

## RESEARCH ARTICLE

# Kernel Canonical Correlation Analysis for Robust Cooperative Spectrum Sensing in Cognitive Radio Networks

J. Manco-Vásquez\*, S. Van Vaerenbergh, J. Vía, I. Santamaría  
Dept. of Communications Engineering, University of Cantabria, Spain

## ABSTRACT

Spectrum sensing is a key operation in Cognitive Radio (CR) systems, where secondary users (SUs) are able to exploit spectrum opportunities by first detecting the presence of primary users (PUs). In a CR network composed of several SUs, the detection accuracy can be much improved by cooperative spectrum sensing (CSS) strategies, which exploit the spatial diversity among SUs. However, cooperative detection strategies, which are typically based on energy sensing, do not perform satisfactorily under impairments such as non-Gaussian noise or interferences. In this paper, we propose a scheme based on kernel canonical correlation analysis (KCCA), which is able to operate in non-ideal scenarios and in a totally blind fashion. This technique is performed at the fusion center (FC) by exploiting the non-linear correlation among the received signals of each SU. In this manner, statistical tests are extracted, allowing the SUs to make decisions either autonomously at each SU or cooperatively at the FC. The performance of the KCCA-based detector is evaluated by means of simulations and over-the-air experiments using a CR testbed composed of several Universal Radio Peripheral (USRPs) nodes. Both the simulations and the measurements show that the KCCA-based detector is able to obtain a significant gain over a conventional energy detector, whose sensing performance is severely degraded by the presence of external interferers.

Copyright © 2014 John Wiley & Sons, Ltd.

## KEYWORDS

Cooperative Spectrum Sensing; Kernel Canonical Correlation Analysis; Hardware Testbed; USRP.

## \*Correspondence

J. Manco-Vásquez, Dept. of Communications Engineering, University of Cantabria, Spain

E-mail: juliocesar@gtas.dicom.unican.es.

## 1. INTRODUCTION

The enormous increase of wireless applications has led to an inefficient use of spectral resources, by leaving empty or overcrowded some parts of the wireless spectrum [1, 2]. This problem is foreseen to be mitigated by Cognitive Radio (CR) technology, under which incumbent or primary users (PUs) and non-legacy or secondary users (SUs) coexist. CR relies on a fast and accurate spectrum sensing process that detects exploitable time-frequency holes, which are subsequently utilized for transmissions by the SUs. Common impairments found in local spectrum sensing, such as fading, shadowing, hidden terminals, and receiver uncertainty, can be overcome by applying cooperative spectrum sensing (CSS) strategies, which exploit the diversity among CR users [3].

However, other impairments such as non-Gaussian noise or the presence of narrowband external interferences

might affect negatively the performance of spectrum sensing techniques: this is the challenging sensing scenario we consider in this paper. Interference, which is sometimes modeled as non-Gaussian noise, may arise from external user operation, either intentionally or unintentionally [4]. As shown in [5, 6], the performance of the energy detector, which is the most common spectrum sensing mechanism, is strongly degraded under interference. Basically, without additional information, the energy detector is unable to distinguish the primary signals from the interference [6]. In [7], several eigenvalue-based cooperative sensing techniques are evaluated under impulsive noise and interference, showing also a significant degradation of their performance and lack of robustness. It is also worth mentioning that, compared to local spectrum sensing, the implementation of CSS strategies might be affected by other impairments such as timing inaccuracies or

synchronization errors among the SUs for simultaneous local sensing when the channel is idle [8, 9].

A typical scenario of current interest where the presence of interference can impair spectrum sensing can be found in Heterogeneous Networks (HetNets), where a macrocell-edge user may experience interference from small cell transmissions using the same radio frequency band. This scenario is considered in [10], where an interference-mitigation scheme close to macrocell/femtocell real-life scenario is experimentally evaluated. Another recent work that takes into account interference in the CR context is [11], where, with the assistance of geolocation information, a sensing scheme is proposed that decomposes the received power into the primary signal power, secondary signal power (treated here as interference), and the device noise power. In this way, after decomposing the total power, the interference power can be canceled prior to PU detection. The impact of interference in underlay cooperative cognitive networks has also been extensively studied [12], [13]. Distinct from these works, we focus in this paper on interweave cooperative cognitive networks without any geolocation assistance or any other statistical prior information, and propose to apply a kernel-based method for detection.

Recently, the introduction of machine learning techniques in CR applications [14] has shown to improve the detection performance of soft-decision approaches. In CR applications, prediction schemes based on machine learning techniques have been also proposed for opportunistic channel selection [15]. In [16, 17], the energy levels measured at each SU are reported to the FC. This set of energy levels, arranged as feature vectors, are fed into a classifier that categorizes them into classes that represent whether the channel is available or not. The classifier first requires a *training phase*, during which it learns from a set of training feature vectors. Then, it can be employed for online detection, in what is typically known as the *test phase*.

In this paper, we propose a KCCA-based technique for robust cooperative spectrum sensing in a scenario exposed to external interferers. We consider a distributed configuration in which the SUs do not communicate with each other and only report their local measurements to a FC. The technique is applied at the FC, and exploits the non-linear learning capabilities of kernel-based methods [18], which have been used previously in the context of cognitive radio networks, for instance in [19]. Previous kernel-based CR detectors follow a supervised approach in which it is assumed that a set of patterns, labeled with the correct decisions, is available for training the classifier. Our approach, however, does not require neither any labeled data set nor any other prior information about the PU signalling format, and thus operates in a completely blind fashion.

More specifically, the proposed scheme only exploits the (possibly) non-linear correlation among the received measurements at the FC during an initial cooperative stage.

This is carried out by extracting non-linear transformations which are employed as statistical tests. The received measurements, reported by each SU, can be composed of different features, such as the kurtosis and the energy of the data acquired at each sensing period. We stress again that these features do not need to be labeled with the corresponding states of the primary signal, and as such no additional prior information is required. In fact, the proposed technique could be easily adapted to a time-varying radio environment by re-training the detector from time to time or continuously while the detection operates normally.

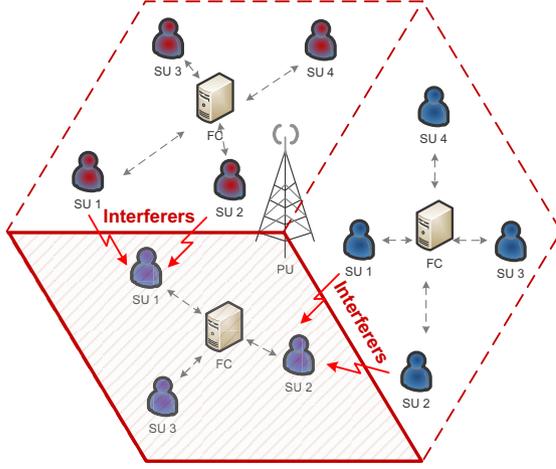
We consider a general setting, where a PU has a large radio coverage, while interferers have a small coverage area and hence each affects a single SU. Some initial results were presented in [20]. In this paper we extend this work and present a more detailed study of the proposed CR detector, as well as a complete experimental evaluation that corroborates the results obtained by simulations. The experiments were conducted in a cognitive radio testbed composed of several USRP devices [21], emulating a scenario where a PU and several SUs, possibly affected by interferences, coexist.

The rest of the paper is organized as follows. In Section 2, we give an overview of the CSS problem. A detailed description of the proposed KCCA-based detector and its operation is presented in Section 3. In Section 4, we analyze the simulation results for different scenarios. The description of the CR testbed and the measurement procedure along with the experimental results are exposed in Section 5. Finally, the paper concludes with a discussion of the obtained results in Section 6.

## 2. COOPERATIVE SPECTRUM SENSING

Let us consider a cooperative spectrum sensing scenario where  $M$  SUs and a PU coexist in the same area [3]. We assume the PU has a large coverage area and then it can be sensed by several SUs. During an initial learning phase, the sensor measurements are sent to the FC, which extracts the local decision functions in a completely unsupervised manner. After this unsupervised learning stage, the SUs are able to operate autonomously, or they can still cooperate by sending their local decisions to the FC, which can subsequently combine them to make a global decision.

In order to take into account the potential presence of local interferers while using a very general signal model, we simply assume the independence of the measurements under the null hypothesis (idle channel). In words, this means that the interferences seen by different SUs are independent of each other. More formally, the binary hypothesis testing problem considered in this paper can be



**Figure 1.** A spectrum sensing problem in a HetNet. Three SUs in a small cell cooperate to detect the presence of a PU, while two of them receive interference from other small cells. The interferences are independent of each other.

formulated as follows:

$$p(\mathbf{r}|\mathcal{H}_1) \neq \prod_{i=1}^M p_i(r_i|\mathcal{H}_1)$$

$$p(\mathbf{r}|\mathcal{H}_0) = \prod_{i=1}^M p_i(r_i|\mathcal{H}_0)$$

where  $r_i$  denotes the received signal at the  $i$ -th SU,  $\mathbf{r}$  is a vector signal composed of all observations,  $\mathcal{H}_1$  denotes the alternative hypothesis (PU active) and  $\mathcal{H}_0$  is the null hypothesis. Notice that the primary, interference and noise signal may follow any distribution, since we do not make any assumptions about them. Thus, the model is rather general and, in particular, is independent of the underlying technology utilized during the transmissions by the PU and the interferers. A particular scenario where these assumptions hold is depicted in Fig. 1, where a small cell (shaded) within a heterogeneous network (HetNet) receives interference from neighboring cells during the time that the channel is considered vacant.

### 3. KERNEL CANONICAL CORRELATION ANALYSIS FOR CSS

The primary purpose of the proposed CSS framework is to correctly determine, locally at the SUs, the channel availability based on a set of features extracted from the local measurements. The main idea of the proposed detector is very simple. Although we cannot obtain the optimal (Neyman-Pearson) detector at each SU, since the distributions under the two hypothesis are unknown, we know that the test statistics of the optimal detectors at each SU will be highly correlated, since the SUs are either all under the null hypothesis or all under the

alternative hypothesis. Therefore, we will look for the non-linear transformations of the measurements providing maximal correlation, which are expected to be monotone transformations of the optimal test statistics. That is, the proposed scheme aims to exploit the non-linear correlation among SUs at the FC to decide if the measurements come from the distribution  $p_i(r_i|\mathcal{H}_1)$  or from  $p_i(r_i|\mathcal{H}_0)$ .

The operation of the proposed sensing paradigm is illustrated in Fig. 2. In an initial cooperative learning stage, the sensor measurements are transmitted to the FC, which extracts the near-optimal local decision functions. These functions are broadcasted to the SUs, which can then operate in one of two modes (cf. Fig. 2):

1. Autonomous testing: Each SU takes independent decisions based on its local test statistic.
2. Cooperative testing: Each SU transmits its local test statistic to the FC, where a global decision is finally made by combining the local test statistics.

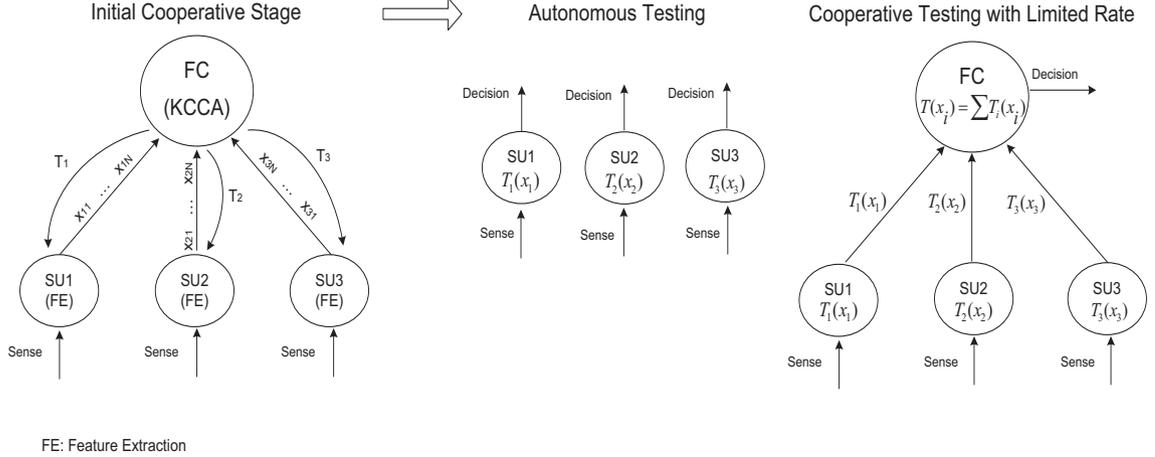
It is interesting to highlight that the transmission of information from the SUs to the FC needed in the cooperative testing mode is very limited. Specifically, each SU only needs to transmit its test statistic (a scalar value) instead of the whole set of measurements or feature vectors. Also, notice that as a byproduct of the process for extracting the local decision function, we obtain a quantitative indicator of the sensing performance of each SU. These indicators can be directly used for selecting a reduced number of sensors in the cooperative operation mode, thus further reducing the communication requirements of the whole procedure.

#### 3.1. Local feature extraction

Feature vectors are extracted from the measurements at each SU and used as input for the KCCA-based detector. We denote a feature vector as  $\mathbf{x}_{in}$ , where  $i$  refers to the  $i$ -th SU, and  $n$  denotes the  $n$ -th sensing period during which  $N_s$  samples of the received signal  $r_i$  are sensed. We denote the feature vector extracted by the  $i$ -th SU during the  $n$ -th sensing period as,

$$\mathbf{x}_{in} = \left( f_{in}^1 \ f_{in}^2 \ \dots \ f_{in}^{N-1} \ f_{in}^N \right)^T$$

where  $f_{in}^j$  is the  $j$ -th feature. For instance, if only the measured energy is considered,  $\mathbf{x}_{in} = f_{in}^1$  will be a scalar value. A wide variety of features can be included into the feature vector such as energy, kurtosis, or cyclic statistics, among others. Finding the optimal features is a challenging problem because of the different trade-offs that exist among the performance, number of features, number of available data, and temporal coherence of the channel. In this paper, we will mainly consider the energy and the kurtosis of the signal as the main features for the detector. A detailed analysis of the optimal feature extraction procedure will be considered in a future work.



**Figure 2.** Operation of the proposed KCCA scheme: In an initial cooperative stage (left side of the picture), the measurements are reported to the FC to extract the local statistical tests  $T_i$ , then it starts operating either following a distributed (Autonomous Testing) or centralized configuration (Cooperative Testing). In the distributed configuration each SU makes a decision after a sensing period, whereas in the centralized configuration all local test statistics are reported to the FC, where a global decision is finally made.

### 3.2. Initial Cooperative Stage

In the initial cooperative stage, the feature vectors extracted at each SU are reported to the FC, where the statistical dependencies among the different SUs are retrieved. In particular, we seek to combine the feature vectors for each SU individually in such a manner that the resulting combinations are maximally correlated among the different SUs.

The technique of canonical correlation analysis (CCA) allows to retrieve the linear projections of the feature vectors that provide maximum correlation among the SUs. In order to allow the optimal projections to be non-linear, we resort to the kernel-based version of CCA, known as KCCA [20, 22]. This procedure consists in mapping the data into a high-dimensional space first, after which standard CCA is performed in the new space.

#### 3.2.1. Kernel-Based Learning

In kernel-based learning (KBL), the data is transformed into a high-dimensional *feature space* [23],

$$\Phi : \mathbf{x}_{in} \rightarrow \Phi(\mathbf{x}_{in}). \quad (1)$$

While explicit calculations in the new space may be hard due to its high dimensionality, for certain feature spaces it is possible to calculate inner products as a positive definite kernel function  $\kappa(\cdot, \cdot)$  in the input space. This is the case when Mercer's condition is satisfied,

$$\kappa(\mathbf{x}_{ij}, \mathbf{x}_{ik}) = \langle \Phi(\mathbf{x}_{ij}), \Phi(\mathbf{x}_{ik}) \rangle. \quad (2)$$

In order to illustrate the concept of the feature space induced by a kernel, we consider a simple polynomial kernel of the form  $\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j)^2$ . Given a two-dimensional feature vector  $\mathbf{x}_i = (f_{i1}, f_{i2})$  that is only

composed of energy levels (where the upper index in  $f_{in}^{index}$  has been omitted for clarity purposes), this kernel can be expanded in individual terms as

$$\begin{aligned} (\mathbf{x}_i^T \mathbf{x}_j)^2 &= (f_{i1}f_{j1} + f_{i2}f_{j2})^2 \\ &= f_{i1}^2 f_{j1}^2 + 2f_{i1}f_{j1}f_{i2}f_{j2} + f_{i2}^2 f_{j2}^2 \\ &= (f_{i1}^2, \sqrt{2}f_{i1}f_{i2}, f_{i2}^2)(f_{j1}^2, \sqrt{2}f_{j1}f_{j2}, f_{j2}^2)^T \\ &= \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j). \end{aligned}$$

In this case, the feature mapping takes the form  $\phi(\mathbf{x}_i) = (f_{i1}^2, \sqrt{2}f_{i1}f_{i2}, f_{i2}^2)^T$ , which corresponds to a three-dimensional feature space. The polynomial kernel is typically used in its more general formulation,

$$\kappa(\mathbf{x}_{ij}, \mathbf{x}_{ik}) = (\mathbf{x}_{ij}^T \mathbf{x}_{ik} + d)^p,$$

where  $p$  and  $d$  are the order of the polynomial kernel and a constant, respectively. In this paper, we consider the standard Gaussian kernel with kernel width  $w_i$ , given by

$$\kappa(\mathbf{x}_{ij}, \mathbf{x}_{ik}) = \exp(-\|\mathbf{x}_{ij} - \mathbf{x}_{ik}\|^2 / 2w_i^2),$$

which induces an infinitely-dimensional feature space [23]. We maintain the subindex  $i$  to indicate that the kernel parameter may be chosen differently for each SU.

The *Gram matrix* (or *kernel matrix*)  $\mathbf{K}_i$ , for the data set obtained at the  $i$ -th SU, contains pairwise kernels of the data as its elements,

$$K_i(j, k) = \kappa(\mathbf{x}_{ij}, \mathbf{x}_{ik}) = \Phi(\mathbf{x}_{ij})^T \Phi(\mathbf{x}_{ik}).$$

#### 3.2.2. Kernel Canonical Correlation Analysis for CSS

Consider a scenario in which  $M$  SUs are present, and each SU produces  $N$  feature vectors,  $\{\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iN}\}$ . In

order to define the correlation between multiple data sets, a summation of the individual correlations of each pair of data sets can be used\*.

The pairwise canonical correlations between the data sets,  $\rho_{ij}$ , are obtained in the context of KCCA as  $\rho_{ij} = \mathbf{z}_i^\top \mathbf{z}_j = \boldsymbol{\alpha}_i^\top \mathbf{K}_i \mathbf{K}_j \boldsymbol{\alpha}_j$  [26], where  $\mathbf{z}_i = \mathbf{K}_i \boldsymbol{\alpha}_i$  is a canonical variate obtained as the projection of the  $i$ -th set of data by means of the canonical vector  $\boldsymbol{\alpha}_i$ . A measure of the correlation associated to the  $i$ -th data set,  $\rho_i$ , can be subsequently obtained as,

$$\rho_i = \frac{1}{M-1} \sum_{\substack{j=1 \\ j \neq i}}^M \rho_{ij}, \quad (3)$$

and a generalized canonical correlation can be obtained as

$$\rho = \frac{1}{M} \sum_{i=1}^M \rho_i. \quad (4)$$

The maximization of  $\rho$  with respect to the canonical vectors  $\boldsymbol{\alpha}_i$  admits a trivial solution, which can be easily avoided by means of the following constraint on the energy of the canonical variates

$$\frac{1}{M} \sum_{i=1}^M \|\mathbf{z}_i\|^2 = \frac{1}{M} \sum_{i=1}^M \boldsymbol{\alpha}_i^\top \mathbf{K}_i \mathbf{K}_i \boldsymbol{\alpha}_i = 1.$$

Analogously, overfitting problems can be avoided by adding a regularization factor,  $c$ , to the norm of the projectors in the previous constraint [22].

$$\frac{1}{M} \sum_{i=1}^M \boldsymbol{\alpha}_i^\top \mathbf{K}_i \mathbf{K}_i \boldsymbol{\alpha}_i + c \boldsymbol{\alpha}_i^\top \mathbf{K}_i \boldsymbol{\alpha}_i = 1 \quad (5)$$

The canonical weights  $\boldsymbol{\alpha}_i$  are obtained by maximizing  $\rho$  subject to the restriction given in equation (5). This can be solved by the method of Lagrange multipliers, yielding the following generalized eigenvalue problem (GEV)

$$\frac{1}{M} \mathbf{R} \boldsymbol{\alpha} = \beta \mathbf{D} \boldsymbol{\alpha}, \quad (6)$$

where  $\mathbf{R}$ , for  $M$  sets of data, is defined as

$$\mathbf{R} = \begin{bmatrix} \mathbf{K}_1 \mathbf{K}_1 & \cdots & \mathbf{K}_1 \mathbf{K}_M \\ \vdots & \ddots & \vdots \\ \mathbf{K}_M \mathbf{K}_1 & \cdots & \mathbf{K}_M \mathbf{K}_M \end{bmatrix}, \quad (7)$$

and  $\mathbf{D}$  is given by

$$\mathbf{D} = \begin{bmatrix} \mathbf{K}_1(\mathbf{K}_1 + c\mathbf{I}) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{K}_M(\mathbf{K}_M + c\mathbf{I}) \end{bmatrix}. \quad (8)$$

\* Several generalizations of CCA to more than two sets of variables can be found in [24, 25].

The solution  $\boldsymbol{\alpha}$  contains the different canonical weights as stacked vectors,  $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2^\top, \dots, \boldsymbol{\alpha}_M^\top]^\top$ , and it is retrieved as the eigenvector corresponding to the largest eigenvalue of the GEV problem (6) [20] [27]. The eigenvalue  $\beta$  relates to the generalized canonical correlation as  $\beta = \frac{1+(M-1)\rho}{M}$ .

The squared norm of each of the canonical variates,  $\mathbf{z}_i$ , indicates the contribution of each of the data sets to the final canonical correlation. Therefore, in the CSS scenario it provides an indication of the reliability of each sensor when implementing, for instance, the centralized cooperative testing at the FC.

### 3.2.3. Data Centering

Canonical Correlation Analysis (CCA) requires the input data to have zero mean. Since KCCA applies CCA in the feature space, the data must be centered in this space [22],

$$\sum_{n=1}^N \Phi(\mathbf{x}_{in}) = 0, \quad i = 1, \dots, M. \quad (9)$$

Since the transformations  $\Phi$  are not necessarily explicitly known, it may be impossible to obtain centered versions of the data in feature space. However, it is possible to find the Gram matrix of the centered data points as

$$\tilde{\mathbf{K}}_i = \mathbf{N}_o \mathbf{K}_i \mathbf{N}_o^\top, \quad (10)$$

where  $\mathbf{N}_o = (\mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top)$ ,  $\mathbf{1}$  is an  $N \times 1$  all-one vector and  $\mathbf{I}$  the  $N \times N$  unit matrix. In order to center a vector of kernel elements,  $\mathbf{k}_i$ , defined as

$$\mathbf{k}_i(k) = [\kappa(\mathbf{x}_{ij}, \mathbf{x}_{ik})]_{k=1, \dots, N}.$$

a similar procedure is followed [28], leading to

$$\tilde{\mathbf{k}}_i = \left( \mathbf{k}_i - \mathbf{1}^T \tilde{\mathbf{K}}_i / N \right) \mathbf{N}_o. \quad (11)$$

### 3.3. Local and Global Tests

As a result of the KCCA learning stage, we obtain the following non-linear local detectors

$$T_i(\mathbf{x}_i) = \sum_{j=1}^N \alpha_{ij} \tilde{\kappa}(\mathbf{x}_i, \mathbf{x}_{ij}) \quad (12)$$

where  $\alpha_{ij}$  refers to the  $j$ -th element of the canonical vector  $\boldsymbol{\alpha}_i$ , and  $\tilde{\kappa}(\cdot, \cdot)$  refers to the kernel function calculated on the centered data. In essence, the statistical tests  $T_i$  constitute a weighted sum of similarities, as measured by the kernel functions. Notice that, since the feature vectors entering the expansion are already available at each sensing device, in order to compute (12) locally the FC only has to transmit to each SU its own canonical vector. This is the only transmission required if an autonomous testing procedure is followed.

On the other hand, the local decisions at the SUs can be easily combined at the FC if a cooperative testing

procedure is preferred. In this case, the global test statistic is simply obtained as

$$T_i(\mathbf{x}) = \sum_{i=1}^M T_i(\mathbf{x}_i), \quad (13)$$

which represents the best one-dimensional approximation of the (norm constrained) canonical variates. As we will see later, this additional cooperative stage results in an improved detection performance.

## 4. SIMULATION RESULTS

In this section, we study the detection performance of the proposed KCCA-based detector. We consider different scenarios in which noise, or noise plus interference are present, and for which different features are extracted during the sensing period. The performance is quantified in terms of probability of detection ( $P_D$ ) and probability of false alarm ( $P_{FA}$ ), by showing the Receiving Operating Characteristic (ROC) curves.

The following examples are evaluated for a number of training data  $N = 300$ , and  $N_s = 50$  samples per sensing period. The selection of the value  $N = 300$  corresponds to a tradeoff between the complexity to solve a GEV problem (recall that each kernel matrix in the GEV problem has dimensions  $N \times N$ ), and the obtained detection performance. Also, it is important that the scenario remains more or less static over the whole training period, which also calls for using a reduced number of sensing periods. On the other hand, we should mention that the value of  $N_s = 50$  does not target a particular application or standard<sup>†</sup>. Again, this value has been chosen mainly for computational reasons, as well as to avoid abrupt changes in the scenario statistics.

For the KCCA-based detector a Gaussian kernel is employed, the kernel width  $w_i$  for each set of data  $\mathbf{x}_i$  is chosen by applying the Silverman's rule [28, 30], and the regularization parameter is set to  $c = 10$ .

### 4.1. Decision functions and ROC Curves

For each example, we plot the estimated probability density function (PDF) of the feature used as input of the test statistic under both hypotheses<sup>‡</sup>, as well as the decision functions  $T_i$ , which represent the projections of the transformed data sets.

This allows us to study how the KCCA decision function is able to separate both hypotheses. In most

cases, the decision function for only one of the SUs is plotted, since similar curves are obtained among all SUs. In addition, a comparison of the ROC curves between an energy detector and the proposed KCCA-based detector is shown for the considered cases. We consider both configurations, a distributed (autonomous testing at each SU) and centralized configuration (cooperative testing at the FC). A stationary channel is considered, and both the PU and the interferers employ orthogonal frequency division multiplexing (OFDM) waveforms during their transmissions.

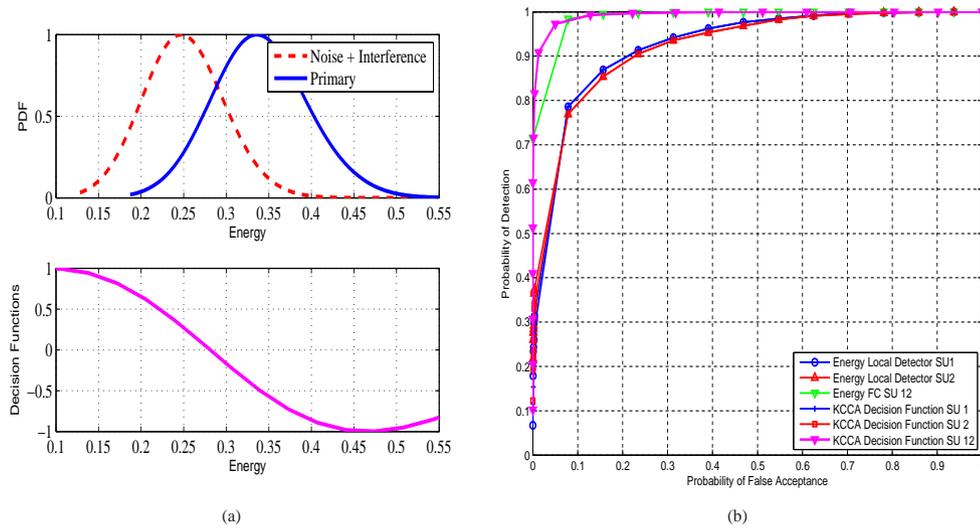
**Example 1.** In Fig. 3(a) a scenario is considered with two SUs ( $M = 2$ ), a PU, and only Gaussian noise under the null hypothesis. For this case, the feature vector is only composed of the measured energy, and therefore its PDF follows chi-squared distributions, which can be approximated by Gaussian distributions. A near-linear decision function is obtained by KCCA, which assigns negative values to the primary signal and positive ones to the noise. In Fig. 3(b), we show the corresponding ROC curves for a distributed and centralized configuration, where similar results are obtained by applying either KCCA or an (optimal) energy detector.

**Example 2.** For the same scenario, we now consider the presence of an interferer under the null hypothesis. Fig. 4(a) shows that the interference power is much higher than the primary signal, thus requiring a more complex decision function. In this example, the obtained KCCA decision function assigns high values to the noise and the interference signal, whereas low values are assigned to the primary signal. Notice also that the use of a Gaussian kernel function is related with the shape of the decision function, and for that reason very low or very high values of energy levels are mapped around zero. Nevertheless, this saturating effect can be avoided by applying a different kernel function or by setting a different kernel width. Furthermore, its impact does not affect the performance since these extreme energy values rarely occur. In fact, this might even increase the robustness of the proposed detector under impulsive noise. As it is depicted by the ROC curves in Fig. 4(b), we observe that the energy detector is clearly outperformed by the proposed KCCA-based scheme, which is able to distinguish the PU and the interference signals based solely on the correlation among test statistics.

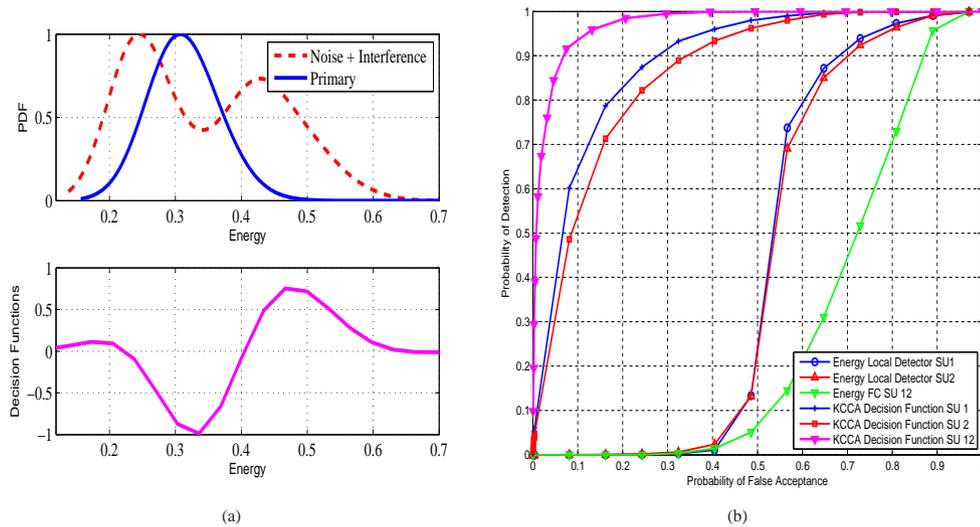
**Example 3.** Finally, in Figs. 5(a) and 5(b) we have considered a scenario with  $M = 3$  SUs, where the advantages of including more information in the feature vector is illustrated. In this example, the interferer utilizes a BPSK single-carrier modulation and we consider a feature vector composed of the energy and the kurtosis estimated over the sensing period. The PDF corresponding to the energy is shown in Fig. 5(a), where we observe that the energy of the primary signal almost overlaps with that of

<sup>†</sup>In practice, much larger sensing periods are typically used. For instance, the requirements of the spectrum sensing of ATSC DTV signals establish that the miss detection should not exceed 0.1 subject to a  $P_{fa} = 0.1$  when the SNR is -20.8 dB, these requirements yield sensing periods of thousands of samples at a sampling rate of 21.52 MHz [29].

<sup>‡</sup>The PDFs for the results shown in this paper are obtained using a Parzen density estimator with a Gaussian kernel [31].



**Figure 3.** (a) Probability density function for the primary and noise signals at SU 1 for a SNR  $\approx -5.3$  dB, and the corresponding KCCA decision function  $T_i$ . (b) The corresponding ROC curves for local decisions (at each SU) and centralized decisions (at the FC) using KCCA and an energy detector.



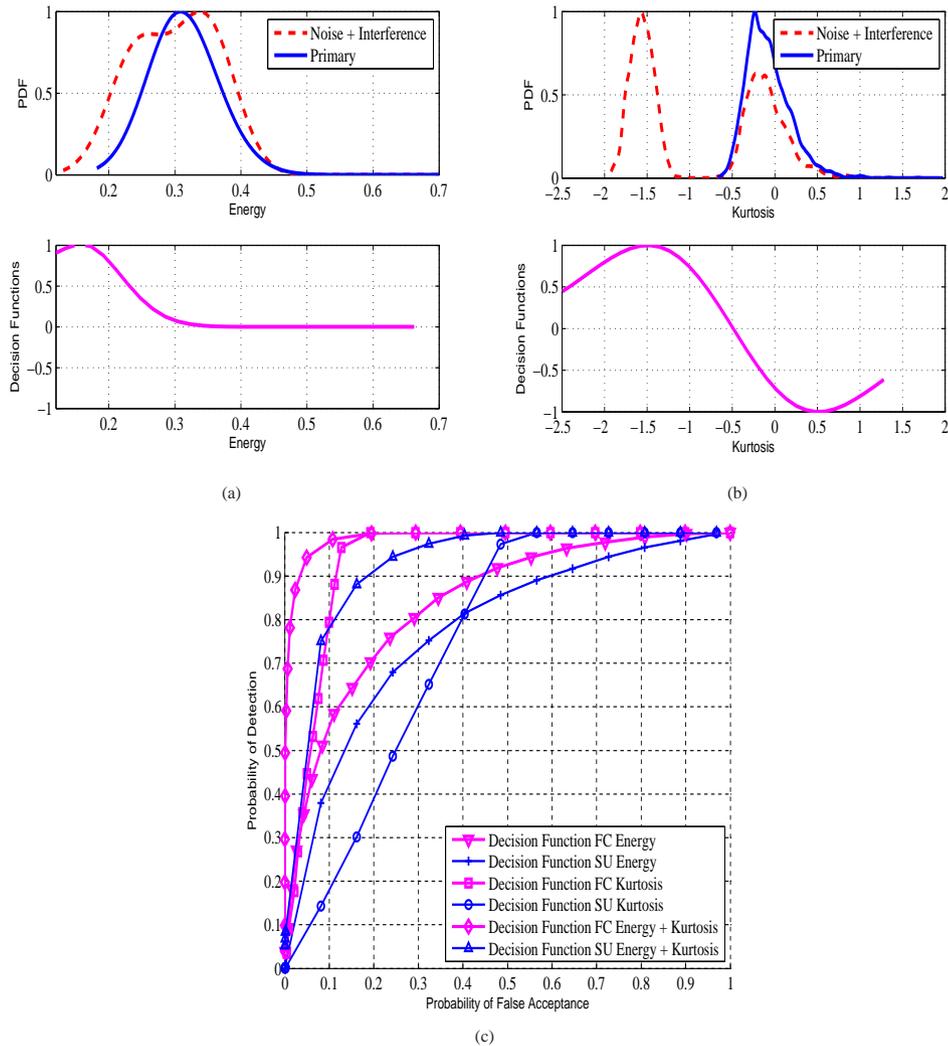
**Figure 4.** (a) Probability density function for the primary and noise-plus-interference signals at SU 1 for a SINR  $\approx -8.5$  dB, and the corresponding KCCA decision function  $T_i$ . (b) The corresponding ROC curves for local decisions (at each SU) and centralized decisions (at the FC) using KCCA and an energy detector.

the interfering signal, and thus this feature alone is not discriminative enough to detect the primary signal. This limitation can be avoided by including into the feature vector the kurtosis [32], which is defined as the normalized fourth-order cumulant,

$$kwr(r_{in}) = \frac{E(|r_{in}|^4) - |E(r_{in}^2)|^2 - 2E^2(|r_{in}|^2)}{E^2(|r_{in}|^2)}.$$

The PDF of the kurtosis is shown in Fig. 5(b), where it can be observed that it is unable to distinguish the primary signal from the noise, since both follow a Gaussian distribution. Therefore, neither the energy nor

the kurtosis alone seem to be able to distinguish the primary signal from the null hypothesis. However, if we use both features the proposed KCCA-based provides a considerable advantage, which is quantified by the corresponding ROC curve in Fig. 5(c).



**Figure 5.** PDF of the received signal under both hypothesis, and the KCCA decision function for a  $\text{SINR} \approx -7.45$  dB at the SU 1 (a) Energy (b) Kurtosis. The corresponding ROC curves for local decisions (at each SU) and centralized decisions (at the FC) with a KCCA-based detector using only one or both features (c).

## 5. EXPERIMENTAL RESULTS

### 5.1. Testbed Description

A cognitive radio platform has been built by integrating several USRP nodes in the laboratory of the Advanced Signal Processing Group at the University of Cantabria. Each of these nodes works with a universal hardware driver (UHD) as a host driver. By default this UHD driver allows us to control only a USRP device, which makes it more difficult to set up more complex scenarios. We have developed a custom Universal Software Architecture for Software Defined Radio (USASDR) that employs the UHD driver to operate simultaneously over several USRP devices from a remote PC running higher level instructions

from Matlab.

The transmitters and receivers are composed of N210 USRP motherboards and Radio Frequency (RF) XCVR2450 daughterboards; and allow us to operate in the ISM bands of 2.4GHz to 2.5GHz, and 4.9GHz to 5.8GHz. For a more detailed description of the node characteristics, the reader is referred to [21].

The processing chain at the transmitter side in our setup is as follows:

- After an instruction from Matlab is executed, the Gigabit Ethernet controller of the host computer transfers the data to the USRP. This received complex signal is upconverted to an analog

Intermediate Frequency (IF) signal and transmitted over the air by the RF transceiver.

On the other hand, the process at the receiver side is as follows:

- The flow of the signal at the receiver side is similar to its counterpart, but in a reverse order. After capturing the data, a Gigabit Ethernet controller is in charge of transferring it to the host computer where the rest of the signal processing tasks are performed. A detailed description of the flow of data with our custom implementation can be found in [33].

In addition, a Pulse Per Second (PPS) signal provided by an external clock is employed for timing synchronization among the nodes in the testbed, it allows the transmission and reception among the USRP nodes simultaneously, as it is shown in Fig. 6(a), where a PU, two SU nodes and an interfering node are configured and synchronized in time by a PPS signal for simultaneous transmission and reception during the measurement procedure.

Notice that the experimental part only considers two SUs, since we aim to show the feasibility of our proposal. An scenario composed of more SUs turns out to be interesting for boosting the performance as more feature vectors are available. However, it also involves higher complexity to solve the GEV problem (eq. 5), along with new approaches which deserve further research before implementing them in complex experimental scenarios.

## 5.2. Measurement Procedure

All the measurements were tested in an indoor quasi-static (the coherence time is rather long in comparison to the measurement time) channel of 4 MHz centered at 5.6 GHz. To recreate a scenario in which the interferences observed by each SU are independent, we divide the 4 MHz channel into 2 sub-channels of 2 MHz each. Each SU senses a different sub-channel, whereas the PU transmits over the whole 4 MHz channel. On the other hand, the interfering node randomly transmits on one of the two sub-channels, or on both simultaneously. Each interfering node follows independent Bernoulli distributions with a probability of sub-channel occupancy  $p = 0.5$ . In this configuration, either both SUs, only one of them, or neither of them will be affected by the interference, while both SUs are able to detect a busy channel when the PU is present. The transmission/sensing cycle is shown in Fig. 6(b), where the transmitted signal is an orthogonal frequency division multiplexing (OFDM) waveform that follows the IEEE 802.11a standard. This waveform is generated with a rate of 9 Mbps using BPSK symbols, and up-sampled to modify the bandwidth of the signal so as to accomplish the described configuration. After multiple sensing periods, two sets of data composed of the estimated energy levels

at each SU, are collected in a central PC acting as a FC. Finally, the canonical weights  $\alpha_i$  are calculated and used to form the statistic  $T_i(x)$ , whose performance is evaluated during an off-line process.

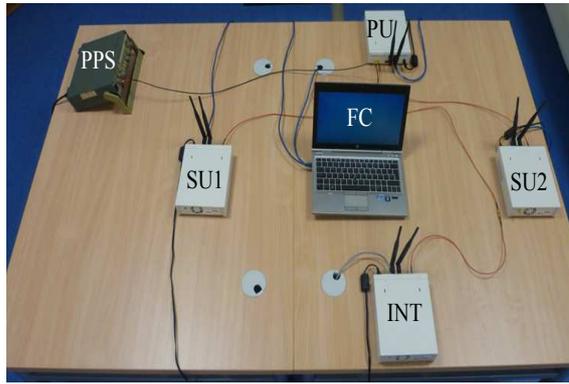
## 5.3. Experimental Measurements

In this section, we describe the experimental results obtained by the proposed procedure, and highlight the more challenging cases where the interference is present during the sensing period. The following results were obtained by a feature vector only composed of energy measurements with  $M = 2$  (number of SUs),  $N_s = 50$  (number of samples during each sensing period),  $N = 300$  (number of training patterns sent to the FC), and the ROC curves were computed after collecting 10000 sensing periods.

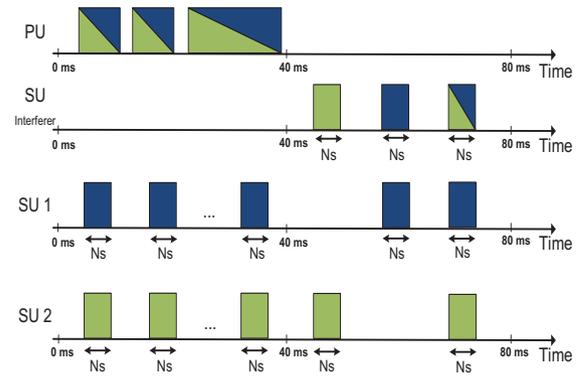
At the transmitter side, the maximum transmission power allowed by the N210 USRP is 5dBm, and it is controlled by applying a constant factor to the signal's amplitude. This allow us to control the measured SNR at the receiver side at baseband. On the other hand, the energy levels indicated in the experimental results correspond to the energy of the acquired discrete-time signal normalized by its maximum value. This normalization step plays the role of an automatic gain control (AGC) system, which is not implemented by the USRP nodes.

As we already mentioned, the measurements correspond to an indoor channel that presents long coherence times in comparison to the time elapsed during each data acquisition. In fact, for the same scenario it was shown in Gutierrez et al [34] that the channel remains almost constant at the band of 5 GHz with coherence times on the order of seconds. Thus, we expect that the PDFs under both hypothesis do not change abruptly, since the measured scenario is almost stationary. For a non-stationary environment, our scheme should include an updating procedure, but this is left as future work.

**Example 1.** In Figs. 7(a) and 7(b) the PDFs of the measured energy levels are shown for each SU under both hypotheses. It can be observed that the primary, the noise and the interfering signal approximately follow Gaussian distributions. For this case, the interference power lies below the received power of the primary signal, and as it is expected from the simulation results, the KCCA-based detector is able to separate the interference and noise from the primary signal, by mapping them to different values of the test statistic. The corresponding ROC curves for this example are shown in Fig. 8(a), where we see that each SU, when operating autonomously, obtains similar results. This can be explained by the fact that both detectors are close to the optimal solution to separate both PDFs. On the other hand, a slight improvement is obtained when the decision is cooperatively taken at the FC, as it employs all feature vectors from both SUs to attain a better performance.



(a)



(b)

**Figure 6.** (a) Two SUs acting as sensing nodes, an interfering node (INT), a PU, and a FC in the middle of them. All USRP are synchronized by a pulse per second signal (PPS) provided by Signal Generator. The SUs are located at approximately 1 m from the PU and the interfering node. (b) Measurement procedure: the PU transmits using two bands of frequency channels represented by two colors (2-4 & 4-6MHz), each SU senses a different band, and the interfering node transmits randomly on any of the channels, or in both.

**Example 2.** A more interesting case is depicted in Figs. 7(c) and 7(d), where the power of the interference signal is high enough to be above the primary signal power, and the primary and noise signals have similar energy levels at a SU. For this case, the energy detector is unable to distinguish between the noise and the primary signal. However, in spite of the degraded measurement at one of the SUs, the KCCA-based detector obtains a significant improvement, as the ROC curves in Fig. 8(b) show. This advantage can be attributed to the fact that the KCCA detector effectively exploits the non-linear correlation between the sensor measurements at the FC. In addition, if some knowledge about the PU signal is available, it could be easily exploited by our framework to boost its performance detection.

**Example 3.** A similar measurement result is shown in Figs. 7(e) and 7(f), and its corresponding ROC curve in Fig. 8(c). In this example, the interference power is above that of the primary signal, and the obtained performance corroborates the simulations results given in Figs. 4(a), 4(b). Moreover, as the interference level increases the proposed technique exhibits a much better performance than that of the energy detector. In fact, our KCCA framework is able to deal with different noise variances found at each SUs, since it learns from the particular feature vectors reported by each SU.

## 6. CONCLUSIONS

In this paper, we have derived a KCCA-based detector for spectrum sensing in a cognitive radio scenario where not only noise, but also interference is taken into account. The proposed detector does not require any prior information and operates in a totally blind fashion. During an initial cooperative stage, the proposed blind technique extracts

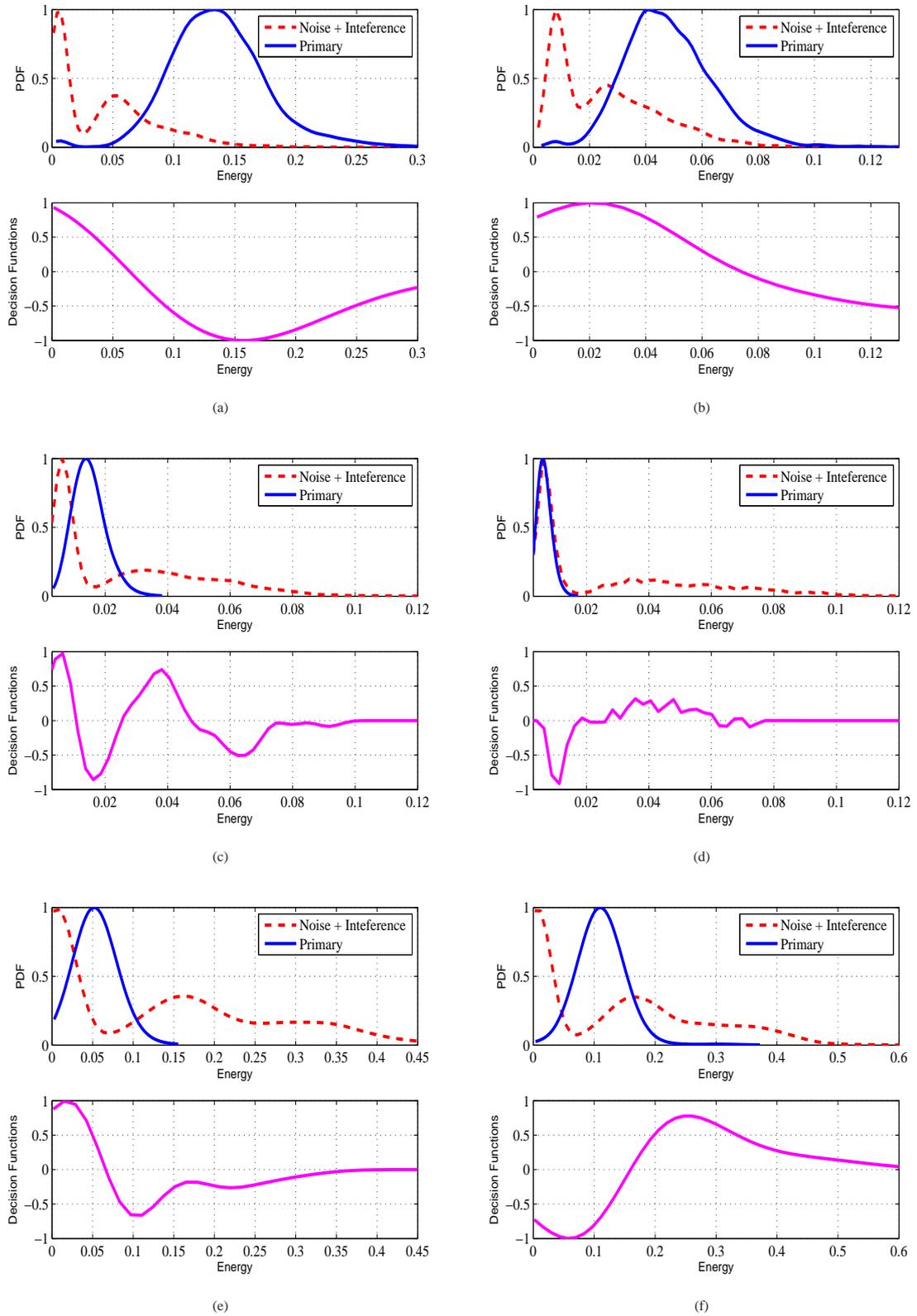
local statistical tests at the fusion center that maximize the non-linear correlation by means of a KCCA approach. These test statistics are then broadcasted to the secondary users for online operation. We have carried out a set of simulations as well as experimental measurements using a CR testbed to assess the performance of the proposed detector. Both the simulations and the experimental results show that the proposed method is robust under the presence of interference, and obtains a considerable advantage with respect to the use of an energy detector either locally or cooperatively.

## ACKNOWLEDGMENT

The research leading to these results has received funding from the Spanish Government (MICINN) under projects TEC2010-19545-C04-03 (COSIMA), TEC2013-47141-C4-3-R (RACHEL), CONSOLIDER-INGENIO 2010 CSD2008-00010 (COMONSENS), and FPI Grant BES-2011-047647.

## REFERENCES

1. Van de Beek J, Riihijarvi J, Achtzehn A, Mahonen P. TV white space in Europe. *IEEE Transactions on Mobile Computing* Feb 2012; **11**(2):178–188.
2. Elshafie H, Faisal N, Abbas M, Hassan WA, Mohamad H, Ramli N, Jayavalan S, Zubair S. A survey of cognitive radio and TV white spaces in Malaysia. *Transactions on Emerging Telecommunications Technologies*. 2014; (doi: 10.1002/ett.2778).
3. Akyildiz IF, F Lo B, Balakrishnan R. Cooperative spectrum sensing in cognitive radio networks: A survey. *Phys. Commun.* 2011; **4**:40–62.
4. Parsa A, Gohari A, Sahai A. Exploiting interference diversity for event-based spectrum sensing. *3rd IEEE*



**Figure 7.** Three considered cases (rows): KCCA decision function and probability density function for the primary, the interfering and noise signal at SU 1 (left) and SU 2 (right). (a) and (b) with an approx. SNR 0.63 dB, (c) and (d) with an approx. SINR -11.4 dB and -9.2 dB respectively, and finally (e) and (f) with an approx. SINR -6.3 dB and -5.1 dB.

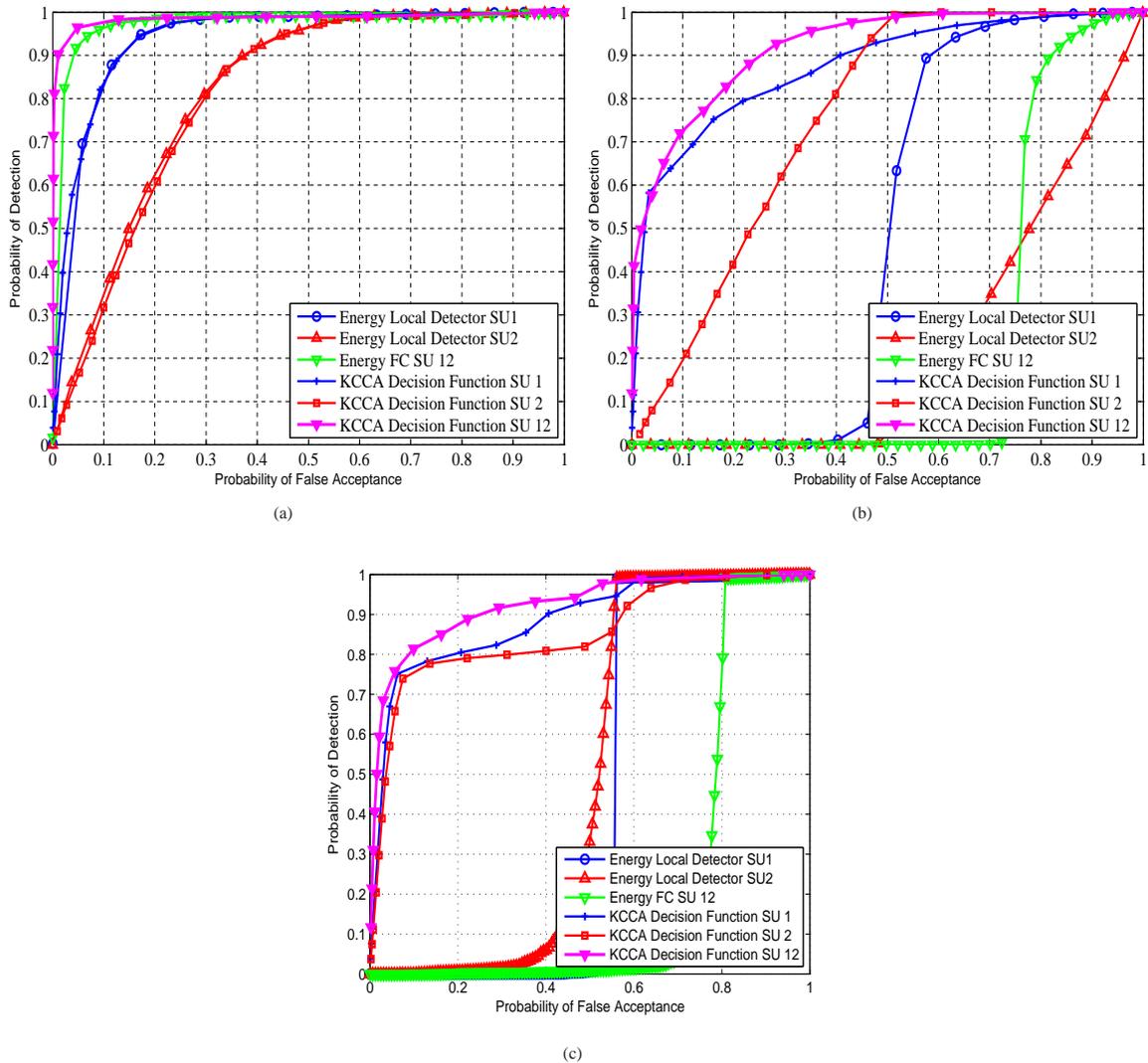


Figure 8. ROC Curves for the KCCA and energy detector and for three considered cases.

*Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN, 2008; 1–12.*

5. Biglieri E. Effect of uncertainties in modeling interferences in coherent and energy detectors for spectrum sensing. *Proc. of IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2011; 2418–2421.
6. Makarfi A, Hamdi K. Interference analysis of energy detection for spectrum sensing. *IEEE Transactions on Vehicular Technology* 2013; **62**(6):2570–2578.
7. Guimarães DA, de Souza RAA, Barreto AN. Performance of cooperative eigenvalue spectrum sensing with a realistic receiver model under impulsive noise. *Journal of Sensor and Actuator Networks* 2013; **2**(1):46–69.
8. Nieminen J, Jantti R, Qian L. Primary user detection in distributed cognitive radio networks under timing

*inaccuracy. Proc. of IEEE Symposium on New Frontiers in Dynamic Spectrum, 2010; 1–8.*

9. Song SH, Hamdi K, Letaief K. Spectrum sensing with active cognitive systems. *IEEE Transactions on Wireless Communications* 2010; **9**(6):1849–1854.
10. Font-Bach O, Bartzoudis N, Pascual-Iserte A, Payaro M, Blanco L, López D, Molina M. Interference management in LTE-based HetNets: a practical approach. *Transactions on Emerging Telecommunications Technologies* 2014; (doi: 10.1002/ett.2833).
11. Zhao B, Sasaki S. Active spectrum sensing for cognitive radio networks. *Transactions on Emerging Telecommunications Technologies* 2013; (doi: 10.1002/ett.2731).
12. Ho-Van K. Exact outage analysis of underlay cooperative cognitive networks with maximum

- transmit-and-interference power constraints and erroneous channel information. *Transactions on Emerging Telecommunications Technologies* 2013; (doi: 10.1002/ett.2719).
13. Ding G, Wu Q, Zou Y, Wang J, Gao Z. Joint spectrum sensing and transmit power adaptation in interference-aware cognitive radio networks. *Transactions on Emerging Telecommunications Technologies* 2014; **25**(2):231–238.
  14. Thilina K, Choi KW, Saquib N, Hossain E. Machine learning techniques for cooperative spectrum sensing in cognitive radio networks. *IEEE Journal on Selected Areas in Communications* 2013; **31**(11):2209–2221.
  15. Tan X, Zhang H, Chen Q, Hu J. Opportunistic channel selection based on time series prediction in cognitive radio networks. *Transactions on Emerging Telecommunications Technologies* 2013; (doi: 10.1002/ett.2664).
  16. Ma J, Zhao G, Li Y. Soft combination and detection for cooperative spectrum sensing in cognitive radio networks. *IEEE Transactions on Wireless Communications* 2008; **7**(11):4502–4507.
  17. Yan Y, Li A, Kayama H. Study on soft decision based cooperative sensing in cognitive radio network. *IEEE 70th Vehicular Technology Conference Fall, 2009*; 1–5.
  18. Harchaoui Z, Bach F, Cappe O, Moulines E. Kernel-based methods for hypothesis testing: A unified view. *IEEE Signal Processing Magazine* July 2013; **30**(4):87–97.
  19. Ding G, Wu Q, Yao YD, Wang J, Chen Y. Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions. *Signal Processing Magazine, IEEE* July 2013; **30**(4):126–136.
  20. Van Vaerenbergh S, Vía J, Manco-Vásquez J, Santamaría I. Adaptive kernel canonical correlation analysis algorithms for maximum and minimum variance. *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, 2013*.
  21. Ettus research LLC. Universal Software Radio. Jun 2012. URL <http://www.ettus.com/>.
  22. Hardoon DR, Szedmak SR, Shawe-taylor JR. Canonical correlation analysis: An overview with application to learning methods. *Neural Comput.* 2004; **16**(12):2639–2664.
  23. Bishop CM. *Pattern recognition and machine learning*, vol. 1. springer New York, 2006.
  24. Kettenring J. Canonical analysis of several sets of variables. *Biometrika* December 1971; **58**(3):433–451.
  25. Vía J, Santamaría I, Pérez J. A learning algorithm for adaptive canonical correlation analysis of several data sets. *Neural Networks* January 2007; **20**(1):139–152.
  26. Bach FR, Jordan MI. Kernel independent component analysis. *J. Mach. Learn. Res.* Mar 2003; **3**:1–48.
  27. Van Vaerenbergh S, Vía J, Santamaría I. Blind identification of SIMO Wiener systems based on kernel canonical correlation analysis. *IEEE Transactions on Signal Processing* May 2013; **61**(9):2219–2230.
  28. Van Vaerenbergh S. Kernel methods for nonlinear identification, equalization and separation of signals. PhD Thesis, University of Cantabria 2010.
  29. Cordeiro C, Ghosh M, Cavalcanti D, Challapali K. Spectrum sensing for dynamic spectrum access of TV bands. *2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom), 2007*.
  30. Silverman BW. *Density estimation for Statistics and Data Analysis*. Chapman & Hall/CRC: London, UK, 1986.
  31. Duda RO, Hart PE, Stork DG. *Pattern Classification*. John Wiley, 2000.
  32. Shalvi O, Weinstein E. New criteria for blind deconvolution of nonminimum phase systems (channels). *IEEE Transactions on Information Theory*. Mar 1990; **36**(2):312–321.
  33. Manco-Vasquez J, Gutierrez J, Perez-Arriaga J, Ibañez J, Santamaría I. Experimental evaluation of multiantenna spectrum sensing detectors using a cognitive radio testbed. *Proc. of IEEE Int. Symposium on Signals, Systems, and Electronics (ISSSE), 2012*.
  34. Gutiérrez J, González Ó, Pérez J, Ramírez D, Vielva L, Ibañez J, Santamaría I. Frequency-domain methodology for measuring MIMO channels using a generic test bed. *IEEE Transactions on Instrumentation and Measurement*. March 2011; **60**(3):827–838.