Noise Variance Estimation Through Penalized Least-Squares for ED-Spectrum Sensing

B.Naveen kumar*, S.Chris Prema*

*Dept. of Avionics, Indian Institute of Space Science and Technology, Thiruvananthapuram, India 695547

buduru.sc14m075@pg.iist.ac.in, chrisprema@iist.ac.in

***Abstract*—Cognitive Radio (CR) is an auspicious solution to current problem of spectrum scarcity due to evaluation of new technologies. These techniques are useful in detecting spectral holes, and allocating them to secondary users. Energy Detection is a predominant method for spectrum sensing due to its low computational complexity and capability of detecting spectrum holes without requiring apriori knowledge of primary signal. The energy based spectrum detectors depends on the precision of threshold chosen to distinguish signal and noise. But, energy detection needs to estimate the noise variance for finding the detection threshold. Most of the conventional techniques use fixed threshold with known noise variance. In practical scenarios noise variance is unknown, so we are proposing a fast computational noise variance estimation algorithm for spectrum sensing using Penalized Least Squares (PLS). We have introduced a smoothing parameter which is determined by Discrete Cosine Transform (DCT) as the penalizing factor. The amount of smoothing is determined by minimizing Generalized Cross Validation (GCV). Simulations were carried out in AWGN and Rayleigh fading channels for the proposed noise variance estimation through which Receiver Operating Characteristics (ROC) are obtained.**

Keywords*—Cognitive Radio, Spectrum Sensing, Energy Detection, Penalized Least Squares, Generalized Cross Validation, Primary User, Secondary User.

I. Introduction

The exponential growth in the wireless technology demands enhancement in spectrum utility. Federation Communication Commission (FCC) has declared different radio spectrum for specified application. According to FCC most of the available spectrum is underutilized [1]. Thus, there is a need to reduce spectrum scarcity and utilize the under utilized spectrum in an efficient manner. CR network solves the issue of spectrum under utilization. CR is a smart radio that is programmed dynamically to detect spectral holes and allocate it to secondary users. Most of the spectrum sensing techniques for CR available in literature are broadly classified as non-coperative spectrum sensing and coperative spectrum sensing [2]. Some of the existing techniques are matched filter [3], cyclostationary feature detection [4], covariance detection [5] and energy detection [6]. The pro's and con's of different existing methods are discussed in [7]. In all those methods Energy detection method is simple as they do not require prior knowledge of the Primary User (PU) and has simple hardware implementation. The challenges in energy detection method is poor performance under low Signal to Noise Ratio (SNR) environment due to noise uncertainty.

In practical situations the signals are distorted due to additive noise and multipath channels. To determine threshold accurately one needs to have exact knowledge of noise power. But in practice, noise power may vary with geographic location and time. Therefore it may not be feasible to compute exact noise power at a particular time and location. Thus, the estimation of noise variance becomes critical for detection of PU signals under low SNR. Since, threshold strongly depends on the accurate estimation of noise variance which changes temporally and spatially [8]. Also, SNR walls define that energy detection becomes unreliable in low SNR conditions [9]. The effect of noise variance estimation in energy detection method is explained through Receiver Operating Characteristic (ROC) curves as mentioned in [10]. The noise variance estimation methods for SS in energy detection method are limited in literature [11] [12]. Yule-walker equations and iterative methods with multiple measurements of a signal from the same radio spectrum have also been used to estimate noise variance from the received signals [13].

In this paper, a low computational complexity of a noise variance estimation method is proposed for spectrum sensing in Additive White Gaussian Noise (AWGN) and fading channels. The noise variance is estimated by minimizing the error between the estimated signal and received signal, allowing penalized smoothed data. The received signals are smoothed with the second order difference, using the DCT with low computational complexity of order $O(n \log n)$ and GCV to identify the best value of smoothing parameter.

The rest of this paper is organized as follows: Section II discusses the system model for CR depending on the different hypothesis. The proposed noise variance estimation algorithm is detailed in section III. Section IV discusses the simulations in detail using AWGN and Rayleigh fading channel to get receiver operating characteristics at low SNR, followed with conclusions in section V.

II. System Model

The transmitted signal is received through multipath channels at the receiver. The multipath channel affect the spectrum of the received signals. The received signal is modeled as

$$y(n) = \begin{cases} w(n) & \text{if } H_0 \\ hx(n) + w(n) & \text{if } H_1 \end{cases}$$

where $x(n)$ is a primary user signal at the n^{th} instant, $w(n)$ is the AWGN which is assumed to be zero mean and variance σ_w^2. Here channel properties are represented by h. For AWGN channel, h is assumed to be one. A binary hypothesis test is considered to detect the presence and absence of the PU signal. Where H_0 and H_1 represent the hypothesis for the presence or absence of the primary user signal. In energy detection method the energy of the received signal is computed as the test statistics using (1) which is sum of the absolute square of the received signal. Thus test statistic is compared with the predefined threshold λ which is a function of noise variance and probability of false alarm to detect the presence or absence of the primary signal.

$$r(n) = \sum_{n=1}^{N} |y(n)|^2 \tag{1}$$

Where $r(n)$ is the computed test statistic. The test statistic follow the chi-square distribution with N degrees of freedom. The probability distribution $f(r)$ of the test statistic is given below [14]

$$f(r) = \begin{cases} \frac{1}{2^{N/2}\Gamma(\frac{N}{2})} r^{\frac{N}{2}-1} exp(-\frac{1}{2}r) & \text{if } H_0 \\ \frac{1}{2}(\frac{r}{\alpha})^{\frac{N-2}{4}} exp[-\frac{1}{2}(r+\alpha)] I_{\frac{N}{2}-1}(\sqrt{\alpha r}) & \text{if } H_1 \end{cases}$$

Where α is a non-centrality parameter which is defined as $\alpha = 2\gamma$ [15]. $I_p(u)$ is a p^{th}-order modified Bessel function.

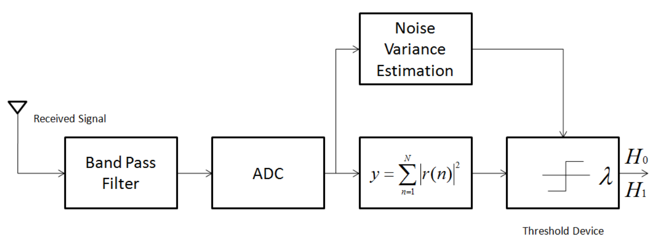

Fig. 1: Block diagram of an Energy Detection Method

The conventional energy detection method with noise variance estimation is shown in Fig 1. The noise variance is calculated from the received signal dynamically to get threshold and steps followed are detailed in the algorithm 1. The expressions for probability of false alarm ($P_f a$) and probability of detection (P_d) is computed from test statistic $r(n)$ is shown below.

$$P_d = Pr(r > \lambda/H_1) = Q_{N/2}(\sqrt{2\gamma}, \sqrt{\lambda}) \tag{2}$$

$$P_{fa} = Pr(r > \lambda/H_0) = \frac{\Gamma(\frac{N}{2}, \frac{\lambda}{2})}{\Gamma(\frac{N}{2})} \tag{3}$$

Where γ is a SNR and Q_u is the genaralaized marcum Q-function. $\Gamma(.)$ and $\Gamma(.,.)$ are the gamma and incomplete gamma functions respectively. Due to central limit theorem, the chi- square distribution is approximated to Gaussian distribution when the number of samples are large. Here, we discuss the probability distributions of received signal for AWGN and Rayleigh channel as given below.

*a) **AWGN channel**:* In communication system AWGN is unavoidable. This noise is added to the receiver structure due to electronic components. The expressions of P_d and P_{fa} for the AWGN channel is given below.

$$P_{fa} = Q(\frac{\lambda - N{\sigma_n}^2}{\sqrt{N}{\sigma_n}^2}) \tag{4}$$

$$P_d = Q(\frac{\lambda - N{\sigma_n}^2(1+\gamma)}{\sqrt{N(1+2\gamma){\sigma_n}^2}}) \tag{5}$$

where Q(.) is the tail probability of standard Normal distribution and it is defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \mathrm{e}^{-\frac{u^2}{2}} \,\mathrm{d}u \tag{6}$$

*b) **Rayleigh fading channel**:* Fading is a common phenomenon in wireless communication system due to scattering from objects that block the transmitted signals. The fading channels comprise of multipath scattering effects, time dispersion, and Doppler-shift that is included because the relative motion between the transmitter and receiver. Most of the multipath channels are modeled with the Rayleigh as they do not consider line of sight propagation which is a more practical channel in wireless communication. As we know that the transmitted signal will reach the receiver via multipath, which will add either constructively or destructively that depends on the phase variations of multipaths of the received signal, the fading model may undergo either flat fading or frequency selective fading. Rayleigh distribution is expressed in terms of SNR as [16]

$$f(\gamma) = \frac{1}{\overline{\gamma}} \exp(-\frac{\gamma}{\overline{\gamma}}) \quad \gamma > 0 \tag{7}$$

Probability of detection of Rayleigh channels are evaluated from the P_d of AWGN channel as

$$\overline{P}_{d,r} = \int_{0}^{\infty} \mathrm{P_d f(\gamma)} \,\mathrm{d}\gamma \tag{8}$$

The simplified expression of (8) is as shown in (9).

$$\begin{aligned} \overline{P}_{d,r} \quad = \quad & \exp(-\frac{\lambda}{2}) \sum_{n=0}^{\frac{N}{2}-2} \frac{1}{n!}(\frac{\lambda}{2})^n \\ & +(\frac{1+\overline{\gamma}}{\overline{\gamma}})^{\frac{N}{2}-1}[\exp(-\frac{\lambda}{2(1+\overline{\gamma})}) \\ & -\exp(-\frac{\lambda}{2}) \sum_{n=0}^{\frac{N}{2}-2} \frac{1}{n!}(\frac{\lambda\overline{\gamma}}{2(1+\overline{\gamma})})^n] \end{aligned} \tag{9}$$

III. Proposed Noise Variance Estimation and Optimal Threshold:

Threshold is selected based on the constant probability of false alarm rate. i.e, for AWGN channel and Rayleigh fading channel it is expressed from equation (4).

$$\lambda = Q^{-1}(P_{fa})\sqrt{N}\sigma^2 + N\sigma^2 \tag{10}$$

where σ^2 is the noise variance. The estimated noise variance $\hat{\sigma_n}^2$ becomes crucial to determine the threshold for detection of PU signal. In the proposed method we estimate the noise variance $\hat{\sigma_n}^2$, so the threshold becomes

$$\lambda^* = Q^{-1}(P_{fa})\sqrt{N}\hat{\sigma_n}^2 + N\hat{\sigma_n}^2 \tag{11}$$

In the low SNR, the optimum threshold for AWGN channel and fading channel is given in [17] by minimizing the probability of error.

$$\lambda = N\sigma_n^2 \tag{12}$$

Substitute equation (12) in the (4) leads to $P_{fa} = 0.5$ which is not preferable in practical condition.

*c) **Noise variance estimation**:* The noise variance estimation is calculated from the idea proposed in [18]. This is based on the discrete cosine transform, which makes robust smoothing of equally spaced data in one and higher dimensions. The author [18] proposed that noise variance can be estimated by modeling a system as

$$y = \hat{y} + \epsilon \tag{13}$$

Where ϵ represents a Gaussian noise with mean zero and unknown variance, and $\hat{y}$ is the smoothed data of y.
Which is further formulated as

$$y = hx + w \tag{14}$$

Where h is the channel response, it is considered as unity in AWGN case. In flat fading channels this is a weighted scalar. Where x is the transmitted data at the transmitter.

The best estimate of $\hat{y}$ is depends on the smoothness of y. Data smoothing is widely carried out by means of parametric and non-parametric regression. Parametric regression requires prior knowledge about regression equations that can represent data well, therefore non-parametric regression is the best option for smoothing of data [19]. The most common approaches to non-parametric regression used in data processing includes the kernel regression like moving average and local polynomial regression. Another classical approach to smoothing is the penalized least square which has been extensively studied in [20]. Here, the idea is to minimize the residual sum-of-squares (RSS) and a cost term (C) that shows irregularities of the smooth data. The objective is to minimize the cost function which is formulated as

$$F(\hat{y}) = RSS + sC(\hat{y}) = ||y - \hat{y}||^2 + sC(\hat{y}) \tag{15}$$

Where s is the smoothing parameter, which has to be positive. The smoothness parameter is directly proportional to smoothness of estimated data. Where C shows the irregularity parameter which is function of second order divided difference for one-dimensional signals as shown in (16).

$$C(\hat{y}) = ||K\hat{y}||^2 \tag{16}$$

where K is a tridiagonal square matrix defined as

$$K_{i,i-1} = \frac{2}{t_{i-1}(t_{i-1}+t_i)}, K_{i,i} = \frac{-2}{t_{i-1}t_i}, K_{i-1,i} = \frac{2}{t_i(t_{i-1}+t_i)}$$

for $2 \leq i \leq n-1$ where n is the number of elements in the $\hat{y}$ and t_i indicates the difference between $\hat{y}_i$ and $\hat{y}_{i+1}$. To get smooth response of the system, we minimize equation (15) with respect to $\hat{y}$

$$\hat{y} = (I_n + sK^TK)^{-1}y \tag{17}$$

Where I_n is the $n \times n$ Identity matrix and $I_n + sK^TK$ is a symmetric pentadiagonal matrix. If the smoothing parameter is chosen properly, the estimation accuracy of $\hat{y}$ is increased. The best estimation of smoothing parameter allows an accurate estimation of the original data. The smoothing parameter is obtained by minimizing GCV score with respect to s. The GCV was introduced in [21] for smoothing spline which assumes the linear equation as follows

$$\hat{y} = R(s)y \tag{18}$$

Where $R(s) = (I_n + sK^TK)^{-1}$. The GCV method chooses the smoothing parameter that minimize the GCV score.

$$s = argmin(GCV)$$

The GCV score is calculated as.

$$GCV(s) = \frac{RSS/n}{(1 - Tr(R)/n)^2} \tag{19}$$

To compute the GCV score, the author [22] have proposed algorithm that involve the determination of R through continual methods of minimization at each step. Such a process is very time-elapsing and can be avoided. Indeed, $Tr(R)$ is modified as [18]

$$Tr(R) = \sum_{i=1}^{n} \frac{1}{1 + s{\lambda_i}^2} \tag{20}$$

Where $\lambda_{i=1,2..n}^2$ are the eigen values of K^TK and GCV score is given as

$$GCV(s) = \frac{n\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{(n - \sum_{i=1}^{n}(1 + s{\lambda_i}^2)^{-1})^2} \tag{21}$$

Computing the s value by minimizing the GCV score using the equation (21) makes the smoothing algorithm completely automated. Because the elements of $\hat{y}$ appear in the GCV score expression, $\hat{y}$ has to be calculated at each step of the minimization process. This can be avoided for equally spaced data and is explained below. For evenly spaced data, the tridiagonal square matrix can also be obtained efficiently using the matrix with $t_i = 1$. Then the tridiagonal matrix K is represented as

$$K = \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{bmatrix} \tag{22}$$

This matrix K can be further Eigen decomposed into $K = U \wedge U^{-1}$ [23]. Where $\wedge$ is a diagonal matrix containing the eigenvalues of K. It is represented as

$$\begin{aligned} \wedge &= diag(\lambda_1, \lambda_2, ..., \lambda_n) \\ with \quad \lambda_i &= -2 + 2cos[(i-1)\pi/n] \end{aligned} \tag{23}$$

Where U is a unitary matrix ($UU^T = I_n$) of order $n \times n$ and it is a type-2 Discrete cosine transform.
i.e,

$$\begin{aligned} \hat{y} &= (I_n + sK^TK)^{-1}y \\ \hat{y} &= U(I_n + s\wedge^2)^{-1}U^Ty = U\Gamma U^Ty \end{aligned} \tag{24}$$

The Γ can be represented in matrix format having diagonal elements $\Gamma_{i,i} = [1 + s(2 - 2cos((i-1)\pi/n)^2)]^{-1}$ and other elements $\Gamma_{i,j} = 0 \; i,j \neq 0$ then smoothed output is given as

$$\hat{y} = IDCT(\Gamma(DCT(y))) \tag{25}$$

The RSS is expressed in terms of DCT as follows

$$RSS = \sum_{i=1}^{n}(\frac{1}{1 + s{\lambda_i}^2} - 1)^2 DCT_i^2(y) \tag{26}$$

Then the GCV score is obtain by substituting RSS value

$$GCV(s) = \frac{\sum_{i=1}^{n}(\frac{1}{1+s\lambda_i} - 1)^2 DCT_i^{\,2}(y)}{(n - \sum_{i=1}^{n}\frac{1}{1+s\lambda_i{}^2})^2} \tag{27}$$

The above equation (25) required less number of computation with order of complexity is $O(n \log n)$ than the method used in (18) with the order of complexity is $O(n^3)$. The complexity of conventional Maximum Likelihood and Least Squares is $O(n)$. The proposed noise variance estimation algorithm is given below.

Algorithm 1 Proposed Algorithm

1: *y=Received signal at the Receiver*
2: *$\hat{y}$=Estimated signal at the Receiver*
3: *s=Smoothing factor*:
4: *Read the data y*
5: *Transform data DCT(y)*
6: *Compute* $\Gamma = \frac{1}{1+s\lambda_i{}^2}$ *i=1,2,...,n*
7: *Determine RSS by using (26)*
8: *Minimization of GCV score by using(27) with respect to s*

IV. Simulation Results

In this section simulation results obtained through MATLAB are illustrated. The results depict the comparison of our method (PLS) with least squares (LS) and maximum likelihood estimation (ML) for noise variance estimation in SS. The experimentation is performed, for 5000 Monte Carlo runs, of the user signal using BPSK modulation with a signal strength of unity. For the Rayleigh case 2-tap channel has been considered with zero mean Gaussian random variable with the exponential power delay profile of $E|h_l|^2 = Aexp(-l)$. Where, A is a parameter to guarantee $E|h_0|^2 + E|h_1|^2 = 1$. The threshold for both AWGN and Rayleigh channels is calculated using Eq.11.

Steps for Simulation:

A) Probability detection over a AWGN channel (h=1)

1) Generate a BPSK signal $x(n)$ with 1 and -1 with N samples and Noise is generated with mean zero and variance as a $\sigma_w^2 = E_x/\gamma$.
2) Received signal is $y(n) = h * x(n) + w(n)$.
3) Noise variance calculated through the proposed algorithm is $\hat{\sigma}_w^2$.
4) Threshold is calculated for each value of P_{fa}.
5) Compare the E_y with the threshold detection which is greater than one can be incremented by one.
6) The above steps from $1-4$ are repeated for 5000 times.

B) Probability detection over a Rayleigh fading channel

1) Generate a Rayleigh fading channel $h = \sqrt{X_1^2 + X_2^2}$ where X_1 and X_2 are two Gaussian signal having mean zero and half variance.
2) Next steps are repeated $1-5$ from subsection A of section-IV.

From the figures it can be clearly seen that our method has better noise estimation capability when compared to LS and ML. From Fig.2 and Fig.4 is also inferred that the performance of Rayleigh channel is higher compared to the AWGN channel. From Fig.2 and Fig.3 is observed that number of samples increases the performance of ROC curves also increased.

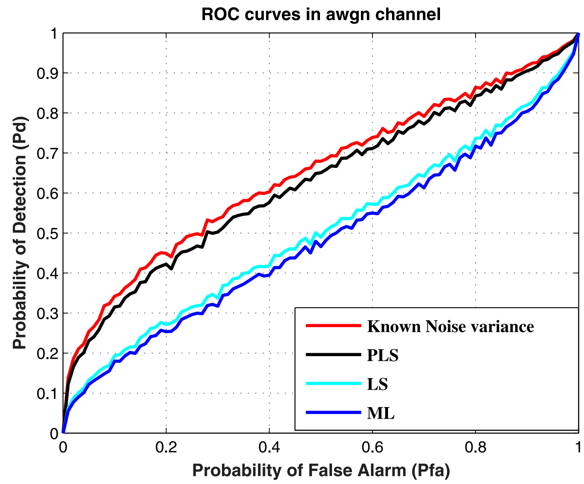

Fig. 2: ROC curves in AWGN with estimated noise power at SNR -15dB with N=500

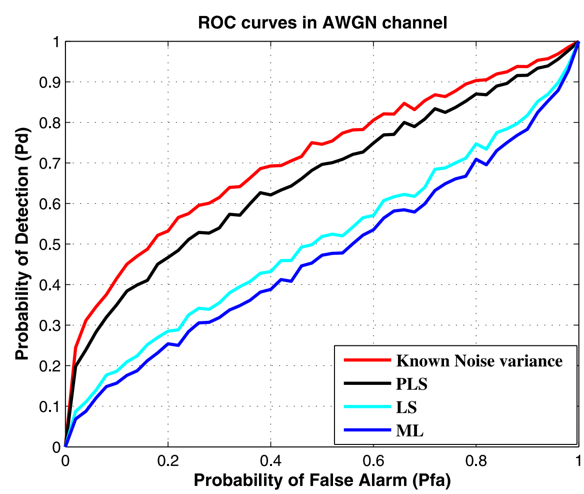

Fig. 3: ROC curves in AWGN with estimated noise power at SNR -15dB with N=1000

V. Conclusion

In this paper noise variance estimation method was proposed in AWGN and fading channel under low SNR condition. The noise variance was estimated by smoothing of received signal. The smoothing factor was determined by minimizing the GCV score. The proposed method has computational complexity more than that of ML and LS but gives better performance for noise variance estimation with low estimated error. The proposed method has a trade-off between computational complexity and accuracy. This can be extended to other fading channels in future.

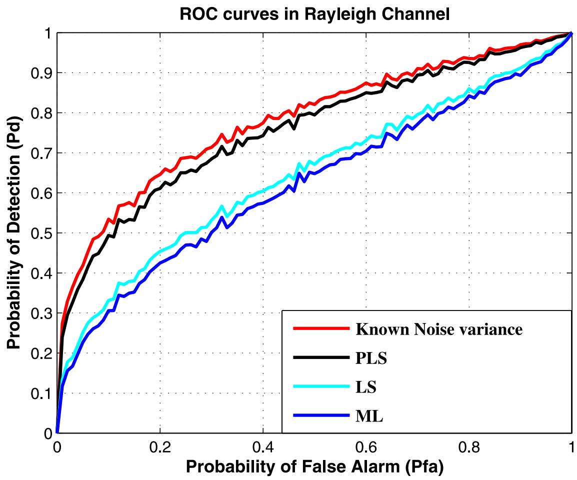

Fig. 4: ROC curves in Rayleigh with estimated noise power at SNR -15dB with N=500

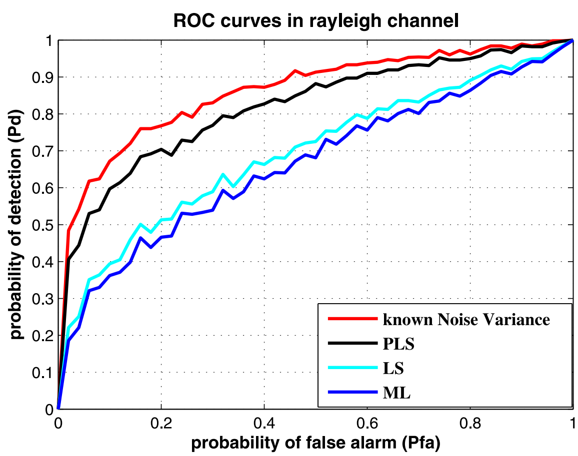

Fig. 5: ROC curves in Rayleigh with estimated noise power at SNR -15dB with N=1000

References

[1] Fedaral Communication Commission,"Spectrum Policy Task Force," Rep. ET Docket no.02-135, Nov–2002.

[2] Ian F. Akyildiz, Brandon F. Lo and Ravikumar Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," Physical Communication 4 (2011) 40-62.

[3] H. S. Chen, W. Gao, and D. G. Daut, Signature based spectrum sensing algorithms for IEEE 802.22WRAN, in Proc. IEEE Int. Conf. Commun. (ICC), Glasgow, Scotland, Jun. 2007, pp. 6487-6492.

[4] P. D. Sutton, K. E. Nolan, and L. E. Doyle, Cyclostationary signature in practical cognitive radio applications, IEEE J. Sel. Areas Commun., vol. 26, no. 1, pp. 13-24, 2008.

[5] Yonghong Zeng and Y.C. Liang, "Covariance Based Signal Detections for Cognitive Radio," New Frontiers in Dynamic Spectrum Access Networks, 2007.

[6] H. Urkowitz, Energy detection of unknown deterministic signals, Proc. IEEE, vol. 55, no. 4,pp. 523-531, Apr. 1967.

[7] Tevfik Yucek and Huseyin Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications," IEEE Communication Surveys & Tutorial, 2009.

[8] Umar, Raza, and Asrar UH Sheikh. "A comparative study of spectrum awareness techniques for cognitive radio oriented wireless networks." Physical Communication 9 (2013): 148-170.

[9] Tandra, Rahul, and Anant Sahai. "SNR walls for signal detection." Selected Topics in Signal Processing, IEEE Journal of 2.1 (2008): 4-17.

[10] Pawan Dhakal, Daniel Rivello and Federico Penna, "Impact of Noise Estimation on Energy Detection and Eigenvalue Based Spectrum Sensing Algorithms," IEEE ICC 2014

[11] Adeel Ahmed, Yim Fun Hu and James M. Noras "Noise Variance Estimation for Spectrum Sensing in Cognitive Radio Networks," 2014 AASRI Conference on Circuit and Signal Processing, Volume 9, 2014, Pages 3743.

[12] Vincent Savaux, Mose Djoko-Kouam and Yves Lout1and Alexandre Skrzypczak. "Application of a Channel Estimation Algorithm to Spectrum Sensing in a Cognitive Radio Context," International Journal of Antennas and Propagation 2014.

[13] Deepak R. Joshi and Dimitrie C. Popescu Octavia A. Dobre. "Adaptive Spectrum Sensing with Noise Variance Estimation for Dynamic Cognitive Radio Systems," Information Sciences and Systems (CISS), 2010.

[14] Steven Kay, "Fundamentals of Statistical Signal Processing: Detection Theory," Volume 2. Prentice Hall 1998.

[15] Fadel F. Digham, Mohammed-Slim Alouini and Marvin K.simmon, "On the Energy Detection of Unknown Signals over Fading Channels," IEEE Trans. on Communications,vol. 55, No. 1, January 2007.

[16] David Tse and Pramod Viswanath, "Fundamentals of Wireless Communication," Cambridge University Press, 2005.

[17] Saman Atapattu, Chintha Tellumbara, and Hai Jaing, "Spectrum Sensing via Energy Detector in Low SNR," IEEE ICC 2011.

[18] Damien Garcia, "Robust smoothing of gridded data in one and higher dimensions with missing values," Computational Statistics & Data Analysis Volume 54, Issue 4, 1 April 2010, Pages 1167 – 1178

[19] Takezawa, K. Introduction to nonparametric regression. John Wiley & Sons, Inc; Hoboken, NJ: 2005

[20] Wahba, Grace. Spline models for observational data. Vol. 59. Siam, 1990.

[21] Gene H. Golub, Michael Heath and Grace Wahba "Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter," Technometrics, Vol. 21, N0. 2, MAY 1979.

[22] Weinert, Howard L. "Efficient computation for WhittakerHenderson smoothing." Computational Statistics & Data Analysis 52.2 (2007): 959-974.

[23] Yueh WC Eigenvalues of several tridiagonal matrices. Applied Mathematics E-Notes. 2005; 5:6674.