Research Article

Blind Detection of Wideband Interference for Cognitive Radio Applications

Alan J. Coulson

Communications Team, Industrial Research Ltd., P.O. Box 31-310, Lower Hutt 5040, New Zealand

Correspondence should be addressed to Alan J. Coulson, a.coulson@irl.cri.nz

Cognitive radio technologies are being developed which allow heterogeneous systems to share spectrum access while minimizing interference to improve the overall efficiency of spectrum usage. Thus, one important function of a cognitive radio is dynamically to avoid transmitting in occupied spectrum by detecting signals received from unknown competing systems. Robust operation requires the detection of multiple wideband interferers of unknown statistics from a single received sample vector. This paper describes the hypothesis tests that must be evaluated to perform detection of such signals and discusses several methods for performing detection. Computer simulation results are presented to show that hidden Markov modelling and power spectral analysis with edge enhancement are more robust than a simple "interference temperature-" based energy detector.

1. Introduction

The recent unprecedented commercial success of unlicensed spectrum systems has motivated the development of new concepts which will improve the ability of wireless users operating in heterogeneous networks to share radio spectrum. Simple existing concepts such as dynamic frequency selection and transmit power control contribute to an important new paradigm in "smart" wireless systems, cognitive radio [1]. Cognitive radio concepts cover a range of intelligent behaviour in wireless systems [2], but it is the potential to flexibly, reliably, and efficiently reuse radio spectrum which has captured the attention of regulators [3] and standards bodies [4]. This specific ability also is termed "agile radio."

An ability to reliably detect interference from other systems competing for spectrum use, colloquially known as "spectrum sensing" [5, 6], is a prerequisite to robust cognitive radio operation. This is necessary both in the cognitive radio receiver to protect the integrity of the wanted signal, and also in the cognitive radio transmitter which must avoid producing interference in other systems. Indeed, future regulatory compliance for cognitive radios may require the "ability of a sensing receiver to detect the presence of other signals" [7].

Several spectrum sensing methods have been proposed which require some knowledge of the potential interferer, including matched filter detection for specific systems [8] and cyclostationarity detectors for known modulations [9, 10] based on spectral correlation theory developed by Gardner [11]. These methods will be helpful for detecting known primary systems.

To detect unknown systems, blind methods are required. As reviewed by Kailath and Poor [12] and references therein, blind detection methods have been studied for more than 50 years for various applications. Much recent attention has focused on energy detectors [13–15] and radiometers [16–18], and methods for extending energy detection using multiple time and/or space sample vectors [6, 19, 20]. It is noted that the terms "energy detector" and "radiometer" frequently are used interchangeably [21]. These methods present challenges in estimating noise variance and in determining a robust detection threshold [22, 23], so several methods have been proposed to augment energy detection performance using techniques such as spectral smoothing and incorporation of some *a priori* information [24]. Other approaches include joint interference detection and decoding for OFDM [22] and detection based on information theoretic criteria [5].

This paper considers dynamic spectrum sensing for multiple interferers in the case where all signal, channel, and noise parameters are unknown, stochastic and not cyclostationary. This case is described as "difficult" by Kay [25, page 349] and "impossible" by Larsson and Regnoli [26] where the channel also is static. The interference model is established in Section 2. Detection theory is developed following a conventional hypothesis testing framework, in Section 3, resulting in the specification of the generalized likelihood ratio test for the modelled interference. In Section 4, it is demonstrated how implicit interference characteristics can be exploited in a formal framework through the application of hidden Markov modeling (HMM). Alternative detection methods are reviewed in Section 5, covering radiometry, power spectrum analysis, and information theoretic criteria. Statistical analyses used by the detection methods are developed in Sections 6 and 7. The performance of the detection methods is compared by computer simulation in Section 8.

2. Interference Model

A conventional receiver structure is assumed as illustrated in Figure 1(a), assuming a stationary receiver employing a single antenna which is omnidirectional in azimuth. In the absence of a desired signal, the received complex baseband signal, bandlimited to bandwidth B, is

$$x(t) = \sum_{i=1}^{I} h_i(t, \tau) \underset{\tau}{\bigstar} s_i(t - \tau) + z(t), \tag{1}$$

where $s_i(t)$ is the ith of I interferers experiencing the ith time-varying multipath channel impulse response $h_i(t, \tau)$, "$\bigstar$" denotes convolution in τ, and $z(t)$ is bandlimited additive white Gaussian noise (AWGN). The received signal amplitude spectrum is the Fourier transform of (1)

$$X(f) = \sum_{i=1}^{I} H_i(t, f) S_i(f) + Z(f), \tag{2}$$

where each term in (2) is the Fourier dual of the counterpart in (1).

This paper focuses on interference detection from a secondary-signal-free measurement interval T during which complex samples are obtained at the Nyquist rate $f_S = B$ sufficient to prevent aliasing of the desired signal. A secondary-signal-free N sample vector $\mathbf{x}$ may be obtained either during an interframe space as illustrated in Figure 1(b), or otherwise such as using a self-cancelling pilot symbol [22] (e.g., the IEEE 802.11a short inter-frame space is 16 microseconds and the IEEE 802.16-2004 transmitter turnaround gap for OFDM TDD systems is up to 100 microseconds). Clearly a detection decision based on N samples can be improved upon through subsequent time averaging; however the focus on this paper is to establish a baseline performance based on the initial N sample vector alone.

Through the method described in the appendix, or otherwise, it is assumed that a good quality spectrum estimate described by the N length column vector

$$\hat{\mathbf{X}} \approx \sum_{i=1}^{I} \mathbf{H}_i \mathbf{S}_i + \mathbf{Z} \tag{3}$$

is available in the receiver, where $\mathbf{H}_i$ is the diagonal $N \times N$ channel matrix for the ith interferer, $\mathbf{S}_i$ is the sampled spectrum of the ith interferer, and $\mathbf{Z}$ is AWGN with variance σ_Z^2 per frequency domain sample.

The following assumptions are made concerning the interference.

(1) The duration of $\mathbf{x}$ is sufficiently short that all interference either is present or absent in every time sample (vector element) and that channel matrices $\mathbf{H}_i$ are not time varying.

(2) $\mathbf{s}_i$ is strictly bandlimited such that

$$\left\langle |S_{i,n}|^2 \right\rangle = \begin{cases} \sigma_i^2, & n \in \mathbf{k}_i \triangleq [a_i, \ldots, b_i] \\ 0, & n \notin \mathbf{k}_i, \end{cases} \tag{4}$$

where $\langle \cdot \rangle$ denotes ensemble expectation, $S_{i,n}$ is the nth element of $\mathbf{S}_i$, $(b_i \geq a_i) \in [1, \ldots, N]$, and $\sigma_i^2 = L_i \bar{\sigma}_i^2 / N$, for $L_i = b_i - a_i + 1$, is the frequency domain dual of $\bar{\sigma}_i^2$, the power of the ith interferer as sampled in $\mathbf{x}$.

(3) There is no restriction on $[a_i, \ldots, b_i] \cap [a_j, \ldots, b_j]$: interferers may overlap fully, partially, or not at all. Nor is there any restriction on L_i: the interference model encompasses narrowband interferers.

(4) Elements of $\mathbf{S}_i$ are independent identically distributed (IID) complex Gaussian with zero mean and autocovariance $\mathbf{C}_{S_i S_i} = \sigma_i^2 \mathbf{I}$, and $\mathbf{s}_i$ is complex Gaussian with autocovariance

$$\mathbf{C}_{s_i s_i} = \left\langle \mathbf{s}_i \mathbf{s}_i^H \right\rangle = \mathbf{D} \left\langle \mathbf{S}_i \mathbf{S}_i^H \right\rangle \mathbf{D}^H = \sigma_i^2 \mathbf{D} \mathbf{W}_i \mathbf{D}^H = \sigma_i^2 \mathbf{T}_i, \tag{5}$$

where $\mathbf{D}$ is the discrete Fourier transform (DFT) matrix, $\mathbf{W}_i$ is a diagonal window matrix such that

$$\mathbf{W}_i = \mathrm{diag}([W_0, W_1, \ldots, W_{N-1}]),$$

$$W_n = \begin{cases} 1, & n \in \mathbf{k}_i, \\ 0, & n \notin \mathbf{k}_i, \end{cases} \tag{6}$$

and $\mathbf{T}_i$ is a Toeplitz matrix where the first row is the circular sinc function having the kth element

$$T_{i,1k} = \exp\left(-\frac{j\pi k b_i}{N}\right) \frac{\sin(\pi k L_i / N)}{\sin(\pi k / N)}. \tag{7}$$

The autocorrelation assumption is equivalent to assuming that $\mathbf{s}_i$ is generated from a white IID data sequence. In the absence of multipath fading, where $\mathbf{H}_i \equiv \mathbf{I}$ for all $i \in [1, \ldots, I]$, these assumptions are worst case for detection as the observed $\mathbf{X}$ is Gaussian being the sum of $I + 1$ Gaussian components and the only differences between the unknown interferers and thermal noise are variance and spectral occupancy.

(a) Interference spectra combining and processing at cognitive radio receiver front end

(b) Typical TDD frame structure showing interframe spaces used for interference detection
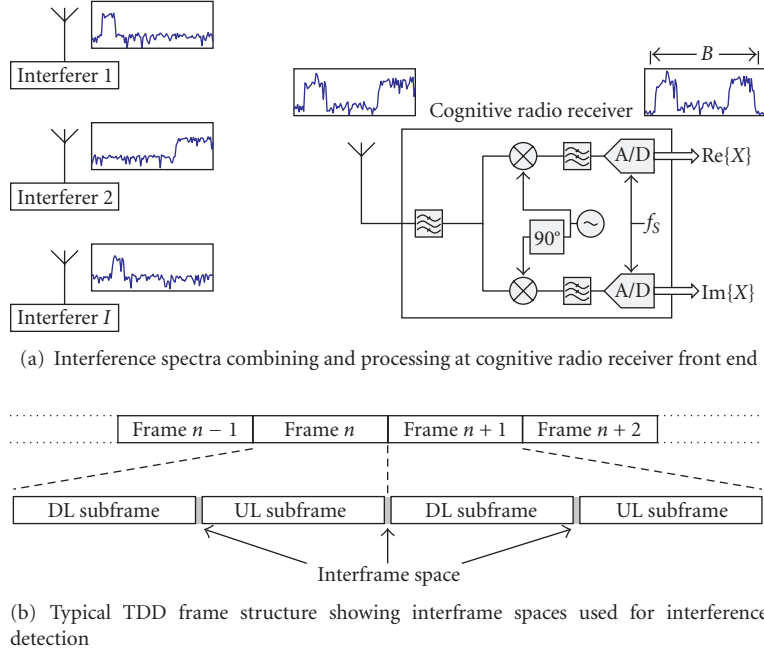
FIGURE 1: System model, showing (a) interference signal spectra at (from left-to-right) interferer transmitters, receiver front end, and after filtering, downcoversion to baseband and digitizing; (b) signal free, interframe spaces during which interference is sampled.

3. Hypothesis Testing and the Generalized Likelihood Ratio Test

Applying classical detection theory, for example, [25, chapter 9], for each frequency bin the receiver must evaluate which of two hypotheses is most likely:

$$\mathcal{H}_0 : X_n = Z_n \quad n \in [1,\dots,N]$$
$$\mathcal{H}_1 : X_n = \sum_{i \in \mathbf{M}(n)} H_{i,nn} S_{i,n} + Z_n \quad n \in [1,\dots,N], \tag{8}$$

where $\mathbf{M}(n) \subseteq [1,\dots,I]$ is the subset of interferers present in the nth frequency bin, and $H_{i,nn}$ denotes the (nn)th element of $\mathbf{H}_i$. Under $\mathcal{H}_1$, frequency sample X_n can comprise any number of interferers between $|\mathbf{M}(n)| = 1$ and $|\mathbf{M}(n)| = I$, where $|\mathbf{M}(n)|$ denotes the cardinality or number of elements in $\mathbf{M}(n)$. For each sample vector $\mathbf{X}$ there will be a finite number $K \le N$ of unique subsets of interferers. These subsets are denoted $\mathbf{M}_1,\dots,\mathbf{M}_K$, noting that $\mathbf{M}_1 \cup \mathbf{M}_2 \cup \cdots \cup \mathbf{M}_K = [1,\dots,I]$. Thus, over the entire signal spectrum vector $\mathbf{X}$, the hypotheses are

$$\mathcal{H}_0 : \mathbf{X} = \mathbf{Z},$$
$$\mathcal{H}_1 : \mathbf{X} = \sum_{k=0}^{K} \sum_{i \in \mathbf{M}_k} \mathbf{H}_{i,\mathbf{l}_k \mathbf{l}_k} \mathbf{S}_{i,\mathbf{l}_k} + \mathbf{Z}_{\mathbf{l}_k}, \tag{9}$$

where $\mathbf{l}_k \subset [0,\dots,N-1]$ is the kth set of indices of adjacent frequency bins with an identical number of interferers, $[\mathbf{l}_1 \cup \mathbf{l}_2 \cup \cdots \cup \mathbf{l}_K] \equiv [\mathbf{k}_1 \cup \mathbf{k}_2 \cup \cdots \cup \mathbf{k}_I]$, and $\mathbf{H}_{i,\mathbf{l}_k \mathbf{l}_k}$ denotes the square diagonal submatrix of $\mathbf{H}_i$ with rows and columns selected by elements of $\mathbf{l}_k$. $\mathbf{l}_0$ is the vector of indices of all

frequency bins, not necessarily adjacent, containing noise only $\mathbf{k}_i$ is defined implicitly in (4) and $\mathbf{M}_0$ is the empty set of interferers.

Hypothesis probabilities may be compared using the generalized likelihood ratio test (GLRT)

$$L_G(\mathbf{X})$$
$$= \frac{p_X\left(\mathbf{X} \,\middle|\, \widehat{\mathbf{M}}_1,\dots,\widehat{\mathbf{M}}_K, \hat{\mathbf{l}}_1,\dots,\hat{\mathbf{l}}_K, \hat{\sigma}_Z^2, \widehat{\mathbf{C}}_{HS_1 HS_1},\dots,\widehat{\mathbf{C}}_{HS_I HS_I}, \mathcal{H}_1\right)}{p_Z\left(\mathbf{X} \mid \hat{\sigma}_Z^2, \mathcal{H}_0\right)}$$

$$> \Gamma_1$$

$$= \max_{\widehat{\mathbf{M}}_1,\dots,\widehat{\mathbf{M}}_K, \hat{\mathbf{l}}_1,\dots,\hat{\mathbf{l}}_K} \frac{p_X\left(\mathbf{X} \mid \hat{\sigma}_Z^2, \widehat{\mathbf{C}}_{HS_1 HS_1},\dots,\widehat{\mathbf{C}}_{HS_I HS_I}, \mathcal{H}_1\right)}{p_Z\left(\mathbf{X} \mid \hat{\sigma}_Z^2, \mathcal{H}_0\right)}, \tag{10}$$

where $\mathbf{C}_{HS_i HS_i}$ is the covariance of $\mathbf{H}_i \mathbf{S}_i$, $p_X(\mathbf{X} \mid \cdots)$ is the multivariate probability density function (PDF) of $\mathbf{X}$ under $\mathcal{H}_1$ and $p_Z(\mathbf{X} \mid \cdots)$ is the noise PDF. One or more interferers are detected when $L_G(\mathbf{X})$ exceeds threshold Γ_1. Evaluation of the GLRT numerator requires estimating how many interferers are present, which interferer is present in each frequency bin and the covariance matrix for each interferer. Thus, complexity of (10) increases exponentially with I.

Provided that interference data is white (interference property 4, above) and the duration of $\mathbf{x}$ is less than the channel coherence time (interference property 1, above), it can be shown that $\mathbf{C}_{HS_i HS_i}$ is diagonal for $\mathbf{s}_i$ produced through any linear modulation, irrespective of any channel conditions additional to the stated assumptions [27]. For $\mathbf{C}_{HS_i HS_i}$ diagonal, $p_X(\mathbf{X} \mid \cdots) = \prod_{n=1}^{N} p_X(X_n \mid \cdots)$ is

simply the likelihood function of the univariate PDF, with dependency on variance not covariance. The GLRT simplifies to

$$
\begin{aligned}
&L_G(\mathbf{X}) \\
&= \frac{\prod_{n=1}^{N} p_X\left(X_n \mid \widehat{\mathbf{M}}_1,\ldots,\widehat{\mathbf{M}}_K,\widehat{\mathbf{l}}_1,\ldots,\widehat{\mathbf{l}}_K,\widehat{\sigma}_Z^2,\widehat{\Omega}_1^2,\ldots,\widehat{\Omega}_I^2,\mathcal{H}_1\right)}{\prod_{n=1}^{N} p_Z\left(X_n \mid \widehat{\sigma}_Z^2,\mathcal{H}_0\right)} \\
&> \Gamma_1 \\
&= \max_{\widehat{\mathbf{M}}_1,\ldots,\widehat{\mathbf{M}}_K,\widehat{\mathbf{l}}_1,\ldots,\widehat{\mathbf{l}}_K} \frac{\prod_{n=1}^{N} p_X\left(X_n \mid \widehat{\sigma}_Z^2,\widehat{\Omega}_1^2,\ldots,\widehat{\Omega}_I^2,\mathcal{H}_1\right)}{\prod_{n=1}^{N} p_Z\left(X_n \mid \widehat{\sigma}_Z^2,\mathcal{H}_0\right)},
\end{aligned}
\tag{11}
$$

where $\widehat{\Omega}_i^2$ is the estimated variance of $\mathbf{H}_{k_i}\mathbf{S}_{k_i}$ as is formally defined in Section 6.2.1. Effectively, $\mathbf{X}$ is partitioned into noise-only samples which have estimated variance $\widehat{\sigma}_Z^2 < \Gamma_1'$, for some threshold Γ_1', and interference plus noise samples which have estimated variance $\widehat{\sigma}_Z^2 + \sum_{i \in \mathbf{M}_k} \widehat{\Omega}_i^2 > \Gamma_1'$. As discussed by Kay [25, page 349] analytical evaluation of Γ_1 for wideband Gaussian signals is infeasible due to the complexity of the PDFs, and it follows that analytical evaluation of the probabilities of false positive detection (type I error) and false negative detection (type II error) also is infeasible.

Pragmatic selection of Γ_1 can be made through Monte Carlo methods. For example, probabilities of type I and type II errors can be found empirically as a function of Γ_1 at some interference to noise ratio (INR) with the aim of producing a constant false alarm rate (CFAR) [28]. For cognitive radio applications, where avoidance of interference may have a higher priority than channel capacity, type I errors may be more acceptable than type II errors and Γ_1 can be set accordingly.

4. Classification Using Hidden Markov Modelling

There are two practical limitations to evaluate the GLRT (10), for this application. Firstly, without constraints on I, the number of interferers, or K, the number of sets of interferers, the GLRT is unable to exploit the implicit characteristic that the ith interferer occupies L_i contiguous frequency bins $X_n,\ldots,X_{n+L_i-1}$. Secondly, the exponential complexity requires consideration. Both limitations can be addressed in a formal context by modelling $\mathbf{X}$ as an observation sequence vector produced by a hidden Markov process. The theory of hidden Markov modelling (HMM) is developed in [29] and applied to the problem of narrowband interference time domain detection for wireless communications in [30]. The following describes how HMM may be applied to perform wideband interference detection in the frequency domain, assuming prior knowledge of the principles of HMM through reference to [29].

Denote the state variable for the nth frequency bin as being in one of $K + 1$ possible states

$$
\mathscr{S}_n \in [\mathbb{S}_0,\ldots,\mathbb{S}_K],
\tag{12}
$$

where state $\mathbb{S}_k$ specifies the set of interferers $\mathbf{M}_k$ plus noise for $k \in [0,\ldots,K]$, noting that $\mathbb{S}_0$ is the noise only state since $\mathbf{M}_0$ is the empty set of interferers. The observation probability for X_n in state $\mathbb{S}_k$ is

$$
\mathcal{P}_k(X_n) \triangleq p(X_n \mid \mathscr{S}_n = \mathbb{S}_k) = p_X(X_n \mid \mathbf{M} = \mathbf{M}_k),
\tag{13}
$$

where $p_X(X_n \mid \cdots)$ is the univariate interference plus noise PDF. A set of N observation probabilities is denoted $\mathcal{P} \triangleq [\mathcal{P}_k(X_n)]$ for $n = 1,\ldots,N$ and the initial state distribution is denoted $\mathscr{I} \triangleq [p(\mathscr{S}_1 = \mathbb{S}_k)]$ for $k = 0,\ldots,K$. Critical to this model is the transition probability between any two states $\mathbb{S}_\ell$ and $\mathbb{S}_k$

$$
\mathcal{T}_{\ell k} \triangleq p(\mathscr{S}_{n+1} = \mathbb{S}_k \mid \mathscr{S}_n = \mathbb{S}_\ell),
\tag{14}
$$

this being the innovation over the GLRT. The posterior transition probabilities are (the posterior transition probability from $\mathbb{S}_\ell$ to any state $\mathbb{S}_k \neq \mathbb{S}_m$ is zero following the definition after (9) that $\mathbf{l}_\ell$ is a set of *adjacent* frequency bins, thus interference state $\mathbb{S}_\ell$ transitions exactly once—to $\mathbb{S}_m$ in the definition scenario)

$$
\mathcal{T}_{\ell k} = \begin{cases} \dfrac{(|\mathbf{l}_\ell| - 1)}{|\mathbf{l}_\ell|}, & k = \ell, \\[2mm] \dfrac{1}{|\mathbf{l}_\ell|}, & k = m, \\[2mm] 0, & k \neq \ell, m, \end{cases}
\tag{15}
$$

for the state sequence $[\cdots \mathscr{S}_{n-|\mathbf{l}_\ell|+1}, \mathscr{S}_{n-|\mathbf{l}_\ell|+2},\ldots, \mathscr{S}_n, \mathscr{S}_{n+1} \cdots]$ given the posterior state assignments $\mathscr{S}_{n-|\mathbf{l}_\ell|+1} = \mathscr{S}_{n-|\mathbf{l}_\ell|+2} = \cdots = \mathscr{S}_n = \mathbb{S}_\ell$, $\mathscr{S}_{n+1} = \mathbb{S}_m$ which define m, and where $\mathbf{l}_\ell$ was introduced in (9). A set of N transition probabilities is denoted $\mathcal{T} \triangleq [\mathcal{T}_{\ell k}]$ for $\ell,k = 0,\ldots,K$. The output of defining the HMM for sample vector $\mathbf{X}$ is a state sequence $\mathscr{S} \triangleq \mathscr{S}_1, \mathscr{S}_2,\ldots,\mathscr{S}_N$ describing whether each frequency domain sample X_n is best classified as noise, $\mathscr{S}_n = \mathbb{S}_0$, or interference $\mathscr{S}_n \in [\mathbb{S}_1,\ldots,\mathbb{S}_K]$. This is achieved by finding the set of model parameters $[\mathcal{T},\mathcal{P},\mathscr{I}]$ which maximise $p(\mathbf{X} \mid \mathcal{T},\mathcal{P},\mathscr{I})$ given observation vector $\mathbf{X}$.

The following sequential process describes one possible implementation for how the HMM described above can be used to estimate state sequence $\mathscr{S}$.

Step 1: Estimate the Noise Variance. This is achieved first by performing a preclassification to partition $\mathbf{X}$ into noise-only samples and interference plus noise samples, as outlined in Section 7.

Step 2: Produce a Nominal Interference Variance. To be able to evaluate $\mathcal{P}_k(X_n)$, it is necessary to estimate interference variance. However, at this point in the process, it is not known how many interferers—if any—are present. So explicit estimation of variance for individual interferers is problematic. Further, estimation of variance for composite interference is of little value as this will tend to produce an estimate dominated by any large interferers, leading to undesirable type II errors for low power interferers. One means of reducing this problem is to produce a

TABLE 1: Optimum threshold parameter for each classification method, being that value for which the probabilities of type I and type II errors were most similar for a standard interferer at INR of 3 dB.

Method	Threshold parameter value	
Generalized log likelihood ratio test (Section 3)	$\Gamma_1 = 0.6$	
Hidden Markov modelling (Section 4)	$r = 6.25$	$\tau_0 = 10$
Radiometry (Section 5.1)	$\Gamma_Z = 0.6$	
Power spectrum analysis (Section 5.2)	$\Gamma_Z = 0.6$	

nominal interference variance by defining an acceptable type I error threshold (CFAR), that is, in absence of any explicit interference estimation. For example, a nominal interference variance

$$\hat{\sigma}_I^2 \triangleq r\hat{\sigma}_Z^2 \tag{16}$$

will produce a type I error for any noise sample for which $|X_n|^2 > 2\hat{\sigma}_Z^2(r/(r-1))\log r$ for Gaussian $\mathbf{Z}$ (see Table 1 for empirical determination of scaling factor r).

Step 3: Perform an Initial Classification by Estimating the HMM State Sequence. To initialise the state sequence $\mathcal{S}$, the two state set $[\mathbb{S}_0, \mathbb{S}_1]$ is considered, where $\mathbb{S}_0$ is the noise-only state and $\mathbb{S}_1$ is the interference-plus-noise state. The observation probabilities $\mathcal{P}$ are evaluated using

$$\mathcal{P}_0(X_n) = p_Z(X_n) \tag{17}$$

for some noise PDF $p_Z(\cdot)$, and

$$\mathcal{P}_1(X_n) = p_X(X_n)\big|_{\mathbf{M}=[1],\sigma_1=\hat{\sigma}_I}, \tag{18}$$

for some interference-plus-noise PDF $p_X(\cdot)$. The set of transition probabilities $\mathcal{T}$ is initialised using some nominal τ_0, such that $\mathcal{T}_{01} = \mathcal{T}_{10} = 1/\tau_0$, $\mathcal{T}_{00} = \mathcal{T}_{11} = (\tau_0 - 1)/\tau_0$, for all $n = [1, \ldots, N]$ (see Table 1 for empirical determination of τ_0). The initial value for $\mathcal{S}$ is determined through evaluation of

$$\mathcal{S} = \arg\max_{\mathcal{S}} p(\mathbf{X} \mid \mathcal{S}, \mathcal{P}, \mathcal{T}, \mathbf{l}) p(\mathcal{S} \mid \mathcal{P}, \mathcal{T}, \mathbf{l}), \tag{19}$$

which may be performed efficiently by applying the Viterbi algorithm [31] to $\mathbf{X}$ with $\mathcal{P}$, $\mathcal{T}$ as above and $\mathbf{l} = [\mathcal{P}_0(X_1), \mathcal{P}_1(X_1)]$. Complexity is set by the Viterbi algorithm window/traceback length and through application of the M-algorithm [32].

Step 4: Perform a Final Classification by Estimating the HMM State Sequence. A first estimate of the actual number of interferers now can be made by sorting the K subsequences of $\mathcal{S}$, with length $|\mathbf{l}_k|$, $k \in [1, \ldots, K]$, for which $\mathcal{S}_{n-|\mathbf{l}_k|+1} = \mathcal{S}_{n-|\mathbf{l}_k|+2} = \cdots = \mathcal{S}_n = \mathbb{S}_1$ by order of increasing (implicit) variance. The K ordered sub-sequences now are reclassified as belonging to $\mathbb{S}_1, \ldots, \mathbb{S}_K$, and the estimated variance for

each interferer in state $\mathbb{S}_k$ is required to define $\mathcal{P}_k(X_n)$. Variance estimation methods for different fading models are described in Section 6. Finally, the state sequence $\mathcal{S}$ then may be updated for the expanded state set $[\mathbb{S}_0, \ldots, \mathbb{S}_K]$, again by evaluating (19) using the Viterbi algorithm. Interference is detected if $\mathcal{S}_n \neq \mathbb{S}_0$ for any $n \in [1, \ldots, N]$.

This process can be iterated again to improve classification reliability, but one iteration empirically was observed to achieve good results in achieving the detection goal of classifying frequency bin samples as being either noise or interference plus noise.

5. Comparative Classification Methods

This section describes three types of interference detection methods, used in Section 8 for performance assessments in comparison to the GLRT and HMM.

5.1. Radiometry [16–18]. The energy detector [13–15] is a low complexity method of classifying $\mathbf{X}$. This method estimates some noise power threshold Γ_Z then decides $\mathcal{H}_1$ if $\max(|\mathbf{X}|) > \Gamma_Z$ [7]. Where $\mathcal{H}_1$ is decided, the nth sample of $\mathbf{X}$ is classified as interference if $|X_n| > \Gamma_Z$. For the simplified case of a diagonal covariance matrix, the posterior probability of type I error for this test is

$$p_{\text{Type I}}(\mathcal{H}_1 \mid |\mathbf{X}|, \mathcal{H}_0) = 1 - \left(\int_0^{\Gamma_Z} p_Z^M(|X_n|) d|X_n| \right)^N, \tag{20}$$

where $p_Z(|X_n|)$ is the noise magnitude PDF. The posterior probability of a type II error then is

$$\begin{aligned} p_{\text{Type II}}(\mathcal{H}_0 \mid |\mathbf{X}|, \mathcal{H}_1) = 1 &- \left(\int_{\Gamma_Z}^{\infty} p_X^M(|X_n|) d|X_n| \right)^{|\mathbf{L}|} \\ &- \left(\int_{\Gamma_Z}^{\infty} p_Z^M(|X_n|) d|X_n| \right)^{N-|\mathbf{L}|}, \end{aligned} \tag{21}$$

where $p_X^M(|X_n|)$ is the interference plus noise magnitude PDF and $|\mathbf{L}|$ is the cardinality of $\mathbf{L} \triangleq [\mathbf{l}_1 \cup \mathbf{l}_2 \cup \cdots \cup \mathbf{l}_K]$. On a per-sample-bin basis, this test can produce a similar result to the GLRT; noting that the GLRT described in (11) performs a nonlinear transformation of $\mathbf{X}$. One method for estimating Γ_Z is by using the nominal interference variance described in (16).

The distinction between the radiometer and the energy detector is that the radiometer typically uses substantial time averaging to produce $\mathbf{X}$, often implemented using a filter bank. Naturally, the performance of any method can be improved by increasing the sample vector duration. For this study, where the duration of $\mathbf{x}$ is strictly limited to N samples, the effect of applying the principle of radiometry was investigated by subdividing $\mathbf{x}$ into a number of overlapping subvectors to obtain multiple estimates of $\mathbf{X}$ which then were averaged. Intuitively, some improvement in estimation accuracy is expected but at the cost of decreased spectral resolution as each sub-vector of $\mathbf{x}$ must be smaller than $\mathbf{x}$ itself.

5.2. Power Spectrum Analysis [24]. Like the GLRT, energy detection is unable to exploit the interference characteristic that the *i*th interferer occupies L_i contiguous frequency bins. A number of methods have been proposed which aim to enhance the performance of energy detection through some form of spectral analysis. The method chosen for reference, reported in [24], first performs low pass filtering on **X** to reduce variation between adjacent frequency bins, performs thresholding to identify interference-affected bins, and then estimates the centre frequency and bandwidth of each detected interferer. The net effect is to perform edge enhancement on each wideband interferer, exploiting the contiguous bin interference characteristic.

5.3. Information Theoretic Criteria [5]. An alternative approach to threshold-based methods is to estimate the composite interference signal spectrum directly from time sequence **x** through eigenanalysis. As with radiometry, this method compromises spectral resolution—in this case through the process of estimating the covariance matrix [5, 33, page 373] used to produce the eigenvalues—as a tradeoff against avoiding the need to determine any energy threshold. However, estimation of the signal space dimension is required, and suggested methods [5] are Akaike information criteria (AIC) [34] and minimum description length (MDL) [34, 35]. The power spectrum estimate is then sorted by power, and the number of signal (interference) bins suggested through AIC or MDL is identified. Effectively, the signal space dimension defines the energy threshold, so this method intuitively would be expected to perform similarly to energy detection.

6. Received Interference Spectrum Statistics

This section develops the received signal statistics required by the interference detection methods above. The AWGN and interferer statistics developed below generally are well known refer, for example, to [36]. Each of the various spectrum sensing methods may operate most effectively using different statistical representations, for example, density of magnitude or joint density of the real and imaginary part. For this reason, each noise and interference statistical model is represented in a variety of forms below, providing consistency of definition and notation for prospective implementers of the spectrum sensing methods.

6.1. AWGN Statistics. The multivariate PDF of the real part of **Z**, denoted $\mathbf{Z}^R \triangleq \operatorname{Re}\{\mathbf{Z}\}$, is the standard Gaussian multivariate PDF

$$p_Z\left(\mathbf{Z}^R\right) \triangleq \frac{1}{\sqrt{2\pi}\sigma_Z}\exp\left(-\frac{1}{2}\left[\frac{\mathbf{Z}^R}{\sigma_Z}\right]^2\right)$$

$$\triangleq \prod_{n=1}^{N}\frac{1}{\sqrt{2\pi}\sigma_Z}\exp\left(-\frac{1}{2}\left[\frac{Z_n^R}{\sigma_Z}\right]^2\right),\tag{22}$$

noting that the PDF of the imaginary part $\mathbf{Z}^I \triangleq \operatorname{Im}\{\mathbf{Z}\}$ has identical form. For the purposes of definition and clarity, the

multivariate joint PDF of noise real and imaginary parts is given here as

$$p_Z^J(\mathbf{Z}) \triangleq p_Z\left(\mathbf{Z}^R, \mathbf{Z}^I\right) = p_Z\left(\mathbf{Z}^R\right)p_Z\left(\mathbf{Z}^I\right)$$

$$= \frac{1}{2\pi\sigma_Z^2}\exp\left(-\frac{1}{2}\left[\frac{|\mathbf{Z}|}{\sigma_Z}\right]^2\right),\tag{23}$$

and the multivariate magnitude PDF of **Z** is

$$p_Z^M(|\mathbf{Z}|) = \frac{|\mathbf{Z}|}{\sigma_Z^2}\exp\left(-\frac{1}{2}\left[\frac{|\mathbf{Z}|}{\sigma_Z}\right]^2\right).\tag{24}$$

Further standard results inform that the first four moments and cumulants of $\mathbf{Z}^R$ are identically zero except the second moment (μ_2) and second cumulant (κ_2) $\mu_2(\mathbf{Z}^R) = \kappa_2(\mathbf{Z}^R) = \sigma_Z^2$ and fourth moment $\mu_4(\mathbf{Z}^R) = 3\sigma_Z^4$. Further, the kurtosis of $\mathbf{Z}^R$ also is zero, where kurtosis is defined as [37]

$$\gamma_2 \triangleq \frac{\mu_4}{\mu_2^2} - 3 \equiv \frac{\kappa_4}{\kappa_2^2}.\tag{25}$$

6.2. Received Interference Plus Noise Statistics. This subsection develops interference plus noise statistics under two channel scenarios: Rayleigh fading and no fading (Gaussian channel), as follows.

6.2.1. Rayleigh Fading Channel. In this model, diagonal elements of $\mathbf{H}_i$ are correlated complex Gaussian random variables, although elements of $\mathbf{H}_i\mathbf{S}_i$ are uncorrelated as discussed in Section 3. The PDF of $U_{i,n}^R \triangleq \operatorname{Re}\{H_{i,nn}S_{i,n}\}$, the real part of an individual, noise-free frequency bin in which interferer *i* is present, is well known as

$$p_U\left(U_{i,n}^R\right) = \frac{1}{(2\Omega_i)}\exp\left(-\frac{\left|U_{i,n}^R\right|}{\Omega_i}\right),\tag{26}$$

where $H_{i,nn}$ is the (nn)th element of $\mathbf{H}_i$ and $S_{i,n}$ is the nth element of $\mathbf{S}_i$, $\Omega_i \triangleq \sigma_i\sigma_{\mathbf{H}_i}$ for $\sigma_{\mathbf{H}_i}^2 = (1/2)\langle|\mathbf{H}_i|^2\rangle$ the power gain of the *i*th multipath channel, and noting that this also is the PDF of $U_{i,n}^I \triangleq \operatorname{Im}\{H_{i,nn}S_{i,n}\}$. From (26), the PDF of $X_n^R \triangleq \operatorname{Re}\{X_n\} = \operatorname{Re}\{\sum_{i\in\mathbf{M}}H_{i,nn}S_{i,n} + Z_n\}$, the real part of a frequency bin in which one or more interferers is present, may be derived by first deriving the first characteristic function of X_n^R

$$\Phi_X(\theta) = \prod_{i\in\mathbf{M}}\frac{1}{1+\Omega_i^2\theta^2}\exp\left(-\frac{1}{2}\sigma_Z^2\theta^2\right).\tag{27}$$

Therefore, it follows that the PDF of X_n^R is given by

$$p_{X1}\left(X_n^R\right) = \sum_{i\in\mathbf{M}}\exp\left(\frac{\sigma_Z^2}{2\Omega_i^2}\right)\frac{\Omega_i}{4\prod_{j\in\mathbf{M},\,j\neq i}\left(\Omega_i^2 - \Omega_j^2\right)}$$

$$\times\left[\exp\left(\frac{X_n^R}{\Omega_i}\right)\operatorname{erfc}\left(\frac{\sigma_Z}{\sqrt{2}\Omega_i} + \frac{X_n^R}{\sqrt{2}\sigma_Z}\right)\right.$$

$$\left. + \exp\left(-\frac{X_n^R}{\Omega_i}\right)\operatorname{erfc}\left(\frac{\sigma_Z}{\sqrt{2}\Omega_i} - \frac{X_n^R}{\sqrt{2}\sigma_Z}\right)\right],$$

$$\tag{28}$$

noting that this also is the PDF of $X_n^I \triangleq \text{Im}\{X_n\} = \text{Im}\{\sum_{i \in \mathbf{M}} H_{i,nn} S_{i,n} + Z_n\}$. The joint PDF

$$p_{X1}^J(X_n) \triangleq p_{X1}(X_n^R, X_n^I) = p_{X1}(X_n^R) p_{X1}(X_n^I) \qquad (29)$$

readily can be evaluated through substitution of (28), but the magnitude PDF

$$p_{X1}^M(|X_n|) = |X_n| \int_{-\pi}^{\pi} p_{X1}(|X_n| \cos \phi) p_{X1}(|X_n| \sin \phi) \, d\phi \qquad (30)$$

requires numerical integration.

Moments and cumulants of X are required for parameter estimation. From the first characteristic function of X_n^R, (27), the second characteristic function easily is derived as

$$\Psi_X(\theta) = -\sum_{i \in \mathbf{M}} \log\left(1 + \Omega_i^2 \theta^2\right) - \frac{1}{2} \sigma_Z^2 \theta^2. \qquad (31)$$

By exploiting the moment generating property of the first characteristic function (27), and cumulant generating property of the second characteristic function (31), the first two even order moments (μ_2, μ_4) and cumulants (κ_2, κ_4) of X_n^R for this channel model can be found to be

$$\mu_2\left(X_n^R\right) \triangleq \left\langle \left(X_n^R\right)^2 \right\rangle = \sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2,$$

$$\kappa_2\left(X_n^R\right) = \sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2,$$

$$\mu_4\left(X_n^R\right) \triangleq \left\langle \left(X_n^R\right)^4 \right\rangle = 3\sigma_Z^4 + 12\sigma_Z^2 \sum_{i \in \mathbf{M}} \Omega_i^2 + 24 \sum_{i \in \mathbf{M}} \Omega_i^4,$$

$$\kappa_4\left(X_n^R\right) = 12 \sum_{i \in \mathbf{M}} \Omega_i^4. \qquad (32)$$

The kurtosis of X_n^R easily is derived from (32) and (25) as

$$\gamma_2\left(X_n^R\right) = \frac{12 \sum_{i \in \mathbf{M}} \Omega_i^4}{\left(\sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2\right)^2}. \qquad (33)$$

6.2.2. Gaussian Channel (No Fading). This model applies either where strong line-of-sight conditions exist or otherwise where multipath fading is negligible, such that $\mathbf{H}_i = \sigma_{H_i} \mathbf{I}$, for all i. Under this assumption, the PDF of X_n^R reduces to

$$p_{X2}\left(X_n^R\right) = p_2\left(X_n^I\right) = \frac{1}{\sqrt{2\pi\sigma_{X2}^2}} \exp\left(-\frac{1}{2}\left[\frac{X_n^R}{\sigma_{X2}}\right]^2\right), \quad (34)$$

where $\sigma_{X2}^2 \triangleq \sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2$. In this case, the joint PDF directly can be written as

$$p_{X2}^J(X_n) \triangleq p_{X2}\left(X_n^R, X_n^I\right)$$

$$= \frac{1}{2\pi\sigma_{X2}^2} \exp\left(-\frac{1}{2}\left[\frac{|X_n|}{\sigma_{X2}}\right]^2\right), \qquad (35)$$

and the magnitude PDF is

$$p_{X2}^M(|X_n|) = \frac{|X_n|}{\sigma_{X2}^2} \exp\left(-\frac{1}{2}\left[\frac{|X_n|}{\sigma_{X2}}\right]^2\right). \qquad (36)$$

The first two even order moments and cumulants of X_n^R for this channel model are

$$\mu_2\left(X_n^R\right) = \sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2,$$

$$\kappa_2\left(X_n^R\right) = \sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2,$$

$$\mu_4\left(X_n^R\right) = 3\left(\sigma_Z^2 + \sum_{i \in \mathbf{M}} \Omega_i^2\right)^2, \qquad (37)$$

$$\kappa_4\left(X_n^R\right) = 0.$$

From (37) and (25), the kurtosis of X_n^R in this case is zero.

7. Estimation of Noise Variance

Estimation of the noise variance of $\mathbf{X}$ is essentially an initial classification of $\mathbf{X}$ into noise and interference and may be achieved as follows.

Step 1. Sort $\mathbf{X}$ by order of increasing magnitude to produce $\vec{\mathbf{X}}$.

Step 2. Find $\hat{\mathbf{X}}_Z \triangleq [\vec{\mathbf{X}}_1, \ldots, \vec{\mathbf{X}}_{N_Z}]$, the sub-vector of $\vec{\mathbf{X}}$ which minimizes distance from Gaussianity, for

$$\hat{N}_Z = \arg\min_{N_Z} \mathrm{D}\left(\vec{\mathbf{X}}_Z\right), \qquad (38)$$

where $\mathrm{D}(\cdot)$ is some measure of Gaussianity, such as kurtosis [37, page 86] and [38] or the Kolmogorov-Smirnov (K-S) distance [39]. From (25), a moment method estimate of kurtosis is

$$D_{\gamma_2} \triangleq \hat{\gamma}_2\left(\vec{\mathbf{X}}_Z\right)$$

$$= \frac{2N_Z \sum_{n=1}^{N_Z} \mathrm{Re}\{X_n\}^4}{\left[\sum_{n=1}^{N_Z} \mathrm{Re}\{X_n\}^2\right]^2} + \frac{2N_Z \sum_{n=1}^{N_Z} \mathrm{Im}\{X_n\}^4}{\left[\sum_{n=1}^{N_Z} \mathrm{Im}\{X_n\}^2\right]^2} - 3. \qquad (39)$$

The K-S distance may be evaluated as

$$\mathrm{D}_{\text{K-S}} \triangleq \max\left\{\left|p_Z\left(\vec{\mathbf{X}}_Z < x\right) - \frac{\mathbf{N}_Z}{N_Z}\right|\right\}, \qquad (40)$$

where $p_Z(\mathbf{X} < x)$ is the cumulative distribution function (CDF) for dummy variable x and $\mathbf{N}_Z \triangleq [1, \ldots, N_Z]$. In general, the K-S distance is more robust, as it involves all moments; whereas kurtosis involves the fourth cumulant (or second and fourth moments) only. Note that, in calculating the K-S distance, it is equally valid to evaluate the distance of the real and imaginary parts of $\vec{\mathbf{X}}_Z$ from Gaussian, the distance of the $|\vec{\mathbf{X}}_Z|$ from Rayleigh, or the distance of $|\vec{\mathbf{X}}_Z|^2$
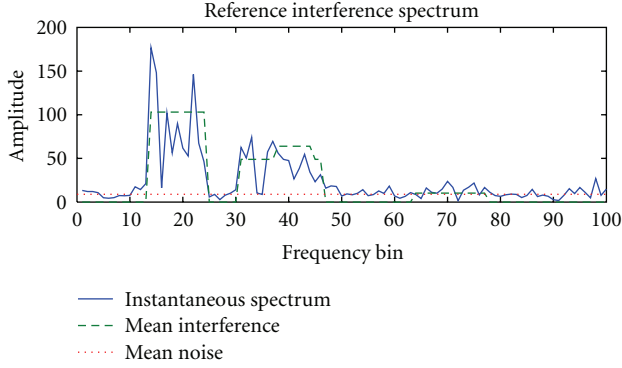
FIGURE 2: An example received interference sample vector spectrum. The instantaneous spectrum of the sample vector is shown in comparison to the ensemble mean spectrum that would be obtained by averaging many sample vectors with identical interference statistics. The ensemble mean noise root variance also is shown for reference. This example sample vector is used as a reference to illustrate the detection methods described in Sections 3 to 5. Note the low power interferer occupying frequency bins 64 to 77.

from exponential. Note also that *truncated* CDFs are required to evaluate the K-S test. For data truncated at $T_X = |\vec{X}_Z|_{N_z+1}$, the truncated Rayleigh CDF is evaluated as the conditional probability

$$p_Z(|X| < x \mid |X| < T_X) = \frac{1 - \exp(-x^2/(2\sigma_Z^2))}{1 - \exp(-T_X^2/(2\sigma_Z^2))}. \quad (41)$$

Step 3. Estimate the noise variance. From (41), the maximum likelihood estimate variance estimate based on the Rayleigh distribution truncated at T_X requires numerical solution of

$$\hat{\sigma}_Z^2 - \frac{T_X^2}{2}\left[\frac{\exp(-T_X^2/(2\hat{\sigma}_Z^2))}{1 - \exp(-T_X^2/(2\hat{\sigma}_Z^2))}\right] = \frac{1}{2N_Z}\sum_{n=1}^{N_Z}\left|\vec{X}_n\right|^2, \quad (42)$$

which reduces to the conventional maximum likelihood estimator for $T_X \gg \sigma_Z$ and as $T_X \to \infty$.

8. Simulation-Based Performance Comparison of Classification Methods

The length $N = 100$ amplitude spectrum of a reference received sample vector, shown in Figure 2, is a typical example produced according to the model described in Section 2. The sample vector (solid line) comprises four interferers of varying bandwidths and ensembles average power spectral densities (PSDs) (dashed line). Note that one interferer—present in frequency bins 38 through to 44—sits entirely within the spectrum of a second interferer, while a third interferer—present in bins 64 through to 77—has mean PSD which is almost identical to the noise PSD (dotted line). In this example both noise and interference per bin were modelled as IID complex Gaussian.

To illustrate the operation of the various classification methods, each was applied to the reference sample
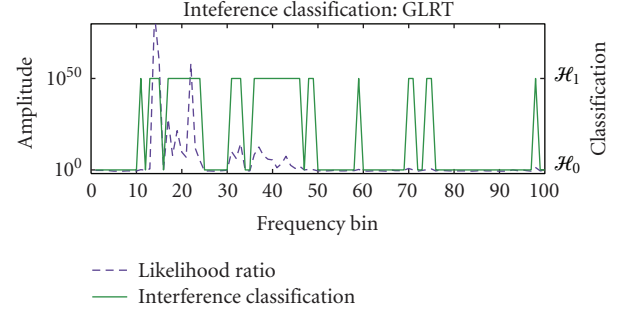

FIGURE 3: An example interference classification produced using the generalized likelihood ratio test on a per bin basis. The classification should be compared against reference sample vector interference spectrum shown in Figure 2. Several type I and type II errors can be seen.
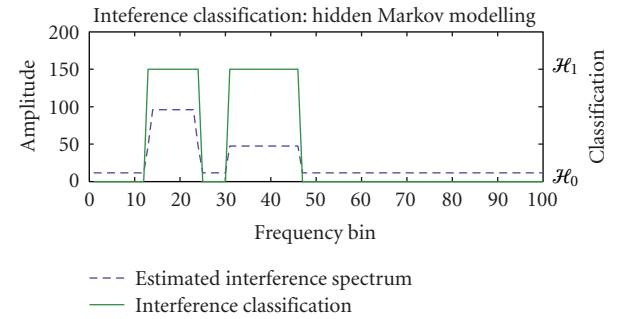

FIGURE 4: An example interference classification produced using hidden Markov modelling. The classification should be compared against reference sample vector interference spectrum shown in Figure 2. In this example, this method failed to detect the low power interferer occupying frequency bins 64 through to 77, but otherwise produced an accurate classification.

vector spectrum, with the results shown in Figures 3–7. In each figure, the dashed line shows the decision vector for that method and the solid line shows the resulting classification, where a classification of zero indicates noise and a nonzero classification indicates signal. The GLRT classification, shown in Figure 3, has a number of type I and type II errors per bin which can be observed by comparison with the reference sample vector in Figure 2. The classification output of the second pass of HMM, shown in Figure 4, performed well for the higher power interferers. The radiometer classification, shown in Figure 5, produced a number of type I and type II errors per bin for the reference sample vector. The power spectrum analysis method classification, shown in Figure 6, correctly classified the high power interferers. The information theoretic classification, shown in Figure 7, produced a classification similar to that of the GLRT, confirming the intuitive expectation discussed in Section 5.3. All methods failed to detect the low power interferer in bins 64 to 77.

8.1. Detection Performance for a Standard, Single, Wideband Interferer. To estimate detection performance as a function of signal-to-interference ratio (INR), and in order to set
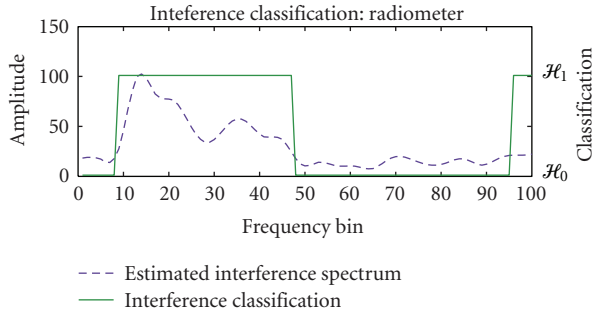
FIGURE 5: An example interference classification produced using the radiometer method. The classification should be compared against reference sample vector interference spectrum shown in **Figure 2**. In this example, this method failed to detect the low power interferer occupying frequency bins 64 through to 77, but otherwise produced an accurate classification.
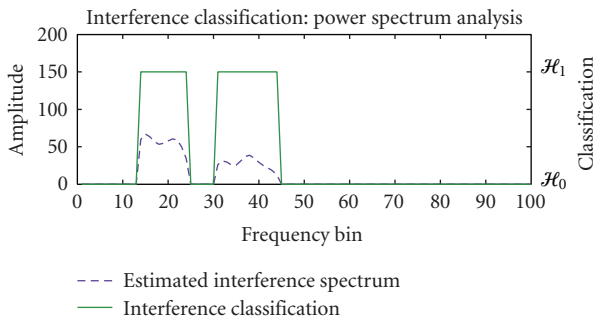

FIGURE 6: An example interference classification produced using power spectrum analysis. The classification should be compared against reference sample vector interference spectrum shown in **Figure 2**. In this example, this method failed to detect the low power interferer occupying frequency bins 64 through to 77, but otherwise produced an accurate classification.
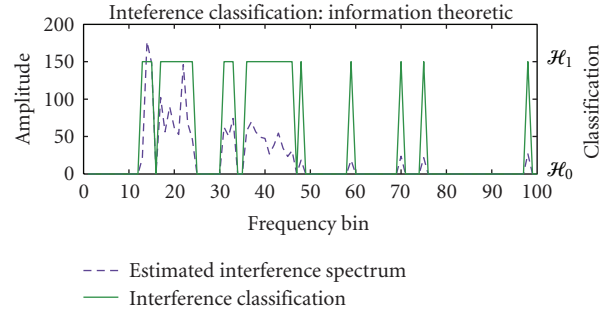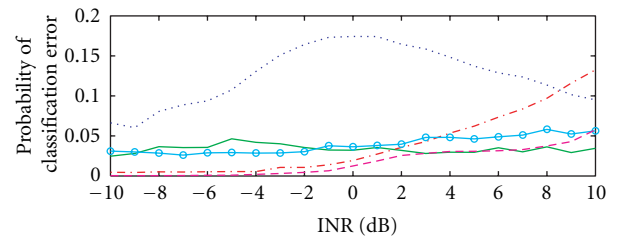

FIGURE 7: An example interference classification produced using information theoretic criteria. The classification should be compared against reference sample vector interference spectrum shown in **Figure 2**. In this example, this method mostly failed to detect the low power interferer occupying frequency bins 64 through to 77 and produced several type I and type II errors.

(a) Type I error (false positive) rate
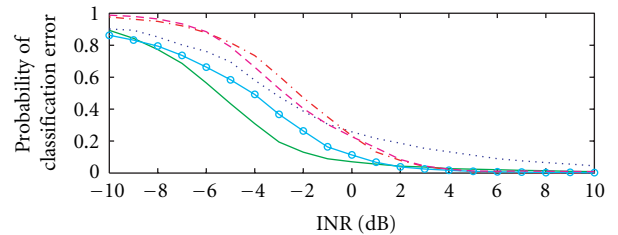
(b) Type II error (false negative) rate

FIGURE 8: Detection performance, showing the probability of classification error per frequency bin averaged over 1000 trials for an $N = 100$ length sample vector. One standardized interferer occupying 20% of the receiver input bandwidth was used for all simulations. For each method, the observed probability of type I errors is shown in (a) and the observed probability of type II errors is shown in (b). The legend in (b) applies to both subfigures.

threshold parameters for the various methods, repeated detection trials were performed for a single standard interferer. The size of the received signal vector was to set $N = 100$, and the interferer was bandlimited to 20% of the received signal bandwidth with a low-pass-equivalent carrier frequency of zero. Thus, from (3), $I = 1$ and, from (4), $L = 20$, $a = 41$, and $b = 60$. The interferer was generated as being IID complex Gaussian in each frequency bin in $\mathbf{k}_1$. Results for each method as a function of INR were obtained by averaging the performance of 1000 detections.

For each method, a family of results for probabilities of type I and type II errors as functions of INR was produced for different values of the respective threshold parameter. The threshold parameter value for which the probability of type I and type II errors was most similar at an INR of 3 dB, given in Table 1, then was used to produce the comparative results shown in Figure 8.

Figure 8(a) indicates that the methods producing the lowest $\sum_{n=1}^{N} p_{\text{TypeI}}(\mathcal{H}_1 \mid X_n, \mathcal{H}_0)$ at all INRs were HMM, power spectrum analysis, and the information theoretic method. Figure 8(b) indicates that the methods producing the lowest $\sum_{n=1}^{N} p_{\text{TypeII}}(\mathcal{H}_0 \mid X_n, \mathcal{H}_1)$ were HMM and power spectrum analysis with HMM performing best at low INR. Note that the two methods which produced the least type I errors at low INR—information theoretic and radiometer—also produced the most type II errors at low INR. HMM and power spectrum analysis produced the lowest average probability of classification error, with HMM being the best performed overall.

(a) ROC

.... GLRT
———— Hidden Markov model
—·—·— Radiometer

—⊙— Power spectrum
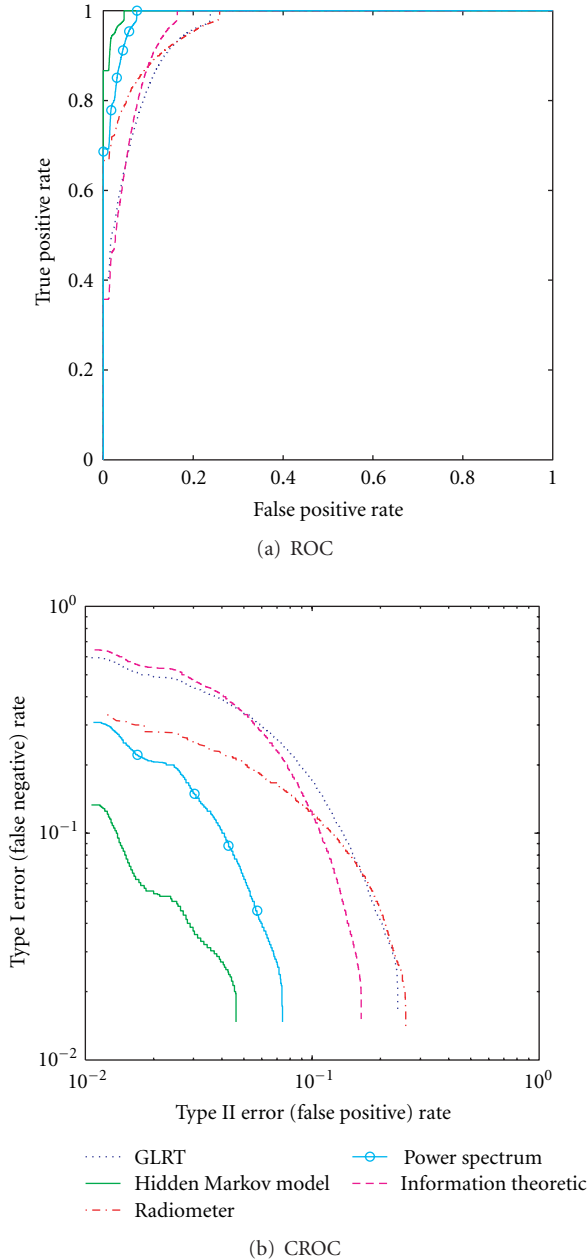— — — Information theoretic

(b) CROC

FIGURE 9: Low INR detection performance, showing (a) the receiver operating characteristic and (b) the complementary receiver operating characteristic, for an $N = 100$ length sample vector produced by aggregating results of 100 000 trials. Four interferers of random bandwidth, power, and demodulated carrier frequency were used in each simulation, with INRs uniformly distributed over the range -10 dB to $+10$ dB. The legend in (a) applies to both subfigures. Note that the best performance region of the ROC (a) is the top left and that the best performance region of the CROC (b) is the bottom left.

8.2. Detection Performance for Multiple Wideband Interferers. A more rigorous performance assessment of the detection methods was performed by simulating four wideband interferers with random and unknown parameters. Each interferer bandwidth was randomly chosen between 0% and 20% of the receive bandwidth, with a low-pass-equivalent

carrier frequency uniformly randomly distributed between -0.8 and $+0.8$ of the sample frequency. The INR for each interferer was uniformly randomly distributed in decibels between -10 dB and $+10$ dB. Interferers were generated as being IID complex Gaussian in each frequency bin in $\mathbf{k}_i$ for each i. Classification errors were recorded for each length 100 sample vector and averaged over 100 000 detections for each method.

Figures 9(a) and 9(b) show, respectively, the receiver operating characteristic (ROC) and the complementary receiver operating characteristic (CROC) for each method. The CROC best illustrates that HMM and power spectrum analysis are the best performed methods, being the only two that meet the IEEE 802.22 target of 0.9 detection probability and 0.1 false alarm rate [40]. HMM can be seen to significantly out-perform power spectrum analysis.

The key to the performance of HMM is that the model framework incorporates into the detection process the assumption that wideband interferers occupy more than one adjacent frequency bin of signal vector spectrum. This is achieved through the HMM transition probability. Thus, a single low observation probability in a sequence may be balanced by the HMM in sequence likelihood by the cost of two state transitions, resulting in fewer state transitions in the final classification.

8.3. Discussion. The simulation results show that all of the methods trialled exhibited relatively high probabilities of producing classification errors averaged over the INR range. Most problematic for each method was the common case of a sample vector in which the instantaneous PSD at either edge of an interferer passband was substantially lower than the noise PSD for that sample vector. For cognitive radio applications, where prevention of interference is paramount, inclusion of a "guard band" around each detected interferer would mitigate against the risk of interference resulting from a type II error, at the expense of increasing the probability of a type I error. Increasing the length of the sample vector and using more than one sample vector to make the detection decision also would improve performance of all methods.

9. Conclusion

This paper has investigated detection of wideband interference ("spectrum sensing") from multiple heterogeneous system sources which are unknown, stochastic, and not cyclostationary. All blind spectrum sensing methods are based on the hypothesis testing framework reviewed here, either implicitly or explicitly. Several methods of performing detection by classifying the received sample vector either as containing interference plus noise or as containing noise only were summarised. Of these, two methods were shown to allow explicit incorporation into the detection process of the assumption that wideband interferers occupy more than one adjacent frequency bin of signal vector spectrum: hidden Markov modelling and power spectrum analysis with edge enhancement. The detection performance of these methods was shown by computer simulation generally to be better

than that of methods which did not allow incorporation of the occupancy assumption. Of the methods considered, hidden Markov modelling exhibited the best overall performance determined through two types of detection test.

Appendix

Spectrum Estimation

The N-length received column vector obtained through windowing and sampling $x(t)$ from (1) is

$$\widehat{\mathbf{x}} = \mathbf{w}\left(\sum_{i=1}^{I} \widetilde{\mathbf{h}}_i \widetilde{\mathbf{s}}_i + \widetilde{\mathbf{z}}\right), \tag{A.1}$$

where $\mathbf{w}$ is a window matrix defined below, $\widetilde{\mathbf{s}}_i$ is the ith interferer experiencing fading channel matrix $\widetilde{\mathbf{h}}_i$, and $\widetilde{\mathbf{z}}$ is AWGN, noting that $\widehat{\cdot}$ indicates windowing and $\widetilde{\cdot}$ indicates the change in dimensionality explained below. It is assumed that the power in each channel impulse response is zero after the T_mth tap and is time-invariant over measurement interval T. So each $\widetilde{\mathbf{h}}_i$ has the circulant structure

$$\widetilde{\mathbf{h}}_i =$$

$$
\left[
\begin{array}{c}
\underbrace{
\begin{array}{c}
\left._{T_m}\right\{ \\
\\
\\
\\
\left._{N-1}\right\{ \\
\\
\\
\\
\\
\end{array}
}
\begin{array}{ccccccc}
h_0 & 0 & \cdots & \cdots & \cdots & 0 & h_{T_m-1} & \cdots & h_1 \\
h_1 & h_0 & \ddots & & & & & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & & & & \ddots & h_{T_m-1} \\
h_{T_m-1} & \cdots & h_1 & h_0 & \ddots & & & & 0 \\
0 & h_{T_m-1} & & \ddots & \ddots & & \ddots & & \vdots \\
\vdots & \ddots & \ddots & & \ddots & h_0 & & \ddots & \vdots \\
\vdots & & \ddots & \ddots & & h_1 & & \ddots & \vdots \\
\vdots & & & \ddots & h_{T_m-1} & \vdots & \ddots & h_0 & 0 \\
0 & \cdots & \cdots & \cdots & 0 & h_{T_m-1} & \cdots & h_1 & h_0
\end{array}
\end{array}
\right],
\tag{A.2}
$$

and $\widetilde{\mathbf{s}}_i$ and $\widetilde{\mathbf{z}}$ are of dimension $N_E \times 1$ reflecting that, through convolution, $N_E \triangleq N + T_m - 1$ elements of $\widetilde{\mathbf{s}}_i$ contribute to $\widehat{\mathbf{x}}$, but noting that only N elements of $\widetilde{\mathbf{z}}$ contribute to $\widehat{\mathbf{x}}$ after windowing. Window matrix $\mathbf{w}$ is of dimension $N \times N_E$ and has the structure

$$
\mathbf{w} = \left[
\begin{array}{ccc|cccc}
0 & \cdots & 0 & w_0 & 0 & \cdots & 0 \\
\vdots & & \vdots & 0 & w_1 & & \vdots \\
\vdots & & \vdots & \vdots & & \ddots & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 & w_{N-1}
\end{array}
\right] \right\} N, \tag{A.3}
$$

where $\mathbf{w}_D \triangleq [w_0, \ldots, w_{N-1}]^T$ is a windowing vector which provides truncation to N elements in $\widehat{\mathbf{x}}$ preserving the causality of (A.1) and, with nonuniform element values, also may be used to produce the (optimal) Blackman-Tukey spectrum estimate [41, page 879].

Application of the $N \times N$ DFT matrix $\mathbf{D}_N$ to $\widehat{\mathbf{x}}$ produces

$$\widehat{\mathbf{X}} = \mathbf{D}_N \widehat{\mathbf{x}} = \mathbf{W}\left(\sum_{i=1}^{I} \widetilde{\mathbf{H}}_i \widetilde{\mathbf{S}}_i + \widetilde{\mathbf{Z}}\right), \tag{A.4}$$

where $\mathbf{W} = \mathbf{D}_N \mathbf{w} \mathbf{D}_{N_E}^H$ is a rectangular matrix producing convolution from $\mathbf{w}$, $\widetilde{\mathbf{H}}_i = \mathbf{D}_{N_E} \widetilde{\mathbf{h}}_i \mathbf{D}_{N_E}^H$ is the diagonal channel transfer function matrix for the ith interferer, $\widetilde{\mathbf{S}}_i$ is the DFT of $\widetilde{\mathbf{s}}_i$, and $\widetilde{\mathbf{Z}}$ is the DFT of AWGN vector $\widetilde{\mathbf{z}}$. The effect of $\mathbf{W}$ is two-fold. Firstly, $\mathbf{W}$ introduces spectral leakage to $\widehat{\mathbf{X}}$ resulting from truncation, where the severity of this depends on N, T_m relative to N and the design of window $\mathbf{w}_D$. Where the sample vector length is large compared to the channel multipath spread ($N \gg T_m$) and $\mathbf{w}_D$ is an appropriately chosen window, $\widehat{\mathbf{X}}$ can be a good estimate of the sampled spectrum $X(n f_S/N)$ for $n \in [0, \ldots, N-1]$. Secondly, $\mathbf{W}$ resamples the signal spectrum from $N + T_m - 1$ frequency domain samples to N frequency domain samples through interpolation. Analyses in this paper assume length N vectors as it is not important that these were obtained through interpolation, provided that the spectrum estimate is of good quality. While production of a good quality spectrum estimate is important to this paper, this is not the principal topic. Interested readers are referred to [41, Chapter 12] for comprehensive treatment, including a discussion of factors which collectively constitute quality in spectrum estimation.

References

[1] J. Mitola, *Cognitive Radio—Model-Based Competence for Software Radios*, Licentiate thesis, KTH, Stockholm, Sweden, 1999.

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

[3] FCC, "Facilitating opportunities for flexible, efficient and reliable spectrum use employing cognitive radio technologies," *ET Docket No. 03-108*, December 2003.

[4] IEEE, "Wireless regional area networks (WRAN)," *802.22 Project Authorization Request*, July 2004.

[5] S. Liu, J. Shen, R. Zhang, Z. Zhang, and Y. Liu, "Information theoretic criterion-based spectrum sensing for cognitive radio," *IET Communications*, vol. 2, no. 6, pp. 753–762, 2008.

[6] Y.-C. Liang, Y. Zeng, E. Peh, and A. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[7] FCC, "In the matter of 'establishment of an interference metric to quantify and manage interference and to expand available unlicensed operation in certain fixed, mobile and satellite frequency ands'," Tech. Rep., Federal Communications Commission, 03-289, November 2003.

[8] Z. Chen and S. Nagaraj, "Entropy based spectrum sensing in cognitive radio," in *Proceedings of the 7th Annual Wireless Telecommunications Symposium (WTS '08)*, pp. 57–61, 2008.

[9] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Spectrum sensing in cognitive radios based on multiple cyclic frequencies," in *Proceedings of the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom '07)*, pp. 37–43, August 2007.

[10] P. Sutton, K. Nolan, and L. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 13–24, 2008.

[11] W. Gardner, "The spectral correlation theory of cyclostationary time-series," *Signal Processing*, vol. 11, no. 1, pp. 13–36, 1986.

[12] T. Kailath and H. V. Poor, "Detection of stochastic processes," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2230–2259, 1998.

[13] E. Chaumette and P. Larzabal, "Cramér-Rao bound conditioned by the energy detector," *IEEE Signal Processing Letters*, vol. 14, no. 7, pp. 477–480, 2007.

[14] Z. Ye, G. Memik, and J. Grosspietsch, "Energy detection using estimated noise variance for spectrum sensing in cognitive radio networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '08)*, pp. 711–716, 2008.

[15] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 21–24, 2007.

[16] Y. Kim, G. Zheng, S. Sohn, and J. Kim, "An alternative energy detection using sliding window for cognitive radio system," in *Proceedings of the 9th International Conference on Advanced Communication Technology (ICACT '08)*, vol. 1, pp. 481–485, February 2008.

[17] H. Sarvanko, M. Mustonen, A. Hekkala, A. Mämmelä, M. Matinmikko, and M. Katz, "Cooperative and noncooperative spectrum sensing techniques using Welch's periodogram in cognitive radios," in *Proceedings of the 1st International Workshop on Cognitive Radio and Advanced Spectrum Management (CogART '08)*, pp. 1–5, February 2008.

[18] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1801–1811, 2008.

[19] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Spatial-spectral joint detection for wideband spectrum sensing in cognitive radio networks," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 2793–2796, April 2008.

[20] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 88–102, 2008.

[21] A. Taherpour, Y. Norouzi, M. Nasiri-Kenari, A. Jamshidi, and Z. Zeinalpour-Yazdi, "Asymptotically optimum detection of primary user in cognitive radio networks," *IET Communications*, vol. 1, no. 6, pp. 1138–1145, 2007.

[22] T. Li, W. Mow, V. Lau, M. Siu, R. Cheng, and R. Murch, "Robust joint interference detection and decoding for OFDM-based cognitive radio systems with unknown interference," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 566–575, 2007.

[23] B. Shen, K. Kwak, L. Huang, and Z. Zhou, "Improved consecutive mean excision algorithm based spectrum sensing for dynamic spectrum access," in *Proceedings if the IEEE International Conference on Communications (ICC '08)*, pp. 513–517, May 2008.

[24] T. Yücek and H. Arslan, "Spectrum characterization for opportunistic cognitive radio systems," in *Proceedings of the IEEE Military Communications Conference (MILCOM '06)*, pp. 1–6, October 2006.

[25] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1998.

[26] E. Larsson and G. Regnoli, "Primary system detection for cognitive radio: does small-scale fading help?" *IEEE Communications Letters*, vol. 11, no. 10, pp. 799–801, 2007.

[27] A. Coulson, "Do wireless data signals exhibit spectral autocorrelation?" in *Proceedings of the Australian Communications Theory Workshop (AUSCTW '05)*, pp. 6–11, January-February 2008.

[28] J. Lehtomäki, J. Vartiainen, M. Juntti, and H. Saarnisaari, "CFAR outlier detection with forward methods," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4702–4706, 2007.

[29] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[30] C. Carlemalm, H. V. Poor, and A. Logothetis, "Suppression of multiple narrowband interferers in a spread-spectrum communication system," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 8, pp. 1365–1374, 2000.

[31] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 13, no. 4, pp. 238–242, 1967.

[32] C. Fragouli, N. Al-Dhahir, S. Diggavi, and W. Turin, "Pre-filtered space-time M-BCJR equalizer for frequency-selective channels," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 742–753, 2002.

[33] S. S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd edition, 1991.

[34] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[35] P. Stoica, Y. Selén, and J. Li, "On information criteria and the generalized likelihood ratio test of model order selection," *IEEE Signal Processing Letters*, vol. 11, no. 10, pp. 794–797, 2004.

[36] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[37] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics Volume 1: Distribution Theory*, Charles Griffin, 2nd edition, 1963.

[38] R. Dwyer, "Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals," *IEEE Journal of Oceanic Engineering*, vol. 9, no. 2, pp. 85–92, 1984.

[39] A. J. Coulson, A. G. Williamson, and R. G. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 494–502, 1998.

[40] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S.J. Shellhammer, and W. Caldwell, "IEEE 802.22: the first cognitive radio wireless regional area network standard," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 130–138, 2009.

[41] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principals, Algorithms and Applications*, McMillan, 2nd edition, 1992.