\Gamma(\frac{M}{2}+k)}\Bigg)dx \\
&= -\frac{1}{\theta_k} exp\Big(-\frac{M\gamma}{2}\Big)\frac{1}{k!}\Big(\frac{M\gamma}{2}\Big)^k \int_0^{\infty} \frac{x^{\frac{M}{2}+k-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}+k}\Gamma(\frac{M}{2}+k)} \\
&\quad \cdot\Bigg[log\Big(\frac{1}{\theta_k}exp\Big(-\frac{M\gamma}{2}\Big)\frac{1}{k!}\Big(\frac{M\gamma}{2}\Big)^k\Big) + log\Big(\frac{x^{\frac{M}{2}+k-1}e^{-\frac{x}{2}}}{2^{\frac{M}{2}+k}\Gamma(\frac{M}{2}+k)}\Big)\Bigg]dx \qquad (5.10) \\
&= \frac{1}{\theta_k}exp\Big(-\frac{M\gamma}{2}\Big)\frac{1}{k!}\Big(\frac{M\gamma}{2}\Big)^k\Bigg[log(\theta_k) + \frac{M(\gamma+1)}{2} + log\Bigg(2e^k k!\Gamma\Big(\frac{M}{2}+k\Big)\Big(\frac{2}{M\gamma}\Big)^k\Bigg) \\
&\quad +\Big(1-\frac{M}{2}-k\Big)\psi\Big(\frac{M}{2}+k\Big)\Bigg]. \qquad (5.11)
\end{aligned}$$

In Equation 5.10, the first term after integral before logarithms is the pdf of chi square distribution and its integral is 1 from zero to infinity. One can notice that, the degree of freedom of chi-square distributin is $M + \frac{k}{2}$, which may not be an integer. Then, its integral between zero and infinity is still 1 due to,

$$\begin{aligned} \int_0^\infty f_0(x)dx &= \int_0^\infty \frac{x^{\frac{a}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{a}{2}}\Gamma(\frac{a}{2})}dx \\ &= \frac{1}{2^{\frac{a}{2}}\Gamma(\frac{M}{2})}\frac{1}{(\frac{1}{2})^{\frac{a}{2}}}\Gamma\Big(\frac{M}{2}\Big) && (5.12) \\ &= 1 && (5.13) \end{aligned}$$

where Equation 5.12 is derived by 3.381.4th Equation of [62] on page 346. This shows a may not be an integer in spite of the definition of chi-square distribution.

<u>5.1.2.1. Optimization of θ_k.</u> For maximizing the lower bound in Equation 5.9, the following optimization problem should be solved:

$$\begin{aligned} \max_{\theta_i} \quad & \sum_{i=0}^{\infty} a_i log(\theta_i) && (5.14) \\ s.t. \quad & \sum_{i=0}^{\infty} \theta_i = 1, \\ & 0 < \theta_i < 1 \end{aligned}$$

where

$$a_i = exp\Big(-\frac{M\gamma}{2}\Big)\frac{1}{i!}\Big(\frac{M\gamma}{2}\Big)^i. \quad (5.15)$$

For solving this problem, the transformation $e^{-\delta_i} = \theta_i$ is employed. The resulting new problem becomes

$$\begin{aligned}\min \quad & \sum_{i=0}^{\infty} a_i\delta_i \\ s.t. \quad & \sum_{i=0}^{\infty} e^{-\delta_i} = 1 \\ & \delta_i > 0.\end{aligned} \tag{5.16}$$

Then, for optimization by Lagrange multipliers

$$\begin{aligned}\frac{\partial L}{\partial \theta_k} &= \frac{\partial L}{\partial \delta_k}\frac{\partial \delta_k}{\partial \theta_k} \\ &= \frac{\partial}{\partial \delta_k}\left(\sum_{i=0}^{\infty} a_i\delta_i + \lambda\Big(\sum_{i=0}^{\infty} e^{-\delta_i} - 1\Big)\right)\frac{\partial}{\partial \theta_k}(-log(\theta_k)) \\ &= (a_k - \lambda e^{-\delta_k})(-e^{\delta_k}) = 0 \\ e^{-\delta_k} &= \theta_k = \frac{a_k}{\lambda}.\end{aligned} \tag{5.17}$$

is derived as the optimum θ_k parameter which depends on λ. By putting this result into the constraint of Equation 5.16,

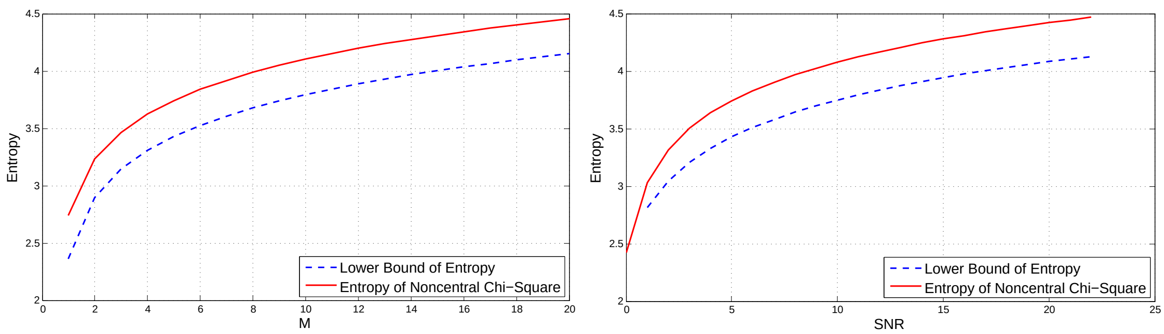
$$\begin{aligned}\lambda &= \sum_{i=0}^{\infty} a_i \\ &= exp\Big(-\frac{M\gamma}{2}\Big)\sum_{i=0}^{\infty}\frac{1}{i!}\Big(\frac{M\gamma}{2}\Big)^i \\ &= exp\Big(-\frac{M\gamma}{2}\Big)exp\Big(\frac{M\gamma}{2}\Big) \\ &= 1.\end{aligned} \tag{5.18}$$

is found. Then, the optimum value of θ_k is

$$\theta_k = \frac{a_k}{\lambda} = a_k = exp\Big(-\frac{M\gamma}{2}\Big)\frac{1}{k!}\Big(\frac{M\gamma}{2}\Big)^k. \tag{5.19}$$

5.1.3. Plots of Lower Bounds

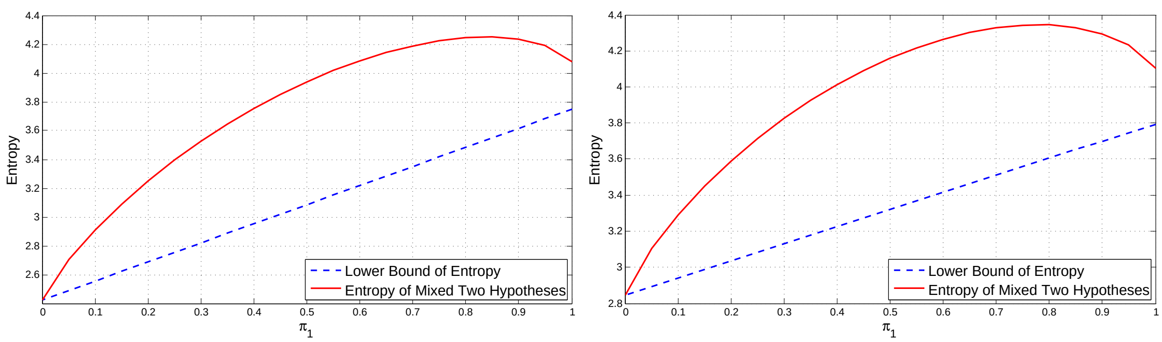
In Figure 5.1, one can examine the difference between the derived lower bound and exact entropy values. Entropy values are approximated by taking the integrals as summation of the areas of rectangulars which have small intervals. This figure shows that the lower bound of noncentral chi-square has approximately the same slope on the true entropy after certain parameters.



(a) Entropy vs. M (b) Entropy vs. γ (SNR)

Figure 5.1. Lower bound of entropy of concentral chi-square distribution.

In Figure 5.2, entropy bounds of two hypotheses together with respect to *a priori* probabilities are shown. One can observe that this bound is too loose in these figures, especially around $\pi_1 = 0.6 - 0.9$. Furthermore, this approach may not be feasible to define number of bits of quantized observations.



(a) $M = 5$, $SNR = 10$ (b) $M = 10$, $SNR = 5$

Figure 5.2. Lower bound of entropy two hypotheses of chi and noncentral chi square distributions.

5.2. Chernoff Bound Approach

A performance measure on detection problems by help of Chernoff bound is derived in [46] as

$$\zeta(s) = \min_{0 \leq s \leq 1} \ln \int f_1(x)^s f_0(x)^{1-s} dx \tag{5.20}$$

which is,

$$P_e = \frac{1}{2}P_F + \frac{1}{2}P_M \leq \frac{1}{2}e^{\zeta(s)}. \tag{5.21}$$

The performance measure in Equation 5.20 depends on the minimization of upper bound on the sum of P_F and P_M. We will firstly derive error bound for quantized distributions and then observe the bound for different parameters.

All observed data are assumed to be independent and identically distributed for each sensing node. For the first derivation, we will assume that there is no fusion by central processor. Namely, the distribution which fusion center encounters is joint distribution of each quantized observation.

$$f_{1,C}(x_1, x_2, \dots) = \prod_{i=1}^{N} f_1(x_i) \quad , \quad f_{0,C}(x_1, x_2, \dots) = \prod_{i=1}^{N} f_0(x_i) \tag{5.22}$$

As a result of the distributions in Equation 5.22, the performance measure becomes,

$$\begin{aligned} \zeta(s) &= \ln \int f_{1,C}(x_1, x_2, \dots)^s f_{0,C}(x_1, x_2, \dots)^{1-s} dx_1 dx_2 \dots dx_N \\ &= \sum_{i=1}^{N} \ln \int f_1(x_i)^s f_0(x_i)^{1-s} dx_i \\ &= \sum_{i=1}^{N} \ln \sum_{j=1}^{m} p_1(q_j)^s p_0(q_j)^{1-s} \end{aligned} \tag{5.23}$$

and then error bound becomes

$$P_{e,C} \leq \frac{1}{2}\left[\sum_{j=1}^{m} p_1(q_j)^s p_0(q_j)^{1-s}\right]^N \tag{5.24}$$

which depends on error measure of a single sensing node. Furthermore, any minimization of $\zeta(s)$ will affect performance, directly.

In the case on data fusion, we assume that the fusion center performs equal gain combining to quantized observations. Performance measure can be derived by Equation 2.42 as,

$$\begin{aligned}
\zeta(s) &= \ln \int \left[\sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m} \dots p_1(q_a)p_1(q_b)p_1(q_c)\dots v(q_a+q_b+q_c+\dots-x)\right]^s \\
&\quad \cdot \left[\sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m} \dots p_0(q_a)p_0(q_b)p_0(q_c)\dots v(q_a+q_b+q_c+\dots-x)\right]^{1-s} dx \\
&= \ln \left[\sum_{a=1}^{m}\sum_{b=1}^{m}\sum_{c=1}^{m} \dots p_1(q_a)^s p_0(q_a)^{1-s} p_1(q_b)^s p_0(q_b)^{1-s} p_1(q_c)^s p_0(q_c)^{1-s} \dots\right] \\
&= N \ln \sum_{j=1}^{m} p_1(q_j)^s p_0(q_j)^{1-s}
\end{aligned} \tag{5.25}$$

which again leads the error bound to

$$P_{e,C} \leq \frac{1}{2}\left[\sum_{j=1}^{m} p_1(q_j)^s p_0(q_j)^{1-s}\right]^N. \tag{5.26}$$

These two derivations show that the Chernoff bound results give the same performance on fusion (summing the test data) and no fusion cases.

$$\min_{0\leq s\leq 1} \sum_{j=1}^{m} p_1(q_j)^s p_0(q_j)^{1-s} \tag{5.27}$$

Another approach to tighten the error bound is to minimize the bound with s parameter. This bound is strictly convex and has only one minimum value due to s. By solving Equation 5.27 with Karush-Kuhn-Tucker (KKT) conditions [59], when relative entropy between two discrete distributions equals zero, s can take the values of 0 or 1. This result shows that s takes values 0 or 1 only when the two distributions are identically distributed. Otherwise, minimization can be done by solving the derivative of bound which is,

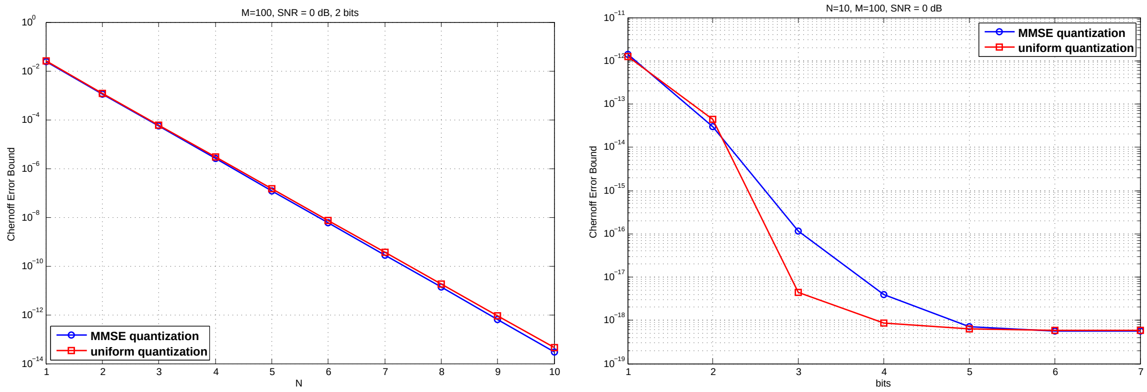
$$\sum_{j=1}^{m} p_0(q_j) ln \left(\frac{p_1(q_j)}{p_0(q_j)} \right) \left(\frac{p_1(q_j)}{p_0(q_j)} \right)^{s} = 0. \tag{5.28}$$

This equation can be solved by numerical methods such as bisection method with low operating time, because there is only one unknown variable and it is restricted between 0 and 1.

5.2.1. Performance of Chernoff Bound

Figure 5.3(a) shows that the Chernoff bound decreases exponentially with increasing of number of users. This fast decrease is a result of Equation 5.26 where N is exponent. Figure 5.3(b) presents the decrease on Chernoff bound due to the increase in the number of bits. Different quantization methods from the Section 4.4 obviously differ in this figure. Another important result is that the bound do not decrease anymore after a certain number of bits. Since most of the quantization levels of uniform quantization at 3-4 bits are between means of two hypotheses, it performs better than $MMSE$ 3-4 bits quantization.

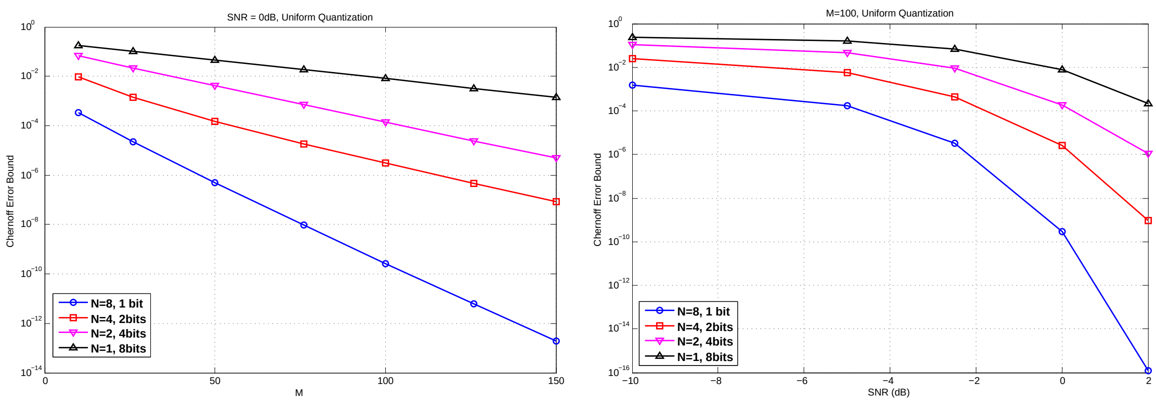
Figure 5.4 shows a reporting channel system which allows 8 bit transmission of sensing data in total for each sensing instance. The motivation for these results is to point out the performance when there is a contraint on the reporting channel resources. In both figures, the curves with more users and fewer quantized bits outperform. The



(a) Effect of number of users (N) (b) Effect of quantized bits ($\log_2(m)$)

Figure 5.3. Effect of number of users and bits on Chernoff error bound.

reason for this may be exponential behavior of number of users on Chernoff bound. Another important aspect is the SNR wall concept [16]. In Figure 5.4(b), curves converge to a certain level of error bound when SNR decreases. So, signal energy may not be detectible for lower SNR values below a certain level.



(a) Effect of number of samples (M) (b) Effect of SNR (dB)

Figure 5.4. Effect of SNR and number of samples on Chernoff Error Bound for reporting chennel which has 8 bit rate for each sensing.

6. CONCLUSION

ED is widely studied spectrum sensing system to detect signal energy due to its simplicity and requirement to none of signal characteristics. This thesis focuses on two main problems on distributed detection of ED. The first problem is threshold finding in the nodes due to achieve constant false alarm probability for binary decisions. The second problem is to derive optimum quantization of node observations due to minimum distortion and to give a expression on number of bits which should nodes quantize their observations to.

For adaptive threshold finding, two algorithms are proposed. The first algorithm depends on minimizing the distance between instantenous P_F and desired P_F by Netwon's method. For analytical implementation, this algorithm finds the exact threshold without execution of inverse of $P_F(\lambda)$ function, if perfect knowledge of noise variance is known. Another step on this algorithm, the estimation of P_F by counting rule. Even though this step adds requirement to long iteration due to achieve good estimation, it is robust to small changes in the noise variance. The second algorithm depends on generating a discrete random variable and comparing it to previous decision in every iteration. By this algorithm, system adjusts its threshold to any changes in the noise variance and works totally independent from $P_F(\lambda)$. Furthermore, this algorithm is executable to any hypothesis testing system without imposing any assumptions about the signal. Also, by updating the threshold addition, variance of the converged threshold can be decreased significantly. As future work on this algorithm, an additional threshold updating rule to increase the addition may be derived for the channels with varying noise variance. With a good design of decrease and increase on threshold addition, algorithm may have less variance when it converges and may react agilely to sudden changes of noise variance.

The second focus of this thesis is on the quantization of observations. Even though optimum quantization is generally studied as maximization of deflection coefficient at distributed detection problems, we study minimum distortion quantization of two hypothesis together and observe its performance. Mathematical derivation for both Gaussian and chi-square hypotheses for Lloyd-Max algorithm is handled. As simulation results point out, the minimum distortion quantization has good performance at low SNR case. Since quantized observations are generally used at low SNR case, this result shows minimum distortion quantization may be reasonable to implement.

Number of quantized bits of observations is another problem which is in the scope of this thesis. As this problem is already studied as rate-distortion theory, a lower bound on rate-distortion function is derived for the signal which consists of two hypotheses. Even though this bound is observed to be loose, the results suggest that when the *a priori* probabilities are around 0.5, entropy of the observation increases. Another approach on number of bits is handled by Chernoff bound. As a result of figure of this bound, number of users has more importance than number of bits for limited bit-rate reporting channel. As future work on Chernoff bound, this bound may be used to find optimum quantization thresholds. This optimum quantization may result a better perfromance than minimum distortion, because it depends on minimization of P_F and P_M.

APPENDIX A: GAMMA FUNCTIONS

The gamma function is defined as [54]

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

which is equal to

$$\Gamma(a) = (a-1)! \tag{A.2}$$

when a has integer value.

Regularized upper incomplete gamma function can be defined as [54]

$$\gamma_u(a,b) = 1 - \gamma_l(a,b) = \frac{1}{\Gamma(a)}\Gamma_u(a,b) = \frac{1}{\Gamma(a)}\int_b^\infty t^{a-1} e^{-t} dt. \tag{A.3}$$

In this equation, $\gamma_l(a,b)$ is the regularized lower incomplete function and $\Gamma_u(a,b)$ is upper incomplete function which is reqularized by dividing to $\Gamma(a)$. After regularizing (dividing to gamma function), the output of both upper and lower regularized gamma functions can take values in [0 1].

A finite sum expansion of the regularized incomplete gamma function for integer a parameter is

$$\begin{aligned}
\gamma_u(a,b) &= \frac{1}{(a-1)!}\int_b^{\infty} t^{a-1}e^{-t}dt \\
&= \frac{1}{(a-1)!}\left[-t^{a-1}e^{-t} + (a-1)\int t^{a-2}e^{-t}dt\right]_b^{\infty} \\
&= \frac{1}{(a-1)!}\left[-t^{a-1}e^{-t} - (a-1)t^{a-2}e^{-t} + (a-1)(a-2)\int t^{a-3}e^{-t}dt\right]_b^{\infty} \\
&\vdots \\
&= \frac{1}{(a-1)!}\left[-\sum_{i=0}^{a-1}\frac{t^i e^{-t}(a-1)!}{i!}\right]_b^{\infty} \\
&= e^{-b}\sum_{i=1}^{a-1}\frac{b^{-i}}{i!} \qquad \text{(A.4)}
\end{aligned}$$

which is calculated by taking integration by parts several times. This result can also be calculated from 8.352.4th Equation of [62] on page 900. Since this is not a infinite sum, the output of this function can be acquired without approximation.

REFERENCES

1. McHenry, M., P. Tenhula, D. McCloskey, D. Roberson and C. Hood, "Chicago Spectrum Occupancy Measurements & Analysis and A Long-term Studies Proposal", *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, p. 1, ACM, 2006.

2. Islam, M., C. Koh, S. Oh, X. Qing, Y. Lai, C. Wang, Y. Liang, B. Toh, F. Chin, G. Tan *et al.*, "Spectrum Survey in Singapore: Occupancy Measurements and Analyses", *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on*, pp. 1–7, IEEE, 2008.

3. López-Benítez, M., A. Umbert and F. Casadevall, "Evaluation of Spectrum Occupancy in Spain for Cognitive Radio Applications", *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, pp. 1–5, IEEE, 2009.

4. Mitola III, J. and G. Maguire Jr, "Cognitive Radio: Making Software Radios More Personal", *Personal Communications, IEEE*, Vol. 6, No. 4, pp. 13–18, 1999.

5. Haykin, S., "Cognitive radio: Brain-empowered Wireless Communications", *IEEE journal on selected areas in communications*, Vol. 23, No. 2, pp. 201–220, 2005.

6. Akyildiz, I., W. Lee, M. Vuran and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey", *Computer Networks*, Vol. 50, No. 13, pp. 2127–2159, 2006.

7. *Federal Communications Commission (FCC)*, 2013, `http://www.fcc.gov`, accessed at January 2013.

8. Yucek, T. and H. Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications", *Communications Surveys & Tutorials, IEEE*, Vol. 11, No. 1, pp. 116–130, 2009.

9. Tang, H., “Some Physical Layer Issues of Wide-band Cognitive Radio Systems”, *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pp. 151–159, IEEE, 2005.

10. Proakis, J., *Digital Communications*, New York, McGraw Hill, 2001.

11. Maeda, K., A. Benjebbour, T. Asai, T. Furuno and T. Ohya, “Recognition Among OFDM-based Systems Utilizing Cyclostationarity-inducing Transmission”, *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pp. 516–523, IEEE, 2007.

12. Yucek, T. and H. Arslan, “Spectrum Characterization for Opportunistic Cognitive Radio Systems”, *Military Communications Conference, 2006. MILCOM 2006. IEEE*, pp. 1–6, IEEE, 2006.

13. Urkowitz, H., “Energy Detection of Unknown Deterministic Signals”, *Proceedings of the IEEE*, Vol. 55, No. 4, pp. 523–531, 1967.

14. Ciftci, S. and M. Torlak, “A Comparison of Energy Detectability Models for Spectrum Sensing”, *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pp. 1–5, IEEE, 2008.

15. Hoven, N., R. Tandra and A. Sahai, “Some Fundamental Limits on Cognitive Radio”, *Proceedings of the Forty-second Allerton Conference on Communication, Control, and Computing held in Monticello, IL, USA*, pp. 131–136, 2005.

16. Tandra, R. and A. Sahai, “SNR Walls for Signal Detection”, *Selected Topics in Signal Processing, IEEE Journal of*, Vol. 2, No. 1, pp. 4–17, 2008.

17. Digham, F., M. Alouini and M. Simon, “On the Energy Detection of Unknown Signals over Fading Channels”, *Communications, 2003. ICC’03. IEEE International Conference on*, Vol. 5, pp. 3575–3579, IEEE, 2003.

18. Digham, F., M. Alouini and M. Simon, "On the Energy Detection of Unknown Signals over Fading Channels", *Communications, IEEE Transactions on*, Vol. 55, No. 1, pp. 21–24, 2007.

19. Ghasemi, A. and E. Sousa, "Opportunistic Spectrum Access in Fading Channels through Collaborative Sensing", *Journal of Communications*, Vol. 2, No. 2, pp. 71–82, 2007.

20. Atapattu, S., C. Tellambura and H. Jiang, "Performance of An Energy Detector over Channels with Both Multipath Fading and Shadowing", *Wireless Communications, IEEE Transactions on*, Vol. 9, No. 12, pp. 3662–3670, 2010.

21. Atapattu, S., C. Tellambura and H. Jiang, "Representation of Composite Fading and Shadowing Distributions by Using Mixtures of Gamma Distributions", *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, pp. 1–5, IEEE, 2010.

22. Kim, S., J. Lee, H. Wang and D. Hong, "Sensing Performance of Energy Detector with Correlated Multiple Antennas", *Signal Processing Letters, IEEE*, Vol. 16, No. 8, pp. 671–674, 2009.

23. Akyildiz, I., W. Lee and K. Chowdhury, "CRAHNs: Cognitive Radio Ad hoc Networks", *Ad Hoc Networks*, Vol. 7, No. 5, pp. 810–836, 2009.

24. Mishra, S., A. Sahai and R. Brodersen, "Cooperative Sensing Among Cognitive Radios", *Communications, 2006. ICC'06. IEEE International Conference on*, Vol. 4, pp. 1658–1663, IEEE, 2006.

25. Aysal, T., S. Kandeepan and R. Piesiewicz, "Cooperative Spectrum Sensing over Imperfect Channels", *GLOBECOM Workshops, 2008 IEEE*, pp. 1–5, IEEE, 2008.

26. Altay, C., H. Yilmaz and T. Tugcu, "Cooperative Sensing Analysis Under Imperfect Reporting Channel", *Computers and Communications (ISCC), 2012 IEEE*

Symposium on, pp. 000770–000775, IEEE, 2012.

27. Zhu, P., J. Li and X. Wang, “Cooperative Spectrum Sensing with Feedback”, *Wireless Communications, Networking and Mobile Computing, 2009. WiCom’09. 5th International Conference on*, pp. 1–4, IEEE, 2009.

28. Varshney, P., *Distributed Detection Theory and Data Fusion.*, Tech. rep., DTIC Document, 1995.

29. Pados, D., K. Halford, D. Kazakos and P. Papantoni-Kazakos, “Distributed Binary Hypothesis Testing with Feedback”, *Systems, Man and Cybernetics, IEEE Transactions on*, Vol. 25, No. 1, pp. 21–42, 1995.

30. Chen, R., J. Park and K. Bian, “Robust Distributed Spectrum Sensing in Cognitive Radio Networks”, *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp. 1876–1884, IEEE, 2008.

31. Ma, J., G. Zhao and Y. Li, “Soft Combination and Detection for Cooperative Spectrum Sensing in Cognitive Radio Networks”, *Wireless Communications, IEEE Transactions on*, Vol. 7, No. 11, pp. 4502–4507, 2008.

32. Quan, Z., S. Cui and A. Sayed, “An Optimal Strategy for Cooperative Spectrum Sensing in Cognitive Radio Networks”, *Global Telecommunications Conference, 2007. GLOBECOM’07. IEEE*, pp. 2947–2951, IEEE, 2007.

33. Peh, E. and Y. Liang, “Optimization for Cooperative Sensing in Cognitive Radio Networks”, *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, pp. 27–32, IEEE, 2007.

34. Zhang, W., R. Mallik and K. Letaief, “Optimization of Cooperative Spectrum Sensing with Energy Detection in Cognitive Radio Networks”, *Wireless Communications, IEEE Transactions on*, Vol. 8, No. 12, pp. 5761–5766, 2009.

35. Quan, Z., S. Cui, A. Sayed and H. Poor, "Optimal Multiband Joint Detection for Spectrum Sensing in Cognitive Radio Networks", *Signal Processing, IEEE Transactions on*, Vol. 57, No. 3, pp. 1128–1140, 2009.

36. Fan, R. and H. Jiang, "Optimal Multi-channel Cooperative Sensing in Cognitive Radio Networks", *Wireless Communications, IEEE Transactions on*, Vol. 9, No. 3, pp. 1128–1138, 2010.

37. Ye, Z., G. Memik and J. Grosspietsch, "Energy Detection Using Estimated Noise Variance for Spectrum Sensing in Cognitive Radio Networks", *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*, pp. 711–716, IEEE, 2008.

38. Mariani, A., A. Giorgetti and M. Chiani, "SNR Wall for Energy Detection with Noise Power Estimation", *Communications (ICC), 2011 IEEE International Conference on*, pp. 1–6, IEEE, 2011.

39. Mansouri, N. and M. Fathi, "Simple Counting Rule for Optimal Data Fusion", *Control Applications, 2003. CCA 2003. Proceedings of 2003 IEEE Conference on*, Vol. 2, pp. 1186–1191, IEEE, 2003.

40. Pados, D., P. Papantoni-Kazakos, D. Kazakos and A. Koyiantis, "On-line Threshold Learning for Neyman-Pearson Distributed Detection", *Systems, Man and Cybernetics, IEEE Transactions on*, Vol. 24, No. 10, pp. 1519–1531, 1994.

41. Gorcin, A., K. Qaraqe, H. Celebi and H. Arslan, "An Adaptive Threshold Method for Spectrum Sensing in Multi-channel Cognitive Radio Networks", *Telecommunications (ICT), 2010 IEEE 17th International Conference on*, pp. 425–429, IEEE, 2010.

42. Lee, J., J. Kim, H. Oh and S. Hwang, "Energy Detector Using Adaptive-fixed Thresholds in Cognitive Radio Systems", *Communications, 2008. APCC 2008. 14th Asia-Pacific Conference on*, pp. 1–4, IEEE, 2008.

43. Kaligineedi, P. and V. Bhargava, "Sensor Allocation and Quantization Schemes for Multi-band Cognitive Radio Cooperative Sensing System", *Wireless Communications, IEEE Transactions on*, Vol. 10, No. 1, pp. 284–293, 2011.

44. Yilmaz, H., T. Tugcu and F. Alagoz, "Uniform Quantizer for Cooperative Sensing in Cognitive Radio Networks", *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, pp. 548–553, IEEE, 2010.

45. Kassam, S., "Optimum Quantization for Signal Detection", *Communications, IEEE Transactions on*, Vol. 25, No. 5, pp. 479–484, 1977.

46. Van Trees, H., "Detection, Estimation, and Modulation: Part I", New York, Wiley, 1968.

47. Picinbono, B. and P. Duvaut, "Optimum Quantization for Detection", *Communications, IEEE Transactions on*, Vol. 36, No. 11, pp. 1254–1258, 1988.

48. Chen, L., J. Wang and S. Li, "Cooperative Spectrum Sensing with Multi-bits Local Sensing Decisions in Cognitive Radio Context", *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*, pp. 570–575, IEEE, 2008.

49. Lee, C. and J. Chao, "Optimum Local Decision Space Partitioning For Distributed Detection", *Aerospace and Electronic Systems, IEEE Transactions on*, Vol. 25, No. 4, pp. 536–544, 1989.

50. Cover, T. and J. Thomas, *Elements of Information Theory*, Wiley-Interscience, Hoboken, New-Jersey, 2006.

51. Kazakos, D., "New Error Bounds and Optimum Quantization for Multisensor Distributed Signal Detection", *Communications, IEEE Transactions on*, Vol. 40, No. 7, pp. 1144–1151, 1992.

52. Shannon, C., "Communication in the Presence of Noise", *Proceedings of the IRE*,

Vol. 37, No. 1, pp. 10–21, 1949.

53. Slepian, D., “Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty-I”, *Bell System Technical Journal*, Vol. 40, pp. 43–64, 1961.

54. Abramowitz, M. and I. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Washington D.C., 1965.

55. Nuttall, A., “Some Integrals Involving the Q_m Function”, *Information Theory, IEEE Transactions on*, Vol. 21, No. 1, pp. 95–96, 1975.

56. Grinstead, C. and J. Snell, *Introduction to Probability*, Amer Mathematical Society, 1997.

57. Chaudhari, S. and V. Koivunen, “Effect of Quantization and Channel Errors on Collaborative Spectrum Sensing”, *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*, pp. 528–533, IEEE, 2009.

58. Arthur Woodward, J. and C. Palmer, “On the Exact Convolution of Discrete Random Variables”, *Applied mathematics and computation*, Vol. 83, No. 1, pp. 69–77, 1997.

59. Boyd, S. and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

60. Max, J., “Quantizing for Minimum Distortion”, *Information Theory, IRE Transactions on*, Vol. 6, No. 1, pp. 7–12, 1960.

61. Shanmugan, K. and A. Breipohl, *Random Signals: Detection, Estimation, and Data Analysis*, Wiley & Sons Inc, 1988.

62. Gradshteyn, I.S.and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 2007.