LEARNING BASED SPECTRUM SENSING IN OFDM COGNITIVE RADIOS

by

Muhammad Umair Muzaffar

A Thesis Presented to the Faculty of the
American University of Sharjah
College of Engineering
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in
Electrical Engineering

Sharjah, United Arab Emirates

May 2012

Approval Signatures

We, the undersigned, approve the Master's Thesis of Muhammad Umair Muzaffar.

Thesis Title: Learning Based Spectrum Sensing in OFDM Cognitive Radios

Signature	**Date of Signature**
___________________________ Dr. Mohamed El-Tarhuni Associate Professor Department of Electrical Engineering Thesis Advisor	_______________
___________________________ Dr. Khaled Assaleh Professor Department of Electrical Engineering Thesis Co-Advisor	_______________
___________________________ Dr. Mohamed Hassan Associate Professor Department of Electrical Engineering Thesis Committee Member	_______________
___________________________ Dr. Arafat Al-Dweik Associate Professor, Communication Engineering Program ECE Department, Khalifa University (Sharjah Campus) Thesis Committee Member	_______________
___________________________ Dr. Mohamed El-Tarhuni Head, Department of Electrical Engineering	_______________
___________________________ Dr. Hany El Kadi Associate Dean, College of Engineering	_______________
___________________________ Dr. Yousef Al-Assaf Dean of College of Engineering	_______________
___________________________ Dr. Khaled Assaleh Director of Graduate Studies	_______________

Acknowledgements

All praise and thanks is due to Allah, the Almighty, for providing me with the strength and determination to complete this thesis and infinite peace and blessings of Allah be upon the beloved Prophet Muhammad. In addition, I gratefully acknowledge my advisors Dr. Mohamed El-Tarhuni and Dr. Khaled Assaleh for their continuous help and support throughout the course of this research. Under their special guidance, I was able to grasp the most difficult concepts of this thesis easily. They were quick to help me whenever I struggled with a topic and prompt in providing different solutions to an ongoing problem. I feel that I was really lucky to have such dedicated and inspiring advisors. Finally, I am also extremely grateful to my parents, brother and sister for their unwavering support through both the easy and tough times. It was their firm faith in me that inspired me to continue working harder and complete the thesis.

To my beloved parents, brother and sister
for their evergreen love and faith in me

Abstract

Cognitive Radio (CR) is an innovative technology introduced to efficiently utilize the spectrum. It allows secondary users to access the licensed portion of the spectrum when it is not occupied by the primary licensed user. Most of the CR applications are expected to operate in channels occupied by OFDM based systems since OFDM is the preferred modulation scheme of most recent wireless technologies. To be able to efficiently utilize the spectrum, CR must be able to properly sense the spectrum. This thesis models the spectrum sensing problem in a Cooperative CR system as a two class pattern recognition problem: signal present or signal absent. The signals from both classes have different characteristics which are learned by a linear classifier during the training phase. Once fully trained, the classifier utilizes this learning to classify any unseen data into one of the classes. The characteristics which differentiate the signals from both classes are called features and are acquired by the linear classifier through the process of feature extraction. In a cooperative CR network, each CR extracts features from its received signal and sends it to a fusion center. At the fusion center, a universal decision is made on spectrum occupancy based on features received from all the CRs. In this thesis, energy, correlation and entropy are used as features to distinguish between the primary OFDM signal and noise. The performance of the spectrum sensing schemes is evaluated in terms of the detection and false alarm probabilities. It is shown that energy and correlation detectors outperform the entropy detector in AWGN channels. However, in a fading channel, the correlation detector outperforms both the energy and entropy detectors due to the degradation of their performance caused by the deep fades in the channel. The performance can be improved by increasing the observation window size and by changing the number of users in the Cooperative CR network.

Search Terms: Cognitive Radios, Learning, Pattern Recognition, Linear Classifier, OFDM, Energy, Correlation, Entropy.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The increasing demand to access the electromagnetic radio spectrum has resulted in congestion and shortage of the natural spectrum resource. Traditionally, a frequency band is allocated to each user by a regulatory authority and the user is given complete access to that frequency band at all times. However, statistical information indicates that not all primary users utilize their licensed frequency spectrum at all times of the day. In fact, the spectrum remains unused most of the time and therefore results in an inefficient utilization of this precious resource. Cognitive Radio (CR) has emerged as a promising technology to opportunistically exploit the unused parts of the licensed electromagnetic spectrum by frequently sensing the spectrum and transmitting only when the licensed primary user is not using the spectrum.

In CR technology, the unlicensed secondary users are allowed to use a licensed part of the spectrum only when the primary licensed user is not accessing it. However, the primary users still have the legal rights to access the spectrum whenever they wish and the CR has to stop transmitting when the primary user starts to transmit. To accomplish this task and to have minimum interference between the primary and secondary users, the CR should be able to properly sense the spectrum which is one of the most important tasks performed by the CR transmitter. Various techniques have been developed to provide better spectrum sensing for different communication technologies. Nevertheless, spectrum sensing has several challenges to be overcome. The most important among these challenges is to avoid interference with the primary user since it has the legal right to use the spectrum at all times. The CR transmitter should vacate the spectrum immediately when it senses the presence of a primary signal.

Several recent wireless communication technologies such as Wi-Fi, WiMax and LTE employ Orthogonal Frequency Division Multiplexing (OFDM). OFDM is a multi-carrier system that is preferred over traditional single-carrier systems for wireless communications because it is not very sensitive to multipath channels that cause Intersymbol Interference (ISI) and, therefore, does not require complex equalization at the receiver. Consequently, most CRs operating in wireless channels are expected to deal with OFDM based primary users and therefore it is imperative to investigate different spectrum sensing techniques for OFDM based CRs.

In this thesis, spectrum sensing in OFDM Cooperative CR network is considered as a pattern recognition problem and different techniques for spectrum sensing are investigated. The signal received at the CRs can belong to two classes: the licensed user signal present or noise only. Different features are extracted from the received signal by each CR and sent to a central node which uses a linear classifier to make a decision on spectrum vacancy. A single CR system is also considered where the CR makes an individual decision on the existence of the primary user signal. A linear classifier is used because the features used in this work effectively separate the data belonging to both classes linearly.

The objectives of this thesis are to:

- Model the CR spectrum sensing problem as a pattern recognition system and identify appropriate features.
- Design a spectrum sensing algorithm using a linear classifier for a single OFDM CR to detect the presence or absence of the primary user signal in an AWGN and fading channel.
- Design a spectrum sensing algorithm using a linear classifier for a Cooperative OFDM CR network to detect the presence or absence of the primary user signal in an AWGN and fading channel.

1.1. Literature Review

The rapid advances in wireless communication technologies have resulted in an increasing demand to access the electromagnetic radio spectrum. The spectrum resources, however, are limited and therefore need to be efficiently utilized and distributed among different candidates of spectrum usage. Each user of the radio spectrum has to obtain a license to use a particular frequency band. The FCC (Federal Communications Commission) regulates and licenses the use of radio spectrum in the United States by assigning part of the spectrum to a specific user, called the primary user, and giving it the legal right to use that portion of the spectrum at all times of the day. It is the responsibility of the regulatory authorities to efficiently utilize this scarce natural resource and to cope with the ever increasing number of prospective spectrum users. In UAE, the equivalent of the FCC is the Telecommunications Regulatory Authority (TRA).

However, according to current practices, each assigned frequency band can only be used by a single user or enterprise at all times in a given geographical area and no interference from other unlicensed users is allowed. Furthermore, there is no obligation by the licensed primary user to use its allocated spectrum at all times. In fact, the licensed spectrum occupancy may be as low as 5% for certain periods [1]. This under-utilized licensed spectrum band by the primary users creates spaces in time and geography which are referred to as spectrum holes [2].

As an attractive solution to the inefficient spectrum usage, the concept of Cognitive Radio (CR) has been introduced to utilize the licensed spectrum whenever it is not used by the primary users [3]. The FCC defines CR as a radio or system that senses its operational electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify system operation, such as maximizing throughput, mitigating interference, facilitating interoperability and access to secondary markets [4]. The CR is considered as a secondary user which is given the right to access the unused portion of the spectrum only when the primary user is not accessing it. It is, therefore, the responsibility of the CR to avoid causing any interference to the licensed primary user. Towards this goal, a CR should be able to continuously sense the spectrum and determine the existence of the so-called spectrum holes. Once a spectrum hole is found, the CR should be able to transmit/broadcast at that frequency band. In addition, it should also be able to immediately vacate the spectrum when the primary user wishes to access it. The secondary user, therefore, needs to have cognitive capabilities to be able to sense the spectrum for the presence of a primary user and then change its radio parameters (i.e. start or stop transmission) accordingly to exploit any unused part of the spectrum [5]. Figure 1 shows a graphical representation of the difference between primary and secondary users in a CR. To sum up, a cognitive radio (CR) should be able to:

- Sense the electromagnetic spectrum over a specific frequency band.
- Detect the presence or absence of primary users.
- Transmit information over the frequency band when a spectrum hole is detected.
- Vacate the spectrum (frequency band) when the primary user starts broadcasting.

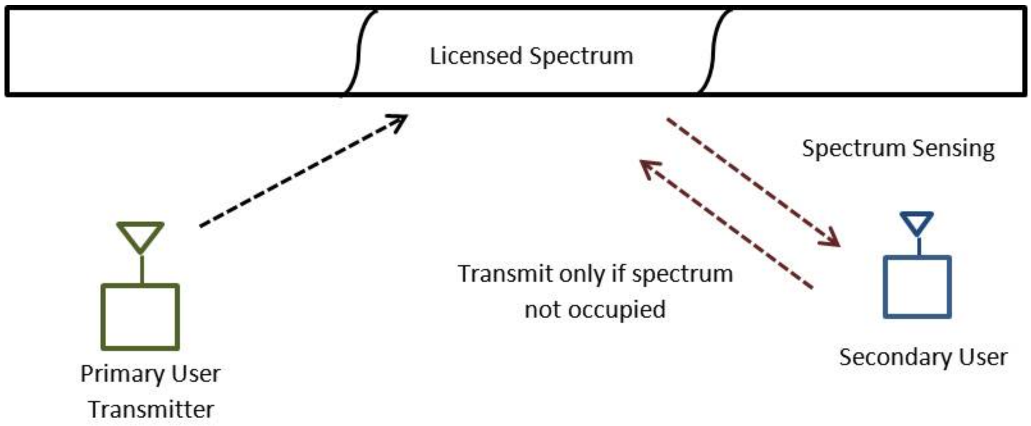

Figure 1 - Primary and Secondary Users

1.1.1. Performance Metric of a CR.

To evaluate the performance of the CR system and quantify the amount of interference between the primary and secondary users, certain performance metrics are defined which are discussed below.

- Detection Probability (P_d): defined as the probability that the secondary user (CR) correctly decides that the target spectrum is busy (i.e. primary user signal is present). The complement of the detection probability is the miss-detection probability (P_m) which is defined as the probability that the CR incorrectly decides that the target spectrum is empty while it is actually busy.
- False Alarm Probability (P_f): defined as the probability that the CR makes an incorrect decision that the spectrum is occupied (by the primary user) while, in fact, it is not.

The objective of any CR system is to maximize the detection probability (P_d) in order to minimize interference between the primary and secondary users and maintain the false alarm probability (P_f) below a certain threshold or constraint in order to maximize spectrum efficiency.

1.1.2. Orthogonal Frequency Division Multiplexing.

OFDM is a multicarrier modulation scheme in which a high data rate signal is sent over multiple carriers unlike the traditional schemes where a single carrier is used to carry the message. The single carrier modulation results in Intersymbol Interference (ISI) among the transmitted symbols due to the large data rates involved. However, multicarrier modulation schemes try to minimize the effect of ISI by dividing the available spectrum into many narrow bands (subcarriers). The frequency response of the channel in each narrowband can be considered to be flat which implies that ISI can be minimized [6]. Each subband is associated with a sinuosidal carrier signal. Multicarrier modulation or OFDM is now widely being used in both wireline and wireless communication systems and has been adopted as the standard modulation scheme for both the WiFi and WiMax systems.

The data to be transmitted is initially modulated using M-ary Quadrature Amplitude Modulation (M-QAM). The series of M-QAM symbols are then divided into several parallel streams and each stream is transmitted on a separate subcarrier. Effectively, a high rate stream of data is split into several low data rate streams each sent on a separate subcarrier. This splitting increases the symbol duration by the number of orthogonally overlapping subcarriers and therefore results in a reduction of ISI [7]. The FFT (Fast Fourier Transform), which is an efficient implementation of DFT (Discrete Fourier Transform) is used to modulate the M-QAM symbols to the different subcarriers. Each row of the DFT (and IDFT) matrix is considered to be a different orthogonal OFDM subcarrier. IFFT (and FFT) is used to implement OFDM systems because it is a computationally efficient orthogonal linear transformation which results in robustness of the OFDM signal in the time domain [8].

A typical OFDM allocates its subcarriers as guard, data (majority of subcarriers are used for this purpose) and pilot subcarriers. The function of each type is discussed below:

- Pilot Subcarriers: Some of the subcarriers are modulated with known data symbols at regular intervals to allow the receiver to perform necessary channel estimation and other physical layer estimation tasks.
- Guard subcarriers (Lower and Upper): These are subcarriers at the edge of the band that are not transmitted at all to reduce adjacent channel interference.
- Data Subcarriers: These are the subcarriers which contain the data to be transmitted.

Figure 2 shows the OFDM modulation and demodulation block diagram. First, the data to be transmitted is converted to parallel streams. If the system uses pilot subcarriers, they are added into these parallel streams. The number of parallel streams, including the data symbols and pilots, are equal to the number of subcarriers in the OFDM system. IFFT is then used to perform the OFDM modulation. The output of the IFFT operator is the OFDM symbol. A cyclic prefix is added to the OFDM symbol by taking the last L samples of the OFDM symbol and appending them to the beginning of the symbol. The cyclic prefix is used as a time guard-band against any remaining interference from other symbols (ISI) [**6**]. Also, it helps preventing the effect of ICI (inter carrier interference) by maintaining orthogonality among the subcarriers **[**9**]**. ICI is caused in a multipath channel when a full OFDM subcarrier cycle is not received resulting in loss of orthogonality among the subcarriers. At the demodulator, the cyclic prefix is removed followed by the FFT operator to recover the data that is converted from parallel to serial and finally delivered to the destination.

The multiple inputs to the IFFT operator in the modulator can be from a single data source or from multiple sources. If the data is from multiple sources, a serial to parallel converter may not be required and data could be directly passed on to the respective subcarrier. Also, based on the application and user requirements, at a particular time, not all the subcarriers are utilized. Some of the subcarriers may be left unused or may be used by other users or data sources.

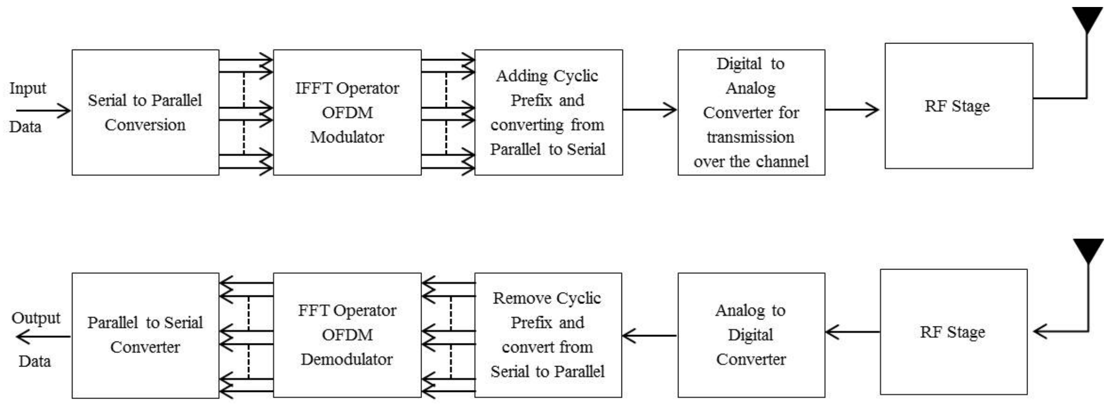

Figure 2 - OFDM Modulation and Demodulation

As mentioned earlier, OFDM is the preferred modulation scheme for several wireless standards today and is used in various applications such as digital television, audio broadcasting, wireless networking and broadband internet access [10]. In fact, DVB-T (Digital Video Broadcasting – Terrestrial), IEEE 802.11a/g, Wireless Local Area Networks (WLANs), IEEE 802.16, IEEE 802.20 are all OFDM based [10]. Therefore, it is reasonable to assume that most primary user signals in a wireless medium will be OFDM based. OFDM is also the best physical layer candidate for a CR system because it allows easy generation of signals that fit into discontinuous and arbitrary sized spectrum segments. Besides, OFDM is the optimal method to achieve Shannon capacity in a spectrum divided into many narrow bands [11]. Recently, a new standard, IEEE 802.22, specifically designed for CR applications was published. The IEEE 802.22 (also known as Wireless Regional Area Network (WRAN)) includes cognitive features such as channel sensing and primary user detection and the modulation scheme employed by this standard is OFDM [7]. Since OFDM is a likely candidate for being the modulation technique used in a CR, this thesis will focus on spectrum sensing techniques in OFDM CR systems only. Table 1 shows the OFDM parameters for the IEEE 802.11a/g and DVB-T standards. The simulation results presented in this thesis are based on the DVB-T standard in 4k mode.

1.1.3. Spectrum Sensing.

From the previous discussion, it is clear that the most important and critical task of a CR is to sense the spectrum to obtain awareness about the spectrum usage and existence of primary users in a geographical area [5]. In fact, the entire operation and performance of the CR depends on spectrum sensing. A CR must regularly survey the portion or band of the spectrum which it wishes to use [12]. If a spectrum hole is found, the CR can engage in communication with other users of the spectrum or broadcast information.

Table 1 – OFDM Parameters for IEEE 802.11a and DVB-T

Parameter	IEEE 802.11a/g	DVB-T 2k, 4k, 8k Mode
Total Subcarriers (FFT size)	64	2048, 4096, 8192
Pilot Subcarriers	4	62, 245
Cyclic Prefix	1/4 of FFT size	1/4, 1/8, 1/16, 1/32 of FFT size
Bits per Symbol	1, 2, 4, 6	1, 2, 4, 6

Spectrum sensing in a CR generally involves obtaining spectrum usage characteristics across multiple dimensions such as time, spatial and/or frequency domain [5]. The spectrum sensing problem can be characterized as a hypotheses test based on the existence of the primary user. This problem is written, in general, as:

$$H_0\text{: } Primary\ user\ is\ absent$$
$$H_1\text{: } Primary\ user\ is\ present, \tag{1}$$

where H_0 is the null hypothesis and H_1 is the alternate hypothesis. The numerical/mathematical definition of the hypothesis test changes based on the type of spectrum sensing technique. Using the hypothesis test definition in (1), the different performance metric probabilities of the CR (discussed in section 1.1.1.) are written in terms of the proposed hypotheses and are shown in Table 2 where $\Pr(.)$ is defined as the probability of the expression in the parenthesis. The detection probability is now defined as $\Pr(H_1|H_1)$ which is the probability of correctly deciding on the alternate hypothesis H_1 i.e. deciding on the presence of primary signal given that the hypothesis H_1 is true.

Table 2 – Performance metrics using Hypothesis Test

Performance Metric	Definition
Detection Probability	$\Pr(H_1\|H_1)$
Miss-Detection Probability	$\Pr(H_0\|H_1)$
False Alarm Probability	$\Pr(H_1\|H_0)$

1.1.4. Spectrum Sensing Techniques.

There are various techniques for spectrum sensing which have been discussed exhaustively in the literature. A brief overview of the most popular techniques is presented in this section.

1.1.4.1. Neyman-Pearson Theorem and Likelihood Ratio Test.

The most famous technique for hypothesis testing is the Neyman-Pearson Theorem. From (1), H_0 is defined as the null hypothesis (when primary user is not present) and H_1 is defined as the alternate hypothesis (primary user is present). The Neyman-Pearson Theorem uses the likelihood ratio test to maximize the detection probability for a fixed false alarm probability. The likelihood ratio test decides on H_1 (alternate hypothesis) when the likelihood, $L(x)$, is greater than a threshold λ such that [13] [14]:

$$L(x) = \frac{p(x|H_1)}{p(x|H_0)} > \lambda, \tag{2}$$

where x is the received signal amplitude, $p(x|H_1)$ is the conditional probability of receiving a signal given the hypothesis H_1 and $p(x|H_0)$ is the conditional probability of receiving a signal given the hypothesis H_0. The threshold λ is determined according to the desired false alarm probability. If the distributions of the signals are known, the necessary parameters are computed and $L(x)$ is calculated using (2). However, if the received signals under the two hypotheses are assumed to be Gaussian, the unknown mean and variance is estimated using the maximum likelihood estimate (MLE) and then the likelihood ratio test is applied [14]. This method is known as the generalized likelihood ratio test (GLRT) and the MLE is defined as:

$$\widehat{\theta_0} = \arg {}^{max}_{\theta_0} p(x|H_0, \theta_0)$$
$$\widehat{\theta_1} = \arg {}^{max}_{\theta_1} p(x|H_1, \theta_1), \tag{3}$$

where θ_0 and θ_1 are the unknown parameters (mean and variance) of the distributions under H_0 and H_1 respectively. After obtaining the estimate of the parameters, the likelihood, $L(x)$, is computed and the presence of primary user is decided (H_1) by comparing it with the threshold value using:

$$L(x) = \frac{p(x|\widehat{\theta_1}, H_1)}{p(x|\widehat{\theta_0}, H_0)} > \lambda. \tag{4}$$

The likelihood ratio test is the optimal algorithm for spectrum sensing but it requires the exact knowledge of the noise variance and other parameters which are generally difficult to obtain [3]. Also, if the assumption of Gaussian signals is not true, then the estimates of the parameters will be incorrect resulting in a degraded performance. On the contrary, if this assumption is correct and the parameters are properly estimated, the Likelihood ratio test will give the best results compared to any other spectrum sensing technique.

1.1.4.2. Energy Based Detection.

Energy detection is the most popular technique for spectrum sensing [15]. It is simple to implement and does not require any knowledge of the primary signal characteristics. The presence or absence of a signal is decided by comparing the output of the energy detector with a threshold that depends on the noise floor. This method works best for the detection of any unknown zero-mean constellation signals [16]. A simple approach is discussed in [5] where the signal is assumed to be corrupted by an Additive White Gaussian Noise (AWGN) noise for the case when primary user is accessing the spectrum. The discrete time received signal is defined as:

$$y[k] = s[k] + w[k], \tag{5}$$

where $s[k]$ is the primary user signal, $w[k]$ is AWGN, and k is the discrete time index. In general, the energy of any received signal is computed as:

$$\epsilon = E[|y[k]|^2]. \tag{6}$$

where $E[.]$ is the expectation operator. However, for a zero mean Ergodic signal, the energy can be estimated as [17] [18]:

$$\epsilon = \sum_{k=1}^{W} |y[k]|^2, \tag{7}$$

where W is the size of the observation window. The energy is then compared against a threshold λ to decide about the hypotheses:

$$\begin{aligned} &H_1: if\ \epsilon \geq \lambda \\ &H_0: otherwise. \end{aligned} \tag{8}$$

Another method to implement the energy detector is to use Fast Fourier Transform (FFT). The received signal is sampled and passed on to a FFT to obtain a discrete frequency-domain signal, $X(m)$ where m is the dicrete frequency index. Using Parseval's theorem, the energy of the signal is computed by summing the energy at each discrete frequency $|X(m)|^2$ over all frequencies and comparing it against the threshold λ:

$$\begin{aligned} &H_1: if\ \sum_{m} |X(m)|^2 \geq \lambda \\ &H_0: otherwise. \end{aligned} \tag{9}$$

Despite being easy to implement, the energy detector still has its disadvantages. One main drawback is that it can only detect those primary signals which have energy levels above the set threshold. This is problematic for the case when the primary signal energy levels are below the noise floor. Also, the selection of the optimal threshold is challenging [16]. It requires the knowledge of noise power to achieve expected performance in terms of probability of detection and probability of false alarm. However, noise power is difficult to be accurately estimated and a rough estimate of noise power is used, which severely degrades the performance of energy detection [17]. Also, noise power varies over adding more error to the estimation process [18]. Typically, spectrum sensing through energy detection uses the system model described in (5) which assumes the signal is corrupted by AWGN only and no fading due to the channel is assumed [5] [17] [18]. Fading causes significant degradation to the performance of an energy detector.

To avoid estimation of noise power to set the threshold for energy detection, different approaches are used. A Bayesian estimation is used which takes measurements in all the sub channels to determine the occupied channels. The occupied channels are estimated by maximizing the likelihood of the measured samples. Out of these channels, the channels with the largest power are assumed to be used by the primary users. Using this technique, degradation caused by lack of knowledge of noise power on energy detection is significantly reduced [17]. A histogram based approach is also used to determine the threshold. A large number of samples of signals belonging to H_0 and H_1 are collected and two histograms are obtained. Based on the histograms, a threshold, λ, is chosen to meet a certain target for probabilities of detection and false alarm [18]. With this technique, the received signal does not need to be modeled statistically and no estimation of noise power is required. Also, the histogram method is independent of the distribution of the primary user signal.

Energy detection is also used in Co-operative CR networks [19]. Two different methods have been proposed: data fusion and decision fusion. In both cases, a fusion center receives data from all CRs in the network and makes a decision on the presence or absence of the primary user signal. For data fusion, each CR amplifies the received signal and forwards it to the fusion center. At the fusion center, different fusion techniques such as Maximal Ratio Combining (MRL) or square-law combining (SLC) are used. On the other hand, in decision fusion, each CR makes a decision on the presence of primary user and sends this decision to the fusion center. The fusion center uses OR, AND or Majority rule to make a decision on the presence of primary user. With more number of CRs in the network, the system makes a more informed decision on the primary user. Therefore, for energy detection with Co-operative CR systems, multipath fading and shadowing have less impact on the performance [19].

1.1.4.3. Correlation and Cyclostationary Detection.

Most signals inherently contain some correlation properties which can be exploited to sense their presence in the spectrum. The autocorrelation, which is correlation of the signal with itself, can be used as a tool to determine the presence of signal when the primary user signal is correlated and is defined as:

$$R_x(\tau) = E[x(t)x^*(t+\tau)], \tag{10}$$

where τ is the delay, (.)* represents conjugate and $E[.]$ is the expectation operator. If the received signal $x(t)$ is the primary user signal, the correlation value $R_x(\tau)$ has a maximum value when $x(t)$ is correlated with a delayed version of itself. However, when the received signal is noise only, any two samples of noise are uncorrelated. Using this fact, the existence of the primary signal is decided. Besides, if the correlation in the signal is known to the CR, the optimal detector uses the Neyman-Pearson theorem for spectrum sensing [13] [20].

A random signal or process is said to be cyclostationary if it is periodic or its statistics such as mean and autocorrelation are periodic. In [5], cyclostationary detection is performed by computing the cyclic autocorrelation function, CAF, for the received signal at the CR as:

$$CAF(\tau, \alpha) = E[x(t+\tau)x^*(t+\tau)e^{-j2\pi\alpha t}], \tag{11}$$

where α is the cyclic frequency. From the CAF, the Cyclic Spectral Density (CSD) can be computed by finding the discrete Fourier transform of the CAF. The CSD gives peak values only when the cyclic frequency is equal to the fundamental frequency of the primary user signal When only noise is present, no peaks are observed. This property is used to determine the presence of the primary user signal by comparing the CSD output with a threshold that maximizes the detection probability for a fixed false alarm. Also, from (12), it can be seen that CAF is a generalization of the autocorrelation function, $R_x(\tau)$ and autocorrelation is, in fact, equivalent to the CAF computed at zero cyclic frequency. The cyclostationary based detection can differentiate between noise and primary user signals even under bad channel conditions [5].

1.1.4.4. Entropy based Detection.

In information theory, entropy is defined as a measure of uncertainty and ambiguity in a discrete random variable, Z [6]. Since knowledge of Z removes all uncertainty about it, the entropy is also a measure of information that is acquired by knowledge of Z and is defined as:

$$H(Z) = -\sum_{z \in Z} P[Z = z] \log_2 P[Z = z]. \tag{12}$$

The entropy of a signal is usually calculated using the histogram based method where the number of states (or levels) the primary user signal (random variable) occupies is equal to the number of bins of the histogram [21]. Using this method, the entropy of any received signal is calculated at the CR and compared with a threshold and a decision on presence or absence of primary user signal is made. It is expected that entropy is low when the known primary signal is received and is high when random noise is received. It is proven in [21] that when computed in the time domain, the entropy values of the primary signal and noise fluctuate around a constant value and are invariant against SNR for both the hypotheses which makes it difficult to distinguish noise from the primary signal. The received signal is therefore first transformed to the frequency domain using DFT (or FFT) and then the entropy is computed.

1.1.4.5. Waveform-based Sensing.

This method of spectrum sensing makes use of the known patterns which are inserted in the primary user's transmitted signal to assist in synchronization. These patterns include preamble, midamble, regularly transmitted pilot patterns, spreading sequence, etc. Preamble is a known sequence transmitted before each burst while midamble is transmitted in the middle of the burst. If it is known that there are such patterns present in the transmitted signal, then sensing can be performed by correlating the received signal with a known copy of the pattern over an observation window of W samples. The output of the correlation operation is compared with a threshold and presence or absence of the signal is decided.

$$R = \sum_{k=0}^{W} y[k]s^*[k], \tag{13}$$

where $s[k]$ is the known copy of the signal pattern and $y[k]$ is the received signal pattern.

Waveform based sensing can only be applied to signals with known patterns, and for this reason this method is also known as coherent sensing. Coherent sensing outperforms energy detector in reliability and convergence time but coherent sensing requires the knowledge of certain patterns in the transmitted signal which are not required by the energy detector [5]. In order to perform well, the CR must be perfectly synchronized with the primary user signal.

1.1.4.6. Matched Filter.

A Matched filter is an optimal detection method as, fundamentally, it maximizes the SNR of the received signal in the presence of AWGN. A matched filter is implemented by correlating a known transmitted signal with the received signal. This is, effectively, equivalent to convolving the received signal with a time reversed version of the transmitted signal. Matched filter, therefore, has the constraint that it requires complete knowledge of the transmitted signal shape and duration [5].

1.1.4.7. Pattern Classification Tools for Spectrum Sensing.

Pattern recognition systems are used to classify data into pre-determined classes. Since the data received at the CR can be either a primary user signal or noise only, the spectrum sensing problem can be considered as a two class pattern recognition problem and different classifiers can be used to solve the problem. However, for proper signal detection, discriminative features have to be extracted from the received signal. An energy detection based co-operative CR network is proposed in [22] where linear and polynomial classifiers are used for spectrum sensing. Each CR extracts energy as a feature from the received signal over an observation window and transmits it to a fusion center. The fusion center uses the data received from all the CRs as features, and feeds it to the classifier which decides on the existence of the primary user signal globally. The system is first trained using a large number of signals belonging to both classes and weights are computed which are used to classify any unknown received signal into one of the two classes. In the testing phase, known data is used to determine the performance of the classifier, in terms of detection probability, while maintaining the false alarm probability below a certain value. It is observed that both the linear and polynomial classifiers had comparable performance in a fading channel; and linear classifier provided a better solution to the spectrum sensing problem due to its reduced complexity compared to a polynomial classifier.

Cyclostationary features of the received signal at the CR are used in [23] to detect the presence of the primary signal in a co-operative CR network employing linear and polynomial classifiers. Each CR uses the cyclostationary property in the received signal over an observation window, by computing the CSD, which is the discrete Fourier transform of CAF, defined in (12). The calculated feature is then transmitted to the fusion center. The fusion center receives features from all CRs in the network and uses them to decide on the presence of the primary user signal. Once again, the system is first trained and weights are computed. Finally, testing is performed and the performance of the classifier is evaluated. It is concluded that the linear and polynomial classifiers have comparable performance and the linear classifier is the preferred solution since it is less complex to implement. However, the primary user signal in both cases is considered as a single carrier signal and OFDM systems have not been considered.

All the aforementioned methods have their advantages and drawbacks. The likelihood ratio test provides the optimal solution for the case when all the signals in the system are Gaussian distributed. Energy detector does not require any knowledge of primary signal while all other techniques require certain amount of knowledge of the signal. However, energy detector cannot detect signals with energy levels below the noise floor. Cyclostationary detection, on the other hand, is robust to noise and interference from other users whereas the matched filter provides the optimal solution to the detection problem when the primary signal is known and the channel is AWGN. However, both these method require either partial or complete knowledge of the primary user signal. Finally, spectrum sensing can also be defined as a two class pattern recognition problem and different classifier models can be used to decide if the received signal at the CR is a primary signal. This method does not require any statistical analysis to compute the thresholds, but large data belonging to both classes is required to train the classifier and compute the weights.

1.1.4.7. Spectrum Sensing in OFDM based CR: Overview.

As discussed earlier, OFDM systems are expected to be the most common primary user signals for CR systems. In this section, different spectrum sensing techniques for OFDM based CRs are discussed.

The spectrum sensing task for an OFDM based CR is slightly simplified by the fact that the CR has some knowledge about the primary user signal. However, in different applications, the amount of knowledge available to the CR may vary. The location of pilot carriers and length of the cyclic prefix may or may not be known. If the CR is operating under some standard system, such as Wi-Fi, then the location of the pilot carriers and cyclic prefix length is determined by the standard and are known to all users. Besides, most OFDM based CRs utilize techniques similar to the aforementioned methods to sense the spectrum with slight modifications to account for the special properties of OFDM. Almost all techniques for spectrum sensing discussed in the literature employ the likelihood ratio rest, described in Section 1.1.4.1, to decide on the presence of the primary user. To be able to correctly exploit the potential of the Neyman-Pearson theorem, the distribution of the data has to be Gaussian. Assuming that the number of subcarriers is large enough, the central limit theorem is invoked and the OFDM signal is considered to be Gaussian distributed with a known variance and zero mean. The variance of the Gaussian distributed primary user signal is estimated using the maximum likelihood estimate (MLE) [10]. Cyclic prefix is another important feature of the OFDM system. The insertion of cyclic prefix at the beginning of each OFDM symbols adds a cyclostationary and correlation property to the system which can be used for spectrum sensing.

In [24], energy detection in presence of AWGN is used for spectrum sensing but the detection is done over the cyclic prefix region of the OFDM signal only. The energy of the received signal is computed and compared to a threshold which is determined using the Neyman-Pearson theorem to maximize the detection probability and keep the false alarm probability fixed. After comparing the received signal energy with the threshold, a decision on the presence or absence of primary signal is made.

In [10], it is proven that the presence of cyclic prefix in the OFDM signal makes the autocorrelation coefficient nonzero at delays $\tau = N$, where N is the number of subcarriers. This implies that the last L samples of the OFDM symbol are correlated with the first L samples. Noise, on the other hand, is uncorrelated. Utilizing this property, a spectrum sensing technique is proposed. Assuming that both noise and the OFDM signal are Gaussian distributed, the MLE is first used to estimate the variance of the received signal. Then, the likelihood ratio test is used to find the autocorrelation coefficient. This coefficient is then compared with a threshold to determine the existence of the primary user signal. The threshold is found using the Neyman-Pearson detector to satisfy a given false alarm probability constraint. A cooperative scheme is also proposed in [10] where the log likelihood ratios of the autocorrelation coefficient from all CRs are sent to a fusion center to determine the existence of the primary user signal.

Another autocorrelation based Cooperative sequential spectrum sensing technique is proposed in [25]. A Cooperative CR network is proposed to reduce degradation in performance which a standalone CR might experience due to shadowing and fading. At the CR, the correlation between the beginning and end of the received signal is computed. If the OFDM signal is present, this value is high; and if there is only noise present, the correlation is small. The correlation output is then compared with a threshold which is determined using the Neyman-Pearson theorem under the assumption that the received signal is Gaussian distributed. In this method, each CR computes the correlation value and compares it with a local threshold to decide on the presence or absence of the primary user signal. The CR then transmits a message to the fusion node. At the fusion node, the messages from all the CRs are compared with a threshold and a final decision is made. The threshold is determined using log likelihood ratio based on the two possible messages: signal present or absent.

A method of spectrum sensing using Pilot Tones in OFDM systems is proposed in [26]. It is assumed that the CR has knowledge of the location of the pilots embedded in the OFDM signal. The pilots are added in frequency domain for synchronization and channel estimation. The spectrum sensing algorithm utilizes the fact that the Time-Domain Symbol Cross-Correlation (TDSC) has peak values if the two OFDM symbols have the same frequency-domain pilot symbols. A decision is made on the presence of the primary user signal by comparing the TDSC with a threshold. The commonly used Neyman-Pearson method is used along with the likelihood ratio test to determine the threshold that gives the maximum detection probability.

1.1.5. Pattern Recognition/Classification.

This section provides an overview of pattern recognition systems. Pattern recognition is the act of taking in raw data and taking an action based on the category of the pattern [27]. In pattern recognition, objects are defined by a set of measurements called attributes or features which are used to classify them into different classes [28]. Figure 3 shows the different stages in the design of a pattern classification system which are discussed below.

Data Collection: A suitable sensor is used to collect data belonging to all classes typically using a transducer. Each data element provides some knowledge about its respective class. Generally, a classification system performs better when a large number of data samples are available as it will accommodate all the possible variations and changes in the data. Besides, the data is labeled and marked according to its class so that it can be used for training and testing the system. The collected data is then divided into training and testing data.

Feature Extraction: After collecting the data through sensors, any unwanted or redundant information is removed from the data to aid in feature extraction. By definition, feature extraction is the process of acquiring features from the input data. A feature is a discriminating characteristic of the data which can be used to identify the data as belonging to a particular class. Sometimes, multiple features are extracted from the data and used for classification. Nevertheless, features have to be carefully selected and only those characteristics which make the data belonging to different classes distinct are chosen. The features are extracted from the data and stored in a feature vector which is plotted on a feature space. The dimensionality of this space is equal to the number of features in the feature vector. If the features are selected correctly, the feature vectors of each class will be clustered together. These clusters should be separable through a linear or non-linear decision boundary which is determined by a classifier.

Classifier Design: A pattern recognition machine, which classifies any input signal into a particular class, is called a classifier. It is very important to select the correct classifier model that fits the statistics of the extracted features. Important parameters which are considered when deciding the model of the classifier include dimensionality of the feature vector and separation of the feature vectors in the feature space (linear or non-linear decision boundaries). After selecting the appropriate model, training data is used to train the classifier and calculate the weights. Finding the weights is equivalent to determining the decision boundaries of classes in the feature space. It has to be ensured that the training data set is significantly large to accommodate all variations in the data, but at the same time is small enough to prevent over-fitting. Over-fitting happens when the training data size is very large and the classifier is extensively calibrated to correctly classify the training data specifically. If this happens, the chance of a classification error for new testing data increases as the classifier has lost generalization.

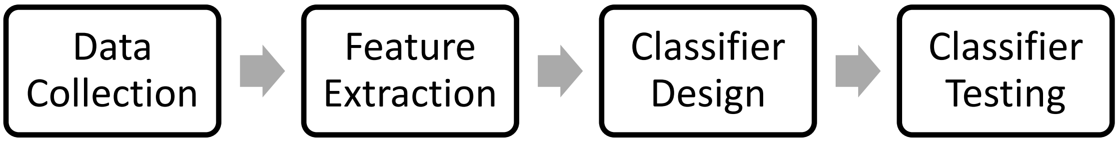

Figure 3 - A Pattern Recognition System

Classifier Testing: After training the classifier and computing its weights, it has to be tested using a different set of data (i.e. testing data). The performance of the classifier is determined by observing the classification rate achieved when testing data is used. Each element of the testing data set is passed through the classifier and a decision on its class is made. Since the testing data is labeled, the decided class is compared to the actual class to determine if it has been correctly classified. The process is repeated for all for all the elements in the testing data set and the classification rate of the classifier is calculated. A high classification rate is achieved if the data belonging to different classes is separable, either through a linear or non-linear decision boundary in the feature space. If the data for different classes overlaps, then the classifier is confused and a poor classification rate is achieved.

Pattern recognition problems are divided into two main types: unsupervised and supervised. Unsupervised learning implies that there is no knowledge of any groups or classes in the data and the system attempts to determine if the data can be classified into groups. This is done by determining which characteristics make the data similar within the same class and different across other groups. In supervised learning, on the other hand, each data sample belongs to one of a known pre-specified number of classes. The classifier needs to be trained properly and discriminating features have to be selected to achieve good classification rates [28].

1.1.5.1. Linear Classifiers.

A linear classifier is a commonly used model to implement a pattern recognition system with linear decision boundaries among classes. A linear classifier is employed when the features of the different classes are linearly separable. A linear discriminant function is a linear decision boundary which separates data from different classes and is defined using the weight vectors. The number of linear discriminant functions is equal to the number of classes in the data set. The weights of the discriminant function are computed during the classifier training stage. Linear classifier is the preferred model for a classifier because it is relatively simple to implement and computationally cheap [29]. In this thesis, a linear classifier is used to classify the incoming received signal into primary user (class 1) or noise (class 2).

1.1.6. Fading Channel.

In an ideal non-bandlimited channel environment, a transmitted signal is corrupted by AWGN (Additive White Gaussian Noise). This implies that a zero-mean unit variance Gaussian noise is added to the received signal. The matched filter is the optimum receiver structure in an AWGN channel which successfully recovers the transmitted signal from the received signal. However, in practical wireless environment, the transmitted signal suffers from rapid fluctuations in its amplitude which is referred to as small-scale fading (or simply fading). Fading is the result of movements of the transmitter, receiver or objects surrounding them. The two main factors which contribute to the fading effect in the channel are discussed below:

Multipath Delay Spread and Coherence Bandwidth: In most communication systems, there is usually no line of sight (LOS) between the transmitter and receiver which means that the received signal contains multiple copies of the transmitter signal from different directions and with different delays. The multiple copies received at the receiver result in interference which could be constructive or destructive depending on the phase of each copy of the signal [9]. The multipath components of the received signal are typically represented using the power delay profile (PDP) which plots the average power output of the channel as a function of time delay. Effectively, the PDP presents the multipath components of the received signal according to their delays. The range of values over which the power is essentially nonzero is called the multipath spread of the channel T_m. Besides, Root Mean Square (RMS) delay spread is another useful parameter which provides a reference of comparison among different multipath fading channels [9]. The RMS delay spread σ_τ is defined as the square root of the second central moment of the PDP and it describes the time dispersion (or spread) of the channel. The coherence bandwidth f_C of the channel is defined as the bandwidth over which the channel frequency response is considered be constant or fixed and is inversely proportional to the RMS delay spread. If the signal bandwidth is greater than the coherence bandwidth of the channel, then different frequencies of the signal will be affected differently by the channel resulting in distortion and ISI. Moreover, the multipath delays in the channel result in time dispersion of the channel.

Doppler Spread and Coherence Time: The time variations in the channel are closely related to the motion of transmitter or receiver which cause a dispersion (or shift) in the carrier frequency of the signal which is referred to as Doppler spread f_D [9]. The Doppler spread depends on the velocity of the mobile transmitter or receiver and angle of incidence of the received signal relative to the direction of motion and also on the carrier frequency [30]. The time variations in the channel are also described using the coherence time of the channel t_C defined as the time over which the channel impulse response is constant or fixed and is the reciprocal of the Doppler spread of the channel. If the signal duration is greater than the coherence time of the channel, the signal will suffer from rapid and fast variations in the channel which result in severe fading.

Based on the parameters and relationships defined earlier, the signal can undergo different type of fading based on the values of RMS delay spread, Doppler spread, and coherence time and bandwidth of the channel. In essence, Multipath (or RMS) delay spread and coherence bandwidth result in time dispersion and frequency selective fading while Doppler spread and coherence time (of the channel) lead to frequency dispersion and time-selective fading. The conditions for the different types of fading are shown in Figure 4 where T_s is the symbol duration and W is the signal bandwidth.

Due to fading, the transmitted signal is attenuated by complex-valued channel coefficients which represent the fading caused by the channel besides being corrupted by AWGN. The channel coefficients are modeled statistically and the envelope of the channel is assumed to be Rayleigh distributed and the phase is uniformly distributed over the interval $(0,2\pi]$. If there is a LOS between the transmitted and the receiver, then the envelope of the channel is assumed to be Rician distributed.

1.2. Research Methodology

In this thesis, the problem of spectrum sensing in a cooperative CR network is approached by first developing the signal and channel models. It is assumed that the primary user signal is OFDM based and the signal passes through a flat fading channel. Spectrum sensing is then formulated as a pattern recognition problem and a classifier is used to decide on the existence of the primary user signal. Finally, extensive simulation results have been presented for several cases.

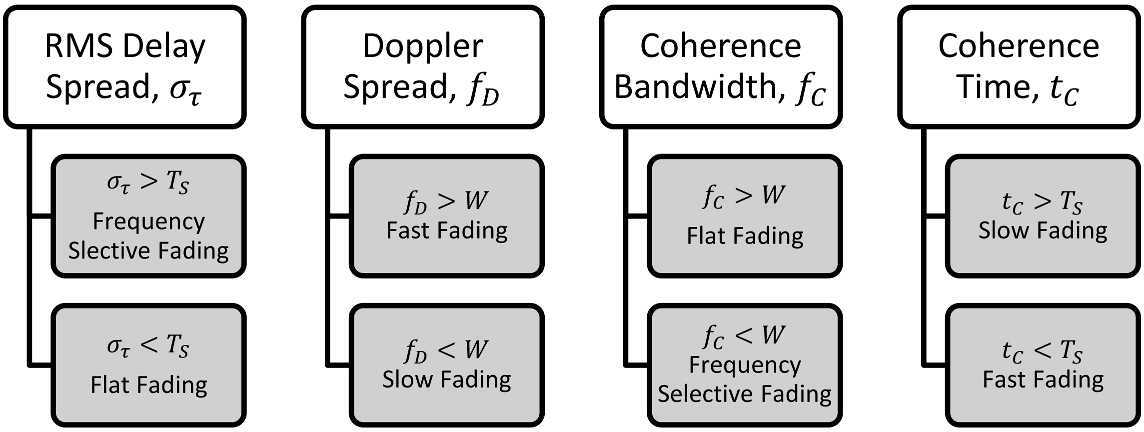

Figure 4 - Different Fading Types

1.3. Thesis Contributions

The main contributions of this thesis can be summarized as:

- Spectrum Sensing as Pattern Recognition problem: The problem of spectrum sensing is modeled as a two-class pattern recognition system. Features are extracted from the signal and fed to a classifier which operates on the features and provides a decision on spectrum occupancy.
- A Cooperative CR network is implemented where several CRs collaborate together to sense the spectrum. A fusion center receives features from all CRs in the network and makes a decision on the existence of the primary user signal.
- A comprehensive performance assessment of three different spectrum sensing schemes is presented. Energy, correlation and entropy are used as features to the linear classifier and their performance is evaluated at different noise levels in a fading channel. Some of the findings of this work were presented in [31].

1.4. Thesis Organization

The rest of the thesis is organized as follows. In Chapter 2, the system model used for simulations is presented. The proposed problem is formulated in Chapter 3 and the features to be used by the classifier are discussed. Chapter 4 illustrates the performance of the proposed system by presenting the simulation results. Finally, Chapter 5 concludes the thesis and discusses possible scope of future work.

Chapter 2

System Model

In this chapter, the mathematical model for the proposed system is developed. First, the model for the OFDM primary user signal is discussed. Then, the channel model used in this thesis and the effect it has on the transmitted signal is presented. Finally, the CR receiver operations are described.

2.1. OFDM Signal Model

The OFDM signal is constructed by feeding N M-QAM symbols to an N-point IFFT operator. Therefore, to form an OFDM symbol, the binary bits of information (1's and 0's) have to be first modulated using M-QAM. In M-QAM, the digital signal can take one of M different levels with varying amplitude and phase. If M is the number of levels that the M-QAM signal can occupy then k is defined as the number of bits needed to represent each symbol:

$$k = \log_2 M. \tag{14}$$

The binary stream of information is divided into several blocks of k bits. Each of these blocks is then converted to one of the M signal levels in the M-QAM constellation space to form a symbol. In practice, two simultaneous blocks, each containing k bits, from the information bit sequence are impressed upon two quadrature carriers $\cos 2\pi f_0 t$ and $\sin 2\pi f_0 t$, where f_0 is the carrier frequency [6]. The M-QAM signal is a complex valued signal which has an amplitude and phase.

After dividing the information bit stream into blocks of k bits each and modulating them into a M-QAM signal, the N M-QAM symbols are fed into an IFFT operator to create the OFDM symbol. Since IFFT is used at the modulator, the frequency domain representation of the OFDM signal is the set of M-QAM symbols which were input to the IFFT.

If $S(m)$ is an M-QAM symbol where m is the frequency index, then N such symbols are fed into the IFFT operator. Therefore, the input to the IFFT operator is

$S(0), S(1), \ldots, S(N-1)$. The output of the N-point IFFT is the OFDM signal in the time domain:

$$s[k] = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} S(m) e^{\frac{j2\pi km}{N}}, \quad k = 0, \ldots, N-1, \tag{15}$$

where k is the discrete time index and m is the discrete frequency index. Thus, N denotes the number of samples in an OFDM data block. The last L samples $s(N - L), s(N - L + 1), \ldots, s(N - 1)$ are added to the front (beginning) of each block as a cyclic prefix to obtain the OFDM symbol of the form:

$$\boldsymbol{s} = [s(N-L), \ldots, s(N-1), s(0), s(1), \ldots, s(N-1)]. \tag{16}$$

The signal in (16) is a discrete time-domain digital signal which is first converted to a continuous analog signal $s(t)$ which is then sent over the channel after up-conversion to the desired radio frequency carrier. This signal interacts with the channel depending on the type of channel. The channel model considered in this thesis is discussed in the next section.

2.2. Channel Model

As the signal propagates from the transmitter to the receiver, it interacts with the channel, i.e. the medium through which it is travelling through, in ways that are determined by different channel models. These models try to relate the physical characteristics of the medium to a statistical definition for the channel. A statistical model can be used to represent the channel since it is a randomly changing medium. The channel is also modeled as a linear filter through which a signal has to pass. Therefore, the response of the channel is described through its impulse response or transfer function. Like a filter, a channel too has an effective bandwidth over which a signal passes without severe attenuation. The signals outside the bandwidth of the channel are attenuated and suppressed.

Additive white Gaussian noise (AWGN) is the simplest channel model which assumes that the signal is corrupted by a white noise process. Physically, white noise is caused by thermal noise from electronic components and amplifiers in the receiver and from interference during transmission **[6]**. Thermal noise is statistically characterized as being Gaussian distributed therefore the name Gaussian noise. Besides, white noise is uncorrelated and has a constant power spectral density over all frequencies.

When a signal passes through a wireless channel, it is not only corrupted by a linear addition of Gaussian noise, but it also suffers from rapid fluctuations in amplitude and phase due to movement of the transmitter, receiver or objects surrounding them. The different causes and types of fading have been discussed in Section 1.1.6. In this thesis, it is assumed that the signal passes through a flat fading slow channel with a low Doppler frequency of 3 Hz. This implies that the channels coherence bandwidth is greater than the bandwidth of the signal and the Doppler frequency is much smaller than the signal bandwidth. The transmitted signal is attenuated by an amount determined by the channel coefficients which are modeled statistically using the Rayleigh distribution. Besides, the signal is also corrupted by Gaussian noise. Therefore, the received signal $x(t)$ is:

$$x(t) = c(t)s(t) + n(t), \quad 0 < t \leq T_0 \tag{17}$$

where $c(t)$ is the channel coefficient at time t, $n(t)$ is the Gaussian distributed noise component with the two-side power spectral density of $\frac{N_0}{2}$ and T_0 is the observation window duration in seconds. Due to the low Doppler frequency, the channel coefficients are considered to be slowly varying (slow fading) over the observation window. However, the channel changes for the next observation window. Figure 5 shows the block diagram of the channel model. When there is no fading in the channel, the channel becomes an AWGN channel and $c(t)$ = 1.

2.3. CR Receiver Operation

At the CR receiver, the signal defined in (17) is received which is first down-converted to baseband. The receiver then performs analog-to-digital conversion to form the discrete digital signal:

$$x[k] = c[k]s[k] + n[k], \qquad k = 1, \dots, W \tag{18}$$

where k is the discrete time and W is the observation window size. After analog-to-digital conversion, the observation window duration T_0 seconds is converted to W samples. At the CR, all computations and signal processing are performed on the signal defined in (18). Since the received signal suffers from fading, the effect of the channel has to be compensated to allow for proper demodulation at the receiver. It is assumed that the receiver has complete knowledge of the channel coefficients and the received signal is divided by the channel coefficients to eliminate the distortion in the signal caused by fading. Dividing by the channel coefficients, however, can result in enhancement of noise when the signal energy is comparable to the noise floor. After channel compensation, the signal is then sent to a feature extraction module followed by the linear classifier for classification. Additionally, poor estimation of the channel coefficients results in a severe degradation in performance as the threshold and weights of the classifier will be computed inaccurately. The process of feature extraction, testing and training the linear classifier is discussed in the next chapter.

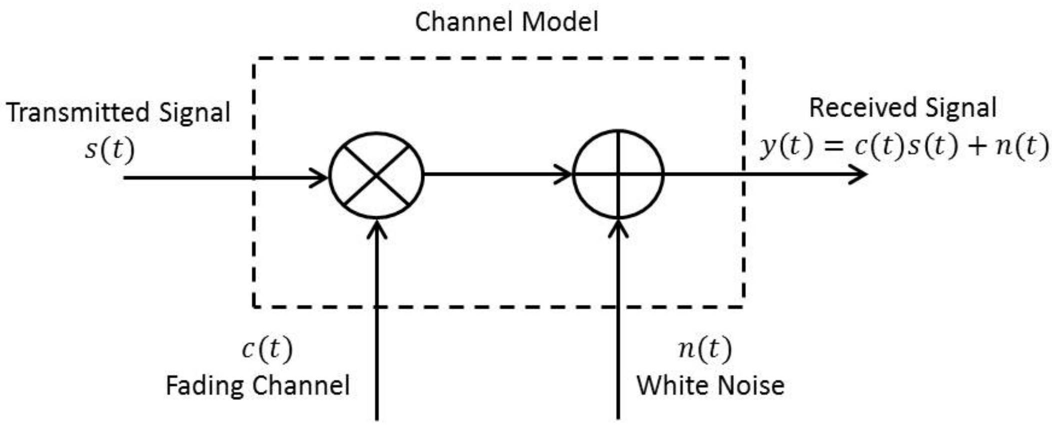

Figure 5 - Fading Channel Model

Chapter 3

Problem Formulation

In this chapter, spectrum sensing is modeled as a pattern recognition problem. It is first shown that spectrum sensing can be considered as a two class pattern recognition problem. Then, the Cooperative CR network architecture used in this thesis is described. Further, the different features that are used to train the linear classifier are defined. Finally, the process of training and testing the linear classifier is discussed.

3.1. Spectrum Sensing as a Pattern Recognition Problem

The most important function of a CR is to perform spectrum sensing to detect whether the target spectrum is occupied by the primary user or not. The CR senses the spectrum over an observation window and makes a decision based on the signal received during this time. The signal received at the CR can belong to one of two types: the primary OFDM signal or noise.

When the OFDM primary signal is being transmitted, the CR receives the primary signal which is attenuated by a value determined by the channel coefficient and corrupted by a linearly added Gaussian noise signal (AWGN). On the other hand, when there is no primary user signal present in the spectrum, the signal received at the CR is random Gaussian noise only. It is clear from this discussion that the received signal belongs to one of two classes only and therefore spectrum sensing can be considered as a pattern recognition problem. A pattern recognition system classifies any data into one of the pre-specified number of classes. In this case, there are two classes and the pattern recognition system decides the class of any received signal by passing it through a classifier. The two classes are defined as:

$$x[k] = \begin{cases} c[k]s[k] + n_1[k]\,; & \text{Class 1} \\ n_2[k] & ;\ \text{Class 2,} \end{cases} \tag{19}$$

where $n_1[k]$ and $n_2[k]$ are different random Gaussian distributed noise components received at the CR. The CR compensates for the channel by dividing the received signal with the channel coefficients. The characteristics which differentiate the signals from both classes are called features and are acquired by the CR through the process of feature extraction. The classifier is able to learn these characteristics during the training phase. Once fully trained, the classifier then utilizes this learning to classify any unseen data into one of the two classes.

At each CR, features are extracted from the received signal using well known techniques such as Energy, Correlation, etc. which are discussed in Section 1.1.4. After the feature extraction stage, the classifier makes a decision on the class of the received signal/feature. In the next section, the CR network architecture is discussed.

3.2. Cooperative Cognitive Radio Architecture

In a Cooperative CR network, several CRs cooperate with each other in the spectrum sensing process. A fusion center, or central node, is used to make the final decision on the presence of the primary user signal. In pattern recognition terms, each CR is equipped with a feature extraction module which extracts the necessary features from the received signal $x[k]$. The extracted features are then sent to the fusion center. At the fusion center, the features received from all the CRs are used as inputs to a classifier which makes a universal decision on the existence of the primary OFDM signal. Figure 6 shows the block diagram of the proposed Cooperative CR system which consists of d CRs.

Each CR senses the spectrum and receives a different signal, $x_i[k]$ where $i \in \{1, 2, \dots, d\}$. If a single CR is used, then d=1. The CR extracts the feature f_i from its received signal. All the extracted features, $f_1, \dots, f_d$, are sent to the fusion center. The fusion center applies the received features to the classifier which computes an output vector $\boldsymbol{q}$ which is used to decide whether the primary OFDM signal is present or not.

If the fusion center decides that the target spectrum is vacant, it allows one CR to access the spectrum for a period of time. The decision on the CR is made based on priority and previous transmission time. If the target spectrum remains vacant, the fusion center assigns another CR to access the spectrum for a period of time and the cycle continues. Consequently, each CR has to wait for its turn to be able to use the vacant spectrum since it is part of a collaborating (cooperative) CR network where a decision on the presence of the OFDM signal is made based on inputs from several CRs. A special case of the Cooperative CR network is the single CR (d = 1) which operates independently without collaboration with neighboring CRs.

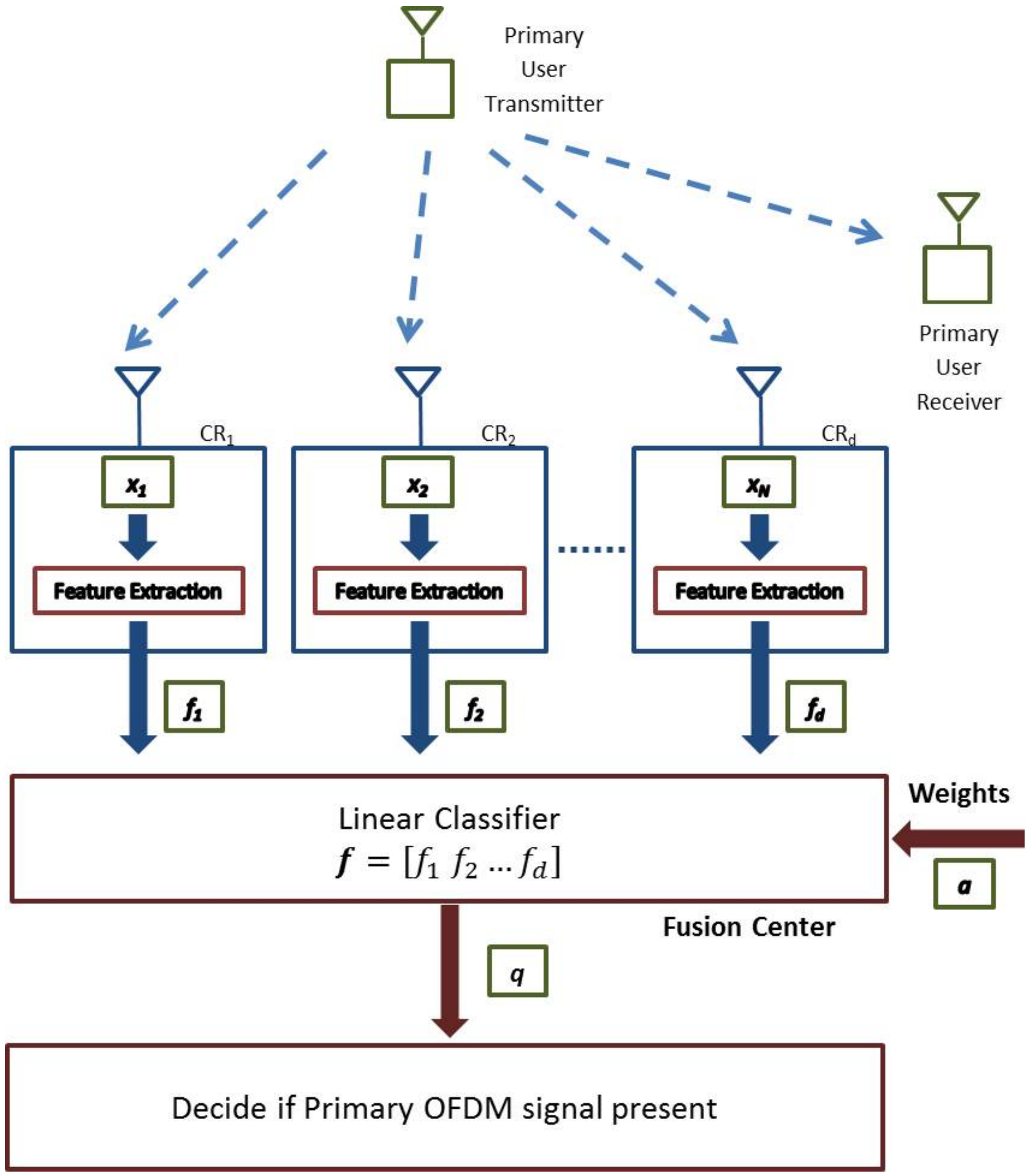

Figure 6 - Cooperative CR System based Spectrum Sensing

3.3. Feature Extraction

After sensing the spectrum, the CR extracts features from the received signal. Various properties of the signal are used as features at the CR which allow for easier classification of the received signal into the primary OFDM signal (Class 1) or noise

(Class 2). In this section, the different features which have been used in this thesis are discussed.

3.3.1. Energy.

One of the most commonly used techniques for spectrum sensing is Energy Detection. The CR senses the spectrum for a period of time and using the received signal energy decides on the presence or absence of the primary signal. However, this type of detection is unreliable in fading environments because the energy of the primary signal severely fluctuates and the signal energy becomes comparable to the noise level. This may happen due to deep fades in the channel or due to the primary signal energy being very small resulting in a very low signal-to-noise ratio (SNR).

When the spectrum sensing technique used is energy detection, the feature extraction process in the CR will compute the energy of the received signal $x[k]$ and pass it on to the linear classifier. At the CR, the time domain signal is received over an observation window of size W. Under the assumption that the signal is Ergodic, the energy of the signal is computed as defined in (7):

$$f_E = \sum_{k=0}^{W-1} |x[k]|^2, \tag{20}$$

where k is the discrete time index. The energy can also be computed in the frequency domain using the Parseval's theorem and W-point FFT as:

$$f_E = \sum_{m=0}^{W-1} |X[m]|^2, \tag{21}$$

where $X[m]$ is the received signal in the frequency domain, m is the discrete frequency index and $|X[m]|^2$ is the energy at the discrete frequency m. The extracted energy feature, f_E, is then used by the linear classifier to make a decision on the class of the received signal $x[k]$.

3.3.2. Correlation.

Energy detection does not require any prior knowledge of the type of primary user signal. This could be considered as an advantage for systems where the primary signal is unknown. However, energy based features result in an inferior performance compared to other schemes that take advantage of certain structures and properties in the OFDM signal. OFDM symbols have an inherent special property; namely the cyclic prefix, which can be utilized to detect the existence of the primary signal. The addition of a cyclic prefix at the beginning of the OFDM symbol means that the first L samples of the OFDM symbol are identical to the last L samples. This implies that the first L samples of the OFDM symbol are correlated with the last L samples even in the presence of noise. This property can be used to sense the spectrum for presence of the OFDM signal. The CR performs correlation between the first and last W samples of the cyclic prefix at the start and end of the OFDM symbol respectively and takes the maximum correlation value. The size of W should always be less than the cyclic prefix size L. If a primary OFDM signal is present, then there will be high correlation. On the other hand, if only noise is present, then any two samples of Gaussian noise are uncorrelated. The correlation feature is extracted using:

$$f_c = max|E[\boldsymbol{gh}^*]| \tag{22}$$

where $\boldsymbol{g} = [x_1, x_2, \dots., x_W]$ is a vector of first W samples of the cyclic prefix at the beginning of the received signal and g_i is the i^{th} element of $\boldsymbol{g}$ where $i \in \{1, \dots, W\}$ while $\boldsymbol{h} = [x_{N-L}, x_{N-L+1}, \dots., x_{N-L+W}]$ is a vector of the last W samples of the cyclic prefix at the end of the OFDM received signal and h_j is the j^{th} element of $\boldsymbol{h}$ where both $i, j \in \{1, \dots, W\}$. Also, $E[.]$ is the expectation operator and $max|.|$ takes the maximum value of the elements inside the argument. The correlation $E[\boldsymbol{gh}^*]$ is computed as:

$$E[\boldsymbol{gh}^*] = R_{\boldsymbol{gh}}(p) = \sum_{q=0}^{W-p-1} g_{p+q} h_q^*, \quad p = 1, \dots, 2W - 1 \tag{23}$$

where the correlation output $E[\boldsymbol{gh}^*]$ is $2W-1$ samples long.

Finally, using correlation f_c as a feature, the linear classifier can then make a decision on whether the received signal $x[k]$ belongs to class 1 or class 2.

3.3.3. Entropy.

Another feature which is used for spectrum sensing is entropy. As discussed in section 1.1.4.4, entropy is defined as the information contained in the received signal. It is expected that the entropy of the received signal when the known primary OFDM signal is present will be considerably lower than the entropy when the received signal is random noise. This is because entropy is higher for random signals and as the randomness decreases, the entropy also decreases. At low SNR values, the entropy of the received primary OFDM signal will also be high as it is severely corrupted by random noise. The entropy of both the classes will, therefore, overlap and the performance will degrade.

To compute the entropy, the CR receives the signal over an observation window of size W samples. The signal is first converted to the frequency domain using W-point FFT and entropy is then calculated as:

$$f_{ent} = -\sum_{i=1}^{K} p_i \log_2 p_i, \tag{24}$$

where K is the number of levels occupied by the received signal in the frequency domain and p_i is the histogram count of the i^{th} level.

In the next section, the process of training the linear classifier to obtain the weights for future classification of new received signal.

3.4. Classification

The extracted features are sent to a classifier which classifies them into their respective classes. In general, the classifier works in two modes: Training and Testing. In the training phase, known data from all classes is input into the classifier to obtain the weights. Once the weights are computed, the performance of the classifier is evaluated in the testing mode. Generally, classifiers can be either liner or non-linear based on the separation of classes. In this thesis, a linear classifier is used to classify any received signal into one of the classes by assuming that the data belonging to the two classes is linearly separable.

3.4.1. Classifier Training.

To classify any received signal into one of the two classes, the linear classifier is initially trained to compute the weights. Figure 7 shows a block diagram of the classifier in training mode. A large set of data belonging to both classes is used to determine the weights. First, a linear discriminant function is defined for each class which is used to separate data of a particular class from data of another class:

$$g_i = \boldsymbol{w}_i' \boldsymbol{f} + w_{i0} \,; \quad i = 1, \dots, N_C, \tag{25}$$

where,

$$\boldsymbol{f} = [f_1, \dots, f_d], \tag{26}$$

where, for the i^{th} class, g_i is the linear discriminant function, $\boldsymbol{w}_i$ is the weigth vector, w_{i0} is the bias or threshold. The vector $\boldsymbol{f}$ is the input feature vector, N_C is the number of classes (for our case, N_C=2), d is the dimension of the feature vector $\boldsymbol{f}$ and also the number of CRs in the Cooperative network and $(.)'$ is the transpose operation. Any incoming feature vector is multiplied by the weights $\boldsymbol{w}_i$ and shifted by the bias w_{i0} to get the linear discriminant function for each class. For a given feature vector $\boldsymbol{f}$ the class which gives the maximum value for g is the class of $\boldsymbol{f}$. To compute the weights for each class, the linear classifier has to be trained using training data. As a first step, the bias w_{i0} is incorporated into the weight vector $\boldsymbol{w}_i$ such that a new weight vector $\boldsymbol{a}_i$ and a new feature vector $\boldsymbol{y}$ are defined:

$$\boldsymbol{a}_i = [w_0 \; \boldsymbol{w}_i'], \tag{27}$$

and,

$$\boldsymbol{y} = [1 \quad \boldsymbol{f}] = [y_0\, y_1 \dots y_d]. \tag{28}$$

The linear discriminant function for class i can now be written as

$$g_i = \boldsymbol{a}_i'\boldsymbol{y}\,;\;\; i = 1, \dots, N_C. \tag{29}$$

The weights of the linear classifier have to be computed using a set of training data which consists of feature vectors belonging to both classes. The training data $\boldsymbol{Y}$ is defined as:

$$\boldsymbol{Y} = [\boldsymbol{y}_{11} \quad \boldsymbol{y}_{12} \quad \cdots \quad \boldsymbol{y}_{1\beta} \quad \boldsymbol{y}_{21} \quad \cdots \quad \boldsymbol{y}_{2\beta}]', \tag{30}$$

where $\boldsymbol{y}_{11} \dots \boldsymbol{y}_{1\beta}$ are feature vectors of data belonging to class 1 (OFDM signal) and $\boldsymbol{y}_{21} \dots \boldsymbol{y}_{2\beta}$ are features vectors of data belonging to class 2 (noise). The first β rows of $\boldsymbol{Y}$ correspond to data belonging to class 1 while the last β rows correspond to data from class 2. The number of elements in $\boldsymbol{Y}$ is $2\beta \times d + 1$. Furthermore, two target vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are defined for each class ($\boldsymbol{t}_1$ for class 1 and $\boldsymbol{t}_2$ for class 2). Each element of $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ is basically a linear discriminant function defined in (29). However, since the data is already known, the values of $\boldsymbol{t}_1$ are set to zero everywhere except for rows belonging to class 1. Similarly $\boldsymbol{t}_2$ is zero everywhere except the rows belonging to class 2. $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are $2\beta \times 1$ dimensional vectors. The first β elements of $\boldsymbol{t}_1$ are 1 while the last β elements of $\boldsymbol{t}_2$ are 1. The target vectors are combined into a matrix $\boldsymbol{T}$ defined as:

$$\boldsymbol{T} = [\boldsymbol{t}_1\, \boldsymbol{t}_2]. \tag{31}$$

Finally, a new weight matrix $\boldsymbol{A}$ is formed whose columns are the weight matrices for each class:

$$\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2]. \tag{32}$$

Therefore, the linear classifier problem now becomes a linear equation with $\boldsymbol{A}$ being the unknown quantity.

$$\boldsymbol{T} = \boldsymbol{Y}\boldsymbol{A}. \tag{33}$$

The weight matrix $\boldsymbol{A}$ is computed using the pseudo-inverse of $\boldsymbol{Y}$:

$$\boldsymbol{A} = (\boldsymbol{Y}'\boldsymbol{Y})^{-1}\boldsymbol{Y}\boldsymbol{T}. \tag{34}$$

The training data has to be large enough to provide a good estimate of the weight matrix $\boldsymbol{A}$. If the data from both classes is linearly separable, a linear classifier will perform well. However, if the data is not linearly separable, the linear classifier may fail to distinguish data belonging to different classes. This may happen when at low SNR values, the signal and noise have comparable levels and their feature values overlap.

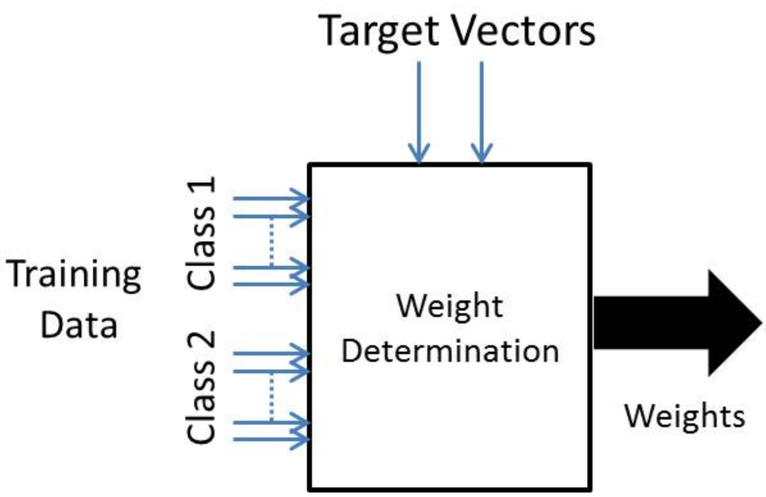

Figure 7 – Classifier Training

3.4.2. Classifier Testing.

After training the linear classifier to compute the weight matrix $\boldsymbol{A}$, the linear classifier has to be tested using test data $\boldsymbol{Y_{test}}$ to evaluate its performance. Figure 8 shows the block diagram of the testing process. Similar to the training data described in (30), the test data consists of feature vectors belonging to class 1 and class 2. The number of elements in $\boldsymbol{Y_{test}}$ is $2\gamma \times d + 1$. The first γ elements of $\boldsymbol{Y_{test}}$ belong to class 1 while the last γ elements belong to class 2. The linear classifier multiplies the test data $\boldsymbol{Y_{test}}$ with the weight matrix, $\boldsymbol{A}$, to get a two column matrix $\boldsymbol{T_{test}}$:

$$\boldsymbol{T_{test}} = \boldsymbol{Y_{test}A}. \tag{35}$$

When the number of classes, N_c = 2, the dimension of the $\boldsymbol{T_{test}}$ matrix is $2\gamma \times 2$. Each row of $\boldsymbol{T_{test}}$ corresponds to a test data vector. Ideally, the first column of $\boldsymbol{T_{test}}$ should be one for the first γ elements (corresponding to class 1) and zero for the rest while the second column of $\boldsymbol{T_{test}}$ should be zero for the first γ elements and one for the last γ elements (corresponding to class 2). However, the obtained values vary around these ideal values when novel data is fed to the classifier [22].

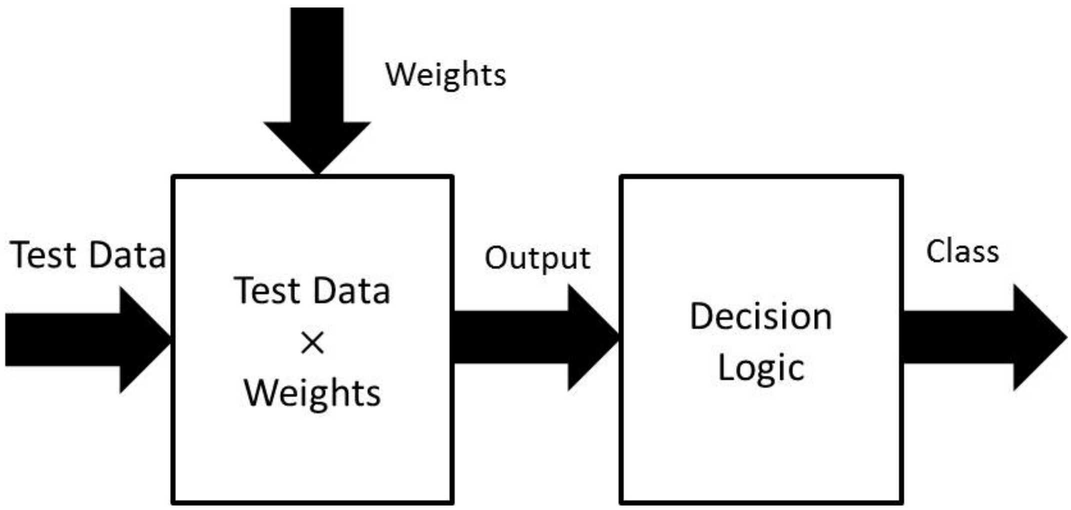

Figure 8 - Classifier Testing

The obtained $\boldsymbol{T_{test}}$ matrix is used to classify the data by comparing the values of each row. Usually, the column which contains the higher value is decided to be the class of that particular feature vector. To maintain the false alarm probability below a certain target value, a threshold is used to distinguish between the two classes. The detection probability of the classifier is then determined by comparing the classified data with the actual classes of the data. The training and threshold setting are usually done offline to reduce the complexity of the CR system [22].

The threshold is initially set to 0.5 and each row of $\boldsymbol{T_{test}}$ is compared with it. The column whose numerical value is greater than the threshold value is chosen to be the class of the received signal. Consequently, a confusion matrix is constructed as follows:

$$\boldsymbol{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}, \tag{36}$$

where c_{11} is the number of data samples classified correctly as belonging to class1, c_{12} is defined as the number of data samples incorrectly classified as class 2 while belonging to class 1, c_{21} is the number of data samples incorrectly classified as class 1 and c_{22} is the number of data samples correctly classified as class 2. The sum of all the elements of $\boldsymbol{C}$ is equal to the number of data samples used as testing data.

Table 3 shows the different performance metrics of the CR using the confusion matrix. Using the definition in Table 3, the false alarm probability is computed. If the achieved false alarm probability is greater than the desired false alarm probability, the threshold is increased by a small value, for example by 0.05, and the entire classification process is repeated. The procedure is repeated iteratively until the false alarm probability is equal to or slightly below the desired value. Finally, the detection probability is computed using the formula defined in Table 3 to evaluate the performance of the spectrum sensing scheme. In a communication system, performance is typically measured based on the amount of noise in the system and Signal-to-Noise ratio (SNR) is used as a tool to determine the signal power relative to the noise power. In this thesis, performance of the linear classifier is investigated over a wide range of SNR values and the corresponding detection probability is determined. This serves as an indicator of the performance of the spectrum sensing scheme. For each SNR value, the linear classifier has to be first trained to compute the weights and then tested using the testing data.

Once the linear classifier is trained and tested, any received signal $x[k]$ is classified by extracting the feature vector $\boldsymbol{f}$ and the weight vector $\boldsymbol{a}$. The output vector is found using:

$$\boldsymbol{q} = \boldsymbol{f}\boldsymbol{a}, \tag{37}$$

where $\boldsymbol{q}$ is a two column vector where column 1 corresponds to class 1 while column 2 corresponds to class 2. The value of each column of $\boldsymbol{q}$ is compared and the column number which has the highest number is chosen as the class of the received signal.

Table 3 - Performance Metrics Using Confusion Matrix

Performance Metric	Definition
Detection Probability	$\frac{c_{11}}{c_{11}+c_{12}+c_{21}+c_{22}}$
False Alarm Probability	$\frac{c_{21}}{c_{11}+c_{12}+c_{21}+c_{22}}$

In the next chapter, an extensive performance evaluation of the proposed schemes for spectrum sensing is presented.

Chapter 4

Performance Evaluation and Simulation Results

In this chapter, the performance of different CR systems employing various features for spectrum sensing is investigated in terms of detection and false alarm probabilities. This chapter is divided into three main sections: Energy, Correlation and Entropy detection. For each type of detection, the performance in AWGN and fading channel is investigated first for a single CR and then for a Cooperative CR network. Finally, a summary of all the results obtained in this thesis is presented.

4.1. Simulation Parameters

As a first step in evaluating the performance of the proposed spectrum sensing system, the linear classifier is initially trained using a random model for the OFDM primary user with 50% spectrum utilization by defining 2000 training data vectors, 1000 belonging to class 1 (primary OFDM signal) and 1000 to class 2 (noise only). This implies that the primary OFDM user occupies the spectrum only 50% of the time. Using this training data set, the weight vector $\boldsymbol{A}$, defined in (34) is computed. After computing the weights, a set of 1000 testing data vectors, equally divided between the two classes, are used to compute the detection probability while maintaining the false alarm probability below a fixed value of 0.1 as per the IEEE 802.22 WRAN standard. The entire process of computing the weights and determining the detection probability is repeated for different SNR values. The SNR values are varied to account for the different types of channel states and noise levels which a CR may face in physical environments. The simulations are repeated 100 times to account for most of the channel variations and the average detection probability is computed.

For illustration purposes, the Digital Video Broadcasting – Terrestrial (DVB-T) standard is used in 4k mode for the simulations. Under this condition, an OFDM signal structure with 4096 subcarriers and the cyclic prefix length $1/8$ of the number of subcarriers, i.e. 512 samples, is used. The modulation level of the M-QAM is set to M=2 (binary). The performance of the linear classifier is evaluated at different energy per bit-to-noise spectral density (E_b/N_0) values when the signal passes through an ideal channel with AWGN only and also when the signal experiences a single path flat fading channel with a low Doppler frequency of 3 Hz. Since it is assumed that CR has complete knowledge of the channel, similar results were obtained for different Doppler frequencies.

4.2. Energy Detection

In this section, the performance of energy detection as a feature to the linear classifier is investigated.

4.2.1. Single CR System.

Energy detector in a single CR system is first tested in an AWGN channel using the parameters defined in the previous section. Figure 9 shows the performance of the energy detector in an AWGN with the observation window size of W=50 samples.

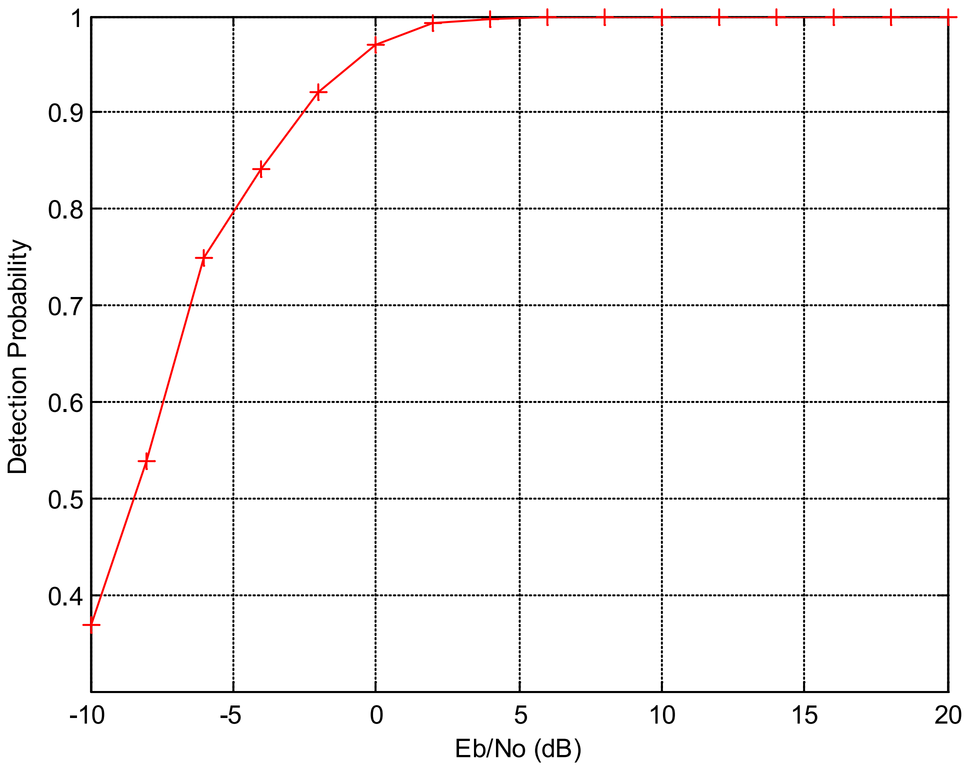

Figure 9 - Energy Detector Performance in AWGN

It can be clearly observed from the figure that the energy detector performs well in an AWGN channel. For instance, 90% detection is achieved at around E_b/N_0 = -3 dB and 100% detection requires E_b/N_0 = 4 dB or above. The false alarm probability for the classifier based energy detector is plotted in Figure 10. From the figure, it is clear that the false alarm probability is always kept below 0.1. At high E_b/N_0 values, when the detection probability reaches 100%, the false alarm probability is 0.

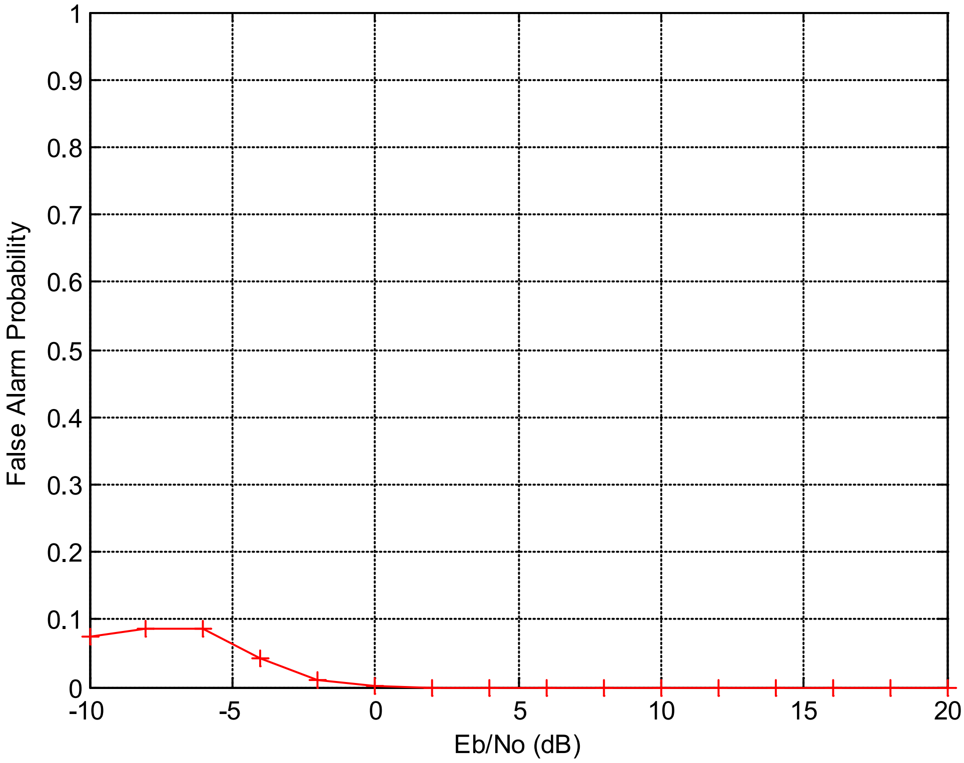

Figure 10 - False Alarm Probability of Energy Detector in AWGN

Furthermore, the performance of the energy detector is expected to improve if the window size is increased as the CR is able to accumulate more energy from the received signal. But, being able to capture more energy also implies that more noise is captured. These characteristics are investigated and the resulting performance is shown in Figure 11 where the window size is varied from 50 to 512 samples. As expected, the performance of the energy detector is directly proportional to the window size in an AWGN channel. A detection probability of 90% is achieved at E_b/N_0 = -3 dB when the observation window size is 50 samples. The same detection is achieved at E_b/N_0 = -5 dB, -7 dB and -9 dB as the observation window size is changed to 100, 200 and 512, respectively. Clearly, increasing the window size results in an improvement in the CR performance.

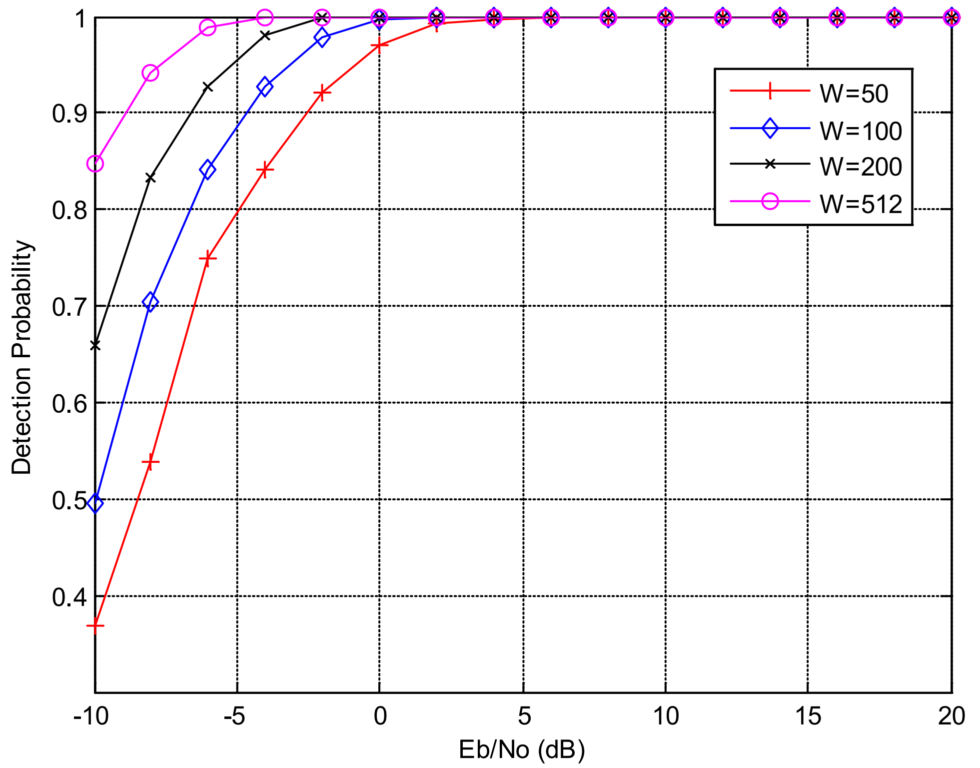

Figure 11 - Energy Detector Performance in AWGN for different Window Sizes

From the simulation results, it is concluded that the energy detector performs excellent in an AWGN channel. However, the performance is expected to degrade in a fading channel where the signal suffers variations in its amplitude and phase and also encounters deep fades. Since energy detector is a non-coherent detection scheme, it does not assume or account for any prior knowledge of the primary user signal or the channel. This means that no channel compensation is performed on the received signal at the CR. Figure 12 shows the performance of an energy detector of window size $W = 50$ in a fading channel and an AWGN channel. A significant degradation in the performance of the energy detector is observed when a fading channel model is used instead of the AWGN channel model. The detection probability remains constant around 60% and no improvement is seen as E_b/N_0 is increased. The variations in the amplitude of the transmitted signal due to fading coupled with the fact that no channel compensation is done at the CR result in a poor performance.

Next, the energy detector performance with varying window size is investigated. In AWGN, increasing the window size resulted in a significant improvement in the performance of the CR. In a fading channel, it is expected that the performance may improve but not as significantly as the improvement in the AWGN case. This is because as the window size is increased, there is a greater chance of the received signal being attenuated severely by more deep fades. The performance of the energy detector for different window sizes in fading channel is shown in Figure 13.

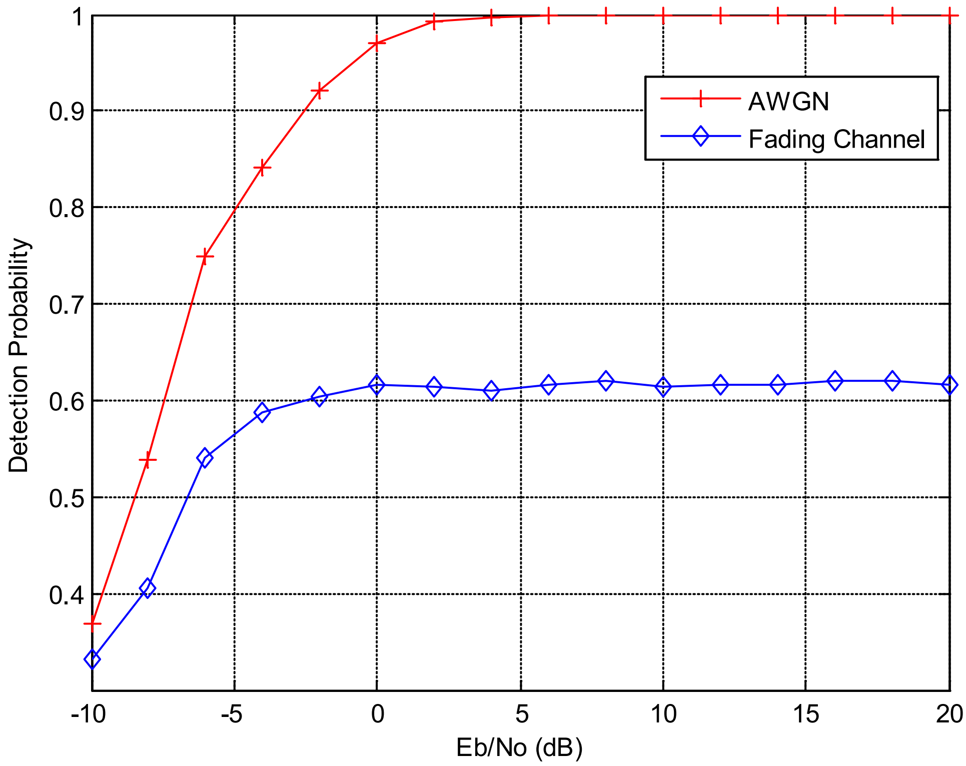

Figure 12 - Energy Detector in AWGN and Fading Channel

The performance improves by around 1 dB when the window size is increased from 50 to 100 and then 200 samples. For example, 50% detection rate is achieved at E_b/N_0 = -3 dB, -2 dB and -1 dB for observation windows of size 50, 100 and 200 samples, respectively. As the window size is increased from 50 to 200 samples, the number of deep fades in the received signal also increases and the CR does not accumulate enough energy to be able to properly distinguish the two classes. However, when the window size is increased further, the performance improves significantly and 50% detection is achieved at E_b/N_0 = -10 dB. This implies that there is at least a 7 dB improvement in performance when the window size is 512 samples compared to the performance when the window size is 200 samples. This is because when the window size is increased above 200 samples, to 512 samples in this case, the CR accumulates sufficient energy to properly estimate the class of the received signal even at low SNR. Nevertheless, the best detection achieved is around 70% which is significantly lower than the 100% detection achieved in an AWGN channel. In conclusion, despite an improvement in the overall performance when the window size is increased, the CR still suffers from severe degradation at high SNR values due to the deep fades in the channel and the non-coherent nature of the detection scheme.

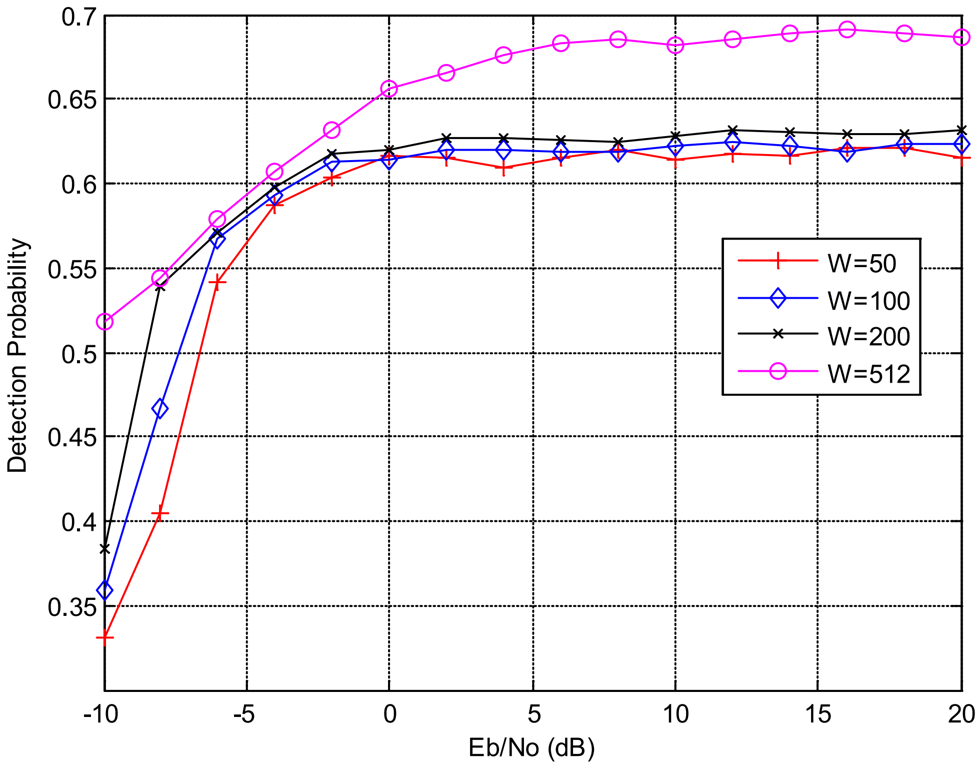

Figure 13 - Energy Detector in Fading Channel for Different Window Sizes

4.2.2. Cooperative CR Network.

In this section, energy detector is used in a Cooperative CR Network of linear classifiers as a feature to classify any received signal into the primary OFDM signal (class 1) or noise (class 2).

Similar to the previous section, the system is first simulated in an AWGN channel. A 5 CR network is used initially with an observation window of 50 samples and compared with the performance obtained using a single CR with a window size of 50. Figure 14 shows this comparison. From the figure, it can be deduced that a cooperative CR network of 5 users performs better than a single CR system. The improvement is clearly visible at low SNR values. For instance, in a single CR system, 90% detection is achieved at E_b/N_0 = -2 dB while the same detection requires only E_b/N_0 = -7 dB in a 5 user CR network which is an improvement in performance by 5 dB. Besides, both systems provide 100% detection for all SNR values above 5 dB.

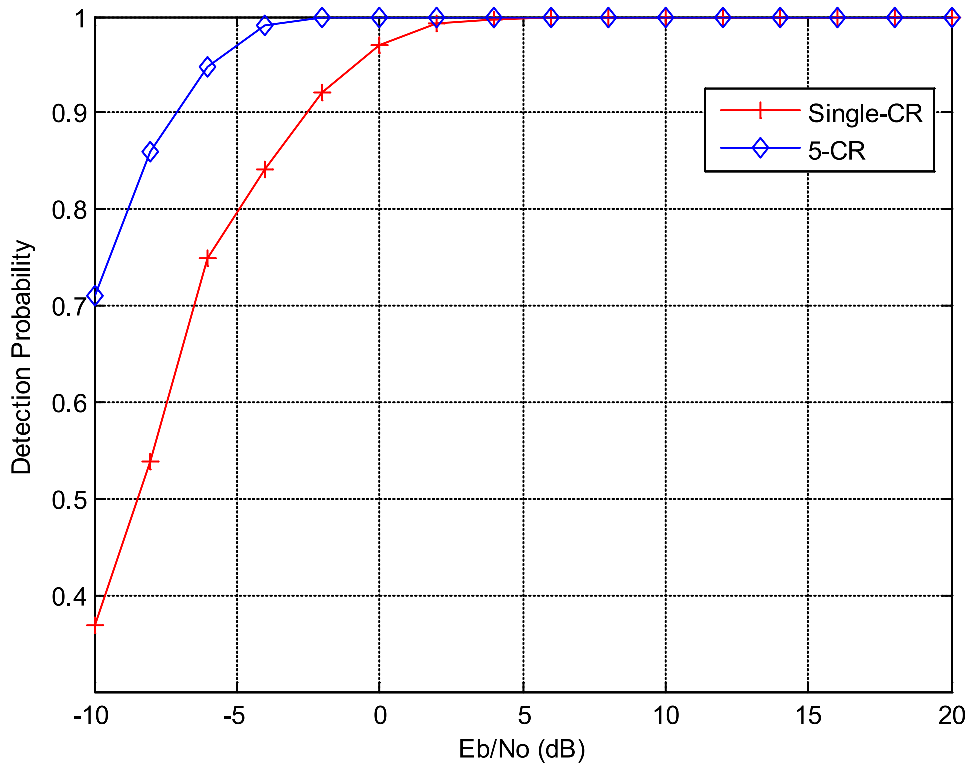

Figure 14 – Single-CR and 5-CR Energy Detector in AWGN

Similar to the single CR case, the observation window size is varied and the changes in the performance are observed. Keeping the number of cooperative CRs to 5, the performance for different window sizes in AWGN channel is shown in Figure 15. Clearly, increasing the window size improves the performance significantly. From Figure 15, the performance improves by around 2 dB for a fixed detection probability at E_b/N_0 values below 0 dB. For instance, 100% detection is achieved at E_b/N_0 = -2 dB, -4 dB, -6 dB and -8 dB for observation window size samples of 50, 100, 200 and 512, respectively. A similar improvement in performance is observed for 7 user Cooperative CRs and 3 user Cooperative CRs.

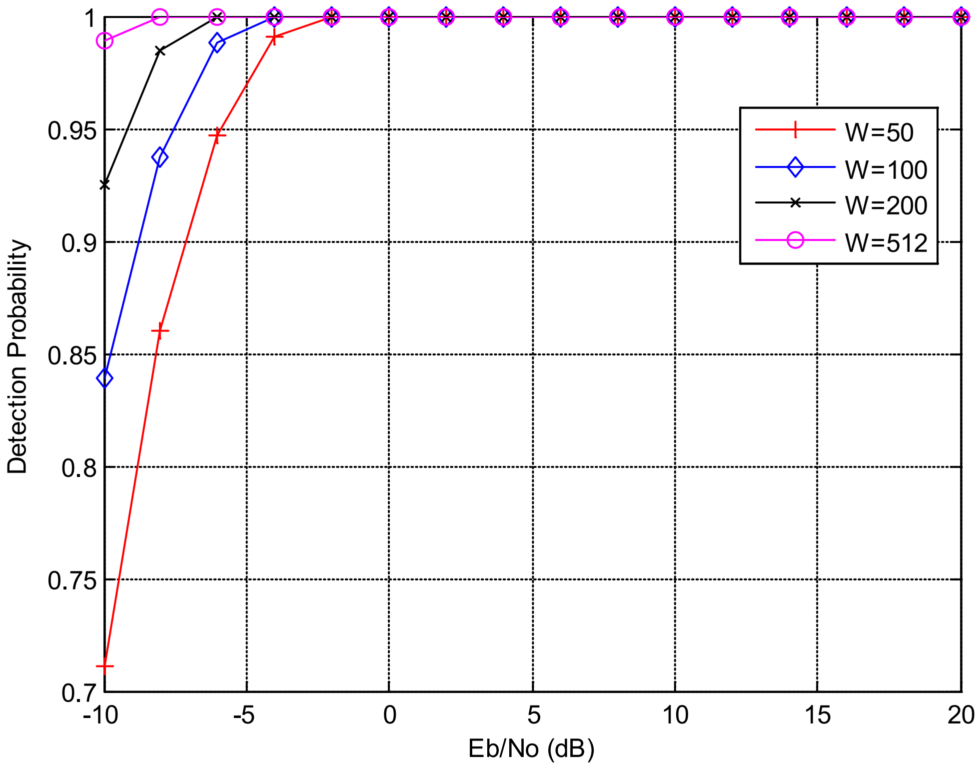

Figure 15 - 5-CR Energy Detector in AWGN for different window sizes

Intuitively, changing the number of CRs in the network should affect the performance of the linear classifier. The number of CRs is changed from 1 to 7 and the performance obtained is shown in Figure 16 for an observation window of 50 samples. It is observed that increasing the number of cooperating CRs in the network results in an improvement in the overall performance of the CR system. However, the improvement seen when the cooperative CRs is increased from 1 to 3 is significantly larger compared to the later increments.

In the next simulation, the performance of a 5 user cooperative CR is investigated in the fading channel environment as shown in Figure 17 and compared with the performance of a single CR system for an observation window of 50 samples. The performance is improved significantly for 5 user Cooperative CRs compared to a single CR system. Unlike the single CR case, the detection probability reaches approximately 90% at E_b/N_0 = 6 dB and above. Figure 18 shows the effect of varying the window size on the performance for a 5 user CR network. From the figure, it is deduced that increasing the window size from 50 to 200 results in a similar overall performance. However, when the window size is increased to 512 samples, the performance improves by around 5 dB. For instance, with windows of 50, 100 and 200 samples, 90% detection is achieved at about 4-6 dB but when the window size is increased to 512 samples, the same detection rate is achieved at E_b/N_0 = 0 dB.

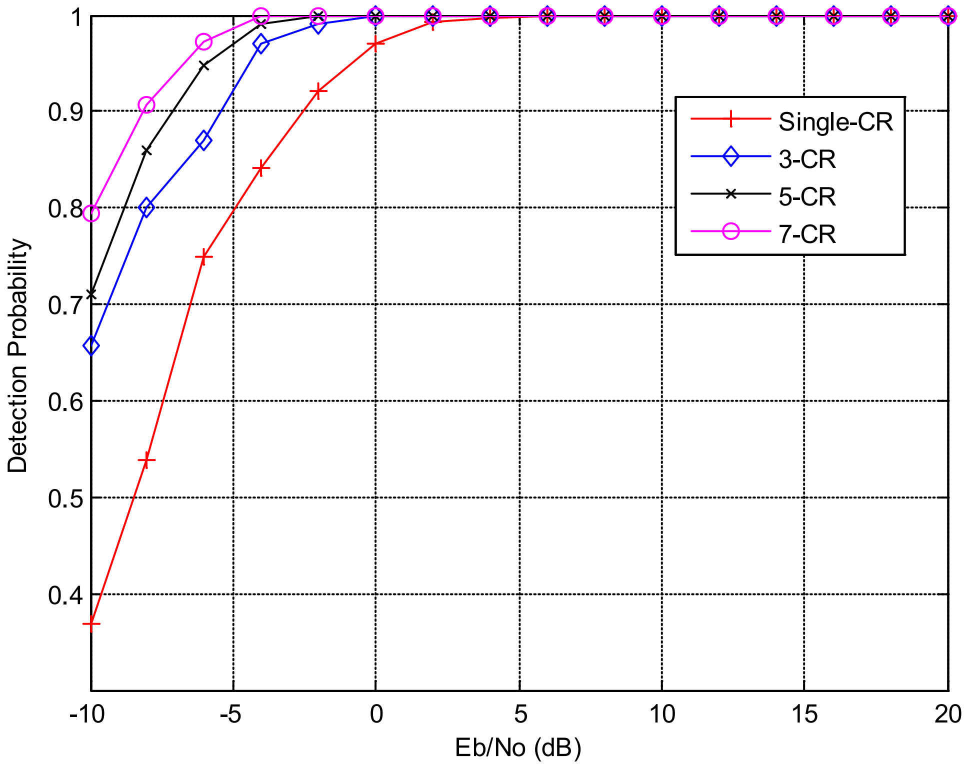

Figure 16 – Cooperative Energy Detector in AWGN

Figure 19 illustrates the performance obtained when the number of users is the CR network is varied under fading conditions. When the number of CR users in the network is changed, the performance improves significantly. For instance, the single CR never reaches beyond 60% detection which requires E_b/N_0 = -4 dB or above. As the number of CRs is increased, the maximum detection rate also increases. The 3-CR reaches its maximum detection rate of 80% at E_b/N_0 = 2 dB. The maximum detection rate of the 5-CR network is 90% which is achieved at around E_b/N_0 = 6 dB. Finally, the 7-CR network has a maximum detection rate of 95% which requires E_b/N_0 = 4 dB.

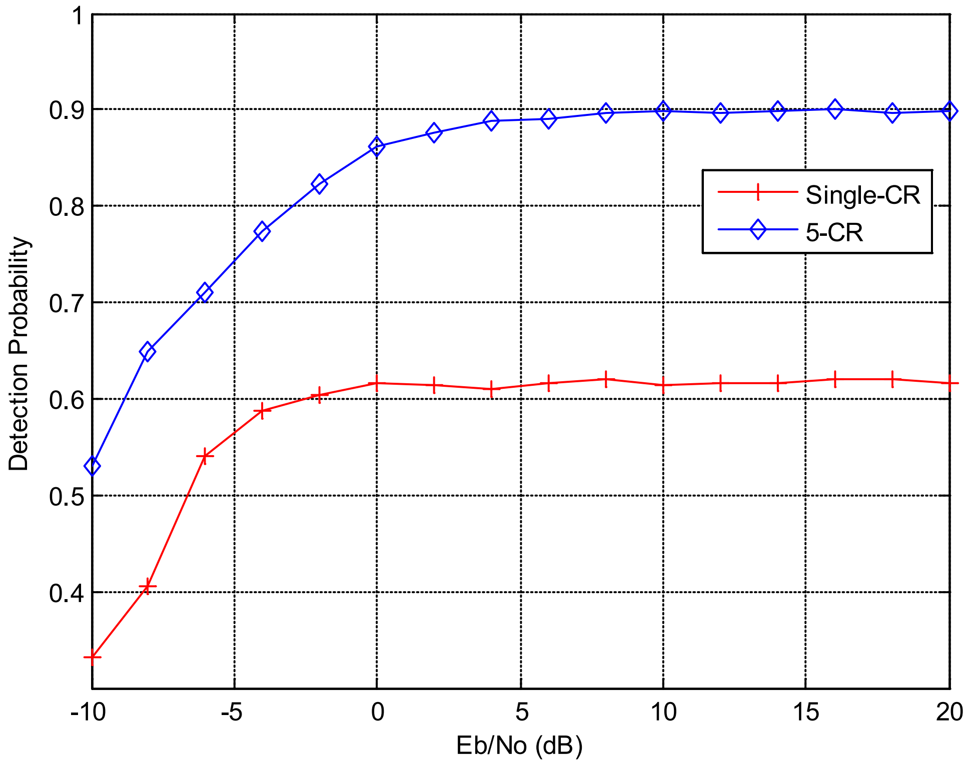

Figure 17 – Single-CR and 5-CR Energy Detector in Fading Channel

From the results and discussions, it can be concluded that increasing the number of CRs in the network and increasing the window size improves the performance of the energy detector in the AWGN channel. However, in a fading channel, the poor performance of the single CR system improves slightly by increasing the window size. The performance also improves when a cooperating CR system is implemented instead of a single CR system. When the number of CRs in the network is increased, the performance of the system also improves. The severe degradation in the performance of the energy detector in a fading channel is observed due to the non-coherent detection scheme which does not compensate for the channel at the CR.

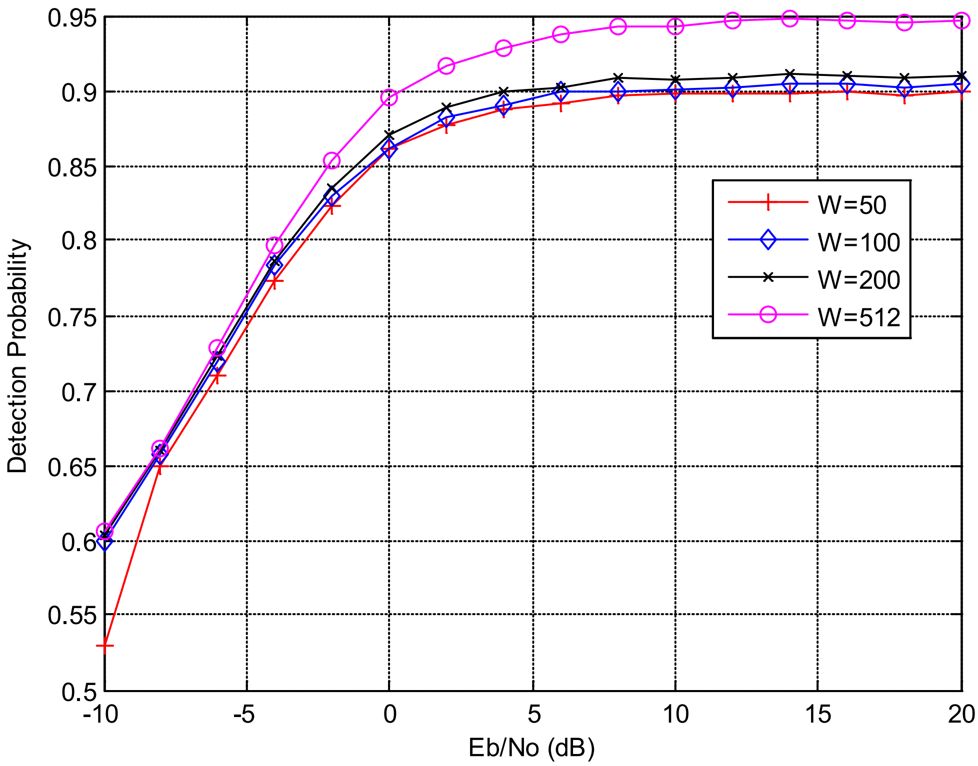

Figure 18 - 5-CR Energy Detector in Fading Channel for Different Window Sizes

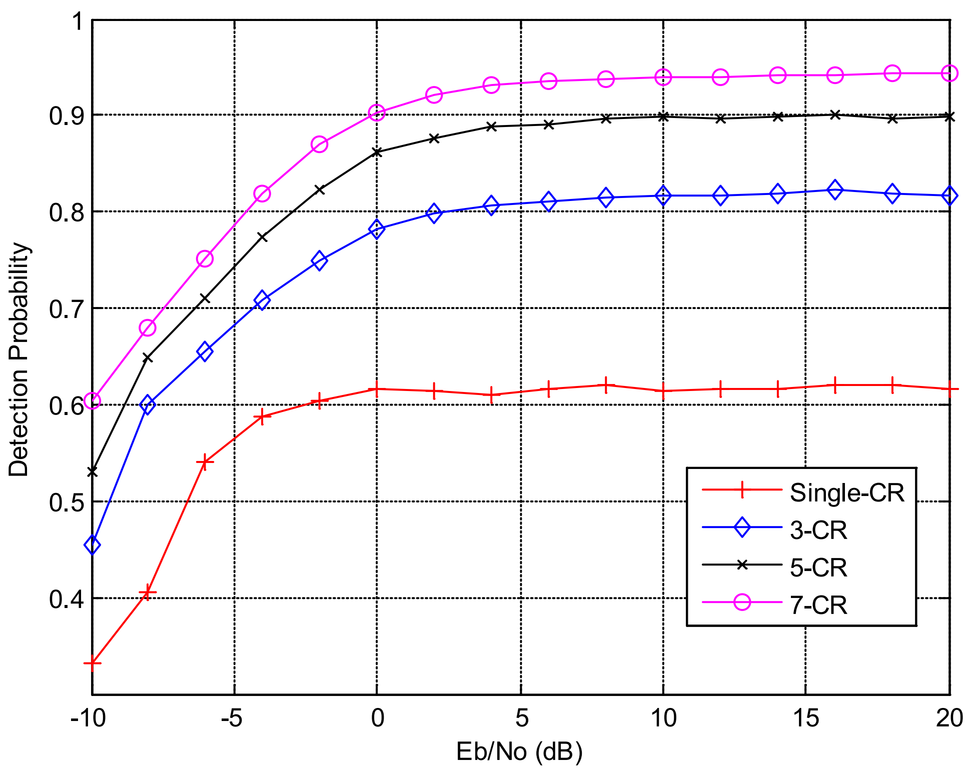

Figure 19 - Cooperative Energy Detector in Fading Channel

4.3. Correlation Detection

This section investigates the performance of the CR when correlation is used as a feature into the linear classifier. Correlation detector utilizes the fact that the last L samples of the OFDM signal are similar to the first L samples due to the insertion of the cyclic prefix. The correlation value is expected to be high when the OFDM signal is present in the spectrum and low when only the uncorrelated noise is received at the CR.

4.3.1. Single CR System.

As a first step, the performance of the correlation detector in a single CR system is investigated. The plot of the detection probability against E_b/N_0 for an observation window of 50 samples in an AWGN channel is shown in Figure 20.

The correlation detector performs well in an AWGN channel and 90% detection is achieved at E_b/N_0 = -2 dB and 100% detection requires E_b/N_0 = 2 dB or above. From the results, it can be concluded that in an AWGN channel, the correlation property of the OFDM signal is preserved and it can be used as a good feature for the linear classifier. When the window size is increased, an improvement in the performance is observed which is shown in Figure 21. As the window size increases from 50 to 100 samples, the performance improves by 2 dB. A similar improvement is seen when the window size is increased further from 100 to 200 samples and then from 200 to 512 samples. While a window size of 50 samples requires E_b/N_0 = -2 dB to achieve 90% detection probability, the same performance is achieved at E_b/N_0 = -3 dB, -5 dB and -7 dB for windows of size 100, 200 and 512 samples, respectively. It is important to note that the window size should not exceed the size of the cyclic prefix.

When the signal passes through a fading channel which causes variations in amplitude, some degradation is expected in the performance of the CR. Correlation is a coherent detection scheme which assumes complete synchronization with the primary user and complete knowledge of the channel. Therefore, the received signal at the CR is divided by the channel coefficients to compensate for the effects of the channel. After compensating for the channel, the correlation parameter is computed and sent to the linear classifier for classification. Figure 22 shows the performance of the system in a fading channel when an observation window size of 50 samples is used at the CR. As expected, there is some degradation in the performance of the system in a fading channel. An additional 5 dB is required to achieve 100% detection in a fading channel compared with the AWGN case. The performance degrades severely at low E_b/N_0 values with detection reaching close to 0% at very small SNR values. Also, 90% detection is achieved at about E_b/N_0 = 3 dB which is around 5 dB higher than required in an AWGN channel.

Figure 23 shows the performance of the correlation detector for different window sizes in the fading channel. As the window size is increased, the detector performance improves slightly. For instance, using an observation window of size 50 requires E_b/N_0 = 3 dB to achieve 90% detection while the same detection probability is achieved at E_b/N_0 = 2 dB when the window size is increased to 100. However, no improvement in the performance is seen when the window size is increased beyond 200.

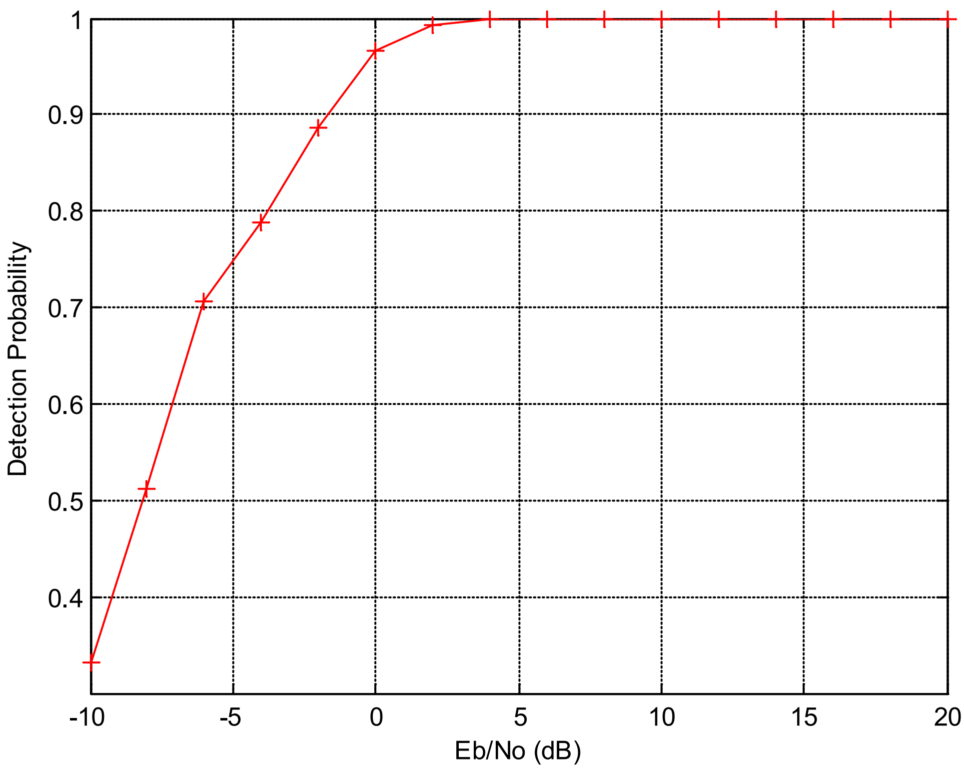

Figure 20 - Correlation Detector in AWGN

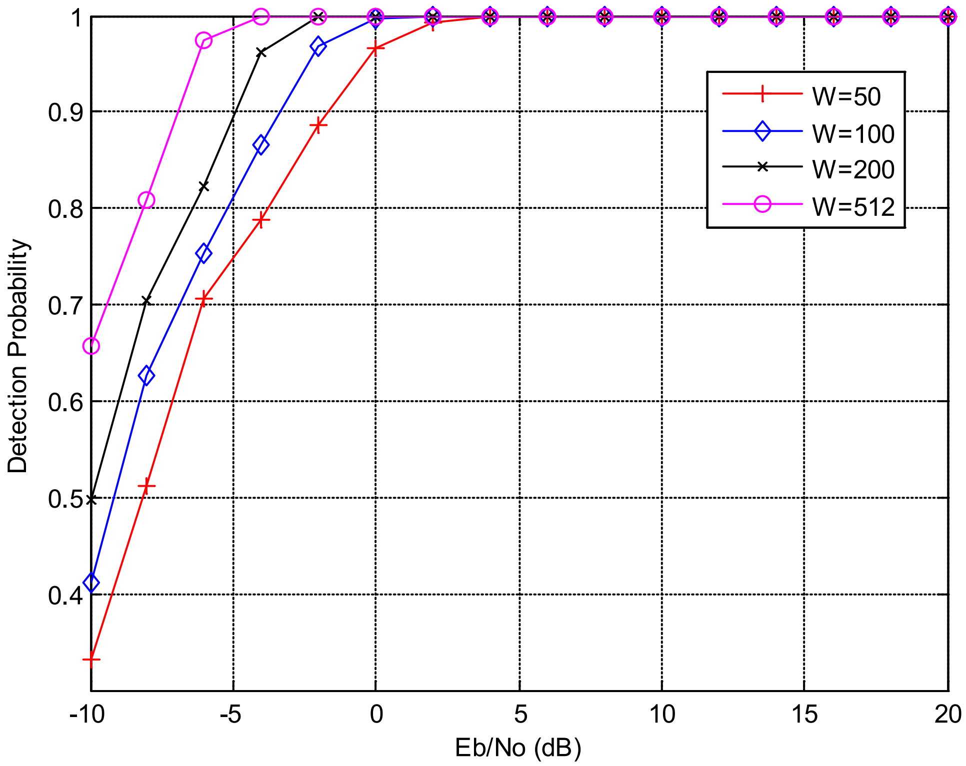

Figure 21 - Correlation Detector in AWGN for different window sizes

4.3.2. Cooperative CR Network.

In this section, the performance of the correlation detector in a Cooperative CR network is evaluated. The system is first simulated using 5 CRs in the network in an AWGN channel with an observation window size of 50 samples. Figure 24 illustrates the performance compared with the corresponding single CR case. From the figure, it can be deduced that increasing the number of users from one to 5 in a CR system results in an improvement in the performance by 5 dB. For example, 90% detection in a single CR system is achieved at E_b/N_0 = -2 dB while the same detection is achieved in a 5-CR network at E_b/N_0 = -7 dB.

A further improvement in performance is seen when the observation window size is increased for the same 5 user cooperative CR network as shown in Figure 25. The 5 user Cooperative CR network performance improves significantly as the window size is increased from 50 to 512. For example, 90% detection is achieved at E_b/N_0 = -7 dB when the window size is 50 samples while E_b/N_0 = -8 dB, -9 dB and -10 dB is required to achieved the same detection rate for windows of size 100, 200 and 512 samples, respectively. The cooperative CR network is further investigated by fixing the window size to 50 samples and changing the number of CRs in the network. The obtained performance is shown in Figure 26. As the number of CRs is increased, there is a small improvement in the performance of the CR. However, improvement in performance decreases as the number of CRs increases.

In the next simulation, the performance of a 5 user CR network is investigated in a fading channel with an observation window of size 50. The performance is shown in Figure 27. The performance of the 5-CR is better compared to the single CR system. The 5-CR network achieves 90% detection at around E_b/N_0 = 2 dB while a single CR requires E_b/N_0 = 4 dB to achieve the same detection rate.

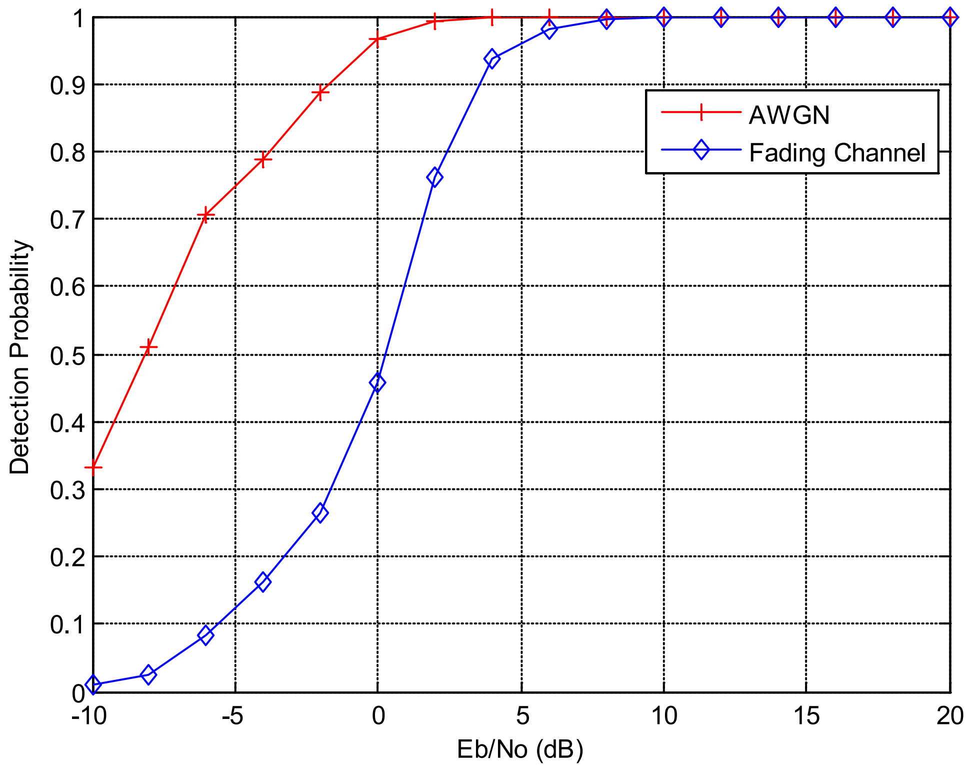

Figure 22 - Correlation Detector in AWGN and Fading Channel

Increasing the window size in the fading channel does not result in a great improvement in the performance of the CR network as can be seen from Figure 28. There is an improvement of around 1 dB when the window size is increased from 50 to 100 samples. A similar improvement is seen when the window size is further increased to 200 samples. However, increasing the window size beyond 200 samples did not result in any significant improvement in the performance. This result is similar to the single-CR fading channel simulations where the performance also did not change when the window size was increased above 200 samples.

Finally, the number of CRs in the network is changed while the observation window size is fixed at 50 samples and the performance of the correlation detector CR system is evaluated under fading. Figure 29 illustrates the performance achieved. From the figure, it is deduced that there is an improvement of around 2 dB when the number of CRs in the network is increased from 1 to 3. A further 1 dB improvement is seen when the CRs are increased from 3 to 5. No significant improvement is seen when the number of CRs in the network is increased above 5. This implies that considerable improvement in performance is seen only when the number of CR users in the network is increased above one. Once a multi-user network is in place, the performance remains quite similar with only small improvements when the number of CRs is increased.

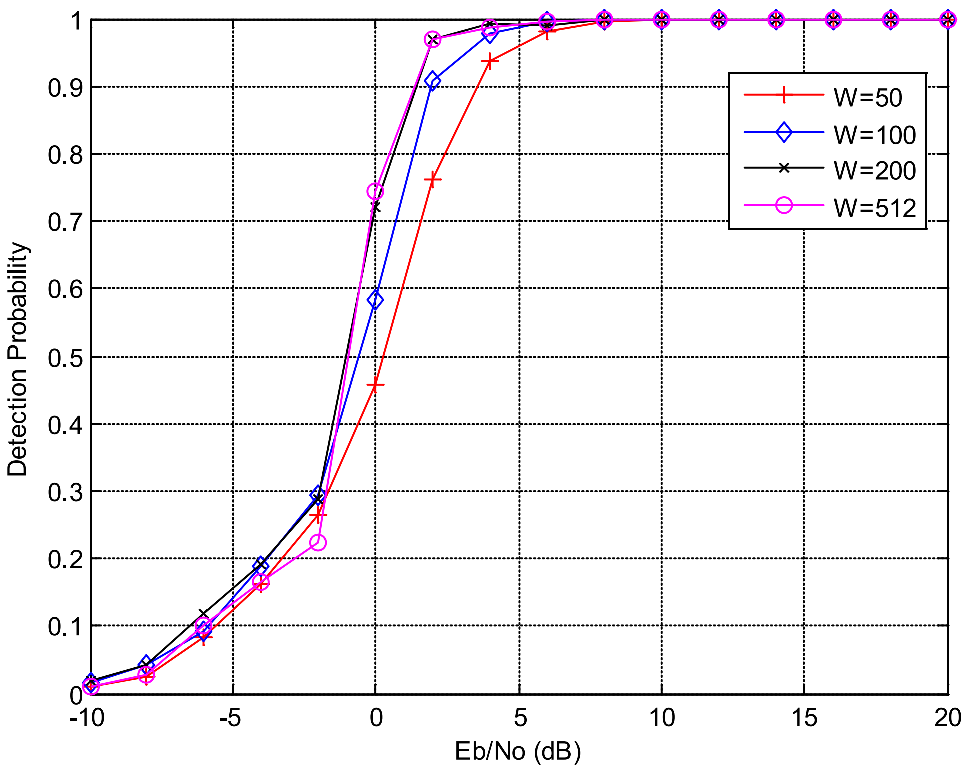

Figure 23 - Correlation Detector in Fading Channel for different window sizes

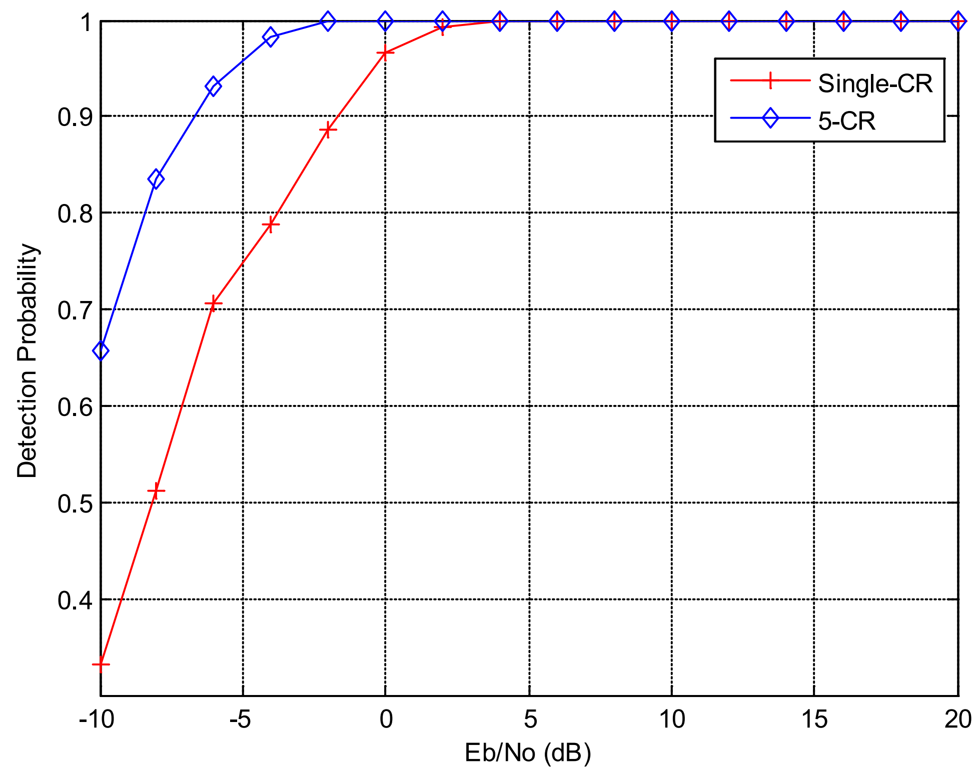

Figure 24 – Single-CR and 5-CR Correlation Detector in AWGN

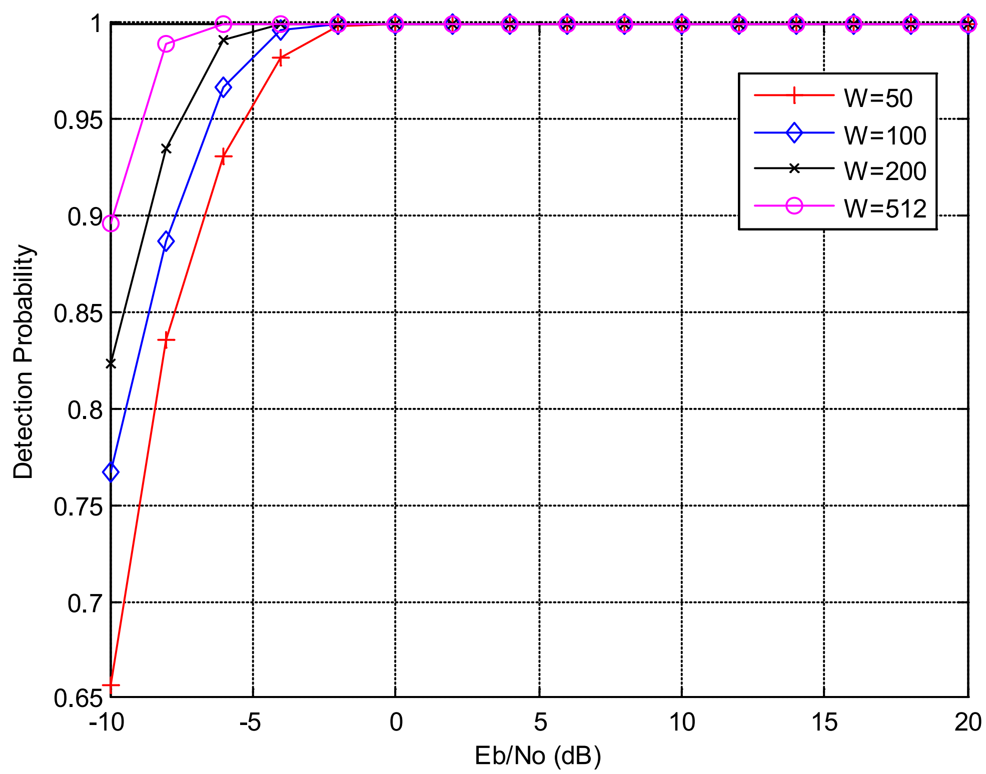

Figure 25 - 5-CR Correlation Detector in AWGN for different window sizes

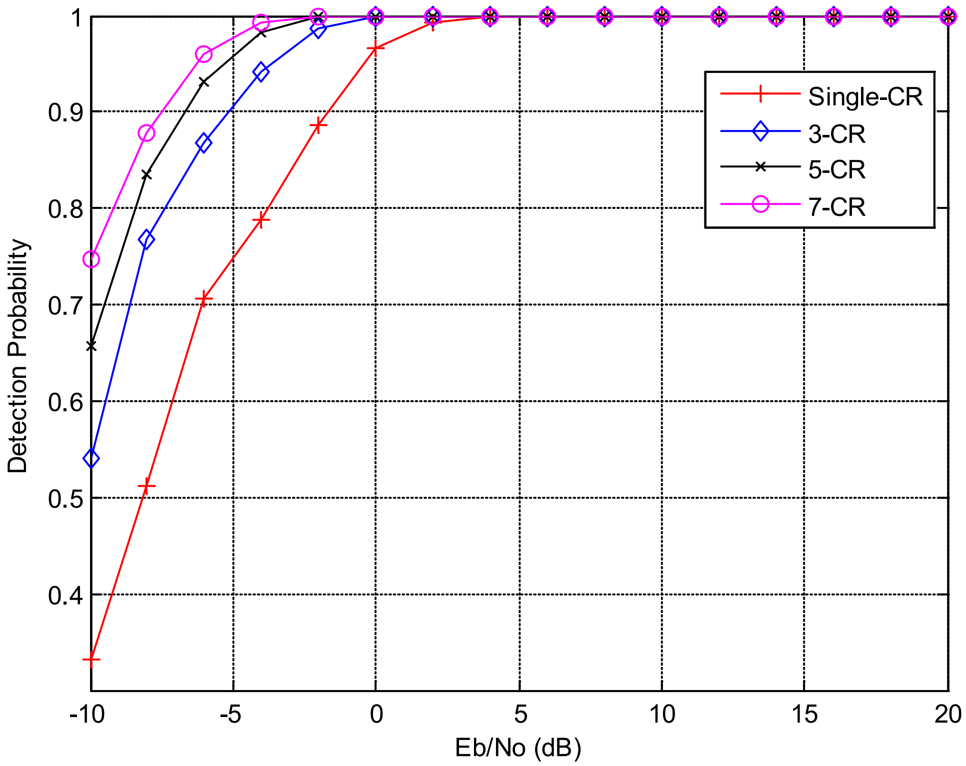

Figure 26 –Cooperative Correlation Detector in AWGN

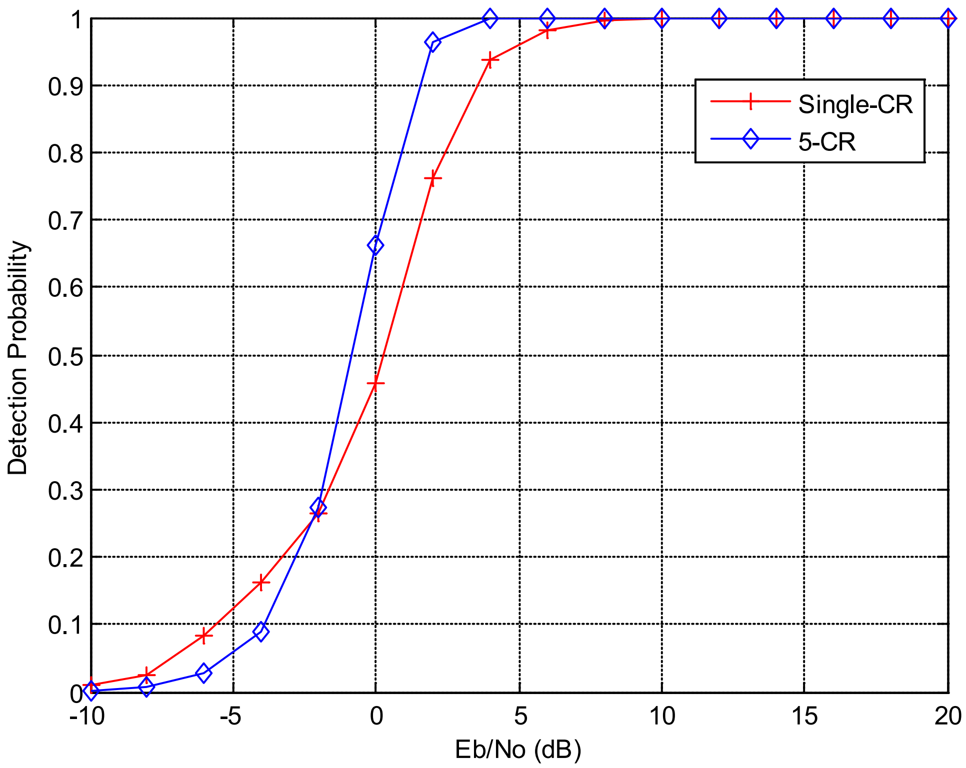

Figure 27 - Single-CR and 5-CR Correlation Detector in Fading Channel

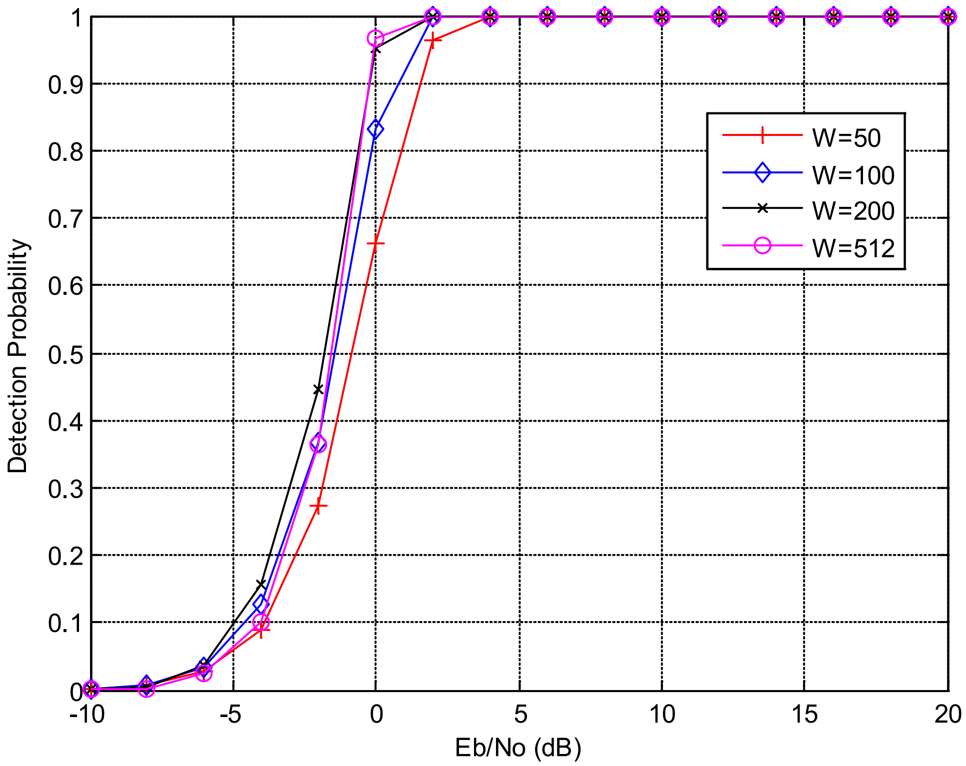

Figure 28 – 5-CR Cooperative Correlation Detector for different Window Sizes

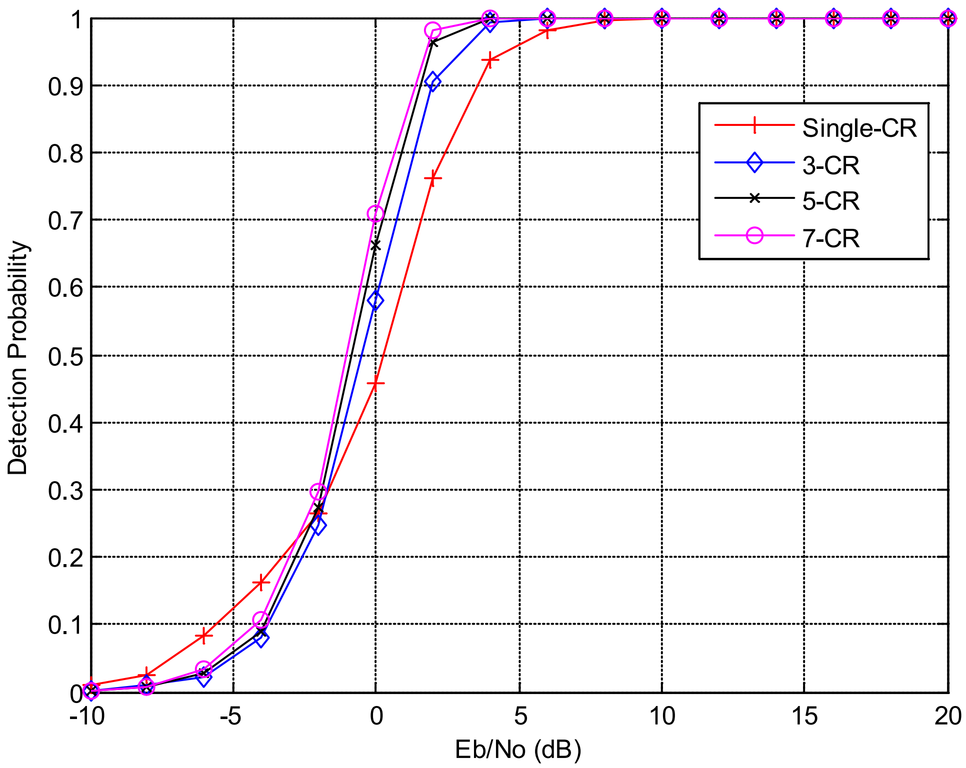

Figure 29 - Cooperative Correlation Detector in Fading Channel

In conclusion, increasing the window size and the number of CRs in the network has a greater impact in the improvement of the overall performance of the system in the AWGN channel as compared to the fading channel. In both cases, the correlation detector still performs relatively better compared to the energy detector. Finally, having more CRs in the network increases the complexity and processing time of the system and should only be used if the gain in the performance is significant.

4.4. Entropy Detection

In this section, entropy detection is used at the CR to detect the presence of the primary OFDM signal and the performance is evaluated. At the CR, the received signal samples are first converted to the frequency domain and the entropy of the received signal is calculated using the histogram method.

At high E_b/N_0 the entropy of the OFDM signal is expected to be significantly lower than the entropy of random noise. Figure 30 shows an example of the histograms of the entropy of the testing data (1000 belonging to class 1 and 1000 belonging to class 2) at E_b/N_0 = 10 dB. Clearly, when entropy is used as a feature, the data from both classes is linearly separable and a linear classifier is expected to perform well. However, at low E_b/N_0 values, the entropy of the OFDM is expected to increases since it is severely corrupted by random noise. In such a case, the entropies of the primary OFDM signal and noise overlap and the performance of the classifier is expected to fail. Figure 31 shows the histograms of the entropy of the testing data at E_b/N_0 = -5 dB.

4.4.1. Single CR System.

The entropy detector is first evaluated in an AWGN channel with observation window of size 50. Figure 32 shows the system performance. The performance is poor at low E_b/N_0 values but improves gradually as E_b/N_0 is increased until 100% detection is achieved at around 6 dB. For example, 90% detection is achieved at E_b/N_0 = 2 dB. Compared with the energy and correlation detector, the performance of entropy detector is clearly inferior. Figure 33 shows the change in performance when the window size is increased. As the window size is increased from 50 to 100, the performance of the system improves by around 3 dB. For instance, when the window size is 50 samples, 90% detection is achieved at E_b/N_0 = -3 dB while E_b/N_0 = 0 dB is required to achieve the same detection probability when the window size is increased to 100 samples. However, as the window size is increased further, the entropy detector behaves in a strange way at low E_b/N_0 values when the signal energy is comparable to the noise floor. This is probably because at low SNR, noise severely corrupts the signal and the entropy values of noise and the primary signal are similar thereby confusing the classifier. A similar observation about the entropy detector was made in [21]. When E_b/N_0 value is increased, the performance once again becomes normal as the signal energy becomes much higher than the noise power.

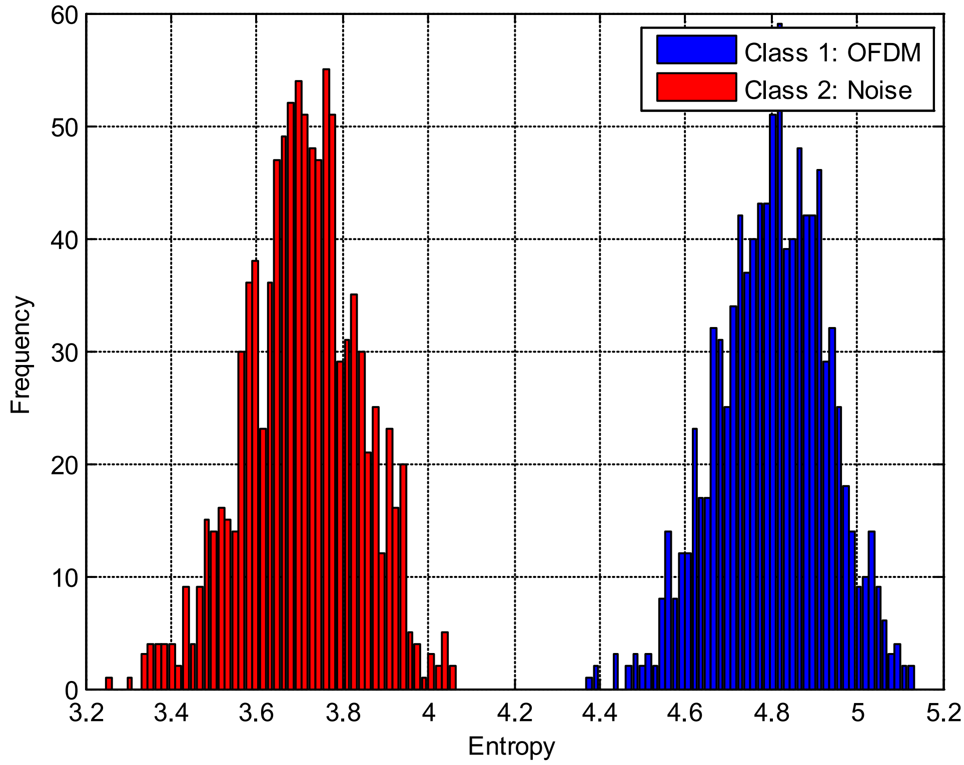

Figure 30 - Entropy Histograms at Eb/N_0 = 10 dB

The entropy detector is then tested in a fading channel and the performance for an observation window size of 50 samples is plotted in Figure 34. Entropy detector is also a coherent detection scheme and the CR compensates for the channel before computing the entropy of the received signal. The performance degrades by around 6 dB in a fading channel compared to the AWGN channel. In the AWGN channel, 90%

detection is achieved at E_b/N_0 = 2 dB while the same detection is achieved at E_b/N_0 = 8 dB in the fading channel. The performance, at low SNR, is extremely poor and the detection probability is 0% for E_b/N_0 = -4 dB and less. This is expected as the entropy detector does not work well when the signal and noise energy are comparable and is it results in an overlap in the entropy values of noise and primary signal. In addition to the corruption by noise, the signal also suffers from deep fades which result in further degradation of the performance at low E_b/N_0 values. Increasing the window size also does not solve the problem as can be seen in Figure 35. The best performance is still obtained when the window size is 50 samples. As the window size is increased, the performance at lower E_b/N_0 values worsens. This is an unexpected result for the fading channel with no clear justification as the performance had improved when the window size was increased previously for an AWGN channel.

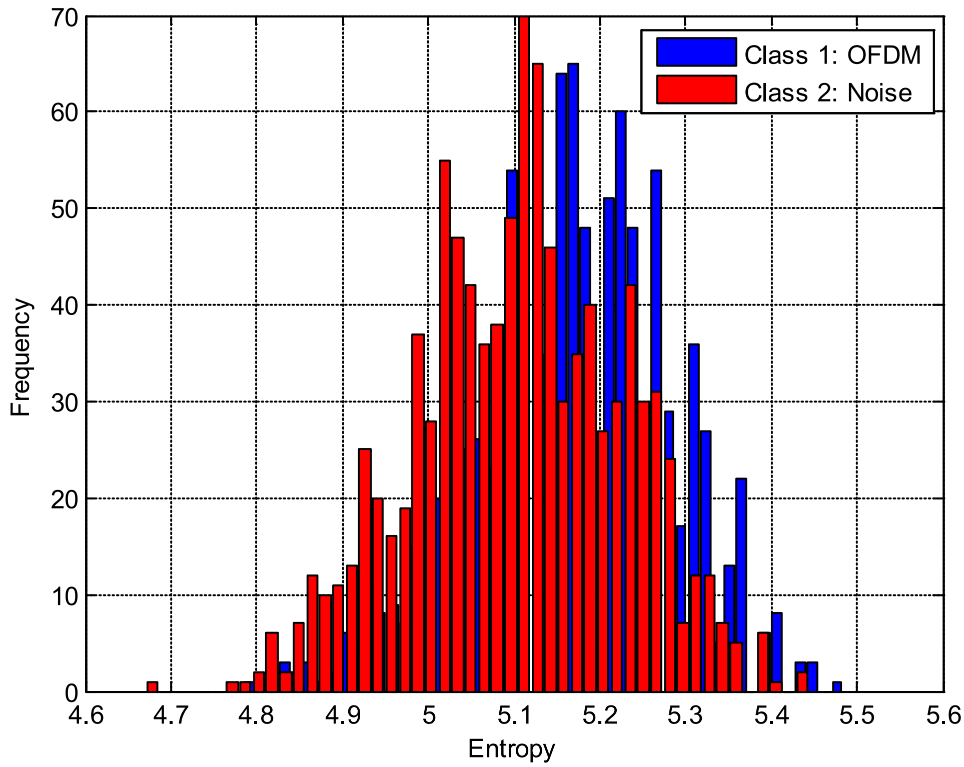

Figure 31 - Entropy Histograms at Eb/N_0 = -5 dB

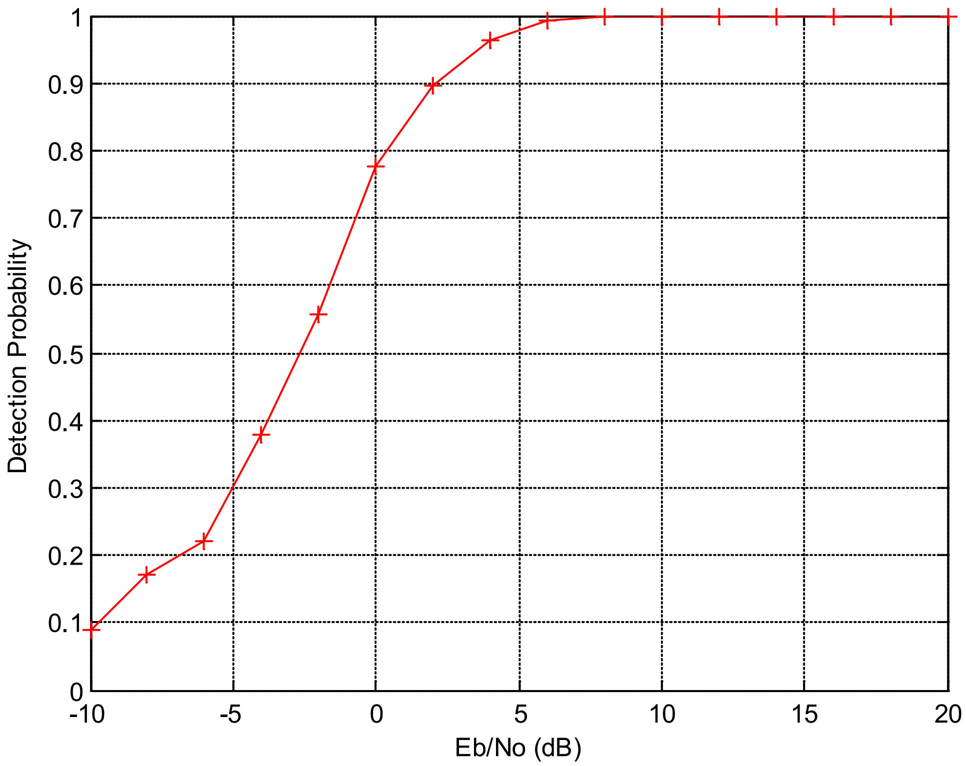

Figure 32 - Entropy Detector in AWGN

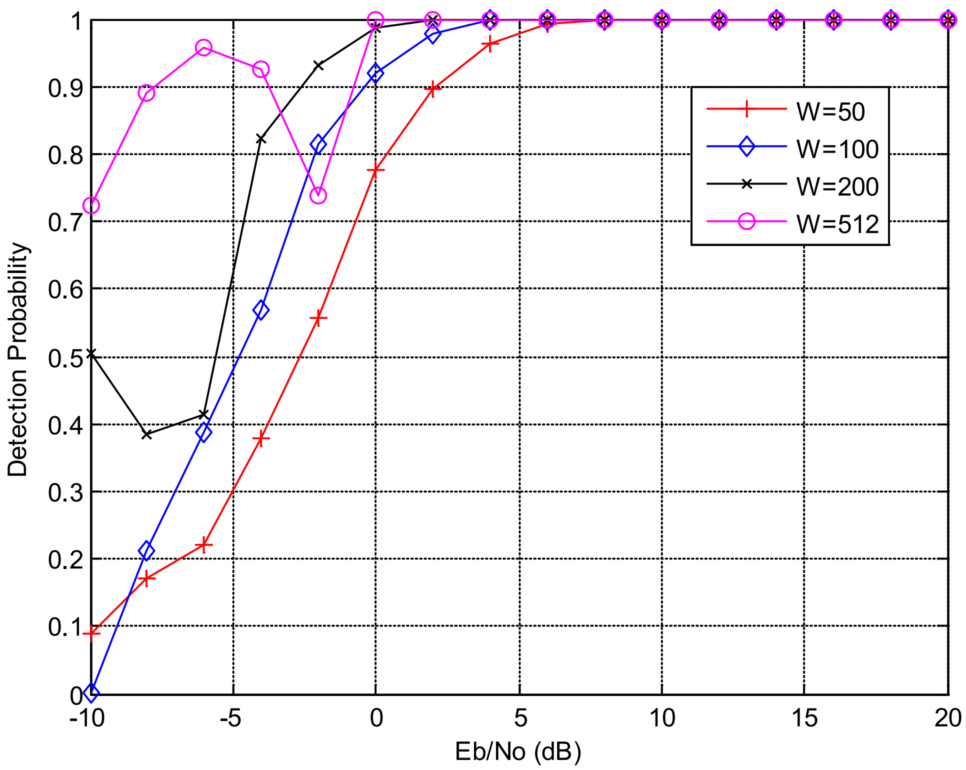

Figure 33 - Entropy Detector in AWGN for different window sizes

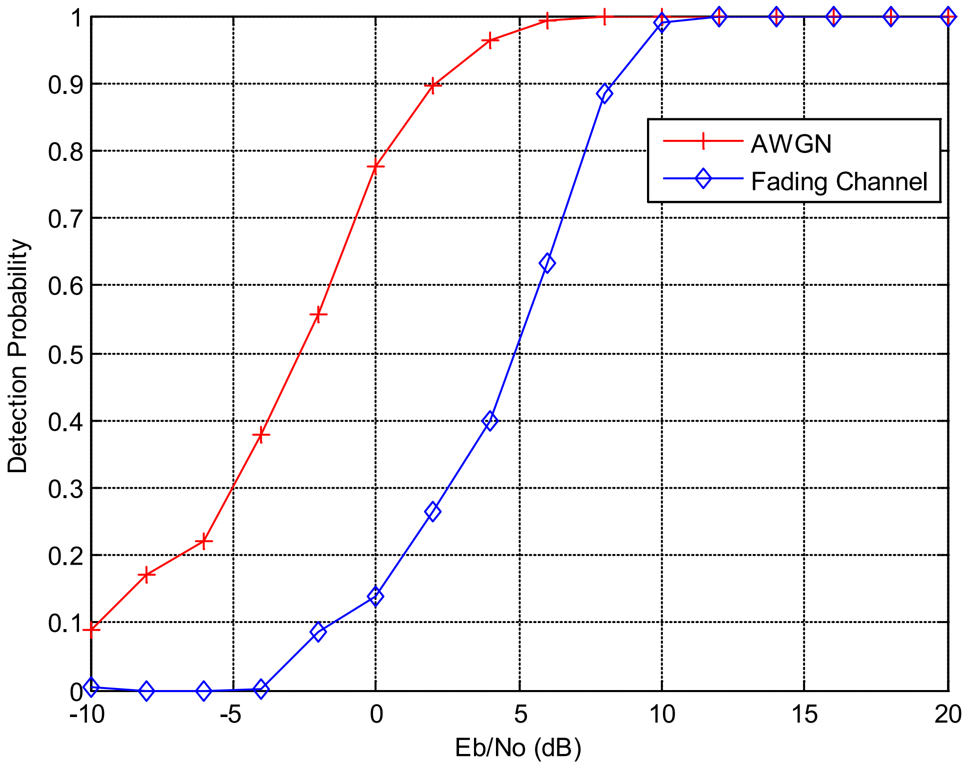

Figure 34 - Entropy Detector in AWGN and Fading Channel

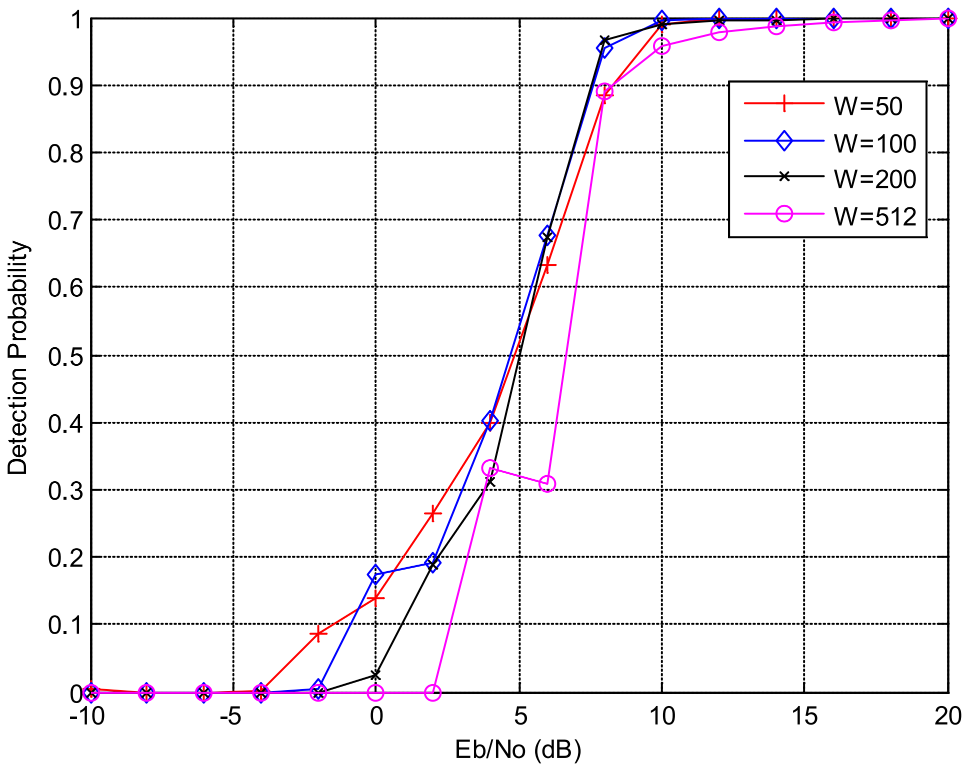

Figure 35 - Entropy Detector in Fading Channel for different window sizes

4.4.2. Cooperative CR Network.

In this section, the entropy detector performance is evaluated in a Cooperative CR network. A 5-CR network is simulated and its performance is compared with the single CR system in the AWGN channel for an observation window of 50 samples and the results are plotted in Figure 36. The figure shows an improvement of around 4 dB when the number of CRs in the network is increased from 1 to 5. For a single CR system, 90% detection is achieved at E_b/N_0 = 2 dB while the 5-CR network requires E_b/N_0 = -2 dB to achieve the same detection.

Next, Figure 37 shows the detection rate when the observation window size is varied for the 5-CR network. When the window size is increased from 50 to 100, an improvement of around 2 dB is seen. However, increasing the window size beyond 100 samples results in an unexpected performance at low SNR values. At low SNR values, when the signal energy and noise power are comparable, the entropies of both the signals, OFDM and noise, overlaps and the linear classifier is unable to differentiate between them. This unexpected performance is usually observed when E_b/N_0 is less than 5 dB. However, when E_b/N_0 is increased beyond 5 dB, the performance becomes normal and increases proportionally with the SNR value. When the window size is 512 samples, the detector performs extremely well and 100% detection is reached at E_b/N_0 = -8 dB. In general, increasing the window size results in an improvement in the overall performance of the system. For example, 90% detection is achieved at E_b/N_0 = -2 dB, -4 dB, -5 dB and less than -10 dB for observation windows of size 50, 100, 200 and 512 respectively. The best performance is clearly seen when the window size is 512 samples.

The performance also improves when the number of CRs in the network is increased. The change in the performance is shown in Figure 38. When the number of CRs in the network is increased from 1 to 3 users, an improvement of about 3 dB is observed. However, when the number of CRs is increased to 5 or 7, an improvement of only 1 dB is seen. Clearly, moving from a single CR system to a CR network significantly improves the performance but increasing the number of CRs in the network does not result in a considerable improvement as compared to the added complexity to the system when the number of CRs is increased. For a single CR system with an observation window of 50 samples in an AWGN channel, 90% detection is achieved at E_b/N_0 = 2 dB while the same detection requires E_b/N_0 = -1 dB, -2 dB and -3 dB when the number of CRs in the network are 3, 5 and 7, respectively.

The Cooperative Entropy detector system is investigated in a fading channel in the next simulation. Figure 39 compares the performance of a single CR and a 5-CR network with an observation window of 50 samples. An overall improvement of 6 dB is seen for SNR values above 0 dB. At low SNR values, on the other hand, the performance is poor at reaches 0% detection below E_b/N_0 = -5 dB due to the large amount of noise and deep fades in the received signal. While the single CR system requires E_b/N_0 = 8 dB to achieve 90% detection, the 5-CR network requires E_b/N_0 = 2 dB to achieve the same performance which is a 6 dB improvement in the performance.

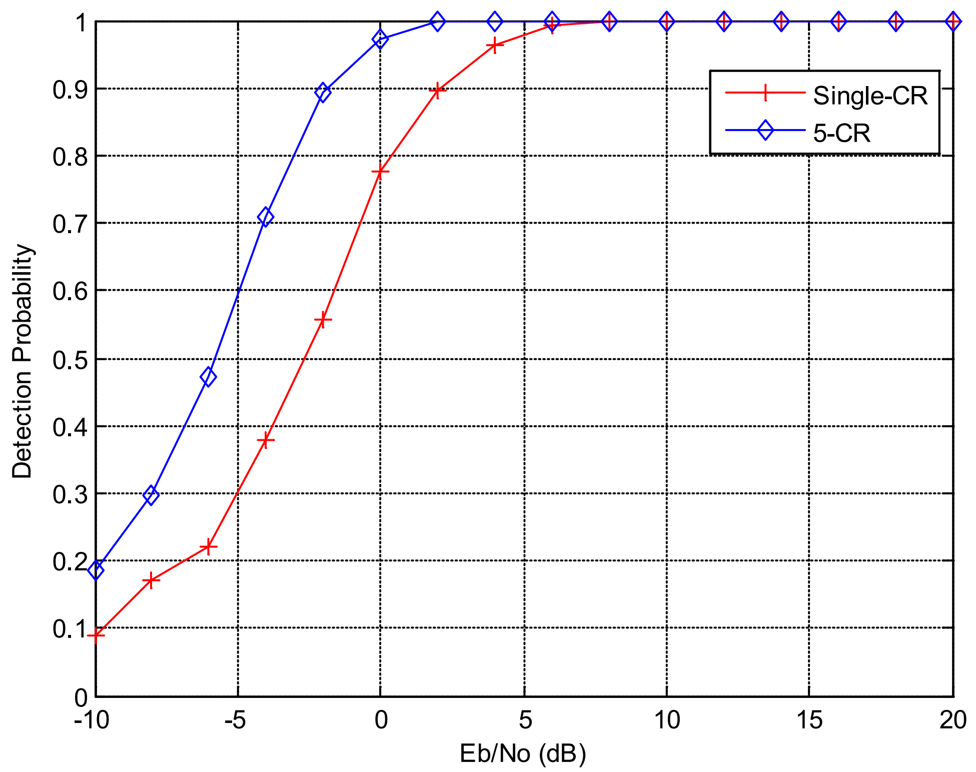

Figure 36 - Single CR and 5-CR Entropy Detector in AWGN

When the window size is increased for a 5-CR network, an unexpected overall degradation in the performance is observed as shown in Figure 40. The system performs poorly at low SNR values for all window sizes. A similar performance is observed for observation windows of size 50 and 100 both of which require $E_b/N_0 =$ 2 dB to achieve 90% detection. However, when the window size is increased to 200 and 512 samples, the performance degrades by around 2 dB (for window size of 200) and a further 4 dB (for a window size of 512). This is again an unexpected result as the performance is expected to improve when the window size is increased, as observed for the AWGN channel. No justification for obtaining such a result was found. For observation windows of size 200 and 512 samples, 90% detection is achieved at E_b/N_0 = 4 dB and 8 dB respectively.

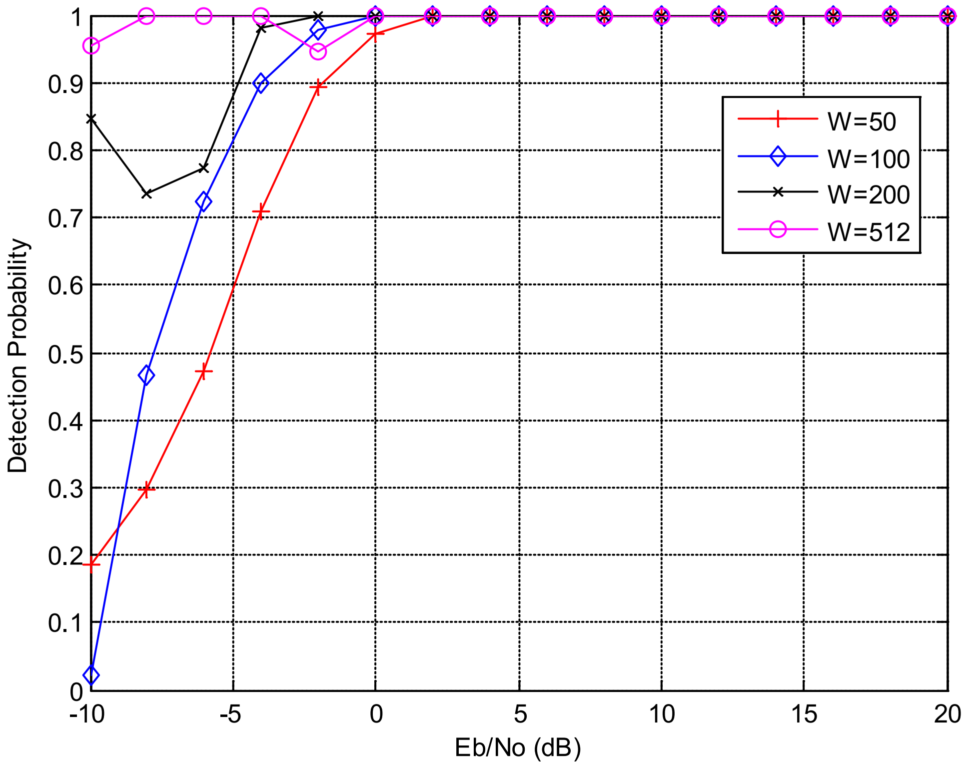

Figure 37 - 5-CR Entropy Detector in AWGN for different window sizes

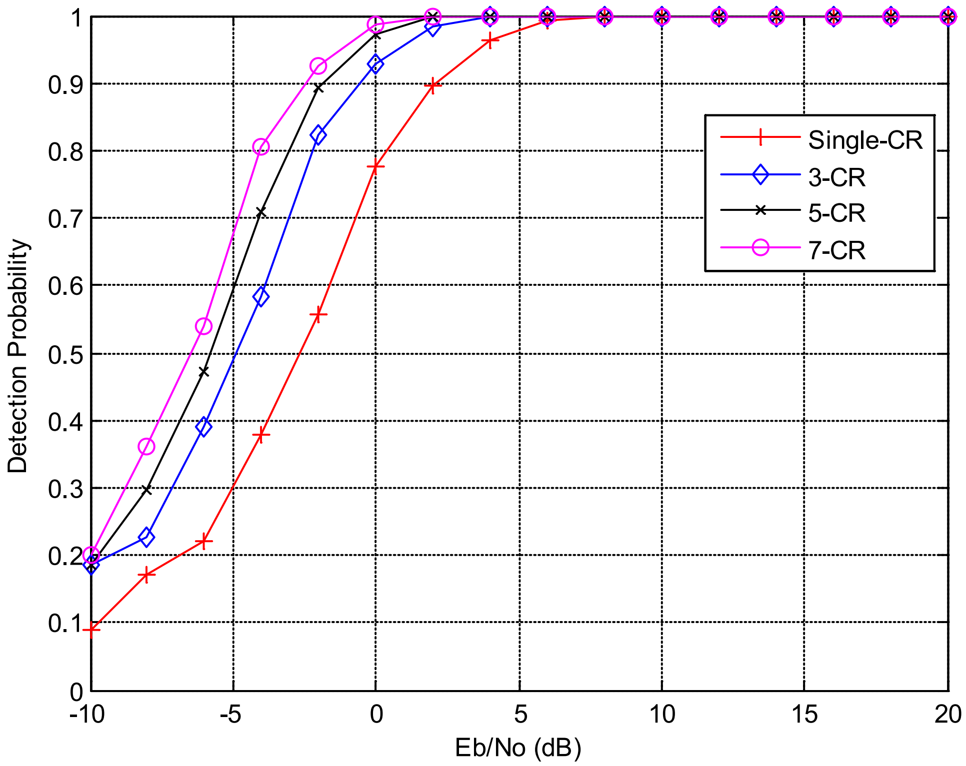

Figure 38 - Cooperative Entropy Detector in AWGN

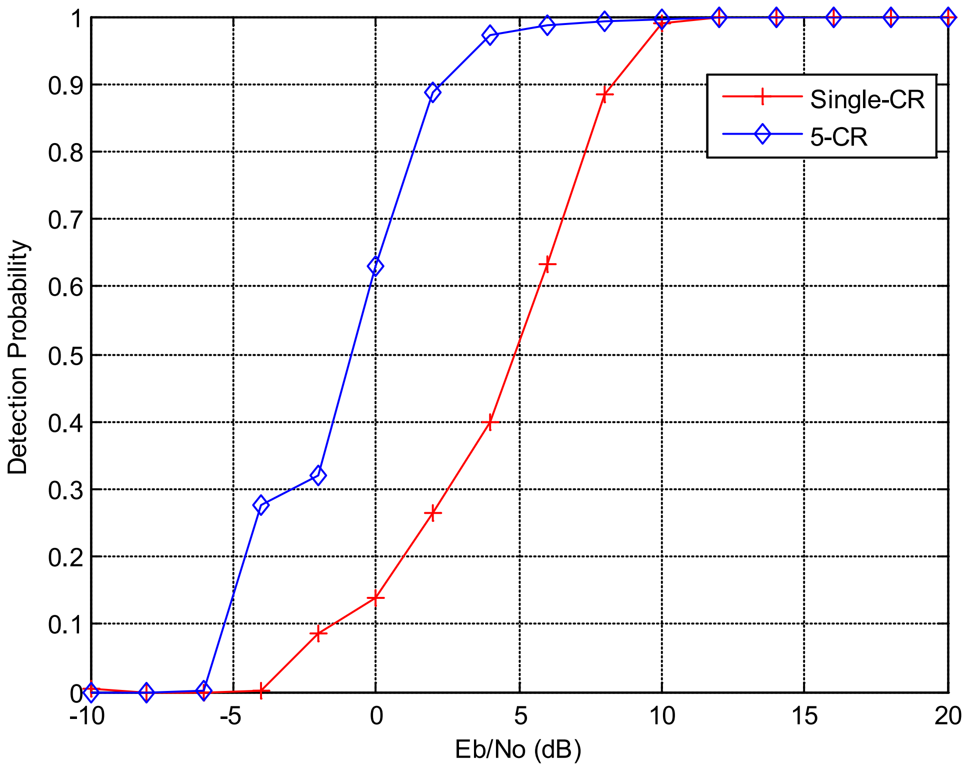

Figure 39 - Single-CR and 5-CR Entropy Detector in Fading Channel

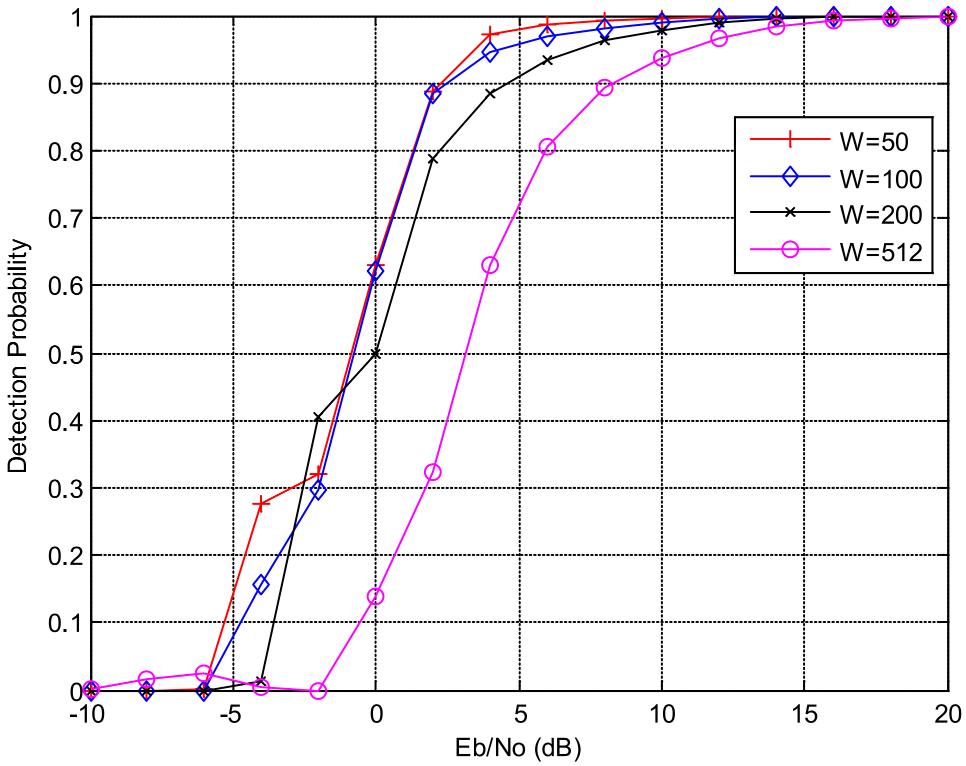

Figure 40 - 5-CR Entropy Detector in Fading Channel for different window sizes

Furthermore, increasing the number of CRs results in an overall improvement in the entropy detector performance in the fading channel too as the linear classifier is able to make a more informed decision about the received signal when the number of CRs in the network is high. Figure 41 shows the performance for different number of CRs in the network operating in a fading channel for a window size of 50 samples. Similar to AWGN case, when the number of CRs is increased from 1 to 3, the improvement in performance is larger compared to the improvement observed when the number is increased from 3 CRs to 5 or 7 CRs. When the number of CRs is increased from 1 CR to 3 CRs, an improvement of around 4 dB is seen. An improvement of 2 dB is seen when the number of CRs is increased from 3 to 5 users. The entropy detector fails once again at low SNR values with 0% detection for E_b/N_0 = -5 dB and below. 90% detection is achieved at E_b/N_0 = 8 dB, 4 dB, 2 dB and 1 dB for a single-CR, 3-CR, 5-CR and 7-CR network, respectively.

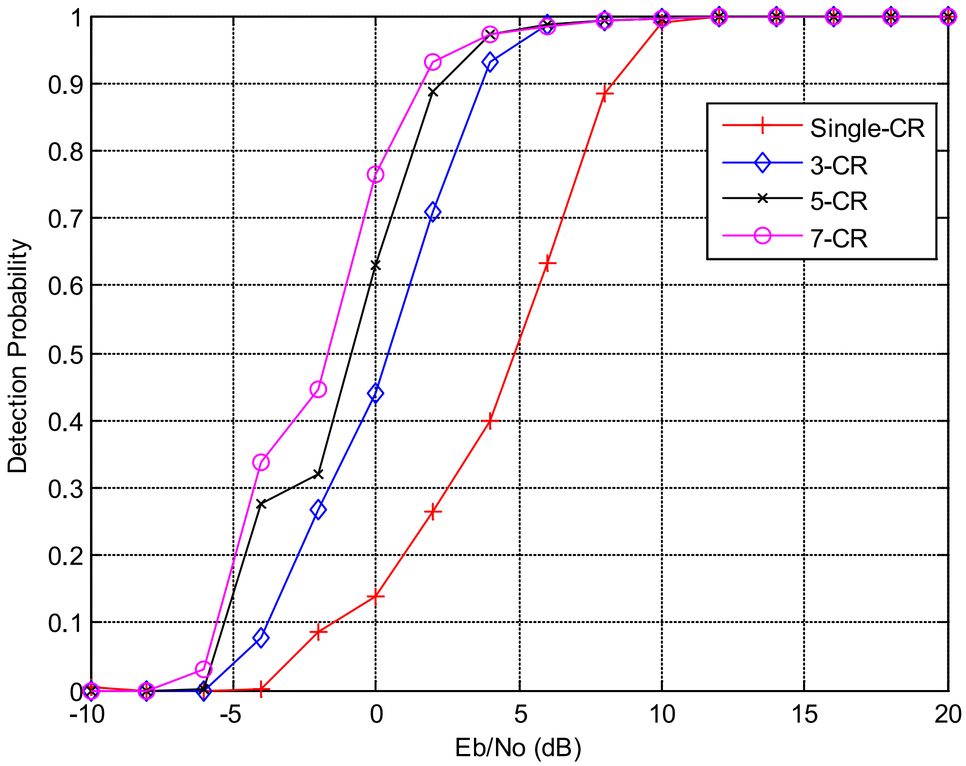

Figure 41 - Cooperative Entropy Detector in Fading Channel

In summary, the entropy detector performs excellent at high SNR values but fails at small SNR. The performance of the system improves when the number of CRs in increased in both the AWGN and the fading channel. However, increasing the window size in an AWGN improves the performance but gives unexpected results in a fading channel.

4.5. Summary of Results

This section presents a comparison of the performance of the three detectors investigated in this thesis: energy, correlation and entropy. The performance of these detectors over an AWGN and fading channel in a single or cooperative CR system is discussed and evaluated.

Figure 42 shows the performance of the detectors in a single CR system over an AWGN channel for an observation window of 50 samples. From the figure, it is observed that both energy and correlation detectors have similar performance in an AWGN channel while the entropy detector suffers at low E_b/N_0 values when the entropy of the signal is lost due to corruption by noise. 90% detection is achieved at E_b/N_0 = -3 dB and -2 dB for energy and correlation detectors respectively while the entropy detector requires E_b/N_0 = 2 dB to achieve the same detection rate. Therefore, due to its simplicity, energy detector is the best solution for spectrum sensing in single CR system operating in an AWGN channel.

Figure 43 shows the performance of the energy, correlation and entropy detector in a fading channel for an observation window of 50 samples. Clearly, in a fading channel, the correlation detector outperforms both the energy and entropy detectors. The maximum detection rate of the energy detector is 60% which requires at least E_b/N_0 = -3 dB. The correlation detector requires E_b/N_0 = 3 dB to achieve 90% detection as compared to E_b/N_0 = 8 dB to attain the same detection for entropy detector. However, energy detector performs better than the correlation and entropy detectors at low E_b/N_0 values. At E_b/N_0 = -10 dB, the detection rate of both the correlation and entropy detectors is almost 0% but the energy detector achieves 34% detection. Both the correlation and entropy detectors are coherent detection schemes in which the received signal is divided by the channel coefficients to compensate for the channel effect. This compensation results in noise enhancement at low E_b/N_0 values and degrades the performance of the system. Energy detector is a non-coherent scheme and therefore does not suffer at low E_b/N_0. Clearly, correlation detector is the best candidate for spectrum sensing in a fading channel at high E_b/N_0 for a single CR system when the observation window size is 50 samples while energy detector is the recommended at low E_b/N_0.

Figure 44 depicts the performance of the three detectors in a 5-CR Cooperative network in an AWGN channel for a window size of 50 samples. Similar to the single CR system in an AWGN, energy and correlation detector once again perform alike while entropy detector shows an improvement in performance as compared to the single CR system. For example, 90% detection is achieved at E_b/N_0 = -7 dB for both the energy and correlation detectors while the same detection rate requires E_b/N_0 = -2 dB for the entropy detector. Once again, the energy detector is the preferred spectrum sensing technique in a Cooperative CR network operating in an AWGN channel due to its ease of implementation.

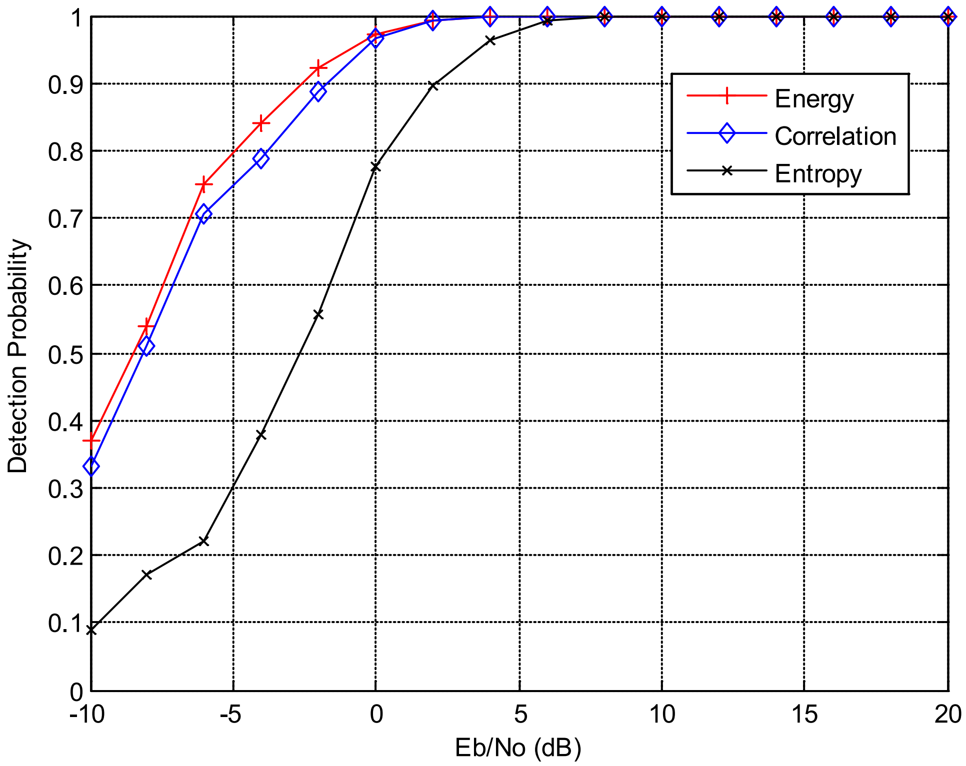

Figure 42 - Energy, Correlation and Entropy Detectors in AWGN

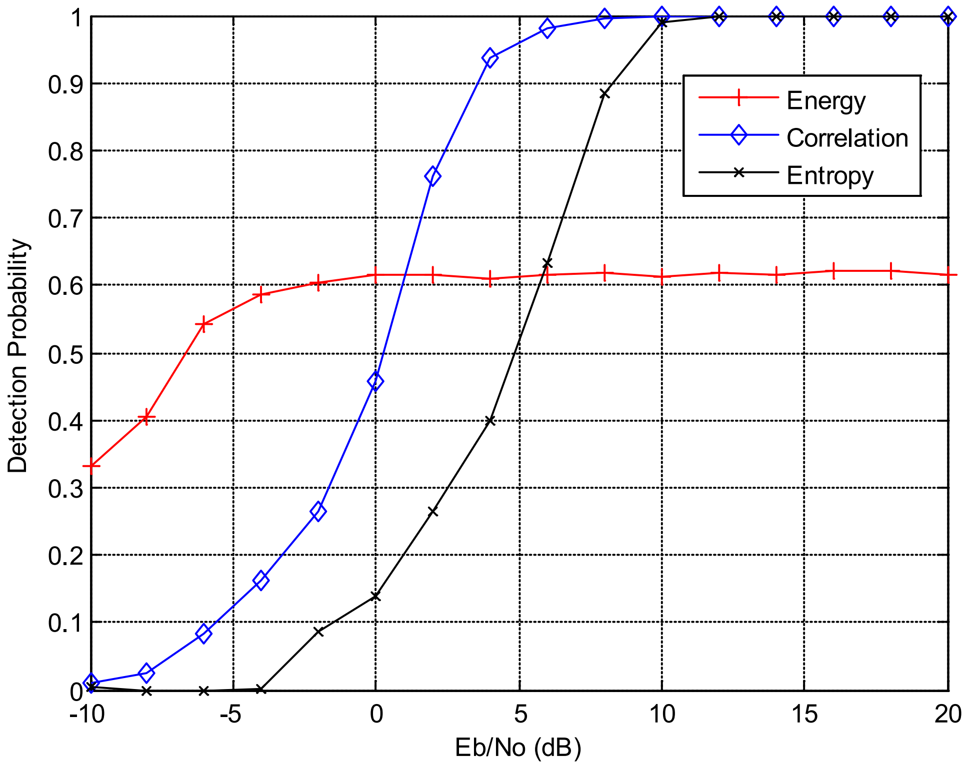

Figure 43 - Energy, Correlation and Entropy Detectors in Fading Channel

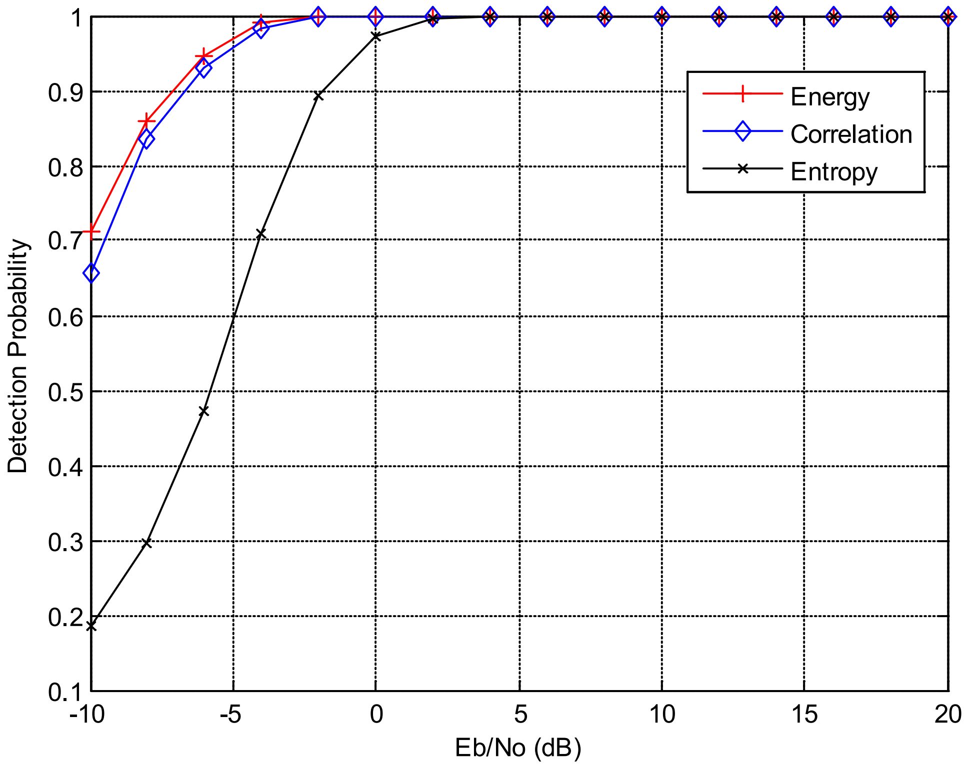

Figure 44 - 5-CR Coop Energy, Correlation and Entropy Detectors in AWGN

Finally, Figure 45 plots the performance of the three detectors in a 5-CR network operating in a fading channel. The energy detector still performs poorly in a fading channel although an improvement is seen as compared to the single CR case. The entropy detector shows the most improvement when the number of users in the CR is increased from one to 5 and its performance is comparable to the correlation detectors performance. For instance, both the correlation and entropy detector require around E_b/N_0 = 2 dB to achieve 90% detection while the energy detector requires E_b/N_0 = 6 dB to achieve the same detection. However, the energy detector outperforms the correlation and entropy detectors at low E_b/N_0 as noise enhancement due to channel compensation degrades their performance. At E_b/N_0 = -10 dB, the detection rate of both the correlation and entropy detector is almost 0% but the energy detector achieves 51% detection. In conclusion, both the correlation and entropy detectors are suitable candidates for spectrum sensing in a 5-CR cooperative network operating in a fading channel at high E_b/N_0 while energy detector is recommended at low E_b/N_0. Table 4 summarizes these observations.

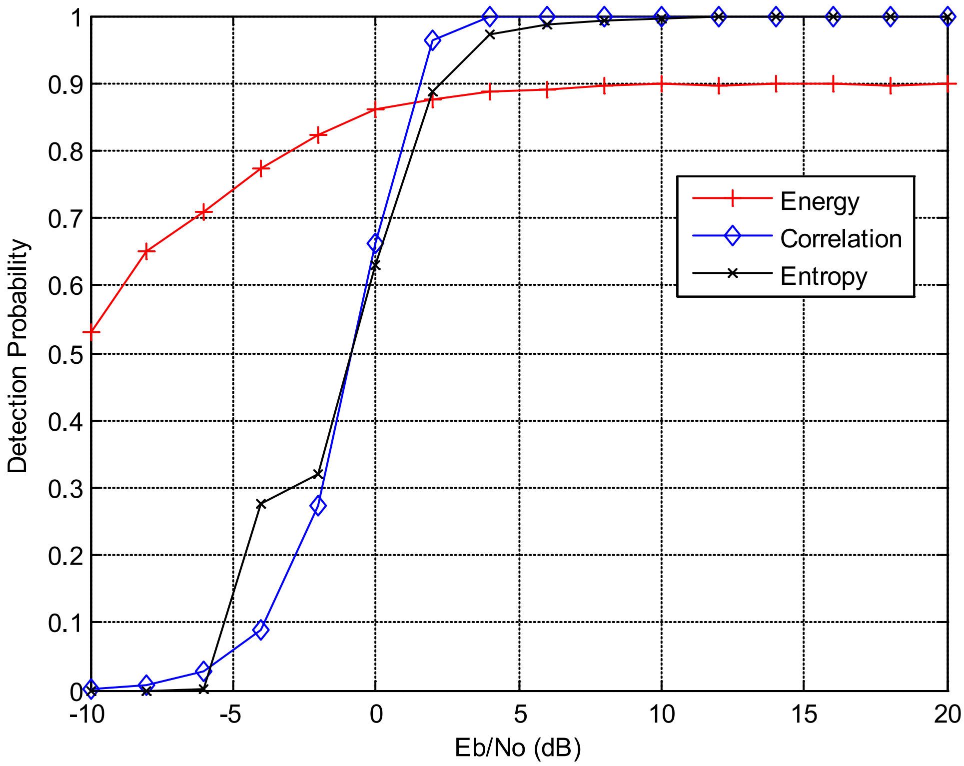

Figure 45 - 5-CR Coop Energy, Correlation, Entropy Detectors in Fading Channel

Table 4 – Recommended Spectrum Sensing Techniques

	AWGN		Fading Channel	
	Low E_b/N_0	High E_b/N_0	Low E_b/N_0	High E_b/N_0
Single CR	Energy	Energy	Energy	Correlation
5-CR Cooperative Network	Energy	Energy	Energy	Correlation or Entropy

4.6. Comparison with Traditional Detection Schemes

In this section, a comparison of the traditional detection schemes with the proposed classifier based schemes is presented. The traditional schemes set a threshold and adjust it iteratively until the target false alarm probability is achieved for each E_b/N_0 value. Using the adjusted threshold, the achieved detection probability is computed.

Figure 46 shows the performance of the traditional energy detector. Comparing the two energy detectors, classifier and non-classifier based, it is concluded that both perform similarly. The main difference is in the number of iterations that are required to adjust the threshold to provide the desired false alarm probability of 0.1. The traditional energy detector requires an average of 130 iterations at E_b/N_0 = -10 dB to achieve the desired false alarm while the classifier based energy detector requires around 3 iterations only to achieve the desired false alarm for the same E_b/N_0 value.

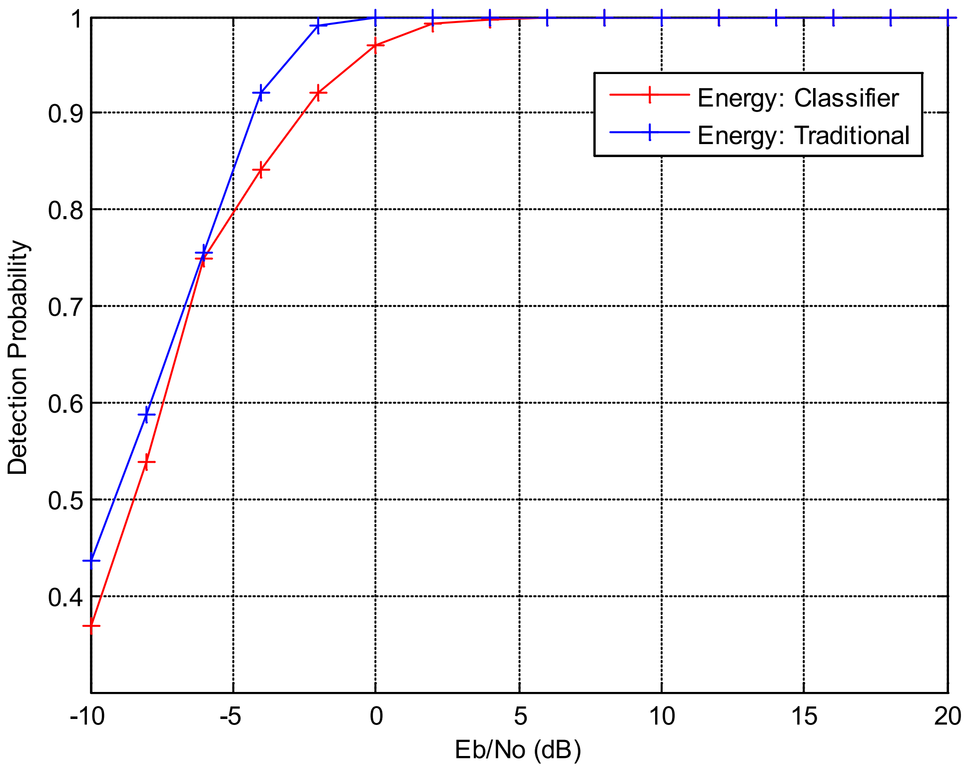

Figure 46 - Comparison of Traditional and Classifier based Energy Detectors

Figure 47 shows the performance of the traditional correlation detector. Comparing the two correlation detectors, classifier and non-classifier based, it is concluded that both perform quite similarly. The traditional correlation detector requires an average of 5 iterations to adjust the threshold and achieve the target false alarm at E_b/N_0 = -10 dB to achieve the desired false alarm while the classifier based correlation detector requires around 2 iterations to achieve the desired false alarm for the same E_b/N_0 value.

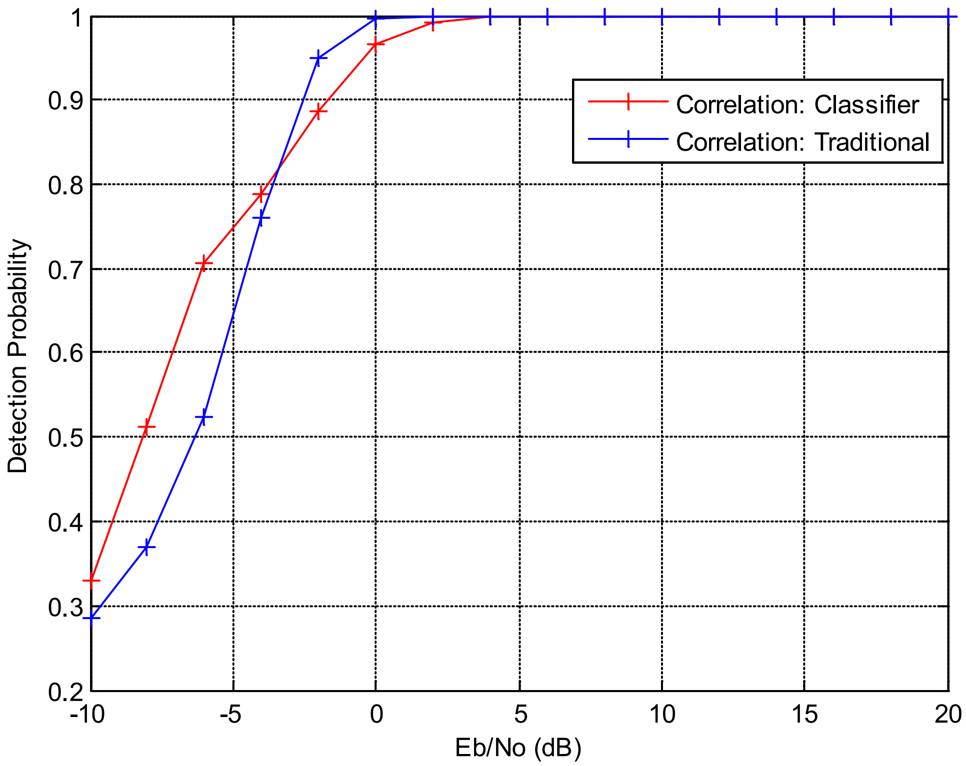

Figure 47 - Comparison of Traditional and Classifier based Correlation Detectors

Finally, Figure 48 shows the performance of the traditional entropy detector. Comparing the two entropy detectors, classifier and non-classifier based, it is concluded that both perform quite similarly. The traditional entropy detector requires an average of 5400 iterations to adjust the threshold and achieve the target false alarm at E_b/N_0 = -10 dB to achieve the desired false alarm which is significantly greater than the 2 iterations required on average by the classifier based entropy detector to achieve the target false alarm at the same E_b/N_0 value.

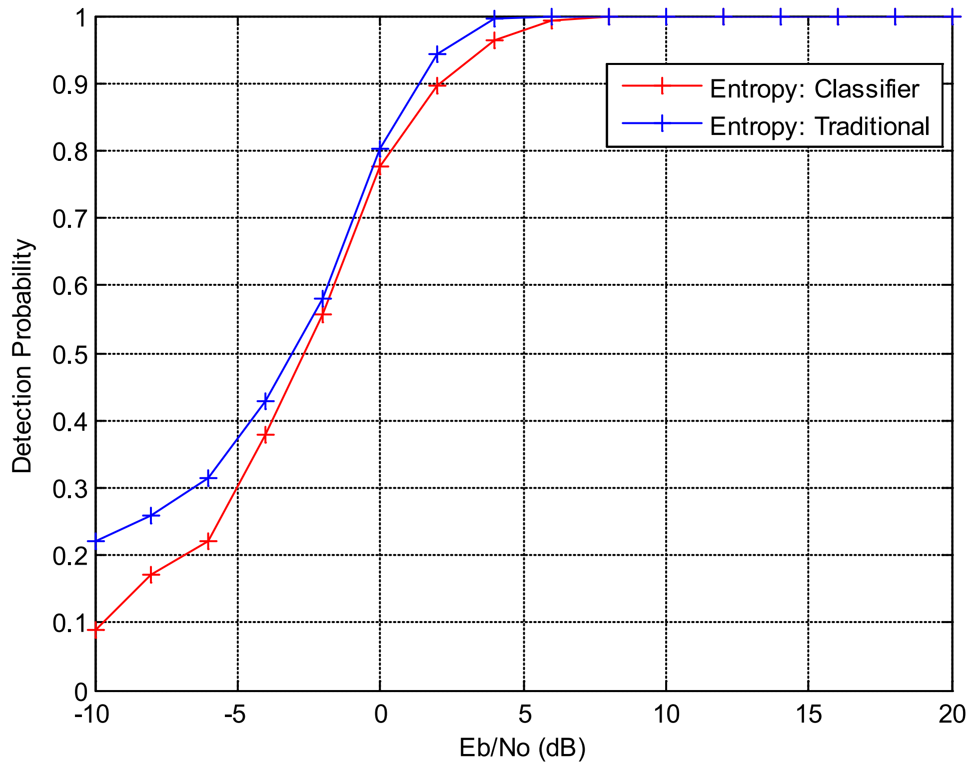

Figure 48 - Comparison of Traditional and Classifier based Entropy Detectors

Chapter 5

Conclusion and Future Work

Cognitive Radios opportunistically exploit the unused portions of the radio spectrum by sensing the spectrum for the presence of the primary licensed user. Most CRs are expected to work in environments where the primary users are OFDM based. In this thesis, spectrum sensing was formulated as a two class pattern recognition problem and a linear classifier was used to classify the received signal into either the primary OFDM signal or noise. The signals from both classes have different characteristics which were exploited for classification. These characteristics, known as features, were learnt by the classifier during the training phase and then used to classify signals into their respective classes.

Three different features were used in this work: energy, correlation and entropy. Energy detector computes the energy of the received signal and does not require any prior knowledge of the primary signal. The correlation detector exploits the inherent correlation in the OFDM signal due to the insertion of the cyclic prefix. The CR assumes complete synchronization and knowledge of the channel when using correlation as a detection scheme. Entropy detection computes the amount of information contained in the received signal. The performance of all these detection schemes was evaluated in an AWGN and fading channels. Additionally, the number of users in the cooperative CR network was changed which resulted in an overall improvement in the performance. The performance of the CR was evaluated in terms of the detection probability obtained when the false alarm probability is kept below a certain level.

In an AWGN channel, it was observed that energy and correlation detectors had similar performances and both were better than the entropy detector. When the observation window size was increased, the performance of all the three detection schemes improved. Introduction of multiple CRs into the system also resulted in an overall improvement in performance. However, adding multiple CRs into the network increases the complexity of the system. In the fading channel, on the other hand, the performance of the three detectors degraded. But, energy detector suffered the most degradation due to its non-coherent nature. Correlation detector performed best in the fading channel followed by the entropy detector. Increasing the window size improved the performance of the correlation and energy detectors but resulted in degrading the entropy detector performance. When the number of CRs was increased, there was an improvement in the performance of all the detectors.

Future work could focus on employing polynomial classifiers to solve the spectrum sensing problem. In a polynomial classifier, a lower order feature vector is converted into a higher order through the process of polynomial expansion. Unlike the linear classifier, which requires the data to be linearly separable, the polynomial classifier can operate on data which is not linearly separable but has a nonlinear decision boundary. Besides, spectrum sensing in OFDMA systems could also be investigated which is another form of OFDM in which different subcarriers are assigned to different users unlike the traditional OFDM system where all subcarriers are used by a single user.

References

[1] "Spectrum Policy Task Force: Report of the Spectrum Efficiency Working Group," Federal Communications Commission, Nov. 2002.

[2] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201-219, Feb. 2005.

[3] S. Bokhariee, H. Nguyen, and E. Shwedyk, "Spectrum Sensing for OFDM-Based Cognitive Radios," in *Vehicular Technology Conference Fall (VTC 2010-Fall)*, 2010, pp. 1-5.

[4] "Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing congnitive radio technologies," Federal Communications Commission, ET Docket No. 03-108, Feb. 2005.

[5] T. Yucek and H. Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 116-130, 2009.

[6] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. Boston: Mcgraw-Hill, 2007.

[7] H. Mahmoud, T. Yucek, and H. Arslan, "OFDM for Cognitive Radio: Merits and Challenges," *IEEE Wireless Communications*, vol. 16, no. 2, pp. 6-15, April 2009.

[8] A. Bahai and B. Saltzberg, *Multi-Carrier Digital Communications Theory and Applications of OFDM*. New York: Kluwer Academic Publishers, 2002.

[9] Y. Cho, J. Kim, W. Yang, and C. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Singapore: John Wiley & Sons (Asia) Ltd., 2010.

[10] S. Chaudhari, V. Koivunen, and H. Poor, "Autocorrelation-Based Decentralized Sequential Detection of OFDM Signals in Cognitive Radios," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2690-2700, July 2009.

[11] C. Hwang, G. Lai, and S. Chen, "Spectrum Sensing in Wideband OFDM Cognitive Radios," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 709-719, Feb. 2010.

[12] B. Fette, *Cognitive Radio Technology*, 2nd ed. Burlington: Academic Press, 2009.

[13] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, 3rd ed.: Prentice Hall, 2012.

[14] T. Lim, R. Zhang, Y. Liang, and Y. Zeng, "GLRT-Based Spectrum Sensing for Cognitive Radios," in *IEEE Global Telecommunications Conference GLOBECOM*, 2008, pp. 1-5.

[15] P. Pawelczak, K. Nolan, S. Oh, and D. Cabric, "Cognitive Radio: Ten Years of Experimentation and Development," *IEEE Communications Magazine*, vol. 55, pp. 523-531, March 2011.

[16] E. Hossain and V. Bhargava, *Cognitive Wireless Communication Networks*. New York: Springer, 2007.

[17] J. Shen, Y. Liu, S. Liu, J. Gao, G. Xie and C. Chi, "Robust Energy Detection Based on Bayesian Estimation for Cognitive Radio," in *IEEE Global Telecommunications Conference GLOBECOM*, 2008, pp. 1-5.

[18] K. Kim, Y. Xin, and S. Rangarajan, "Energy Detection Based Spectrum Sensing for Cognitive Radio: An Experimental Study," in *IEEE Global Telecommunications Conference GLOBECOM*, 2010, pp. 1-5.

[19] S. Atapattu, C. Tellambura, and H. Jiang, "Energy Detection Based Cooperative Spectrum Sensing in Cognitive Radio Networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1232-1241, April 2011.

[20] R. Sharma and J. Wallace, "Correlation-Based Sensing for Cognitive Radio Networks: Bounds and Experimental Assessment," *IEEE Sensors Journal*, vol. 11, no. 3, pp. 657-666, March 2011.

[21] Y. Zhang, Q. Zhang, and S. Wu, "Entropy-based robust spectrum sensing in cognitive radio," *IET Communications*, vol. 4, no. 4, pp. 428-436, April 2010.

[22] Y. Hassan, M. El-Tarhuni, and K. Assaleh, "Comparison of Linear and Polynomial Classifiers for Co-operative Cognitive Radio Networks," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications PIMRE*, 2010, pp. 797-802.

[23] Y. Hassan, M. El-Tarhuni, and K. Assaleh, "Knowledge Based Cooperative Spectrum Sensing Using Polynomial Classifiers in Cognitive Radio Networks," in *4th International Conference on Signal Processing and Communication Systems*, 2010, pp. 1-6.

[24] G. Noh, J. Lee, H. Wang, S. You, and D. Hong, "A New Spectrum Sensing Scheme using Cyclic Prefix for OFDM-based Cognitive Radio Systems," in *IEEE Vehicular Technology Conference VTC*, 2008, pp. 1891-1895.

[25] A. Jayaprakasam, V. Sharma, C. Murthy, and P. Narayanan, "Cyclic Prefix Based Cooperative Sequential Spectrum Sensing Algorithms for OFDM," in *IEEE International Conference on Communications ICC*, 2010, pp. 1-6.

[26] H. Chen, W. Gao, and D. Daut, "Spectrum Sensing for OFDM Systems Employing Pilot Tones," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5862-5870, Dec. 2009.

[27] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.

[28] L. Kuncheva, *Combining Pattern Classifiers Methods and Algorithms*. United States of America: John Wiley & Sons, 2004.

[29] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. United States of America: Academic Press, 2009.

[30] H. Schulze and C. Luders, *Theory and Applications of OFDM and CDMA*. West Sussex: John Wiley & Sons Ltd., 2005.

[31] M. Muzaffar, M. El-Tarhuni, and K. Assaleh, "Learning-based Spectrum Sensing in OFDM Cognitive Radios," in *The Second International Conference on Advances in Cognitive Radio*, Chamonix / Mont Blanc, 2012, pp. 57-62.

Vita

Muhammad Umair Muzaffar was born in 1989 in Dubai, United Arab Emirates. He completed O-levels and A-levels from Arab Unity School in Dubai. He was awarded the Sheikh Maktoum bin Muhammad bin Rashid Al Maktoum award for Outstanding Academic Achievement in Cambridge International Examinations in 2005. He received his Bachelors degree in Electrical Engineering with Magna Cum Laude from American University of Sharjah in 2010. In Fall 2010, Umair joined the Masters of Electrical Engineering program in American University of Sharjah where he worked as a Graduate Teaching Assistant for two years. During this time, he published a conference paper in the Second International Conference on Advances in Cognitive Radio held at France in 2012.