

# Sensing and Recognition When Primary User Has Multiple Transmit Power Levels

Feifei Gao, *Senior Member, IEEE*, Jiachen Li, Tao Jiang, *Senior Member, IEEE*, and Wen Chen, *Senior Member, IEEE*

**Abstract**—In this paper, a new cognitive radio (CR) scenario called Multiple Primary Transmit Power (MPTP) is investigated where the primary user (PU) could possibly work on more than one discrete transmit power levels. Different from most existing literatures where PU is assumed to operate with a constant transmit power only, this new consideration well matches the practical standards, e.g., IEEE 802.11 Series, GSM, LTE, LTE-A, etc., as well as the adaptive powering concept that a user would vary its transmit power under different situations. The primary target of CR under MPTP may still be detecting the presence of PU. However, there emerges a secondary target as to identify the PU's transmit power level. Compared to the existing sensing strategy where the secondary user (SU) only detects the “on-off” status of PU, recognizing the transmit power level of PU achieves more “cognition” and makes the CR more intelligent. Meanwhile, SU could utilize the power level information of PU and make the subsequent design. We derive quite many closed-form results for either the threshold expressions or the performance analysis in this new CR scenario, from which many interesting points and discussions are raised. We then study the cooperative sensing schemes under MPTP and demonstrate their significant differences from traditional cooperative algorithms. Lastly, numerical examples are provided to corroborate the proposed studies.

**Index Terms**—Cooperative sensing, majority decision, modified multiple hypothesis testing, multiple primary transmit power (MPTP), power-mask, spectrum sensing.

## I. INTRODUCTION

COGNITIVE RADIO (CR) has been recognized as a promising solution to spectrum scarcity and spectrum under-utilization [1] by allowing the secondary users (SUs) to

Manuscript received August 01, 2014; revised January 19, 2015; accepted March 08, 2015. Date of publication August 01, 2014; date of current version April 15, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rui Zhang. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2013CB336600, 2012CB316100, by the National Natural Science Foundation of China under Grant 61422109, 61428104, 61401169, by the Beijing Natural Science Foundation under Grant 4131003, and by the Tsinghua University Initiative Scientific Research Program under Grant 20121088074.

F. Gao and J. Li are with the Department of Automation, Tsinghua University, State Key Lab of Intelligent Technologies and Systems, Tsinghua National Laboratory for Information Science and Technology (TNList) Beijing 100084, China (e-mail: feifeigao@ieee.org; jiacli@outlook.com).

T. Jiang is with the School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan, 430074, China (e-mail: Tao.Jiang@ieee.org).

W. Chen is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China (e-mail: wenchen@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2015.2415751

access the spectrum of the primary user (PU) when the latter is idle. In this context, spectrum sensing, which helps SU to detect the occupancy status of a specific frequency band, forms a key component of CR.

Popular spectrum sensing techniques include matched filter detection [2], energy detection [3]–[9], and cyclostationary detection [10], [11], among which energy detection has received intensive attention because it requires the least prior knowledge of PU and is very simple to implement. When SU is equipped with multiple antennas, a promising sensing technology was designed in [12] where the eigenvalues of the receive covariance matrix are used to judge the status of PU. It is shown that the performance of multiple antennas based spectrum sensing is much better than that from a single antenna because the former fully utilizes the correlation among antennas. On the other side, when multiple antennas are not available due to the size limitation of the wireless terminal, one can refer to the cooperative spectrum sensing from more than one SUs to enhance the sensing performance [13]–[16].

There are also quite a number of works discussing different issues of CR over the past decade, for example, parameter uncertainty based spectrum sensing [17], design with imperfect sensing [18], sensing throughput tradeoff [19], spectrum sharing [20], as well as many other hybrid schemes between CR and other technologies, e.g., games among SUs [21], sensing in OFDM system [22], sensing in relay network [23], [24], sensing with MAC layer protocol [25] and so on.

It is not difficult to notice that most existing spectrum sensing techniques [2]–[16] as well as other related studies [17]–[24] assume PU is either absent or transmits with a constant power level. However, it can be easily known from the current standards, e.g., IEEE 802.11 series [26], GSM [27], and the future standards, e.g., LTE [28], LTE-A [29] that the licensed users could be working under different transmit power levels in order to cope with different situations, e.g., environment, rate, etc. A typical example is in CDMA [30] uplink scenario when the users are subjected to power control in order to deal with the near-far effects. Actually, varied transmit power has been extensively considered in many existing literatures studying the power allocation problem [31], [32]. Therefore, the traditional spectrum sensing techniques, which only deal with a constant power level of PU, cannot adequately match the practical situations or the theoretical demands.

On the industry aspect, FCC has specified the interference protection requirements for TV white space, and the requirements for different powered services, e.g., full-power digital TV, full-power analog TV, low-power analog TV, low-power digital TV, etc., are different [33]. It then enlightens us that by detecting

the power levels of PU, SU may know PU's protection requirement and then could adjust its transmission strategy accordingly. On theoretical aspect, an important scheme that combines spectrum sensing and spectrum sharing to further enhance the spectrum efficiency was proposed recently [34]–[36], known as sensing based sharing scheme. The key idea is that SU should sense the presence of PU first and then transmit with full power if PU is absent but with limited power if PU is present. When PU has multiple transmit power levels, as usually the case, it is possible that PU's interference temperature could vary under different transmit power levels. Hence, by detecting the power levels of PU, SU may know PU's interference temperature requirement and could then adjust its transmit power accordingly. A vivid example where SU could adjust its transmission strategy according to different PU's transmission power levels in order to achieve certain optimal goal can be found in [37].

Therefore, the sensing target when PU has multiple power levels should not only be detecting the “on-off” status of PU but also to identify its power level, by which we can also achieve more “cognition” and make the CR more intelligent,<sup>1</sup> as originally desired in [1]. In fact, spectrum sensing under multiple primary transmit power (MPTP) scenario has already been presented in [39], where the authors briefly describe the new sensing strategy but focused more on the optimal power allocation of SU after power-level recognition. Unfortunately, [39] does not fully address the fundamental issues of spectrum sensing in MPTP, such as the necessary analysis in sensing process, power-mask effect, and the detailed evaluation framework. Moreover, [39] possesses some careless results like the expression of the decision probability and the definition of the discrimination probability.

In this paper, we provide a thorough investigation over the spectrum sensing problem in MPTP scenario, where PU randomly chooses a power level from a discrete power set according to certain prior probability. We propose two different but relevant spectrum sensing strategies, i.e., *detection before recognition* and *recognition before detection*, for local secondary user, where the former applies a modified hypothesis testing approach while the later relies on the standard hypothesis testing approach. We fully discuss a *power-mask* effect that is shown to be a unique phenomenon under MPTP scenario. We derive closed-form expressions of decision regions and also provide many remarks explaining the fundamental reasoning behind the multiple hypothesis detection under MPTP. To improve the sensing performance, we further propose two cooperative sensing schemes, which exhibit critical differences from the traditional cooperative binary sensing algorithms. Moreover, the closed-form performance analysis of all these sensing algorithms are derived, and various numerical examples are provided to corroborate the proposed studies.

The rest of this paper is organized as follows: Section II presents the system model of MPTP scenario. In Section III, we propose two spectrum sensing strategies with different detection and recognition order, and analyze their relationship. In Section IV, we investigate cooperative sensing in MPTP scenario and derive two different algorithms based

on the majority voting and the Bayes approach, respectively. In Section V, we provide simulation results to evaluate the designed algorithms. Finally, conclusions are drawn in Section VI.

## II. SYSTEM MODEL AND ASSUMPTIONS

Consider the classic CR model that consists of one PU and one SU. As regulated in IEEE 802.11 series, GSM, LTE, and LTE-A, PU could either be absent or operate under one transmit power level  $P_i$  from a predetermined set  $\{P_1, P_2, \dots, P_N\}$ . It is reasonable to assume that a power level, once being chosen, will be used for a certain period during which SU could perform sensing as well as the subsequent transmission.

Without loss of generality, we assume  $P_{i+1} > P_i > 0, \forall i$ , and the  $l$ th received sample at SU can be expressed as:

$$x_l = \begin{cases} n_l & \mathcal{H}_0 \\ \sqrt{P_i} \sqrt{\gamma} e^{j\phi} s_l + n_l & \mathcal{H}_i, i = 1, 2, \dots, N \end{cases} \quad (1)$$

where  $\mathcal{H}_0$  denotes the hypothesis that PU is absent while  $\mathcal{H}_i$  indicates PU is operating under transmit power  $P_i$ ;  $s_l$  is the  $l$ th symbol transmitted from PU, which is assumed to follow complex Gaussian distribution with zero mean and unit variance, i.e.,  $s_l \sim \mathcal{CN}(0, 1)$ ;  $\sqrt{\gamma}$  is the channel gain and  $\phi$  is the channel phase;  $n_l$  is the additive noise that follows  $\mathcal{CN}(0, \sigma_n^2)$  for all cases. Note that when  $N = 1$ , the signal model (1) reduces exactly to the traditional binary spectrum sensing in [8], [15], [19].

If we define  $P_0 = 0$  as the *equivalent* transmit power level when PU is absent, then a unified distribution of  $x_l$  can be expressed as

$$x_l \sim \mathcal{CN}(0, \gamma P_i + \sigma_n^2), \quad \forall \mathcal{H}_i, i \in \{0, 1, \dots, N\}. \quad (2)$$

Interestingly, the unknown phase  $\phi$  does not affect the signal statistics.

Let us denote the prior probability of PU taking on each power level as  $\Pr(\mathcal{H}_i)$  while denote the presence state of PU as  $\mathcal{H}_{\text{on}} = \bigcup_{i=1}^N \mathcal{H}_i$ . Obviously,  $\mathcal{H}_{\text{on}}$  has the prior probability  $\Pr(\mathcal{H}_{\text{on}}) = \sum_{i=1}^N \Pr(\mathcal{H}_i)$ . Meanwhile, the absence state of PU, denoted by  $\mathcal{H}_{\text{off}} = \mathcal{H}_0$ , has the probability  $\Pr(\mathcal{H}_{\text{off}}) = \Pr(\mathcal{H}_0)$ .

As one of the first trial studies in MPTP scenario, we make the following assumptions that has also been adopted in many conventional binary spectrum sensing works [6], [8], [13]–[15], [19], [39]:

*Assumption 1:* We assume that SU has the knowledge of transmit power levels of PU as they are normally the deterministic values regulated by the standards. To make the discussion more complete, we further assume that the corresponding priori probabilities  $\Pr(\mathcal{H}_i)$  is known to SU.<sup>2</sup>

*Assumption 2:* We assume that the channel gain  $\sqrt{\gamma}$  is known at SU while the phase  $\phi$  can be unknown.<sup>3</sup>

<sup>2</sup>Or, the user's behavior can be statistically learned from a long run.

<sup>3</sup>Some energy detection algorithms [8], [19], [15] assume AWGN channel, which is the equivalent to assuming  $\sqrt{\gamma} = 1$ . Nevertheless, a new result that could release the necessities of channel information and PU's power level information was recently published in [40].

<sup>1</sup>In some situations [38], it is assumed that multiple PUs each has constant power level, while by detecting the power levels at SU side one may predict the number of PUs that are present.

More considerable situations, e.g., power level uncertainty, partially known channels, unknown channels, etc., could serve as separate future research topics and will not be involved in this paper.

### III. SPECTRUM SENSING STRATEGIES

In MPTP scenario, we define the primary target of spectrum sensing as detecting the presence of PU, while define the secondary target as recognizing the power-level of PU.

*Sensing Strategy-I: Detection Before Recognition:* To achieve the primary task first, we may verify the hypothesis  $\mathcal{H}_{\text{on}}/\mathcal{H}_{\text{off}}$ . Once  $\mathcal{H}_{\text{on}}$  is detected, then the next step is to recognize which  $\mathcal{H}_i, i \geq 1$  is true.

Suppose SU receives a total number of  $M$  samples during the sensing period, denoted as  $\mathbf{x} = [x_1, x_2, \dots, x_M]^T$ . The ratio of the posterior probabilities between two hypothesis can be written as

$$\begin{aligned} \eta(\mathbf{x}) &= \frac{\Pr(\mathcal{H}_{\text{on}}|\mathbf{x})}{\Pr(\mathcal{H}_{\text{off}}|\mathbf{x})} = \frac{\sum_{i=1}^N \Pr(\mathcal{H}_i|\mathbf{x})}{\Pr(\mathcal{H}_0|\mathbf{x})} = \sum_{i=1}^N \frac{p(\mathbf{x}|\mathcal{H}_i) \Pr(\mathcal{H}_i)}{p(\mathbf{x}|\mathcal{H}_0) \Pr(\mathcal{H}_0)} \\ &= \sum_{i=1}^N \frac{\Pr(\mathcal{H}_i)}{\Pr(\mathcal{H}_{\text{off}})} \left( \frac{\sigma_n^2}{\gamma P_i + \sigma_n^2} \right)^M \exp \left\{ \frac{\gamma P_i \sum_{l=1}^M |x_l|^2}{\sigma_n^2 (\gamma P_i + \sigma_n^2)} \right\}. \end{aligned} \quad (3)$$

It is easily seen that  $\eta(\mathbf{x})$  is strictly increasing over  $y \triangleq \sum_{l=1}^M |x_l|^2$ , i.e., the received energy, and the decision can be alternatively made through

$$y \underset{\mathcal{H}_{\text{off}}}{\overset{\mathcal{H}_{\text{on}}}{\geq}} \theta, \quad (4)$$

where  $\theta$  is a pre-determined threshold. Hence, the optimal detector is the energy detector and we may re-represent  $\eta(\mathbf{x})$  as  $\eta(y)$ . The parameter  $\theta$  is used to control the detection performance. For example if  $\theta$  is used to control the false alarm probability, then the detection follows Neyman-Pearson rule; If  $\theta$  is set such that  $\eta(y) = 1$ , then the detection follows the maximum a posterior (MAP) rule.

The probability density functions (pdf) of  $y$  conditioned on  $\mathcal{H}_i, i \in \{0, 1, \dots, N\}$  and  $\mathcal{H}_{\text{on}}$  can be derived as

$$\begin{aligned} p(y|\mathcal{H}_i) &= \frac{y^{M-1} e^{-\frac{y}{\gamma P_i + \sigma_n^2}}}{\Gamma(M) (\gamma P_i + \sigma_n^2)^M}, \quad (5) \\ p(y|\mathcal{H}_{\text{on}}) &= \sum_{i=1}^N p(y|\mathcal{H}_i) \frac{\Pr(\mathcal{H}_i)}{\Pr(\mathcal{H}_{\text{on}})} \\ &= \frac{1}{\Pr(\mathcal{H}_{\text{on}})} \sum_{i=1}^N p(y|\mathcal{H}_i) \Pr(\mathcal{H}_i), \end{aligned} \quad (6)$$

where  $\Gamma(\cdot)$  denotes the gamma function.

Similar to the conventional CR, we could resort to the false alarm probability and the detection probability to describe the performance of the detection, separately calculated as

$$P_{\text{fa}}(\theta) = \Pr(\hat{\mathcal{H}}_{\text{on}}|\mathcal{H}_{\text{off}}) = \int_{\theta}^{\infty} p(y|\mathcal{H}_{\text{off}}) dy = \frac{\Gamma\left(M, \frac{\theta}{\sigma_n^2}\right)}{\Gamma(M)}, \quad (7)$$

$$\begin{aligned} P_{\text{d}}(\theta) &= \Pr(\hat{\mathcal{H}}_{\text{on}}|\mathcal{H}_{\text{on}}) = \frac{1}{\Pr(\mathcal{H}_{\text{on}})} \sum_{i=1}^N \int_{\theta}^{\infty} p(y|\mathcal{H}_i) \Pr(\mathcal{H}_i) dy \\ &= \sum_{i=1}^N \frac{\Gamma\left(M, \frac{\theta}{\gamma P_i + \sigma_n^2}\right)}{\Gamma(M)} \frac{\Pr(\mathcal{H}_i)}{\Pr(\mathcal{H}_{\text{on}})}, \end{aligned} \quad (8)$$

where  $\hat{\mathcal{H}}_{\text{on}}$  represents the detection result, and  $\Gamma(\cdot, \cdot)$  denotes the upper incomplete gamma function.

In the following part, we use  $\theta_{\text{on/off}}$  to denote the decision threshold for the ‘‘on-off’’ status of PU. Moreover, we prefer to compute  $\theta_{\text{on/off}}$  according to MAP criterion for the consistency with the later power recognition part. From (3), we know that  $\theta_{\text{on/off}}$  can be obtained from  $\eta(y) = 1$  and is equivalent to the root of  $\Phi(\theta) = 0$ , where  $\Phi(\theta)$  is simplified from (3) and has the expression

$$\Phi(\theta) \triangleq \sum_{i=1}^N \frac{\Pr(\mathcal{H}_i)}{\left(\frac{\gamma P_i}{\sigma_n^2} + 1\right)^M} e^{\frac{\gamma P_i}{\sigma_n^2 (\gamma P_i + \sigma_n^2)} \theta} - \Pr(\mathcal{H}_0). \quad (9)$$

*Remark 1:* It can be easily checked that  $\frac{\partial \Phi(\theta)}{\partial \theta} > 0$  and  $\Phi(0) < 0$  as long as  $M$  is sufficiently large. Hence  $\Phi(\theta)$  is strictly increasing over  $\theta$  and the solution that makes  $\Phi(\theta) = 0$  (i.e.,  $\theta_{\text{on/off}}$ ) must exist and is definitely unique as well.

If the received energy satisfies  $y > \theta_{\text{on/off}}$ , then PU is claimed to be present and the next step is to recognize which power-level of PU is in use. A natural approach is to formulate multiple hypothesis testing [41] and apply the optimal MAP detection, where for a hypothesis pair  $(\mathcal{H}_i, \mathcal{H}_j), \forall i, j \geq 1, \mathcal{H}_i$  beats  $\mathcal{H}_j$  if

$$\Pr(\mathcal{H}_i|\mathbf{x}, \hat{\mathcal{H}}_{\text{on}}) > \Pr(\mathcal{H}_j|\mathbf{x}, \hat{\mathcal{H}}_{\text{on}}). \quad (10)$$

Here, we use  $\hat{\mathcal{H}}_{\text{on}}$  to denote that the presence detection has been made already.<sup>4</sup> From Bayes rule, there is

$$\Pr(\mathcal{H}_i|\mathbf{x}, \hat{\mathcal{H}}_{\text{on}}) = \frac{p(\mathbf{x}|\mathcal{H}_i, \hat{\mathcal{H}}_{\text{on}}) \Pr(\mathcal{H}_i|\hat{\mathcal{H}}_{\text{on}})}{p(\mathbf{x}|\hat{\mathcal{H}}_{\text{on}})}. \quad (11)$$

Let us define the equivalent region of  $\mathbf{x}$  to  $\{y > \theta_{\text{on/off}}\}$  as  $\{\mathbf{x} \in \mathcal{X}\}$ , and equivalently represent  $\hat{\mathcal{H}}_{\text{on}}$  as  $\mathbf{x} \in \mathcal{X}$ . Then (11) can be rewritten as

$$\begin{aligned} \Pr(\mathcal{H}_i|\mathbf{x}, \hat{\mathcal{H}}_{\text{on}}) &= \frac{1}{p(\mathbf{x}|\mathbf{x} \in \mathcal{X})} \cdot p(\mathbf{x}|\mathcal{H}_i, \mathbf{x} \in \mathcal{X}) \cdot \Pr(\mathcal{H}_i|\mathbf{x} \in \mathcal{X}) \\ &= \frac{1}{p(\mathbf{x}|\mathbf{x} \in \mathcal{X})} \cdot \frac{p(\mathbf{x}|\mathcal{H}_i)}{\Pr(\mathbf{x} \in \mathcal{X}|\mathcal{H}_i)} \\ &\quad \cdot \frac{\Pr(\mathbf{x} \in \mathcal{X}|\mathcal{H}_i) \Pr(\mathcal{H}_i)}{\Pr(\mathbf{x} \in \mathcal{X})} \\ &= \frac{p(\mathbf{x}|\mathcal{H}_i) \Pr(\mathcal{H}_i)}{p(\mathbf{x}|\mathbf{x} \in \mathcal{X}) \Pr(\mathbf{x} \in \mathcal{X})}. \end{aligned} \quad (12)$$

Note that the following equality holds from the definition of probability density function

$$p(\mathbf{x}|\mathcal{H}_i, \mathbf{x} \in \mathcal{X}) = \frac{p(\mathbf{x}|\mathcal{H}_i)}{\Pr(\mathbf{x} \in \mathcal{X}|\mathcal{H}_i)}, \quad \mathbf{x} \in \mathcal{X}. \quad (13)$$

<sup>4</sup>Please note the difference between  $\hat{\mathcal{H}}_{\text{on}}$  and  $\mathcal{H}_{\text{on}}$ .

which is used to derive (12). We place  $\mathbf{x} \in \mathcal{X}$  in (13) to represent that (13) holds only for domain  $\mathbf{x} \in \mathcal{X}$ .

Therefore, the MAP detection (10) is simplified to

$$p(\mathbf{x}|\mathcal{H}_i) \Pr(\mathcal{H}_i) > p(\mathbf{x}|\mathcal{H}_j) \Pr(\mathcal{H}_j), \quad \mathbf{x} \in \mathcal{X}. \quad (14)$$

*Remark 2:* From (14), we know the MAP detection for power levels is not related with how  $\hat{\mathcal{H}}_{\text{on}}$  is detected, i.e., we can apply either MAP detection or Neyman-Pearson to check the presence of PU without affecting (14). Nevertheless, the way to detect the presence of PU will affect the value of  $\Pr(\mathcal{H}_i|\mathbf{x}, \hat{\mathcal{H}}_{\text{on}})$  as seen from (12).

Hence, the MAP detection of the power level can be simply described as

$$\hat{i} = \arg \max_{i \in \{1, \dots, N\}} p(\mathbf{x}|\mathcal{H}_i) \Pr(\mathcal{H}_i), \quad \mathbf{x} \in \mathcal{X}. \quad (15)$$

Let us then define the ratio

$$\begin{aligned} \xi(\mathbf{x}) &= \frac{p(\mathbf{x}|\mathcal{H}_i) \Pr(\mathcal{H}_i)}{p(\mathbf{x}|\mathcal{H}_j) \Pr(\mathcal{H}_j)} \\ &= \frac{\Pr(\mathcal{H}_i)}{\Pr(\mathcal{H}_j)} \left( \frac{\gamma P_j + \sigma_n^2}{\gamma P_i + \sigma_n^2} \right)^M \exp \left\{ \frac{\gamma(P_i - P_j) \sum_{l=1}^M |x_l|^2}{(\gamma P_i + \sigma_n^2)(\gamma P_j + \sigma_n^2)} \right\}. \end{aligned} \quad (16)$$

Obviously,  $\xi(\mathbf{x})$  is purely determined by the energy  $y = \sum_{l=1}^M |x_l|^2$  (other variables are constants). Hence, the energy detector is again optimal when recognizing the power levels of PU, and we can represent  $\xi(\mathbf{x})$  by  $\xi(y)$ . Since  $\xi(\mathbf{x})$  is an increasing function of  $y$  when  $P_i > P_j$ , one can easily know that the decision region of  $\mathcal{H}_i$ , denoted as  $\mathcal{R}(\mathcal{H}_i)$ , must be a continuous region of  $y$ , and  $\mathcal{R}(\mathcal{H}_i)$  must stay on the right side of  $\mathcal{R}(\mathcal{H}_j)$  if  $P_i > P_j$ .

*Theorem 1:* The decision regions of hypothesis  $\mathcal{H}_i, i \in \{1, 2, \dots, N\}$  are computed as in (17), shown at the bottom of the page, where  $\Theta(i, j)$  is a constant defined as

$$\begin{aligned} \Theta(i, j) &\triangleq \frac{(\gamma P_i + \sigma_n^2)(\gamma P_j + \sigma_n^2)}{\gamma(P_i - P_j)} \ln \left[ \left( \frac{\gamma P_i + \sigma_n^2}{\gamma P_j + \sigma_n^2} \right)^M \frac{\Pr(\mathcal{H}_j)}{\Pr(\mathcal{H}_i)} \right]. \end{aligned} \quad (18)$$

*Proof:* Substituting (16) into (14) yields

$$\begin{aligned} &\frac{y \cdot \gamma(P_i - P_j)}{(\gamma P_i + \sigma_n^2)(\gamma P_j + \sigma_n^2)} \\ &> \ln \left[ \left( \frac{\gamma P_i + \sigma_n^2}{\gamma P_j + \sigma_n^2} \right)^M \frac{\Pr(\mathcal{H}_j)}{\Pr(\mathcal{H}_i)} \right], \quad \forall i, j \geq 1. \end{aligned}$$

If  $P_i > P_j$ , i.e.,  $i > j$ , then there is

$$\begin{aligned} y &> \frac{(\gamma P_i + \sigma_n^2)(\gamma P_j + \sigma_n^2)}{\gamma(P_i - P_j)} \\ &\ln \left[ \left( \frac{\gamma P_i + \sigma_n^2}{\gamma P_j + \sigma_n^2} \right)^M \frac{\Pr(\mathcal{H}_j)}{\Pr(\mathcal{H}_i)} \right], \quad \forall i > j; \end{aligned}$$

If  $P_i < P_j$ , i.e.,  $i < j$ , then there is

$$\begin{aligned} y &< \frac{(\gamma P_i + \sigma_n^2)(\gamma P_j + \sigma_n^2)}{\gamma(P_i - P_j)} \\ &\ln \left[ \left( \frac{\gamma P_i + \sigma_n^2}{\gamma P_j + \sigma_n^2} \right)^M \frac{\Pr(\mathcal{H}_j)}{\Pr(\mathcal{H}_i)} \right], \quad \forall i < j. \end{aligned}$$

Then for  $1 < i < N$ , the lower bound of  $\mathcal{R}(\mathcal{H}_i)$  should be  $y > \max_{1 \leq j < i} \Theta(i, j)$  and the upper bound should be  $y < \min_{i < j \leq N} \Theta(i, j)$ . Moreover, the MAP detection is defined on the domain  $\mathbf{x} \in \mathcal{X}$ , i.e.,  $y > \theta_{\text{on/off}}$ , so all decision regions of non-zero power should stay in  $(\theta_{\text{on/off}}, +\infty)$ . Bearing in mind that  $\theta_{\text{on/off}}$  may be greater than  $\max_{1 \leq j < i} \Theta(i, j)$  for some  $i$ , the proof is completed. ■

*Remark 3:* The decision region of  $\mathcal{H}_0$ , i.e., the absence of PU can be expressed in a unified way as

$$\mathcal{R}(\mathcal{H}_0) := y \in (0, \theta_{\text{off}}^{\text{an}}). \quad (19)$$

To further unify our discussion, let us use  $\theta_i, i \in \{1, 2, \dots, N\}$  to denote the threshold between  $\mathcal{R}(\mathcal{H}_{i-1})$  and  $\mathcal{R}(\mathcal{H}_i)$ . Besides, let us also define  $\theta_0 \triangleq 0, \theta_{N+1} \triangleq +\infty$  for completeness, and  $\theta_1 \triangleq \theta_{\text{on/off}}$  for consistence.

*Remark 4:* Compared to the traditional binary sensing that has only one threshold, the new scenario MPTP needs multiple thresholds to separate different power levels, as shown in Fig. 1. The result is generally simple to understand while one contribution of this work lies in deriving the closed-form solutions of all thresholds.

To characterize the performance of the spectrum sensing in MPTP, purely resorting to  $P_d(\theta)$  and  $P_{\text{fa}}(\theta)$  cannot be adequate.

$$\mathcal{R}(\mathcal{H}_i) := \begin{cases} y \in \left( \theta_{\text{off}}^{\text{an}}, \min_{1 < j \leq N} \Theta(1, j) \right) & i = 1 \\ y \in \left( \max \left\{ \theta_{\text{off}}^{\text{an}}, \max_{1 \leq j < i} \Theta(i, j) \right\}, \min_{i < j \leq N} \Theta(i, j) \right), & 1 < i < N \\ y \in \left( \max \left\{ \theta_{\text{off}}^{\text{an}}, \max_{1 \leq j < N} \Theta(N, j) \right\}, +\infty \right), & i = N \end{cases} \quad (17)$$

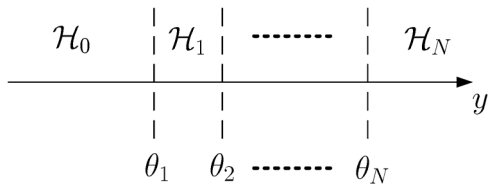


Fig. 1. Multiple power levels detection from multiple thresholds.

We should calculate all the probabilities when SU makes the decision as hypothesis  $\mathcal{H}_j$  while PU is actually transmitting with  $P_i$ , that is

$$\begin{aligned} & \Pr(\mathcal{H}_j|\mathcal{H}_i) \\ &= \int_{\theta_j}^{\theta_{j+1}} p(y|\mathcal{H}_i) dy \\ &= \frac{\Gamma\left(M, \frac{\theta_j}{\gamma P_i + \sigma_n^2}\right)}{\Gamma(M)} - \frac{\Gamma\left(M, \frac{\theta_{j+1}}{\gamma P_i + \sigma_n^2}\right)}{\Gamma(M)}, \quad \forall i, j, \end{aligned} \quad (20)$$

which is also defined as *decision probability* in this paper.

Then,  $P_d$  and  $P_{fa}$  can also be easily obtained from the summations of the corresponding  $\Pr(\mathcal{H}_j|\mathcal{H}_i)$ 's, i.e.,

$$\begin{aligned} P_d &= \frac{1}{\Pr(\mathcal{H}_{on})} \sum_{j=1}^N \sum_{i=1}^N \Pr(\mathcal{H}_j|\mathcal{H}_i) \Pr(\mathcal{H}_i), \\ P_{fa} &= \sum_{i=1}^N \Pr(\mathcal{H}_i|\mathcal{H}_0). \end{aligned}$$

For this new MPTP scenario, we would also introduce a new performance metric

$$P_{dis1} = \frac{1}{\Pr(\mathcal{H}_{on})} \sum_{i=1}^N \Pr(\mathcal{H}_i|\mathcal{H}_i) P(\mathcal{H}_i), \quad (21)$$

named as *discrimination probability*, to describe the recognition capability of our strategy.

#### A. Power-Mask Effect of Sensing Strategy-I

In MPTP scenario, an interesting and special phenomenon happens when the computed lower bound of a specific decision region  $\mathcal{R}(\mathcal{H}_{i_0})$  is greater than the upper bound, i.e.,

$$\max \left\{ \theta_{\frac{on}{off}}, \max_{1 \leq j < i_0} \Theta(i_0, j) \right\} > \min_{i_0 < j \leq N} \Theta(i_0, j) \quad (22)$$

holds for some specific  $1 < i_0 < N$ . Once this happens, the decision region  $\mathcal{R}(\mathcal{H}_{i_0})$  will be empty and the corresponding power level  $P_{i_0}$  can never be detected. We call this new phenomenon in MPTP as *power-mask* effect. Hence in Fig. 1, the number of the thresholds may be less than or equal to  $N$ . Moreover, if  $P_{i_0}$  cannot be detected, we will set  $\theta_{i_0} = \theta_{i_0+1}$  so that the corresponding decision region  $[\theta_{i_0}, \theta_{i_0+1})$  is empty. By doing this, the power-mask effect will have no influence over our previous discussions (20) to (21).

The reasons of power-mask effect may vary a lot, and some intuitive explanations are provided here. First note that the

bounds in the decision region  $\mathcal{R}(\mathcal{H}_i)$  are affected by many parameters, i.e.,  $P_i, \sigma_n^2, \Pr(\mathcal{H}_i)$ . If the priori probability  $\Pr(\mathcal{H}_{i_0})$  is very small, i.e., the power level  $P_{i_0}$  is seldom used by PU. Then  $P_{i_0}$  may easily be “ignored” by SU and is then masked. Another example is that, if  $P_{i_0}$  is close to  $P_{i_0-1}$  and  $P_{i_0+1}$ , and if  $\sigma_n^2$  is relatively large, then it is very likely that  $P_{i_0}$  will be masked by  $P_{i_0-1}$  or  $P_{i_0+1}$  due to the large uncertainty caused by the noise.

We have the following observations about power-mask effect:

- If  $\theta_{on/off} > \max_{1 \leq j < i_0} \Theta(i_0, j) > \min_{i_0 < j \leq N} \Theta(i_0, j)$ , then  $P_{i_0}$  is masked from left by  $P_0$ ;
- If  $\max_{1 \leq j < i_0} \Theta(i_0, j) > \theta_{on/off} > \min_{i_0 < j \leq N} \Theta(i_0, j)$ , then  $P_{i_0}$  is masked from both sides by  $P_{i_0-1}$  and  $P_{i_0+1}$ ;
- The leftmost level  $P_0$  and the rightmost level  $P_N$  cannot be masked and are always detectable.

*Remark 5:* Due to the effect of power-mask, it is then of great interest to check whether the decision regions for  $\mathcal{H}_i$ 's are continuously connected for two consecutively detectable indices  $1 < i_0 < i_0 + 1 < N$ , i.e., for those not masked power level check whether  $\min_{i_0 < j \leq N} \Theta(i_0, j) = \max_{1 \leq j < i_0+1} \Theta(i_0 + 1, j)$  holds. Unfortunately, due to the discrete nature of the power-mask effect, we cannot mathematically prove this property. Nevertheless, it can be easily known that, for any  $y$ , there is always a corresponding decision according to the MAP detection (15). Hence, there should be no gap between any two consecutive decision regions.

A special case that affects the power-mask effect appears when  $\Pr(\mathcal{H}_i) = \Pr(\mathcal{H}_j), \forall i, j \in \{1, 2, \dots, N\}$ .

*Lemma 1:* If  $\Pr(\mathcal{H}_i) = \Pr(\mathcal{H}_j)$  holds for  $\forall i, j \in \{1, 2, \dots, N\}$ , then  $\Theta(i, j)$  is not related to  $\Pr(\mathcal{H}_i)$  and is an increasing function over  $P_j$  for any  $P_i$ .

*Proof:* In this case, the partial derivative of  $\Theta(i, j)$  over  $P_j$  can be computed as

$$\begin{aligned} & \frac{\partial \Theta(i, j)}{\partial P_j} \\ &= \frac{M(\gamma P_i + \sigma_n^2)^2}{\gamma(P_j - P_i)^2} \left\{ \frac{\gamma(P_j - P_i)}{\gamma P_i + \sigma_n^2} - \ln \left[ 1 + \frac{\gamma(P_j - P_i)}{\gamma P_i + \sigma_n^2} \right] \right\}. \end{aligned} \quad (23)$$

Redefining  $z = (\gamma(P_j - P_i))/(\gamma P_i + \sigma_n^2)$ , it can be easily known that  $z \in (-1, +\infty)$  for any  $i$ . We then obtain

$$\frac{\partial \Theta(i, j)}{\partial P_j} = \frac{M\gamma}{z^2} [z - \ln(1 + z)].$$

It is also clear that  $z - \ln(1 + z) \geq 0$  for  $z > -1$ . Hence,  $\frac{\partial \Theta(i, j)}{\partial P_j} \geq 0$  holds<sup>5</sup> for all possible  $P_j$  and  $\Theta(i, j)$  is an increasing function over  $P_j$ . ■

According to Lemma 1, when  $\Pr(\mathcal{H}_i) = \Pr(\mathcal{H}_j), i, j \geq 1$  there is

$$\max_{1 \leq j < i} \Theta(i, j) = \Theta(i, i-1) < \Theta(i, i+1) = \min_{i < j \leq N} \Theta(i, j). \quad (24)$$

<sup>5</sup>Note that for  $z = 0$  the value of  $\frac{\partial \Theta(i, j)}{\partial P_j}$  is obtained from Hospital's rule as  $(M\gamma/2) > 0$ .

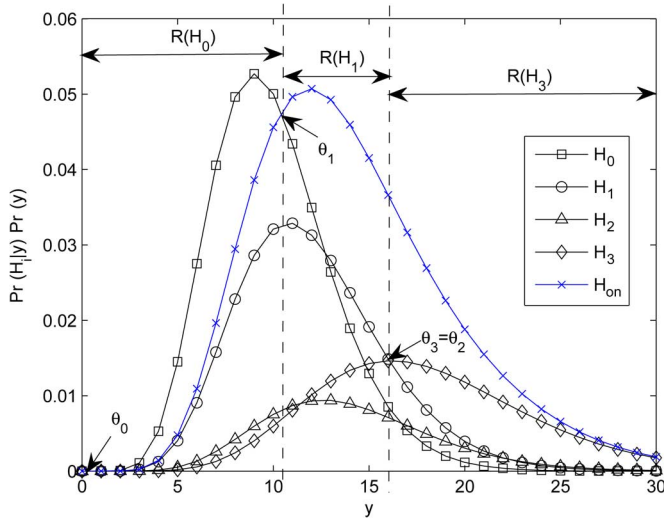


Fig. 2. Illustration for power-mask effect, decision thresholds and decision regions in Sensing Strategy-I.

Therefore, the non-zero power levels cannot mask each other, while the power-mask effect may only happen when  $P_0$  mask the power levels on its right side.

We here present one example to illustrate the power-mask effect in Fig. 2. Three power levels  $P_1, P_2, P_3$  are used and the corresponding priori probabilities  $\Pr(P_i), i = 0, 1, 2, 3$  are taken as 0.4, 0.3, 0.1, 0.2. The decision regions for each hypothesis, i.e.  $\mathcal{H}_i, i = 1, 2, 3$  are demonstrated in the figure. It can be seen that due to the power-mask effect,  $P_2$  cannot be detected. As mentioned earlier, when  $P_2$  is masked, we set the threshold  $\theta_2 = \theta_3$  so that the decision region  $\mathcal{R}(\mathcal{H}_2)$  is empty.

*Remark 6:* When power-mask happens for a specific  $\mathcal{H}_{i_0}$ , it means that the  $\Pr(\mathcal{H}_{i_0}|\mathbf{x})$  cannot beat any other  $\Pr(\mathcal{H}_i|\mathbf{x}), i \neq i_0$  but  $\Pr(\mathcal{H}_{i_0}|\mathbf{x})$  could still possess non-zero value. Hence, it is possible to design some sophisticated approach which considers this “soft” information and removes the power-mask effect. Nevertheless, the corresponding study is out of the scope of this paper and will be left for future research.

### B. Sensing Strategy-II: Recognition Before Detection

Another reasonable approach is to directly detect the power level of PU (including the zero power level) from standard multiple hypothesis testing approach if one treats  $P_0$  as an *equivalent* power level as other non-zero  $P_i, i \geq 1$ . The presence or the absence can be immediately found after the power level index is detected.

From MAP based multiple hypothesis testing, the optimal detection can be stated as

$$\hat{i} = \arg \max_{i \in \{0, 1, \dots, N\}} \Pr(\mathcal{H}_i|\mathbf{x}) = \arg \max_{i \in \{0, 1, \dots, N\}} \Pr(\mathbf{x}|\mathcal{H}_i)p(\mathcal{H}_i). \quad (25)$$

Since the expression of (25) is, mathematically, the same as (15) but includes one more index 0, the previous results can be

immediately modified here. For example, the decision region is computed as

$$\mathcal{R}(\mathcal{H}_i) := \begin{cases} y \in \left(0, \min_{0 < j \leq N} \Theta(0, j)\right) & i = 0 \\ y \in \left(\max_{0 \leq j < i} \Theta(i, j), \min_{i < j \leq N} \Theta(i, j)\right), & 0 < i < N \\ y \in \left(\max_{0 \leq j < N} \Theta(N, j), +\infty\right), & i = N \end{cases} \quad (26)$$

where  $\Theta(i, j)$  is given in (18). Hence, the power-mask effect also exists if  $\max_{1 \leq j < i_0} \Theta(i_0, j) > \min_{i_0 < j \leq N} \Theta(i_0, j)$  for some  $i_0$ . Let us use  $\phi_i$  to represent the thresholds separating  $\mathcal{R}(\mathcal{H}_{i-1})$  and  $\mathcal{R}(\mathcal{H}_i)$  in sensing strategy-II. Then the decision probability is given in (20) with  $\theta_i$  being replaced by  $\phi_i$ .

Due to the similarity between the decision regions of the two sensing approaches, i.e., (17) and (26), one natural and interesting question arises: are all the thresholds or parts of the thresholds the same in these two sensing approaches? Of all the thresholds, the first one that separates the absence decision and presence decision is of particular importance. We then provide the following lemma.

*Lemma 2:* Let  $\phi_{\text{on/off}} \triangleq \phi_1$  be the threshold between “absence” and “presence” of PU in sensing strategy-II. There is  $\phi_{\text{on/off}} > \theta_{\text{on/off}}$ .

*Proof:* In sensing strategy-II,  $\phi_{\text{on/off}}$  can be computed from  $\Pr(\mathcal{H}_0|\mathbf{x}) = \Pr(\mathcal{H}_{j_0}|\mathbf{x})$  where  $j_0$  is the index of the first non-zero power that is not masked. From (16), we can derive a unique

$$\phi_{\text{on/off}} = \frac{\sigma_n^2 (\gamma P_{j_0} + \sigma_n^2)}{\gamma P_{j_0}} \cdot \ln \left[ \frac{\Pr(\mathcal{H}_0)}{\Pr(\mathcal{H}_{j_0})} \left( \frac{\gamma P_{j_0}}{\sigma_n^2} + 1 \right)^M \right]. \quad (27)$$

Let us then compute  $\Phi(\phi_{\text{on/off}})$  from (9), which yields

$$\begin{aligned} \Phi(\phi_{\text{on/off}}) &= \sum_{i=1}^N \frac{\Pr(\mathcal{H}_i)}{\left(\frac{\gamma P_i}{\sigma_n^2} + 1\right)^M} e^{\frac{\gamma P_i}{\sigma_n^2 (\gamma P_i + \sigma_n^2)} \cdot \phi_{\text{on/off}}} - \Pr(\mathcal{H}_0) \\ &= \sum_{i \neq j_0}^N \frac{\Pr(\mathcal{H}_i)}{\left(\frac{\gamma P_i}{\sigma_n^2} + 1\right)^M} e^{\frac{\gamma P_i}{2\sigma_n^2 (\gamma P_i + \sigma_n^2)} \cdot \phi_{\text{on/off}}} > 0. \end{aligned}$$

Since  $\Phi(\theta_{\text{on/off}}) = 0$  and  $\Phi(\theta)$  is an increasing function over  $\theta$ , it is obvious that  $\phi_{\text{on/off}} > \theta_{\text{on/off}}$ . ■

*Remark 7:* Since we treat  $P_0$  as an equal state as other non-zero  $P_i$ 's in sensing strategy-II, the discrimination probability here could also be defined as

$$P_{\text{dis2}} = \sum_{i=0}^N \Pr(\mathcal{H}_i|\mathcal{H}_i)P(\mathcal{H}_i), \quad (28)$$

which is the value of correct detection and will be maximized by the MAP based multiple hypothesis testing [41].

### C. Power-Mask Effect of Sensing Strategy-II

Lemma 2 suggests that sensing strategy-II claims more absence of PU than sensing strategy-I. In fact, when we compare (26) and (17), it is easy to note that  $\max_{0 \leq j < i} \Theta(i, j)$  may be greater than  $\max_{1 \leq j < i} \Theta(i, j)$  if  $\Theta(i, 0)$  is bigger than other

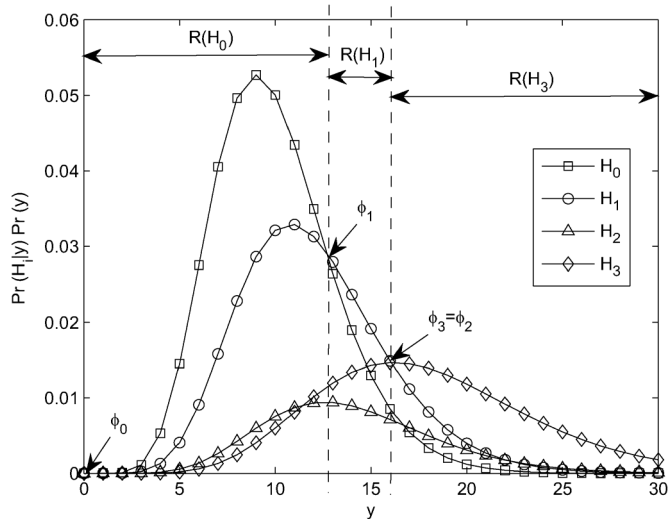


Fig. 3. Illustration for power-mask effect, decision thresholds and decision regions in Sensing Strategy-II.

$\Theta(i, j), 0 < j < i$  for a specific  $i_0$ . When this happens, it also means that  $P_0$  masks all power levels smaller than  $P_{i_0+1}$  in sensing strategy-II. Hence, the lower bound of  $\mathcal{R}(\mathcal{H}_{i_0})$  in sensing strategy-II is  $\phi_{\text{on/off}}$  while the lower bound of  $\mathcal{R}(\mathcal{H}_{i_0})$  in sensing strategy-I is  $\max\{\theta_{\text{on/off}}, \max_{1 \leq j < i_0} \Theta(i_0, j)\}$ . Combining Lemma 2, it is then clear that the lower bound of  $\mathcal{R}(\mathcal{H}_{i_0})$  in sensing strategy-II is bigger than that of sensing strategy-I. Nevertheless, when  $\Theta(i, 0)$  is not the dominant one in  $\max_{0 \leq j < i} \Theta(i, j)$ , i.e.,  $\Pr(\mathcal{H}_0|\mathbf{x})$  is not the biggest among all  $\Pr(\mathcal{H}_i|\mathbf{x})$ , then the thresholds for the two sensing strategies are the same.

*Remark 8:* In sensing strategy-II, once again, both  $P_0$  and  $P_N$  cannot be masked and are always detectable. Also, a special case happens when  $\Pr(\mathcal{H}_i) = \Pr(\mathcal{H}_j), \forall i, j$ . In this situation, (24) holds and  $\Theta(i, j)$  is not related with  $\Pr(\mathcal{H}_i)$ . Hence, except for  $\theta_{\text{on/off}}$  and  $\phi_{\text{on/off}}$ , all the other thresholds from both sensing strategies are the same, i.e.,  $\theta_i = \phi_i, i \geq i_0$  where  $i_0$  is the first power level that is not masked by  $P_0$ .

We then present another example to demonstrate the power mask effect of sensing strategy-II in Fig. 3, where all the settings are the same as those in Fig. 2. Since we directly plot  $\Pr(\mathcal{H}_i|y)\Pr(y)$ , the cross points of different curves will be  $\phi_i$  immediately. Again, the power-level  $P_2$  is masked. However, the decision region of  $\mathcal{H}_0$ , i.e.,  $\mathcal{R}(\mathcal{H}_0)$  in sensing strategy-II is larger than that in sensing strategy-I, which matches Lemma 2 well. Meanwhile, there is  $\phi_3 = \theta_3$ , which matches our previous analysis.

#### D. Fundamental Rationale Behind Two Spectrum Sensing Strategies

After presenting two different spectrum sensing approaches, both seemingly reasonable, a natural and interesting question arises: which one is better and why? Let us answer from MAP detection point of view.

When MAP detection is applied in the first step of sensing strategy-I, the so obtained  $\theta_{\text{on/off}}$  is optimal in terms of minimizing the following error

$$\begin{aligned} & \arg \min \Pr(\mathcal{H}_{\text{off}}|\mathcal{H}_{\text{on}})\Pr(\mathcal{H}_{\text{on}}) + \Pr(\mathcal{H}_{\text{on}}|\mathcal{H}_{\text{off}})\Pr(\mathcal{H}_{\text{off}}) \\ & = \arg \min \sum_{i=1}^N \Pr(\mathcal{H}_0|\mathcal{H}_i)\Pr(\mathcal{H}_i) + \sum_{i=1}^N \Pr(\mathcal{H}_i|\mathcal{H}_0)\Pr(\mathcal{H}_0). \end{aligned} \quad (29)$$

On the other side, applying MAP in sensing strategy-II that directly detects the power level is optimal in terms of minimizing the following error

$$\arg \min \sum_{j \neq i} \Pr(\mathcal{H}_j|\mathcal{H}_i)\Pr(\mathcal{H}_i). \quad (30)$$

Therefore, sensing strategy-I does not count in the errors  $\Pr(\mathcal{H}_j|\mathcal{H}_i), i, j \geq 1$ , i.e., the errors among non-zero power levels, when detecting the presence of PU. On the other hand, sensing strategy-II takes into account of all error probabilities  $\Pr(\mathcal{H}_j|\mathcal{H}_i), \forall i, j$  all at once. Hence, the two strategies are, in fact, applied for different criteria and thus are incomparable, although look very similar.

In general, if identifying the presence of PU is a more important target than discriminating the power levels of PU, then sensing strategy-I is preferable. However, for sensing based sharing CR scheme where the penalty due to the wrong interference protection is very high, then sensing strategy-II could be preferable.

*Remark 9:* In fact, both the strategies falls into the Baye's Risk based multiple hypothesis test where the optimization criterion is to minimize

$$\sum_{i,j} C_{i,j} \Pr(\mathcal{H}_j|\mathcal{H}_i)\Pr(\mathcal{H}_i), \quad (31)$$

where  $C_{i,j}$  is the price or the cost for detecting  $\mathcal{H}_j$  when  $\mathcal{H}_i$  is true. Obviously, the value of  $C_{i,j}$  should be set according to practical requirements and can be different in different applications.

#### IV. COOPERATIVE SPECTRUM SENSING SCHEMES

Now assume that there are  $K$  SUs that cooperative to detect the presence of PU as well as the power levels. After performing the local spectrum sensing, either with sensing strategy-I or sensing strategy-II, SU- $k$  makes its own decision as  $\mathcal{H}_{i_k}$  and then forwards the index  $j_k$  to the fusion center who combines these results into a  $K \times 1$  vector  $\mathbf{b} = [j_1, j_2, \dots, j_K]$ . Note that, the "on-off" information of PU is automatically embedded in the index  $j_k$  and does not need separate feedback. The probability of any specific  $\mathbf{b}$  can be easily computed as

$$\Pr(\mathbf{b}|\mathcal{H}_i) = \prod_{k=1}^K \Pr(\mathcal{H}_{j_k}^{(k)}|\mathcal{H}_i), \quad (32)$$

where the superscript  $(\cdot)^{(k)}$  represents the user index, while the total number of possible  $\mathbf{b}$  can be calculated as  $(N+1)^K$ .

Unlike the conventional cooperative spectrum sensing [13]–[16] where PU is assumed to have only “on-off” status, the cooperative sensing in MPTP scenario exhibits many differences, for which the existing *Logic-AND* (LA), *Logic-OR* (LR) and their general form *k-out-of-N* (KON) based fusion rules are no longer applicable. For example, assuming PU has two power levels and the local detection results from five SUs are  $\{0,0,1,1,2\}$ . In this case, AND and OR rule cannot be used since they are defined over binary states. Similarly, the *k-out-of-N* rule cannot be applied either. Say, if  $k = 2$ , then both 0 and 1 get two votes and the final decision cannot be made. While if  $k = 3$  is used, then no decision can be made because no power level (including 0 power) gets 3 votes.

Hence, it is necessary to design new cooperative sensing schemes for MPTP scenario. We here propose two different fusion rules, i.e., the modified majority fusion and the optimal fusion.

#### A. Modified Majority Decision Fusion (MMDF)

From a given  $\mathbf{b}$ , we can immediately formulate a voting pool  $\vec{d} = (d_0, \dots, d_N)$ , where  $d_i$  denotes the number of SUs that claim  $\mathcal{H}_i$ . Obviously, there is  $\sum_{i=0}^N d_i = K$ . Define the mapping function from  $\mathbf{b}$  to  $\vec{d}$  as  $\mathcal{M}(\mathbf{b}) = \vec{d}$ , which can be easily obtained from an offline manner once  $N$  and  $K$  are fixed. It is not difficult to find that the total number of possible  $\vec{d}$  is  $\binom{K+N}{N} = \frac{(K+N)!}{K!N!}$ . The probability of any specific  $\vec{d}$  can then be computed as

$$\Pr(\vec{d}|\mathcal{H}_i) = \sum_{\mathbf{b}:\mathcal{M}(\mathbf{b})=\vec{d}} \Pr(\mathbf{b}|\mathcal{H}_i). \quad (33)$$

*Remark 10:* If we assume all SUs have the same detection capability, i.e., the same  $\Pr(\mathcal{H}_j|\mathcal{H}_i)$ ,<sup>6</sup> then the expression of  $\Pr(\vec{d}|\mathcal{H}_i)$  can be simplified into (34) shown on top of this page.

$$\begin{aligned} \Pr(\vec{d}|\mathcal{H}_i) &= \binom{K}{d_0} \Pr(\mathcal{H}_0|\mathcal{H}_i)^{d_0} \\ &\times \binom{K-d_0}{d_1} \Pr(\mathcal{H}_1|\mathcal{H}_i)^{d_1} \dots \\ &\times \binom{K-\sum_{l=0}^{N-1} d_l}{d_N} \Pr(\mathcal{H}_N|\mathcal{H}_i)^{d_N} \\ &= \frac{K!}{\prod_{l=0}^N d_l!} \prod_{n=0}^N \Pr(\mathcal{H}_n|\mathcal{H}_i)^{d_n} \end{aligned} \quad (34)$$

A simple yet reasonable way to make the decision fusion is to count the majority claims from SUs, i.e., pick

$$\hat{i} = \arg \max_i d_i. \quad (35)$$

However, a special case happens when  $d_0 = \arg \max_i d_i$  while  $d_0 < \sum_{i=1}^N d_i$ , for example  $\vec{d} = (0, 0, 1, 2, 3)$ . In this example, (35) will output  $\hat{i} = 0$  and claims the absence of PU, but in fact more users claim the presence of PU. Therefore, we

<sup>6</sup>For example, the received signal at each SU experiences almost identical path loss and noise variance [13], which approximately holds when the distance between any two SUs is small compared to the distance from PU to any one of the SUs.

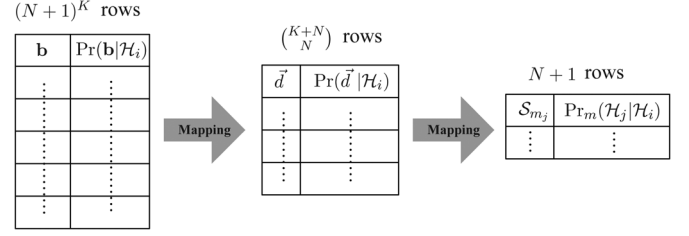


Fig. 4. Illustration of mapping from  $\mathbf{b}$  to  $\vec{d}$  and then to  $\mathcal{S}_{m_j}$ .

should check the presence of PU before applying the majority rule when the primary target is to detect the “on-off” status of PU, and we call this newly designed rule as Modified Majority Decision Fusion (MMDF).

Let us define  $d_{\text{off}} \triangleq d_0$  and  $d_{\text{on}} \triangleq \sum_{i=1}^N d_i$ . Then, the decision rule can be expressed as

$$d_{\text{on}} \underset{\mathcal{H}_{\text{off}}}{\overset{\mathcal{H}_{\text{on}}}{\geq}} d_{\text{off}}, \quad (36)$$

which can be simplified as

$$d_{\text{on}} \underset{\mathcal{H}_{\text{off}}}{\overset{\mathcal{H}_{\text{on}}}{\geq}} K/2. \quad (37)$$

Note that, when  $K$  is even and when  $d_{\text{on}} = K/2$ , the final decision can either be made as “on” or “off” because they are equally probable. In the rest of the discussion, we claim “on” if  $d_{\text{on}} = K/2$ .

If PU is detected to be present, the next step is to discriminate which power level is in use by majority law

$$\hat{i} = \arg \max_{i \geq 1} d_i. \quad (38)$$

Let us then define  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$  as the decision probability of MMDF. It can be computed that

$$\Pr_m(\mathcal{H}_j|\mathcal{H}_i) = \sum_{\vec{d} \in \mathcal{S}_{m_j}} \Pr(\vec{d}|\mathcal{H}_i), \quad (39)$$

where the set  $\mathcal{S}_{m_j}$  is defined as

$$\mathcal{S}_{m_j} = \begin{cases} \left\{ \vec{d} \mid d_0 > \frac{K}{2} \right\} & \text{if } j = 0 \\ \left\{ \vec{d} \mid d_j = \max\{d_1, \dots, d_N\}, d_0 \leq \frac{K}{2} \right\} & \text{if } j \geq 1 \end{cases} \quad (40)$$

*Remark 11:* There exist special cases when more than one state  $\hat{i}$  simultaneously achieve the maximum number of votes. In this situation, one can choose any of them as the final decision since they are equally probable. In this paper, we always choose the largest value of  $\hat{j}$  as the final decision if this special case happens, and all the theoretical and numerical results in the rest of the paper are based on this consideration.

The set  $\mathcal{S}_{m_j}$  can be obtained from a mapping function in an offline manner. Hence, we can easily build tables and mapping functions as illustrated in Fig. 4, and compute  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$  a priori. Fortunately, a more explicit expression for  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$  in majority voting can be derived from the following theorem.



*Theorem 2:* The MMDF has the decision probability as (41) shown on the bottom of the page, where  $\lceil \cdot \rceil$ , and  $\lfloor \cdot \rfloor$  denote the ceiling function and the floor function, respectively. Moreover,  $\alpha_n$  and  $\beta_n$  are defined as

$$\alpha_n = \begin{cases} K - \sum_{i=0}^{n-1} d_i - (N-n)d_j + N - j & \text{if } 1 \leq n < j \\ K - \sum_{i=0}^{n-1} d_i - (N-n)d_j + N - n & \text{if } j < n \leq N \end{cases}$$

$$\beta_n = \begin{cases} K - \sum_{i=0}^{n-1} d_i - d_j & \text{if } 1 \leq n < j \\ K - \sum_{i=0}^{n-1} d_i & \text{if } j < n \leq N \end{cases} \quad (42)$$

*Proof:* To calculate the decision probability from (39), we need to find all candidates  $\vec{d}$  in  $\mathcal{S}_{m_j}$ . In other words, we need to determine the range of elements  $d_j$ 's,  $j = 0, 1, \dots, N$  in  $\vec{d}$ .

Let us start from  $j \geq 1$ . As shown in  $\mathcal{S}_{m_j}$ ,  $\vec{d}$  must satisfy  $d_0 \leq K/2$ , so the range of  $d_0$  should be from 0 to  $\lfloor K/2 \rfloor$ . Moreover, since  $d_j = \max\{d_1, \dots, d_N\}$ , the lower bound of  $d_j$  must be no less than  $K - d_0/N$ , otherwise there will always be another  $d_{j_1} \neq d_j$  but satisfies  $d_{j_1} > d_j$ . Therefore the range of  $d_j$  is from  $\lceil K - d_0/N \rceil$  to  $K - d_0$ . We then separately determine the range of  $d_n$  for  $1 \leq n < j$  and  $j < n \leq N$  respectively by assigning the remaining votes from  $K$  SUs, which is  $K - d_0 - d_j$ .

- 1) When  $n < j$ , the upper bound of  $d_n$  must be less than or equal to the unassigned votes of  $K$  SUs, which is  $K$  minus all the values that have already been assigned to  $d_i, 0 \leq i < n$  and  $d_j$ , i.e.,  $K - \sum_{i=0}^{n-1} d_i - d_j$ . Bearing in mind the constraint  $d_n \leq d_j$ , the upper bound can be expressed as

$$d_n^{\text{upper}} = \min \left\{ d_j, K - \sum_{i=0}^{n-1} d_i - d_j \right\}, \quad n < j.$$

As for the lower bound,  $d_n$  should not be too small to allow any other undetermined  $d_k$  ( $n < k \leq N$  and  $k \neq j$ ) be greater than  $d_j$ . The extreme case happens when all the undetermined  $d_k$ 's get their highest values, i.e.,  $d_k$  being equal to  $d_j$  for  $n < k < j$  while  $d_k$  being  $d_j - 1$  for  $j < k \leq N$ . Then the summation of these  $d_k$ 's is  $(j - n - 1)d_j + (N - j)(d_j - 1)$ . Combining this result as well as the constraint  $d_n \geq 0$ , we get the lower bound of  $d_n$  as

$$d_n^{\text{lower}} = \max \left\{ K - d_j - \sum_{i=0}^{n-1} d_i - (j - n - 1)d_j - (N - j)(d_j - 1), 0 \right\}, \quad n < j.$$

- 2) When  $n > j$ , the maximum value of  $d_n$  can only be  $d_j - 1$ , thus the upper bound of the summand changes to

$$d_n^{\text{upper}} = \min \left\{ d_j - 1, K - \sum_{i=0}^{n-1} d_i \right\}, \quad n > j.$$

Similar to the previous discussion, the maximum summation of all the undetermined  $d_k$  ( $n < k \leq N$ ) in this situation is  $(N - n)(d_j - 1)$ , so the lower bound is

$$d_n^{\text{lower}} = \max \left\{ K - \sum_{i=0}^{n-1} d_i - (N - n)(d_j - 1), 0 \right\}, \quad n > j.$$

When we use  $\alpha_n$  and  $\beta_n$  in (42) to simplify the expression of the range of  $d_n$ 's, the equality (41a) for  $j \geq 1$  is proved.

As for  $j = 0$ , the only constraint for  $\vec{d}$  is  $d_0 > K/2$  from  $\mathcal{S}_{m_j}$ , so the summation range of  $d_0$  must be from  $\lfloor K/2 \rfloor + 1$  to  $N$ . All the others  $d_n, n \in \{1, 2, \dots, N\}$  can be freely chosen as long as  $\sum_{i=0}^N d_i = K$ . If we assign the values for  $d_i$ 's one by one, then for any  $d_n$ , its lowest possible value is 0 while its highest possible value is  $K - \sum_{i=0}^{n-1} d_i$ . Note that,  $d_N$  is fixed when all the previous  $d_i$ 's,  $0 \leq i < N$  are chosen and does not need to be included in the summand. Then, the equality (41b) for  $j = 0$  is proved. ■

Once  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$  is derived, then the false alarm, the detection probability, as well as the discrimination probability for majority cooperation can be immediately obtained as the summation of the corresponding  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$ 's.

In Fig. 5, we provide an example to verify several analytically derived  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$  with  $N = 4, K = 5$  and an average SNR = -12 dB. It is clearly seen that the numerical results match the theoretical ones very well. It needs to be mentioned that  $\Pr_m(\mathcal{H}_1|\mathcal{H}_0)$  is close to zero because the number of samples starts from 3000, which is sufficient to provide a very small false alarm provability.

*Remark 12:* Though the majority decision fusion rule has been widely accepted in many research areas, the analytical approach to study its performance, e.g., obtaining  $\Pr_m(\mathcal{H}_j|\mathcal{H}_i)$  from  $\Pr(\mathcal{H}_j|\mathcal{H}_i)$ , has not been fully discussed to the best of the authors' knowledge.

### B. Optimal Decision Fusion (ODF)

Though MMDF is very simple and effective, it does have some drawbacks which limit the performance of cooperative decision. For example, when  $\Pr(\mathcal{H}_i) \gg \Pr(\mathcal{H}_j)$ , even if the

$$\Pr_m(\mathcal{H}_j|\mathcal{H}_i) = \begin{cases} \sum_{d_0=0}^{\lfloor \frac{K}{2} \rfloor} \sum_{d_j=\lceil \frac{K-d_0}{N} \rceil}^{K-d_0} \min\{d_j, \beta_1\} \dots \sum_{d_{j-1}=\max\{0, \alpha_{j-1}\}}^{\min\{d_j, \beta_{j-1}\}} \sum_{d_{j+1}=\max\{0, \alpha_{j+1}\}}^{\min\{d_j-1, \beta_{j+1}\}} \dots \\ \sum_{d_N=\max\{0, \alpha_N\}}^{\min\{d_j-1, \beta_N\}} \Pr(\vec{d}|\mathcal{H}_i), & j \geq 1 \end{cases} \quad (41a)$$

$$\sum_{d_0=\lfloor \frac{K}{2} \rfloor + 1}^K \sum_{d_1=0}^{K-d_0} \sum_{d_2=0}^{K-d_0-d_1} \dots \sum_{d_{N-1}=0}^{K-\sum_{l=0}^{N-2} d_l} \Pr(\vec{d}|\mathcal{H}_i), \quad j = 0 \quad (41b)$$

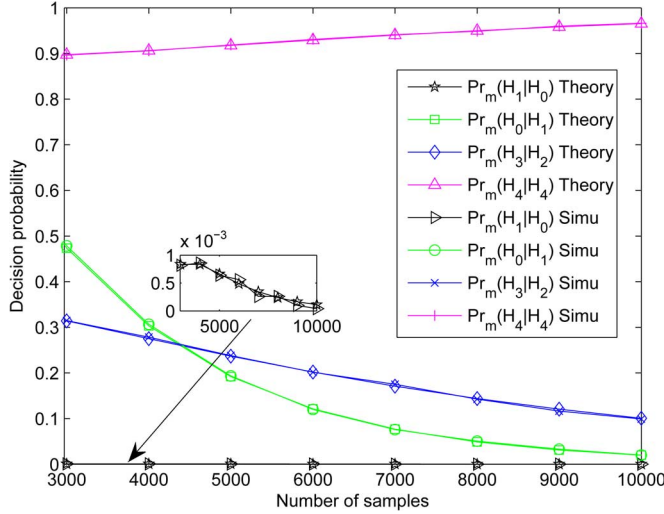


Fig. 5. Theoretical analysis and numerical results for decision probability under MMDF.

detection result is  $d_i < d_j$ , it is still possible that  $\mathcal{H}_i$  is truer than  $\mathcal{H}_j$ . The reason is that majority decision is a type of “hard” decision and is not “soft” enough to count in the priori probability of each hypothesis.

From the probabilistic point of view, the optimal decision fusion with the observation  $\vec{d}$  should follow MAP criterion [41]. Similar to majority decision fusion, we need to first to make a decision about the presence of PU before recognizing the power levels, i.e.,

$$\Pr(\mathcal{H}_{\text{off}}|\vec{d}) \underset{\mathcal{H}_{\text{on}}}{\overset{\mathcal{H}_{\text{off}}}{\geq}} \Pr(\mathcal{H}_{\text{on}}|\vec{d}). \quad (43)$$

From Bayes rule, there is

$$\Pr(\mathcal{H}_n|\vec{d}) = \frac{\Pr(\vec{d}|\mathcal{H}_n)\Pr(\mathcal{H}_n)}{\Pr(\vec{d})}. \quad (44)$$

Hence, (43) can be simplified to

$$\Pr(\vec{d}|\mathcal{H}_0)\Pr(\mathcal{H}_0) \underset{\mathcal{H}_{\text{on}}}{\overset{\mathcal{H}_{\text{off}}}{\geq}} \sum_{i=1}^N \Pr(\vec{d}|\mathcal{H}_i)\Pr(\mathcal{H}_i). \quad (45)$$

If PU is detected to be present, we continue to recognize the power level of PU. Following the similar steps from (10) to (14), the detection rule is

$$\hat{i} = \arg \max_{i \geq 1} \Pr(\mathcal{H}_i|\vec{d}) = \arg \max_{i \geq 1} \Pr(\vec{d}|\mathcal{H}_i)\Pr(\mathcal{H}_i). \quad (46)$$

The decision probability of the optimal decision fusion can be expressed as

$$\Pr_o(\mathcal{H}_j|\mathcal{H}_i) = \sum_{\vec{d} \in \mathcal{S}_{o_j}} \Pr(\vec{d}|\mathcal{H}_i) \quad (47)$$

where the set  $\mathcal{S}_{o_j}$  is defined as

$$\mathcal{S}_{o_j} = \begin{cases} \left\{ \vec{d} \mid \text{those } \vec{d} \text{ that claim} \right. \\ \left. \text{absence from (43)} \right\} & \text{if } j = 0 \\ \left\{ \vec{d} \mid \text{those } \vec{d} \text{ that result in } \hat{i} = j \text{ in (46)} \right. \\ \left. \text{but claim presence from (43)} \right\} & \text{if } j \geq 1 \end{cases}$$

The elements in  $\mathcal{S}_{o_j}$  is an implicit function of  $\Pr(\mathcal{H}_i)$  and  $\Pr(\mathcal{H}_j|\mathcal{H}_i)$ , which makes it difficult to obtain an explicit expression of  $\Pr_o(\mathcal{H}_j|\mathcal{H}_i)$ . Nevertheless,  $\Pr_o(\mathcal{H}_j|\mathcal{H}_i)$  is the summation of those  $\Pr(\vec{d}|\mathcal{H}_i)$  whose  $\vec{d}$  could result in the decision of  $j$ , while these  $\vec{d}$  can be found from a predetermined mapping in an offline manner, as did in Fig. 4.

*Remark 13:* Since  $\Pr(\mathcal{H}_i)$  and  $\Pr(\mathcal{H}_j|\mathcal{H}_i)$  are real continuous values, the probability for obtaining more than one maximum index from (46) is 0, and the corresponding discussion is not necessary.

*Remark 14:* If we assume the same detection capability for all SUs, as did in [13], and apply (34), then a more concise form of the decision (46) can be obtained as

$$\begin{aligned} \hat{i} &= \arg \max_{i \geq 1} \Pr(\mathcal{H}_i) \prod_{n=0}^N \Pr(\mathcal{H}_n|\mathcal{H}_i)^{d_n} \\ &= \arg \max_{i \geq 1} \log \Pr(\mathcal{H}_i) + \sum_{n=0}^N d_n \log \Pr(\mathcal{H}_n|\mathcal{H}_i). \end{aligned} \quad (48)$$

## V. SIMULATIONS

In this section, we resort to numerical examples to evaluate the proposed studies. Four levels of primary transmit power are assumed, while the corresponding prior probabilities are set as  $\Pr(\mathcal{H}_0) = 0.5$ , and  $\Pr(\mathcal{H}_i) = 0.125, i = 1, 2, 3, 4$ . The noise variances is taken as unit. The power levels satisfy  $P_1 : P_2 : P_3 : P_4 = 3 : 5 : 7 : 9$ , and the average SNR is defined as  $(1/4) \sum_{i=1}^4 P_i / \sigma_n^2$ .

### A. Sensing With a Single SU

We first demonstrate the performance loss of the traditional binary spectrum sensing algorithm when PU is actually working with multiple power levels. Since the conventional binary sensing algorithm has to assume one constant power level, it cannot be directly applied to MPTP scenario. We then consider the following three different cases: 1) binary sensing algorithm assumes the lowest power level of PU as the constant power level in its algorithm; 2) binary sensing algorithm assumes the average power level of PU as the constant power level in its algorithm; 3) binary sensing algorithm assumes the highest power level of PU as the constant power level in its algorithm; In Fig. 6, we choose the figure of the merit as  $P_{\text{miss}} + P_f$ , where  $P_{\text{miss}} = 1 - P_d$  is the miss detection probability, and compare the performance of different algorithms. The PU’s average is taken as -12 dB. It can be seen that the conventional binary spectrum sensing always suffers from performance loss due to the model mismatch, which supports our consideration of MPTP scenario.

In Fig. 7, we evaluate the performance of detecting the presence of PU versus the number of samples for the proposed sensing strategies. It is seen that sensing strategy-I works better than sensing strategy-II, especially when the sampling number is small, which matches our discussion in Section III-E that  $\theta_{\text{on/off}} < \phi_{\text{on/off}}$ . Nevertheless, the gaps between the two sensing strategies reduce when the number of samplings becomes larger or the PU’s SNR becomes higher. This implies that when the sensing conditions become better, then difference

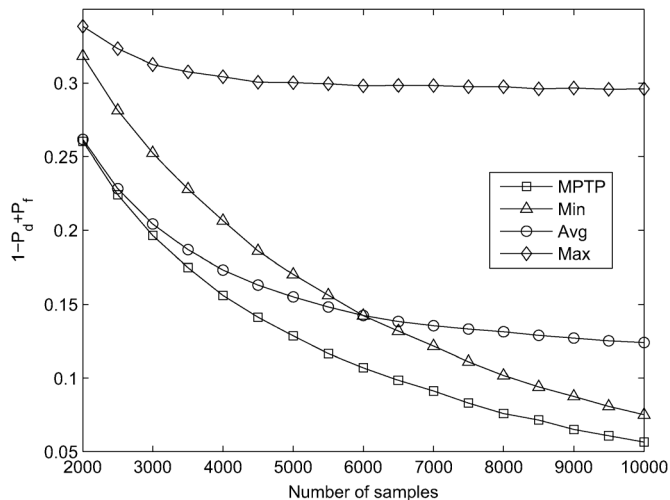


Fig. 6. The performance loss of binary spectrum sensing compared to the proposed algorithm at average SNR = -12 dB.

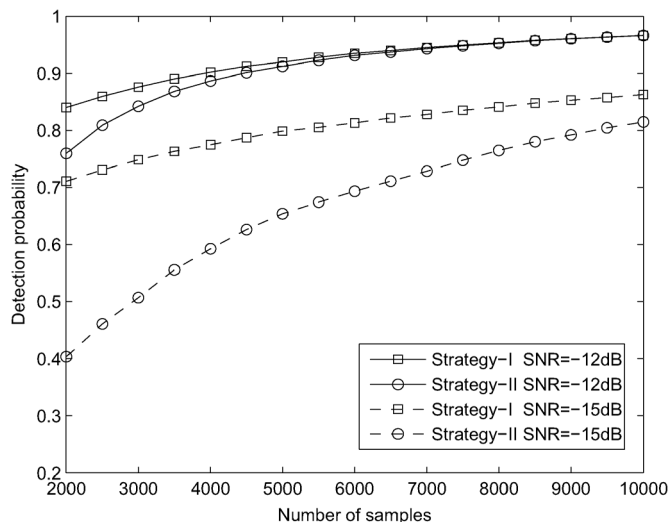


Fig. 7. The detection probability of local SU versus the number of samples.

of the two sensing strategies gradually diminishes and the choice of sensing strategies becomes less important.

On the other hand, Fig. 8 displays the performance of discriminating the power level versus the number of samples of the proposed sensing strategies. The discrimination probability follows the definition in (28), i.e., we treat absence as an equivalent power level with 0 value. From Fig. 8, we see that sensing strategy-II works slightly better than sensing strategy-I but the difference diminishes when SNR becomes higher. This phenomenon right fits our analysis in Section III-E that sensing strategy-II takes into account all error cases.

The average detection error probabilities versus the number of samples for different power offsets are displayed in Fig. 9 with the average SNR = -12 dB, where  $\delta$  denotes the offset of between the true power level and the detected power level. For example, detection error probability with  $\delta = 1$  is the summation of those  $\Pr(\mathcal{H}_j|\mathcal{H}_i)\Pr(\mathcal{H}_i)$  satisfying  $|i - j| = 1$ , i.e.,  $\sum_{|i-j|=1} \Pr(\mathcal{H}_j|\mathcal{H}_i)\Pr(\mathcal{H}_i)$ . And the curve with  $\delta \geq 3$  means we sum up all the error probabilities for  $\delta$

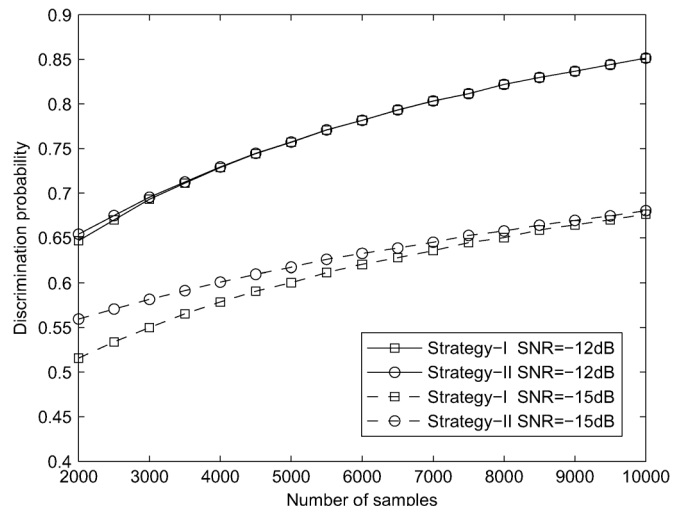


Fig. 8. The discrimination probability of local SU versus the number of samples.

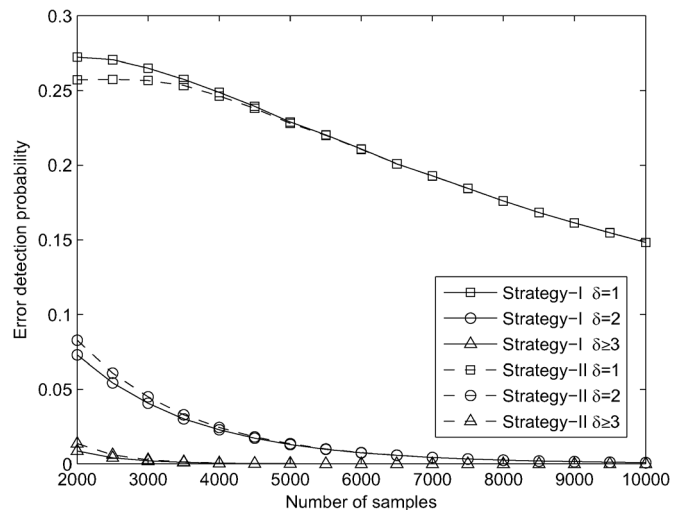


Fig. 9. Error detection probability versus the number of samples with average SNR = -12 dB.

$\geq 3$ . These curves indicate the trend of error distribution as the received sample number increases. Moreover, it is seen that the error probability decreases extensively when  $\delta$  increases. Especially when  $\delta \geq 3$ , the error probability is almost ignorable compared with that of  $\delta = 1$ . This is not unexpected because the chances for making a wrong decision to the farer power level should be smaller. A very important indication to practical design is that, SU may only pay attention to those errors with smaller  $\delta$  and set the corresponding cost values in Bayes Risk detection. Furthermore, sensing strategy-I performs worse than sensing strategy-II for  $\delta = 1$  while performs better than sensing strategy-II for  $\delta \geq 2$ .

Next we demonstrate the sensing performance versus average SNR in Fig. 10 for both the detection probability and the discrimination probability defined in (21). It is seen that sensing strategy-I outperforms sensing strategy-II in terms of both detection probability and discrimination probability. For sensing strategy-I, the difference between detection probability and discrimination probability is very large at low SNR. The reason is

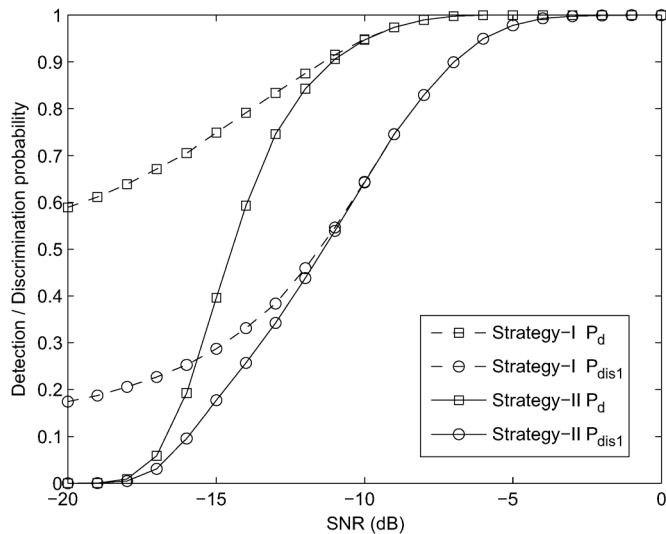


Fig. 10. The detection and discrimination probability of local SU versus SNR with  $M = 3000$ .

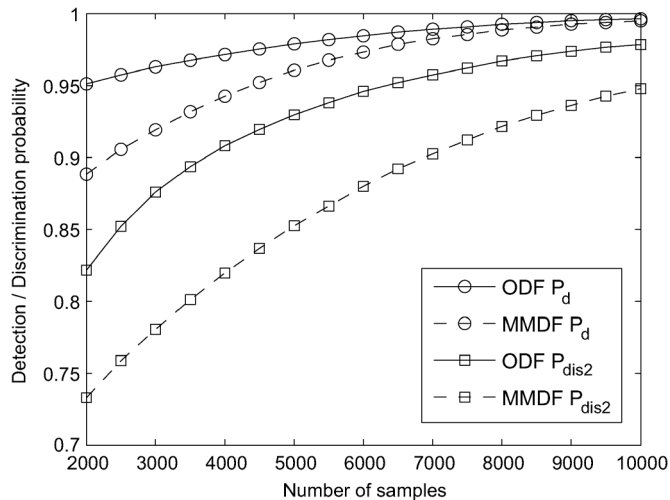


Fig. 11. Detection and discrimination probability versus number of samples for these two cooperative sensing methods with  $K = 5$ ,  $\text{SNR} = -12$  dB.

that even if PU is detected to be present, the sensing strategy-I makes many mistakes about PU's actual power level. However, for sensing strategy-II,  $\mathcal{H}_0$  may mask all the other states when SNR is low, which almost ruins the detecting ability. Nevertheless, since sensing strategy-II is originally defined for discriminating all  $\mathcal{H}_i, i \geq 0$ , once we include  $\mathcal{H}_0$ , the discrimination probability could outperform sensing strategy-I as has been demonstrated in Fig. 8.

### B. Cooperative Sensing

For all cooperative sensing examples, the local sensing results of all secondary users are made from sensing strategy-I.

Then, in Fig. 11, we show the detection probability as well as the discrimination probability versus the number of received samples when five SUs cooperate to make the final decision. The average SNR is taken as  $-12$  dB. Comparing with the sensing performance of single local SU in Fig. 7 and Fig. 8, it is clearly seen that the performance is greatly improved when cooperative

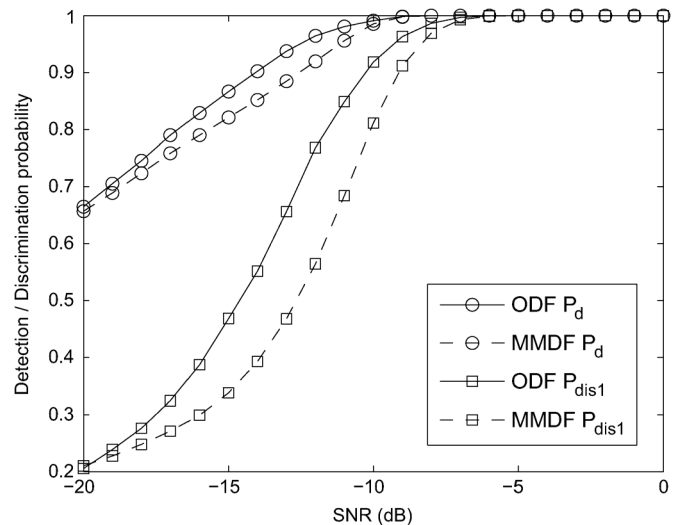


Fig. 12. Detection probability and discrimination probability in cooperative sensing versus SNR with  $K = 5$ ,  $M = 3000$ .

scheme is applied. Moreover, ODF outperforms MMDF at all sample numbers. Nevertheless, ODF needs to dynamically build the mapping function and is not as simple as MMDF.

In the last example, we show the performance of cooperative sensing versus average SNR in Fig. 12. Not surprisingly, ODF rule outperforms the MMDF in all SNR range. Besides, the discrimination probabilities for both sensing rules also get closer to the detection probabilities as the SNR grows. Moreover, the gaps between the ODF and MMDF diminish when SNR becomes larger.

## VI. CONCLUSIONS

In this paper, we thoroughly investigated a new CR scenario when PU works with different discrete power levels, which both matches the practical transmission and fits the theoretical demands of adapting the transmit power. We designed two spectrum sensing strategies which are shown to possess different but correlated optimization criteria. Most results, e.g., threshold expressions, decision probabilities, are derived in closed-forms. We present fruitful discussions over all aspects of the new spectrum sensing strategies, including the power-mask effect and its reasoning, the new definition of performance metrics, as well as the rationales behind. Moreover, we developed two cooperative sensing schemes which are shown to be very different from the traditional cooperative sensing algorithms. Various simulations are provided to corroborate the proposed studies. From the proposed studies, we believe that there exist many new problems in MPTP waiting for exploitation, while in the mean time many existing results from traditional CR deserve re-investigation.

## REFERENCES

- [1] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 201–220, Feb. 2005.
- [2] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 7–10, 2004, pp. 772–776.

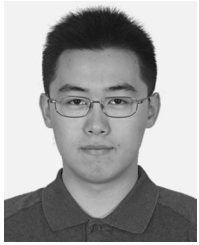
- [3] Y. Zeng, Y.-C. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Process. Lett.*, vol. 15, pp. 649–652, Oct. 2008.
- [4] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, 2009, First Quart.
- [5] H. Sun, W.-Y. Chiu, J. Jiang, A. Nallanathan, and H. V. Poor, "Wideband spectrum sensing with sub-Nyquist sampling in cognitive radios," *IEEE Signal Process. Lett.*, vol. 60, no. 11, pp. 6068–6073, Nov. 2012.
- [6] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1232–1241, Apr. 2011.
- [7] Y. Xin, H. Zhang, and L. Lai, "A low-complexity sequential spectrum sensing algorithm for cognitive radio," *J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 387–399, Mar. 2014.
- [8] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. IEEE Int. Conf. Wireless Netw., Commun. Mobile Comput.*, Maui, HI, USA, Jun. 2005, vol. 1, pp. 464–469.
- [9] G. Ding, J. Wang, Q. Wu, L. Zhang, Y. Zou, Y.-D. Yao, and Y. Chen, "Robust spectrum sensing with crowd sensors," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3129–3143, Sept. 2014.
- [10] P. Urriza, E. Rebeiz, and D. Cabric, "Multiple antenna cyclostationary spectrum sensing based on the cyclic correlation significance test," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 362–373, Oct. 2013, vol. 31, no. 11, pp. 2185–2195.
- [11] P. D. Sutton, K. E. Nolan, and L. E. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 13–24, Jan. 2008.
- [12] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.
- [13] W. Zhang, R. K. Mallik, and K. B. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5761–5766, Dec. 2009.
- [14] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Cooperative spectrum sensing in cognitive radio networks with weighted decision fusion schemes," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3838–3947, Dec. 2010.
- [15] S.-Q. Liu, B.-J. Hu, and X.-Y. Wang, "Hierarchical cooperative spectrum sensing based on double thresholds energy detection," *IEEE Commun. Lett.*, vol. 16, pp. 1096–1099, 2012.
- [16] W. Han, J. Li, Z. Li, and J. Si, "Efficient soft decision fusion rule in cooperative spectrum sensing," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1931–1943, Mar. 2013.
- [17] J. Font-Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2137–2146, Jul. 2010.
- [18] Z. Zhang, H. Jiang, P. Tan, and J. Slevinsky, "Channel exploration and exploitation with imperfect spectrum sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 429–441, Feb. 2013.
- [19] Y.-C. Liang, Y. Zeng, C. IY. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 1326–1337, Mar. 2008.
- [20] Y. J. Zhang and A. M.-C. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 362–373, Jan. 2011.
- [21] G. Wu, P. Ren, and Q. Du, "Recall-based dynamic spectrum auction with the protection of primary users," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 2070–2081, Oct. 2012.
- [22] Y. Zeng, Y.-C. Liang, and H. Pham, "Spectrum sensing for OFDM signals using pilot induced auto-correlations," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 353–363, Mar. 2013.
- [23] W. Yin, P. Ren, F. Li, and Q. Du, "Joint sensing and transmission for AF relay assisted PU transmission in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2249–2261, Nov. 2013.
- [24] Y. Zou, Y.-D. Yao, and B. Zheng, "Cooperative relay techniques for cognitive radio systems: Spectrum sensing and secondary user transmissions," *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 98–103, Apr. 2012.
- [25] P. Ren, Y. Wang, and Q. Du, "CAD-MAC: A channel-aggregation diversity based MAC protocol for spectrum and energy efficient cognitive ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 237–250, Feb. 2014.
- [26] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11, 1999.
- [27] European digital cellular telecommunication system (Phase 2) ETSI, 1995.
- [28] 3GPP TS 36.213, Evolved Universal Terrestrial Radio Access (E-UTRA), 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] D. Catrein, L. A. Imhof, and R. Mathar, "Power control, capacity, duality of uplink and downlink in cellular CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1777–1785, Oct. 2004.
- [31] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [32] Q. Cao, Y. Jing, and H. V. Zhao, "Power allocation in multi-user wireless relay networks through bargaining," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2870–2882, Jun. 2013.
- [33] In the Matter of Unlicensed Operation in the TV Broadcast Bands, ET Docket No. 04-186, Notice of Proposed Rulemaking, FCC OET May 2004.
- [34] X. Kang, Y. Liang, H. K. Garg, and L. Zhang, "Sensing-based spectrum sharing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4649–4654, Oct. 2009.
- [35] S. Chouh, H. Park, and T. Hwang, "Optimal beamforming and power allocation for sensing-based spectrum sharing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 412–417, Jan. 2014.
- [36] Z. Chen, X. Wang, and X. Zhang, "Continuous power allocation strategies for sensing-based multiband spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2409–2419, Nov. 2013.
- [37] R. Zhang, "On active learning and supervised transmission of spectrum sharing based cognitive radios by exploiting hidden primary radio feedback," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2960–2970, Oct. 2010.
- [38] K. M. Thilina *et al.*, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2209–2221, Nov. 2013.
- [39] C. Zhong, F. Gao, X.-D. Zhang, 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, June 2013.
- [40] 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.
- [41] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall PTR, 1993.



**Feifei Gao** (M'09–SM'14) received the B.Eng. degree from Xian Jiaotong University, Xi'an, China, in 2002, the M.Sc. degree from McMaster University, Hamilton, ON, Canada, in 2004, and the Ph.D. degree from National University of Singapore, Singapore, in 2007.

He was a Research Fellow with the Institute for Infocomm Research (I2R), A\*STAR, Singapore, in 2008 and an Assistant Professor with the School of Engineering and Science, Jacobs University, Bremen, Germany, from 2009 to 2010. In 2011, he joined the Department of Automation, Tsinghua University, Beijing, China, where he is currently an Associate Professor. His research areas include communication theory, signal processing for communications, array signal processing, and convex optimizations, with particular interests in MIMO techniques, multicarrier communications, cooperative communication, and cognitive radio networks. He has authored/coauthored more than 60 refereed IEEE journal papers and more than 80 IEEE conference proceeding papers, which have been cited more than 2500 times from Google Scholar.

Prof. Gao has served as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS LETTERS, *International Journal on Antennas and Propagations*, and *China Communications*. He has also served as the Symposium Co-Chair for 2015 IEEE Conference on Communications (ICC), 2014 IEEE Global Communications Conference (GLOBECOM), 2014 IEEE Vehicular Technology Conference Fall (VTC), as well as Technical Committee Members for many other IEEE conferences.



**Jiachen Li** received the B.E. degree from Tsinghua University, Beijing, China, in 2014.

He is currently pursuing the M.S. degree in computational data science at School of Computer Science from Carnegie-Mellon University, Pittsburgh, PA, USA. His current research interests include statistical analysis, machine learning, and data mining.



**Tao Jiang** (M'06–SM'10) received the B.S. and M.S. degrees in applied geophysics from China University of Geosciences, Wuhan, P. R. China, in 1997 and 2000, respectively, and the Ph.D. degree in information and communication engineering from Huazhong University of Science and Technology, Wuhan, in April 2004.

He is currently a Chair Professor in the School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan. From August 2004 to December 2007,

he was with universities, including Brunel University and the University of Michigan-Dearborn, respectively. He has authored or coauthored more than 200 technical papers in major journals and conferences and six books/chapters in the areas of communications and networks.

Dr. Jiang served or is serving as Symposium Technical Program Committee member of some major IEEE conferences, including INFOCOM, GLOBECOM, and ICC, etc. He was invited to serve as TPC Symposium Chair for the IEEE GLOBECOM 2013 and IEEE WCNC 2013. He has served or is serving as Associate Editor of some technical journals in communications, including the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE INTERNET OF THINGS JOURNAL. He is a recipient of the NSFC for Distinguished Young Scholars Award in P. R. China.



**Wen Chen** (M'03–SM'11) received the B.S. and M.S. degrees from Wuhan University, Wuhan, P.R. China, in 1990 and 1993, respectively, and the Ph.D. degree from The University of Electro-Communications, Tokyo, Japan, in 1999.

From 1991 to 2001, he was a Researcher with the Japan Society for the Promotion of Sciences. In 2001, he was with the University of Alberta, Edmonton, AB, Canada, first as a Postdoctoral Fellow with the Information Research Laboratory and then as a Research Associate with the Department of

Electrical and Computer Engineering. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, where he is also the Director of the Institute for Signal Processing and Systems. His research interests include network coding, cooperative communications, cognitive radio, and multiple-input-multiple-output orthogonal frequency-division multiplexing systems. In this area, he has published 60 IEEE journal papers and more than 100 IEEE conference papers.