Sensing and Recognition for Multiple Primary Power Level Scenario with Noise Uncertainty

Chen Qian, Feifei Gao, Han Qian, and Tao Zhang

***Abstract*—In this paper, we consider the spectrum sensing in a new cognitive radio (CR) scenario where the primary user (PU) could possibly transmit with multiple power levels and the noise power at the secondary user (SU) side is uncertain. The target for SU is not only to detect the presence of PU, as did in the conventional "on-off" CR scenario, but also to identify the PU's transmitting power level. With the aid of the generalized likelihood ratio test (GLRT), we prove that the energy detection (ED) is still the optimal detection technique, and then derive the closed-form decision thresholds as well as the corresponding analytical performance. Interestingly, a unique phenomenon called power ambiguity arises in this new CR scenario that demands for careful investigation. Moreover, the signal-to-noise ratio (SNR) wall in the conventional binary CR also exists here and separates all power levels into two groups: one power group leads to robust recognition while the other does not. To make the discussion complete, we further design a cooperative spectrum sensing scheme and also derive the corresponding analytical performance expression. In the end, we validate our study through various numerical results.**

***Index Terms*—Spectrum sensing, multiple primary transmit power (MPTP), noise uncertainty, SNR wall, majority rule.**

I. Introduction

Cognitive radio (CR) is a promising technology to solve the spectrum scarcity problem and to cope with the inefficient spectrum usage [1]. A secondary (unlicensed) user (SU) is allowed to access the spectrum licensed to a primary user (PU) when PU is absent or when the interference to PU is below an acceptable threshold. Naturally, spectrum sensing that can detect the vacant spectrum becomes the key technology of CR [2]. The most commonly used spectrum sensing techniques include matched filter detection [3], [4], energy detection [5], and cyclostationary detection [6]. Matched filter detection is the optimal when PU's signals are perfectly known to SU; Energy detection requires least prior information about the PU's signal and is widely studied because of its lower complexity; Cyclostationary detection requires primary signal to possess cyclic features but is of high complexity and needs a large number of samples to formulate the cyclostationarity behavior.

One important assumption of these sensing methods is that PU is either idle or transmitting with a constant power, which makes the task of spectrum sensing as detecting the binary status of PU. However, it is readily seen from many current standards, e.g., IEEE 802.11 series, GSM, and the future standards, e.g., LTE, LTE-A that the licensed users could and should be working under different transmit power levels in order to cope with different situations. The spectrum sensing problem when PU has multiple transmit power levels (MPTP) was recently investigated in [20], where both the PU's on-off status as well as the PU's power levels could be recognized. It was also demonstrated in [15] that by detecting the multiple power levels of PU, SU could improve the power allocation strategy and achieve much higher throughput than mistakenly assuming binary status of PU.

On the other hand, many conventional sensing works also assume perfect knowledge of noise variance. However, in a practical system the noise variance cannot be perfectly measured due to initial calibration error, temperature variation, changes in low noise amplifier gain by thermal variation, and interference, etc [26]. Noise uncertainty has negative effects on the precision of spectrum sensing and sometimes even fails the sensing. Such influence has been studied in several works, including [8], [10], [11], [16], [17]. In [8], the authors presented the fundamental limits on energy detection in low signal-to-noise ratio (SNR) under noise uncertainty. Following this work, [9] designed energy detectors based on various types of noise uncertainty models, and [10], [11] proposed several spectrum sensing methods using the generalized likelihood ratio test (GLRT) when signals, noise and channels contain uncertainty. Especially, Tandra and Sahai [26] found that when the signal-to-noise ratio (SNR) is below a certain value, the performance of detection cannot be improved by increasing the sensing time. This interesting phenomena is well known as *SNR wall* and has attracted much attention [18], [19].

In this paper, we study the spectrum sensing in MPTP scenario considering the noise uncertainty at SU side. With multiple hypotheses testing and GLRT, we first prove that the energy detection is still the optimal for both PU detection and power level recognition. However, we point out that the uncertain noise power would bring a new and unique phenomenon in MPTP scenario, named as *power ambiguity*. We carefully investigate the effect of power ambiguity and derive the closed-form solutions for the decision region of each power level. The corresponding closed-form decision probabilities are also computed. Similar to binary sensing [26], SNR wall effect also exists for MPTP scenario but with a modified definition. We further derive the exact position of SNR wall. To make the proposed study complete, we design a cooperative sensing scheme based on the majority law to combat the power ambiguity effect. The closed form

C. Qian is with Beijing University of Posts and Telecommunicaitons, Beijing, P. R. China (email: qianchen94era@bupt.edu.cn)

F. Gao, H. Qian, and Tao Zhang are with Department of Automation, Tsinghua University, State Key Lab of Intelligent Technologies and Systems, Tsinghua National Laboratory for Information Science and Technology (TNList) Beijing, P. R. China (email: feifeigao@ieee.org).

The conference version of the paper has been accepted as [32].

expression of the decision probability for cooperative sensing is also obtained. Finally, simulating results are provided to corroborate the proposed studies.

The rest of this paper is organized as follows. In Section II, we describe the CR model in MPTP scenario and formulate the spectrum sensing problem. In Section III, we design the sensing and recognition algorithm and derive the decision regions as well as the decision probability. In the same section, the power ambiguity and the SNR wall phenomena are investigated. In Section IV, we propose a cooperative sensing scheme based on the majority law and derived its analytical performance expression. Simulation results are provided in Section V and conclusions are made in Section VI.

II. System Descriptions

A. Signal Model

Consider a simple CR network with a pair of PU and SU, where PU could either be absent or operating under any power level $P_k, k \in \{1, 2, \ldots, M\}$ satisfying $P_{k+1} > P_k > 0$. To unify the notation, let us define $P_0 = 0$ as the power when PU is absent. The received signal at SU is given by

$$\mathcal{H}_k : x_n = \sqrt{P_k}\sqrt{\gamma}e^{j\phi}s_n + \omega_n, k = 0, \cdots, M, \quad (1)$$

where $\mathcal{H}_k$ denotes the hypothesis that PU is operating under power level P_k; s_n denotes the n-th sample transmitted from PU which follows circularly symmetric complex Gaussian (CSCG) distribution with zero mean and unit variance; $\sqrt{\gamma}$ is the channel gain and ϕ is the channel phase; ω_n denotes the additive white Gaussian noise with zero mean and variance σ^2.

Hence, x_n also follows CSCG distribution, i.e.,

$$x_n|\mathcal{H}_k \sim \mathcal{CN}(0, P_k\gamma + \sigma^2), \quad \forall \mathcal{H}_k. \quad (2)$$

Define $\Pr(\mathcal{H}_k)$ as the prior probability of the state $\mathcal{H}_k$. Then the presence state of PU is denoted as $\mathcal{H}_{\text{on}} = \bigcup_{k=1}^{M} \mathcal{H}_k$, and the prior probability of $\mathcal{H}_{\text{on}}$ is $\Pr(\mathcal{H}_{\text{on}}) = \sum_{k=1}^{M} \Pr(\mathcal{H}_k)$. The absence state of PU can also be defined $\mathcal{H}_{\text{off}} = \mathcal{H}_0$ and has the prior probability $\Pr(\mathcal{H}_{\text{off}}) = \Pr(\mathcal{H}_0)$.

Similar to [20], we here make the following assumptions:

1) SU has knowledge power levels $\{P_0, P_1, \cdots, P_M\}$ as they are normally the deterministic values regulated by the standards.
2) We assume that the channel gain $\sqrt{\gamma}$ is known at SU while the phase ϕ can be unknown.[1]

B. Noise Uncertainty

Noise is an aggregation of various sources, among which four factors mainly contribute to the noise power uncertainty [21]: initial calibration error, temperature variation, changes in low noise amplifier gain due to the thermal variation, and interferers. The uncertainty caused by calibration error and thermal variation can be overcome by noise power estimation because thermal changes slowly. Indeed this part of the noise power is typically stationary for a few minutes [22]. However, the interference from other SUs or other opportunistic systems changes too fast to be tracked with periodical estimations. This dynamic background RF energy causes noise uncertainty in CR systems, i.e., σ^2 in system (1) is unknown.

[1]Traditional spectrum sensing algorithms [18], [19], [26], [27], [30] assume AWGN channel, which is the equivalent to assuming $\sqrt{\gamma} = 1$. Nevertheless, the recent result in [31] demonstrated that the channel gain knowledge can be released from machine learning technique.

III. Sensing and Recognition Algorithms

Suppose that SU receives a total number of N samples during the sensing period, and define $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T$, $\mathbf{s} = [s_1, s_2, \cdots, s_N]^T$, $\mathbf{w} = [w_1, w_2, \cdots, w_N]^T$. Due to the independence between $\mathbf{s}$ and $\mathbf{w}$, the probability density function (pdf) of $\mathbf{x}$ under $\mathcal{H}_k$ is given by

$$p(\mathbf{x}|\mathcal{H}_k) = \frac{1}{[\pi(\sigma^2 + P_k\gamma)]^N} \exp\left(-\frac{\|\mathbf{x}\|^2}{\sigma^2 + P_k\gamma}\right). \quad (3)$$

The unknown noise variance σ^2 leads to an employment of composite hypothesis test. Let us first apply the generalized likelihood ratio test (GLRT) [24] and obtain the maximum likelihood estimate (MLE) of σ^2 under $\mathcal{H}_k$ as

$$\begin{aligned}\hat{\sigma}_k^2 &= \arg\max_{\sigma^2} p(\mathbf{x}|\mathcal{H}_k, \sigma^2) \\ &= \arg\max_{\sigma^2} \frac{1}{[\pi(\sigma^2 + P_k\gamma)]^N} \exp\left(-\frac{\|\mathbf{x}\|^2}{\sigma^2 + P_k\gamma}\right). \quad (4)\end{aligned}$$

To find the most possible power level of PU, we use GLRT to compare each hypothesis pair $(\mathcal{H}_i, \mathcal{H}_j), \forall i, j \in \{0, 1, \cdots, M\}$

$$\xi_{i,j}(\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{H}_i, \hat{\sigma}_i)}{p(\mathbf{x}|\mathcal{H}_j, \hat{\sigma}_j)} \underset{\mathcal{H}_j}{\overset{\mathcal{H}_i}{\gtrless}} 1, \quad (5)$$

and the decision of the PU's power level should be

$$\hat{k} = \arg\max_{i \in \{0,1,\cdots,M\}} p(\mathbf{x}|\mathcal{H}_k, \hat{\sigma}_k^2). \quad (6)$$

In the worst case, SU has no knowledge of noise variance and the MLE of σ^2 under each $\mathcal{H}_k$ is derived from (4) as

$$\hat{\sigma}_k^2 = (T(\mathbf{x}) - P_k\gamma)^+, \quad (7)$$

where $(\cdot)^+$ denotes $\max\{0, \cdot\}$, and $T(\mathbf{x}) \triangleq \|\mathbf{x}\|^2/N$ represents the mean energy of the received samples. Without loss of generality, assume $T(\mathbf{x})$ stays between the value of $P_m\gamma$ and $P_{m+1}\gamma$ for a certain m, i.e., $P_0\gamma < \cdots < P_m\gamma < T(\mathbf{x}) < P_{m+1}\gamma < \cdots < P_M\gamma$, for one test. Then, the estimation (7) can be expanded as

$$\hat{\sigma}_k^2 = \begin{cases} T(\mathbf{x}) - P_k\gamma, & k = 0, 1, \ldots, m \\ 0, & k = m+1, \ldots, M, \end{cases}$$

and the corresponding pdf is

$$p(\mathbf{x}|\mathcal{H}_k, \hat{\sigma}_k^2) = \begin{cases} \dfrac{1}{(\pi T(\mathbf{x}))^N} \exp(-N), & 0 \le k \le m \\ \dfrac{1}{(\pi P_k\gamma)^N} \exp\left(-\dfrac{N}{P_k\gamma} T(\mathbf{x})\right), & m+1 \le k \le M. \end{cases} \quad (8)$$

It is seen from (8) that all the $m+1$ hypotheses $\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_m$ share the same likelihood pdf. Moreover, since $T(\mathbf{x}) < P_{m+1}\gamma$, it can be readily check that $p(\mathbf{x}|\mathcal{H}_0, \hat{\sigma}_0^2) = \ldots = p(\mathbf{x}|\mathcal{H}_m, \hat{\sigma}_m^2)$ are the maximum pdf

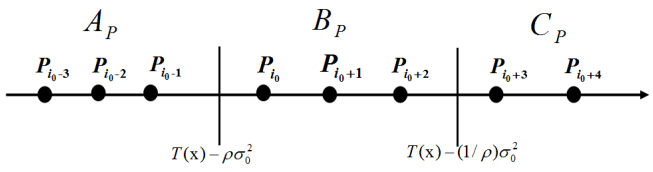

Fig. 1. Distribution of P_k for a given $T(\mathbf{x})$.

among all $p(\mathbf{x}|\mathcal{H}_k, \hat{\sigma}_k^2), k \in \{0, 1, \cdots, M\}$. Hence, the detection decision contains $m+1$ hypotheses.

Definition 1: If the decision include more than one hypothesises, we call this phenomenon *power ambiguity*.

Remark 1: Power ambiguity is a unique phenomenon for MPTP scenario with unknown noise variance while does not exist for the conventional binary spectrum sensing. Interestingly, another unique phenomenon for MPTP scenario with known noise variance is identified as *power mask* [20] in which some power levels hide below other power levels. Interestingly, power mask effect does not exist when the noise variance is unknown while power ambiguity effect does not exist when the noise variance is known.

Practically, a stationary noise power can be estimated by taking a large number of samples, while some residual uncertainty caused by interference still exists. For this reason, one may treat the noise power as ranging from a lower bound to an upper bound [11]. We then adopt the uncertainty model from [26] and assume the value of σ^2 belongs to $[(1/\rho)\sigma_0^2, \rho\sigma_0^2]$, where σ_0^2 denotes the nominal noise variance and $\rho > 1$ is a parameter that quantifies the size of uncertainty.

With the constraint $(1/\rho)\sigma_0^2 \leqslant \sigma^2 \leqslant \rho\sigma_0^2$, the MLE of σ^2 under hypothesis $\mathcal{H}_k$ is recomputed from (4) as

$$\hat{\sigma}_k^2 = \min\{\max\{(1/\rho)\sigma_0^2, T(\mathbf{x}) - P_k\gamma\}, \rho\sigma_0^2\}. \quad (9)$$

Let us then define the sets: $A_P \triangleq \{k|P_k\gamma < T(\mathbf{x}) - \rho\sigma_0^2\}$, $B_P \triangleq \{k|T(\mathbf{x}) - \rho\sigma_0^2 \leqslant P_k\gamma \leqslant T(\mathbf{x}) - \frac{1}{\rho}\sigma_0^2\}$, and $C_P \triangleq \{k|P_k\gamma > T(\mathbf{x}) - \frac{1}{\rho}\sigma_0^2\}$ for a given $T(\mathbf{x})$. Fig. 1 helps illustrate the division of the primary power levels into A_P, B_P and C_P with $\gamma = 1$.

Then the estimation from (9) can be explicitly expressed as

$$\hat{\sigma}_k^2 = \begin{cases} \rho\sigma_n^2, & k \in A_P \\ T(\mathbf{x}) - P_k\gamma, & k \in B_P \\ \frac{1}{\rho}\sigma_n^2, & k \in C_P. \end{cases} \quad (10)$$

With (10), we further obtain

$$p(\mathbf{x}|\mathcal{H}_k, \hat{\sigma}_k^2) = \begin{cases} \dfrac{1}{[\pi(P_k\gamma + \rho\sigma_n^2)]^N} \\ \qquad \times \exp\left[-\frac{NT(\mathbf{x})}{P_k\gamma + \rho\sigma_n^2}\right], & k \in A_P, \\ \dfrac{1}{(\pi T(\mathbf{x}))^N}\exp(-N), & k \in B_P, \\ \dfrac{1}{[\pi(P_k\gamma + \frac{1}{\rho}\sigma_0^2)]^N} \\ \qquad \times \exp\left[-\frac{NT(\mathbf{x})}{P_k\gamma + \frac{1}{\rho}\sigma_0^2}\right] & k \in C_P. \end{cases} \quad (11)$$

where $\mathcal{H}_k$'s, $k \in B_P$, share the same pdf.

To find the maximum likelihood probability $p(\mathbf{x}|\mathcal{H}_k, \hat{\sigma}_k^2)$, we compare the two hypothesis pairs $(\mathcal{H}_{k_A}, \mathcal{H}_{k_B})$ and $(\mathcal{H}_{k_B}, \mathcal{H}_{k_C})$ using the likelihood ratio (5) and obtain

$$\ln \xi_{k_B,k_A}(\mathbf{x}) = \frac{1}{N}\left[\ln \frac{\rho\sigma_0^2 + P_{k_A}\gamma}{T(\mathbf{x})} + \frac{T(\mathbf{x})}{\rho\sigma_0^2 + P_{k_A}\gamma} - 1\right], \quad (12)$$

$$\ln \xi_{k_B,k_C}(\mathbf{x}) = \frac{1}{N}\left[\ln \frac{\frac{1}{\rho}\sigma_0^2 + P_{k_C}\gamma}{T(\mathbf{x})} + \frac{T(\mathbf{x})}{\frac{1}{\rho}\sigma_0^2 + P_{k_C}\gamma} - 1\right], \quad (13)$$

where $k_A \in A_P$, $k_B \in B_P$, $k_C \in C_P$. Since the function $f(y) = y + \ln\frac{1}{y}$ has the minimum value $f(1) = 1$ in the interval $(0, +\infty)$, the inequalities $\xi_{k_B,k_A}(\mathbf{x}) > 1$ and $\xi_{k_B,k_C}(\mathbf{x}) > 1$ hold regardless of the value of P_{k_A} and P_{k_C}. Hence, $\mathcal{H}_k$ with $k \in B_P$ all have the maximum likelihood probability. Once again, we cannot discriminate among these power levels, and the power ambiguity happens when B_P contains more than one power levels.

Remark 2: In this paper, we consider the power recognition as fail if the power ambiguity happens. This consideration will add a strong restriction on the decision region as will be seen later.

A. Decision Region

Since all $\mathcal{H}_k$'s falling in set B_P have the largest pdf, our discussion will start from those power levels staying in B_P.

For a given $T(\mathbf{x})$, the sufficient and necessary condition for a specific hypothesis $\mathcal{H}_k$ to be contained in set B_P is $T(\mathbf{x}) - \rho\sigma_0^2 < P_k\gamma < T(\mathbf{x}) - \frac{1}{\rho}\sigma_0^2$, or equivalently $T(\mathbf{x}) \in (P_k\gamma + \sigma_0^2/\rho, P_k\gamma + \rho\sigma_0^2)$. This interval is named as the *potential decision region* (PDR) for hypothesis $\mathcal{H}_k$ (because the final decision has not been made yet). Let us then define PDR for all power level $P_k, k \in \{0, 1, \cdots, M\}$ as

$$\mathcal{R}_1(\mathcal{H}_k) = (P_k\gamma + \sigma_0^2/\rho, P_k\gamma + \rho\sigma_0^2) = (\phi_{lk}, \phi_{uk}), \quad (14)$$

where ϕ_{lk} and ϕ_{uk} are the corresponding items.

From (14), we know that the width of $\mathcal{R}_1(\mathcal{H}_k)$ is determined by ρ, i.e., the size of noise uncertainty. When ρ is small, $\mathcal{R}_1(\mathcal{H}_k)$'s are separated by certain gaps, as shown in Fig. 2(a). When ρ becomes larger, $\mathcal{R}_1(\mathcal{H}_k)$ becomes wider and the gap between neighbouring $\mathcal{R}_1(\mathcal{H}_k)$ gradually shrinks. When ρ is large enough, $\mathcal{R}_1(\mathcal{H}_k)$'s will overlap with each other as shown in Fig. 2(b) and Fig. 2(c).

Case I, $T(\mathbf{x})$ falling in overlapping areas between $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$, as shown by Fig. 3: In this case, $\mathcal{H}_k$ and $\mathcal{H}_{k+1}$ are contained in B_P, i.e. power ambiguity happens. Thus, the decision region of $\mathcal{H}_k$ under Case I is always $\emptyset$. The boundary of this overlapping area is computed as $(\phi_{l(k+1)}, \phi_{uk}) = (P_{k+1}\gamma + \frac{1}{\rho}\sigma_0^2, P_k\gamma + \rho\sigma_0^2)$. In general, if B_P consists of $P_k, P_{k+1}, \cdots, P_{k+m}$,[2] then $T(\mathbf{x})$ must fall into

$$(\phi_{l(k+m)}, \phi_{uk}) = (P_{k+m}\gamma + \sigma_0^2/\rho, P_k\gamma + \rho\sigma_0^2). \quad (15)$$

Lemma 1: If the power ambiguity between $\mathcal{H}_k$ and $\mathcal{H}_{k+m_1}(m_1 > 0)$ happens, then ambiguity between $\mathcal{H}_k$ and $\mathcal{H}_{k+m_2}(m_2 < m_1)$ must happen.

[2]The power ambiguity must only happens for consecutive power levels.

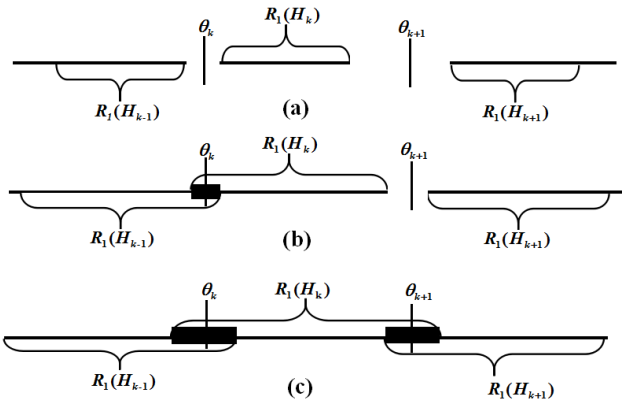

Fig. 2. Three different distributions of $\mathcal{R}_1(\mathcal{H}_{k-1})$, $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$.

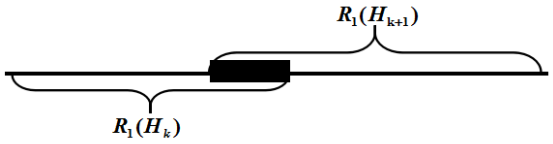

Fig. 3. A possible distribution of PDR for $\mathcal{H}_i$ and $\mathcal{H}_{i+1}$.

Proof: If the power ambiguity between $\mathcal{H}_k$ and $\mathcal{H}_{k+m_1}$ happens, then $T(\mathbf{x})$ falls into $(P_{k+m_1}\gamma+\frac{1}{\rho}\sigma_0^2, P_k\gamma+\rho\sigma_0^2)$. As $P_M > P_{M-1} > \cdots > P_0$, we have $P_{k+m_2}\gamma+\frac{1}{\rho}\sigma_0^2 < P_{k+m_1}\gamma+\frac{1}{\rho}\sigma_0^2$. Thus $T(\mathbf{x})$ also falls into $(P_{k+m_2}\gamma+\frac{1}{\rho}\sigma_0^2, P_k\gamma+\rho\sigma_0^2)$, causing power ambiguity between $\mathcal{H}_k$ and $\mathcal{H}_{k+m_2}$. ■

Case II, $T(\mathbf{x})$ falling in regions included in $\mathcal{R}_1(\mathcal{H}_k)$ that does not overlap with other PDR: In this case, P_k is the only power level contained in B_P and the decision at SU is made as P_k. Then, the region included in $\mathcal{R}_1(\mathcal{H}_k)$ but does not overlap with any other PDR is the decision region of $\mathcal{H}_k$. Assume $\mathcal{R}_1(\mathcal{H}_k)$ overlaps with $\mathcal{R}_1(\mathcal{H}_i)$, $i = m_1, \ldots m_2 (m_1 < k < m_2)$. Then, the decision region of $\mathcal{H}_k$ can be computed as

$$\mathcal{R}_1(\mathcal{H}_k) - \sum_{i=m_1, i\neq k}^{m_2} \mathcal{R}_1(\mathcal{H}_{k-1}) \tag{16}$$

$$= \Big(\max(P_k+\frac{\sigma_0^2}{\rho}, P_{k-1}+\rho\sigma_0^2, P_{k-2}+\rho\sigma_0^2, \cdots, P_{m_1}+\rho\sigma_0^2),$$
$$\min(P_k+\rho\sigma_0^2, P_{k+1}+\frac{\sigma_0^2}{\rho}, P_{k+2}+\frac{\sigma_0^2}{\rho}, P_{m_2}+\frac{\sigma_0^2}{\rho})\Big). \tag{17}$$

Since P_k's are in increasing order, (17) can be simplified to

$$\mathcal{R}_1(\mathcal{H}_k) - \sum_{i=m_1, i\neq k}^{m_2} \mathcal{R}_1(\mathcal{H}_{k-1})$$
$$= \Big(\max(P_k+\frac{\sigma_0^2}{\rho}, P_{k-1}+\rho\sigma_0^2), \min(P_k+\rho\sigma_0^2, P_{k+1}+\frac{\sigma_0^2}{\rho})\Big), \tag{18}$$

which says that overlapping with more PDR does not change the expression of decision region.

Remark 3: In Case II, if $\mathcal{R}_1(\mathcal{H}_k)$ does not overlap with both $\mathcal{R}_1(\mathcal{H}_{(k+1)})$ and $\mathcal{R}_1(\mathcal{H}_{(k-1)})$, then the decision region for $\mathcal{H}_k$ is simply $\mathcal{R}_1(\mathcal{H}_k)$.

Case III, $T(\mathbf{x})$ falling out of any PDR and stay in the blank region, as shown in Fig. 2(a) or Fig. 2(b): Then no power level is contained in B_P, i.e., no power level is contained in $(T(\mathbf{x})-\rho\sigma_0^2, T(\mathbf{x})-\frac{1}{\rho}\sigma_0^2)$. In this case, there must be an $m \in \{0, 1, \cdots, M-1\}$ satisfying $P_m\gamma < T(\mathbf{x})-\rho\sigma_0^2 < T(\mathbf{x})-\frac{1}{\rho}\sigma_0^2 < P_{m+1}\gamma$, or equivalently $P_m\gamma+\rho\sigma_0^2 < T(\mathbf{x}) < P_{m+1}\gamma+\frac{1}{\rho}\sigma_0^2$. Now that B_P is empty, the MLE of σ^2 is simplified to

$$\hat{\sigma}_k^2 = \begin{cases} \rho\sigma_0^2, & k = 0, 1, \ldots, m \quad (19a) \\ \frac{1}{\rho}\sigma_0^2, & k = m+1, \ldots, M, \quad (19b) \end{cases}$$

which upon substitution into the likelihood pdf yields

$$\begin{cases} p(\mathbf{x}|\mathcal{H}_k, \hat{\sigma}_k^2) = \frac{1}{[\pi(P_k\gamma+\rho\sigma_0^2)]^N} \exp\left(-\frac{NT(\mathbf{x})}{P_k\gamma+\rho\sigma_n^2}\right) \\ \qquad\qquad k = 0, 1, \ldots, m, \quad (20a) \\ \frac{1}{[\pi(P_k\gamma+\frac{1}{\rho}\sigma_n^2)]^N} \exp\left(-\frac{NT(\mathbf{x})}{P_k\gamma+\frac{1}{\rho}\sigma_n^2}\right) \\ \qquad\qquad k = m+1, \ldots, M. \quad (20b) \end{cases}$$

Considering the monotonicity of function $h(y) = \frac{1}{y^N}\exp(-Na/y)$ and the condition $P_0\gamma+\rho\sigma_n^2 < \ldots < P_m\gamma+\rho\sigma_n^2 < T(\mathbf{x}) < P_{m+1}\gamma+\frac{1}{\rho}\sigma_n^2 < \ldots < P_M\gamma+\frac{1}{\rho}\sigma_n^2$, (20a) has the maximum value when $k = m$, and (20b) has the maximum value when $k = m+1$. Then, we only need to compare the hypothesis pair $(\mathcal{H}_{m+1}, \mathcal{H}_m)$, whose likelihood ratio is given by

$$\frac{1}{N}\ln\xi_{m+1,m}(\mathbf{x}) = T(\mathbf{x})\frac{(\sigma_0^2/\rho+P_{m+1}\gamma)-(\rho\sigma_0^2+P_m\gamma)}{(\sigma_0^2/\rho+P_{m+1}\gamma)(\rho\sigma_0^2+P_m\gamma)} + \ln\frac{\rho\sigma_0^2+P_m\gamma}{\sigma_0^2/\rho+P_{m+1}\gamma}. \tag{21}$$

Hence we obtain the detection threshold as

$$T(\mathbf{x}) \underset{\mathcal{H}_m}{\overset{\mathcal{H}_{m+1}}{\gtrless}} \frac{(\sigma_0^2/\rho+P_{m+1}\gamma)(\rho\sigma_0^2+P_m\gamma)}{(\sigma_0^2/\rho+P_{m+1}\gamma)-(\rho\sigma_0^2+P_m\gamma)}\ln\frac{\sigma_0^2/\rho+P_{m+1}\gamma}{\rho\sigma_0^2+P_m\gamma}. \tag{22}$$

Let us define $\mathcal{R}_2(\mathcal{H}_k)$

$$\mathcal{R}_2(\mathcal{H}_k) = (\theta_k, \theta_{k+1}), \tag{23}$$

where

$$\theta_k = \begin{cases} 0, & k = 0 \\ \frac{(\sigma_0^2/\rho+P_k\gamma)(\rho\sigma_0^2+P_{k-1}\gamma)}{(\sigma_0^2/\rho+P_k\gamma)-(\rho\sigma_0^2+P_{k-1}\gamma)} & \\ \qquad \times \ln\frac{\sigma_0^2/\rho+P_k\gamma}{\rho\sigma_0^2+P_{k-1}\gamma}, & k = 1, 2, \cdots, M \\ +\infty, & k = M+1. \end{cases} \tag{24}$$

Before computing the decision region for $\mathcal{H}_k$ in Case III, it is necessary to check the relationship between $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_2(\mathcal{H}_k)$.

Lemma 2: If $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k-1})$ do not overlap, i.e., $(\sigma_0^2/\rho+P_k\gamma)-(\rho\sigma_0^2+P_{k-1}\gamma) > 0$, then there are $\theta_k < \phi_{lk}$ and $\theta_k > \phi_{u(k-1)}$. On the other side, if $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k-1})$ overlap, then there are $\theta_k > \phi_{lk}$ and $\theta_k < \phi_{u(k-1)}$.

Proof: We first investigate the situation where $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k-1})$ do not overlap, i.e. $(\sigma_0^2/\rho + P_k\gamma) - (\rho\sigma_0^2 + P_{k-1}\gamma) > 0$. It is known:

$$f_1(y) = \ln y + \frac{1}{y} > f_1(1) = 1, \forall y \neq 1 \tag{25}$$

$$f_2(y) = \ln y - y < f_2(1) = -1, \forall y \neq 1. \tag{26}$$

Substituting $y = \dfrac{P_k\gamma + \sigma_n^2/\rho}{P_{k-1}\gamma + \rho\sigma_n^2}$ into (25) and (26) yields $\theta_k > P_{k-1}\gamma + \rho\sigma_n^2$, $\theta_k < P_k\gamma + \sigma_n^2/\rho$ respectively, i.e. $\theta_k < \phi_{lk}$ and $\theta_k > \phi_{u(k-1)}$. By applying the same method on the situation where $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k-1})$ overlap, i.e. $(\sigma_0^2/\rho + P_k\gamma) - (\rho\sigma_0^2 + P_{k-1}\gamma) > 0$, we get $\theta_k > \phi_{lk}$ and $\theta_k < \phi_{u(k-1)}$. ∎

Lemma 3: The decision region for $\mathcal{H}_k$ in Case III is:

$$T(\mathbf{x}) \in (\theta_k, \max(\theta_k, \phi_{lk})] \bigcup [\min(\theta_{k+1}, \phi_{uk}), \theta_{k+1}]. \tag{27}$$

Proof: See Appendix A ∎

After obtaining the separate decision regions for Case I, Case II and Case III, the final decision region for $\mathcal{H}_k$ is obtained by summing the decision region of Case I, Case II and Case III:

$$\begin{aligned}
\mathcal{DR}(\mathcal{H}_k) =& (\theta_k, \max(\theta_k, \phi_{lk})] + [\min(\theta_{k+1}, \phi_{uk}), \theta_{k+1}] \\
&+ \big(\max(\phi_{lk}, \phi_{u(k-1)}), \min(\phi_{uk}, \phi_{l(k+1)})\big) \\
=& (\theta_k, \theta_{k+1}] - (\max(\theta_k, \phi_{lk}), \min(\theta_{k+1}, \phi_{uk})) \\
&+ \big(\max(\phi_{lk}, \phi_{u(k-1)}), \min(\phi_{uk}, \phi_{l(k+1)})\big)\big) \\
=& (\theta_k, \theta_{k+1}] - \big(\max(\theta_k, \phi_{lk}), \max(\phi_{lk}, \phi_{u(k-1)})\big] \\
&- \big[\min(\phi_{uk}, \phi_{l(k+1)}), \min(\theta_{k+1}, \phi_{uk})\big] \\
=& (\theta_k, \theta_{k+1}] - \big(\theta_k, \max(\theta_k, \phi_{u(k-1)})\big] \\
&- \big[\min(\theta_{k+1}, \phi_{l(k+1)}), \theta_{k+1}\big].
\end{aligned} \tag{28}$$

Remark 4: The decision region (28) may not be consecutive. When $T(\mathbf{x})$ falls out of any $\mathcal{DR}(\mathcal{H}_k)$, then power ambiguity happens and the recognition fails.

Moreover, the overall power recognition strategy is summarized in Algorithm 1.

Algorithm 1 Strategy for recognition in MPTP scenario

1) Compute the average energy $T(\mathbf{x}) \triangleq \dfrac{\|\mathbf{x}\|^2}{N}$, where $\mathbf{x}$ is the received signal.
2) Formulate the set B_P by checking $P_k\gamma \in (T(\mathbf{x}) - \rho\sigma_0^2, T(\mathbf{x}) - \frac{1}{\rho}\sigma_0^2)$. If B_P contains only one power level P_k, then the sensing ends and decision $\mathcal{H}_k$ is made. If two or more power levels satisfy the requirement, the recognition is claimed to fail. If there is no power level in B_P, then proceed to the next step.
3) If $T(\mathbf{x})$ belongs to $\mathcal{R}_2(\mathcal{H}_k)$, then the decision $\mathcal{H}_k$ is made.

B. Performance Analysis

In this subsection, we examine both the performance with and without power ambiguity. According to [27], the test statistic $T(\mathbf{x})$ under hypothesis $\mathcal{H}_j$ approximately follows a Gaussian distribution when N is relatively large. The corresponding mean and variance are computed as $\mu_t = P_j\gamma + \sigma^2$ and $\sigma_t^2 = \frac{1}{N}(P_j\gamma + \sigma^2)^2$, respectively. Thus, the pdf of $T(\mathbf{x})$ is given by

$$p(t|\mathcal{H}_j) = \frac{\sqrt{N}}{\sqrt{2\pi}(P_j\gamma + \sigma^2)} \exp\left(-\frac{N(t - P_j\gamma - \sigma^2)^2}{2(P_j\gamma + \sigma^2)^2}\right). \tag{29}$$

Let us define $\Pr(\mathcal{A}_{i_0,i_1}|\mathcal{H}_j)(i_0 < i_1)$ as the probability of power ambiguity happening between $\mathcal{H}_{i_0}, \mathcal{H}_{i_0+1}, \cdots, \mathcal{H}_{i_1}$. According to Lemma 1, $\Pr(\mathcal{A}_{i_0,i_1}|\mathcal{H}_j)$ can be computed as the probability of $T(\mathbf{x})$ falling into the overlapping area between $\mathcal{R}_1(\mathcal{H}_{i_0})$ and $\mathcal{R}_1(\mathcal{H}_{i_1})$, i.e.,

$$\begin{aligned}
&\Pr(\mathcal{A}_{i_0,i_1}|\mathcal{H}_j) = \Pr\{\phi_{li_1} < T(x) \leqslant \phi_{ui_0}|\mathcal{H}_j\} \\
&= \mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{li_1}}{P_j\gamma + \sigma_j^2} - 1\right)\right] - \mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{ui_0}}{P_j\gamma + \sigma_j^2} - 1\right)\right].
\end{aligned} \tag{30}$$

Similarly, define $\Pr(\mathcal{H}_i|\mathcal{H}_j)$ as the probability when the final decision is $\mathcal{H}_i$ while the actual primary transmitting power is P_j. From (28), we directly obtain

$$\begin{aligned}
&\Pr(\mathcal{H}_i|\mathcal{H}_j) = \Pr\{\theta_i < T(x) \leqslant \theta_{i+1}|\mathcal{H}_j\} \\
&\quad - \Pr\{\min(\theta_{i+1}, \phi_{l(i+1)}) < T(x) \leqslant \theta_{i+1}|\mathcal{H}_j\} \\
&\quad - \Pr\{\theta_i < T(x) \leqslant \max(\theta_i, \phi_{u(i-1)})|\mathcal{H}_j\} \\
&= \left\{\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_i}{P_j\gamma + \sigma_j^2} - 1\right)\right] - \mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_{i+1}}{P_j\gamma + \sigma_j^2} - 1\right)\right]\right\}^+ \\
&\quad - \left\{\mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{l(i+1)}}{P_j\gamma + \sigma_j^2} - 1\right)\right] - \mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_{i+1}}{P_j\gamma + \sigma_j^2} - 1\right)\right]\right\}^+ \\
&\quad - \left\{\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_i}{P_j\gamma + \sigma_j^2} - 1\right)\right] - \mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{u(i-1)}}{P_j\gamma + \sigma_j^2} - 1\right)\right]\right\}^+.
\end{aligned} \tag{31}$$

We can also define the detection probability and false alarm probability as in the conventional binary sensing case, which can be easily obtained from the summations of the corresponding $\Pr(\mathcal{H}_i|\mathcal{H}_j)$, i.e.,

$$\mathrm{P_d} = \frac{1}{\Pr(\mathcal{H}_{\text{on}})} \sum_{i=1}^{M} \sum_{j=1}^{M} \Pr(\mathcal{H}_i|\mathcal{H}_j) \Pr(P_j)$$

$$\mathrm{P_{fa}} = \sum_{i=1}^{M} \Pr(\mathcal{H}_i|\mathcal{H}_0).$$

For MPTP scenario, we may also introduce the discrimination probability as

$$\mathrm{P_{dis}} = \sum_{i=0}^{M} \Pr(\mathcal{H}_i|\mathcal{H}_i) \Pr(P_i) \tag{32}$$

to describe the performance of power level recognition. In Fig. 4, we provide two examples to verify the theoretical expression (32) with $N = 4000$. It is observed that the theoretical value well matches the numerical result.

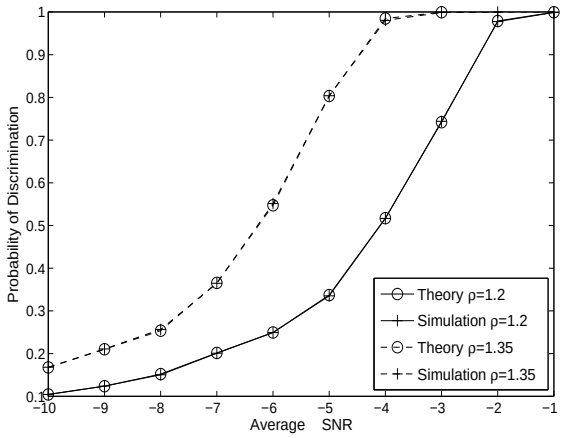

Fig. 4. Comparison between the theoretical value and the numerical one of discrimination probability.

C. SNR Wall

Different from the conventional SNR wall that is only related to false alarm and missing detection possibility [26], the SNR wall in MPTP scenario will be linked with multiple decision probabilities. Hence, we need to slightly modify the definition of SNR wall effect. Let us define the following two probabilities for each P_k:

$$\mathrm{Pr}_H^{(k)} = \mathcal{Q}\left[\sqrt{N}\left(\frac{\min(\theta_{k+1},\phi_{l(k+1)})}{P_k\gamma+\sigma^2}-1\right)\right], \tag{33}$$

$$\mathrm{Pr}_L^{(k)} = 1-\mathcal{Q}\left[\sqrt{N}\left(\frac{\max(\theta_k,\phi_{u(k-1)})}{P_k\gamma+\sigma^2}-1\right)\right], \tag{34}$$

where $\mathrm{Pr}_L^{(k)}$ represents the probabilities that the decision statistics $T(\mathbf{x})$ falls left to $\mathcal{DR}(\mathcal{H}_k)$, while $\mathrm{Pr}_H^{(k)}$ represents the probabilities that $T(\mathbf{x})$ falls right to $\mathcal{DR}(\mathcal{H}_k)$. Denote the maximal tolerable thresholds for $\mathrm{Pr}_L^{(k)}$ and $\mathrm{Pr}_H^{(k)}$ at SU as two constant values $\mathrm{Pr}_{LS}^{(k)}$ and $\mathrm{Pr}_{HS}^{(k)}$, respectively.

Definition 2: We claim the power recognition as *robust* only when $\mathrm{Pr}_L^{(k)}$ and $\mathrm{Pr}_H^{(k)}$ meet the requirement $\mathrm{Pr}_L^{(k)} \le \mathrm{Pr}_{LS}^{(k)}$ and $\mathrm{Pr}_H^{(k)} \le \mathrm{Pr}_{HS}^{(k)}$ for any $\sigma^2 \in (\frac{1}{\rho}\sigma_0^2, \rho\sigma_0^2)$.

Case I, $\mathcal{R}_1(\mathcal{H}_k)$ not overlapping with $\mathcal{R}_1(\mathcal{H}_{k-1})$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$: In this case, according to Lemma 2, we have $\min(\theta_{k+1},\phi_{l(k+1)}) = \theta_{k+1}$ and $\max(\theta_k,\phi_{u(k-1)}) = \theta_k$. Since σ^2 swings between $\frac{1}{\rho}\sigma_0^2$ and $\rho\sigma_0^2$, According to Definition 2 , the recognition is robust if

$$\mathrm{Pr}_{HS}^{(k)} \ge \max_{\sigma^2\in(\frac{1}{\rho}\sigma_0^2,\rho\sigma_0^2)}\left\{\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_{k+1}}{P_k\gamma+\sigma^2}-1\right)\right]\right\}, \tag{35}$$

$$\mathrm{Pr}_{LS}^{(k)} \ge \max_{\sigma^2\in(\frac{1}{\rho}\sigma_0^2,\rho\sigma_0^2)}\left\{1-\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_k}{P_k\gamma+\sigma^2}-1\right)\right]\right\}. \tag{36}$$

Since $\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_{k+1}}{P_k\gamma+\sigma^2}-1\right)\right]$ and $\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_k}{P_k\gamma+\sigma^2}-1\right)\right]$ are both monotonic increasing function of σ^2, (35) and (36) can be equivalently converted into

$$\mathrm{Pr}_{HS}^{(k)} \ge \left\{\mathcal{Q}\left[\sqrt{N}(\frac{\theta_{k+1}}{P_k\gamma+\rho\sigma_0^2}-1)\right]\right\}, \tag{37}$$

$$\mathrm{Pr}_{LS}^{(k)} \ge \left\{1-\mathcal{Q}\left[\sqrt{N}\left(\frac{\theta_k}{P_k\gamma+\frac{1}{\rho}\sigma_0^2}-1\right)\right]\right\}. \tag{38}$$

Obviously, the only variable that could be adjusted by SU to keep (37) and (38) is the sensing length N. It can be directly computed that the sensing time N should satisfy

$$N \ge \max\left\{\left[\frac{\mathcal{Q}^{-1}(\mathrm{Pr}_{HS}^{(k)})}{\theta_{k+1}/(P_k\gamma+\rho\sigma_0^2)-1}\right]^2, \left[\frac{\mathcal{Q}^{-1}(1-\mathrm{Pr}_{LS}^{(k)})}{\theta_k/(P_k\gamma+\frac{1}{\rho}\sigma_0^2)-1}\right]^2\right\} \tag{39}$$

in order to achieve the robust recognition. We could obtain the lower boundary of N from (39), and if this lower bound is infinite, then the detection is non-robust. Similar to [26], this phenomenon is named as *SNR wall*.

It is then interesting to derive the conditions under which SNR wall happens such that one can choose the system parameters in the real application to avoid SNR wall. Let us define a new variable $\Delta\mathrm{SNR}_{(k)} \triangleq \frac{(P_{k+1}-P_k)\gamma}{\sigma_0^2}$. Referring to (39), if either of the two terms in the parenthesis is infinite, then the SNR wall effect appears. Since $\mathcal{Q}(\mathcal{P}_{HS}^{(k)})$ is a constant, $\left[\frac{\mathcal{Q}^{-1}(\mathrm{Pr}_{HS})}{\theta_{k+1}/(P_k\gamma+\rho\sigma_0^2)-1}\right]^2$ will go to infinity only when its denominator $(\frac{\theta_{k+1}}{(P_k\gamma+\rho\sigma_0^2)}-1)$ approaches zero, which can be expanded as

$$\begin{aligned}\frac{\theta_{k+1}}{P_k\gamma+\rho\sigma_0^2}-1 &= \frac{P_{k+1}\gamma+\frac{1}{\rho}\sigma_0^2}{(P_{k+1}\gamma+\frac{1}{\rho}\sigma_0^2)-(P_k\gamma+\rho\sigma_0^2)}\ln\left(\frac{P_{k+1}\gamma+\frac{1}{\rho}\sigma_0^2}{P_k\gamma+\rho\sigma_0^2}\right)-1\\ &= \frac{P_{k+1}\gamma+\frac{1}{\rho}\sigma_0^2}{(P_{k+1}\gamma-P_k)-(\rho-\frac{1}{\rho})\sigma_0^2}\ln\left[1+\frac{(P_{k+1}-P_k)\gamma-(\rho-\frac{1}{\rho})\sigma_0^2}{P_k\gamma+\rho\sigma_0^2}\right]-1.\end{aligned} \tag{40}$$

Let us now prove that (40) approaches zero when $(P_{k+1}-P_k)\gamma$ drops from infinite to $\rho\sigma_0^2-\frac{1}{\rho}\sigma_0^2$, i.e. $\Delta\mathrm{SNR}_{(k)} \to \rho-\frac{1}{\rho}$. For easier explanation, denote $(P_{k+1}-P_k)\gamma-(\rho\sigma_0^2-\frac{1}{\rho}\sigma_0^2)$ by A. Due to the property $\ln(1+x)\sim O(x)$, there is:

$$\begin{aligned}&\lim_{A\to 0}\left\{\frac{P_{k+1}\gamma+\frac{1}{\rho}\sigma_0^2}{A}\ln\left[1+\frac{A}{P_k\gamma+\rho\sigma_0^2}\right]-1\right\}\\ =&\frac{P_{k+1}\gamma+\frac{1}{\rho}\sigma_0^2}{A}\frac{A}{P_k\gamma+\rho\sigma_0^2}-1=0.\end{aligned} \tag{41}$$

In turn, we know that $\left[\frac{\mathcal{Q}^{-1}(\mathrm{Pr}_{HS})}{\theta_{k+1}/(P_k\gamma+\rho\sigma_0^2)-1}\right]^2$ has an infinite value when there is $\Delta\mathrm{SNR}_{(k)} = \rho-\frac{1}{\rho}$. Similarly, we can prove that when $(P_k-P_{k-1})\gamma-(\rho\sigma_0^2-\frac{1}{\rho}\sigma_0^2)$ approaches zeros, i.e. $\Delta\mathrm{SNR}_{(k-1)} \to \rho-\frac{1}{\rho}$, $\left[\frac{\mathcal{Q}^{-1}(1-\mathrm{Pr}_{LS})}{\theta_k/(P_k\gamma+\frac{1}{\rho}\sigma_0^2)-1}\right]^2$ becomes infinite too.

Case II, $\mathcal{R}_1(\mathcal{H}_k)$ overlapping with $\mathcal{R}_1(\mathcal{H}_{k-1})$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$: In this case, there is $(P_k-P_{k-1})\gamma-(\rho\sigma_0^2-\frac{1}{\rho}\sigma_0^2)$ and $(P_{k+1}-P_k)\gamma-(\rho\sigma_0^2-\frac{1}{\rho}\sigma_0^2)$ i.e., $\Delta\mathrm{SNR}_{(k-1)} < \rho-\frac{1}{\rho}$ and $\Delta\mathrm{SNR}_{(k)} < \rho-\frac{1}{\rho}$. According to Lemma 2, we have

$\theta_{k+1} > \phi_{l(k+1)}$. Hence, with (37) and definition 2, to achieve robust recognition there must be

$$\mathrm{Pr}_{HS}^{(k)} \geq \left\{ \mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{l(k+1)}}{P_k\gamma + \rho\sigma_0^2} - 1 \right) \right] \right\}. \quad (42)$$

Note that $\frac{\phi_{l(k+1)}}{P_k\gamma+\rho\sigma_0^2} - 1 = \frac{(P_{k+1}-P_k)\gamma - (\rho - \frac{1}{\rho})\sigma_0^2}{P_k\gamma+\rho\sigma_0^2} < 0$. Thus, when N increases, $\left\{\mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{l(k+1)}}{P_k\gamma+\rho\sigma_0^2} - 1\right)\right]\right\}$ increases too, which may make the inequality (42) not hold. Moreover, since $\left\{\mathcal{Q}\left[\sqrt{N}\left(\frac{\phi_{l(k+1)}}{P_k\gamma+\rho\sigma_0^2} - 1\right)\right]\right\}$ has the minimum value of 0.5 at $N = 0$, any general $\mathrm{Pr}_{HS}^{(k)}$ (normally very small) cannot satisfy (42). Hence, when $\Delta\mathrm{SNR}_{(k)} < \rho - \frac{1}{\rho}$ holds, the recognition for $\mathcal{H}_k$ is not robust. By similar analysis it could be proved that when $\Delta\mathrm{SNR}_{(k-1)} < \rho - \frac{1}{\rho}$ holds, the recognition for $\mathcal{H}_k$ is not robust either.

Combining Case I and Case II, we may define $(\rho - \frac{1}{\rho})$ as the position of SNR wall. For a given $\mathcal{H}_k$, in order to achieve robust recognition, SNR wall should be prevented and there should be:

$$\Delta\mathrm{SNR}_{(k)} > \rho - \frac{1}{\rho}, \quad (43)$$

$$\Delta\mathrm{SNR}_{(k-1)} > \rho - \frac{1}{\rho}. \quad (44)$$

Interestingly, the SNR wall sets restrictions on ΔSNR from both sides of a specific power level P_k. Fig. 5 displays the required number of samples to achieve robust recognition, with $\mathrm{Pr}_{HS}^{(k)}$ and $\mathrm{Pr}_{LS}^{(k)}$ set as 0.1, and the size of noise uncertainty ρ taken as 1.1,1.3,1.5. The average SNR is defined as:

$$\mathrm{SNR}_{av} = \sum_{i=0}^{M-1} \Delta\mathrm{SNR}_i \Pr(\mathcal{H}_i). \quad (45)$$

It is seen that as the noise uncertainty decreases, the position of SNR wall will be pushed to the left as shown in Fig. 5. When the noise uncertainty disappears, the SNR wall will be pushed to minus infinity and SNR wall effect disappears too.

Remark 5: The effect of SNR wall has two-fold meanings: (i) if $\Delta\mathrm{SNR}_{(k-1)}$ or $\Delta\mathrm{SNR}_{(k)}$ approaches SNR wall from right side, then the required sensing time to achieve robust recognition increases; (ii) if either $\Delta\mathrm{SNR}_{(k-1)}$ or $\Delta\mathrm{SNR}_{(k)}$ drops below SNR wall, then the recognition is not robust.

Remark 6: It is interesting to notice that the position of SNR wall coincides with the boundary condition of power ambiguity. Specifically, if SNR wall does not happen, i.e., $\Delta\mathrm{SNR}_{(k)} > \rho - \frac{1}{\rho}$ for any k, then there is no overlapping areas between any two $\mathcal{R}_1(\mathcal{H}_k)$'s, and thus the power ambiguity does not happen. Only when SNR wall happens, then there will exist overlapping areas between certain two $\mathcal{R}_1(\mathcal{H}_k)$'s, into which $T(\mathbf{x})$ falling leads to power ambiguity.

Remark 7: Traditional binary spectrum sensing only consider two power levels, P_0, $P_1 =$ and SNR wall is only related with $\Delta\mathrm{SNR}_0 = (P_1 - P_0)$. However in MPTP scenario, there are different $(P_{k+1} - P_k)$ for different k, and it is possible that some power levels satisfy the requirements in (43) and (44), while the others do not. Hence, when PU transmits by those power levels not satisfying (43) or (44), the recognition is not robust. On the other side, when PU transmits by power levels satisfying (43) and (44), the power recognition is robust.

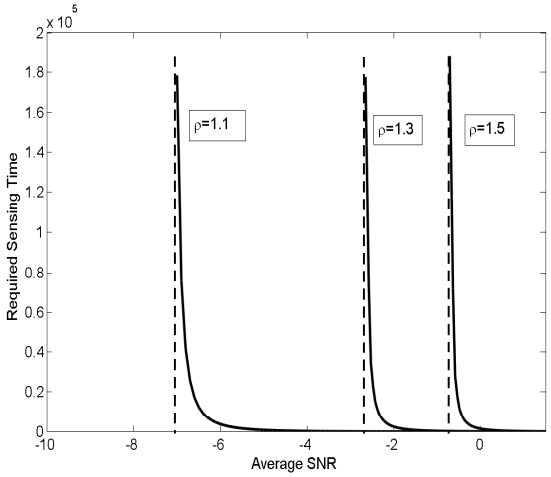

Fig. 5. The required number of samples to achieve robust recognition versus average SNR.

IV. Cooperative Spectrum Sensing via Majority Voting

In this section we examine the distributed cooperative spectrum sensing, where K SUs cooperate to detect the presence of PU as well as the power levels by forwarding their decision results $\{0, \ldots, M\}$ to a fusion center. It is natural to apply the majority law when there are multiple SUs and the decision candidates are multiple.

Cooperative sensing might be hopeful to counteract the power ambiguity since different SUs might vote for different bunches of power levels and the selection probability of the true one is greatly enhanced. The conventional fusion rules, e.g., AND, OR, and KON rules are defined on the binary decision while are not applicable for MPTP scenario [20]. To better illustrate the proposed cooperative sensing scheme, we assume that the power ambiguity at each SU only happens between two neighboring power levels, while extension to more complicated situations is straightforward.

The fusion center builds a $K \times (M+1)$ matrix $\mathbf{A} = [\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_K]^T$ based the decisions from all K SUs, where $\vec{a}_k = (a_{k1}, \cdots, a_{k(M+1)})$ denotes the k-th SU's decision for all $M+1$ power levels. The element $a_{kj} = 1$ denotes that the k-th SU votes for P_{j-1}, otherwise $a_{kj} = 0$. Meanwhile, let us formulate a voting pool $\vec{d} = (d_0, \cdots, d_M)$, where d_m denotes the number of SUs that claim $\mathcal{H}_m$. For a given $\mathbf{A}$, there is

$$d_m = \sum_{i=1}^{K} a_{i(m+1)}, \forall m \in \{0, \cdots, M\}. \quad (46)$$

According to the majority law, the power level corresponding to the largest value in $\vec{d}$ will be made as the decision:

$$\hat{m} = \arg\max_{m \geq 0} \ d_m. \quad (47)$$

We then analyze the power recognition performance of cooperative sensing. Let us introduce a $(2M+1) \times 1$ vector $\vec{c} = (c_0, \cdots, c_M, c_{0,1}, c_{1,2}, \ldots, c_{M-1,M})$ to compute the probability of $\vec{d}$, where c_i denotes the number of SUs that claim only one power level P_i, while $c_{i,i+1}$ denotes the number

of SUs that claim both power levels P_i, P_{i+1}. Obviously, there is $\sum_{i=0}^{M} c_i + \sum_{i=0}^{M-1} c_{i,i+1} = K$.

In order to provide a concise result we assume that each SU has the same decision probabilities $\Pr(\mathcal{H}_j|\mathcal{H}_i)$ and $\Pr(\mathcal{A}_{j,j+1}|\mathcal{H}_i)$ [23], while a general assumption merely leads to tedious computation. Then, the probability of any specific $\vec{c}$ can be computed as

$$\begin{aligned}\Pr(\vec{c}|\mathcal{H}_k) &= \binom{K}{c_0}\Pr(\mathcal{H}_0|\mathcal{H}_k)^{c_0}\binom{K-c_0}{c_1}\Pr(\mathcal{H}_1|\mathcal{H}_k)^{c_1}\cdots \\ &\binom{K-\sum_{l=0}^{M-1} c_l}{c_M}\Pr(\mathcal{H}_M|\mathcal{H}_k)^{c_M}\binom{K-\sum_{l=0}^{M} c_l}{c_{0,1}}\Pr(\mathcal{A}_{0,1}|\mathcal{H}_k)^{c_{0,1}}\cdots \\ &\binom{K-\sum_{l=0}^{M} c_l-\sum_{i=0}^{M-2} c_{i,i+1}}{c_{M-1,M}}\Pr(\mathcal{A}_{M-1,M}|\mathcal{H}_k)^{c_{M-1,M}} \\ &= \frac{K!}{\prod_{l=0}^{M} c_l \prod_{i=0}^{M-1} c_{i,i+1}} \prod_{m=0}^{M}\Pr(\mathcal{H}_m|\mathcal{H}_k)^{c_m}... \\ &\qquad \times \prod_{u=0}^{M-1}\Pr(\mathcal{A}_{u,u+1}|\mathcal{H}_k)^{c_{u,u+1}}. \quad (48)\end{aligned}$$

Theorem 1: The probability of a given $\vec{d}$ is

$$\begin{aligned}\Pr(\vec{d}|\mathcal{H}_i) = \sum_{c_0=0}^{d_0} \sum_{c_1=\max\{0,l_1\}}^{\min\{d_1,u_1\}} \cdots \sum_{c_m=\max\{0,l_m\}}^{\min\{d_m,u_m\}} \cdots \\ \sum_{c_{M-1}=\max\{0,l_{M-1}\}}^{\min\{d_{M-1},u_{M-1}\}} \Pr(\vec{c}|\mathcal{H}_i), \quad (49)\end{aligned}$$

where

$$\begin{cases} u_m = d_m + (-1)^m \sum_{n=0}^{m-1}(-1)^n(d_n - c_n) \\ l_m = -d_{m+1} + u_m \end{cases} \quad (50)$$

for $m = 1, 2, \cdots, M-1$.

Proof: See Appendix B. ∎

Define $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$ as the decision probability at the fusion center.

Theorem 2:

$$\begin{aligned}\Pr_m(\mathcal{H}_j|\mathcal{H}_i) = \sum_{d_j=\lceil \frac{D}{M+1} \rceil}^{\lfloor D \rfloor} \sum_{d_0=\max\{0,\alpha_0\}}^{\min(d_j,\beta_0)} \sum_{d_1=\max\{0,\alpha_1\}}^{\min\{d_j,\beta_1\}} \cdots \\ \sum_{d_M=\max\{0,\alpha_M\}}^{\min\{d_j,\beta_M\}} \Pr(\vec{d}|\mathcal{H}_i) \quad (51)\end{aligned}$$

where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the ceiling function and floor function, respectively, and $D = \sum_{m=0}^{M} d_m = \sum_{i=1}^{K}\sum_{j=1}^{M+1} a_{ij}$ is the total number of hypotheses that all SUs claim. Moreover, α_i's and β_i's are defined as

$$\begin{aligned}\alpha_m &= \begin{cases} D - \sum_{i=0}^{m-1} d_i - (M-m)d_j, & 0 \le m < j \\ D - \sum_{i=0}^{m-1} d_i - (M-m-1)d_j, & j < m \le M \end{cases} \\ \beta_m &= \begin{cases} D - \sum_{i=0}^{m-1} d_i - d_j, & 0 \le m < j \\ D - \sum_{i=0}^{m-1} d_i, & j < m \le M. \end{cases} \end{aligned} \quad (52)$$

Proof: See Appendix C. ∎

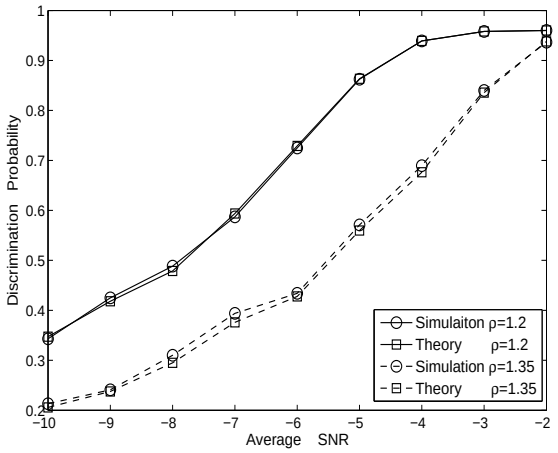

Fig. 6. Comparison between theoretical and numerical probability of cooperative sensing.

In Fig. 6, we provide two examples to verify the derived $\Pr(\vec{d}|\mathcal{H}_i)$. The number of SUs is taken as $K = 3$ and the sample length is $N = 4000$. The size of noise uncertainty ρ is taken as 1.2 and 1.35. Clearly, the theoretical value matches the numerical quite well.

V. Simulations

In this section we provide simulation results to evaluate the performance of the proposed spectrum sensing algorithms. Four positive power levels are assumed for PU, and the prior probabilities of all $\mathcal{H}_i$'s, $i \ge 0$, are set as 0.5, 0.2, 0.125, 0.075, 0.1, respectively. The average SNR is defined in (45).

We first evaluate the performance of discriminating PU's transmitting power level in Fig. 7 with four different sizes of noise uncertainty: $\rho = 1.1, 1.2, 1.3, 1.4$. The number of the sensing sample is taken as $N = 4000$. It can be observed that the detection probability improves with the increase of average SNR. Meanwhile, when noise uncertainty becomes severer, the detection probability suffers from huge loss. It is notable that the gap between each curve shrinks from the top to bottom, which implies that the performance loss caused by augmentation of noise uncertainty is larger when ρ is relatively small.

The discrimination probability versus the number of samples is displayed in Fig. 8. We set $\rho = 1.2$, and the SNR wall could be computed as -4.35dB. Consistent with Fig. 7, the decision performance suffers from the increase of noise uncertainty. When average SNR is -3dB, the discrimination probability gains a lot from longer sensing time, while the discrimination probability gains little when the average SNR is -4.5dB and -5dB. For the case the average SNR is -6dB, the discrimination probability almost remains unchanged for different sensing time, which complies to the fact that the recognition performance cannot be improved through increasing the sampling length when SNR wall happens. It should be emphasized that SNR wall is derived based worst recognition performance (see the usage of max and min in (35) and 36). Hence in Fig. 8, the discrimination probability is still about 0.8

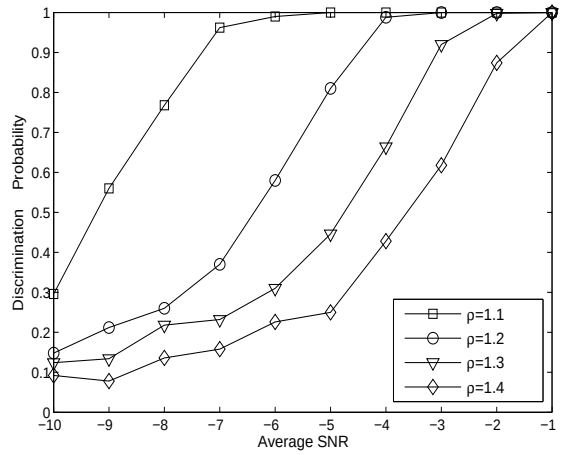

Fig. 7. Discrimination probability under different sizes of noise uncertainty.

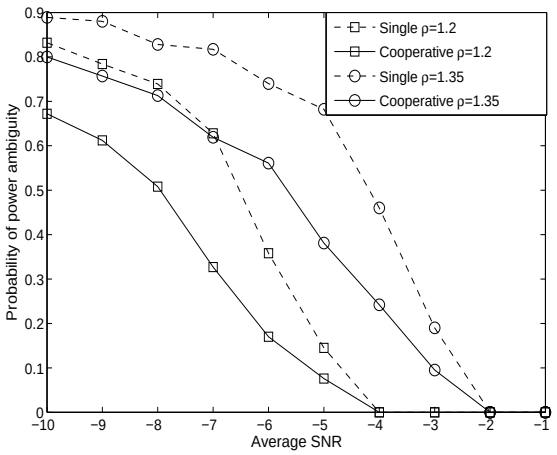

Fig. 9. Probability of power ambiguity happening for cooperative sensing and single sensing.

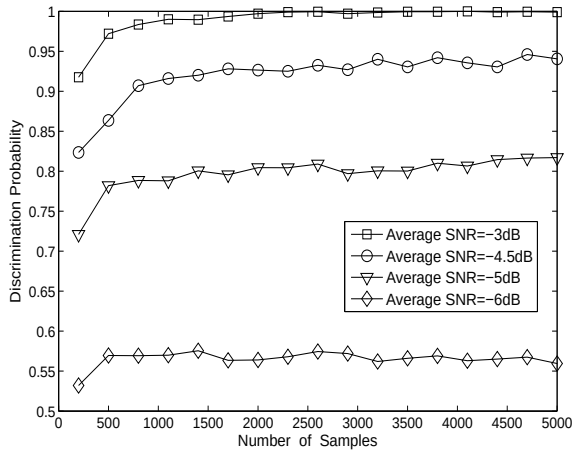

Fig. 8. Discrimination probability versus number of samples.

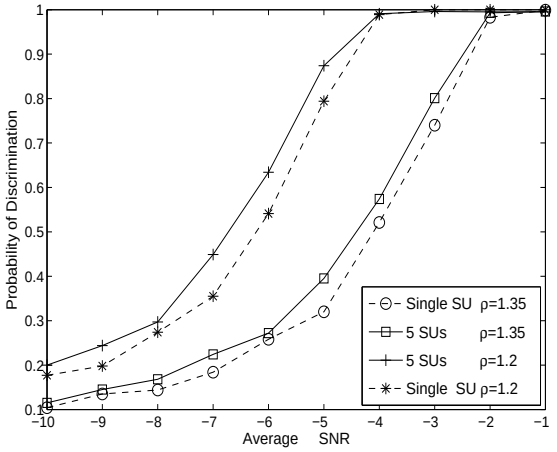

Fig. 10. The discrimination probability of cooperative sensing with 5 SUs and single SU sensing under different sizes of noise uncertainty.

when average SNR is -5dB. Only when SNR is much lower than the theoretical SNR wall, say -6dB, will the SNR wall effect completely appears in all the simulation runs and the performance fails.

Lastly, we compare the sensing performance of cooperative sensing and single SU sensing. Five SUs are used, and N is set as 4000. The noise uncertain size is taken as $\rho = 1.2, 1.35$, respectively. In Fig. 9, we demonstrate the appearance probability of the power ambiguity. It is observed that the probability of power ambiguity happening drops when cooperative sensing is applied. In Fig. 10, we show the discrimination probability for different sensing. It is observed that cooperative sensing performs better than single sensing, especially in the situation where the average SNR is relatively low.

Notice that Fig. 9 shows that the probability of power ambiguity is 0 for both cooperative and single SU sensing when there is $\text{SNR}_{av} > -4$ and $\rho = 1.2$. However, the discrimination probability is still higher when cooperative sensing is applied as shown by Fig. 10 shows. Thus, we claim that even in situation where no power ambiguity happens, cooperative sensing still improves the sensing performance.

VI. Conclusions

In this paper, we studied the spectrum sensing and power recognition in MPTP scenario with noise uncertainty. The GLRT paradigm is applied to estimate the noise variance and the energy detection is, again, proved to be the optimal approach. We derived the closed-form decision region for each power levels as well as the corresponding closed form decision probability. In MPTP case, the larger noise uncertainty could cause power ambiguity effect which was not observed in binary sensing case. With modified definition, we further investigated the SNR wall effect and, interestingly, the happening condition coincides with that of the power ambiguity effect. We also designed a majority law based cooperative sensing scheme and compute the closed-form expression of the corresponding decision probability. Finally, simulation results are provided to corroborate the proposed studies.

APPENDIX A
PROOF FOR LEMMA 3

The decision region for $\mathcal{H}_k$ in Case III is made of part of the blank region between $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$ and part of the blank region between $\mathcal{R}_1(\mathcal{H}_{k-1})$ and $\mathcal{R}_1(\mathcal{H}_k)$. Note that $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$ may overlap, under which case no blank region lies between $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$. Thus, the formulation of the decision regions in Case III is dependent on whether these overlapping areas exist. Let us proceed the discussion in three circumstances:

1) $\mathcal{R}_1(\mathcal{H}_{k-1})\bigcap\mathcal{R}_1(\mathcal{H}_k)=\emptyset$ and $\mathcal{R}_1(\mathcal{H}_{k+1})\bigcap\mathcal{R}_1(\mathcal{H}_k)=\emptyset$: Mathematically, there is $\phi_{u(k-1)}<\phi_{lk}$ and $\phi_{uk}<\phi_{l(k+1)}$ (shown as Fig. 2(a)). In this case, $(\phi_{u(k-1)},\phi_{lk})$ is the blank region between $\mathcal{R}_1(\mathcal{H}_{k-1})$ and $\mathcal{R}_1(\mathcal{H}_k)$, while $(\phi_{uk},\phi_{l(k+1)})$ is the blank region between $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$. We know from Lemma 1 that $\theta_k<\phi_{lk}$ and $\theta_{k+1}>\phi_{uk}$. Moreover, (22) indicates that if $T(\mathbf{x})$ falls in (θ_k,ϕ_{lk}) or $(\phi_{uk},\theta_{k+1}]$, then $\mathcal{H}_k$ is made as the decision. Hence, the decision region is $(\theta_k,\phi_{lk})\bigcup(\phi_{uk},\theta_{k+1}]$.

2) $\mathcal{R}_1(\mathcal{H}_{k-1})\bigcap\mathcal{R}_1(\mathcal{H}_k)\neq\emptyset$ and $\mathcal{R}_1(\mathcal{H}_{k+1})\bigcap\mathcal{R}_1(\mathcal{H}_k)=\emptyset$: Mathematically, there is $\phi_{u(k-1)}>\phi_{lk}$ and $\phi_{u(k)}<\phi_{l(k+1)}$ (shown as Fig.2(b)). In this case, blank region only exists between $\mathcal{R}_1(\mathcal{H}_k)$ and $\mathcal{R}_1(\mathcal{H}_{k+1})$. We know from Lemma 2 that $\theta_k<\phi_{u(k-1)}$ and $\theta_{k+1}>\phi_{uk}$ under this circumstance. As the result, the the decision region is $(\phi_{uk},\theta_{k+1}]$.

3) $\mathcal{R}_1(\mathcal{H}_{k-1})\bigcap\mathcal{R}_1(\mathcal{H}_k)\neq\emptyset$ and $\mathcal{R}_1(\mathcal{H}_{k+1})\bigcap\mathcal{R}_1(\mathcal{H}_k)\neq\emptyset$: Mathematically, there is $\phi_{u(k-1)}>\phi_{lk}$ and $\phi_{uk}>\phi_{l(k+1)}$ (shown as Fig.2(c)). In this case, there is no blank region on either side of $\mathcal{R}_1(\mathcal{H}_k)$, and so the decision region is $\emptyset$.

Combining all the above three circumstances, the decision region of Case III can be written as (27).

APPENDIX B
PROOF FOR THEOREM 1

It is obvious that $\vec{d}$ and $\vec{c}$ have the following mapping relationship:

$$d_i=\begin{cases}c_0+c_{0,1}, & i=0\\ c_i+c_{i-1,i}+c_{i,i+1}, & 1\le i\le M-1\\ c_M+c_{M-1,M}, & i=M,\end{cases}\tag{53}$$

and hence, all elements in $\vec{c}$ can be determined by its first M elements $c_0,c_1,\cdots,c_{M-1}$ as

$$c_{m-1,m}=(-1)^{m-1}\sum_{n=0}^{m-1}(-1)^n(d_n-c_n).\tag{54}$$

For a given $\vec{d}$, we need to find all possible $\vec{c}$'s that contribute to the same $\vec{d}$. First, we know c_0 ranges from 0 to d_0. For a specific c_0, $c_{0,1}$ is determined by $c_0+c_{0,1}=d_0$ from (53). Then referring to the condition $c_1+c_{0,1}+c_{1,2}=d_1$ we obtain $c_{1,2}=d_1-c_1-c_{0,1}$. According to (53), $c_2+c_{1,2}+c_{2,3}=d_2$, so there is $c_{1,2}<d_2$. Hence, we have

$$0\le c_{1,2}=d_1-c_1-c_{0,1}\le d_2,$$

or equivalently that c_1 must satisfy

$$-d_2+d_1-d_0+c_0\le c_1\le d_1-d_0+c_0.$$

Considering the inside condition that $0\le c_1\le d_1$, the upper bound of c_1 is $\min\{d_1,u_1\}$, while the lower bound is $\max\{0,l_1\}$.

After the range of c_1 is determined, $c_{1,2}$ is determined. Similarly when the range of c_{m-1} is determined, $c_{m-1,m}$ is also determined, and $c_{m,m+1}=d_m-c_m-c_{m-1,m}$ must satisfy

$$0\le c_{m,m+1}=d_m-c_m-c_{m-1,m}\le d_{m+1},$$

which gives

$$-d_{m+1}+d_m-c_{m-1,m}\le c_m\le d_m-c_{m-1,m},$$

where $c_{m-1,m}$ is given in (54). Meanwhile, considering $0\le c_m\le d_m$, we obtain the upper and lower bounds of c_m as shown in (49). The proof is completed.

APPENDIX C
PROOF FOR THEOREM 2

Consider the situation where $\mathcal{H}_{k_0}$ wins out. Then, the number of votes for P_k must exceed $\frac{D}{M+1}$, which is the mean value of votes. Meanwhile, d_{k_0} cannot exceed the total number of votes, i.e., $d_{k_0}\le D$.

After the range of d_{k_0} is determined, we could compute the range of $d_0,\cdots,d_{k-1},d_{k+1},\cdots,d_M$. For the lower bound of d_0, besides the obvious restriction $d_0>0$, it should satisfy $\frac{D-d_0-d_{k_0}}{M-1}<d_{k_0}$ or equivalently $d_0>D-Md_{k_0}$. If $\frac{D-d_0-d_{k_0}}{M-1}\le d_{k_0}$, then there must be one or more $d_i(i\neq k)$ satisfies $d_i>d_{k_0}$. For the upper bound of d_0, besides $d_0<d_{k_0}$, there must be $d_0<D-d_{k_0}$. Hence, the range of d_0 is determined.

Then, let us derive the range of $d_m(m<k_0)$. In addition to $0<d_i<d_{k_0}$ and $d_m\le D-\sum_{i=0}^{m-1}-d_{k_0}$, there must be $\frac{D-\sum_{i=0}^{m}d_i-d_{k_0}}{M-m-1}\le d_{k_0}$, which is equivalent to $d_m>D-\sum_{i=0}^{m-1}d_i-(M-m)d_{k_0}$.

We next derive the range of $d_m(m>k_0)$. In addition to $0<d_i<d_{k_0}$ and $d_m\le D-\sum_{i=0}^{m-1}d_i$, there must be $\frac{D-\sum_{i=0}^{m}d_i}{M-m}\le d_{k_0}$, which is equivalent to $d_m>D-\sum_{i=0}^{m-1}d_i-(M-m-1)d_{k_0}$.

Then, for $d_1,\cdots,d_{k-1},d_{k+1},\cdots,d_M$, applying similar analysis could easily derive their range. The proof is completed.

REFERENCES

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[2] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," in *Proc. IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.

[3] A. Sahai and D. Cabric,"A tutorial on spectrum sensing: Fundamental limits and practical challenges," in *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD, Nov. 2005.

[4] H. S. Chen, W. Gao, and D. G. Daut, "Signature based spectrum sensing algorithms for IEEE 802.22 WRAN," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, Scotland, June 2007

[5] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. Asilomar Conf. Signals, Syst. Computers (ACSSC)*, Pacific Grove, CA, Nov. 2004.

[6] P. D. Sutton, K. E. Nolan, and L. E. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 13–24, Jan. 2008.

[7] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 32–39, Apr. 2008.

[8] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. Wireless Commun.*, Maui, HI, June 2005.

[9] W.Lin and Q. Zhang, "A Design of Energy Detector in Cognitive Radio under Noise Uncertainty," in *Proc. IEEE Int. Conf. Commun. Sys. (ICCS)*, Singapore, 2008 , pp. 213–217.

[10] T. J. Lim, R. Zhang, Y.-C. Liang, and Y. Zeng, "GLRT-based spectrum sensing for cognitive radio," in *Proc. IEEE GLOBECOM*, New Orleans, LA, USA, 2008, pp. 1–5.

[11] J. Font-Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2137C-2146, Jul. 2010.

[12] IEEE 802 LAN/MAN Standards Committee, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," IEEE Standard 802.11, 1999 edition, 1999.

[13] 3GPP TS 36.213, Evolved Universal Terrestrial Radio Access (EUTRA), "User Equipment (UE) Radio Transmission and Reception,"(release 8).

[14] 3GPP TR 36.913, "Requirements for Further Advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced)," 3GPP, Tech. Rep. v. 10.0.0, Mar. 2011.

[15] Z. Chen, F. Gao, X.-D. Zhang, J. C. F. Li, and M. Lei, "Sensing and Power Allocation for Cognitive Radio with Multiple Primary Transmit Powers," *IEEE Wireless Commun. Lett.*,vol. 2, no. 3, pp. 319–322, Jun. 2013.

[16] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. WirelessCom 05 Symposium on Signal Processing*, June. 2005, vol. 1, pp. 464–469.

[17] H.V. Poor and D.P. Looze, "Minimax state estimation for linear stochastic systems with noise uncertainty," *IEEE Trans. Auroniatic Conrml*, vol. AC–26, pp. 902–906, 1981.

[18] M. Oude Alink, A. Kokkeler, E. Klumperink, G. Smit, and B. Nauta, "Lowering the SNR wall for energy detection using cross-correlation," *IEEE Trans. Veh Technol*, vol. 60, no. 8, pp. 3748–3757, Oct. 2011.

[19] A. Mariani, A. Giorgetti, and M. Chiani,"SNR wall for energy detection with noise power estimation," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 1–6, June. 2011.

[20] F. Gao, J. Li, T. Jiang, C. Wen. "Sensing and recognition when primary user has multiple transmit power levels". *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2704–2717, Mar. 2015.

[21] S. Shellhammer and G. Chouinard, "Spectrum sensing requirements summary," IEEE P802.22–06/0089r1, Tech. Rep., June 2006.

[22] D. Torrieri, The radiometer and its practical implementation, in *Proc. 2010 IEEE Military Commun. Conf. (MILCOM)*, San Jose, CA, Oct. 2010, pp. 304–310.

[23] W. Zhang, R. Mallik, and K. Letaief, "Optimization of Cooperative Spectrum Sensing with Energy Detection in Cognitive Radio Networks," *IEEE Trans. Wireless Comm.*, vol. 8, no. 12, pp. 5761–5766, Dec. 2009.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998, vol. 2.

[25] S. S. Wilks, *Mathematical Statistics*. Read Brooks, 2007.

[26] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, Feb. 2008.

[27] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, "Sensing throughput tradeoff in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[28] 3GPP TS 36.213, Evolved Universal Terrestrial Radio Access (EUTRA), "User Equipment (UE) Radio Transmission and Reception,"(release 8).

[29] 3GPP TR 36.913, "Requirements for Further Advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced)," 3GPP, Tech. Rep. v. 10.0.0, Mar. 2011.

[30] S. Liu, B. Hu, and X. Wang, "Hierarchical cooperative spectrum sensing based on double thresholds energy detection," *IEEE COMMUN LETT*, Vol. 16, No. 7, pp. 1096–1099, Jul. 2012.

[31] K. Zhang, J. Li, and F. Gao, Machine learning techniques for spectrum sensing when primary user has multiple transmit power, in *Proc. IEEE Int. Conf. Commun. Sys. (ICCS)*, Macau, China, Nov. 2014, pp. 137–141.

[32] C. Qian, H. Qian and F. Gao, "Spectrum sensing and SNR walls when primary user has multiple power levels", to appear in *Proc. IEEE Int. Conf. Commun. China (ICCC)*, Shenzhen, China, Nov. 2015.