

Article

Wideband Spectrum Sensing Based on Riemannian Distance for Cognitive Radio Networks

Qiuyuan Lu *, Shengzhi Yang and Fan Liu

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China; adrien_yang@bit.edu.cn (S.Y.); liufan92@bit.edu.cn (F.L.)

* Correspondence: luqiuyuan@bit.edu.cn; Tel.: +86-10-6891-1963

Academic Editor: Huseyin Arslan

Received: 4 January 2017; Accepted: 20 March 2017; Published: 23 March 2017

Abstract: Detecting the signals of the primary users in the wideband spectrum is a key issue for cognitive radio networks. In this paper, we consider the multi-antenna based signal detection in a wideband spectrum scenario where the noise statistical characteristics are assumed to be unknown. We reason that the covariance matrices of the spectrum subbands have structural constraints and that they describe a manifold in the signal space. Thus, we propose a novel signal detection algorithm based on Riemannian distance and Riemannian mean which is different from the traditional eigenvalue-based detector (EBD) derived with the generalized likelihood ratio criterion. Using the moment matching method, we obtain the closed expression of the decision threshold. From the considered simulation settings, it is shown that the proposed Riemannian distance detector (RDD) has a better performance than the traditional EBD in wideband spectrum sensing.

Keywords: cognitive radio; wideband spectrum sensing; information geometry; Riemannian distance; Riemannian mean; moment matching

1. Introduction

In today's increasingly crowded wireless spectrum environment, cognitive radio (CR) networks are considered a promising technology to mitigate the contradiction between fixed spectrum allocation and efficient utilization, which has received sustained attention in recent years [1,2]. In order to access the spectrum holes without interfering with the primary user (PU) in the network, the secondary user (SU) is required to perform accurate spectrum sensing.

As a key problem in CR networks, spectrum sensing technology has been extensively studied. Several classical spectrum sensing methods, such as energy detection (ED), matched filtering (MF) detection and cyclostationary detection have been proposed. Although ED is easy to implement, it is sensitive to noise uncertainty [3]. The MF method requires waveform information about the primary signal [4] and cyclostationary detection needs to know the cyclic frequencies of the primary signal [5]. Compared with the classical methods, eigenvalue-based spectrum sensing methods for multi-antenna systems require less prior information about noise and signal. Due to the correlation of the primary signal in multi-antenna reception, some spectrum sensing algorithms are designed according to the covariance matrix of the received data vectors, and the corresponding test statistics are constructed by applying the eigenvalues of the covariance matrix. To date, some typical detectors have been proposed in the literature, including the largest eigenvalue detector (LED) [6], maximum-minimum eigenvalue (MME) detector [7], the scaled largest eigenvalue (SLE) detector [8], and the arithmetic to geometric mean (AGM) detector [9]. In practical application scenarios, the CR system may suffer from unknown interference in the wideband spectrum and the imperfections of the sensor hardware [10–12]. Thus the independent and identically distributed (i.i.d.) noise assumption in

eigenvalue-based algorithms may not be realistic. Moreover, it is possible for a number of asynchronous independent PUs to access the same frequency band at the same time. Accordingly, the CR sensor needs to sense the wideband spectrum in the case of the unknown noise characteristics and unknown number of PUs.

Meanwhile, in CR networks, spectrum sensing for wideband is also an important topic worthy of study. In [13], the wideband spectrum sensing (WSS) problem is partitioned into four basic elements to study, namely, system modeling, performance metrics, sampling schemes and detection algorithms. In the context of this paper, we choose the detection probability of the primary signal in each subband of the wideband spectrum as the performance metric. We assume independent PU occupancy across subbands and that the detections of primary signals in each subband are independent with each other. For the above independent channel-by-channel detection mode [13], we can employ the more easily achieved partial band Nyquist sampling scheme to complete the wideband spectrum data acquisition.

In this paper, unlike traditional spectrum sensing algorithms based on a generalized likelihood ratio test (GLRT) [9,14], we consider the rethinking of WSS problem from the perspective of information geometry. As a promising cutting-edge discipline, information geometry studies the problems in the field of statistics and information science by applying modern differential geometry method on Riemannian manifolds. It has been widely used in machine learning, medical imaging, radar signal processing, signal classification and other research areas [15–20]. Inspired by these studies, we consider using the information geometry theory to design the spectrum sensing method, rather than the normed linear space theory. From the view of Lie group theory, the covariance matrix of sensing data is a Toeplitz Hermitian positive definite matrix, which forms the negative curvature space [21]. By constructing statistical models on the manifold, we can exploit the Riemannian distance for hypothesis testing, regardless of the specific characteristic assumptions about signal and noise. On the other hand, in multiband detection, there may be a plurality of vacant subbands available for estimating noise. Considering the geometric features of the matrix manifold, we use the Riemannian mean instead of the arithmetic mean for the joint estimation with multiple covariance matrices. To the best of the authors' knowledge, this is the first time the information geometry is used to solve the WSS problem.

The rest of the paper is organized as follows: in Section 2, we describe the system model in detail, introduce the key concepts of information geometry—Riemannian distance and Riemannian mean, and propose the Riemannian distance-based test statistic. Section 3 presents theoretical analysis and finds thresholds for the proposed detector using random matrix theory. In Section 4, we show the numerical results of the proposed algorithm. Finally, the main results of this paper are drawn in Section 5.

2. System Model and the Proposed Detection Scheme

2.1. System Model

Consider a CR system performing spectrum sensing on a spectrum of B Hz equipped with K antennas in each sensing node, as depicted in Figure 1. In the context of this article, we focus on the spectrum sensing of a single SU. It is worth noting, however, that in the cooperative spectrum sensing model, our methods are still applicable if the data of different nodes can be collected at the fusion center. It is assumed that the bandwidths of subbands are known to the CR receiver and the whole band is divided into L subbands. In the sensing time T_s , the receiver collects N samples in each subbands, the n -th observed data in the i -th subband is $\mathbf{x}_i[n]$, $n = 0, \dots, N-1$. Every subband sampling vector is composed of K -antenna data $\mathbf{x}_i[n] = [x_i^1[n], \dots, x_i^K[n]]^T$. Then, the spectrum sensing problem in each subband can be expressed by the following two hypotheses:

$$\begin{aligned} \mathcal{H}_{0,i} : \quad & \mathbf{x}_i[n] = \mathbf{w}_i[n], \quad n = 0, \dots, N-1 \\ \mathcal{H}_{1,i} : \quad & \mathbf{x}_i[n] = \mathbf{H}_i \mathbf{s}_i[n] + \mathbf{w}_i[n], \quad n = 0, \dots, N-1, \end{aligned} \quad (1)$$

where $\mathcal{H}_{0,i}$ (null hypothesis) stands for the absence of PUs and $\mathcal{H}_{1,i}$ for the presence of PUs. Here, $\mathbf{w}_i[n]$ is the n -th noise sample in the i -th subband over the K antennas. In our sensing scenario assumption, the number of concurrent PU transmissions is P which is unknown to the sensing node. The $K \times P$ matrix $\mathbf{H}_i = [\mathbf{h}_i^{(1)}, \dots, \mathbf{h}_i^{(P)}]$ represents the channels between the P PUs and the receiving antennas in the i -th subband. The $P \times 1$ vectors $\mathbf{s}_i[n] = [s_i^{(1)}[n], \dots, s_i^{(P)}[n]]^T$ denote the n -th samples of the transmitted signal from the PUs in the i -th subband. We assume that the transmitted samples $s_i^{(p)}[n]$ follow an i.i.d. zero mean complex Gaussian distribution, and are independent from the noise.

The entire collected observations of i -th subband is defined as a $K \times N$ matrix $\mathbf{X}_i = [\mathbf{x}_i[0], \dots, \mathbf{x}_i[N-1]]$, where the noise matrix is the $K \times N$ matrix $\mathbf{W}_i = [\mathbf{w}_i[0], \dots, \mathbf{w}_i[N-1]]$, and the signal matrix is the $P \times N$ matrix $\mathbf{S}_i = [\mathbf{s}_i[0], \dots, \mathbf{s}_i[N-1]]$.

By the above assumptions, we define the covariance matrix of i -th subband as $\mathbf{R}_i = \mathbb{E}[\mathbf{x}_i[n]\mathbf{x}_i^H[n]]$, where $(\cdot)^H$ denotes the conjugate transpose. With N received samples in a finite sensing time, the CR receiver calculates the $K \times K$ sample covariance matrix $\hat{\mathbf{R}}_i = \mathbf{X}_i \mathbf{X}_i^H$ to complete the sensing problem.

In our sensing problem assumption, the noise samples subject to zero mean complex Gaussian distribution, and the noise covariance matrix in i -th subband is assumed to be Ψ_i . Under the hypothesis $\mathcal{H}_{0,i}$, the sample covariance matrix is a complex Wishart matrix subject to $\mathcal{W}_K(N, \Psi_i)$. Under the hypothesis $\mathcal{H}_{1,i}$, the sample covariance matrix is denoted as:

$$\mathcal{H}_{0,i} : \hat{\mathbf{R}}_i = \mathbf{W}_i \mathbf{W}_i^H = \Psi_i \quad (2)$$

$$\mathcal{H}_{1,i} : \hat{\mathbf{R}}_i = \mathbf{H}_i \mathbf{S}_i \mathbf{S}_i^H \mathbf{H}_i^H + \Psi_i. \quad (3)$$

With the uncorrelated assumption about $\mathbf{S}_i$, and the transmission power of the p -th PU defined as $\eta_i^{(p)} = \mathbb{E}(s_i^{(p)}[n]s_i^{(p)*}[n])$, Equation (3) can be written as:

$$\hat{\mathbf{R}}_i = \sum_{p=1}^P \eta_i^{(p)} \mathbf{h}_i^{(p)} \mathbf{h}_i^{(p)H} + \Psi_i. \quad (4)$$

The sensing performance is evaluated by the detection probability P_D and the false alarm probability P_F , respectively, corresponding to the correct detection of the presence of primary signal at hypothesis $\mathcal{H}_1$ and wrongly claiming of the presence of primary signal at hypothesis $\mathcal{H}_0$.

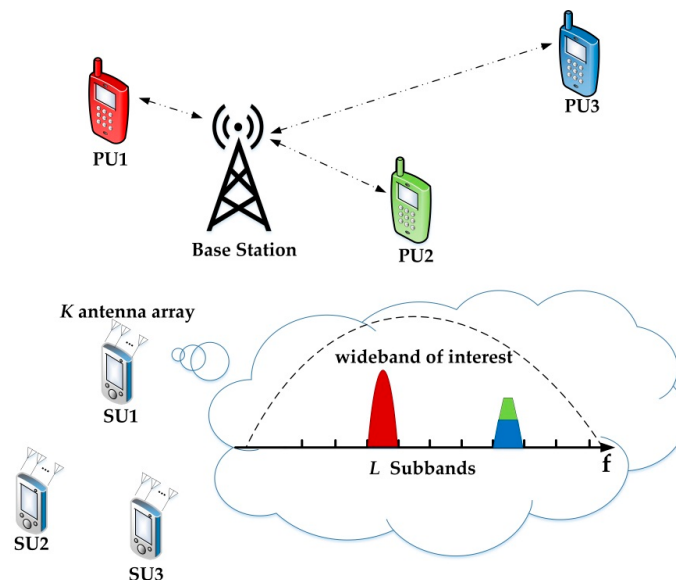


Figure 1. The CR system model with multiple PUs and k antenna SUs.

2.2. Riemannian Distance and Riemannian Mean

In information geometry theory, we consider that $\mathcal{F}$ is a set of probability density functions $p(\mathbf{x}|\boldsymbol{\theta})$, where $\mathbf{x}$ is a sample of n dimensional complex random variable $\mathbf{X}$, i.e., $\mathbf{x} \in \mathbf{X} \subseteq \mathcal{C}^n$. And $\boldsymbol{\theta}$ is the m dimensional parameter vector, i.e., $\boldsymbol{\theta} \in \Theta \subseteq \mathcal{C}^m$. Generally, the probability distribution space can be described by its parameter set Θ . The statistical model S is expressed as:

$$S = \{p(\mathbf{x}|\boldsymbol{\theta}) | \boldsymbol{\theta} \in \Theta \subseteq \mathcal{C}^m\}, \quad (5)$$

Under certain topological structures, S can form a differentiable manifold, called statistical manifold $\mathcal{M}$, where $\boldsymbol{\theta}$ is defined as the natural coordinate of the statistical manifold [22]. From the point of view of information geometry, the probability distribution $\mathcal{W}_k(N, \hat{\mathbf{R}}_i)$ can be parameterized separately by the respective covariance matrix. Then $\mathcal{W}_k(N, \hat{\mathbf{R}}_i)$ can be considered to be located on the statistical manifold which takes covariance matrix as coordinate. In particular, $\hat{\mathbf{R}}_i$ is a Toeplitz Hermitian positive definite matrix with the noise and signal models under $\mathcal{H}_0$ and $\mathcal{H}_1$. The set of covariance matrices constitutes a complex symmetric positive definite (SPD) matrix space denoted by $Sym(n, \mathcal{C})$, which is also defined as SPD manifold. The parameter space Θ of zero-mean multivariate Gaussian distribution and $Sym(n, \mathcal{C})$ are isomorphic. Therefore the statistical manifold can be described by $Sym(n, \mathcal{C})$ due to the mapping relationship between them [23]. Many articles have focused on the research of the geometry structure of complex symmetric positive definite matrix manifold. It is a completely connected, complete Riemannian manifold with non-positive sectional curvature, called the Cartan-Hadamard manifold, whose geodesic exists and is unique [21]. In all curves connecting the two points $\boldsymbol{\theta}_A, \boldsymbol{\theta}_B$ on the manifold, the geodesic is the shortest one. This shortest distance, called the Riemannian distance between $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$, can be used to describe the similarity of the two distributions. The Riemannian distance between two elements in $Sym(n, \mathcal{C})$ is given in [19] as follows:

$$\begin{aligned} D^2(\mathbf{R}_1, \mathbf{R}_2) &= \left\| \log(\mathbf{R}_1^{-1/2} \mathbf{R}_2 \mathbf{R}_1^{-1/2}) \right\|^2 \\ &= \left\| \log(\mathbf{R}_1^{-1} \mathbf{R}_2) \right\|^2 \\ &= \text{Tr} \left[\log^2(\mathbf{R}_1^{-1} \mathbf{R}_2) \right] \\ &= \sum_{i=1}^n \log^2(\lambda_i), \end{aligned} \quad (6)$$

where $\|\cdot\|$ is the Frobenius norm and λ_i is the n eigenvalues of the matrix $\mathbf{R}_1^{-1} \mathbf{R}_2$. Compared with Kullback-Leibler divergence, D^2 has better properties, such as symmetry, satisfying triangle inequality and so on.

Let $\bar{\mathbf{R}}$ be the midpoint between two points $\mathbf{R}_1$ and $\mathbf{R}_2$ in space. In the normed linear space, the midpoint calculation corresponds to the arithmetic mean $\bar{\mathbf{R}} = (\mathbf{R}_1 + \mathbf{R}_2)/2$. On the manifold, it is found that for $Sym(n, \mathcal{C})$, the local curvature is not constant and not positive, so the calculation of the midpoint must depend on the corresponding geometric mean rather than the arithmetic mean. Using the Riemannian distance we can define the midpoint satisfying $D^2(\mathbf{R}_1, \bar{\mathbf{R}}) = D^2(\mathbf{R}_2, \bar{\mathbf{R}})$. The formula for calculating the midpoint with the geometric means is given as $\bar{\mathbf{R}} = \mathbf{R}_1^{1/2} (\mathbf{R}_1^{-1/2} \mathbf{R}_2 \mathbf{R}_1^{-1/2})^{1/2} \mathbf{R}_1^{1/2}$ in [19]. For the N points on $Sym(n, \mathcal{C})$, the Riemannian mean is defined by Riemannian distance:

$$\bar{\mathbf{R}} = \arg \min_{\mathbf{R} \in \text{Sym}(n, \mathcal{C})} \frac{1}{N} \sum_{k=1}^N D^2(\mathbf{R}_k, \mathbf{R}). \quad (7)$$

The geometric representation of the Riemann mean is given in Figure 2. In the field of information geometry research, Riemannian mean computation is an important problem. In [17], a Riemannian mean computation method based on the gradient descent algorithm is proposed. The Riemannian mean is iteratively computed by using the gradient descent algorithm along the geodesic direction of the manifold:

$$\bar{\mathbf{R}}_{t+1} = \bar{\mathbf{R}}_t^{1/2} e^{-\varepsilon \sum_{k=1}^N \log(\bar{\mathbf{R}}_t^{-1/2} \mathbf{R}_k^{1/2} \bar{\mathbf{R}}_t^{-1/2})} \bar{\mathbf{R}}_t^{1/2}. \quad (8)$$

In (8), $\varepsilon \geq 0$ controls the iteration speed and $\mathbf{R}_k (k=1, 2, \dots, N)$ are the N matrices on the manifold. $\bar{\mathbf{R}}_t$ is the estimate of the Riemannian mean calculated from t iterations. The detailed process of iterative computation is given in Algorithm 1.

Algorithm 1: Iterative Calculation of the Riemannian Mean By a Gradient Descent Algorithm

Input: $\mathbf{R}_k (k=1, 2, \dots, N)$ and ε .

Output: Estimates of Riemannian mean $\bar{\mathbf{R}}$.

Initialize: $t = 1$; $\bar{\mathbf{R}}_1 = \mathbf{R}_1$.

repeat

 Compute gradient of objective function $\nabla f = \sum_{k=1}^N \log(\bar{\mathbf{R}}_t^{-1/2} \mathbf{R}_k^{1/2} \bar{\mathbf{R}}_t^{-1/2})$;

 Obtain $\bar{\mathbf{R}}_{t+1} = \bar{\mathbf{R}}_t^{1/2} e^{-\varepsilon \nabla f} \bar{\mathbf{R}}_t^{1/2}$;

 Update $t = t + 1$;

until convergence.

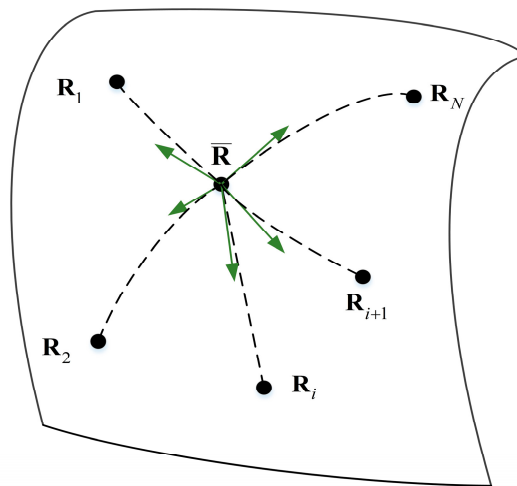


Figure 2. The Riemannian mean of N points on the manifold.

2.3. The Riemannian Distance Based Test Statistic

In our sensing scenario assumptions, the specific form of Ψ_i is arbitrary and unknown, but the Ψ_i of different subbands are identical. This assumption of noise means the different subbands suffer from the same unknown interference or the imperfections of the sensor hardware. It is difficult to design the test statistic according to the probability distributions of the received data under $\mathcal{H}_0$ under such assumptions. We consider the sample covariance matrix of the vacant channel in the wideband spectrum as an estimate of the unknown Ψ_i , called the reference matrix. Then we calculate the Riemannian distance between the covariance matrix of the channel under test

and the reference matrix. Thus, the solution to the detection problem is transformed from the traditional statistical inference method (which is accomplished by hypothesis testing the probability distribution of the statistical model via the N observations of $\mathbf{x}$) into the information geometry method (which is accomplished by computing the geodesic between the two points on manifold).

There are several ways to get the reference matrix. One is to extract noise-only data samples in the system's agreed free time and frequency bands. If the interval between the free time and the sensing time is short enough, or the free bands is close enough to the bands under test, they can be considered to have the same noise covariance characteristics. Besides, we can select one or more subbands with the lower power spectral density (PSD) as the vacant noise-only subbands as detailed in [24]. The sample covariance matrix of the vacant subband is set to be the reference matrix for the Riemann distance based detection method.

If the reference matrix is $\hat{\mathbf{R}}_V$, then the test statistic is defined as:

$$\begin{aligned} T_{RD} &= \left\| \log \left(\hat{\mathbf{R}}_V^{-1/2} \hat{\mathbf{R}}_i \hat{\mathbf{R}}_V^{-1/2} \right) \right\|^2 \\ &= \left\| \log \left(\hat{\mathbf{R}}_V^{-1} \hat{\mathbf{R}}_i \right) \right\|^2 \\ &= \sum_{k=1}^K \log^2 (\lambda_k), \end{aligned} \quad (9)$$

where $\hat{\mathbf{R}}_i$ is the $K \times K$ sample covariance matrix of the subband to be tested, and λ_k ($k = 1, \dots, K$) is the eigenvalue of the matrix $\hat{\mathbf{R}}_V^{-1} \hat{\mathbf{R}}_i$. The proposed Riemannian distance detector is:

$$T_{RD} \underset{\mathcal{H}_{1,i}}{\overset{\mathcal{H}_{0,i}}{\leq}} \gamma, \quad (10)$$

where γ is the pre-set detection threshold.

In summary, we propose the process of spectrum sensing algorithm based on Riemannian distance in wideband:

1. Determine the vacant subband. If there is only one vacant (noise-only) subband, then its sample covariance matrix $\Psi_V = \mathbf{X}_V \mathbf{X}_V^H$ is calculated and used as a reference matrix $\hat{\mathbf{R}}_V = \Psi_V$. If there are multiple vacant subbands, such as A , then the Riemannian mean $\bar{\Psi}_V$ of the noise covariance matrices Ψ_i ($i = 1, \dots, A$) of the multiple vacant subbands can be used as the reference matrix:

$$\hat{\mathbf{R}}_V = \bar{\Psi}_V = \arg \min_{\Psi \in \text{Sym}(n, \mathbb{C})} \frac{1}{A} \sum_{i=1}^A D^2(\Psi_i, \Psi). \quad (11)$$

2. Compute the sample covariance matrix $\hat{\mathbf{R}}_i = \mathbf{X}_i \mathbf{X}_i^H$ of the i -th subband to be tested.
3. Obtain the test statistic $T_{RD, i\text{-th subband}} = D^2(\hat{\mathbf{R}}_V, \hat{\mathbf{R}}_i) = \sum_{k=1}^K \log^2(\lambda_k)$, where λ_k is the eigenvalue of $\hat{\mathbf{R}}_V^{-1} \hat{\mathbf{R}}_i$ with ordered $0 \leq \lambda_1 \leq \dots \leq \lambda_K \leq \infty$.
4. Compare the test statistic with the threshold, and get the sensing result:

$$T_{RD, i\text{-th subband}} \underset{\mathcal{H}_{1,i}}{\overset{\mathcal{H}_{0,i}}{\leq}} \gamma. \quad (12)$$

In the above sensing method, the main computational complexity lies in the covariance matrix inversion calculation. The difficulty of inversion depends on the value of K , which is the number of antennas. As we will present in the Section 4, even with small K , the proposed Riemannian distance detector still shows good detection performance. If K is large, the matrix inversion calculation may not be easy. Then the alternative methods noted in [19] may be a better choice. It should be noticed that in Equation (6), we can calculate the square root matrix instead of the inverse matrix. In [19], a

square root of positive definite matrix computing method is given. Using the Schulz iteration, the intractable matrix inversion can be avoided.

3. Threshold and Probability of False Alarm

The P_F performance and P_D performance of the proposed Riemannian distance detector depend on the probability distribution of the test statistic Equation (9) under $\mathcal{H}_0$ and $\mathcal{H}_1$. However, the analytical forms of the probability distribution under both hypotheses are hard to be obtained. In this section we derive the accurate closed-form approximation of P_F to get the decision threshold. And the P_D performance will be presented and discussed through numerical simulations in Section 4.

3.1. Moments of Test Statistics under $\mathcal{H}_0$

We consider a moment matching method to approximate the probability density function (PDF) of the test statistics. Therefore, we need to obtain the exact moments of T_{RD} under $\mathcal{H}_0$:

$$\mathbb{E}(T_{RD}^p) = \mathbb{E}\left\{\left(\sum_{k=1}^K \log^2(\lambda_k)\right)^p\right\}. \quad (13)$$

In this section, we will give a detailed calculation process of the moments of test statistics under $\mathcal{H}_0$. First, we show the joint PDF of the eigenvalues in Equation (9) according to the random matrix theory. Then we use the joint PDF to solve the moments of test statistics which can be regarded as the eigenvalue function. The calculation of the first order moments will be presented as an example. And the closed-form solution to the p -th moment will be given at last.

Consider the case where the number of reference matrices is $A = 1$, then the reference covariance matrix and the sample covariance matrix of the i -th subband in the expression of the test statistic T_{RD} under $\mathcal{H}_0$ follow the distributions respectively:

$$\hat{\mathbf{R}}_v \sim \mathcal{W}_K(N, \mathbf{\Psi}_v), \quad (14)$$

$$\hat{\mathbf{R}}_i \sim \mathcal{W}_K(N, \mathbf{\Psi}_v), \quad \text{under } \mathcal{H}_0, \quad (15)$$

where $\hat{\mathbf{R}}_v$ is independent of $\hat{\mathbf{R}}_i$. The joint PDF of the eigenvalues λ_k ($k = 1, \dots, N$) of the matrix $\hat{\mathbf{R}}_v^{-1}\hat{\mathbf{R}}_i$ has been given under the assumption that the dimensions of two complex Wishart matrices are equal [25,26], which is:

$$f(\lambda_1, \lambda_2, \dots, \lambda_K) = \frac{\Gamma_N(2K)}{\Gamma_N(K)\Gamma_N(K)\Gamma_N(N)} \prod_{k=1}^K \frac{\lambda_k^{N-K}}{(1+\lambda_k)^{2K}} \prod_{k < l}^K (\lambda_l - \lambda_k)^2. \quad (16)$$

In the case of a known joint probability density of λ_k , we can refer to the moment calculation method in [27] to solve the moments of the eigenvalue function $T(\lambda_1, \lambda_2, \dots, \lambda_K) = \sum_{k=1}^K \log^2(\lambda_k)$. In general, if the joint probability density (16) can be written as:

$$f(\lambda) = C_0 |\Phi(\lambda)| |\Psi(\lambda)| \prod_{k=1}^K \xi(\lambda_k). \quad (17)$$

Then the expectation of the function $\beta(\lambda_k)$ for λ_k can be solved as follows:

$$\mathbb{E}\left\{\prod_{k=1}^K \beta(\lambda_k)\right\} = C_0 |U|, \quad (18)$$

where $\Phi(\lambda)$, $\Psi(\lambda)$, $\mathbf{U}$ are $K \times K$ matrices. The (i, j) elements in $\Phi(\lambda)$ and $\Psi(\lambda)$ are $\Phi_i(\lambda_j)$ and $\Psi_j(\lambda_i)$, respectively. The (i, j) element in $\mathbf{U}$ can be expressed as:

$$u_{i,j} = \int_0^\infty \Phi_i(\lambda) \Psi_j(\lambda) \xi(\lambda) \beta(\lambda) d\lambda. \quad (19)$$

By introducing the function $\omega(x) = e^{sx}$, we can use Equations (17) and (18) to solve the required moments:

$$\mathbb{E} \left\{ e^{s \sum_{k=1}^K \log^2(\lambda_k)} \right\} = \mathbb{E} \left\{ \prod_{k=1}^K e^{s \log^2(\lambda_k)} \right\}. \quad (20)$$

By p -th order derivatives of (20) at $s = 0$, we have the p -th moment of $T(\lambda_1, \lambda_2, \dots, \lambda_K)$:

$$\mathbb{E} \left\{ \left(\sum_{k=1}^K \log^2(\lambda_k) \right)^p \right\} = \frac{d^p}{ds^p} \left\{ \mathbb{E} \left(e^{s \sum_{k=1}^K \log^2(\lambda_k)} \right) \right\} \Big|_{s=0}. \quad (21)$$

We take $C_0 = \frac{\Gamma_N(2K)}{\Gamma_N(K) \Gamma_N(K) \Gamma_N(N)}$ as the normalized coefficients determined by N , K , and choose $\Phi(\lambda) = \Psi(\lambda) = \mathbf{V}(\lambda)$, where $\mathbf{V}(\lambda)$ is a Vandermonde matrix. Choosing $\xi(\lambda_k) = \frac{\lambda_k^{N-K}}{(1 + \lambda_k)^{2K}}$, $\beta(\lambda_k) = e^{s \log^2(\lambda_k)}$, we can get:

$$\mathbb{E} \left\{ \prod_{k=1}^K e^{s \log^2(\lambda_k)} \right\} = \mathbb{E} \left\{ \prod_{k=1}^K \beta(\lambda_k) \right\} = C_0 |\mathbf{U}(s)|. \quad (22)$$

The element $u(s)_{i,j}$ in matrix $\mathbf{U}(s)$ corresponds to the sub-item in (19) as follows that:

$$\begin{aligned} \Phi_i(\lambda) &= V_i(\lambda) = \lambda^{i-1}, \\ \Psi_j(\lambda) &= V_j(\lambda) = \lambda^{j-1}, \\ \xi(\lambda) &= \frac{\lambda^{N-K}}{(1 + \lambda)^{2N}}, \\ \beta(\lambda) &= e^{s \log^2(\lambda)}. \end{aligned} \quad (23)$$

Then (19) can be written as:

$$u(s)_{i,j} = \int_0^\infty \frac{\lambda^{N-K+i+j-2}}{(1 + \lambda)^{2N}} e^{s \log^2(\lambda)} d\lambda. \quad (24)$$

For the differential operation in (22), we can use the rules for the matrix determinant in [28]:

$$\frac{d}{ds} |\mathbf{U}(s)| = \sum_{k=1}^K |\mathbf{U}_{(k)}(s)|, \quad (25)$$

where $\mathbf{U}_{(k)}(s)$ is the matrix that coincides with $\mathbf{U}(s)$ except that every entry in the k -th row (equivalently, columns could be used) is differentiated with respect to s .

Then we take the first moment calculation as an example, for $p = 1$:

$$\mathbb{E}(T_{RD}) = \frac{d}{ds} \{ C_0 |\mathbf{U}(s)| \} \Big|_{s=0} = C_0 \sum_{k=1}^K |\mathbf{U}_{(k)}(0)|. \quad (26)$$

$u(0)_{i,j}$ and $\frac{d}{ds} u(s)_{i,j} \Big|_{s=0}$ are the two types of matrix elements in $\mathbf{U}_{(k)}(0)$:

$$u(0)_{i,j} = \int_0^{\infty} \frac{\lambda^{N-K+i+j-2}}{(1+\lambda)^{2N}} d\lambda = B(N-K+i+j-1, N+K-i-j+1), \quad (27)$$

$$\frac{d}{ds} u(s)_{i,j} \Big|_{s=0} = \int_0^{\infty} \frac{\lambda^{N-K+i+j-2}}{(1+\lambda)^{2N}} \log^2(\lambda) d\lambda, \quad (28)$$

where $B(x, y)$ is the Beta function. Thus the (i, j) element of $\mathbf{U}_{(k)}(0)$ can be denoted as and the function $I_{k,j}(x)$ is defined by:

$$[\mathbf{U}_{(k)}(0)]_{i,j} = \int_0^{\infty} \frac{\lambda^{N-K+i+j-2}}{(1+\lambda)^{2N}} I_{k,j}(\log^2(\lambda)) d\lambda. \quad (29)$$

$$I_{k,j}(x) \triangleq \begin{cases} x, & \text{if } k = j \\ 1, & \text{if } k \neq j \end{cases}. \quad (30)$$

Similarly, the result of p -th moment can be obtained. There will be an integral term with the following form when computing the p -th moment:

$$\frac{d^p}{ds^p} u(s)_{i,j} \Big|_{s=0} = \int_0^{\infty} \frac{\lambda^{N-K+i+j-2}}{(1+\lambda)^{2N}} \log^{2p}(\lambda) d\lambda. \quad (31)$$

In Appendix A we give a detailed derivation for the integral form such as $\int_0^{\infty} \frac{t^{x-1}}{(1+t)^{x+y}} \log^q(t) dt$.

To sum up, we can get the exact analytical p -th moment.

3.2. Gamma Approximation Approach

With the exact analytical p -th moment obtained in Section 3.1, we can use the gamma distribution function to approximate the test statistic under $\mathcal{H}_0$ according to the moment matching method in [29].

Using the gamma approximation, we only need to compute the first and second moments to obtain the mean and variance of the test statistic as follows:

$$\begin{aligned} \mu_T &\triangleq \mathbb{E}(T_{RD}) = M(1) \\ \sigma_T^2 &\triangleq \mathbb{E}(T_{RD}^2) - [\mathbb{E}(T_{RD})]^2 = M(2) - M(1)^2, \end{aligned} \quad (32)$$

where $M(p)$ is the p -th moment of T_{RD} . Suppose the test statistic satisfies the gamma distribution with shape parameter k_T and scale parameter θ_T and $F_{T_{RD}}$ is denoted as the cumulative distribution function (CDF) of T_{RD} . According to the CDF of a gamma distribution defined in [30], $F_{T_{RD}}$ is derived as:

$$\begin{aligned} k_T &= \frac{\mu_T^2}{\sigma_T^2} = \frac{M(1)^2}{M(2) - M(1)^2} \\ \theta_T &= \frac{\sigma_T^2}{\mu_T} = \frac{M(2) - M(1)^2}{M(1)}, \end{aligned} \quad (33)$$

$$F_{T_{RD}}(x; k_T, \theta_T) = 1 - \frac{\Gamma(k_T, x/\theta_T)}{\Gamma(k_T)}, \quad (34)$$

where $\Gamma(k_T, x/\theta_T)$ is the upper incomplete gamma function.

Denoting γ as the decision threshold, then the probability of false alarm under $\mathcal{H}_0$ is:

$$P_F = 1 - F_{T_{RD}}(\gamma; k_T, \theta_T). \quad (35)$$

So we can use the inverse of the gamma cumulative distribution function to compute the decision threshold:

$$\gamma = F_{T_{RD}}^{-1}(1 - P_F; k_T, \theta_T). \quad (36)$$

Although the threshold calculation method we derived seems to be complex enough, such complexity may not be a serious problem in the spectrum sensing process. As we can see in the deduction process above, the threshold depends on the desired P_F and the system parameters N , K . These parameters should be predefined in the CR node. Therefore threshold calculation does not need to be real-time. Even if the system is required to operate under different parameter set, we can still calculate the threshold of different parameter sets in advance and make a query table for real-time mode changing.

4. Numerical Results

In this section, we give the performance evaluation of the proposed algorithms by means of Monte Carlo simulations. First, we examine the accuracy of the P_F approximation method and the decision threshold calculation. Then we compare the detection performance of RDD with some conventional EBDs in different scenarios. In addition, we show the performance of RDD when Riemannian mean method is applied in multiple vacant subbands situation. In this section of the simulation, we choose the value of N , K according to the practical system. Due to the limited sensing time and sampling rate, N generally takes tens to hundreds, while K is generally less than eight because of hardware size. We choose the exponential correlation model [31] as the noise covariance matrix:

$$(\Psi_0)_{i,j} = \rho^{|i-j|}, \quad \rho \in [0, 1] \quad (37)$$

where ρ denotes the degree of noise correlation. And each subband has the same noise covariance matrix, i.e., $\Psi_i = \Psi_0$ for $i = 1, 2, \dots, L$.

In addition, we assume that the channel matrix $\mathbf{H}_i$ of the subband is composed of independent Gaussian random variables and satisfies the normalization conditions $\|\mathbf{H}_i\|^2 = 1$. In each Monte Carlo realization, the channel matrix is generated randomly. The PU signals in the simulation follow the i.i.d. zero mean complex Gaussian distribution, and are independent from the noise.

The received SNR of p -th primary signal in the i -th subband is defined as:

$$SNR_i^{(p)} = \frac{\eta_i^{(p)} \|\mathbf{H}_i\|^2}{\nu_i} \quad (38)$$

where $\eta_i^{(p)}$ is the transmission power of the p -th PU, and ν_i is the noise power in the i -th subband, which satisfies $\nu_i = \text{Tr}(\Psi_i)/K$. In the case of multiple PUs, we set different SNRs for different PUs due to the fact that the distance between multiple PUs and SU are not the same in the practical sensing situations.

4.1. Decision Threshold and P_F

In Table 1, we present the numerical simulation of the moments computing method proposed in Section 3.1. It is worth noting that the joint PDF of the eigenvalues defined by (16) holds for any value of the covariance matrix parameter Σ in two complex Wishart distributions $\mathcal{W}_K(N, \Sigma)$. Hence, in the simulation, we can specify some combinations of the sample size N and antenna number K , and then generate the complex Wishart matrix with arbitrary covariance matrix to

calculate the test statistic under $\mathcal{H}_0$. In Table 1, the accuracy of the theoretical calculation is verified by comparing the simulated and theoretical values of the first and second moment of the test statistic T_{RD} under $\mathcal{H}_0$.

Table 1. Numerical results of T_{RD} .

(N,K)	$\mathbb{E}(T_{RD})$ Simulated	$\mathbb{E}(T_{RD})$ Analytical	$\mathbb{E}(T_{RD}^2)$ Simulated	$\mathbb{E}(T_{RD}^2)$ Analytical
(30,4)	1.1708	1.1695	1.5489	1.5474
(50,4)	0.6750	0.6756	0.5147	0.5151
(50,8)	2.8708	2.8667	8.5091	8.4914
(80,8)	1.7135	1.7424	3.1125	3.14

In Figure 3, we give the simulation values of the CDF of the test statistic and the theoretical value of the gamma distribution approximation at the specified (N,K) . This figure shows that the gamma distribution approximation based on the moment matching method proposed in Section 3.2 achieves good performance. Moreover, our approximation algorithm are perfectly matched to the simulation values for different (N,K) .

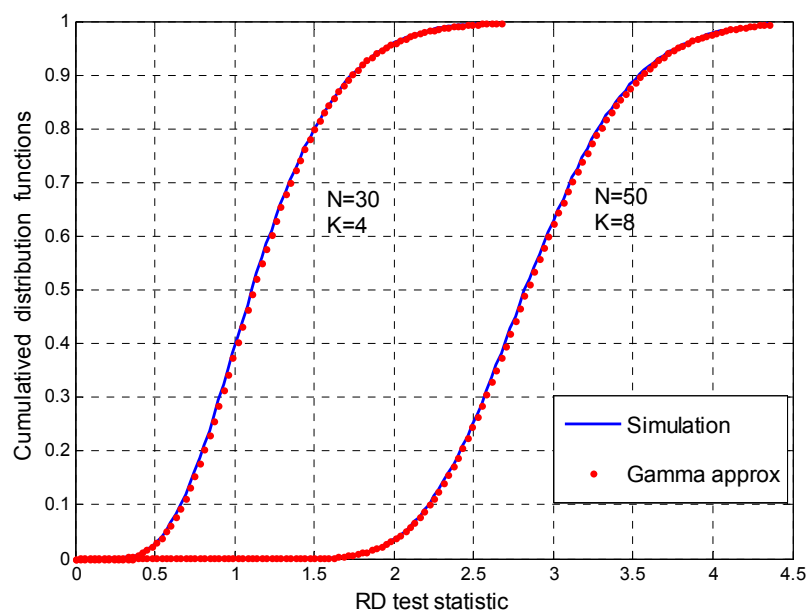


Figure 3. Cumulative distribution functions (CDFs) of the test statistics for $(N,K) = (30,4)$ and $(N,K) = (50,8)$.

We plot the decision threshold as a function of P_F for the specified (N,K) in Figure 4. The simulation value curve is obtained by Monte Carlo method. Meanwhile, the theoretical value of P_F can be obtained directly by (35). The figure shows the perfect agreement between the theoretical computing and simulation results. Therefore, the sensing algorithm can get the corresponding decision threshold under the specified (N,K) and the desired P_F . In this way, the detector can satisfy the constant false alarm rate (CFAR) requirements of the CR system.

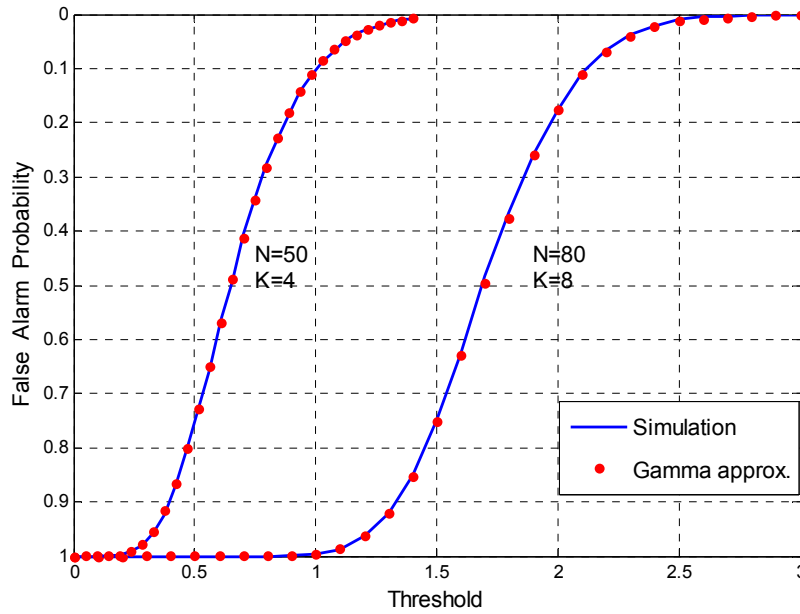


Figure 4. Decision threshold γ vs. P_F for $(N,K) = (50,4)$ and $(N,K) = (80,8)$.

4.2. Detection Performance

First, the Receiver Operating Characteristic (ROC) curve is used to compare the performance of the proposed detector and other detectors when the number of reference matrices is $A = 1$. The ROC curve gives the detection probability P_d as a function of the false alarm probability P_F . By changing the threshold γ , the operating point of the detector can be chosen anywhere along its ROC curve.

Considering that the noise variance, noise covariance matrix and the number of PUs are both unknown in our assumption, blind detection is necessary. Among the conventional EBDs, the SLE detector and the AGM detector are two typical blind detectors [32]. According to the classification of EBD in [29], they correspond to arithmetic mean detector (ARMD) and arithmetic-geometric mean detector (AR-GEMD), respectively. Both detectors operate in a single-band spectrum sensing scenario and do not require noise-only subband for detection. In order to deploy these two EBDs in multi-band detection, we use the pre-whitening technique [33] to improve their sensing performance in multi-band with unknown noise. By replacing $\hat{\mathbf{R}}_i$ by $\hat{\mathbf{R}}_V^{-1}\hat{\mathbf{R}}_i$ for the EBD in the i -th subband, we can compare the performance of EBD and the proposed RDD in the simulation.

The ROC curves of the detectors when there is one primary signal are plotted in Figure 5 where $SNR^{(1)} = -3dB$, $\rho = 0.4$ and $(N,K) = (100,4)$. It is clear that the proposed detector outperforms the two EBDs under the assumption of correlated noise.

Figures 6 and 7 show the ROC curves when the number of PUs is assumed to be $P = 3$ and $P = 5$ respectively. The SNRs of multiple PUs are defined as $SNR^{(i+1)} = SNR^{(i)} - 1dB$. In the case of different combinations of N and K , we can see the same performance differences between RDD and EBDs as in Figure 5. It should be seen that ARMD performs better than AR-GEMD in the case of single PU but worse in the case of multiple PUs. As noted in [9,14], the rank of received covariance matrix is assumed in the derivation of the GLR-based detector. In our signal model, the number of PUs has an effect on the rank of the covariance matrix, resulting in the EBD method being selective to P , as shown in the figures. By contrast, RDD is not sensitive to the number of PUs because we do not make any assumptions about P in the proposed algorithm. From the geometric point of view, the detection performance of RDD depends on the Riemannian distance between $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{R}}_V$. Although the combination of different primary signals will lead to the position change of $\hat{\mathbf{R}}_i$ on the manifold, the distance-based detection method remains effective.

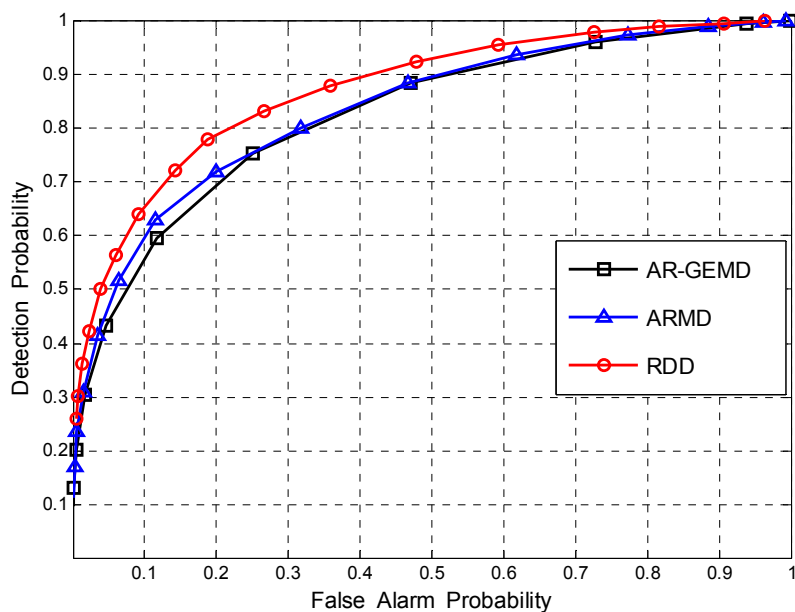


Figure 5. The ROCs comparison for $(N,K) = (100,4)$. Assuming $P = 1$, $\rho = 0.4$, and $SNR^{(1)} = -3dB$.

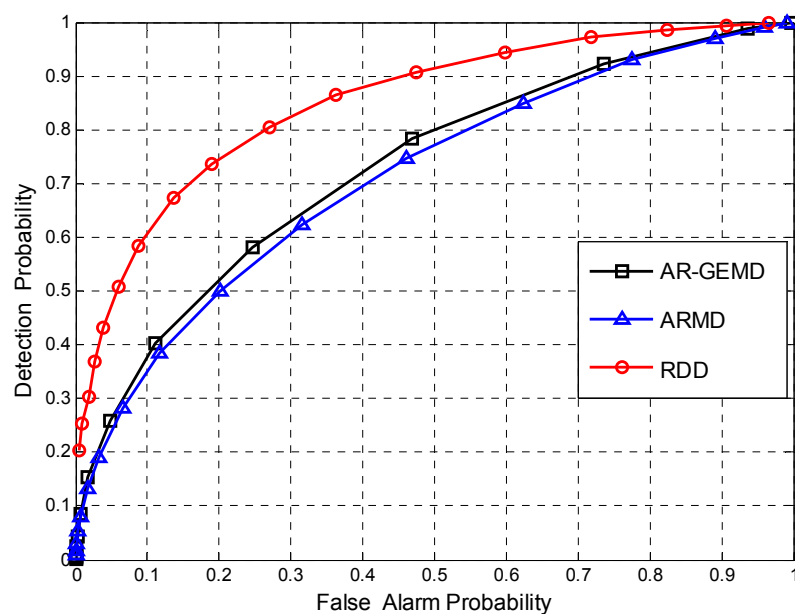


Figure 6. The ROCs comparison for $(N,K) = (100,4)$. Assuming $P = 3$, $\rho = 0.4$, and $SNR^{(1)} = -6dB$.

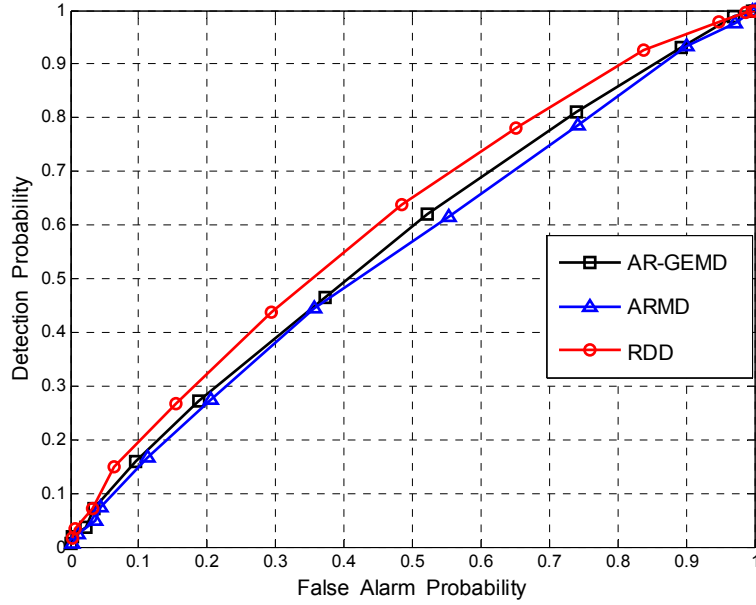


Figure 7. The ROCs comparison for $(N,K) = (80,8)$. Assuming $P = 5$, $\rho = 0.4$, and $SNR^{(1)} = -8dB$.

In practical, communication systems often require a constant P_F , as required by the IEEE 802.22 standard to be 10% [34]. We plot P_d at different average received SNR in Figure 8 and choose $(N,K) = (50,4)$ and $\rho = 0.4$, while the decision threshold has been set to achieve $P_F = 0.1$. And the number of PUs is one in Figure 8a and three in Figure 8b. The figures illustrate that the RDD can achieve higher detection probability than EBD for the same SNR level.

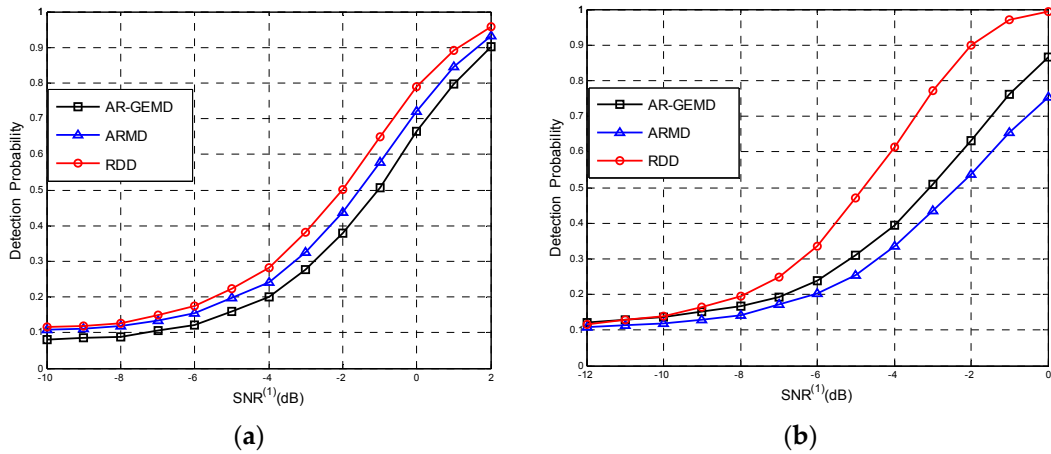


Figure 8. The probability of detection versus SNR for $(N,K) = (50,4)$ and $\rho = 0.4$ in different PU numbers (a) $P = 1$; (b) $P = 3$.

4.3. Multiband Detection with Riemannian Mean

In this section, we consider the performance of RDD when multiple vacant subbands are present, i.e., $A > 1$. First, we would like to evaluate the performance of matrix mean estimation between the Riemannian mean (RM) method and the arithmetic mean (AM) method. In the simulation, the matrix dimension parameter is $(N,K) = (80,8)$, the matrix type is the correlated noise matrix, and the correlation coefficient ρ satisfying the uniform distribution $\mathcal{U}(0,1)$ is randomly generated in each Monte Carlo implementation. The objective function in the Riemannian mean expression (7), which is the mean distance, can be written as follows:

$$f(\mathbf{R}_k, \mathbf{R}_t) = \frac{1}{A} \sum_{k=1}^A D^2(\mathbf{R}_k, \mathbf{R}_t). \quad (39)$$

The trend of the mean distance with the iteration number t is shown in Figure 9. When calculating (8), we choose $\varepsilon = 0.1$ to control the iteration speed. Meanwhile, in Figure 9 we show the comparison of RM and AM estimation performance when the number of reference matrices is $A = 4$ and $A = 8$. It can be observed that the objective function of the RM calculation based on the gradient descent algorithm decreases as the number of iterations increases, and tends to be stable after a certain number of iterations. It should be noted that the iterative computation is not required in AM method, so the mean distance of AM in the figure remains unchanged. The figure shows that for different number of reference matrices, the mean distance of the converged RM method is less than that of the AM method.

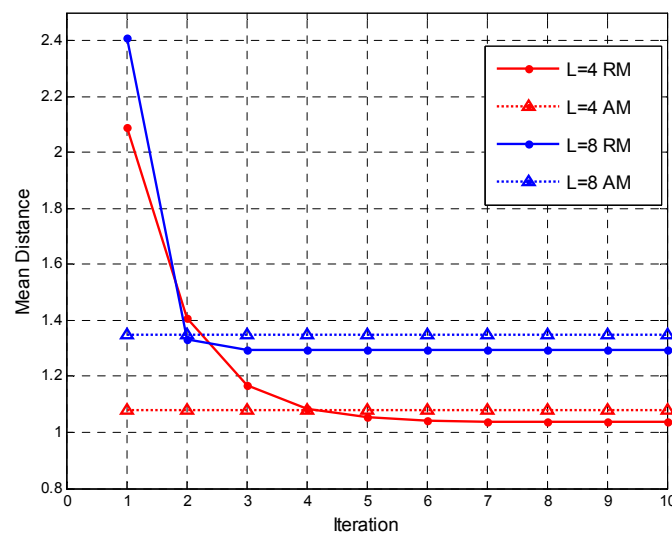


Figure 9. Estimation Performance Comparison for $(N,K) = (80,8)$ and random ρ when the number of reference matrices is $A = 4$ and $A = 8$.

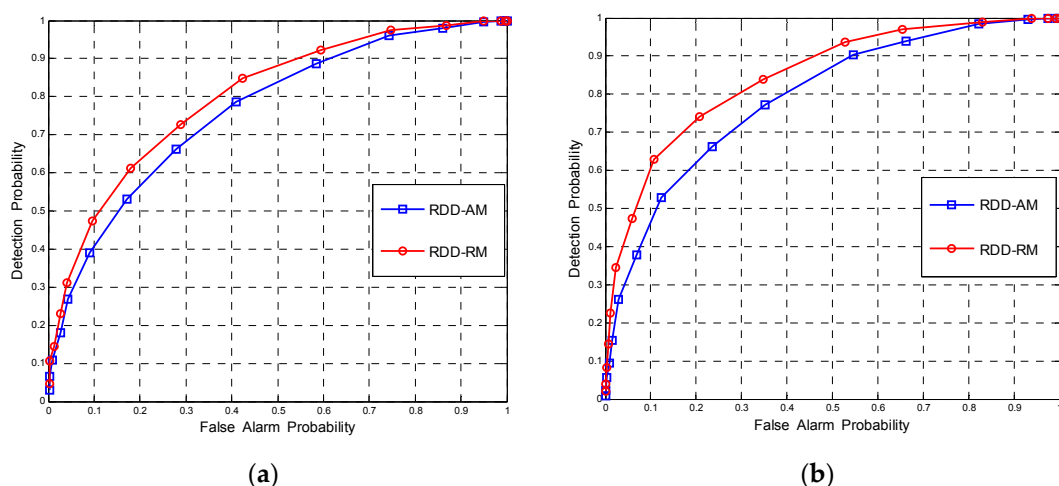


Figure 10. The ROCs of RDD using AM method and RM method for $(N,K) = (80,8)$, $\rho = 0.4$, $P = 3$ and $SNR^{(1)} = -5$ with different number of reference matrices (a) $A = 2$; (b) $A = 4$.

Finally, we compare the ROC curves of RM method and the AM method in the case of multiple vacant subbands. We choose $(N,K) = (80,8)$, $\rho = 0.4$, $P = 3$ and $SNR^{(1)} = -5$ for the simulation. Figure 10a,b plot the ROC curves with reference matrices $A = 2$ and $A = 4$, respectively, where the iteration number of RM method is 8. We can see that for RDD, RM estimation outperforms the AM

estimation, and when the number of reference matrices is large, the gap is more obvious. The above simulation results illustrate that the sensing performance can be improved if the RM matrix of reference matrices is used for the RDD when multiple vacant subbands are available in WSS.

5. Conclusions

In this paper, we propose a novel WSS detector based on Riemannian distance for multi-antenna CR. For a wideband spectrum divided into several subbands, we compute the Riemannian distance of covariance matrices between the vacant subband and the subband to be detected. Through theoretical analysis, we obtain the exact closed expression of the decision threshold using the moment matching method. Unlike the traditional EBD derived with the GLR criterion, the proposed RDD is derived from information geometry theory. By applying the geometric method, we do not have to make too many assumptions about noise and primary signals like the traditional methods do. Therefore we obtain a detector which is blind to noise statistical characteristics and number of PUs. The simulation results show that the proposed detector exhibits better performance than the conventional EBD method in the correlated noise model and is robust to the number of PUs. Moreover, we propose a matrix mean estimation method based on RM. In the presence of multiple vacant subbands, the RM method can better estimate the noise distribution than the AM method. Thus it is more suitable for RDD due to the better use of the wideband spectrum information for sensing.

Acknowledgments: This study is funded by the National Basic Research Program of China (973 Program, No. 613314). The fund received covers the costs to publish in open access.

Author Contributions: Q.L. and F.L. contributed to the conception of the study; Q.L., S.Y. and F.L. contributed significantly to theoretical analysis and manuscript preparation; Q.L. performed the analysis with constructive discussions and made simulation experiment; Q.L. and S.Y. wrote and revised the paper; All authors approved the final version.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

According to the definition of Beta function [30] we have:

$$B(x, y) = \int_0^1 \frac{t^{x-1}}{(1+t)^{x+y}} dt, \quad (A1)$$

$$\frac{\partial^q}{\partial x^q} B(x, y) = \int_0^1 \frac{t^{x-1}}{(1+t)^{x+y}} \log^q \left(\frac{t}{1+t} \right) dt, \quad (A2)$$

$$\frac{\partial^q}{\partial y^q} B(x, y) = \int_0^1 \frac{t^{x-1}}{(1+t)^{x+y}} [-\log(1+t)]^q dt. \quad (A3)$$

By transformation as:

$$\left\{ \log \left(\frac{t}{1+t} \right) - [-\log(1+t)] \right\}^q = \log^q t \quad (A4)$$

We can easily obtain:

$$\int_0^1 \frac{t^{x-1}}{(1+t)^{x+y}} \log^q t dt = \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right)^q B(x, y). \quad (A5)$$

In calculating the first and second moments of the test statistic, there are $q = 2$ and $q = 4$ in the integral equation to be solved. They can be expressed as follows:

$$\int_0^1 \frac{t^{x-1}}{(1+t)^{x+y}} \log^2 t dt = B_{2,0}(x, y) + B_{0,2}(x, y) - 2B_{1,1}(x, y), \quad (A6)$$

$$\int_0^{\infty} \frac{t^{x-1}}{(1+t)^{x+y}} \log^4 t \, dt = B_{4,0}(x, y) + B_{0,4}(x, y) - 4B_{3,1}(x, y) - 4B_{1,3}(x, y) + 6B_{2,2}(x, y) \quad (A7)$$

where $\frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} B(x, y) = B_{m,n}(x, y)$. The required $B_{m,n}(x, y)$ in the computing can be written as a combination of digamma functions, such as:

$$B_{1,0}(x, y) = B(x, y) (\psi(x) - \psi(x+y)), \quad (A8)$$

$$B_{2,0}(x, y) = B(x, y) (\psi(1, x) - \psi(1, x+y) + (\psi(x) - \psi(x+y))^2). \quad (A9)$$

Similarly, we can derive other $B_{m,n}(x, y)$ using digamma function expansion.

References

1. Mitola, J.; Maguire, G.Q. Cognitive radio: Making software radios more personal. *IEEE Pers. Commun.* **1999**, *6*, 13–18.
2. Hong, X.; Wang, J.; Wang, C.X.; Shi, J. Cognitive radio in 5G: A perspective on energy-spectral efficiency trade-off. *IEEE Commun. Mag.* **2014**, *52*, 46–53.
3. Tandra, R.; Sahai, A. SNR Walls for Signal Detection. *IEEE J. Sel. Top. Signal Process.* **2008**, *2*, 4–17.
4. Sahai, A.; Cabric, D.; Hoven, N.; Mishra, S.; Tandra, R. Spectrum sensing: Fundamental limits and practical challenges. In Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Baltimore, MD, USA, 8–11 November 2005.
5. Sutton, P.D.; Nolan, K.E.; Doyle, L.E. Cyclostationary Signatures in Practical Cognitive Radio Applications. *IEEE J. Sel. Areas Commun.* **2008**, *26*, 13–24.
6. Zeng, Y.; Koh, C.L.; Liang, Y.C. Maximum Eigenvalue Detection: Theory and Application. In Proceedings of the 2008 IEEE International Conference on Communications, Beijing, China, 19–23 May 2008; pp. 4160–4164.
7. Zeng, Y.; Liang, Y.C. Eigenvalue based Spectrum Sensing Algorithms for Cognitive Radio. *IEEE Trans. Commun.* **2008**, *57*, 1784–1793.
8. Zeng, Y.; Liang, Y.C.; Zhang, R. Blindly Combined Energy Detection for Spectrum Sensing in Cognitive Radio. *AEU—Int. J. Electron. Commun.* **2012**, *66*, 649–652.
9. Zhang, R.; Teng, J.L.; Liang, Y.C.; Zeng, Y. Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach. *IEEE Trans. Commun.* **2010**, *58*, 84–88.
10. Sacuto, F.; Labeau, F.; Agba, B.L. Wide Band Time-Correlated Model for Wireless Communications under Impulsive Noise within Power Substation. *IEEE Trans. Wirel. Commun.* **2013**, *13*, 1449–1461.
11. Cacciapuoti, A.; Caleffi, M.; Paura, L. Optimal Strategy Design for Enabling the Coexistence of Heterogeneous Networks in TV White Space. *IEEE Trans. Veh. Technol.* **2016**, *65*, 7361–7373.
12. Leshem, A.; Veen, A.J.V.D. Multichannel detection of Gaussian signals with uncalibrated receivers. *IEEE Signal Process. Lett.* **2001**, *8*, 120–122.
13. Sun, Z.; Laneman, J.N. Performance Metrics, Sampling Schemes, and Detection Algorithms for Wideband Spectrum Sensing. *IEEE Trans. Signal Process.* **2014**, *62*, 5107–5118.
14. Besson, O.; Kraut, S.; Scharf, L.L. Detection of an unknown rank-one component in white noise. *IEEE Trans. Signal Process.* **2006**, *54*, 2835–2839.
15. Amari, S. *Information Geometry and Its Applications*; Springer: Tokyo, Japan, 2016; pp. 315–353.
16. Lenglet, C.; Rousson, M.; Deriche, R.; Faugeras, O. Statistics on the Manifold of Multivariate Normal Distributions: Theory and Application to Diffusion Tensor MRI Processing. *J. Math. Imaging Vis.* **2006**, *25*, 423–444.
17. Barbaresco, F. Interactions between symmetric cone and information geometries: Bruhat-tits and siegel spaces models for high resolution autoregressive doppler imagery. In Proceedings of the Emerging Trends in Visual Computing 2008, Palaiseau, France, 18–20 November 2008; pp. 124–163.
18. Arnaudon, M.; Barbaresco, F.; Yang, L. Riemannian Medians and Means with Applications to Radar Signal Processing. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 595–604.

19. Barbaresco, F. Innovative Tools for Radar Signal Processing Based on Cartan's Geometry of SPD Matrices & Information Geometry. In Proceedings of the 2008 IEEE Radar Conference, Rome, Italy, 26–30 May 2008; pp. 1–6.
20. Li, Y.; Wong, K.M. Riemannian Distances for Signal Classification by Power Spectral Density. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 655–669.
21. Nielsen, F.; Bhatia, R. *Matrix Information Geometry*; Springer: Heidelberg, Germany, 2013; pp. 199–207.
22. Amari, S.I. Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory* **2001**, *47*, 1701–1711.
23. Arsigny, V.; Fillard, P.; Pennec, X.; Ayache, N. Geometric Means in A Novel Vector Space Structure on Symmetric Positive-Definite Matrices. *SIAM J. Matrix Anal. Appl.* **2006**, *29*, 328–347.
24. An, T.; Song, I.; Lee, S.; Min, H.K. Detection of Signals with Observations in Multiple Subbands: A Scheme of Wideband Spectrum Sensing for Cognitive Radio with Multiple Antennas. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 6968–6981.
25. Khatri, C.G. Classical Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution. *Ann. Math. Stat.* **1965**, *36*, 98–114.
26. Muirhead, R.J. *Aspects of Multivariate Statistical Theory*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1982; pp. 450–451.
27. Chiani, M.; Win, M.Z.; Zanella, A. On the capacity of spatially correlated MIMO Rayleigh-fading channels. *IEEE Trans. Inf. Theory* **2003**, *49*, 2363–2371.
28. Horn, R.A. *Topics in Matrix Analysis*; Cambridge University Press: New York, NY, USA, 1991; pp. 491–492.
29. Shakir, M.Z.; Rao, A.L.; Alouini, M.S. Generalized Mean Detector for Collaborative Spectrum Sensing. *IEEE Trans. Commun.* **2013**, *61*, 1242–1253.
30. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; U.S. Department Commerce: Washington, DC, USA, 1964; pp. 255–263.
31. Shiu, D.S.; Faschini, G.J.; Gans, M.J.; Kahn, J.M. Fading correlation and its effect on the capacity of multi-element antenna systems. *IEEE Trans. Commun.* **2000**, *48*, 502–513.
32. Zhang, W.; Wang, C.-X.; Tao, X.; Patcharamaneepakorn, P. Exact Distributions of Finite Random Matrices and Their Applications to Spectrum Sensing. *Sensors* **2016**, *16*, doi:10.3390/s16081183.
33. Zeng, Y.; Liang, Y.C. Spectrum-Sensing Algorithms for Cognitive Radio Based on Statistical Covariances. *IEEE Trans. Veh. Technol.* **2009**, *58*, 1804–1815.
34. Cordeiro, C.; Challapali, K.; Birru, D.; Sai, S.N. IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive Radios. *J. Commun.* **2006**, *1*, 38–47.

