Joint Learning-Detection Framework: an Empirical Analysis

Wassim Jouini and Amor Nafkha
SCEE/IETR, SUPELEC,
Avenue de la Boulaie, CS 47601,
35576 Cesson Sévigné Cedex, France.
wassim.jouini@supelec.fr

Miguel López-Benítez and Jordi Pérez-Romero
Dept. Signal Theory and Communications,
Universitat Politècnica de Catalunya,
08034 Barcelona, Spain.
miguel.lopez@tsc.upc.edu

***Abstract*—Within a context of Cognitive Radio, we explore the possibility to detect the presence of users on a given band with no a priori knowledge on the environment. This paper ventures the analysis of a joint learning-detection framework where the detector improves its detecting abilities online relying on past gathered information on the probed frequency band. Simulations conducted on real data to support the suggested approach show interesting results. These results are then discussed and further possible improvements are proposed.**

I. Introduction

During the last century, most of the meaningful frequency bands were licensed to emerging wireless applications, where the static frequency allocation policy combined with a growing number of spectrum demanding services led to a spectrum scarcity. However, several measurements conducted in the United-States [1], first, and then in numerous other countries and/or contexts [2][3][4][5], showed a chronic underutilization of the frequency band resources revealing substantial communication opportunities in these licensed bands.

To alleviate the spectrum scarcity, the (Federal Communications Commission) FCC [1] suggested opening the licensed bands dedicated to primary users (PUs) to unlicensed users and services (usually referred to as secondary users, SUs), allowing them to exploit these bands if unoccupied at a particular time in a particular geographical area. Moreover, Cognitive Radio has been proposed as a promising technology to enable a harmless coexistence of PUs and SUs on a same frequency band.

In a nutshell, Cognitive Radio (CR) equipment is a communication device aware of its environment as well as of its operational abilities and capable of using them intelligently to fulfill its tasks. SUs based CR equipment are, thus characterized by their ability to gather information through their sensors and to use them to adapt their behavior to their nearby primary network [6]. In a generic scenario, a SU should first probe its environment to detect spectrum holes, and then access the available bands if needed.

In order to exploit as many opportunities as possible while ensuring minimum interference with PUs, spectrum sensing appears as a crucial step where SUs are required to perform quick and accurate detection of PUs while having minimum (or no) *a priori* information on PUs activity or the channels characteristics. To that purpose, extensive work has been provided by the Cognitive Radio community to design blind and efficient spectrum sensing tools [7][8].

In this paper, we suggest a joint learning-detection framework to alleviate the lack of knowledge on the noise level. Section II introduces the usually considered theoretical model as well as the commonly used energy detector in case of a perfect knowledge on the noise level. Section III and Section IV suggest using a learning algorithms to alleviate the absence of information on the noise level. Information on the used real data for the experiment are briefly described in Section V. Then, first experimental results on the performances of the introduced ratio detector are presented and discussed in Section VI. Finally, Section VII concludes the paper.

II. System model

A. Network assumption

We consider a receiver willing to gather information on a pool of bands allocated to a primary network. For the sake of simplicity, we represent time as a discrete sequence of slots. In every slot, the CR can acquire a given number of samples depending on the characteristics of the receiver. Every time the CR collects samples on a specific band, only two observation outcomes are possible: the channel can be sensed either idle or busy. In the rest of the paper, we will associate the numerical value 1 to a busy channel and 0 to an idle channel. We assume that the SU has no *a priori* knowledge on the characteristics of the bands (e.g., occupied or not).

Let $\mathbf{r}_t = [r_{t,0}, r_{t,1}, \cdots, r_{t,M-1}]$ be M independent and identically distributed (i.i.d.) samples gathered by the CR receiver at the current slot t. The outcome of the sensing process can be modeled as a binary hypothesis test described as follows:

$$\mathbf{r}_t = \begin{cases} \mathbf{n}_t, & \mathbb{H}_0 \\ \mathbf{x}_t + \mathbf{n}_t, & \mathbb{H}_1 \end{cases}$$

where hypotheses $\mathbb{H}_0$ and $\mathbb{H}_1$ refer respectively to the case of an absent or a present signal on the sensed channel. On the one hand, $\mathbf{x}_t = [x_{t,0}, x_{t,1}, \cdots, x_{t,M-1}]$ refers to the source signal where every sample $x_{t,k}$ is perceived as an i.i.d. realization of a Gaussian stochastic distribution $\mathcal{N}(0, \sigma^2_{x,t})$. On the other hand, $\mathbf{n}_t = [n_{t,0}, n_{t,1}, \cdots, n_{t,M-1}]$ refers to i.i.d. additive white

Gaussian noise (AWGN) samples $\mathcal{N}(0, \sigma_{n,t}^2)$. Moreover, $\mathbf{x}_t$ and $\mathbf{n}_t$ are assumed to be independent. Thus, we consider the following Gaussian received signals under either hypothesis $\forall r_{t,i}\ i \in \{0, \cdots, M-1\}$:

$$\begin{cases} \mathbb{H}_0 : r_{t,i} \sim \mathcal{N}(0, \sigma_{n,t}^2) \\ \mathbb{H}_1 : r_{t,i} \sim \mathcal{N}(0, \sigma_{x,t}^2 + \sigma_{n,t}^2) \end{cases}$$

The previously introduced network considerations will be referred to as *Assumption* 1. Within this context, the detection outcome can be modeled as the output of a decision making policy π that maps the current samples $\mathbf{r}_t$ into a binary value $d_t = \pi(\mathbf{r}_t)$, $d_t \in \{0, 1\}$.

In the next subsection, we summarize the usually used criteria to evaluate the performance of a signal detection policy.

B. Performance evaluation of a detection policy π

Under the previously considered binary hypothesis test, one can define two probabilities that characterize the performance of the detection policy π at the slot number t: The probability of false alarm ($\mathbb{P}_{fa,t}$) and the probability of detection ($\mathbb{P}_{d,t}$):

$$\begin{cases} \mathbb{P}_{fa,t} = \mathbb{P}\left(d_t = 1 | \mathbb{H}_0\right) \\ \mathbb{P}_{d,t} \ = \mathbb{P}\left(d_t = 1 | \mathbb{H}_1\right) \end{cases}$$

Usually, constraints impose to fix the $\mathbb{P}_{fa,t}$ under a given level α, such that $\mathbb{P}_{fa,t} \leq \alpha$. The most powerful decision policy is then defined as the one having the largest $\mathbb{P}_{d,t}$ value for a given $\mathbb{P}_{fa,t} = \alpha$.

C. Neyman-Pearson energy detector

The Neyman-Pearson energy detector (NP-ED, also known as *radiometer*) is a commonly used spectrum sensor. It has been extensively analyzed for its proprieties as a semi-blind low complexity spectrum sensor, since it ignores the characteristics of the received signals and only relies on the perceived energy of the signal. It assumes known the noise level $\sigma_{n,t}^2$ at every slot number t. As a first approximation, we consider in the rest of the paper a constant noise level for all t, $\sigma_{n,t}^2 = \sigma_n^2$. Under these assumption, the NP-ED is proven to be the most powerful test. Hence, NP-ED appears as an efficient approach when no information other than the noise level is available at the receiver side.

NP-ED relies on the computation of the received energy statistic $\mathcal{T}_t$ at the slot number t defined such as:

$$\mathcal{T}_t = \sum_{i=0}^{M-1} |r_{t,i}|^2$$

The decision policy π_{NP-ED} is a simple Heaviside function that depends only on the evaluation of the statistic $\mathcal{T}_t$ on the current slot t:

$$d_t = \pi_{NP-ED}(r_t) \Longleftrightarrow d_t = \mathcal{H}(\mathcal{T}_t - \xi_t(\alpha))$$

where $\xi_t(\alpha)$ is the selected threshold to guaranty $\mathbb{P}_{fa} \leq \alpha$. Such policies are usually described using the following notation:

$$\mathcal{T}_t \lessgtr_{\mathcal{H}_1}^{\mathcal{H}_0} \xi(\alpha)$$

The following two equations briefly remind the expressions of the $\mathbb{P}_{fa,t}$ and the $\mathbb{P}_{d,t}$ as well as their approximations for large M.

$$\begin{cases} \mathbb{P}_{fa,t} = 1 - F_{\chi_M^2}\left(\frac{\xi_t(\alpha)}{\sigma_n^2}\right) \approx Q\left(\sqrt{\frac{M}{2}}\left(\frac{\xi_t(\alpha)/M}{\sigma_n^2} - 1\right)\right) \\ \mathbb{P}_{d,t} = 1 - F_{\chi_M^2}\left(\frac{\xi_t(\alpha)}{\sigma_n^2 + \sigma_{x,t}^2}\right) \approx Q\left(\sqrt{\frac{M}{2}}\left(\frac{\xi_t(\alpha)/M}{\sigma_n^2 + \sigma_{x,t}^2} - 1\right)\right) \end{cases}$$

where $F_{\chi_M^2}(\cdot)$ refers to the cumulative distribution function of a χ^2-distribution with M degrees of freedom, and $Q(\cdot)$ is the complementary cumulative distribution function of Gaussian random variable (also known as Marcum function), formally defined as:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(\frac{-t^2}{2}\right) dt$$

NP-ED provides satisfactory behavior when σ_n is known (or accurately estimated). Unfortunately, when such knowledge is unavailable, its performances degrade very quickly.

III. Learning based ratio statistic

In this section, we aim at introducing general notations to define a statistic as a ratio of transformations of both the currently gathered samples and the past collected information. Mainly the ratio expression is introduced to alleviate the lack of knowledge on the noise level σ_n^2. We argue that the key enabling concept to guaranty given performances within an unknown environment is learning over past information. Thus, it is a joint learning-detection framework.

Let $\mathbf{h}_t = [h_{t,0}, h_{t,1}, \cdots, h_{t,N-1}]$ denote a set of N samples collected in the past slots and memorized. We refer to $\mathbf{h}_t$ as the information or history vector. It is mainly introduced to enable the exploitation of past information on the environment to enhance the behavior of the Cognitive Radio decision making engine. Consequently, the detection outcome can be seen as the output of a decision making policy π that maps the current and past information, represented respectively, by the vectors $\mathbf{r}_t$ and $\mathbf{h}_t$ into a binary value $d_t = \pi(\mathbf{r}_t, \mathbf{h}_t),\ d_t \in \{0, 1\}$.

Let us introduce the following statistic $\mathcal{F}_t$ and two real functions $f(\cdot)$ and $g(\cdot)$ such that:

$$\mathcal{F}_t = \frac{f(\mathbf{r}_t)}{g(\mathbf{h}_t)}$$

Mainly, the function $f(\cdot)$ represents a transformation of the currently received signals at the slot number t, while, the function $g(\cdot)$ aims at extracting information on the noise of the sensed channel using past information. Since $g(\mathbf{h}_t)$ provides an estimation of the noise level, it is well known that the estimation uncertainty implies a less effective detection. Namely, there exist an *SNR wall* under which it is theoretically impossible to detect primary users' activity [9].

In the rest of this paper, we consider the following well known form of the functions $f(\cdot)$:

$$f(\mathbf{r}_t) = \frac{\mathcal{T}_t}{M}$$

IV. Learning model and tools

A. History vector modeling

Let us consider the particular case of a history vector $\mathbf{h}_t$ containing all collected samples. Moreover let us assume that $\mathbb{P}(\mathbb{H}_0) \neq \{0,1\}$. In other words, the probed channel does not remain in the same state (busy or idle) indefinitely. In this case, a fair assumption would be to consider the vector $\mathbf{h}_t$ as a mixture of Gaussian distributions. Thus the history vector of observations is assumed to be drawn form a stochastic distribution θ formalized as a Gaussian Mixture Model (GMM) with K components (to specify later on real data). Let $\theta_{k\in\{0,\cdots,K-1\}}$ denote the distribution of the k^{th} component of the distribution θ, and $p_{\theta_k}(\cdot)$ its associated probability density function (*pdf*):

$$p_{\theta_k}(h) = \frac{1}{\sqrt{2\pi\mathbb{V}_k}} e^{-\frac{(h-\mu_k)^2}{2\mathbb{V}_k}}$$

where $\mu_k \triangleq \mathbb{E}[\theta_k]$ and $\mathbb{V}_k \triangleq \mathbb{V}[\theta_k]$. $\mathbb{E}$ and $\mathbb{V}$ refer, respectively, to the expectation and the variance of the considered distribution.

Moreover, let us denote by $a_{k\in\{0,\cdots,K-1\}}$ the proportion of samples drawn from the k^{th} distribution and $\mathcal{A}_k$ the set of samples drawn from the distribution θ_k. Then we can define a_k such that: $a_k \triangleq \frac{\#\{\mathcal{A}_k\}}{N}$. Finally, the *pdf*, $p_\theta(\cdot)$, of the observed random data h can be written as:

$$p_\theta(h) = \sum_{k=0}^{K-1} a_k p_{\theta_k}(h)$$

Notice that depending on the spectrum sensing scenario, only a subset of these parameters might be of interest. As a matter of fact, on the one hand, in the case of extensive offline spectrum sensing, the decision maker, aims at evaluating the occupation pattern of the band. In this case, the parameters a_k appear crucial. On the other hand, in the case of online learning, evaluating the parameters related to the noise level and/or signal level would be prioritized.

The next subsection will briefly present different possible machine learning approaches to evaluate the parameters of Gaussian mixtures. Anticipating the empirical experiments on real observed data, we will only discuss algorithms that tackles two component ($K = 2$) Gaussian mixtures.

B. Learning algorithms

When dealing with a mixture of two Gaussian distributions, several alternative approaches can be suggested by the machine leaning community. Among which the most famous and usually implemented: the *Expectation-Maximization* algorithm, the *K-means* algorithm and the *Moment Method*.

1) Expectation-Maximization (EM): The EM algorithm is an iterative algorithm that generalizes the maximum likelihood concept to GMMs[10][11]. It is known to be very efficient compared to other algorithms for a same training set. Unfortunately it can become computationally burdensome when the training set or the number of mixtures grow large. It mainly computes the *a posteriori* distribution of the evaluated parameters. This algorithm is usually presented as a soft clustering tool where all samples are used, with specific weights depending on the cluster, to evaluate the parameters of all clusters.

2) K-means: The K-means algorithm is usually presented as a particular case of EM algorithms where every element of the training set is allocated to only one cluster (i.e., one Gaussian component).

3) Moment-Method (MM): The MM algorithm only relies on the empirical moments of the training set. It is a very low complexity algorithm compared to the EM algorithm. It is, unfortunately, also less efficient for a same number of training samples[12].

Although, in general, learning algorithms have nice mathematical proprieties in theory (e.g., in terms of convergence), many practical difficulties can occur when implementing them. Briefly we can stress out, first, the usually high complexity of these algorithms to reach their solution (especially true for the EM algorithm). Moreover, if the stopping criteria are not carefully selected, they could lead to a non convergence of the algorithms. As a consequence, “wrong” estimations of the parameters’ distributions could occur. Last but not least, if the algorithms rely on specific hypothesizes on the data distributions (e.g., Gaussian mixtures in the case of EM and MM algorithms), not fulfilling these assumptions might significantly affect their behavior.

In the rest of this paper, we chose to implement the EM algorithm and to present first results of its performances on real data collected during a spectrum measurement campaign. The goal being to characterize the distribution of the ratio statistic $\mathcal{F}_t$. As a matter of fact, this would enable the design of a learning based detector where the probabilities of false alarm and miss detection could be evaluated and controlled.

V. Overview of the measurement campaign setup

In this Section, we briefly describe the experimental setup, equipment used during the spectrum measurement campaign as well as the probed channels characteristics.

The measurement platform employed in this study is based on the Universal Software Radio Peripheral (USRP) [13] and GNU Radio [14] architecture. USRP is an openly designed inexpensive Software Designed Radio (SDR) hardware platform that provides radio front-end functionalities, Analogical to Digital and Digital to Analogical Conversion (ADC/DAC), decimation/interpolation with filtering and a Universal Serial Bus 2 (USB2) interface to connect to an off-the-shell Personal Computer (PC). The PC runs the GNU Radio software, a free and open source toolkit that provides a library of signal processing blocks for building SDRs. In addition, it also provides blocks for communicating with the USRP. The general scheme of the measurement platform is illustrated in Figure 1.

The primary signal of interest is captured with an omnidirectional discone-type antenna AOR DN753 that covers the frequency range 75–3000 MHz. The USRP Radio Frequency (RF) front-ends are provided in form of daughter boards that can be plugged to the USRP main board. In this study we

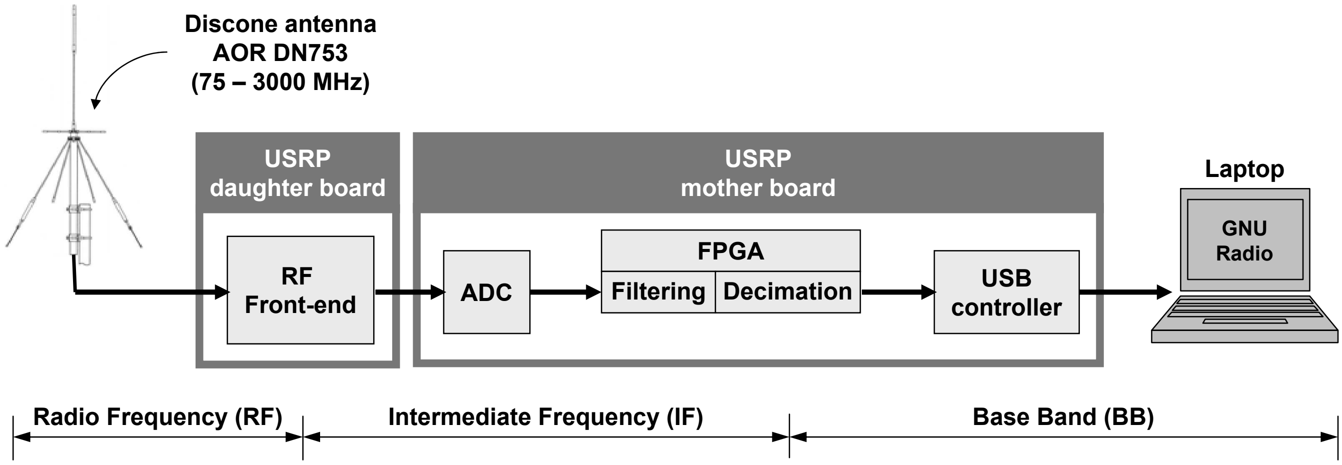

Fig. 1. Measurement platform employed in this study.

have employed two receiver-only daughter boards: TVRX (50–860 MHz, 8 dB typical noise figure) and DBSRX (800-2400 MHz, 3-5 dB typical noise figure). The daughter boards are employed to tune to the carrier frequency of the desired primary signal and perform down-conversion to the Intermediate Frequency (IF) at which the USRP main board operates. The USRP main board includes 12-bit ADCs working at $64 \cdot 10^6$ samples per second to digitize the received signal and a Field Programmable Gate Array (FPGA) to perform filtering and digital down-conversion (decimation) from the IF band to the Base Band (BB). Decimation is required in order to adapt the incoming data rate to the USB2 and PC computing capabilities. A USB controller sends the digital signal samples to the PC in 16-bit I and 16-bit Q complex data format (4 bytes per complex sample), resulting in a maximum rate of $8 \cdot 10^6$ complex samples per second. The maximum RF bandwidth that can be handled is therefore 8 MHz (narrower bandwidths can be selected by adjusting the decimation rate). The host PC runs the GNU Radio's `usrp_rx_cfile.py` script, which simply collects the digital signal samples sent by the USRP board through the USB2 interface and saves the received BB digital signal sequence to a file in the host PC's hard drive for off-line post-processing.

The aforementioned measurement equipment was used to collect digital signal samples of real channels for various radio technologies. To this end, the measurement platform was placed on a building rooftop in urban Barcelona, Spain (41º 23' 20" N, 2º 6' 43" E, 175 meters of altitude) with direct line-of-sight to several transmitters located a few tens or hundreds of meters away from the antenna and without buildings blocking the radio propagation. This measurement scenario enabled us to reliably capture the desired signals under high SNR conditions. Table I summarizes the main characteristics of the measured channels.

For each channel, the channel's radio frequency range, the radio bandwidth, the selected USRP decimation rate and the resulting sampled BandWidth (BW) as well as the USRP RF front-end gain factor are shown. The gain factor was chosen so as to maximize the received signal level (and hence the receiving SNR) without incurring in saturation. For most of the channels the optimum gain value was 70 dB and in the particular case of TV (both analogical and digital) the gain was drastically reduced due to the proximity of the TV station ($\approx$ 3 km). The captured signal sequences were filtered in software with Matlab. The normalized cut-off frequencies for each channel are shown in Table I, resulting in passbands equal to or greater than the signal BW, except for TV channels where some BW was required to accommodate the filter's transient band (for DAB-T the RF BW is 1.712 MHz but the signal information is confined within a BW of 1.54 MHz).

VI. Simulation performances and discussion

To evaluate the distribution of the statistic $\mathcal{F}_t$, we conducted series of simulations on the measurements obtained using the USRP. The considered bands are the following: channels Ch08A, Ch10A and Ch11B of the DAB-T standard. The data collected in these cases contain either noise or signal (i.e., $\mathbb{P}(\mathbb{H}_0) = i, i \in \{0,1\}$). For every band and each hypothesis ($\mathbb{H}_0$, $\mathbb{H}_1$) we collected 10 million samples.

A quick evaluation of the used data shows that in both cases, whether the samples follow hypothesis $\mathbb{H}_0$ or $\mathbb{H}_1$, and as a first approximation, we can indeed assume the samples as Gaussian as described in Section II. However, the Gaussian assumption is not perfect. The Gaussianity of the samples becomes more satisfactory if the data is further down-sampled in software. Since we are limited in the amount of date used, in this study we chose to down-sample the data by a factor 2.

Thus, three training sets composed of a random mixtures (K=2) of noise and signal collected from each of these bands are created. Then sequentially, the EM algorithm estimates the noise level (previously referred to as $g(\mathbf{h}_t)$) on 400 training sets of 25000 mixed samples of noise and signal. Thus, The learning process uses every time a set of 25000 samples (which

TABLE I

CHANNELS MEASURED IN THIS STUDY: ANALOGICAL/DIGITAL TV, TERRESTRIAL TRUNKED RADIO (TETRA), TERRESTRIAL DIGITAL AUDIO BROADCASTING (DAB-T), EXTENDED GLOBAL SYSTEM FOR MOBILE COMMUNICATIONS 900 DOWNLINK (E-GSM 900 DL), DIGITAL CELLULAR SYSTEM 1800 DOWNLINK (DCS 1800 DL) AND UNIVERSAL MOBILE TELECOMMUNICATIONS SYSTEM FREQUENCY-DIVISION DUPLEX DOWNLINK (UMTS FDD DL).

System	Channel number	f_{start} (MHz)	f_{center} (MHz)	f_{stop} (MHz)	Signal BW (MHz)	Decimation rate	Sampled BW (MHz)	Gain (dB)	Cut-off frequency	Pass band (MHz)
Analogical TV	23 29 34 38	486 534 574 606	490 538 578 610	494 542 582 614	8	8	8	10	0.94	7.52
Digital TV	26 48 61 67	510 686 790 838	514 690 794 842	518 694 798 846	8	8	8	10	0.94	7.52
TETRA	37 44 45 47 53	420.8875 421.0625 421.0875 421.1375 421.2875	420.900 421.075 421.100 421.150 421.300	420.9125 421.0875 421.1125 421.1625 421.3125	0.025	256	0.25	70	0.1	0.03
DAB-T	08A 10A 11B	195.080 209.080 217.784	195.936 209.936 218.640	196.792 210.792 219.496	1.712	32	2	70	0.8	1.6
E-GSM 900 DL	60 113 975	946.8 957.4 925.0	947.0 957.6 925.2	947.2 957.8 925.4	0.2	64	1	70	0.3	0.3
DCS 1800 DL	546 771 786	1811.8 1856.8 1859.8	1812.0 1857.0 1860.0	1812.2 1857.2 1860.2	0.2	64	1	70	0.3	0.3
UMTS FDD DL	10588 10663 10738	2115.1 2130.1 2145.1	2117.6 2132.6 2147.6	2120.1 2135.1 2150.1	5	8	8	70	0.625	5

could be considered, depending on the standards, as few slots: ≤ 20 slots).

We only consider, for the next results, the ratio statistic under hypothesis $\mathbb{H}_0$. Evaluating this distribution would allow to develop a joint learning-detection framework, where the new detector could improve its detection capabilities while ensuring a given false alarm.

Figures 2, 3 and 4 show the learning results, respectively, for the channels Ch08A, Ch10A and Ch11B. In all figures, three curves appears: the empirical ratio distribution, the theoretical ratio distribution and Fisher-Snedecor distribution. First the empirical distribution represent, the ratio distribution where the function $f(\cdot)$ is described in Section III with $M = 100$ (however, different values for M were tested leading to equivalent results). The theoretical ratio distribution presents the simulated distribution of a ratio of independent scaled-Chi-square and scale-inverse Chi-square distributions with parameters respectively $M = 100$ and $N = 25000$. This is due to the fact that the EM algorithm computes the posterior distribution of the variance, which is known to follow a scale-inverse Chi-square distribution. Finally, a Fisher-Snedecor distribution with parameters $M = 100$ and $N = 25000$ is also presented as possible approximation. This latter is usually employed in the analysis of variance. Since it only depends on known parameters, this approximation might be convenient to design a future detector.

Notice that in all cases, the empirical distribution can be fit, as a first approximation, by any of the other distributions. More specifically, in Figure 4 the empirical ratio distribution seems particularly well fitted. The biased observed can be explained by two factors: on the one hand, as noticed before, the samples are not perfectly Gaussian. On the other hand the EM algorithm stops sometimes before convergence which leads to a bad evaluation of the noise level. This results seems to suggest that is indeed possible to design a detector that relies on *a priori* known parameters M and N. Unfortunately although, these curves are interesting, they cannot be generalized yet. To confirm these results and to design an adapted joint learning-detector, a rigorous theoretical analysis supported by extensive measurements are needed.

VII. CONCLUSION

We ventured, in this paper, the analysis of an energy detector with no *a priori* information on the noise level. It relies on the computation of a statistic based on the ratio of a function of the currently collected samples and of some information provided by a learning algorithm that exploits past information on the environment. Thus a joint learning-detection framework was suggested. A first empirical analysis showed that it might be possible to design a detector that depends only on two known parameters: the number of samples in a slot and the number of samples used for the training set. Although this research show interesting results, it is still in its infancy and these results still need to be confirmed. To that purpose a theoretical analysis supported by extensive empirical measurements are currently, under investigation.

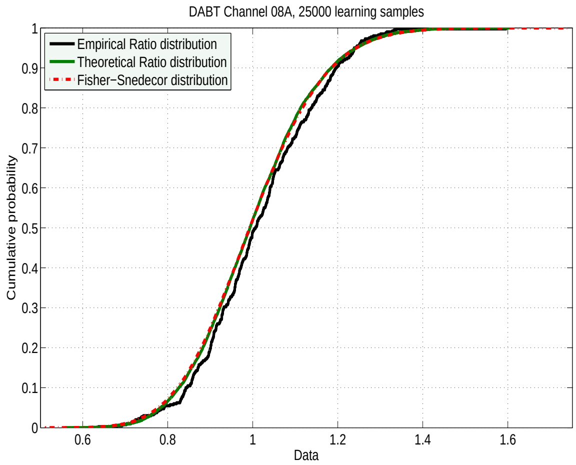

Fig. 2. Learning results tested on the channel 08A of the DAB-T standard. The curves compare the empirical ratio statistic to both the theoretical ratio distribution and Fisher-Snedecor distribution.

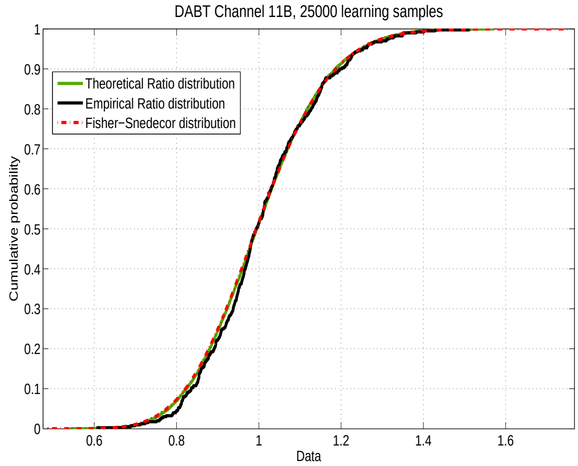

Fig. 4. Learning results tested on the channel 11B of the DAB-T standard. The curves compare the empirical ratio statistic to both the theoretical ratio distribution and Fisher-Snedecor distribution.

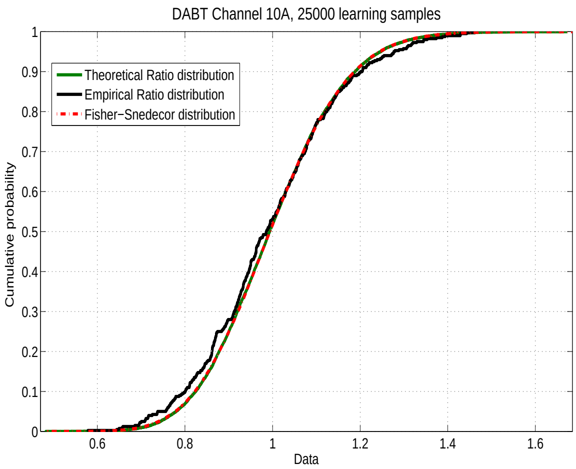

Fig. 3. Learning results tested on the channel 10A of the DAB-T standard. The curves compare the empirical ratio statistic to both the theoretical ratio distribution and Fisher-Snedecor distribution.

ACKNOWLEDGMENT

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM++ (contract no. 216715).

REFERENCES

[1] Federal Communications Commission. Spectrum policy task force report. November 2002.

[2] M. A. McHenry et al. Spectrum occupancy measurements. Technical report, Shared Spectrum Company, Jan 2004 - Aug 2005. Available at: http://www.sharedspectrum.comhttp://www.sharedspectrum.com.

[3] R. I. C. Chiang, G. B. Rowe, and K. W. Sowerby. A quantitative analysis of spectral occupancy measurements for cognitive radio. In *Proceedings of the IEEE 65th Vehicular Technology Conference (VTC 2007 Spring)*, pages 3016–3020, 2007.

[4] M. Wellens, J. Wu, and P. Mähönen. Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio. In *Proceedings of the Second International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrowCom 2007)*, pages 1–8, 2007.

[5] M. López-Benítez, F. Casadevall, A. Umbert, J. Pérez-Romero, J. Palicot, C. Moy, and R. Hachemani. Spectral occupation measurements and blind standard recognition sensor for cognitive radio networks. In *Proceedings of the 4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2009)*, pages 1–9, 2009.

[6] J. Mitola. Cognitive radio: An integrated agent architecture for software defined radio. *PhD Thesis, Royal Inst. of Technology (KTH)*, 2000.

[7] T. Yucek and H. Arslan. A survey of spectrum sensing algorithms for cognitive radio applications. *In IEEE Communications Surveys and Tutorials*, 11, no.1, 2009.

[8] Leus G. Axell, E. and G. E. Larsson. Overview of spectrum sensing for cognitive radio. *CIP2010 IAPR Workshop on Cognitive Information Processing, pp. 322-327*, 11, no.1, 2010.

[9] R. Tandra and A. Sahai. SNR walls for signal detection. *IEEE Journal of Selected Topics in Signal Processing*, 2(1):4–17, Feb. 2008.

[10] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. In *Journal of the Royal Statistical Society. Series B 39(1)*, pages 1–38, 1977.

[11] G.J. McLachlan and T. Krishnan. The em algorithm and its extensions. In *Wiley*, 1997.

[12] A. Clifford Cohen. Estimation in mixtures of two normal distributions. *Technometrics, Vol. 9, No.1*, pages 15–28, 1967.

[13] Ettus Research LLC. Universal Software Radio Peripheral (USRP). Available at: http://www.ettus.com.

[14] E. Blossom. Exploring GNU Radio, November 2004. Available at: http://www.gnu.org/software/gnuradio/doc/exploring-gnuradio.html.