Locally Best Invariant Test for Multiple Primary User Spectrum Sensing

Lu Wei, Prathapasinghe Dharmawansa and Olav Tirkkonen
Department of Communications and Networking,
Aalto University
P. O. Box 13000, Aalto-00076, Finland
Email: {lu.wei, olav.tirkkonen}@aalto.fi, prathapakd@ieee.org

Abstract—We consider multi-antenna cooperative spectrum sensing in cognitive radio networks, when there may be multiple primary users. A noise-uncertainty-free detector that is optimal in the low signal to noise ratio regime is analyzed in such a scenario. Specifically, we derive the exact moments of the test statistics involved, which lead to simple and accurate analytical formulae for the false alarm probability and the decision threshold. Simulations are provided to examine the accuracy of the derived results, and to compare with other detectors in realistic sensing scenarios. From the considered simulation settings, performance gain over several known detection algorithms is observed in scenarios with relatively low signal to noise ratio.

Index Terms—Cognitive radio; multi-antenna spectrum sensing; multiple primary users; locally best invariant test.

I. INTRODUCTION

In Cognitive Radio (CR) networks, dynamic spectrum access is implemented to mitigate spectrum scarcity. Namely, a secondary (unlicensed) user is allowed to utilize the spectrum resources when it does not cause intolerable interference to the primary (licensed) users. Spectrum sensing is the first key step towards this dynamic spectrum access scenario.

Prior work on cooperative spectrum sensing predominately employ the assumption of a single active primary user. Based on this assumption, several eigenvalue based sensing algorithms have been proposed recently [1–8]. These algorithms are non-parametric, i.e. they do not require information of the primary user, in contrast to e.g. feature detection. Moreover, they achieve optimality under different assumptions on the knowledge of the parameters. The assumption of a single primary user is made as the investigations in the literature have mainly focussed on CR networks, where the primary users are TV or DVB systems. In these systems the single active primary user assumption is, to some extent, justifiable. In addition, assuming a single primary user leads to analytically tractable problems.

The single primary user assumption may fail to reflect the situation in forthcoming CR networks, where the primary system could be a cellular network, and the existence of more than one primary user would be the prevailing condition. Using existing single primary user detection algorithms in such a scenario will induce performance loss. Despite the need to understand multiple primary user detection, the results in this direction are rather limited. A heuristic detection algorithm based on the ratio of the extreme eigenvalues is investigated in [9, 10], but its detection performance turns out to be sub-optimal [6]. An optimal detection algorithm in the presence of multiple primary users, based on the spherical test, has been proposed in [11] and subsequently studied in [12]. However, numerical evidence suggests that this detector does not perform particularly well when Signal-to-Noise Ratio (SNR) is relatively low [12]. Spectrum sensing in the low SNR regime is a practical and challenging issue in cooperative spectrum sensing. For example, recent FCC regulations require that the secondary devices must be able to detect signals with SNR as low as -18 dB [13]. To address this challenge, in this paper we consider a multiple primary user detector that is optimal in the low SNR regime. In particular, we investigate its detection performance by deriving a closed-form moment expression of the test statistics. Using the derived moments, approximations to the false alarm probability and the decision threshold are constructed. The derived approximations are easily computable and simulations show that they are accurate for the considered sensor sizes and number of samples.

The rest of this paper is organized as follows. In Section II we propose the optimal low SNR detector for multiple primary user sensing after outlining the signal model. Performance analysis of the proposed detection algorithm is addressed in Section III. Section IV presents numerical examples to verify the derived results and to study the detection performance in diverse scenarios. Finally in Section V we conclude the main results of this paper.

II. PROBLEM FORMULATION

A. Signal Model

Consider the standard model for K-sensor cooperative detection in the presence of P primary users,

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \sigma\mathbf{n} \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^K$ is the received data vector. The K sensors may be e.g. K receive antennas in one secondary terminal or K secondary devices each with a single antenna, or any combination of these.[1] The $K \times P$ matrix $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_P]$

---

[1]This collaborative sensing scenario is more relevant when the K sensors are in one device, since for multiple collaborating devices, accurate time synchronization between devices are needed and communications to the fusion center becomes an issue.

represents the channels between the P primary users and the K sensors. The $P \times 1$ vector $\mathbf{s} = [s_1, \ldots, s_P]'$ denotes zero mean transmitted signals from the primary users. The $K \times 1$ vector $\sigma \mathbf{n}$ is the complex Gaussian noise with zero mean and covariance matrix $\sigma^2 \mathbf{I}_K$, where the scalar σ^2 is the noise power.

We collect N i.i.d observations from model (1) to a $K \times N$ matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$. The problem of interest is to use the data matrix $\mathbf{X}$ to decide whether there are primary users. For ease of analysis we make the following assumptions

1) The channel $\mathbf{H}$ is constant during sensing time.
2) The primary user's signal follows an i.i.d zero mean Gaussian distribution and is uncorrelated with the noise.

In the absence of primary users, the sample covariance matrix $\mathbf{R} = \mathbf{X}\mathbf{X}^\dagger$ follows an uncorrelated (white) complex Wishart distribution $\mathcal{W}_K(N, \boldsymbol{\Sigma})$ with population covariance matrix

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

In the presence of primary users, by assumptions 1) and 2), the sample covariance matrix $\mathbf{R}$ follows a correlated complex Wishart distribution. The correlation is induced by the presence of the signals, and it is characterized by the population covariance matrix

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

where $\gamma_i := \mathbb{E}[s_i s_i^\dagger]$ defines the transmission power of the i-th primary user. The received SNR of primary user i across the K sensors is $\mathrm{SNR}_i := \gamma_i \|\mathbf{h}_i\|^2 / \sigma^2$. Finally, we denote the ordered eigenvalues of the sample covariance matrix $\mathbf{R}$ by $0 \leq \lambda_K \leq \ldots \leq \lambda_1 < \infty$.

B. Test Statistics

The differences between the population covariance matrices (2) and (3) can be explored to detect the primary users. This detection problem can be formulated as a binary hypothesis test, where hypothesis $\mathcal{H}_0$ denotes the absence of primary users and hypothesis $\mathcal{H}_1$ denotes the presence of primary users. Declaring wrongly $\mathcal{H}_0$, or declaring correctly $\mathcal{H}_1$, defines the false alarm probability P_{fa}, and the detection probability P_{d}, respectively. Since the sample covariance matrix $\mathbf{R}$ is a Wishart matrix, it is a sufficient statistics for the population covariance matrix $\boldsymbol{\Sigma}$ [14]. This leads to various test statistics as functions of $\mathbf{R}$ with different assumptions on the number of primary users P, and the knowledge of the noise power σ^2.

In the case of *a single* primary user, $P = 1$, the hypothesis test can be expressed as

$$\mathcal{H}_0 \quad : \quad \boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_K \tag{4}$$
$$\mathcal{H}_1 \quad : \quad \boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_K + \gamma_1 \mathbf{h}_1 \mathbf{h}_1^\dagger. \tag{5}$$

Assuming known noise power σ^2, the Largest Eigenvalue based (LE) detection $T_{\mathrm{LE}} := \lambda_1$ is shown to be optimal under the Generalized Likelihood Ratio Test (GLRT) criterion [2]. Here we concentrate on unknown noise power. Then the

optimal detector in the GLRT sense is the Scaled Largest Eigenvalue based (SLE) detection $T_{\mathrm{SLE}} := \lambda_1 / \sum_{i=1}^{K} \lambda_i$. The SLE detector is first proposed in the context of spectrum sensing in [4] and analyzed in [5–8].

It can be verified that the matrix $\sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^\dagger$ in (3) is positive definite, i.e. $\sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^\dagger \succ \mathbf{0}$. Considering this fact, in the presence of possibly *multiple* primary users, when $P \geq 1$ but not known a priori, the hypothesis test is expressed as

$$\mathcal{H}_0 \quad : \quad \boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_K \tag{6}$$
$$\mathcal{H}_1 \quad : \quad \boldsymbol{\Sigma} \succ \sigma^2 \mathbf{I}_K, \tag{7}$$

where the noise power σ^2 is assumed to be unknown. Essentially, we are testing a null hypothesis $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_K$ against all the other possible alternatives of $\boldsymbol{\Sigma}$, i.e. the hypothesis test is blind to P. In the statistics literature, this hypothesis test is known as the sphericity test. The corresponding optimal detector under the GLRT criterion is the so-called Spherical Test based detection $T_{\mathrm{ST}} := \prod_{i=1}^{K} \lambda_i / \left(\sum_{i=1}^{K} \lambda_i / K \right)^K$. In the context of spectrum sensing, the ST detection is first proposed in [11] and analyzed in [12]. Although in general the ST detector achieves good performance, it is not the best one in the low SNR regime. A test statistics that is optimal in detecting small deviations from $\mathcal{H}_0$ is the so-called John's detection

$$T_{\mathrm{J}} := \frac{\mathrm{tr}(\mathbf{R}^2)}{(\mathrm{tr}(\mathbf{R}))^2} = \frac{\sum_{i=1}^{K} \lambda_i^2}{\left(\sum_{i=1}^{K} \lambda_i \right)^2}, \tag{8}$$

which is first considered by S. John [15]. A more rigorous derivation of the test statistics (8) can be found in [16], where the resulting test procedure is

$$T_{\mathrm{J}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \zeta, \tag{9}$$

ζ being a threshold. The optimality property of the T_{J} detector in detecting small derivations is known as the locally best invariant property. Mathematically, it means that for every σ^2 and for every other test T (say), there is a neighborhood of $\sigma^2 \mathbf{I}_K$ such that T_{J} achieves no worse performance than T does [16].

Besides the ST and John's detectors, another detector in the presence of multiple primary users is the Eigenvalue Ratio based (ER) detection $T_{\mathrm{ER}} = \lambda_1 / \lambda_K$ [9, 10]. The ER detector is not constructed from any optimality considerations, thus its performance is substantially worse than those of the ST and John's detectors [6, 12]. Finally, we note that no eigenvalue decomposition is needed in implementing John's detector as opposed to most of other eigenvalue based detectors.

III. PERFORMANCE ANALYSIS

In this section we first derive an exact expression for the moment of T_{J}. Based on this result, we construct an approximation to the distribution of T_{J}, which leads to closed-form formulae for the false alarm probability and the decision threshold.

A. Exact Moment Expression

Under $\mathcal{H}_0$, by following the similar argument for the real Wishart case [17] it can be easily verified that the random variable $\left(\sum_{i=1}^{K} \lambda_i\right)^2$ is independent of the random variable of interest

$$T_{\mathrm{J}} = \frac{\sum_{i=1}^{K} \lambda_i^2}{\left(\sum_{i=1}^{K} \lambda_i\right)^2} \in [1/K, 1]. \qquad (10)$$

By this independence, the m-th moment of $\sum_{i=1}^{K} \lambda_i^2$ equals

$$\mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i^2\right)^m\right] = \mathbb{E}[T_{\mathrm{J}}^m]\mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i\right)^{2m}\right], \qquad (11)$$

and thus

$$\mathbb{E}[T_{\mathrm{J}}^m] = \mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i^2\right)^m\right] \Bigg/ \mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i\right)^{2m}\right]. \qquad (12)$$

The random variable $2\operatorname{tr}(\mathbf{R}) = 2\sum_{i=1}^{K} \lambda_i$ follows a Chi-square distribution with $2KN$ degrees of freedom. By using the moment expression for Chi-square distribution [18] (Eq. (2.35)), the $2m$-th moment of $\sum_{i=1}^{K} \lambda_i$ is obtained as

$$\mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i\right)^{2m}\right] = \frac{\Gamma(2m + KN)}{\Gamma(KN)}. \qquad (13)$$

The next step is to calculate the moment of $\sum_{i=1}^{K} \lambda_i^2$, which is given by the following result.

Proposition 1. *The m-th non-negative integer moment of the random variable $\sum_{i=1}^{K} \lambda_i^2$ equals*

$$\mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i^2\right)^m\right] = \sum_{a_1 + \cdots + a_K = m} \frac{m! C}{a_1! \cdots a_K!} \times$$
$$\prod_{1 \le i < j \le K} (2a_j - 2a_i + j - i) \prod_{i=1}^{K} \Gamma(2a_i + N - K + i), \quad (14)$$

where the sum is over all the non-negative integer solutions of $a_1 + \cdots + a_K = m$ and the constant $C = \left(\prod_{i=1}^{K} \Gamma(N - i + 1)\Gamma(K - i + 1)\right)^{-1}$.

The proof of Proposition 1 is in Appendix A. Inserting (14) and (13) into (12), the m-th moment of random variable T_{J}, denoted by $\mathcal{M}_m$, equals

$$\mathcal{M}_m := \frac{C \Gamma(KN)}{\Gamma(2m + KN)} \sum_{a_1 + \cdots + a_K = m} \frac{m!}{a_1! \cdots a_K!} \times$$
$$\prod_{1 \le i < j \le K} (2a_j - 2a_i + j - i) \prod_{i=1}^{K} \Gamma(2a_i + N - K + i). \quad (15)$$

The sum over the partition $a_1 + \cdots + a_K = m$ can be implemented by normal sums as $\sum_{a_1=0}^{m} \sum_{a_2=0}^{m-a_1} \cdots \sum_{a_{K-1}=0}^{m-a_1-\cdots-a_{K-2}}$ with a_K replaced by $m - \sum_{i=1}^{K-1} a_i$ in the summand. Note that for real Wishart matrix, up to the second, fourth and sixth moment of T_{J}

under $\mathcal{H}_0$ can be found in [17], [21] and [22], respectively. To the best of our knowledge, the derived moment expression of T_{J} (15) for complex Wishart matrix, which is valid for arbitrary non-negative moment, is new.

B. Moment Based Approximation

It is a standard technique in statistics to approximate some unknown distribution by a known one having the same support and moments. Motivated by the results for real Wishart case [22], in this work we choose a generalized Beta distribution with the same support as that of T_{J} to approximate the distribution of T_{J}. Specifically, the linear transform $x = \frac{(K-1)z+1}{K}$ on a standard Beta density[2] $z^{\alpha-1}(1-z)^{\beta-1}/B(\alpha,\beta)$, $z \in [0,1]$ leads to a generalized Beta density

$$C'\left(x - \frac{1}{K}\right)^{\alpha-1}(1-x)^{\beta-1}, \qquad (16)$$

with the support $x \in [1/K, 1]$ and the constant $C' = K^{\alpha+\beta-1}/\left(B(\alpha,\beta)(K-1)^{\alpha+\beta-1}\right)$. Since the m-th moment of a standard Beta random variable equals $\mathbb{E}[z^m] = (\alpha)_m/(\alpha+\beta)_m$, where $(\alpha)_m = \Gamma(\alpha+m)/\Gamma(\alpha)$ defines the Pochhammer symbol, the m-th moment of the generalized Beta random variable is obtained by binomial expansion as

$$\mathbb{E}[x^m] = \mathbb{E}\left[\left(\frac{(K-1)z+1}{K}\right)^m\right] \qquad (17)$$
$$= \frac{1}{K^m} \sum_{i=0}^{m} \binom{m}{i} (K-1)^i \mathbb{E}[z^i] \qquad (18)$$
$$= \frac{1}{K^m} \sum_{i=0}^{m} \binom{m}{i} \frac{(K-1)^i (\alpha)_i}{(\alpha+\beta)_i}, \qquad (19)$$

where $\binom{m}{i}$ denotes the binomial coefficient. In particular, the first two moments are

$$\frac{\alpha K + \beta}{(\alpha+\beta)K}, \quad \frac{(\alpha K+\beta)^2 + \alpha K^2 + \beta}{(\alpha+\beta)(\alpha+\beta+1)K^2}, \qquad (20)$$

by matching them to the first two moments of T_{J} (15), the parameters α and β of the generalized Beta density (16) can be obtained

$$\alpha = \frac{(K\mathcal{M}_1 - 1)(K\mathcal{M}_1 - K\mathcal{M}_2 + \mathcal{M}_1 - 1)}{(K-1)K(\mathcal{M}_2 - \mathcal{M}_1^2)}, \qquad (21)$$
$$\beta = \frac{(\mathcal{M}_1 - 1)(K\mathcal{M}_1 - K\mathcal{M}_2 + \mathcal{M}_1 - 1)}{(K-1)(\mathcal{M}_1^2 - \mathcal{M}_2)}. \qquad (22)$$

As a result, the two-moment-based approximation to the CDF of T_{J} under $\mathcal{H}_0$ is

$$F_{\mathrm{J}}(y) \approx C' \int_{1/K}^{y} \left(x - \frac{1}{K}\right)^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x \qquad (23)$$
$$= 1 - \frac{B\left(\frac{K(1-y)}{K-1}; \beta, \alpha\right)}{B(\alpha,\beta)}, \qquad (24)$$

where $y \in [1/K, \infty)$ and $B(x; a, b) = \int_0^x z^{a-1}(1-z)^{b-1}\mathrm{d}z$ is the lower incomplete Beta function.

---

[2] $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ defines the Beta function.

By (9), for a given threshold ζ the two-moment-based approximation to the false alarm probability equals

$$P_{\text{fa}}(\zeta) = 1 - F_{\text{J}}(\zeta) \approx \frac{B\left(\frac{K(1-\zeta)}{K-1}; \beta, \alpha\right)}{B(\alpha, \beta)}. \qquad (25)$$

To implement the proposed spectrum sensing algorithm, a decision threshold needs to be determined for a given detection requirement in a computationally affordable manner. Using the derived approximation to P_{fa}, an approximative decision threshold can be obtained by numerically inverting (25). Note that both α and β are elementary functions of the sensor size K and sample size N through (14). Moreover, if we further approximate α and β to their respective nearest integers, (25) reduces to a finite sum of polynomials in ζ. Thus the computational complexity of threshold calculation becomes quite affordable for on-line implementations.

Here we note that the proposed two-moment-based Beta approximation corresponds to the simplest form of a general Jacobi polynomials based approximation. In the general framework, up to any n-th degree of Jacobi polynomials matching the corresponding n moments of T_{J} would be used. Since the random variable T_{J} is of a finite support, the Jacobi polynomials expansion for the distribution of T_{J} is exact according to Weierstrass approximation theorem [23]. Namely, when n goes to infinity the Jacobi polynomials based approximation represents the exact distribution of T_{J}. In practise, the choice of n reflects a trade-off between the approximation accuracy and the implementation complexity. In light of the good accuracy as shown in the next section, we consider $n = 2$ in this paper. The general n-moment-based approximation, including the error analysis, can be easily obtained by following the procedures in [22, 24].

IV. NUMERICAL RESULTS

In this section we first investigate the accuracy of the derived approximative false alarm probability by simulations. Then we compare the performance of John's detector with those of several detectors in realistic scenarios. The considered values of the parameters K and N in this section reflect practical spectrum sensing situations. The sample size N can be as large as a couple of hundred whereas the number of sensors K is at most eight due to physical constraints of the device size.

A. Accuracy of the Approximative False Alarm Probability

In Figure 1 we plot the approximative P_{fa} using (25) and the simulated P_{fa} as a function of the threshold. To quantitatively show the approximation accuracy, we also calculate the average approximation error[3] of the proposed P_{fa} with respect to the exact P_{fa} as resulting from simulations. The results, summarized in the caption of Figure 1, show that the derived analytical P_{fa} matches the simulations well.

[3]Defined as $\left(\sum_{i=1}^{n} |P_{\text{fa}}(\zeta_i) - \tilde{P}_{\text{fa}}(\zeta_i)|\right)/n$, where $\tilde{P}_{\text{fa}}$ denotes the approximative false alarm probability, i.e. the RHS of (25), and n is the sampling size. In Figure 1 we assume uniform sampling in $\zeta \in [0.125, 0.3]$ with $n = 10^7$.
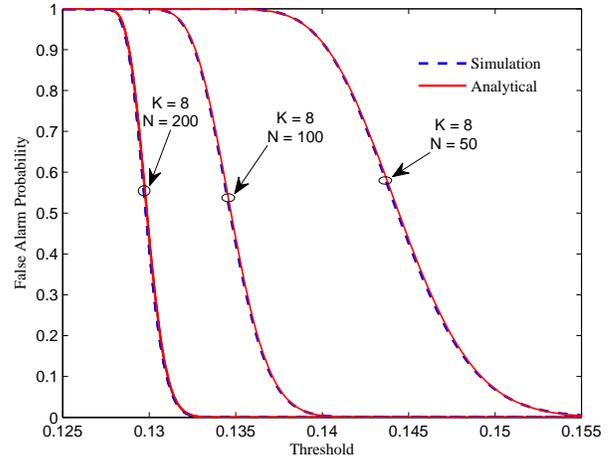
Fig. 1. False alarm probability: analytical versus simulations. For $K = 8$, $N = 50$, 100 and 200, the average approximation error on false alarm probability is respectively 5.94×10^{-8}, 5.03×10^{-8} and 5.07×10^{-8}.
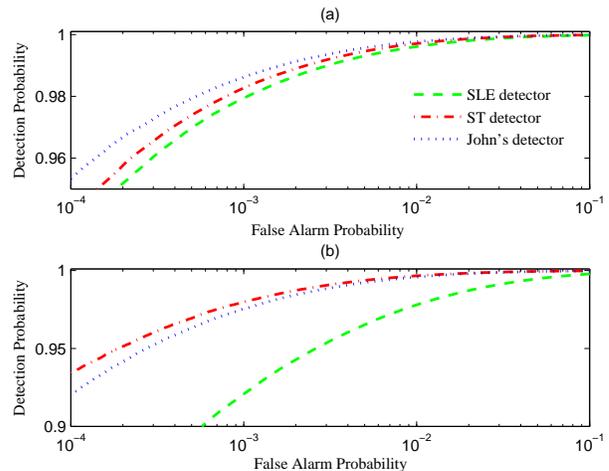
Fig. 2. ROCs for $P = 3$. Subplot (a): $\text{SNR}_1 = -6$ dB, $\text{SNR}_2 = -5$ dB, $\text{SNR}_3 = -4$ dB with $K = 4$, $N = 400$; subplot (b): $\text{SNR}_1 = 1$ dB, $\text{SNR}_2 = 2$ dB, $\text{SNR}_3 = 3$ dB with $K = 4$, $N = 50$.

B. Detection Performance

We compare the detection performance of John's detector with those of other known detectors by means of ROC curve. Since a ROC curve shows the achieved detection probability as a function of the false alarm probability, it reflects the overall detection performance for a given detector. Our focus here is detection in the presence of multiple primary users, thus we consider for comparison the ST detector. In addition the SLE detector, which is optimal for single primary user detection, is considered for comparison as well. Comparisons with the non-optimal ER detector and detectors that are sensitive to noise uncertainty [25], such as the LE detector and the energy detector [26], are excluded in this paper. For results in this direction, the readers are referred to [6, 7, 12].

In Figure 2 we consider a scenario of three simultane-

ously transmitting primary users with relatively low SNRs ($\text{SNR}_1 = -6$ dB, $\text{SNR}_2 = -5$ dB, $\text{SNR}_3 = -4$ dB using $K = 4$, $N = 400$) in subplot (a) and relatively high SNRs ($\text{SNR}_1 = 1$ dB, $\text{SNR}_2 = 2$ dB, $\text{SNR}_3 = 3$ dB using $K = 4$, $N = 50$) in subplot (b). Without loss of generality, we assume unit powers for the zero mean Gaussian signal and noise. The entries of the channel matrix $\mathbf{H}$, which are fixed during sensing, are independently drawn from a standard complex Gaussian distribution. The same channel matrix is used in both subplots. The channel vector for each primary user is normalized as $\mathbf{u}_i = \mathbf{h}_i / \|\mathbf{h}_i\|$. As a result, the population covariance matrix $\mathbf{\Sigma}$ now equals $\mathbf{\Sigma} = \mathbf{I}_K + \sum_{i=1}^{P} \text{SNR}_i \mathbf{u}_i \mathbf{u}_i^\dagger$. For the specific channel realizations considered in Figure 2, the eigenvalues[4] of $\mathbf{\Sigma}$ in subplot (a) are $[1.6225, 1.2217, 1.1213, 1]$ in subplot (b) are $[4.0417, 2.2375, 1.56, 1]$. From Figure 2 (a) we observe that John's detector achieves the best detection performance in the low SNR case considered. However, when the SNRs increase we see from Figure 2 (b) that the ST detector performs better than John's detector. In both subplots, it is seen that the ST and John's detectors outperform the SLE detector. This is intuitively clear since the SLE detector is optimized for single primary user detection. Moreover, we see that as the SNRs increase the performance gap between the SLE detector and the multiple primary user detectors becomes larger, as expected.

V. CONCLUSION

In this paper, we investigated the sensing performance of John's detector, which a candidate detector in the presence of multiple primary users. John's detector is optimal in detecting small deviations of the covariance matrix from a matrix proportional to identity. Analytical formulae have been derived for the false alarm probability and decision threshold of John's detector. The derived results are simple to calculate and yield an almost exact fit to simulations. From the simulation setting considered, performance gain over several detection algorithms is observed in the low SNR regime.

Characterizing T_J distribution under $\mathcal{H}_1$, which leads to analytical results for detection probability, is the work in progress.

APPENDIX A
THE NON-NEGATIVE INTEGER MOMENT OF $\sum_{i=1}^{K} \lambda_i^2$

Under $\mathcal{H}_0$, the joint density of the unordered eigenvalues $\lambda_i \in [0, \infty)$ for the sample covariance matrix $\mathbf{R}$ reads [19]

$$\Phi(\lambda_1, \ldots, \lambda_K) := \frac{C}{K!} |\Delta(\lambda)|^2 \prod_{i=1}^{K} \lambda_i^{N-K} \mathrm{e}^{-\lambda_i}, \quad (26)$$

[4]For the considered detectors, the test statistics depend on $\mathbf{\Sigma}$ only through the eigenvalues of $\mathbf{\Sigma}$.

where $|\Delta(\lambda)|$ defines the determinant of Vandermonde matrix with i,j-th entry $(\Delta(\lambda))_{i,j} = \lambda_i^{j-1}$, $i,j = 1, \ldots, K$ and the constant $C = \left(\prod_{i=1}^{K} \Gamma(N-i+1)\Gamma(K-i+1) \right)^{-1}$.

Before we prove Proposition 1, we need the following two lemmas.

Lemma 1. *The average value of the function $\prod_{i=1}^{K} \lambda_i^{2a_i}$, where a_is are non-negative integers, equals*

$$\int_{[0,\infty)^K} \left(\prod_{i=1}^{K} \lambda_i^{2a_i} \right) \Phi(\lambda_1, \ldots, \lambda_K) \mathrm{d}\lambda_1 \cdots \mathrm{d}\lambda_K =$$
$$\frac{C}{K!} \sum_{\nu} |\Gamma(2a_{\nu_i} + N - K + i + j - 1)|_{i,j=1,\ldots,K}, \quad (27)$$

where $\nu = \nu_1, \ldots, \nu_K$ defines a permutation of the integers $1, \ldots, K$ and the sum is over all the $K!$ permutations.

Proof: Using the fact that a product of determinants equals the determinant of the matrix product and the fact that the determinant remains unchanged under transpose operation, we have

$$|\Delta(\lambda)|^2 = |\Delta(\lambda)\Delta'(\lambda)| = \left| \sum_{l=1}^{K} \lambda_l^{i+j-2} \right|_{i,j=1,\ldots,K}. \quad (28)$$

By invoking the multi-linearity property of determinants, the Hankel determinant (28) above can be written as a sum of two determinants, where the first rows are $[1, \cdots, \lambda_1^{K-1}]$ and $[K-1, \cdots, \sum_{i=2}^{K-1} \lambda_i^{K-1}]$ with the respect second to the last rows remain unchanged. By repeated use of the multi-linearity property, (28) can be written as sum of K^K determinants, out of which $K!$ determinants with non-identical λ index in rows give non-zero contribution, namely,

$$\left| \sum_{l=1}^{K} \lambda_l^{i+j-2} \right|_{i,j=1,\ldots,K} = \sum_{\nu} \left| \lambda_{\nu_i}^{i+j-2} \right|_{i,j=1,\ldots,K}, \quad (29)$$

where $\nu = \nu_1, \ldots, \nu_K$ defines a permutation of the integers $1, \ldots, K$ and the sum is over all the $K!$ permutations. Inserting (29) into LHS of (27) and disregarding the constant $C/K!$ which will be re-installed, we have

$$\int_{[0,\infty)^K} \sum_{\nu} \left| \lambda_{\nu_i}^{i+j-2} \right|_{i,j=1,\ldots,K} \prod_{i=1}^{K} \lambda_i^{2a_i + N - K} \mathrm{e}^{-\lambda_i} \mathrm{d}\lambda_i$$
$$= \sum_{\nu} \int_{[0,\infty)^K} \left| \lambda_{\nu_i}^{2a_{\nu_i} + N - K + i + j - 2} \mathrm{e}^{-\lambda_{\nu_i}} \right|_{i,j=1,\ldots,K} \prod_{i=1}^{K} \mathrm{d}\lambda_i$$
$$= \sum_{\nu} |\Gamma(2a_{\nu_i} + N - K + i + j - 1)|_{i,j=1,\ldots,K},$$

where in the first equality we multiplied each $\lambda_i^{2a_i + N - K} \mathrm{e}^{-\lambda_i}$ with the row of $\left| \lambda_{\nu_i}^{i+j-2} \right|_{i,j=1,\ldots,K}$ having the same λ index and the second equality is achieved by first expanding the determinant using Leibniz formula, integrating term-wise and rewriting the integration results as a determinant. This completes the proof. ∎

Note that Lemma 1 can be considered as an extension to the Selberg type integral considered in [20] (Eq. (17.6.5) and Eq. (17.8.1)).

Lemma 2. *The following determinant can be simplified to*

$$|\Gamma(b_i + j - 1)|_{i,j=1,\ldots,K} = \prod_{1 \le i < j \le K} (b_j - b_i) \prod_{i=1}^{K} \Gamma(b_i), \quad (30)$$

where b_i is a positive integer.

Proof: We first realize that from each row the term $\Gamma(b_i)$ can be factored out, namely,

$$|\Gamma(b_i + j - 1)|_{i,j=1,\ldots,K} = |(b_i)_{j-1}|_{i,j=1,\ldots,K} \prod_{i=1}^{K} \Gamma(b_i), \quad (31)$$

where $(b)_j = \prod_{k=0}^{j-1}(b+k)$. By extracting from the i-th column a suitable linear combination of previous $i-1$ columns, the determinant

$$|(b_i)_{j-1}|_{i,j=1,\ldots,K} = \left| b_i^{j-1} \right|_{i,j=1,\ldots,K}, \quad (32)$$

which is a Vandermonde determinant $\prod_{1 \le i < j \le K}(b_j - b_i)$. This completes the proof. ∎

We are now in a position to prove Proposition 1. By using the multinomial expansion

$$\left(\sum_{i=1}^{K} \lambda_i^2 \right)^m = \sum_{a_1 + \cdots + a_K = m} \frac{m!}{a_1! \cdots a_K!} \prod_{i=1}^{K} \lambda_i^{2a_i} \quad (33)$$

and Lemma 1, we have

$$\mathbb{E}\left[\left(\sum_{i=1}^{K} \lambda_i^2 \right)^m \right] = \sum_{a_1 + \cdots + a_K = m} \frac{m!}{a_1! \cdots a_K!} \frac{C}{K!} \times$$
$$\sum_{\nu} |\Gamma(2a_{\nu_i} + N - K + i + j - 1)|_{i,j=1,\ldots,K}. \quad (34)$$

For any given permutation ν, it is observed that the sum over $a_1 + \cdots + a_K = m$ is symmetric in the sense that one can arbitrarily permute the index of a without changing the value of this sum, in particular the following permutation of a holds,

$$\sum_{a_1 + \cdots + a_K = m} \frac{m! |\Gamma(2a_{\nu_i} + N - K + i + j - 1)|_{i,j=1,\ldots,K}}{a_1! \cdots a_K!}$$
$$= \sum_{a_1 + \cdots + a_K = m} \frac{m! |\Gamma(2a_i + N - K + i + j - 1)|_{i,j=1,\ldots,K}}{a_1! \cdots a_K!}.$$

Since the number of possible permutations is $K!$, (34) now equals

$$\sum_{a_1 + \cdots + a_K = m} \frac{m! C |\Gamma(2a_i + N - K + i + j - 1)|_{i,j=1,\ldots,K}}{a_1! \cdots a_K!}. \quad (35)$$

Using Lemma 2 with $b_i = 2a_i + N - K + i$ completes the proof of Proposition 1.

REFERENCES

[1] Y. Zeng, C. L. Koh, and Y. C. Liang, "Maximum eigenvalue detection: Theory and application," in *Proc. IEEE Int. Conf. Commun.,* May 2008.

[2] A. Taherpour, M. N. Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.,* vol. 9, no. 2, pp. 814-823, Feb. 2010.

[3] L. Wei and O. Tirkkonen, "Cooperative spectrum sensing of OFDM signals using largest eigenvalue distributions," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.,* Sep. 2009.

[4] Y. Zeng, Y. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Process. Lett.,* vol. 15, pp. 649-652, 2008.

[5] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Trans. Veh. Technol.,* vol. 59, no. 4, pp. 1791-1800, May 2010.

[6] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Trans. Inf. Theory,* vol. 57, no. 4, pp. 2400-2419, Apr. 2011.

[7] B. Nadler, F. Penna, and R. Garello, "Performance of eigenvalue-based signal detectors with known and unknown noise power," in *Proc. IEEE Int. Conf. Commun.,* June 2011.

[8] L. Wei and O. Tirkkonen, "Analysis of scaled largest eigenvalue based detection for spectrum sensing," in *Proc. IEEE Int. Conf. Commun.,* June 2011.

[9] Y. Zeng and Y. C. Liang, "Eigenvalue based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.,* vol. 57, no. 6, pp. 1784-1793, June 2009.

[10] F. Penna, R. Garello, and M. A. Spirito, "Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in Wishart matrices," *IEEE Commun. Lett.,* vol. 13, issue 7, pp. 507-509, July 2009.

[11] 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, no. 1, pp. 84-88, Jan. 2010.

[12] L. Wei and O. Tirkkonen, "Spectrum sensing in the presence of multiple primary users," *IEEE Trans. Commun.,* to appear. Available: http://arxiv.org/abs/1201.0737

[13] FCC, "In the matter of unlicensed operation in the TV broadcast bands: second memorandum opinion and order, Federal Communications Commission," FCC 10-174, Sep. 2010.

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

[15] S. John, "Some optimal multivariate tests," *Biometrika,* vol. 58, no. 1, pp. 123-127, Apr. 1971.

[16] N. Sugiura, "Locally best invariant test for sphericity and the limiting distributions," *Annals of Mathematical Statistics,* vol. 43, no. 4, pp. 1312-1316, Aug. 1972.

[17] S. John, "The distribution of a statistic used for testing sphericity of normal distributions," *Biometrika,* vol. 59, no. 1, pp. 169-173, Apr. 1972.

[18] M. K. Simon, *Distributions Involving Gaussian Random Variables.* New York: Springer, 2002.

[19] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics,* 35, 475-501, 1964.

[20] M. L. Mehta, *Random Matrices.* 3rd Edition, Elsevier, 2004.

[21] S. John, "Fitting sampling distribution agreeing in support and moments and tables of critical values of sphericity criterion," *Journal of Multivariate Analysis,* 6, pp. 601-607, 1976.

[22] R. J. Boik, "Algorithm AS 284: Null distribution of a statistics for testing sphericity and additivity: a Jacobi polynomial expansion," *Journal of the Royal Statistical Society, series C (Applied Statistics),* vol. 42, no. 3, pp. 567-576, 1993.

[23] H. Hochstadt, *Special Functions of Mathematical Physics.* Holt, Rinehart and Winston, New York, 1961.

[24] H. T. Ha, *Advances in Moment-Based Density Approximation Methods.* Ph.D thesis, University of Western Ontario, 2006.

[25] R. Tandra and A. Sahai, "SNR walls for signal detetion," *IEEE J. Sel. Topics in Signal Process.,* vol. 2, no. 1, Feb. 2008.

[26] F. F. Digham, M. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," in *Proc. IEEE Int. Conf. Commun.,* May 2003.