國 立 交 通 大 學

電子工程學系 電子研究所碩士班

碩 士 論 文

以統計分佈距離為基礎之感知無線電合作式頻譜偵測技術

Distribution Distance based Collaborative Spectrum Detection in Cognitive Radio Networks

研 究 生：陳信宏

指導教授：簡鳳村 博士

中 華 民 國 九 十 八 年 八 月

以統計分佈距離為基礎之感知無線電合作式頻譜偵測技術

Distribution Distance based Collaborative Spectrum Detection in Cognitive Radio Networks

研 究 生：陳信宏　　Student: Shin-Horng Chen
指導教授：簡鳳村 博士　　Advisor: Dr. Feng-Tsun Chien

國 立 交 通 大 學

電子工程學系 電子研究所碩士班

碩士論文

A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
in
Electronics Engineering
August 2009
Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 八 年 八 月

以統計距離分布為基礎之感知無線電合作式頻譜偵測技術

研究生: 陳信宏　　　　　　　　　　指導教授:簡鳳村博士

國立交通大學電子工程學系　　電子研究所碩士班

摘要

在此論文中,我們討論了在感知無線電的環境中,使用了距離量測 (Distance measure)在機率分佈下之頻譜感測。使用距離量測當作效能之度量的原因是基於我們並不能輕易的從 log likelihood ratio 中得到近似解。所以在每一個次要使用者 (Secondary users)中,我們採用距離量測作為度量。

這篇論文主要分為兩部分。在第一部分中,我們考慮集中式偵測的方式:每一個次要使用者傳送沒有量化過的訊號到共同接收器 (Fusion center)來偵測是否有頻譜洞 (Spectrum hole)存在。在此,我們使用兩種距離量測方法,J-divergence 及 L2 distance 來設計決策方法(Decision rule)。實際上,我們嘗試使用最佳功率分配(Optimal power allocation)及最佳線性組合(Optimal linear combination)兩種方法使偵測頻譜洞之偵測機率最大同時維持主要使用者收到的干擾在一定的程度之內。從模擬結果可以看出,使用距離量測下所得到的偵測機率的確比相同功率分布(Equal power allocation)及相同加權組合(Equal weighting combination)來的好。

在第二部分中,我們考慮了使用設限(Censoring)之非集中式偵測。設限代表了次要使用者只傳有資訊的資料到共同接收器,否則便不傳任何資料。因為高斯

混合模型(Gaussian mixture model)的關係，依然很難最大化偵測機率。所以我們一樣使用距離量測的方式當作效能之度量。從模擬結果中我們可以看到，在設限方法下的偵測機率，的確比原本沒有任何限制的方法還好。

Distribution Distance based Collaborative Spectrum Detection in Cognitive Radio Networks

Student: Shin-Horng Chen　　　　　Advisor: Dr. Feng-Tsun Chien

Department of Electronics Engineering
Institute of Electronics
National Chiao Tung University

Abstract

In this thesis, we discuss the problem of collaborative spectrum sensing in cognitive radio networks from the perspective of distance measures between probability distributions. The rationale behind using the distance measures as the performance metric lies on the difficulty of having a closed-form expression for the log likelihood ratio. We adopt the distance measure as the metric to design the decision criterion in each of the secondary users in the cooperative environment.

The thesis is mainly consisted of two parts. In the first part, we consider the case of centralized detection in which every secondary user sends un-quantized signal to the fusion center for the ultimate detection of the spectrum hole. We use two distance measures, J-divergence and L2 distance, to design the local decision rule. In particular, we attempt to devise a power allocation scheme among secondary users, as well as a combination scheme to gather received signals in the fusion center, for maximizing the probability of detection of a spectrum hole while keeping the interference observed by the primary user within a predetermined level. The analytical and simulated results show that we can improve the detection probability by optimizing the distance measures as compared to the equal power allocation and equal weight combination.

In the second part, we consider the case of decentralized detection with censoring. The censoring means the secondary users only transmit informative observations to the fusion center or keep silent. In this case, it's also hard to maximize detection probability because of the underlying Gaussian mixture model (GMM). We again use the distance measures as the performance metric. Simulation results show that the detection probability of the censoring scheme is better than that of the non-censoring approach.

誌謝

本篇論文的完成，誠摯地感謝我的指導老師 簡鳳村 博士，從進入交通大學電子所開始，多虧老師的指導，無論上是在課業、研究上，都給我許多的幫助。老師也時常教導我們要如何在社會上應對進退，在各種的場合下各種禮儀、發言，我想我在這方面都學到許多。在此謹向老師表達我最高的感謝。

另外要感謝的是實驗室的同學重佑，熱心的幫我解決了許多問題，讓我省去不少的麻煩。還有遠在台大的志寧，常常給我一些有用的研究上需要的論文以及跟我討論研究上的問題。還有大學的前室友們，幫我看了論文的一小部分以及給了一些建議。

還有真的很感謝通訊電子與訊號處理實驗室(commlab)，從進入開始便給我很多軟硬體上的資源，這都是我在大學部的時候難以比擬的，也讓我做研究有強大的後盾。感謝朝雄、鴻志、家揚等博班學長，他們幫助我處理了很多實驗室上的問題以及各種學校方面的雜務。以及其他實驗室的成員，每每在實驗室窩到凌晨時，還有人陪著的感覺真好，讓我知道自己不是最晚的。還有感謝世榮，在我口試預演以及各種需要幫忙的場合，我都感受到他的熱心。

最後，感謝我的家人及女朋友佩蓉，陪我度過一切的難關。家人的支持真的讓我無後顧之憂地在新竹念書、做研究，以及佩蓉在我煩躁時適時地給予鼓勵跟幫助。也感謝所有研究生生涯有幫到我忙的人。

謝謝所有幫助過我、陪我走過這一段歲月的師長、同儕與家人。謝謝！

誌於 2009.8 風城交大

信宏

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cognitive radio has been viewed as a promising technology in next generation wireless communication networks striving for a better utilization of the wireless spectrum. In many countries, most of frequency bands are assigned to different wireless services. But some frequency bands are under-utilized. In [1], it has been shown that 70% of the allocated spectrum in the U.S. are not efficiently utilized. If we allow secondary users (unlicensed users) to use the frequency band of primary users (licensed users) when the primary users are idle, the utilization of spectrum will be enhanced. In other words, the main purpose of the cognitive radio systems is to utilize the spectrum and limit the interference to primary users in a efficient and intelligent manner.

- **R**elated work:

  There are many papers in cognitive radio. The work in [2]-[5] discuss the basic concepts and limitations of the cognitive radios. In [6], the authors propose 3 local sensing methods, the matched filter detection, the energy detection, and the cyclostationary feature detection. The detection probability of matched filter detection is optimal. In local sensing, it's hard to distinguish between the noise and the weak signal because of deep fading. To improve the spectrum detection performance, cooperative spectrum sensing has been proposed in [7]-[19].

  - Centralized detection

    The centralized detection is that the secondary users transmit the observation without quantization. In [10] and [11], the authors propose a cooperation scheme among secondary users to mitigate possible deep fading or shadowing, and thereby improve the detection probability. The work in [12], discuss several optimization methods in a linear combination system, which combines signals from all the antennas in the fusion center, with orthogonal channel in centralized detection. The work discusses 3 systems, conservative system, aggressive system, and hostile system, and a detection performance measure, modified deflection coefficient(MDC). But in this paper, it only concerns the orthogonal channel and doesn't have any power constraint on secondary users.

– Distributed detection

  The distributed detection is that the secondary users transmit the observation with quantization. Because of bandwidth constraint, it is often desirable to quantize data before transmitting. Many papers in distributed detection use only 1 bit. In distributed detection, we can reduce the data rate between fusion center and secondary users. But the detection performance of distributed detection is worse than centralized detection. The work in [13], discuss the semidefinite programming in distributed system in linear combination. The authors in [19] and [18] use distance measures to solve the power allocation problem. The work in [19] considers the approximated J-divergence, instead of the likelihood function, as the performance matric in distributed sensor network with multiple input multiple output(MIMO) channels, for the reason that the log likelihood ratio(LLR) does not have a closed-form expression and thus an explicit decision rule does not exist. In [18], the authors propose using approximated J-divergence to approximate the J-divergence of gaussian mixture model. To have better detection probability, the authors adopt the element distance measure, instead of the approximated J-divergence, as the performance matric in [19]. The work in [14] discusses the trade-off between the sensing time and throughput. If sensing time is shorter, the transmitting time is longer and the throughput is higher. The authors propose a multi-slot spectrum sensing scheme to maximize the detection probability of local sensing. The multi-slot spectrum sensing approach divides the sensing time into M mini-slots and uses those slots to maximize the detection probability. This work also discusses the centralized detection and distributed detection to maximize the detection probability in cooperative spectrum sensing.

– Distance measure

In [17], the L2 distance approach is applied in the problem of speech recognition. In addition, the work [20] and [21] discuss about fundamental properties of the distance measures. In [20], the J-divergence and B-divergence can lead to, respectively, a lower bound and an upper bound of the Bayesian error probability.

– Censoring scheme

The work [22]-[23] propose a censoring scheme in which secondary users only transmit informative observation to the fusion center. The censoring scheme can reduce the interference to primary users. In [22], it transmits LLR to fusion center. When LLR is greater than a threshold, the user will transmit this LLR to fusion center or keep silent. In this paper, it discusses that the detection performance of one threshold is equal to the performance of two thresholds. In [23], it proposes a simple censoring scheme. By censoring the observation, only the users with enough information will transmit their local bit decision (0 or 1) to the fusion center. The detection probability and false alarm probability of spectrum sensing are investigated for both perfect and imperfect reporting channel.

- **M**otivation:

In this thesis, we still focus on the spectrum sensing problems in cognitive radio systems. When primary users use frequency bands, secondary users should be able to detect the existence of primary users. When secondary users transmit data or receive data, they can't cause intolerable interference on the primary users, if accurate detection fails.

In much of the previous work, they only constraint the interference to primary users when transmitting. However, when secondary users send signals to the fusion center for spectrum sensing, they should also have power constraint. In this thesis, we focus on the power constraint when secondary users are in the sensing phase. Here we use target signal to interference plus noise ratio(SINR) as the performance constraint. The SINR of the primary users should be greater than the target SINR. We are interested in finding the weighting factor that maximizes the probability of detection while restricting the interference on the primary user within a predetermined level. In summary, our objectives in this research are,

  – Maximize the detection probability.
  – Satisfy the SINR constraint of primary users.

- **C**ontributions:
  – We use a distance measure of the probability distribution as the metric, instead of the likelihood measure, to find the detection rule at the fusion center.
  – We find the optimal power allocation scheme among secondary users and optimal linear combination approach in the fusion center, both to maximize the detection probability.
  – We propose a censoring scheme in secondary users when transmitting signals to the fusion center, attempting to lower interference while achieving acceptable performance.

In the first part of the thesis, we consider two schemes, namely the optimal power allocation and the optimal linear combination, in centralized cooperative spectrum sensing. Optimal power allocation scheme is that we control the power of every secondary user and optimize the detection probability. Optimal linear combination is that we combine the signal of every antenna by different weighting to maximize the detection probability. In centralized detection, it's not easy to have the closed-form expression of detection probability and false alarm probability. But we can approach the detection probability, as promised by the likelihood detection rule, by the use of distance measures. Analytical results show if the distance is larger, the detection probability is higher. In the second part of the thesis, we propose a censoring scheme in which secondary users transmit only the informative data to the fusion center. In censoring scheme, we use J-divergence and L2 distance to maximize the detection probability. Analytical results show that the censoring scheme have better detection probability than non-censoring schemes.

- **O**rganization of the Thesis:

  The thesis is organized as follows. In chapter 2, we talk about the fundamental concept of spectrum sensing in cognitive radio networks. In chapter 3, we introduce two different distance measures, the J-divergence and the L2 distance, in the optimal power allocation and optimal linear combination schemes. And the simulation results show that the performance are better than equal power allocation and equal linear combination. In chapter 4, we use the censoring method in optimal power allocation and optimal linear combination. And we show the simulation of the censoring method. Finally, chapter 5 gives the conclusion.

Chapter 2

Spectrum Sensing and System Model

Cognitive Radio is the solution of the spectrum utilization. The users who have license are primary users. In some frequency bands, the spectrum utilization is not high. It means that primary users seldom use the frequency bands. Those frequency bands are under low utilization. To enhance the spectrum utilization, primary users should share their frequency bands to other users who don't have the license. The users who want to use the frequency bands are secondary users. But one important thing is that secondary users should return the spectrum to primary users as soon as possible when primary users want to use and secondary users don't cause much interference on primary users. Therefore, spectrum sensing is the key technique of cognitive radio. Secondary users should sense whether the primary users use the spectrum or not. If they are sure that primary users don't use the spectrum now, they could use it. But there are still many problems in the spectrum sensing.

2.1 The Problems of Spectrum Sensing

The detection of the existence of primary user can be viewed as binary hypothesis. When primary users exist, it's under hypothesis H_1. When primary users don't exist and secondary users can use the spectrum, it's under hypothesis H_0.

The main problems we want to solve in spectrum sensing are:

- Maximize the detection probability P_d

- Minimize the false alarm probability P_f
- Minimize the sensing time

2.1.1 Detection Probability and False Alarm Probability

The definition of detection probability is

$$P_d = P(\hat{H} = H_1|H_1). \tag{2.1}$$

And the definition of false alarm probability is

$$P_f = P(\hat{H} = H_1|H_0). \tag{2.2}$$

$\hat{H}$ means secondary users make a decision which hypothesis is possible. P_d means the probability that the secondary users decide H_1 hypothesis and primary users is really using the spectrum. P_f means the probability the secondary users decide H_0 hypothesis and primary users is not really using the spectrum. Obviously, the system performance will be better if P_d is higher and P_f is lower. When P_d is higher, secondary users won't use the spectrum when primary users exist. When P_f is lower, it means the spectrum utilization is higher.

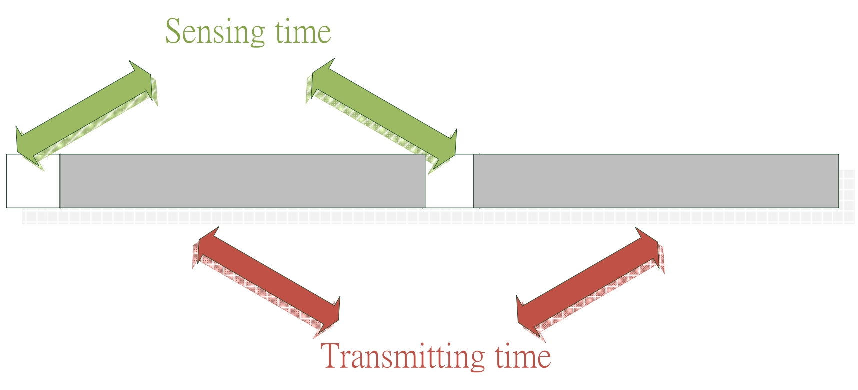


Figure 2.1: Sensing time and transmitting time

2.1.2 Sensing Time

Because we can't exactly know when primary users want to use the spectrum, we periodically sense the spectrum to solve this problem. Fig. 2.1 shows that it can be separated into the sensing time and the transmitting time. It is obvious that the sensing time should be short. In other words, if the sensing time is too long, the transmitting time will be short and the average data rate will be low. But in our system model, we assume a fixed sensing time and periodically sense the spectrum.

2.2 Maximize the Detection Probability

Why should we maximize the detection probability? The reason is that secondary users should return the spectrum when the primary users exist. The definition of detection probability is that we decide primary users is using the spectrum and primary users is really using the spectrum. Therefore, when secondary users detect the primary users is using the spectrum, secondary users will stop transmitting data. When the detection probability is low, secondary users will frequently use the spectrum to transmit data when primary users exists, like in Fig. 2.2. It may cause intolerance interference on primary users.

2.2.1 Local Sensing

Local sensing means every secondary user makes his own decision without cooperation. In [6], there are 3 methods to implement the local detection. For example, matched filter detection. The matched filter detection maximizes the signal-to-noise ratio. But the matched filter detection needs to know the prior knowledge of primary users signal at both PHI and MAC layers. It's not practical. Energy detection is another method. It collects the energy of every sensing sample. But the energy detection still has drawbacks. For example, how to decide the threshold of the decision rule. Obviously, the local sensing will depends on the power of primary users, the channel between primary users and secondary users, and so on. For example, if the signal of primary users is weak, the detection proba-

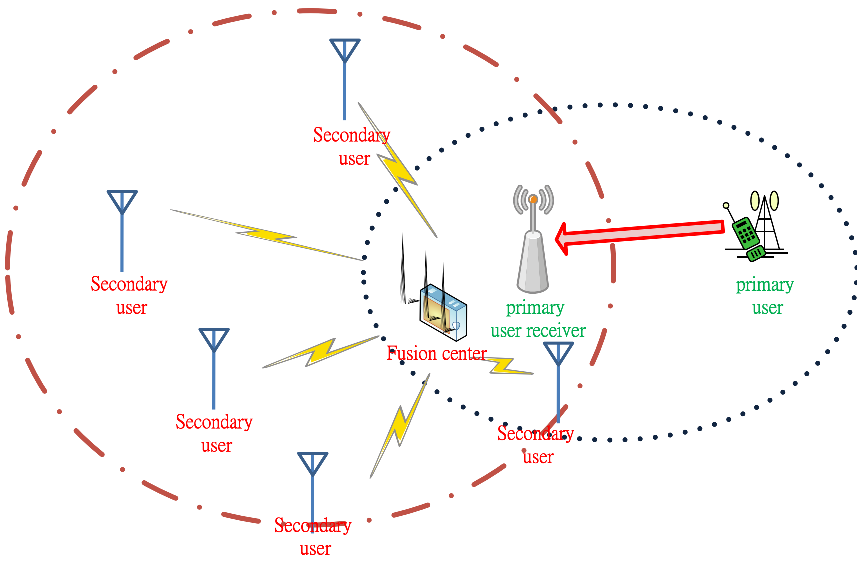


Figure 2.2: Interference of primary users

bility of secondary users will be low and the secondary users will transmit data frequently. To avoid this problem, we should consider cooperative spectrum sensing to enhance the detection probability.

2.2.2 Cooperative Sensing

Cooperative sensing means some secondary users cooperate with each other to enhance the detection probability. Many papers use cooperative scheme because of poor performance of local sensing. In the cooperative spectrum sensing, every secondary user sends the data to the fusion center. The fusion center can be viewed as a common receiver of secondary users. Because the fusion center has more data than every secondary user, it can have more accuracy detection. There are two schemes to solve the cooperative spectrum sensing, centralized detection and distributed detection. The centralized detection is that every secondary user transmits unquantized observation to the fusion center. Then the fusion center uses the observations to make the decision. The distributed detection is that every secondary user transmits quantized signal to the fusion center, like ”-1” and ”1”. ”-1” means H_0 hypothesis and ”1” means H_1 hypothesis. In [14], it uses the majority decision method. If the number of ”-1” at the fusion center is greater than the number of ”1” at the fusion center, the fusion center will decide H_0 hypothesis. Obviously, the performance of the centralized detection is better than the distributed detection. And in [10], secondary users can sense lower power of primary users than local sensing because of cooperation. We can enhance the performance by the use of the cooperative spectrum sensing. The following section will discuss our system model of the cooperative spectrum sensing.

2.3 System Model

Fig. 2.3 shows the system model. Firstly, secondary users sense data for detecting the existence of the primary user. Here we use multiple input multiple output (MIMO) channel between secondary users and the fusion center. Every user sends the data to the fusion center for the cooperative spectrum sensing and the fusion center has multiple antennas. Every secondary user has only one antenna. It is reasonable because secondary users often have the power constraint. There is only one primary user and the primary users has only one antenna. Assume every secondary user uses the same spectrum in sensing time and in transmitting time. Obviously, it will produce a problem. When secondary users transmit their sensing data to the fusion center for spectrum sensing and primary users want to access spectrum in the same time, the data of secondary users transmitting may cause intolerance interference to the primary user. This problem should be avoided. We list the problems we should solve:

- Maximize the detection probability
- Minimize the false alarm probability
- Minimize the interference when cooperative spectrum sensing
- Minimize sensing time

But in our system model, the sensing time will be a fixed time for simplicity. The detection probability and the interference of primary users should be minimized. Assume the power of the primary user is known by the fusion center. When secondary users

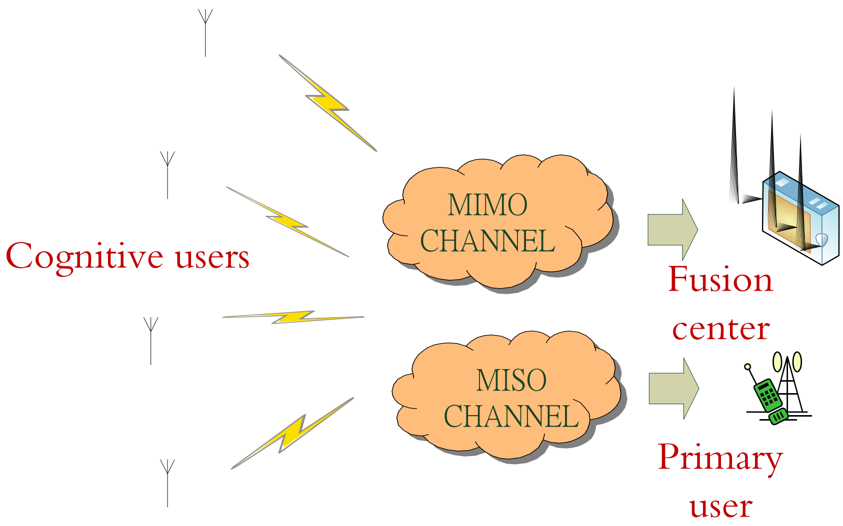


Figure 2.3: System model of MIMO channel

transmit the data to the fusion center and the primary user exists, we can set a target SINR, $SINR_t$, for the primary user and the SINR of the primary user should be greater than $SINR_t$.

The equation of SINR is

$$SINR_p = \frac{P_P}{\sum_i^N h_{p_i} P_{s_i} + \sigma_n^2}, \tag{2.3}$$

where P_{s_i} is the transmitting power of ith secondary user, h_{p_i} is the channel between secondary users and the primary user, N is the number of secondary users, and σ_n^2 is the variance of noise.

Therefore, the capacity of the primary user is

$$C_P = \frac{1}{2}\log(1 + SINR_p) \geq \frac{1}{2}\log(1 + SINR_t), \tag{2.4}$$

where C_p is the capacity of the primary user.

If the $SINR_t$ is high, it means that the primary user can't allow much interference from secondary users and the secondary users can't use much power for transmission. In other words, we should satisfy the quality of service (QoS) requirement of the primary user. Therefore, it should have power constraint on secondary users. And the final objective is to optimize the detection probability under the SINR constraint. In the following chapters, we will discuss how to satisfy the target SINR and optimize the detection probability.

Chapter 3

Centralized Detection in Optimal Power Allocation and Optimal Linear Combination

3.1 Introduction

In cooperative spectrum sensing, every secondary user transmits the observation to the fusion center. The fusion center can use those observations to have accuracy detection. If all secondary users send their data to the fusion center without quantization, this is called centralized detection. For example, consider a simple centralized detection scheme. Fig. 3.1 shows the system model. Every secondary user transmits the observation to the fusion center. From Fig. 3.1, the received signal of the fusion center is

$$y = \sum_{i=1}^{N} x_i, \tag{3.1}$$

where x_i is the signal from the ith secondary user and there are N secondary users.

Then the fusion center uses this signal, y, to make the decision. A simple decision method is that whether the value of y is greater than γ or not. If y is greater than γ, we can decide that it is under hypothesis H_1.

In this thesis, we consider two optimization schemes, optimal power allocation and optimal linear combination.

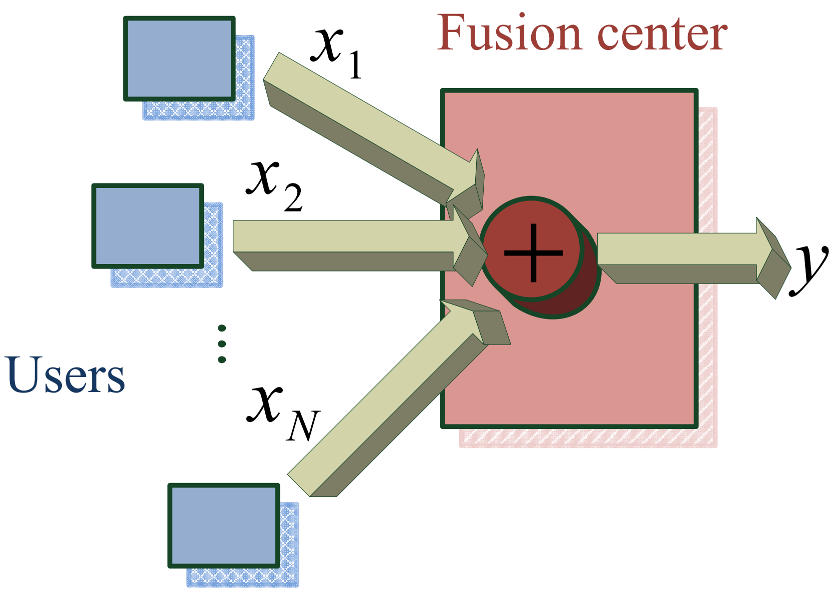


Figure 3.1: Simple centralized detection

3.1.1 Processing of Transmitting Signal

Assume that secondary users sense their signals in a sensing time, τ, and the sampling rate is δ. Therefore, the number of samples is $n = \tau\delta$. For the secondary users, if they transmit their signal directly without any signal processing, the detection probability will be low. Before transmitting, every secondary user sums the square of every sample and transmits it to the fusion center. The equation of the sensing signal of ith secondary user in kth sample is

$$y_i(k) = \begin{cases} v_i(k), & \text{under } H_0 \\ h_{pi}s(k) + v_i(k), & \text{under } H_1, \end{cases} \tag{3.2}$$

where $y_i(k)$ is the signal sensed by ith secondary user, $v_i(k)$ is the sensing noise and its distribution is $N(0, \sigma_{v_i}^2)$, $s(k)$ is the signal of the primary user, k means kth sample, and h_{pi} is the channel between ith secondary user and the primary user. Assume all sensing samples in one secondary user are independent.

The overall power of primary user in n samples is

$$E_s = \sum_{k=1}^{n} |s(k)|^2. \tag{3.3}$$

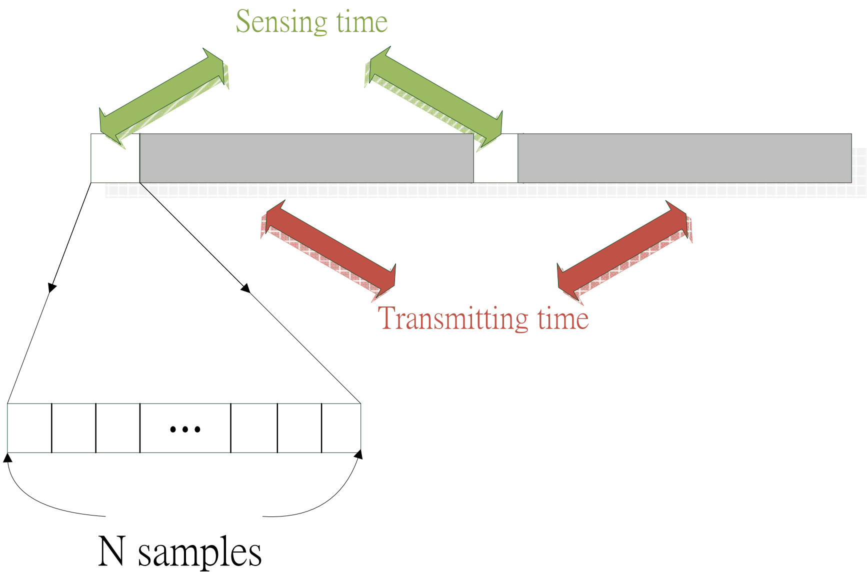


Figure 3.2: N samples in the sensing time

u_i is the transmitting signal of ith secondary user. The equation of u_i is

$$u_i = \sum_{k=1}^{n} |y_i(k)|^2. \tag{3.4}$$

The n samples are identical and independent distribution in one secondary user. The distribution of u_i can be viewed as a non-centralized Chi-square distribution with n degrees. If n is large enough, it can be asymptotical to a Gaussian distribution by the central limit theorem.

After some computation, the asymptotical distribution of u_i is:

$$E[u_i] = \begin{cases} n\sigma_{v_i}^2, & \text{under } H_0 \\ (n+\eta_i)\sigma_{v_i}^2, & \text{under } H_1 \end{cases} \tag{3.5}$$

and

$$Var[u_i] = \begin{cases} 2n\sigma_{v_i}^4, & \text{under } H_0 \\ 2(n+2\eta_i)\sigma_{v_i}^4, & \text{under } H_1, \end{cases} \tag{3.6}$$

where $\eta_i = \frac{|h_{pi}|^2 E_s}{\sigma_{v_i}^2}$. Therefore, the distribution can be represented as a Gaussian random variable.

In the simulation result, it will prove the detection probability will be higher after the signal processing.

3.2 Optimal Power Allocation

In optimal power allocation scheme, we multiply weighting factor to the transmitting signal of every secondary user. There are two reasons why we multiply weighting factor to transmitting signal.

- Maximize the detection probability
- Satisfy the power constraint

In the following sections, we will discuss the optimal power allocation.

3.2.1 System Model

In the cognitive radio environment, assume that there are N secondary users and the fusion center has M receiving antennas. To simplify the system, there is only one primary user. The primary user and every secondary user only has one antenna.

Our objective is to maximum the global detection probability, P_D, and reduce the interference to primary users when secondary users transmit data to the fusion center. In other words, when the primary users use the frequency bands, the secondary users can't produce the intolerance interference to primary users. The Fig. 3.3 shows the system model. In the centralized method, secondary users send their observation to the fusion center for the detection. But we should guarantee the QoS requirement of primary users.

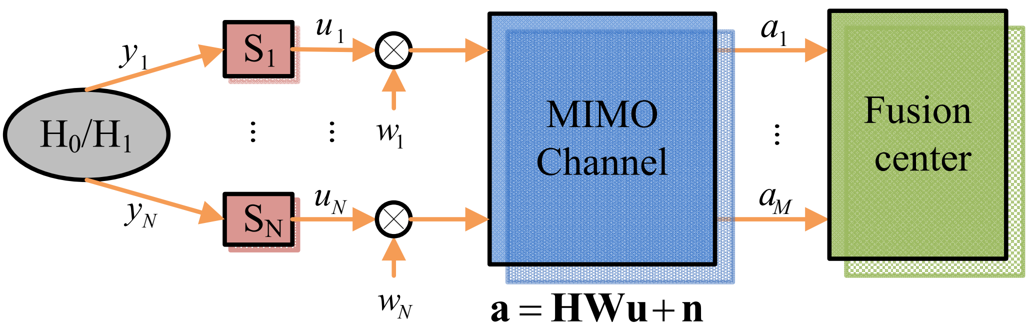


Figure 3.3: System model of optimal power allocation

In other words, we should guarantee the SINR of the primary users should be greater than one determined value, $SINR_t$. The secondary users can not transmit their signal directly because of the power constraint. They should have the weighting factors for the power control.

The received signal of the fusion center is

$$\mathbf{a} = \mathbf{HWu} + \mathbf{n}, \tag{3.7}$$

where

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}, \mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1} & h_{M2} & \dots & h_{MN} \end{bmatrix}$$

$$, \mathbf{W} = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & w_N \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}, \mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}$$

$\mathbf{u}$ is the signal vector transmitted by secondary users, $\mathbf{W} = diag([w_1, w_2, ..., w_N]^T)$ is the power control, $\mathbf{H}$ is channel matrix between the fusion center and the secondary users, $\mathbf{n} \sim N(0, \sigma_n^2 \mathbf{I}_{M\times M})$ is the noise received by the fusion center, and $\mathbf{a}$ is the signal received of the fusion center.

Assume that the signals sensed by the secondary users are independent. Therefore, we can have the distribution of $\mathbf{u}$.

$$p(\mathbf{u}|H_i) = \frac{1}{|2\pi\sigma^2_{\mathbf{u}_i}|}^{\frac{1}{2}} \exp[-\frac{1}{2}(\mathbf{u} - \mathbf{E}_i)^T {\sigma^2_{\mathbf{u}_i}}^{-1} (\mathbf{u} - \mathbf{E}_i)], \tag{3.8}$$

where

$$\sigma^2_{\mathbf{u}_1} = \begin{bmatrix} 2(n+2\eta_1)\sigma_v^4 & 0 & \dots & 0 \\ 0 & 2(n+2\eta_2)\sigma_v^4 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 2(n+2\eta_N)\sigma_v^4 \end{bmatrix}$$

$$, \sigma^2_{\mathbf{u}_0} = \begin{bmatrix} 2n\sigma_v^4 & 0 & \dots & 0 \\ 0 & 2n\sigma_v^4 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 2n\sigma_v^4 \end{bmatrix}$$

$$, \mathbf{E}_1 = \begin{bmatrix} (n+\eta_1)\sigma_v^2 \\ (n+\eta_2)\sigma_v^2 \\ \vdots \\ (n+\eta_N)\sigma_v^2 \end{bmatrix}, \mathbf{E}_0 = \begin{bmatrix} n\sigma_v^2 \\ n\sigma_v^2 \\ \vdots \\ n\sigma_v^2 \end{bmatrix}$$

From the equation (3.8), we can have the distribution of $\mathbf{a}$.

$$p(\mathbf{a}|H_i) = \frac{1}{|2\pi\mathbf{\Sigma}_i|^{\frac{1}{2}}} \exp[-\frac{1}{2}(\mathbf{a}-\mu_i)^T\mathbf{\Sigma}_i^{-1}(\mathbf{a}-\mu_i)], \tag{3.9}$$

where

$$\mathbf{\Sigma}_i = \mathbf{HW}\sigma^2_{\mathbf{u}_i}\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I} \tag{3.10}$$

and

$$\mu_i = \mathbf{HWE}_i. \tag{3.11}$$

Then we can discuss the detection method by the use of the likelihood ratio test.

3.2.2 The Performance Matric of Detection Probability

The final goal of our scheme is to maximum the detection probability. In the Neyman-Pearson detection, we should calculate likelihood ratio firstly. By the use of the likelihood ratio, we can have the detection probability.

$$L(\mathbf{a}) = \frac{p(\mathbf{a}|H_1)}{p(\mathbf{a}|H_0)} = \frac{\frac{1}{|2\pi\mathbf{\Sigma}_1|^{\frac{1}{2}}}\exp[-\frac{1}{2}(\mathbf{a}-\mu_1)^T\mathbf{\Sigma}_1^{-1}(\mathbf{a}-\mu_1)]}{\frac{1}{|2\pi\mathbf{\Sigma}_0|^{\frac{1}{2}}}\exp[-\frac{1}{2}(\mathbf{a}-\mu_0)^T\mathbf{\Sigma}_0^{-1}(\mathbf{a}-\mu_0)]} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{3.12}$$

where γ is the detection threshold.

Therefore, the log likelihood ratio (LLR) is

$$\begin{aligned} l(\mathbf{a}) &= \ln \frac{|\Sigma_0|^{\frac{1}{2}}}{|\Sigma_1|^{\frac{1}{2}}} + (-\tfrac{1}{2}(\mathbf{a}-\mu_{\mathbf{1}})^T \boldsymbol{\Sigma}_1^{-1}(\mathbf{a}-\mu_{\mathbf{1}})) \\ &\quad -(-\tfrac{1}{2}(\mathbf{a}-\mu_{\mathbf{0}})^T \boldsymbol{\Sigma}_0^{-1}(\mathbf{a}-\mu_{\mathbf{0}})) \gtrless_{H_0}^{H_1} \ln\gamma. \end{aligned}$$

From equation (3.10) and equation (3.11), the log likelihood ratio is

$$\begin{aligned} l(\mathbf{a}) &= \ln \frac{|\boldsymbol{\Sigma}_0|^{\frac{1}{2}}}{|\boldsymbol{\Sigma}_1|^{\frac{1}{2}}} + (-\tfrac{1}{2}(\mathbf{a}-\mu_1)^T \boldsymbol{\Sigma}_1^{-1}(\mathbf{a}-\mu_1)) - (-\tfrac{1}{2}(\mathbf{a}-\mu_0)^T \boldsymbol{\Sigma}_0^{-1}(\mathbf{a}-\mu_0)) \\ &= \ln \frac{|\boldsymbol{\Sigma}_0|^{\frac{1}{2}}}{|\boldsymbol{\Sigma}_1|^{\frac{1}{2}}} + \tfrac{1}{2}\mathbf{a}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{a} + \mathbf{a}^T(\boldsymbol{\Sigma}_1^{-1}\mu_1 - \boldsymbol{\Sigma}_0^{-1}\mu_0) + \tfrac{1}{2}\mu_1^T\boldsymbol{\Sigma}_0^{-1}\mu_1 - \tfrac{1}{2}\mu_0^T\boldsymbol{\Sigma}_0^{-1}\mu_0 \\ &= \ln \frac{|\mathbf{HW}\sigma_{\mathbf{u}_0}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I}|^{\frac{1}{2}}}{|\mathbf{HW}\sigma_{\mathbf{u}_1^2}\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I}|^{\frac{1}{2}}} \\ &\quad + \tfrac{1}{2}\mathbf{a}^T((\mathbf{HW}\sigma_{\mathbf{u}_0}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1} - (\mathbf{HW}\sigma_{\mathbf{u}_1}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1})\mathbf{a} \\ &\quad + \mathbf{a}^T((\mathbf{HW}\sigma_{\mathbf{u}_1}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{HWE}_1) - (\mathbf{HW}\sigma_{\mathbf{u}_0}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{HWE}_0)) \\ &\quad + \tfrac{1}{2}(\mathbf{HWE}_1)^T(\mathbf{HW}\sigma_{\mathbf{u}_1}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{HWE}_1) \\ &\quad - \tfrac{1}{2}(\mathbf{HWE}_0)^T(\mathbf{HW}\sigma_{\mathbf{u}_0}^2\mathbf{W}^T\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{HWE}_0). \end{aligned} \tag{3.13}$$

Therefore, the detection probability, P_D, is

$$P_D = P(l(\mathbf{a}) > \ln\gamma | H_1). \tag{3.14}$$

The optimization problem is

$$\begin{aligned} \max_{\mathbf{W}} \quad & P_D \\ s.t. \quad & P_{si} \leq P_c \\ & SINR_p \geq SINR_t. \end{aligned} \tag{3.15}$$

P_{si} is the power of ith secondary user and P_c is the power constraint of secondary users.

The equation of $SINR_p$ is

$$SINR_p = \frac{P_p}{\sum_i h_{p_i}^2 w_i^2 P_{si} + \sigma_{n_p}^2}, \tag{3.16}$$

where P_p is the power of the primary user, h_{p_i} is the channel between the primary user and ith secondary user, and $\sigma_{n_p}^2$ is the noise power of the primary users.

Because $l(\mathbf{a})$ is a nonlinear combination of $\mathbf{a}$, it is difficult to calculate the distribution of $l(\mathbf{a})$. Obviously, it's hard to have the closed-from expression of the detection probability. And from the equation (3.13), it's also hard to optimize the detection probability.

Therefore, it's hard to use the likelihood ratio based detection for optimization. But we can use the performance matric, such as the distance measure, instead of calculating the detection probability directly. We use this performance matric to optimize the detection probability. In the following sections, we will introduce the distance based detection by the use of J-divergence in the optimal linear combination.

J-divergence

J-divergence is a distance measure. For example, if we have two random variables, the distance measure can be viewed as the distance between this two random variables. If the distance between them is long, we can easily distinguish them and the detection probability will be high. Like in the Fig. 3.4, if distance between them is small, it's not easy to justify which hypothesis really exists. Under this condition, it's easy to make a fault decision. In the distance measure method, the objective is to maximize the distance between two distributions.

The error probability is defined as

$$P_e = P(H_0)P(l(\mathbf{a}) > \ln\gamma|H_0) + P(H_1)P(l(\mathbf{a}) < \ln\gamma|H_1) \tag{3.17}$$

, and J-divergence can provide the lower bound of P_e[20].

$$P_e > P(H_0)P(H_1)e^{(-\frac{J}{2})}, \tag{3.18}$$

where $P(H_0)$ is the probability that primary users don't use the spectrum and $P(H_1)$ is the probability that primary users use the spectrum.

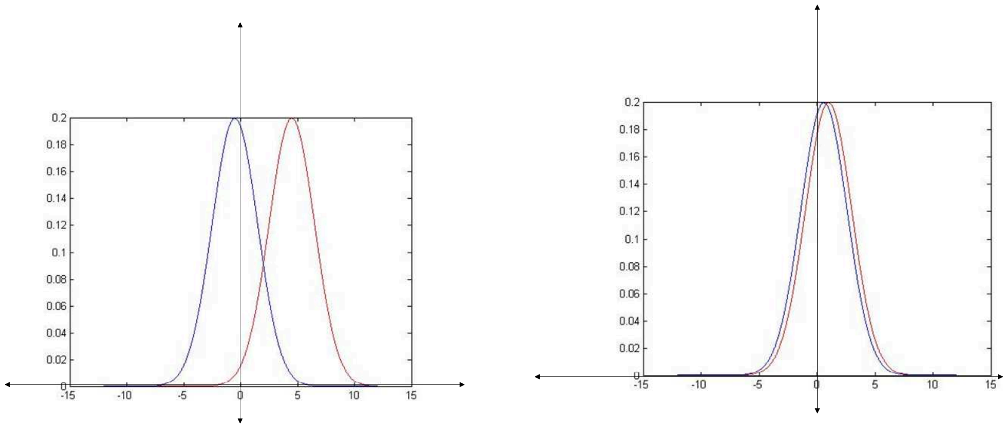
From the above equation, we can use the J-divergence for replacing calculate the error probability directly. The definition of J-divergence is

$$\begin{aligned} J &= E_1[(L-1)\ln L] \\ &= \int_{-\infty}^{+\infty}[(L-1)\ln L]p_1(\mathbf{x})d\mathbf{x} \end{aligned} \tag{3.19}$$

where $L(x) = \frac{p_0(x)}{p_1(x)}$.

In our case, $L(x) = \frac{P(\mathbf{a}|H_0)}{P(\mathbf{a}|H_1)}$. J-divergence is the symmetric form of the Kullback-Leibler (KL) distance.

$$J(F,G) = D_{KL}(F||G) + D_{KL}(G||F), \tag{3.20}$$



(a) Long distance (b) Small distance

Figure 3.4: Two different distances

where $D_{KL}(.)$ is the KL distance. The definition of the KL distance is

$$\begin{aligned} D_{KL}(F||G) &= -\sum f(x)\log g(x) + \sum f(x)\log f(x) \\ &= H(F,G) + H(F), \end{aligned} \tag{3.21}$$

where $H(F,G)$ is the cross entropy and $H(F)$ is the entropy of F. The physical mean of the KL distance is expected number of extra bits using a code based on G rather than F. If the distance is larger, it will use more bits to transmit.

By adopting J-divergence as the performance matric, the optimization problem becomes to

$$\begin{aligned} \max_{\mathbf{w}} \quad & J(p(\mathbf{a}|H_0), p(\mathbf{a}|H_1)) \\ s.t. \quad & P_{si} \leq P_c \\ & SINR_p \geq SINR_t. \end{aligned} \tag{3.22}$$

After computation, we can using the variance and mean of $\mathbf{a}$ to calculate the value of J-divergence,

$$\begin{aligned} & J(p(\mathbf{a}|H_0), p(\mathbf{a}|H_1)) = \\ & \tfrac{1}{2}\mathrm{Tr}[\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1} + (\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Sigma}_1^{-1})(\mu_1 - \mu_0)(\mu_1 - \mu_0)^T] - M. \end{aligned} \tag{3.23}$$

Therefore, we can use the optimization method to find the maximum value of J-divergence in the optimal power allocation.

Here we discuss a special condition. Assume the received SNR of the fusion center is low. Let $\mathbf{E} = (\mathbf{E}_1 - \mathbf{E}_0)$ and $\mathbf{HW}\sigma^2_{\mathbf{u}_i}\mathbf{W}^T\mathbf{H}^T \ll \sigma^2_n\mathbf{I}$. Therefore, $\mathbf{\Sigma}_i \simeq \sigma^2_n\mathbf{I}$. The equation (3.23) becomes to

$$\begin{aligned}
&J(p(\mathbf{a}|H_0), p(\mathbf{a}|H_1)) \\
&= \frac{1}{2}\mathrm{Tr}[\frac{\sigma_n^2\mathbf{I}}{\sigma_n^2\mathbf{I}} + \frac{\sigma_n^2\mathbf{I}}{\sigma_n^2\mathbf{I}} + (2(\sigma_n^2\mathbf{I})^{-1}(\mathbf{HWE})(\mathbf{HWE})^T)] - M \\
&= \frac{1}{\sigma_n^2}\mathrm{Tr}(\mathbf{HWEE}^T\mathbf{W}^T\mathbf{H}^T) \\
&= \frac{1}{\sigma_n^2}\mathrm{Tr}(\mathbf{E}^T\mathbf{W}^T\mathbf{H}^T\mathbf{HWE}) \\
&= \frac{1}{\sigma_n^2}\mathbf{E}^T\mathbf{W}^T\mathbf{H}^T\mathbf{HWE}.
\end{aligned} \tag{3.24}$$

Let $\mathbf{w} = [w_1, w_2, ..., w_N]^T$. The equation (3.24) becomes to

$$\begin{aligned}
&J(p(\mathbf{a}|H_0), p(\mathbf{a}|H_1)) \\
&= \frac{1}{\sigma_n^2}\mathbf{w}^T diag(\mathbf{E})\mathbf{H}^T\mathbf{H} diag(\mathbf{E})\mathbf{w}.
\end{aligned} \tag{3.25}$$

Obviously, the equation (3.25) is a convex function. Therefore, when the receiving SNR of the fusion center is low, J-divergence is a convex function. Here we can't prove the J-divergence only has the global optimal solution. We may only find the local maximum by the optimization method. We only know that J-divergence will be a convex function under low SNR condition. But in simulation, the optimal power allocation by the use of J-divergence has much better detection probability than the equal power allocation.

3.3 Optimal Linear Combination

In the optimal linear combination, we combine the received signal by every antenna with weighting factors, like in the Fig. 3.5. Therefore, the signal which we want to detect becomes to

$$\mathbf{a} = \sqrt{g}\mathbf{H}\mathbf{u} + \mathbf{n}, \tag{3.26}$$

where $\mathbf{a} = [a_1, a_2, ..., a_M]^T$, $H = [\tilde{\mathbf{h}}_1, \tilde{\mathbf{h}}_2, ..., \tilde{\mathbf{h}}_M]$, and $\tilde{\mathbf{h}}_i$ is $N \times 1$ channel vector, $i = 1, 2, ..., M$. $\sqrt{g}$ is the power control and it is a constant. The power summation of all secondary users can not exceed P_c. For simplicity, let the power summation of all secondary users are equal to P_c. Then the value of g is

$$g\sum_{i=1}^{N} h_{pi}^2 P_{si} = P_c. \tag{3.27}$$

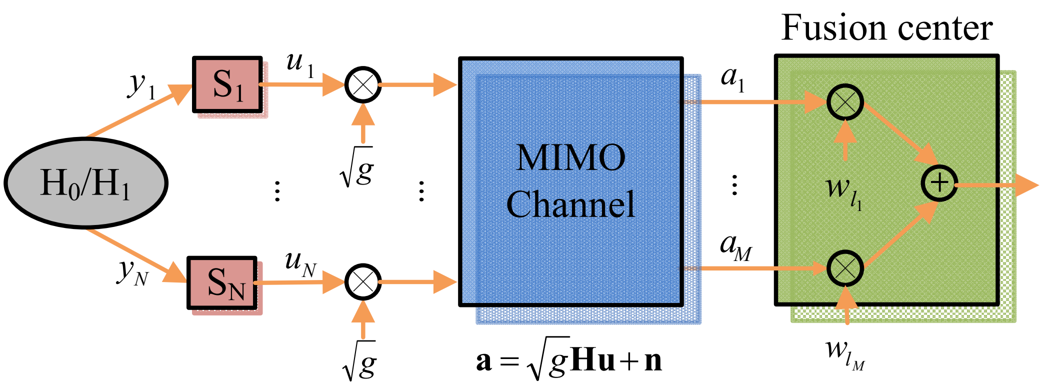


Figure 3.5: System model of optimal power allocation

From the above equation, we can calculate g. The value of P_c is decided by $SINR_t$. In other words, $g = w_1 = w_2 = ... = w_N$ in the optimal linear combination.

a_i is the received signal by the ith antenna. It is the Gaussian distribution and its mean and variance are

$$E_{a_i|H_j} = \sqrt{g}\tilde{\mathbf{h}}_i^T \mathbf{E}_j \tag{3.28}$$

and

$$Var_{a_i|H_j} = g\tilde{\mathbf{h}}_i^T \sigma^2_{\mathbf{u_j}} \tilde{\mathbf{h}}_i + \sigma^2. \tag{3.29}$$

The fusion center can detect the primary users by the signal, r.

$$r = \sum_{i}^{M} w_{l_i} a_i = \mathbf{w}_l^T \mathbf{a}, \tag{3.30}$$

where $\mathbf{w}_l = [w_{l_1} w_{l_2} ... w_{l_M}]^T$. r is a Gaussian random variable and its mean and variance are

$$E_r = \begin{cases} \sqrt{g}\mathbf{w}_l{}^{\mathbf{T}}\mathbf{H}\mathbf{E}_0, & \text{under } H_0 \\ \sqrt{g}\mathbf{w}_l{}^{\mathbf{T}}\mathbf{H}\mathbf{E}_1, & \text{under } H_1 \end{cases} \tag{3.31}$$

and

$$Var_r = \begin{cases} \mathbf{w}_l^T (g\mathbf{H}\sigma^2_{\mathbf{u_0}}\mathbf{H}^T + \sigma_n^2 \mathbf{I})\mathbf{w}_l, & \text{under } H_0 \\ \mathbf{w}_l^T (g\mathbf{H}\sigma^2_{\mathbf{u_1}}\mathbf{H}^T + \sigma_n^2 \mathbf{I})\mathbf{w}_l, & \text{under } H_1. \end{cases} \tag{3.32}$$

Here we use another distance measure, L2 distance, in the optimal linear combination. J-divergence can't optimize in the optimal linear combination. The definition of L2 distance is

$$D_{L2} = \int (f_1 - f_2)^2 d\mathbf{x}, \tag{3.33}$$

where f_1 and f_2 are two distribution functions.

Here we use the a_i in L2 distance instead of r. The L2 distance scheme can be viewed as that the fusion center linearly combines the distribution of every antenna. Obviously, the distribution of r is not the linear combination of the distribution of every antenna. But in the simulation result, L2 distance will have better detection probability than other distance measure.

Rewrite the equation (3.33) as

$$\int (\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}} - \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}})^2 d\mathbf{x}, \tag{3.34}$$

where $\mathbf{w}_l = [w_{l_1}, w_{l_2}, ..., w_{l_M}]^T$ and $\mathbf{P}_{\mathbf{a}_{H_i}} = [p(a_1|H_i), p(a_2|H_i), ..., p(a_M|H_i)]^T$.

From equation (3.28) and (3.29), a_i is a Gaussian random variable. The L2 distance can be write as

$$\begin{aligned} D_{L2}(\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}}, \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}}) &= \int (\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}} - \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}})^2 d\mathbf{x} \\ &= \int [(\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}})^2 - 2\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}} \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}} + (\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}})^2] d\mathbf{x} \\ &= \sum_i \sum_j w_{l_i} w_{l_j} \int p(a_i|H_1) p(a_j|H_1) d\mathbf{x} \\ &\quad -2 \sum_i \sum_j w_{l_i} w_{l_j} \int p(a_i|H_1) p(a_j|H_0) d\mathbf{x} \\ &\quad + \sum_i \sum_j w_{l_i} w_{l_j} \int p(a_i|H_0) p(a_j|H_0) d\mathbf{x}. \end{aligned} \tag{3.35}$$

Assume two Gaussian random variables have the means, μ_a and μ_b, and the variances, σ_a^2 and σ_b^2. The integration of this two random variables is

$$\int N(\mathbf{x}, \mu_a, \sigma_a^2) N(\mathbf{x}, \mu_b, \sigma_b^2) d\mathbf{x} = \frac{1}{\sqrt{det(2\pi(\sigma_a^2 + \sigma_b^2))}} e^{-\frac{1}{2}(\mu_a - \mu_b)^T (\sigma_a^2 + \sigma_b^2)^{-1} (\mu_a - \mu_b)}. \tag{3.36}$$

Therefore, we can rewrite all the multiplications of two Gaussian random variables by the use of three matrices, $\mathbf{M}^{11}$, $\mathbf{M}^{10}$, and $\mathbf{M}^{00}$.

$$\mathbf{M}_{ij}^{11} = \int p(a_i|H_1) p(a_j|H_1) d\mathbf{x}. \tag{3.37}$$

$$\mathbf{M}_{ij}^{10} = \int p(a_i|H_1) p(a_j|H_0) d\mathbf{x}. \tag{3.38}$$

$$\mathbf{M}_{ij}^{00} = \int p(a_i|H_0)p(a_j|H_0)d\mathbf{x}. \tag{3.39}$$

The equation of L2 distance which we want to optimize is

$$\begin{aligned} D_{L2}(\mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_1}}, \mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_0}}) &= \sum_i \sum_j w_{l_i} w_{l_j} \mathbf{M}_{ij}^{11} - 2\sum_i \sum_j w_{l_i} w_{l_j} \mathbf{M}_{ij}^{10} \\ &+ \sum_i \sum_j w_{l_i} w_{l_j} \mathbf{M}_{ij}^{00} \\ &= \mathbf{w}_l^T \mathbf{M}^{11} \mathbf{w}_l - 2\mathbf{w}_l^T \mathbf{M}^{10} \mathbf{w}_l + \mathbf{w}_l^T \mathbf{M}^{00} \mathbf{w}_l \\ &= \mathbf{w}_l^T (\mathbf{M}^{11} - 2\mathbf{M}^{10} + \mathbf{M}^{00}) \mathbf{w}_l. \end{aligned} \tag{3.40}$$

Obviously, the equation we want to optimize is a convex problem. If we don't have any constraint on $\mathbf{w}_l$, the value of the L2 distance will become to infinity. To avoid this problem, let $\mathbf{w}_l^T\mathbf{w}_l = 1$. Therefore, the optimization problem becomes to

$$\begin{aligned} \max_{\mathbf{w}_l} \quad & D_{L2}(\mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_1}}, \mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_0}}) = \mathbf{w}_l^T(\mathbf{M}^{11} - 2\mathbf{M}^{10} + \mathbf{M}^{00})\mathbf{w}_l \\ s.t. \quad & w_{l_i} \geq 0, i = 1, 2, ..., M \\ & \mathbf{w}_l^T\mathbf{w}_l = 1. \end{aligned} \tag{3.41}$$

The optimization problem can be easily solved by the optimization method. Here we use the active set method. The active set method is that checking the inequality constraints are active or not. If the inequality constraints are active, we can view the inequality constraints as the equality constraints and use the Lagrange multiplier to solve this problem. If all the computation results satisfy the Karush-Kuhn-Tucker (KKT) condition, it's one iteration. After several iterations, we can find the optimal solution.

Then we discuss a simple decision rule:

$$r > t, \tag{3.42}$$

where t is the detection threshold. The equation (3.31) and (3.32) show the distribution of r. It's easy to calculate the detection probability and the false alarm probability. The equations of the detection probability and the false alarm probability are

$$P_D = Q\left(\frac{t - E_{r|H_1}}{\sqrt{Var_{r|H_1}}}\right) \tag{3.43}$$

and

$$P_f = Q\left(\frac{t - E_{r|H_0}}{\sqrt{Var_{r|H_0}}}\right). \tag{3.44}$$

The optimal linear combination is simpler than the optimal power allocation. The optimal linear combination uses g for the power control. Every secondary user uses the same scalar, g. Therefore, the fusion center can broadcast g to secondary users. But the optimal power allocation should tell secondary users their own weighting factors. In the L2 distance method, the antenna which has the better L2 distance, its weighting will be larger. For example, if there are 2 antennas, the L2 distance of the first antenna is greater than the second antenna, its weighting is larger.

3.4 Simulation Results

In this section, we present the simulation by the use of the distance measures. Assume the MIMO channel and the noise are all random and the power of the primary user is known. The SNR of the primary user is 5. The probability that primary users use the frequency band is 0.2 and the probability that don't use the frequency band is 0.8 because the primary users seldom exist. In the following, we will discuss simulation of the optimal power allocation and the optimal linear combination.

3.4.1 Optimal Power Allocation

In the optimal power allocation scheme, the distance measure is J-divergence . In this scheme, the decision rule is Neyman-Pearson detection at the fusion center. We set that the false alarm probability is 0.4. It means that the spectrum utilization is 60% in this simulation environment. As we mentioned before, it's hard to calculate and optimize the detection probability directly because LLR is not a linear combination of $\mathbf{a}$. Therefore, we simulate the detection probability by the Monte Carlo method.

Fig. 3.6 shows two different transmitting signals of secondary users. If the signal doesn't have any signal processing and transmits it directly to the fusion center, its detection probability is worse than the signal with the signal processing. Obviously, it's trade off between the sensing time and the detection probability because the signal with the signal processing needs more sensing samples.

Fig. 3.7(a) shows when the target SINR is high, the value of J-divergence is low. When the target SINR is high, it means that the total power of secondary users can use is low. In the equal power allocation, the detection probability of J-divergence is lower than the optimal power allocation. But only Fig. 3.7(a) can't prove the J-divergence can be viewed as a performance matric.

Fig. 3.7(b) shows the detection probability. In Fig. 3.7(a) and Fig. 3.7(b), we can observe that when the value of J-divergence is higher, the detection probability is higher. Therefore, J-divergence can be a performance matric in the optimal power allocation. And in Fig. 3.7(b), the detection probability of the optimal power allocation is much better than equal power allocation.

Fig. 3.8(a) shows four cases, 2x2, 2x4, 4x2, and 4x4. 4x2 means there are 4 receive antennas and 2 secondary users in the simulation, and so on. Obviously, the more antennas or the more secondary users, the better performance. Therefore, the fusion center can have benefit by using the MIMO channel in the optimal linear combination. Then compare the 2x4 case and the 4x2 case. In this figure, the 2x4 case is better than the 4x2 case. The reason is that it optimizes the detection probability by multiplying the weighting factors to secondary users in the optimal power allocation. If there are more users, we can have more degrees of freedom. Therefore, if there are more users or more receiving antennas of the fusion center in cognitive radio system, the detection probability will be better. But increasing the number of secondary users is better than increasing the number of receiving antenna of the fusion center. Fig. 3.9 shows the error probability and the lower bound. If the target SINR is lower, the error probability will also be lower. From the equation (3.18), $P_e > P(H_0)P(H_1)e^{(-\frac{J}{2})}$, the error probability has a lower bound. Fig. 3.10 shows this property.

3.4.2 Optimal Linear Combination

In this simulation, we also use the Neyman-Pearson decision rule in the fusion center. And in the later section, we will compare the performance between the optimal power allocation and the optimal linear combination.

Fig. 3.11(a) shows the simulation result of the optimal linear combination by using L2 distance. In this figure, when target SINR is high, it means the power that secondary

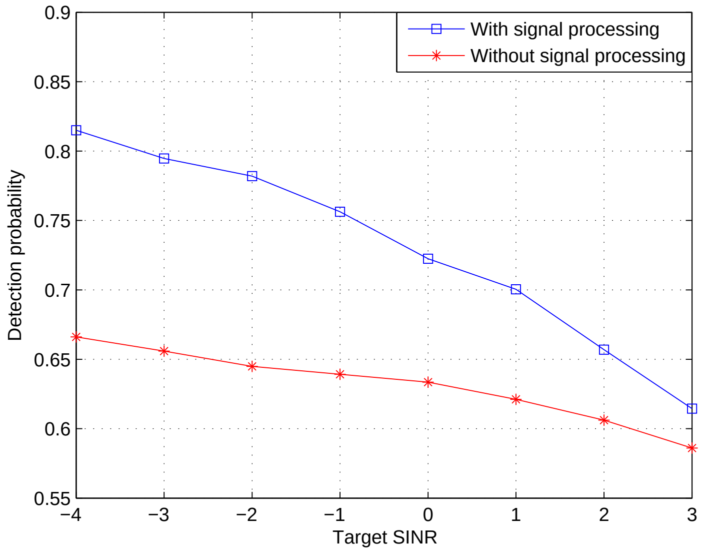


Figure 3.6: Signal processing and without signal processing

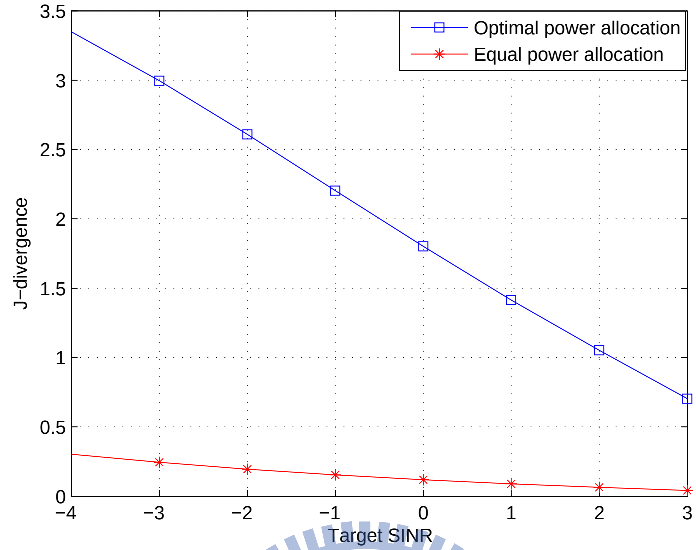
users can use for transmitting is low and the L2 distance is small.

Fig. 3.11(b) shows the detection probability. This figure can tell us that L2 distance can be really as a performance matric. When the L2 distance is larger, the detection probability is larger. Therefore, we can use this distance measure to maximize the detection probability in the linear combination.

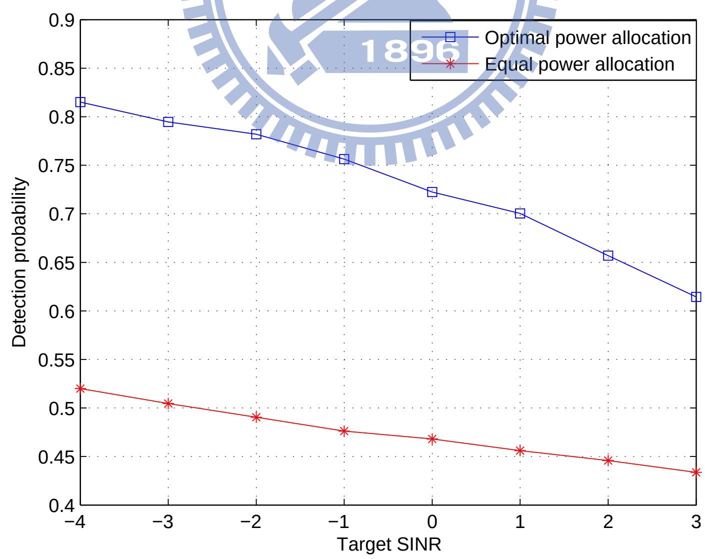
Fig. 3.12(a) shows the four cases of optimal linear combination, 2x2, 2x4, 4x2, 4x4. In this figure, the 4x2 case is better than the 2x4 case. The reason is that it optimizes the detection probability by the use of the weighting factor of every receive antenna in the optimal linear combination. Therefore, if there are more antennas in the fusion center, the degree of freedom is high. Increasing the number of receiving antennas of the fusion center is better than increasing the number of secondary users. Fig. 3.12(b) shows the detection probability. In this figure, when L2 distance is high, the detection probability will also be high.

3.4.3 Comparison

In this subsection, we will compare 2 schemes, the optimal power allocation and the optimal linear combination.

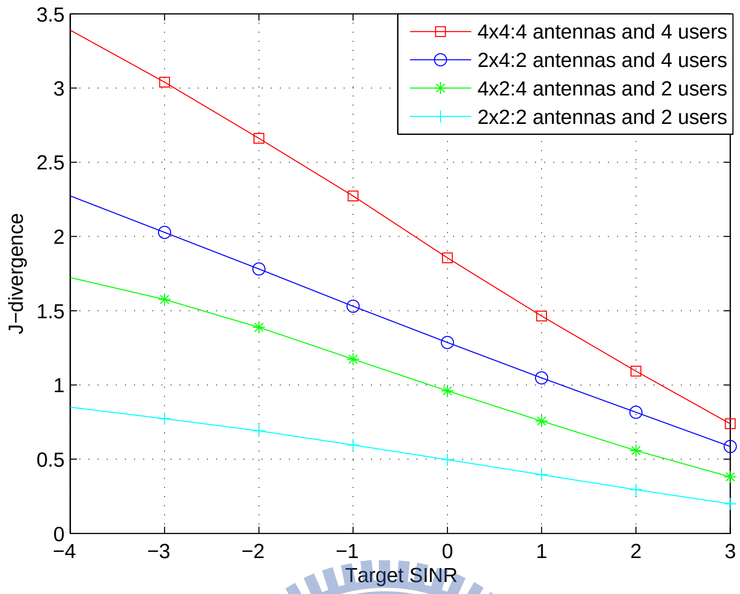


(a) J-divergence

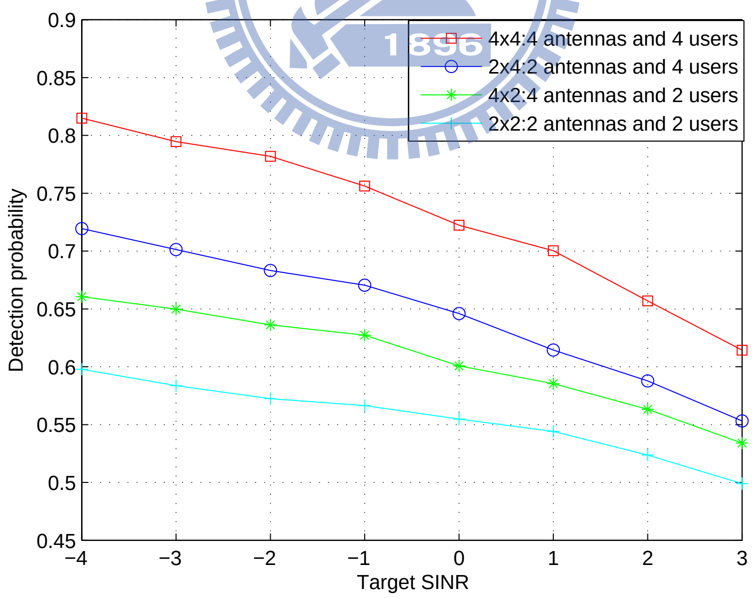


(b) Detection probability

Figure 3.7: J-divergence and the detection probability of optimal power allocation and equal power allocation



(a) J-divergence



(b) Detection probability

Figure 3.8: Detection probability and J-divergence of four cases in optimal power allocation

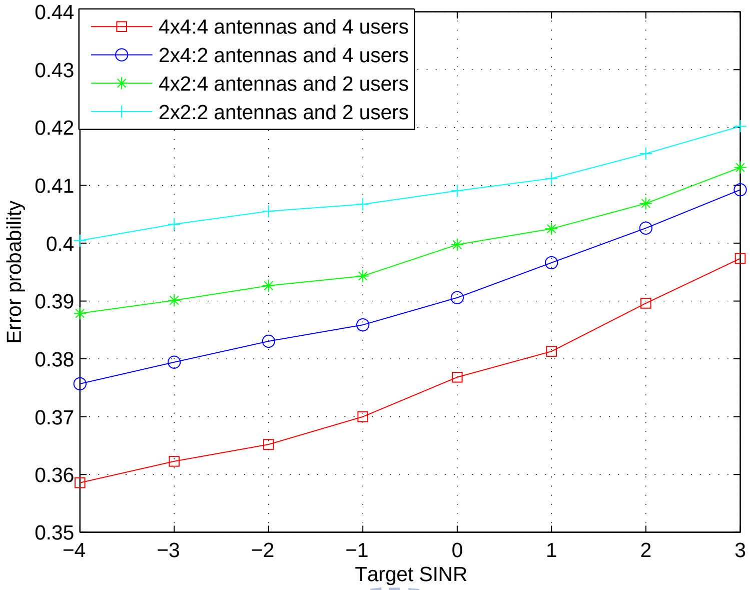


Figure 3.9: The error probability of four cases in optimal power allocation

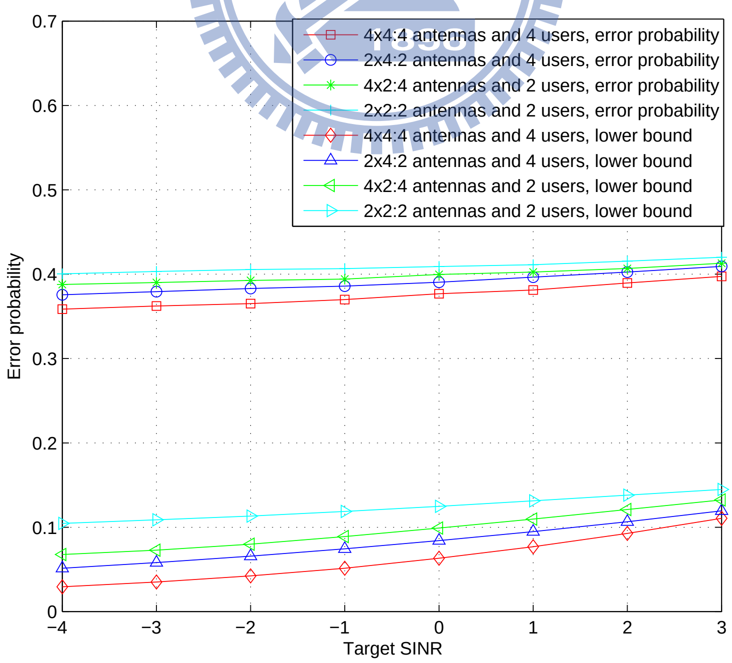
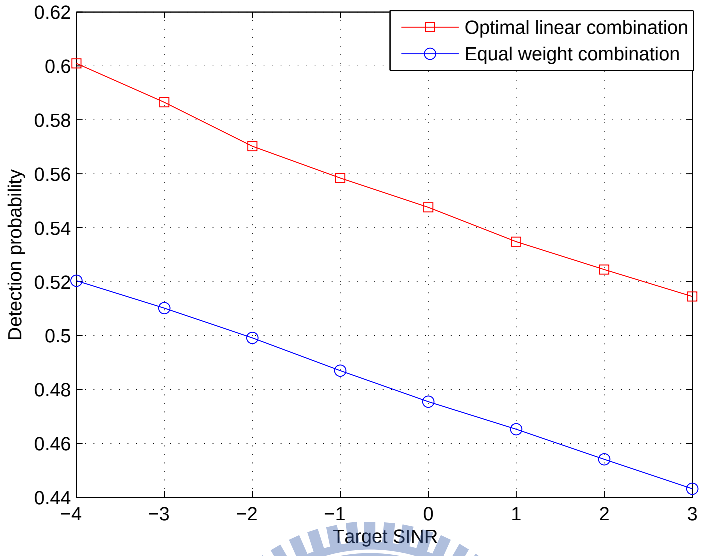
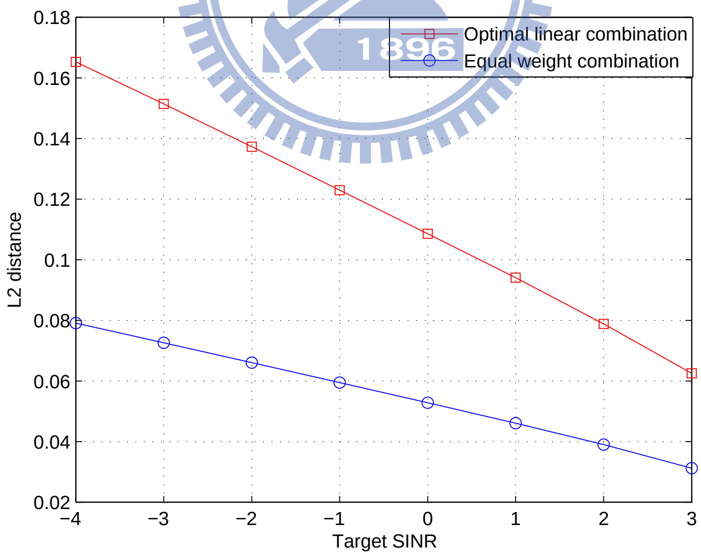


Figure 3.10: The error probability and the lower bound of four cases in optimal power allocation

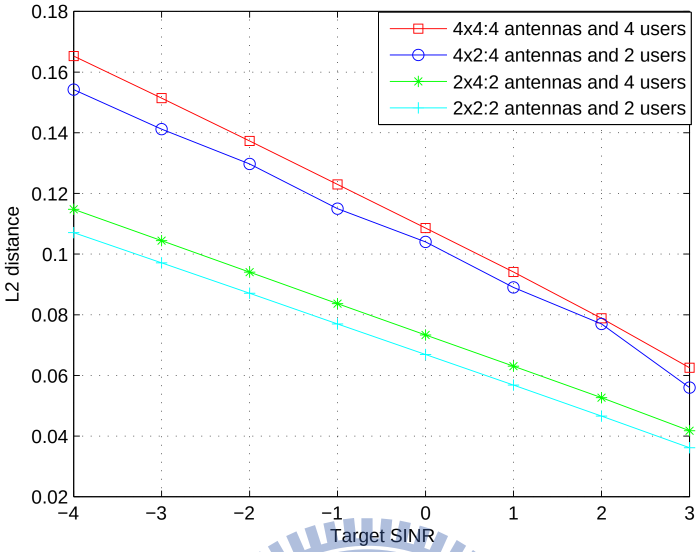


(a) L2 distance

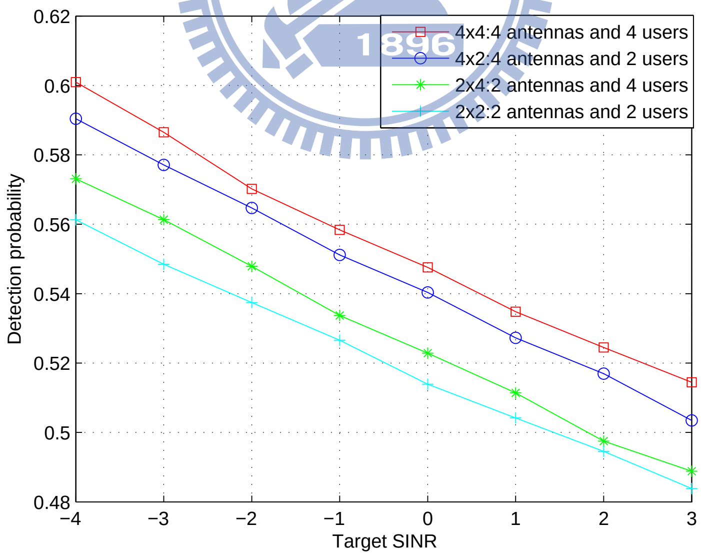


(b) Detection probability

Figure 3.11: The L2 distance and detection probability of optimal linear combination and equal weighting combination



(a) L2 distance



(b) Detection probability

Figure 3.12: The L2 distance and detection probability of four cases in optimal linear combination

By the equation (3.43) and (3.44), we can calculate P_D and P_f by the use of a simple detection method in the optimal linear combination. In [12], it focuses on the optimization of optimal linear combination in the cognitive radio and proposes 3 optimization methods in 3 different systems, conservative system, aggressive system, and hostile system. In this paper, we can know the optimization of the equation (3.43) is hard. Only in the aggressive system, the equation (3.43) is a convex problem. But in the final part of this paper, it proposes a modified deflection coefficient (MDC) method. In the MDC method, it's also a distance measure and combines the receiving data of every antenna in the fusion center. In other words, it's also a distance measure in the optimal linear combination. Let $\mathbf{h}' = [h_{11}^2 h_{22}^2 ... h_{NN}^2]^T$.

The definition of the MDC is

$$d_m^2(\mathbf{w}_l) = \frac{(E_s \mathbf{h}'^T \mathbf{w}_l)^2}{\mathbf{w}_l^T Var_r \mathbf{w}_l}. \tag{3.45}$$

And the optimization problem is

$$\begin{aligned} \max_{\mathbf{w}_l} \quad & d_m^2(\mathbf{w}_l) \\ s.t. \quad & \mathbf{w}_l^T \mathbf{w}_l = 1. \end{aligned} \tag{3.46}$$

and let $\mathbf{w}_l' = Var_{r|H_1}^{-\frac{T}{2}} \mathbf{h}'$. The optimal solution of equation (3.46) is

$$\mathbf{w}_{l_{opt}} = \frac{Var_{r|H_1}^{-\frac{1}{2}} \mathbf{w}_l'}{\|Var_{r|H_1}^{-\frac{1}{2}} \mathbf{w}_l'\|_2}. \tag{3.47}$$

Fig. 3.13 shows the comparison between the MDC method and the L2 distance method. Consider a simple decision rule, $r > t$. In this decision rule, we can fix the value of P_f and maximize the detection probability. In Fig. 3.13, the detection probability of the L2 distance method is little better than the detection probability of the MDC method under the same false alarm probability. But in [12], the system model of the MDC method is in the orthogonal channel. But L2 distance can also use not only in the orthogonal channel. The L2 distance method is more general than the MDC method.

Fig. 3.14 shows the detection probability of 3 schemes, optimal power allocation, optimal linear combination, and MDC. This simulation uses 2 secondary users and 2 receiving antennas in the fusion center. In this simulation, the decision criterion in the fusion center is the Neyman-Pearson detection. The performance of the optimal linear combination is the best and the MDC is the worst. The optimal power allocation is the best because it controls the power of secondary users. If one of secondary users has lower sensing noise or receives a stronger signal from primary users, it will have more transmitting power and the fusion center can be more sure which hypothesis is right. The optimal linear combination only combines the received data, so it can't have benefit on the data directly. Therefore, the detection probability of the optimal linear combination is worse than the detection probability of the optimal power allocation.

The MDC method also uses in the optimal linear combination, like L2 distance. But comparing to L2 distance, the MDC method only uses variance and mean for the maximization, it's worse than the L2 distance method in the simulation results.

Fig. 3.15 and Fig. 3.16 show that the simulated optimal detection probability and the distance based detection probability. In this figure, we can know the distance based detection probability will smaller than the real optimal detection probability. But the computation complexity of the distance based detection method is much lower than finding the optimal detection probability exhaustedly.

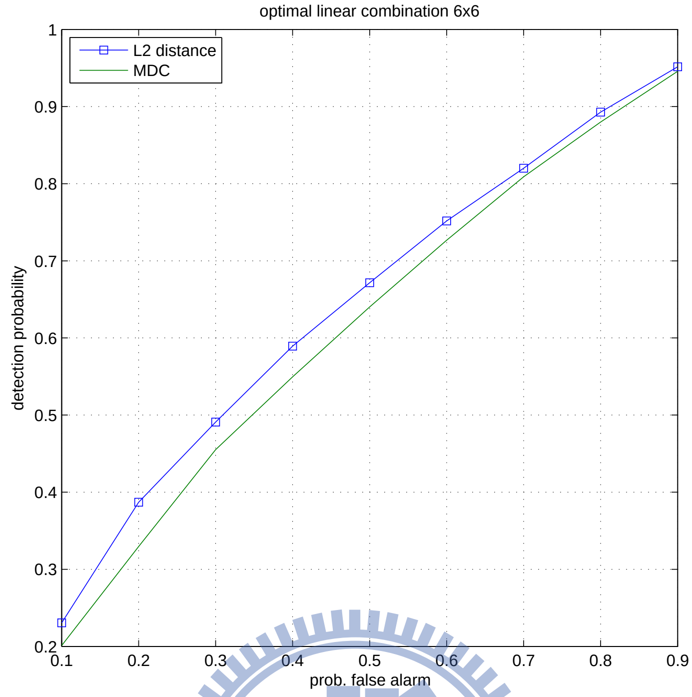


Figure 3.13: MDC and L2 distance with fixed false alarm probability by the use of the simple detection

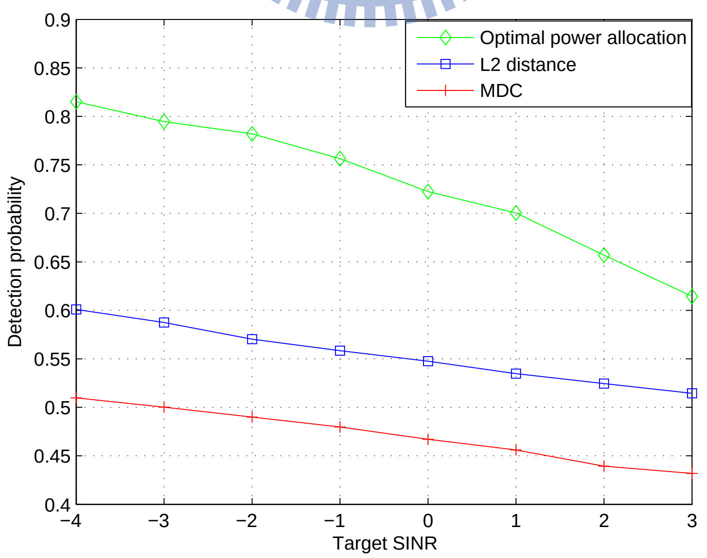


Figure 3.14: Comparison between optimal power allocation, L2 distance, and MDC by the use of Neyman-Pearson detection

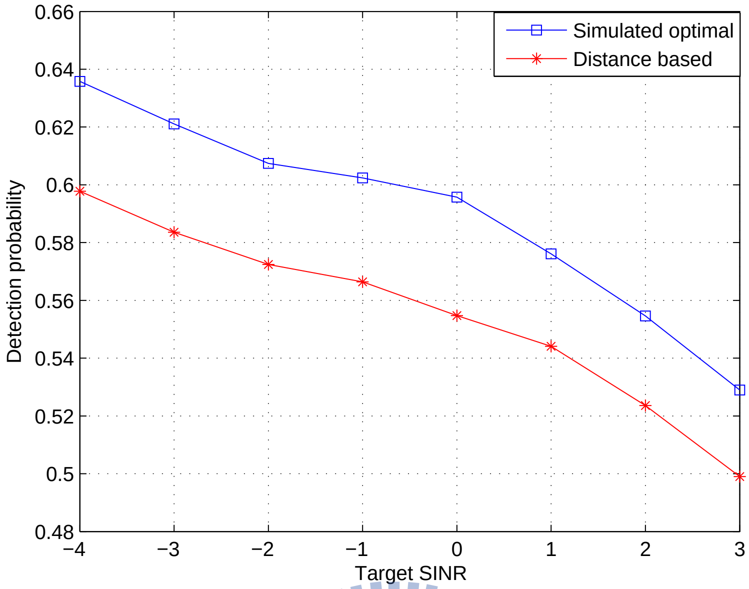


Figure 3.15: Simulated optimal detection probability and distance based detection probability in optimal power allocation

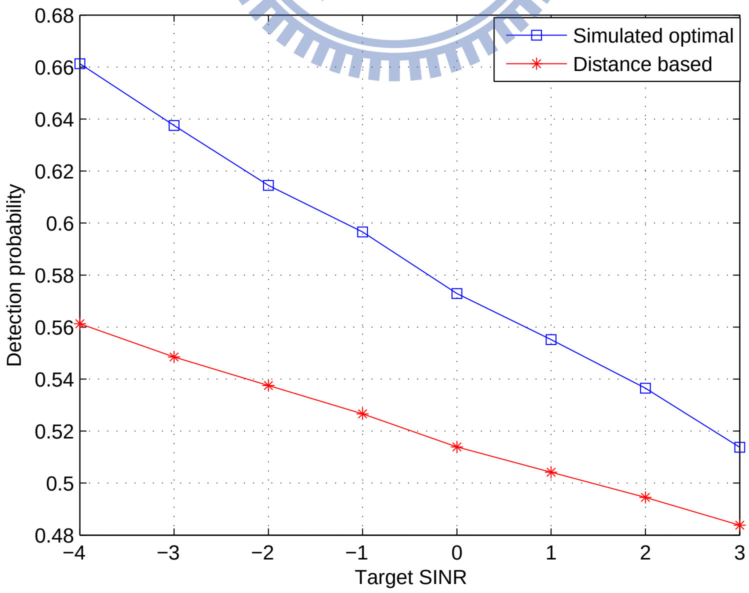


Figure 3.16: Simulated optimal detection probability and distance based detection probability in optimal linear combination

3.5 Summary

In this chapter, we discuss the optimal power allocation and the optimal linear combination. We adopt J-divergence as a performance matric in the optimal power allocation and adopt L2 distance as a performance matric in the optimal linear combination. In the optimal power allocation, the detection probability is better than the detection probability in the equal power allocation. The detection probability of the optimal linear combination is better than the detection probability of the equal weighting combination, too. If two schemes use the same decision rule, such as the Neyman-Pearson detection rule, the detection probability of the optimal linear combination is worse than the detection probability of the optimal power allocation. But as mentioned before, the optimal linear combination is simpler than the optimal power allocation.

Chapter 4

Censoring Scheme in Centralized Detection

4.1 Introduction

In the cognitive environment, when secondary users transmit their observations to the fusion center, they should reduce interference when primary users exists. Here we propose a new method, censoring, to reduce the interference of primary users. Before secondary users transmit their observations to the fusion center, they can judge their observations whether the observations are worthy to transmit or not. For example, if the observation is too small or too high, it may not be worth to transmit and the secondary users can keep silence. In this chapter, we will discuss a censoring scheme.

The censoring scheme means that every secondary user doesn't transmit all data they sensing. In [22], the users transmit the log likelihood ratio (LLR) to the fusion center and the fusion center uses received LLR to make a decision. But we don't transmit the LLR to the fusion center. Because every secondary user uses the energy detection method, it's not easy to have the distribution of the LLR of transmitting signal. Therefore, we consider another simple method. For example, when the signal sensed by the secondary user is great than γ, the secondary user will transmit this signal to the fusion center. If the signal is less than γ, this secondary user will keep silent. Therefore, the distribution of the signal transmitted by secondary users can be viewed as a truncated Gaussian distribution, like

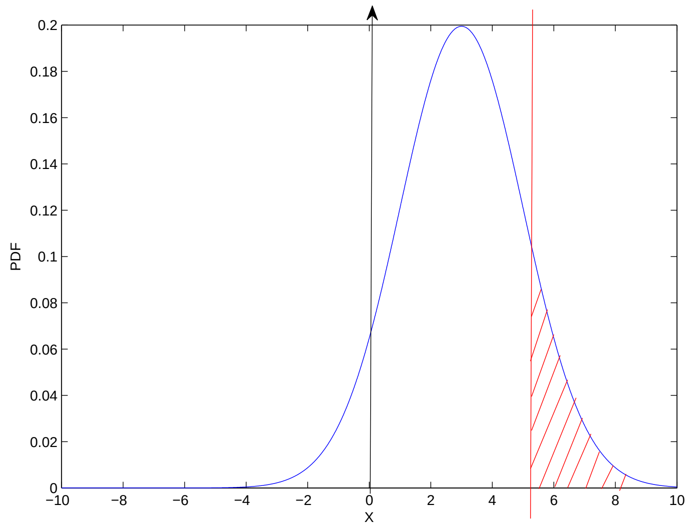


Figure 4.1: Truncated gaussian

in Fig. 4.1. In Fig. 4.1, when the observation is greater than γ, this observation will be transmitted. Like in chapter 3, here we discuss two schemes, the optimal power allocation and the optimal linear combination.

4.2 Optimal Power Allocation

Here we still use the optimal power allocation to satisfy the power constraint. We still have two objectives:

- Maximize the detection probability
- Satisfy the power constraint

When all secondary users transmit their observations to the fusion center at the same time, they can not produce much interference on primary users and maximize the detection probability.

4.2.1 System Model

Fig. 4.2 shows the system model of the optimal power allocation. For simplicity, it's the orthogonal MIMO channel between the fusion center and secondary users. In Fig. 4.1, the

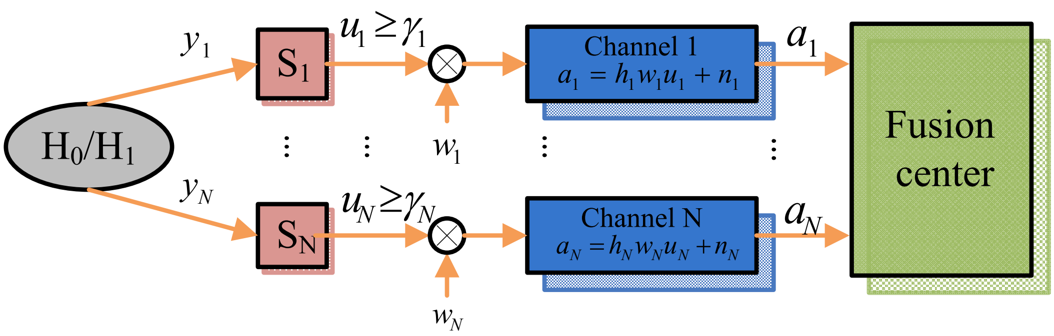


Figure 4.2: System model of censoring scheme

transmitting signal of every secondary user is a truncated Gaussian. If sensing signal is greater than γ, the secondary users will transmit it. By using moment generating function, we can have the mean and the variance of the truncated Gaussian. Assume the distribution of the signal of the ith secondary user sensed is f_{t_i} and its mean and variance are μ_{t_i} and $\sigma_{t_i}^2$. From Fig. 4.2, the truncated signal, u_i, is added with a Gaussian noise, n. Therefore, the received signal of fusion center is

$$\mathbf{a} = \mathbf{HWu} + \mathbf{n}. \tag{4.1}$$

Therefore, the distribution of $\mathbf{a}$ is

$$f_{a_j}(x) = \int_{-\infty}^{+\infty} f_{t_j}(x-\tau) f_n(\tau) d\tau, \tag{4.2}$$

where f_{t_j} is the distribution of a truncated Gaussian, f_n is the distribution of noise received by fusion center, and f_{a_j} is the distribution received by jth antenna.

Because the equation (4.2) doesn't have the closed form expression, we approximate this distribution to the Gaussian distribution. Assume the mean and variance of the signal sensed by ith secondary user are μ_{t_i} and $\sigma_{t_i}^2$. We calculate the moment generating function of a truncated Gaussian random variable with threshold γ_i. Q-function here is used for calculating the probability when distribution is a Gaussian random variable, define as:

$$Q(\gamma) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(-\frac{x^2}{2})} dx. \tag{4.3}$$

The distribution of a truncated Gaussian with threshold γ is

$$\begin{cases} f(x) = \frac{1}{\sqrt{2\pi}\sigma_{t_i} Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} e^{(-\frac{(x-\mu_{t_i})^2}{2\sigma_{t_i}^2})}, & x \geq \gamma_i \\ f(x) = 0, & x < \gamma_i. \end{cases} \tag{4.4}$$

The moment generating function is

$$\begin{aligned} E[e^{tx}] &= \int_{-\infty}^{\infty} e^{tx} f(x) dx \\ &= \int_{\gamma_i}^{\infty} e^{tx} f(x) dx \\ &= \int_{\gamma_i}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma_{t_i} Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} e^{(-\frac{(x-\mu_{t_i})^2}{2\sigma_{t_i}^2})} dx \\ &= \frac{1}{Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} \int_{\gamma_i - \mu_{t_i}}^{\infty} e^{t(\tau + \mu_{t_i})} \frac{1}{\sqrt{2\pi\sigma_{t_i}}} e^{(-\frac{(\tau)^2}{2\sigma_{t_i}^2})} d\tau \\ &= \frac{e^{\mu_{t_i} t + \frac{\sigma_{t_i}^2 t^2}{2}}}{Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} \int_{\gamma_i - \mu_{t_i}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2})(\frac{\tau^2}{\sigma {t_i}^2} - 2\frac{\sigma t_i t\tau}{\sigma t_i} + \sigma {t_i}^2 t^2)} \frac{d\tau}{\sigma_{t_i}} \\ &= \frac{e^{\mu_t t + \frac{\sigma_{t_i}^2 t^2}{2}}}{Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} \int_{\gamma_i - \mu_{t_i}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2})(\frac{\tau}{\sigma t_i} - \sigma t_i t)^2} \frac{d\tau}{\sigma_{t_i}} \\ &= \frac{e^{\mu_{t_i} t + \frac{\sigma_{t_i}^2 t^2}{2}}}{Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} \int_{\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}} - \sigma_{t_i} t}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2})z^2} dz \\ &= \frac{Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}} - \sigma_{t_i} t)}{Q(\frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}})} e^{\mu_{t_i} t + \frac{\sigma_{t_i}^2 t^2}{2}}, \end{aligned} \tag{4.5}$$

where $\tau = x - \mu_{t_i}$ and $z = \frac{\tau}{\sigma_{t_i}} - \sigma_{t_i} t$. From the above, the moment generating function is

$$M_{t_i}(t) = \frac{Q(\gamma_i' - \sigma_{t_i} t)}{Q(\gamma_i')} e^{\mu_{t_i} t + \frac{\sigma_{t_i}^2 t^2}{2}}, \tag{4.6}$$

where $\gamma_i' = \frac{\gamma_i - \mu_{t_i}}{\sigma_{t_i}}$.

From $M_{t_i}(t)$, the mean and the variance of the signal that ith secondary user transmitting are

$$\begin{aligned} \mu_{t_i}' &= M_{t_i}'(t)|_{t=0} \\ &= \frac{\frac{\sigma_{t_i}}{\sqrt{2\pi}} e^{-\frac{(\gamma_i' - \sigma_{t_i} t)^2}{2}}}{Q(\gamma_i')} e^{\mu_{t_i} t + \frac{\sigma_{t_i}^2 t^2}{2}} |_{t=0} \\ &+ \frac{Q(\gamma_i' - \sigma_{ti} t)}{Q(\gamma_i')} (\mu_{t_i} + \sigma_{t_i}^2 t) e^{\mu_{t_i} t + \frac{\sigma_{t_i}^2 t^2}{2}} |_{t=0} \\ &= \frac{\frac{\sigma_{t_i}}{\sqrt{2\pi}} e^{-\frac{(\gamma_i')^2}{2}}}{Q(\gamma_i')} + \mu_{t_i} \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} M''_{t_i}(t)|_{t=0} &= \frac{\frac{\sigma_{t_i}^2}{\sqrt{2\pi}}e^{-\frac{(\gamma'_i-\sigma_{t_i}t)^2}{2}}}{Q(\gamma'_i)}e^{\mu_{t_i}t+\frac{\sigma_{t_i}^2t^2}{2}}(\gamma'_i-\sigma_{t_i}t)|_{t=0} \\ &+\frac{\frac{\sigma_{t_i}}{\sqrt{2\pi}}e^{-\frac{(\gamma'_i-\sigma_{t_i}t)^2}{2}}}{Q(\gamma'_i)}e^{\mu_{t_i}t+\frac{\sigma_{t_i}^2t^2}{2}}(\mu_{t_i}+\sigma_{t_i}^2t)|_{t=0} \\ &+\frac{\frac{\sigma_{t_i}^2}{\sqrt{2\pi}}e^{-\frac{(\gamma'_i-\sigma_{t_i}t)^2}{2}}}{Q(\gamma'_i)}(\mu_{t_i}+\sigma_{t_i}^2t)e^{\mu_{t_i}t+\frac{\sigma_{t_i}^2t^2}{2}}|_{t=0} \\ &+\frac{Q(\gamma'_i-\sigma_ti t)}{Q(\gamma'_i)}(\sigma_{t_i}^2)e^{\mu_{t_i}t+\frac{\sigma_{t_i}^2t^2}{2}}|_{t=0} \\ &+\frac{Q(\gamma'_i-\sigma_ti t)}{Q(\gamma'_i)}(\mu_{t_i}+\sigma_{t_i}^2t)e^{\mu_{t_i}t+\frac{\sigma_{t_i}^2t^2}{2}}(\mu_{t_i}+\sigma_{t_i}^2t)|_{t=0} \\ &= \frac{\frac{\sigma_{t_i}^2\gamma'_i}{\sqrt{2\pi}}e^{-\frac{\gamma_i'^2}{2}}}{Q(\gamma'_i)}+\frac{\frac{\sigma_{t_i}}{\sqrt{2\pi}}e^{-\frac{\gamma_i'^2}{2}}}{Q(\gamma'_i)}\mu_{t_i}+\frac{\frac{\sigma_{t_i}}{\sqrt{2\pi}}e^{-\frac{\gamma_i'^2}{2}}}{Q(\gamma'_i)}\mu_{t_i}+\sigma_{t_i}^2+\mu_{t_i}^2. \end{aligned} \tag{4.8}$$

Form equation (4.7) and equation (4.8), the variance is

$$\begin{aligned} \sigma_{t_i}'^2 &= M''_{t_i}(t)|_{t=0}-(M'_{t_i}(t)|_{t=0})^2 \\ &= \frac{\sigma_{t_i}'^2\gamma'_i}{\sqrt{2\pi}Q(\gamma'_i)}e^{-\frac{\gamma_i'^2}{2}}-\frac{\sigma_{t_i}^2}{2\pi Q(\gamma')^2}e^{-\gamma_i'^2}+\sigma_{t_i}^2. \end{aligned} \tag{4.9}$$

Therefore, the mean and the variance of transmitting signal of ith secondary user are

$$\mu_{i|H_j,u_i}=\begin{cases}\mu'_{t_{i|H_j}}, & \text{under } u_i\neq 0, H_j\\ 0, & \text{under } u_i=0, H_j\end{cases} \tag{4.10}$$

and

$$\sigma_{i|H_j,u_i}^2=\begin{cases}\sigma_{t_{i|H_j}}'^2, & \text{under } u_i\neq 0, H_j\\ 0, & \text{under } u_i=0, H_j.\end{cases} \tag{4.11}$$

The receiving signal of ith antenna of fusion center is $a_i = h_iw_iu_i+n$. The noise distribution is a Gaussian distribution, $N(0,\sigma_n^2)$.

From the equation (4.10) and the equation (4.11), the mean and the variance of antenna a_i under H_j hypothesis are

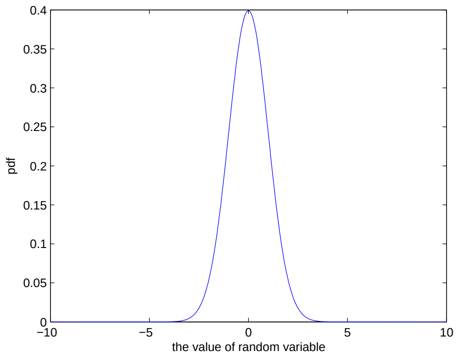
$$\mu_{a_{i|H_j,u_i}}=\begin{cases}h_iw_i\mu'_{t_{i|H_j}}, & \text{under } u_i\neq 0, H_j\\ 0, & \text{under } u_i=0, H_j\end{cases} \tag{4.12}$$

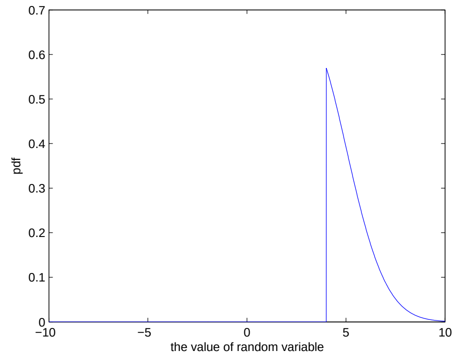
and

$$\sigma_{a_{i|H_j,u_i}}^2=\begin{cases}h_i^2w_i^2\sigma_{t_{i|H_j}}'^2+\sigma_n^2, & \text{under } u_i\neq 0, H_j\\ \sigma_n^2, & \text{under } u_i=0, H_j.\end{cases} \tag{4.13}$$

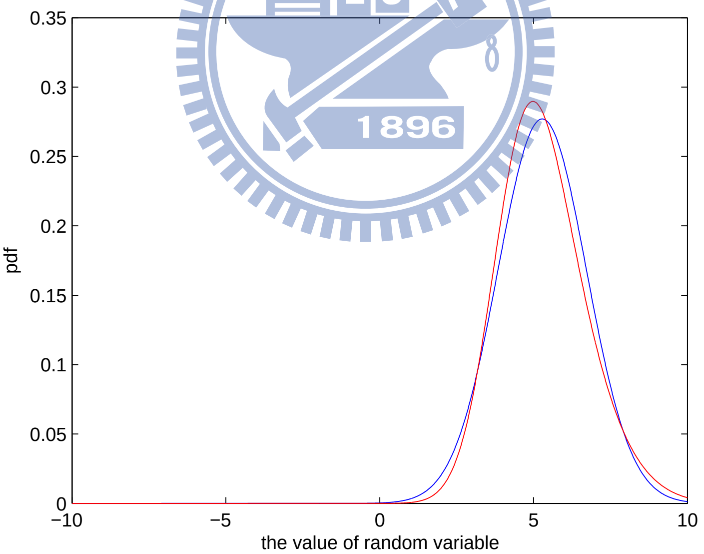
From the above equation, we can have the approximated distribution of $\mathbf{a}$.

Fig. 4.3(a) is the noise distribution and Fig. 4.3(b) is the truncated distribution. In Fig 4.3(c), the red line is real distribution of a_i and the blue line is the approximation





(a) noise distribution (b) truncated distribution



(c) approximated distribution

Figure 4.3: Approximation of truncated Gaussian

of a_i distribution. In Fig. 4.3(c), the approximated distribution of $\mathbf{a}$ is close to the real distribution.

Here we use u_i to present the data of ith secondary user transmitting to the fusion center. When $u \neq 0$, it means the secondary users send the data to the fusion center. When $u = 0$, it means the secondary user doesn't send the data to the fusion center. Therefore, there is 2^N combinations of $\mathbf{u}$. Assume the probability that ith secondary user sends data to the fusion center is $Psend_{i,j}$ and j means that it's under hypothesis H_j. The equation of $p(\mathbf{u}_k|H_j)$ is

$$p(\mathbf{u}_k|H_j) = \prod_{u_m \neq 0} P_{send_{m,j}} \prod_{u_t = 0} (1 - P_{send_{t,j}}), \tag{4.14}$$

where $k = 1, 2, ..., 2^N$ and k is kth combination.

Therefore, the distribution of $\mathbf{a}$ under different hypothesis is

$$p(\mathbf{a}|H_i) = \sum_k p(\mathbf{a}|\mathbf{u}_k, H_i) p(\mathbf{u}_k|H_i). \tag{4.15}$$

Obviously, $p(\mathbf{a}|H_i)$ is a Gaussian mixture model(GMM) because $\sum_k p(\mathbf{u}_k|H_i) = 1$ and $p(\mathbf{a}|\mathbf{u}_k, H_i)$ is a Gaussian random variable. The likelihood ratio is

$$\begin{aligned} L(\mathbf{a}) &= \frac{p(\mathbf{a}|H_1)}{p(\mathbf{a}|H_0)} \\ &= \frac{\sum_k p(\mathbf{a}|\mathbf{u}_k, H_i) p(\mathbf{u}_k|H_1)}{\sum_k p(\mathbf{a}|\mathbf{u}_k, H_i) p(\mathbf{u}_k|H_0)}. \end{aligned} \tag{4.16}$$

Obviously, solving this problem by $L(\mathbf{a})$ is hard. We still need the distance measure to solve the Gaussian mixture model problem.

4.2.2 GMM J-divergence

Here we consider the optimal power allocation. But it doesn't have the closed-form expression of J-divergence in the Gaussian mixture model(GMM). Therefore, we must modify the equation of J-divergence in GMM.

Consider two different Gaussian mixture models(GMM),

$$\sum_i \alpha_i f_i = \alpha^T \mathbf{f}, \tag{4.17}$$

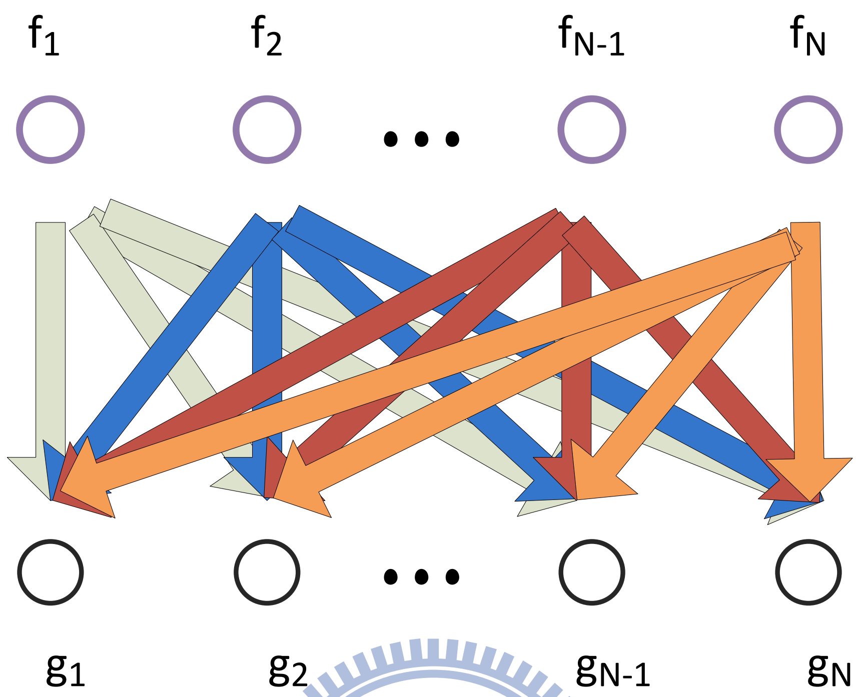


Figure 4.4: Structure of GMM J-divergence

where $\alpha = [\alpha_1, \alpha_2, ..., \alpha_N]^T$ and $\mathbf{f} = [f_1, f_2, ..., f_N]^T$

and

$$\sum_j \beta_j g_j = \beta^T \mathbf{g}, \tag{4.18}$$

where $\beta = [\beta_1, \beta_2, ..., \beta_N]^T$ and $\mathbf{g} = [g_1, g_2, ..., g_N]^T$.

Define J-divergence in the GMM. As Fig. 4.4 showing, the structure of GMM J-divergence is

$$\sum_i \sum_j \alpha_i \beta_j J(f_i, g_j). \tag{4.19}$$

In Fig. 4.4, we can know the GMM J-divergence adds the J-divergence of every f_i and g_i together. $\sum_i \alpha_i f_i$ is equal to $\sum_i p(\mathbf{a}|\mathbf{u}_i, H_1)p(\mathbf{u}_i|H_1)$ and $\sum_j \beta_j g_i$ is equal to $\sum_j p(\mathbf{a}|\mathbf{u}_j, H_0)p(\mathbf{u}_j|H_0)$. Therefore, we can rewrite equation

$$J_{GMM}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0)) = \sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)J(p(\mathbf{a}|\mathbf{u}_i, H_1), p(\mathbf{a}|\mathbf{u}_j, H_0)). \tag{4.20}$$

The optimization problem becomes to

$$\begin{array}{ll} \max_{\mathbf{w}} & J_{GMM}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0)) \\ s.t. & SINR_p \geq SINR_t. \end{array} \tag{4.21}$$

Obviously ,this is the linear combination of all J-divergence. We can also use the optimization method to solve this problem.

4.2.3 GMM L2 Distance

L2 distance can also use the in optimal power allocation. From equation (4.15), we know that $p(\mathbf{a}|H_i)$ is a Gaussian mixture model. The L2 distance compares the distance between $p(\mathbf{a}|H_1)$ and $p(\mathbf{a}|H_0)$. The equation of GMM L2 distance is

$$\begin{aligned} D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0)) &= \int (p(\mathbf{a}|H_1) - p(\mathbf{a}|H_0))^2 d\mathbf{x} \\ &= \int (\sum_i p(\mathbf{u}_i|H_1)p(a|\mathbf{u}_i, H_1) - \sum_j p(\mathbf{u}_j|H_0)p(a|\mathbf{u}_j, H_0))^2 d\mathbf{x} \\ &= \sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1) \int p(a|\mathbf{u}_i, H_1)p(a|\mathbf{u}_j, H_1) d\mathbf{x} \\ &\quad -2\sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0) \int p(\mathbf{a}|\mathbf{u}_i, H_1)p(\mathbf{a}|\mathbf{u}_j, H_0) d\mathbf{x} \\ &\quad +\sum_i \sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0) \int p(\mathbf{a}|\mathbf{u}_i, H_0)p(\mathbf{a}|\mathbf{u}_j, H_0) d\mathbf{x} \\ &= \mathbf{P_{u,H_1}}^T \mathbf{M_a}^{11} \mathbf{P_{u,H_1}} \\ &\quad -2\mathbf{P_{u,H_1}}^T \mathbf{M_a}^{10} \mathbf{P_{u,H_0}} \\ &\quad +\mathbf{P_{u,H_0}}^T \mathbf{M_a}^{00} \mathbf{P_{u,H_0}}, \end{aligned} \tag{4.22}$$

where

$$\begin{aligned} \mathbf{P_{u,H_1}} &= [p(\mathbf{u}_1|H_1), p(\mathbf{u}_2|H_1), ..., p(\mathbf{u}_N|H_1)]^T \\ \mathbf{P_{u,H_0}} &= [p(\mathbf{u}_1|H_0), p(\mathbf{u}_2|H_0), ..., p(\mathbf{u}_N|H_0)]^T \\ \mathbf{M}_{\mathbf{a}ij}^{11} &= \int p(\mathbf{a}|\mathbf{u}_i, H_1)p(\mathbf{a}|\mathbf{u}_j, H_1) d\mathbf{x} \\ \mathbf{M}_{\mathbf{a}ij}^{10} &= \int p(\mathbf{a}|\mathbf{u}_i, H_1)p(\mathbf{a}|\mathbf{u}_j, H_0) d\mathbf{x} \\ \mathbf{M}_{\mathbf{a}ij}^{00} &= \int p(\mathbf{a}|\mathbf{u}_i, H_0)p(\mathbf{a}|\mathbf{u}_j, H_0) d\mathbf{x} \end{aligned}$$

and

$$\begin{aligned} \mathbf{M}_{\mathbf{a}ij}^{mn} &= \int p(\mathbf{a}|\mathbf{u}_i, H_m)p(\mathbf{a}|\mathbf{u}_j, H_n) \\ &= \frac{1}{\sqrt{\det(2\pi(\Sigma_{\mathbf{a}_{im}} + \Sigma_{\mathbf{a}_{jn}}))}} e^{-\frac{1}{2}(\mu_{\mathbf{a}_{im}} - \mu_{\mathbf{a}_{jn}})^T (\Sigma_{\mathbf{a}_{im}} + \Sigma_{\mathbf{a}_{jn}})^{-1} (\mu_{\mathbf{a}_{im}} - \mu_{\mathbf{a}_{jn}})}, \end{aligned} \tag{4.23}$$

where

$$\Sigma_{\mathbf{a}_{im}} = \begin{bmatrix} w_1^2 h_1^2 \sigma^2_{1|H_m,\mathbf{u}_i} + \sigma_n^2 & 0 & \dots & 0 \\ 0 & w_1^2 h_1^2 \sigma^2_{2|H_m,\mathbf{u}_i} + \sigma_n^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_1^2 h_1^2 \sigma^2_{N|H_m,\mathbf{u}_i} + \sigma_n^2 \end{bmatrix}$$

$$\mu_{\mathbf{a}_{im}} = \begin{bmatrix} w_1 h_1 \mu_{1|H_m,\mathbf{u}_i} & 0 & \dots & 0 \\ 0 & w_1 h_1 \mu_{2|H_m,\mathbf{u}_i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_1 h_1 \mu_{N|H_m,\mathbf{u}_i}. \end{bmatrix}$$

Let $w_i \geq 0$ for $i = 0, 1, 2, ..., N$. Because of the orthogonal channel, we can rewrite the equation (4.23) to:

$$\begin{aligned} \mathbf{M}_{\mathbf{a}_{ij}}^{mn} &= (\textstyle\prod_k 2\pi(w_k^2 h_k^2 \sigma^2_{k|H_m,\mathbf{u}_i} + \sigma_n^2 + w_k^2 h_k^2 \sigma^2_{k|H_n,\mathbf{u}_j} + \sigma_n^2))^{-\frac{1}{2}} \\ &\times \exp(-\tfrac{1}{2} \textstyle\sum_k \frac{w_k^2 h_k^2 (\mu_{k|H_m,\mathbf{u}_i} - \mu_{k|H_n,\mathbf{u}_j})^2}{w_k^2 h_k^2 \sigma^2_{k|H_m,\mathbf{u}_i} + \sigma_n^2 + w_k^2 h_k^2 \sigma^2_{k|H_n,\mathbf{u}_j} + \sigma_n^2}) \\ &= (\textstyle\prod_k 2\pi(w_k^2 h_k^2 \sigma^2_{k|H_m,\mathbf{u}_i} + w_k^2 h_k^2 \sigma^2_{k|H_n,\mathbf{u}_j} + 2\sigma_n^2))^{-\frac{1}{2}} \\ &\times \exp(-\tfrac{1}{2} \textstyle\sum_k \frac{w_i^2 h_i^2 (\mu_{k|H_m,\mathbf{u}_i} - \mu_{k|H_n,\mathbf{u}_j})^2}{w_k^2 h_k^2 (\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j}) + 2\sigma_n^2}). \end{aligned} \quad (4.24)$$

For simplicity, we consider 2 receiving antennas and 2 secondary users.

4.2.4 2x2 Case

Considering two special cases, w_i are all equal to zero and w_i are all very big.

When w_i are all equal to zero, the equation (4.24) becomes to

$$\begin{aligned} \mathbf{M}_{\mathbf{a}_{ij}}^{mn} &= (\textstyle\prod_k^2 2\pi(w_k^2 h_k^2 \sigma^2_{k|H_m,\mathbf{u}_i} + w_k^2 h_k^2 \sigma^2_{k|H_n,\mathbf{u}_j} + 2\sigma_n^2))^{-\frac{1}{2}} \exp(-\tfrac{1}{2} \textstyle\sum_k \frac{w_k^2 h_k^2 (\mu_{k|H_m,\mathbf{u}_i} - \mu_{k|H_n,\mathbf{u}_j})^2}{w_k^2 h_k^2 (\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j}) + 2\sigma_n^2}) \\ &= (\textstyle\prod_k^2 \pi(2\sigma_n^2))^{-\frac{1}{2}}. \end{aligned} \quad (4.25)$$

Every $\mathbf{M}_{\mathbf{a}_{ij}}^{mn}$ is a constant when w_i are all equal to zero. Therefore, the L2 distance becomes to

$$
\begin{aligned}
D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0)) &= \sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M_a}_{ij}^{11} \\
&-2\sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M_a}_{ij}^{10} \\
&+\sum_i \sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M_a}_{ij}^{00} \\
&= \sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)(\prod_k 2\pi(2\sigma_n^2))^{-\frac{1}{2}} \\
&-2\sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)(\prod_k 2\pi(2\sigma_n^2))^{-\frac{1}{2}} \\
&+\sum_i \sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)(\prod_k 2\pi(2\sigma_n^2))^{-\frac{1}{2}}.
\end{aligned}
\tag{4.26}
$$

And we know

$$
\begin{aligned}
\sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1) = 1 \\
\sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0) = 1 \\
\sum_i \sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0) = 1.
\end{aligned}
\tag{4.27}
$$

From equation (4.26) and equation (4.27), the $D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0))$ are equal to 0 when w_i are all equal to zero.

When w_1 and w_2 are smaller than σ_n^2, we can rewrite the equation (4.24) as

$$
\begin{aligned}
\mathbf{M_a}_{ij}^{mn} &= (\prod_k^2 2\pi(w_k^2h_k^2\sigma_{k|H_m,\mathbf{u}_i}^2 + w_k^2h_k^2\sigma_{k|H_n,\mathbf{u}_j}^2 + 2\sigma_n^2))^{-\frac{1}{2}} \exp(-\frac{1}{2}\sum_k \frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{w_k^2h_k^2(\sigma_{k|H_m,\mathbf{u}_i}^2+\sigma_{k|H_n,\mathbf{u}_j}^2)+2\sigma_n^2}) \\
&= (\prod_k^2 2\pi(2\sigma_n^2))^{-\frac{1}{2}} \exp(-\frac{1}{2}\sum_k \frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2}).
\end{aligned}
\tag{4.28}
$$

Therefore,

$$
\begin{aligned}
&\frac{dD_{L2}(p(\mathbf{a}|H_1),p(\mathbf{a}|H_0))}{dw_l}\\
&=\sum_i\sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\frac{d\mathbf{M_a}_{ij}^{11}}{dw_l}\\
&-2\sum_i\sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\frac{d\mathbf{M_a}_{ij}^{10}}{dw_l}\\
&+\sum_i\sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\frac{d\mathbf{M_a}_{ij}^{00}}{dw_l}\\
&=\sum_i\sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)(\prod_k^2 4\pi\sigma_n^2)^{-\frac{1}{2}}e^{(-\frac{1}{2}\sum_k\frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})}(-w_l\frac{h_l^2(\mu_{l|H_m,\mathbf{u}_i}-\mu_{l|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})\\
&-2\sum_i\sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)(\prod_k^2 4\pi\sigma_n^2)^{-\frac{1}{2}}e^{(-\frac{1}{2}\sum_k\frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})}(-w_l\frac{h_l^2(\mu_{l|H_m,\mathbf{u}_i}-\mu_{l|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})\\
&+\sum_i\sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)(\prod_k^2 4\pi\sigma_n^2)^{-\frac{1}{2}}e^{(-\frac{1}{2}\sum_k\frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})}(-w_l\frac{h_l^2(\mu_{l|H_m,\mathbf{u}_i}-\mu_{l|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})\\
&=-w_l(\sum_i\sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)(\prod_k^2 4\pi\sigma_n^2)^{-\frac{1}{2}}e^{(-\frac{1}{2}\sum_k\frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})}\frac{h_l^2(\mu_{l|H_m,\mathbf{u}_i}-\mu_{l|H_n,\mathbf{u}_j})^2}{2\sigma_n^2}\\
&-2\sum_i\sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)(\prod_k^2 4\pi\sigma_n^2)^{-\frac{1}{2}}e^{(-\frac{1}{2}\sum_k\frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})}\frac{h_l^2(\mu_{l|H_m,\mathbf{u}_i}-\mu_{l|H_n,\mathbf{u}_j})^2}{2\sigma_n^2}\\
&+\sum_i\sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)(\prod_k^2 4\pi\sigma_n^2)^{-\frac{1}{2}}e^{(-\frac{1}{2}\sum_k\frac{w_k^2h_k^2(\mu_{k|H_m,\mathbf{u}_i}-\mu_{k|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})}\frac{h_l^2(\mu_{l|H_m,\mathbf{u}_i}-\mu_{l|H_n,\mathbf{u}_j})^2}{2\sigma_n^2})\\
&=-w_lC.
\end{aligned}
\tag{4.29}
$$

In the equation (4.29), it's easy to know $C \le 0$. Because there are many terms becomes to 0 in $\frac{d\mathbf{M_a}11}{dw_l}$ and $\frac{\mathbf{M_a}^{00}}{dw_l}$ including diagonal terms. But most terms in $\frac{\mathbf{M_a}^{10}}{dw_l}$ are not zero. But when all $w_i \to 0$, $D_{L2}(p(\mathbf{a}|H_1),p(\mathbf{a}|H_0))$ must be an increasing function. If it's not an increasing function, $D_{L2}(p(\mathbf{a}|H_1),p(\mathbf{a}|H_0))$ will be negative and it's impossible. Therefore, $D_{L2}(p(\mathbf{a}|H_1),p(\mathbf{a}|H_0))$ is an increasing function when all w_i approach to zero.

Now consider $w_k^2h_k^2\sigma_k^2 \gg \sigma_n^2$. The table (4.1) shows the structure of $\mathbf{M_a}_{ij}^{11}$, $\mathbf{M_a}_{ij}^{10}$, and $\mathbf{M_a}_{ij}^{00}$

$I_1 I_2$	11	10	01	00
11	o	o	o	o
10	o	x	o	x
01	o	o	x	x
00	o	x	x	x

Table 4.1: Table of I_1 and I_2

When $w_k^2 h_k^2 \sigma_k^2 \gg \sigma_n^2$, we can write the following equation.

$$\begin{aligned}
\mathbf{M}_{\mathbf{a}_{ij}}^{mn} &\approx (\prod_k 2\pi(w_k^2 h_k^2 \sigma^2_{k|H_m,\mathbf{u}_i} + w_k^2 h_k^2 \sigma^2_{k|H_n,\mathbf{u}_j}))^{-\frac{1}{2}} \exp(-\frac{1}{2}\sum_k \frac{w_k^2 h_k^2 (\mu_{k|H_m,\mathbf{u}_i} - \mu_{k|H_n,\mathbf{u}_j})^2}{w_k^2 h_k^2 (\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j})}) \\
&= (\prod_k w_k^{-1})(\prod_k 2\pi h_k^2(\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j})))^{-\frac{1}{2}} \exp(-\frac{1}{2}\sum_k \frac{(\mu_{k|H_m,\mathbf{u}_i} - \mu_{k|H_n,\mathbf{u}_j})^2}{\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j}}) \\
&= (\prod_k w_k^{-1}) \mathbf{M}_{\mathbf{a}_{ij}}'^{mn},
\end{aligned} \tag{4.30}$$

where $\mathbf{M}_{\mathbf{a}_{ij}}'^{mn} = (\prod_k 2\pi h_k^2(\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j})))^{-\frac{1}{2}} \exp(-\frac{1}{2}\sum_k \frac{(\mu_{k|H_m,\mathbf{u}_i} - \mu_{k|H_n,\mathbf{u}_j})^2}{\sigma^2_{k|H_m,\mathbf{u}_i} + \sigma^2_{k|H_n,\mathbf{u}_j}})$

In table (4.1), $I_m = 1$ represents $u_m \neq 0$ and $I_m = 0$ represents $u_m = 0$. ”o” means we can use equation (4.30) without any problem and ”x” means $\sigma^2_{k|H_m,\mathbf{u}_i}$ and $\sigma^2_{k|H_n,\mathbf{u}_j}$ are both equal to zero. When $\sigma^2_{k|H_m,\mathbf{u}_i}$ and $\sigma^2_{k|H_n,\mathbf{u}_j}$ are both equal to zero, we can't eliminate σ_n^2 in equation (4.30). Now consider an example, $\mathbf{M}_{\mathbf{a}_{44}}^{11}$. The value of $\mathbf{M}_{\mathbf{a}_{44}}^{11}$ is

$$\mathbf{M}_{\mathbf{a}_{44}}^{11} = (\prod_k^2 2\pi w_k^2 h_k^2(\sigma^2_{k|H_1,\mathbf{u}_4} + \sigma^2_{k|H_1,\mathbf{u}_4}) + 2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\sum_k^2 \frac{w_k^2 h_k^2 (\mu_{k|H_1,\mathbf{u}_4} - \mu_{k|H_1,\mathbf{u}_4})^2}{w_k^2 h_k^2 (\sigma^2_{k|H_1,\mathbf{u}_4} + \sigma^2_{k|H_1,\mathbf{u}_4}) + 2\sigma_n^2}). \tag{4.31}$$

But from the table (4.1), we know that $\sigma^2_{k|H_1,\mathbf{u}_4}$ and $\mu_{k|H_1,\mathbf{u}_4}$ are all equal to zero. Then the equation (4.31) becomes to

$$\mathbf{M}_{\mathbf{a}_{44}}^{11} = (\prod_k^2 2\pi(2\sigma_n^2))^{-\frac{1}{2}}. \tag{4.32}$$

$\mathbf{M}_{\mathbf{a}_{44}}^{11}$ is a constant .

Then we consider $\mathbf{M}_{\mathbf{a}_{24}}^{11}$. $\sigma^2_{2|H_1,\mathbf{u}_2}$,$\sigma^2_{1|H_1,\mathbf{u}_4}$, and $\sigma^2_{2|H_1,\mathbf{u}_4}$ are equal to zero but $\sigma^2_{1|H_1,\mathbf{u}_2}$ is not equal to zero. The value of $\mathbf{M}_{\mathbf{a}_{24}}^{11}$ is

$$\begin{aligned}
\mathbf{M}_{\mathbf{a}_{24}}^{11} &= (4\pi^2(w_1^2 h_1^2 \sigma^2_{1|H_1,\mathbf{u}_2} + 2\sigma_n^2)2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{w_1^2 h_1^2 \mu^2_{1|H_1,\mathbf{u}_2}}{w1^2 h_1^2 \sigma^2_{1|H_1,\mathbf{u}_2} + 2\sigma_n^2}) \\
&= (4\pi^2(w_1^2 h_1^2 \sigma^2_{1|H_1,\mathbf{u}_2})2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{w_1^2 h_1^2 \mu^2_{1|H_1,\mathbf{u}_2}}{w_i^2 h_1^2 \sigma^2_{1|H_1,\mathbf{u}_2}}) \\
&= (4\pi^2(w_1^2 h_1^2 \sigma^2_{1|H_1,\mathbf{u}_2})2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{\mu^2_{1|1,\mathbf{u}_2}}{\sigma^2_{1|1,\mathbf{u}_2}}) \\
&= w_1^{-1}(4\pi^2 h_1^2 \sigma^2_{1|H_1,\mathbf{u}_2} 2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{\mu^2_{1|H_1,\mathbf{u}_2}}{\sigma^2_{1|H_1,\mathbf{u}_2}}) \\
&= w_1^{-1}\mathbf{M}_{\mathbf{a}_{24}}'^{11}.
\end{aligned} \tag{4.33}$$

Then we consider $\mathbf{M}_{\mathbf{a}_{11}}^{34}$. $\sigma^2_{1|H_1,\mathbf{u}_3}$,$\sigma^2_{1|H_1,\mathbf{u}_4}$, and $\sigma^2_{2|H_1,\mathbf{u}_4}$ are equal to zero but $\sigma^2_{2|H_1,\mathbf{u}_3}$

is not equal to zero. The value of ${\mathbf{M}_\mathbf{a}}_{24}^{11}$ is

$$\begin{aligned} {\mathbf{M}_\mathbf{a}}_{34}^{11} &= (4\pi^2(w_2^2h_2^2\sigma^2_{2|H_1,\mathbf{u}_3} + 2\sigma_n^2)2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{w_2^2h_2^2\mu^3_{2|H_1,\mathbf{u}_2}}{w1^2h_1^2\sigma^2_{2|H_1,\mathbf{u}_3}+2\sigma_n^2}) \\ &= (4\pi^2(w_2^2h_2^2\sigma^2_{2|H_1,\mathbf{u}_3})2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{w_2^2h_2^2\mu^2_{2|H_1,\mathbf{u}_3}}{w_i^2h_1^2\sigma^2_{2|H_1,\mathbf{u}_3}}) \\ &= (4\pi^2(w_2^2h_2^2\sigma^2_{2|H_1,\mathbf{u}_3})2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{\mu^2_{2|H_1,\mathbf{u}_3}}{\sigma^2_{2|H_1,\mathbf{u}_3}}) \\ &= w_2^{-1}(4\pi^2h_2^2\sigma^2_{2|H_1,\mathbf{u}_3}2\sigma_n^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{\mu^2_{2|H_1,\mathbf{u}_3}}{\sigma^2_{2|H_1,\mathbf{u}_3}}) \\ &= w_2^{-1}{\mathbf{M}_\mathbf{a}}_{34}^{\prime 11}. \end{aligned} \tag{4.34}$$

From the equation (4.30), (4.32), (4.33), and (4.34), we can separate all ${\mathbf{M}_\mathbf{a}}_{ij}^{mn}$ into four sets. S_1 means the 1st secondary user and the 2nd secondary user both transmit signal to the fusion center. S_2 means only the 1st secondary user transmits signal to the fusion center but the 2nd secondary user keeps silent. S_3 means only the 2st secondary user transmits signal to the fusion center but the 1nd secondary user keeps silent. S_4 means both the 1st secondary user and the 2nd secondary user keep silent.

When all $w_k \to \infty$, from the above equation, all ${\mathbf{M}_\mathbf{a}}_{ij}^{mn}$ in S_1, S_2, and S_3 are equal to zero. For the 2x2 case, the equation (4.30) becomes to

$$\begin{aligned} & D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0)) \\ &= p(\mathbf{u}_4|H_1)p(\mathbf{u}_4|H_1){\mathbf{M}_\mathbf{a}}_{44}^{11} - 2p(\mathbf{u}_4|H_1)p(\mathbf{u}_4|H_0){\mathbf{M}_\mathbf{a}}_{44}^{10} + p(\mathbf{u}_4|H_0)p(\mathbf{u}_4|H_0){\mathbf{M}_\mathbf{a}}_{44}^{00} \\ &= (\textstyle\prod_k^2 2\pi(2\sigma_n^2))^{-\frac{1}{2}}(p_4(\mathbf{u}|H_1) - p_4(\mathbf{u}|H_0))^2. \end{aligned} \tag{4.35}$$

From the above equation, when all $w_i \to \infty$, $D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0))$ will be a constant.

But when $w_k^2 h_k^2 \sigma^2_{k|H_m,\mathbf{u}_i} \gg \sigma_n^2$, the equation of L2 distance becomes to

$$\begin{aligned}
&D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0))\\
&= \sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M}_{\mathbf{a}ij}^{11}\\
&-2\sum_i \sum_j p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M}_{\mathbf{a}ij}^{10}\\
&+\sum_i \sum_j p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M}_{\mathbf{a}ij}^{00}\\
&\simeq w_1^{-1}w_2^{-1}(\textstyle\sum_{ij\in S_1} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11}\\
&-2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10}\\
&+p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00})\\
&+w_1^{-1}(\textstyle\sum_{ij\in S_2} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11}\\
&-2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10}\\
&+p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00})\\
&+w_2^{-1}(\textstyle\sum_{ij\in S_3} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11}\\
&-2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10}\\
&+p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00})\\
&+(\textstyle\sum_{ij\in S_4} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11}\\
&-2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10}\\
&+p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00}).
\end{aligned} \tag{4.36}$$

Let

$$\begin{aligned}
C_1 &= \textstyle\sum_{ij\in S_1} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11} - 2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10} + p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00}\\
C_2 &= \textstyle\sum_{ij\in S_2} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11} - 2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10} + p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00}\\
C_3 &= \textstyle\sum_{ij\in S_3} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11} - 2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10} + p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00}\\
C_4 &= \textstyle\sum_{ij\in S_4} p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_1)\mathbf{M'}_{\mathbf{a}ij}^{11} - 2p(\mathbf{u}_i|H_1)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{10} + p(\mathbf{u}_i|H_0)p(\mathbf{u}_j|H_0)\mathbf{M'}_{\mathbf{a}ij}^{00}.
\end{aligned} \tag{4.37}$$

The equation (4.36) becomes to

$$D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0)) \simeq w_1^{-1}w_2^{-1}C_1 + w_1^{-1}C_2 + w_2^{-1}C_3 + C_4. \tag{4.38}$$

From the equation (4.38), we can have the following equations.

$$\frac{dD_{L2}}{dw_1} = -w_1^{-2}w_2^{-1}C_1 - w_1^{-2}C_2 = -w_1^{-2}(w_2^{-1}C_1 + C_2) \tag{4.39}$$

and

$$\frac{dD_{L2}^2}{d^2w_1} = 2w_1^{-3}w_2^{-1}C_1 + 2w_1^{-3}C_2 = 2w_1^{-3}(w_2^{-1}C_1 + C_2). \tag{4.40}$$

From the equation (4.37), we can know the value of C_i is composed of many integrations of two Gaussian distributions. When two Gaussian distributions are the same, the value of the integration will be maximum. The diagonal terms in M_{aij}^{11} and M_{aij}^{00} will dominate the value of C_1, C_2, C_3, and C_4 and they are all larger than zero. Therefore, we can know that C_1, C_2, C_3, and C_4 are greater than zero. If $C_1 \geq 0$, $C_2 \geq 0$, and $C_3 \geq 0$, the equation (4.38) is a decreasing function and it is a convex problem.

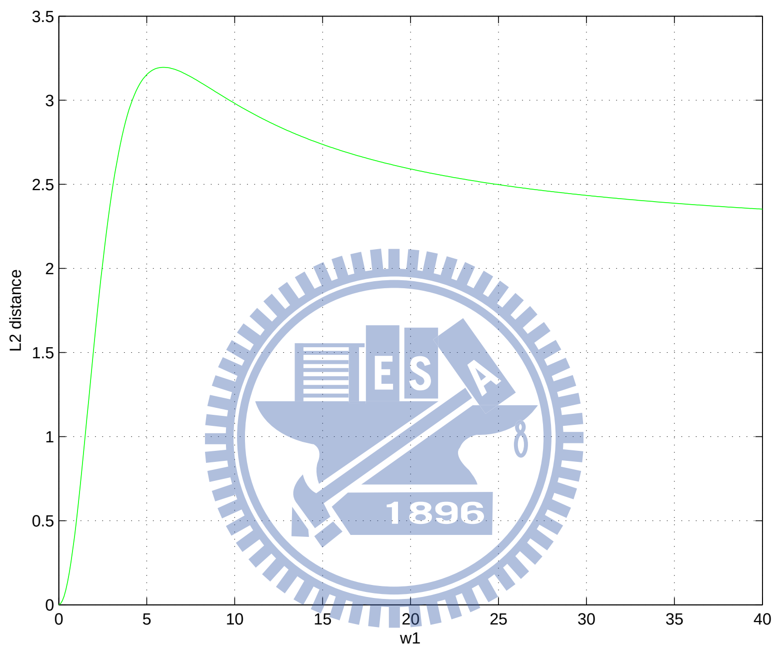


Figure 4.5: L2 distance optimal diagram

From the above discussion, we can know:

- When all w_k approach to 0, $D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0))$ is a increasing function
- When all $w_k^2 h_k^2 \sigma_{k|H_m,\mathbf{u}_i}^2 \gg \sigma_n^2$, $D_{L2}(p(\mathbf{a}|H_1), p(\mathbf{a}|H_0))$ is a decreasing function

Therefore, the L2 distance may have one optimal solution. Like in Fig. 4.5, when w_1 is small, the value of L2 distance approaches 0. But when $w_1 \rightarrow \infty$, the value of L2 distance is close to a constant. In Fig. 4.5, it only has one optimal solution.

4.3 Optimal Linear Combination

4.3.1 System Model

The system model of the optimal linear combination is like in the chapter 3. But here we still use the orthogonal channel and approximate the convolution of the truncated Gaussian distribution and the Gaussian distribution into the Gaussian distribution. Fig. 4.6 shows the system model.

In the optimal linear combination, we combine the signal received by every antenna with weighting factors, like in Fig. 4.6. The signal which we want to detect becomes to

$$\mathbf{a} = \sqrt{g}\mathbf{H}\mathbf{u} + \mathbf{n}, \tag{4.41}$$

where $\mathbf{a} = [a_1, a_2, ..., a_M]^T$ and $H = diag([h_1, h_2, ..., h_N])$. $\sqrt{g}$ is the power control and it is a constant. The power summation of all secondary users can not exceed P_c. The value of g is

$$g \sum_{i=1}^{N} h_{pi}^2 P_{si} = P_c. \tag{4.42}$$

The value of P_c can be calculated by the target SINR.

From the above equation, we can calculate the value of g. But like in chapter 3, it's not easy to optimize the detection probability. Here we still use L2 distance as the performance matric.

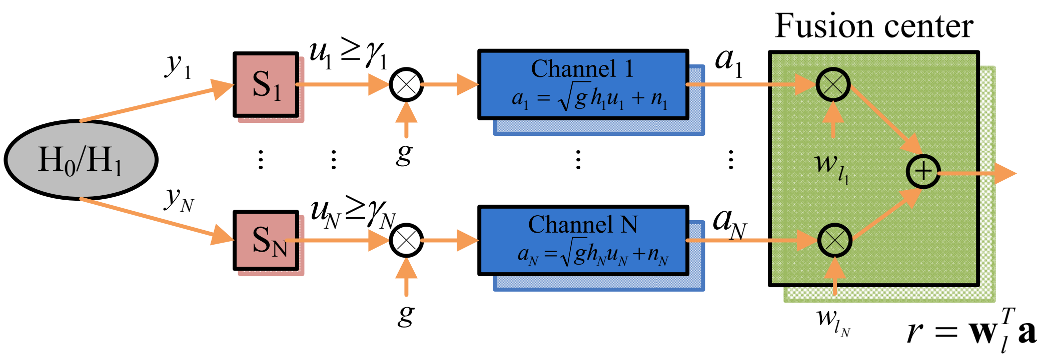


Figure 4.6: System model of censoring scheme of optimal linear combination

4.3.2 GMM L2 Distance

The signal of every antenna received is the GMM distribution. First of all, consider the signal received by ith antenna, a_i. The distribution of the signal received by a_i is

$$\begin{aligned} p(a_i|H_j) &= \sum_{i=1}^{2} p(a_i|u_i, H_j)p(u_i|H_j) \\ &= P_{send_{i,j}} p(a_i|u_i \neq 0, H_j) + (1 - P_{send_{i,j}})p(a_i|u_i = 0, H_j), \end{aligned} \tag{4.43}$$

where $j = 0, 1$ and $i = 1, 2, ..., N$. $p(a_i|u_i \neq 0, H_j)$ and $p(a_i|u_i = 0, H_j)$ are Gaussian distributions. Therefore, $p(a_i|H_j)$ is also GMM.

The equation of L2 distance in GMM is

$$\int (\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}} - \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}})^2 d\mathbf{x}, \tag{4.44}$$

where $\mathbf{w}_l = [w_{l_1}, w_{l_2}, ..., w_{l_N}]^T$ and $\mathbf{P}_{\mathbf{a}_{H_j}} = [p(a_1|H_j), p(a_2|H_j), ..., p(a_N|H_j)]^T$.

The equation (4.44) becomes to

$$\begin{aligned} D_{L2}(\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}}, \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}}) &= \int (\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}} - \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}})^2 dx \\ &= \int [(\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}})^2 - 2\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_1}} \mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}} + (\mathbf{w}_l^T \mathbf{P}_{\mathbf{a}_{H_0}})^2] d\mathbf{x} \\ &= \sum_i \sum_j w_{l_i} w_{l_j} \int p(a_i|H_1)p(a_j|H_1) d\mathbf{x} \\ &\quad -2 \sum_i \sum_j w_{l_i} w_{l_j} \int p(a_i|H_1)p(a_j|H_0) d\mathbf{x} \\ &\quad + \sum_i \sum_j w_{l_i} w_{l_j} \int p(a_i|H_0)p(a_j|H_0) d\mathbf{x}. \end{aligned} \tag{4.45}$$

The integration of two Gaussian mixture models is

$$\begin{aligned} C &= \int (\textstyle\sum_i \alpha_i f_i)(\sum_j \beta_j g_j) d\mathbf{x} \\ &= \int \textstyle\sum_i \sum_j \alpha_i f_i \beta_j g_j d\mathbf{x} \\ &= \textstyle\sum_i \sum_j \int \alpha_i f_i \beta_j g_j d\mathbf{x}, \end{aligned} \tag{4.46}$$

where $\sum_i \alpha_i f_i$ and $\sum_j \beta_j g_j$ are Gaussian mixture models.

From the equation (4.46), the value of $\int p(a_i|H_m)p(a_j|H_n)d\mathbf{x}$ is

$$\begin{aligned} \int p(a_i|H_m)p(a_j|H_n) &= \int (\textstyle\sum_{i=1}^{2} p(a_i|u_i, H_m)p(u_i|H_m))(\sum_{j=1}^{2} p(a_j|u_j, H_n)p(u_j|H_n)) d\mathbf{x} \\ &= \int \textstyle\sum_{i=1}^{2} \sum_{j=1}^{2} p(a_i|u_i, H_m)p(u_i|H_m)p(a_j|u_j, H_n)p(u_j|H_n) d\mathbf{x} \\ &= \textstyle\sum_{i=1}^{2} \sum_{j=1}^{2} \int p(a_i|u_i, H_m)p(u_i|H_m)p(a_j|u_j, H_n)p(u_j|H_n) d\mathbf{x} \\ &= \textstyle\sum_{i=1}^{2} \sum_{j=1}^{2} p(u_i|H_m)p(u_j|H_n) \int p(a_i|u_i, H_m)p(a_j|u_j, H_n) d\mathbf{x}. \end{aligned} \tag{4.47}$$

From the equation (4.47), we can have the value of $\int p(a_i|H_m)p(a_j|H_n)$.

Like in chapter 3, let

$$\begin{aligned} \mathbf{M_{GMM}}_{ij}^{11} &= \int p(a_i|H_1)p(a_j|H_1)d\mathbf{x} \\ \mathbf{M_{GMM}}_{ij}^{10} &= \int p(a_i|H_1)p(a_j|H_0)d\mathbf{x} \\ \mathbf{M_{GMM}}_{ij}^{00} &= \int p(a_i|H_0)p(a_j|H_0)d\mathbf{x}. \end{aligned} \tag{4.48}$$

Therefore, the equation (4.45) becomes to

$$\begin{aligned} D_{L2}(\mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_1}}, \mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_0}}) &= \sum_i \sum_j w_{l_i} w_{l_j} \mathbf{M_{GMM}}_{ij}^{11} - 2\sum_i \sum_j w_{l_i} w_{l_j} \mathbf{M_{GMM}}_{ij}^{10} \\ &+ \sum_i \sum_j w_{l_i} w_{l_j} \mathbf{M_{GMM}}_{ij}^{00} \\ &= \mathbf{w}_l^T \mathbf{M_{GMM}}^{11} \mathbf{w}_l - 2\mathbf{w}_l^T \mathbf{M_{GMM}}^{10} \mathbf{w}_l + \mathbf{w}_l^T \mathbf{M_{GMM}}^{00} \mathbf{w}_l \\ &= \mathbf{w}_l^T (\mathbf{M_{GMM}}^{11} - 2\mathbf{M_{GMM}}^{10} + \mathbf{M_{GMM}}^{00}) \mathbf{w}_l. \end{aligned} \tag{4.49}$$

The optimization problem is

$$\begin{aligned} \max_{\mathbf{w}_l} \quad & D_{L2}(\mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_1}}, \mathbf{w}_l^T\mathbf{P}_{\mathbf{a}_{H_0}}) = \mathbf{w}_l^T (\mathbf{M_{GMM}}^{11} - 2\mathbf{M_{GMM}}^{10} + \mathbf{M_{GMM}}^{00}) \mathbf{w}_l \\ s.t. \quad & w_{l_i} \geq 0, i = 1, 2, ..., N \\ & \mathbf{w}_l^T \mathbf{w}_l = 1. \end{aligned} \tag{4.50}$$

This problem can be solved by the active set method.

4.4 Simulation Result

In the following simulation results, they have 2 secondary users and 2 receiving antennas. The SNR of the primary user is 5. For convenience, we set the transmitting probability of all secondary users is 0.3 under H_0 hypothesis. The MIMO channel and sensing noise are all random. Here we still use Neyman-Pearson detection in the fusion center and the false alarm probability is 0.4.

4.4.1 Optimal Power Allocation

Fig. 4.7(a) and Fig. 4.7(b) show the simulated GMM J-divergence and detection probability in the optimal power allocation and the equal power allocation. In Fig. 4.7(a), when the target SINR is high, the value of GMM J-divergence will be low. And in Fig. 4.7(b) shows the detection probability in the optimal power allocation and the equal power allocation. In Fig. 4.7(b), when the target SINR is low, the detection probability will be high. And the detection probability of the optimal power allocation is much better than the equal power allocation. Fig. 4.7(a) and Fig. 4.7(b) can prove the GMM J-divergence can be viewed as a performance matric for optimizing the detection probability.

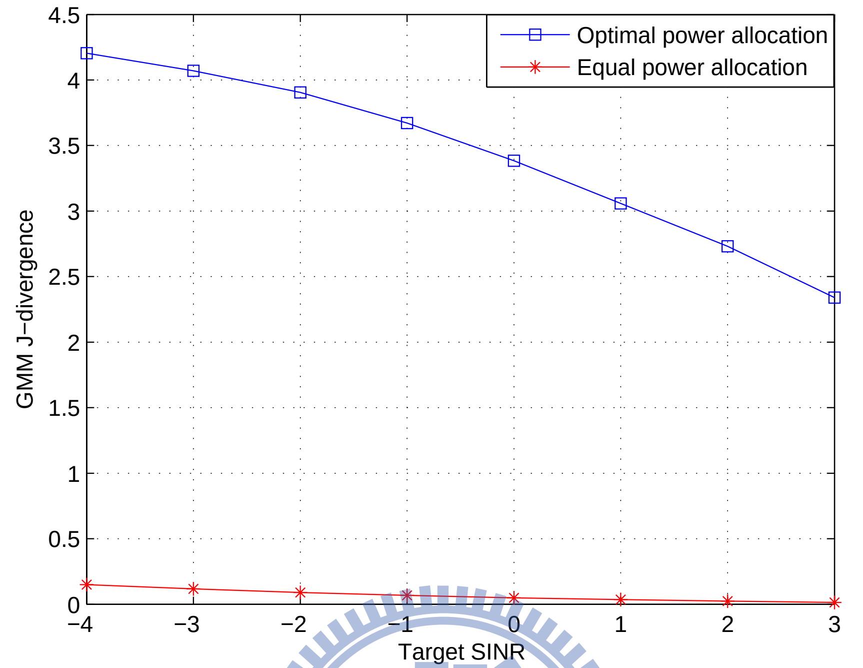
Fig. 4.8(a) and Fig. 4.8(b) show the simulated GMM L2 distance and detection probability in the optimal power allocation and the equal power allocation. In Fig. 4.8(a), when the target SINR is high , the value of GMM L2 distance will be low. And the Fig. 4.8(b) shows the detection probability of GMM L2 distance in optimal power allocation. In Fig. 4.8(b), when the target SINR is low, the detection probability will be high. And the detection probability of the optimal power allocation is much better than the equal power allocation. Fig. 4.8(a) and Fig. 4.8(b) can prove the GMM L2 distance can be viewed as a performance matric for maximizing the detection probability.

4.4.2 Optimal Linear Combination

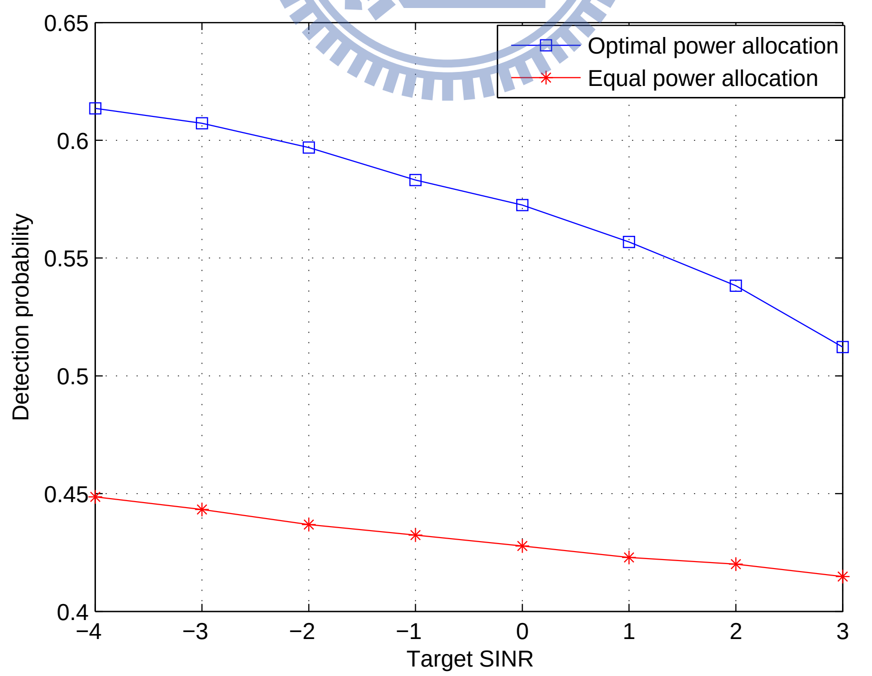
Fig. 4.9(a) shows the optimal linear combination by the use of GMM L2 distance. When target SINR is high , the value of GMM L2 distance will be low. And Fig. 4.9(b) shows the detection probability of GMM in the optimal linear combination and the equal weighting combination. In Fig. 4.9(b), when the target SINR is low, the detection probability will be high. And the detection probability in the optimal linear combination is better than in the equal weighting combination. Comparing two figures can prove the GMM L2 distance can be viewed as a performance matric in the optimal linear combination.

4.4.3 Comparison

Fig. 4.10 shows that the detection probability of the censoring scheme is better than the non-censoring scheme in the optimal power allocation. But the GMM L2 method is a little better than the GMM J-divergence method. In Fig. 4.12(a), the interference of primary users by the use of the censoring method is lower than by the use of non-censoring. From

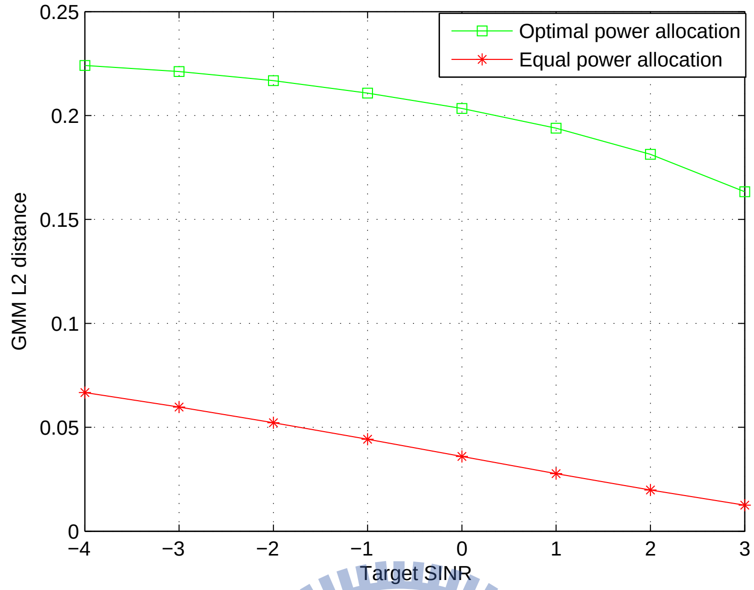


(a) GMM J-divergence of optimal power allocation and equal power allocation

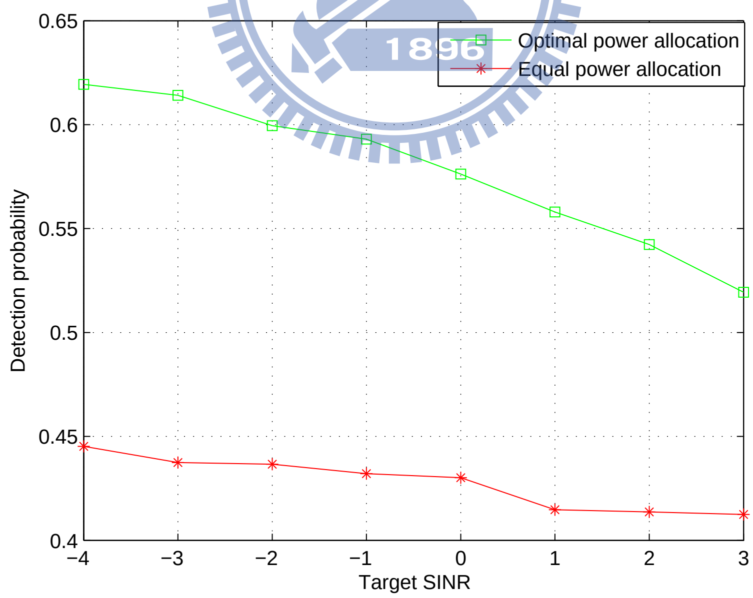


(b) Detection probability of optimal power allocation and equal power allocation

Figure 4.7: The GMM J-divergence and detection probability of optimal power allocation censoring scheme

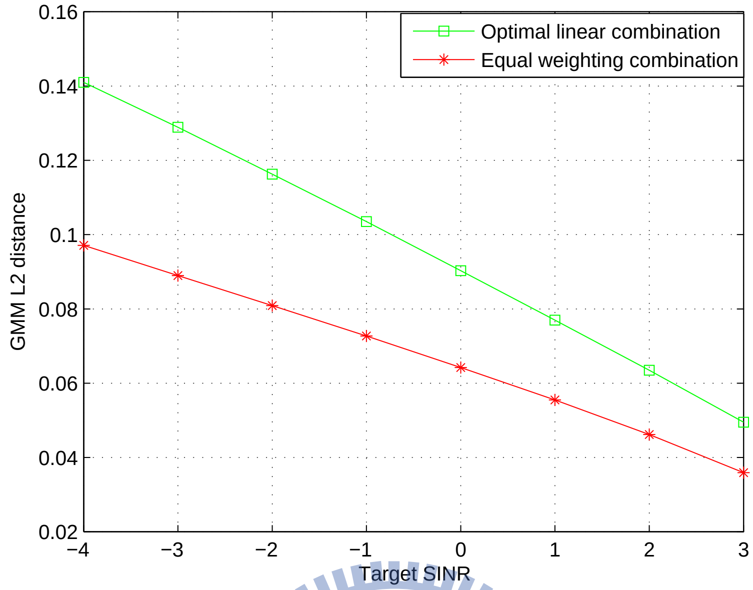


(a) L2 distance of optimal power allocation

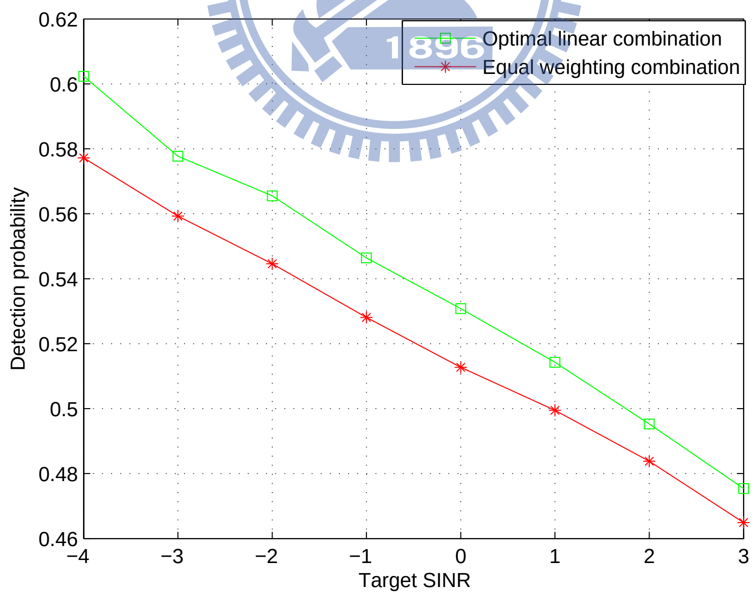


(b) Detection probability of L2 distnance of optimal power allocation

Figure 4.8: The GMM L2 distance and detection probability of optimal power allocation censoring scheme



(a) GMM L2 distance



(b) Detection probability

Figure 4.9: The L2 distance and detection probability of optimal linear combination censoring scheme

Fig. 4.10 and Fig. 4.12(a), we can know that the censoring scheme has lower interference and higher detection probability. In Fig. 4.10, the detection probability of L2 distance is greater than the detection probability of GMM j-divergence. But in Fig. 4.12(a), the interference of GMM L2 distance is little higher than the interference of GMM J-divergence. Therefore, GMM L2 distance has the better detection probability and produces the lower interference in the optimal power allocation. The GMM L2 distance method is better than the GMM J-divergence method.

Fig. 4.11 shows that the detection probability of the censoring scheme is also better than the non-censoring scheme in the optimal linear combination. Like in the optimal power allocation scheme, the secondary users only transmit the informative data to the fusion center. And as we mentioned before, the censoring scheme will keep silent when the signal is not informative enough. Therefore, the received interference of the primary users in the censoring scheme will be little lower than the interference in the non-censoring scheme. Fig. 4.12(b) shows this result.

In [19], it uses the distributed detection in the MIMO channel in the optimal power allocation and it's also a censoring scheme. In this system, every user makes his own decision and transmits this decision to the fusion center. In this paper, the secondary user transmits "1" when this user decides that it's under H_1 and keeps silent when this user decides that it's under H_0. We let the transmitting probability of the proposed scheme is equal to the transmitting probability of the distributed detection and have the same constraint. And in [19], it approximates the GMM into the Gaussian distribution in the distributed detection.

In Fig. 4.13, it shows the comparison between the distributed detection and the our proposed censoring scheme. In this figure, the proposed scheme is better than the distributed scheme. In the distributed detection, it quantizes the observation and lose the information. Therefore, the detection probability of the proposed scheme will be better.

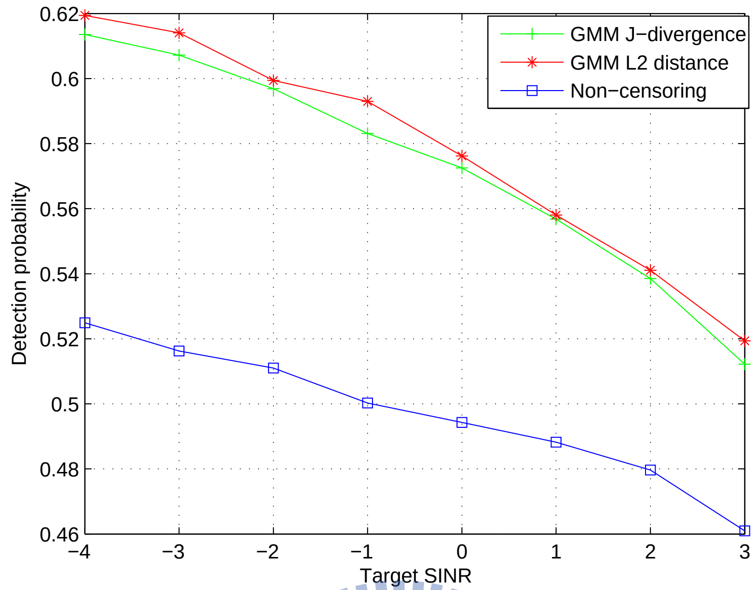


Figure 4.10: Comparison between censoring and non-censoring in the optimal power allocation

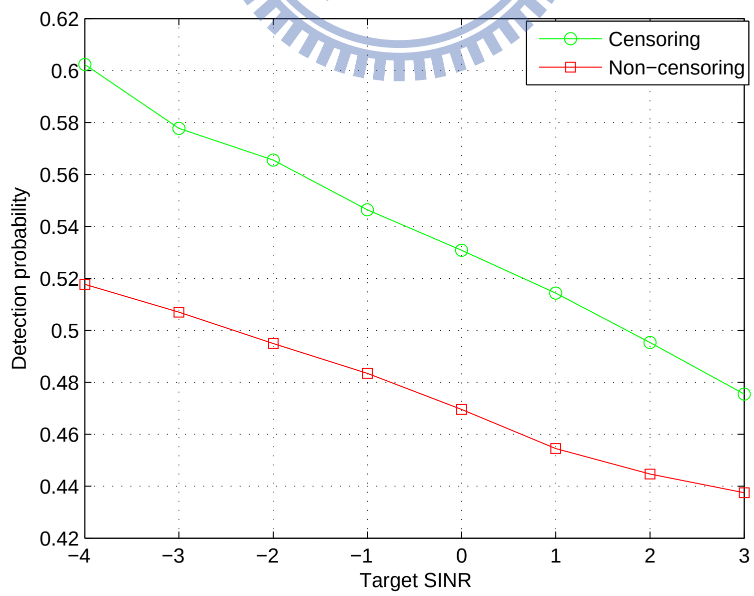
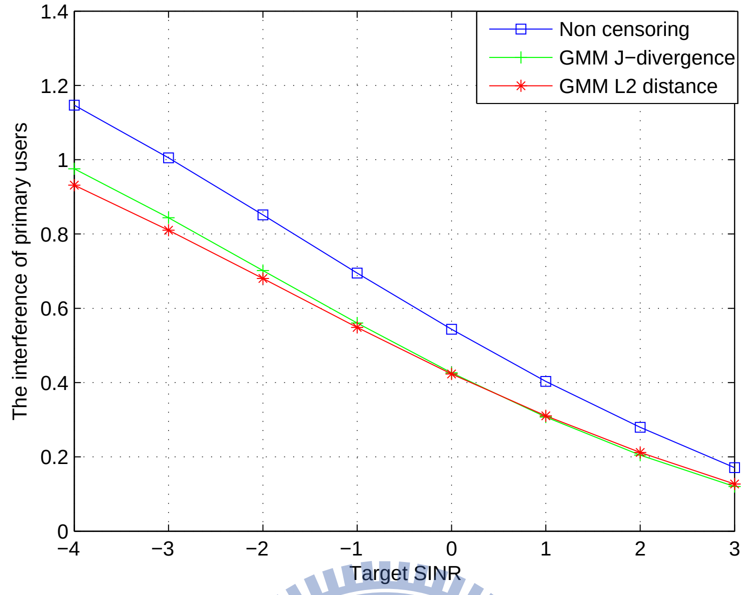
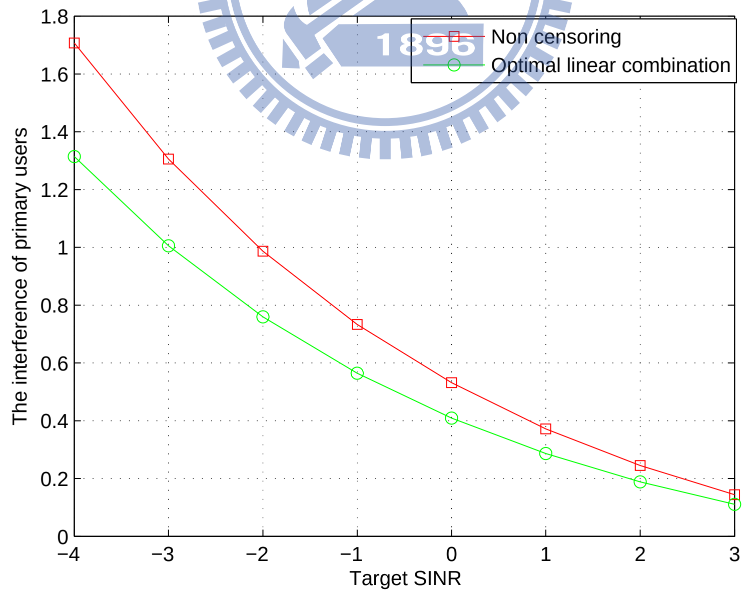


Figure 4.11: Comparison between censoring and non-censoring in the optimal linear combination



(a) The interference of the primary user in optimal power allocation



(b) The interference of the primary user in optimal linear combination

Figure 4.12: Comparison between the power of censoring scheme and the power of non-censoring scheme

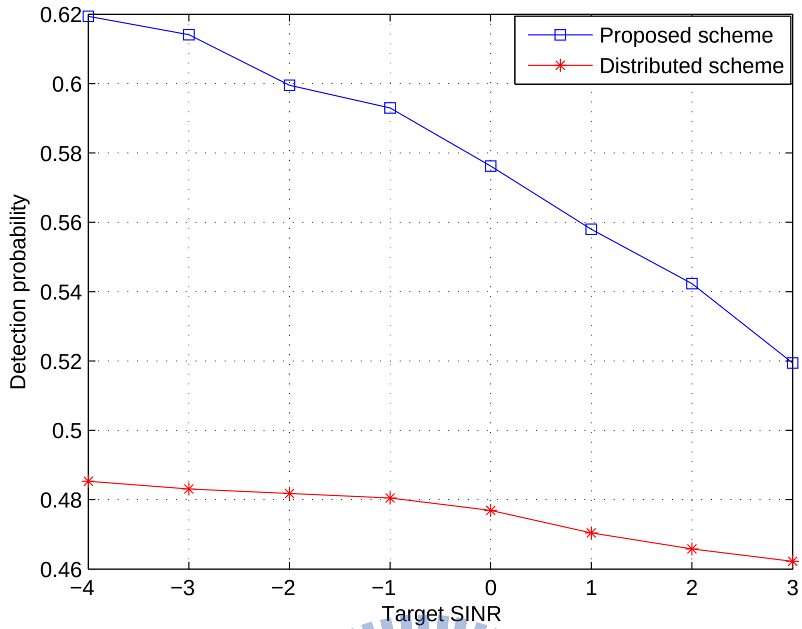


Figure 4.13: Comparison between distributed scheme and proposed scheme

4.5 Summary

In this chapter, we use the censoring scheme to reduce the average interference of the primary users. The censoring scheme is that every secondary user transmits the observation with enough information to the fusion center or keeps silent. The distribution of transmitting signal is the truncated Gaussian. Because the convolution of the truncated Gaussian distribution and a Gaussian distribution doesn't have the closed-form expression, we use the Gaussian distribution to approximate it. We use the GMM J-divergence and the L2 distance as the performance metrics in the optimal linear combination. In the optimal linear combination, we still use the L2 distance. In simulation results of the optimal power allocation and the optimal linear combination, the detection probability of the censoring scheme is better than the detection probability of the non-censoring scheme and the interference of primary users is also lower than non-censoring scheme.

Chapter 5

Conclusion and Future Work

In this thesis, we discuss two different schemes, the optimal power allocation and the optimal linear combination, by the use of the cooperative spectrum sensing in the cognitive radio network. In part 1, we adopt the centralized detection in the fusion center. We focus on maximizing the detection probability and satisfying the target SINR of primary users. Because the LLR doesn't have a closed-form expression of the detection probability, we use distance measures to optimize the detection probability. In this thesis, we use two distance measures, J-divergence and L2 distance, to maximize the detection probability. In simulation results, we can know that the detection probability of the optimal power allocation is better than the equal power allocation and the detection probability of the optimal linear combination is better than the equal weighting combination. In part 2, we propose another method, the censoring method, to reduce the average interference of primary users but has better detection probability than non-censoring scheme.

In this thesis, we assume that the channel and the power of primary users are known. But in practical system, they may not be known by the fusion center or secondary users. We could estimate these parameters and observe the affect of the system performance. And in censoring scheme, we don't discuss the affect of transmitting thresholds, γ_i, of secondary users. Obviously, the thresholds will affect the detection probability in censoring scheme. We can try to find the relation between the thresholds and the detection probability.

Bibliography

[1] F. C. Commission, "Spectrum policy task force report, FCC 02-155," Nov. 2002.

[2] Mitola, J., III, Maguire, G.Q., Jr., "Cognitive radio: Making software radios more personal ," *IEEE Personal Communications*, vol. 6, pp. 13-18, Aug. 1999.

[3] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE J. Selet. Areas Commun.*, vol. 23, pp. 201-220, 2005.

[4] Natasha Devroye, Patrick Mitran and V. Tarokh, "Limits on communication in a cognitive radio channel ," *IEEE Commun. Mag.*, vol.44, pp. 44-49, June 2006.

[5] A. Sahai, N. Hoven and R. Tandra, "Some fundamental limits in cognitive radio," *in Proc. of Allerton Conf.*, Oct. 2004.

[6] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," *in Proc. Asilomar Conf. on Signals, Syst., and Comput.*, 2004, pp. 772-776.

[7] G. Ghurumuruhan and Y. Li, "Agility improvement through cooperative diversity in cognitive radio," *in Proc. IEEE GLOBECOM*, pp. 2505V2509, Nov. 2005.

[8] Z. Quan, S. Cui, and A. H. Sayed, "An optimal strategy for cooperative spectrum sensing in cognitive radio networks," *in Proc. IEEE GLOBECOM*, Nov. 2007.

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

[10] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part I: two user networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2204-2213, 2007.

[11] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part II: multiuser networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2214-2222, 2007.

[12] Q. Zhi, C. Shuguang, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 28-40, 2008.

[13] Q. Zhi, C. Shuguang, and A. H. Sayed, "Optimal linear fusion for distributed spectrum sensing via semidefinite programming," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3629-3632 , 2009.

[14] L. Ying-Chang, Z. Yonghong, E. C. Y. Peh, and H. Anh Tuan, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1326-1337, Apr. 2008.

[15] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *Signal Processing Magazine, IEEE*, vol. 8, pp. 14-36, 1991.

[16] A. Ghasemi and E. S. Sousa, "Interference aggregation in spectrum sensing cognitive wireless networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 41-56, 2008.

[17] Raab, Martin and Schreiner, Olaf and Herbig, Tobias and Gruhn, Rainer and Nöth, Elmar, "Optimal projections between gaussian mixture feature spaces for multilingual speech recognition, " *International Conference on Acoustics*, 2009.

[18] Rajan, K. and Natarajan, B., "A distance based comparison of optimal power allocation to distributed sensors in a wireless sensor network," *IEEE International Conference on Communications*, pp.4391-4395, May. 2008.

[19] Xin Zhang and Vincent, P.H. and Mung Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4124-4140, 2008.

[20] H. Kobayashi and J. B. Thomas, "Distance measures and related criteria," *in Proc. 5th Annu. Allerton Conf. Circuit System Theory*, Oct. 1967, pp. 491-500.

[21] H. Kobayashi, "Distance measures and asymptotic relative efficiency," *IEEE Trans. Inf. Theory*, vol. 16, no. 3, pp. 288-291, May 1970.

[22] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 55-568, Apr. 1996.

[23] Chunhua Sun and Wei Zhang and Letaief, K.B., "Cooperative spectrum sensing for cognitive radios under bandwidth constraints," *Wireless Communications and Networking Conference*, pp. 1-5, 2007.