Blind Spectrum Sensing by Information Theoretic Criteria for Cognitive Radios

Rui Wang and Meixia Tao, *Member, IEEE*

Abstract—Spectrum sensing is a fundamental and critical issue for opportunistic spectrum access in cognitive radio networks. Among the many spectrum-sensing methods, the information theoretic criteria (ITC)-based method is a promising blind method that can reliably detect the primary users while requiring little prior information. In this paper, we provide an intensive treatment on the ITC sensing method. To this end, we first introduce a new overdetermined channel model constructed by applying multiple antennas or oversampling at the secondary user to make ITC applicable. Then, a simplified ITC sensing algorithm is introduced, which needs to compute and compare only two decision values. Compared with the original ITC sensing algorithm, the simplified algorithm significantly reduces the computational complexity with no loss in performance. Applying the recent advances in random matrix theory, we then derive closed-form expressions to tightly approximate both the probability of false alarm and the probability of detection. Based on the insight derived from the analytical study, we further present a generalized ITC sensing algorithm that can provide a flexible tradeoff between the probability of detection and the probability of false alarm. Finally, comprehensive simulations are carried out to evaluate the performance of the proposed ITC sensing algorithms. Results show that they considerably outperform other blind spectrum-sensing methods in certain cases.

Index Terms—Cognitive radio (CR) networks, information theoretic criteria (ITC), random matrix theory, spectrum sensing.

I. INTRODUCTION

DUE TO THE increasing popularity of wireless devices in recent years, the radio spectrum has been an extremely scarce resource. By contrast, 90% of the existing licensed spectrum remains idle, and the usage geographically and temporally varies, as reported by the Federal Communication Commission (FCC) [1]. This indicates that the fixed frequency regulation policy drastically conflicts with the high demand for frequency resource. Cognitive radio (CR) is one of the most promising technologies to deal with such irrational frequency regulation policy [2], [3] and has received lots of attention. In CR networks, secondary (unlicensed) users first reliably sense the

primary (licensed) channel and then opportunistically access it without causing harmful interference to primary users [4]. By doing this, the spectrum utilization of existing wireless communication networks can be tremendously improved. The FCC has issued a Notice of Proposed Rule Making to allow the unlicensed CR devices to operate in the unused channel [5]. The IEEE has also formed the 802.22 working group to develop the standard for wireless regional area networks, which will operate on unused very-high-frequency/ultrahigh-frequency TV bands based on CR technology. Both of these activities will significantly change the current wireless communication situation.

As mentioned above, the secondary users need to opportunistically access the unused licensed channel while causing negligible interference to the primary users. As a result, the detection of presence of primary users is a fundamental and critical task in CR networks. Although the detection of presence of signals is known as a classical problem in signal processing, sensing the presence of primary users in a complicated communication environment, particularly a CR-based network, is still a challenging problem from the practice perspective. This is mainly due to the following two limiting factors: First, it is very difficult, if not possible, for the secondary user to obtain the necessary prior information about the signal characteristics of the primary user for most of the traditional detection techniques to apply. Second, the CR devices should be capable of sensing the very weak signals that are transmitted by the primary user. For instance, the standard released by the FCC has required that spectrum-sensing algorithms need to reliably detect the transmitted TV signals at a very low SNR of at least -18 dB [4].

Thus far, there are mainly four types of spectrum-sensing methods: 1) energy detection (ED) [6], [7], 2) matched filtering (coherent detection) [8], 3) feature detection [9], and 4) eigenvalue-based detection [10]–[12]. Among them, ED is optimal if the secondary user only knows the local noise power [13]. The matched-filtering-based coherent detection is optimal for maximizing the detection probability, but it requires the explicit knowledge of the transmitted signal pattern (e.g., pilot, training sequence, etc.) of the primary user. The feature detection, which is often referred to as the cyclostationary detection, exploits the periodicity in the modulation scheme, which, however, is difficult to determine in certain scenarios. By constructing the decision variables based on the eigenvalues of the sampled covariance matrix to detect the presence of the primary user, the eigenvalue-based sensing methods presented in [10]–[12] do not need to estimate the power of the noise and, hence, are more practical in most CR networks. Recently, several new spectrum-sensing schemes by incorporating system-level design parameters have been introduced, such as

Manuscript received January 13, 2010; revised June 6, 2010; accepted July 22, 2010. Date of publication August 12, 2010; date of current version October 20, 2010. This paper will be presented in part at the IEEE Global Telecommunications Conference, Miami, FL, December 2010. This work was supported in part by the Ministry of Science and Technology of China under Grant 2008ZX03003-004 and Grant 2008BAH30B09, by the Inter-Government Technology Collaborative Research Project of Shanghai under Grant 10220712000, and by the Shanghai Pujiang Talent Program under Grant 09PJ1406000. The review of this paper was coordinated by Prof. B. Hamdaoui.

The authors are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: liouxingrui@sjtu.edu.cn; mxtao@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/TVT.2010.2065250

throughput maximization [14]–[16] and cooperative sensing using multiple nodes [17]–[20]. Nevertheless, the aforementioned four types of sensing techniques are still treated as a basic component in these new schemes.

In this paper, we study a blind spectrum-sensing method based on information theoretic criteria (ITC), which is an approach originally for model selection introduced by Akaike [21], [22] and by Schwartz [23] and Rissanen [24]. Applying ITC for spectrum sensing was firstly introduced in [25]–[28]. This paper provides a more intensive study on the ITC sensing algorithm and its performance. The main contributions of this paper are concluded as follows.

- First of all, to make ITC applicable, a new overdetermined channel model is constructed by introducing multiple antennas or oversampling at the secondary user.
- Then, a simplified ITC (SITC) sensing algorithm, which only involves the computation of two decision values, is presented. Compared with the original ITC (OITC) sensing algorithm in [25], the SITC sensing algorithm is much less complex and yet almost has no performance loss. Simulation results also demonstrate that the proposed SITC-based spectrum sensing outperforms the eigenvalue-based sensing algorithm in [10] and almost obtains similar performance with [11]. The proposed sensing algorithm also enables a more tractable analytical study on the detection performance.
- Applying the recent advances in random matrix theory, we then derive closed-form expressions for both the probability of false alarm and the probability of detection, which can approximate the actual results in simulation very well.
- Finally, based on the insight derived from the analytical study, we further present a generalized ITC (GITC) sensing algorithm. By involving an adjustable threshold, the proposed GITC can provide flexible tradeoff between the probability of detection and the probability of false alarm to supply different system requirements.

The rest of this paper is organized as follows. In Section II, the preliminary on ITC is provided. The proposed overdetermined system model is presented in Section III. Section IV gives the proposed SITC sensing algorithm and the theoretical analysis of its detection performance, followed by the GITC sensing algorithm in Section V. Extensive simulation results are illustrated in Section VI. Finally, Section VII offers some concluding remarks.

Notations: $\mathcal{E}[\cdot]$ denotes expectation over the random variables within the brackets. $\text{Tr}(\mathbf{A})$ stands for the trace of matrix $\mathbf{A}$. Superscripts $(\cdot)^T$ and $(\cdot)^\dagger$ denote transpose and conjugate transpose, respectively.

II. PRELIMINARY ON INFORMATION THEORETIC CRITERIA

ITC is an approach that was originally implemented for model selection and was introduced by Akaike [21], [22] and by Schwartz [23] and Rissanen [24]. There are two well-known criteria that have been widely used: the Akaike information criterion (AIC) and the minimum description length (MDL) criterion. One of the most important applications of ITC is to

estimate the number of source signals in array signal processing [29]. Consider a system model described as

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \boldsymbol{\mu} \quad (1)$$

where $\mathbf{x}$ is the $p \times 1$ complex observation vector, $\mathbf{A}$ is a $p \times q$ ($p > q$) complex system matrix, $\mathbf{s}$ denotes the $q \times 1$ complex source modulated signals, and $\boldsymbol{\mu}$ is the additive complex white Gaussian noise vector. It is noted that the definite parameters q , $\mathbf{A}$, and σ^2 are all unknown. The resulting cost functions of the AIC and the MDL have the following form [29]:

$$\begin{aligned} \text{AIC}(k) = & -2 \log \left(\frac{\prod_{i=k+1}^p l_i^{1/(p-k)}}{\frac{1}{p-k} \sum_{i=k+1}^p l_i} \right)^{N(p-k)} \\ & + 2k(2p - k) + 2 \end{aligned} \quad (2)$$

$$\begin{aligned} \text{MDL}(k) = & -\log \left(\frac{\prod_{i=k+1}^p l_i^{1/(p-k)}}{\frac{1}{p-k} \sum_{i=k+1}^p l_i} \right)^{N(p-k)} \\ & + \left(\frac{1}{2}k(2p - k) + \frac{1}{2} \right) \log N \end{aligned} \quad (3)$$

where N signifies the observation times, and l_i denotes the i th decreasing ordered eigenvalue of the sampled covariance matrix. The estimated number of source signals is determined by choosing the minimum (2) or (3). That is

$$\hat{k}_{\text{AIC}} = \arg \min_{j=0,1,\dots,p-1} \text{AIC}(j) \quad (4)$$

$$\hat{k}_{\text{MDL}} = \arg \min_{j=0,1,\dots,p-1} \text{MDL}(j). \quad (5)$$

III. SYSTEM MODEL

We consider a multipath fading channel model and assuming that there is only one primary user in the cognitive radio network. Let $x(t)$ be a continuous-time baseband received signal at the secondary user's receiver. Spectrum sensing can be formulated as a binary hypothesis test between the following two hypotheses:

$$\mathcal{H}_0 : \quad x(t) = \mu(t) \quad (6)$$

$$\mathcal{H}_1 : \quad x(t) = \int_0^T h(\ell)s(t-\ell)d\ell + \mu(t) \quad (7)$$

where $s(t)$ denotes the signal that is transmitted by the primary user, $h(t)$ is the continuous channel response between the primary transmitter and the secondary receiver, $\mu(t)$ denotes the additive white noise, and the parameter T signifies the duration of the channel. The channel response is also assumed to remain invariant during each observation. To obtain the discrete representation, we assume that the received signal is sampled at rate f_s , which is equal to the reciprocal of the baseband symbol duration T_0 . For notation simplicity, we define $x(n) = x(nT_0)$, $s(n) = s(nT_0)$, and $\mu(n) = \mu(nT_0)$.

Hence, the corresponding received signal samples under the two hypotheses are described as

$$\mathcal{H}_0: x(n) = \mu(n) \quad (8)$$

$$\mathcal{H}_1: x(n) = \sum_{i=0}^{L-1} h(i)s(n-i) + \mu(n) \quad (9)$$

where $h(i)$ ($0 \leq i \leq L-1$) denotes the discrete channel response of $h(t)$, and L denotes the order of the discrete channel (L taps). Let each observation consist of M received signal samples. Then, (8) and (9) can be rewritten in a matrix form as

$$\mathcal{H}_0: \mathbf{x}_i = \boldsymbol{\mu}_i \quad (10)$$

$$\mathcal{H}_1: \mathbf{x}_i = \mathbf{H}\mathbf{s}_i + \boldsymbol{\mu}_i \quad (11)$$

where $\mathbf{H}$ is an $M \times (L+M-1)$ circular channel matrix, which is defined as

$$\mathbf{H} = \begin{bmatrix} h(L-1) & h(L-2) & \cdots & h(0) \\ & h(L-1) & h(L-2) & \cdots & h(0) \\ & & \ddots & \ddots & \\ & & & h(L-1) & h(L-2) & \cdots & h(0) \end{bmatrix}$$

$\mathbf{x}_i$, $\mathbf{s}_i$, and $\boldsymbol{\mu}_i$ are the $M \times 1$ observation vector, the $(L+M-1) \times 1$ source signal vector, and the $M \times 1$ noise vector, respectively, and are defined as

$$\mathbf{x}_i = [x(iM-M+1), x(iM-M+2), \dots, x(iM)]^T \quad (12)$$

$$\mathbf{s}_i = [s(iM-M-L+2), s(iM-M-L+3), \dots, s(iM)]^T \quad (13)$$

$$\boldsymbol{\mu}_i = [\mu(iM-M+1), \mu(iM-M+2), \dots, \mu(iM)]^T \quad (14)$$

Now, comparing (11) with the array signal processing model (1), we find that a major difference is that the $\mathbf{H}$ in our considered system model is an underdetermined matrix, i.e., the order of the column is larger than the order of the row. Therefore, ITC are not directly applicable here [29].

To construct an overdetermined channel matrix $\mathbf{H}$ as in (1), one needs to enlarge the observation space. Obviously, simply increasing the observation window M does not work. Here, we propose expanding the observation space using one of the following two methods. One method increases the spatial dimensionality by employing multiple receive antennas at the secondary user, and the other is increases the time dimensionality by oversampling the received signals. It turns out that the two methods are similar to each other. Hence, we shall focus on the multiple-antenna approach hereafter. The difference for the oversampling method will be discussed at the end of this section. Specifically, suppose that the detector at the secondary user is equipped with K antennas. Redefine (12) and (14) as

$$\mathbf{x}_i = [x_1^i(1), x_2^i(1), \dots, x_K^i(1), x_1^i(2), \dots, x_K^i(2), \dots, x_1^i(M), \dots, x_K^i(M)]^T \quad (15)$$

$$\boldsymbol{\mu}_i = [\mu_1^i(1), \mu_2^i(1), \dots, \mu_K^i(1), \mu_1^i(2), \dots, \mu_K^i(2), \dots, \mu_1^i(M), \dots, \mu_K^i(M)]^T \quad (16)$$

where $\mathbf{x}_k^i = [x_k^i(1), x_k^i(2), \dots, x_k^i(M)]^T$ represents the $M \times 1$ observation vector at the k th antenna at the i th observation as

in (12), and $\boldsymbol{\mu}_k^i = [\mu_k^i(1), \mu_k^i(2), \dots, \mu_k^i(M)]^T$ is the corresponding noise vector at the k th antenna at the i th observation as in (14). Then, the new channel matrix $\mathbf{H}$ becomes an $MK \times (M+L-1)$ matrix

$$\mathbf{H} = \begin{bmatrix} h_1(L-1) & h_1(L-2) & \cdots & h_1(0) \\ \vdots & \vdots & \vdots & \vdots \\ h_K(L-1) & h_K(L-2) & \cdots & h_K(0) \\ & h_1(L-1) & h_1(L-2) & \cdots & h_1(0) \\ & \vdots & \vdots & \vdots & \vdots \\ & h_K(L-1) & h_K(L-2) & \cdots & h_K(0) \\ & & \ddots & \ddots & \end{bmatrix} \quad (17)$$

Here, $h_k(i)$, for $i = 0, \dots, L-1$, denotes the i th channel tap observed at the k th antenna. To ensure that $\mathbf{H}$ is now an overdetermined matrix (the order of the row is larger than the order of the column), we need to have

$$K > \frac{L+M-1}{M}$$

or alternatively

$$M > \frac{L-1}{K-1}.$$

Furthermore, we assume that the noise samples coming from different antennas are independent with zero mean and $\mathcal{E}(\boldsymbol{\mu}_i \boldsymbol{\mu}_i^H) = \sigma^2 \mathbf{I}_{MK}$. Then, we can exactly ensure that the system model under multiple antennas satisfies the overdetermined condition specified in [29]. For ease of presentation, we define $p = MK$ and $q = L+M-1$ in (11) for the rest of this paper.

As mentioned earlier, the second approach to construct the overdetermined channel model is for the secondary user to oversample the received signals. Suppose that the oversampling factor is given by K . That is, the received baseband signal is sampled K times in one symbol. Then, a similar system model as in (15)–(17) can be obtained, except that $\mathbf{x}_i$ and $\boldsymbol{\mu}_i$ should be replaced with

$$\mathbf{x}_i = [x(iMK-MK+1), x(iMK-MK+2), \dots, x(iMK)]^T \quad (18)$$

$$\boldsymbol{\mu}_i = [\mu(iMK-MK+1), \mu(iMK-MK+2), \dots, \mu(iMK)]^T \quad (19)$$

and $h_k(i)$, for $i = 0, \dots, L-1$, becomes the k th oversampling point of the i th channel tap. It can be verified that $h_k(i)$'s are different for different k [30]. The major difference between the oversampling approach and the multiple-antenna approach is that the oversampled noise samples in (19) are correlated, which contradicts the primary assumption of independent noise samples. Nevertheless, the prewhiting technique can be used to whiten the correlated noises based on the known correlation matrix. The details are shown in Appendix A.

Before leaving this section, it is noted that, although the proposed overdetermined model is based on the assumption that there is only one primary user in the cognitive network, it is also

applicable to the scenario where there exist multiple primary users. An alternative approach to construct the overdetermined model in the presence of multiple primary users is to use the cooperative sensing technique as in [28] by using multiple detectors.

IV. SIMPLIFIED INFORMATION THEORETIC CRITERIA SENSING ALGORITHM AND PERFORMANCE ANALYSIS

Since the binary hypothesis test in the spectrum sensing is equivalent to the special case of the source number estimation problem, the ITC method can be directly applied to conduct spectrum sensing, as first proposed in [25]–[28]. The basic idea is when the primary user is absent, the received signal $\mathbf{x}_i$ is only the white noise sample. Therefore, the estimated number of source signals via ITC (AIC or MDL) should be zero. Otherwise, when the primary user is present, the number of source signals must be larger than zero. Hence, by comparing the estimated number of source signals with zero, the presence of the primary user can be detected. It is noted that the estimation of the number of source signals by using (4) and (5) needs very little prior information about the primary user. In particular, it does not require the knowledge of channel state information, synchronization, pilot design, or modulation strategy. Moreover, it does not need the estimation of noise power. Hence, we argue that the ITC method is a blind spectrum sensing similar to [10]–[12], and it is robust and suitable for practical applications.

However, it is known that signal detection is much easier than signal estimation. Therefore, using the estimation method to conduct the detection as in [25]–[28] may lead to unnecessary computational complexity overhead. Meanwhile, it makes it difficult to carry out analytical study on the detection performance. Here, we propose an SITC algorithm to conduct the spectrum sensing. It can significantly reduce the computational complexity while having almost no performance loss, as will be illustrated in Section VI. It also enables a more tractable analytical study on the detection performance.

A. SITC Sensing Algorithm

Before presenting the SITC sensing algorithm in detail, we have the following lemma.

Lemma 1: If there is one value $\hat{k}(> 0)$ that minimizes the AIC metric in (2) [MDL metric in (3)], then $\text{AIC}(0) > \text{AIC}(1)$ ($\text{MDL}(0) > \text{MDL}(1)$) with high probability.

Proof: See Appendix B. ■

The outline of the proposed simplified sensing algorithm is as follows.

Algorithm 1: SITC sensing algorithm

Step 1. Compute the sampled covariance matrix of received signals, i.e., $\mathbf{R}_x = (1/N) \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\dagger$, where $\mathbf{x}_i$'s are received vectors as described in (15) or (18), and N denotes the number of observations.

Step 2. Obtain the eigenvalues of $\mathbf{R}_x$ through an eigenvalue decomposition technique and denote them as $\{l_1, l_2, \dots, l_p\}$ with $l_1 \geq l_2, \dots, \geq l_p$.

Step 3. Calculate the decision values $\text{AIC}(0)$ and $\text{AIC}(1)$ [$\text{MDL}(0)$ and $\text{MDL}(1)$] according to (2) [see (3)]. Then, the detection decision metric is

$$\mathcal{T}_{\text{SITC-AIC}}(\mathbf{L}_x) : \text{AIC}(0) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \text{AIC}(1)$$

if AIC is adopted or

$$\mathcal{T}_{\text{SITC-MDL}}(\mathbf{L}_x) : \text{MDL}(0) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \text{MDL}(1)$$

if MDL is adopted, where $\mathbf{L}_x$ denotes the set of eigenvalues $\{l_i, i = 1, 2, \dots, p\}$.

Note that in the OITC sensing algorithm [25], one needs to find the exact value of $\hat{k}$ from 0 to $p - 1$ to minimize the AIC in (2) or the MDL in (3). In the proposed SITC algorithm, only two decision values at $k = 0$ and 1 should be computed and compared. Thus, the computational complexity is significantly reduced. In Section IV-B, based on the proposed SITC algorithm, we present the analytical results on the detection performance. Since, from Lemma 1, the SITC algorithm almost obtains the same performance as the OITC algorithm, we claim that our analytical results are also applicable for evaluating the performance of the OITC algorithm.

B. Performance Analysis

Since spectrum sensing is actually a binary hypothesis test, the performance we focus on is the probability of detection P_d (the probability for identifying the signal when the primary user is present) and the probability of false alarm P_f (the probability for identifying the signal when the primary user is absent). As no threshold value is involved in the ITC sensing algorithm, P_d is not directly related to P_f . The two probabilities are presented separately. For ease of presentation, we shall take the AIC criterion for example to illustrate the analysis throughout this section. The extension to the MDL criterion is straightforward if not mentioned otherwise.

1) *Probability of False Alarm:* According to the sensing steps in Algorithm 1, the false alarm occurs when $\text{AIC}(0)$ is larger than $\text{AIC}(1)$ at hypothesis $\mathcal{H}_0$. The probability of false alarm can be expressed as

$$P_{f\text{-AIC}} = \Pr(\text{AIC}(0) > \text{AIC}(1) | \mathcal{H}_0). \quad (20)$$

Since the primary user is absent, the received signal $\mathbf{x}_i$ only contains the noises. The sampled covariance matrix $\mathbf{R}_x$ in Algorithm 1 thus turns to $\mathbf{R}_\mu$, which is defined as

$$\mathbf{R}_\mu = \frac{1}{N} \sum_{i=1}^N \boldsymbol{\mu}_i \boldsymbol{\mu}_i^\dagger. \quad (21)$$

Hence, the eigenvalues in (2) become the eigenvalues of the sampled noise covariance matrix $\mathbf{R}_\mu$ in (21), which is a Wishart random matrix [31]. By applying the recent advances on the eigenvalue distribution for Wishart matrices, a closed-form expression for the probability of false alarm can be obtained.

Proposition 1: The probability of false alarm of the proposed spectrum-sensing algorithm can be approximated as

$$P_f \approx F_2 \left(\frac{pN - (\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p}) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) - F_2 \left(\frac{(p - \alpha_1)N - (\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p}) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) + F_2 \left(\frac{(p - \alpha_2)N - (\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p}) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) - F_2 \left(\frac{-(\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p}) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) \quad (22)$$

where $F_2(\cdot)$ is the cumulative distribution function (cdf) of the Tracy–Widom distribution of order 2 [31], and α_1 and α_2 with $\alpha_1 < \alpha_2$ are the two real positive roots of the function in (28) if the AIC is applied or (31) if the MDL is applied.

Proof: Recall the definition in (20), to compute the probability of false alarm is to compute the probability

$$P_{f\text{-AIC}} = \Pr(\text{AIC}(0) - \text{AIC}(1) > 0 | \mathcal{H}_0). \quad (23)$$

According to the cost function of the AIC defined in (2), we have

$$\begin{aligned} \text{AIC}(0) - \text{AIC}(1) &= -2 \log \left[\frac{\prod_{i=1}^p l_i^{1/p}}{\frac{1}{p} \sum_{i=1}^p l_i} \right]^{pN} \\ &\quad + 2 \log \left[\frac{\prod_{i=2}^p l_i^{1/(p-1)}}{\frac{1}{p-1} \sum_{i=2}^p l_i} \right]^{(p-1)N} - (4p - 2). \end{aligned}$$

Then, we can rewrite (23) as

$$P_{f\text{-AIC}} = \Pr \left(\log \left[\frac{\left(\frac{1}{p} \sum_{i=1}^p l_i \right)^p}{\left(\frac{1}{p-1} \sum_{i=2}^p l_i \right)^{p-1} l_1} \right] > \frac{4p-2}{2N} \middle| \mathcal{H}_0 \right). \quad (24)$$

Note here that the sum of the eigenvalues of the sampled covariance matrix, i.e., $(1/p) \sum_{i=1}^p l_i$, is equivalent to $(1/pN) \text{Tr}(\sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\dagger)$. At hypothesis $\mathcal{H}_0$, where the received vector involves only the noise samples, $(1/pN) \text{Tr}(\sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\dagger)$ is the unbiased estimation of the covariance of the white noise. Therefore, when N is sufficiently large, we have

$$\frac{1}{p} \sum_{i=1}^p l_i \approx \sigma^2. \quad (25)$$

Substituting (25) into (24) yields

$$P_{f\text{-AIC}} \approx \Pr \left[\frac{(\sigma^2)^p}{\left(\frac{p}{p-1} \sigma^2 - \frac{l_1}{p-1} \right)^{p-1} l_1} > \exp \left(\frac{2p-1}{N} \right) \middle| \mathcal{H}_0 \right]. \quad (26)$$

From (26), it is seen that the probability of false alarm is only dependent on the largest eigenvalue of the noise sampled covariance matrix $\mathbf{R}_\mu$. Since $\mathbf{R}_\mu$ is actually a Wishart random matrix, its largest eigenvalue l_1 satisfies the Tracy–Widom distribution of order 2 [31]. To apply this result, we rewrite (26) as

$$\begin{aligned} P_{f\text{-AIC}} &\approx \Pr \left[\frac{l_1}{\sigma^2} \left(p - \frac{l_1}{\sigma^2} \right)^{p-1} < \frac{(p-1)^{p-1}}{\exp \left(\frac{2p-1}{N} \right)} \middle| \mathcal{H}_0 \right] \\ &= \Pr \left[x^p - px^{p-1} + \frac{(p-1)^{p-1}}{\exp \left(\frac{2p-1}{N} \right)} > 0 \middle| \mathcal{H}_0 \right] \end{aligned} \quad (27)$$

where $x \triangleq p - (l_1/\sigma^2)$.

Define a function

$$f(x) \triangleq x^p - px^{p-1} + \frac{(p-1)^{p-1}}{\exp \left(\frac{2p-1}{N} \right)}. \quad (28)$$

We next find the real roots of this function.

Taking the differentiation of $f(x)$ and equating it to zero, we obtain

$$\frac{df(x)}{dx} = px^{p-1} - p(p-1)x^{p-2} = px^{p-2} [x - (p-1)] = 0.$$

Clearly, $f(x)$ has two stationary points, which are $x = p-1$ and $x = 0$. In the following, two scenarios with p being even or odd are considered, respectively. When p is even, it is easily found that the function $f(x)$ monotonically decreases over $(-\infty, p-1)$ and monotonically increases over $(p-1, \infty)$. Simultaneously, we can verify that

$$f(p-1) = (p-1)^p - p(p-1)^{p-1} + \frac{(p-1)^{p-1}}{\exp \left(\frac{2p-1}{N} \right)} < 0$$

$$f(0) = f(p) = \frac{(p-1)^{p-1}}{\exp \left(\frac{2p-1}{N} \right)} > 0.$$

Therefore, there must be two real roots within $(0, p)$ and around $p-1$ for function $f(x)$. Let α_1 and α_2 , with $\alpha_1 < \alpha_2$, denote the two real roots; then, (27) is converted into an equivalent form

$$P_{f\text{-AIC}} \approx \Pr[x < \alpha_1 | \mathcal{H}_0] + \Pr[\alpha_2 < x | \mathcal{H}_0]. \quad (29)$$

When p is odd, we can find that $f(x)$ monotonically decreases over $(0, p-1)$, whereas it is the monotonic increasing function over both $(-\infty, 0)$ and $(p-1, \infty)$. According to the fact that $f(-\infty) < 0$, $f(0) > 0$, $f(p-1) < 0$, and $f(p) > 0$, we conclude that $f(x)$ has three real roots, which are denoted as α_0 , α_1 , and α_2 , with $\alpha_0 < 0$ and $0 < \alpha_1 < \alpha_2$, respectively. Then, (27) can be rewritten as

$$P_{f\text{-AIC}} \approx \Pr[\alpha_0 < x < \alpha_1 | \mathcal{H}_0] + \Pr[\alpha_2 < x | \mathcal{H}_0]. \quad (30)$$

However, it is noted that as N is large enough, the largest eigenvalue of the sampled noise covariance matrix, i.e., l_1 , is just slightly larger than the true covariance of noise σ^2 . Hence, from the definition, x can be reasonably limited in $(0, p)$. Therefore, both the probabilities of (29) and (30) can be summarized as the following form:

$$P_{f\text{-AIC}} \approx \Pr[0 < x < \alpha_1 | \mathcal{H}_0] + \Pr[\alpha_2 < x < p | \mathcal{H}_0].$$

In other words

$$P_{f\text{-AIC}} \approx \Pr \left[p - \alpha_1 < \frac{l_1}{\sigma^2} < \left| \mathcal{H}_0 \right| \right] + \Pr \left[0 < \frac{l_1}{\sigma^2} < p - \alpha_2 \mid \mathcal{H}_0 \right].$$

Applying the distribution of the largest eigenvalue for the Wishart matrix in random matrix theory [31], the variable $N(l_1/\sigma^2)$ satisfies the distribution of Tracy–Widom of order 2, i.e.,

$$\frac{N \frac{l_1}{\sigma^2} - \left(\sqrt{N} + \sqrt{p} \right)^2}{\left(\sqrt{N} + \sqrt{p} \right) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \sim W_2 \sim F_2.$$

Here, W_2 and F_2 denote the probability density function and the cdf for the distribution of Tracy–Widom of order 2, respectively. Therefore, the probability of false alarm of the AIC can be concluded as (22).

Similar to the above derivation, when the MDL criterion is applied, we only need to modify the step in (27) as

$$P_{f\text{-MDL}} \approx \Pr \left[x^p - px^{p-1} + \frac{(p-1)^{p-1}}{\exp \left(\frac{(p-0.5) \log N}{N} \right)} > 0 \mid \mathcal{H}_0 \right]$$

and redefine the function $f(x)$ in (28) as

$$f(x) \triangleq x^p - px^{p-1} + \frac{(p-1)^{p-1}}{\exp \left(\frac{(p-0.5) \log N}{N} \right)}. \quad (31)$$

From Proposition 1, it is found that the probability of false alarm is independent with noise covariance σ^2 . Therefore, the proposed SITC sensing algorithm is robust in practical applications. It is also noted that P_f depends on the product of M and K , i.e., $p = MK$, rather than the individual values of M and K .

2) *Probability of Detection*: When the primary user is present, the event of detection also occurs when $\text{AIC}(0) > \text{AIC}(1)$. The probability of detection is, thus, described as

$$P_{d\text{-AIC}} = \Pr (\text{AIC}(0) > \text{AIC}(1) \mid \mathcal{H}_1). \quad (32)$$

Since, at $\mathcal{H}_1$, the received vector $\mathbf{x}_i$ involves the signals that are transmitted by the primary user, the sampled covariance matrix $\mathbf{R}_x$ can be written as

$$\mathbf{R}_x = \frac{1}{N} \sum_{i=1}^N (\mathbf{H} \mathbf{s}_i + \boldsymbol{\mu}_i)(\mathbf{H} \mathbf{s}_i + \boldsymbol{\mu}_i)^\dagger. \quad (33)$$

Note that $\mathbf{R}_x$ is no longer a Wishart random matrix. The exact distribution of its eigenvalues is unknown and difficult to find, and hence, so is P_d . In the following, we resort to deriving a closed-form expression for the conditional probability of detection given the channel matrix $\mathbf{H}$. The average probability of detection can then be obtained using a hybrid analytical-simulation approach.

Proposition 2: Let $\mathbf{R}_s$ denote the covariance matrix of $\mathbf{s}_i$ given in (13) and $\{\delta_1, \delta_2, \dots, \delta_p\}$ be the eigenvalues of matrix $\mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger$ (with $\delta_1 \geq \delta_2 \geq \dots \geq \delta_p$). Then, there exists a value

of ρ , for $\delta_p \leq \rho \leq \delta_1$, such that the probability of detection given $\mathbf{H}$ can be approximated as $P_{d|H} \approx Q(\rho)$, where the function $Q(\cdot)$ is

$$\begin{aligned} Q(\delta) = & F_2 \left(\frac{pN - \left(\sqrt{N} + \sqrt{p} \right)^2}{\left(\sqrt{N} + \sqrt{p} \right) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) \\ & - F_2 \left(\frac{\left(\frac{(p-\pi_1)\epsilon-\delta}{\sigma^2} \right) N - \left(\sqrt{N} + \sqrt{p} \right)^2}{\left(\sqrt{N} + \sqrt{p} \right) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) \\ & + F_2 \left(\frac{\left(\frac{(p-\pi_2)\epsilon-\delta}{\sigma^2} \right) N - \left(\sqrt{N} + \sqrt{p} \right)^2}{\left(\sqrt{N} + \sqrt{p} \right) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) \\ & - F_2 \left(\frac{- \left(\sqrt{N} + \sqrt{p} \right)^2}{\left(\sqrt{N} + \sqrt{p} \right) \left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}} \right)^{\frac{1}{3}}} \right) \end{aligned} \quad (34)$$

where $\epsilon = (1/p) \text{Tr}(\mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger) + \sigma^2$, and π_1 and π_2 (with $\pi_1 < \pi_2$) denote the two positive roots of the function (38) for the AIC or (40) for the MDL. Furthermore, upper and lower bounds can be obtained as $Q(\delta_p) \leq P_{d|H} \leq Q(\delta_1)$.

Proof: See Appendix C. ■

From Proposition 2, we find that P_d is not only related to N and p but also depends on (ϵ/σ^2) , which is the ratio of the signal strength of the primary user to the noise variance. The exact value of $\rho \in [\delta_p, \delta_1]$ in Proposition 2 is difficult to determine in an analytical way, since it is related to both the channel response $\mathbf{H}$ and the covariance matrix of source signal $\mathbf{R}_s$. In practice, we can simply choose $\rho_{\text{AIC}} = (1/2)(\delta_p + \delta_1)$ and $\rho_{\text{MDL}} = (3/4)(\delta_p + \delta_1)$. It will be demonstrated later in Section VI that the analytical $P_{d|H}$ based on this choice of ρ can approximate the Monte Carlo results very well in most cases.

V. GENERALIZED INFORMATION THEORETIC CRITERIA SENSING ALGORITHM

As mentioned in Section IV, the probability of detection and the probability of false alarm of the proposed SITC sensing algorithm are not directly related to each other, as the algorithm involves no threshold (same for the OITC algorithm in [25]). According to the analytical results given in (22) and (34), to satisfy different system requirements, a proper set of values for the parameters M , K , and N in model (11) should be chosen, which is inconvenient for practical application. Here, based on the analytical discussion in Section IV, we propose a GITC sensing algorithm, which can provide a flexible trade-off between P_d and P_f according to different system design requirements.

From the expression given in (24) and (37), we found that the sensing decision for the SITC algorithm is actually based on the decision variable

$$\log \left[\frac{\left(\frac{1}{p} \sum_{i=1}^p l_i \right)^p}{\left(\frac{1}{p-1} \sum_{i=2}^p l_i \right)^{p-1} l_1} \right].$$

TABLE I
PROBABILITY OF FALSE ALARM WITH DIFFERENT $p = M \times K$ AT $N = 10000$

	$6 = 2 \times 3$	$12 = 3 \times 4$	$20 = 5 \times 4$	$24 = 4 \times 6$	$35 = 5 \times 7$
Simulation results for SITC-AIC	0.0948	0.0770	0.0594	0.0541	0.0460
Simulation results for OITC-AIC	0.0972	0.0773	0.0597	0.0541	0.0470
Analytical results for SITC-AIC	0.1360	0.1036	0.0791	0.0711	0.0550
Simulation results for SITC-MDL	0	0	0	0	0
Simulation results for OITC-MDL	0	0	0	0	0
Analytical results for SITC-MDL	0	0	0	0	0

Thus, we generalize the decision rule as

$$\mathcal{T}_{\text{GITC}}(\mathbf{L}_x) : \frac{\left(\frac{1}{p} \sum_{i=1}^p l_i\right)^p}{\left(\frac{1}{p-1} \sum_{i=2}^p l_i\right)^{p-1} l_1} \stackrel{\mathcal{H}_1}{\geq} \gamma \quad (35)$$

where γ is a preset threshold. It is seen that if we set $\gamma = \exp(2p - 1/N)$, the decision rule given in (35) turns into the AIC-based SITC sensing algorithm presented in Algorithm 1. If we fix $\gamma = \exp((p - 0.5) \log N/N)$, the algorithm becomes the MDL-based SITC sensing algorithm. Furthermore, it is easy to find that the analytical results obtained in Section IV are applicable for the GITC sensing algorithm. The only change that needs to be made is to replace α_1 and α_2 in (22) [or π_1 and π_2 in (34)] by two real positive roots that are generated by the following:

$$f(x) \triangleq x^p - px^{p-1} + \frac{(p-1)^{p-1}}{\gamma}. \quad (36)$$

Thus, the outline of the proposed GITC sensing algorithm can be summarized as follows.

Algorithm 2: GITC sensing algorithm

Steps 1 and 2. The same as Algorithm 1 in Section IV.

Step 3. According to the system requirement on P_f , choose a proper threshold γ based on (22) and (36).

Step 4. Conduct the decision based on (35).

According to the decision variable presented in (35), we find that the proposed GITC sensing algorithm is actually also an eigenvalue-based method similar to [10]–[12]. The advantage of the proposed GITC over the algorithms in [11] and [12] is that it enables us to analytically obtain the explicit decision threshold γ according to the system requirement on P_f before the actual deployment.

VI. SIMULATION RESULTS AND DISCUSSIONS

Here, we present some numerical examples to demonstrate the effectiveness of the proposed sensing schemes and to confirm the theoretical analysis.

A. Comparison Between Simulation and Analytical Results for Both SITC and OITC

In our first set of examples, we compare the simulation results with analytical results given in Propositions 1 and 2. The comparison between SITC and OITC is also presented.

TABLE II
PROBABILITY OF FALSE ALARM WITH DIFFERENT N AT $p = MK = 20$

	$N = 1000$	$N = 5000$	$N = 10000$
Simulation results for SITC-AIC	0.0421	0.0558	0.0594
Simulation results for OITC-AIC	0.0421	0.0561	0.0597
Analytical results for SITC-AIC	0.0581	0.0744	0.0791
Simulation results for SITC-MDL	0	0	0
Simulation results for OITC-MDL	0	0	0
Analytical results for SITC-MDL	0	0	0

In the simulation, the channel taps are random numbers with a zero-mean complex Gaussian distribution. All the results are averaged over 1000 Monte Carlo realizations. For each realization, random channel, random noise, and random binary phase-shift keying modulated inputs are generated. We define the SNR as the ratio of the average received signal power to the average noise power, i.e.,

$$\text{SNR} = \frac{\mathcal{E} [\|\mathbf{x}_i - \boldsymbol{\mu}_i\|^2]}{\mathcal{E} [\|\boldsymbol{\mu}_i\|^2]}.$$

The comparison of simulation and analytical results for P_f is demonstrated in Tables I and II. According to Proposition 1, P_f is independent with the noise variance and, thus, remains constant over different SNRs. Hence, we average multiple values over different SNRs as the simulated P_f and compare it with the analytical P_f . From Table I, we first observe that SITC and OITC perform almost the same. It is also seen that, for the AIC, the analytical results are slightly larger than the simulation results, particularly when $p = MK$ is small. Nevertheless, the analytical approximation is accurate enough to evaluate the performance of the proposed sensing scheme. It is also found that $P_{f\text{-AIC}}$ gradually decreases as $p = MK$ increases, whereas $P_{f\text{-MDL}}$ remains zero in both simulation and analytical results. We conclude that the MDL method has excellent false-alarm performance. From Table II, we find that the probability of false alarm increases very slowly as N increases. In fact, our simulation shows that $P_{f\text{-AIC}}$ is still below 0.1, even when $N = 10^{15}$ at $M = 5$ and $K = 4$.

Figs. 1–4 show P_d at different system parameters. In Fig. 1, we first compare the detection performance obtained by simulation between SITC and OITC. It is seen that the proposed SITC sensing algorithm does not lead to any performance loss compared with the OITC algorithm. Then, comparing the semianalytical results obtained from Proposition 2 with the simulation results, one can observe a very good match between them, particularly for the MDL method. Thus, Proposition 2

Fig. 1. Simulation and analytical results about the probability of detection at different (M, K, N) for both SITC and OITC.

Fig. 2. Probability of detection for different K at $M = 5$ and $N = 10000$.

Fig. 3. Probability of detection for different M at $K = 4$ and $N = 10000$.

Fig. 4. Probability of detection for different N at $M = 5$ and $K = 4$.

Fig. 5. Comparison with the eigenvalue-based methods and the energy-detection method at $M = 5$, $K = 4$, and $N = 10000$.

is validated. Figs. 2–4 present the simulation results of P_d for variables K (at $M = 5$, $N = 10000$), M (at $K = 4$, $N = 10000$), and N (at $K = 4$, $M = 5$), respectively. It is found that the performance is improved as any of these parameters increases.

B. Comparison Between SITC and Other Sensing Algorithms

Thus far, a few efficient sensing algorithms have been proposed in the literature, with each requiring distinct prior information. Here, for fair comparison, we only choose the eigenvalue-based methods proposed in [10]–[12] and the ED method since they both need little prior information. It should be mentioned that the proposed SITC-AIC and SITC-MDL algorithms are equivalent to the GITC algorithm provided in Section V via setting $\gamma_{AIC} = \exp((2p - 1)/N)$ and $\gamma_{MDL} = \exp((p - 0.5) \log N/N)$. Therefore, we omit the performance comparison with the GITC. In the simulation, we fix the order of channel $L=10$ as in [10] and choose $N=10000$, $K=4$, and $M = 5$. Fig. 5 shows the comparison with the

Fig. 6. Comparison with energy detection with noise uncertainty at $M = 5$, $K = 4$, and $N = 10\,000$.

ED method (with perfect estimation of noise covariance) and the four eigenvalue-based methods, namely, the maximum minimum eigenvalue detection (EV-MME) and the energy with minimum eigenvalue detection (EV-EME) [10], the blindly combined ED (EV-BCED) [11], and the arithmetic to geometric mean (EV-AGM) [12]. We see that, under almost the same P_f , ED performs the best, followed by the EV-AGM method, the proposed SITC-AIC method and the EV-BCED method, and then the EV-MME and EV-EME methods. Among the proposed SITC-AIC and four eigenvalue-based methods, the SITC-AIC almost obtains the same performance with the EV-BCED method, and they both outperform EV-MME and EV-EME while being slightly inferior to the EV-AGM method. Although the proposed SITC-MDL method performs the worst in P_d , it is the best among all the considered schemes in terms of P_f performance.

The comparison with ED with noise uncertainty is presented in Fig. 6, where “ED- x dB” means that the noise uncertainty in ED is x dB, as defined in [10]. It is observed that, although the proposed method performs worse than the ED method with accurate noise covariance estimation, it significantly outperforms in both P_d and P_f when there exists noise uncertainty. This clearly demonstrates the robustness of the ITC-based blind-sensing algorithm.

C. Performance of the GITC Algorithm

Results for the GITC sensing algorithm at different threshold values are demonstrated in Fig. 7. It is assumed that we should choose proper thresholds to make $P_f = 0.1$, $P_f = 0.05$, and $P_f = 0.01$. According to Proposition 1 and the discussion in Section V, we choose three thresholds $\gamma = 1.0372$, $\gamma = 1.0393$, and $\gamma = 1.0429$ (note that since the analytical results are slightly larger than the simulation results, the thresholds that we choose should make theoretic P_f larger than required P_f by about 0.02). From the plots, it is found that the P_f requirements are satisfied very well. One can also see that the probability of false alarm is very sensitive to the threshold. Hence, the GITC

Fig. 7. Simulation results for the GITC algorithm for different P_f at $M = 5$, $K = 4$, and $N = 1000$.

sensing algorithm is flexible for system design with different requirements.

VII. CONCLUSION

In this paper, we have provided an intensive study on the ITC-based blind spectrum-sensing method. Based on the prior work on the related study, we first proposed the SITC sensing algorithm. This algorithm significantly reduces the computational complexity without losing any detection performance compared with the existing ITC-based sensing algorithm. Moreover, it enables a more trackable analytical study on the detection performance. Thereafter, applying the recent advances in random matrix theory, we have derived closed-form expressions for both the probability of false alarm and the probability of detection, which can tightly approximate the actual results in simulation. We have further generalized the SITC sensing algorithm to an eigenvalue-based sensing algorithm, which strikes the balance between the probabilities of detection and false alarm by involving an adjustable threshold. Simulation results demonstrate that the proposed blind-sensing algorithm outperforms the existing eigenvalue-based sensing algorithms in certain scenarios.

APPENDIX A

WHITENING THE OVERSAMPLED NOISES

At the secondary receiver, the received continuous signal is usually filtered by a low-pass filter. Therefore, the noise $\mu(t)$ in (6) and (7) should be correlated. We assume that the white noise before the filter is $\hat{\mu}(t)$, and the system function of the low-pass filter is $g(t)$, which is known at the secondary receiver. In the following, we only consider the real-value case, since, in the communication system, the complex-value signal is just the combination of two orthogonal real-value signals. As we have known, $\mu(t)$ can be described by $\hat{\mu}(t)$ and $g(t)$ as

$$\mu(t) = g(t) \otimes \hat{\mu}(t) = \int_0^{t_{\max}} g(\ell) \hat{\mu}(t - \ell) d\ell$$

where $(0, t_{\max})$ represents the time span of $g(t)$, and $\otimes$ denotes the convolution operator. Thus, the autocorrelative function of $\mu(t)$ denoted by $\phi_\mu(\tau)$ can be expressed as

$$\phi_\mu(\tau) = \phi_g(\tau) \otimes \phi_{\hat{\mu}}(\tau)$$

where $\phi_g(\tau)$ and $\phi_{\hat{\mu}}(\tau)$ are the autocorrelative functions of $g(t)$ and $\hat{\mu}(t)$, respectively. Note that $\phi_{\hat{\mu}}(\tau)$ should be equal to $\sigma^2\delta(\tau)$, since $\hat{\mu}(t)$ is white [here, the covariance of $\hat{\mu}(t)$ is assumed to be σ^2]. Therefore, we derive that

$$\phi_\mu(\tau) = \sigma^2 \phi_g(\tau) = \sigma^2 \int_0^{t_{\max}} g(\ell)g(\tau-\ell)d\ell, \quad 0 \leq \tau \leq 2t_{\max}.$$

Thus, if the received signal is oversampled at rate Kf_s , where f_s is the reciprocal of the baseband symbol duration T_0 and K is the oversampling factor, the covariance matrix of the noise vector μ_i given in (19) becomes

$$\mathbf{R}_\mu = \sigma^2 \mathbf{Q}$$

with $\mathbf{Q}$ having entries $q_{i,j} = \phi_g(|i-j|(T_0/K))$. Note that $\mathbf{Q}$ is a positive-definite symmetric matrix. It can be decomposed into $\mathbf{Q} = \tilde{\mathbf{Q}}^2$, where $\tilde{\mathbf{Q}}$ is also a positive-definite symmetric matrix. Hence, to obtain the independent noise samples in the oversampling scheme, we can prewhiten the oversampled noise samples μ_i as

$$\tilde{\mu}_i = \tilde{\mathbf{Q}}^{-1} \mu_i.$$

Then, the covariance matrix of $\tilde{\mu}_i$ transforms into

$$\mathbf{R}_{\tilde{\mu}_i} = \tilde{\mathbf{Q}}^{-1} \mathbf{R}_\mu \tilde{\mathbf{Q}}^{-1} = \sigma^2 \mathbf{I}_p.$$

Now, noise samples μ_i are whitened. It is noted that $\tilde{\mathbf{Q}}$ is only related to the low-pass filter and oversampling factor K and is independent to the signal and noise. Therefore, the prewhitening process can be used blindly.

APPENDIX B PROOF OF LEMMA 1

We prove the lemma from two aspects. First, it has been shown in [32] and [33] that most of the estimation errors of the AIC and the MDL tightly occur around the true numbers. According to this finding, at hypothesis $\mathcal{H}_0$ (the true number of the source signal is zero), if there exists $\hat{k} > 0$ minimizing (2) or (3), then we have $\hat{k} = 1$ with high probability. Hence, Lemma 1 holds for the case of the false alarm. Next, we prove that Lemma 1 succeeds at hypothesis $\mathcal{H}_1$. Since the primary user is present, the eigenvalues l_i of the sampled covariance matrix are distinct at least for $i = 1, 2, \dots, q$ (here, q is the true source number). For $i = q+1, q+2, \dots, p$, the eigenvalues are actually the estimation of noise variance σ^2 . They may be equal to each other when N is large enough. According to the expression of the AIC and the MDL, it is found that the second terms in (2) and (3) are monotonically increasing functions of k . To make the cost function in (2) or (3) minimum at $\hat{k} \in [1, p-1]$, we must have that the first terms in (2) and (3) are monoton-

ically decreasing for $k = 0, 1, \dots, \hat{k}$. We next prove this statement.

We focus on the AIC criterion, and the extension to the MDL is straightforward. Supposing that $k' \in [2, \hat{k}]$ and

$$f_{\text{AIC}}(k) = -2 \log \left(\frac{\prod_{i=k+1}^p l_i^{1/(p-k)}}{\frac{1}{p-k} \sum_{i=k+1}^p l_i} \right)^{(p-k)N}$$

we have

$$f_{\text{AIC}}(k' - 1) - f_{\text{AIC}}(k') = 2N \log \frac{\left(\frac{1}{p-k'+1} \sum_{i=k'+1}^p l_i \right)^{p-k'+1}}{\left(\frac{1}{p-k'} \sum_{i=k'+1}^p l_i \right)^{p-k'} l'_K}.$$

Since

$$\begin{aligned} & \left(\frac{1}{p-k'+1} \sum_{i=k'+1}^p l_i \right)^{p-k'+1} \\ &= \left(\frac{1}{p-k'} \frac{p-k'}{p-k'+1} \sum_{i=k'+1}^p l_i + \frac{1}{p-k'+1} l_{k'} \right)^{p-k'+1} \\ &\geq \left[\left(\frac{1}{p-k'} \sum_{i=k'+1}^p l_i \right)^{\frac{p-k'}{p-k'+1}} l'_K{}^{\frac{1}{p-k'+1}} \right]^{p-k'+1} \\ &= \left(\frac{1}{p-k'} \sum_{i=k'+1}^p l_i \right)^{p-k'} l'_K \end{aligned}$$

here, the arithmetic-mean geometric-mean inequality $a_1 x_1 + a_2 x_2 \geq x_1^{a_1} x_2^{a_2}$ with $a_1 + a_2 = 1$ is applied. Thus, we conclude that

$$\frac{\left(\frac{1}{p-k'+1} \sum_{i=k'+1}^p l_i \right)^{p-k'+1}}{\left(\frac{1}{p-k'} \sum_{i=k'+1}^p l_i \right)^{p-k'} l'_K} \geq 1.$$

It further means that

$$f_{\text{AIC}}(k' - 1) - f_{\text{AIC}}(k') > 0$$

i.e., $f_{\text{AIC}}(k)$ is a monotonic decreasing function. Hence, we have

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{\text{AIC}(0) - \text{AIC}(1)}{N} \\ &= 2 \log \frac{\left(\frac{1}{p} \sum_{i=1}^p l_i \right)^p}{\left(\frac{1}{p-1} \sum_{i=2}^p l_i \right)^{p-1} l_1} + \lim_{N \rightarrow \infty} \frac{-4p+2}{N} > 0. \end{aligned}$$

If N is finite but larger enough, we claim that Lemma 1 holds with a high probability. The high probability is also contributed by the fact that, due to the property of the singular value decomposition technique, the first eigenvalue l_1 is always much larger than other eigenvalues. Therefore, $2N \log(((1/p) \sum_{i=1}^p l_i)^p / ((1/p-1) \sum_{i=2}^p l_i)^{p-1} l_1)$ is large

enough to make Lemma 1 succeed at hypothesis $\mathcal{H}_1$. Thus, we complete the proof of Lemma 1.

APPENDIX C PROOF OF PROPOSITION 2

We firstly derive the derivation of the probability of misdetection P_m (the probability for misdetecting the presence of a primary user at hypothesis H_1) and then obtain the probability of detection P_d through $1 - P_m$. Without loss of generality, the following derivation is also based on the AIC. According to (32), we have

$$P_{m\text{-AIC}|H} = \Pr[\text{AIC}(0) - \text{AIC}(1) < 0 | \mathcal{H}_1].$$

Similar to the process described in the proof of Proposition 1, we can rewrite $P_{m\text{-AIC}|H}$ as

$$P_{m\text{-AIC}|H} = \Pr \left(\log \left[\frac{\left(\frac{1}{p} \sum_{i=1}^p l_i \right)^p}{\left(\frac{1}{p-1} \sum_{i=2}^p l_i \right)^{p-1} l_1} \right] < \frac{4p-2}{2N} \middle| \mathcal{H}_1 \right) \quad (37)$$

where $\{l_1, l_2, \dots, l_p\}$ are the decreasing ordered eigenvalues of the sampled covariance matrix $\mathbf{R}_x$ in (33). When the number of observations N is large enough, we obtain the approximation

$$\frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\dagger \approx \mathcal{E}(\mathbf{x}_i \mathbf{x}_i^\dagger) = \mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger + \sigma^2 \mathbf{I}_p.$$

Thus

$$\frac{1}{p} \sum_{i=1}^p l_i \approx \frac{1}{p} \text{Tr}(\mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger) + \sigma^2.$$

Hence, (37) turns to

$$\begin{aligned} P_{m\text{-AIC}|H} &\approx \Pr \left[\frac{l_1}{\epsilon} \left(p - \frac{l_1}{\epsilon} \right)^{p-1} > \frac{(p-1)^{p-1}}{\exp\left(\frac{2p-1}{N}\right)} \middle| \mathcal{H}_1 \right] \\ &= \Pr \left[y^p - p y^{p-1} + \frac{(p-1)^{p-1}}{\exp\left(\frac{2p-1}{N}\right)} < 0 \middle| \mathcal{H}_1 \right] \end{aligned}$$

where $\epsilon = (1/p) \text{Tr}(\mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger) + \sigma^2$, and $y \triangleq p - (l_1/\epsilon)$.

Assume that π_1 and π_2 (with $\pi_1 < \pi_2$) are two real roots within $(0, p)$ of the following function:

$$g(y) = y^p - p y^{p-1} + \frac{(p-1)^{p-1}}{\exp\left(\frac{2p-1}{N}\right)}. \quad (38)$$

As described in the proof of Proposition 1, the probability of misdetection is concluded as

$$P_{m\text{-AIC}|H} \approx \Pr[\pi_1 < y < \pi_2 | \mathcal{H}_1]$$

that is

$$P_{m\text{-AIC}|H} \approx \Pr[(p - \pi_2)\epsilon < l_1 < (p - \pi_1)\epsilon | \mathcal{H}_1]. \quad (39)$$

Note that l_1 is the largest eigenvalue of the sampled variance matrix $\mathbf{R}_x$. Given the channel matrix, $\mathbf{R}_x$ can be approximated as

$$\begin{aligned} \mathbf{R}_x &\approx \frac{1}{N} \left[\mathbf{H} \sum_{i=1}^N \mathbf{s}_i \mathbf{s}_i^\dagger \mathbf{H}^\dagger \right] + \frac{1}{N} \sum_{i=1}^N \boldsymbol{\mu}_i \boldsymbol{\mu}_i^\dagger \\ &\approx \mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger + \frac{1}{N} \sum_{i=1}^N \boldsymbol{\mu}_i \boldsymbol{\mu}_i^\dagger \end{aligned}$$

when N is larger enough.

Let $\{\delta_1, \delta_2, \dots, \delta_p\}$ and $\{\chi_1, \chi_2, \dots, \chi_p\}$ be the decreasing ordered eigenvalues of $\mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger$ and $(1/N) \sum_{i=1}^N \boldsymbol{\mu}_i \boldsymbol{\mu}_i^\dagger$, respectively. Applying Weyl's inequality theorem in [34], the largest eigenvalue of $\mathbf{R}_x$, i.e., l_1 , satisfies

$$\chi_1 + \delta_p \leq l_1 \leq \chi_1 + \delta_1.$$

Equivalently, χ_1 satisfies

$$l_1 - \delta_1 \leq \chi_1 \leq l_1 - \delta_p.$$

Therefore, there must exist a constant ρ satisfying $\delta_p \leq \rho \leq \delta_1$, which makes $l_1 - \rho$ equal to χ_1 . Then, (39) is rewritten as

$$P_{m\text{-AIC}|H} \approx \Pr[(p - \pi_2)\epsilon - \rho < \chi_1 < (p - \pi_1)\epsilon - \rho | \mathcal{H}_1]$$

that is

$$\begin{aligned} P_{d\text{-AIC}|H} &\approx \Pr \left[\frac{(p - \pi_1)\epsilon - \rho}{\sigma^2} < \frac{\chi_1}{\sigma^2} < p | H_1 \right] \\ &\quad + \Pr \left[0 < \frac{\chi_1}{\sigma^2} < \frac{(p - \pi_2)\epsilon - \rho}{\sigma^2} \middle| \mathcal{H}_1 \right] \end{aligned}$$

where we use the similar constraint for (χ_1/σ^2) as in the proof of Proposition 1. Since χ_1 converges to the Tracy–Widom distribution of order 2, we conclude that

$$P_{d\text{-AIC}|H} \approx Q(\rho)$$

where $Q(\cdot)$ is defined in Proposition 2. Simultaneously, based on (39), the upper and lower bounds for $P_{m\text{-AIC}|H}$ is

$$1 - Q(\delta_1) \leq P_{m\text{-AIC}|H} \leq 1 - Q(\delta_p).$$

Therefore, the upper and lower bounds of $P_{d\text{-AIC}|H}$ can be straightforwardly obtained as

$$Q(\delta_p) \leq P_{d\text{-AIC}|H} \leq Q(\delta_1).$$

The proof for the MDL criterion is the same, except that the function $g(y)$ in (38) is redefined as

$$g(y) = y^p - p y^{p-1} + \frac{(p-1)^{p-1}}{\exp\left(\frac{(p-0.5) \log N}{N}\right)}.$$

Proposition 2 is, thus, proved.

REFERENCES

- [1] "Spectrum Policy Task Force Report," Fed. Commun. Commission, Washington, DC, Tech. Rep. 02-135, Nov. 2002, ET Docket.
- [2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [3] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.
- [4] "FCC-08-260," Fed. Commun. Commission, Washington, DC, Tech. Rep. 08-260, Nov. 2008.
- [5] "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies, notice of proposed rule-making and order," Fed. Commun. Commission, Washington, DC, Tech. Rep. 03-322, Dec. 2003.
- [6] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, Apr. 1967.
- [7] F. F. Digham, M. S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21–24, Jan. 2007.
- [8] B. H. Juang, G. Y. Li, and J. Ma, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.
- [9] A. V. Dandawate and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2355–2369, Sep. 1994.
- [10] 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.
- [11] 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, 2008.
- [12] R. Zhang, T. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.
- [13] S. M. Kay, *Fundamental of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [14] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.
- [15] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.
- [16] J. Shen, T. Jiang, S. Liu, and Z. Zhang, "Maximum channel throughput via cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5166–5175, Oct. 2009.
- [17] G. Ganesan, Y. Li, B. Bing, and S. Li, "Spatiotemporal sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 5–12, Jan. 2008.
- [18] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, Part I: Two user networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2204–2213, Jun. 2007.
- [19] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, Part II: Multiuser networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2214–2222, Jun. 2007.
- [20] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4502–4507, Nov. 2008.
- [21] H. Akaike, "Information theory and an estimation of the maximum likelihood principle," in *Proc. 2nd Int. Symp. Inf. Theory, Problems Control Inf. Theory*, 1973, pp. 267–281.
- [22] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.
- [23] G. Schwartz, "Estimating the dimension of a model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.
- [24] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978.
- [25] B. Zayen, A. M. Hayar, and K. Kansanen, "Blind spectrum sensing for cognitive radio based on signal space dimension estimation," in *Proc. IEEE ICC*, Jun. 14–18, 2009, pp. 1–5.
- [26] B. Zayen, A. M. Hayar, and D. Nussbaum, "Blind spectrum sensing for cognitive radio based on model selection," in *Proc. 3rd CrownCom*, May 15–17, 2008, pp. 1–4.
- [27] M. Haddad, A. M. Hayar, M. H. Fetoui, and M. Debbah, "Cognitive radio sensing information-theoretic criteria based," in *Proc. 2nd CrownCom*, Aug. 1–3, 2007, pp. 241–244.
- [28] B. Zayen and A. M. Hayar, "Cooperative spectrum sensing technique based on sub space analysis for cognitive radio networks," in *Proc. COGIS*, Paris, France, Nov. 2009, pp. 16–18.
- [29] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.
- [30] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [31] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *Ann. Statist.*, vol. 29, no. 2, pp. 295–327, Apr. 2001.
- [32] W. Xu and M. Kaveh, "Analysis of the performance and sensitivity of eigendecomposition-based detectors," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1413–1426, Jun. 1995.
- [33] E. Fishler, M. Grossmann, and H. Messer, "Detection of signals by information theoretic criteria: General asymptotic performance analysis," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1027–1036, May 2002.
- [34] R. Bhatia, *Matrix Analysis*. New York: Springer-Verlag, 1997.

Rui Wang received the B.S. degree in electronic engineering from Anhui Normal University, Wuhu, China, in 2006 and the M.S. degree in electronic engineering from Shanghai University, Shanghai, China, in 2009. He is currently working toward the Ph.D. degree with the Institute of Wireless Communication Technology, Shanghai Jiao Tong University.

His research interests include digital image processing, cognitive radio, and signal processing for wireless cooperative communication.

Meixia Tao (S'00–M'04) received the B.S. degree in electronic engineering from Fudan University, Shanghai, China, in 1999 and the Ph.D. degree in electrical and electronic engineering from Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2003.

From August 2003 to August 2004, she was a Member of the Professional Staff with Hong Kong Applied Science and Technology Research Institute Co. Ltd. From August 2004 to December 2007, she was with the Department of Electrical and Computer

Engineering, National University of Singapore, Singapore, as an Assistant Professor. She is currently an Associate Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. Her current research interests include cooperative transmission, physical-layer network coding, resource allocation of orthogonal frequency-division multiplexing networks, and multiple-input-multiple-output techniques.

Dr. Tao is an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Associate Editor for the IEEE COMMUNICATIONS LETTERS, and an Editor for the *Journal of Communications and Networks*. She served as a Track/Symposium Cochair for the 15th Asia-Pacific Conference on Communications, the Fourth International ICST Conference on Communications and Networking in China in 2009, the 16th IEEE International Conference on Computer Communications and Networks, and the Third IEEE International Conference on Communications, Circuits, and Systems. She has also served as a technical program committee member for various conferences, including the IEEE Global Telecommunications Conference, the IEEE International Conference on Communications, the IEEE Wireless Communication Networking Conference, and the IEEE Vehicular Technology Conference. She was a recipient of the Publication Awards from the Institution of Engineers in Singapore in 2005 and the IEEE Communications Society Asia-Pacific Outstanding Young Researcher Award in 2009.