Robust, Non-Gaussian Wideband Spectrum Sensing in Cognitive Radios

Mario Bkassiny, *Member, IEEE* and Sudharman K. Jayaweera, *Senior Member, IEEE*

***Abstract*—In this paper, we propose detectors (both parametric and robust) for wideband spectrum sensing in cognitive radios (CR's). The proposed detectors are able to detect spectral activity over a wide frequency range, while assuming little knowledge about the signals of interest. The parametric detector is based on a locally optimal (LO) Neyman-Pearson (NP) test and assumes a known non-Gaussian noise distribution. The corresponding decision statistic of the LO NP detector is expressed in frequency domain, allowing to identify the active channels within the wide frequency band of interest. On the other hand, for situations in which the noise distribution is only *approximately* known, we propose a robust signal detector that is immune to deviations of the noise model from a certain nominal distribution. The proposed wideband robust detector is based on a robust spectral estimator and is formulated as a non-linear regression. This regression problem can be solved using a fixed-point iteration algorithm at a quadratic computational complexity, in contrast with the Newton's method which would have cubic complexity order. The simulation results show that the proposed detectors can achieve better detection performance in the presence of non-Gaussian noise, compared to existing detectors under the same conditions.**

***Index Terms*—Cognitive radio, Huber cost function, locally most powerful test, ridge regression, robust detection, spectral estimation, wideband spectrum sensing.**

I. Introduction

Accurate spectrum sensing is crucial to the successful operation of cognitive radios (CR's) [1]–[5]. In the context of dynamic spectrum sharing CR's, various sensing techniques have been proposed for detecting primary signals [4]–[6]. In these mostly narrowband applications, a CR senses a particular channel to identify the existence of a particular signal. In this case, the decision-making reduces to a binary hypothesis testing to determine whether a particular channel is idle or busy [7]–[9]. In the context of wideband CR's (W-CR's), however, a CR not only has to determine the existence of active signals, but it has to identify the location of those detected signals within a wide frequency range [10]–[12]. As noted in [6], a CR sensing a wide frequency band should be able to 1) estimate the number of signals in the sensed wideband, 2) determine the operating frequency band of each signal and 3) estimate the power level of each signal.

As long as noise is Gaussian distributed, the energy detection is optimal if there is no further knowledge on the signals in the RF environment [13]–[15]. Thus, commonly the power spectral density (PSD) of the received signal is computed using the fast Fourier transform (FFT), and a Neyman-Pearson (NP) test is applied to detect the active signals within the sensed spectrum from the estimated PSD [11], [12], [16]. However, the assumption of W-CR's requires a rethinking of this spectrum sensing architecture. Indeed, a wideband spectrum sensing detector may be subjected to wide range of heterogeneous spectral activities, giving rise to outliers and interferers. Such electromagnetic interference that results from a variety of sources has been shown to be non-Gaussian [17]. Thus, the commonly assumed Gaussian noise model may not be justifiable, requiring detector designs that take into account possible deviations of the noise model from a certain nominal distribution [17], [18]. In particular, signal detectors have been proposed for cases in which the noise model has a *known* non-Gaussian distribution such as the Gaussian-Laplace mixture distribution [19], [20], and the Middleton Class A and Class B models [21]. On the other hand, for situations in which the nominal noise model is subject to an *unknown* contaminating distribution, robust detectors are proposed to combat such noise uncertainty [15], [22]. In this case, noise can be modeled using an ϵ-contaminated distribution F denoted as $F = (1-\epsilon)P + \epsilon M$, where $0 \leq \epsilon < 1$ is a known parameter, P is a known nominal distribution (e.g. Gaussian), and M is an unknown contaminating distribution [15], [22]. Usually, these detectors are formulated in time domain assuming narrowband signals, whereas in W-CR applications one needs to consider signals scattered over a range of frequencies [6], [10]–[12].

Although literature on signal detection and spectral estimation is quite mature [15], [18], [23], [24], it has not attacked a problem as challenging as attempting to estimate a very wide power spectrum containing multiple signals of different characteristics in near real-time, as may be the case in W-CR's. There are wideband spectral estimation approaches that are mostly aimed at non-real-time operation [25]. In CR communications, however, the wideband spectral estimation must be performed in real-time. Compressive sensing has been considered as a promising solution for wideband spectrum sensing in the past [26], [27]. However, most of the compressive sensing formulations are not aimed at non-Gaussian and contaminated noise models. For example, the wideband spectral estimation based on compressive sensing proposed in [27], [28] are based on a least-squares spectral estimation which is not suitable under non-Gaussian noise models [13], [29]. In addition, compressive sensing methods assume sparse signals, which may not always be a justifiable assumption in practice [26].

M. Bkassiny is with the Department of Electrical and Computer Engineering, State University of New York at Oswego, Oswego, NY, USA, Email: mario.bkassiny@oswego.edu.

S. K. Jayaweera is with the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, USA, Email: jayaweera@ece.unm.edu.

In this paper, in order to account for non-Gaussian noise, we first propose a signal detector by assuming that the signals of interest are subject to a noise process having a *known* arbitrary distribution [15]. This is a parametric non-Gaussian detector since it assumes a known noise distribution that is completely specified by a finite number of parameters. The corresponding detection criterion is based on a locally optimal (LO) NP test, which does not require explicit knowledge of the signals of interest, making it appropriate for wideband CR applications [10]–[12]. Another advantage of the derived LO detector is that it reduces to a decision statistic that is a second order statistic of the data, which can easily be estimated from the signal periodogram.

In contrast with [18], [30], the derived decision statistic of our proposed LO NP detector is next transformed into frequency domain by expressing it as a correlation between a spectral estimation function and the PSD of the assumed signal. Hence, by exploiting the spectral characteristics of the decision statistic, it becomes possible to identify the operating frequencies of the detected signals, as required in wideband spectrum sensing [6]. We also show that the spectral estimation is to be performed following a non-linear transformation of the signal samples, which depends on the assumed known non-Gaussian noise distribution. However, in many scenarios the exact noise distribution may not be known a priori, requiring a robust approach to address the noise uncertainty.

Hence, for situations in which the noise distribution is only *approximately* known [15], [22], we propose a robust wideband signal detection method based on robust spectral estimation. This spectral estimation is formulated using a modified ridge regressor, in which we replace the quadratic loss function by the Huber cost function [22]. The Huber cost function can lead to the *most robust estimator* when the contaminating noise distribution belongs to the set of all symmetric distributions [22]. We formulate the modified ridge regressor as a non-linear convex optimization which can be solved numerically using Newton's method at a cubic complexity order (due to the inverse operation of the Hessian matrix) [31]. We propose a fixed-point iteration algorithm which reduces the computational complexity to quadratic order. Using the Banach fixed-point theorem, we show that the proposed algorithm indeed converges to the optimal solution [32]. We also compute the receiver operating characteristics (ROC's) of both parametric and robust signal detectors and show their superior performance in the presence of non-Gaussian noise, compared to similar existing detection methods.

Before delving into the discussion of robust spectral estimation and signal detection, we first outline a framework for wideband spectrum sensing [10], [11], [33]. The RF front-end of the CR is assumed to be capable of sensing several sub-bands within the wide spectrum band of interest. The sub-bands can be specified based on a spectrum segmentation method which determines, for example, the bandwidth of each sub-band as a function of hardware constraints such as the maximum rate of the analog-to-digital converters (ADC's).

The remainder of this paper is organized as follows: In Section II, we present an RF front-end design for wideband spectrum scanning in W-CR's. A wideband signal detector for non-Gaussian noise models is proposed in Section III. Next, in Section IV, we propose a robust spectral estimator to perform signal detection in the presence of uncertainty about the noise distribution. Simulations results are presented in Section V and we conclude the paper in Section VI.

II. A Wideband Spectrum Sensing Front-end

Spectrum sensing in W-CR's are aimed at monitoring a wide frequency band spanning several possibly non-contiguous spectrum portions. Thus, the assumed CR architecture consists of an RF front-end with wideband capabilities and a signal processing module. Both RF front-end and signal processing module are controlled by a cognitive engine (CE) that is equipped with learning and decision-making capabilities, as proposed in the Radiobot architecture in [10].

In order to scan a large spectrum range, we may use reconfigurable antennas that can cover different frequency bands $W_1, \cdots, W_R \subset \mathbb{R}$ [34], [35]. When reconfigured to sense frequency band W_l ($l \in \{1, \cdots, R\}$), we denote the bandwidth of the scanned signal as $|W_l| > 0$. Since the bandwidths $|W_l|$ are considered to be wideband, the direct processing of the corresponding signals might not be practical. This requires us to segment those frequency bands into smaller sub-bands prior to processing. The sensing reconfigurable antenna is connected to a bank of band-pass filters allowing proper segmentation of each of the frequency bands $W_1, \cdots, W_R$ in to sets of sub-bands. Each sub-band is still wide enough to contain many communication channels belonging to possibly different systems. At each time instant, the W-CR is assumed to be able to sense and process a single sub-band. The problem of how to select which sub-band to be sensed is beyond the scope of this paper and is addressed in [33], [36]. The focus of this paper is on how to process a scanned sub-band signal to detect any, and all, spectral activities located at different frequency locations within the sub-band when noise characteristics over such a wide range are uncertain and, specifically, non-Gaussian.

There are several issues that need to be taken into account in determining the optimal number of sub-bands in each frequency band. These include the maximum sampling rate of the ADC, the required quantization accuracy and the available a priori knowledge of traffic characteristics on the spectrum bands of interest. This problem, however, is out of the scope of this paper and will be addressed in future research work.

III. Wideband Signal Detection in Non-Gaussian Noise

Signal detection is one of the most challenging problems in CR's since it requires accurate estimation of the ongoing RF activity in real-time. In many such problems, it is assumed that both signal and noise are Gaussian, a condition which may not be justified in general. For example, when sensing a wide frequency band, the detector might be subject to different types of interference, jamming and other impulsive noise [15], [18]. Any misrepresentation of such RF activities may lead to degradation in the detection performance [5]. In order to address this issue, we resort to non-Gaussian distributions

that can better model outliers and interferes in the sensed wide frequency band of interest. We are particularly interested in noise models whose distributions decay at a lower rate, compared to that of the Gaussian, such as the Laplace, Cauchy and Gaussian-Laplace mixture noise distributions [19], [20].

In order to address these issues, we consider a sensed sub-band signal $\mathbf{Y} = [Y_1, \cdots, Y_N]^T$ that is subject to non-Gaussian noise, such that:

$$\mathbf{Y} = \theta \mathbf{S} + \mathbf{W},$$

where $\mathbf{W} = [W_1, \cdots, W_N]^T \sim f_{\mathbf{W}}(\mathbf{w})$ is the noise vector and $\mathbf{S} = [S_1, \cdots, S_N]^T \sim f_{\mathbf{S}}(\mathbf{s})$ is the signal vector. The signal detection problem can thus be formulated as the following composite hypothesis testing:

$$\begin{array}{rcll} \theta \in \Theta_0 & = & \{\theta_0\} & \text{under } \mathcal{H}_0 \\ \theta \in \Theta_1 & = & (\theta_0, \infty) & \text{under } \mathcal{H}_1 \end{array} , \tag{1}$$

where we set $\theta_0 = 0$ such that $\mathcal{H}_0$ and $\mathcal{H}_1$ correspond, respectively, to the signal absent and present hypotheses [15]. We denote the mean and covariance matrix of the signal $\mathbf{S}$ by $\bar{\mathbf{s}}$ and $\Sigma_{\mathbf{S}}$, respectively. It is important to emphasize that beyond these second order moments we do not make any other assumptions on the signal, and the noise distribution $f_{\mathbf{W}}$ is completely arbitrary, but must be *known*. Note that, θ is used to account for the fluctuations in the received signal magnitude, and it is assumed to be an unknown parameter. We define $L(\mathbf{y}, \theta) \triangleq \frac{f(\mathbf{y}|\theta \in \Theta_1)}{f(\mathbf{y}|\theta \in \Theta_0)}$ as the likelihood ratio (LR) between the simple hypothesis $\mathcal{H}_0$ and the composite hypothesis $\mathcal{H}_1$. In this case, since Θ_0 is a singleton, a simple solution for the composite hypothesis testing problem can be obtained, as shown below.

We use the NP criterion to find a decision rule δ that maximizes the detection probability $P_D(\delta, \theta)$ such that the false alarm probability $P_F(\delta, \theta)$ is less than a certain threshold $\alpha \in [0, 1]$:

$$\max_{\delta} \quad P_D(\delta, \theta) \quad \text{subject to} \quad P_F(\delta, \theta) \leq \alpha \ . \tag{2}$$

Since detection probability $P_D(\delta, \theta)$ is called the *power* of δ, the goal of the NP criterion is to find the most powerful α-level test of $\mathcal{H}_0$ versus $\mathcal{H}_1$ [15]. However, it may not be always possible to find a uniformly most powerful (UMP) test for all values of θ, since, in general, the critical region $\Gamma_\theta \triangleq \{\mathbf{y} \in \mathbb{R}^N | L(\mathbf{y}, \theta) > \tau\}$ may not be independent of θ (where τ is chosen to give a false alarm rate of α) [15].

An alternative solution is to find a locally most powerful (LMP) test (also referred to as the LO test). An LMP test can be motivated by expanding $P_D(\delta, \theta)$ in a Taylor series around $\theta = \theta_0$:

$$\begin{aligned} P_D(\delta, \theta) = P_D(\delta, \theta_0) + (\theta - \theta_0) P'_D(\delta, \theta_0) + \\ + \frac{1}{2}(\theta - \theta_0)^2 P''_D(\delta, \theta_0) + \mathcal{O}\left((\theta - \theta_0)^3\right) \ , \end{aligned} \tag{3}$$

where we have denoted

$$P'_D(\delta, \theta_0) = \left. \frac{\partial}{\partial \theta} \left[P_D(\delta, \theta) \right] \right|_{\theta = \theta_0} , \tag{4}$$

and

$$P''_D(\delta, \theta_0) = \left. \frac{\partial^2}{\partial \theta^2} \left[P_D(\delta, \theta) \right] \right|_{\theta = \theta_0} . \tag{5}$$

Note that the first term in the Taylor series expansion $P_D(\delta, \theta_0)$ is the false alarm probability and is constant for all θ. Hence, maximizing $P_D(\delta, \theta)$ near $\theta = \theta_0$ is equivalent to maximizing the derivative terms, in particular, $P'_D(\delta, \theta_0)$. It can be shown that the LMP decision rule that indeed maximizes a first-order approximation of the detection probability in (4) is equivalent to [15]:

$$\delta_{LO}(\mathbf{y}) = \left\{ \begin{array}{lll} 1 & & > \\ \gamma_{LO} & \text{if } \frac{\partial}{\partial \theta} L(\mathbf{y}, \theta)\big|_{\theta = \theta_0} & = \ \eta \\ 0 & & < \end{array} \right. \tag{6}$$

The parameters $\eta \geq 0$ and $0 \leq \gamma_{LO} \leq 1$ are chosen so that $P_F(\delta_{LO}) = \alpha$. We can show that the LO decision statistic of the detector (6) for the problem (1) is given by:

$$T_1^{(LO)}(\mathbf{y}) = \frac{\partial}{\partial \theta} L(\mathbf{y}, \theta) \mid_{\theta = 0} = -\frac{\bar{\mathbf{s}}^T \mathbf{f}'_{\mathbf{W}}(\mathbf{y})}{f_{\mathbf{W}}(\mathbf{y})} \ , \tag{7}$$

where $\mathbb{E}\{\mathbf{S}\} = \bar{\mathbf{s}}$ and $\mathbf{f}'_{\mathbf{W}} = \frac{\partial f_{\mathbf{W}}}{\partial \mathbf{w}} = \left(\frac{\partial f_{\mathbf{W}}}{\partial w_1}, \frac{\partial f_{\mathbf{W}}}{\partial w_2}, \cdots, \frac{\partial f_{\mathbf{W}}}{\partial w_N} \right)^T$. Note that, if the decision statistic $T_1^{(LO)}(\mathbf{Y})$ has a continuous probability distribution, then $Pr\left\{ T_1^{(LO)}(\mathbf{Y}) = \eta \right\} = 0$ and the randomization γ_{LO} in (6) can be chosen arbitrarily [15].

On the other hand, the threshold η in (6) can be defined such that $Pr\left\{ T_1^{(LO)}(\mathbf{Y}) > \eta | \mathcal{H}_0 \right\} = \alpha$. However, if the probability distribution of $T_1^{(LO)}(\mathbf{Y})$ under $\mathcal{H}_0$ cannot be obtained analytically, the threshold η can be computed numerically, as described in Algorithm 1.

From (7), it follows that if $\bar{\mathbf{s}} = \mathbf{0}$, the first order LO test (6) does not work since $T_1^{(LO)}(\mathbf{y}) = 0$ for all $\mathbf{y}$. In this case, however, a similar LO decision statistic can be obtained by maximizing instead the second order derivative $P''_D(\delta, \theta_0)$ in (3) [18]. It can be shown that the general second order LO decision statistic is given by:

$$T_2^{(LO)}(\mathbf{y}) = \left. \frac{\partial^2}{\partial \theta^2} L(\mathbf{y}, \theta) \right|_{\theta = 0} = \frac{\mathrm{Tr}\left(\mathbf{F}''_{\mathbf{W}}(\mathbf{y}) \left(\Sigma_{\mathbf{S}} + \bar{\mathbf{s}} \bar{\mathbf{s}}^T \right) \right)}{f_{\mathbf{W}}(\mathbf{y})} \ ,$$

where we have denoted $\mathbf{F}''_{\mathbf{W}}(\mathbf{w}) = \frac{\partial}{\partial \mathbf{w}} \mathbf{f}'_{\mathbf{W}}(\mathbf{w})$, and we have used the fact that $\mathbb{E}\left\{ \mathbf{S}\mathbf{S}^T \right\} = \Sigma_{\mathbf{S}} + \bar{\mathbf{s}}\bar{\mathbf{s}}^T$. If $\bar{\mathbf{s}} = \mathbf{0}$, then $\mathbb{E}\left\{ \mathbf{S}\mathbf{S}^T \right\} = \Sigma_{\mathbf{S}}$ and the decision statistic $T_2^{(LO)}(\mathbf{y})$ depends only on the probability distribution of the noise samples W_k and the second order statistics of the signal $\mathbf{S}$, without requiring the knowledge of the exact distribution of $\mathbf{S}$.

Furthermore, if we assume that the noise samples are independent (as in [18]–[20], [29], [37]), $T_2^{(LO)}(\mathbf{y})$ reduces to the following [18]:

$$\begin{aligned} T_2^{(LO)}(\mathbf{y}) &= \sum_{k=1}^{N} h_k^{(LO)}(y_k) \rho_{k,k}^{(s)} + \\ &+ \sum_{k=1}^{N} \sum_{j=1, j \neq k}^{N} g_j^{(LO)}(y_j) g_k^{(LO)}(y_k) \rho_{j,k}^{(s)} \ , \end{aligned} \tag{8}$$

where $\rho_{j,k}^{(s)}$ is the (j,k)-th element of the covariance matrix $\Sigma_{\mathbf{S}}$ and

$$g_k^{(LO)}(x) = -\frac{f'_{W_k}(x)}{f_{W_k}(x)}, \quad (9)$$

and

$$h_k^{(LO)}(x) = \frac{f''_{W_k}(x)}{f_{W_k}(x)}. \quad (10)$$

Furthermore, if the noise sequence W_k is i.i.d., then we may drop the indices k from both $g_k^{(LO)}(x)$ and $h_k^{(LO)}(x)$ and denote:

$$g^{(LO)}(x) = -\frac{f'_W(x)}{f_W(x)}, \quad (11)$$

and

$$h^{(LO)}(x) = \frac{f''_W(x)}{f_W(x)}. \quad (12)$$

In this case, $T_2^{(LO)}(\mathbf{y})$ in (8) can be also expressed as a function of $g^{(LO)}(x)$ and $h^{(LO)}(x)$. In Appendix A, we derive the expressions of $g^{(LO)}(x)$ and $h^{(LO)}(x)$ for both Gaussian-Laplace and Gaussian mixture noise models.

An efficient implementation of the non-Gaussian detector based on decision statistic $T_2^{(LO)}(\mathbf{y})$ can be facilitated by expressing it in frequency domain. For this, we assume that the signal $\mathbf{S}$ is wide-sense stationary (WSS) with a mean $\bar{\mathbf{s}} = \mathbf{0}$ and the covariance matrix $\Sigma_{\mathbf{S}}$ is Toeplitz so that $\rho_{k,l}^{(s)} = \rho_{k-l,0}^{(s)} \triangleq \rho_{k-l}^{(s)}$. This is usually a valid assumption for communication signals for which a PSD exists [15], [38]. With these assumptions, we can express the second order LO decision statistic $T_2^{(LO)}(\mathbf{y})$ of (8) as:

$$T_2^{(LO)}(\mathbf{y}) = T_{2,0}(\mathbf{y}) + \sum_{k=1}^{N}\sum_{l=1}^{N} \tilde{y}_l \tilde{y}_k \rho_{l-k}^{(s)}, \quad (13)$$

where $T_{2,0}(\mathbf{y}) \triangleq \rho_0^{(s)} \left[\sum_{k=1}^{N} \left(\hat{y}_k - \tilde{y}_k^2\right)\right]$, $\tilde{y}_j \triangleq g_j^{(LO)}(y_j)$ and $\hat{y}_j \triangleq h_j^{(LO)}(y_j)$.

Let us denote by $\phi_s(F)$ the Fourier transform of the sequence $\{\rho_k^{(s)}\}_{k=0}^{N-1}$, so that $\rho_k^{(s)} = \int_{-1/2}^{1/2} \phi_s(F) e^{j2\pi Fk} dF$, for $k = 0, \cdots, N-1$. Substituting this in (13) we obtain:

$$\begin{aligned} T_2^{(LO)}(\mathbf{y}) &= T_{2,0}(\mathbf{y}) + \sum_{k=1}^{N}\sum_{l=1}^{N} \tilde{y}_l \tilde{y}_k \int_{-1/2}^{1/2} \phi_s(F) e^{j2\pi F(l-k)} dF \\ &= T_{2,0}(\mathbf{y}) + \\ &+ \int_{-1/2}^{1/2} \left[\left(\sum_{k=1}^{N} \tilde{y}_k e^{-j2\pi Fk}\right)\left(\sum_{l=1}^{N} \tilde{y}_l e^{-j2\pi Fl}\right)^*\right] \phi_s(F) dF \\ &= T_{2,0}(\mathbf{y}) + \int_{-1/2}^{1/2} |\tilde{Y}(F)|^2 \phi_s(F) dF, \end{aligned} \quad (14)$$

where (14) follows by assuming that $\tilde{y}_k \in \mathbb{R}$ and denoting $\tilde{Y}(F) \triangleq \sum_{k=0}^{N-1} \tilde{y}_{k+1} e^{-j2\pi Fk}$ for $-1/2 \le F \le 1/2$. Hence, the decision statistic of the LO detector can be expressed as:

$$T_2^{(LO)}(\mathbf{y}) = T_{2,0}(\mathbf{y}) + \int_{-1/2}^{1/2} \hat{\phi}(F) \phi_s(F) dF, \quad (15)$$

where $\hat{\phi}(F) = |\tilde{Y}(F)|^2$ is the periodogram of the transformed observation sequence $\{\tilde{y}_k\}_{k=1}^N$. From (15) we see that the non-Gaussian detector can be implemented by applying a certain non-linearity to the signal sequence $\{y_k\}_{k=1}^N$ and then computing the corresponding periodogram $|\tilde{Y}(F)|^2$. The detector essentially correlates this periodogram of the transformed observations with the known signal spectrum.

Note that, if we assume an i.i.d. Gaussian noise W_k with a pdf $f_W(x) = \mathcal{N}(0, \sigma^2)(x)$, we can express:

$$g^{(LO)}(x) = \frac{x}{\sigma^2}, \quad (16)$$

and

$$h^{(LO)}(x) = \left(\frac{x^2}{\sigma^4} - \frac{1}{\sigma^2}\right). \quad (17)$$

Under this noise assumption, we obtain $\tilde{Y}(F) = \frac{1}{\sigma^2} Y(f)$ and $T_{2,0}(\mathbf{y}) = -\frac{N\rho_0^{(s)}}{\sigma^2}$. In this case, the decision statistic in (15) reduces to:

$$T_2^{(LO)}(\mathbf{y}) = -\frac{N\rho_0^{(s)}}{\sigma^2} + \frac{1}{\sigma^4} \int_{-1/2}^{1/2} |Y(F)|^2 \phi_s(F) dF. \quad (18)$$

Hence, in the case of Gaussian noise, the decision statistic will be based on the conventional periodogram $|Y(F)|^2$ of the original data sequence $\{y_k\}_{k=1}^N$. This is expected since energy detection (or the periodogram-based detection) is known to be optimal when noise is Gaussian and no information beyond second order statistics are available about the signal [14], [15].

However, when the noise model deviates from the Gaussian assumption, the energy-based detection becomes non-optimal. Thus, more sophisticated detectors, such as the one characterized in the non-linear decision statistic in (15), are then required. As a particular example, let us consider the Laplace noise model with $\epsilon = 1$ in (33). We obtain $\tilde{Y}(F) = \lambda \sum_{k=1}^{N} sgn(y_k) e^{-j2\pi Fk}$ and $T_{2,0}(\mathbf{y}) = -2\lambda\rho_0^{(s)} \sum_{k=1}^{N} \delta^{(D)}(y_k)$. The observations y_k's are continuous random variables with $\Pr\{y_k = 0\} = 0$ under the Laplace noise assumption, making $\Pr\{T_{2,0}(\mathbf{y}) = 0\} = 1$. The decision statistic is then based on the discrete Fourier transform (DFT) of the sequence of signs of y_k's. Hence, clearly the conventional periodogram $|Y(F)|^2$ is not suitable for signal detection under Laplace noise assumption.

Similarly, it can be seen from (36) and (37) that, under the Gaussian mixture noise distribution, the LO detector requires us to modify the conventional periodogram. Note that the above formulation assumed that the second order statistics of the signal are all known. This also implies that we know the signal power spectra $\phi_s(F)$. Thus, the optimal spectral activity detection of (15) can be done by correlating the estimated periodogram (or spectral estimation function) $\hat{\phi}(F)$ with each signals known power spectrum $\phi_s(F)$ and comparing to a threshold. The threshold design is to be according to the NP criterion subject to a maximum false alarm rate (as discussed in Section V).

However, in the context of wideband spectrum sensing, it is more likely that the signal covariance structure, and correspondingly the power spectrum $\phi_s(F)$, may not be known.

In this case, we can systematically develop approximations to the PSD $\phi_s(F)$ by making certain reasonable assumptions on signals. Note that, this is perhaps the justification for many existing approaches that essentially search for the peaks in the estimated periodogram. It is clear that one may arrive at that method by making the assumption that $\phi_s(F)$ is a set of impulses characterizing a set of sinusoids.

Another case is the assumption of piecewise constant $\phi_s(F)$. For example, we may not know the covariance structure exactly (i.e. we do not know the location or the shape of the spectrum), but we may know that the signal bandwidth is less than B. Then a reasonable detector can be motivated as a sliding window PSD integrator of width B followed by thresholding, similar to [11], [12]. In Section V, we will consider both impulse and rectangular shapes for $\phi_s(F)$ and analyze the detection performance under each assumption.

Note that, the sliding-window technique can improve the detection rate if the sliding window size is smaller than the signal bandwidth. In particular, if we consider a signal with average power P_s and a normalized bandwidth $B\Delta f$ (where $\Delta f = 1/N$ is the frequency increment and B is an even integer), the detection probability of the corresponding periodogram-based detector can be obtained as in (47) of Appendix B:

$$P_D = Q_L\left(\sqrt{\frac{N\min\{L, B+1\}P_s}{(B+1)\sigma^2}}, \sqrt{\eta}\right), \tag{19}$$

where L (odd integer) is the sliding window length and σ^2 is the average Gaussian noise power. Thus, if $L \leq B+1$, the sliding-window technique can improve the detection probability since it enhances the signal power by a factor L. On the other hand, if the window length is larger than the signal bandwidth (i.e. $L > B+1$), then the detection probability will decrease with L due to the increase in the threshold level η in (19). These results are illustrated in Fig. 4.

An important implementation issue is the computation of the robust spectral estimation function $\hat{\phi}(F)$. As we have shown above, this function is equal to $|\tilde{Y}(F)|^2$ and can be computed as the periodogram of the transformed sequence $\{\tilde{y}_k\}_{k=1}^N$. In practice, however, $\tilde{Y}(F)$ can be evaluated at discrete frequencies using the DFT:

$$\tilde{Y}_i = \sum_{k=0}^{N-1} \tilde{y}_{k+1} e^{-j2\pi i \frac{k}{N}} \text{ , for } i = 0, \cdots, N-1, \tag{20}$$

where we assume a known parametric distribution for the noise model to compute $\tilde{y}_k$'s. Such parametric assumptions may not, however, be valid, in general, since the noise model may be completely unknown or subject to an unknown contamination [15], [22]. In this case, we may not be able to evaluate the transformed sequence $\tilde{y}_k = \mathbf{g}^{(LO)}(y_k)$ in (20). Hence, in the next section, we consider a robust approach to obtain a spectral estimation for $\hat{\phi}(F)$ by assuming an ϵ-contaminated noise distribution [22].

IV. Robust Wideband Spectrum Sensing

As we have seen from the development in the previous section, closed-form expressions for $\hat{\phi}(F)$ can be obtained only for known parametric noise models (e.g. the Gaussian-Laplace noise mixture for which an expression for $g_k^{(LO)}(x)$ can be obtained as in (34)). Even then, such spectral estimation might be distorted if the actual noise distribution deviates from the assumed nominal model [15]. To address such uncertainty, one may resort to robust spectral estimation approaches. For instance, the periodogram provides a powerful non-parametric tool in spectral analysis [39]. However, its lack of robustness against outliers and heavy-tailed noise is well-known [39]. In particular, the periodogram of a given observation vector $\mathbf{y} \in \mathbb{C}^N$ can be obtained using the following least-squares regression [39]:

$$\mathbf{v}^* = \arg\min_{\mathbf{v}\in\mathbb{C}^N} \|\mathbf{y} - \mathbf{X}^H\mathbf{v}\|^2, \tag{21}$$

where $\mathbf{v}^* \triangleq [v_1^*, \cdots, v_N^*]^T$ and $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]$ with $\mathbf{x}_k = \frac{1}{N}[1, e^{-j2\pi\frac{(k-1)}{N}}, \cdots, e^{-j2\pi\frac{(N-1)(k-1)}{N}}]^T$ for $k = 1, \cdots, N$. The resulting periodogram $|v_i^*|^2$ (for $i = 1, \cdots, N$) is sensitive to outliers due to the quadratic loss function in (21) [39]. This quadratic loss function can be modified, for example, by including a regularization term to improve the smoothness and generalization ability of the optimal solution, as in the support vector machine (SVM) and ridge regression formulations [40]. This can help to reduce the statistical variability of the spectral estimation [39]. In particular, the ridge regressor computes the optimal vector $\mathbf{v}^*$ that minimizes the weighted sum of $\|\Lambda\mathbf{v}\|^2$ (where Λ is a regularization matrix or Tikhonov matrix) and the squared residual error $\|\mathbf{y} - \mathbf{X}^H\mathbf{v}\|^2$, such that:

$$\mathbf{v}^* = \arg\min_{\mathbf{v}\in\mathbb{C}^N} \|\Lambda\mathbf{v}\|^2 + \|\mathbf{y} - \mathbf{X}^H\mathbf{v}\|^2 \ . \tag{22}$$

However, still the ridge regressor-based spectral estimation (22) is suitable only for Gaussian noise models since it minimizes the L_2-norm of the residuals. Robustness in non-Gaussian environments may be achieved by replacing the L_2-norm with more suitable cost functions. In particular, by replacing the L_2-norm in (21) by L_1-norm, we may obtain the so-called Laplace periodogram which can be robust against heavy-tailed noise distributions such as Laplace and Cauchy noise [39], [41]. One may obviously generalize this to L_p-norm regression models [39]. However, all these approaches assume arbitrary cost functions which can only be robust against the noise model corresponding to the assumed cost function.

On the other hand, under the assumption of noise uncertainty, we have to account for deviations of the noise model from a nominal distribution. Hence, we may consider an ϵ-contaminated noise model in which the probability distribution of the noise process is denoted by $F = (1-\epsilon)P + \epsilon M$, where $0 \leq \epsilon < 1$ is a known parameter, P is a known Gaussian distribution and M is an unknown contaminating noise distribution that is only assumed to be symmetric and otherwise arbitrary [22][1]. In this case, we may require certain

[1]In this section, we only consider ϵ-contaminated noise models with Gaussian nominal distributions. Generalizing this assumption to non-Gaussian nominal distributions requires further investigation, which will be considered in future. Note that, the nominal Gaussian assumption is a common model assumed in modeling non-Gaussianity using mixture models such as the Gaussian-Laplace mixture and Gaussian mixture models [17], [19], [20].

cost functions that can minimize the impact of contaminating noise on the spectral estimation. In particular, we will consider a regressor with a Huber cost function, which can guarantee robustness against ϵ-contaminated Gaussian noise models [22]. The Huber cost (or loss) function can be defined as:

$$\ell(x) = \begin{cases} x^2/2 & \text{if} \quad |x| \leq \delta_H \\ \delta_H(|x| - \delta_H/2) & \text{if} \quad |x| > \delta_H \end{cases}, \qquad (23)$$

for a given threshold $\delta_H > 0$. This function assigns a quadratic cost to small residuals while imposing a linear cost to larger ones [22], [31]. Furthermore, it is considered to be *the most robust estimator* for ϵ-contaminated noise models, which makes it a good candidate under the assumption of noise uncertainty [22].

Hence, in order to ensure both smoothness and robustness of the periodogram, we will consider a regression model with a multi-objective cost function including both a regularization term and the Huber cost function. The proposed robust spectral estimator can thus be defined as a modified ridge regressor, in which we replace the quadratic cost function by the Huber cost function:

$$\mathbf{v}^* = \arg\min_{\mathbf{v}\in\mathbb{C}^N} \|\Lambda\mathbf{v}\|^2 + c\left(\mathbf{y} - \mathbf{X}^H\mathbf{v}\right), \qquad (24)$$

where $c : \mathbb{C}^N \to \mathbb{R}$ is defined as $c(\mathbf{v}) = \gamma\sum_{i=1}^N \ell\left(\Re\{v_i\}\right) + \ell\left(\Im\{v_i\}\right)$, for all $\mathbf{v} = [v_1, \cdots, v_N]^T \in \mathbb{C}^N$ and for some $\gamma > 0$. Note that, by modifying the regularization matrix Λ in (24), we can balance between robustness and smoothness of the obtained solution. Obviously, the *Huber-optimal* solution would be obtained for $\Lambda = \mathbf{0}$, which is a special case of (24).

A. A Fixed-Point Iteration Algorithm for Robust Spectral Estimation

The model in (24) is a nonlinear optimization problem and does not have a closed-form solution, in general. In addition, it involves an optimization over a complex domain (since $\mathbf{v} \in \mathbb{C}^N$). Hence, we may decompose (24) into its real and imaginary components in order to obtain efficient spectral estimation for a complex discrete-time signals $\mathbf{y} \in \mathbb{C}^N$. The problem can thus be stated as:

$$\min_{\mathbf{v},e_{1,R},\cdots,e_{N,R},e_{1,I},\cdots,e_{N,I}} \frac{1}{2}\|\Lambda\mathbf{v}\|^2 + \frac{\gamma}{2}\sum_{k=1}^N \ell(e_{k,R}) + \ell(e_{k,I}), \qquad (25)$$

subject to:

$$\begin{aligned} e_{k,R} &= y_{k,R} - \left(\mathbf{x}_{k,R}^T\mathbf{v}_R + \mathbf{x}_{k,I}^T\mathbf{v}_I\right), \forall k = 1, \cdots, N, \\ e_{k,I} &= y_{k,I} - \left(-\mathbf{x}_{k,I}^T\mathbf{v}_R + \mathbf{x}_{k,R}^T\mathbf{v}_I\right), \forall k = 1, \cdots, N, \end{aligned}$$

where $y_k \triangleq y_{k,R} + jy_{k,I}$, $\mathbf{x}_k \triangleq \mathbf{x}_{k,R} + j\mathbf{x}_{k,I}$, $\mathbf{v} \triangleq \mathbf{v}_R + j\mathbf{v}_I$, $e_k = y_k - \mathbf{x}_k^H\mathbf{v} \triangleq e_{k,R} + je_{k,I}$. It can be shown that (25) is a convex function of $\mathbf{v}$ since it is a nonnegative weighted sum of convex functions. The problem in (25) can be solved numerically using the Newton's method, which requires a cubic complexity order due to the inversion of a $2N \times 2N$ Hessian matrix at each iteration. However, below we propose a fixed-point iteration algorithm that can solve the Lagrange equations of (25) by avoiding the inverse operation, thus helping to reduce the computational complexity to a quadratic order.

Note that, the norm $\|\Lambda\mathbf{v}\|^2$ can be decomposed as:

$$\begin{aligned} \|\Lambda\mathbf{v}\|^2 &= \mathbf{v}^H\Lambda^H\Lambda\mathbf{v} \\ &= (\mathbf{v}_R + j\mathbf{v}_I)^H Q(\mathbf{v}_R + j\mathbf{v}_I) \\ &= \mathbf{v}_R^T Q\mathbf{v}_R + \mathbf{v}_I^T Q\mathbf{v}_I + j(-\mathbf{v}_I^T Q\mathbf{v}_R + \mathbf{v}_R^T Q\mathbf{v}_I) \\ &= \mathbf{v}_R^T Q\mathbf{v}_R + \mathbf{v}_I^T Q\mathbf{v}_I + j(-\mathbf{v}_R^T Q\mathbf{v}_I + \mathbf{v}_R^T Q\mathbf{v}_I) \\ &= \|\Lambda\mathbf{v}_R\|^2 + \|\Lambda\mathbf{v}_I\|^2, \end{aligned}$$

where we define $Q \triangleq \Lambda^H\Lambda \in \mathbb{R}^{N\times N}$ to be a non-singular matrix. The Lagrangian of (25) can thus be written as:

$$\begin{aligned} \mathcal{L} = \frac{1}{2}\|\Lambda\mathbf{v}_R\|^2 + \frac{1}{2}\|\Lambda\mathbf{v}_I\|^2 + \frac{\gamma}{2}\sum_{k=1}^N \ell(e_{k,R}) + \ell(e_{k,I}) + \\ + \sum_{k=1}^N \alpha_k\left[y_{k,R} - \left(\mathbf{x}_{k,R}^T\mathbf{v}_R + \mathbf{x}_{k,I}^T\mathbf{v}_I\right) - e_{k,R}\right] + \\ + \sum_{k=1}^N \beta_k\left[y_{k,I} - \left(-\mathbf{x}_{k,I}^T\mathbf{v}_R + \mathbf{x}_{k,R}^T\mathbf{v}_I\right) - e_{k,I}\right], \end{aligned}$$

where α_k's and β_k's are the Lagrange multipliers.

By setting $\nabla\mathcal{L} = \mathbf{0}$, we obtain:

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\alpha_k} = 0 \quad &\Rightarrow \quad e_{k,R} = y_{k,R} - \left(\mathbf{x}_{k,R}^T\mathbf{v}_R + \mathbf{x}_{k,I}^T\mathbf{v}_I\right), \forall k = 1, \cdots, N \\ \frac{\partial\mathcal{L}}{\partial\beta_k} = 0 \quad &\Rightarrow \quad e_{k,I} = y_{k,I} - \left(-\mathbf{x}_{k,I}^T\mathbf{v}_R + \mathbf{x}_{k,R}^T\mathbf{v}_I\right), \forall k = 1, \cdots, N \\ \frac{\partial\mathcal{L}}{\partial e_{k,R}} = 0 \quad &\Rightarrow \quad \alpha_k = \frac{\gamma}{2}\ell'(e_{k,R}), \forall k = 1, \cdots, N \\ \frac{\partial\mathcal{L}}{\partial e_{k,I}} = 0 \quad &\Rightarrow \quad \beta_k = \frac{\gamma}{2}\ell'(e_{k,I}), \forall k = 1, \cdots, N \\ \nabla_{\mathbf{v}_R}\mathcal{L} = \mathbf{0} \quad &\Rightarrow \quad \mathbf{v}_R = Q^{-1}\sum_{k=1}^N \alpha_k\mathbf{x}_{k,R} - \beta_k\mathbf{x}_{k,I} \qquad (26) \\ \nabla_{\mathbf{v}_I}\mathcal{L} = \mathbf{0} \quad &\Rightarrow \quad \mathbf{v}_I = Q^{-1}\sum_{k=1}^N \alpha_k\mathbf{x}_{k,I} + \beta_k\mathbf{x}_{k,R} \qquad (27) \end{aligned}$$

The solution of this problem can be obtained by solving the following set of nonlinear equations:

$$\alpha_k = \frac{\gamma}{2}\ell'\left(y_{k,R} - \sum_{j=1}^N p_{j,k}\alpha_j\right), k = 1, \cdots, N, \qquad (28)$$

and

$$\beta_k = \frac{\gamma}{2}\ell'\left(y_{k,I} - \sum_{j=1}^N p_{j,k}\beta_j\right), k = 1, \cdots, N, \qquad (29)$$

where $p_{j,k} = q_{j,k}^R + q_{j,k}^I$, $q_{j,k}^R = \mathbf{x}_{j,R}^T Q^{-1}\mathbf{x}_{k,R}$ and $q_{j,k}^I = \mathbf{x}_{j,I}^T Q^{-1}\mathbf{x}_{k,I}$. Hence, the weight vector $\mathbf{v}$ can be obtained by replacing α_k's and β_k's in (26) and (27).

By letting $\mathbf{P} = [p_{j,k}]$ [2], (28) and (29) can be written as:

$$\boldsymbol{\alpha} = \frac{\gamma}{2}\boldsymbol{\ell}'\left(\Re\{\mathbf{y}\} - \mathbf{P}\boldsymbol{\alpha}\right) \triangleq h_R\left(\boldsymbol{\alpha}\right), \qquad (30)$$

[2]Note that, we can write $\mathbf{P} = \mathbf{X}_R^T Q^{-1}\mathbf{X}_R + \mathbf{X}_I^T Q^{-1}\mathbf{X}_I$, where $\mathbf{X}_R = \Re\{\mathbf{X}\}$ and $\mathbf{X}_I = \Im\{\mathbf{X}\}$.

and

$$\boldsymbol{\beta} = \frac{\gamma}{2}\boldsymbol{\ell}'\left(\Im\{\mathbf{y}\} - \mathbf{P}\boldsymbol{\beta}\right) \triangleq h_I\left(\boldsymbol{\beta}\right), \tag{31}$$

where $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_N]^T$, $\boldsymbol{\beta} = [\beta_1, \cdots, \beta_N]^T$, $\boldsymbol{\ell}'(\mathbf{u}) \triangleq [\ell'(u_1), \cdots, \ell'(u_N)]^T$ for all vectors $\mathbf{u} = [u_1, \cdots, u_N]^T$, $h_R : \mathbb{R}^N \to \mathbb{R}^N$ and $h_I : \mathbb{R}^N \to \mathbb{R}^N$. Note that, the computational complexity of both (30) and (31) are dominated by the terms $\mathbf{P}\boldsymbol{\alpha}$ and $\mathbf{P}\boldsymbol{\beta}$. However, they only require quadratic complexity order in contrast with the cubic complexity order that would be required for the Newton's method.

It can be shown that (30) and (31) admit a unique fixed-point solution if h_R and h_I are, respectively, a contraction mapping [32]. Furthermore, based on the Banach Fixed-Point theorem, the fixed-point solution of (30) and (31) can be found using a fixed-point iteration algorithm [32]. Hence, in the followings we establish the conditions under which (30) and (31) represent a contraction mapping.

Theorem 1. *If $f : \mathbb{R}^N \to \mathbb{R}^N$ and $g : \mathbb{R}^N \to \mathbb{R}^N$ are Lipschitz continuous with respective Lipschitz constants p and q, and if $Kpq < 1$ for some positive scalar $K > 0$, then $h = K(f \circ g)$ is a contraction mapping on $\mathbb{R}^N$.*

Proof. If g is Lipschitz continuous on $\mathbb{R}^N$ with a Lipschitz constant q, then $\|g(x) - g(y)\| \leq q\|x - y\|$, $\forall x, y \in \mathbb{R}^N$. Similarly, $\|f(x) - f(y)\| \leq p\|x - y\|$, $\forall x, y \in \mathbb{R}^N$. But since $g(x), g(y) \in \mathbb{R}^N$, then $\|f \circ g(x) - f \circ g(y)\| \leq p\|g(x) - g(y)\| \leq pq\|x - y\|$, $\forall x, y \in \mathbb{R}^N$.
Therefore, $\|Kf \circ g(x) - Kf \circ g(y)\| \leq Kpq\|x - y\|$, $\forall x, y \in \mathbb{R}^N$. Furthermore, if $Kpq < 1$, then $h = K(f \circ g)$ is a contraction mapping on $\mathbb{R}^N$. □

In (30), let $g(\boldsymbol{\alpha}) \triangleq \mathbf{y} - \mathbf{P}\boldsymbol{\alpha}$ with $g : \mathbb{R}^N \to \mathbb{R}^N$. The norm

$$\|g(\boldsymbol{\alpha}_1) - g(\boldsymbol{\alpha}_2)\|_2 = \|\mathbf{P}\boldsymbol{\alpha}_2 - \mathbf{P}\boldsymbol{\alpha}_1\|_2 \leq \|\mathbf{P}\|_2\|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2\|_2 \ ,$$

$\forall \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathbb{R}^N$. Then, g is Lipschitz continuous with a Lipschitz constant equal to $q = \|\mathbf{P}\|_2$. Furthermore, if $\boldsymbol{\ell}'$ is Lipschitz continuous with a Lipschitz constant p and if $\frac{\gamma}{2}p\|\mathbf{P}\|_2 < 1$, then, using Theorem 1, h_R in (30) is a contraction mapping on $\mathbb{R}^N$. In this case, based on the Banach Fixed-Point Theorem, the solution of (30) can be obtained using fixed-point iterations [32].

Note that, since the regularization parameter γ can be chosen arbitrarily, the condition $\frac{\gamma}{2}p\|\mathbf{P}\|_2 < 1$ can always be satisfied, provided that $p < \infty$ and $\|\mathbf{P}\|_2 < \infty$. This implies that we can always guarantee convergence of the proposed fixed-point iteration algorithm using both (30) and (31). In particular, if ℓ is the Huber cost function, it can be shown that the Lipschitz constant of ℓ' is $p = 1$. Furthermore, if we assume that $\Lambda = \mathbf{I}$, then $\mathbf{P} = \mathbf{X}^H\mathbf{X} = \frac{1}{N}\mathbf{I}$, implying that $\|\mathbf{P}\|_2 = \frac{1}{N}$. In this case, the fixed-point iteration algorithm of (30) and (31) converges if $\gamma < 2N$.

V. Simulation Results

In this section, we consider a sinusoidal signal $s(t) = \sqrt{2P_s}\cos(2\pi f_c t)$, with transmit power P_s and carrier frequency $f_c = 20MHz$[3]. In our simulation, we set $\theta = 1$ under

[3]For illustration purposes, we consider a sinusoidal signal with a frequency of $f_c = 20MHz$. However, the simulation results can be scaled arbitrarily to any frequency.

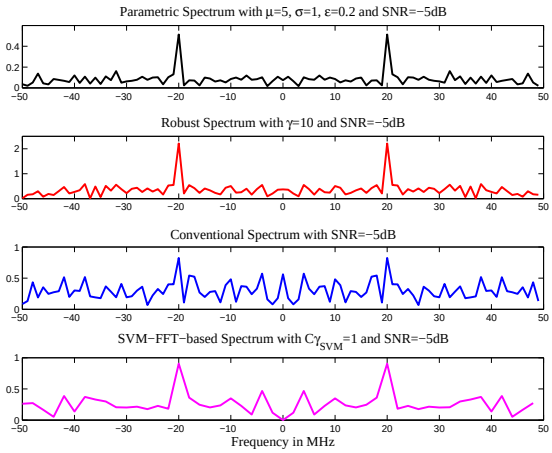


Fig. 1. Magnitude spectra of a sinusoidal signal transmitting at $20MHz$ and subject to Gaussian-Laplace noise with $\epsilon = 0.2$ and $N = 100$.

$\mathcal{H}_1$, and assume that the received signal $y(t)$ is subject to a Gaussian-Laplace mixture noise with a Laplace contamination rate of $\epsilon = 0.2$. The parameters of the Laplace and Gaussian noise are $\lambda = 0.2$ and $\sigma = 1$, respectively. We assume an SNR of $-5dB$, where $SNR = \frac{P_s}{(1-\epsilon)\sigma^2 + 2\epsilon/\lambda^2}$, and show, in Fig. 1, the magnitude spectra computed using both parametric and robust spectral estimation methods (i.e. $\frac{1}{N}|\tilde{Y}(F)|$ and $|v_i^*|$ in (24), respectively), as well as the conventional spectrum $\frac{1}{N}|Y(F)|$ and the SVM-FFT-based spectrum of [42]. The robust method, however, is implemented using $\gamma = 10$ and $\delta_H = 1$. On the other hand, the SVM-FFT method is computed using $C = 10$ and $\gamma_{SVM} = 1$, as defined in [42]. Figure 1 shows the detected peaks at $\pm 20MHz$. In order to compare the performance of the four detection methods, we will compute the corresponding ROC's, as follows.

After obtaining the spectral estimations of the sensed wideband signals, signal detection can be applied to determine the center frequencies of active signals within the wideband of interest. In many CR applications, signal detection is based on the NP criterion [11]. This leads to a threshold η on the spectrum (or the periodogram) to determine the active channels [11], [12]. The performance of such detectors is characterized by the ROC which shows the detection probability for a certain false alarm rate [15]. An analytical closed-form expression for the ROC can be derived if the probability distribution of the decision statistic is known. In our case, however, it is hard to obtain a closed-form expression for the decision statistic $T_2^{(LO)}(\mathbf{y})$ in (15) in the case of Gaussian-Laplace noise model assumed in the example of Fig. 1. Similarly, it is difficult to derive the probability distribution of the robust spectral estimator since it is obtained using an iterative process. Hence, in the following we resort to numerical simulations to compute the desired probability distributions and the corresponding ROC's.

In the case of parametric detection, we can compute the threshold η for a given false alarm probability α, as described in Algorithm 1. We also assume that $\phi_s(F)$ is an impulse

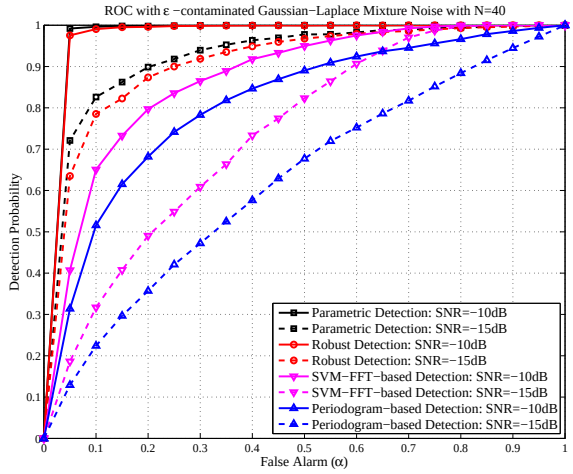


Fig. 2. Comparison between the receiver operating characteristics (ROC's) of our proposed detection methods, as well as the SVM-FFT and the periodogram-based methods with $\epsilon = 0.2$, $\mu = 5$ and $\sigma = 1$.

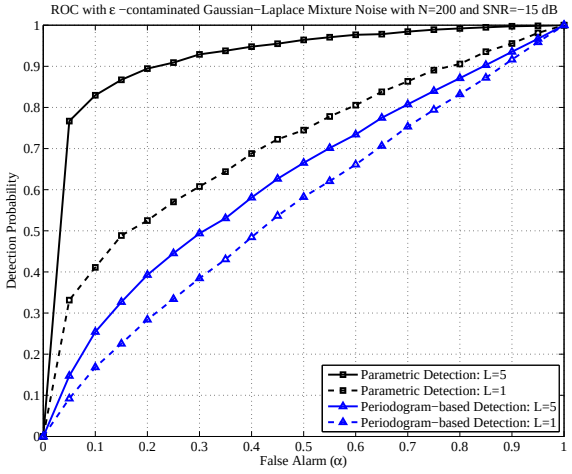


Fig. 3. Receiver operating characteristics of the proposed parametric detection method and the periodogram-based method, assuming a rectangular $\phi_s(F)$ with length $L\Delta f$ with $\epsilon = 0.2$, $\mu = 5$ and $\sigma = 1$.

Algorithm 1 Numerical method for computing the decision threshold η.

Generate M i.i.d. random vectors $\{\mathbf{y}_k : k = 1, \cdots, M\}$ such that $\mathbf{y}_k \sim f_{\mathbf{W}}$ (under $\mathcal{H}_0$).
Compute the corresponding decision statistics $\mathcal{T} = \left\{T_2^{(LO)}(\mathbf{y}_k) : k = 1, \cdots, M\right\}$.
Compute the histogram of $\mathcal{T}$ to approximate the probability distribution of $T_2^{(LO)}(\mathbf{Y})$.
Using numerical integration, compute the complementary cumulative density function (ccdf) of $T_2^{(LO)}(\mathbf{Y})$ from the histogram of $\mathcal{T}$.
Using the ccdf, determine η such that $Pr\left\{T_2^{(LO)}(\mathbf{Y}) > \eta | \mathcal{H}_0\right\} = \alpha$.

function centered at a frequency f of interest such that $\phi_s(F) = \delta^{(D)}(F - f)$, leading to $T_2^{(LO)}(\mathbf{y}) = T_{2,0}(\mathbf{y}) + \hat{\phi}(f)$.

Similarly, we can obtain the decision thresholds of the robust, the SVM-FFT and the periodogram-based spectral estimations by computing the histograms of their corresponding spectral estimations under $\mathcal{H}_0$ hypothesis. After computing the decision thresholds for each case, we generate random signals $\mathbf{y}$ and compute the corresponding spectral estimations. For each random vector $\mathbf{y}$ (and for each detection method), a signal is detected if the corresponding spectral estimation exceeds the threshold at the frequency of interest f. The average detection probability is thus computed as the proportion of random signals whose spectral estimations exceed the decision threshold at frequency f. These results are represented using the ROC's for the above four methods, as shown in Fig. 2. These ROC's show that both parametric and robust detection methods achieve significantly better detection performance, compared to those of the SVM-FFT and periodogram-based approaches.

Next, we analyze the performance of the proposed parametric method using an ideal rectangular PSD $\phi_s(F)$. The bandwidth of $\phi_s(F)$ is assumed to be $L\Delta f$, where Δf is the discrete frequency increment. We assume a binary phase-shift keying (BPSK) signal with carrier frequency $f_c = 5MHz$ and bandwidth $500kHz$. The signal is sampled at $f_s = 20MHz$ and we assume $N = 200$ samples, resulting in a discrete frequency increment $\Delta f = 100kHz$ (i.e. spectral resolution). The ROC of the parametric detection method assuming a rectangular PSD $\phi_s(F)$ is shown in Fig. 3 along with the smoothed periodogram method that was proposed in [11], [12]. The periodogram is smoothed using a rectangular window having a bandwidth $L\Delta f$. Figure 3 shows that the detection probability increases by increasing the window length L, similar to [12]. However, the parametric method still achieves better performance, compared to the smoothed periodogram-based method of [12]. Note that, increasing the window length L reduces the spectral resolution of the spectral estimations and thus should only be subjected to the required minimum spectral discrimination.

Finally, we verify the detection performance of the sliding window detection method as obtained in (19). Thus, we consider a real bandlimited signal s_k with bandwidth $B\Delta f = B/N$ and centered at a normalized frequency F_c such that:

$$s_k = \sqrt{\frac{2P_s}{B+1}} \sum_{m=-B/2}^{B/2} \cos\left(2\pi k\left(F_c + \frac{m}{N}\right)\right), \qquad (32)$$

for $k = 0, \cdots, N-1$. The signal is subject to Gaussian noise with average power σ^2. By applying the sliding-window periodogram-based detection, we compute the corresponding detection probability using both simulated and analytical methods, as shown in Fig. 4. The results show perfect match between both simulated and analytical ROC curves. Furthermore, if $L \leq B + 1 = 3$, a higher detection probability can be achieved with a larger window size. However, if $L > B + 1 = 3$, then the detection probability decreases for larger L.

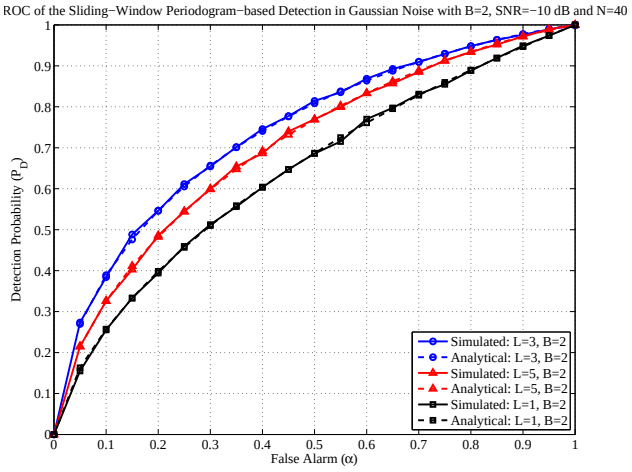


Fig. 4. Receiver operating characteristics (ROC's) of the periodogram-based detection method with a sliding window of length L and subject to Gaussian noise, with $SNR = -10dB$ and $N = 40$.

VI. Conclusion

In this paper, we have proposed both parametric and robust signal detection methods for wideband spectrum sensing in CR's. The parametric detector assumed a known non-Gaussian noise distribution, which can account for outliers and interferers that may occur in the detected signals. The decision statistic of the parametric detector led to a spectral estimation function that can be computed as the periodogram of the transformed signal sequence. On the other hand, if the noise distribution is *approximately* known, we followed a robust approach to obtain a robust signal detector by assuming an ϵ-contaminated noise model. The robust detector was based on a robust spectral estimation method that was formulated as a modified ridge regressor in which the quadratic cost function was replaced by the Huber function. A fixed-point iteration algorithm was proposed to solve the Lagrange equations of the robust spectral estimator, thus reducing the computational complexity to a quadratic order. We computed the ROC curves of both parametric and robust detectors and showed that they can both achieve better detection performance in the presence of Gaussian-Laplace mixture noise models, compared to similar existing detectors.

References

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

[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201 – 220, Feb. 2005.

[3] J. Mitola, "Cognitive radio architecture evolution," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 626 –641, Apr. 2009.

[4] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, Quarter 2009.

[5] E. Biglieri, "Effect of uncertainties in modeling interferences in coherent and energy detectors for spectrum sensing," in *IEEE International Symposium on Information Theory Proceedings (ISIT '11)*, Aug. 2011, pp. 2418 –2421.

[6] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *1st International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Mykonos Island, Greece, June 2006, pp. 1–5.

[7] J. Unnikrishnan and V. V. Veeravalli, "Algorithms for dynamic spectrum access with learning for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 750 –760, Feb. 2010.

[8] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[9] Y. Li, S. Jayaweera, M. Bkassiny, and K. Avery, "Optimal myopic sensing and dynamic spectrum access in cognitive radio networks with low-complexity implementations," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2412 –2423, July 2012.

[10] S. K. Jayaweera and C. G. Christodoulou, "Radiobots: Architecture, algorithms and realtime reconfigurable antenna designs for autonomous, self-learning future cognitive radios," University of New Mexico, Technical Report EECE-TR-11-0001, Mar. 2011. [Online]. Available: http://repository.unm.edu/handle/1928/12306

[11] M. Bkassiny, S. K. Jayaweera, Y. Li, and K. A. Avery, "Wideband spectrum sensing and non-parametric signal classification for autonomous self-learning cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2596 –2605, July 2012.

[12] ——, "Blind cyclostationary feature detection based spectrum sensing for autonomous self-learning cognitive radios," in *IEEE International Conference on Communications (ICC '12)*, Ottawa, Canada, June 2012.

[13] R. Martin and D. Thomson, "Robust-resistant spectrum estimation," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1097 – 1115, Sep. 1982.

[14] A. Sonnenschein and P. Fishman, "Radiometric detection of spread-spectrum signals in noise of uncertain power," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 654 –660, July 1992.

[15] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer, 1998.

[16] Y. M. Kim, G. Zheng, S. H. Sohn, and J. M. Kim, "An alternative energy detection using sliding window for cognitive radio system," in *10th International Conference on Advanced Communication Technology (ICACT '08)*, vol. 1, Gangwon-Do, South Korea, Feb. 2008, pp. 481 – 485.

[17] D. W. J. Stein, "Detection of random signals in gaussian mixture noise," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1788–1801, 1995.

[18] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. Springer-Verlag, 1988.

[19] J. Miller and J. Thomas, "The detection of signals in impulsive noise modeled as a mixture process," *IEEE Transactions on Communications*, vol. 24, no. 5, pp. 559 – 563, May 1976.

[20] ——, "Robust detectors for signals in non-gaussian noise," *IEEE Transactions on Communications*, vol. 25, no. 7, pp. 686 – 690, July 1977.

[21] D. Middleton, "Non-gaussian noise models in signal processing for telecommunications: new methods an results for class a and class b noise models," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[22] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964. [Online]. Available: http://www.jstor.org/stable/2238020

[23] A. Blomqvist and B. Wahlberg, "On the relation between weighted frequency-domain maximum-likelihood power spectral estimation and the prefiltered covariance extension approach," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 384 –389, Jan. 2007.

[24] D. Bell and J. Gowdy, "Power spectral estimation via nonlinear frequency warping," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 5, pp. 436 – 441, Oct. 1978.

[25] S. Cheung and J. Lim, "Combined multiresolution (wide-band/narrow-band) spectrogram," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 975–977, 1992.

[26] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21 –30, Mar. 2008.

[27] Z. Tian and G. Giannakis, "Compressed sensing for wideband cognitive radios," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 4, Honolulu, HI, Apr. 2007, pp. IV–1357 –IV–1360.

[28] H. Wu and S. Wang, "An efficient and robust approach for wideband compressive spectrum sensing," in *IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC '12)*, Hong Kong, China, Aug. 2012.

[29] S. Oh and R. Kashyap, "A robust approach for high resolution frequency estimation," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 627–643, 1991.

[30] S. Kassam and H. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433 – 481, Mar. 1985.

[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[32] S. Banach, "Sur les operations dans les ensembles abstraits et leur application aux equations integrales," *Fundamenta Mathematicae*, vol. 3, pp. 133–181, 1922.

[33] Y. Li, S. K. Jayaweera, C. Ghosh, and M. Bkassiny, "Learning-aided sensing scheduling for wide-band cognitive radios," in *Worskshop on Wideband Cognitive Radio Communications and Networks (WCRCN) at IEEE Vehicular Technology Conference (VTC-Fall '13)*, Las Vegas, NV, Sep. 2013.

[34] Y. Tawk, J. Costantine, K. Avery, and C. Christodoulou, "Implementation of a cognitive radio front-end using rotatable controlled reconfigurable antennas," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 5, pp. 1773 –1778, May 2011.

[35] Y. Tawk, S. K. Jayaweera, C. G. Christodoulou, and J. Costantine, "A comparison between different cognitive radio antenna systems," in *International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS '11)*, Chiangmai, Thailand, Dec. 2011.

[36] Y. Li, S. K. Jayaweera, M. Bkassiny, and C. Ghosh, "Learning-aided sub-band selection algorithms for spectrum sensing in wide-band cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 2012–2024, Apr. 2014.

[37] R. Martin and S. Schwartz, "Robust detection of a known signal in nearly gaussian noise," *IEEE Transactions on Information Theory*, vol. 17, no. 1, pp. 50–56, Jan. 1971.

[38] Y. Zeng and Y.-C. Liang, "Covariance based signal detections for cognitive radio," in *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '07)*, Dublin, Ireland, Apr. 2007, pp. 202–207.

[39] T.-H. Li, "A nonlinear method for robust spectral analysis," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2466–2474, 2010.

[40] M. Ramon, T. Atwood, S. Barbin, and C. Christodoulou, "Signal classification with an svm-fft approach for feature extraction in cognitive radio," in *SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC '09)*, Belem, Brazil, Nov. 2009, pp. 286 –289.

[41] T.-H. Li, "Laplace periodogram for time series analysis," *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 757–768, 2008.

[42] J. Rojo-Alvarez, M. Martinez-Ramon, A. Figueiras-Vidal, A. Garcia-Armada, and A. Artes-Rodriguez, "A robust support vector algorithm for nonparametric spectral analysis," *IEEE Signal Processing Letters*, vol. 10, no. 11, pp. 320 –323, Nov. 2003.

[43] Z. Liang, R. Jaszczak, and R. Coleman, "Parameter estimation of finite mixtures using the EM algorithm and information criteria with application to medical image processing," *IEEE Transactions on Nuclear Science*, vol. 39, no. 4, pp. 1126–1133, Aug 1992.

APPENDIX

A. Non-linearity functions $g^{(LO)}(x)$ and $h^{(LO)}(x)$ for Gaussian-Laplace mixture and Gaussian mixture noise models

1) Example 1: If we assume an i.i.d. noise with a Gaussian-Laplace mixture distribution, such that:

$$f_W(x) = \frac{1-\epsilon}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} + \frac{\lambda\epsilon}{2} e^{-\lambda|x|}, \tag{33}$$

then, the common non-linearity $g^{(LO)}(x)$ can be expressed as [19], [20]:

$$g^{(LO)}(x) = \frac{\frac{\lambda^2\epsilon}{2} sgn(x) e^{-\lambda|x|} + \frac{(1-\epsilon)x}{\sigma^3\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}}}{\frac{\lambda\epsilon}{2} e^{-\lambda|x|} + \frac{1-\epsilon}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}}, \tag{34}$$

where $sgn(x)$ is the signum function, ϵ is the Laplace contamination rate, λ is the parameter of the Laplace noise and σ^2 is the variance of the zero-mean Gaussian mixture component[4]. We can also express $h^{(LO)}(x)$ as:

$$h^{(LO)}(x) = \frac{\frac{\lambda^2\epsilon}{2}\left(\lambda - 2\delta^{(D)}(x)\right) e^{-\lambda|x|} + \frac{1-\epsilon}{\sigma^3\sqrt{2\pi}}\left(\frac{x^2}{\sigma^2} - 1\right) e^{-\frac{x^2}{2\sigma^2}}}{\frac{\lambda\epsilon}{2} e^{-\lambda|x|} + \frac{1-\epsilon}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}}, \tag{35}$$

where $\delta^{(D)}(x)$ is the Dirac delta function. As noted in [19], the Gaussian-Laplace mixture is a convenient representation of impulsive noise models due to its slower decaying tail, compared to the Gaussian distribution. More importantly, it was shown that detectors having a non-linearity with the hard-limiting behavior characteristic of the Laplace distribution are robust against errors in noise distribution measurement [19], [37].

Note that, if we assume an i.i.d. Laplacian noise model (i.e. $\epsilon = 1$) such that $f_W(w) = \frac{\lambda}{2} e^{-\lambda|w|}$, we obtain $g^{(LO)}(x) = \lambda sgn(x)$ and $h^{(LO)}(x) = \lambda^2 - 2\lambda\delta^{(D)}(x)$.

2) Example 2: Gaussian mixture distribution is another model that has been previously used as a non-Gaussian noise model [17]. A Gaussian mixture distribution with m mixture components can be characterized by the pdf $f_W(x) = \sum_{i=1}^{m} \kappa_i \mathcal{N}(\mu_i, \sigma_i^2)(x)$, where $\mathcal{N}(\mu, V)$ denotes the pdf of a Gaussian distribution with mean μ and variance V, $\{\kappa_i\}_{i=1}^m$ are the mixing proportions, $\mu_i = a_i\mu$ and $\sigma_i^2 = a_i^2\sigma^2$, for some parameters a_i's, μ and σ [17]. Similar to the Gaussian-Laplace mixture above, the non-linear transformations $g^{(LO)}(x)$ and $h^{(LO)}(x)$ can be obtained in closed-form for the Gaussian mixture model as:

$$g^{(LO)}(x) = \frac{\sum_{i=1}^{m} \frac{\kappa_i(x-\mu_i)}{\sigma_i^3} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}}{\sum_{i=1}^{m} \frac{\kappa_i}{\sigma_i} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}}, \tag{36}$$

and

$$h^{(LO)}(x) = \frac{\sum_{i=1}^{m} \frac{\kappa_i}{\sigma_i^3}\left[\frac{(x-\mu_i)^2}{\sigma_i^2} - 1\right] e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}}{\sum_{i=1}^{m} \frac{\kappa_i}{\sigma_i} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}}. \tag{37}$$

B. Derivation of the ROC for the Sliding-Window Detection Method

In order to analyze the performance of the sliding-window-based detection and its impact on the detection accuracy, we consider the following observation model:

$$y_k = \begin{cases} w_k \quad (k = 0, \cdots, N-1) & \text{under } \mathcal{H}_0 \\ s_k + w_k \quad (k = 0, \cdots, N-1) & \text{under } \mathcal{H}_1 \end{cases}, \tag{38}$$

where y_k is the detected signal, s_k is the transmit signal and w_k is the noise signal. The sliding-window technique applies a smoothing window to the spectral estimation function. An analytical performance measure of the sliding-window detection can be obtained under certain regularity conditions. That is, we will assume that $w_k \sim \mathcal{N}(0, \sigma^2)$ and s_k is a bandlimited real signal centered at a normalized frequency $F_c = n_c\Delta f$ with a

[4]The parameters of the assumed finite-mixture distribution can be estimated using Expectation-Maximization (EM) algorithms based on the maximum-likelihood (ML) criterion, as proposed in [43].

normalized bandwidth $B\Delta f$, where $\Delta f = \frac{1}{N}$ is the spectral resolution (or discrete frequency increment)[5]. By applying the DFT to y_k under $\mathcal{H}_1$, we obtain:

$$Y_n = \begin{cases} S_n + W_n & \text{if } n \in \mathcal{B} \\ W_n & \text{otherwise} \end{cases}, \quad (39)$$

where $\mathcal{B} = \{n_c - \frac{B}{2}, \cdots, n_c + \frac{B}{2}\} \bigcup \{N - n_c - \frac{B}{2}, \cdots, N - n_c + \frac{B}{2}\}$, and Y_n, S_n and W_n ($n = 0, \cdots, N-1$) are the DFT's of y_k, s_k and w_k, respectively. If we assume that $\mathbb{E}\{s_k^2\} = P_s$, then we can express Y_n as:

$$Y_n = \begin{cases} N\sqrt{\frac{P_s}{2(B+1)}} + W_{R,n} + jW_{I,n} & \text{if } n \in \mathcal{B} \\ W_{R,n} + jW_{I,n} & \text{otherwise} \end{cases}, \quad (40)$$

where $W_{R,n} = \Re\{W_n\} \sim \mathcal{N}(0, \frac{N\sigma^2}{2})$ and $W_{I,n} = \Im\{W_n\} \sim \mathcal{N}(0, \frac{N\sigma^2}{2})$ [12].

We let $T_n = \sqrt{\frac{2}{N\sigma^2}} Y_n$, $\tilde{W}_{R,n} = \sqrt{\frac{2}{N\sigma^2}} W_{R,n}$ and $\tilde{W}_{I,n} = \sqrt{\frac{2}{N\sigma^2}} W_{I,n}$ such that $\tilde{W}_{R,n} \sim \mathcal{N}(0,1)$ and $\tilde{W}_{I,n} \sim N(0,1)$. Thus, we obtain:

$$T_n = \begin{cases} \sqrt{\frac{NP_s}{(B+1)\sigma^2}} + \tilde{W}_{R,n} + j\tilde{W}_{I,n} & \text{if } n \in \mathcal{B} \\ \tilde{W}_{R,n} + j\tilde{W}_{I,n} & \text{otherwise} \end{cases}. \quad (41)$$

From (41) we can obtain the distribution of $|T_n|^2$ (under $\mathcal{H}_1$) as follows:

$$|T_n|^2 \sim \begin{cases} \chi^2\left(2, \frac{NP_s}{(B+1)\sigma^2}\right) & \text{if } n \in \mathcal{B} \\ \chi^2(2) & \text{otherwise} \end{cases}, \quad (42)$$

where $\chi^2(k, \lambda)$ is the non-central Chi-squared distribution with $k > 0$ degrees of freedom and $\lambda > 0$ non-centrality parameter, and $\chi^2(k)$ is the Chi-squared distribution with $k > 0$ degrees of freedom. A sliding window of size L (odd integer) will be applied to $|T_n|^2$, resulting in the decision statistic $Z_n = \sum_{l=-(L-1)/2}^{(L-1)/2} |T_{n+l}|^2$. We are particularly interested in the value of Z_n for $n = n_c$ in order to obtain a decision statistic for the periodogram-based method. The distribution of Z_{n_c} depends on the window length L, relative to the signal bandwidth.

In particular, if the window size is smaller than the signal bandwidth (i.e. $L \leq B + 1$), then the distribution of Z_{n_c} will be given by:

$$Z_{n_c} \sim \begin{cases} \chi^2\left(2L, \frac{NLP_s}{(B+1)\sigma^2}\right) & \text{under } \mathcal{H}_1 \\ \chi^2(2L) & \text{under } \mathcal{H}_0 \end{cases}. \quad (43)$$

This results in a Neyman-Pearson (NP) detection threshold $\eta = 2\gamma^{-1}(L; (1-\alpha)\Gamma(L))$, where $\alpha \in [0,1]$ is the false-alarm rate, γ^-1 is the inverse lower incomplete Gamma function, and Γ is the Gamma function [12]. The corresponding detection probability can be obtained as:

$$P_D = P\left\{\chi^2\left(2L, \frac{NLP_s}{(B+1)\sigma^2}\right) > \eta\right\} \quad (44)$$

$$= Q_L\left(\sqrt{\frac{NLP_s}{(B+1)\sigma^2}}, \sqrt{\eta}\right) \quad (45)$$

where $Q_L(a,b)$ is the Marcum Q-function.

[5] B is an even integer and n_c is an integer.

On the other hand, if the window length is larger than the signal bandwidth (i.e. $L > B + 1$), then the distribution of Z_{n_c} will be given by:

$$Z_{n_c} \sim \begin{cases} \chi^2\left(2L, \frac{NP_s}{\sigma^2}\right) & \text{under } \mathcal{H}_1 \\ \chi^2(2L) & \text{under } \mathcal{H}_0 \end{cases}, \quad (46)$$

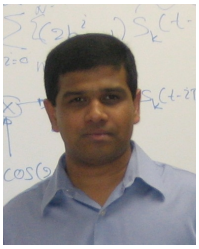
resulting in a detection probability $P_D = Q_L\left(\sqrt{\frac{NP_s}{\sigma^2}}, \sqrt{\eta}\right)$.

Thus, in general, the detection probability of the sliding-window periodogram-based detector can be obtained as:

$$P_D = Q_L\left(\sqrt{\frac{N\min\{L, B+1\}P_s}{(B+1)\sigma^2}}, \sqrt{\eta}\right). \quad (47)$$

Mario Bkassiny (S'06, M'14) received the B.E. degree in Electrical Engineering with High Distinction and the M.S. degree in Computer Engineering from the Lebanese American University, Lebanon, in 2008 and 2009, respectively. He received his PhD degree in Electrical Engineering from the University of New Mexico, Albuquerque, NM, USA in 2013. He is currently an Assistant Professor at the Department of Electrical and Computer Engineering at State University of New York (SUNY) at Oswego, Oswego, NY, USA. He worked as a research assistant at the Communication and Information Sciences Laboratory (CISL), Department of Electrical and Computer Engineering at the University of New Mexico from 2009-2013.

Dr. Bkassiny served as a technical co-chair of the Workshop on Wideband Cognitive Radio Communication and Networks (WCRCN) at the IEEE Vehicular Technology Conference (VTC-Fall 2013). His current research interests are in cognitive radios, wideband spectrum sensing, signal classification and robust signal detection.



Sudharman K. Jayaweera (S'00, M'04, SM'09) was born in Matara, Sri Lanka. He completed his high school education at the Rahula College, Matara, and worked as a science journalist at the Associated Newspapers Ceylon Limited (ANCL) till 1993. Later, he received the B.E. degree in Electrical and Electronic Engineering with First Class Honors from the University of Melbourne, Australia, in 1997 and M.A. and PhD degrees in Electrical Engineering from Princeton University, USA in 2001 and 2003, respectively. He is currently an Associate Professor in Electrical Engineering at the Department of Electrical and Computer Engineering at University of New Mexico, Albuquerque, NM where he is the Associate Chair of the Department and the Director of the Graduate Program. Dr. Jayaweera held an Air Force Summer Faculty Fellowship at the Air Force Research Laboratory, Space Vehicles Directorate (AFRL/RVSV) from 2009-2011.

Dr. Jayaweera is currently an associate editor of IEEE Transactions on Vehicular Technology. He has also served on organization and Technical Program Committees of numerous IEEE conferences. Most recently, he served as the Tutorial and Workshop Chair of the 2013 Fall IEEE Vehicular Technology Conference, General Chair of the First Workshop on Wideband Mobile Cognitive Radios (WMCR) at the IEEE VTC Fall 2013 and the Publicity Chair of the 10th IEEE Broadband Wireless Access Workshop (BWA) at 2014 IEEE Globecom conference. His current research interests include cooperative and cognitive communications, machine learning, information theory of networked-control systems, statistical signal processing and smart-grid control and optimization.