2018 IEEE Wireless Communications and Networking Conference (WCNC)

Novel Distributed Spectrum Sensing Techniques for Cognitive Radio Networks

Peter J. Smith*, Rajitha Senanayake†, Pawel A. Dmochowski * and Jamie S. Evans†
* Victoria University of Wellington, New Zealand
† University of Melbourne, Australia

***Abstract*—We consider distributed spectrum sensing in cognitive radio networks with multiple primary and secondary user terminals. Two novel techniques based on the sphericity test, namely, the multisample sphericity test and meta analysis, are analysed in such a scenario. Instead of sending all the raw data received at the secondary user terminals, as in the case with centralized spectrum sensing, in the multisample sphericity and meta analysis tests only one or two real numbers are required to be sent to the central processor to make a decision about the presence of primary users. Accurate analytical expressions on the false alarm probability are derived for both techniques and numerical examples are provided to verify their accuracy. Receiver operating characteristic (ROC) curves are also presented to compare the performance of the proposed methods and other simple fusion techniques.**

***Index Terms*—Cognitive radio, cooperative spectrum sensing, sphericity test, meta analysis.**

I. Introduction

In view of the radio spectrum's scarcity, it is necessary to find new techniques for its efficient use in wireless communications. Cognitive radio is a promising technology that can be used to efficiently utilize the radio spectrum by allowing unlicensed secondary users to share the spectrum resources of licensed primary users [1], [2]. One principal requirement of this is that secondary user transmission does not cause intolerable interference to the primary users. This is achieved by the secondary users' ability to detect the presence of primary users, which is commonly referred to as spectrum sensing. If no primary users are detected, the secondary users are allowed to utilise the primary user's licensed spectrum.

In the cooperative spectrum sensing literature, many optimal sensing techniques have been proposed and analysed under the assumption of a single active primary user [3], [4]. However, in cellular systems the existence of multiple, simultaneously transmitting primary users is a prevailing condition. For such systems, a novel spectrum sensing algorithm, based on the sphericity test, has been proposed in [5]. The authors use the optimal generalized likelihood ratio test (GLRT) paradigm and simulate the false alarm and detection probabilities. However, no analytical derivation relating to the performance measures was presented. In [6], the authors adopt the same model as in [5] and derive analytical formulae to approximate the false alarm and detection probabilities in the presence of multiple primary users. The approximation was derived by matching the moments of test statistics to a Beta distribution. For the special case of only two secondary users the approximations in [6] reduce to the exact performance measures. In [7] and [8], the authors adopt the same optimal GLRT paradigm and derive different approximations for the false alarm and detection probabilities based on a Gaussian distribution.

While the optimal GLRT technique considered in [5]–[8] exhibits high detection performance, it requires a central processor that gathers and processes the raw data received by all the secondary user terminals. This centralized architecture limits the applicability of the GLRT technique to large networks as it requires the exchange and the processing of a massive amount of raw data at the central processor.

In this paper, we consider a general cognitive radio system with an arbitrary number of primary and secondary user terminals. Based on the sphericity test and the GLRT statistic we propose two new distributed cooperative spectrum sensing techniques, namely the multisample sphericity test and meta analysis, that require only partial information to be sent to the central processor. In the multisample sphericity test each secondary user calculates and sends only two real numbers, specifically, the trace and the determinant of the sample covariance matrix. In meta analysis each secondary user calculates and sends only one real number, i.e., the probability value, more commonly known as the p-value, of the current test statistic. We derive accurate analytical expressions to measure the probability of false alarm for both techniques. Using the receiver operating characteristics (ROC), curves we compare the performance of the proposed techniques. We observe that both the multisample sphericity test and meta analysis have similar performance even though meta analysis requires only half the amount of information sent to the central processor. Furthermore, we compare the performance of the novel techniques with other simple binary fusion methods, where each secondary user terminal makes a binary decision about the presence of primary users and all the decisions are simply combined through binary AND or OR operation at the central processor.

II. System Model

Consider a wireless communications system where M secondary user terminals are tasked with cooperatively determining the presence of P primary users. The secondary

This work is supported in part by the Australian Research Council (ARC) through the Discovery Early Career Researcher (DECRA) Award DE180100501.

978-1-5386-1734-2/18/$31.00 ©2018 IEEE

user terminal m is equipped with Q_m antennas such that, $K = \sum_{m=1}^{M} Q_m$. In this system, each secondary user terminal is connected to a central processor that cooperatively detects the presence of primary users. This detection problem can be formulated as a binary hypothesis test, where the null hypothesis, $\mathcal{H}_0$, denotes the absence of primary users and the non-null hypothesis, $\mathcal{H}_1$, denotes the presence of primary users. First, let us assume that all the signals received at the secondary user terminals are available at the central processor. The n-th sample vector, $\mathbf{x}_n$, received at the central processor can then be modelled as

$$\begin{aligned} \mathcal{H}_0 :& \quad \mathbf{x}_n = \mathbf{w}_n && \text{no primary users present} \\ \mathcal{H}_1 :& \quad \mathbf{x}_n = \mathbf{H}\mathbf{s} + \mathbf{w}_n && \text{primary users present,} \end{aligned} \tag{1}$$

where $\mathbf{s} = [s_1, s_2, \ldots, s_P]^T$ is the $P \times 1$ data vector, which contains the zero mean transmitted symbols from the P primary users, $\mathbf{w}_n$ is the $K \times 1$ additive white Gaussian noise vector at the n-th sample with $\mathbb{E}[\mathbf{w}_n \mathbf{w}_n^\dagger] = \sigma^2 \mathbf{I}_K$ and, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_P]$ is the $K \times P$ channel matrix between the P primary users and K receive antennas. We collect N independent and identically distributed sample vectors such that the observed received signal matrix is $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$. During this time we assume that the channel $\mathbf{H}$ remains constant. Note that unless otherwise specified, the results in this paper do not assume a specific distribution for $\mathbf{H}$. We also assume that the primary users signal follows an independent and identically distributed zero mean Gaussian distribution and is uncorrelated with the noise.

A. Sphericity Test

Let us define the sample covariance matrix $\mathbf{R} = \mathbf{X}\mathbf{X}^\dagger$ such that when no primary users are present $\mathbf{R}$ follows an uncorrelated complex Wishart distribution with population covariance matrix

$$\boldsymbol{\Sigma} = \mathbb{E}[\mathbf{X}\mathbf{X}^\dagger]/N = \sigma^2 \mathbf{I}_K. \tag{2}$$

When primary users are present, for a given $\mathbf{H}$, $\mathbf{R}$ follows a correlated complex Wishart distribution with population covariance matrix

$$\boldsymbol{\Sigma} = \sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^\dagger + \sigma^2 \mathbf{I}_K, \tag{3}$$

where $\gamma_i = \mathbb{E}[s_i s_i^\dagger]$ defines the transmission power of the i-th primary user. Thus, the hypothesis testing in (1) can be rearranged to test the structure of the population covariance matrix $\boldsymbol{\Sigma}$ as

$$\begin{aligned} \mathcal{H}_0 :& \quad \boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_K && \text{no primary users present} \\ \mathcal{H}_1 :& \quad \boldsymbol{\Sigma} \succ \sigma^2 \mathbf{I}_K && \text{primary users present,} \end{aligned} \tag{4}$$

where $\mathbf{A} \succ \mathbf{B}$ denotes that $\mathbf{A} - \mathbf{B}$ is a positive definite matrix. Note that no a priori knowledge on $\mathbf{H}, P$ and σ^2 is assumed at the secondary user terminals. The most critical information for the secondary user is whether or not there are active primary users and the number of active users is not relevant. As such, we reject $\mathcal{H}_0$ if we have reason to believe that $\boldsymbol{\Sigma}$ departs from the spherical structure of $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_K$, which is the well-known sphericity test [6], [9].

The generalized likelihood ratio used for determining $\mathcal{H}_0$ or $\mathcal{H}_1$ can be written as

$$\mathcal{L} = \frac{\max_{\sigma^2 \in \mathbb{R}^+} L(\mathbf{X}|\sigma^2 \mathbf{I}_K)}{\max_{\mathbf{H} \in \mathcal{C}^{K \times P}, (\gamma_1, \ldots, \gamma_P, \sigma^2) \in \mathbb{R}^+} L(\mathbf{X}|\mathbf{H}, \gamma_1, \ldots, \gamma_P, \sigma^2 \mathbf{I}_K)}, \tag{5}$$

where $L(\mathbf{X}|\sigma^2 \mathbf{I}_K)$ is the likelihood function of the observation matrix under hypothesis $\mathcal{H}_0$ and $L(\mathbf{X}|\mathbf{H}, \gamma_1, \ldots, \gamma_P, \sigma^2 \mathbf{I}_K)$ is the likelihood function of the observation matrix under hypothesis $\mathcal{H}_1$. Finding the maximum likelihood estimates of σ^2 under $\mathcal{H}_0$, and of $\gamma_1, \ldots, \gamma_P, \sigma^2$ and $\mathbf{H}$ under $\mathcal{H}_1$ results in the GLRT statistic [10]

$$\mathrm{TS} = \frac{|\mathbf{R}|}{\left(\frac{1}{K}\mathrm{tr}(\mathbf{R})\right)^K}. \tag{6}$$

Now, to determine $\mathcal{H}_0$ or $\mathcal{H}_1$ admits

$$\mathrm{TS} \gtrless_{\mathcal{H}_1}^{\mathcal{H}_0} \zeta, \tag{7}$$

where ζ is a user-specified detection threshold. More details on the derivation of TS can be found in [6], [10].

B. Performance Measures

To evaluate the performance of the GLRT statistic in (6) we define the false alarm probability P_{fa} and the detection probability P_d. The false alarm probability is the probability of wrongly declaring $\mathcal{H}_0$, i.e., $\mathcal{H}_1$ is chosen given that $\mathcal{H}_0$ is the true hypothesis. It can be defined as

$$P_{\text{fa}} = \Pr[\mathrm{TS}_{\mathcal{H}_0} \le \zeta], \tag{8}$$

where $\mathrm{TS}_{\mathcal{H}_0}$ is TS defined under hypothesis $\mathcal{H}_0$. From (8), note that P_{fa} is the CDF of $\mathrm{TS}_{\mathcal{H}_0}$, evaluated at ζ which can be denoted by $F_{\mathrm{TS}_{\mathcal{H}_0}}(\zeta)$. The threshold ζ should be carefully chosen such that P_{fa} is small and it can be calculated by numerically inverting $F_{\mathrm{TS}_{\mathcal{H}_0}}(\zeta)$, i.e.,

$$\zeta = F^{-1}_{\mathrm{TS}_{\mathcal{H}_0}}(P_{\text{fa}}). \tag{9}$$

The detection probability is the probability of declaring correctly $\mathcal{H}_1$, i.e., $\mathcal{H}_1$ is chosen given that $\mathcal{H}_1$ is the true hypothesis. It can be defined as

$$P_{\text{d}} = \Pr[\mathrm{TS}_{\mathcal{H}_1} \le \zeta] = F_{\mathrm{TS}_{\mathcal{H}_1}}(\zeta), \tag{10}$$

where $\mathrm{TS}_{\mathcal{H}_1}$ is TS defined under hypothesis $\mathcal{H}_1$ and $F_{\mathrm{TS}_{\mathcal{H}_1}}(\zeta)$ is the CDF of $\mathrm{TS}_{\mathcal{H}_1}$ evaluated at ζ. In this paper, the focus is only on analysing the P_{fa} and when necessary P_d is evaluated numerically.

Note that while testing the GLRT statistic in (6) is the optimal approach in terms of maximizing the likelihood ratio, it is not very suitable for practical implementations of cognitive radio networks. This is because in order to calculating TS requires the sharing of all raw data received at the M secondary user terminals with the central processor, producing considerable network overhead. The focus of the paper is to find new distributed spectrum sensing techniques that reduce such overhead. Based on the sphericity test and the GLRT statistic, in the following we propose two new spectrum sensing techniques namely, the multisample sphericity test and meta analysis, that require the exchange of only one or two real numbers, with the central processor.

III. Multisample Sphericity Test

In this section we present the multisample sphericity test for cooperative spectrum sensing. Under this distributed setup we define the two hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$, for sphericity as

$$\begin{aligned} \mathcal{H}_0: \ & \boldsymbol{\Sigma}_m = \sigma^2 \mathbf{I}_{Q_m} \quad \text{for all } m \in \{1,2,\dots,M\}, \\ \mathcal{H}_1: \ & \boldsymbol{\Sigma}_m \succ \sigma^2 \mathbf{I}_{Q_m} \quad \text{for at least one } m \in \{1,2,\dots,M\}, \end{aligned} \tag{11}$$

where $\boldsymbol{\Sigma}_m = \mathbb{E}[\mathbf{X}_m \mathbf{X}_m^\dagger]/N$ is the population covariance matrix with $\mathbf{X}_m$ denoting the $Q_m \times N$ observed data matrix at the secondary user terminal m. Following the derivation methodology in [11, Theorem 8.3.2], the modified GLRT statistic for testing the above hypothesis can be obtained as

$$\text{MSTS} = \frac{\prod_{m=1}^{M} |\mathbf{R}_m|}{\left(\frac{1}{K}\sum_{m=1}^{M} \text{tr}(\mathbf{R}_m)\right)^K}, \tag{12}$$

where $\mathbf{R}_m = \mathbf{X}_m \mathbf{X}_m^\dagger$ is the sample covariance matrix at the secondary user terminal m. Note that the MSTS in (12) reduces to TS in (6) when $M = 1$. Thus, instead of sending the whole matrix $\mathbf{X}_m$, which is the case with the optimal GLRT statistic, TS, the multisample sphericity test requires the secondary user terminal m to calculate $|\mathbf{R}_m|$ and $\text{tr}(\mathbf{R}_m)$ and share with the central processor the two real numbers. The central processor collects all matrix determinants and traces from the M secondary user terminals and calculates the global MSTS according to (12). We determine $\mathcal{H}_0$ or $\mathcal{H}_1$ in the multisample sphericity test as

$$\text{MSTS} \gtrless_{\mathcal{H}_1}^{\mathcal{H}_0} \zeta. \tag{13}$$

In order to analyse the performance of the multisample sphericity test, we next proceed to derive an analytical expression for P_{fa} based on (12).

A. Probability of False Alarm

Based on [6], we first write the n-th moment of the N-th power of the MSTS as

$$\begin{aligned} & \text{E}[(\text{MSTS}^N)^n] \\ & = \frac{M^{nNM}\Gamma(NM)}{\Gamma(NM+nNM)} \prod_{m=1}^{M} \prod_{q=1}^{Q_m} \frac{\Gamma(N+nN+1-q)}{\Gamma(N+1-q)}, \end{aligned} \tag{14}$$

where $\Gamma(\cdot)$ defines the gamma function. Whilst not presenting all the steps here due to page limitation, recognizing that these moments have the same form[1] as [10, Section 8.5.1, eq. (1)], we follow the expansion method in [10, Section 8.5.1] and derive an approximation to the CDF of $-2\rho\ln(\text{MSTS})$ under the null hypothesis as

$$\Pr[-2\rho\ln(\text{MSTS}) \le x] \approx \Pr[\chi_f^2 \le x], \tag{15}$$

where $f = \sum_{m=1}^{M} Q_m^2 - 1$, $\rho = N - \frac{1+K^2-2K\sum_{m=1}^{M} Q_m^3}{6K(1-\sum_{m=1}^{M} Q_m^2)}$ and χ_f^2 denotes a chi-squared distribution with f degrees of freedom. Note that the full expression for the CDF of $-2\rho\ln(\text{MSTS})$ in [10] contains additional correction terms written in-terms of higher order chi-squared distributions. For simplicity, we ignore these correction terms and approximate $\Pr[-2\rho\ln(\text{MSTS}) \le x]$ by the chi-squared distribution in (15). Following simple mathematical manipulations (15) can be expressed as

$$\Pr[\text{MSTS} \ge \exp(-x/2\rho)] \approx \Pr[\chi_f^2 \le x]. \tag{16}$$

We can derive an approximation for the CDF of MSTS as

$$\Pr[\text{MSTS} \le x] \approx 1 - \Pr[\chi_f^2 \le -2\rho\ln(x)]. \tag{17}$$

Thus, P_{fa} for the multisample sphericity test can be approximated as

$$P_{\text{fa}} \approx 1 - \Pr[\chi_f^2 \le -2\rho\ln(\zeta)]. \tag{18}$$

In the following, we use numerical examples to illustrate that the expression in (18) provides a very accurate approximation of P_{fa} for different network scenarios.

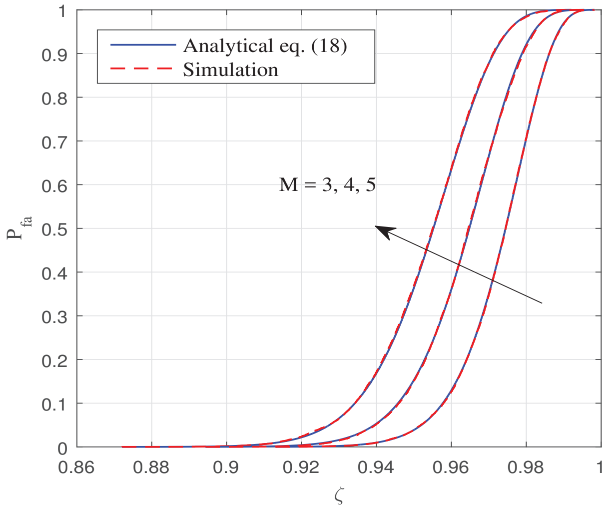

Fig. 1. Probability of false alarm vs the detection threshold when $N = 200$, $Q_m = 2$ and $M = 3$, 4 and 5.

B. Numerical Examples

Fig. 1 plots P_{fa} versus the detection threshold ζ for different numbers of secondary user terminals, by setting $M = 3, 4$ and 5. The secondary user terminals are allocated $Q_m = 2$ antennas each and make $N = 200$ observations. For each simulation trial, the noise components are drawn from an independent complex Gaussian distribution. The simulation curves, illustrated in dashed lines, are generated based on

[1] Note that the expression in [10, Section 8.5.1, eq. (1)] reduces to (14) when $K = \Gamma(NK)\left(\prod_{m=1}^{M}\prod_{q=1}^{Q_m}\Gamma(N+1-q)\right)^{-1}$, $b=1$, $a=K$, $y_1 = NK$, $x_k = N \ \forall\, k \in \{1,2,\dots,M\}$, $\eta_1 = 0$ and $\{-\xi_1, -\xi_2, \dots, -\xi_K\} = \{0,\dots,Q_1-1, 0,\dots,Q_2-1,\dots,0,\dots,Q_M-1\}$.

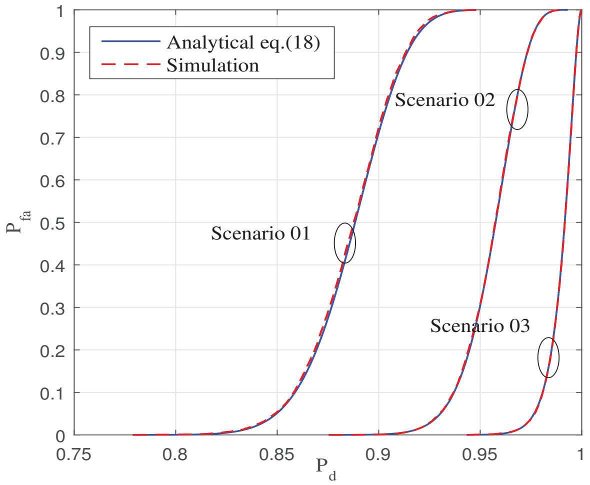

Fig. 2. Probability of false alarm vs the detection threshold for scenario 01, 02 and 03 with N = 200, $M = 5$.

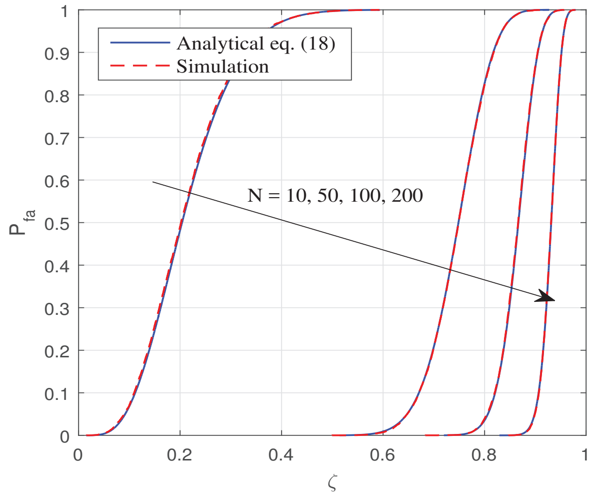

Fig. 3. Probability of false alarm vs the detection threshold when M = 4, $Q_1 = 3$, $Q_2 = 1$ $Q_3 = 2$, $Q_4 = 4$ and N = 50, 100 and 200.

Monte-Carlo simulation, while the analytical curves, illustrated in solid lines, are generated using the approximation in (18). The figure illustrates extremely close agreement between the approximation and the simulation curves and, for a given detection threshold, an increase in P_{fa} with increasing M.

Fig. 2 plots P_{fa} versus the detection threshold ζ for different number of antennas at each secondary user terminal. For fixed $M = 5$, $N = 200$, we consider the following three scenarios. In scenario 01, $Q_1 = 3$, $Q_2 = 1$ $Q_3 = 2$, $Q_4 = 4$ and $Q_5 = 2$. In scenario 02, $Q_1 = 2$, $Q_2 = 1$ $Q_3 = 1$, $Q_4 = 3$ and $Q_5 = 2$, and in scenario 03, $Q_m = 1, \forall m \in \{1, 2, \ldots 5\}$. As such, the root mean square (RMS) value of Q_1, Q_2 and Q_3 are 2.6, 1.67 and 1, respectively. The figure illustrates that the approximation is very accurate for all three scenarios and for a given detection threshold the P_{fa} increases with the increasing RMS value of Q_m.

Finally, Fig. 3 plots P_{fa} versus the detection threshold ζ for different number of sample sizes N. Here we select an example where $M = 4$, $Q_1 = 3$, $Q_2 = 1$ $Q_3 = 2$, $Q_4 = 4$ and examine $N = 10, 50, 100$ and 200. Once again, the figure illustrates that for all four cases the approximation in (18) is very accurate. Note that the accuracy of (18) decreases for small N, since the approximation ignores higher order terms of inverse powers of N. Despite this, however, even for a very small value of $N = 10$, the approximation remains tight, as witnessed in Fig. 3. The figure also illustrates that for a given detection threshold the P_{fa} decreases with increasing N.

IV. Meta Analysis

In this section we present the technique of meta analysis for cooperative spectrum sensing. Meta analysis refers to the synthesis of data from multiple independent tests. Several techniques of meta analysis are available in the literature. In this paper, we use the Fisher's method which combines the results from several independent tests bearing upon the same overall hypothesis to produce a global test statistic [12].

Similar to the multisample sphericity test, under this distributed setup we can define the hypothesis $\mathcal{H}_0$ for sphericity given in (11). Based on the measurements over the N time slots, each secondary user terminal calculates the extreme value probabilities, commonly known as p-values. The p-value at the secondary user terminal m can be written as

$$p_m = \Pr[\text{TS}_m \leq \overline{\text{TS}}_m | \mathcal{H}_0] \tag{19}$$

where

$$\text{TS}_m = \frac{|\mathbf{R}_m|}{\left(\frac{1}{Q_m}\text{tr}(\mathbf{R}_m)\right)^{Q_m}}, \tag{20}$$

is the random variable representing the local GLRT at the secondary user terminal m and $\overline{\text{TS}}_m$ is the current value of this test statistic. It is important to note that the MSTS expression in (12) reduces to TS_m when $M = 1$. As such, we can follow the same steps as in the derivation of P_{fa} in (18) to find an analytical approximation for p_m as

$$p_m \approx 1 - \Pr[\chi^2_{f_m} \leq -2\rho_m \ln(\overline{\text{TS}}_m)], \tag{21}$$

where $f_m = Q_m^2 - 1$, $\rho_m = N - \frac{1+Q_m^2-2Q_m^4}{6Q_m(1-Q_m^2)}$.

Having computed the p-value , each secondary user terminal sends it to the central processor. Thus, instead of sending the whole matrix $\mathbf{X_m}$, meta analysis requires each secondary user terminal to send a single real number. The central processor collects all the p-values from the M secondary user terminals and combines them according to Fisher's method to produce a global test statistic [12]

$$\text{MATS} = -2\sum_{m=1}^{M} \ln(p_m). \tag{22}$$

We determine $\mathcal{H}_0$ or $\mathcal{H}_1$ in the meta analysis test as

$$\text{MATS} \lessgtr^{\mathcal{H}_0}_{\mathcal{H}_1} \zeta, \tag{23}$$

where we reject the null hypothesis if MATS is larger than the threshold ζ and reject the alternative hypothesis if MATS is smaller than ζ. In order to analyse the performance of the meta analysis sphericity test, we now proceed to derive analytical expressions for P_{fa} based on (23).

A. Probability of False Alarm

While the CDF of the test statistic in (22) is commonly available in the meta analysis literature [12], for the sake of completeness we summarize the derivation of P_{fa} in the following. As the test statistic MATS is a sum of the logarithms of p-values, in order to derive the CDF of MATS we first focus on the CDF of $-\ln(p_m)$. Following [12] we use the inverse transform method and write

$$\overline{\text{TS}}_m = F^{-1}_{\text{TS}_{m,\mathcal{H}_0}}(u), \tag{24}$$

where $F_{\text{TS}_{m,\mathcal{H}_0}}$ is the CDF of TS_m under the hypothesis $\mathcal{H}_0$ and u is uniformly distributed between 0 and 1. Based on (24) we can express the p-value in (19) as $p_m = F_{\text{TS}_{m,\mathcal{H}_0}}(F^{-1}_{\text{TS}_{m,\mathcal{H}_0}}(u)) = u$. Hence, p_m is also uniformly distributed between 0 and 1 and the CDF of $-\ln(p_m)$ has an exponential distribution. As the sum of M independent and exponentially distributed random variables has a standard chi-squared distribution with $2M$ degrees of freedom, we derive the CDF of MATS as

$$\Pr[\text{MATS} \leq x] = \Pr[\chi^2_{2M} \leq x]. \tag{25}$$

Thus, the exact closed-form expression for P_{fa} in meta analysis can be derived as

$$P_{\text{fa}} = 1 - \Pr[\chi^2_{2M} \leq \zeta]. \tag{26}$$

From (26) we learn that P_{fa} for meta analysis only depends on M and ζ, i.e., unlike in (18) for the multisample sphericity test, it is not a function of the sensing duration N nor the number of secondary user antennas Q_m. In the following, we use numerical examples to illustrate the accuracy of (26).

While in this paper we consider only the GLRT statistic to calculate the p-values, it is important to note that meta analysis can be implemented with secondary user terminals operating with different test statistics. For example, the approach can be used when one secondary user terminal uses GLRT statistic and another uses a simple power detector.

B. Numerical Examples

Fig. 4 plots the meta analysis P_{fa} versus the detection threshold ζ for different numbers of secondary user terminals, specifically $M = 3, 4$ and 5. In each case the secondary user terminals are allocated unequal number of antennas as noted in the figure. The observation window is $N = 200$. We note that the parameters Q_m and N do not effect the P_{fa} in meta analysis. This is because under the null hypothesis the p-values have a uniform distribution that varies between zero and one. The distribution of these tail probabilities do not change with Q_m and N. Once again, the results obtained using the analytical expression in (26) agree closely with the Monte-Carlo simulations. Note that while the probability of false alarm expression is exact, the p-values used to compute MATS use the chi-squared approximation in (21), which is the cause of the slight discrepancy between the simulated and analytical results.

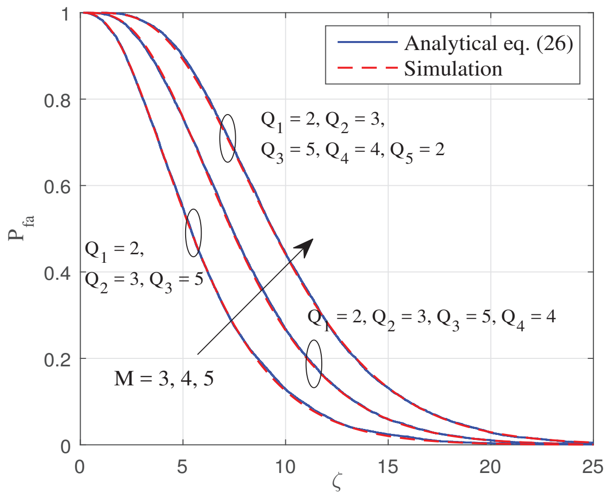

Fig. 4. Probability of false alarm vs the detection threshold when M = 3, 4 and 5.

V. Receiver operating characteristics

In this section we present a comparison of the above cooperative spectrum sensing techniques. In addition to the multisample sphericity test and meta analysis we also consider a basic binary fusion technique which is described in the following.

Binary Fusion: In the case of binary fusion each secondary user terminal performs a GLRT based on its own measurements and makes a binary decision about the presence of primary users. Based on (6), the individual test statistic of secondary user terminal m as given in (20), which is then used to declare either $\mathcal{H}_0$ or $\mathcal{H}_1$ using (7). Each secondary user terminal then sends its binary decisions to the central processor. As such, instead of sending the whole matrix $\mathbf{X}_m$, in binary fusion the secondary user terminal m sends only one bit of information. The central processor collects all M decisions and performs a binary AND or OR operation to form the final decision.

If the binary AND operation is used, the central processor declares $\mathcal{H}_0$ only if all M secondary user terminals have declared $\mathcal{H}_0$. Assuming independence of test statistics formed by each secondary user terminal, the overall P_{fa} is given by

$$P_{\text{fa}} = \prod_{m=1}^{M} \Pr[\text{TS}_{m\mathcal{H}_0} < \zeta], \tag{27}$$

where $\text{TS}_{m\mathcal{H}_0}$ is TS_m defined under hypothesis $\mathcal{H}_0$. Based on (18) we can derive an analytical approximation for P_{fa} in (27) as

$$P_{\text{fa}} \approx \prod_{m=1}^{M} [1 - \Pr[\chi^2_{f_m} \leq -2\rho_m \ln(\zeta)]], \tag{28}$$

where $f_m = Q_m^2 - 1$, $\rho_m = N - \frac{1+Q_m^2-2Q_m^4}{6Q_m(1-Q_m^2)}$. Similarly, if the binary OR operation is used, the central processor declares $\mathcal{H}_0$ if at least one secondary user terminal has declared $\mathcal{H}_0$. As such, we can derive P_{fa} as

$$P_{\text{fa}} = 1 - \prod_{m=1}^{M} [1 - \Pr[\text{TS}_{m\mathcal{H}_0} < \zeta]] \,. \tag{29}$$

Based on (18) we can derive an analytical approximation for P_{fa} in (29) as

$$P_{\text{fa}} \approx 1 - \prod_{m=1}^{M} \Pr[\chi^2_{f_m} \leq -2\rho_m \ln(\zeta)]. \tag{30}$$

In Fig. 5 we present a ROC curve where we plot P_d versus the P_{fa} for the multisample sphericity test, meta analysis and binary fusion. While P_{fa} for the multisample sphericity test, meta analysis, binary AND fusion and binary OR fusion is generated based on our analytical results in (18), (25), (28) and (30), respectively, the corresponding P_d results are generated numerically using Monte-Carlo simulations. We fix $M = 4$, $Q = 2$, $P = 3$ and $N = 200$. The noise variance $\sigma^2 = 1$ and $\gamma_1 = \gamma_2 = \gamma_3 = -10$ dB.

As expected binary OR and binary AND fusion perform worse than the other two, with binary OR fusion considerably outperforming binary AND fusion. Importantly, we observe a very small performance difference between that of multisample sphericity test and meta analysis. This represents a promising result, as the meta analysis approach halves the feedback overhead of the multisample approach (one real number per secondary terminal versus two real numbers). In the tail regime, the performance of all four techniques seems to coincide in the linear scale presented in Fig. 5. However, in the log scale we could clearly observe that multisample sphericity test and meta analysis outperforms binary fusion. It is also important to note that the plots in Fig. 5 were generated for one random instance of channel values. The performance of all four techniques vary when we change the channel values. However, we observe the same pattern in all channel instances.

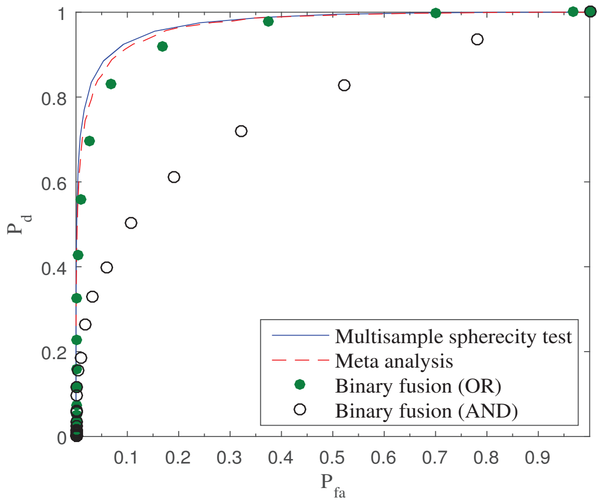

Fig. 5. The detection probability vs the probability of false alarm when M = 4, Q = 2, P = 3 and N = 200.

VI. Conclusion

Based on the sphericity test and the GLRT statistic, two new distributed spectrum sensing techniques are presented for a cognitive radio network with multiple primary and secondary user terminals. In these novel techniques only partial information, instead of raw data, needs to be sent from the secondary user terminals to the central processor. More specifically, in the multisample sphericity test each secondary user terminal sends only two real numbers, and in meta analysis they send only one real number to the central processor. This allows the application of these distributed spectrum sensing techniques to larger cognitive radio networks. We derive rigorous analytical expressions to measure the false alarm probability of each technique and use numerical examples illustrate the accuracy of the derived results. Based on ROC curves we also compare the performance of these novel techniques with each other and with other fusion techniques such as binary fusion.

References

[1] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 201–220, Feb. 2005.

[2] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, “A survey on spectrum management in cognitive radio networks,” *IEEE Commun. Mag.*, vol. 46, pp. 40–48, Jan. 2008.

[3] S. Kritchman and B. Nadlerg, “Non-parametric detections of the number of signals: hypothesis testing and random matrix theory,” *IEEE Trans. Signal Process.*, vol. 57, pp. 3930–3941, Oct. 2009.

[4] A. Taherpour, M. N. Kenari, and S. Gazor, “Multiple antenna spectrum sensing in cognitive radios,” *IEEE Trans. Wireless Commun.*, vol. 9, pp. 814–823, Feb. 2010.

[5] R. Zhang, T. J. Lim, Y. C. Liang, and Y. Zeng, “Multi-antenna based spectrum sensing for cognitive radios: a GLRT approach,” *IEEE Trans. Commun.*, vol. 58, pp. 84–88, Jan. 2010.

[6] L. Wei and O. Tirkkonen, “Spectrum sensing in the presence of multiple primary users,” *IEEE Trans. Commun.*, vol. 60, pp. 1268–1277, May 2012.

[7] R. H. Y. Louie, M. R. McKay, and Y. Chen, “Multiple-antenna signal detection in cognitive radio networks with multiple primary user signals,” in *IEEE Int. Conf. Commun. (ICC) 2014*, Sydney, Australia, Jun. 2014, pp. 4951–4956.

[8] D. M. Jimenez, R. H. Y. Louie, M. R. McKay, and Y. Chen, “Analysis and design of multiple-antenna cognitive radios with multiple primary user signals,” *IEEE Trans. Signal Process.*, vol. 63, pp. 4925–4939, Sep. 2015.

[9] J. W. Mauchly, “Significance test for sphericity of a normal n-variate distribution,” *Annals of Mathematical Statistics*, vol. 11, pp. 204–209, Jan. 1940.

[10] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. Wiley, 2003.

[11] R. J. Murihead, *Aspects of multivariate statistical theory*. Hoboken, New Jersey: John Wiley and Sons, Inc, 1982.

[12] R. A. Fisher, *Statistical Methods for Research Workers*, 5th ed. Tweeddale Court, Edinburgh: Oliver and Boyd, 1925.