

# ADAPTIVE EM-BASED ALGORITHM FOR COOPERATIVE SPECTRUM SENSING IN MOBILE ENVIRONMENTS

Jesus Perez, Ignacio Santamaria and Javier Via

Advanced Signal Processing Group, University of Cantabria, Spain, <https://gtas.unican.es>

## ABSTRACT

In this work we propose a new adaptive algorithm for cooperative spectrum sensing in dynamic environments where the channels are time varying. We assume a cooperative sensing procedure based on the soft fusion of the signal energy levels measured at the sensors. The detection problem is posed as a composite hypothesis testing problem. Then, we consider the Generalized Likelihood Ratio Test approach where the maximum likelihood estimate of the unknown parameters (which are the signal-to-noise ratio under the different hypotheses) are obtained from the most recent energy levels at the sensors by means of the Expectation-Maximization algorithm. We derive simple closed-form expressions for both, the E and the M steps. The algorithm can operate even when only a subset of sensors report their energy estimates, which makes it suited to be used with any sensor selection strategy (active sensing). Simulation results show the feasibility and efficiency of the method in realistic slow-fading environments.

**Index Terms**— Cooperative spectrum sensing, energy detection, likelihood ratio test, EM algorithm, maximum likelihood estimation, fading channels

## 1. INTRODUCTION

Spectrum sensing is a key operation in cognitive radio. Through spectrum sensing the cognitive radios (CRs) try to detect frequency bands that are not being used by the primary network. The performance of spectrum sensing is limited by shadowing and multi-path fading effects in the sensing channels between the primary users (PUs) and the CRs. By using cooperative spectrum sensing (CSS) the impact of those effects can be mitigated efficiently by the inherent multiuser/spatial diversity of the CR network [1], [2].

This work focuses on centralized CSS based on the soft fusion of the signal energy levels measured by the CRs [3], [4], [5], [6]. Each CR estimates the energy level at its location in the frequency band of interest and reports it to a fusion center (FC) through a control channel. Then, the FC makes a decision on the presence or absence of primary signals in the channel. In practice any of the CRs can act as a FC.

The likelihood ratio test (LRT) is the optimal test when the FC knows the signal-to-noise ratio (SNR) at the CRs. In [3] and [4] the authors show that, when there is a single PU, the LRT leads to a linear fusion rule where the test statistic is a linear combination of the energy levels with appropriate weights.

Unfortunately, knowing the SNR at multiple CRs is in general difficult and entails a large overhead in mobile scenarios therefore, we have to deal with unknown parameters and the

detection problem becomes a composite hypothesis test [7]. Generalized-LRT based algorithms have been proposed in [8] and [9] to solve this problem, whereas, [10] and [11] derive detection algorithms based on the Rao test and the locally most powerful test. All these works assume a single PU.

In this work, we propose a different approach to the CSS problem. First, we pose it as a composite hypothesis testing problem and consider the GLRT. Then, we derive a new adaptive algorithm that allows the FC to dynamically estimate the unknown test parameters from the most recent energy estimates reported by the neighbouring CRs. The algorithm computes the maximum likelihood (ML) estimate of the LRT parameters by means of the Expectation-Maximization (EM) algorithm [12], [13]. We derive closed-form expressions for both the  $E$  and the  $M$  steps that take into account the constraints and relationships among the problem variables.

A similar approach to the CSS problem has been proposed in [5] and [6], where the authors use different Machine Learning techniques. However, these techniques are batch and, therefore, difficult to apply in mobile scenarios. In addition, they do not exploit the constraints and the relationship among the unknown parameters of the LRT.

It is well known that the adequate selection of a subset of active CRs is a key issue in CSS because it permits to alleviate the overhead without degrading the detection performance (see [1] and the references therein). This means that, at each time, the FC only fuses the energy levels from a subset of active sensors. Unlike the methods proposed in the previous references, our method is able to deal with incomplete energy vectors (i.e. missing data), which makes it suited to be used with any CRs selection method.

## 2. SYSTEM MODEL

### 2.1. Primary network state

We consider a general model for the primary network activity where more than one PU can transmit simultaneously [6]. Let  $s_i \in \{0, 1\}$  indicate the state of  $PU_i$ , where  $s_i = 1$  when it is transmitting (active) and  $s_i = 0$  when it is inactive. The PU network state is given by the binary vector  $\mathbf{s} = [s_1 \ s_2 \ \dots \ s_{N_U}]^T$ , where  $N_U$  is the number of PUs. Therefore, the channel is idle when  $\mathbf{s} = \mathbf{0}$ , whereas it is occupied when  $\mathbf{s} \neq \mathbf{0}$ .

### 2.2. Energy detection

Let  $W$  denote the channel bandwidth. The CRs perform energy detection [14] for a time duration of  $\tau$ , and hence they take  $M = W\tau$  signal samples during  $\tau$ . The normalized received energy estimate at  $CR_j$  is

$$e_j = \frac{2}{\eta_j} \sum_{m=1}^M |x_j[m]|^2,$$

where  $x_j[m]$  denotes the  $m$ -th signal sample and  $\eta_j$  is the noise variance at the  $CR_j$ .

---

This work has been funded by SODERCAN and Programa Operativo FEDER under grant CAIMAN - 12.JU01.64661, and by the Ministerio de Economía, Industria y Competitividad (MINECO) of Spain, and AEI/FEDER funds of the E.U., under grants TEC2017-86921-C2-1-R (CAIMAN), TEC2013-47141-C4-R (RACHEL) and TEC2016-75067-C4-4-R (CARMEN).

### 2.3. Distribution of the energy estimates

The distribution of the energy estimates depends on the PU network state. Let  $e_j|s$  be the received energy estimate at CR<sub>*j*</sub> conditioned to the primary network state  $s$ . According to the central limit theorem, if  $M$  is large enough (e.g.  $M \geq 20$  in practice),  $e_j|s$  is approximately normal distributed [6] with mean and variance given by

$$\mu_{s,j} = 2M(1 + \sum_{i=1}^{N_U} s_i g_{i,j}), \quad \sigma_{s,j}^2 = 4(\mu_{s,j} - M), \quad (1)$$

where  $g_{i,j}$  denotes the SNR at CR<sub>*j*</sub> when PU<sub>*i*</sub> is the only active user. Note that the variance is completely determined by the mean. In addition, when  $s = \mathbf{0}$ ,  $\mu_{\mathbf{0}} = 2M$  and  $\sigma_{\mathbf{0}} = 4M$ . Therefore, the distribution of  $e_j|\mathbf{0}$  only depends on  $M$ , which is assumed to be known by the FC.

Every time that the CRs report their energy estimates to the FC, a new energy vector  $\mathbf{e} = [e_1 \ e_2 \ \dots \ e_{N_S}]^T$  is generated, where  $N_S$  is the number of CRs. From (1), and assuming that the energy estimates from different CRs are conditionally independent, the distribution of the energy vectors will be

$$f(\mathbf{e}|\boldsymbol{\mu}_s) = \prod_{j=1}^{N_S} \left[ \frac{1}{\sqrt{8\pi}(\mu_{s,j} - M)} \exp\left(-\frac{(e_j - \mu_{s,j})^2}{8(\mu_{s,j} - M)}\right) \right]^{\mathbf{v}_j}, \quad (2)$$

where  $\mathbf{v} = [v_1 \ v_2 \ \dots \ v_{N_S}]^T$  is a binary vector that indicates which CRs have reported their energy estimate to the FC.

### 3. GLRT DETECTOR

The CSS detection problem can be posed as a binary hypothesis test, where the null hypothesis ( $H_0$ ) is  $s = \mathbf{0}$  and the alternative hypothesis ( $H_1$ ) is  $s \neq \mathbf{0}$ . Then, for a given observed energy vector  $\mathbf{e}$ , the LRT is

$$\frac{f(\mathbf{e}|H_1)}{f(\mathbf{e}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \quad (3)$$

where  $\lambda$  is a suitable detection threshold [7], and the likelihoods of  $\mathbf{e}$  under both hypotheses are given by

$$\begin{aligned} f(\mathbf{e}|H_0) &= f(\mathbf{e}|\boldsymbol{\mu}_{\mathbf{0}}), \\ f(\mathbf{e}|H_1) &= \sum_{s \neq \mathbf{0}} \frac{\pi_s}{1 - \pi_{\mathbf{0}}} f(\mathbf{e}|\boldsymbol{\mu}_s), \end{aligned} \quad (4)$$

where  $\boldsymbol{\mu}_s = [\mu_{s,1} \ \mu_{s,2} \ \dots \ \mu_{s,N_S}]^T$ . As it was mentioned, the distribution  $f(\mathbf{e}|\boldsymbol{\mu}_{\mathbf{0}})$  only depends on  $M$ . On the other hand, the distribution of  $\mathbf{e}|H_1$  is a mixture of  $2^{N_U} - 1$  Gaussian densities given by (2). The terms  $\{\pi_s\}$  are the prior probabilities of the PU network states. They are assumed to be known. If they were unknown, a reasonable choice would be identical probability for all states of  $H_1$ .

The application of the LRT requires the FC to estimate the unknown parameters  $\boldsymbol{\mu} = \{\boldsymbol{\mu}_s\}_{s \neq \mathbf{0}}$ . According to the GLRT approach they are replaced by their ML estimates. In this point, the FC takes advantage of the fact that the energy vectors arrive sequentially so, in order to estimate the unknown parameters, the FC considers the most recent energy vectors reported by the CRs. Note that those energy estimates may have been acquired under different PU network states.

### 4. ADAPTIVE ML ESTIMATION

In mobile environments the sensing channels are time varying and therefore the means  $\boldsymbol{\mu}$  (and consequently also the distributions of the energy vectors  $f(\mathbf{e}|\boldsymbol{\mu}_s)$ ) change over time. The FC

applies a sliding window containing the  $N$  most recent energy vectors. Each time a new energy vector arrives to the FC, it is stored, and the oldest one is removed from the window. Then, using the new window of energy vectors, the FC updates the estimates of the means. This way, the FC may compute the LRT (3), at any time, with the current estimates of  $\boldsymbol{\mu}$ . This procedure is depicted in Figure 1 where  $\mathbf{e}(n)$  denotes the  $n$ -th energy vector in the sliding window.

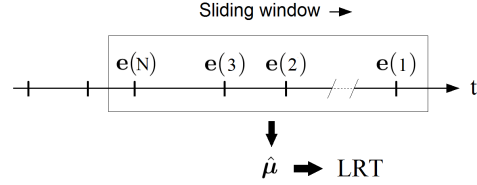


Fig. 1. Sliding window over the energy vectors.

The window size  $N$  is a key parameter. The larger  $N$ , the more accurate the estimates. On the other hand, the channels (and therefore the means  $\boldsymbol{\mu}$ ) must remain practically constant during the total acquisition time of the energy vectors. Consecutive energy vectors in the window can be incomplete. This occurs when only a subset of CRs report their energy estimates to the FC. This makes the estimation algorithm to be able to deal with incomplete (missing) data.

#### 4.1. ML estimation

The EM algorithm is well suited to estimate the parameters of a mixture of Gaussian densities [13]. However, our problem has some peculiarities that precludes the application of the standard version of the EM algorithm:

1. The mean of the component  $s = \mathbf{0}$ , is known in advance:  $\boldsymbol{\mu}_{\mathbf{0}} = 2M\mathbf{1}$ .
2. The means are constrained as follows  $\mu_{s,j} \geq 2M$ .
3. The covariance matrices of the Gaussian components are diagonal and the variances of the Gaussian components are fully determined by their means (1).
4. The mixing coefficients  $\{\pi_s\}$  are known.
5. The FC fuses the energy estimates from a subset of the CRs, in other words the data vectors can be incomplete.

These peculiarities require an EM algorithm specifically designed for this problem, which is derived in the following.

Let  $\mathbf{E} = \{\mathbf{e}(n)\}_{n=1}^N$  the set of energy vectors (possibly incomplete) in the sliding window at a given time. Then, the ML estimate of the means  $\boldsymbol{\mu}$ ,

$$\hat{\boldsymbol{\mu}}_{ML}(\mathbf{E}) = \underset{\boldsymbol{\mu} \in \Psi}{\operatorname{argmax}} \log L(\boldsymbol{\mu}), \quad (5)$$

where  $\Psi$  denotes the space of the means with the constraints  $\mu_{s,j} \geq 2M$ , and the log-likelihood function is given by

$$\log L(\boldsymbol{\mu}) = \sum_{n=1}^N \log \sum_s \pi_s f(\mathbf{e}(n)|\boldsymbol{\mu}_s) \quad (6)$$

#### 4.2. EM algorithm

Let  $\mathbf{Z} = \{z_s(n)\}$  a set of hidden variables that indicate the PU network state when the energy vectors were acquired, that is,  $z_s(n) = 1$  if  $\mathbf{e}(n)$  was acquired when the network was in state

s, and  $z_s(n) = 0$  otherwise. If the hidden variables were known the (complete) log-likelihood function would be

$$\log L_C(\boldsymbol{\mu}) = \sum_{n=1}^N \sum_{\mathbf{s}} z_{\mathbf{s}}(n) \log(\pi_{\mathbf{s}} f(\mathbf{e}(n)|\boldsymbol{\mu}_{\mathbf{s}})). \quad (7)$$

Each iteration of the EM algorithm involves two steps:

- E-step:  $Q(\boldsymbol{\mu}, \boldsymbol{\mu}^{(k)}) = E_{\mathbf{Z}} [\log L_C(\boldsymbol{\mu}) | \mathbf{E}, \boldsymbol{\mu}^{(k)}]$ ,
- M-step:  $\boldsymbol{\mu}^{(k+1)} = \underset{\boldsymbol{\mu} \in \Psi}{\operatorname{argmax}} Q(\boldsymbol{\mu}, \boldsymbol{\mu}^{(k)})$ ,

where  $k$  denotes the iterations and  $\boldsymbol{\mu}^{(k)}$  denotes the estimate after the  $k$ -th iteration.

#### 4.2.1. E-step:

Considering (7),

$$Q(\boldsymbol{\mu}, \boldsymbol{\mu}^{(k)}) = \sum_{n=1}^N \sum_{\mathbf{s}} \gamma_{\mathbf{s}}^{(k)}(n) \log(\pi_{\mathbf{s}} f(\mathbf{e}(n)|\boldsymbol{\mu}_{\mathbf{s}})), \quad (8)$$

where

$$\gamma_{\mathbf{s}}^{(k)}(n) = \frac{\pi_{\mathbf{s}} f(\mathbf{e}(n)|\boldsymbol{\mu}_{\mathbf{s}}^{(k)})}{\sum_{\mathbf{r}} \pi_{\mathbf{r}} f(\mathbf{e}(n)|\boldsymbol{\mu}_{\mathbf{r}}^{(k)})}. \quad (9)$$

#### 4.2.2. M-step:

Considering (8) and (2), the M-step can be broken down into a set of uncoupled optimization problems of the form

$$\begin{aligned} \mu_{\mathbf{s},j}^{(k+1)} &= \underset{\mu_{\mathbf{s},j}}{\operatorname{argmin}} \sum_{n=1}^N \gamma_{\mathbf{s}}^{(k)}(n) v_j(n) \\ &\quad \left( \log(\mu_{\mathbf{s},j} - M) + \frac{(e_j(n) - \mu_{\mathbf{s},j})^2}{4(\mu_{\mathbf{s},j} - M)} \right), \\ \text{s.t.} \quad &\mu_{\mathbf{s},j} \geq 2M \end{aligned}$$

Setting the derivative of the objective function with respect to  $\mu_{\mathbf{s},j}$  to zero, and considering the constraint, leads to

$$\begin{aligned} \mu_{\mathbf{s},j}^{(k+1)} &= M - 2 \\ &+ \sqrt{4 + \frac{\sum_{n=1}^N \gamma_{\mathbf{s}}^{(k)}(n) v_j(n) (e_j(n) - M)^2}{\sum_{n=1}^N \gamma_{\mathbf{s}}^{(k)}(n) v_j(n)}}. \end{aligned} \quad (10)$$

---

#### Algorithm 1 : ML estimate of $\boldsymbol{\mu}$

---

- 1: **input:**  $\mathbf{E}, \boldsymbol{\mu}^{(0)}, \{\pi_{\mathbf{s}}\}, M$
  - 2: Initialize iterations:  $k = 0$
  - 3: **repeat**
  - 4: E-step: Compute the terms,  $\gamma_{\mathbf{s}}^{(k)}(n)$ , with (9)
  - 5: M-step: Update the mean estimates,  $\mu_{\mathbf{s},j}^{(k+1)}$ , with (10)
  - 6: Compute the log-likelihood function of  $\boldsymbol{\mu}^{(k+1)}$  with (6)
  - 7:  $k = k + 1$
  - 8: **until** the log-likelihood function converges
  - 9: **return:**  $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}^{(k)}$
- 

The resulting algorithm is summarized in **Algorithm 1**. Iterations stop when there is no further change in the log-likelihood function. Since each iteration increases its value, the convergence of the algorithm is ensured [13]. However, the algorithm may converge to a local maximum.

The algorithm is run to update the estimate of  $\boldsymbol{\mu}$  whenever a new energy vector enters the sliding window. Each run is initialized with the last estimate of  $\boldsymbol{\mu}$  obtained with the previous window of energy vectors. Therefore, the initial values of  $\boldsymbol{\mu}$  will be close to the new means. As a result, the algorithm requires very few iterations to converge.

## 5. SIMULATION RESULTS

In this section we obtain the ROC (Receiver Operating Characteristic) curve to illustrate the performance of the proposed algorithm in different scenarios. Each ROC curve averages the LRT results over  $4 \cdot 10^5$  energy vectors. Before applying the LRT, we compute the ML estimate of the parameters  $\boldsymbol{\mu}$  by means of **Algorithm 1**, considering a sliding window with the  $N$  most recent energy vectors.

Unless otherwise indicated, we have considered the following assumptions and parameter values in the simulations,

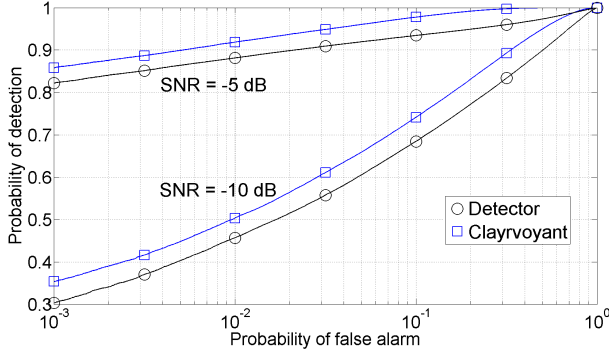
- The channel bandwidth is  $W = 5$  MHz.
- The noise spectral density is  $\eta = -174$  dBm at all CRs.
- The channels between different pairs PU-CR are independent and identically Rayleigh distributed. For each sensing channel we generate time-correlated realizations using the Jakes' model [15], for a given maximum Doppler shift  $f_D = 25$  Hz. Therefore, the coherence time of the sensing channels  $T_C \approx \frac{0.4}{f_D} = 16$  ms [16].
- The number of signal samples for energy estimation (2.2) is  $M = 100$ , therefore, the sensing time is  $\tau = \frac{M}{W} = 20 \mu\text{s}$
- The PU network state can change between two consecutive energy vectors. The activity of the PUs is modeled as independent and identically distributed homogeneous Markov chains with two states: inactive ( $s_i(n) = 0$ ) and active ( $s_i(n) = 1$ ). In the simulations we assume that the transition probabilities are  $p_{0,0} = P(s_i(n) = 0 | s_i(n-1) = 0) = 0.75$  and  $p_{1,1} = P(s_i(n) = 1 | s_i(n-1) = 1) = 0.5$ . Accordingly, for a PU network with  $N_U = 3$  users, the probability that the channel is idle is  $\pi_0 = 0.3$ .
- All PUs transmit identical power. Then, the average SNR is the same for all CRs (averaging over the sensing channel realizations and the PU's activity). The assumption is mainly made to facilitate the interpretability of the results. Unless otherwise indicated, we assume that the average SNR is  $g_{i,j} = -5$  dB,  $\forall i, j$ .
- The CRs (or a subset of them) sense the channel periodically with constant sensing period  $T_F = 1$  ms. This value requires the window size to be  $N \leq \frac{T_C}{T_F} = 16$ .
- The control channels between the CRs and the FC are error-free.

An ideal detector that assumes perfect knowledge of the unknown parameters is referred to as a clairvoyant detector [7]. Its ROC will be shown in some figures as an upper bound. The performance gap between the clairvoyant detector and our detector is the performance loss due to the parameter estimation errors.

The total number of energy estimates in the sliding window is  $N_o = \sum_{n=1}^N \sum_{j=1}^{N_S} v_j(n)$ , which can take values in the interval  $N \leq N_o \leq N_S N$ . In the simulations, when  $N_o < N_S N$ , the CRs that report their energy estimates are selected randomly. Obviously, if an optimal CRs selection strategy were used, the detection performance would be better.

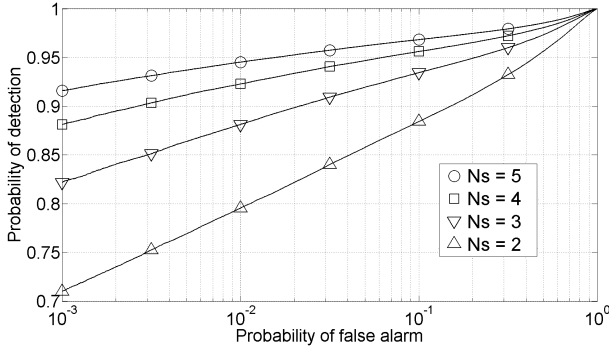
Figure 2 shows the ROC curves for different SNR values when  $N_U = 3$  and  $N_S = 3$ . In this scenario all energy estimates are available at the FC, and hence  $N_o = N_S N$ . The ROC curves of the clairvoyant detector are also shown. It can be observed that the performance loss due to the estimation errors is relatively low.

Figure 3 shows the ROC curves for a network with  $N_U = 3$  and different number of CRs. Again, we assume that all energy estimates are available at the FC ( $N_o = N_S N$ ). Obviously, the



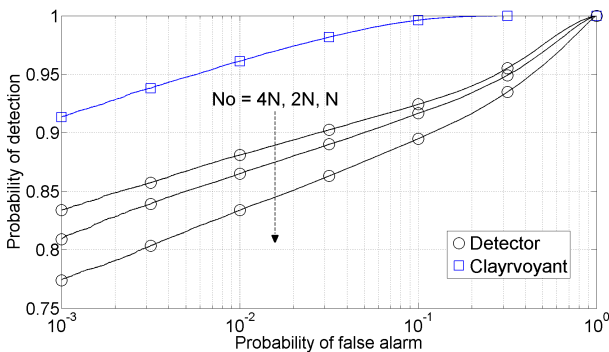
**Fig. 2.** ROC curves for  $N_U = 3$ ,  $N_S = 3$  and different values of average SNR.

higher the number of sensors the better the performance. As it is also expected, the gain by increasing  $N_S$  is higher when the number of CRs is low.



**Fig. 3.** ROC curves for  $N_U = 3$  and different number of CRs.

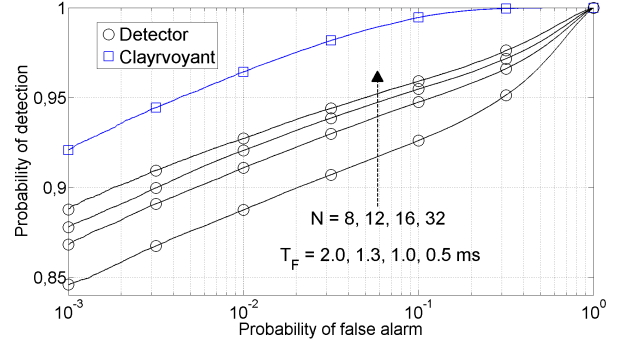
Figure 4 compares the ROC curves for different number of energy estimates in the sliding windows ( $N_o$ ). In this example the number of PUs and CRs are  $N_U = 3$  and  $N_S = 4$ , respectively. In the case  $N_o = 4N$  all CRs report their energy estimates to the FC, and hence the energy vectors do not have missing entries, whereas when  $N_o = N$  only a single CR reports its energy level each time. As it is expected, the higher the  $N_o$ , the better the performance.



**Fig. 4.** ROC curves for  $N_U = 2$ ,  $N_S = 4$  and different number of energy estimates at the FC.

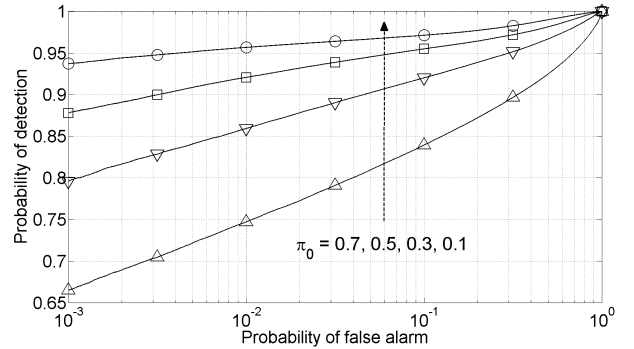
Figure 5 shows the ROC curves for different values of the sensing period  $T_F$ . The window size  $N$  is chosen to guarantee that the energy vectors in the sliding windows have been obtained for a constant channel:  $N \leq T_C/T_F$ . The lower the  $N$  the less accurate are the estimates of the LRT parameters. The number of PUs and CRs are  $N_U = 3$  and  $N_S = 4$ , respec-

tively, but only the energy estimates from two CRs are available for clustering each time ( $N_o = 2N$ ).



**Fig. 5.** ROC curves for  $N_U = 3$ ,  $N_S = 4$ ,  $N_o = 2N$  and different sliding window sizes.

Finally, Figure 6 shows the ROC curves for PU networks with different activity characteristics. The activity of the network is characterized by parameter  $\pi_0$ , that is, is the probability that none of the PUs is transmitting. The lower the  $\pi_0$ , the better the performance. This is because the EM algorithm has to estimate the parameters of the distributions associated with all network states with active PUs,  $s \neq 0$ , whereas the parameters of the distribution for  $s = 0$  are known in advance. Therefore, the more energy vectors are generated under active network states ( $s \neq 0$ ), the better the algorithm estimates the parameters of their distributions. The different values of  $\pi_0$  were obtained by choosing the appropriate values of the transition probability  $p_{0,0}$  while keeping  $p_{1,1} = 0.5$  fixed.



**Fig. 6.** ROC curves for  $N_U = 3$ ,  $N_S = 4$ ,  $N_o = 2N$  and different probability that the channel is idle.

## 6. CONCLUSIONS

In this work we have presented a novel method for cooperative spectrum sensing in mobile environments where the sensors or cognitive radios perform energy detection. We have posed the detection problem as a Generalized Likelihood Ratio Test problem where the maximum likelihood estimates of the unknown parameters are computed by means of an Expectation-Maximization algorithm. The input to the estimation algorithm are the most recent local energy levels at the sensors. Due to its adaptive nature, the algorithm is suited for time varying scenarios. In addition, the algorithm is able to deal with missing values, which makes it suited to be applied with any sensor selection strategy to reduce the overhead in the control network. The simulation results show that the proposed method is feasible and efficient in slow fading environments.

## 7. REFERENCES

- [1] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: a survey," *Physical Communications*, vol. 4, no. 1, pp. 40–62, March 2011.
- [2] Y. Zeng, Y. C. Liang, A. T. Hoang Liu, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, January 2010.
- [3] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4502–4507, December 2008.
- [4] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 28–40, February 2008.
- [5] K. W. Choi, E. Hossain, and D. I. Kim, "Cooperative spectrum sensing under a random geometric primary user network model," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1932–1944, June 2011.
- [6] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 10, pp. 2209–2221, November 2013.
- [7] S. M. Kay, *Fundamentals of Statistical Signal Processing. Detection Theory*, Prentice Hall. Signal Processing Series, 1998.
- [8] T. J. Lim, R. Zhang, Y. C. Liang, and Y. H. Zeng, "GLRT-based spectrum sensing for cognitive radio," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, December 2008, pp. 4391–4395.
- [9] J. Font Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2137–2146, July 2010.
- [10] S. Zarrin and T. J. Lim, "Composite hypothesis testing for cooperative spectrum sensing in cognitive radio," in *Proceedings of the IEEE International Conference on Communications (ICC)*, June 2009, pp. 1–5.
- [11] S. Zarrin and T. J. Lim, "Cooperative spectrum sensing in cognitive radios with incomplete likelihood functions," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3272–3281, June 2010.
- [12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood for incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [13] G. J. McLachlan and T. Krishnan, *The EM algorithm and extensions*, John Wiley and Sons, 2008.
- [14] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, April 1967.
- [15] W. C. Jakes, *Microwave Mobile Communications*, John Wiley and Sons, 1975.
- [16] T. S. Rappaport, *Wireless Communications*, Prentice Hall, 2002.