Multi-Frequency GLRT Spectrum Sensing for Wideband Cognitive Radio

Josep Font-Segura, Gregori Vázquez and Jaume Riba
Signal Theory and Communications Department, Technical University of Catalonia (UPC)
D5-{204, 214, 116}, Campus Nord UPC, Jordi Girona 1-3, 08034 Barcelona, Spain
Email: {josep.font-segura, gregori.vazquez, jaume.riba}@upc.edu

***Abstract*—The problem of spectrum sensing in multi-frequency cognitive radio systems is addressed. We show that as the sensed bandwidth increases, the primary user detection is governed by a low signal-to-noise ratio (low-SNR) regime. By means of low-SNR approximations, we show that the optimal generalized likelihood ratio test (GLRT) only depends on the second order statistics of the observations and on a shaping kernel that highlights the relevant parameters required for detection. Furthermore, the ML estimates of the unknown model parameters are derived for multi-frequency systems, which allow closed-form expressions for the GLRT statistic. The detection performance and the kernel interpretation are supported with simulation results.**

***Index Terms*—Cognitive radio, spectrum sensing, GLRT, wideband/low-SNR regime, frequency-domain.**

I. Introduction

Cognitive radio is a new wireless communication paradigm that utilizes advanced signal processing along with novel dynamic spectrum policies to support new users who wish to opportunistically communicate in the existing congested spectrum without degrading the established users [1]. A cognitive radio is an adaptive wireless communication system that takes advantage of side information on the network. Interweave cognitive radio [2] is motivated by opportunistic transmission of secondary users over the available spectrum gaps or holes in given time and geographical location conditions.

A primary function of interweave cognitive radios is to reliably identify the available spectrum resources temporally unused by the primary users. This awareness can be obtained through a database, by using beacons, or by local spectrum sensing [3]. In this paper, we focus on spectrum sensing performed at the cognitive radio receivers as it constitutes a broader solution and has less infrastructure requirements. On the one hand, the energy detector, pilot-based methods [4], cyclostationarity feature detection [5], and match-filtering [6] are the most commonly employed techniques for spectrum sensing. However, the performance of such detectors is severely degraded with inaccurate prior information on the model features such as the noise variance in the energy detector, or the cyclic frequencies in the cyclostationarity detector [7]. On the other hand, spectrum sensing detectors based on the generalized likelihood ratio test (GLRT) have received recent attention, e.g. [8], [9], as the GLRT statistic is optimal in the Neyman-Pearson sense [10] and natively incorporates joint parameter maximum likelihood (ML) estimation for inaccurate model parameters. ML estimation in wideband cognitive radios is especially challenging because wideband regimes are characterized by close to zero spectral efficiency and very low signal-to-noise ratios (SNRs) [11]. Furthermore, the method of ML, despite its theoretical appeal, is often difficult to implement, and analytical solutions are not available in many circumstances. However, we show the tractability of the ML formulation in asymptotically low-SNR regimes, and identify that the second-order statistics of the observations are sufficient statistics for the spectrum detection problem when the noise level is high.

In this paper, we derive the optimal ML estimates for GLRT spectrum sensing detection in multi-frequency scenarios, and show through low-SNR approximations that GLRT detection depends on a kernel operator and the sample covariance matrix of the observations, asymptotically as the SNR tends to zero. We further derive the optimal GLRT detectors and assess their performance comparison by means of simulation results.

II. Signal Model and Problem Statement

We consider the spectrum sensing problem for multi-frequency wideband cognitive radio networks consisting of a set of secondary users, each equipped with one receiving antenna with the purpose of individually monitoring the activity of the primary users', denoted by the wide-sense stationary signal $S(t)$, which accounts for the superposition of the primary services over the sensed spectrum of bandwidth B. In multi-frequency systems, the primary users' services employ frequency-division multiplexing (FDM) with predetermined channelization. The sensed multi-frequency system is characterized by K adjacent channels, i.e.,

$$X(t) \doteq \sum_{k=1}^{K} S_k(t) + W(t), \tag{1}$$

where $S_k(t)$ denotes the primary users' signal located at the k-th channel, and $W(t)$ is the double-sided complex zero-mean additive white Gaussian noise with spectral density $N_0/2$. The N-dimensional discrete-time received signal is defined

This work has been partially funded by the Spanish Government under TEC2010-21245-C02-01 (DYNACS), CONSOLIDER INGENIO CSD2008-00010 (COMONSENS), CENIT CEN-20101019 (THOFU), and the Catalan Government (DURSI) under Grant 2009SGR1236 and Fellowship FI-2010.

as $\mathbf{x}[m] \doteq [X(t_1^m), \ldots, X(t_N^m)]^T$, where the sampling instants satisfy Nyquist-rate uniform sampling and piece-wise stacking, i.e., $t_n^m = t_0 + (mN+n)\frac{1}{B}$. As a result, each cognitive radio acquires an observations data record of size $N \times M$ given by $\mathbf{X} \doteq (\mathbf{x}[1], \ldots, \mathbf{x}[M])$. We similarly define $\mathbf{S}_k$ and $\mathbf{W}$ for the signal and noise components.

Let $\mathcal{H}_{0,k}$ be the hypothesis that the primary users are not transmitting over the k-th channel. Similarly, let $\mathcal{H}_{1,k}$ denote the event in which there is activity on the k-th channel during the sensing interval. The spectrum sensing problem may be therefore cast as the binary hypotheses testing problem

$$\begin{array}{lcl} \mathcal{H}_{0,k} & : & \mathbf{X} = \sum_{l \neq k} \mathbf{S}_l + \mathbf{W} \\ \mathcal{H}_{1,k} & : & \mathbf{X} = \sum_{l=1}^{K} \mathbf{S}_l + \mathbf{W}, \end{array} \quad (2)$$

for $1 \leq k \leq K$. In (2), the column entries of $\mathbf{S}_k$ and $\mathbf{W}$ are complex multivariate and zero-mean Gaussian distributed with correlation matrices $\mathbf{R}_{s_k} = \gamma_k \mathbf{R}_k$, and $\mathbf{R}_w = \sigma^2 \mathbf{I}$, respectively, where $\mathbf{R}_k$ is the normalized Toeplitz correlation matrix of the primary users' signal on the k-th channel, with $\mathrm{tr}(\mathbf{R}_k) = N$ and where γ_k stands for the received power level on the k-th channel. If the sensed multi-frequency system employs homogeneous services, the signal statistics across channels further accomplish $\mathbf{R}_k = \mathbf{R}_0 \odot (\mathbf{e}(\omega_k)\mathbf{e}^H(\omega_k))$, where $\mathbf{R}_0$ is the baseband basic modulation format, and $\mathbf{e}(\omega)$ is the steering vector at ω, i.e., $\mathbf{e}^H(\omega) \doteq \left[1 \; e^{\jmath\omega} \; \ldots \; e^{\jmath\omega(N-1)}\right]$. For notation purposes, we define $\mathcal{M} = \{\mathbf{R}_k\}$ as the set of normalized multi-frequency correlation matrices, and $\boldsymbol{\gamma} \doteq (\gamma_1, \ldots, \gamma_K)^T$. We further define the SNR at the k-th channel as $\rho_k \doteq \frac{\gamma_k}{\sigma^2}$ and $\rho_0 \doteq \sum_{k=1}^{K} \rho_k$.

In the problem in hand, it is a valid assumption that both noise and signal are normally distributed. While facilitating the analysis, this is reasonable because usually there is no line-of-sight (LOS) path between the cognitive radio receiver and the primary user transmitter. Hence, the resulting signal is the superposition of no-LOS signals and approximates Gaussian distribution as the number of observations is sufficiently large, according to the central limit theorem. Moreover, it has recently proved that Gaussian ML estimation provides, asymptotically as $\rho_0 \to 0$, the optimum second-order estimator [12].

III. Multi-Frequency GLRT

We are interested in detecting the presence of the signal observations $\mathbf{S}_k$ based on the local observations $\mathbf{X}$ in (2). It is known that the GLRT is asymptotically optimal under the Neyman-Pearson criterion, i.e., to maximize the probability of detection, for a given probability of false alarm level when the number of observations tends to infinity [10]. Recently, the finite-sample optimality of GLRT has been established in [13]. When treating the multi-frequency cognitive radio problem as a joint multiple-hypotheses test, the complexity of the spectrum sensing detectors grows exponentially with the number of channels, which becomes impractical due to the limited processing capabilities of the wideband cognitive radios. Conversely, when casting the cognitive radio problem as in (2), we are dealing with K nuisance-parameter binary tests [10]. Let $\Psi_{0,k}$ and $\Psi_{1,k}$ denote the sets of unknown model parameters under $\mathcal{H}_{0,k}$ and $\mathcal{H}_{1,k}$, respectively. The nuisance-hypotheses testing problem at the k-th channel is given by

$$L_k(\mathbf{X}, \Theta_k) = \frac{p(\mathbf{X}|\hat{\Psi}_{1,k}, \Theta_{1,k}, \mathcal{H}_{1,k})}{p(\mathbf{X}|\hat{\Psi}_{0,k}, \Theta_{0,k}, \mathcal{H}_{0,k})} \geq \lambda_k, \quad (3)$$

where the ML estimates of the unknown parameters are given by $\hat{\Psi}_{1,k} = \arg\max_{\Psi} p(\mathbf{X}|\Psi, \Theta_{1,k}, \mathcal{H}_{1,k})$ and $\hat{\Psi}_{0,k} = \arg\max_{\Psi} p(\mathbf{X}|\Psi, \Theta_{0,k}, \mathcal{H}_{0,k})$. In (3), Θ_k denotes the set of a priori known parameters, and $\Theta_{0,k}, \Theta_{1,k} \subset \Theta_k$. The threshold λ_k sets the decision level for which the test statistic $L_k(\mathbf{X}, \Theta_k)$ decides for $\mathcal{H}_{1,k}$, and for $\mathcal{H}_{0,k}$ otherwise, and is selected to satisfy the false alarm level $\mathbb{P}(\mathcal{H}_{1,k}|\mathcal{H}_{0,k}) = \mathbb{P}\left[L_k(\mathbf{X}, \Theta_k) \geq \lambda_k | \mathcal{H}_{0,k}\right] = \alpha_k$. In the sequel, we set $\alpha_k = \alpha$ for all k.

Despite its theoretical appeal, the ML estimation and GLRT detection in (3) are often difficult to implement, and analytical solutions are not available in many circumstances. However, we show that GLRT spectrum sensing detection for wideband cognitive radio is encompassed in a low-complex unified framework, which we state in the following Theorem.

Theorem 1. *Consider the wideband cognitive radio spectrum sensing problem* (2). *Under the Gaussian assumption, the multi-frequency GLRT spectrum sensing detector* (3) *is, asymptotically as $\rho_k \to 0$ and $N \to \infty$, given by*

$$T_k(X, \Theta_k) \doteq \int_B \mathcal{K}_k(\omega, \Theta_k) P(\omega) d\omega \geq \lambda_k, \quad (4)$$

where X is the asymptotic data record, Θ_k is the set of known a priori parameters, $\mathcal{K}_k(\omega, \Theta_k)$ is a kernel associated to each detector, $P(\omega)$ is the continuous-frequency periodogram of X, and λ_k is the detection threshold. Furthermore, the multi-frequency GLRT spectrum sensing detector only depends, asymptotically as $\rho_k \to 0$, on the second-order statistics of the observations.

Proof: See Appendix A. ∎

From (4), it is deduced that the multi-frequency GLRT spectrum sensing detection is based on the product between the second-order statistics of the observations and a shaping kernel that highlights the model features which are relevant for detection. Moreover, we show that frequency-domain asymptotic kernels derived from GLRT spectrum sensing detectors as $\rho_k \to 0$ have a common inner structure that depends on the signal and noise-plus-interference statistics. Let $\phi_s(\omega)$ and $\phi_\nu(\omega)$ denote the PSD of the signal to be detected and the noise-plus-interference, respectively. The detection kernel in (4) has an internal structure given by

$$\mathcal{K}_0\left[\phi_s(\omega), \phi_\nu(\omega)\right] \doteq \frac{1}{\phi_\nu(\omega)} \frac{\phi_s(\omega)}{\phi_s(\omega) + \phi_\nu(\omega)}. \quad (5)$$

For the cognitive radio problem (2), we note that $\phi_\nu(\omega) = \sum_{l \neq k} \phi_{s_l}(\omega) + \phi_w(\omega)$. On the one hand, as $\mathcal{K}_0\left[\phi_s(\omega), \phi_\nu(\omega)\right]$ approaches to the unity, the detector asymptotically behaves as the energy detector. On the other hand, we show that the kernel becomes proportional to the primary users' signal power spectral density (PSD) as the noise level increases, and the detector performs the spectral correlation between the averaged periodogram of the observations and the signal PSD. This result is indeed well-known as the locally optimum detector for the cognitive radio problem (2), obtained through directly expanding the optimum quadratic statistic in the low-SNR limit [14]. When the signal and noise-plus-interference statistics are not perfectly characterized, the expression of $\mathcal{K}_0 \left[\phi_s(\omega), \phi_\nu(\omega)\right]$ depends on the ML estimates of the unknown parameters. Therefore, the performance of the GLRT spectrum sensing detectors is related to the derivation of the ML estimates in (3). Even though parameter estimation is especially challenging in wideband regimes, in what follows, we show that analytical solutions are obtained asymptotically for $\rho_k \to 0$. The expressions of the ML estimates provide further insight on the fundamentals of the spectrum sensing detection problem (2).

IV. Optimal Multi-Frequency Wideband Detectors

We next discuss the optimal GLRT spectrum sensing detectors for the multi-frequency model (1) with known and unknown noise level. We assume that the frequencies ω_k and the baseband modulation format $\mathbf{R}_0$ are perfectly known.

A. Multi-Frequency Detector

Assume that each cogitative radio device has perfect knowledge on the noise variance σ^2, as well as the primary system multi-frequency structure $\mathcal{M}$.

Theorem 2. *For a given multi-frequency system $\mathcal{M}$, the optimal GLRT spectrum sensing detector at the k-th channel in the wideband regime with known noise variance is given by*

$$T_k(\mathbf{X}|\mathcal{M}, \sigma^2) = \mathrm{tr}\left(\mathbf{\Xi}_k^{-1}\hat{\gamma}_k\mathbf{R}_k(\hat{\gamma}_k\mathbf{R}_k + \mathbf{\Xi}_k)^{-1}\mathbf{R}_x\right) \geq \lambda_k, \tag{6}$$

where $\mathbf{\Xi}_k \doteq \sum_{l\neq k}\hat{\gamma}_l\mathbf{R}_l + \sigma^2\mathbf{I}$, and the ML estimates of the signal levels are given by

$$\begin{pmatrix} \mathrm{tr}(\mathbf{R}_1^2) & \dots & \mathrm{tr}(\mathbf{R}_1\mathbf{R}_K) \\ \vdots & & \vdots \\ \mathrm{tr}(\mathbf{R}_K\mathbf{R}_1) & \dots & \mathrm{tr}(\mathbf{R}_K^2) \end{pmatrix} \times \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_K \end{pmatrix} = \begin{pmatrix} \mathrm{tr}(\mathbf{R}_1\mathbf{R}_x) \\ \vdots \\ \mathrm{tr}(\mathbf{R}_K\mathbf{R}_x) \end{pmatrix} - \sigma^2 N\mathbf{1}, \tag{7}$$

where $\mathbf{1}$ is the all-ones column vector.

Proof: See Appendix B. ■

We note that when the number of available samples is low, the orthogonality between channels is not preserved and, in general, the system of equations (7) is coupled because $\mathrm{tr}(\mathbf{R}_k\mathbf{R}_l) \neq 0$, for $l \neq k$. However, for large data-records, the system matrix in (7) becomes diagonal, and the ML estimates at each channel are independent on the other channels, giving $\mathrm{tr}(\mathbf{R}_k^2)\hat{\gamma}_k = \mathrm{tr}(\mathbf{R}_k\mathbf{R}_x) - \sigma^2 N$. In both cases, the system matrix in (7) can be computed off-line. The frequency-domain asymptotic kernel for the spectrum sensing detection at the k-th channel is given by

$$\mathcal{K}_k(\omega, \mathcal{M}, \sigma^2) = \mathcal{K}_0\left[\hat{\gamma}_k\phi_k(\omega), \sum_{l\neq k}\hat{\gamma}_l\phi_l(\omega) + \sigma^2\right]. \tag{8}$$

We see that the detector employs the occupation on the remaining frequencies as interference for sensing the k-th channel. As expected, the performance of (6) is affected by the signal-to-interference-plus-noise ratio (SINR) based on the cross-correlation that arise from the adjacent channels.

B. Multi-Frequency and Noise Level Detector

We finally consider the detection of multi-frequency systems with unknown noise variance.

Theorem 3. *For a given multi-frequency system $\mathcal{M}$, the optimal GLRT spectrum sensing detector at the k-th channel in the wideband regime with unknown noise variance is given by*

$$T_k(\mathbf{X}|\mathcal{M}) = \mathrm{tr}\left(\mathbf{\Xi}_k^{-1}\hat{\gamma}_k\mathbf{R}_k(\hat{\gamma}_k\mathbf{R}_k + \mathbf{\Xi}_k)^{-1}\mathbf{R}_x\right) \geq \lambda_k, \tag{9}$$

where $\mathbf{\Xi}_k \doteq \sum_{l\neq k}\hat{\gamma}_l\mathbf{R}_l + \hat{\sigma}_1^2\mathbf{I}$, and the ML estimates of the signal and noise levels are given by

$$\begin{pmatrix} \mathrm{tr}(\mathbf{R}_1^2) & \dots & \mathrm{tr}(\mathbf{R}_1\mathbf{R}_K) & N \\ \vdots & & \vdots & \vdots \\ \mathrm{tr}(\mathbf{R}_K\mathbf{R}_1) & \dots & \mathrm{tr}(\mathbf{R}_K^2) & N \\ N & \vdots & N & N \end{pmatrix} \times \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_K \\ \hat{\sigma}_1^2 \end{pmatrix} = \begin{pmatrix} \mathrm{tr}(\mathbf{R}_1\mathbf{R}_x) \\ \vdots \\ \mathrm{tr}(\mathbf{R}_K\mathbf{R}_x) \\ \mathrm{tr}(\mathbf{R}_x) \end{pmatrix}. \tag{10}$$

Proof: See Appendix C. ■

The main advantage of the test statistic (3) is that all the information on the sensed bandwidth is exploited for joint detection and estimation at a given band. Whereas filter-bank based detectors may suffer from adjacent channel leakage, the nuisance-parameter formulation allows the detector to take advantage of the multi-frequency structure $\mathcal{M}$ for estimating both signal and noise levels. The frequency-domain asymptotic interpretation of the associated kernel

$$\mathcal{K}_k(\omega, \mathcal{M}) = \mathcal{K}_0\left[\hat{\gamma}_k\phi_k(\omega), \sum_{l\neq k}\hat{\gamma}_l\phi_l(\omega) + \hat{\sigma}_1^2\right] \tag{11}$$

shows that the ML estimate of the noise variance flourishes together with the remaining bands as interference.

V. Numerical Results

We evaluate the performance of the GLRT spectrum sensing detectors for multi-frequency systems. We consider a cognitive radio network with primary systems based on the terrestrial digital video broadcasting (DVB-T) standard in the 2k-mode

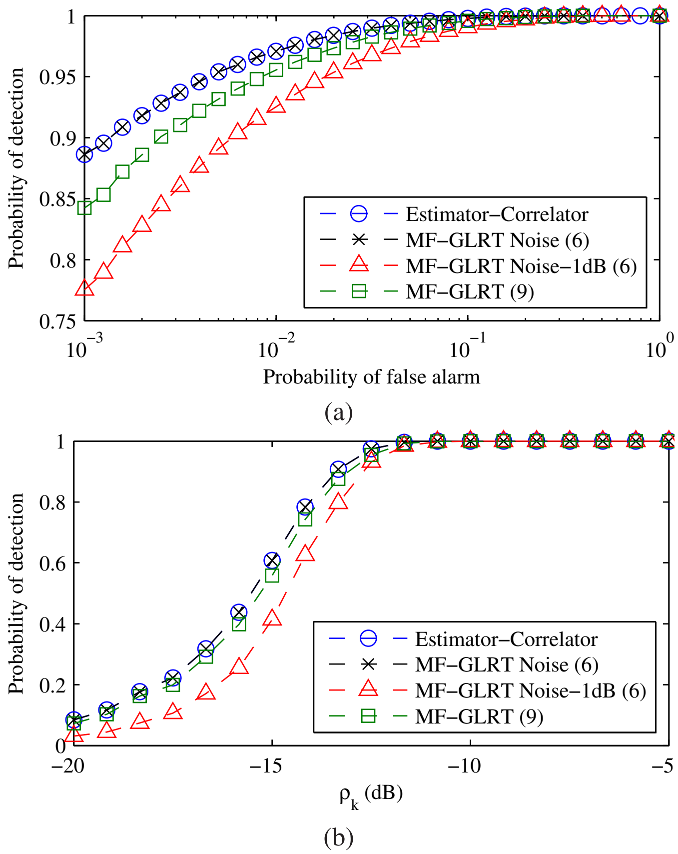

Fig. 1. (a) Receiver operating characteristics (ROC) of the multi-frequency spectrum sensing detectors at $\rho_k = -12.5$ dB, and (b) Probability of detection of the multi-frequency spectrum sensing detectors versus average ρ_k with false alarm level $\alpha = 0.01$.

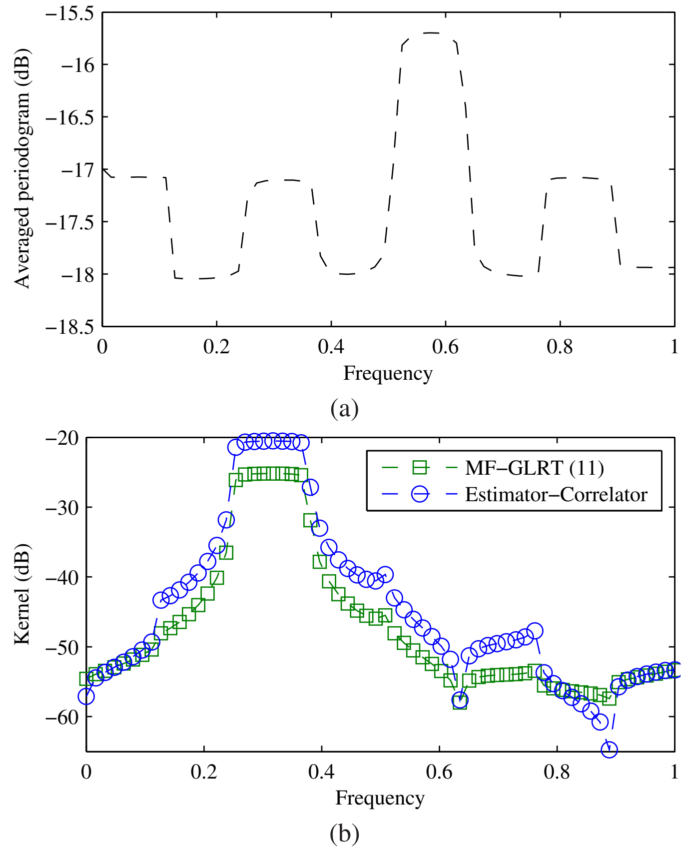

Fig. 2. (a) Observations averaged periodogram, and (b) the frequency-domain interpretation of the kernels in multi-frequency spectrum sensing detectors at average $\rho_k = -12.5$ dB.

in an example with $K = 8$ channels when sensing an arbitrary channel. The observation size is $N = 32$, and $M = 2N$. For comparison reasons, we add the estimator-correlator, i.e., the test statistic (6) with perfect knowledge on $(\gamma_1, \ldots, \gamma_K)$, and the multi-frequency GLRT (6) with noise uncertainty of 1 dB.

On the one hand, the sensing performance of the multi-frequency GLRT spectrum sensing detectors is depicted in Fig. 1. The ROC at $\rho_k = -12.5$ dB is analyzed in Fig. 1(a), whereas the probability of detection versus average SNR is shown in 1(b). It can be highlighted that, analogous to the wideband detectors, the prior knowledge on the normalized correlation matrices $\mathbf{R}_k$ is the most informative statistic on the primary users' signal as the multi-frequency detector (6) incurs no sensing loss in estimating the signal levels when the noise variance is perfectly known. However, it is seen in both figures that the multi-frequency detector with known noise variance suffers from a significant degradation in performance when there is a mismatch in the value of σ^2. Contrarily, the multi-frequency detector (9) achieves performance similar to that of the estimator-correlator as it is robust to unknown noise variance. The estimation of σ^2 in (10) degrades the performance in roughly only 1 dB in SNR, whereas the ROC for very restrictive false alarm levels incurs a larger penalty due to the sensitivity in the computation of the threshold.

We further show in Fig. 2 how the kernels actuate over the averaged periodogram of the observations in the multi-frequency detector when sensing the 3rd band with $\gamma_1 = \gamma_3 = \gamma_7$, γ_5 with slightly more power, and $\gamma_k = 0$ otherwise. As it can be appreciated, $\mathcal{K}_3(\omega, \mathcal{M})$ keeps the spectral shape of the estimator-correlator detector with a small shift and scaling. Both kernels show how they are affected by the spectral information outside the sensing frequency, because they are incorporated in $\phi_\nu(\omega)$ in (5), along with the additive thermal noise. As an example, we emphasize that the kernel focus on diminishing the contribution of ω_5 and ω_7, because it has a limited small number of degrees of freedom (i.e., N is small). For asymptotically large N, it is expected that $\mathcal{K}_k(\omega, \mathcal{M})$ will tend to the trivial solution, i.e., a narrowband filter at ω_3.

VI. Conclusions

In this paper we have derived a unified structure and closed-form solutions of wideband GLRT spectrum sensing detectors for multi-frequency cognitive radio networks by means of low-SNR approximations. The asymptotic interpretation of the detectors provides a physical insight on the main factors that determine the sensing performance. Numerical results provide performance comparison between the derived tests.

Appendix A
Proof of Theorem 1

We first proof that, asymptotically for $\rho_k \to 0$, the GLRT (3) is of the form $T_k(\mathbf{X}, \Theta_k) = \mathrm{tr}(\mathbf{K}_k \mathbf{R}_x)$. For clarity purposes, we make the following definitions: $\mathcal{M} \doteq \{\mathbf{R}_l\}_l$, $\boldsymbol{\gamma} \doteq (\gamma_1, \ldots, \gamma_K)^T$, $\boldsymbol{\gamma}_{\bar{k}} \doteq (\gamma_1, \ldots, \gamma_{k-1}, 0, \gamma_{k+1}, \gamma_K)^T$, and $\mathbf{R}_s[\boldsymbol{\gamma}] \doteq \sum_{l=1}^{K} \gamma_l \mathbf{R}_l$. Consider now the GLRT (3) for the spectrum sensing problem (2) under the Gaussian assumption, i.e., $L_k(\mathbf{X}, \mathcal{M}) = \frac{p(\mathbf{X}|\hat{\boldsymbol{\gamma}}, \hat{\sigma}_1^2, \mathcal{H}_{1,k})}{p(\mathbf{X}|\hat{\boldsymbol{\gamma}}_{\bar{k}}, \hat{\sigma}_0^2, \mathcal{H}_{0,k})}$, where $p(\mathbf{X}|\hat{\boldsymbol{\gamma}}, \hat{\sigma}_1^2, \mathcal{H}_{1,k}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_s[\hat{\boldsymbol{\gamma}}] + \hat{\sigma}_1^2 \mathbf{I})$, and $p(\mathbf{X}|\hat{\boldsymbol{\gamma}}_{\bar{k}}, \hat{\sigma}_0^2, \mathcal{H}_{0,k}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_s[\hat{\boldsymbol{\gamma}}_{\bar{k}}] + \hat{\sigma}_0^2 \mathbf{I})$. We further define $\boldsymbol{\Xi}_k \doteq \mathbf{R}_s[\hat{\boldsymbol{\gamma}}] + \hat{\sigma}_1^2$. By taking the logarithm and defining $\mathbf{R}_x \doteq \frac{1}{M}\mathbf{X}\mathbf{X}^H$, after grouping terms we obtain that $\frac{1}{M}\log L_k(\mathbf{X}, \mathcal{M}) = \mathrm{tr}\left(\left((\mathbf{R}_s[\hat{\boldsymbol{\gamma}}_{\bar{k}}] + \hat{\sigma}_0^2 \mathbf{I})^{-1} - (\hat{\gamma}_k \mathbf{R}_k + \boldsymbol{\Xi}_k)^{-1}\right) \mathbf{R}_x\right) + \log\det\left(\frac{\hat{\mathbf{R}}_s[\hat{\boldsymbol{\gamma}}_{\bar{k}}] + \sigma_0^2}{\boldsymbol{\Xi}_k}\right) - \log\det\left(\mathbf{I} + \boldsymbol{\Xi}_k^{-1}\hat{\gamma}_k \mathbf{R}_k\right)$. On the one hand, we consider the approximations $\hat{\sigma}_0^2 \approx \hat{\sigma}_1^2$ and $\log\det\left(\mathbf{I} + \boldsymbol{\Xi}_k^{-1}\hat{\gamma}_k \mathbf{R}_k\right) \approx \mathrm{tr}(\boldsymbol{\Xi}_k^{-1}\hat{\gamma}_k \mathbf{R}_k) \approx 0$, asymptotically as $\rho_k \to 0$. On the other hand, by recalling the matrix inversion lemma $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1}$ we obtain that $(\hat{\gamma}_k \mathbf{R}_k + \boldsymbol{\Xi}_k^{-1}) = \boldsymbol{\Xi}_k^{-1} - \boldsymbol{\Xi}_k^{-1}\hat{\gamma}_k \mathbf{R}_k(\hat{\gamma}_k \mathbf{R}_k + \boldsymbol{\Xi}_k)^{-1}$. Applying this last result to the GLRT, we can approximate

$$\frac{1}{M}\log L_k(\mathbf{X}, \mathcal{M}) \approx \mathrm{tr}\left(\mathbf{K}_k \mathbf{R}_x\right). \tag{12}$$

where $\mathbf{K}_k \doteq \boldsymbol{\Xi}_k^{-1}\hat{\gamma}_k \mathbf{R}_k(\hat{\gamma}_k \mathbf{R}_k + \boldsymbol{\Xi}_k)^{-1}$. For the second part of the proof, we recall that for large data records (i.e., as $N \to \infty$) in (2), it is established in [10, Ch. 5, Sec. 5] though approximating the PDF of $\mathbf{X}$ that log-likelihood decision statistic (12) can be approximated as (4), where $\mathcal{K}_k(\omega, \Theta)$ is the asymptotic continuous-frequency transform of $\mathbf{K}_k$, and $P(\omega)$ is the continuous-frequency periodogram of $\mathbf{X}$.

Appendix B
Proof of Theorem 2

The ML estimate of the power levels $\boldsymbol{\gamma} \doteq (\gamma_1, \ldots, \gamma_K)$ for the multi-frequency model (1) reads $\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \log\det\left(\sum_{k=1}^{K} \gamma_k \mathbf{R}_k + \sigma^2 \mathbf{I}\right) + \mathrm{tr}\left(\left(\sum_{k=1}^{K} \gamma_k \mathbf{R}_k + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_x\right)$, with the additional constraint $\boldsymbol{\gamma} \succeq \mathbf{0}$. The problem is convex on $\boldsymbol{\gamma}$ and we take the derivative of the objective with respect to $\gamma_{k'}$ and equal it to zero, which gives the following equation $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\gamma}] + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_{k'}\right) - \mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\gamma}] + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_{k'} \left(\mathbf{R}_s[\boldsymbol{\gamma}] + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_x\right) = 0$. Solving for $\boldsymbol{\gamma}$ requires the use of iterative algorithms which do not provide interpretation insight on the solution. However, in the wideband regime, we further make use of the following approximations $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\gamma}] + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_{k'}\right) \approx \frac{1}{\sigma^2}\mathrm{tr}(\mathbf{R}_{k'}) - \frac{1}{\sigma^4}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\gamma}]\mathbf{R}_{k'}\right)$, and $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\gamma}] + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_{k'} \left(\mathbf{R}_s[\boldsymbol{\gamma}] + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_x\right) \approx \frac{1}{\sigma^4}\mathrm{tr}(\mathbf{R}_{k'}\mathbf{R}_x) - \frac{2}{\sigma^6}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\gamma}]\mathbf{R}_{k'}\mathbf{R}_x\right)$, as $\frac{\gamma_n}{\sigma^2} \to 0$. Applying the former approximations, we obtain that the ML estimate of $\mathbf{R}_s$ accomplishes $\frac{2}{\sigma^6}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\gamma}]\mathbf{R}_{k'}\mathbf{R}_x\right) - \frac{1}{\sigma^4}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\gamma}]\mathbf{R}_{k'}\right) = \frac{1}{\sigma^4}\mathrm{tr}(\mathbf{R}_{k'}\mathbf{R}_x) - \frac{1}{\sigma^2}\mathrm{tr}(\mathbf{R}_{k'})$. We further make use of the low-SNR approximation with the left-hand side of the former equation. Noting that $\mathrm{tr}(2\mathbf{R}_x - \sigma^2 \mathbf{I}) \approx \sigma^2 \mathbf{I}$, it reduces, after multiplying both sides by σ^2, to $\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\gamma}]\mathbf{R}_{k'}) = \mathrm{tr}(\mathbf{R}_{k'}\mathbf{R}_x) - \sigma^2\mathrm{tr}(\mathbf{R}_{k'})$. By placing the K equations in a matrix form, we proof (7). By taking (12) with $\boldsymbol{\Xi}_k = \sum_{l \neq k} \hat{\gamma}_l \mathbf{R}_l + \sigma^2 \mathbf{I}$, we prove (6).

Appendix C
Proof of Theorem 3

The ML estimates of the power levels $\boldsymbol{\gamma}$ and the noise variance under $\mathcal{H}_1$, $\hat{\sigma}_1^2$, are given by the optimization problem $\hat{\boldsymbol{\gamma}}, \hat{\sigma}_1^2 = \arg\min_{\boldsymbol{\gamma}} \log\det\left(\sum_{k=1}^{K} \gamma_k \mathbf{R}_k + \sigma^2 \mathbf{I}\right) + \mathrm{tr}\left(\left(\sum_{k=1}^{K} \gamma_k \mathbf{R}_k + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{R}_x\right)$. Under the low-SNR assumption, the derivative with respect to the power levels is given, similarly to Appendix B, by the set of equations $\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\gamma}]\mathbf{R}_{k'}) + \sigma^2\mathrm{tr}(\mathbf{R}_{k'}) = \mathrm{tr}(\mathbf{R}_{k'}\mathbf{R}_x)$. To the K former equations, we add the derivative with respect to the noise variance, which after analogous manipulations, leads to $\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\gamma}]) + \sigma^2 N = \mathrm{tr}(\mathbf{R}_x)$. The $K+1$ equations in vectorial notation are then given by (10). Finally, taking (12) with $\boldsymbol{\Xi}_k = \sum_{l \neq k} \hat{\gamma}_l \mathbf{R}_l + \hat{\sigma}_1^2 \mathbf{I}$, proves (9).

References

[1] I. Mitola, J. and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[3] A. Ghasemi and E. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 32–39, Apr. 2008.

[4] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *Proc. IEEE Int. Symp. New Frontiers on Dyn. Spectr. Access (DySPAN)*, Nov. 2005, pp. 151–159.

[5] P. Sutton, K. Nolan, and L. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 13–24, Jan. 2008.

[6] J. G. Proakis, *Digital Communications*, 4th ed., ser. Science/Engineering/Math. McGraw-Hill, Aug. 2000.

[7] A. Sahai and D. Cabric, "Spectrum sensing: fundamental limits and practical challenges," in *Proc. IEEE Int. Symp. New Frontiers in Dynamic Spectr. Access Networks (DySPAN)*, Nov. 2005.

[8] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitve radios: a GLRT approach," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.

[9] 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.

[10] S. M. Kay, *Fundamentals of statistical signal processing*. Upper Saddle River, NJ: Prentice Hall, 1998, vol. 2 (detection theory).

[11] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[12] J. Villares and G. Vazquez, "The Gaussian assumption in second-order estimation problems in digital communications," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4994–5002, Oct. 2007.

[13] G. Moustakides, "Finite sample size optimality of GLR tests," *IEEE Trans. Inf. Theory*, 2009, submitted. [Online]. Available: http://arxiv.org/abs/0903.3795

[14] H. V. Poor, *An introduction to signal detection and estimation*. New York, NY: Springer-Verlag, 1994.