Learning-based Spectrum Sensing in OFDM Cognitive Radios

Muhammad Umair Muzaffar
Dept. of Electrical Engineering
American University of Sharjah
Sharjah, UAE
b00025553@aus.edu

Mohamed El-Tarhuni
Dept. of Electrical Engineering
American University of Sharjah
Sharjah, UAE
mtarhuni@aus.edu

Khaled Assaleh
Dept. of Electrical Engineering
American University of Sharjah
Sharjah, UAE
kassaleh@aus.edu

***Abstract*—In this paper, spectrum sensing in OFDM-based cognitive radio systems is modeled as a pattern recognition problem. The proposed scheme uses a linear classifier to decide on when the spectrum is busy (class 1) or not busy (class 2). Two types of feature vectors are compared in this work, namely energy estimates and cross-correlation estimates using the cyclic prefix of the OFDM signal. Simulation results indicate that the energy-based linear classifier provides excellent performance in terms of detection probability over AWGN channels but suffers significant degradation if the channel undergoes flat Rayleigh fading conditions. On the other hand, the correlation-based features offer a more robust performance under both AWGN and fading conditions with a detection rate of about 90% at a signal-to-noise ratio of -3 dB.**

Keywords- cognitive radio; OFDM; linear classifier; energy detection; correlation detection.

I. INTRODUCTION

The radio spectrum is one of the most expensive resources in wireless communication systems. Service providers and users of the radio spectrum are generally required to obtain a license in order to use a particular frequency band. However, these users do not use the assigned spectrum at all times of the day and spectrum holes are created when the licensed user is not using its allotted spectrum resulting in an inefficient use of the radio spectrum [1]. To counter this problem, cognitive radio technology has been introduced which allows secondary users to access the spectrum only when it is not being used by the licensed user. Intuitively, the cognitive radio (CR) should be able to sense the spectrum to detect the presence or absence of the licensed primary user. By definition, spectrum sensing is the task of obtaining awareness about the spectrum usage and determining the existence of primary users in a geographical area [2].

The optimal algorithm for spectrum sensing is the likelihood ratio test (LRT) [3][4] and several techniques have been proposed in the literature which employ the LRT using energy detection [5][6], autocorrelation [7], cyclostationarity [8] and pilot tones [9] to sense the spectrum. In addition, CR has also been considered as a pattern recognition problem where spectrum sensing is done using linear or polynomial classifiers [10][11]. This is because the signal received at the CR can be either the primary user signal or noise, both of these signals have different characteristics which a classifier can learn during the training phase and then utilize this learning to classify any unseen data into one of two classes: the primary signal (class 1) or noise (class 2). Any incoming signal has to be classified into one of these classes by the linear classifier. However, Orthogonal Frequency Division Multiplexing (OFDM) based CRs were not investigated in this research.

OFDM has rapidly developed into the preferred modulation scheme for most wireless standards such as IEEE 802.11a/g, IEEE 802.16 and IEEE 802.20 [7]. Consequently, cognitive radios operating in wireless channels are expected to be OFDM based. In addition, OFDM is the best physical layer candidate for cognitive radios because it allows for generation of signals which fit into discontinuous and arbitrary sized spectrum segments [12].

The performance of a CR is measured using detection probability which is defined as the probability with which the CR (or secondary user) correctly decides that the target radio spectrum is occupied by the primary user. Another important parameter is the false alarm probability defined as the probability with which the CR incorrectly decides on the presence of a primary signal thereby not allowing the CR to transmit while, in fact, it is eligible to.

As mentioned earlier, most of the existing techniques employ the LRT to decide on the presence or absence of the primary OFDM signal. In [7], the autocorrelation coefficient is computed at the CR which is zero when no signal is present and is a function of different parameters such as the energy per bit-to-noise power spectral density (E_b/N_0), subcarriers, and cyclic prefix when the primary signal is received. However, the variance of the received signal is unknown and maximum likelihood estimate (MLE) is used to compute it. The LRT is then applied and its result is compared with a threshold, which depends directly on the autocorrelation function of the OFDM signal, to make a decision on presence of the primary signal. Alternatively, pilot tones in the OFDM signal can also be used to sense the spectrum [9]. The time-domain symbol cross-correlation (TDSC) of two OFDM symbols is computed which has a nonzero constant value only if both the symbols have same pilots. Comparing the TDSC with a threshold determines the presence or absence of the signal.

In this paper, spectrum sensing technique for an OFDM based CR is proposed using a linear classifier instead of the traditionally used likelihood ratio test. The linear classifier

 ISBN: 978-1-61208-197-7

receives an input signal and decides whether the input signal belongs to one of two classes: Class 1: OFDM primary signal and Class 2: Noise.

The rest of the paper is organized as follows: In Section II, a system model for the OFDM CRs is presented. Section III discusses the proposed spectrum sensing technique and the features to be used for sensing. Section IV illustrates the performance of the proposed system through simulation results and Section V concludes the paper.

II. System Model

In an OFDM system, the available frequency band is divided into N overlapping but orthogonal narrow sub-bands each associated with a sinusoidal subcarrier. For high data rate transmission, each subcarrier is used to carry a small part of data and, due to the narrow band nature, does not suffer from channel distortion caused by Intersymbol Interference (ISI). This is considered as the main advantage of OFDM signal since there is no need for complex equalization schemes to mitigate ISI as in single-carrier systems.

The data to be transmitted using M-QAM or M-PSK modulation is converted into N parallel streams each to be transmitted over a separate subcarrier. An Inverse Fast Fourier Transform (IFFT) block is used to modulate the N subcarriers with the N parallel symbol streams. Since the sinusoidal subcarriers are orthogonal, they do not cause interference among adjacent bands. However, due to channel delays and frequency offsets, the orthogonality among the subcarriers may be lost. To maintain this orthogonality, a cyclic prefix is added to the OFDM signal where the last L samples of the signal are copied and appended to the beginning to form the cyclic prefix.

As discussed above, the OFDM signal is constructed by feeding N symbols (or streams of symbols) to IFFT operator. Assume that $S(0), S(1), \ldots, S(N-1)$ are N complex QAM or PSK symbols, the output of the IFFT is:

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

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

$$\boldsymbol{s} = [s(N-L), \ldots, s(N-1), s(0), s(1), \ldots, s(N-1)]. \qquad (2)$$

The signal in (2) is first converted from digital to analog to form $s(t)$ and is then sent over the channel after up-conversion to the desired radio frequency carrier.

At the CR, the following signal will be received:

$$x(t) = c(t)s(t) + n(t), \qquad (3)$$

where $c(t)$ is the channel coefficient at time t and $n(t)$ is the additive white Gaussian noise, with zero mean and two-side power spectral density of $N_0/2$, which corrupts the transmitted signal. The CR will first down-convert the received signal $x(t)$ and then performs analog-to-digital conversion to get the following digital signal

$$x[k] = c[k]s[k] + n[k], \qquad (4)$$

where $c[k]$ is the discrete channel coefficient. At the CR, all the computations are done on the signal defined in (4).

III. Spectrum Sensing in Cognitive Radios

As discussed earlier, spectrum sensing can be considered a two class pattern recognition problem [10]. The main objective of a pattern recognition system is to assign any input signal or data to one of a number of known classes (or categories) based on features extracted from the input signal. The process of acquiring features from the input signal is called feature extraction. In this paper, pattern recognition is used at the CR to classify the received signal as primary signal or noise such that maximum detection probability is achieved while keeping the false alarm probability below a certain threshold. A block diagram of the proposed system is shown in Fig. 1.

The feature extracted from the received input signal can be one of the many techniques used for spectrum sensing such as Energy, Correlation, etc. In Fig. 1, the input to the CR is the vector of received signal samples, $\{x[k]\}$. The CR then extracts the features, $\boldsymbol{f}$, from this signal which are then input to the linear classifier. The classifier computes an output vector $\boldsymbol{T}$ which is used to classify the input signal based on the features into one of the two classes:

$$x[k] = c[k]s[k] + n[k] \text{ ; Class 1 (Spectrum busy)} \qquad (5)$$

$$x[k] = n[k] \text{ ; Class 2 (Spectrum available)} \qquad (6)$$

where class 1 is the case when the primary OFDM signal is present and the spectrum is occupied and class 2 is the case when no primary signal is present and the spectrum is available.

A. Energy Detection

One of the most commonly used techniques for spectrum sensing is Energy Detection. With this technique, the CR does not require any prior knowledge of the primary signal and, therefore, is very easy to implement. The CR senses the spectrum for a period of time and compares the received signal energy with a defined threshold to decide on the presence or absence of the primary signal. However, this type of detection is unreliable in fading environments where the energy of the primary signal has been severely degraded (attenuated) since 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). In such cases, the selection of a suitable threshold to decide whether the primary signal is present or not becomes a challenging task.

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. When the CR estimates the energy of the received signal over an observation window of size W, the energy, in time-domain, of the detected signal is computed as:

$$f_E = \sum_{k=0}^{W-1} |x[k]|^2. \tag{7}$$

The extracted energy feature, f_E, is then used by the linear classifier to make a decision on the class of the received signal, $x[k]$. To improve the performance of the energy detector, the CR can sense the spectrum more than once (each time for a window of W samples) and compute the energy of the received signal each time and store it as a feature. The linear classifier will now have multiple features and since the energy is computed for different instances of time, it will make a better and more informed decision on the presence or absence of the primary signal.

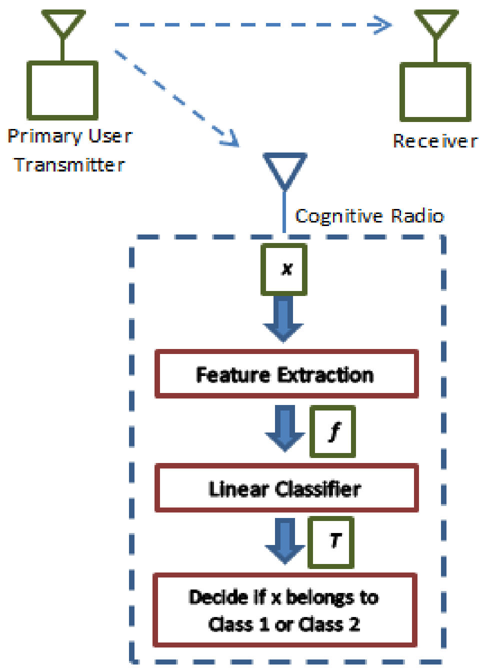

Fig. 1. Block diagram of the proposed system

B. Correlation Detection

Energy detection does not require any prior knowledge of the type of primary user signal. This could be considered as an advantage for such scheme but it results in inferior performance compared to other schemes that take advantage of certain structure in the OFDM signal. OFDM symbols have an inherent special property; namely the cyclic prefix, which can be utilized to sense the presence 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 similar to the last L samples. In the case when there is no distortion due to noise or channel, the first L samples are exactly the same as the last L samples. This implies that the first L samples of the OFDM symbol are highly correlated with the last L samples and this property can be used to sense the spectrum for presence of the signal. The CR performs correlation between the first W samples of the cyclic prefix at the start and end of the OFDM symbol 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 at the CR is computed as:

$$f_C = max|E[x_B x_E{}^*]|, \tag{8}$$

where $x_B = [x_1, x_2, \dots, x_W]$ is a vector of first W samples of the cyclic prefix at the beginning of the received signal and $x_E = [x_{N-L}, x_{N-L+1}, \dots, x_{N-L+W}]$ is a vector of the last W samples of the cyclic prefix at the end of the OFDM signal, $E[.]$ is the expectation operator and $max|.|$ takes the maximum value of the elements inside the argument. Finally, using the 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.

C. Training the Linear Classifier

For a linear classifier, a linear discriminant function is defined for each class which is used to separate data of a particular class from data of another class. A linear discriminant function is defined as:

$$g_i = \boldsymbol{w}_i^t \boldsymbol{f} + w_{i0}\,; \quad i = 1, \dots, N_C, \tag{9}$$

and,

$$\boldsymbol{f} = [f_1 \dots f_d], \tag{10}$$

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 $\boldsymbol{f}$ (for our case d=1) and t 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 incorportated 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:

$$\boldsymbol{a_i} = [w_0 \; \boldsymbol{w}_i^t], \tag{11}$$

and,

$$\boldsymbol{y} = [1 \quad \boldsymbol{f}] = [y_0 \; y_1 \dots y_d]. \tag{12}$$

The linear discriminant function for class i can be written as

$$g_i = \boldsymbol{a}_i^t \boldsymbol{y}\,; \qquad i = 1, \dots, N_C. \tag{13}$$

 ISBN: 978-1-61208-197-7

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:

$$\boldsymbol{Y} = [\boldsymbol{y}_{11} \quad \boldsymbol{y}_{12} \quad \cdots \quad \boldsymbol{y}_{1K} \quad \boldsymbol{y}_{21} \quad \cdots \quad \boldsymbol{y}_{2K}]^t, \qquad (14)$$

where $\boldsymbol{y}_{11}$ ... $\boldsymbol{y}_{1K}$ are feature vectors of data belonging to class 1 (OFDM signal) and $\boldsymbol{y}_{21}$ ... $\boldsymbol{y}_{2K}$ are features vectors of data belonging to class 2 (noise). The first K rows of $\boldsymbol{Y}$ correspond to data belonging to class 1 while the last K rows correspond to data from class 2. The number of elements in $\boldsymbol{Y}$ is $2K \times d + 1$. 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 (13). 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. Similary $\boldsymbol{t_2}$ is zero everywhere except the rows belonging to class 2. $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ are $2K \times 1$ dimensional vectors. The first K elements of $\boldsymbol{t_1}$ are 1 while the last K elements of $\boldsymbol{t_2}$ are 1. The target vectors are combined into a matrix $\boldsymbol{T}$ defined as:

$$\boldsymbol{T} = [\boldsymbol{t_1}\ \boldsymbol{t_2}]. \qquad (15)$$

In addition, a weight matrix, $\boldsymbol{A}$, is formed whose columns are the weight matrices for each class.

$$\boldsymbol{A} = [\boldsymbol{a_1}\ \boldsymbol{a_2}]. \qquad (16)$$

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

$$\boldsymbol{T} = \boldsymbol{YA}. \qquad (17)$$

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

$$\boldsymbol{A} = (\boldsymbol{Y^tY})^{-1}\boldsymbol{YT}. \qquad (18)$$

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, linear classifier will perform really well. However, if the data is not linearly separable, the linear classifier may fail. This can happen when at low SNR values when the signal and noise have comparable levels.

D. Testing the Linear Classifier

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. Similar to the training data described in (14), the test data consists of feature vectors belonging to class 1 and class 2. The first Z elements of $\boldsymbol{Y_{test}}$ belong to class 1 while the last Z elements belong to class 2. The linear classifier multiplies the test data, $\boldsymbol{Y_{test}}$, with the weight matrix, $\boldsymbol{A}$, to get a matrix, $\boldsymbol{T_{test}}$ with two columns. Ideally, the first column of $\boldsymbol{T_{test}}$ should be one for the first Z elements (corresponding to class 1) and zero for the rest while the second column of $\boldsymbol{T_{test}}$ should be zero for the first Z elements and one for the last Z elements (corresponding to class 2). However, the obtained values vary around these ideal values when novel data is fed to the classifier [10].

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. However, 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 [10].

IV. SIMULATION RESULTS

In this section, the performance of the linear classifier is determined using test data belonging to class 1 and class 2. Simulations were used to obtain results due to the complexity of analytical evaluation of the proposed technique. The signal is received by the CR and energy detection is performed by the CR and a decision is made on the presence or absence of the primary user signal. In addition to the energy detector, simulation results are also presented for the correlation detector where the CR uses the correlation between the cyclic prefix at the beginning and the end of the OFDM symbol as a feature to decide on the availability of the spectrum. The transmitted signal is modulated using M-QAM for different values for modulation level $\boldsymbol{M}$. For illustration purposes, the Digital Video Broadcasting – Terrestrial (DVB-T) standard is used in 4k mode. Under this condition, an OFDM signal structure with 4096 subcarriers and the cyclic prefix length $1/8$ of the number of subcarriers is used. The performance of the linear classifier is evaluated at different E_b/N_0 values when the signal passes through an ideal channel with AWGN only and also when the signal experiences flat channel fading with a low Doppler frequency of 3 Hz. Before testing the linear classifier, for all cases, the weight vector $\boldsymbol{A}$, defined in (18) is obtained using a random model for the primary user with 50% spectrum utilization and defining 2000 training data vectors, 1000 belonging to class 1 (primary signal) and 1000 to class 2 (noise only). This implies that the primary user occupies the spectrum only 50% of the time.

Fig. 2 shows the detection probability achieved by the CR using a linear classifier, while maintaining the false alarm probability below 0.1, for different values of E_b/N_0 using an observation window of W=50 samples and modulation level of M=2. The performance is shown for the cases when there is no fading and when there is slow fading using the energy detector and correlation detector. All results are averaged over 100 simulation runs. It can be seen that when there is no fading in the channel, the energy detector performs very well as it can accumulate enough energy to detect the presence of the signal. The correlation detector has a similar performance but falls behind at very low SNR conditions. On the other hand, when fading is present, the energy detector performance is severely degraded while the correlation detector exhibits a very small degradation in performance. This is because flat fading causes significant attenuation in the received signal energy resulting in degradation in performance of the energy detector. The correlation detector, however, depends on the repetitiveness in the received signal and is therefore less affected by flat fading. It is noted that, under AWGN, a detection probability of about 90% is achieved at about –4 dB and –3 dB for the energy and correlation detectors, respectively. Fading degrades the energy detector performance by about 15 dB while the correlation detector suffers around 6 dB degradation for the same detection probability. Ideally, the spectrum utilization can reach 100% where the secondary users utilize the spectrum whenever the primary user is not active. However, a reduction in the spectrum utilization is incurred in the event of a false alarm where secondary users decide that the spectrum is busy while the primary user is not transmitting. For the simulation example used in this paper, the primary user has a utilization of 50% and hence the secondary users can ideally achieve a utilization of 50% but since the false alarm rate was fixed to 0.1 then the actual utilization for the secondary users will be about 45%. Therefore, the total spectrum utilization by the primary and secondary users will be about 95%. Note that further improvements in spectrum utilization could be obtained by reducing the false alarm probability but this may result in reducing the detection probability leading to more interference from the secondary users to the primary user and hence reducing the overall spectrum utilization.

The performance of the classifier can be improved further by increasing the window size W. However, the window size of the correlation detector should not exceed the length of the cyclic prefix. Fig. 3 shows the performance of the correlation detector for different window sizes in a flat fading channel. The modulation level used is M =16. An improvement in performance is seen as the observation window size is increased for the correlation detector. For instance, using an observation window of size 10 requires E_b/N_0 = -1 dB to achieve 90% detection while the same detection probability is achieved at E_b/N_0 = -6 dB when the window size is increased to 100. However, no significant improvement can be seen when the observation window size is increased beyond 200. For window size of 200 and above, 90% detection is reached at around E_b/N_0 = -7 dB.

V. CONCLUSION

In this paper, spectrum sensing in a CR is modeled as a pattern recognition problem with two classes: the primary user signal and noise. Energy detection and correlation detection are used as features which are input to the linear classifier that decides on the presence or absence of the primary signal while maintaining the false alarm probability below a certain value. At the CR, training data is used to compute the optimal weight matrix. Simulation results show that energy detection provides excellent results only when there is no fading by the channel. However, in presence of flat fading, the energy detector suffers significant degradation in the detection performance while the correlation detector maintains good performance for most E_b/N_0 values. It is also shown that increasing the observation window size results in an improvement in the performance of the CR.

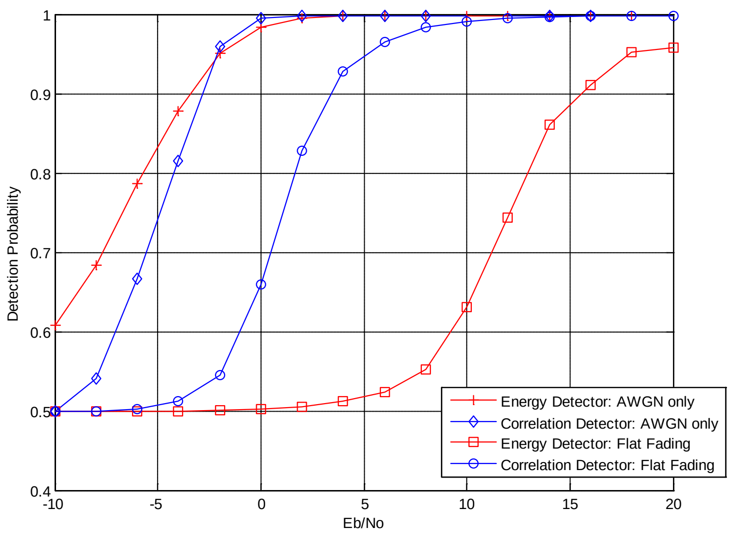

Fig. 2. CR performance in AWGN and flat fading

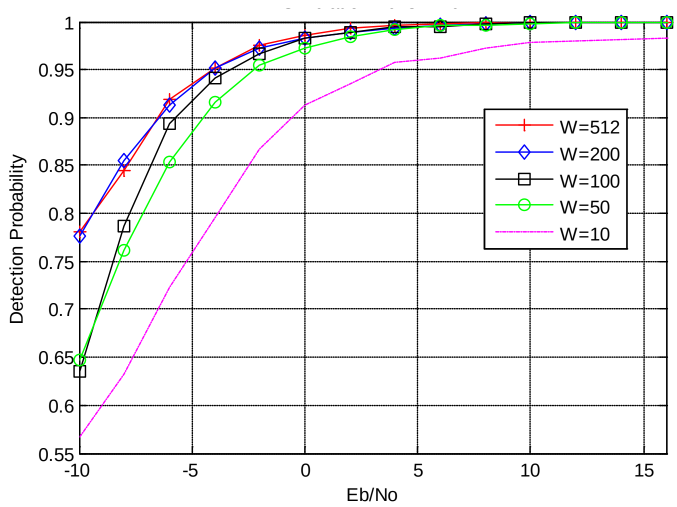

Fig. 3. Correlation detection performance for different window sizes

REFERENCES

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

[2] 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.

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

[4] T. J. Lim, R. Zhang, Y. C. Liang and Y. Zeng, "GLRT-Based Spectrum Sensing for Cognitive Radio," in IEEE Global Telecommunications Conference, 2008, pp. 1-5.

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

[6] H. Urkowitz, “Energy detection of unknown deterministic signals,” Proc. IEEE, vol. 55, pp. 523–531, Apr. 1967.

[7] S. Chaudhari, V. Koivunen and H. V. Poor, "Autocorrelation-Based Decentralized Sequential Detection of OFDM Signals in Cognitive

Radios," IEEE Transactions on Signal Processing, vol. 57, no. 7, pp. 2690-2700, 2009.

[8] W. A. Gardner and C. M. Spooner, “Signal interception: Performance advantages of cyclic-feature detectors,” IEEE Trans. Commun., vol. 40, pp. 149–159, 1992.

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

[10] 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, 2010, pp. 797-802.

[11] 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.

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