Blind Spectrum Sensing by Information Theoretic Criteria

Rui Wang and Meixia Tao
Department of Electronic Engineering
Shanghai Jiao Tong University, Shanghai, 200240, P. R. China
Emails: {liouxingrui, mxtao}@sjtu.edu.cn

Abstract—Information theoretic criteria (ITC) based spectrum sensing is a promising blind method which can reliably detect the primary users while requiring little prior information in cognitive radio networks. In this paper, we provide an intensive treatment on the ITC sensing. We first introduce a new over-determined channel model constructed by applying multiple antennas in order to make the 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 (OITC) sensing algorithm, the simplified algorithm significantly reduces the computational complexity without losing any performance. Furthermore, applying the recent advances in random matrix theory, we derive closed-form expressions to tightly approximate both the probability of false alarm and probability of detection. Finally, comprehensive simulations are carried out to evaluate the performance of the proposed ITC sensing algorithms. Results show that they considerably outperform existing blind spectrum sensing methods in certain cases.

I. INTRODUCTION

Cognitive radio (CR) is one of the most promising technologies to deal with irrational frequency regulation policy and has received lots of attention. In cognitive radio networks, secondary (unlicensed) users first reliably sense the primary (licensed) channel and then opportunistically access it without causing harmful interference to primary users. As a result, the detection of presence of primary users is a fundamental and critical task in the cognitive radio networks.

Thus far, there are mainly four types of spectrum sensing methods: energy detection [1], matched filtering (coherent detection) [2], feature detection [3] and eigenvalue-based algorithm [4], each of which requires different prior information and achieves different performance. Among them, energy detection is optimal if the secondary user only knows the local noise power. 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, often referred to as the cyclostationary detection, exploits the periodicity in the modulation scheme which, however, is difficult to determine in certain scenarios. By utilizing the ratio of the maximum and minimum eigenvalues of the sampled covariance matrix to detect the presence of the primary user, the eigenvalue-based sensing method [4] does not need to estimate the power of the noise and hence is a practical scheme in most CR networks.

In this paper, we study a blind spectrum sensing method based on information theoretic criteria (ITC), an approach originally for model selection introduced by Akaike [5] and by Schwartz [6] and Rissanen [7]. Applying ITC for spectrum sensing was firstly introduced in [8]–[11]. This work provides a more intensive study on the ITC sensing algorithm and its performance. First of all, to make the information theoretic criteria applicable, a new over-determined channel model is constructed by introducing multiple antennas. Then, a simplified information theoretic criteria (SITC) sensing algorithm which only involves the computation of two decision values is presented. Compared to the original information theoretic criteria (OITC) sensing algorithm in [8], SITC is much less complex and yet almost has no performance loss.

Furthermore, applying the recent advances in random matrix theory, we then derive closed-form expressions for both the probability of false alarm and probability of detection, which can approximate the actual results in simulation very well. Through simulation, it is also demonstrated that the proposed SITC sensing algorithm considerably outperforms the existing eigenvalue based sensing algorithm.

II. PRELIMINARY ON THE INFORMATION THEORETIC CRITERIA

There are two well-known information theoretic criteria that have been widely used: Akaike information criterion (AIC) and minimum description length (MDL) criterion. One of the most important applications of information theoretic criteria is to estimate the number of source signals in array signal processing [12]. Consider a system model described as

$$x = As + \mu,$$

where $x$ is the $p \times 1$ complex observation vector, $A$ is a $p \times q$ ($p > q$) complex system matrix, $s$ denotes the $q \times 1$ complex source modulated signals and $\mu$ is the additive complex white Gaussian noise vector. It is noted that the definite parameters $q$, $A$ and $\sigma^2$ are all unknown. The resulting cost functions of

This work is supported in part by the Ministry of Science and Technology (MOST) of China under grants 2006ZX03003-004 and 2008BAH30B09, the Doctoral Fund of Ministry of Education of China under grant 200802481002, and Shanghai Pujiang Talent Program under grant 09PJ1406000.
AIC and MDL have the following form [12]:

\[
\text{AIC}(k) = -2\log \left( \frac{\prod_{p=1}^{P} l_{i}^{1/(p-k)}}{\sum_{i=1}^{k} l_{i}} \right) N^{(p-k)} + 2k(2p-k) + 2,
\]

\[
\text{MDL}(k) = -\log \left( \frac{\prod_{p=1}^{P} l_{i}^{1/(p-k)}}{\sum_{i=1}^{k} l_{i}} \right) N^{(p-k)}
\]

\[
+ \left( \frac{1}{2} k(2p - k) + 1 \right) \log N,
\]

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,

\[
\hat{k}_{\text{AIC/MDL}} = \arg_{j=0,1,\ldots,p-1} \min \text{AIC/MDL}(j).
\]

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 \( x(n) \) be the discrete baseband received signal sampled at a rate equal to the reciprocal of the baseband symbol duration at the secondary user’s receiver. Spectrum sensing can be formulated as a binary hypothesis test between the following two hypotheses

\[
\mathcal{H}_0 : \quad x(n) = \mu(n),
\]

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

where \( s(n) \) denotes the symbol transmitted by the primary user, \( h(i) \) \((0 \leq i \leq L - 1)\) denotes the discrete channel response with \( L \) being the order of the discrete channel (\( L \) taps), and \( \mu(n) \) denotes the additive white noise. Let each observation consist of \( M \) received signal samples. The channel response is also assumed to remain invariant during each observation. Then (4) and (5) can be rewritten in matrix form as:

\[
\mathcal{H}_0 : \quad x_i = \mu_i,
\]

\[
\mathcal{H}_1 : \quad x_i = Hs_i + \mu_i,
\]

where \( H \) is an \( M \times (L+M-1) \) circular channel matrix, \( x_i, s_i \), and \( \mu_i \) are the \( M \times 1 \) observation vector, \( (L+M-1) \times 1 \) source signal vector and \( M \times 1 \) noise vector, respectively, and are defined as

\[
x_i = [x(iM-M+1), x(iM-M+2), \ldots, x(iM)]^T,
\]

\[
s_i = [s(iM-M-L+2), s(iM-M-L+3), \ldots, s(iM)]^T,
\]

\[
\mu_i = [\mu(iM-M+1), \mu(iM-M+2), \ldots, \mu(iM)]^T.
\]

To construct an over-determined channel matrix \( H \) as in (1), one needs to enlarge the observation space. Obviously, simply increasing the observation window \( M \) does not work. Here we propose to expand the observation space by employing multiple receive antennas at the secondary user to increase the spatial dimensionality. In specific, suppose that the detector at the secondary user is equipped with \( K \) antennas. Redefine (8) and (10) as

\[
x_i = [x_1^i(1), x_2^i(1), \ldots, x_K^i(1), x_1^i(2), \ldots, x_K^i(2), \ldots, x_1^i(M), x_2^i(M), \ldots, x_K^i(M)]^T,
\]

\[
\mu_i = [\mu_1^i(1), \mu_2^i(1), \ldots, \mu_K^i(1), \mu_1^i(2), \ldots, \mu_K^i(2), \ldots, \mu_1^i(M), \ldots, \mu_K^i(M)]^T,
\]

where \( x_i^k = [x_1^k(1), x_2^k(2), \ldots, x_K^k(M)]^T \) represents the \( M \times 1 \) observation vector at the \( k \)-th antenna at the \( i \)-th observation as in (8), and \( \mu_i^k = [\mu_1^k(1), \mu_2^k(1), \ldots, \mu_K^k(M)]^T \) is the corresponding noise vector at the \( k \)-th antenna at the \( i \)-th observation as in (10). Then, the new channel matrix \( H \) becomes an \( MK \times (M+L-1) \) matrix:

\[
H = \begin{bmatrix}
h_1(L-1) & h_1(L-2) & \ldots & h_1(0) \\
& \vdots & \ddots & \vdots \\
h_K(L-1) & h_K(L-2) & \ldots & h_K(0) \\
& \vdots & \ddots & \vdots \\
h_1(L-1) & h_1(L-2) & \ldots & h_1(0) \\
& \vdots & \ddots & \vdots \\
h_K(L-1) & h_K(L-2) & \ldots & h_K(0)
\end{bmatrix}.
\]

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

\[
K > \frac{L+M-1}{M},
\]

or, alternatively,

\[
M > \frac{L-1}{K-1}.
\]

Furthermore, we assume the noise samples observed from different antennas are independent with zero mean and \( \mathcal{E}(\mu_i \mu_i^H) = \sigma^2 I_{MK} \). Then we can exactly ensure that the system mode under multiple antennas satisfies the over-determined condition specified in [12]. For ease of presentation, we define \( p = MK \) and \( q = L+M-1 \) in (7) for the rest of the paper.

IV. SIMPLIFIED INFORMATION THEORETIC CRITERIA SENSING ALGORITHM

Since the binary hypothesis test in the spectrum sensing is equivalent to the special case of source number estimation problem, the information theoretic criteria method can be directly applied to conduct spectrum sensing as firstly proposed in [8]–[11]. The basic idea is when the primary user is absent, the received signal \( x_i \) is only the white noise samples. Therefore, the estimated number of source signals via
information theoretic criteria (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.

However, it is known that signal detection is much easier than signal estimation. Therefore, using the estimation method to conduct the detection as in [8]–[11] may lead to unnecessary computational complexity overhead. Meanwhile, it makes it difficult to carry out analytical study on the detection performance. In this section, we propose a simplified ITC 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.

Before presenting the simplified ITC sensing algorithm in detail, we have the following lemma, the proof of which can be found in [13] and is omitted here due to space limitation.

**Lemma 1:** If there is one value \( k (> 0) \) which minimizes the AIC metric in (2) (MDL metric in (3)), then AIC(0) > AIC(1) (MDL(0) > MDL(1)) with high probability.

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

**Algorithm 1: SITC sensing algorithm**

1. Compute the sampled covariance matrix of received signals, i.e., \( R_x = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^\dagger \), where \( x_i \)'s are the received vectors as described in (11) and \( N \) denotes the number of the observations.
2. Obtain the eigenvalues of \( R_x \) through eigenvalue decomposition technique, and denote them as \( \{l_1, l_2, \ldots, l_p\} \) with \( l_1 \geq l_2, \ldots, \geq l_p \).
3. Calculate the decision values AIC(0) and AIC(1) (MDL(0) and MDL(1)) according to (2)(3). Then the detection decision metric is

\[
T_{SITC\_AIC/MDL}(L_x) : AIC/MDL(0) \begin{cases} > \sqrt{\mu} & \text{AIC/MDL(1)} \\
\end{cases}
\]

where \( L_x \) denote the set of eigenvalues \( \{l_i, i = 1, 2, \ldots, p\} \).

Note that in the OITC sensing algorithm [8], one needs to find the exact value of \( k \) from 0 to \( p \) to minimize the AIC in (2) or 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 the next section, based on the proposed SITC algorithm, we present the analytical results on the detection performance. Since from the Lemma 1, the SITC algorithm almost obtains the same performance as OITC algorithm, we claim that our analytical results are also applicable for evaluating the performance of OITC algorithm.

**V. PERFORMANCE ANALYSIS**

For ease of presentation, we shall take the AIC criterion for example to illustrate the analysis throughout this section. The extension to MDL criterion is straightforward if not mentioned otherwise.

**A. Probability of false alarm**

According to the sensing steps in Algorithm 1, the false alarm occurs when AIC(0) is larger than AIC(1) at hypothesis \( H_0 \). The probability of false alarm can be expressed as

\[
P_{f \_AIC} = \Pr(\text{AIC(0)} > \text{AIC(1)} | H_0).
\]

Since the primary user is absent, the received signal \( x_t \) only contains the noises. The sampled covariance matrix \( R_x \) in Algorithm 1 thus turns to \( R_{\mu} \) defined as

\[
R_{\mu} = \frac{1}{N} \sum_{i=1}^{N} \mu_i \mu_i^\dagger.
\]

Hence, the eigenvalues in (2) become the eigenvalues of the sampled noise covariance matrix \( R_{\mu} \) in (13), which is a Wishart random matrix [14]. 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:

\[
P_f \approx F_2 \left( \frac{pN - (\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p})(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}})^2} \right) - F_2 \left( \frac{(p - \alpha_1)N - (\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p})(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}})^2} \right) + F_2 \left( \frac{(p - \alpha_2)N - (\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p})(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}})^2} \right) - F_2 \left( \frac{-(\sqrt{N} + \sqrt{p})^2}{(\sqrt{N} + \sqrt{p})(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}})^2} \right)
\]

where \( F_2(\cdot) \) is the cumulative distribution function (CDF) of Tracy-Widom distribution of order two [14], \( \alpha_1 \) and \( \alpha_2 \) with \( \alpha_1 < \alpha_2 \) are the two real roots of the following function

\[
f(x) \triangleq x^p - px^{p-1} + \frac{(p - 1)x^{p-1}}{\exp \left( \frac{2p-1}{N} \right)},
\]

if AIC is applied, or

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

if MDL is applied.

**Proof:** Please refer to [13].

From Proposition 1, it is found that the probability of false alarm is independent with noise covariances \( \sigma^2 \). 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 \( K \).
B. Probability of detection

When the primary user is present, the probability of detection is described as

\[ P_{d−AIC} = \Pr(\text{AIC}(0) > \text{AIC}(1)|\mathcal{H}_1). \] \hspace{1cm} (14)

Since at \( \mathcal{H}_1 \), the received vector \( x \) involves the signals transmitted by the primary user, the sampled covariance matrix \( \textbf{R}_x \) can be written as

\[ \textbf{R}_x = \frac{1}{N} \sum_{i=1}^{N} (\textbf{H}s_i + \mu_i)(\textbf{H}s_i + \mu_i)^\dagger. \] \hspace{1cm} (15)

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

**Proposition 2:** Let \( \textbf{R}_x \) denote the covariance matrix of \( s_i \) given in (9) and \( \{\delta_1, \delta_2, \ldots, \delta_p\} \) be the eigenvalues of matrix \( \textbf{H}\textbf{R}_x\textbf{H}^\dagger \) (with \( \delta_1 \geq \delta_2 \geq \ldots \geq \delta_p \)). Then there exists a value of \( \rho \), for \( \delta_p \leq \rho \leq \delta_1 \), such that the probability of detection given \( \textbf{H} \) can be approximated as \( P_{d|H} \approx Q(\rho) \), where the function \( Q(\cdot) \) is

\[
Q(\delta) = F_2\left(\frac{pN - (\sqrt{N} + \sqrt{\rho})^2}{(\sqrt{N} + \sqrt{\rho})(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{\rho}})^\frac{3}{2}}\right) - F_2\left(\frac{(\frac{p-1}{p})\epsilon - \delta}{\sqrt{p}}\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{\rho}}\frac{1}{\sqrt{\rho}}\right) + F_2\left(\frac{(\frac{p-1}{p})\epsilon - \delta}{\sqrt{p}}\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{\rho}}\frac{1}{\sqrt{\rho}}\right) - F_2\left(-\frac{d}{\sqrt{N} + \sqrt{\rho}}\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{\rho}}\frac{1}{\sqrt{\rho}}\right),
\]

where \( \epsilon = \frac{1}{p}\text{Tr}(\textbf{H}\textbf{R}_x\textbf{H}^\dagger) + \sigma^2 \) and \( \pi_1, \pi_2 \) (with \( \pi_1 < \pi_2 \)) denote the two real roots of the function

\[
g(y) = y^p - py^{p-1} + \frac{(p-1)y^{p-1}}{\exp\left(\frac{2p-1}{N}\right)}\] \hspace{1cm} (16)

for AIC, or

\[
g(y) = y^p - py^{p-1} + \frac{(p-1)y^{p-1}}{\exp\left(\frac{p-0.5}{N}\log N\right)}\] \hspace{1cm} (17)

for MDL.

**Proof:** Please refer to [13].

From Proposition 2, we find that \( P_d \) is not only related to \( N \) and \( p \), but also depends on \( \sigma^2 \), which is the ratio of the signal strength of 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. In practice, we can simply choose \( \rho_{AIC} = \frac{1}{\pi}(\delta_p + \delta_1) \) and \( \rho_{MDL} = \frac{3}{4}(\delta_p + \delta_1) \). It will be demonstrated later in Section VI that the analytical \( P_{d|H} \) based on this choice of \( \rho \) can approximate the Monte Carlo results very well in most of cases.

**VI. SIMULATION RESULTS AND DISCUSSIONS**

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

\[
SNR = \frac{\mathcal{E}[\|x_i - \mu_i\|^2]}{\mathcal{E}[\|\mu_i\|^2]}.
\]

**A. Comparison between simulation and analytical results**

The comparison of simulation and analytical results for \( P_f \) is demonstrated in Table I. We observe that for AIC (since \( P_f \) for SITC and OITC are almost equal to each other, we here use AIC to denote both of them for simplicity), the analytical results are slightly larger than the simulation results. Nevertheless, the analytical approximation is accurate enough to evaluate the performance of the proposed sensing scheme.

It is also found that \( P_f − AIC \) gradually decreases as \( p = MK \) increases while the \( P_f − MDL \) remains zero in both simulation and analytical results. We conclude that the MDL method has excellent false alarm performance.

The theoretical analysis for \( P_d \) is validated in Fig. 1. It is first seen that the proposed SITC sensing algorithm do not lead to any performance loss compared to OITC algorithm.

Then, comparing the semi-analytical results obtained from

![Fig. 1. Simulation and theoretic results about probability of detection at different \((M, K, N)\).](image-url)
see that the proposed method significantly outperforms the energy detection method in both $P_d$ and $P_f$ when there exists noise uncertainty. This clearly demonstrated the robustness of information theoretic criteria based blind sensing algorithm.

VII. CONCLUSIONS

This paper provides an intensive study on the information theoretic criteria based blind spectrum sensing method. After introducing one simplified ITC sensing algorithm without losing any detection performance compared with the existing ITC algorithm, we then derived the closed-form expressions to tightly approximate both the probability of false alarm and probability of detection based on the recent advances in random matrix theory. Simulation results showed that the proposed blind sensing algorithm with AIC criterion achieves higher probability of detection than the eigenvalue-based sensing algorithms while having the same probability of false alarm. Meanwhile, the proposed algorithm with MDL criterion achieves the lowest probability of false alarm among all the considered blind sensing methods.

REFERENCES