Optimal GLRT-Based Robust Spectrum Sensing for MIMO Cognitive Radio Networks With CSI Uncertainty

Adarsh Patel, *Student Member, IEEE*, Sinchan Biswas, *Student Member, IEEE*, and Aditya K. Jagannatham, *Member, IEEE*

***Abstract*—In this paper, we develop generalized likelihood ratio test (GLRT)-based detectors for robust spectrum sensing in multiple-input multiple-output (MIMO) cognitive radio networks considering uncertainty in the available channel state information (CSI). Initially, for a scenario with known CSI uncertainty statistics, we derive the novel robust estimator-correlator detector (RECD) and the robust generalized likelihood detector (RGLD), which are robust against the uncertainty in the available estimates of the channel coefficients. Subsequently, for a scenario with unknown CSI uncertainty statistics, we develop a generalized likelihood ratio test (GLRT) based composite hypothesis robust detector (CHRD) for spectrum sensing. Closed form expressions are presented for the probability of detection (P_D) and the probability of false alarm (P_{FA}) to characterize the detection performance of the proposed robust spectrum sensing schemes. Further, a deflection coefficient based optimization framework is also developed and solved to derive closed form expressions for the optimal beacon sequences. Simulation results are presented to demonstrate the performance improvement achieved by the proposed robust spectrum sensing schemes and to verify the analytical results derived.**

***Index Terms*—Cognitive radio, spectrum sensing, multiple-input multiple-output (MIMO), channel state information (CSI) uncertainty.**

I. Introduction

THE proliferation of wireless multimedia applications has lead to a natural increase in the demand for a higher data rate in current and upcoming 3G/4G wireless communication systems. The radio frequency (RF) spectrum is a valuable resource for wireless communication applications and its usage is regulated by agencies such as the Federal Communications Commission (FCC) of the United States etc. However, recent reports such as [1] on the temporal and geographic RF spectrum utilization patterns point to an extremely low efficiency of current spectrum utilization. Thus, to cope with the tremendous increase in the demand for RF spectrum, coupled with the motivation for efficient spectrum utilization, the FCC has recently proposed the cognitive radio paradigm [2]–[4] which allows a set of unlicensed/secondary users to opportunistically access unused spectrum bands licensed to primary users. This in turn leads to an increase in the efficiency of spectrum utilization. This strategic reuse of the licensed spectrum by the unlicensed secondary users in cognitive radio networks necessitates the need to reliably detect the presence of vacant spectral bands, termed as *spectral holes* or *white spaces*, without causing a significant interference to the licensed primary users. This key task of spectral occupancy detection in cognitive radio networks is termed as *spectrum sensing*.

Several spectrum sensing techniques have been proposed in literature [5] for the detection of primary user signals in cognitive radio networks. Among these, the energy detector [6] has a simple structure and does not require any prior knowledge of the primary user signal. However, the non-coherent energy detector has a poor performance compared to other techniques for spectrum sensing [7]–[9]. Other popular spectrum sensing techniques exploit statistical properties of the primary user signal for the detection of spectral holes. These techniques can be extended to wideband spectrum sensing, with Nyquist and sub-Nyquist sampling, as described in [10], [11] and to a distributed wideband sensing scenario with finite feedback in [12]. For instance, the cyclostationarity based detection scheme [5] requires the existence and knowledge of the cyclic frequency of the primary user signal. Similarly, the matched-filter based detection scheme [5] requires perfect channel state information (CSI) of the primary user, in the absence of which its performance degrades drastically [13]. Further, the work in [14] demonstrates that soft combination based schemes, such as equal gain combining (EGC) and maximal ratio combining (MRC), exhibit a significant improvement in the detection performance over conventional hard decision fusion schemes. However, it is shown in [15], that the improvement in the detection performance obtained from such soft-decision based spectrum sensing schemes depends significantly on the accuracy of the CSI at the secondary user. An adaptive spectrum sensing scheme based on a finite state Markov channel (FSMC) model is presented in [16]. However, the FSMC with finite states for the channel coefficient does not capture the continuous variation of the channel.

On the other hand, the generalized likelihood ratio test (GLRT) based approach exploits the prior information of the signal of the licensed user. One major advantage of the GLRT based approach is that it achieves joint primary user detection and unknown parameter estimation. In literature, several spec-

Manuscript received November 20, 2014; revised April 15, 2015 and July 15, 2015; accepted October 18, 2015. Date of publication November 13, 2015; date of current version February 16, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Joseph Cavallaro. This work was supported in part by the TCS Research Scholarship Program, DST-SERB and DST-IUATC Project Grants.

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India 208016 (e-mail: adarsh@iitk.ac.in; sinchan@iitk.ac.in; adityaj@iitk.ac.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2015.2500183

trum sensing algorithms have been proposed for cognitive radio scenarios, which compute the maximum likelihood estimates (MLE) of the unknown system parameters towards employing GLRT based detection schemes. For example, [17]–[20] compute the MLE of the noise variance and the signal correlation matrix, whereas [21]–[23] compute the MLE only for the noise variance in order to employ the GLRT framework for primary user signal detection. However, it is shown [7], [8] that the noise power uncertainty agnostic detectors, such as the energy detector, are limited by a signal-to-noise ratio (SNR) wall below which the detector fails to detect the presence of the primary user signal, irrespective of the sensing duration. Similarly, [24] shows the existence of the sampling wall, i.e., the sampling density below which the performance of the detector can not be guaranteed at a given SNR, regardless of the number of samples of the primary signal acquired by the secondary user. Further, [17], [18], [25] employ the GLRT framework with the unknown parameter being the CSI between the primary transmitter and the secondary receiver. The various contributions of these works are listed below.

- The work in [17] presents a GLRT based framework for spectrum sensing with unknown channel gains and considering a single antenna at the primary user and multiple antennas at the secondary user
- The work in [18] presents various partial GLRT schemes with each of the channel gain, noise variance and signal variance components unknown considering block as well as fast fading channel scenarios
- The work in [25] presents a framework for GLRT based spectrum sensing with Bernoulli nonuniform sampling (BNS) considering unknown signal power.

However, a significant shortcoming of all these works above is that they do not consider CSI uncertainty, which is of significance in communication scenarios where frequently nominal CSI is available at the receiver. The work in [26] considers the Wilks' detector based spectrum sensing for a scenario with unknown noise covariance. However, the authors therein consider a simplistic scenario with a deterministic channel matrix and analyze the instantaneous detection performance without averaging over the channel statistics.

Obtaining the true channel coefficients in a cognitive radio network is challenging, owing to estimation/quantization errors, limited feedback, Doppler shift etc, in practical wireless scenarios. These uncertainties in the system parameters can potentially lead to a performance degradation of uncertainty agnostic spectrum sensing schemes. Therefore, this paper considers a GLRT based approach for spectrum sensing, unlike conventional techniques such as matched filtering, energy detection and soft combining etc described in [5], [6], [14] respectively. Works such as [10]–[12], exploit only second order statistical information. Further, the motivation of this work is to develop detection schemes which are robust against the performance degradation caused by the uncertainty in the available CSI estimates, unlike the conventional GLRT schemes [17]–[23], [25] which consider completely unknown CSI, noise variance or other parameters. We propose novel primary user detection schemes, namely the robust estimator-correlator detector (RECD) and the generalized likelihood ratio test (GLRT) based robust generalized likelihood detector (RGLD), which are robust against the uncertainty in the nominal estimates of the channel coefficients for a scenario with known uncertainty statistics. An analytical framework is developed to characterize the theoretical performance of the proposed RECD scheme and derive the expressions for the probability of detection (P_D) and probability of false alarm (P_{FA}). Also, the work in this paper considers a general scenario with multiple antennas, while [10]–[12], [14]–[16], [23] are restricted to single antenna scenarios. Next we present the composite hypothesis based robust detector (CHRD), for primary user detection in scenarios with unknown CSI uncertainty statistics. Further, we also derive closed form expressions to determine the probability of detection (P_D) and probability of false alarm (P_{FA}) performance of the proposed CHRD scheme. Also, the analysis for the P_{FA}, P_D in our work is carried out by averaging with respect to the distributions of both the CSI uncertainty and also the available nominal channel estimate, unlike the work in [26]. The subsequent section presents a novel deflection coefficient based optimization framework to derive the optimal beacon matrices for the proposed robust detection schemes, with known/unknown uncertainty covariance statistics, which further enhance the detection performance. Simulation results demonstrate the improvement in the detection performance achieved by the proposed spectrum sensing schemes and also validate the analytical results.

This paper is organized as follows. Section II describes the multiuser MIMO system model for cognitive radio networks. Section III presents the robust estimator-correlator detector (RECD) and the robust generalized likelihood detector (RGLD) with known CSI uncertainty statistics and the associated detection and false alarm probabilities. Next, in Section IV we present the composite hypothesis based robust detector (CHRD) with unknown CSI uncertainty statistics. Closed form expressions for the P_D and P_{FA} to analytically characterize the detection performance of the proposed scheme are also derived therein. Section V describes the optimization framework to derive the optimal beacon sequence with both known and unknown CSI uncertainty statistics. Simulation results are given in Section VI followed by the conclusion in Section VII.

Throughout this paper we use boldface uppercase/lowercase letters to denote matrices/vectors, respectively. All the vectors are column vectors. The operations $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $\mathrm{E}\{\cdot\}$ denote the conjugate, transpose, conjugate transpose and expectation operators. The random vector $\mathbf{x}$, when defined as $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ follows a complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\mathbf{R}$. Similarly, $\mathbf{x} \sim \chi_n^2$ implies that $\mathbf{x}$ follows a central chi-squared distribution with n degrees of freedom. The L_2 norm of a vector and the trace of a matrix are represented by $\|\cdot\|$ and $\mathrm{Tr}(\cdot)$ respectively. The matrix $\mathcal{D}(\mathbf{a})$ denotes the diagonal matrix with elements of the vector $\mathbf{a}$ along its principal diagonal. The identity matrix of dimension $N \times N$ is denoted by $\mathbf{I}_N$ and the various detector test statistics are denoted by $T(\cdot)$. Finally, the functions $\Gamma(\cdot)$, $\Gamma(\cdot,\cdot)$ and $Q(\cdot)$ denote the Gamma function, the incomplete Gamma function and the Gaussian Q-function respectively.

II. System Model

Consider a spectrum sensing scenario with a primary user base-station and a secondary user. Further, we assume a multiple-input multiple-output (MIMO) cognitive radio network with N_t transmit antennas at the primary user base-station and N_r receive antennas at the secondary user. The baseband system model for the scenario described above at the nth time instant is given as,

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \boldsymbol{\eta}(n),$$

where $\mathbf{y}(n) \in \mathbb{C}^{N_r \times 1}$ is the receive signal vector at the secondary user corresponding to the primary user base-station broadcast beacon signal $\mathbf{x}(n) \in \mathbb{C}^{N_r \times 1}$ and the vector $\boldsymbol{\eta}(n) \in \mathbb{C}^{N_r \times 1}$ is the additive spatio-temporally white Gaussian noise at the secondary user with covariance $\mathbf{R} = \mathrm{E}\{\boldsymbol{\eta}(n)\boldsymbol{\eta}^H(n)\} = \sigma^2\mathbf{I}_{N_r}$. Each $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, is the MIMO channel matrix between the primary user base-station and the secondary user. The system model considered follows in the same spirit as those in works such as [17], [18], [22], [25]. From the Cognitive Radio system model described above, the signal $y_k^*(n)$ for the kth receive antenna can be equivalently expressed as,

$$y_k^*(n) = \mathbf{x}^H(n)\mathbf{h}_k + \boldsymbol{\eta}_k^*(n),$$

where $(\cdot)^*$ denotes the complex conjugate and the vectors $\mathbf{h}_k^H \in \mathbb{C}^{1 \times N_t}, 1 \leq k \leq N_r$ constitute the rows of the channel matrix $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N_r}]^H$. Let the L broadcast beacon vectors $\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(L)$ be concatenated to form a beacon matrix $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(L)]^H \in \mathbb{C}^{L \times N_t}$. The concatenated signal $\mathbf{y}_k$ corresponding to the L symbols can be equivalently represented as,

$$\mathbf{y}_k = \mathbf{X}\mathbf{h}_k + \boldsymbol{\eta}_k, \tag{1}$$

where $\mathbf{y}_k = [y_k(1), y_k(2), \ldots, y_k(L)]^H \in \mathbb{C}^{L \times 1}$ is the signal at the kth receive antenna corresponding to the broadcast beacon matrix $\mathbf{X}$. Similarly, the concatenated noise vector $\boldsymbol{\eta}_k = [\eta_k(1), \eta_k(2), \ldots, \eta_k(L)]^H \in \mathbb{C}^{L \times 1}$ has covariance matrix $\mathbf{R}_\eta = \mathrm{E}\{\boldsymbol{\eta}_k\boldsymbol{\eta}_k^H\} = \sigma_\eta^2\mathbf{I}_L$. In practical wireless scenarios it is significantly challenging to obtain accurate CSI as described previously. Hence, we model the channel matrix $\mathbf{H}$ by incorporating the CSI uncertainty as,

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{U}, \tag{2}$$

where $\hat{\mathbf{H}} = [\,\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_{N_r}\,]^H$ is the available nominal CSI and the matrix $\mathbf{U} \in \mathbb{C}^{N_r \times N_t}$ denotes uncertainty in the channel matrix $\mathbf{H}$. The uncertainty matrix $\mathbf{U} = [\,\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{N_r}\,]^H$, where each row vector $\mathbf{u}_k^H \in \mathbb{C}^{1 \times N_t}, 1 \leq k \leq N_r$ follows a complex Gaussian distribution i.e., $\mathbf{u}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_u)$ with the uncertainty covariance matrix $\mathbf{R}_u = \mathrm{E}\{\mathbf{u}_k\mathbf{u}_k^H\} \in \mathbb{C}^{N_t \times N_t}$. The concatenated system model in (1), incorporating the uncertainty model described in (2), can be equivalently obtained as,

$$\mathbf{y}_k = \mathbf{X}(\hat{\mathbf{h}}_k + \mathbf{u}_k) + \boldsymbol{\eta}_k. \tag{3}$$

In the next section we describe robust spectrum sensing schemes, which consider CSI uncertainty, for primary user detection in multiuser MIMO cognitive radio networks, with known uncertainty covariance statistics.

III. Robust Spectrum Sensing With Known Uncertainty Statistics

Let the beacon matrix $\mathbf{X}_i = [\mathbf{x}_i(1), \mathbf{x}_i(2), \ldots, \mathbf{x}_i(L)]^H \in \mathbb{C}^{L \times N_t}$ where $\mathbf{x}_i(n) \in \mathbb{C}^{N_t \times 1}, 1 \leq n \leq L$ denote the beacon signals from the primary user base-station and $i = 0, 1$ correspond to the absence, presence of the primary users respectively. From (3) it follows that the signal $\mathbf{y}_k$ at the kth receive antenna is given as,

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{y}_k = \mathbf{X}_0(\hat{\mathbf{h}}_k + \mathbf{u}_k) + \boldsymbol{\eta}_k\\ \mathcal{H}_1 &: \mathbf{y}_k = \mathbf{X}_1(\hat{\mathbf{h}}_k + \mathbf{u}_k) + \boldsymbol{\eta}_k,\end{aligned} \tag{4}$$

where the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ correspond to the absence and presence of the primary user respectively for the binary hypothesis testing problem towards primary user detection. For example, in a practical wireless scenario, the beacon matrix $\mathbf{X}_0 = \mathbf{0}_{L \times N_t}$ denotes the absence of primary transmission i.e., when the spectral band is vacant. This non-antipodal signaling model is well suited to the context of a cognitive radio scenario where the non-zero beacon matrix $\mathbf{X}_1$ denotes the presence of the primary user signal while the all zero beacon matrix $\mathbf{X}_0$ denotes the absence of the primary user signal. We now present the RECD scheme for robust spectrum sensing in MIMO cognitive radio scenarios.

A. Robust Estimator-Correlator Detection (RECD)

The observation vectors $\mathbf{y}_k$ in (4), corresponding to the hypotheses $\mathcal{H}_0, \mathcal{H}_1$ are distributed as,

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{y}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\eta)\\ \mathcal{H}_1 &: \mathbf{y}_k \sim \mathcal{CN}(\mathbf{X}_1\hat{\mathbf{h}}_k, \boldsymbol{\Gamma}),\end{aligned}$$

where $\boldsymbol{\Gamma} = \mathbf{X}_1\mathbf{R}_u\mathbf{X}_1^H + \mathbf{R}_\eta \in \mathbb{C}^{L \times L}$ and $\mathbf{R}_u \in \mathbb{C}^{N_t \times N_t}$ is the uncertainty covariance matrix. Let the concatenated observation matrix $\mathbf{Y}$ corresponding to the N_r receive antennas be defined as $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_r}] \in \mathbb{C}^{L \times N_r}$. The likelihood ratio $L(\mathbf{Y})$ corresponding to the concatenated observation matrix $\mathbf{Y}$ can be expressed as,

$$\begin{aligned}L(\mathbf{Y}) &= \log\left(\frac{\prod_{\mathrm{k}=1}^{\mathrm{N_r}} \mathrm{p}(\mathbf{y}_{\mathrm{k}}; \mathcal{H}_1)}{\prod_{\mathrm{k}=1}^{\mathrm{N_r}} \mathrm{p}(\mathbf{y}_{\mathrm{k}}; \mathcal{H}_0)}\right),\\ &= \sum_{k=1}^{N_r}\left(\mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{y}_k - (\mathbf{y}_k - \mathbf{X}_1\hat{\mathbf{h}}_k)^H\boldsymbol{\Gamma}^{-1}(\mathbf{y}_k - \mathbf{X}_1\hat{\mathbf{h}}_k)\right) + c\\ &\doteq \sum_{k=1}^{N_r}\mathbf{y}_k^H\left(\mathbf{R}_\eta^{-1} - \hat{\mathbf{h}}_k^H\mathbf{X}_1^H\boldsymbol{\Gamma}^{-1}\mathbf{X}_1\hat{\mathbf{h}}_k\right)\mathbf{y}_k + 2\mathbf{y}_k^H\boldsymbol{\Gamma}^{-1}\mathbf{X}_1\hat{\mathbf{h}}_k,\end{aligned} \tag{5}$$

where the constant $c = N_r\log(\frac{|\mathbf{R}_\eta|}{|\boldsymbol{\Gamma}|})$ and the operator $\doteq$ represents an equivalence to a constant factor. Using the matrix inversion identity from [27] in (5), the likelihood ratio $L(\mathbf{Y})$ can be simplified to obtain the robust estimator-correlator detector (RECD) test statistic,

$$T_{\mathrm{RECD}}(\mathbf{Y}) = \sum_{k=1}^{N_r}\left(2\mathbf{y}_k^H\boldsymbol{\Gamma}^{-1}\mathbf{X}_1\hat{\mathbf{h}}_k + \mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{X}_1\mathbf{R}_u\mathbf{X}_1^H\boldsymbol{\Gamma}^{-1}\mathbf{y}_k\right). \tag{6}$$

Therefore the optimal Neyman-Pearson criterion based LRT, which maximizes the probability of detection P_D for a given probability of false alarm P_{FA} can be obtained as,

$$T_{\mathrm{RECD}}(\mathbf{Y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{7}$$

where the presence/absence of the primary user signal is determined depending on whether the test statistic $T_{\text{RECD}}(\mathbf{Y})$ exceeds or falls short of the detection threshold γ. The proposed RECD exploits the statistical properties of the CSI uncertainty, thereby leading to an improvement in the accuracy of primary user detection. Consider the specific case when the CSI uncertainty covariance matrix $\mathbf{R}_u$ is given as $\mathbf{R}_u = \sigma_u^2 \mathbf{I}_{N_t}$ and the beacon matrix $\mathbf{X}_1$ corresponding to the alternative hypothesis is an orthogonal matrix with $\mathbf{X}_1\mathbf{X}_1^H = L\mathbf{I}$. Under these assumptions, the matrix $\mathbf{\Gamma}$ in (5) is seen to be given as, $\mathbf{\Gamma} = \sigma_\gamma^2 \mathbf{I}$ with $\sigma_\gamma^2 = L\sigma_u^2 + \sigma_\eta^2$. Similarly, let the matrix $\mathbf{V} = \mathbf{R}_\eta^{-1}\mathbf{X}_1\mathbf{R}_u\mathbf{X}_1^H$. Therefore $\mathbf{V} = \sigma_v^2\mathbf{I}$ with $\sigma_v^2 = L\sigma_u^2\sigma_\eta^{-2}$. The Lemma below characterizes the asymptotic P_D versus P_{FA} performance of the RECD scheme described above for spectrum sensing.

Lemma 1: The probabilities of detection and false alarm, denoted by P_D and P_{FA} respectively, for the RECD detector in (7) towards robust spectrum sensing for a large number of receive antennas N_r are given as,

$$P_D = Q\left(\frac{\frac{\gamma'}{\sigma_\gamma^2} - (\theta + 2LN_r)}{\sqrt{2(\theta + 2LN_r)}}\right) \tag{8}$$

$$P_{FA} = Q\left(\frac{\frac{\gamma'}{\sigma_\eta^2} - 2LN_r}{\sqrt{4LN_r}}\right), \tag{9}$$

where $\theta = \sum_{k=1}^{N_r}\sum_{l=1}^{L} \sigma_\gamma^{-2}(1+\sigma_v^{-2})^2|\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k|^2$.

Proof: It can be seen that the test statistic $T_{\text{RECD}}(\mathbf{Y})$ in (5) can be equivalently written as,

$$\begin{aligned}
&T_{\text{RECD}}(\mathbf{Y})\\
&= \sum_{k=1}^{N_r}\left(2\mathbf{y}_k^H\mathbf{\Gamma}^{-1}\mathbf{X}_1\hat{\mathbf{h}}_k + \mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{X}_1\mathbf{R}_u\mathbf{X}_1^H\mathbf{\Gamma}^{-1}\mathbf{y}_k\right)\\
&= \sum_{k=1}^{N_r}\left(2\mathbf{y}_k^H(\sigma_\gamma^{-2}\mathbf{I})\mathbf{X}_1\hat{\mathbf{h}}_k + \mathbf{y}_k^H(\sigma_v^2\sigma_\gamma^{-2})\mathbf{I}\mathbf{y}_k\right)\\
&= \frac{\sigma_v^2}{\sigma_\gamma^2}\sum_{k=1}^{N_r}\sum_{l=1}^{L}\Big(\left(y_{kl}+\sigma_v^{-2}\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k\right)^*\left(y_{kl}+\sigma_v^{-2}\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k\right)\\
&\qquad - \sigma_v^{-2}\sigma_\gamma^2\left|\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k\right|^2\Big)\\
&\doteq \sum_{k=1}^{N_r}\sum_{l=1}^{L}\underbrace{\left(y_{kl}+\sigma_v^{-2}\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k\right)^*\left(y_{kl}+\sigma_v^{-2}\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k\right)}_{T_{\text{RECD}}(y_{kl})},
\end{aligned} \tag{10}$$

where y_{kl} is the lth component of the vector $\mathbf{y}_k$ defined as $\mathbf{y}_k = [y_{k1}, \ldots, y_{kl}, \ldots, y_{kL}]^T$. The component test statistic $T_{\text{RECD}}(y_{kl})$ defined above follows a chi-squared distribution with two degrees of freedom, described as,

$$\begin{aligned}
\mathcal{H}_0 &: \frac{1}{\sigma_\eta^2}T_{\text{RECD}}(y_{kl}) \sim \chi_2^2\\
\mathcal{H}_1 &: \frac{1}{\sigma_\gamma^2}T_{\text{RECD}}(y_{kl}) \sim \chi_2'^2(\theta_l),
\end{aligned}$$

respectively. The quantity $\chi_2'^2(\theta_l)$ denotes the non-central chi-squared distribution with two degrees of freedom, and non-centrality parameter $\theta_l = \sigma_\gamma^{-2}(1+\sigma_v^{-2})^2|\mathbf{x}_1^H(l)\hat{\mathbf{h}}_k|^2$. The distributions of the scaled test statistic $T_{\text{RECD}}(\mathbf{Y})$ for the N_r received vectors stacked as $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_r}]$ corresponding to the two hypotheses $\mathcal{H}_0, \mathcal{H}_1$ can be equivalently derived as,

$$\begin{aligned}
\mathcal{H}_0 &: \frac{1}{\sigma_\eta^2}T_{\text{RECD}}(\mathbf{Y}) = \sum_{k=1}^{N_r}\sum_{l=1}^{L}\frac{1}{\sigma_\eta^2}T_{\text{RECD}}(y_{kl}) \sim \chi_{2LN_r}^2\\
\mathcal{H}_1 &: \frac{1}{\sigma_\gamma^2}T_{\text{RECD}}(\mathbf{Y}) = \sum_{k=1}^{N_r}\sum_{l=1}^{L}\frac{1}{\sigma_\gamma^2}T_{\text{RECD}}(y_{kl}) \sim \chi_{2LN_r}'^2(\theta),
\end{aligned} \tag{11}$$

where the non-centrality parameter $\theta = \sum_{k=1}^{N_r}\sum_{l=1}^{L}\theta_l$. The chi-squared distribution corresponding to the two hypotheses in (11) has $2LN_r$ degrees of freedom. Using the asymptotic property of the chi-squared random variable for a large number of receive antennas N_r [6], the moments of the scaled test statistic $T_{\text{RECD}}(\mathbf{Y})$ corresponding to the two hypotheses in (11) can be equivalently obtained as,

$$\begin{aligned}
\mathrm{E}\left\{\frac{1}{\sigma_\eta^2}T_{\text{RECD}}(\mathbf{Y}); \mathcal{H}_0\right\} &= 2LN_r\\
\mathrm{var}\left\{\frac{1}{\sigma_\eta^2}T_{\text{RECD}}(\mathbf{Y}); \mathcal{H}_0\right\} &= 4LN_r\\
\mathrm{E}\left\{\frac{1}{\sigma_\gamma^2}T_{\text{RECD}}(\mathbf{Y}); \mathcal{H}_1\right\} &= \theta + 2LN_r\\
\mathrm{var}\left\{\frac{1}{\sigma_\gamma^2}T_{\text{RECD}}(\mathbf{Y}); \mathcal{H}_1\right\} &= 2(\theta + 2LN_r).
\end{aligned} \tag{12}$$

Using the results in (12), the analytical expression for the detection probability P_D of the proposed RECD can be equivalently obtained as,

$$\begin{aligned}
P_D &= Pr\left\{T_{\text{RECD}}(\mathbf{Y}) > \gamma'; \mathcal{H}_1\right\}\\
&= Pr\left\{\frac{1}{\sigma_\gamma^2}T_{\text{RECD}}(\mathbf{Y}) > \frac{\gamma'}{\sigma_\gamma^2}; \mathcal{H}_1\right\}\\
&= Q\left(\frac{\frac{\gamma'}{\sigma_\gamma^2} - (\theta + 2LN_r)}{\sqrt{2(\theta + 2LN_r)}}\right),
\end{aligned}$$

where γ' is the detection threshold and $Q(\cdot)$ denotes the standard Gaussian Q-function [6], [28]. Similarly, the probability of false alarm P_{FA} can be obtained as,

$$\begin{aligned}
P_{FA} &= Pr\left\{T_{\text{RECD}}(\mathbf{Y}) > \gamma'; \mathcal{H}_0\right\}\\
&= Pr\left\{\frac{1}{\sigma_\eta^2}T_{\text{RECD}}(\mathbf{Y}) > \frac{\gamma'}{\sigma_\eta^2}; \mathcal{H}_0\right\}\\
&= Q\left(\frac{\frac{\gamma'}{\sigma_\eta^2} - 2LN_r}{\sqrt{4LN_r}}\right).
\end{aligned}$$

□

B. Robust Generalized Likelihood Detector (RGLD)

We now employ the generalized likelihood ratio test (GLRT) paradigm to develop the robust generalized likelihood detector (RGLD) for primary user detection in MIMO cognitive radio networks with CSI uncertainty. Let the vector $\mathbf{r}_k$ be defined as $\mathbf{r}_k = \mathbf{y}_k - \mathbf{X}_1\hat{\mathbf{h}}_k$ corresponding to the alternative hypothesis $\mathcal{H}_1$. Thus, the signal described in (3) corresponding to the L

concatenated sensed samples at the kth receive antenna of the total N_r receive antennas can be equivalently written as,

$$\mathbf{r}_k = \mathbf{X}_1\mathbf{u}_k + \boldsymbol{\eta}_k, \\ = \underbrace{\left[\mathbf{X}_1 \quad \mathbf{I}_L\right]}_{\mathbf{A}} \underbrace{\begin{bmatrix}\mathbf{u}_k \\ \boldsymbol{\eta}_k\end{bmatrix}}_{\mathbf{z}_k}, \tag{13}$$

where $\mathbf{z}_k = [\mathbf{u}_k^T, \boldsymbol{\eta}_k^T]^T \in \mathbb{C}^{(L+N_t)\times 1}$ denotes the concatenated unknown random vector and the identity matrix $\mathbf{I}_L$ is of dimension $L \times L$.

Lemma 2: The GLRT-based test statistic $T_{\text{GLRT}}(\mathbf{Y})$ corresponding to the uncertainty covariance matrix $\mathbf{R}_u$, is given as,

$$T_{\text{RGLD}}(\mathbf{Y}) = \sum_{k=1}^{N_r} -\hat{\mathbf{z}}_k^H \mathbf{R}_z^{-1}\hat{\mathbf{z}}_k + \mathbf{y}_k^H \mathbf{R}_\eta^{-1}\mathbf{y}_k, \tag{14}$$

where $\mathbf{z}_k = [\mathbf{u}_k^T, \boldsymbol{\eta}_k^T]^T$ and $\mathbf{R}_z \in \mathbb{C}^{L+N_t\times L+N_t}$ is the block diagonal matrix with $\mathbf{R}_u, \mathbf{R}_\eta$ along its principal diagonal.

Proof: The likelihood of the observation vector $\mathbf{y}_k$ parameterized by $\mathbf{z}_k$ corresponding to the alternative hypothesis $\mathcal{H}_1$ can be derived as,

$$p(\mathbf{y}_k; \mathbf{z}_k, \mathcal{H}_1) = \frac{1}{\pi^{L+N_t}\,|\mathbf{R}_z|} \exp\left(-\mathbf{z}_k^H \mathbf{R}_z^{-1}\mathbf{z}_k\right), \tag{15}$$

where $\mathbf{R}_z \in \mathbb{C}^{L+N_t\times L+N_t}$ denotes the covariance matrix of $\mathbf{z}_k$, given as,

$$\mathbf{R}_z = \mathrm{E}\left\{\mathbf{z}_k\mathbf{z}_k^H\right\} = \begin{bmatrix}\mathbf{R}_u & \mathbf{0}_{N_t\times L} \\ \mathbf{0}_{L\times N_t} & \mathbf{R}_\eta\end{bmatrix},$$

and $|\mathbf{R}_z|$ denotes the determinant of the matrix $\mathbf{R}_z$. The parameterized vector $\mathbf{z}_k$ is obtained on maximizing the likelihood $p(\mathbf{y}_k; \mathbf{z}_k, \mathcal{H}_1)$ in (15) which can be formulated as the standard weighted minimum L_2 norm optimization problem [27],

$$\begin{aligned} \min.\quad & \mathbf{z}_k^H \mathbf{R}_z^{-1}\mathbf{z}_k \\ \text{s.t.}\quad & \mathbf{r}_k = \mathbf{A}\mathbf{z}_k. \end{aligned}$$

The solution of the above convex optimization problem yields the estimate of $\hat{\mathbf{z}}_k$ as,

$$\hat{\mathbf{z}}_k = \mathbf{R}_z\mathbf{A}^H(\mathbf{A}\mathbf{R}_z\mathbf{A}^H)^{-1}\mathbf{r}_k.$$

Employing the GLRT framework, the RGLD test statistic $T_{\text{RGLD}}(\mathbf{Y})$ for the primary user detection problem in cognitive radio scenarios can be derived as,

$$\begin{aligned} T_{\text{RGLD}}(\mathbf{Y}) &= \log\left(\frac{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \hat{\mathbf{z}}_k, \mathcal{H}_1)}{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \mathcal{H}_0)}\right) \\ &= \log\left(\frac{\prod_{k=1}^{N_r} \frac{1}{\pi^{L+N_t}|\mathbf{R}_z|}\exp(-\mathbf{z}_k^H\mathbf{R}_z^{-1}\mathbf{z}_k)}{\prod_{k=1}^{N_r} \frac{1}{\pi^{L+N_t}|\mathbf{R}_\eta|}\exp(-\mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{y}_k)}\right) \\ &\doteq \sum_{k=1}^{N_r} -\hat{\mathbf{z}}_k^H\mathbf{R}_z^{-1}\hat{\mathbf{z}}_k + \mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{y}_k, \end{aligned}$$

where the operator $\doteq$ denotes equivalence to a constant factor. This completes the proof. □

The inherent ability of the proposed RECD and RGLD schemes to exploit the CSI uncertainty for primary user detection in spectrum sensing scenarios provides a performance edge over the conventional uncertainty agnostic matched filter detector. In the next section we further relax the CSI uncertainty model and consider the uncertainty covariance statistics to be unknown. We then derive the corresponding composite hypothesis based robust detector (CHRD) for spectrum sensing in MIMO cognitive radio scenarios.

IV. Robust Detection With Unknown Uncertainty Covariance Statistics

A. Composite Hypothesis Based Robust Detector (CHRD)

In this section we compute a generalized likelihood ratio test based composite hypothesis testing framework with unknown uncertainty statistics. The primary user detection problem from (1) can be equivalently recast as the hypothesis testing problem,

$$\begin{aligned} &\mathcal{H}_0\colon \mathbf{I}_{N_t}\mathbf{h}_k = 0 \\ &\mathcal{H}_1\colon \mathbf{I}_{N_t}\mathbf{h}_k \neq 0, \end{aligned}$$

where $\mathbf{I}_{N_t}$ is the $N_t \times N_t$ identity matrix and the vector parameters $\mathbf{h}_k, 1 \le k \le N_r$ form the rows of the channel matrix $\mathbf{H}$. Let the vectors $\hat{\mathbf{h}}_{k|0}, \hat{\mathbf{h}}_{k|1}$ denote the maximum likelihood estimate (MLE) of the vector parameter $\mathbf{h}_k$ corresponding to the null hypothesis $\mathcal{H}_0$, alternative hypothesis $\mathcal{H}_1$ respectively, which can be obtained [6] as,

$$\begin{aligned} \hat{\mathbf{h}}_{k|1} &= \left(\mathbf{X}_1^H\mathbf{X}_1\right)^{-1}\mathbf{X}_1^H\mathbf{y}_k \\ \hat{\mathbf{h}}_{k|0} &= \hat{\mathbf{h}}_{k|1} - \left(\mathbf{X}_1^H\mathbf{X}_1\right)^{-1}\mathbf{I}^H\left[\mathbf{I}\left(\mathbf{X}_1^H\mathbf{X}_1\right)^{-1}\mathbf{I}^H\right]^{-1}\left(\mathbf{I}_{N_t}\hat{\mathbf{h}}_{k|1}\right) \\ &= \mathbf{0}. \end{aligned} \tag{16}$$

The concatenated received vector $\mathbf{y}_k$, defined in (1), follows a complex Gaussian distribution given as,

$$\begin{aligned} &\mathcal{H}_0\colon \mathbf{y}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\eta) \\ &\mathcal{H}_1\colon \mathbf{y}_k \sim \mathcal{CN}(\mathbf{X}_1\hat{\mathbf{h}}_{k|1}, \mathbf{R}_\eta), \end{aligned}$$

corresponding to the hypotheses $\mathcal{H}_0, \mathcal{H}_1$. Let the matrix $\mathbf{Y}$ be obtained by stacking the received signal vectors $\mathbf{y}_k$, as $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots \mathbf{y}_{N_r}]$. Hence, the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ for the composite hypothesis testing based primary user detection problem can be computed by applying the generalized likelihood ratio test (GLRT) as,

$$\begin{aligned} &T_{\text{CHRD}}(\mathbf{Y}) \\ &= \log\left(\frac{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \hat{\mathbf{h}}_{k|1}, \mathcal{H}_1)}{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \mathcal{H}_0)}\right) \qquad (17) \\ &= \log\left(\frac{\prod_{k=1}^{N_r} \exp\left(-(\mathbf{y}_k - \mathbf{X}_1\hat{\mathbf{h}}_{k|1})^H\mathbf{R}_\eta^{-1}(\mathbf{y}_k - \mathbf{X}_1\hat{\mathbf{h}}_{k|1})\right)}{\prod_{k=1}^{N_r}\exp(-\mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{y}_k)}\right) \\ &= \sum_{k=1}^{N_r} \hat{\mathbf{h}}_{k|1}^H\mathbf{X}_1^H\mathbf{R}_\eta^{-1}\mathbf{X}_1\hat{\mathbf{h}}_{k|1} \\ &\doteq \sum_{k=1}^{N_r} \hat{\mathbf{h}}_{k|1}^H\hat{\mathbf{h}}_{k|1}, \qquad (18) \end{aligned}$$

where the last equality holds due to the orthogonality property of the beacon matrix, i.e., $\mathbf{X}_1^H\mathbf{X}_1 = L\mathbf{I}_{N_t}$. The test statistic $T_{\text{CHRD}}(\mathbf{Y})$ obtained above yields the primary user detection rule that is robust against the uncertainty in the estimate of the MIMO channel matrix. We now consider a general scenario with non-isotropic CSI uncertainty covariance, i.e., the uncertainty covariance is not necessarily of the form $\mathbf{R}_u = \sigma_u^2\mathbf{I}$. It can be noted that the non-isotropic CSI uncertainty covariance matrix considered in our formulation is more general and practical in comparison to the isotropic model considered in works such as [29], [30]. We now characterize the theoretical performance of the CHRD for composite hypothesis testing based primary user detection in MIMO cognitive radio scenarios. Using (3) and the orthogonality property of the beacon matrix in (16), the maximum likelihood estimate $\hat{\mathbf{h}}_{k|1}$ of the column vectors $\mathbf{h}_k, 1 \leq k \leq N_r$ of the channel matrix $\mathbf{H}$, corresponding to the alternative hypothesis $\mathcal{H}_1$, can be equivalently written as,

$$\begin{aligned}\hat{\mathbf{h}}_{k|1} &= \left(\mathbf{X}_1^H\mathbf{X}_1\right)^{-1}\mathbf{X}_1^H\mathbf{y}_k \\ &= \left(\mathbf{X}_1^H\mathbf{X}_1\right)^{-1}\mathbf{X}_1^H\mathbf{X}_1(\hat{\mathbf{h}}_k + \mathbf{u}_k) + \mathbf{X}_1^H\boldsymbol{\eta}_k \\ &= \hat{\mathbf{h}}_k + \mathbf{u}_k + \mathbf{n}_k, \end{aligned} \tag{19}$$

where $\mathbf{n}_k$ is defined as $\mathbf{n}_k = \mathbf{X}_1^H\boldsymbol{\eta}_k$ with covariance $\mathbf{R}_n = E\{\mathbf{n}_k\mathbf{n}_k^H\} = L\sigma_\eta^2\mathbf{I}_{N_t}$. The CSI estimate $\hat{\mathbf{h}}_k \in \mathbb{C}^{N_t\times 1}$ follows a complex Gaussian distribution $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_t})$. Thus, this implies that the true channel coefficient vector $\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{u}_k$ is also complex Gaussian distributed, leading to a Rayleigh fading wireless channel, which is a standard assumption for fading wireless scenarios. It follows from (19) that the maximum likelihood estimate $\hat{\mathbf{h}}_{k|1}$ has a complex Gaussian distribution, with $\hat{\mathbf{h}}_{k|1} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Lambda})$ where $\boldsymbol{\Lambda} = (1 + L\sigma_\eta^2)\mathbf{I}_{N_t} + \mathbf{R}_u$. Let the eigenvalue decomposition of $\boldsymbol{\Lambda}$ be $\mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^H$, with the eigenvalue matrix $\boldsymbol{\Sigma} = \mathcal{D}([\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N_t}^2]^T)$, where $\mathcal{D}(\mathbf{a})$ denotes a diagonal matrix with the elements of vector $\mathbf{a}$ along the principal diagonal. Therefore, the lth element $\hat{h}_{kl|1}$ of the vector $\hat{\mathbf{h}}_{k|1} = [\hat{h}_{k1|1}^T, \ldots, \hat{h}_{kl|1}^T, \ldots, \hat{h}_{kN_t|1}^T]^T \in \mathbb{C}^{N_t\times 1}$ follows a Gaussian distribution. Let z_{kl} be defined as $z_{kl} = \frac{1}{\sigma_l}\hat{h}_{kl|1}$. The test statistic in (18) can be equivalently described as,

$$\begin{aligned} T_{\text{CHRD}}(\mathbf{Y}) &= \sum_{k=1}^{N_r}\hat{\mathbf{h}}_{k|1}^H(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1})\hat{\mathbf{h}}_{k|1} \\ &= \sum_{k=1}^{N_r}\sum_{l=1}^{N_t}\sigma_l^2\frac{|\hat{h}_{kl|1}|^2}{\sigma_l^2} \\ &= \sum_{k=1}^{N_r}\sum_{l=1}^{N_t}\sigma_l^2|z_{kl}|^2, \end{aligned} \tag{20}$$

where each z_{kl} is Gaussian distributed with $z_{kl} \sim \mathcal{CN}(0, 1)$. Therefore, $|z_{kl}|^2, 1 \leq k \leq N_r, 1 \leq l \leq N_t$ is distributed as a χ_2^2 random variable. Employing the above results, we now characterize the performance of the composite hypothesis detector corresponding to the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ in (20).

Theorem 1: The probability of false alarm P_{FA} and probability of detection P_D of the CHRD detector based on the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ in (20), for primary user detection in MIMO cognitive radio networks can be derived as,

$$P_{FA} = Q_{\chi^2_{2N_rN_t}}\left(\frac{\gamma}{\sigma_\eta^2}\right), \tag{21}$$

$$P_D = \kappa\sum_{l=1}^{L}\sum_{k=1}^{N_r}A_{l,k}\frac{(-2\sigma_l^2)^k}{(k-1)!}\,\Gamma\left(k, \frac{\gamma}{2\sigma_l^2}\right), \tag{22}$$

where the coefficients $A_{l,k}$ are the partial fraction constants obtained using the residue method, $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function defined in (30) and the constant κ is defined as $\kappa = \prod_{m=1}^{N_t}(\frac{-1}{2\sigma_m^2})^{N_r}$.

Proof: The probability of false alarm P_{FA} of the GLRT test statistic $T_{\text{CHRD}}(\mathbf{Y})$ can be derived as,

$$\begin{aligned} P_{FA} &= Pr\left\{T_{\text{CHRD}}(\mathbf{Y}) > \gamma; \mathcal{H}_0\right\} \\ &= Pr\left\{\frac{1}{\sigma_\eta^2}T_{\text{CHRD}}(\mathbf{Y}) > \frac{\gamma}{\sigma_\eta^2}; \mathcal{H}_0\right\} \\ &= Pr\left\{\chi^2_{2N_rN_t} > \frac{\gamma}{\sigma_\eta^2}\right\} \\ &= Q_{\chi^2_{2N_rN_t}}\left(\frac{\gamma}{\sigma_\eta^2}\right). \end{aligned}$$

Next, we derive the probability of detection P_D for the test statistic $T_{\text{CHRD}}(\mathbf{Y})$. This can be expressed as,

$$P_D = Pr\left\{T_{\text{CHRD}}(\mathbf{Y}) > \gamma; \mathcal{H}_1\right\} = \int_\gamma^\infty p_T(t)dt, \tag{23}$$

where $p_T(t)$ denotes the probability density function of the test statistic $T_{\text{CHRD}}(\mathbf{Y})$. Let $\Phi_T(\omega)$ denote the characteristic function of the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ corresponding to the alternative hypothesis $\mathcal{H}_1$. The probability density function $p_T(t)$ can be expressed in terms of $\Phi_T(\omega)$ as,

$$p_T(t) = \begin{cases} \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_T(\omega)\exp^{-j\omega t}d\omega & t \geq 0 \\ 0 & t < 0. \end{cases} \tag{24}$$

The characteristic function $\Phi_T(\omega)$ for the CHRD test statistic corresponding to the alternative hypothesis $\mathcal{H}_1$ is derived as,

$$\begin{aligned} \Phi_T(\omega) &= E\left\{\exp(j\omega T_{\text{CHRD}}(\mathbf{Y}))\right\} \\ &= E\left\{\exp\left(j\omega\sum_{l=1}^{N_t}\sum_{k=1}^{N_r}\sigma_l^2 z_{kl}^\star z_{kl}\right)\right\} \qquad (25) \\ &= \prod_{l=1}^{N_t}\prod_{k=1}^{N_r}E\left\{\exp(j\omega\sigma_l^2|z_{kl}|^2)\right\} \\ &= \prod_{l=1}^{N_t}\frac{1}{(1-2j\sigma_l^2\omega)^{N_r}}, \qquad (26) \end{aligned}$$

where (25) follows from the simplification of the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ described in (20). The characteristic function $\Phi_T(\omega)$ obtained in (26) follows from the fact that each $|z_{kl}|^2$ follows a chi-squared distribution with 2 degrees of freedom as described earlier. Employing (26) above, the probability density function $p_T(t)$ of the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ defined in (24) can be equivalently expressed as,

$$\begin{aligned} p_T(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\prod_{l=1}^{N_t}\frac{1}{(1-2j\sigma_l^2\omega)^{N_r}}\exp(-j\omega t)\,d\omega \\ &= \prod_{m=1}^{N_t}\left(\frac{-1}{2\sigma_m^2}\right)^{N_r}\mathcal{F}_{-t}^{-1}\left\{\underbrace{\prod_{l=1}^{N_t}\frac{1}{(j\omega-\frac{1}{2\sigma_l^2})^{N_r}}}_{\psi(\omega)}\right\} \\ &= \kappa\,\mathcal{F}_{-t}^{-1}\left\{\sum_{l=1}^{N_t}\sum_{k=1}^{N_r}\frac{A_{l,k}}{(j\omega-\frac{1}{2\sigma_l^2})^k}\right\} \\ &= \kappa\sum_{l=1}^{N_t}\sum_{k=1}^{N_r}A_{l,k}\mathcal{F}_{-t}^{-1}\left\{\frac{1}{(j\omega-\frac{1}{2\sigma_l^2})^k}\right\}, \end{aligned} \tag{27}$$

where the coefficients $A_{l,k}$ correspond to the partial fraction expansion of $\psi(\omega)$ and the constant κ is defined as $\kappa = \prod_{m=1}^{N_t}(\frac{-1}{2\sigma_m^2})^{N_r}$. The coefficients $A_{l,k}$ can be obtained using the standard residue method for partial fraction expansion [31] and can be explicitly expressed as,

$$A_{l,k} = \beta_k \lim_{j\omega = \frac{1}{2\sigma_l^2}} \left\{ \frac{d^{N_r-k}}{dj\omega^{N_r-k}} \left(\frac{\left(j\omega - \frac{1}{2\sigma_l^2}\right)^{N_r}}{\prod_{n=1}^{L}\left(j\omega - \frac{1}{2\sigma_n^2}\right)^{N_r}} \right) \right\}, \quad (28)$$

where $\beta_k = \frac{1}{(N_r-k)!}$. Computing the inverse Fourier transform of the expression in (27), the probability density function $p_T(t)$ of the test statistic $T_{\mathrm{CHRD}}(\mathbf{Y})$ in (20), corresponding to the alternative hypothesis $\mathcal{H}_1$, can be obtained as,

$$p_T(t) = \kappa \sum_{l=1}^{N_t}\sum_{k=1}^{N_r} \frac{A_{l,k}(-1)^k t^{k-1} e^{-\frac{t}{2\sigma_l^2}} u(t)}{(k-1)!}, \quad (29)$$

where $u(t)$ is the unit step function defined as,

$$u(t) = \begin{cases} 1 & t > 0 \\ 0 & t \le 0. \end{cases}$$

Hence, the probability of detection P_D for the composite hypothesis detector can be derived by substituting the above expression for $p_T(t)$ in (23) as,

$$\begin{aligned} P_D &= \int_{\gamma}^{\infty} \kappa \sum_{l=1}^{N_t}\sum_{k=1}^{N_r} \frac{A_{l,k}(-1)^k t^{k-1} e^{-\frac{t}{2\sigma_l^2}} u(t)}{(k-1)!}\, dt \\ &= \kappa \sum_{l=1}^{N_t}\sum_{k=1}^{N_r} \frac{A_{l,k}(-1)^k}{(k-1)!} \int_{\gamma}^{\infty} t^{k-1} e^{-\frac{t}{2\sigma_l^2}} u(t)\, dt \\ &= \kappa \sum_{l=1}^{N_t}\sum_{k=1}^{N_r} A_{l,k} \frac{(-2\sigma_l^2)^k}{(k-1)!} \Gamma\left(k, \frac{\gamma}{2\sigma_l^2}\right), \end{aligned}$$

where γ is the threshold for detection and $\Gamma(.,.)$ denotes the incomplete Gamma function [28], defined as,

$$\Gamma(s,\gamma) = \int_{\gamma}^{\infty} t^{s-1} e^{-t} u(t) dt. \quad (30)$$

□

Further, we now derive the probability of detection P_D for the restrictive case of an isotropic CSI uncertainty covariance matrix, i.e., $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_{N_t}$ with $\sigma_i^2 = \sigma^2, \forall 1 \le i \le N_t$.

Lemma 3: For an isotropic covariance matrix i.e., with $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$, the probability of detection for the detector in (18) reduces to,

$$P_D = Q_{\chi^2_{2N_rN_t}}\left(\frac{\gamma}{\sigma^2}\right). \quad (31)$$

Proof: The characteristic function $\Phi_T(\omega)$ in (26) and the corresponding probability density function $p_T(t)$ of the test statistic $T_{\mathrm{CHRD}}(\mathbf{Y})$ in (20), for the covariance matrix $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$, can be obtained as,

$$\begin{aligned} \Phi_T(\omega) &= \left(\frac{1}{1 - 2j\sigma^2\omega}\right)^{N_rN_t}, \\ p_T(t) &= \kappa_0 \sum_{k=1}^{N_rN_t} A_k \mathcal{F}_{-t}^{-1}\left\{\frac{1}{(j\omega - \frac{1}{2\sigma^2})^k}\right\} \\ &= \kappa_0 \sum_{k=1}^{N_rN_t} \frac{A_k(-1)^k t^{k-1} e^{-\frac{t}{2\sigma^2}} u(t)}{(k-1)!}, \end{aligned} \quad (32)$$

respectively, where the constant κ_0 is defined as $\kappa_0 = (\frac{-1}{2\sigma^2})^{N_rN_t}$. The partial fraction constants A_k in (32) can now be equivalently obtained as,

$$\begin{aligned} A_k &= \beta_k \lim_{j\omega = \frac{1}{2\sigma^2}} \left\{ \frac{d^{N_rN_t-k}}{dj\omega^{N_rN_t-k}} \left(\frac{1}{\left(j\omega - \frac{1}{2\sigma^2}\right)^{N_rN_t}} \right) \right\} \\ &= \begin{cases} 1, & k = N_rN_t \\ 0, & k \neq N_rN_t, \end{cases} \end{aligned} \quad (33)$$

where the constant β_k is defined as $\beta_k = \frac{1}{(N_rN_t-k)!}$. Using the expression for $p_T(t)$ from (32), the corresponding probability of detection P_D for this scenario with a diagonal covariance matrix $\mathbf{R}_u$ can be derived as,

$$P_D = \kappa_0 \sum_{k=1}^{N_rN_t} A_k \frac{(-2\sigma)^k}{(k-1)!} \Gamma\left(k, \frac{\gamma}{2\sigma^2}\right).$$

Substituting the expression for the coefficient A_k from (33), the above expression for P_D can be simplified as,

$$\begin{aligned} P_D &= \frac{1}{(N_rN_t - 1)!} \Gamma\left(N_rN_t, \frac{\gamma}{2\sigma}\right) \\ &= e^{-\frac{\gamma}{2\sigma^2}} \sum_{k=0}^{N_rN_t-1} \frac{1}{k!}\left(\frac{\gamma}{2\sigma^2}\right)^k \\ &= Q_{\chi^2_{2N_rN_t}}\left(\frac{\gamma}{\sigma^2}\right), \end{aligned} \quad (34)$$

where (34) follows from the following property of the Gamma function [28],

$$\Gamma(q,s) = (q-1)!\, e^{-s} \sum_{m=0}^{q-1} \frac{s^m}{m!}.$$

□

In the next section we develop the framework to obtain the optimal beacon sequence which can further enhance the detection performance of the proposed spectrum sensing schemes.

V. Optimal Beacon Formulation

This section presents a deflection coefficient based optimization framework to derive the optimal beacon sequence $\mathbf{X}_1$, which can further improve the primary user detection performance for scenarios with known/unknown CSI uncertainty statistics.

A. Known Uncertainty Covariance $\mathbf{R}_u$

Let the stacked vector $\tilde{\mathbf{r}} \in \mathbb{C}^{LN_r \times 1}$ be defined as $\tilde{\mathbf{r}} = [\mathbf{r}_1^T, \ldots \mathbf{r}_k^T, \ldots, \mathbf{r}_{N_r}^T]^T$ for $1 \le k \le N_r$ corresponding to the N_r receive antennas. The equivalent system model considering this stacked observation vector $\tilde{\mathbf{r}}$ can be derived as,

$$\tilde{\mathbf{r}} = \underbrace{(\mathbf{I}_{N_r} \otimes \mathbf{X}_1)}_{\tilde{\mathbf{X}}} \mathrm{vec}(\mathbf{U}^H) + \tilde{\boldsymbol{\eta}}, \quad (35)$$

where $\tilde{\mathbf{X}} = (\mathbf{I}_{N_r} \otimes \mathbf{X}_1) \in \mathbb{C}^{LN_r \times N_tN_r}$, the identity matrix $\mathbf{I}_{N_r}$ has dimensions $N_r \times N_r$ and $\otimes$ denotes the matrix Kronecker product. The vector $\mathrm{vec}(\mathbf{U}^H) \in \mathbb{C}^{N_rN_t \times 1}$ is the column vector obtained by stacking the columns of the uncertainty matrix $\mathbf{U}^H$ and has the covariance matrix $\mathbf{R}_U = \mathrm{E}\{\mathrm{vec}(\mathbf{U}^H)(\mathrm{vec}(\mathbf{U}^H))^H\} = \mathbf{I}_{N_r} \otimes \mathbf{R}_u \in \mathbb{C}^{N_rN_t \times N_rN_t}$. Similarly, the concatenated noise vector $\tilde{\boldsymbol{\eta}}$ is obtained as

$\tilde{\boldsymbol{\eta}} = [\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \ldots, \boldsymbol{\eta}_{N_r}^T]^T$ with the noise covariance matrix $\mathbf{R}_{\tilde{\eta}} = \mathrm{E}\{\tilde{\boldsymbol{\eta}}\tilde{\boldsymbol{\eta}}^H\} = \mathbf{I}_{N_r} \otimes \mathbf{R}_\eta \in \mathbb{C}^{N_r L \times N_r L}$. The result below derives the optimal beacon matrix $\mathbf{X}_1$ for this scenario.

Theorem 2: The optimal beacon matrix $\mathbf{X}_1$ for primary user detection towards spectrum sensing, for a known uncertainty covariance $\mathbf{R}_u$, can be obtained as a solution of the optimization problem,

$$\begin{aligned} \max. \quad & \mathrm{Tr}(\mathbf{X}_1 \mathbf{W} \mathbf{X}_1^H) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{X}_1 \mathbf{X}_1^H) \leq P_0, \end{aligned} \tag{36}$$

where $\mathbf{W} = \sum_{k=1}^{N_r} (\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H) - \mu N_r \mathbf{R}_u \in \mathbb{C}^{N_t \times N_t}$, the quantity μ is an appropriate non-negative constant and P_0 denotes the total transmit beacon power.

Proof: To obtain the optimal beacon sequence, we employ the deflection coefficient $d^2(\mathbf{X}_1)$ for the binary hypothesis testing based primary user detection problem [6], defined as,

$$\begin{aligned} d^2(\mathbf{X}_1) &\triangleq \frac{\|\mathrm{E}\{\mathbf{Y}; \mathcal{H}_1\} - \mathrm{E}\{\mathbf{Y}; \mathcal{H}_0\}\|_2^2}{\mathrm{Tr}(\mathrm{cov}\{\mathbf{Y}; \mathcal{H}_1\})}, \\ &= \frac{\left\| \tilde{\mathbf{X}} \,\mathrm{vec}\,(\hat{\mathbf{H}}^H) \right\|_2^2}{\mathrm{Tr}\left(\tilde{\mathbf{X}} \mathbf{R}_U \tilde{\mathbf{X}}^H + \mathbf{R}_{\tilde{\eta}}\right)}, \\ &= \frac{1}{N_r} \frac{\sum_{k=1}^{N_r} \left\| \mathbf{X}_1 \hat{\mathbf{h}}_k \right\|_2^2}{\mathrm{Tr}(\mathbf{X}_1 \mathbf{R}_u \mathbf{X}_1^H + \mathbf{R}_\eta)}, \end{aligned} \tag{37}$$

where $\mathrm{E}\{\mathbf{Y}; \mathcal{H}_0\}$, $\mathrm{E}\{\mathbf{Y}; \mathcal{H}_1\}$ denote the expected values of the observation vector $\mathbf{Y}$ under the null hypothesis $\mathcal{H}_0$, the alternative hypothesis $\mathcal{H}_1$ respectively and $\mathrm{cov}\{\mathbf{Y}; \mathcal{H}_1\}$ denotes the covariance of $\mathbf{Y}$ under the alternative hypothesis $\mathcal{H}_1$. However, the direct optimization of the deflection coefficient above is intractable since it is non-convex. Therefore, we consider a simplified convex problem of maximizing the weighted difference of the distance between the hypothesis means and the trace of the covariance. This yields a tractable problem which can be solved to yield the closed form expressions for the optimal beacon matrices as shown below. Hence, the deflection coefficient based optimization framework to obtain the optimal beacon matrix $\mathbf{X}_1$ for a scenario with known uncertainty covariance, towards maximization of the primary user detection performance, can be formulated as,

$$\begin{aligned} \max. \quad & \sum_{k=1}^{N_r} \left\| \mathbf{X}_1 \hat{\mathbf{h}}_k \right\|_2^2 - \mu(N_r) \mathrm{Tr}\left(\mathbf{X}_1 \mathbf{R}_u \mathbf{X}_1^H + \mathbf{R}_\eta\right) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\mathbf{X}_1 \mathbf{X}_1^H\right) \leq P_0. \end{aligned} \tag{38}$$

It can be observed that the optimization framework above is a bi-criterion optimization problem [32] where the choice of μ allows for a tradeoff between the uncertainty variance and the separation between the vectors corresponding to the two hypotheses. The optimization framework in (38) above can be equivalently reduced to,

$$\begin{aligned} \max. \quad & \mathrm{Tr}(\mathbf{X}_1 \mathbf{W} \mathbf{X}_1^H) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{X}_1 \mathbf{X}_1^H) \leq P_0, \end{aligned}$$

where $\mathbf{W} = \sum_{k=1}^{N_r} (\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H) - \mu N_r \mathbf{R}_u$. The solution to the above optimization problem is given by the beacon vectors $\mathbf{x}_1(i), 1 \leq i \leq L$, defined as $\mathbf{x}_1(i) = \sqrt{\frac{P_0}{L}} \nu_{max}(\mathbf{W})$, where $\nu_{max}(\mathbf{W})$ denotes the unit-norm principal eigenvector of the matrix $\mathbf{W}$. □

B. Unknown Uncertainty Covariance $\mathbf{R}_u$

Let $\mathbf{q}_k \in \mathbb{C}^{L \times 1}$ be defined as $\mathbf{q}_k = \mathbf{y}_k - \mathbf{X}_1 \hat{\mathbf{h}}_{k|1}$ where $\hat{\mathbf{h}}_{k|1}$ is the MLE of $\mathbf{h}$ corresponding to the alternative hypothesis $\mathcal{H}_1$ derived in (16). Hence, the equivalent system model for the spectrum sensing scenario with the concatenated vector $\tilde{\mathbf{q}} = [\mathbf{q}_1^T, \mathbf{q}_2^T, \ldots, \mathbf{q}_{N_r}^T]^T \in \mathbb{C}^{LN_r \times 1}$ for the N_r receive antennas can be equivalently written as,

$$\tilde{\mathbf{q}} = \underbrace{(\mathbf{I}_{N_r} \otimes \mathbf{X}_1)}_{\tilde{\mathbf{X}}} \mathrm{vec}(\mathbf{U}^H) + \tilde{\boldsymbol{\eta}},$$

where the matrix $\tilde{\mathbf{X}}$ and the concatenated noise vector $\tilde{\boldsymbol{\eta}}$ are as defined in (35).

Lemma 4: The optimal beacon matrix $\mathbf{X}_1$ for a CHRD based robust spectrum sensing scenario with unknown CSI statistics can be obtained as the solution of the optimization problem,

$$\begin{aligned} \max. \quad & \mathrm{Tr}\left(\mathbf{X}_1 \mathbf{Z} \mathbf{X}_1^H\right) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\mathbf{X}_1 \mathbf{X}_1^H\right) \leq P_0, \end{aligned} \tag{39}$$

where $\mathbf{Z} = \sum_{k=1}^{N_r} (\hat{\mathbf{h}}_{k|1} \hat{\mathbf{h}}_{k|1}^H) \in \mathbb{C}^{N_t \times N_t}$.

Proof: Similar to the procedure in (37), the deflection coefficient $d_{MLE}^2(\mathbf{X}_1)$ for the composite hypothesis based primary user detection problem with an unknown uncertainty covariance matrix, can be determined as,

$$d_{MLE}^2(\mathbf{X}_1) = \frac{\left\| \tilde{\mathbf{X}} \,\mathrm{vec}\,(\hat{\mathbf{H}}_{MLE}^H) \right\|_2^2}{\mathrm{Tr}(\mathbf{R}_{\tilde{\eta}})} = \frac{\sum_{k=1}^{N_r} \left\| \mathbf{X}_1 \hat{\mathbf{h}}_{k|1} \right\|_2^2}{N N_r \mathrm{Tr}(\mathbf{R}_\eta)}, \tag{40}$$

where $\hat{\mathbf{H}}_{MLE} \in \mathbb{C}^{N_r \times N_t}$ is defined as $\hat{\mathbf{H}}_{MLE} = [\hat{\mathbf{h}}_{1|1}, \ldots, \hat{\mathbf{h}}_{k|1}, \ldots, \hat{\mathbf{h}}_{N_r|1}]^H$. The optimization problem to obtain the optimal beacon matrix $\mathbf{X}_1$ that maximizes the performance of the proposed detection scheme can be readily derived as,

$$\begin{aligned} \max. \quad & \mathrm{Tr}\left(\mathbf{X}_1 \mathbf{Z} \mathbf{X}_1^H\right) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\mathbf{X}_1 \mathbf{X}_1^H\right) \leq P_0, \end{aligned}$$

where $\mathbf{Z} = \sum_{k=1}^{N_r} (\hat{\mathbf{h}}_{k|1} \hat{\mathbf{h}}_{k|1}^H)$. The above optimization problem is a quadratic constrained quadratic program (QCQP) [32] which can be solved by aligning each beacon vector as $\mathbf{x}_1(i) = \sqrt{\frac{P_0}{L}} \nu_{max}(\mathbf{Z})$, $1 \leq i \leq L$, along the principal unit-norm eigenvector $\nu_{\max}(\mathbf{Z})$ of the matrix $\mathbf{Z}$ corresponding to the largest eigenvalue. This yields the optimal beacon matrix $\mathbf{X}_1$ for the MIMO cognitive radio spectrum sensing scenario with unknown CSI statistics. □

VI. Simulation Results

This section presents simulation results to illustrate the performance of the spectrum sensing schemes proposed above for MIMO cognitive radio scenarios. We consider a system where the secondary user has $N_r = 2$ receive antennas and the primary user base-station has $N_t = 2$ transmit antennas. In Figs. 1–7, we consider a non-antipodal signaling system with the beacon matrix $\mathbf{X}_0 \in \mathbb{C}^{L \times 2}$ corresponding to the null hypothesis set as $\mathbf{X}_0 = \mathbf{0}$ and the beacon matrix $\mathbf{X}_1 \in \mathbb{C}^{2 \times L}$ corresponding

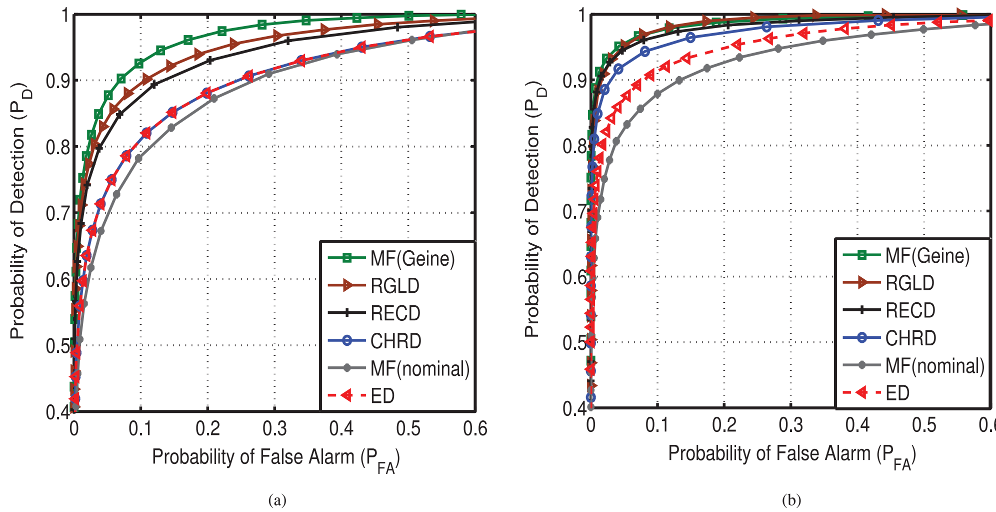


Fig. 1. Receiver operating characteristic (ROC) curves for the genie aided matched filter detector (MF genie), robust generalized likelihood detector (RGLD), robust estimator-correlator detector (RECD), composite hypothesis based robust detector (CHRD), energy detector (ED) in (43) and nominal estimate based matched filter detector (MF) in (42) with SNR $= -3$ dB and $\sigma_u^2 = 1$ for (a) $L = 2$, (b) $L = 4$. (a) Case I, (b) Case II.

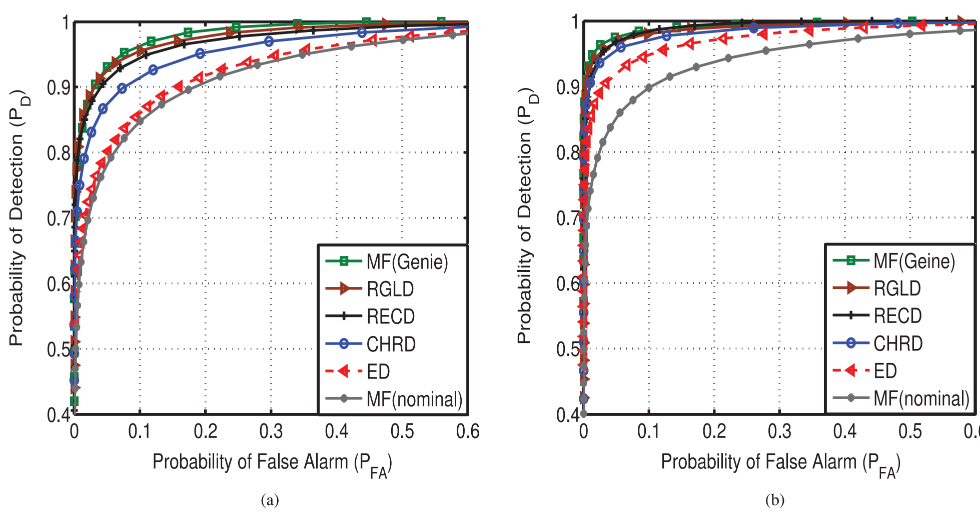


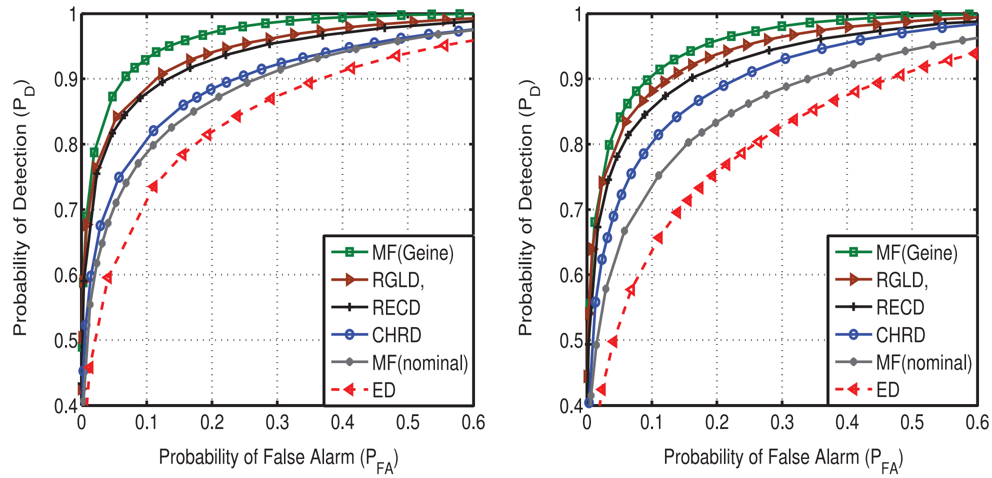
Fig. 2. ROC curves for the genie aided matched filter detector (MF genie), robust generalized likelihood detector (RGLD), robust estimator-correlator detector (RECD), composite hypothesis based robust detector (CHRD), energy detector (ED) in (43) and nominal estimate based matched filter detector (MF) in (42) with $L = 4$, $\sigma_u^2 = 1$ for (a) SNR $= -4$ dB, (b) SNR $= -2$ dB. (a) Case I, (b) Case II.

to the alternative hypothesis set as an orthogonal beacon matrix, i.e., satisfying $\mathbf{X}_1^H\mathbf{X}_1 = L\mathbf{I}_{2\times 2}$. In our simulations we consider different levels of CSI uncertainty with $\sigma_u^2 \in \{1, 0.8, 0.6\}$ in the CSI uncertainty covariance matrix $\mathbf{R}_u = \sigma_u^2\mathcal{D}([1, 0.9]^T)$. We present the probability of detection P_D versus the probability of false alarm P_{FA} for the proposed RECD, RGLD and CHRD robust MIMO spectrum sensing schemes towards primary user detection. The performance of the uncertainty agnostic matched filter detector corresponding to the nominal CSI estimate is also given in the figures, similar to the comparison between the robust and the non-robust techniques in [30], [33]. The proposed schemes are also compared with the energy detector (ED). Additionally, in our simulations we present comparisons of the proposed robust detection schemes with the genie aided matched filter detector (MF genie) with perfect knowledge of the CSI uncertainty, i.e., with knowledge of the true MIMO channel matrix $\mathbf{H}$. This serves as an upper bound for the proposed detection techniques illustrating the best detection performance achievable for the corresponding scenario. We now describe the genie aided matched filter and the nominal CSI based matched filter detector employed to benchmark the performance of the proposed schemes.

A. Genie-Aided Matched Filter Detector (MF Genie)

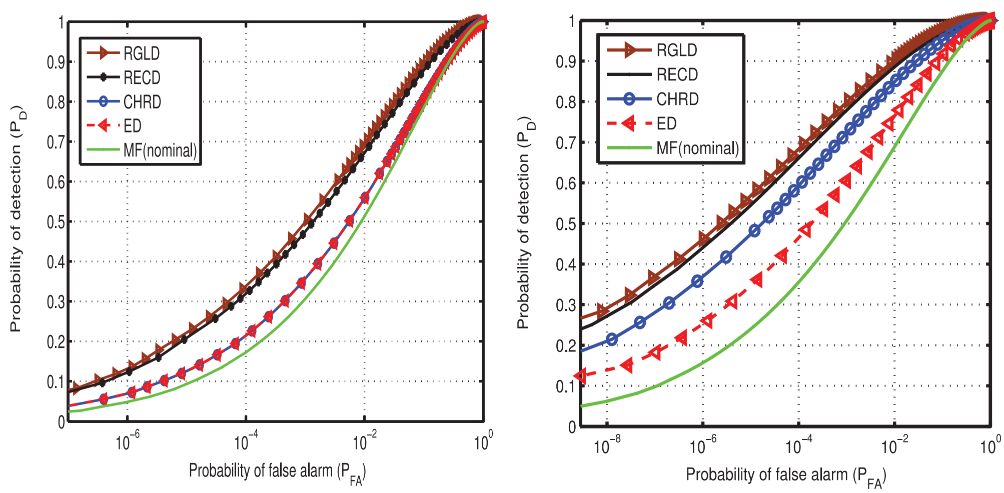
The optimal detector with perfect CSI for an additive white Gaussian noise scenario is given by the standard matched filter detector [6]. This can be derived employing the likelihood ratio test as,

$$L_{\text{genie}}(\mathbf{Y}) = \log\left(\frac{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \mathcal{H}_1)}{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \mathcal{H}_0)}\right) \doteq \sum_{k=1}^{N_r}\left(\mathbf{y}_k^H\mathbf{R}_\eta^{-1}\mathbf{y}_k - (\mathbf{y}_k - \mathbf{X}\mathbf{h}_k)^H\mathbf{R}_\eta^{-1}(\mathbf{y}_k - \mathbf{X}\mathbf{h}_k)\right).$$



(a) (b)

Fig. 3. ROC curves for the genie aided matched filter detector (MF genie), robust generalized likelihood detector (RGLD), robust estimator-correlator detector (RECD), composite hypothesis based robust detector (CHRD), nominal estimate based matched filter detector (MF) in (42) and energy detector (ED) in (43) for SNR $= -6$ dB, $L = 4$ and (a) $\sigma_u^2 = 1$, (b) $\sigma_u^2 = 0.6$. (a) Case I, (b) Case II.



(a) (b)

Fig. 4. Receiver operating characteristic (ROC) curves in logarithmic scale for the robust generalized likelihood detector (RGLD), robust estimator-correlator detector (RECD), composite hypothesis based robust detector (CHRD), energy detector (ED) in (43) and nominal estimate based matched filter detector (MF) in (42) with SNR $= -3$ dB, $\sigma_u^2 = 1$, (a) $L = 2$ and (b) $L = 4$. (a) Case I, (b) Case II.

From the likelihood ratio $L_{\text{genie}}(\mathbf{Y})$ above, the Neyman-Pearson (NP) based matched filter detector and the test statistic $T_{MF(\text{genie})}(\mathbf{Y})$ are given as,

$$T_{MF(\text{genie})}(\mathbf{Y}) = \sum_{k=1}^{N_r} \mathbf{y}_k^H \mathbf{R}_\eta^{-1} \mathbf{X} \mathbf{h}_k \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \tag{41}$$

B. Matched Filter Detector (MF)

The joint log-likelihood ratio test for the matched filter detector ignoring CSI uncertainty can be derived as,

$$\begin{aligned} &L_{\text{MF}}(\mathbf{Y}) \\ &= \log\left(\frac{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \mathcal{H}_1)}{\prod_{k=1}^{N_r} p(\mathbf{y}_k; \mathcal{H}_0)}\right) \\ &\doteq \sum_{k=1}^{N_r} \left(\mathbf{y}_k^H \mathbf{R}_\eta^{-1} \mathbf{y}_k - (\mathbf{y}_k - \mathbf{X}\hat{\mathbf{h}}_k)^H \mathbf{R}_\eta^{-1} (\mathbf{y}_k - \mathbf{X}_k \hat{\mathbf{h}}_k)\right). \end{aligned}$$

Therefore, the test statistic $T_{MF}(\mathbf{Y})$ for the uncertainty agnostic matched filter detector can be equivalently obtained as,

$$T_{MF}(\mathbf{Y}) = \sum_{k=1}^{N_r} \mathbf{y}_k^H \mathbf{R}_\eta^{-1} \mathbf{X} \hat{\mathbf{h}}_k. \tag{42}$$

C. Energy Detector (ED)

The test statistic T_{ED} for the energy detector [5]–[7] for the MIMO cognitive radio system model described in (1), is given as,

$$T_{\text{ED}}(\mathbf{Y}) = \sum_{k=1}^{N_r} \mathbf{y}_k^H \mathbf{y}_k. \tag{43}$$

Consider the test statistic obtained for the CHRD in (18). Let the beacon matrix $\mathbf{X}_1 \in \mathbb{C}^{L \times N_t}$ corresponding to the alternative hypothesis be of dimension 2×2, i.e., the number of transmit

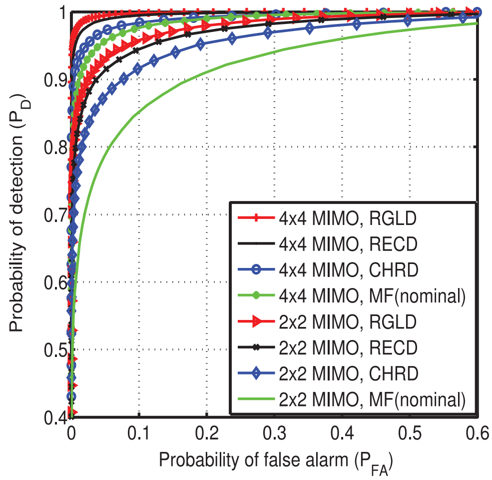


Fig. 5. Receiver operating characteristic (ROC) curves in logarithmic scale for the robust generalized likelihood detector (RGLD), robust estimator-correlator detector (RECD), composite hypothesis based robust detector (CHRD) and nominal estimate based matched filter detector (MF) in (42) with $L = 2$, SNR $= -1$ dB for 2×2 MIMO with $\mathbf{R}_u = \mathcal{D}([1, 0.9]^T)$ and for 4×4 MIMO with $\mathbf{R}_u = \mathcal{D}([1, 0.9, 0.6, 0.4]^T)$.

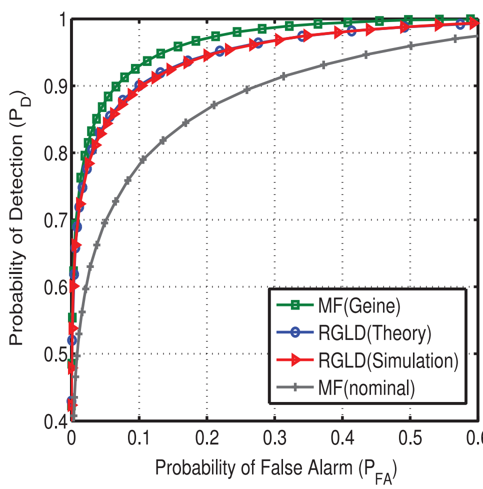


Fig. 6. ROC curves for the simulation based performance of the robust estimator-correlator detector (RECD simulation) and the corresponding plots from analytical expressions (RECD theory) for SNR $= -3$ dB, $L = 2$, and $\mathbf{R}_u = \mathcal{D}(1, 1)$.

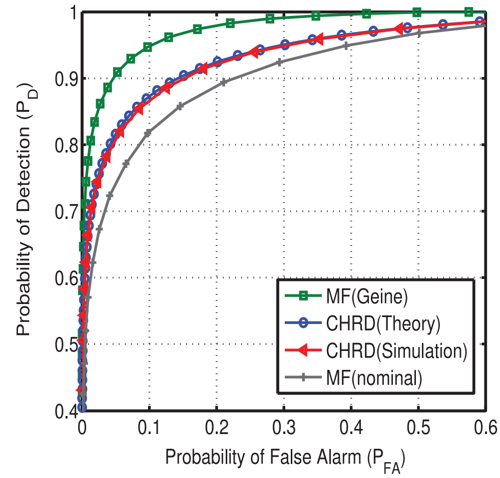


Fig. 7. ROC curves for the simulation based performance of the composite hypothesis based robust detector (CHRD simulation) and the corresponding plots from analytical expressions (CHRD theory) for SNR $= -2$ dB, $L = 2$, and $\mathbf{R}_u = \mathcal{D}(1, 0.9)$.

antennas $N_t = 2$ and the number of beacon vectors $L = 2$. This makes the beacon matrix an orthogonal square matrix with $\mathbf{X}_1^H \mathbf{X}_1 = \mathbf{X}_1 \mathbf{X}_1^H = 2\mathbf{I}_2$ and the resultant MLE $\hat{\mathbf{h}}_{k|1}$ of the vector $\mathbf{h}_k$ corresponding to the alternative hypothesis $\mathcal{H}_1$ reduces to $\hat{\mathbf{h}}_{k|1} = \frac{1}{2}\mathbf{X}_1^H \mathbf{y}_k$. Hence the test statistic $T_{\text{CHRD}}(\mathbf{Y})$ for the CHRD in (18) can be equivalently written as,

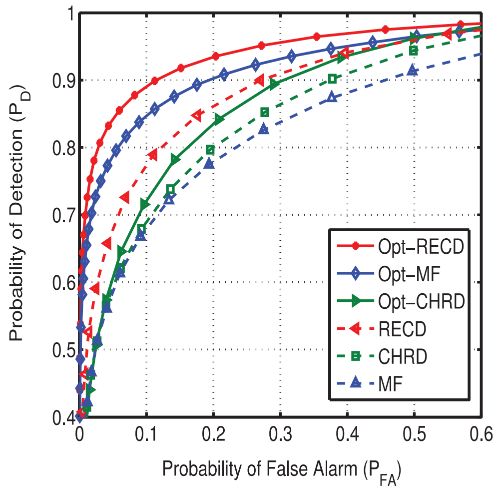
$$T_{\text{CHRD}}(\mathbf{Y}) = \sum_{k=1}^{N_r} \hat{\mathbf{h}}_{k|1}^H \hat{\mathbf{h}}_{k|1} \doteq \sum_{k=1}^{N_r} \mathbf{y}_k^H \mathbf{y}_k, \quad (44)$$

which is identical to that of the conventional energy detector given in (43).

In Figs. 1(a),(b), we compare the primary user detection performance of the proposed robust estimator-correlator detector (RECD) in (6), the robust generalized likelihood detector (RGLD) in (14), the composite hypothesis based robust detector (CHRD) in (18), with the energy detector (ED) in (43), the nominal channel estimate based matched filter detector (MF) in (42) and the true channel coefficient based genie aided matched filter detector (MF genie) in (41) for $L \in \{2, 4\}$. The proposed robust detection schemes can be seen to lead to an improved detection performance in comparison to the uncertainty agnostic matched filter detector. Further, RGLD demonstrates a performance edge over the other schemes. Across the figures, the primary user detection performance of the proposed schemes improves with an increase in the number of beacon vectors L, thereby decreasing the performance gap with respect to the genie aided matched filter detector. Simulation results also demonstrate that the detection performance of the CHRD and the ED is identical for $N_t = L = 2$. This is due to the fact that ED is a special case of the proposed CHRD for $N_t = L = 2$ as shown in Section VI-C.

Figs. 2(a), (b) present a performance comparison of the competing spectrum sensing schemes for various SNR values in the set $\{-4, -2\}$ dB and $L = 4$ beacon symbols. Similarly, Figs. 3(a), (b) present the P_D versus P_{FA} performance with different levels of CSI uncertainty considering the uncertainty covariance matrices $\mathbf{R}_u = \sigma_u^2 \mathcal{D}([1, 0.9]^T)$ with $\sigma_u^2 \in \{1, 0.6\}$. The simulation results demonstrate a similar trend in the detection performance of the proposed robust detection schemes. It can also be observed across the figures that the performance gap between the proposed robust schemes and the uncertainty agnostic matched filter detector along with the energy detector (ED) widens with increasing CSI uncertainty. Further, Fig. 4(a) and (b) present a P_{FA} versus P_D performance comparison of the proposed robust detection schemes with the uncertainty agnostic matched filter (MF) detector and the energy detector (ED) on a logarithmic scale for improved resolution and show a trend similar to Fig. 1.

Fig. 5 presents a performance comparison of the proposed detection schemes for various values of the number of receive and transmit antennas. We consider 2×2 MIMO scenario with CSI uncertainty covariance matrix $\mathbf{R}_u = \mathcal{D}([1, 0.9]^T)$ and 4×4 MIMO scenario with CSI uncertainty covariance matrix



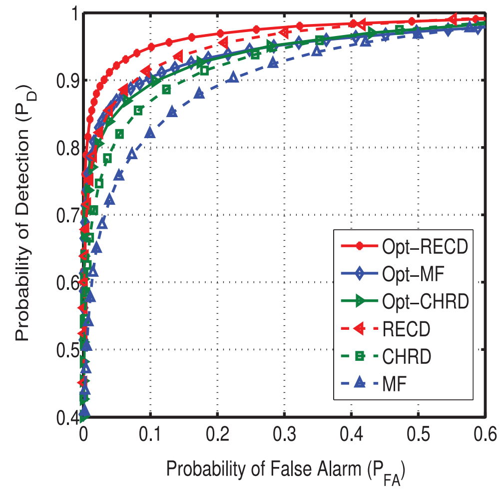


Fig. 8. ROC curves for the optimal beacon matrix versus the orthogonal beacon matrix for the composite hypothesis based robust detector (CHRD) and matched filter (MF) in (42) with SNR $= -5$ dB, and $\mathbf{R}_u = \mathcal{D}(1, 0.9)$ for (a) $L = 2$, (b) $L = 4$. (a) Case I, (b) Case II.

$\mathbf{R}_u = \mathcal{D}([1, 0.9, 0.6, 0.4]^T)$. The detection performance of the proposed robust detection schemes significantly improves with the increase in the number of receive and transmit antennas.

In Fig. 6 we plot the probability of detection (P_D) versus the probability of false alarm (P_{FA}) and compare the simulated detection performance with the analytical performance curves generated from the derived expressions for P_D in (8) and P_{FA} in (9) for the RECD with $L = 2$, SNR $= -3$ dB and $\mathbf{R}_u = \mathcal{D}([1, 1]^T)$. It is evident from the figure that the detection performance of the proposed RECD technique obtained through simulations is in close agreement with the results obtained via theory. Similarly, in Fig. 7 we compare the detection performance from the expressions for P_D in (22) and P_{FA} in (21) corresponding to the CHRD with the simulation results for a scenario with $L = 2$, SNR $= -1$ dB and $\sigma_u^2 = 1$. The simulated detection performance of the CHRD scheme coincides with the analytical results.

Fig. 8(a), (b) present the P_D versus P_{FA} performance considering the optimal beacon matrix derived in Section V for the scenarios with a known uncertainty covariance and unknown uncertainty covariance. The detection performance for the sub-optimal orthogonal beacon sequence is also given therein. Fig. 8 clearly demonstrate a performance improvement for the derived optimal beacon sequence based spectrum sensing in comparison to the orthogonal beacon sequence.

VII. Conclusion

This paper considers the problem of primary user detection for MIMO cognitive radio scenarios with CSI uncertainty. In this context, novel detection schemes such as the robust estimator-correlator detector (RECD) and the robust generalized likelihood detector (RGLD), which are robust against CSI uncertainty, have been proposed for scenarios with known uncertainty statistics. Further, for the scenario with unknown CSI uncertainty statistics, we developed a GLRT based composite hypothesis robust detector (CHRD) for spectrum sensing in MIMO cognitive radio networks. Closed form analytical expressions have been derived to characterize the theoretical detection performance of the proposed RECD and CHRD schemes. Subsequently, an optimization framework has also been presented to obtain the optimal beacon sequences which further enhance the performance of the proposed detectors. Simulation results were presented to illustrate the improved detection performance of the proposed robust detection schemes which consider CSI uncertainty in MIMO cognitive radio networks. It has also been shown that the optimal beacon matrix significantly boosts the detection performance towards MIMO spectrum sensing. The proposed framework can be further extended considering other additional challenging aspects such as noise uncertainty in future works.

References

[1] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, “Chicago spectrum occupancy measurements and analysis and a long-term studies proposal,” in *Proc. ACM 1st Int. Workshop Technol. Policy Accessing Spectrum (TAPAS’06)*, New York, NY, USA, 2006.

[2] J. Mitola, J. Maguire, and G. Q. , “Cognitive radio: Making software radios more personal,” *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[3] J. Mitola, “Cognitive radio: An integrated agent architecture for software defined radio,” Ph.D. dissertation, Royal Inst. Technol. (KTH), Stockholm, Sweden, 2000.

[4] S. Haykin, “Cognitive radio: Brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[5] E. Axell, G. Leus, E. Larsson, and H. Poor, “Spectrum sensing for cognitive radio : State-of-the-art and recent advances,” *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Englewood, NJ, USA: Prentice-Hall PTR, Jan. 1998.

[7] A. Sonnenschein and P. Fishman, “Radiometric detection of spread-spectrum signals in noise of uncertain power,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 654–660, 1992.

[8] A. Sahai, N. Hoven, and R. Tandra, “Some fundamental limits on cognitive radio,” pp. 1–11, 2004 [Online]. Available: http://www.eecs.berkeley.edu/~sahai/Papers/cognitive_radio_preliminary.pdf

[9] E. Axell and E. Larsson, “Optimal and sub-optimal spectrum sensing of OFDM signals in known and unknown noise variance,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 290–304, 2011.

[10] H. Sun, W.-Y. Chiu, J. Jiang, A. Nallanathan, and H. Poor, “Wideband spectrum sensing with sub-Nyquist sampling in cognitive radios,” *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6068–6073, Nov. 2012.

[11] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, “Wideband spectrum sensing for cognitive radio networks: A survey,” *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.

[12] O. Mehanna and N. Sidiropoulos, "Maximum likelihood passive and active sensing of wideband power spectra from few bits," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1391–1403, Mar. 2015.
[13] D. Cabric, "Addressing feasibility of cognitive radios," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 85–93, 2008.
[14] J. Ma and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," in *Proc. Conf. Global Telecommun. Conf. (GLOBECOM'07)*, Nov. 2007, pp. 3139–3143.
[15] S. Qing-Hua, W. Xuan-Li, S. Xue-Jun, and Z. Nai-Tong, "Cooperative spectrum sensing under imperfect channel estimation," in *Proc. Int. Conf. Commun. Mobile Comput. (CMC)*, Apr. 2010, vol. 2, pp. 174–178.
[16] H. He, G. Li, and S. Li, "Adaptive spectrum sensing for time-varying channels in cognitive radios," *IEEE Wireless Commun. Lett.*, vol. 2, no. 2, pp. 1–4, Apr. 2013.
[17] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814–823, Feb. 2010.
[18] J. Font-Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2137–2146, Jul. 2010.
[19] 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.
[20] T. J. Lim, R. Zhang, Y. C. Liang, and Y. Zeng, "GLRT-based spectrum sensing for cognitive radio," in *Proc. Global Telecommun. Conf. (IEEE GLOBECOM)*, Nov. 2008, pp. 1–5.
[21] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria, "Detection of rank-p signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.
[22] A. Taherpour, S. Gazor, and M. Nasiri-Kenari, "Invariant wideband spectrum sensing under unknown variances," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2182–2186, May 2009.
[23] A. Taherpour, S. Gazor, and M. Nasiri-Kenari, "Wideband spectrum sensing in unknown white Gaussian noise," *IET Commun.* , vol. 2, no. 6, pp. 763–771, Jul. 2008.
[24] J. Font-Segura, G. Vazquez, and J. Riba, "Nonuniform sampling walls in wideband signal detection," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 44–55, Jan. 2014.
[25] G. Vazquez-Vilar, R. Lopez-Valcarce, and J. Sala, "Multiantenna spectrum sensing exploiting spectral a priori information," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4345–4355, Dec. 2011.
[26] L. Wei, O. Tirkkonen, and Y.-C. Liang, "Multi-source signal detection with arbitrary noise covariance," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5907–5918, Nov. 2014.
[27] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, Aug. 1999.
[28] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. Amsterdam: Elsevier/Academic, 2007.
[29] S. Vorobyov, A. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.
[30] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, Sep. 2004.
[31] D. Eustice and M. S. Klamkin, "On the coefficients of a partial fraction decomposition," *Amer. Math. Month.*, vol. 86, no. 6, pp. 478–480, 1979.
[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, Mar. 2004.
[33] R. Lorenz and S. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

Adarsh Patel (S'12) received the B.Tech. degree in electronics and communication engineering from Integral University, Lucknow, India, in 2008 and the M.Tech. degree in computer science and engineering (digital communication) from ABV-Indian Institute of Information Technology and management, Gwalior, in 2010.

He is currently working toward the Ph.D. degree in electrical engineering at the Indian Institute of Technology, Kanpur. His main research interests include detection and estimation theory within the areas of wireless communication and specially spectrum sensing aspect of cognitive radio.

Mr. Patel is a recipient of the TCS Research Fellowship.

Sinchan Biswas (S'14) was born in India in 1988. He received the M.Tech. degree from Department of Electrical Engineering, Indian Institute of Technology (IIT), Kanpur, in 2013.

He is currently pursuing the Ph.D. degree with the Department of Engineering Sciences, Uppsala University, Sweden. From 2013 to 2014, he was a Research Assistant with the IIT Kanpur. His research interest includes signal processing, information theory, and stochastic control.

Aditya K. Jagannatham (M'09) received the Bachelors degree from the Indian Institute of Technology (IIT), Bombay, and the M.S. and Ph.D. degrees from the University of California, San Diego, USA.

From April 2007 to May 2009, he was a Senior Wireless Systems Engineer at Qualcomm, Inc., San Diego, where he worked on developing 3G UMTS/WCDMA/HSDPA mobile chipsets as part of the Qualcomm CDMA technologies division. His research interests are in the area of next-generation wireless communications and networking, sensor and *ad-hoc* networks, digital video processing for wireless systems, wireless 3G/4G cellular standards and CDMA/OFDM/MIMO wireless technologies. He has contributed to the 802.11n high throughput wireless LAN standard. Since 2009, he has been with Electrical Engineering Department, IIT Kanpur, where he is currently an Associate Professor, and is also associated with the BSNL-IITK Telecom Center of Excellence (BITCOE).

Dr. Jagannatham was awarded the CAL(IT)2 fellowship for pursuing graduate studies at the University of California San Diego and, in 2009, received the Upendra Patel Achievement Award for his efforts towards developing HSDPA/HSUPA/HSPA+ WCDMA technologies at Qualcomm. At IIT Kanpur, he has been awarded the P.K. Kelkar Young Faculty Research Fellowship (June 2012 to May 2015) for excellence in research and the Gopal Das Bhandari Memorial Distinguished Teacher Award for the year 2012–2013 for excellence in teaching. He has also delivered a set of video lectures on Advanced 3G and 4G Wireless Mobile Communications for the Ministry of Human Resource Development funded initiative National Programme on Technology Enhanced Learning (NPTEL).