The Extended Neyman-Pearson Hypotheses Testing Framework and its Application to Spectrum Sensing in Cognitive Wireless Communication

An Jiang, Harry Leib,

Department of Electrical Engineering

McGill University

Montreal, Quebec, Canada,

Email: an.jiang@mail.mcgill.ca, harry.leib@mcgill.ca

Abstract

Neyman Person (NP) testing for two hypotheses is widely used in radar system, spectrum sensing, medical detection and many other applications. Traditional NP testing maximizes the probability of detection under a probability of false alarm constraint. This paper consider an extended form of the NP test that is suitable for spectrum sensing when there might be different type of primary signals.

Index Terms

Neyman Pearson, Hypothesis Testing, Spectrum Sensin, Cognitive Radio

I. INTRODUCTION

The increasing radio spectrum demand for wireless communication is driving the development of new approaches for its usage, such as Cognitive Radio (CR)[1]. The concept behind CR networks is the unlicensed use of radio spectrum while ensuring no interference to the licensed users. [2]. To achieve this goal, a CR system must have "cognitive" abilities to identify opportunities for communication[3]. In CR systems, the spectrum sensing component brings in such capabilities [4].

Spectrum sensing has been a subject for extensive studies in the last years[5]. Common techniques for spectrum sensing are based on energy detection, exploitation of cyclostationarity properties of the signal being sensed, and also could use preamble sequences that are embedded in the signals [6]. In [7], the authors point out that for OFDM signals when the variances of the noise and signal are known, the performance of energy based spectrum sensing is very close to optimal. Cooperative spectrum sensing [8], employing multiple sensors that are geographically scattered, are effective against shadowing. Different sensing algorithms and DSP techniques have been introduced for specific situations, such as in [9] where the authors study algorithms for wide band spectrum sensing focusing on sub-Neyquist sampling techniques. In [10] and [11] the authors propose compressed sensing and sub-Nyquist sampling techniques for wide band spectrum sensing.

In order to detect a free frequency channel, a spectrum sensing scheme solves a binary hypotheses testing problem, where the null hypothesis refers to the event that a user is not using the channel and the alternative hypothesis refers to the event that a user is occupying the channel. If the a-prior probabilities of null and alternative hypotheses are known, then such a problem can be solved using a Bayesian framework [12]. The situation when Baysian framework could be used is considered in [13] and the performance of such scheme is analyzed. Even though the Baysian framework can be used in some situations, in more often cases, the a-prior probabilities are not available. In such situations Neyman Pearson testing is employed [12].When there is a single type of primary user (two hypotheses), NP testing will achieve the largest probability of detecting a vacant channel when no primary user is present under a constraint on the probability that the channel will be declared vacant when in fact a primary user is present. Traditional NP testing has been extensively used for two hypotheses problem. In such cases, the performance is characterized by the Receiver Operating Characteristic (ROC) curve, which represents the relationship between probability of detection and probability of false alarm [12].

In spectrum sensing application, there may be more than one type of primary users. For example in IEEE 802.22 [14] the channel can be occupied by Analog-TV Digital-TV and wireless microphone. In such cases, an extended NP (ENP) test can provide important advantages [15]. A detector based on ENP testing can ensure the largest probability of detection under a separate constraint of false alarm for each primary user type [16].

With an ENP detector however, the achievable false alarm probabilities are limited. In this work we analyze the properties of the ENP test and propose the Modified Extended Neyman Pearson (MENP) test, that circumvents the limitations of the false alarm probabilities. The structure of the remaining of this paper is [TO BE ADDED].

II. Extended Neyman Pearson Test

A. Introduction to Extended Neyman Pearson Test

The theories of hypotheses testing have been a subject of continuous studies, and have found applications in various fields such as radar systems, spectrum sensing for cognitive communication systems, and in medical science. One type of hypotheses testing problem can be abstracted as follows: assume $M+1$ hypotheses H_0, H_1, ..., H_M, inducing $M+1$ Probability Density Functions (PDFs) on the observable Y.

$$\begin{aligned} H_0: &\qquad Y \sim f_0(y) \\ H_1: &\qquad Y \sim f_1(y) \\ &\qquad \ldots\ldots \\ H_M: &\qquad Y \sim f_M(y) \end{aligned} \tag{1}$$

Based on y, a realization of Y, the detector needs to decide whether or not it comes from $f_0(y)$. A framework for solving this problem for $M = 1$ was introduced in [17] and it is commonly known as Neyman Pearson (NP) testing. The theory of NP testing was further developed in [18]. In [19], the theory of NP testing was expanded also for $M > 2$. A comprehensive exposition of such generalized NP testing can be found in [16].

The Extended Neyman Pearson (ENP) Lemma: *Let $f_0(x), f_1(x), ..., f_m(x)$ be real Borel measurable functions defined on finite dimensional Euclidean space $\mathcal{R}$ such that $\int_R |f_i(x)|\mathrm{d}x < \infty (i = 0, 1, ..., M)$. Suppose that for given constants $c_1, ..., c_M$ there exists a class of subsets $\mathcal{S}$, denoted $\mathcal{C}_\mathcal{S}$, such that for every $\mathcal{S} \in \mathcal{C}_\mathcal{S}$ we have*

$$\int_\mathcal{S} f_i(x)\mathrm{d}x = c_i, \qquad i = 1, ..., M \tag{2}$$

Then:

(i) *Among all members of $\mathcal{C}_\mathcal{S}$ there exists one that maximizes*

$$\int_\mathcal{S} f_0(x)\mathrm{d}x.$$

(ii) *A sufficient condition for a member of $\mathcal{C}_\mathcal{S}$ to maximize*

$$\int_\mathcal{S} f_0(x)\mathrm{d}x.$$

is the existence of constants $k_1, ..., k_M$ such that

$$f_0(x) > \sum_{j=1}^{M} k_j f_j(x) \quad \textit{when } x \in \mathcal{S} \tag{3}$$

$$f_0(x) < \sum_{j=1}^{M} k_j f_j(x) \quad \textit{when } x \notin \mathcal{S} \tag{4}$$

(iii) *If a member of $\mathcal{C}_\mathcal{S}$ satisfies* (3) *and* (4) *with $k_1, ..., k_M \geq 0$, then it maximizes*

$$\int_\mathcal{S} f_0(x)\mathrm{d}x \tag{5}$$

among all $\mathcal{S} \in \mathcal{C}_{\mathcal{S}}$ satisfying

$$\int\limits_{\mathcal{S}} f_i(x)\mathrm{d}x \leq c_i, \quad i = 1, ..., M. \tag{6}$$

(iv) *The set M of points in M-dimensional space whose coordinates are*

$$(\int_{\mathcal{S}} f_1(x)\mathrm{d}x, ..., \int_{\mathcal{S}} f_M(x)\mathrm{d}x) \tag{7}$$

for any $\mathcal{S}$ is convex and closed. If $(c_1, ..., c_M)$ is an inner point of M, then a necessary and sufficient condition for a member of $\mathcal{C}_{\mathcal{S}}$ to maximize

$$\int\limits_{\mathcal{S}} f_0(x)\mathrm{d}x.$$

is that there exist M constants $k_1, ..., k_M$ such that (3) (4) *holds a.e..*

The associated probability of detection, P_d and false alarms P_{f_i} for a certain subset $\mathcal{S}$ are defined as [17], $P_d = P(H_0|H_0) = \int_{\mathcal{S}} f_0(x)\mathrm{d}x, P_{f_i} = P(H_0|H_i) = \int_{\mathcal{S}} f_i(x)\mathrm{d}x \;\; i = 1, ..., M$. Define the step function

$$u(x) = \begin{cases} 0 & x < 0 \\ 0.5 & x = 0 \\ 1 & x > 0\,, \end{cases} \tag{8}$$

Then for a subset $\mathcal{S}$ satisfying (3) and (4) we have:

$$\begin{aligned} P_d &= \int_{-\infty}^{\infty} u(f_0(x) - \sum_{j=1}^{M} k_j f_j(x)) f_0(x) \mathrm{d}x\,, \\ P_{f_i} &= \int_{-\infty}^{\infty} u(f_0(x) - \sum_{j=1}^{M} k_j f_j(x)) f_i(x) \mathrm{d}x \quad i = 1, 2, ..., M\,. \end{aligned} \tag{9}$$

The relationship between P_d and P_{f_i} can be represented by Receiver Operating Characteristic (ROC) surface [16].

From to (3) (4), the ENP Decision rule δ is

$$\delta : \sum_{j=1}^{M} k_j \frac{f_j(x)}{f_0(x)} \underset{\bar{H}_0}{\overset{H_0}{\lesseqgtr}} 1 \tag{10}$$

From the **ENP Lemma**, δ achieves the largest P_d under the constraints $P_{f_i} = c_i (i = 1, 2, ..., M)$. When $M = 1$, it achieves the largest P_d under the constraint $P_f \leq c$ [16], which is the well known and commonly used form.

For applications in spectrum sensing, H_0 denotes the hypothesis that the channel is free and H_m $(m = 1, ..., M)$ corresponds to the hypothesis that the channel is occupied by the m-th primary signal. Although we have M hypotheses, we intend to determine if the channel is free or not. Hence the problem is to find a binary test of deciding H_0 versus $\bar{H}_0$ such that P_d is maximized under the constraints $P_{f_m} \leq c_m$ $m = 1, ..., M$. In context of spectrum sensing, $1 - P_{f_m}$ can be interpreted as the protection level of the $m-$th primary signal. The larger is this protection level, the smaller is the probability that when the $m-$th signal is active, the test will not detect it and will declare the channel free. In context of spectrum sensing the solution of the ENP problem maximizes the probability of detecting a free channel under a constraint on the protection level for each primary signal. The protection levels of primary signals can be different and they are guaranteed.

B. Properties of Extended Neyman Pearson Test

We consider now several properties for the ENP test, embodied by three lemmas with proof placed in the appendix A, B and C.

Condition 1 *Let* $f_i(x)$ $\;i = 0, 1, ..., M$ *be the PDF induced by hypothesis* H_i*, and define* $g(x) = f_0(x) - \sum_{j=1}^{M} k_j f_j(x)$ *where* k_i *(*$i = 1, 2, ..., M$*) are real numbers. Let* $\mathcal{D} \in \mathbb{R}$ *be an open set such that* $\int_{\bar{\mathcal{D}}} f_i(x) = 0$ $\;i = 1, 2, ..., M$*. Furthermore, if* x_0 *is a solution for* $g(x) = 0$ $\;(x \in \mathcal{D})$*, there exists an integer* n *such that the* n *order derivative of* $g(x_0)$ *is not equal to zero (*$g^{(n)}(x_0) \neq 0$*).*

Lemma 1 *Under* **Condition 1** *, let* $\mathbf{P}$ *be a point with coordinate* $(P_d, P_{f_1}, ..., P_{f_M})$ *on the ROC surface of the EPN test. If there exists a tangent hyperplane at* $\mathbf{P}$*, then its normal is parallel to the vector* $\mathbf{n} = (-1, k_1, ..., k_M)$*, where* k_i *are the parameters of the ENP test achieving* $\mathbf{P}$.

Lemma 2 *Under* **Condition 1** *, let* $\mathbf{P}$ *be a point on the ROC surface,* $\left.\frac{\partial P_d}{\partial P_{f_i}}\right|_P = k_i$*, where* k_i *are the parameters of ENP test achieving* $\mathbf{P}$.

Lemma 3 *Let* f_0*,* f_1*, ...,* f_M *be PDFs defined on set* $\mathcal{D}$ *and* $f_i(x) \neq 0$ *holds a.e.* $\mathcal{D}$ *(*$i = 0, 1, ..., M$*). Suppose that for given constants* $c_1, ..., c_M \in (0, 1]$ *there exists a class of decision rules* δ*, denote* $\mathcal{C}_\delta$*, such that for every* $\delta \in \mathcal{C}_\delta$*, we have* $P_{f_i} \leq c_i$*. Then:*

If δ^* *is a member of* $\mathcal{C}_\delta$ *that maximum* P_d*, then there exists non-negative constants* $k_1, ..., k_M$ *such that*

$$x \in H_0 \quad \textit{when} \quad f_0(x) > \sum_{i=1}^{M} k_i f_i(x)$$

$$x \notin H_0 \quad \textit{when} \quad f_0(x) < \sum_{i=1}^{M} k_i f_i(x)$$

C. Modified Extended Neyman Pearson Algorithm

Before we proceed, we will define some notations for presentation. Define $\mathbf{c}^T = [c_1, c_2, ..., c_M]$, $\mathbf{a}^T = [a_1, a_2, ..., a_M]$, $\mathbf{k}^T = [k_1, k_2, ..., k_M]$ and $\mathbf{P}_f^T = [P_{f_1}, P_{f_2}, ..., P_{f_M}]$. Let $F(\mathbf{a})$ denote the largest P_d under the constraints $P_{f_i} = a_i \;\; i = 1, ..., M$. $G(\mathbf{c})$ denote the largest P_d under the constraints $P_{f_i} \le c_i \;\; i = 1, ..., M$. Also define the set $\mathcal{A}_{\mathbf{c}} = \{\mathbf{P}_f | 0 \le P_{f_i} \le c_i \;\; i = 1, 2, ..., M\}$. and set $\alpha^+ \triangleq \{\mathbf{P}_f | P_{f_i} = \int_{-\infty}^{\infty} u(f_0(x) - \sum_{j=1}^{M} k_j f_j(x)) f_i(x) \mathrm{d}x, \text{where } k_i \ge 0 \;\; i = 1, ..., M\}$. Let $\mathbf{A}$ and $\mathbf{B}$ denote two vectors. By $\mathbf{A} \le \mathbf{B}$, $\mathbf{A} = \mathbf{B}$ and $\mathbf{A} \ge \mathbf{B}$ we mean that every element of $\mathbf{A}$ is no larger than, equal to and no smaller than its corresponding element of $\mathbf{B}$, respectively.

In practice (e.g. spectrum sensing in CR communication system), the following problem needs to be solved,

$$\begin{aligned} \max \quad & P_d \\ \text{s.t.} \quad & \mathbf{P}_f \le \mathbf{c} \end{aligned} \tag{11}$$

According to ENP Lemma, this problem can be solved by an ENP test only when there are parameters $k_i \ge 0, \;\; i = 1, ..., M$ such that

$$c_i = \int_{-\infty}^{\infty} u(f_0 - \sum_{j=1}^{M} k_j f_j(x)) f_i(x) \mathrm{d}x \;\; i = 1, ..., M \tag{12}$$

The case when the given $c_i \;\; i = 1, 2, ..., M$ do not satisfy (12) was not considered so far.

Next we present the Modified Extended Neyman Pearson Test (MENP) for solving (11).

Modified Extended Neyman Pearson Test

(i) *Assume* $\mathbf{c} \in \alpha^+$*. Then there must be a* $\mathbf{k}^0 = [k_1^0, k_2^0, ..., k_M^0]^T$ *with* $k_i \geq 0 \;\; i = 1, ..., M$ *satisfying*

$$P_{f_i}^0 = \int_{-\infty}^{\infty} u(f_0(x) - \sum_{j=1}^{M} k_j^0 f_j(x)) f_i(x) \mathrm{d}x = c_i \;\; (i = 1, 2, ..., M)\,. \tag{13}$$

and the decision rule δ *solving* (11) *is:*

$$\delta: \quad \frac{f_0(x)}{\sum_{i=1}^{M} k_i^0 f_i(x)} \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} 1 \tag{14}$$

(ii) *Assuming* $\mathbf{c} \notin \alpha^+$*, Let* $\mathcal{C} = \mathcal{A}_{\mathbf{c}} \cap \alpha^+$*, and* $\mathbf{a}^0 = [a_1^0, a_2^0, ..., a_M^0]^T \in \mathcal{C}$ *be such that*

$$\max_{\mathbf{a} \in \mathcal{C}} \quad F(\mathbf{a}) = F(\mathbf{a}^0) \tag{15}$$

Since $\mathbf{a}^0 \in \mathcal{A}_{\mathbf{c}} \cap \alpha^+$*, from* (i) *we have that there exists a vector* $\mathbf{k}^0$ *such that* (14) *maximizes* P_d *under the constraints* $P_{f_i} = a_i^0, \; i = 1, ..., M$*. Here since* $a_i^0 \leq c_i^0, \; i = 1, ..., M$ *this decision rule solves* (11).

In Appendix D, we show MENP Test can provide an optimal solution for (11).

Next we introduce a method of finding $\mathbf{a}^0$. For a specific $\mathbf{c}$, in order to get $\mathbf{a}^0$, we need to determine the set of α^+. We can conduct an exhaustive search over a grid of $k_1, ..., k_M$ depending on the desired accuracy, and store these results $(P_d, \mathbf{P}_f, \mathbf{k})$ into a lookup table T_1.

After that, we iterate every item in T_1 and put the item into table T_2 if the item satisfies: $\mathbf{P}_f \leq \mathbf{c}$. At last, we iterate every item in T_2 and get the largest P_d and its decision rule $\mathbf{k}$.

III. The ROC surface of MENP

The ROC surface of MENP (M-ROC) depicts the relationship between P_d and $c_1, c_2, ..., c_M$. On one hand, it illustrates the largest P_d can that can be achieved under the constraints $P_{f_i} \leq c_i (i = 1, 2, ..., M)$, on the other hand, it provides the range of $\mathbf{c}$ for a given P_d. Points $(P_d, c_1, c_2, ..., c_M)$ on M-ROC surface can be divided into two types:

1. Those with $[c_1, c_2, ..., c_M] \in \alpha^+$. Define this set of points as M_0;

2. Those with $[c_1, c_2, ..., c_M] \notin \alpha^+$. Define this set of points as M_1.

Obviously points in M_0 can be achieved by MENP (i) and points in M_1 can be achieved by MENP (ii). Next we consider a property of the M-ROC surface.

Property 1 *Assume hypotheses given as:*

$$
\begin{aligned}
H_0: &\quad X \sim f_0(x) \\
H_1: &\quad X \sim f_1(x) \\
&\quad ...... \\
H_M: &\quad X \sim f_M(x)
\end{aligned}
\tag{16}
$$

define $g(x) = \frac{\sum_{i=1}^{M} k_i f_i(x)}{f_0(x)}$ *and let* $F_i(x)$ *to represent the CDF of hypothesis* H_i *(*$i = 1, ..., M$*).*

If $g(x)$ is a monotonically increasing function of x for any non-negative k_i $(i = 1, ..., M)$ and $F_i(x)$ is monotonically increasing function, then we have:

(1)The region achieved by ENP test with non-negative parameters degenerates to a curve and the decision rule for M_0 is $x \underset{H_0}{\overset{\bar{H}_0}{\gtrless}} x_0$.

(2)For a specific false alarm constraints $P_{f_i} \leq c_i$ $(i = 1, ..., M)$, the decision rule for MENP test is $x \underset{H_0}{\overset{\bar{H}_0}{\gtrless}} x_0$, where $x_0 = \min(F_1^{-1}(c_1), ..., F_M^{-1}(c_M))$.

(3)The expression of P_d, P_{f_1}, ..., P_{f_M} can be written as

$$
\begin{aligned}
P_d &= \text{Pr}(X \leq x_0 | H_0) = F_0(x_0) \\
P_{f_1} &= \text{Pr}(X \leq x_0 | H_1) = F_1(x_0) \\
&\ldots\ldots \\
P_{f_M} &= \text{Pr}(X \leq x_0 | H_2) = F_M(x_0)
\end{aligned}
\tag{17}
$$

The proof is given in the appendix E.

We consider two examples that illustrate the properties of M-ROC surface for $M = 2$.

A. Gaussian Hypotheses

Assume three hypotheses given as:

$$\begin{aligned} H_0: &\qquad X \sim \mathcal{N}(-1, 1) \\ H_1: &\qquad X \sim \mathcal{N}(0, 1) \\ H_2: &\qquad X \sim \mathcal{N}(1, 10)\,, \end{aligned} \tag{18}$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian PDF with mean μ and variance σ^2. To form the M-ROC surface, we first consider points in region M_0.The decision rule in region M_0 is given by (10) with $M = 2, \quad k_1, k_2 \geq 0$, and the expression for P_d, P_{f_1} and P_{f_2} are given by (9). According to **Neyman Pearson Lemma**, if $k_1, k_2 \geq 0$, then $P_{f_1} = c_1$ and $P_{f_2} = c_2$.

We use Matlab to illustrate the region M_0. The values of k_1 and k_2 range from 0 to 100 in steps of 0.01. Substituting the value of k_1 and k_2 into (9), results in the corresponding P_d P_{f_1} and P_{f_2}. The set M_0 is illustrated in Figure 1. Figure 2 presents the projection of Figure 1 on the c_1, c_2 plane. In Figure 2, region N_0 is the projection of region M_0 on the c_1, c_2 plane. Since M_0 is the set of points with $[c_1, c_2] \in \alpha^+$, N_0 is the region of points that belong to α^+. Define curve L_1 as the set of points that satisfy the conditions: (1) $(c_1, c_2) \in N_0$; (2) $(c_1, c_2 + \epsilon) \notin N_0$, for any $\epsilon \geq 0$. Define curve L_2 as the set of points that satisfy the conditions: (1) $(c_1, c_2) \in N_0$; (2) $(c_1 + \epsilon, c_2) \notin N_0$, for any $\epsilon \to 0$. N_1 denotes the region enclosed by line $c_1 = 0$, c_2; line c_1, $c_2 = 1$ and curve L_1. N_2 denotes the region enclosed by line $c_1 = 1$, c_2; line c_1, $c_2 = 0$ and curve L_2. The regions of N_0 N_1 and N_2 are shown in Figure 2.

Conclusion 1

(1) All points belonging to region N_1 or curve L_1, if they have the same c_1, they have the same decision rule and same P_d.

(2) All points belonging to region N_2 or curve L_2, if they have the same c_2, they have the same decision rule and same P_d.

The proof is given in the appendix F.

Since we already know P_d for points in region N_0 and curves L_1 and L_2 belong to N_0, we can get P_d for points in N_1 and N_2. M-ROC surface for this example is given in Figure 3 and the contour of M-ROC surface is given in Figure 4.

Next we consider a more general case of $M+1$ hypotheses given as

$$\begin{aligned}
H_0 : &\quad X \sim \mathcal{N}(\mu_0, \sigma_0^2) \\
H_1 : &\quad X \sim \mathcal{N}(\mu_1, \sigma_1^2) \\
&\quad \ldots\ldots \\
H_M : &\quad X \sim \mathcal{N}(\mu_M, \sigma_M^2)
\end{aligned} \tag{19}$$

We will prove when $\sigma_0^2 = \sigma_1^2 = ... = \sigma_M^2$ and $\mu_0 < \mu_i (i = 1, ..., M)$, the region achieved by ENP test with $k_i \geq 0 (i = 1, ..., M)$ degenerates to a curve.

Consider

$$g(x) = \sum_{i=1}^{M} k_i \frac{f_i(x)}{f_0(x)}. \tag{20}$$

Since

$$f_i(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp(-\frac{(x-\mu_i)^2}{2\sigma_i^2}). \tag{21}$$

we have

$$g(x) = \sum_{i=1}^{M} \frac{\frac{1}{\sqrt{2\pi\sigma_i^2}} \exp(-\frac{(x-\mu_i)^2}{2\sigma_i^2})}{\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp(-\frac{(x-\mu_0)^2}{2\sigma_0^2})} \tag{22}$$

Using the condition $\sigma_i^2 = \sigma_0^2 (i = 1, ..., M)$ in (22), we have

$$g(x) = \sum_{i=1}^{M} k_i \exp(\frac{(\mu_i - \mu_0)(2x - \mu_i - \mu_0)}{2\sigma_0^2}) \tag{23}$$

Defining $c_i = \frac{\mu_i - \mu_0}{2\sigma_0^2}$, (23) can be written as

$$g(x) = \sum_{i=1}^{M} k_i \exp(c_i(2x - \mu_0 - \mu_i) \tag{24}$$

From the condition $\mu_0 < \mu_i (i = 1, ..., M)$, we know $c_i > 0$. Hence it can be conclude that $g(x)$ is a monotonically increasing function with x. Hence from **Property 1** we have that the region M_0 (region achieved by ENP test with $k_i \geq 0(i = 1, ..., M)$) degenerates to a curve. For a specific $\mathbf{c}$, the decision rule is

$$x \underset{\bar{H}_0}{\overset{H_0}{\lesseqgtr}} x_0$$

where $x_0 = \min\{F_1^{-1}(c_1), ..., F_M^{-1}(c_M)\}$.

B. Chi-Square Hypotheses

Assume three hypotheses given as:

$$\begin{aligned}
H_0 : \qquad & \frac{X}{\sigma_0^2} \sim \mathcal{X}^2(2N) \\
H_1 : \qquad & \frac{X}{\sigma_1^2} \sim \mathcal{X}^2(2N) \\
& \ldots\ldots \\
H_M : \qquad & \frac{X}{\sigma_2^M} \sim \mathcal{X}^2(2N) \,,
\end{aligned} \tag{25}$$

where $\mathcal{X}^2(2N)$ is the Chi-square distribution with $2N$ degree freedom(N is an integer, $\sigma_0^2 < \sigma_1^2, ..., \sigma_M^2$ and $\sigma_i^2 \neq \sigma_j^2$ if $i \neq j$). By a random variable transformation space [20], we can get the PDFs for the hypotheses:

$$H_0: \qquad f_0(x) = \frac{1}{\sigma_0^2 2^N \Gamma(N)} \left(\frac{x}{\sigma_0^2}\right)^{N-1} \exp\left(-\frac{x}{2\sigma_0^2}\right)$$

$$H_1: \qquad f_1(x) = \frac{1}{\sigma_1^2 2^N \Gamma(N)} \left(\frac{x}{\sigma_1^2}\right)^{N-1} \exp\left(-\frac{x}{2\sigma_1^2}\right) \tag{26}$$

......

$$H_M: \qquad f_2(x) = \frac{1}{\sigma_M^2 2^N \Gamma(N)} \left(\frac{x}{\sigma_M^2}\right)^{N-1} \exp\left(-\frac{x}{2\sigma_M^2}\right)$$

.

In the following, we will prove that in this example the region achieved by ENP test with **k** degenerates to a curve.

Consider

$$g(x) = \frac{\sum_{i=1}^{M} k_i f_i(x)}{f_0(x)} \quad k_i \geq 0 \tag{27}$$

Substituting $f_i(x)(i = 1, ..., M)$ from (26) into (27), we get:

$$g(x) = \sum_{i=1}^{M} k_i' \exp\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_i^2}\right)x \tag{28}$$

where $k_i' = k_i(\frac{\sigma_0}{\sigma_i})^{2N}, i = 1, ..., M$. Define $p_i = \frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_i^2}, i = 1, ..., M$. Hence $g(x) =$

$\sum_{i=1}^{M} k_i' \exp p_i x$.

The parameters $k_i'(i = 1, ..., M)$ are always non-negative when $k_i(i = 1, ..., M)$ are such, and from the condition $\sigma_0^2 \leq \sigma_i^2(i = 1, ..., M)$ we can conclude $p_i(i = 1, ..., M)$ are positive. Hence $g(x)$ is a monotonically increasing function with x. From **Property 1**, we have that the region achieved by ENP test with $k_i(i = 1, ..., M)$ degenerates to a curve. For a specific **c**, the decision rule is

$$x \underset{H_0}{\overset{\bar{H}_0}{\gtrless}} x_0 \quad , \tag{29}$$

where $x_0 = \min\{F_1^{-1}(c_1), ..., F_M^{-1}(c_M)\}$, and the corresponding $P_d = F_0(x_0)$.

Figure 5 shows the M-ROC surface for $M = 2$, $\sigma_0^2 = 1$, $\sigma_1^2 = 1.1$, $\sigma_2^2 = 1.15$ and $N = 120$. The bold curve is the region M_0.

IV. Determine the decision rule Under MENP framework

In this section, we consider the method to derive the decision rule under MENP framework. Except for certain cases (like the Chi-Square Hypotheses Example), P_{f_i} $(i = 1, 2, ..., M)$ and P_d are generally not explicit function of k_i. Hence a variational approach to this problem is not applicable. One possible method is Conducting an exhaustive search over a grid of k_i $(i = 1, 2, ..., M)$ depending on desired accuracy and store these results for a table to lookup operation. However as the number of M increasing, the lookup table will be prohibitively large, which make it not plausible for a spectrum sensing device to store these data. To solve this issue, in [15], the authors proposed a method under ENP framework. In the following, we will prove the method proposed by [15] can also be applied to MENP framework.

In [15], the author approached an search algorithm which could get the decision rule satisfy

$$\begin{aligned} \max \quad & P_d \\ \text{s.t.} \quad & \mathbf{P}_f \le \mathbf{c} \\ & \mathbf{k} \ge 0 \end{aligned} \tag{30}$$

where $\mathbf{c}$ is a specific constraints for $\mathbf{P}_f$. When $\mathbf{k} \ge 0$, $\mathbf{P}_f \in \alpha^+$ so we can write (31) as

$$\begin{aligned} \max \quad & P_d \\ \text{s.t.} \quad & \mathbf{P}_f \le \mathbf{c} \\ & \mathbf{P}_f \in \alpha^+ \end{aligned} \tag{31}$$

For the situation $\mathbf{c} \in \alpha^+$, the decision rule for ENP and MENP are the same, so we can use this search algorithm to get the ENP (MENP) parameters ($\mathbf{k}$). In the following, we focus on the situation when $\mathbf{c} \notin \alpha^+$.

Assume $\mathbf{c} \notin \alpha^+$, under MENP framework, the decision rule satisfies (15). Since $\mathcal{C} =$

$\mathcal{A}_c \cap \alpha^+$ and $F(\mathbf{a}) = P_d$, (15) can be written as

$$\begin{aligned} \max \quad & P_d \\ \text{s.t.} \quad & \mathbf{P}_f \leq \mathbf{c} \\ & \mathbf{P}_f \in \alpha^+ \end{aligned} \tag{32}$$

It can be observed that (31) and (32) has the same form. Hence the searching algorithm approached in [15] can also be used to derive the decision rule for $\mathbf{c} \notin \alpha^+$.

The following example illustrate the operation of the algorithm when $\mathbf{c} \notin \alpha^+$. Assume three hypotheses given as (18). We consider the decision rule for $c_1 = 0.2, c_2 = 0.4$. The algorithm carries eleven iterations of parameter adjustments to achieve the decision rule. The corresponding (P_{f_1}, P_{f_2}) points are plotted and connected by dotted lines in Fig. 7. The (P_{f_1}, P_{f_2}) point of the first and last iteration is plotted as a 'o' and '□', respectively. The change of P_d after each iteration is plotted in Figure 8. The ENP parameters of the final solution are

$$k_1 = 1.4072$$

$$k_2 = 0$$

and the performance measures of the solution are

$$P'_{f_1} = 0.2$$

$$P'_{f_2} = 0.2705$$

$$P'_d = 0.5629$$

In Figure 3, the probability of detection for $c_1 = 0.2$ and $c_2 = 0.4$ is 0.5629, which is equal to P'_d. The simulation result supports the view that the search algorithm proposed in [15] also work under MENP framework.

V. Energy Based Spectrum Sensing for two Primary Users

In this section, we consider an example for applying the MENP test to energy based spectrum sensing for two primary users.

A. System Model

Consider a cognitive radio system where the frequency can be occupied by one of the two distinct primary signals $\{s_1, s_2\}$ or it could be vacant. Let H_0 denote the hypothesis that the channel is vacant, and $H_i, i = 1, 2$, denote the hypothesis that the channel is occupied by primary user signal $s_i, i = 1, 2$.

A block diagram of the system is illustrated in Figure 6. The system consists of a measuring device followed by a testing device. The role of the measuring device is to output a suitable test statistics. The role of the testing device is to decide whether the channel is free based on the output of the measuring device. The input of the measuring device is

$$x[n] = \begin{cases} \omega[n] & \text{when no primary user is occupying the channel} \\ \omega[n] + s_i[n] & \text{when primary user } i \text{ is occupying the channel}. \end{cases} \quad (33)$$

for $n = 1, 2, ..., N$, and $i = 1, 2$. We also assume that $s_i[n]$ and $\omega[n]$ are i.i.d. circular symmetric complex Gaussian(CSCG) random variables with variance $2\sigma_{s_i}^2$ and $2\sigma_\omega^2$, in other words, $s[n] \sim \mathcal{CN}(0, 2\sigma_{s_i}^2)$ and $\omega[n] \sim \mathcal{CN}(0, 2\sigma_\omega^2)$. Since the noise and the signal are independent, $s_i[n] + \omega[n] \sim \mathcal{CN}(0, 2(\sigma_\omega^2 + \sigma_{s_i}^2))$.

When the testing device is a energy detector, its output is

$$X = \sum_{n=1}^{N} |x[n]|^2 = \sum_{n=1}^{N} (x_R[n]^2 + x_I[n]^2), \quad (34)$$

where x_R and x_I are the real and imaginary part of signal x. Substituting (33) into (34) and defining $\sigma_i^2 = \sigma_\omega^2 + \sigma_{s_i}^2$ we have

$$
\begin{aligned}
\mathcal{H}_0: &\quad \frac{X}{\sigma_\omega^2} \sim \mathcal{X}_{2N} \\
\mathcal{H}_1: &\quad \frac{X}{\sigma_1^2} \sim \mathcal{X}_{2N} \\
\mathcal{H}_2: &\quad \frac{X}{\sigma_2^2} \sim \mathcal{X}_{2N}
\end{aligned}
\tag{35}
$$

Our goal is to test H_0 against $\bar{H}_0$.

B. Using Minimax Hypothesis Decision Rule

Since there being multiple hypotheses and lacking prior information, a conventional method for solving this issue would be using minimax hypothesis testing.

Let δ denote a decision rule in the context of spectrum sensing. Let C_{ij} denote the cost incurred by choosing hypothesis H_i when hypothesis H_j is true. Let Ω_M be the set of all minimax decision rule in the context of spectrum sensing. Let Ω_{NP} be the set of all ENP decision rule with non-negative parameters in the context of spectrum sensing. Define $P_i(\mathcal{C}) = \int_{\mathcal{C}} f_i(x) \mathrm{d}x$.

First we will show $\Omega_M = \Omega_{NP}$.

Assume there are $M+1$ hypotheses (H_0, ..., H_M) and a minimax decision rule will be in form of:

$$\delta = \begin{cases} H_0 \quad \text{is true} \quad \text{if } \ x \in \mathcal{C}_0 \\ H_1 \quad \text{is true} \quad \text{if } \ x \in \mathcal{C}_1 \\ \ldots\ldots \\ H_M \quad \text{is true} \quad \text{if } \ x \in \mathcal{C}_M \,, \end{cases} \tag{36}$$

where H_0 denotes the channel is free and H_i $(i = 1, ..., M)$ denotes the channel is occupied by primary user i. In the context of spectrum sensing, following events will jeopardize the system performance:

(1) The system miss a spectrum hole, the lost is represented by C_{i0} $(i = 1, ..., M)$. In spectrum sensing, the performance lost caused by choosing H_k when H_0 is true (C_{k0}) is equal to the lost caused by choosing H_j when H_0 is true (C_{j0}), where $k \neq j$ and $k, j \neq 0$, hence we set $C_{10} = C_{20} = ...C_{M0} = a_0$ $(a_0 > 0)$.

(2) Primary user i is interfered by secondary user while transferring data, the lost is represented by C_{0i}. We set $C_{0i} = a_i$ $(a_0 > 0)$.

For C_{ij} other than mentioned above, we set them to zero.

The conditional risk for each hypothesis can be written as

$$
\begin{aligned}
R_0(\delta) &= C_{00}P_0(\mathcal{C}_0) + C_{10}P_0(\mathcal{C}_1) + ... + C_{M0}P_0(\mathcal{C}_M) = a_0P_0(\mathcal{C}_1) + a_0P_0(\mathcal{C}_2) + ... + a_0P_0(\mathcal{C}_M) \\
&= a_0P_0)(\mathcal{C}_1 \cup \mathcal{C}_2 \cup ... \cup \mathcal{C}_M) \\
&= a_0P_0(\bar{\mathcal{C}_0}) \\
&= a_0(1 - P_0(\mathcal{C}_0)) \\
R_1(\delta) &= C_{01}P_1(\mathcal{C}_0) + C_{11}P_1(\mathcal{C}_1) + ... + C_{M1}P_1(\mathcal{C}_M) = a_1P_1(\mathcal{C}_0) \\
&......\\
R_M(\delta) &= C_{0M}P_M(\mathcal{C}_0) + C_{1M}P_M(\mathcal{C}_1) + ... + C_{2M}P_M(\mathcal{C}_2) = a_MP_M(\mathcal{C}_0)\,.
\end{aligned}
\tag{37}
$$

For any possible prior probability π_0, π_1, ..., π_M, the total cost function is

$$
\begin{aligned}
r(\delta) &= \pi_0R_0(\delta) + \pi_1R_1(\delta) + ... + \pi_MR_M(\delta) \\
&= \pi_0a_0 - (\pi_0a_0P_0(\mathcal{C}_0) - \pi_1a_1P_1(\mathcal{C}_0 - ... - \pi_Ma_MP_M(\mathcal{C}_0))) \\
&= \pi_0a_0 - \int_{\mathcal{C}_0} \pi_0a_0f_0(x) - \pi_1a_1f_1(x) - ... - \pi_Ma_Mf_M(x)\mathrm{d}x\,.
\end{aligned}
\tag{38}
$$

and thus we see that $r(\delta)$ is a minimum over all $\mathcal{C}_0$ if we choose

$$
\begin{aligned}
\mathcal{C}_0 &= \{x \in \mathcal{C}_0 | \pi_0a_0f_0(x) - \pi_1a_1f_1(x) - ... - \pi_Ma_Mf_M(x) \geq 0\} \\
&= \{x \in \mathcal{C}_0 | \frac{f_0(x)}{(\pi_1a_1/\pi_0a_0)f_1(x) + ... + (\pi_Ma_M/\pi_0a_0)f_M(x)} \geq 0\}\,.
\end{aligned}
\tag{39}
$$

We can also observe that it is impossible to decide the range of $\mathcal{C}_1$, $\mathcal{C}_2$..., $\mathcal{C}_M$ through $r(\delta)$, so δ cannot discriminate among H_i $(i = 1, ..., M)$. In other words, δ can only distinguish between H_0 and $\bar{H}_0$. Hence (36) can be written as:

$$\delta : \quad \frac{f_0(x)}{(\pi_1 a_1/\pi_0 a_0) f_1(x) + ... + (\pi_M a_M/\pi_0 a_0) f_M(x)} \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} 1 \tag{40}$$

Let $k_i = \frac{\pi_i a_i}{\pi_0 a_0}$, since $a_i > 0$, we have $k_i \geq 0$. δ has following form:

$$\frac{f_0(x)}{k_1 f_1(x) + ... + k_M f_M(x)} \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} 1 \quad k_i \geq 0 \tag{41}$$

We can see (41) has the same form of ENP decision rule with non-negative parameters. In other words, $\forall \delta \in \Omega_M$, we have $\delta \in \Omega_{NP}$. We can conclude $\Omega_M \subseteq \Omega_{NP}$.

Consider an ENP decision rule

$$\delta : \quad \frac{f_0(x)}{k_1 f_1(x) + ... + k_M f_M(x)} \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} 1 \quad k_i \geq 0 \tag{42}$$

Under (42), the expression of P_d and P_{f_i} are

$$\begin{aligned} P_d &= P_0(\mathcal{C}_0) \\ P_{f_i} &= P_i(\mathcal{C}_0)\,, \end{aligned} \tag{43}$$

where $\mathcal{C}_0$ is defined as

$$\mathcal{C}_0 = \{x \in \mathcal{C}_0 | \frac{f_0(x)}{k_1 f_1(x) + ... + k_M f_M(x)} \geq 1\} \,. \tag{44}$$

In the following we will show (42) is a minimax decision rule with $a_0 = 1$ and $a_i = \frac{1-P_d}{P_{f_i}}$.

Since by using decision rule (42) we have

$$R_0(\delta) = R_1(\delta) = ... = R_M(\delta) = 1 - P_d \,, \tag{45}$$

we can see (42) is also a minimax decision rule. In other words, $\forall \delta \in \Omega_{NP}$, we have $\delta \in \Omega_M$. We can conclude $\Omega_{NP} \subseteq \Omega_M$.

Since $\Omega_M \subseteq \Omega_{NP}$ and $\Omega_{NP} \subseteq \Omega_M$, we proved $\Omega_M = \Omega_{NP}$.

Next we consider using minimax decision rule to solve (35) in Matlab. P_d, P_{f_1} and P_{f_2} will be used to depict the performance. The value of a_0, a_1, a_2 change from 0.01 to 100 in step of 0.05. For each a_0, a_1 and a_2, we iterate every decision rule in form of (41) and choose the one that has the minimum value of $\max\{R_0, R_1, R_2\}$ as the minimax decision rule. Fig 9 is the ROC under minimax hypotheses testing. We can see the curve in Fig 9 has the same shape as the bold curve in Fig 5, which is achieved by ENP with non-negative parameters. The simulation result supports our theoretical conclusion that $\Omega_{NP} = \Omega_M$.

C. Using MENP Framework

Next we test H_0 against $\overline{H}_0$ using MENP framework. The problem can be abstracted into following optimization problem:

$$\begin{aligned} \max \qquad & P_d\,, \\ \text{s.t.} \qquad & P_{f_i} \leq c_i \quad (i = 1, 2)\,. \end{aligned} \tag{46}$$

Our object is to plot the maximum P_d vs. c_1, c_2 (M-ROC) and find the specific decision rule for a given c_1, c_2.

Since $\sigma_\omega^2 < \sigma_1^2, \sigma_2^2$, this problem has the same form as that of the Chi-Square example given in last section. Hence for any given c_1, c_2 the decision rule has the form

$$\delta : \qquad x \underset{\overline{H}_0}{\overset{H_0}{\lesseqgtr}} x_0\,, \tag{47}$$

where $x_0 = \min(F_1^{-1}(c_1), F_2^{-1}(c_2))$. The performance analysis is given in Fig 5.

D. Comparison between Minimax framework and MENP framework

By comparing Fig 9 with Fig 5, we can observe:

(1) The performance of minimax hypotheses testing is reflected by the relationship between P_d, P_{f_1} and P_{f_2}, while the performance of MENP framework is reflected by the relationship between P_d, c_1 and c_2. In the context of spectrum sensing, the constraints are usually given

in form of $\mathbf{P}_f \leq \mathbf{c}$, hence the latter one is more feasible in practice.

(2) The achievable region for MENP framework is larger. It can be seen from Fig 9, the achievable region for minimax hypotheses testing is a curve, while the achievable region for MENP framework is the whole c_1, c_2 plane.

(3) It is easier to design the detector's parameter under MENP framework. For a specific c_1 and c_2, it is easy to get the MENP decision rule by (47). While for minimax decision rule, there is no direct relationship between a_i and P_{f_i}.

Hence we can conclude in the context of spectrum sensing, MENP framework is more feasible than minimax hypotheses testing.

VI. CONCLUSIONS

In this paper, we explore the new field of spectrum sensing with emphasis on providing different levels of protections for various primary users. The proposed MENP test could achieve a value approaches to the largest probability of detection detection under multiple constraints of probability of false alarms. After that, we approach M-ROC surface to represent the relationship between the probability of detection and the constraint conditions. Two examples respectively concerning Gaussian distribution and Chi-Square distribution are given to show the properties of M-ROC surface. One method to achieve the MENP decision rule is provided with simulation. We also proposed an energy based detector for two different types of primary users. For simplicity, we assume the signal samples of the primary users subject to iid CSCG distribution and the variance of the signal and the noise are known to the detector. In this work, we did not consider randomized decision rule. This could be a future topic.

APPENDIX

A. Proof for Lemma 1

Define $\mathbf{k} = [k_1, k_2, ..., k_M]^T$ and $\mathbf{P}_f = [P_{f_1}, P_{f_2}, ..., P_{f_M}]^T$. Since both P_d and $\mathbf{P}_f$ are functions of $\mathbf{k}$, $\mathbf{P}_f(\mathbf{k}_0)$ denotes the value of $\mathbf{P}_f$ when $\mathbf{k} = \mathbf{k}_0$ and $P_d(\mathbf{k}_0)$ denotes the value of P_d when $\mathbf{k} = \mathbf{k}_0$. Using Taylor's expansion for $\mathbf{P}_f$ and P_d,

$$P_d = P_d(\mathbf{k}_0) + \left.\frac{\mathrm{d}P_d}{\mathrm{d}\mathbf{k}^T}\right|_{\mathbf{k}=\mathbf{k_0}} (\mathbf{k} - \mathbf{k}_0) + o(\mathbf{k} - \mathbf{k}_0) \tag{48}$$

$$\mathbf{P}_f = \mathbf{P}_f(\mathbf{k}_0) + \left.\frac{\mathrm{d}\mathbf{P}_f}{\mathrm{d}\mathbf{k}^T}\right|_{\mathbf{k}=\mathbf{k_0}} (\mathbf{k} - \mathbf{k}_0) + o(\mathbf{k} - \mathbf{k}_0) \tag{49}$$

here $\mathbf{k} \to \mathbf{k}_0$.

Consider the hyperplane y as a function of $\mathbf{x}$ defined by

$$\mathbf{x} = \mathbf{P}_f(\mathbf{k}_0) + \left.\frac{\mathrm{d}\mathbf{P}_f}{\mathrm{d}\mathbf{k}^T}\right|_{\mathbf{k}=\mathbf{k_0}} (\mathbf{z} - \mathbf{k}_0) \tag{50}$$

$$y = P_d(\mathbf{k}_0) + \frac{\mathrm{d}P_d}{\mathrm{d}\mathbf{k}^T}\bigg|_{\mathbf{k}=\mathbf{k_0}}(\mathbf{z} - \mathbf{k}_0) \tag{51}$$

The above equations construct a tangent hyperplane for the ROC surface at point $(P_d(\mathbf{k}_0), \mathbf{P}_f^T(\mathbf{k}_0))$. Combining both equations we get

$$y = P_d(\mathbf{k}_0) + \frac{\mathrm{d}P_d}{\mathrm{d}\mathbf{k}^T}\bigg|_{\mathbf{k}=\mathbf{k_0}}(\frac{\mathrm{d}\mathbf{P}_f}{\mathrm{d}\mathbf{k}^T}\bigg|_{\mathbf{k}=\mathbf{k_0}})^{-1}(\mathbf{x} - \mathbf{P}_f(\mathbf{k}_0)) \tag{52}$$

Hence the normal for point $(P_d(\mathbf{k}_0), \mathbf{P}_f^T(\mathbf{k}_0))$ on ROC surface can be written as

$$[-1, \frac{\mathrm{d}P_d}{\mathrm{d}\mathbf{k}^T}\bigg|_{\mathbf{k}=\mathbf{k_0}}(\frac{\mathrm{d}\mathbf{P}_f}{\mathrm{d}\mathbf{k}^T}\bigg|_{\mathbf{k}=\mathbf{k_0}})^{-1}]. \tag{53}$$

In the following, we will prove $\frac{\mathrm{d}P_d}{\mathrm{d}\mathbf{k}^T}(\frac{\mathrm{d}\mathbf{P}_f}{\mathrm{d}\mathbf{k}^T})^{-1} = \mathbf{k}^T$, which can be written as

$$\frac{\mathrm{d}P_d}{\mathrm{d}\mathbf{k}^T} = \mathbf{k}^T\frac{\mathrm{d}\mathbf{P}_f}{\mathrm{d}\mathbf{k}^T} \tag{54}$$

Previous equation can be written in component form as

$$\frac{\partial P_d}{\partial k_i} - \sum_{n=1}^{M} k_n \frac{\partial P_{f_n}}{\partial k_i} = 0 \quad (i = 1, 2, ..., M). \tag{55}$$

Calculating the partial derivatives results in

$$\frac{\partial P_{f_n}}{\partial k_i} = -\int_{\mathcal{D}} \delta(f_0(x) - \sum_{j=1}^{M} k_j f_j(x)) f_i(x) f_n(x) \mathrm{d}x \,, \tag{56}$$

$$\frac{\partial P_d}{\partial k_i} = -\int_{\mathcal{D}} \delta(f_0(x) - \sum_{j=1}^{M} k_j f_j(x)) f_i(x) f_0(x) \mathrm{d}x \,, \tag{57}$$

where $\delta(\bullet)$ is Dirac's delta function defined as following,

$$\delta(x) = \lim_{\epsilon \to 0} \begin{cases} \frac{1}{\epsilon} & \text{when } \; x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2}) \\ 0 & \text{otherwise} \end{cases} \tag{58}$$

Defining $g(x) = f_0(x) - \sum_{j=1}^{M} k_j f_j(x)$, (55) can be written as $\int_{\mathcal{D}} \delta(g(x)) g(x) f_n(x) \mathrm{d}x = 0, n = 0, 1, ..., M$.

When $g(x) \neq 0$, we have $\delta(g(x)) = 0$ and $\delta(g(x)) g(x) f_i(x) = 0$. When $g(x) = 0$, we can solve the equation according to the definition of $\delta(\bullet)$ and consider

$$\int_{\{x | g(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})\}} \frac{1}{\epsilon} g(x) f_n(x) \mathrm{d}x \quad n = 0, 1, ..., M \tag{59}$$

Since when $g(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})$, $|g(x)| < \frac{\epsilon}{2}$,

$$|\int_{\{x | g(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})\}} \frac{1}{\epsilon} g(x) f_i(x) \mathrm{d}x| < \int_{\{x | g(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})\}} \frac{1}{2} f_i(x) \mathrm{d}x \tag{60}$$

Assume x_s is one of the zero point of $g(x)$, also assume $g'(x_s)$, $g^{(2)}(x_s)$, ..., $g^{(n-1)}(x_s)$

are zero but $g^{(n)}(x_s) \neq 0$ (here $n = 1, 2, ...$). Use Taylor expansion near point x_s,

$$g(x) = \frac{g^{(n)}(x_s)}{n!}(x - x_s)^n + o((x - x_s)^n) \, . \tag{61}$$

From equation (61), it can be seen when $g(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})$, $x \in \left(x_s - \left(\frac{n!\varepsilon}{2|g^{(n)}(x_s)|}\right)^{\frac{1}{n}}, x_s + \left(\frac{n!\varepsilon}{2|g^{(n)}(x_s)|}\right)^{\frac{1}{n}}\right)$. Define $\triangle x = \left(\frac{n!\varepsilon}{2|g^{(n)}(x_s)|}\right)^{\frac{1}{n}}$, when $\epsilon \to 0$, $\triangle x \to 0$ and

$$g(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2}) \Leftrightarrow x \in (x_s - \triangle x, x_s + \triangle x) \tag{62}$$

Hence when $\epsilon \to 0$ we have

$$\int_{\{x|g(x)\in(-\frac{\epsilon}{2},\frac{\epsilon}{2})\}} \frac{1}{2} f_i(x)\mathrm{d}x \to f_i(x_s) \triangle x \to 0 \tag{63}$$

Using the above two conclusions for $g(x) = 0$ and $g(x) \neq 0$, we get

$$\int_{\mathcal{D}} \delta(g(x))g(x)f_i(x)\mathrm{d}x = 0 \tag{64}$$

In this way, we prove the normal for the point $\mathbf{P}$ on the ROC is $(-1, k_1, k_2, ..., k_M)$.

B. Proof for Lemma 2

The expression of tangent hyper surface for point $(P_d^0, P_{f_1}^0, ..., P_{f_M}^0)$ on the ROC hyper surface can be written as

$$P_d = P_d^0 + \sum_{i=1}^{M} \frac{\partial P_d}{\partial P_{f_i}}\bigg|_{P_{f_i}=P_{f_i}^0} (P_{f_i} - P_{f_i}^0)\,. \tag{65}$$

Hence the normal at this point is

$\mathbf{n} = [-1, \frac{\partial P_d}{\partial P_{f_1}}, \frac{\partial P_d}{\partial P_{f_2}}, ..., \frac{\partial P_d}{\partial P_{f_M}}]$. Since we have proved that the normal for this point is $\mathbf{n} = [-1, k_1, k_2, ..., k_M]$, we must have

$$\frac{\partial P_d}{\partial P_{f_i}}\bigg|_{P} = k_i \tag{66}$$

C. Proof for Lemma 3

This lemma is used in [21] without serious math proof. In this part, we will use a similar proof as it was used in [16], [19] to prove Extended Neyman Pearson (iv).

Before we proceed, we will define some notations for presentation. Define $\mathbf{c}^T = [c_1, c_2, ..., c_M]$, $\boldsymbol{\mu}_0^T = [\mu_1, ..., \mu_M]$, $\mathbf{k}^T = [k_1, k_2, ..., k_M]$ and $\mathbf{P}_f^T = [P_{f_1}, P_{f_2}, ..., P_{f_M}]$, obviously $\mathbf{c}$, $\boldsymbol{\mu}_0$, $\mathbf{k}$ and $\mathbf{P}_f$ are vectors in M dimensional Euclidean space. Let $\boldsymbol{\mu}^T = [\boldsymbol{\mu}_0, \mu_{M+1}]$ denote a vector in $M+1$ dimensional Euclidean space. Let $F(\boldsymbol{\mu}_0)$ denote the largest P_d under the constraints $P_{f_i} = \mu_i \quad i = 1, ..., M$. Let $G(\boldsymbol{\mu}_0)$ denote the largest P_d under the constraints

$P_{f_i} \leq \mu_i \;\; i = 1, ..., M$. Let $P_d(\delta)$, $P_{f_i}(\delta)$ denote the P_d and P_{f_i} achieved by using decision rule δ. By $\mathbf{A} \leq \mathbf{B}$, $\mathbf{A} = \mathbf{B}$ and $\mathbf{A} \geq \mathbf{B}$ we mean that every element of $\mathbf{A}$ is no larger than, equal to and no smaller than its corresponding element of $\mathbf{B}$, respectively. By $\mathbf{A} \in (0, 1)$ and $\mathbf{A} \in [0, 1]$, we mean that every element of $\mathbf{A}$ belongs to $(0, 1)$, $[0, 1]$respectively. Let G' denote the hyper surface achieved by $G(\boldsymbol{\mu}_0) = \mu_{M+1}$, where $\boldsymbol{\mu}_0 \in [0, 1]$. Let G_s denote the set of points in $M + 1$ Euclidean space such that

$$\{(\boldsymbol{\mu}_0, \mu_{M+1}) \in G_s | \boldsymbol{\mu}_0 \in [0, 1] \text{and} \quad \mu_{M+1} \in [0, G(\boldsymbol{\mu_0})]\}$$

we can see G_s is a space enclosed by G', $\mu_i = 1$ and $\mu_{M+1} = 0$ $(i = 1, ..., M)$. Let $\mathcal{N}$ represents the set of points in $M + 1$ dimensional Euclidean space such that

$$\begin{aligned} \{(\mu_1, \mu_2, ..., \mu_{M+1}) \in \mathcal{N} | \mu_i = \int_{\mathcal{S}} f_i(x) \mathrm{d}x \;\; i = 1, ..., M, \\ \mu_{M+1} = \int_{\mathcal{S}} f_{M+1}(x) \mathrm{d}x \text{for any subset } \mathcal{S} \text{ of } \mathcal{D}\} \end{aligned} \tag{67}$$

we can see, $\mathcal{N}$ is the set of $(\boldsymbol{\mu}_0, \mu_{M+1}) = (\mathbf{P}_f(\delta), P_d(\delta))$, where δ could be any possible decision rule.

The whole proof will be divided into following parts: we will first prove $G(\boldsymbol{\mu}_0)$ is a convex, non-decreasing function; we will also prove when $\boldsymbol{\mu}_0 \neq 0$, $G(\boldsymbol{\mu}_0) \neq 0$; then we will prove G_s is a convex set and $\mathcal{N} \subseteq G_s$; after that we will prove for any points $(\mu_1^0, \mu_2^0, ..., \mu_{M+1}^0) \in G'$, there exists a non-negative $\mathbf{k}$ such that for any $(\mu_1, \mu_2, ..., \mu_{M+1}) \in G_s$, $\mu_{M+1} - \sum_{i=1}^{M} k_i \mu_i \leq \mu_{M+1}^0 - \sum_{i=1}^{M} k_i \mu_i^0$ holds; in the end, we will prove for any $\boldsymbol{\mu}_0$, there exists non-negative $\mathbf{k}$ such that the optimal decision rule to achieve the largest P_d under constraint $\mathbf{P}_f \leq \boldsymbol{\mu}_0$ can be written in form of

$$\delta^* : \quad f_0(x) \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} \sum_{i=1}^{M} k_i f_i(x) \,.$$

Firstly we will prove $G(\boldsymbol{\mu}_0)$ is a convex non-decreasing function for $\boldsymbol{\mu}_0 \in [0,1]$. Let $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$ be two points on G' with coordinates $(\boldsymbol{\mu}_0^1, \mu_{M+1}^1)$ and $(\boldsymbol{\mu}_0^2, \mu_{M+1}^2)$, i.e. $\mu_{M+1}^1 = G(\boldsymbol{\mu}_0^1)$ and $\mu_{M+1}^2 = G(\boldsymbol{\mu}_0^2)$. Let δ_1 be the decision rule which can achieved the largest P_d under the constraint $\mathbf{P}_f \leq \boldsymbol{\mu}_0^1$ and δ_2 be the decision rule which can achieved the largest P_d under the constraint $\mathbf{P}_f \leq \boldsymbol{\mu}_0^2$. So we can see $\mu_{M+1}^1 = P_d(\delta^1) = G(\boldsymbol{\mu}_0^1)$, $\mu_{M+1}^2 = P_d(\delta^2) = G(\boldsymbol{\mu}_0^1)$, $P_f(\delta^1) \leq \boldsymbol{\mu}_0^1$ and $P_f(\delta^2) \leq \boldsymbol{\mu}_0^2$.

Construct a new randomized new test δ^3, where δ^1 and δ^2 are used with equal probability. With decision rule δ^3, we have $P_d(\delta^3) = 0.5P_d(\delta^1) + 0.5P_d(\delta^2) = 0.5G(\boldsymbol{\mu}_0^1) + 0.5G(\boldsymbol{\mu}_0^2)$ and $\mathbf{P}_f(\delta^3) = 0.5\mathbf{P}_f(\delta^1) + 0.5\mathbf{P}_f(\delta^2) \leq 0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2$. Let δ' denote the optimal decision rule for

$$\begin{aligned} & \max \quad P_d \\ & \text{s.t.} \quad \mathbf{P}_f \leq 0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2 \end{aligned} \tag{68}$$

then obviously $P_d(\delta') \geq P_d(\delta^3)$ (otherwise δ' cannot be the optimal decision rule), i.e. $G(0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2) > P_d(\delta^3)$. Since $P_d(\delta^3) = 0.5G(\boldsymbol{\mu}_0^1) + 0.5G(\boldsymbol{\mu}_0^2)$, the equation can be written as

$$G(0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2) \geq 0.5G(\boldsymbol{\mu}_0^1) + 0.5G(\boldsymbol{\mu}_0^2) \tag{69}$$

Hence $G(\boldsymbol{\mu}_0)$ is a convex function for $\boldsymbol{\mu}_0 \in [0,1]$.

According to the definition of $G(\boldsymbol{\mu}_0)$, when μ_i increases, the value of $G(\boldsymbol{\mu}_0)$ will not decrease, we can conclude $G(\boldsymbol{\mu}_0)$ is also a non-decreasing function of $\boldsymbol{\mu}_0$.

Next we will prove when $\boldsymbol{\mu}_0 \neq 0$, $G(\boldsymbol{\mu}_0) \neq 0$.

Let δ^* be the optimal decision rule for $\boldsymbol{\mu}_0$, by optimal decision rule we mean this decision rule can get the largest P_d under the constraint $\mathbf{P}_f \leq \boldsymbol{\mu}_0$. According to [21], there exists at least one $P_{f_i}(\delta^*) = \mu_i$ $(i = 1, ..., M)$. Without losing generality, let $P_{f_l}(\delta^*) = \mu_l$, i.e. $\int_{\mathcal{S}^*} f_l(x)\mathrm{d}x = \mu_l \neq 0$. Hence we can conclude the measurement of $\mathcal{S}^*$ is not zero. Since we know $f_0(x) \neq 0$ a.e. on $\mathcal{D}$, we can conclude $\int_{\mathcal{S}^*} f_0(x)\mathrm{d}x > 0$, i.e. $G(\boldsymbol{\mu}_0) \neq 0$.

In the following, we will consider the property of G_s.

Let $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$ be two points in G_s with coordinates $(\boldsymbol{\mu}_0^1, \mu_{M+1}^1)$ and $(\boldsymbol{\mu}_0^2, \mu_{M+1}^2)$. According to the definition of G_s, we have $0 \leq \boldsymbol{\mu}_0^1 \leq 1$, $0 \leq \boldsymbol{\mu}_0^2 \leq 1$, $0 \leq \mu_{M+1}^1 \leq G(\boldsymbol{\mu}_0^1)$ and $0 \leq \mu_{M+1}^2 \leq G(\boldsymbol{\mu}_0^2)$.

Let $\boldsymbol{\mu}^3$ be the middle point between $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$ with coordinate $(\boldsymbol{\mu}_0^3, \mu_{M+1}^3)$, where $\boldsymbol{\mu}_0^3 =$

$0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2$ and $\mu_{M+1}^3 = 0.5\mu_{M+1}^1 + 0.5\mu_{M+1}^2$.

Since $\boldsymbol{\mu}_0^1, \boldsymbol{\mu}_0^2 \in [0,1]$, we can conclude $\boldsymbol{\mu}_0^3 = 0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2 \in [0,1]$. Since $0 \le \mu_{M+1}^1 \le G(\boldsymbol{\mu}_0^1)$, $0 \le \mu_{M+1}^2 \le G(\boldsymbol{\mu}_0^2)$, we can conclude

$$0 \le \mu_{M+1}^3 = 0.5\mu_{M+1}^1 + 0.5\mu_{M+1}^2 \le 0.5G(\boldsymbol{\mu}_0^1) + 0.5G(\boldsymbol{\mu}_0^2) \le G(0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2) = G(\boldsymbol{\mu}_0^3)\,.$$

$0.5G(\boldsymbol{\mu}_0^1) + 0.5G(\boldsymbol{\mu}_0^2) \le G(0.5\boldsymbol{\mu}_0^1 + 0.5\boldsymbol{\mu}_0^2)$ comes form the fact that $G(\boldsymbol{\mu}_0)$ is a convex function. We can see $\boldsymbol{\mu}_0^3 \in [0,1]$ and $\mu_{M+1}^3 \in [0, G(\boldsymbol{\mu}_0^3)]$. Hence $\boldsymbol{\mu}^3 \in G_s$, i.e. G_s is a convex set.

Next we will prove $\mathcal{N} \subseteq G_s$, i.e. $\forall(\mu_0, ..., \mu_{M+1}) \in \mathcal{N}$, this point also belongs to set G_s. According to the definition of $\mathcal{N}$, $\forall(\mu_1^0, ..., \mu_{M+1}^0) \in \mathcal{N}$, we have $\mu_i^0 \in [0,1]$ $(i = 1, ..., M)$ and $\mu_{M+1}^0 \in [0, F(\boldsymbol{\mu}_0^0)]$, where $\boldsymbol{\mu}_0^0 = [\mu_1^0, ..., \mu_M^0]$. From the definition of $F(\boldsymbol{\mu}_0^0)$ and $G(\boldsymbol{\mu}_0^0)$, we can see $G(\boldsymbol{\mu}_0^0) \ge F(\boldsymbol{\mu}_0^0)$. Hence we can conclude $\mu_{M+1}^0 \in [0, G(\boldsymbol{\mu}_0^0)]$ and $(\mu_1^0, \mu_2^0, ..., \mu_{M+1}^0) \in G_s$, i.e. $\mathcal{N} \subseteq G_s$.

In the following we will prove for any points $(\mu_1^0, \mu_2^0, ..., \mu_{M+1}^0)$ belong to G', there exists a non-negative $\mathbf{k}$ such that for any $(\mu_1, \mu_2, ..., \mu_{M+1}) \in G_s$, $\mu_{M+1} - \sum_{i=1}^{M} k_i\mu_i \le \mu_{M+1}^0 - \sum_{i=1}^{M} k_i\mu_i^0$ holds. A similar proof is given in [16], [19].

Assume $(\mu_0^0, \mu_1^0, ..., \mu_{M+1}^0)$ is a point on the G' surface. Since for any positive ϵ, point $(\mu_0^0, \mu_1^0, ..., \mu_{M+1}^0 + \epsilon) \notin G_s$, we can conclude $(\mu_1^0, ..., \mu_{M+1}^0)$ is a boundary point of G_s.

Since G_s is a convex set and $(\mu_0^0, \mu_1^0, ..., \mu_{M+1}^0)$ is a boundary point of G_s, there exists an $M+1$-dimensional hyperplane Π through this point such that Π contains only boundary points of G_s and G_s lies entirely on one sides of Π [16], [19]. Since Π is a hyperplane in $M+1$-dimensinal space through point $(\mu_1^0, ..., \mu_{M+1}^0)$, there exists k_i^0 $(i = 1, ..., M+1)$ such that the expression of Π can be written as

$$k_{M+1}^0 \mu_{M+1} - \sum_{i=1}^{M} k_i^0 \mu_i = k_{M+1}^0 \mu_{M+1}^0 - \sum_{i=0}^{M} k_i^0 \mu_i^0 \tag{70}$$

We have proved when $\mu_i^0 \neq 0$ $(i = 1, ..., M)$, $\mu_{M+1}^0 > 0$, i.e. there exists a $\mu'_{M+1} = \frac{\mu_{M+1}^0}{2}$ such that $\mu'_{M+1} \in (0, \mu_{M+1}^0)$. From the definition of G_s, it is easy to know $(\mu_1^0, ..., \mu_M^0, \mu'_{M+1})$ is also belongs to set G_s. Besides that, since $\mu_i^0 \in (0, 1)$ $(i = 1, ..., M)$ and $\mu'_{M+1} \in (0, \mu_{M+1}^0)$, we can see $(\mu_1^0, \mu_2^0, ..., \mu'_{M+1})$ is an inner point of G_s, i.e. point $(\mu_1^0, ..., \mu_M^0, \mu'_{M+1})$ is not contained by hyperplane Π. We can conclude

$$k_{M+1}^0 \mu'_{M+1} - \sum_{i=1}^{M} k_i^0 \mu_i^0 \neq k_{M+1}^0 \mu_{M+1}^0 - \sum_{i=0}^{M} k_i^0 \mu_i^0$$

$$\therefore k_{M+1}^0 \mu'_{M+1} \neq k_{M+1}^0 \mu_{M+1}^0$$

From above equation we can conclude $k_{M+1}^0 \neq 0$ and Π can be written as

$$\mu_{M+1} - \sum_{i=1}^{M} k_i \mu_i = \mu_{M+1}^0 - \sum_{i=0}^{M} k_i \mu_i^0 \tag{71}$$

where $k_i = \frac{k_i^0}{k_{M+1}}$.

Since G_s lies entirely on one side of (71), and since when $(\mu_1, ..., \mu_M, \mu_{M+1}) = (\mu_1^0, ..., \mu_M^0, 0)$, the left hand member of (71) is smaller than the right hand member, hence we can conclude there exists k_i $(i = 1, ..., M)$ such that

$$\mu_{M+1} - \sum_{i=1}^{M} k_i \mu_i \leq \mu_{M+1}^0 - \sum_{i=0}^{M} k_i \mu_i^0 \tag{72}$$

holds for all points belong to G_s.

Even though we can not acquire the value of k_i, we can still achieve its property by analysing Π and G_s. Now consider a point on the G' with coordinate $(\mu_1^0, ..., \mu_l^0 + \epsilon, ..., \mu_M^0, \mu'_{M+1})$, where $\epsilon > 0$. Since $\mu_l^0 \in (0, 1)$, there always exists a positive ϵ satisfying $\mu_l^0 + \epsilon \in (0, 1]$. Since both points $(\mu_1^0, ..., \mu_l^0, ..., \mu_M^0, \mu_{M+1}^0)$ and $(\mu_1^0, ..., \mu_l^0 + \epsilon, ..., \mu_M^0, \mu'_{M+1})$ lies on G', we can conclude $G([\mu_1^0, ..., \mu_l^0, ..., \mu_M^0]) = \mu_{M+1}^0$ and $G([\mu_1^0, ..., \mu_l^0 + \epsilon, ..., \mu_M^0]) = \mu'_{M+1}$. As we have proved $G(\boldsymbol{\mu}_0)$ is a non-decreasing function, it is easy to see $\mu'_{M+1} \geq \mu_{M+1}^0$.

Substitute point $(\mu_1^0, ..., \mu_i^0 + \epsilon, ..., \mu_M^0, \mu'_{M+1})$ into equation (72), we have

$$\mu'_{M+1} - k_l(\mu_l^0 + \epsilon) \leq \mu_{M+1}^0 - k_l \mu_l^0 \tag{73}$$

$$\mu'_{M+1} - \mu_{M+1}^0 \leq k_l \epsilon \tag{74}$$

$$\therefore k_l \geq \frac{\mu'_{M+1} - \mu^0_{M+1}}{\epsilon} \geq 0$$

In the same way, we can prove $k_i \geq 0$ ($i = 1, ..., M$), i.e. for any point $(\mu^0_1, \mu^0_2, ..., \mu^0_{M+1})$ belongs to G_s, there exists a non-negative $\mathbf{k}$ such that for any $(\mu_1, \mu_2, ..., \mu_{M+1}) \in G_s$, $\mu_{M+1} - \sum_{i=1}^{M} k_i\mu_i \leq \mu^0_{M+1} - \sum_{i=1}^{M} k_i\mu^0_i$ holds.

Finally we consider the optimal decision rule for a given $\boldsymbol{\mu}^1_0 \in (0, 1)$, by optimal decision rule we mean this decision rule can get the largest P_d under the constraint $\mathbf{P}_f \leq \boldsymbol{\mu}^1_0$. Assume δ^* be the optimal decision rule for $\boldsymbol{\mu}^1_0$, obviously we have $\mathbf{P}_f(\delta^*) \leq \boldsymbol{\mu}^1_0$ and $G(\boldsymbol{\mu}^1_0) = P_d(\delta^*)$. According to the definition of G', we know point $(\boldsymbol{\mu}^1_0, G(\boldsymbol{\mu}^1_0))$ is on hyper surface G'. As we proved, there exists non-negative $\mathbf{k}$ such that for any $(\mu_0, ..., \mu_{M+1}) \in G_s$, following equation holds

$$\mu_{M+1} - \sum_{i=1}^{M} k_i\mu_i \leq \mu^1_{M+1} - \sum_{i=0}^{M} k_i\mu^1_i \tag{75}$$

Since $P_{f_i}(\delta^*) \leq \mu^1_i$ (for $i = 1, ..., M$) and $k_i \geq 0$, we know $\mu^1_{M+1} - \sum_{i=0}^{M} k_i\mu^1_i \leq \mu_{M+1} - \sum_{i=1}^{M} k_iP_{f_i}(\delta^*)$. (75) can be written as

$$\mu_{M+1} - \sum_{i=1}^{M} k_i\mu_i \leq \mu^1_{M+1} - \sum_{i=0}^{M} k_i\mu^1_i \leq \mu^1_{M+1} - \sum_{i=0}^{M} k_iP_{f_i}(\delta^*) \tag{76}$$

$$\therefore \mu_{M+1} - \sum_{i=1}^{M} k_i\mu_i \leq \mu^1_{M+1} - \sum_{i=0}^{M} k_iP_{f_i}(\delta^*) \tag{77}$$

Substitute $\mu_{M+1} = P_d(\delta^*)$ into (77)

$$\mu_{M+1} - \sum_{i=1}^{M} k_i \mu_i \leq P_d(\delta^*) - \sum_{i=0}^{M} k_i P_{f_i}(\delta^*) \tag{78}$$

holds for any point $(\mu_0, ..., \mu_{M+1}) \in G_s$.

Since $\mathcal{N} \subseteq G_s$, (78) also hold for any point belongs to $\mathcal{N}$. According to the definition of $\mathcal{N}$, the expression of μ_i can be written as $\int_{\mathcal{S}} f_i(x)\mathrm{d}x$ and μ_{M+1} can be written as $\int_{\mathcal{S}} f_0(x)\mathrm{d}x$, where $\mathcal{S}$ can be any subset of domain $\mathcal{D}$. $P_{f_i}(\delta^*)$ can be written as$\int_{\mathcal{S}^*} f_i(x)\mathrm{d}x$ and $P_d(\delta^*)$ can be written as $\int_{\mathcal{S}^*} f_0(x)\mathrm{d}x$. Hence (78) can be written in following form:

$$\int_{\mathcal{S}} f_0(x)\mathrm{d}x - \sum_{i=1}^{M} k_i \int_{\mathcal{S}} f_i(x)\mathrm{d}x \leq \int_{\mathcal{S}^*} f_0(x)\mathrm{d}x - \sum_{i=1}^{M} k_i \int_{\mathcal{S}^*} f_i(x)\mathrm{d}x$$

$$\int_{\mathcal{S}} (f_0(x) - \sum_{i=1}^{M} k_i f_i(x))\mathrm{d}x \leq \int_{\mathcal{S}^*} (f_0(x) - \sum_{i=1}^{M} k_i f_i(x))\mathrm{d}x \tag{79}$$

(79) can be fulfilled only if $\mathcal{S}^*$ satisfies

$$x \in \mathcal{S}^* \quad \text{if} \quad f_0(x) - \sum_{i=1}^{M} k_i f_i(x) > 0$$

$$x \notin \mathcal{S}^* \quad \text{if} \quad f_0(x) - \sum_{i=1}^{M} k_i f_i(x) < 0$$

Hence we can conclude for any given $\boldsymbol{\mu_0} \in (0, 1)$, there exists a set of non-negative $\mathbf{k}$

such that the optimal decision rule can be written as

$$f_0(x) \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} \sum_{i=1}^{M} k_i f_i(x)\,.$$

Q.E.D.

D. Proof for MENP Test

MENP (i) is a direct conclusion from ENP Lemma, we will consider MENP (ii). Assume $\mathbf{a}^0$ satisfy

$$\max_{\mathbf{a}\in\mathcal{C}} F(\mathbf{a}) = F(\mathbf{a^0})$$

and $\mathbf{a^0}$ can be achieved by decision rule

$$\delta^0: \quad \frac{f_0(x)}{\sum_{i=1}^{M} k_i^0 f_i(x)} \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} 1$$

Assume δ^0 is not the optimal decision rule for (11). Let δ^* be the optimal decision rule for (11). Obviously we have $P_d(\delta^*) > P_d(\delta^0)$ and $\mathbf{P}_f(\delta^*) \leq \mathbf{c}$. According to Lemma 3, δ^* must satisfy

$$x \in H_0 \quad \text{when} \;\; f_0(x) > \sum_{i=1}^{M} k_i' f_i(x)$$

$$x \in \bar{H}_0 \quad \text{when} \;\; f_0(x) < \sum_{i=1}^{M} k_i' f_i(x)$$

a.e. on $\mathcal{D}$. Let δ' be a decision rule such that

$$\frac{f_0(x)}{\sum_{i=1}^{M} k_i' f_i(x)} \underset{\bar{H}_0}{\overset{H_0}{\gtrless}} 1$$

Since we do not consider random decision rule, we can see $P_d(\delta') = P_d(\delta^*)$ and $\mathbf{P}_f(\delta') = P_f(\delta^*)$. We have $\mathbf{P}_f(\delta') \in \mathcal{C}$ and $F(\mathbf{P}_f(\delta')) \leq F(\mathbf{a^0})$, i.e., $P_d(\delta') < P_d(\delta^0)$. Hence we have $P_d(\delta^*) < P_d(\delta^0)$. This is contradictory with our assumption.

E. Proof for Property 1

According to the definition, M_0 is the region achieved by ENP test with $k_i \geq 0 (i = 1, ..., M)$. The ENP decision rule is

$$\frac{\sum_{i=1}^{M} k_i f_i(x)}{f_0(x)} \underset{H_0}{\overset{\bar{H}_0}{\gtrless}} 1 \tag{80}$$

which can be written as

$$g(x) \underset{H_0}{\overset{\bar{H}_0}{\gtrless}} 1 \tag{81}$$

Since $g(x)$ is a monotonically increasing function with x, hence $g^{-1}(x)$ exists and (81) can be written as

$$x \underset{H_0}{\overset{\bar{H}_0}{\gtrless}} x_0 \tag{82}$$

where $x_0 = g^{-1}(1)$. Under decision rule (82), the expression of P_d, P_{f_1}, ..., P_{f_M} can be written as

$$
\begin{aligned}
P_d &= \text{Pf}(X \le x_0 | H_0) = F_0(x_0) \\
P_{f_1} &= \text{Pf}(X \le x_0 | H_1) = F_1(x_0) \\
&\cdots\cdots \\
P_{f_M} &= \text{Pf}(X \le x_0 | H_M) = F_M(x_0)
\end{aligned}
\tag{83}
$$

where F_0, F_1, ..., F_M are the CDFs of X under H_0, H_1, ..., H_M. From (83), P_{f_1} determines x_0 that in turn determines P_{f_2}, ..., P_{f_M}, P_d. Hence for a given P_{f_1}, there is only one corresponding $P_{f_i}(i = 1, ..., M)$. Hence the ENP ROC with $k_i \ge 0(i = 1, ..., M)$ is a curve in this case.

From (83), (11) can be written as

$$
\begin{aligned}
\max \quad & P_d = F_0(x_0) \\
\text{s.t.} \quad & P_{f_i} = F_i(x_0) \le c_i \quad i = 1, 2, ..., M
\end{aligned}
\tag{84}
$$

Since $P_d = F_0(x_0)$ is an increasing function of x_0, (84) can be represented in following form:

$$
\begin{aligned}
\max \quad & x_0 \\
\text{s.t.} \quad & x_0 \le F_i^{-1}(c_i) \quad i = 1, 2, ..., M
\end{aligned}
\tag{85}
$$

Hence the expression of x_0 can be written as

$$x_0 = \min(F_1^{-1}(c_1), F_2^{-1}(c_2), ..., F_M^{-1}(c_M))\,. \tag{86}$$

F. Proof for Conclusion 1

First we prove $F\mathbf{a}$ is a non-decreasing function of $\mathbf{a}$ when $\mathbf{a} \in \alpha^+$.

Since when $\mathbf{a} \in \alpha^+$ we have

$$\begin{aligned} \max_{\mathbf{a}\in\alpha^+} \quad & P_d = F(\mathbf{a}) \\ \text{s.t.} \quad & \mathbf{P}_f = \mathbf{a} \end{aligned} \tag{87}$$

According to ENP Lemma, $F(\mathbf{a})$ also satisfies

$$\begin{aligned} \max_{\mathbf{a}\in\alpha^+} \quad & P_d = F(\mathbf{a}) \\ \text{s.t.} \quad & \mathbf{P}_f \leq \mathbf{a} \end{aligned} \tag{88}$$

In (88) when one of $a_i, \quad i = 1, 2, ..., M$ increases and other a_i remain the same, the constraints become more relax, hence P_d will not decrease. Therefore $F(\mathbf{a})$ is a non-decreasing function of $\mathbf{a}$ when $\mathbf{a} \in \alpha^+$. Consider points in region N_1. Since $(c_1, c_2) \notin N_0$, we use MENP (ii) to achieve P_d. As it is shown in Figure 2, A is a point in N_1 with coordinate (x, y). According to MENP (ii), $\mathcal{C}$ is the set of points lie in region enclosed by curve L_1 L_2 and line $c_1 = x$. For all points belongs to $\mathcal{C}$, point B has the largest c_1 and c_2 value. As we proved function

$F(c_1, c_2)$ is non-decreasing, point B has the largest P_d among all points belong to region $\mathcal{C}$. Hence we proved point A and point B have the same decision rule and same P_d. Since point in region N_1 with $c_1 = x$ has the same decision rule and P_d with point B, we can conclude: for points belong to region N_1 or curve L_1, if they have the same c_1 value, they have the same decision rule and same P_d.

In the same way, we can prove: For points belong to region N_2 or curve L_2, if they have the same c_2 value, they have the same decision rule and same P_d.

REFERENCES

[1] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *Selected Areas in Communications, IEEE Journal on*, vol. 23, pp. 201 – 220, feb. 2005.

[2] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, “Breaking spectrum gridlock with cognitive radios: An information theoretic perspective,” *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, 2009.

[3] M. M. Buddhikot, “Understanding dynamic spectrum access: Models, taxonomy and challenges,” in *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pp. 649–663, IEEE, 2007.

[4] R. Tandra, S. M. Mishra, and A. Sahai, “What is a spectrum hole and what does it take to recognize one?,” *Proceedings of the IEEE*, vol. 97, no. 5, pp. 824–848, 2009.

[5] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, “Spectrum sensing for cognitive radio: State-of-the-art and recent advances,” *Signal Processing Magazine, IEEE*, vol. 29, no. 3, pp. 101–116, 2012.

[6] D. Cabric, S. M. Mishra, and R. W. Brodersen, “Implementation issues in spectrum sensing for cognitive radios,” in *Signals, systems and computers, 2004. Conference record of the thirty-eighth Asilomar conference on*, vol. 1, pp. 772–776, IEEE, 2004.

[7] E. Axell and E. G. Larsson, “Optimal and sub-optimal spectrum sensing of ofdm signals in known and unknown noise variance,” *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 2, pp. 290–304, 2011.

[8] G. Ganesan and Y. Li, “Cooperative spectrum sensing in cognitive radio networks,” in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pp. 137–143, IEEE, 2005.

[9] Z. Tian and G. B. Giannakis, “Compressed sensing for wideband cognitive radios,” in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 4, pp. IV–1357, IEEE, 2007.

[10] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, “Wideband spectrum sensing for cognitive radio networks: a survey,” *Wireless Communications, IEEE*, vol. 20, no. 2, pp. 74–81, 2013.

[11] H. Sun, W.-Y. Chiu, J. Jiang, A. Nallanathan, and H. V. Poor, "Wideband spectrum sensing with sub-nyquist sampling in cognitive radios," *arXiv preprint arXiv:1302.1847*, 2013.

[12] H. V. Poor, *An introduction to signal detection and estimation*. Springer, 1994.

[13] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, p. 2, 2010.

[14] S. J. Shellhammer *et al.*, "Spectrum sensing in ieee 802.22," *IAPR Wksp. Cognitive Info. Processing*, pp. 9–10, 2008.

[15] Q. Zhang, P. K. Varshney, and Y. Zhu, "Design of extended neyman-pearson hypothesis tests," in *AeroSense'99*, pp. 378–389, International Society for Optics and Photonics, 1999.

[16] E. L. Lehmann, *Testing statistical hypotheses*. Wiley, 1959.

[17] J. Neyman and E. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933.

[18] A. Wald, "Contributions to the theory of statistical estimation and testing hypotheses," *The Annals of Mathematical Statistics*, pp. 299–326, 1939.

[19] G. Dantzig and A. Wald, "On the fundamental lemma of neyman and pearson," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 87–93, 1951.

[20] B. L. Mark and W. Turin, *Probability, Random Processes, and Statistical Analysis*. Cambridge University Press Textbooks, 2011.

[21] Q. Zhang, "Efficient computational algorithms for the design of distributed detection networks," 2000.

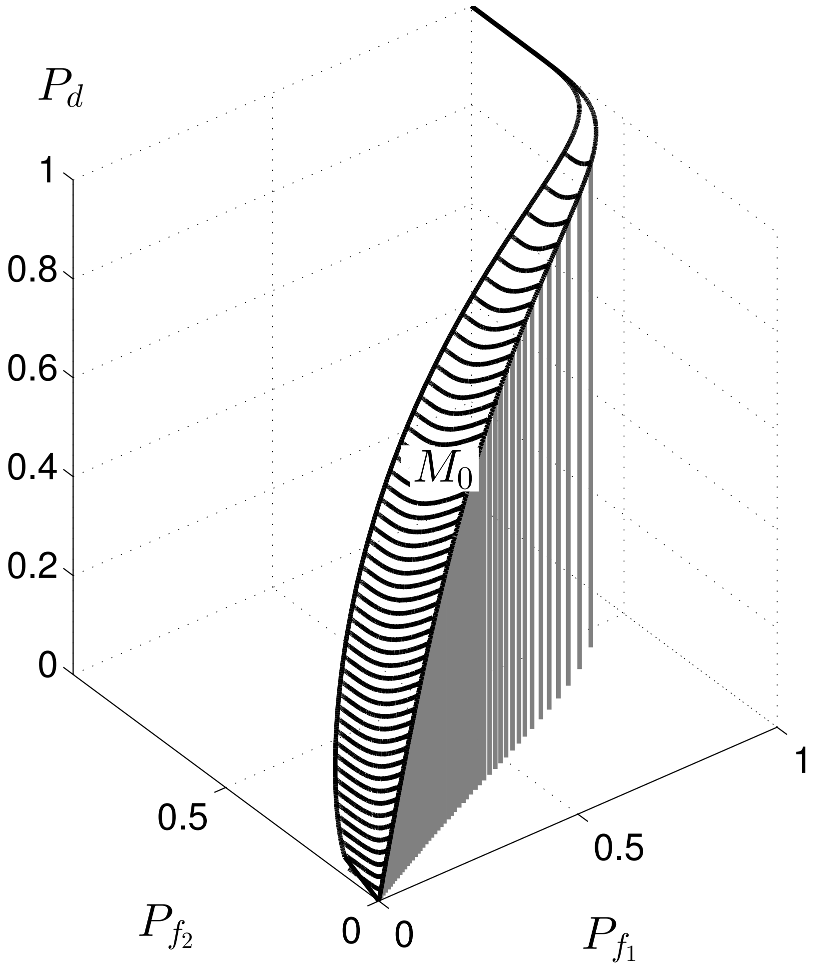

Fig. 1. Region that can be achieved by Neyman Pearson testing with $k_i \geq 0 (i = 1, ..., M)$.

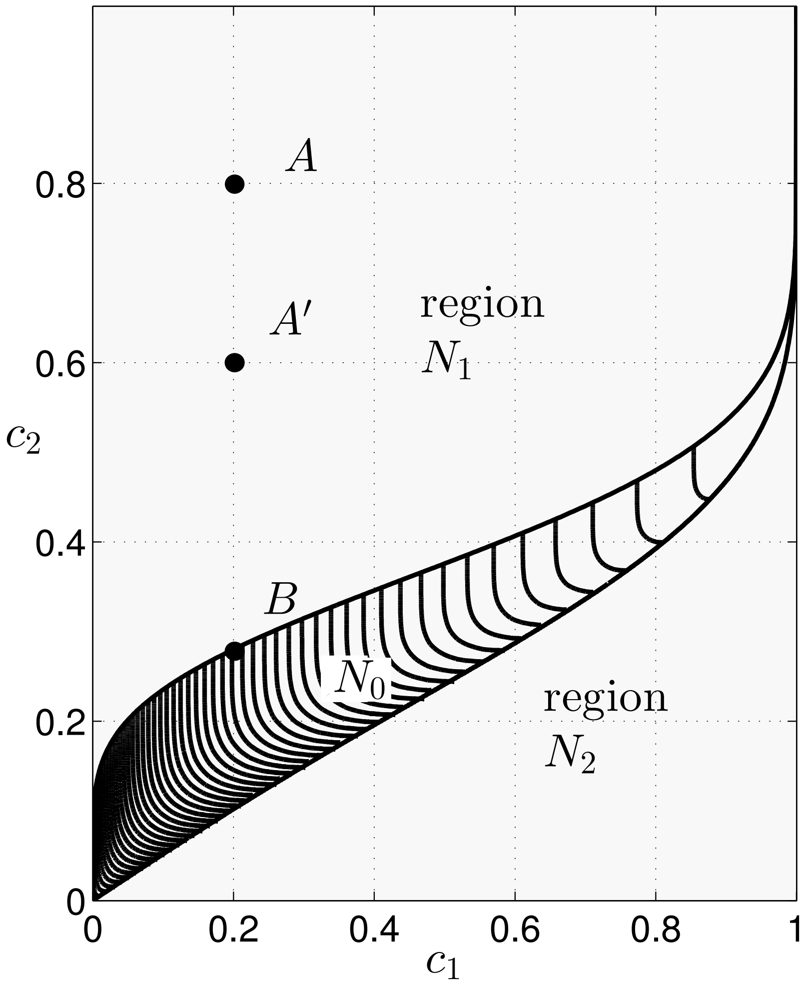

Fig. 2. Region that can be achieved by Neyman Pearson testing with $k_i \geq 0 (i = 1, ..., M)$.

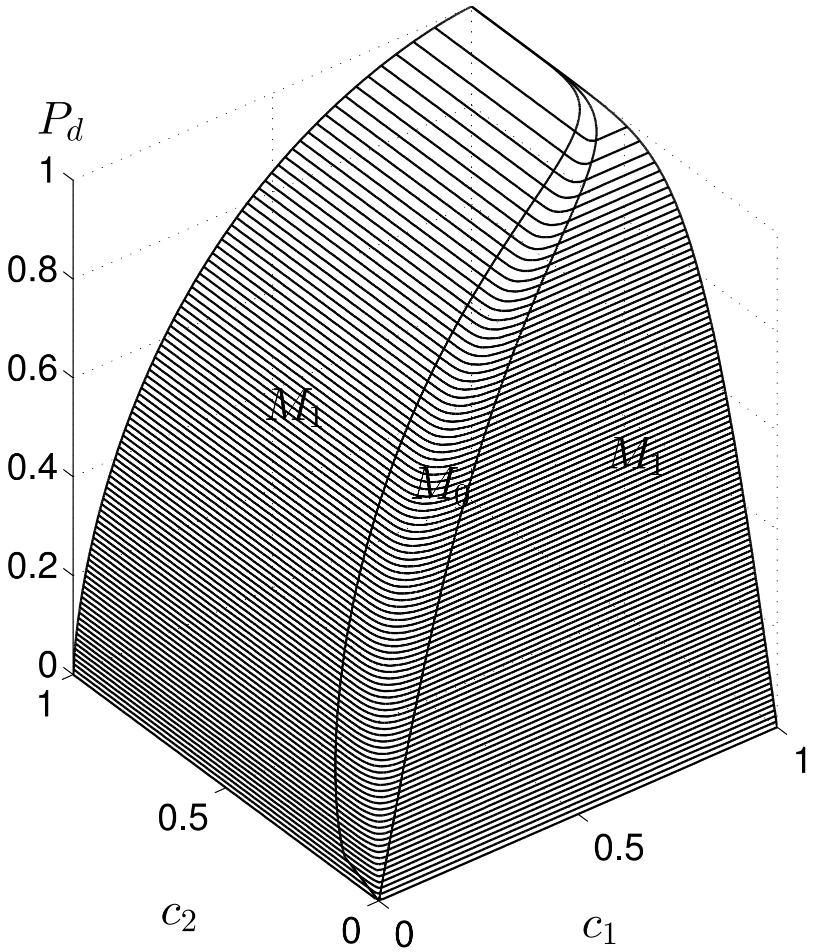

Fig. 3. The M-ROC surface for Gaussian Hypotheses.

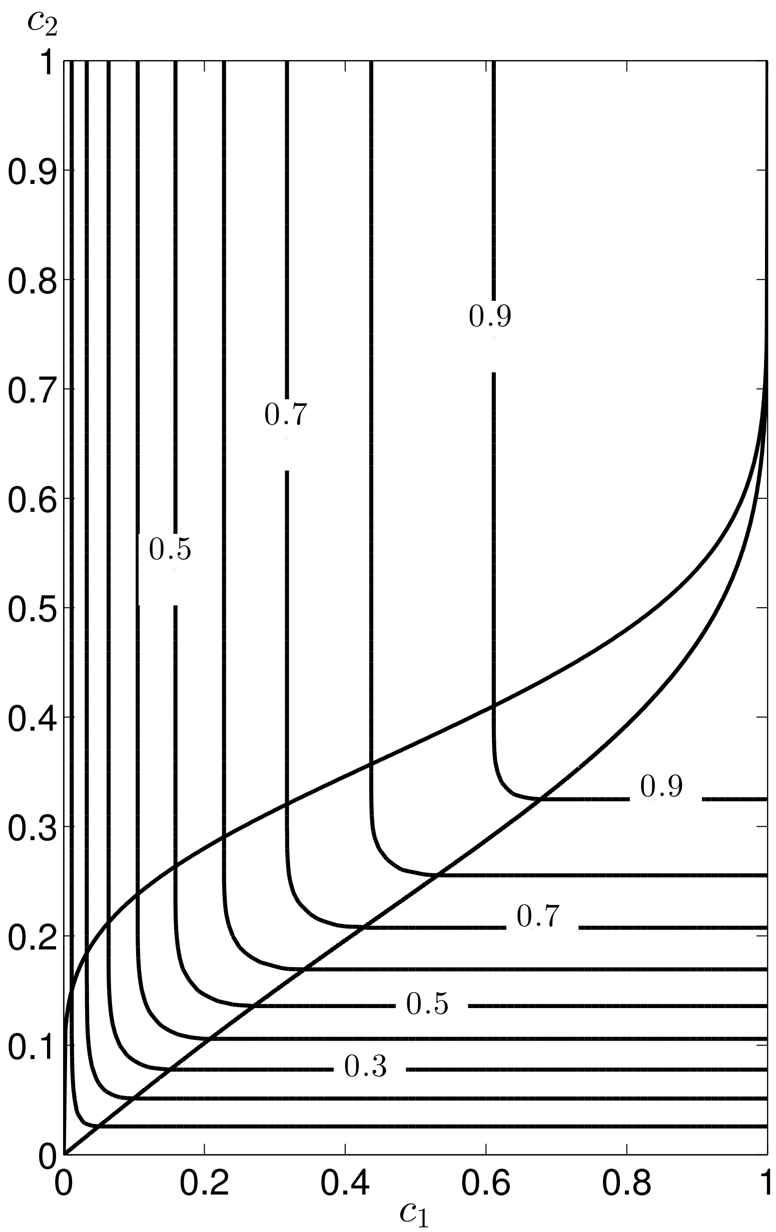

Fig. 4. Contour for M-ROC surface.

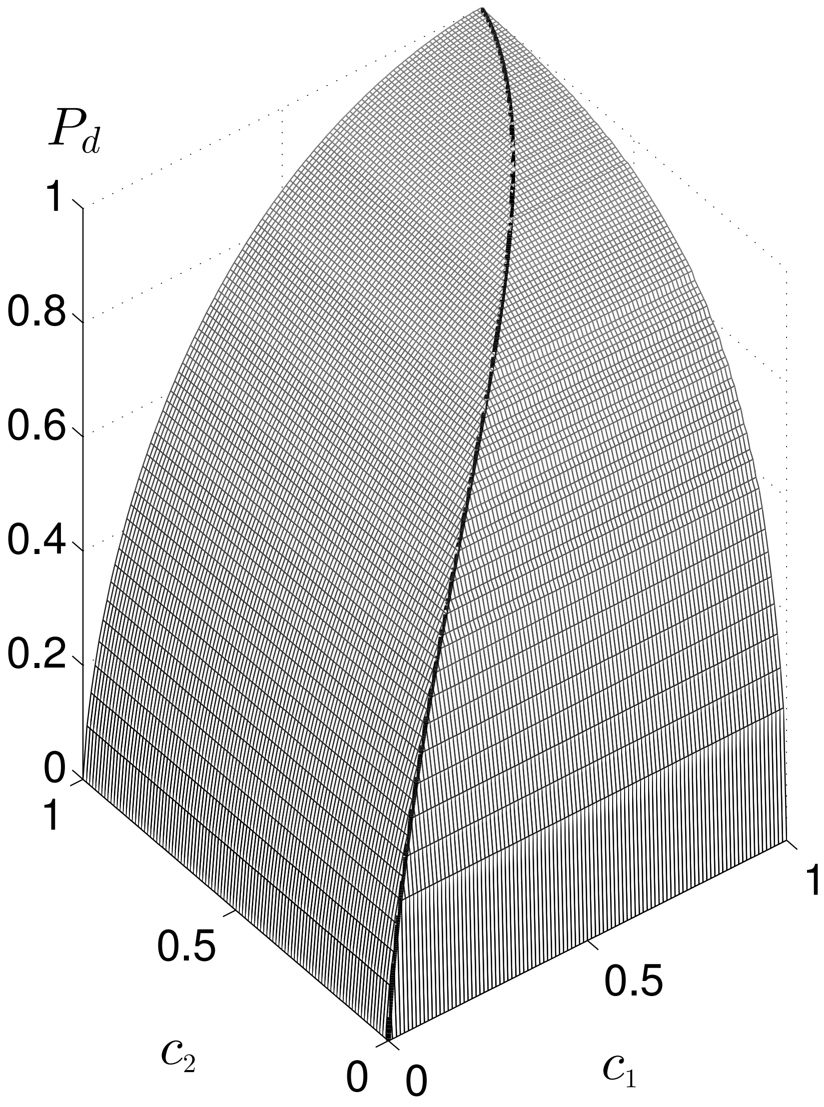

Fig. 5. The M-ROC for the Chi-square example.

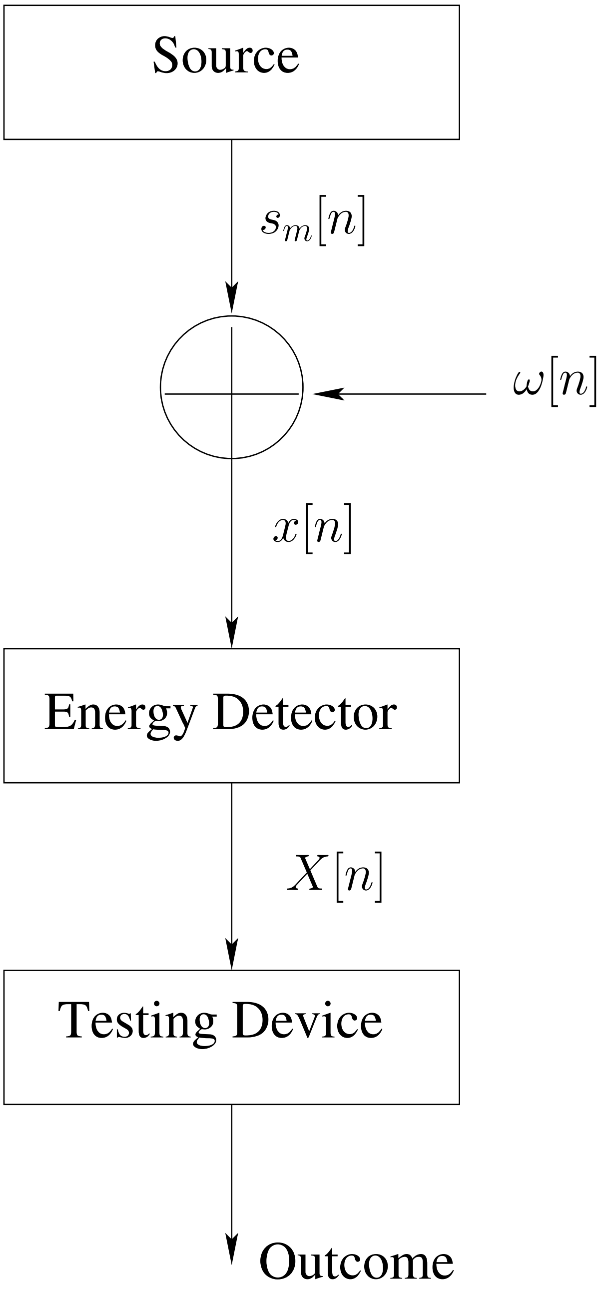

Fig. 6. The Block Diagram for Energy Based Spectrum Sensing.

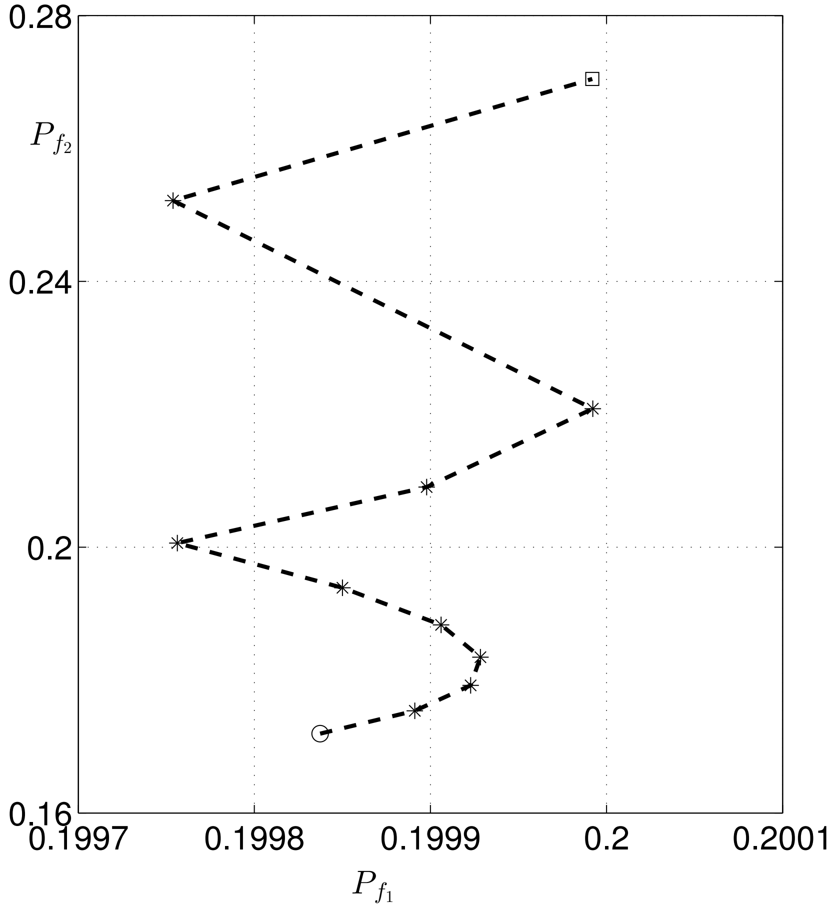

Fig. 7. Change of P_{f_1} and P_{f_2} after each iteration.

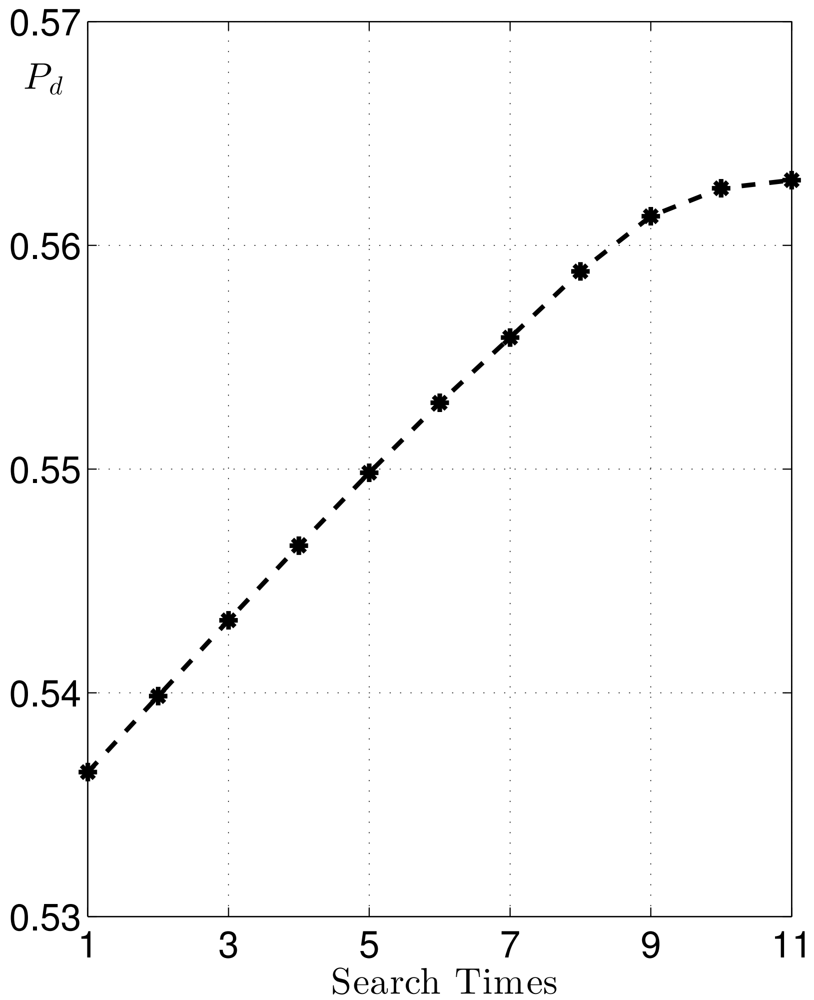

Fig. 8. Change of P_d after each iteration.

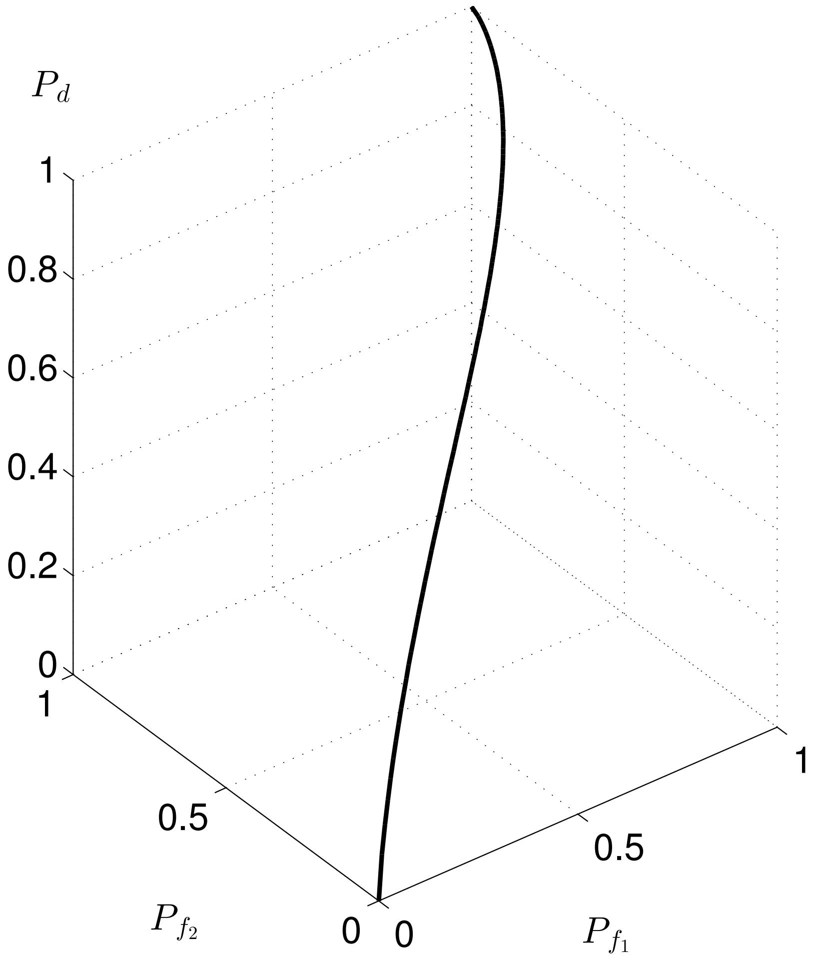

Fig. 9. ROC for minimax decision rule.