BAYESIAN MULTI-ANTENNA SENSING IN COGNITIVE RADIO NETWORKS USING FRACTIONAL BAYES FACTOR

Mohannad H. Al-Ali and K. C. Ho

ECE Department, University of Missouri, Columbia, MO 65211, USA
mohannadalalicom@yahoo.com, hod@missouri.edu

ABSTRACT
This paper proposes a Bayesian detector for spectrum sensing in a multi-antenna cognitive radio (CR) network in which no channel state information (CSI) is available. The Bayesian approach for detection necessitates a prior distribution of the CSI in terms of the spatial covariance matrix, and unfortunately it is often improper and cannot be applied directly. We shall introduce the use of the Fractional Bayes Factor (FBF) approach to handle improper prior, which in turn yields a well-defined Bayes factor as the test statistic for detection. A number of priors of the CSI are examined and a closed-form expression for the test statistics is derived. The developed Bayesian detector is compared with those by using the conjugate priors for both hypotheses and the generalized likelihood ratio test (GLRT), and it yields considerable improvement in detection performance.

Index Terms— Bayesian detection, Fractional Bayes factor, multi-antenna, prior distribution, spectrum sensing

1. INTRODUCTION

Opportunistic access for cognitive radio (CR) network can provide an efficient use of the limited spectrum resources and allow interweaving between heterogeneous networks [1]. Essentially, the secondary user (SU) in a CR network seeks opportunistic access to the spectral band of a licensed primary network when the primary user (PU) is idle. Spectrum sensing in the interweave paradigm is an essential component for the design of a CR network [2]. The capabilities of SU to detect the presence of PU can be enhanced significantly through incorporating multiple antennas at the terminals of SU and PU [2–14].

Spectrum sensing in the interweave multi-antenna CR requires a detector that generates a defined test statistic to be compared against a specific threshold value to attain a reliable belief about the activity of PU. Such a detector can be separated into the deterministic and Bayesian categories. For the deterministic category, the test statistic can be based on energy [3], multivariate cyclostationary [4], eigenvalues of the sample covariance matrix [5–8], or the generalized likelihood ratio test (GLRT) [9–11]. The first three kinds have limitations as they assume, respectively, the noise power is accurately known, the PU signal has a format with its cyclic frequency known, or the PU signal does not have any structure. The GLRT detector does not guarantee in general optimality for the employed test statistic [15, Ch. 6]. On the other hand, the Bayesian approach avoids estimating the unknown parameters through introducing prior distributions for them and marginalizes the likelihood function [12–15]. In the Bayesian framework, the Bayes factor is used as the test statistic and it can be considered as the odds of one hypothesis to another provided by the data. In this work, we shall regard the Bayesian detection for spectrum sensing, where the unknown parameters are the spatial covariance matrix that represents the channel state information (CSI) from PU to SU.

Several approaches have been proposed in the statistics community for the evaluation of the Bayes factor, please refer to [16] for a good overview. Obtaining the Bayes factor is not straightforward, due to the choice of prior, the improper prior behavior and the integration for marginalization. Indeed, utilizing proper priors for hypothesis testing is crucial to ensure a well behaved Bayes factor [16]. The conjugate prior is proper but it often yields inadequate results for Bayesian detection [12] as indicated in Section 4. In this paper, we shall introduce the fractional Bayes factor (FBF) approach to define the Bayes factor for the spectrum sensing problem that can provide better performance [17, 18].

FBF can work with different priors that can be improper, and it can transform improper priors into proper ones through the concept of training samples. Specifically, FBF uses a fraction of the likelihood function to make the priors become proper and the remaining for hypothesis evaluation. Consequently, it avoids the need in determining the rather difficult non-informative or objective conventional priors (CPs) [19]. Furthermore, it is computationally attractive and does not need averaging over all the possible training data as compared to the intrinsic Bayes factor [20].

In conjunction with the use of FBF, we shall introduce a class of improper priors for the CSI under the hypothesis that PU is active in transmitting signal [21]. It includes the Jeffreys’, independence Jeffreys’ [22], Geisser and Cornfield’s [23], right- and left-Haar measure, and reference priors [24]. This set of priors has found particularly desirable in (sparse) Gaussian graphical models [25] and provided better results but it is not proper. For the hypothesis that PU is not active, we shall use the conjugate prior for the unknown noise parameters, which unfortunately becomes improper as the hyperparameters go to zero [26]. In either hypothesis, FBF will handle the improperness of the priors naturally without difficulty.

We would like to outline the contributions as follows. The use of FBF to produce the Bayes factor with improper priors has not been explored in signal processing. Indeed, the only available work we found in the engineering literature on FBF is to use it for associating proper priors of the unknown parameters [29]. To the best of our knowledge, the proposed class of improper priors for CR spectrum sensing is new. Previous attempts in the literature are limited to the conjugate priors [12] or CPs [27, 28]. Furthermore, we have derived closed-form expressions for the marginal likelihoods and the FBF test statistic.

The organization of this paper is as follows. Section 2 formulates the problem and defines the Bayes factor for spectrum sensing.
Section 3 presents the FBF technique, introduces prior distributions for the unknown CSI parameters, and determines the expression for the associated FBF. Section 4 compares the proposed Bayesian detector with the state-of-art detectors and Section 5 concludes the paper.

Notation: Upper-case and lower-case bold-face letters denote matrices and column vectors. The matrix functions $|A|$, $etr(A)$, and $Diag(A)$ denote the determinant, exp(trace), and an operator that takes the diagonal elements of $A$ to form a diagonal matrix. $E[·]$ and $(·)'$ denotes the expectation operator and the transpose. $\mathbb{R}^{n \times n}$ is the space of $n \times n$ real matrices. $\mathcal{N}(m, C)$ represents multivariate Gaussian distribution with mean $m$ and covariance $C$.

2. PROBLEM FORMULATION

2.1. Problem Setup

We shall consider spectrum sensing for an interweave CR network in which SU is able to exploit the spectrum resources of the primary network whenever PU is not active. Fig. 1 depicts the considered interweave CR model. We assume PU has $l$ transmit and SU has $p$ receive antennas. The transmitted signal from PU is $x_k \in \mathbb{R}^{l \times 1}$, $k = 1, \ldots, N$, where $N$ is the number of samples available for detection. We shall follow [9–13] and consider $x_k$ is Gaussian distributed with $\mathcal{N}[x_k] = \mathbf{0}$. The signal propagates through the channel represented by the matrix $\mathbf{H} \in \mathbb{R}^{p \times N}$ that is assumed static during the $N$ sample period and reaches SU. The observed signal at SU is $y_k = x_k + w_k$, $k = 1, \ldots, N$. The collections of the transmitted and received samples form the matrices $\mathbf{X} = [x_1 \cdots x_N]$ and $\mathbf{Y} = [y_1 \cdots y_N]$.

Spectrum sensing in CR can be cast as a detection problem that intends to distinguish between the following two hypotheses:

$$H_0: \mathbf{Y} = \mathbf{W},$$

$$H_1: \mathbf{Y} = \mathbf{HX} + \mathbf{W},$$

where $\mathbf{W} = [w_1 \cdots w_N]$, $w_k \in \mathbb{R}^{p \times 1}$, is the zero-mean additive white Gaussian noise matrix. The hypotheses $H_0$ and $H_1$ in (1) correspond to the null (noise) model and to the data model.

It is reasonable to consider the received signal $y_k$ is IID. Thus the detection problem becomes one of choosing between two multivariate normal distributions from the observations [2, 11–13].

$$H_0: y_k \sim \mathcal{N}(\mu, \mathbf{D}), \quad k = 1, \ldots, N$$

$$H_1: y_k \sim \mathcal{N}(\mu, \Sigma), \quad k = 1, \ldots, N$$

where $\mu \in \mathbb{R}^{p \times 1}$ is the mean vector, $\mathbf{D} = \mathbb{E}[ww']$ is a diagonal matrix of positive diagonal elements. We do not restrict the diagonal elements to be identical to account for uncalibrated multi-antenna receivers [9, 12]. $\Sigma \in \mathbb{R}^{p \times p}$ is a positive definite matrix that is equal to $\Sigma = \mathbf{H}\mathbf{H}' + \mathbf{D}$.

The CSI for the detection problem (2) is not available and $\mathbf{D}$, $\mathbf{H}$, and $\Sigma$ are unknown parameters. In the development follows, we shall denote the likelihood function for the hypothesis $H_0$ as $f_0(\mathbf{Y}/\mu, \mathbf{D})$ and that for $H_1$ as $f_1(\mathbf{Y}/\mu, \Sigma)$.

2.2. Bayesian Detection

Let $\pi_0(\mathbf{D})$ and $\pi_1(\mu, \Sigma)$ be the objective (non-informative) prior distributions for the unknown parameters $\mu, \mathbf{D}$ and $\mu, \Sigma$. The marginal likelihood functions for the hypotheses are

$$m_0(\mathbf{Y}) = \int f_0(\mathbf{Y}/\mu, \mathbf{D})\pi_0(\mu, \mathbf{D})d\mu d\mathbf{D},$$

$$m_1(\mathbf{Y}) = \int f_1(\mathbf{Y}/\mu, \Sigma)\pi_1(\mu, \Sigma)d\mu d\Sigma.$$
3.2. Prior Distribution under \( H_0 \)

For calculation tractability, we shall use the following conjugate proper prior for \( (\mu, \mathbf{D}) \)

\[
\pi_0(\mu, \mathbf{D}) \overset{\text{indp}}{\sim} \prod_{j=1}^{p} IG(h_j/2, \delta_jj/2)
\]

(8a)

\[
= \prod_{j=1}^{p} \left( \frac{\delta_jj/2}{\Gamma(h_j/2)} \right)^{h_j/2} d_j^{-(\frac{j+1}{2})} \exp \left( \frac{-\delta_jj}{2d_jj} \right) \quad \text{(8b)}
\]

\[
= \left( \frac{\Delta^{(h/2)}}{2^{h/2} \Gamma(h/2)} \right) |\mathbf{D}|^{-(\frac{j+1}{2})} \exp \left( \frac{-1}{2} \Delta \mathbf{D}^{-1} \right) \quad \text{(8c)}
\]

where \( IG(\cdot, \cdot) \) stands for the inverse Gamma distribution, \( h_j/2 \) and \( \delta_jj/2, \ j = 1, \cdots, p \) are the shape and scale parameters of \( IG(\cdot, \cdot) \) and \( \Gamma(\cdot) \) denotes the gamma function. The matrix \( \Delta \) is diagonal formed by having \( \delta_jj > 0, \ j = 1, \cdots, p \), as the diagonal elements. Note that in (8c) the prior for \( \mu \) is uniform and independent of the prior for \( \mathbf{D} \).

In the absence of prior information about \( \delta_jj \), it is customary to make the prior in (8) having non-influential effect by setting the hyperparameters \( (h_j, \delta_jj) \) to small values such as \((0.001, 0.001)\). We should note however that the prior \( IG(\epsilon, \epsilon) \) becomes improper as \( \epsilon \to 0 \) [26].

Define \( h = h + N \) and the sample covariance matrix \( \mathbf{S} \) as

\[
\mathbf{S} = \sum_{k=1}^{N} (\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^T
\]

(9)

where \( \bar{\mathbf{y}} = (1/N) \sum_{k=1}^{N} \mathbf{y}_k \) is the sample mean. It can be shown that the marginal likelihood \( m_0(\mathbf{Y}) \) defined in (3a) under the prior distribution (8c) has the following closed form expression [14]

\[
m_0(\mathbf{Y}) = \frac{1}{\pi^{Np/2}} \frac{\Gamma_p(\hat{h}/2)}{\Gamma_p(h/2)} |\Delta + \text{Diag}(\mathbf{S})|^{h/2} \quad \text{(10)}
\]

3.3. Prior Distributions under \( H_1 \)

We shall introduce a collection of possible improper priors for the covariance matrix \( \mathbf{\Sigma} \). It would be convenient to express \( \mathbf{\Sigma}^{-1} \) in a unique decomposition form through the Cholesky factorization

\[
\mathbf{\Sigma}^{-1} = \Psi \Psi^T
\]

(11)

where \( \Psi \in \mathbb{R}^{p \times p} \) is an upper triangular matrix with positive diagonal elements. The off-diagonal elements are denoted by \( \psi_{jk}, k > j \).

Rather than using priors on the whole \( \mathbf{\Sigma} \) as in [12, 13], we shall follow [21] and apply priors on the elements of \( \Psi \). The general class of priors for \( (\mu, \Psi) \) that we propose is

\[
\pi_{a}(\mu, \Psi) = \prod_{j=1}^{p} \frac{1}{\psi_{jj}}
\]

(12)

Finally, using \( a_j = 1 \) reduces (12) to the reference prior \( \pi_{RH} \) that is defined in [24].

We shall next obtain \( m_1(\mathbf{Y}) \) described in (3b) for the prior (12). Let the Cholesky factorization of the sample covariance matrix given in (9) be \( \mathbf{S} = \mathbf{V}^T \mathbf{V} \), where \( \mathbf{V} \in \mathbb{R}^{p \times p} \) is an upper triangular matrix with diagonal elements \( v_{jj} > 0, \ j = 1, \cdots, p \). Define the upper triangular matrix \( \mathbf{T} = \mathbf{V}^T \mathbf{W} \) whose diagonal elements are \( t_{jj} > 0, \ j = 1, \cdots, p \). Through the Jacobian of the transformation from \( \Psi \) to \( \mathbf{V} \) and from \( \Psi \) to \( \mathbf{T} \), we have \( d\Psi/d\mathbf{V} = 2^p \prod_{j=1}^{p} \psi_{jj}^{2(\nu+1)+j} \) and \( d\Psi/d\mathbf{T} = \prod_{j=1}^{p} v_{jj}^{2p-j+1} \). As a result, the likelihood function \( f_1(\mathbf{Y}/\mu, \Psi) \) becomes \( f_1(\mathbf{T}/\mu, \mathbf{V}) \). Since \( \psi_{jj} = t_{jj}v_{jj} \), it can be shown that \( f_1(\mathbf{T}/\mu, \mathbf{V}) \) has the following expression [14],

\[
f_1(\mathbf{T}/\mu, \mathbf{V}) = 2^p (2\pi)^{-Np/2} \prod_{j=1}^{p} \frac{v_{jj}^{c_j}}{t_{jj}^{d_j}} \exp \left( \frac{-1}{2} \mathbf{T}^T \mathbf{T} \right) \quad \text{(13)}
\]

where \( c_j = 3(p+1) - 2j - N \) and \( d_j = -(2p+1) + j \).

We have shown in [14] that the off-diagonal elements of \( \mathbf{T} \) follows the unit normal distribution and the diagonal elements the chi-square distribution with \( \tilde{\nu}_j = d_j + 1 \) degrees of freedom (dof).

Consequently, \( m_1(\mathbf{Y}) \) can be evaluated explicitly as

\[
m_1(\mathbf{Y}) = 2^p (2\pi)^{-Np/2} \prod_{j=1}^{p} \frac{v_{jj}^{c_j}}{t_{jj}^{d_j}} \exp \left( \frac{-1}{2} \mathbf{T}^T \mathbf{T} \right) \quad \text{(14)}
\]

where \( O = (p+1)/2, \tilde{d}_j = d_j - a_j, \tilde{\nu}_j = d_j + 1, \) and \( \tilde{c}_j = c_j + a_j \).

For further details, please refer to [14].

3.4. Bayes Factor Evaluation

The prior for \( \mathbf{D} \) is improper as \( h, \delta \to 0 \), so does the class of priors (12) for the matrix \( \mathbf{\Sigma} \). The following proposition provides a closed form expression for the test statistic \( B_{10}^p \).

Proposition 1 [14]: The FBF \( B_{10}^p \) can be evaluated according to (7) as

\[
B_{10}^p = K \prod_{j=1}^{p} 2^{p-\frac{j}{2}} \frac{\Gamma_p(\tilde{\nu}_j/2)}{\Gamma_p(\tilde{\nu}_j/2)} \left( \frac{\Delta + \text{Diag}(\mathbf{S})}{\Delta + b \text{Diag}(\mathbf{S})} \right)^{\frac{1}{2}}
\]

(15)

where the parameters \( \tilde{b}_{0,j} = b_{0,j} + a_{j}, \tilde{b}_{j} = b_{j} - a_{j}, \tilde{h}_0 = h_0 + b_{h}N \)

and \( K = 2^{p(2-b)/2(N/2)^{p} h_0/(h/2)} \).

Proof: We should first substitute (8c) and use \( \pi_{a} \) for \( \pi_{RH} \) from (12) for \( \pi_{a}(\mu, \Sigma) \) in (7). Then we obtain from (10) and (14) the factor \( B_{10}^p \) defined in (4). Simplifying yields the FBF expression (15). Please refer to [14] for the details.

Since the gamma function \( \Gamma(\tilde{\nu}_j + 1)/2 \) in (15) has to be a positive value, the following condition must be satisfied for the fraction \( b \)

\[
b > \frac{p - 1}{-2(p + 1) + N + 1},
\]

where the number of samples should be \( N > 2(p + 1) - 1 \).

4. NUMERICAL RESULTS

This section presents numerical results for the hypothesis testing problem of spectrum sensing in interweave CR. We use \( 10^3 \) realizations to generate data in each hypothesis to evaluate the probability of detection and the probability of false alarm according to (5) and (6). The detection threshold \( \gamma \) in (6) for a given \( P_{fa} \) value is

3576
determined experimentally as the probability $P_{FA}$ does not admit a closed-form expression. The number of samples to obtain the test statistics for detection is $N = 50$.

We compare the performance of the proposed FBF detector (15) with the ones that are based on the conjugate prior $\pi_C$ in both hypotheses [12] and on the generalized likelihood ratio test (GLRT) [5]. Their test statistics are provided in Appendices A and B.

Using $l = 4$ PU transmit antennas and $p = 4$ SU receive antennas, Fig. 2 shows the average probability of detection as the signal-to-noise ratio (SNR) at SU increases while keeping a fixed $P_{FA}$ at $10^{-3}$. The fraction $b$ for the proposed FBF is set to $b = 0.1$. We examine all members of the class of improper priors: $\pi_I$, $\pi_{GC}$, $\pi_{RH}$, and $\pi_R$ for FBF. They behave similarly and the difference occurs at very low SNR where the $\pi_{RH}$ prior shows the lowest $P_D$ while the $\pi_R$ prior has the best detection performance. Any of the FBF outperforms the conjugate prior [12] and the GLRT test [5] detector considerably. For example to reach a $P_D = 0.5$, the FBF detectors require an SNR of $-2.5$ dB while the other two detectors need 1 dB and 6 dB, respectively.

Fig. 3 shows the average probability of detection versus the probability of false alarm. The SNR is $-4$ dB and the numbers of antennas are $l = 5, p = 5$. The fraction for FBF is $b = 0.12$. We observe again that the behaviors of the different priors defined in (12) for the CSI in $H_I$ are similar. They all provide superior performance to the detectors using the conjugate prior $\pi_C$ and GLRT test. At $P_{FA} = 10^{-2}$, the corresponding probability of detection $P_D$ for the FBF detector with $\pi_{RH}$ prior, and the $\pi_C$ prior and GLRT based detectors are 0.67, 0.16 and $1.1 \times 10^{-2}$.

5. CONCLUSION

In this paper, we have developed a Bayesian detector for spectrum sensing in interweave CR networks. The proposed detector employs FBF to handle the problem of improper prior distributions in generating the Bayes factor for detection. We have introduced a new class of improper priors for the covariance matrix that represents the CSI between PU and SU. Numerical results show that the proposed FBF Bayesian detector has superior performance to that with the conjugate prior and to the GLRT test.

Appendix A

The conjugate priors for the distributions (2a) and (2b) are (8c) and the Wishart distributions, respectively. Let $u$ and $\Upsilon$ be the dof and the scale matrix of the Wishart distribution for (2b). It can be verified that the Bayes factor (4) corresponding to the conjugate priors $B^{(8c)}_{10}$ has the following form

$$B^{(8c)}_{10} = \frac{\Gamma(\tilde{u}/2)^{\rho(h/2)}}{\Gamma(u/2)^{h/2}} \frac{|\Y^{\tilde{u}/2}| |\Delta + \text{Diag}(\S)|^{h/2}}{|\Y + \S|^{u/2}} \frac{|\Delta^{h/2}}{\Delta^{h/2}}}, \quad (17)$$

where $\tilde{u} = u + N$. We set $\Upsilon = I$ to indicate that there is no CSI about the covariance matrix $\Sigma$.

Appendix B

The GLRT test statistic from [5] is

$$T_{\text{GLRT}} = (C (1 - \kappa/p)^{p-1})^{-N} \quad (18)$$

where $C = (1 - 1/p)^{p-1}$ and $\kappa = \lambda_1 / \sum_{i=1}^p \lambda_i$ with $\lambda_i$ being the eigenvalues of the covariance matrix $\frac{1}{N} \S$ arranged in non increasing order.

6. REFERENCES


