Cooperative Sequential Binary Hypothesis Testing Using Cyclostationary Features

K. Aditya Mohan
Dept. of ECE, National Institute of Technology
Surathkal 575025, Karnataka, INDIA
Email: adityamnk@gmail.com

Chandra R. Murthy
Dept. of ECE, Indian Institute of Science
Bangalore 560012, Karnataka, INDIA
Email: cmurthy@ece.iisc.ernet.in

***Abstract*—A cognitive radio should be capable of finding the best spectrum band for communication depending on primary transmissions, the ambient noise level and interference. The first step in achieving this goal is to sense the existence of primary and secondary transmitters in various channels. In addition to the problem of signal detection, there is a need to distinguish between different signals at very low SNR. In this paper, the spectral correlation function is used for hypothesis testing. The sufficient statistic for feature vector based detection in the presence of timing uncertainty is derived. Sequential detection is used to decrease the average number of samples required for testing. Theoretical expressions for the stopping time at low SNR are derived for the AWGN channel. In a fading environment, the performance is evaluated using an approximate expression for the distribution of spectral correlation function. Monte-Carlo simulations verify the accuracy of the theoretical expressions.**

***Index Terms*—Cognitive Radio, Detection and Estimation, Cyclostationarity, Sequential tests**

I. Introduction

A Cognitive Radio (CR) [1] is aware of its radio environment and chooses its transmission and reception parameters depending on the other users, their transmit powers, bandwidths, quality of service requirements, etc. Spectrum Sensing (SS), a topic that has been very well studied in recent years, involves detection of signals in a particular band. The decision statistics considered in this paper are cyclostationary features extracted from the received data. Spectral correlation makes it possible to reject noise and interference for signal detection and extraction. It is known that signal detection in low SNR environment is best done using cyclic features [2].

In a Sequential Detector (SD), one allows the number of samples used for detection to vary to achieve the required performance [3]. The average sample size for a given target detection performance depends on the distribution of the samples. In [4], a multiple cyclic frequency based Generalized Likelihood Ratio Test (GLRT) in a cooperative environment is discussed. In [5], [6], a single Spectral Correlation Function (SCF) value is used as the feature of interest for several variants of the Fixed Sample Size Detector (FSSD). [7] and [8] discuss the SD performance of a QAM and an OFDM signal, respectively, and compare their performance with the FSSD. A hidden Markov-model based signal classification technique using cyclic frequency profile as feature vector is discussed in [9]. Cyclostationarity based cooperative and distributed techniques of signal detection are presented in [10], [11]. In [12], sequential detection of cyclostationary signals is proposed. The distribution is derived empirically and a multiple sequential probability ratio test based technique of signal detection is proposed to improve performance. A comprehensive comparison of cyclostationarity, energy and matched-filter based detection is presented in [13].

The features of the SD discussed in this paper, in contrast to the existing works is discussed below:

- The use of multiple cyclostationary features in conjunction with SD is proposed, to significantly reduce the number of samples required to achieve the same performance as with an FSSD.
- The existing multiple cyclostationary feature vector based detection techniques are not amenable to theoretical analysis in the sequential setup. In this paper, by using SCF feature vectors in a cooperative sequential detection scenario, this paper derives expressions for stopping time and the probability of detection.
- The theory discussed in this paper is applicable to any narrowband cyclostationary signal. It uses the magnitude of the SCF as the feature of interest for either detection of signal *vs.* noise or detection between two signal types. The deflection coefficient [14] is used as a criterion to identify features.
- The feature vector used in this paper comprises of SCF values at different frequencies f and cyclic frequencies α. It is shown that in a cooperative scenario with timing uncertainty, the magnitude of SCF is a sufficient feature vector.
- Typically, the effect of fading on performance is evaluated as the probability of detection of a detector designed for an AWGN channel [6] and/or using a fusion rule [15]. In contrast, this paper derives an approximate distribution for the SCF in a fading channel, and uses it to design a detector that achieves near-optimal performance.

The sequential detector discussed in this paper is compared with the fixed sample size Maximum Likelihood Detector (MLD). The theoretical expressions are validated using Monte-Carlo simulations.

In the next section, an introduction to cyclostationarity is provided, and the system model is explained. In Sec.III, the design of SD is presented and in Sec.(IV), hypothesis testing using MLD is discussed. In Sec.V, simulation results are presented.

II. Background and System Model

A. Introduction to Cyclostationarity

1) Spectral Correlation Function: The cyclic autocorrelation function $R_x^\alpha(\tau)$ is defined as [16]

$$R_x^\alpha(\tau) \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t+\tau/2)x^*(t-\tau/2)e^{-i2\pi\alpha t}dt \quad (1)$$

A signal exhibits second-order periodicity when $R_x^\alpha(\tau)$ is nonzero for some nonzero frequency α, called the *cyclic frequency*. The Fourier transform of the cyclic autocorrelation function is called the *Spectral Correlation Function* (SCF):

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau)e^{-i2\pi f\tau}d\tau = S_{uv}(f) \quad (2)$$

Here, $u(t) = x(t)e^{-i\pi\alpha t}$ and $v(t) = x(t)e^{+i\pi\alpha t}$ are frequency shifted versions of $x(t)$, and $S_{uv}(f)$ is the cross spectral density of $u(t)$ and $v(t)$. $S_x^\alpha(f)$ is thus the cross spectral density of frequency shifted signals $u(t)$ and $v(t)$.

B. System Model

Initially, consider the scenario where one wishes to determine whether or not a primary signal is present. When the primary signal is absent, the received signal $x(t)$ is modeled as a complex AWGN process $w(t)$ with variance σ_w^2. When the primary signal is present, $x(t)$ is modeled as a pure signal $\hat{z}(t)$, corrupted by the independent AWGN $w(t)$. The transmitted signal $\hat{z}(t)$ is assumed unit power and narrow band, and $x(t)$ can be written as

$$x(t) = h\hat{z}(t - t_0) + w(t) \quad (3)$$

where t_0 is the unknown delay in sampling and h is the fade, assumed frequency flat and constant for a measurement duration of computation of the SCF. The cooperative sensing model used in this paper is the one adopted in [7], where each Secondary User (SU) computes the Log Likelihood Ratio (LLR) and sends it to the fusion center for detection. The computation of LLR is explained in Sec. III. Noise does not exhibit any cyclostationarity, and hence its SCF is $S_w^\alpha(f) = 0 \quad \forall\alpha \neq 0$. When the signal is present, the cyclic spectrum of $x(t)$ is

$$S_x^\alpha(f) = |h|^2 S_{\hat{z}}^\alpha(f) + S_w^\alpha(f) \quad (4)$$

where $S_{\hat{z}}^\alpha(f)$ is the cyclic spectrum of $\hat{z}(t)$. Thus, evaluating $S_x^\alpha(f)$ at an appropriately chosen $\alpha \neq 0$ helps separate the signal from the purely stationary AWGN.

The SCF is estimated as [17] $S_{x_{N'}}^{\alpha_i}(\hat{k}, m_i)_N =$

$$\frac{1}{PN'}\sum_{l=0}^{P-1} X_{N'}(\hat{k} + lN', m_i + \frac{\alpha_i}{2})X_{N'}^*(\hat{k} + lN', m_i - \frac{\alpha_i}{2}) \quad (5)$$

where $X_{N'}(\hat{k} + lN', m_i)$ is the FFT coefficient at frequency m_i computed in the discrete time window, $[\hat{k} + lN', \hat{k} + (l + 1)N']$, P is the number of FFT windows used for frequency smoothing, $\hat{k}$ is the discrete time sample, $m_i = f_iN'/f_s$ is the discrete frequency corresponding to the frequency f_i, α_i is the discrete analog of the cyclic frequency and is even.[1]

III. Sequential Binary Hypothesis Testing

The problem of interest in this paper is to classify between two different signals or a signal and noise based on the received signal. The feature vectors can be written as $Z_k = [z_{k,1}z_{k,2}\ldots z_{k,\nu}]$ where ν is the total number of features. Here $z_{k,i}$ corresponds to the i^{th} feature and k^{th} SU. Thus, the sequential detection is performed by accumulating mutually independent Log-Likelihood Ratios (LLRs) that could either come from different SUs, or from a single SU using samples that are more than the channel coherence time apart. The joint pdf of $z_{k,i}$ is $p(Z_k) = p(z_{k,1})p(z_{k,2})\ldots p(z_{k,\nu})$, assuming the features are independent. If $p_1(Z_k)$ and $p_2(Z_k)$ are the probability densities of Z_k under hypothesis H_1 and hypothesis H_2 respectively, then the Log-Likelihood Ratio (LLR) is

$$l_k = \log\left(\frac{p_2(Z_k)}{p_1(Z_k)}\right). \quad (6)$$

[1]The notation used here is borrowed from [6] [17].

In terms of the LLRs, the SD scheme can be written as [3]

$$\sum_{k=1}^{K} l_k > a \rightarrow H_2, \quad \sum_{k=1}^{K} l_k < b \rightarrow H_1 \quad (7)$$

Otherwise, the test continues by taking the next LLR and adding it to the previous sum. This test is called the Sequential Probability Ratio Test (SPRT). The $k = K$ at which a decision is made is the converging sample size and the thresholds a and b are set as

$$a = \log\left(\frac{1-\beta}{\alpha}\right) \quad \text{and} \quad b = \log\left(\frac{\beta}{1-\alpha}\right), \quad (8)$$

where α and β are the probability of miss detection under H_1 and H_2, respectively, which are our design targets.

The Expected Converging Sample Size (ECSS) of the binary classifier is [3]

$$\begin{aligned} E(K|H_2) &\approx \frac{1}{\gamma_2}\left[\beta \ln\frac{\beta}{1-\alpha} + (1-\beta)\ln\frac{1-\beta}{\alpha}\right] \\ E(K|H_1) &\approx \frac{1}{\gamma_1}\left[(1-\alpha)\ln\frac{\beta}{1-\alpha} + \alpha\ln\frac{1-\beta}{\alpha}\right] \end{aligned} \quad (9)$$

where $\gamma_1 = E(l_k|H_1)$ and $\gamma_2 = E(l_k|H_2)$. For small α and β, $\alpha \approx e^{-\gamma_2 E(K|H_2)}$ and $\beta \approx e^{\gamma_1 E(K|H_1)}$. It is shown in the Appendix B that a sufficient statistic for the present problem is the magnitudes of SCF values at different cyclic frequencies, i.e., the feature vector is

$$Z_k = \left[\left|S_{x_{N'}}^{\alpha_1}(k, m_1)\right|, \ldots, \left|S_{x_{N'}}^{\alpha_\nu}(k, m_\nu)\right|\right]$$

The time domain samples of a cyclostationary signal are not i.i.d., but the SCFs at different frequencies f and cyclic frequencies α are independent. This independence follows from the independence of FFT values at different frequency bins. However, the necessary frequency and cyclic frequency resolution need to be ensured as described in [17]. The SCF values of the pure signal is assumed to be known at the receiver.

A. The AWGN Channel

In an AWGN channel, the channel gain h is deterministic and known. It is shown in the Appendix A that the pdf of $z_{k,i}$ under hypothesis $H_j, j = 1, 2$ is $z_{k,i} \sim R(v_{ji}, \sigma_{ji}^2)$ where R stands for the Rician distribution, i.e.,

$$p(z_{k,i}) = \frac{z_{k,i}}{\sigma_{ji}^2} e^{-\frac{z_{k,i}^2 + v_{ji}^2}{2\sigma_{ji}^2}} I_0\left(\frac{z_{k,i}v_{ji}}{\sigma_{ji}^2}\right), \quad (10)$$

where $v_{ji} \triangleq |h|^2\mu_{ji}$, I_0 is the modified Bessel function of first kind of order 0, $\mu_{ji} = |S_{\hat{z}_{jN'}}^{\alpha_i}(\hat{k}, m_i)_N|$ and σ_{ji}^2 is given by

$$\sigma_{ji}^2 = \frac{\sigma_w^4}{2P}\left(1 + \frac{|h|^2 S_{\hat{z}_{jN'}}(m_i + \frac{\alpha_i}{2}) + |h|^2 S_{\hat{z}_{jN'}}(m_i - \frac{\alpha_i}{2})}{\sigma_w^2}\right).$$

The noise variance σ_w^2 is assumed to be known at the receiver. Thus, the LLR for the hypothesis test is

$$\begin{aligned} \log(\lambda(Z_k)) &= \sum_{i=1}^{\nu}\left(2\log\left(\frac{\sigma_{1i}}{\sigma_{2i}}\right) - \frac{z_{k,i}^2 + v_{2i}^2}{2\sigma_{2i}^2} + \frac{z_{k,i}^2 + v_{1i}^2}{2\sigma_{1i}^2}\right), \\ &+ \sum_{i=1}^{\nu}\left(\log\left(I_0\left(\frac{z_{k,i}v_{2i}}{\sigma_{2i}^2}\right)\right) - \log\left(I_0\left(\frac{z_{k,i}v_{1i}}{\sigma_{1i}^2}\right)\right)\right) \end{aligned} \quad (11)$$

where $v_{ji} \triangleq |h|^2\mu_{ji}$. The expected value of (11) under either hypothesis is derived in Appendix B as:

$$\gamma_j = \sum_{i=1}^{\nu} \left(2\log\left(\frac{\sigma_{1i}}{\sigma_{2i}}\right) - \frac{2\sigma_{ji}^2 + v_{ji}^2 + v_{2i}^2}{2\sigma_{2i}^2} + \frac{2\sigma_{ji}^2 + v_{ji}^2 + v_{1i}^2}{2\sigma_{1i}^2} \right)$$
$$+ \sum_{i=1}^{\nu} \left(\frac{(2\sigma_{ji}^2 + v_{ji}^2)v_{2i}^2}{4\sigma_{2i}^4} - \frac{(2\sigma_{ji}^2 + v_{ji}^2)v_{1i}^2}{4\sigma_{1i}^4} \right) \quad (12)$$

From (12), it can be shown that $\gamma_2 > 0$ and $\gamma_1 < 0$, which is necessary for the SPRT to work.

At low SNR, $\sigma_{1i}^2 \approx \sigma_{2i}^2 \propto \sigma_w^4/P$ and hence $\gamma_j \propto P/\sigma_w^4$. Thus, for sufficiently large K, $\alpha \approx (C_1)^{-PE(K|H_2)}$ and $\beta \approx (C_2)^{-PE(K|H_1)}$ where C_1 and C_2 are constants. That is, the performance of SD depends on the expected number of FFTs, i.e., $PE(K|H_j)$, enabling one to choose P depending on α and β to meet the performance requirement. In an AWGN channel, therefore, one can trade-off P with K without losing performance. Thus, the right choice of P and $E(K|H_i)$ is dependent on the traffic cost (due to transmission of LLRs to fusion center), detection time constraint and the number of SUs.

B. Rayleigh Fading Channel

In a Rayleigh fading channel, h is a random variable whose magnitude is Rayleigh distributed. In Appendix A, it is shown that the pdf of the magnitude of SCF in Rayleigh fading channel is

$$p(z_i) = \frac{z_i}{2\mu_i\sigma_i\sigma_f^2}\exp\left(\frac{-z_i^2}{2\sigma_i^2}\right)\exp\left(\frac{\sigma_i^2}{16\mu_i^2\sigma_f^4}\right)$$
$$\sum_{l=0}^{\infty}\left(\frac{z_i^2}{4\sigma_i^2}\right)^l \frac{(2l)!}{(l!)^2} D_{-(2l+1)}\left(\frac{\sigma_i}{2\mu_i\sigma_f^2}\right) \quad (13)$$

where σ_f^2 is the variance of fade and $D_k(.)$ is the parabolic cylinder function. In the above, the hypothesis index j and the SU index k have been dropped for notational simplicity. Note that (13) is an approximation which is valid for $\mu_i \neq 0$. If $\mu_i = 0$, then fading has no effect on the SCF and its distribution will be same as that in the case of AWGN. Further, $D_l(.)$ falls rapidly with l and at low SNRs the first few terms are sufficient for a good approximation. At low values of μ_i, the distribution is well approximated by (10), and hence fading can be neglected. The expected value of LLR can be estimated empirically and using (9), the expected sample size can be evaluated.

The next step is the selection of the frequency and cyclic frequency at which to compute the SCFs for detection. The deflection coefficient [14] is a simple, convenient metric to pick features that best separate the hypotheses. With this, the design procedure for the SD is summarized below:

- For the given pair of signals, select n cyclic features with the maximum deflection coefficient
$$\frac{(E(z_{k,i}|H_2) - E(z_{k,i}|H_1))^2}{(Var(z_{k,i}|H_2) + Var(z_{k,i}|H_1))}$$
.
- Compute the LLR as given by (6) and add it to the previously computed cumulative sum of LLRs. In an AWGN channel, the LLR simplifies to (11).
- Compare the cumulative sum to the thresholds (8).
- The detector stops when the cumulative sum of LLRs cross the threshold, and the decision rule is (7).

IV. Maximum Likelihood Detector (MLD)

The hypothesis test of FSSD (here, also called the MLD) in terms of LLR is

$$\sum_{k=1}^{K} l_k \geq \tau \rightarrow H_2, \quad \sum_{k=1}^{K} l_k < \tau \rightarrow H_1 \quad (14)$$

where all the terms are as explained in Sec. III. The (fixed) sample size K and threshold τ are chosen suitably to satisfy the required detection performance under both hypotheses. As in the case of SD, the independent LLRs $l_k = \log(p_2(Z_k)/p_1(Z_k))$ are obtained from different SUs and accumulated in the fusion center.

V. Simulation Results

In this section, Monte-Carlo simulation results are presented to illustrate the performance of the proposed detectors and corroborate the theoretical analysis. For all simulations, $P = N/N' = 8$,where N' is the FFT size and N is the total number of samples. N' is chosen to be 32, to provide the frequency resolution required to distinguish the cyclic spectral peaks of the signals under consideration. All signals used in simulation are normalized to unit power. The required SNR is obtained by scaling the noise variance. Note that the binary detectors designed in this paper can be used either to distinguish between two different signal types or for detecting signal *vs.* noise; examples of both applications are illustrated here.

The cyclic spectrum of all the signals used in this paper is derived in [18]. However, in this work, we use pulse shaping using raised cosine filter for Binary-Phase Shift Keying (BPSK) and Minimum Phase Shift Keying (MSK) signals. The performance of the BPSK vs. MSK SD, designed for a target probability of detection of $P_d(H_1) = P_d(H_2) = 0.9$ is shown in Table I. The SD employed here uses two feature vectors. Note that more than two feature vectors can be chosen but it would increase the computational complexity without a significant increase in performance. $E_{th}(K)$ is the theoretical value of the ECSS and $E_{pr}(K)$ is the empirical estimate. We see that the theoretical and experimental values of the ECSS match, and the target detection performance is nearly satisfied. The expected number of *FFT samples* can be computed as $PE(K)$, with $P = 8$. For comparison with the FSS MLD, the SD is designed to distinguish between H_2:BPSK and H_1:noise. It uses a single SCF as its feature vector. Note that, noise as the input signal to the detector at different values of SNR is modeled by scaling the variance of the noise relative to a hypothetical signal at unit power. The distribution of SCF of the BPSK signal in fading is shown in Fig. 1, and the theoretical and empirical distribution are found to match well. Figure 2 plots the number of FFT samples required to achieve the target performance in presence of hypothesis H_2. The graph for H_1 is similar and is omitted. The SNR is fixed at -20dB. In fading environment, since the SNR is random, its expected value, $10\log\left(2\sigma_f^2/\sigma_w^2\right)$, is fixed at

H_2: MSK vs. H_1: BPSK					
	AWGN Channel			Fading Channel	
	P_d	$E_{th}(K)$	$E_{pr}(K)$	P_d	E(K)
H_2	0.88	45.68	44.54	0.85	55.30
H_1	0.88	46.63	46.87	0.85	67.13

TABLE I
MSK vs. BPSK SD at -20dB SNR

-20dB. The superior performance in fade when compared to that in AWGN in Fig. 2 is due to the diversity advantage in combining independent LLRs. Moreover, the LLR grows very quickly with fade value, and hence the average LLR with independently faded SCFs is much higher than that with AWGN, leading to a faster detection time. In Fig. 2, for the SD, K, the number of LLRs, is a random variable. MLD1 is the MLD with $P = 8$ and K chosen to equal the ECSS of the SD, and MLD2 is the MLD with P equal to the ECSS of the SD multiplied by 8 and and $K = 1$. The FFT samples in the y-axis is the value $PE(K)$. The superior performance of the SD is clear from the graph. To distinguish between two signals in AWGN, a knowledge of channel gain is required, since the true value of SCF is a function of the channel gain. However, in a fading channel, only the mean channel gain needs to be known at the detector.

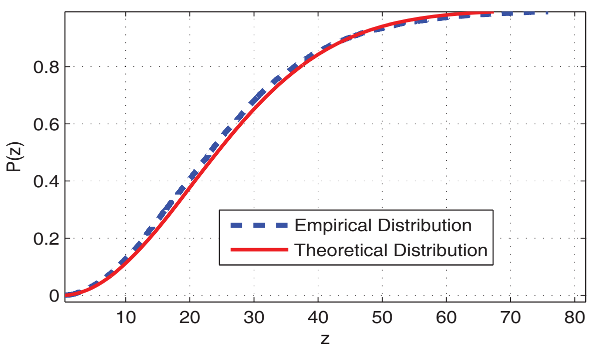

Fig. 1. Distribution of SCF of BPSK in a fading channel.

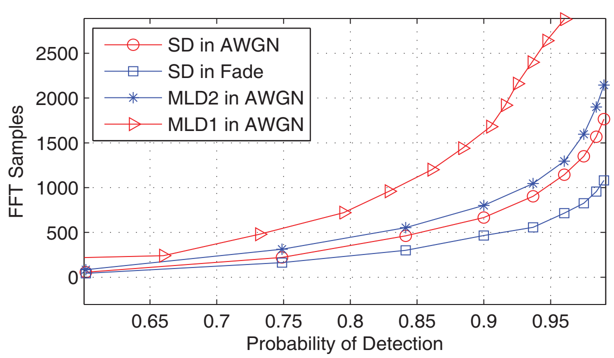

Fig. 2. FFT sample size i.e., $PE(K)$ for different detectors

In summary, this paper investigated binary sequential detection using cyclostationary properties of the signals, and presented a design methodology that can be used either for classifying a signal versus noise or between different signal types. The SD discussed in this paper allows one to flexibly trade-off between the number of samples i.e., P at each SU and K, the number of SUs (or independent SCF values), to achieve a desired performance. The MLD was found to require a significantly larger number of samples compared to SD, to achieve the same performance. Detection using independent SCF values provides spatial diversity and parallel accumulation of LLRs, which can be subsequently processed in the fusion center, leading to significant savings in detection time. Future work can consider the use of overlapping samples in computation of SCF. A possible solution to the problem of overlapped processing is to use a correction factor ϵ to the variance term in (10), as discussed in [6]. In this case, the optimum SD will involve an optimization over ϵ, P and K.

Appendix A
Derivation of the Distribution of the SCF

In (5), by property of Fourier transform, the FFT coefficients $X_{N'}(\hat{k}+lN', m_i+\frac{\alpha_i}{2})$ and $X_{N'}(\hat{k}+lN', m_i-\frac{\alpha_i}{2})$ are independent if $|m_i+\frac{\alpha_i}{2}| \neq |m_i-\frac{\alpha_i}{2}|$. Thus, if $m_i = 0$, the two frequency coefficients will be dependent and $X_{N'}(-\alpha_i/2) = X^*_{N'}(\alpha_i/2)$. Since the time domain signal at the receiver can be written as in (3), $X_{N'}(\hat{k}+lN', m_i+\frac{\alpha_i}{2}) = hZ(m_i+\alpha_i/2)+W(m_i+\alpha_i/2)+iY(m_i+\alpha_i/2)$ and $X^*_{N'}(\hat{k}+lN', m_i-\frac{\alpha_i}{2}) = h^*Z^*(m_i-\alpha_i/2)+W^*(m_i-\alpha_i/2)-iY^*(m_i-\alpha_i/2)$. Here, $Z(m)$ and $W(m)/Y(m)$ are the FFTs of the raw signal and real/imaginary parts of (time-domain) noise, respectively, and are computed in the time window $[\hat{k}+lN', \hat{k}+(l+1)N']$. If the noise $w(t)$ is zero mean and has a variance σ_w^2 then $E(|W(m)|^2) = \sigma_w^2/2$, $E(W^2(m)) = 0$, $E(|W(m)|^2 W(m)) = 0$ and $E(|W(m)|^4) = \sigma_w^4/2$, and the same results hold for $Y(m)$. Further, $Z(m_i+\alpha_i/2)Z^*(m_i-\alpha_i/2) = S^{\alpha_i}_{\hat{z}_{N'}}(\hat{k}, m_i)_N$. Using these properties, it can be shown that $E(X_{N'}(\hat{k}+lN', m_i+\frac{\alpha_i}{2})X^*_{N'}(\hat{k}+lN', m_i-\frac{\alpha_i}{2})) = |h|^2 S^{\alpha_i}_{\hat{z}_{N'}}(\hat{k}, m_i)_N$. Similarly, the variance of the product can also be derived. Using the central limit theorem, it can be shown that the SCF has a complex Gaussian distribution with mean $|h|^2 S^{\alpha_i}_{\hat{z}_{N'}}(\hat{k}, m_i)_N$ and variance of $2\sigma_i^2$, where

$$\sigma_i^2 = \frac{\sigma_w^4}{2P}\left(1+\frac{|h|^2\left(S_{\hat{z}_{N'}}(m_i+\frac{\alpha_i}{2})+S_{\hat{z}_{N'}}(m_i-\frac{\alpha_i}{2})\right)}{\sigma_w^2}\right) \quad (15)$$

$S_{\hat{z}_{N'}}(m_i)$ is the power spectral density of the pure signal evaluated at discrete frequency m_i. In an AWGN channel, the gain h is deterministic. In a fading channel, the distribution of SCF is a function of the random variable h. We assume $|h|$ to be Rayleigh distributed with parameter σ_f^2. A similar method is adopted in [6] where they derive the distribution of the SCF for a multi-antenna fixed sample size technique. At low SNR, the variance σ_i^2 can be approximated by $\sigma_w^4/2P$, i.e., it is independent of the fade h. This approximation is used to design the (near-optimal) SD in this paper.

The phase of the SCF is a random variable which is dependent on the time of sampling. In [16], it has been shown any time delay in sampling translates to an equivalent phase rotation of the SCF. Assuming a uniform distribution for the phase rotation, it has been shown in the next subsection that the magnitude of SCF is a sufficient feature vector.

Since the SCF is Gaussian distributed, its magnitude is Rician distributed with parameters $v_i = |h|^2|S^{\alpha_i}_{\hat{z}_{N'}}(\hat{k}, m_i)_N| = |h|^2\mu_i$ and σ_i^2 as given by (15). Since $|h|$ (denoted by r below) is a random variable the distribution of $z_i = |S^{\alpha_i}_{x_{N'}}(\hat{k}, m_i)_N|$ is given by

$$p(z_i) = \int_{r=0}^{\infty}\frac{z_i}{\sigma_i^2}e^{-\frac{z_i^2+r^4\mu_i^2}{2\sigma_i^2}} I_0\left(\frac{r^2\mu_i z_i}{\sigma_i^2}\right)\frac{r}{\sigma_f^2}e^{-r^2/(2\sigma_f^2)}dr$$

where $I_0(.)$ is the modified Bessel function of first kind of order 0. Replacing the Bessel function by its infinite series expansion we get

$$p(z_i) = \frac{A}{\sigma_f^2}\sum_{l=0}^{\infty}\frac{c^{2l}}{4^l(l!)^2}\int_0^\infty r^{4l+1}e^{-br^4-r^2/(2\sigma_f^2)}dr$$

where $A = \frac{z_i}{\sigma_i^2}e^{-z_i^2/(2\sigma_i^2)}$, $b = \frac{\mu_i^2}{2\sigma_i^2}$ and $c = \frac{\mu_i z_i}{\sigma_i^2}$. Thus,

$$p(z_i) = \frac{A}{2\sigma_f^2}\sum_{l=0}^{\infty}\frac{c^{2l}}{4^l(l!)^2}\int_0^\infty r^{2l}e^{-br^2-r/(2\sigma_f^2)}dr$$

The above definite integral has a solution in terms of the parabolic cylinder function $D_\nu(x)$,[19]. After making suitable substitutions the probability density function of the SCF can be shown to be

$$\begin{aligned} p(z_i) &= \frac{z_i}{2\mu_i\sigma_i\sigma_f^2}\exp\left(\frac{-z_i^2}{2\sigma_i^2}\right)\exp\left(\frac{\sigma_i^2}{16\mu_i^2\sigma_f^4}\right) \\ & \sum_{l=0}^{\infty}\left(\frac{z_i^2}{4\sigma_i^2}\right)^l\frac{(2l)!}{(l!)^2}D_{-(2l+1)}\left(\frac{\sigma_i}{2\mu_i\sigma_f^2}\right) \end{aligned} \quad (16)$$

APPENDIX B
SUFFICIENCY OF THE FEATURE VECTOR IN AN AWGN CHANNEL

In the analysis that follows, the channel considered will be an AWGN channel. The extension to fading channel is straightforward and is omitted due to lack of space. Let $\mu_r = Re(S^{\alpha}_{\hat{z}_{N'}}(k,m)_N)$, $\mu_i = Im(S^{\alpha}_{\hat{z}_{N'}}(k,m)_N)$, $x = Re(S^{\alpha}_{x_{N'}}(k,m)_N)$ and $y = Im(S^{\alpha}_{x_{N'}}(k,m)_N)$ then $\mu = \sqrt{\mu_r^2+\mu_i^2}$ where the subscript related to feature vector is dropped. Then $x+iy \sim CN(\mu_r + i\mu_i, 2\sigma^2)$ as derived in Appendix A.

Assuming a sampling delay of t_0 at the receiver, the pure SCF will be given by [16] $|S^{\alpha}_{\hat{z}_{N'}}(k,m)_N|e^{j(\phi-2\pi\alpha t_0)} = \mu e^{j\theta}$. The time delay is unknown to the receiver, and as sensing can start at any instant of time at the different SUs, the phase of SCF is uniformly distributed on $[0,2\pi)$. Let $\theta = \phi - 2\pi\alpha t_0$. The effective distribution of the SCF is given by the integral

$$\begin{aligned} p(x,y) &= \int_0^{2\pi}\frac{1}{(2\pi)^2\sigma^2}\exp\left(-\frac{\left|x+iy-\mu e^{i\theta}\right|^2}{2\sigma^2}\right)d\theta \\ p(x,y) &= A\int_0^{2\pi}\exp\left(b\cos(\theta)+c\sin(\theta)\right)d\theta \\ &= A\int_0^{2\pi}\sum_{k=0}^{\infty}\frac{(b\cos(\theta)+c\sin(\theta))^k}{k!}d\theta \end{aligned} \quad (17)$$

$$\text{where} \qquad A = \frac{1}{(2\pi)^2\sigma^2}\exp\left(-\frac{x^2+y^2+\mu^2}{2\sigma^2}\right),$$

$b = \mu x/\sigma^2$, and $c = \mu y/\sigma^2$. The integral above has a closed form solution as given by (3.661(1),3.661(2)) in [19]. Further, using the relation $(2k)!!(2k-1)!! = (2k)!$ and $(2k)!! = 2^k k!$ we can show that

$$\begin{aligned} p(x,y) &= 2\pi A\sum_{k=0}^{\infty}\frac{(2k-1)!!}{(2k)!(2k)!!}(b^2+c^2)^k \\ &= \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2+\mu^2}{2\sigma^2}\right)I_0\left(\frac{\mu}{\sigma^2}\sqrt{x^2+y^2}\right) \end{aligned}$$

The expected value of any function of x^2+y^2 is given by

$$E(f(x^2+y^2)) = \int_0^{\infty}\int_0^{\infty}f(x^2+y^2)p(x,y)dxdy$$

By translating the integral to polar coordinates with $z = \sqrt{x^2+y^2}$, we get

$$E(f(z^2)) = \int_0^{\infty}f(z^2)\frac{z}{\sigma^2}\exp\left(-\frac{z^2+\mu^2}{2\sigma^2}\right)I_0\left(\frac{\mu z}{\sigma^2}\right)dz$$

Thus, since the SCF is complex Gaussian with independent real and imaginary parts, the magnitude of SCF, z will be Rician distributed. The correlation coefficient of the real and imaginary parts of SCF can be proved to be equal to zero. The above results show that taking the magnitude of the SCF is the same as evaluating the LLR using the SCF with unknown timing (phase), and averaging over the pdf of the unknown phase angle. The stopping time of the SD is a function of the expected value of the LLR. Therefore, choosing the magnitude of the SCF as the feature of interest results in no loss of performance.

The next step is to find a theoretical approximation for the expected value of the LLR. Using a Taylor series expansion, it can be shown that $\log(I_0(x)) \approx x^2/4$. At low SNR, the value of v_{ji}/σ_{ji} in (11) will be small, as it is directly proportional to the SNR. Further, $E(z_{ji}^2) = 2\sigma_{ji}^2 + v_{ji}^2$ [20]. Thus, the expected value of LLR under either hypothesis can be derived to be (12).

REFERENCES

[1] J. Mitola-III Jr, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology, Sweden, May 2000.

[2] W. A. Gardner and C. M. Spooner, "Detection and source location of weak cyclostationary signals: simplifications of the maximum-likelihood receiver," *IEEE Transactions on Communications*, vol. 41, no. 6, June 1993

[3] H. V. Poor, "An introduction to signal detection and estimation," Springer, 2nd Ed., 1994.

[4] J. Lunden, V. Koivunen, A. Huttunen, and H. V. Poor, "Spectrum sensing in cognitive radios based on multiple cyclic frequencies," *Proc. of Int. Conf. on Cognitive Radio Oriented Wireless Networks and Commn.*, pp. 37–43, Aug. 2007.

[5] H.-S. Chen, W. Gao, and D. G. Daut, "Spectrum sensing using cyclostationary properties and application to IEEE 802.22 WRAN," *Proc. of Global Telecommunications Conference*, pp. 3133–3138, Nov. 2007.

[6] X. Chen, W. Xu, Z. He, and X. Tao, "Spectral correlation based multi-antenna spectrum sensing technique," *IEEE Wireless Communication and Networking Conference*, 2008.

[7] S. Chaudhari, V. Koivunen and H. V. Poor, "Distributed autocorrelation-based sequential detection of OFDM signals in cognitive radios," *Cognitive Radio Oriented Wireless Networks and Commn.*, 2008.

[8] Y. C. Lin, C. C. J. Ibob, "Classification of quadrature amplitude modulated (QAM) signals via sequential probability ratio test (SPRT)," *Elsevier* (1997) Signal Processing, 60 (3), pp. 263-280.

[9] K. Kim, I. A. Akbar, K. K. Bae, J.-S. Um, C. M. Spooner, and J. H. Reed, "Cyclostationary approaches to signal detection and classification in cognitive radio," *Proc. of IEEE Int. Symp. on New Frontiers in Dynamic Spectrum Access Networks*, pp. 212–215, Apr. 2007.

[10] H. Sadeghi and P. Azmi, "Cyclostationarity-based cooperative spectrum sensing for cognitive radio networks," *Proc. of Internatioal Symposium on Telecommunications*, pp. 429–434, Aug. 2008.

[11] C. R. C. M. da Silva, B. Choi, and K. Kim, "Distributed spectrum sensing for cognitive radio systems," *Proc. of Info. Theory and Appl. Workshop*, pp. 120–123, Feb. 2007.

[12] K. W. Choi, W. S. Jeon and D. G. Jeong, "Sequential Detection of Cyclostationary Signal for Cognitive Radio Systems," *IEEE transactions on wireless communications*, vol. 8, no. 9, Sept. 2009

[13] B. Deepa, "Design and analysis of cyclostationarity based spectrum sensing for cognitive radio," *M. E. Thesis, Indian Institute of science, India*, May 2009.

[14] S. M. Kay, "Fundamentals of statistical signal processing: detection theory," *Prentice Hall Signal Processing Series*

[15] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2005

[16] W. A. Gardner, "Exploitaion of spectral redundancy in cyclostationary signals," *IEEE Sig. Proc. Magazine*, pp. 14–35, Apr. 1991.

[17] R. S. Roberts, W. A. Brown, and H. H. Loomis, "Computationally efficient algorithms for cyclic spectral analysis," IEEE Sig. Proc. Magazine, pp. 3848, Apr. 1991

[18] W. A. Gardner, W. A. Brown and C. K. Chen, "Spectral Correlation of Modulated Signals: Part II-Digital Modulation," *IEEE Transactions on Communications* vol. com-35. no.6, June 1987

[19] I. H. Gradshteyn and I. M. Ryzhik, "Table of integrals series and products," 7th ed., *Academic Press*

[20] A. Papoulis and S. U. Pillai, "Probability, random variables and stochastic Processes," 4th ed., *TATA McGraw-Hill*