Cooperative Spectrum Sensing Using Random Matrix Theory

Leonardo S. Cardoso and Merouane Debbah and Pascal Bianchi
SUPELEC
Gif-sur-Yvette, France
{leonardo.cardoso, merouane.debbah and pascal.bianchi}@supelec.fr

Jamal Najim
ENST
Paris, France
jamal.najim@enst.fr

***Abstract*—In this paper, using tools from asymptotic random matrix theory, a new cooperative scheme for frequency band sensing is introduced for both AWGN and fading channels. Unlike previous works in the field, the new scheme does not require the knowledge of the noise statistics or its variance and is related to the behavior of the largest and smallest eigenvalue of random matrices. Remarkably, simulations show that the asymptotic claims hold even for a small number of observations (which makes it convenient for time-varying topologies), outperforming classical energy detection techniques.**

I. Introduction

It has already become a common understanding that current mobile communication systems do not make full use of the available spectrum, either due to sparse user access or to the system's inherent deficiencies, as shown by a report from the Federal Communications Commission (FCC) Spectrum Policy Task Force [1]. It is envisioned that future systems will be able to opportunistically exploit those spectrum 'left-overs', by means of knowledge of the environment and cognition capability, in order to adapt their radio parameters accordingly. Such a technology has been proposed by Joseph Mitola in 2000 and is called cognitive radio [2]. Due to the fact that recent advances on micro-electronics and computer systems are pointing to a -not so far- era when such radios will be feasible, it is of utmost importance to develop good performing sensing techniques.

In its simplest form, spectrum sensing means looking for a signal in the presence of noise for a given frequency band (it could also encompass being able to classify the signal). This problem has been extensively studied before, but it has regained attention now as part of the cognitive radio research efforts. There are several classical techniques for this purpose, such as the energy detector (ED) [3]–[5], the matched filter [6] and the cyclostationary feature detection [7]–[9]. These techniques have their strengths and weaknesses and are well suited for very specific applications.

Nevertheless, the problem of spectrum sensing as seen from a cognitive radio perspective, has very stringent requirements and limitations, such as,

- no prior knowledge of the signal structure (statistics, noise variance value, etc...);
- the detection of signals in the shortest time possible;
- ability to detect reliably even over heavily faded environments;

The works by Cabric et al. [7], Akyildiz et al. [10] and Haykin [11] provide a summary of these classical techniques from the cognitive network point of view. It is clear from these works, that none can fully cope with all the requirements of the cognitive radio networks.

In simple AWGN (Additive White Gaussian Noise) channels, most classical approaches perform very well. However, in the case of fast fading, these techniques are not able to provide satisfactory solutions, in particular to the hidden node problem [12]. To this end, several works [13]–[16] have looked into the case in which cognitive radios cooperate for sensing the spectrum. These works aim at reducing the probability of false alarm by adding extra redundancy to the sensing process. They also aim at reducing the number of samples collected, and thus, the estimation times by the use parallel measuring devices. However, even though one could exploit the spatial dimension efficiently, these works are based on the same fundamental techniques, which require a priori knowledge of the signal.

In this work, we introduce an alternative method for blind (in the sense that no a priori knowledge is needed) spectrum sensing. This method relies on the use of multiple receivers to infer on the structure of the received signals using random matrix theory (RMT). We show that we can estimate the spectrum occupancy reliably with a small amount of received samples.

The remainder of this work is divided as follows. In section II, we formulate the problem of blind spectrum sensing. In section III, we introduce the proposed approach based on random matrix theory. In section IV, we present some practical results which confirm that the asymptotic assumptions hold even for a small amount of samples. Then, in section V, we show the performance results of the proposed method. Finally, in section VI, we draw the main conclusions and point out further studies.

II. Problem Formulation

The basic problem concerning spectrum sensing is the detection of a signal within a noisy measure. This turns out to be a difficult task, especially if the received signal power is very low due to pathloss or fading, which in the blind spectrum sensing case is unknown. The problem can be posed as a hypothesis test such that [3]:

978-1-4244-1653-0/08/$25.00 © 2008 IEEE

$$y(k) = \begin{cases} n(k): & H_0 \\ h(k)s(k) + n(k): & H_1 \end{cases}, \qquad (1)$$

where $y(k)$ is the received vector of samples at instant k, $n(k)$ is a noise (not necessarily gaussian) of variance σ^2, $h(k)$ is the fading component, $s(k)$ is the signal which we want to detect, such that $E\left[\mid s(k) \mid^2\right] \neq 0$, and H_0 and H_1 are the noise-only and signal hypothesis, respectively. We suppose that the channel h stays constant during N blocks ($k = 1..N$).

Classical techniques for spectrum sensing based on energy detection compare the signal energy with a known threshold V_T [3]–[5] derived from the statistics of the noise and channel. The following is considered to be the decision rule

$$\text{decision} = \begin{cases} H_0, & \text{if } E\left[\mid y(k) \mid^2\right] < V_T \\ H_1, & \text{if } E\left[\mid y(k) \mid^2\right] \geq V_T \end{cases},$$

where $E[\mid y(k) \mid^2]$ is the energy of the signal and V_T is usually taken as the noise variance. One drawback of this approach is that neither the noise/channel distribution nor V_T are known a priori. In real life scenarios V_T depends on the radio characteristics and is hard to be estimated properly. Moreover, in the case of fading and path loss, the energy of the received signal can be of the order of the noise, making it difficult to be detected all the more as the number of samples N may be very limited. Indeed, $E\left[\mid y(k) \mid^2\right]$ is estimated by

$$\frac{1}{N}\sum_{k=1}^{N} \mid y(k) \mid^2,$$

which is not a good estimator for the small sample size case.

In the following, we provide a cooperative approach for cognitive networks to detect the signal from a primary system without the need to know the noise variance using results from random matrix theory.

III. Random Matrix Theory for Spectrum Sensing

Consider the scenario depicted in Figure 1, in which primary users (in white) communicate to their dedicated (primary) base station. Secondary base stations $\{BS_1, BS_2, BS_3, ..., BS_K\}$ are cooperatively sensing the channel in order to identify a white space and exploit the medium.

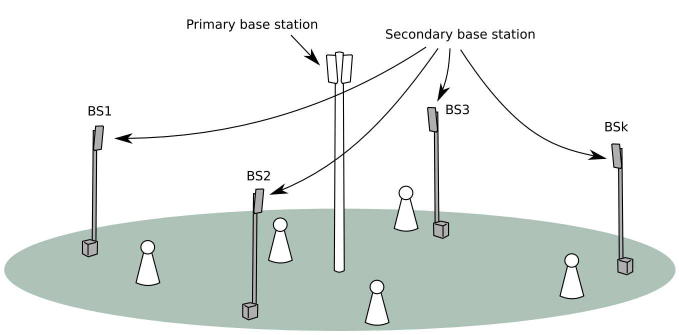

Fig. 1. Considered scenario for spectrum sensing.

Before going any further, let us assume the following:

- The K base stations in the secondary system share information between them. This can be performed by transmission over a wired high speed backbone.
- The base stations are analyzing the same portion of the spectrum.

Let us consider the following $K \times N$ matrix consisting of the samples received by all the K secondary base stations ($y_i(k)$ is the sample received by base station i at instant k):

$$\mathbf{Y} = \begin{bmatrix} y_1(1) & y_1(2) & \cdots & y_1(N) \\ y_2(1) & y_2(2) & \cdots & y_2(N) \\ y_3(1) & y_3(2) & \cdots & y_3(N) \\ \vdots & \vdots & & \vdots \\ y_K(1) & y_K(2) & \cdots & y_K(N) \end{bmatrix}.$$

The goal of the random matrix theory approach is to perform a test of independence of the signals received by the various base stations. Indeed, in the presence of signal (H_1 case), all the received samples are correlated, whereas when no signal is present (H_0 case), the samples are decorrelated whatever the fading situation. Hence, in this case, for a fixe K and $N \to \infty$, the sample covariance matrix $\frac{1}{N}\mathbf{Y}\mathbf{Y}^H$ converge $\sigma^2\mathbf{I}$. However, in practice, N can be of the same order of magnitude than K and therefore one can not infer directly $\frac{1}{N}\mathbf{Y}\mathbf{Y}^H$ independence of the samples. This can be formalized using tools from random matrix theory [17]. In the case where the entries of $\mathbf{Y}$ are independent (irrespectively of the specific probability distribution, which corresponds to the case where no signal is transmitted - H_0) results from asymptotic random matrix theory [17] state that:

Theorem. Consider an $K \times N$ matrix $\mathbf{W}$ whose entries are independent zero-mean complex (or real) random variables with variance $\frac{\sigma^2}{N}$ and fourth moments of order $O(\frac{1}{N^2})$. As $K, N \to \infty$ with $\frac{K}{N} \to \alpha$, the empirical distribution of $\mathbf{W}\mathbf{W}^H$ converges almost surely to a nonrandom limiting distribution with density

$$f(x) = (1 - \frac{1}{\alpha})^+ \delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi\alpha x}$$

where

$$a = \sigma^2(1 - \sqrt{\alpha})^2 \qquad \text{and} \qquad b = \sigma^2(1 + \sqrt{\alpha})^2.$$

Interestingly, when there is no signal, the support of the eigenvalues of the sample covariance matrix (in Figure 2, denoted by M̌P) is finite, whatever the distribution of the noise. The Marchenko-Pastur law thus serves as a theoretical prediction under the assumption that matrix is "all noise". Deviations from this theoretical limit in the eigenvalue distribution should indicate non-noisy components i.e they should suggest information about the matrix.

In the case in which a signal is present (H_1), $\mathbf{Y}$ can be rewritten as

$$\mathbf{Y} = \begin{bmatrix} h_1 & \sigma & & 0 \\ \vdots & & \ddots & \\ h_K & 0 & & \sigma \end{bmatrix} \begin{bmatrix} s(1) & \cdots & s(N) \\ z_1(1) & \cdots & z_1(N) \\ \vdots & & \vdots \\ z_K(1) & \cdots & z_K(N) \end{bmatrix},$$

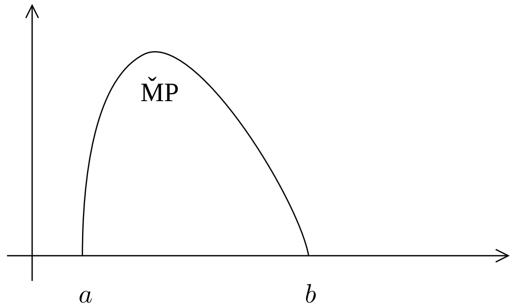

Fig. 2. The Marchenko-Pastur support (H_0 hypothesis).

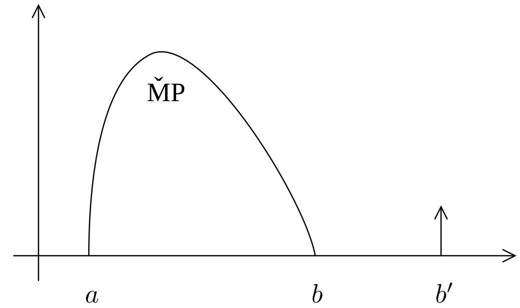

Fig. 3. The Marchenko-Pastur support plus a signal component.

where $s(i)$ and $z_k(i) = \sigma n_k(i)$ are respectively the independent signal and noise with unit variance at instant i and base station k. Let us denote by $\mathbf{T}$ the matrix:

$$\mathbf{T} = \begin{bmatrix} h_1 & \sigma & & 0 \\ \vdots & & \ddots & \\ h_K & 0 & & \sigma \end{bmatrix}.$$

$\mathbf{TT}^H$ has clearly one eigenvalue $\lambda_1 = \sum |h_i|^2 + \sigma^2$ and all the rest equal to σ^2. The behavior of the eigenvalues of $\frac{1}{N}\mathbf{YY}^H$ is related to the study of the eigenvalue of large sample covariance matrices of spiked population models [18]. Let us define the signal to noise ratio (SNR) ρ in this work as

$$\rho = \frac{\sum |h_i|^2}{\sigma^2}.$$

Recent works of Baik et al. [18], [19] have shown that, when

$$\frac{K}{N} < 1 \qquad \text{and} \qquad \rho > \sqrt{\frac{K}{N}} \tag{2}$$

(which are assumptions that are clearly met when the number of samples N are sufficiently high), the maximum eigenvalue of $\frac{1}{N}\mathbf{YY}^H$ converges almost surely to

$$b' = (\sum |h_i|^2 + \sigma^2)(1 + \frac{\alpha}{\rho}),$$

which is superior to $b = \sigma^2(1+\sqrt{\alpha})^2$ seen for the H_0 case.

Therefore, whenever the distribution of the eigenvalues of the matrix $\frac{1}{N}\mathbf{YY}^H$ departs from the Marchenko-Pastur law (Figure 3), the detector knows that the signal is present. Hence, one can use this interesting feature to sense the spectrum.

Let λ_i be the eigenvalues of $\frac{1}{N}\mathbf{YY}^H$ and $G = [a, b]$, the cooperative sensing algorithm works as follows:

A. *Noise distribution unknown, variance known*

In this case, the following criteria is used:

$$\text{decision} = \begin{cases} H_0: & \text{, if } \lambda_i \in G \\ H_1: & otherwise \end{cases} \tag{3}$$

Note that refinements of this algorithm (where the probability of false alarm is taken into account in the non-asymptotic case) can be found in [20]. The results are based on the computation of the asymptotic largest eigenvalue distribution in the H_0 and H_1 case.

B. *Both noise distribution and variance unknown*

Note that the ratio of the maximum and the minimum eigenvalues in the H_0 hypothesis case does not depend on the noise variance. Hence, in order to circumvent the need for the knowledge of the noise, the following criteria is used:

$$\text{decision} = \begin{cases} H_0: & \text{, if } \frac{\lambda_{\max}}{\lambda_{\min}} \leq \frac{(1+\sqrt{\alpha})^2}{(1-\sqrt{\alpha})^2} \\ H_1: & otherwise \end{cases} \tag{4}$$

It should be noted that in this case, one needs to still take a sufficiently high number of samples N such that the conditions in Eq. (2) are met. In other words, the number of samples scales quadratically with the inverse of the signal to noise ratio. Note moreover that the test H_1 provides also a good estimator of the SNR ρ. Indeed, the ratio of largest eigenvalue (b') and smallest (a) of $\frac{1}{N}\mathbf{YY}^H$ is related solely to ρ and α i.e

$$\frac{b'}{a} = \frac{(\rho+1)(1+\frac{\alpha}{\rho})}{(1-\sqrt{\alpha})^2}$$

To our knowledge, this estimator of the SNR has never been put forward in the literature before.

IV. Performance Analysis

The previous theoretical results have shown that one is able to distinguish a signal from noise by the use of only a limiting ratio of the highest to the smallest eigenvalue of the sample covariance matrix. For finite dimensions, the operating region for such an algorithm is still an issue and is related to the asymptotic distribution of a scaling factor of the ratio [20]. This section provides some characterization of this region through the analysis of the ratio between λ_{max} and λ_{min} of $\frac{1}{N}\mathbf{YY}^H$ for various matrix sizes.

Figures 4 and 5 present the $\lambda_{max}/\lambda_{min}$ for various sizes of $\mathbf{Y}$ in the pure noise case, with $\alpha = 1/2$ and $\alpha = 1/10$, respectively. From the figures we see that both cases provide a good approximation of the asymptotic ratio even with small matrix sizes. If one takes, for example, $N = 100$ ($K = 50$ for $\alpha = 1/2$ and $K = 10$ for $\alpha = 1/10$), it can be seen that the simulated cases are respectively equal to 81% percent and 83% of the asymptotic limit for $\alpha = 1/2$ and $\alpha = 1/10$.

As expected, for a larger $\mathbf{Y}$ matrix size, the empirical ratio approaches the asymptotic one.

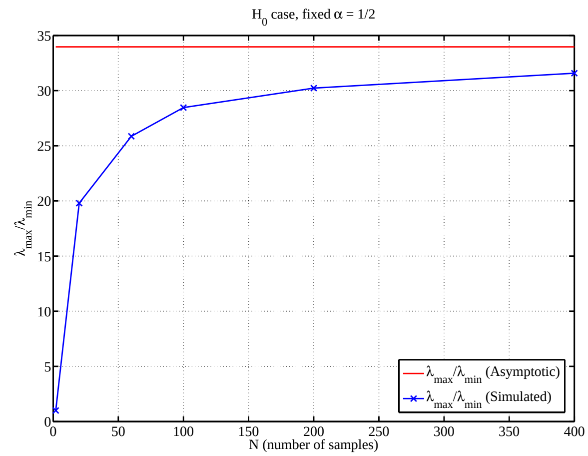

Fig. 4. Behavior of $\lambda_{max}/\lambda_{min}$ for increasing N (case H_0, $\alpha = 1/2$).

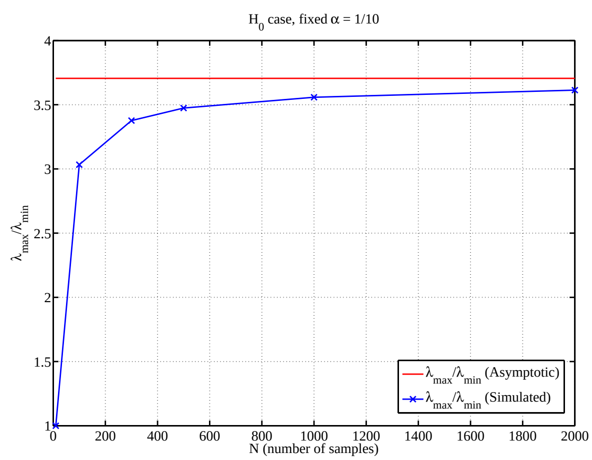

Fig. 5. Behavior of $\lambda_{max}/\lambda_{min}$ for increasing N (case H_0, $\alpha = 1/10$).

Figures 6 and 7 show the behavior of the $\lambda_{max}/\lambda_{min}$ for the signal plus noise case for $\alpha = 1/2$ and $\alpha = 1/10$, respectively. In both cases, $\sigma^2 = 1/\rho$ (with a ρ of -5 dB) with $\sum |h_i|^2 = 1$ (which holds under the criteria in Eq. (2)). In this case, $\frac{\lambda_{max}}{\lambda_{min}} = \frac{b'}{a}$, for the pure signal case. Interestingly, for $N = 100$ ($K = 50$ for $\alpha = 1/2$ and $K = 10$ for $\alpha = 1/10$), it can be seen that the simulated case is approximately 70% percent and 83% of the asymptotic limit for $\alpha = 1/2$ and $\alpha = 1/10$, respectively. As expected, the larger the $\mathbf{Y}$ matrix sizes, the closer one gets to the asymptotic ratio. A good approximation was obtained for values of N as low as 100 samples.

V. RESULTS

Simulations were carried out to establish the performance of the random matrix theory detector scheme in comparison to the cooperative energy detector scheme based on voting [15], [16]. The framework for the energy detector is exposed in section II,

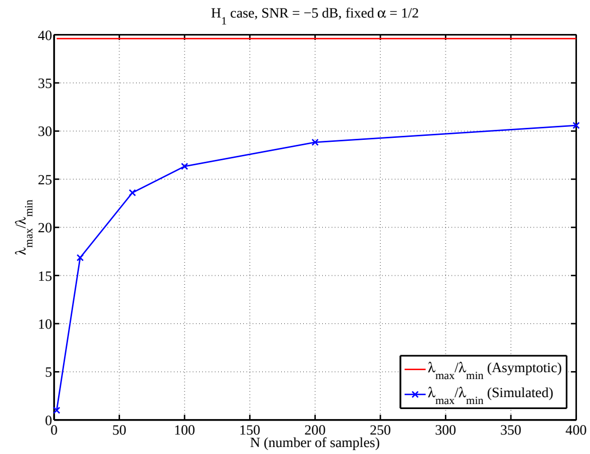

Fig. 6. Behavior of $\lambda_{max}/\lambda_{min}$ for increasing N (case H_1, $\alpha = 1/2$).

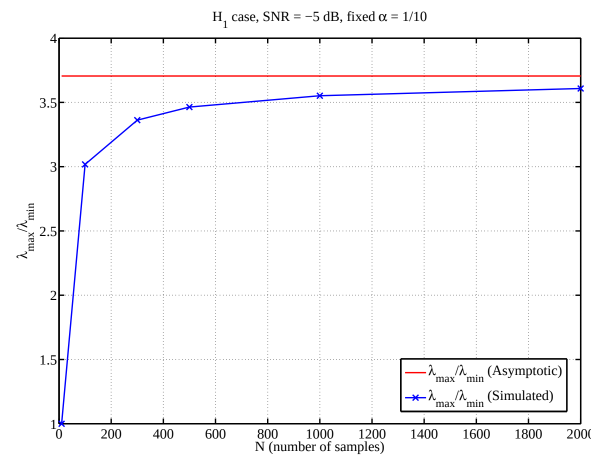

Fig. 7. Behavior of $\lambda_{max}/\lambda_{min}$ for increasing N (case H_1, $\alpha = 1/10$).

with $h(k)$ modeled as a rayleigh multipath fading of variance $1/K$. The variance is normalized to take into account the fact that the energy does not increase without bound as the number of base stations increases due to the path loss. A total of 10 secondary base stations were simulated. For the voting scheme, the decision rule is the following: one considers the overall spectrum occupancy decision to be the one chosen by most of the secondary base stations. The threshold V_T is taken as σ^2 (for the known noise variance case). For the random matrix theory based scheme, a fixed total of ($K = 10$) base stations were adopted. Note that the algorithms can be optimized for the voting and random matrix theory based rules by adopting decision margins [20].

Figure 8 depicts the performance of the energy detector scheme along with the random matrix theory one for $N = \{10, 20, ..., 60\}$ samples and a known noise variance of σ^2 at SNR equal to -5dB. It is important to stress that since K is fixed, α is not constant as in the previous section. As clearly shown, the random matrix theory scheme outperforms the cooperative energy detector case for all number of samples

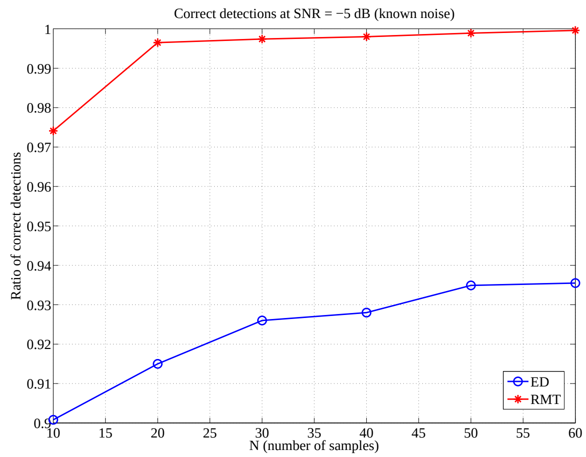

Fig. 8. Comparison between the ED and random matrix theory approach ($\rho = -5\ dB$).

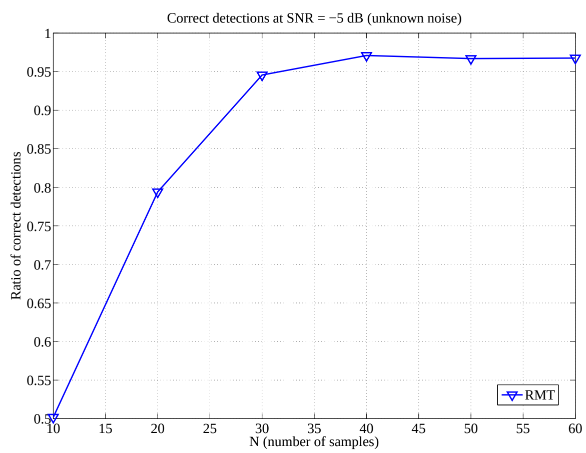

Fig. 9. Random matrix theory approach for an unknown noise variance.

due to its inherent robustness.

Figure 9 plots the performance of the random matrix theory scheme for an unknown noise variance (the voting scheme can not be compared as it relies on the knowledge of the noise variance). One can see that, indeed, even without the knowledge of a noise variance, one is still able to achieve a very good performance for sample sizes greater than 30.

VI. Conclusions

In this paper, we have provided a new spectrum sensing technique based on random matrix theory and shown its performance in comparison to the cooperative energy detector scheme for both a known and unknown noise variance. Remarkably, the new technique is quite robust and does not require the knowledge of the signal or noise statistics. Moreover, the asymptotic claims turn out to be valid even for a very low number of dimensions. The method can be enhanced (see [20]) by adjusting the threshold decision, taking into account the number of samples though the derivation of the probability of false alarm of the limiting ratio of the largest to the smallest eigenvalue.

Acknowledgment

This work was funded by the Alcatel-Lucent Chair in Flexible Radio. The authors would like to thank prof. Walid Hachem for fruitful discussions on the topic.

References

[1] Spectrum Efficiency Working Group. Report of the Spectrum Efficiency Working Group. Technical report, FCC, November 2002.

[2] J. Mitola. *Cognitive Radio An Integrated Agent Architecture for Software Defined Radio*. PhD thesis, Royal Institute of Technology (KTH), May 2000.

[3] H. Urkowitz. Energy detection of unknown deterministic signals. *Proceedings of the IEEE*, 55:523–531, 1967.

[4] V. I. Kostylev. Energy detection of a signal with random amplitude. 2002.

[5] F. F. Digham, M. S. Alouini, and M. K. Simon. On the energy detection of unknown signals over fading channels. 2003.

[6] A. Sahai, N. Hoven, and R. Tandra. Some fundamental limits on cognitive radio. *Allerton Conference on Communication, Control, and Computing*, 2003.

[7] D. Cabric, S. M. Mishra, and R. W. Brodersen. Implementation issues in spectrum sensing for cognitive radios. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, volume 1, pages 772–776 Vol.1, 2004.

[8] A. Fehske, J. Gaeddert, and JH Reed. A new approach to signal classification using spectral correlation and neural networks. *Proc. IEEE Int. Symp. New Frontiers Dynamic Spectr. Access Networks*, 1:144–150.

[9] H. Tang. Some physical layer issues of wide-band cognitive radio systems. *Proc. IEEE Int. Symp. New Frontiers Dynamic Spectr. Access Networks*, 1:151–159.

[10] I. F. Akyildiz, W. Y. Lee, M.C. Vuran, and S. Mohanty. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Comput. Networks*, 50(13):2127–2159, 2006.

[11] S. Haykin. Cognitive radio: brain-empowered wireless communications. *Selected Areas in Communications, IEEE Journal on*, 23(2):201–220, 2005.

[12] C.L. Fullmer and JJ Garcia-Luna-Aceves. Solutions to hidden terminal problems in wireless networks. *Proceedings of the ACM SIGCOMM'97 conference on Applications, technologies, architectures, and protocols for computer communication*, pages 39–49, 1997.

[13] S. M. Mishra, A. Sahai, and R. W. Brodersen. Cooperative sensing among cognitive radios. In *Communications, 2006. ICC '06. IEEE International Conference on*, volume 4, pages 1658–1663, 2006.

[14] G. Ganesan and Y. Li. Cooperative spectrum sensing in cognitive radio networks. In *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pages 137–143, 2005.

[15] A. Ghasemi and E. S. Sousa. Collaborative spectrum sensing for opportunistic access in fading environments. *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pages 131–136, 2005.

[16] A. Ghasemi and E. S. Sousa. Spectrum sensing in cognitive radio networks: the cooperation-processing tradeoff. *Wireless Communications and Mobile Computing - WILEY, 2007.*, pages 1049–1060, 2007.

[17] V. A. Marchenko and L. A. Pastur. Distributions of eigenvalues for some sets of random matrices. *Math USSR-Sbornik*, 1:457–483, 1967.

[18] J. Baik, G. B. Arous, and S. Peche. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab*, 33(5):1643–1697, 2005.

[19] J. Baik and J. W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivar. Anal.*, 97(6):1382–1408, 2006.

[20] L. S. Cardoso, M. Debbah, P. Bianchi, J. Najim, and W. Hachem. Cooperative Spectrum Sensing for Opportunistic Wireless Access Using Random Matrix Theory (under preparation). 2008.