UNIVERSITY OF CANTABRIA

DEPARTMENT OF
COMMUNICATIONS ENGINEERING

PhD DISSERTATION

EXPERIMENTAL EVALUATION OF NEW DETECTION TECHNIQUES FOR COGNITIVE RADIO

AUTHOR : Julio César Manco Vásquez

SUPERVISORS : Ignacio Santamaría Caballero
Javier Vía Rodríguez

2015

Abstract

The rapid growing demand on the deployment of new wireless services has contributed to a poor spectrum utilization due to fixed spectrum assignment that wastes valuable frequency resources. This problem has been addressed by the Federal Communications Commission (FCC) that allows the use of unlicensed devices in licensed bands. Toward this end, an envisioned Cognitive Radio (CR) technology is expected to improve the efficient use of the wireless spectrum with the detection of spectral holes not utilized by primary users or legacy users, i.e. by filling them with transmissions from secondary users or non-legacy users.

This modern technology faces major challenges in the transition to smarter radio devices that are supposed to operate in any available spectrum band, at any time, in any location and without requiring any prior knowledge about the operating Radio Frequency (RF) environment. For this reason, it is expected that the CR devices will incorporate learning abilities, so that they can operate efficiently in any available part of the spectrum through the sensing and adaption to its surrounding environment.

One of the key components with this technology is the spectrum sensing where numerous approaches have been proposed. In spite of the vast amount of works in the literature, the introduction of learning mechanisms in the development towards more intelligent radios have been scarcely investigated. In this thesis, we focus on the evaluation of novel detection schemes that incorporate learning techniques to accomplish the requirements imposed by this emerging technology.

These new approaches require to be tested under more realistic environments where some departures from reported theoretical results can be revealed. Hence, the experimental evaluations allow us to take into account several aspects of the implementation often overlooked when a proposal is formulated. In this dissertation, we also address the experimental assessment of the proposed detection schemes, and with this objective in mind, the development of a cognitive radio testbed is another major contribution of this thesis. For each detection scheme, different software-based PHY-layer implementations are tested in this platform. In this way, the numerical results obtained in simulation-like environments are corroborated with those obtained in more realistic scenarios.

Concretely, we propose a Bayesian framework and a kernel canonical correlation analysis (KCCA) scheme for the detection problem in CR networks. In the

first scheme, Bayesian inference is applied at each sensing period and the posterior probabilities for the channel occupancy are utilized for detection. These posteriors summarize what we have learned so far, and are employed as prior for the next sensing period after a suitable approximation. In this manner, it constitutes the basis of the proposed Bayesian learning mechanism.

In the second scheme, a robust detection for a cooperative spectrum sensing scenario exposed to external interferences is formulated within a KCCA framework. Kernel methods are employed at the fusion center (FC) to exploit the non-linear correlations among the received signals at each secondary user (SU) during a training stage, after which statistical tests are performed. With the proposed approach, the SUs are able to make decisions either autonomously (independently at each SU) or cooperatively (at the FC). This learning ability makes this approach suitable to face the aforementioned challenges in a CR scenario.

Resumen

El rápido crecimiento de la demanda en el despliege de nuevos servicios inalámbricos ha contribuido a una pobre utilización del espectro debido a una asignación de espectro fija que desecha valiosos recursos frequenciales. Este problema ha sido abordado por organismos reguladores como la Comision Federal de las Comunicaciones (FCC) los cuales han permitido el uso de dispositivos no licenciados en bandas licenciadas. De esta manera, se espera que la tecnología de Radio Cognitiva mejore el uso eficiente del espectro inalámbrico con la detección de bandas frecuenciales no utilizadas por usuarios primarios o usuarios con licencia, esto es ocupando dichas bandas con transmisiones de usuarios secundarios o usuarios sin licencia.

Esta moderna tecnología afronta grandes retos en la transición hacia dispositivos de radio inteligente que puedan operar en cualquier banda espectral disponible, en cualquier momento, en cualquier localización y sin requerir ningún conocimiento a priori acerca del entorno de radio frecuencia. Por esta razón, se espera que los dispositivos de radio cognitiva incorporen habilidades de aprendizaje, de modo que puedan operar eficientemente en cualquier parte disponible del espectro a traves del sensado y adaptación a su entorno.

Uno de los componentes claves con esta tecnología es el sensado espectral donde numerosas soluciones han sido propuestas. A pesar de la gran cantidad de trabajos en la literatura, la introdución de mecanismos de aprendizaje en el desarrollo hacia radios más inteligentes ha sido escasamente investigada. En esta tesis, nos centramos en la evaluación de nuevos esquemas de detección que incorporan técnicas de aprendizaje para cumplir con los requerimientos impuestos por esta tecnología emergente.

Estas nuevas soluciones requieren ser probadas en un entorno más realista donde algunas diferencias respecto a resultados teóricos reportados pueden ser reveladas. Por tanto, las evaluaciones experimentales nos permiten tener en cuenta varios aspectos de la implementación, que por lo general no se consideran cuando se formula una propuesta teórica. En esta tesis, abordamos la evaluación experimental de los esquemas de detección propuestos, y con este objetivo en mente, se desarrolla e implementa una plataforma de radio cognitiva. Por cada esquema de detección, diferentes implementaciones en la capa física basadas en software son llevadas a cabo en la plataforma, donde corroboramos los resultados numéricos obtenidos en los entornos de simulacion con aquellos obtenidos en escenarios más realistas.

Concretamente, como primer contribución de esta tesis proponemos un esquema Bayesiano y un esquema KCCA (Kernel Canonical Correlation Analysis) para el problema de detección en redes de Radio Cognitiva. En el primero de ellos, un mecanismo de inferencia Bayesiana es aplicada en cada periodo de sensado y las probabilidades a posteriori de la ocupación del canal son utilizados como estadísticos para la detección. Estas probabilidades a posteriori resumen lo que hemos aprendido y son empleados como probabilidades a priori en el siguiente periodo de sensado tras aplicar una adecuada approximación. Esta idea constituye la base del mecanismo de aprendizaje Bayesiano propuesto.

Como segunda contribución de la tesis, una detección robusta para un escenario de sensado espectral cooperativo expuesto a interferencias externas es formulado en un esquema KCCA. Los métodos kernel son empleados en el centro de fusión para explotar las correlaciones no-lineales entre las señales recibidas en cada SU durante una etapa de entrenamiento, tras el cual los tests estadísticos son aplicados. Mediante estos tests los SU pueden tomar decisiones de manera autónoma (en cada SU) o cooperativa (en el centro de fusión). Esta capacidad de aprendizaje hace que esta solución sea adecuada para afrontar los retos mencionados en un escenario de radio cognitiva.

Contents

Chapter 1
Introduction

1.1 Cognitive radio

Nowadays, the emerging technologies and the growing demand of wireless services require a more efficient usage of the spectrum because of the spectrum scarcity [Federal Communications Commission - FCC, 2002]. In this context, Cognitive Radio (CR) has been foreseen as a promising approach to improve the efficiency of the wireless spectrum [Mitola and Maguire, 1999, Haykin, 2005, Akyildiz et al., 2006], where legacy or primary users (PU) and non-legacy or secondary users (SU) coexist by sharing the same spectrum. This is carried out by detecting holes in the spectrum and filling them with transmissions from SUs. In this way, the improvement of the usage of frequency bands is attained while avoiding interference to primary users.

For that end, CR is defined as an intelligent radio that can sense and adapt to its environment, where this environment may vary over the time. It is expected that the radios are able to learn from and adapt to this unknown and time-varying environment in real time. In this respect, learning engines that can remember lessons learned in the past and act quickly in the future have been proposed to be employed with this technology [Haykin, 2005]. This learning mechanism allows, for example, to carry out changes in some radio parameters, such as the transmit power, the carrier frequency, and the modulation strategy among others. In fact, this learning ability is what make them truly intelligent. In doing so, three main components can be identified, i.e. perception, learning and reasoning. The first of them is achieved through the sensing measurements of the spectrum, which allows to identify ongoing RF activities in its surrounding environment. In a learning stage, a radio employs learning algorithms, e.g. machine learning techniques, in order to, for instance, classify and organize the observations into suitable categories. Finally, the reasoning ability allows the radio to use the knowledge acquired through learning to achieve its objectives where only the current state of the environment is observed and decisions are made while ignoring the past information. In other words, a CR device is supposed to provide opportunistic access to the unoccupied frequency bands, and with this aim learning algorithms are employed to operate efficiently and optimally in any RF environment.

Spectrum sensing is a key element of this technology that allows the SUs to detect the presence of PUs. In fact, several approaches have addressed the detection problem [Yucek and Arslan, 2009, Cabric et al., 2006, Lu et al., 2012, Tandra and Sahai, 2008, Ramirez et al., 2011, Wang et al., 2010]. Notwithstanding, the ability to learn from past decisions and the capacity to adapt to a continuously changing environment have not been so well explored yet. This motivates the main theoretical contributions of this thesis, where the introduction of learning techniques in the schemes of detections are foreseen to obtain significant advantages.

With the new requirements imposed by this CR technology, the feasibility and validation of the novel approaches also require to be examined with experimental results that corroborate the simulation results. With this purpose, software defined radio (SDR) systems have been proposed for the evaluation of innovative PHY-layer algorithms which are implemented with the existing hardware technology available in the market. The experimental evaluations allow to establish the feasibility of a proposal, and hence they contribute to fill the gap between the theory and the real operation. In addition, realistic wireless channels and hardware impairments as well as its limitations are taken into account in the performance of new proposals [Cabric and Brodersen, 2005, Mubaraq et al., 2004]. This experimental assessment constitutes another major contribution of this thesis where the development of a CR testbed is carried out to validate the proposed schemes.

1.1.1 Thesis scope

In this thesis, we address the detection problem in CR networks, and its experimental evaluation. In this way, we propose and validate the feasibility of novel spectrum sensing schemes. These PHY-layer algorithms employ learning techniques that turn out to be useful and provide significant advantages. On the other hand, the feasibility of the proposed approaches is addressed with the development of a CR testbed where software-based PHY-layer implementations are carried out for each proposal under study. The contributions of this thesis can be divided in two main subjects where each of them entails simulations and experimental results. Here we briefly summarize each of the addressed subjects.

- Firstly, we propose an adaptive spectrum sensing scheme for multiantenna receivers. In this scheme, the ability of learning from previous decisions and, at the same time, of adapting to a changing environment is incorporated within a Bayesian framework. The behavior of a CR network is formulated probabilistically, and each time that new observations are obtained, a Bayesian inference is applied and posterior probabilities are computed. These posterior probabilities that summarize what has been learnt are employed as priors in a next Bayesian inference to provide a learning mechanism.

Our scheme is intended to operate not only in stationary environments, but also in non-stationary environments. For that end, a forgetting mechanism is devised to exploit adequately previous statistical information. The simulation results shows that our Bayesian detector is able to obtain significant gains in comparison to generalized likelihood ratio test (GLRT) detectors that lacks of a learning mechanism.

In addition, this scheme is evaluated by means of experimental measurements carried out in the developed CR testbed. The testbed is implemented using multiantenna nodes prototyped with Universal Software Radio Peripheral (USRP) devices. Finally, our experimental results corroborate the obtained gains shown in simulations.

- Secondly, we propose a robust cooperative spectrum sensing (CSS) scheme for an scenario exposed to external interferences. Our detection scheme aims to exploit the correlations among the measured signals at each SU. With this aim, a kernel canonical correlation analysis (KCCA) scheme is devised [Manco-Vásquez et al., 2014b, Van Vaerenbergh et al., 2013, Manco-Vásquez et al., 2014c]. This scheme comprises two stages: a first one during which the measured signals of each SU are collected at the fusion center (FC), and where a kernel-based learning method is employed to maximize the possible non-linear correlations among the collected measurements. This method allows us to obtain non-linear transformations that provide the maximal correlation. Then, these transformations are employed as statistical tests in a second stage.

  With the obtained tests during the first stage, this scheme can operate in two modes i.e. making decisions autonomously (autonomous testing) at each SU or making decisions at the FC (cooperative testing). In addition, the proposed scheme exhibits many features some of which we list here and which make it suitable for CR applications:

  - The scheme operates in a blind manner.
  - Prior information is not required.
  - Its capacity to retrain from time to time (during an initial stage) allows the scheme to adapt itself to the surrounding environment.
  - The detection performance can be boosted by including additional features (if available) in the reported measurements of each SU such as the kurtosis, the cyclic statistics, among others.

  The simulation results show that our scheme is able to obtain significant gains over a conventional energy detector, in scenarios exposed to interference signals. In fact, better performance is attained as the interference level increases.

In a similar way to the Bayesian scheme, the KCCA-based detector is evaluated experimentally using single-antenna USRP devices, where significant gains are also exhibited. In fact, more challenging cases not taken into account in our simulation environment are tested, where, for instance, different levels of noise and interference signal are considered in the detection of the primary signal.

1.2 Outline

The structure of this thesis is based on published papers that were written during the doctoral period. In the following chapter, we summarize the overall contribution of this dissertation. We provide a comprehensive overview of the proposed schemes in Chapter 2, which summarizes the context within which the formulation of proposals takes place and its validation with software-based implementations.

The rest of the chapters are self-contained publications where each journal or conference paper represents a chapter. Each of these chapters encompasses, the state of art, the proposal, as well as the final conclusions. We organize the structure of the thesis in the following way:

- The Chapter 2 introduces the detection problem as well as the development of a CR testbed. Within this scope, two detection schemes that incorporate learning techniques are presented in this chapter. It starts providing a general framework for CR networks tailored for spectrum sensing, after which the development of a CR testbed for the experimental evaluation of the proposals is exposed, and finally the evaluations of the proposed detection schemes are presented at the end of the chapter. Specifically, we describe the evaluation of a Bayesian framework and a kernel canonical correlation analysis (KCCA) scheme for the detection problem.

- The Chapters 3, 4, and 5 comprise a set of attached papers, i.e. the representative papers related to the Bayesian detection scheme [Manco-Vásquez et al., 2014a] and [Manco-Vasquez et al., 2012a] are annexed in Chapters 3 and 4 respectively, and a journal paper that encompasses the proposal of a robust CSS scheme and its experimental evaluation [Manco-Vásquez et al., 2014c] is annexed in Chapter 5. These papers correspond to the following publications:

  - J. Manco-Vásquez, Miguel Lázaro, David Ramírez, J. Vía, I. Santamaría, "A Bayesian Approach for Adaptive Multiantenna Sensing in Cognitive Radio Networks" Signal Processing Elsevier, Volume 96, 2014.

  - J. Manco Vásquez, J. Gutiérrez, J. Pérez, J. Ibáñez, and I. Santamaría, "Experimental Evaluation of Multiantenna Spectrum Sensing Detectors using a Cognitive Radio Testbed". IEEE International Symposium on Signals, Systems

and Electronics (ISSSE). Potsdam, Germany October 2012.

- J. Manco-Vásquez, S. Van Vaerenbergh, J. Vía, I. Santamaría, "Kernel Canonical Correlation Analysis for Robust Cooperative Spectrum Sensing in Cognitive Radio Networks". Transactions on Emerging Telecommunications Technologies, Oct 2014.

- Finally, some concluding remarks are summarized in Chapter 6, along with the suggestion of future research directions.

Chapter 2

Spectrum sensing for cognitive radio networks

In this chapter, we summarize the contributions of this thesis. First, a set of preliminaries related to the detection problem in CR networks are provided in Section 2.1. Then, we describe the development of experimental assessments tailored for communication systems in CR networks at the physical layer. For that end, we introduce some notions in the implementation and prototyping of a communication system in Section 2.2, and provide a detailed description of the employed platform in Section 2.3. Finally, we summarize the results obtained in the evaluation of the new detection schemes and point out the main contributions for each of the proposed scheme in Section 2.4.

2.1 Spectrum sensing

In cognitive radio networks, it is understood that a CR device should incorporate the ability of sensing, learning, and being aware of the parameters related to the radio channel characteristics. In this context, spectrum sensing plays a key role in the detection of spectrum opportunities which entails the learning and exploitation of the spectrum holes. In spite of the vast amount of related works in the literature, various aspects in the signal detection still remain unexplored. For instance, the detection of a PU transmitting spread spectrum signals turns out to be a more difficult task since the power of the PU is distributed over a wide frequency range. In a similar way, with the advent of multi-antenna technologies, several users can share the same frequency channel at the same time, while being placed within the same geographical area. In this context, it is not straightforward to design algorithms that can exploit these additional spatial dimensions of the radio space.

Many challenges need to be addressed. The period of sensing is one of them, since it should be short to transmit payload data but at the same time long enough to provide the required reliability. The failure in the detection of a primary transmitter due to its location may provoke interference to some PU receivers (also known as the PU hidden problem) which require to consider new strategies of detection.

Cooperative schemes have also been devised in the literature as solutions to the noise uncertainty, fading, and shadowing. In these schemes the sensing information is shared among the SUs to make a decision, and it can be implemented in two modes, i.e. a centralized and a distributed manner. The way in which the sensing information is employed to maximize the detection probability while keeping the false alarm bounded still requires to be further explored for the new CR scenarios.

The aforementioned techniques also require to be evaluated under more realistic conditions. In this regard, the development of testbeds to study the feasibility of new detection schemes is being currently pursued in the research community. Nevertheless, it also imposes additional requirements related to the hardware devices. High sampling rate, high resolution of analog to digital converters (ADCs) as well as high speed signal processors (DSPs or FPGAs) constitute some of these new hardware requirements. Furthermore, the radio frequency (RF) components (e.g. antennas and power amplifiers) in the CR terminals should allow to operate over a much wider range of frequency as these devices should identify the spectrum opportunities by capturing and analyzing a large band of frequency.

2.1.1 Preliminaries

In spectrum sensing, the detection problem consists in deciding between two or more possible hypotheses based on an observed data set. Specific detection problems can be formulated depending on the available information. For instance, if the waveform is known under the alternative hypothesis, a decision is made taking into account this knowledge, whereas when such a waveform is unknown, we resort to other properties of the signal to distinguish among the hypotheses.

The energy detector is probably the most widely employed detector in communication systems. It turns out to be the optimal detector for a model where under an hypothesis there only exists Gaussian noise, whereas under the other hypothesis the observations are composed of an unknown Gaussian signal plus additive noise. It does not require any prior information, and basically consists in comparing the energy of the observations with respect to a threshold [Kay, 1998], to determine the presence of a signal. Nevertheless, if some information is available, it could be employed to improve the performance of the detector. For instance, the cyclostationarity property commonly encountered in the signals of a communication system can be exploited to distinguish between an stationary noise and the cyclostacionarity of the signal [Dandawate and B. Giannakis, 1994].

Nevertheless, the uncertainty of the noise, the fading and shadowing phenomena degrade the performance of these approaches. An energy detector requires to estimate the noise variance to establish the detection threshold, where any estimation error degrades severely its performance. On the other hand, a cyclostacionarity-based detector requires some level of synchronization which is difficult due to the fading

and shadowing phenomena that provoke very low SNR conditions [Cabric, 2008]. To alleviate these problems, asynchronous detectors robust to noise uncertainty have been proposed in the literature. For instance, the use of multiple-antenna sensors allows us to exploit the spatial and temporal structure of the observations, without requiring any prior knowledge of the transmitted waveform, to enhance the detection performance [Ramirez et al., 2011]. On the other hand, the fading and shadowing of the channel have been addressed by cooperative spectrum sensing (CSS) schemes that exploit the spatial diversity to mitigate these effects.

In spite of the existing approaches in the literature, the unknown time-changing environment in a CR scenario requires to incorporate other aspects such as the learning and adaptation in the detection of a primary user. These scenarios can not be tailored to a particular signal model, in fact the lack of available information is often assumed in CR. In the next subsection, we describe some main concepts in the detection problem that will be employed as reference to study and evaluate the proposed approaches.

Statistical tests

An optimal detector for a detection problem with simple hypotheses is given the Neyman-Pearson (NP) test. This optimal detector maximizes the probability of detection for a given false alarm probability, where the probability of detection is defined as $P_D = p(\hat{\mathcal{H}}_1/\mathcal{H}_1)$ and the probability of false alarm is given by $P_{FA} = p(\hat{\mathcal{H}}_1/\mathcal{H}_0)$, with $p(\hat{\mathcal{H}}_i/\mathcal{H}_j)$ as the probability of deciding the i-th hypothesis when the actual hypothesis is the j-th hypothesis. This detector is given by a ratio of likelihoods,

$$T_{NP}(\mathbf{x}) = \frac{p(\mathbf{x};\theta_1)}{p(\mathbf{x};\theta_0)}$$

where $\mathbf{x}$ is the set of available observations, $p(\mathbf{x};\theta_i)$ is the probability under $\mathcal{H}_i$ that depends on the set of known parameters θ_i, for $i = \{0,1\}$ that is for the null and alternative hypothesis, $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.

This statistical test, when possible, can be considered as reference to assess the performance of new detectors. However, these likelihoods depend on a set of parameters which are often unknown θ_i. When the parameters are unknown the hypotheses are named composite, and typically these detection problems are addressed by two approaches: a Bayesian approach where the unknown parameters are considered realizations of random variables which are assigned a prior distribution, and a GLRT approach where the unknown parameters are replaced by their maximum likelihood estimates.

Under a Bayesian philosophy, prior PDFs $p(\theta_o)$ and $p(\theta_1)$ are assigned to the unknown parameters θ_o and θ_1, and the optimal NP detector is given by,

$$T_{NP}(\mathbf{x}) = \frac{\int p(\mathbf{x}|\theta_1;\mathcal{H}_1)p(\theta_1)d\theta_1}{\int p(\mathbf{x}|\theta_0;\mathcal{H}_0)p(\theta_0)d\theta_0}$$

where $p(\mathbf{x}|\theta_i;\mathcal{H}_i)$ is the conditional PDF of $\mathbf{x}$ conditioned on θ_i, assuming that $\mathcal{H}_i$ is true. A prior PDFs should reflect our prior knowledge about the unknown parameters, and their choice can prove difficult. Indeed, if one does have some prior knowledge, then one should use it. Otherwise, a non-informative prior should be employed. On the other hand, under a GLRT approach less restrictive assumptions as well as an ease of implementation are allowed. This statistical test is given by,

$$T_{GLRT}(\mathbf{x}) = \frac{p(\mathbf{x};\hat{\theta}_1)}{p(\mathbf{x};\hat{\theta}_0)}$$

where $\hat{\theta}_1$ and $\hat{\theta}_0$ are the maximum likelihood estimates corresponding to the θ_1 and θ_0 parameters respectively. Finally, the outcome of a detector is obtained by comparing the ratio of probabilities with a threshold,

$$Selected\ \ Hypothesis = \begin{cases} \mathcal{H}_1, & T_{NP}(\mathbf{x}) \geq \eta \\ \mathcal{H}_0, & T_{NP}(\mathbf{x}) < \eta \end{cases}$$

We have provided a set of preliminaries to the detection problem as well as the development of optimal detectors for simple, and composite hypotheses where Bayesian and GLRT approaches have been studied. Nevertheless, aspects of learning have not been well explored yet in the literature, which however play a key role in the design of new detectors tailored for CR networks. More specifically a Bayesian framework and kernel-based learning methods are proposed in this thesis to provide significant advantages in the detection performance of cognitive radio networks.

Learning techniques

Learning is about the construction and study of systems that can learn from data, where learning methods are employed to do better a task in the future based on what was experienced in the past. There exists several forms of learning, such as supervised learning, unsupervised learning, and reinforcement learning. In a supervised learning, the algorithms are trained on labeled examples, i.e. a set of inputs where the desired output is known. On the other hand, unsupervised learning algorithms operate on unlabeled examples, so that it tries to discover some kind of structure in the data. Finally, reinforcement learning addresses the maximization of some reward. With this technique, intelligent agents execute actions to acquire knowledge about the unknown environment and select those actions that maximize a cumulative reward.

A CR device is supposed to operate in any available spectrum band and it may not have any prior knowledge on the operating RF environment such as the noise distribution, the interference levels, or the primary user traffic. Instead, it should be able to apply autonomous learning algorithms that reveal the underlying nature of the environment and its components. In this way, a CR device can incorporate learning methods to be aware of its surrounding environment and adapt its internal

states to statistical variations of the environment.

If some prior information about the environment is available, this knowledge should be exploited to enhance its performance. For instance, if the waveform characteristics are known by the CR before operating, some training techniques could be employed to improve the detection of the primary signal (e.g. supervised learning techniques). Nevertheless, in most CR applications, unsupervised learning techniques turn out to be a perfect candidate for the learning problem. These techniques permit exploring the environment characteristics and applying self-adapting actions without having any prior knowledge [Thilina et al., 2013].

2.2 Experimental assessment of wireless communication systems

The evolution of wireless communications systems involve the development of innovative ideas at different levels, from its theoretical conceptualization and modeling, to its design and implementation in order to validate the innovative proposal. Typically research studies are based on computer simulations which provides a timely evaluation of novel solutions. For this reason, it constitutes an essential step at an early stage of applied research. However, most innovative signal processing solutions that impulse the development of wireless communication systems are conceived as mathematical models that overlook aspects of its implementation in a real word. This lead us to some mismatches between the expected theoretical performance and its experimental assessment. Hence, this evident gap between fundamental and applied research motivates the research and development of experimental evaluations impulsed, at the same time, by research institutions and industrial entities.

To place the scope of this gap in context, researchers typically employ programming languages at higher level to describe complex mathematical models while saving costs in the required expertise to simulate innovative proposals. MATLAB is probably the most employed language for signal processing modeling, since it provides powerful mathematical tools to evaluate complex communication systems, that at the same time, provide scalable solutions where additional modifications are easily included without implying an additional development cost. Nevertheless, in the process of modeling communication systems several assumptions and simplifications about its implementation are often done, and the omission of these aspects lead us to misleading conclusions in the performance of new innovative proposals.

In fact, most part of the digital implementation of a communication system is abstracted away by using libraries based on higher-level programming languages. For instance, to ease the labor of the implementation of PHY-layer applications, most of the Digital Signal Processors (DSP) boards provides DSP libraries that are utilized

without taking into account the underlying processes of the implemented routines.

For that end, it is important to evaluate the proposed schemes under more realistic conditions in an experimental testbed where those omitted aspects of the PHY-layer are taken into account. In doing so, several assumptions and limitations should be considered when assessing its performance, some of which we list here:

- The obtained performance in simulation environments can be affected by the specifications and limitations in some components of a radio transceiver. For instance, the allowed dynamic range of the ADC/DAC, the clock sample rate of ADC/DAC may provoke a mismatch between the theoretical and experimental performance.
- The hardware impairments degrade severely the performance of an approach, and often require to be considered to avoid misleading results. Most of them comes from the digital front-end (DFE), such as the inaccuracies in the local oscillator (LO), the direct current (DC) levels, the imbalance in the in-phase and quadrature (I/Q), filters with some transient states, non-linearities at the RF part, etc.
- Some aspects of the PHY-layer require to be characterized taking into account experimental measurements (e.g. the phase noise, the correlation among the antennas) since some departures from adopted ideal signal models can be found.
- An experimental evaluation provides realistic channels, which can present practical characteristics not considered in the simulation analysis. For instance non expected statistical distributions of the channels, fast fading effects, etc.

Furthermore, the apparition of new emerging technologies motivates not only the development of theoretical approaches, but also its experimental evaluation, where the new requirements imposed by these technologies are still under study. This motivates even more the development of experimental evaluations.

2.2.1 Implementation and prototyping

The implementation of PHY-layer solutions follows designs that employ languages at low-level (i.e. hardware description languages) and/or programming languages at high-level. The hardware description languages (HDL) aim at designing the digital logic of the system such as the exchange of data between registers (e.g. VHSIC Hardware description language - VHDL). By employing HDLs, low-level architectures that allows more control to the design of its components are implemented. However, it also implies a large time in acquiring expertise to implement at low-level.

On the other hand, low-level details are omitted and/or hidden to the researcher

when employing programming languages at high-level (e.g. C/C++). A low development cost is attained with this programming language, since it does not demand a lengthy acquisition of expertise in the utilization of additional tools to implement low-level features. Furthermore, its flexibility allows us to extend or make modifications without introducing noticeable changes.

The time of development in implementing low-level details can be avoided by generating HDL from programming language at high-level, as it is provided in the market for some hardware devices which include specific design tools for this purpose. Nevertheless, most requirements of a proposal do not adjust to the set of low-level routines generated with these tools.

In any case, the choice of a programming language also depends on the available hardware, i.e. General Purpose Processors - GPP, Digital Signal Processor DSP, Field-Programmable Gate Arrays - FPGAs, Application-Specific Integrated Circuit - ASIC, and/or some sort of hybrid composed by these hardware devices. GPPs are tailored for general purpose computers such as PCs, and its cost and power-consumption are usually higher than DSPs. Unlike GPPs, DSP are designed for fast real-time applications. It possesses an internal architecture optimized for digital signal processing providing a more efficient, and cost effective solution. A DSP is typically used for processing continuous real-world analog signals, and has a set of instructions for its operations that are utilized within an integrated development environment (IDE) tool. Both GPPs and DSPs utilize PL at high-level to facilitate the development of complex systems. On the other hand, implementations based on HDL requires a deep knowledge of its architecture and its development framework as it is the case with FPGA and ASIC where usually VHDL or Verilog are employed respectively.

The experimental evaluation of wireless communication systems typically employs radio transceivers composed of an analog front-end, and a baseband chain. In fact, the market provides GPP-based DSP solutions where its analog front-end is interfaced through a FPGA-based board with a PC-based host, and where the communication with GPP-based DSP board is carried out by means of real-time extension of the Linux Operating Systems (OS). For this reason, the implementation and prototyping involves working with programming languages at low and high levels, and different hardware devices.

Software defined radio (SDR) systems have proved to be an ideal vehicle for the evaluation of CR techniques. In these radio communication systems, most part of the physical layer functions are implemented in software instead of hardware. Its open architecture consisting of a generic hardware platform that allows to install different software applications which are highly modular with a high degree of portability and reuse. In fact, new standards and technologies can be easily implemented due to its programmability and reconfigurability. This flexibility has motivated its use in the efficient management of the spectrum for which it turns out

to be a promising candidate for evaluating CR techniques. Furthermore, countless proposed SDR architectures are currently in circulation. Ettus research LLC is one of the companies specialized in providing low-cost, high-quality SDR systems for the evaluation and exploration of novel approaches. Its Universal Software Radio Peripheral (USRP) boards have been widely employed by the research community working in CR.

2.3 Hardware platform for cognitive radio networks

The selection of the hardware is probably one of the first and most important decisions to be made when aiming at building a testbed. The price and capabilities of the boards are intimately related each other and their prices may vary considerable in the market. Hence, the chosen board should fulfill the specific requirements of the target applications. With the acquisition of USRP boards, we aim at building of a cognitive radio platform. Each of these nodes are transceivers composed of two main parts: a baseband part and a RF component. The baseband part is basically composed of dual analog-to-digital converters (ADC) and digital-to-analog (DAC) converters connected to an FPGA where the digital up/down conversion takes place. On the other hand, the RF part is composed of elements such as low noise amplifiers (LNA), variable gain amplifiers (VGA), mixers, among others, and is in charge of analog operations such as the up/down conversion. For some types of USRPs, the RF part is provided as pluggable RF front-ends named daughterborads, whereas for other types it is already incorporated on the same board. These products typically allow to cover the Industrial, Scientific and Medical (ISM) bands, from 50MHz to 2.9GHz, and from 4.9GHz to 5.85GHz, and here we describe the specifications of the employed boards during our experimental assessments:

- N210 USRP

  Known as motherboard, it includes an FPGA-Xilinix Spartan 3A-DSP, where the processes of decimation and interpolation are carried out. The FPGA is connected to a 100 MS/s dual ADC, and a 400 MS/s dual DAC for the transmission and reception of RF signals. It can stream data to and from host PC processors up to 50 Msamples/s by employing a Universal Hardware Driver (UHD) that communicates the board with a Gigabit Ethernet interface of a host PC. In addition, these USRPs can implement MIMO nodes by connecting two N210 USRP through an expansion port, and can be synchronized using a Global Position System (GPS) signal. To clarify the mentioned architecture, we show in Fig. 2.1 the main hardware components of this board.

  The RF part of this board is implemented with a board (also known as daughterboard), whose characteristics are described next:

  – Baseband board

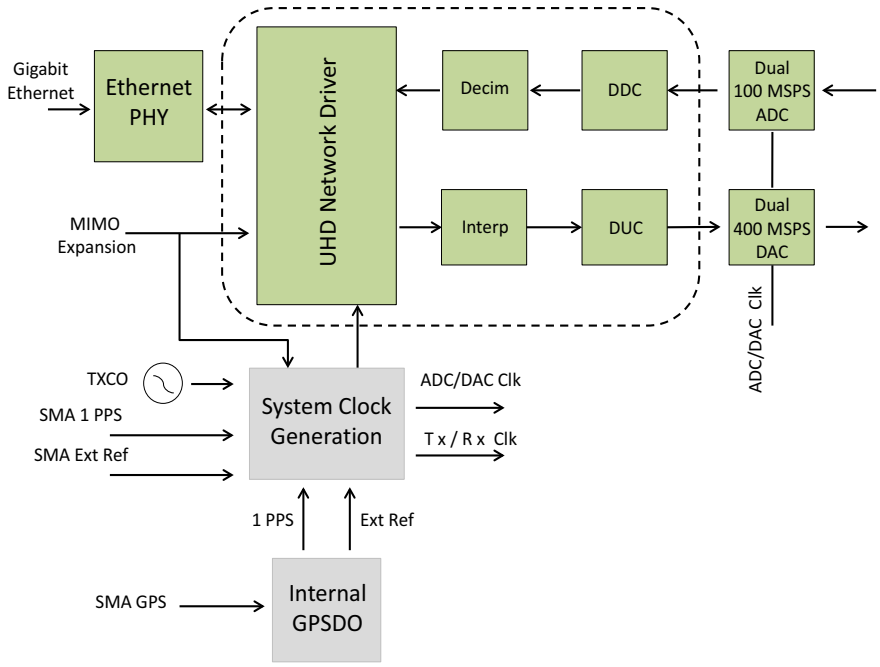
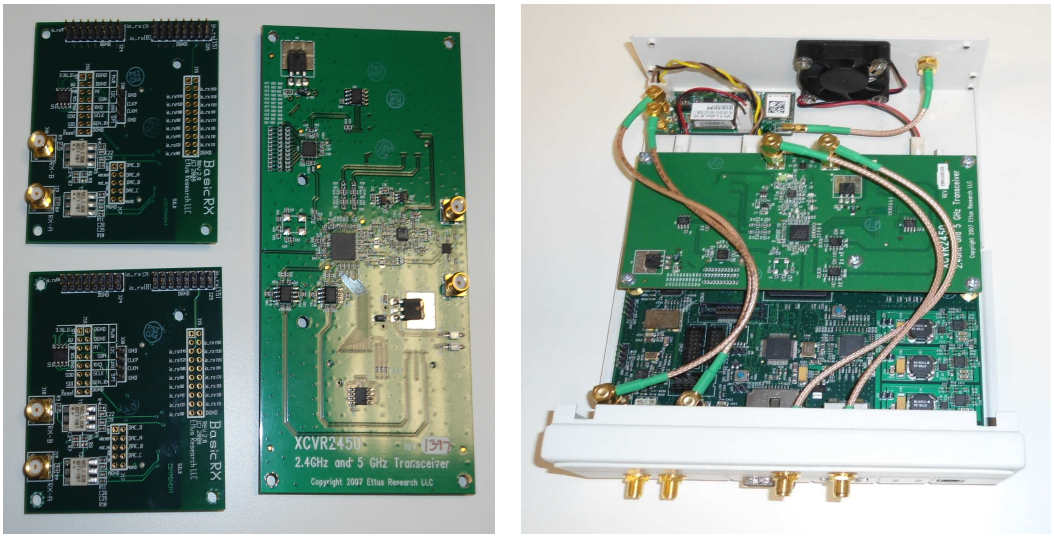


Figure 2.1: Hardware scheme of the N210 USRP

These boards are plugged into the motherboard, and provide transmission/reception capabilities in a range that varies between 1 to 250 MHz, i.e. at baseband. More specifically, these daughterboards match the dual DAC outputs (or dual ADC inputs) of a motherboard to an antenna port through SubMiniature version A (SMA) connections. These have been designed in two types, one for the transmission and another for the reception, and each of them provide two outputs for the phase (I) and the quadrature signal (Q). In our work, these boards have been employed for the debugging of some initial implementations.

– XCVR2450 board

This RF board is a dual band transceiver intended for operations in the ISM bands, and have two quadrature frontends for the transmission and reception that operate as half-dulplex transceiver. It allow us to configure some of the RF components (VGA and LNA) which makes them useful in an experimental evaluation. In Fig. 2.2, we show the described baseband and RF boards, and a transceiver composed of a N210 USRP with a XCVR250 dautherboard at the right side.



(a) (b)

Figure 2.2: (a) Baseband dautherboards (at the left side) and a XCVR2450 daughterboard (at the right side), (b) A N210 USRP with a XCVR2450 daughterboard often employed in the experimental assessments.

- B210 USRP

  With this USRP, both the baseband and RF part are placed at the same board. In a similar way, it has a FPGA-Xilinix Spartan 6 XC6SLX150 that is connected to 100 MS/s dual ADCs and 400 MS/s dual DACs for the transmission and reception of RF signals. Its RF part covers a wide range of frequencies, i.e. from 70 MHz to 6 GHz, and allows to implement fully-coherent 2x2 MIMO nodes. Its instantaneous bandwidth is up to 56 MHz and 30.72 MHz when employing a single antenna (1x1 mode) and two-antennas (2x2 mode), respectively. Moreover, it is also controlled through a UHD driver (with common functionalities to that of a N210 USRP) that connects to a host PC by means of a USB 3.0 port. In Fig. 2.3, we show the main components of the mentioned architecture.

These boards are controlled from a PC (see Fig. 2.4), and employed for transmission and reception of signals where the processing chain at the transmitter and receiver side can be summarized as follows: After an instruction of transmission is executed from the host PC, the order and the data to be transmitted are transferred to the USRP. The received complex signal is upconverted to an analog Intermediate Frequency (IF) signal and transmitted over the air by the RF part. On the other hand, when an instruction of reception is executed, this order is transferred to the USRP board to start acquiring an RF signal. This complex signal is downconverted to a baseband signal which is stored at the host PC, where the rest of signal processing tasks are performed.

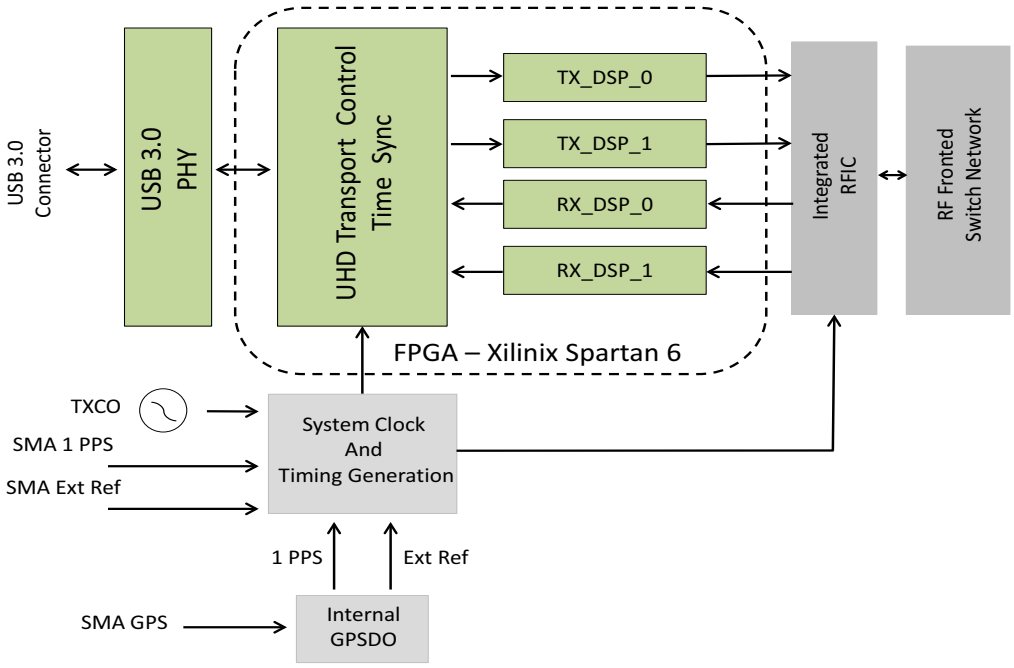


Figure 2.3: Main hardware components of a B210 USRP.

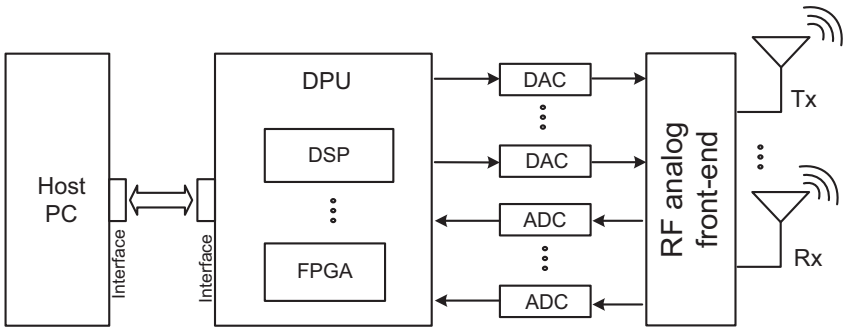


Figure 2.4: Operation of a USRP transceiver

2.3.1 Software for USRPs

For the configuration and operation of a USRP, the manufacturer Ettus provides specific firmware and FPGA images for each series of the USRP. These images are loaded onto the motherboard and along with Universal Hardware Driver (UHD) make possible the communication with a PC. The UHD is compatible with operating systems (OS) such as Windows and Linux, and provides a set of basic functions to start exploring the features of the USRP.

With these basic functions, one can develop applications on the host PC for controlling the USRP boards, where these basic functions can be encapsulate by high-level functions. In fact, third parties (e.g. Matlab and Simulink Packages)

provide a limited set of utilities through the use of application programming interface (API) for some series of USRPs. Here we describe some of the efforts in the development of software that take into account implementations with USRP boards:

- GNU

  GNU Radio [Too, 2015] is a set of tools for the development of software that aims at implementing software defined radio (SDR) systems. It comprises a free set of files and applications joined in libraries that allows to perform signal-processing operations. GNU Radio can be used with low-cost external RF hardware (e.g. USRP) or without it in a simulation-like environment.

  It is based on a layer of signal processing blocks which belong to the *gnuradio-core* library where the blocks are defined by means of C++, or Python classes [Pyt, 2013]. These classes are accessible from a higher-level environment by using a Simple Wrapper and Interface Generator (SWIG) [Swi, 2013] library which is written in Python. Modules are connected in this environment to form a flow graph which represents the transceiver in software, and these flow graphs are executed by high-level Python processes or executables developed for this purpose in C++.

  In addition, it also includes a graphical application called GNU Radio Companion (GRC) that provides an intuitive user interface for the development of systems using flow graphs. It is based on blocks and similar to Simulink in Matlab which consists of source blocks (audio files and binary signals), processing blocks (modulators, filters, multipliers and amplifiers), and sink blocks (FFT sink, constellation sink, oscilloscope sink). This graphical environment alleviates the task of the researcher to perform the experimental evaluation.

- IRIS

  One of the major advantages with Iris is its flexibility and reconfigurability. It is composed by a set of libraries written in C++ to build SDR systems [Doyle et al., 2010, Sutton et al., 2010], and each library does a specific job by providing API for the main program in Iris. These libraries or plugins can be dynamically loaded at runtime, and the main ones (also called components) are in charge of processing streams of data that runs within an engine.

- ASGARD

  Asgard is a framework for building SDR using a set of libraries written in C++ that runs in a Linux environment. This framework consists of four basic building blocks: Components, Modules, Communication Objects (CO), and Applications; and its Asgard API relies on the Poco [Poc, 2013], Boost [Boo, 2013], and Intel Threading Building Blocks [Int, 2013] libraries, that defines the basic functionalities of the framework.

- SDR4All

  SDR4ALL (Software Defined Radio for All) is SDR for education and research purposes, and it has been introduced by the LANEAS group at Supélec university. This software architecture consists in introducing numerical and reconfigurable treatments as close as possible to the radio antenna. It has been tested with real transmissions using USRP as hardware devices. However, its development has been limited to basic experimental evaluations where, for instance, the control of several USRP has not been considered. In fact, no more software releases have been reported since its first inception.

- USASDR

  Universal Software Architecture for Software Defined Radio (USASDR) is an effort to build an architecture to control several USRP from a remote PC running MATLAB. This initial architecture has been part of a project with the GTEC group at the ACoruña university, which aims to evaluate more complex schemes and provide a first prototype for more advanced platforms suitable for experimental evaluations.

Although many developments for SDR can be found available on the web [Tonelli and Buthler, 2013, Pla, 2015b, Asg, 2015, Too, 2015, Mishra et al., 2005, Pla, 2015a]. Some of them require a large expertise such as GNU Radio, and other solutions are not tailored to our required implementations. For this reason, we decided to follow the development of USASDR, which was developed during an initial stage.

2.3.2 Description of the USASDR framework

A USASDR aims to provide a general strategy for the control of several USRPs by allowing their configuration and their control from a unique remote PC. This software architecture operates under Linux OS and is composed by three components or layers, i.e. Usasdr, Controller and Node, as it is depicted in Fig. 2.5. With this architecture, the execution of an instruction from Matlab is transferred through these components until it gets to the USRP. Here we describe the main roles involved in each layer:

- USASDR

  In this class, the implemented methods allows to receive instructions from Matlab and transmit them to an instance of the class controller. It is implemented in Java, and runs as client. This layer allows parsing the received instruction from Matlab.

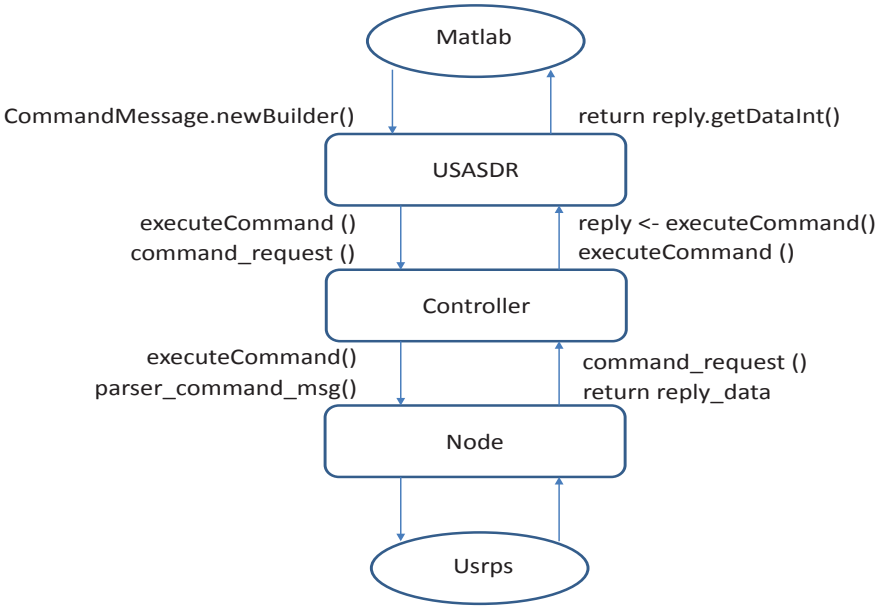


Figure 2.5: Components of the employed software architecture: USASDR, Controller and Node.

- Controller

  With this class, the communication between a client program (USASDR) and the node is possible. The implemented methods in this class include as many threads as instance nodes are considered. It allows to control many process at the same time. For instance, functions such as start_tx (for an order of transmission), or start_rx (for a reception instruction), are transferred simultaneously to many nodes after receiving a single instruction from a client program. The Controller and the USASDR are implemented in Java.

- Node

  The purpose of this layer is to make possible the communication between the UHD driver and the high-level commands. An instance of this class acts as a link between a PC and the hardware device, so that it should be executed at the same PC where the USRP is connected. It is implemented in C++, and most low-level features are implemented here. For instance, this layer allows the configuration of the transceiver by setting the corresponding values of the sampling rate, carrier frequency, among others.

With this software architecture, the basic functions provided by UHD are integrated and encapsulated. Now, we can use Matlab to control several USRP nodes by writing an unique script. However, before start working with the USASDR framework, some initial configurations corresponding to the node and controller layer are required.

These are specified in configuration files, i.e. Node files and Configuration files, which are described as follows:

- *Configuration in a node file*

  In this file, the number of USRPs and their corresponding IP addresses as well as the port number where the node layer is listening to receiving instructions (from the controller program) are specified. The following file “node2marte.cfg" summarizes the mentioned configuration:

  # The configuration file for the node layer

  usrps_count = 1
  addr = 192.168.10.2
  tag = MARTE
  port = 1665

  where “addr" is the IP address of the USRP node, and “port" the number of port from where it receives instructions. In addition, this file allows us to load the configuration of many USRPS, e.g. for a MIMO implementation.

- *Configuration in a controller file*

  A controller configuration file works with all the nodes described in the node configuration file. It requires to specify the IP address of the PC and the port where an instance of the node object is listening. A typical controller configuration file “cfg_tx.multi" is described as follows,

  # The configuration file for the controller layer

  node_count = 1
  addr = 193.144.201.249
  port = 1665
  port_client = 2665

  where “addr" and “port" are the IP address of the local host and port where it is listening (node object), whereas “port_client" is the number of port that a client instance (USASDR) utilizes to communicate with the controller.

Thanks to this architecture, we are able to build a testbed composed of many USRP nodes, which facilitates the experimental evaluation of the proposed schemes. A configuration often employed is depicted in Fig. 2.6, where a remote PC houses the client and controller programs, and a local PC runs the node component which controls two USRP nodes, each of them connected to a network card.

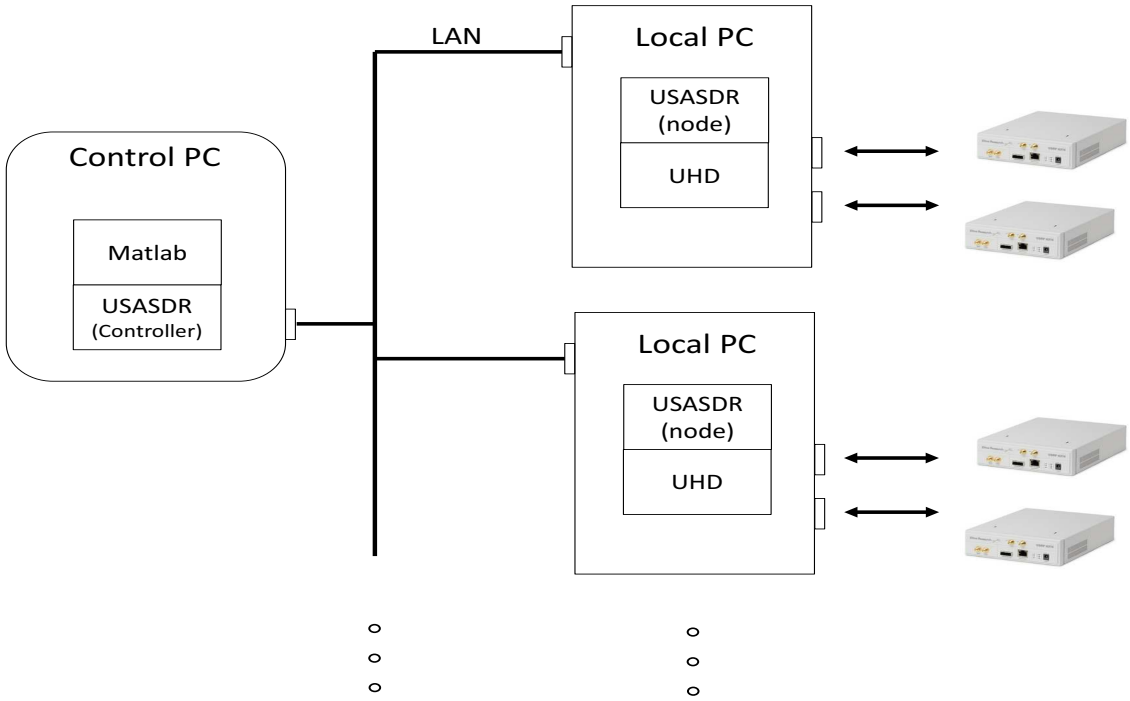


Figure 2.6: A testbed utilizing the USASDR framework.

In the following subsection, we provide a general example using USASDR. Nevertheless, it is worth mentioning that the implementation of this basic example required to devote time and manpower to make it work, since some hardware impairments as well as the misleading documentation in the manuals and sample programs provided by the manufacturer did not permit the rapid development of high-level functions. For instance, some overflow problems arise due to some missing packets between the PC and the USRP which involve time to get them resolved as they often require additional configurations. On the other hand, the hardware impairments introduce undesirable effects in the signals, as it is the case with an initial transient state provoked by an input filter when acquiring a signal. The solution of these problems among others has not been discussed as take part of the implementations detailed in this chapter. However, they constitute part of the development carried out in the platform.

2.3.3 General example using the USASDR

Let us suppose that we have two USRP nodes, a transmitter and a receiver, where each of them is connected to each network card of a PC which are controlled from a remote PC running MATLAB. We desire transmitting a sinusoidal tone indefinitely, while performing some measurements at the receiver side. In this scenario, the node layer is initiated at the local PC (one for each USRP), whereas the remote PC houses

the controller layer, the USASDR layer, and MATLAB as it is shown in Fig. 2.6. The transmission and reception of sinusoidal signal using the USASDR system at the remote PC can be summarized as follows:

First, we load an initial configuration in the file “init_env_local" where the path corresponding to the usasdr libraries are installed. Then, an usasdr object is created with a IP address “127.0.0.1" and port “2665" from where it receives instructions, and this created object, usasdrtx, controls the USRP transmitter.

```
clear all;
close all;

init_env_local % To initiate the Matlab configuration

usasdrtx=Usasdr('usrp_tx','127.0.0.1',2665);
nodes_tx=usasdrtx.getTxNodes(); % Get the list of available nodes
node_tx=char(nodes_tx)
```

Next, we proceed with the parameter setting of the USRP transmitter.

```
% Setting USRP parameters for a transmission
usasdrtx.setFreq(node_tx,2.5e9);   % Carrier frequency 2.5 GHz
usasdrtx.setTxRate(node_tx,20e6);   % Sample rate 20 Msps
usasdrtx.setTxGainBB(node_tx,2);   % Gain at baseband 2 dB
usasdrtx.setTxGainVGA(node_tx,15);   % Variable gain amplifer 15 dB
```

The data to be transmitted require to be stored at the PC. In this implementation, this signal is stored at the PC directly connected to the USRP.

```
% The waveform to be transmitted
fc=1e6;
fs=20e6;
nSamples_tx=1000;
matR = sin(2π(0:nSamples_tx−1)fc/fs);
matI = zeros(1,nSamples_tx);
matC = UsasdrMatrix(matR,matI);
% The stored data at the PC
usasdrtx.write(node_tx,matC);
```

For the transmission of a signal, we are required to specify a reference of time which is carried out by functions of synchronization. With this function the internal reference of the USRP is set to zero. After that, the USRP starts transmitting indefinitely and after three seconds (as it is specified by the fourth parameter in the start function, "3”).

```
% Transmission of the signal
usasdrtx.synchronize(node_tx,' ');
usasdrtx.start(node_tx,' ',0,3); % USRP transmit after three seconds.
```

On the other hand, at the receiver side a similar procedure is followed. In this case, a space of memory (buffer) is created before USRP starts acquiring a signal in order to store the signal. The acquisition is performed once, and after four seconds.

```
% Receiver
usasdrx=Usasdr('usrp_rx','127.0.0.1',2664);   different port number
nodes_rx=usasdrx.getRxNodes();
node_rx=char(nodes_rx)
usasdrx.setFreq(node_rx,2.5e9);
usasdrx.setRxRate(node_rx,20e6);
usasdrx.setRxGainLNA(node_rx,10);    10 dB
usasdrx.setRxGainVGA(node_rx,30);    30 dB
usasdrx.setRxNumSamples(node_rx,nSamples_tx);   % the space of memory is reserved
usasdrx.synchronize(' ',node_rx); % counter clock sets to zero
usasdrx.start(' ',node_rx,1,4); % a receiver operation after four seconds
```

Finally, the stored signal is retrieved at the remote PC and set to a MATLAB format.

```
% Read data from node_rx
matRead=usasdrx.read(node_rx); % To read the stored signal
rx_data=complex(matRead.getReal(),matRead.getImag());
matReadrx=usasdrx.read(node_rx);   % In a MATLAB format
```

In this example, we have created two objects: usasdrtx and usasdrx that are used to access to the functionalities of the USRPs through these identifiers. For the transmission, the usasdrtx object is employed to configure the USRP transmitter, where the sampling rate and the carrier frequency are selected according to the desired application, whereas the selected values of VGA and gain at baseband adequate the signal into a suitable range. In a similar way, the usasdrx object allows us to set the parameters such as carrier frequency, the sampling rate, and the gain values of the LNA and VGA for the USRP receiver.

Note that we are required to allocate an space of memory at the PC for the data to be transmitted or received, since this space is not provided by the USRP. In addition, both the transmission and reception of a signal require to establish a reference of time at each USRP. It is carried out by setting to zero the internal clock of

the USRP, so that one can specify the starting time for the transmission and reception.

In this case, the USRP nodes are controlled in an asynchronous mode. This is the case when a USRP transmits indefinitely and another one starts acquiring a signal after the transmission has started. Thus, we need to guarantee that the transmitter USRP is already transmitting a signal over the air before measuring a signal. Nevertheless, it is also possible to perform experimental evaluations in a synchronized way where a common reference of time is shared among several USRPs. This subject is addressed in the following subsection.

2.3.4 Synchronization among several USRPs

An experimental evaluation has to deal with ignored problems in simulation-based environments. For instance, the synchronization in time and frequency are important tasks for the proper detection and demodulation of a frame, which are taken for granted in a simulation-based environment. For that end, additional procedures are required, which often demand effort and time to know how to get it resolved when implementing them in a testbed.

Each USRP board provides its own internal clock for its synchronization, as well as inputs for external references. Thus, the synchronization in time and frequency among USRPs could be implemented by sharing a common reference source. With this aim, a procedure is required via software to indicate how to carry out this synchronization, and we focus on the time synchronization since it is the first relevant task in the detection of frames, which is also of interest for the evaluation of our proposals.

The UHD provides commands such as “set_time_now" and “set_time_unknown_pps" to have the times synchronized. The first sets the time registers on the USRP immediately. However, it requires to apply this command at the same time for all the USRPs which could be not so affordable if some latencies are introduced between the hardware device and the PC, or if they are connected through a network. On the other hand, the second command waits for a pulse to catch the edge, and then sets the time at the next PPS. This option allows to execute these commands within certain margin of time, since the synchronization does not starts when the instruction is received by the USRP, but after detecting the edge of next incoming PPS which is supposed to be the same for all USRPs. For this reason, the synchronization using the edge of the PPS is advisable and employed in the deployment of our scenarios, and it is implemented at the node class of the USASDR.

To provide a PPS signal, we evaluate the reliability of sharing a common source of clock with different devices. One of these was implemented by connecting a GPS to each USRP. However, this solution requires placing GPSs outside our indoor environment, as well as awaiting large times to capture and lock the satellite signals.

Furthermore, a drift in the edge of a PPS between two USRPs was found after some preliminary measurements. These disadvantages lead us to discard this solution because of its difficulties in indoor scenarios and its lack of precision. The best source for the synchronization in time and frequency was provided by signal generators based on rubidium oscillators. Nonetheless, due to its high expense it does not turn to be an affordable solution. A more practical solution with very good results can be found with a signal generator available in any laboratory, and which is employed in our experimental evaluations. In the next lines, the synchronization among three USRP nodes is implemented using USASDR.

```
% Synchronization among USRPs:

usasdr.synchronize(node_tx,node_rx1 node_rx2);

% Simultaneous transmission and reception after 4 seconds:

usasdr.start(node_tx,node_rx1 node_rx2,1,4);
```

In this case, a single object usasdr is used to control three USRPs, one transmitter node labeled as node_tx, and two receivers nodes: node_rx1 and node_rx2. All of them share the same reference (PPS signal) provided by a signal generator, and are synchronized after executing the function of synchronization. This function encapsulates the described commands set_time_unknown_pps which are executed at the node layer of the USASDR. Finally, the transmission and acquisition of data is carried out simultaneously at the same time as it is specified by the start function that indicates the transmission (node_tx), and the acquisition of data (node_rx1 and node_rx2) after four seconds.

MIMO synchronization

The N210 USRPs allows us to build MIMO nodes by connecting two single-antenna USRPs using a MIMO cable. With this implementation, one of the motherboards is labeled as master board and the other one as slave board. A clock configuration is created and assigned to the slave board where the information on how to get a 10 MHz reference and a PPS clock is provided. For this configuration, a set_time_now() command is advisable, since the two boards are connected through a dedicated cable that does not introduce latencies. Furthermore, this simplifies the synchronization procedure since it is only applied onto the master board, and the slave board is supposed to follow this timing reference. This configuration is described in the following lines:

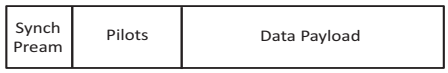


Figure 2.7: A data packet composed of a header (a synchronization preamble and pilots) and the data.

```
if ((*usrps)->get_num_mboards() == 2) {

// In case of detecting two USRP do :

// Specify the clock configuration
uhd::clock_config_t clock_config;
clock_config.ref_source = uhd::clock_config_t::REF_MIMO;
clock_config.pps_source = uhd::clock_config_t::PPS_MIMO;

// Set the clock configuration for the slave (mboard 0)
(*usrps)->set_clock_config(clock_config,1);

// Set time on the master (mboard 0)
(*usrps)->set_time_now(uhd::time_spec_t(0.0), 0);

// Time to do the synchronization
boost::this_thread::sleep(boost::posix_time::milliseconds(100));
}
```

Although this procedure guarantees the correct synchronization between two USRPs that compose a MIMO node, an scenario with several MIMO nodes requires to apply the aforementioned procedure between USRPs (using set_time_unknown_pps). In addition, the transmission and reception with MIMO nodes require to allocate space of memory for two streams of data, and consequently the implemented code should be flexible to any required configuration i.e. for single and/or multiantenna USRPs configurations.

With the prototyped testbed and described configurations, the transmission of data and its demodulation at the receiver side can be evaluated. For instance, a large acquisition of data may contain several data packets, each of them with its own header in order to be demodulated, as it is shown in Fig. 2.7. In this structure, a header is composed of a synchronization preamble and pilots, and the payload is demodulated (see. Fig. 2.8) after the detection of a preamble, a compensation of the carrier frequency offset (CFO), and a channel estimation stage.

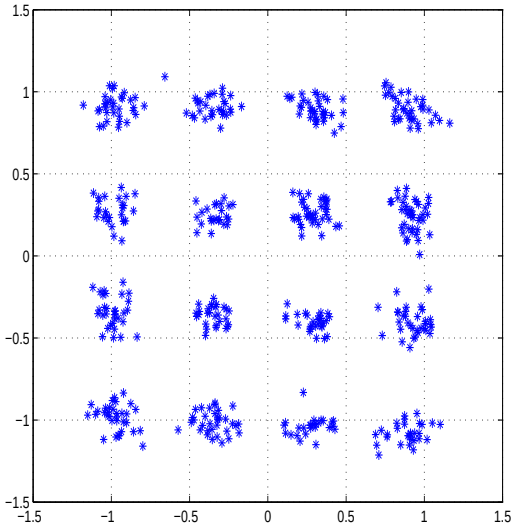


Figure 2.8: Received constellation obtained from data packets with 16-QAM symbols using a single-antenna N210 USRP with a XCVR2450 daughterboard at a carrier frequency of 5.6 GHz

2.3.5 Some considerations

The characterization of several parts in a communications system requires to collect enough measurements for reliable results. This implies the realization of an experiment several times which may not turn out to be practical if some delays and/or latencies are introduced during the evaluation. For this reason, a synchronous procedure where the transmission and reception of signals are executed at the same time is employed during our experimental evaluations. In comparison to an asynchronous procedure, i.e. a sequential execution of the transmission and reception of a signal, a synchronous procedure saves considerable time, and provides a more realistic and efficient PHY-layer implementation. For instance, some experimental evaluations require to have the estimation of a parameter (e.g. channel estimation) at the transmitter side after being computed at the receiver side, where this feedback should be carried out in short periods of time.

With the USASDR architecture, we optimize the set of instructions repeated at each iteration in order to avoid introducing additional delays. In doing so, the initial configuration is placed at the beginning of our program, whereas the involved functions in the acquisition of data are only repeated. This is shown in Fig. 2.9, where the flow of instructions repeated during the experimental evaluation are depicted, i.e. the synchronization, transmission and/or reception of the signal, and the retrieval of the data in a remote PC.

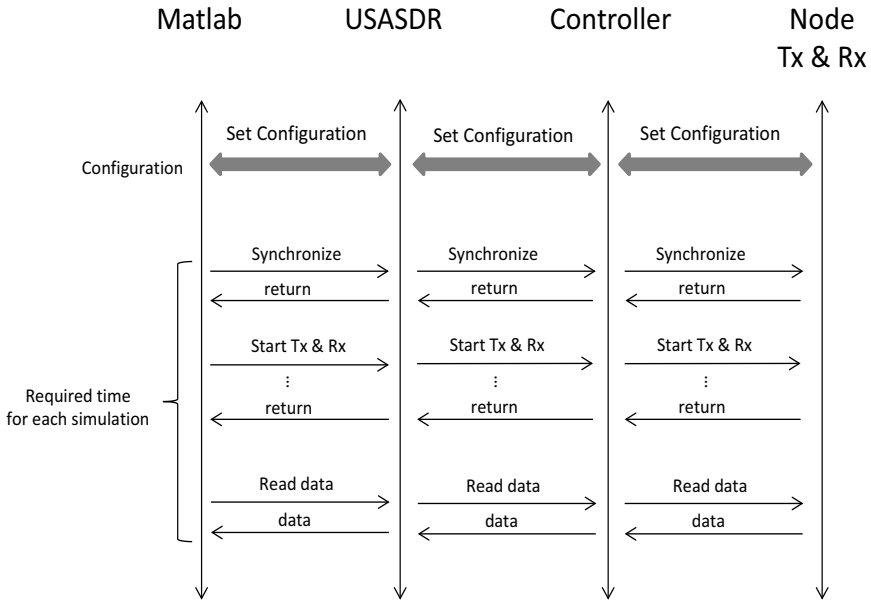


Figure 2.9: Flow of instructions in the experimental evaluation: Each execution of an instruction implies waiting for an acknowledgment at each layer of the software architecture, i.e. Usasdr, Controller and Node.

2.4 Experimental assessments of the proposed schemes

So far, we have presented the main features of the USRP platform, as well as the initial development carried out to start working with these boards. More specifically, the programming and configuration of a software architecture has been addressed, as it will be employed for the implementation of more demanding and complex testbeds. Thus, it is important to highlight that one of the contributions of this thesis is encountered in the development of the USRP platform, since it corresponds to a backbone upon which the proposed schemes have been evaluated.

In fact, a first step towards the implementation of a SDR-based testbed is the understanding of the general requirements to be satisfied by either the hardware and the software platform, since different types of applications may lead to totally different requirements. Some of these aspects can be, for instance, type and number of nodes, network topology, stationary/non-stationary environments, single or multi-antenna/multi-sensor node, processing capabilities available at each node, and scalability with respect to all previous features.

Here, we present some preliminary considerations and challenges encountered in the assessments of our proposals, after which we proceed with the design and implementation of each of these cases.

- First of all, the experiments have been conducted in the Laboratory of the Advanced Signal Processing Group (GTAS) at the University of Cantabria (UC), with a clear line of sight (LOS) between the transmitters and the receivers, where the captured close-to-real-world signals are used in an off-line manner, i.e. as inputs to the proposed schemes under study after being captured.

- The measurements at the GTAS laboratory typically correspond to indoor channels with long coherence times, as it was pointed out by previous measurement campaigns carried out at 5 GHz band [Gutiérrez et al., 2011]. These long coherence times are longer than the time devoted to each experiments, for which this scenario represents a stationary environment. Thus, the evaluation of algorithms in time-varying environments requires additional procedures or alternative solutions. In practice, this means that all the hardware equipment needs to be loaded over movable structures or mounted into vehicles that provide agility and connectivity.

- Some limitations of the USRP such as the latencies for establishing a configuration, the lack of memory, or the restricted bandwidth, among others, are considered when addressing the feasibility of an approach. For instance, a testbed comprised of several nodes implies unnecessary resources and manpower if what we want to validate does not depend on the number of nodes, for which a simplification of the scenario is implemented.

- The performance of a USRP is affected by some restrictions imposed by the host PC. For example, the Gigabit Ethernet allows to stream up to 25 Msps (for 16-bit samples) which translate to 20 MHz of usable bandwidth. For this reason, each USRP is directly connected to a network card instead of sharing this bandwidth with more USRPs using, for instance, a switch. For that end, in each PC three network cards were installed, two of them devoted for being connected to two USRPs, and a third one for allowing a remote control. In addition, in some cases it is required to pass transport parameters in the underlying transport layer between the host PC and the USRP to attain an adequate performance. In fact, UHD driver allows resizing the size of a single buffer in bytes (either for the transmission and the reception) for such a purpose. This, in turn, may lead us to reconsider the employed bandwidth in our experimental considerations.

- Calibration of the RF hardware is often required when doing measurements since imperfections caused by e.g. the digital up converter, digital filters, or the power amplifier, can cause the measured values to differ throughout the deployed nodes. This calibration can be done in the software by correcting the values of the received or transmitted signal based on a table of measured corrections factors. A simple calibration procedure can be made by connecting a signal generator directly to the board, and set up the corresponding gains.

On the other hand, the calibration procedure can be a time consuming process if done manually and there is chance that the new drivers or firmware on the hardware can render the calibration factor useless. For instance B210 USRPs provides many calibration procedures in comparison to the N210 USRPs.

- Misleading results can be obtained if hardware impairments are not taken into account. Even worse, many times its effects are mixed making more difficult to identify the problem and consequently to propose solutions. For instance the effects of spikes at particular frequencies, inadequate gains for the LNA and VGA components, and filter transients.

2.4.1 Contribution I: Bayesian spectrum sensing scheme

In this work, an adaptive Bayesian framework for multiantenna spectrum sensing is presented. In this scheme, we formulate our knowledge about the scenario employing distributions that can accommodate most of the general behavior of a CR network. Prior probability distributions are assigned to the unknown parameters, and posterior probability distributions are computed for these parameters. With these posterior distributions, some predictions are performed to improve the performance of our Bayesian detector. Furthermore, we include a learning mechanism that allows us to improve its detection performance by learning from previous sensing periods. Finally, this Bayesian scheme is evaluated experimentally employing multiantenna USRPs that operate under stationary and non-stationary environments.

State of art

The use of multiple antennas has been widely explored in previous works such as [Tugnait, 2012, Zhang et al., 2010, Taherpour et al., 2010, Wang et al., 2010, Sala-Alvarez et al., 2012, Ramirez et al., 2011, Ramirez et al., 2010] where the spatial correlation of the received signal at the different antennas, is exploited. Typically, a GLRT approach has been followed to find one-shot detectors [Wang et al., 2010, Sala-Alvarez et al., 2012, Ramirez et al., 2011, Ramirez et al., 2010]. For instance in [Ramirez et al., 2010], frequency and time-domain GLRTs that exploit the spatial correlation induced by the presence of a PU are proposed, whereas in [Ramirez et al., 2011] and [Sala-Alvarez et al., 2012] the additional structure in the spatial covariance matrix is addressed for the detection of a rank-P PU signal. However, the characteristics of the channel or the noise between consecutive sensing periods show smooth changes which are not taken into account by these one-shot detectors. In addition, the channel access patterns for PUs have been characterized as slowly time-varying in [Clancy and Walker, 2006] and [Lopez, 2011]. This motivate us to propose detectors that can exploit these slowly time-varying scenarios by learning from past sensing periods to improve its performance.

Some Bayesian approaches can be found in the literature. In [Bidon et al., 2008, Maio and Conte, 2010, Sohn et al., 2007] adaptive Bayesian

detectors for radar applications assume the availability of a training set of data for the estimation of noise statistics. This assumption, however, makes them inadequate for CR scenarios where noise-only data is not always available. Bayesian detectors specifically tailored for CR have also been proposed in [Font-Segura and Wang, 2010, Couillet and Debbah, 2010, Axel and Larsson, 2009, Couillet and Debbah, 2009, Jin, 2012]. These works assume prior distributions for the unknown parameters and apply Bayesian inference to come up with improved parameter estimates and, consequently, more reliable detectors. Nevertheless, we stress here that our Bayesian detector places priors directly on the spatial covariance matrices under both hypotheses. Moreover, our detection framework includes learning and forgetting steps that allow to track the variations of the channel and noise characteristics from frame to frame, which has not been explored in the aforementioned works. In doing so, we start with some initial results presented in [Manco-Vasquez et al., 2012b], which were later extended and published in [Manco-Vásquez et al., 2014a].

Eventually, this scheme has been validated by means of experimental evaluations in [Manco-Vasquez et al., 2012a]. In this respect, the feasibility of a detection technique involves meeting the requirements on sensitivity, dynamic range of the RF front-end, among other features, as it was shown in [Cabric and Brodersen, 2005]. One of the first publications in addressing the detection performance of single-antenna detectors is undoubtedly [Mubaraq et al., 2004] where detectors such as: a matched filtering, a cyclostationary-based detector and an energy detector are examined considering the noise variance uncertainty and impairments encountered in the receiver. Since then, several measuring campaigns have been carried out. More recently, the performance of a blind detector that exploit the symmetry property of cyclic autocorrelation function is reported in [Safatly et al., 2014]. Nevertheless, few works have addressed the feasibility of multi-antenna detectors, and even worse for CR networks. In fact, the introduction of low-cost SDR devices makes possible the evaluation, for instance, of several covariance-based detectors using USRPs as it is addressed in [Mate et al., 2011] where GNU radio is employed. In [Grimm et al., 2011], the evaluation of a detector which exploits the spatial resource among several antennas that sense in multiple bands and different directions is carried out. These approaches, however, do not incorporate learning techniques, and consequently motivates our work in the following subsection.

Preliminaries

We consider a cognitive receiver equipped with L antennas that acquires $n = \{0, \ldots, N-1\}$ snapshots denoted by $\mathbf{x}_t[n] \in \mathbb{C}^L$ at the t-th sensing period. The received signal during this sensing period is stacked in a matrix: $\mathbf{X}_t = [\mathbf{x}_t[0], \ldots, \mathbf{x}_t[N-1]]$, and the detection problem amounts to testing between two different structures

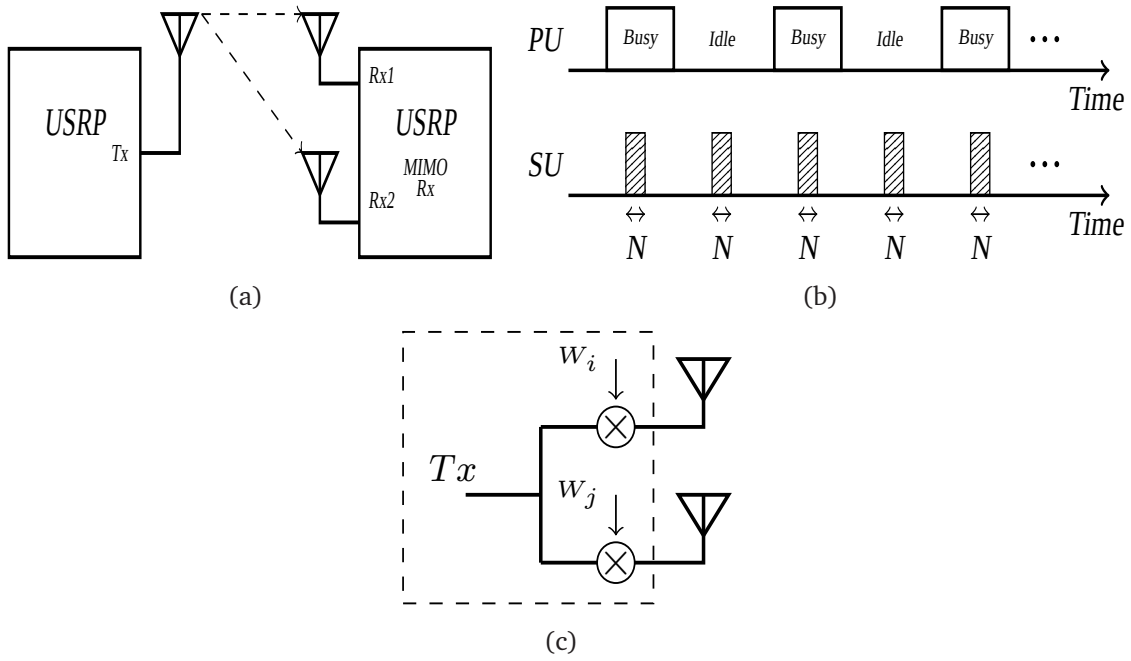


Figure 2.10: (a) The considered setup, (b) Sensing procedure: the CR device acquires N samples during the sensing period under each hypothesis, (c) Time-varying environment.

for the covariance matrix of $\mathbf{x}_t[n]$, as follows:

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}_t[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_t),\ n = 0, \ldots, N-1, \\ \mathcal{H}_0 &: \mathbf{x}_t[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{D}_t),\ n = 0, \ldots, N-1. \end{aligned} \tag{2.1}$$

Operation of the Bayesian scheme

A Bayesian inference is performed over a single sensing frame (at the t-th sensing period) by placing prior distributions to the covariance matrices under both hypotheses, as well as to the probability of the channel occupancy. With this Bayesian inference, we determine the availability of the channel if the collision probability is below some desired threshold. In addition, we introduce a learning mechanism where the posterior distribution under both hypothesis are updated from priors and the likelihood obtained from $\mathbf{X}_t$. In this learning mechanism, the obtained posteriors at each sensing period become the priors for the next sensing period, and uninformative priors at $t = 0$ are assumed. We provide two approximations for the exact posteriors within the family of its corresponding priors, a thresholding-based approximation and a Kullback-Leibler approximation. In the first of them, the posterior of the channel occupancy is truncated to either 0 or 1, whichever it is closer, whereas for the second approximation, the posteriors are obtained by minimizing the Kullback-Leibler distance.

Finally, a forgetting mechanism to forget past data is introduced to operate in non-stationary environments. It is based on a Bayesian λ-forgetting [Kulhavý and Zarrop, 1993] idea where the prior distributions at the frame $t+1$ are given by a "smoothed" version of the posterior distributions obtained after processing the frame t and the original prior distributions. With this forgetting step, a λ factor allows us to control how much information can be taken into account from previous data. For instance, with $\lambda = 0$, all previous data is forgotten and each sensing frame is considered independently, whereas with $\lambda = 1$, no forgetting occurs and the new posterior corresponds to the standard Bayesian posterior when $\mathbf{D}_t$ and $\mathbf{R}_t$ remains constant $\forall t$.

Conclusions

A Bayesian framework that employs a forgetting mechanism where the posterior for the unknown parameters $\mathbf{R}_t$ and $\mathbf{D}_t$ are used as priors for the next Bayesian inference is first studied by means of simulations. We evaluate this scheme under a stationary channel, slowly time-varying channel, and fast time-varying channel. For these scenarios, our Bayesian spectrum sensing scheme is able to exploit previous statistical information from past sensing frames to improve its performance. In fact, it is observed that it takes some sensing frames to reach a steady state, after which we evaluate its performance. Our evaluation can be summarized for two scenarios:

- For stationary environments, we observe that a $\lambda = 1$, provides the best detection performance, since in this static environment the unknown covariance matrices ($\mathbf{R}_t$ and $\mathbf{D}_t$) remains constant, then the posterior estimates are correctly updated. In fact, a Kullback-Leibler approximation provides a best performance in comparison to the thresholding-based approximation, at the cost of higher computational complexity.

- In non-stationary environments, the best detection performance is attained by selecting forgetting factor values less than $\lambda = 1$. In this case, the covariance matrices do not remain constant, and it is convenient to forget some previous data to attain better estimates. The accurate estimate provided by the Kullback-Leibler approximation is limited by the variations of the environment, and a coarse approximation (thresholding-based approach) attain a better performance.

Our implementation takes into account an scenario comprised of a PU and a two-antenna SU as it is shown in Fig. 2.10(a). This is prototyped with a single-antenna node (See Fig. 2.2(b)) and a MIMO node composed of two single-antenna USRP connected through a MIMO cable. Each node consists of a N210 USRP motherboard and a XCVR2450 RF daughterboard.

In each experiment, a single-antenna PU accesses the channel according to a predefined pattern, and a SU senses periodically the medium. The sensing procedure is shown in detail in Fig. 2.10(b), where N samples are acquired at each sensing

period and stored in a $2 \times N$ matrix format. In this case, the predefined pattern is known, so that each time the environment is sensed we know the true hypothesis and consequently the P_D and P_{FA} can be estimated for a given threshold.

We evaluate the performance of our detector in regimes of low SNR, where it is more challenging to attain the specified requirements with a CR technology. This SNR is measured from the received signal at baseband and is controlled in our setup by the transmitter power. Finally, for both considered scenarios, i.e. stationary and non-stationary environments, the obtained performance is compared with those of one-shot GLRT detectors (Sphericity and Hadamard detectors) which have been widely reported in the literature. Our observations regarding these scenarios (corresponding to those evaluated in a simulation environment) are summarized as follows:

- For a stationary environment, the measurements in our indoor scenario provides the necessary environment, as long as the transceivers remain at the same position (see [Gutiérrez et al., 2011]). In this case, our results (quantified in P_D and P_{FA}) show that our Bayesian detector achieves an steady-state performance, after processing some sensing periods which corresponds to the required time in estimating the posterior parameters. Moreover, the selection of a higher value for the forgetting factor, provides better results for which it is convenient to set this variable to its maximum value. The obtained gains of one-shot detectors (Sphericity and Hadammard detectors) show to be almost identical since the noise variance at each receiver antenna was very similar, for which both detectors turn out to be the same. On the other hand, a comparison among the evaluated detectors shows a significant gain of our Bayesian scheme over the one-shot detectors.

- A more challenging scenario is addressed with the experimental evaluation in a non-stationary environment. The recreation of this scenario is provided by the mobility of the USRP. Nonetheless, this option was discarded due to the complexity to accomplish this requirement with our testbed. Instead, we resort to emulate time-varying channels by transmitting the PU signal using a time-varying beamformer to provide correlated Rayleigh processes [Baddour and Beaulieu, 2005]. It consists in multiplying the PU signal by a 1×2 time-varying beamformer obtained according to an autoregressive model for fading channel 2.10(c). The SU follows the same described procedure in the Fig. 2.10(b), while the PU follows a predefined pattern with a weighted transmitter signal.

  For this non-stationary environment, the obtained results also show to be better than those obtained by the GLRT approaches. However, a suitable forgetting factor, in this case, does not correspond to the maximum value since a small degradation in its performance is observed. Instead, a bit less than the maximum value provides the best performance as it is shown in the simulation results.

To sum up, we have evaluated and validated a novel Bayesian detector. More particularly, we show the feasibility of learning efficiently the posteriors parameters to detect a PU signal under stationary and non-stationary environments, which lead us to significant gains over the performance of widely employed one-shot GLRT detectors.

2.4.2 Contribution II: Cooperative spectrum sensing approach

In this second contribution, a KCCA-based detector tailored for CSS scenarios is first examined in a simulation-like environment. More specifically, we address an scenario where a PU has a large radio coverage, while some interferers with a small coverage affect a single SU. Our detector is based on a KCCA technique that allows us to exploit the non-linear correlation among the received signals of each SUs. In doing so, the measurements at each SU are reported to the fusion center (FC) where kernel methods are applied in order to maximize the non-linear correlations among the received measurements. Next, the statistical tests are derived allowing the SUs to make decisions either autonomously at each SU or cooperatively at the FC. The proposed scheme operates in a totally blind fashion, and adapts itself to a time-varying radio environment by retraining from time to time or continuously while detection operates normally. Furthermore, our scheme is able to work with different features extracted from data acquired by each SU. Finally, this CSS scheme is validated in our platform, where we show that our KCCA detector is able to operate in non-ideal scenarios where impairments such as non-Gaussian noise and/or interference may appear.

State of art

There are many works related to the detection problem under interfering signals. An energy detector under interference is unable to distinguish the PU signals from the interference, and it shows a poor performance [Makarfi and Hamdi, 2013, Biglieri, 2011]. In fact, these non-ideal scenarios have also been studied for CSS schemes. In [Guimarães et al., 2013], several eigenvalue-based approaches are evaluated under impulsive noise and interference. In addition, other impairments inherent to the implementation of CSS strategies such as timing inaccuracies or the synchronization errors among the SUs may provoke interfering signals that degrade the detection performance as it is shown in [Nieminen et al., 2010, Song et al., 2010].

In Heterogeneous Networks (HetNets) a macrocell-edge user may experience interference from small cell transmissions using the same radio frequency band, and thus the performance of the spectrum sensing schemes are also degraded. Some interference-mitigation schemes have been proposed in [Font-Bach et al., 2014, Zhao and Sasaki, 2013]. In [Zhao and Sasaki, 2013], the authors propose a sensing scheme that cancels the interference power prior to PU detection by decomposing the received power into primary signal power,

interference signal power, and the noise power. Moreover, the impact of interference in underlay cooperative cognitive networks has also been extensively studied [Ho-Van, 2013], [Ding et al., 2014].

Nevertheless, we propose an CSS scheme that employs learning techniques and that does not require any statistical prior information. In this context, learning techniques have shown to improve the detection performance of soft-decision approaches [Ma et al., 2008, Yan et al., 2009]. For instance, kernel-based learning (KBL) methods have been recently tailored for CR networks where it covers a wide range of topics such as the development of detection procedures [Harchaoui et al., 2013], robust signal classification, and spectrum occupancy online prediction [Ding et al., 2013]. These latest approaches motivate us to explore the application of kernel methods in the detection of a primary user. More specifically, in a CSS scenario where the primary signal is supposed to be correlated among the received signals at each secondary user, we aim to look for non-linear transformations of the received data that provide the maximal correlation, and resort to KCCA methods that provides the retrieval of optimal projections (often non-linear transformation) [Hardoon et al., 2004, Van Vaerenbergh et al., 2013], that could attain an improvement in the detection performance. More recently, in [Thilina et al., 2013], the energy levels reported by each SU are categorized into classes that represent the availability of channel, i.e. occupied or empty. This classifier requires a *training phase*, during which it learns from a set of training feature vectors, after which it can be employed for online detection (i.e. *test phase*). In our approach, however, we do not require any labeled data set nor any other prior information about the PU signaling format, since it operates in a completely blind fashion. Some preliminary ideas were devised in [Van Vaerenbergh et al., 2013], and later developed and published in [Manco-Vásquez et al., 2014c].

On the other hand, because of the inability to model all noise sources, impairments as well as interference sources, some efforts in the deployment of a cooperative detection by means of experimental evaluations have been carried out [Chaudhari et al., 2014, Guimarães et al., 2013, Chen and Qiu, 2011, Kremo et al., 2014, Yagi et al., 2010, Chowdhury et al., 2011]. For instance, in [Chaudhari et al., 2014] a practical testbed composed of mobile sensors is used to verify the cooperative gains when applying different fusion rules. In [Guimarães et al., 2013] an eigenvalue-based detector tailored for cooperative spectrum sensing is evaluated under impulsive noise, and more recently, in [Font-Bach et al., 2014] the interference between a macrocell and femtocell in an heterogeneous network is addressed employing an interference management scheme. This scheme is implemented in an FPGA-based architecture and validated in a testbed that emulates the coexistence of a macrocell/femtocell. However, these works do not address the experimental evaluation of CSS scenarios under external interferences and they do not incorporate learning techniques.

Preliminaries

For our scenario, we consider M single-antenna SUs and a PU that coexist in the same area [Akyildiz et al., 2011]. We assume that the PU signal can be sensed at each SU, while some interferences can be experience independently at each SU during the time that the channel remains idle. The binary hypothesis testing problem is formulated as follows:

$$p(\mathbf{r}|\mathcal{H}_1) \neq \prod_{i=1}^{M} p_i(r_i|\mathcal{H}_1)$$
$$p(\mathbf{r}|\mathcal{H}_0) = \prod_{i=1}^{M} p_i(r_i|\mathcal{H}_0)$$

This model does not make any assumption about the distributions of the primary, interference and noise signal, and is independent of employed transmission schemes (i.e. for the PU or the interferers). On the other hand, we exploit the fact that the optimal detectors at each SU will be highly correlated, i.e. if SUs are either all under the null hypothesis or all under the alternative hypothesis. For that end, the proposed scheme aims to find the non-linear transformations of the measurements that provides maximal correlation. These non-linear transformations are employed to decide if the measurements come from the distribution $p(\mathbf{r}|\mathcal{H}_1)$ or from $p(\mathbf{r}|\mathcal{H}_0)$.

Operation of the KCCA scheme

Our scheme starts with an initial cooperative learning stage where the sensors measurements are transmitted to the FC, and the near-optimal local decision functions (local statistics) are extracted. Then, these local statistics are broadcasted to the SUs, which can operate in one of two modes:

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.

The input to the KCCA-based detector is given by feature vectors extracted from the measurements at each SU, where this feature vector is composed of a wide variety of features such as energy, kurtosis, among others. During the initial cooperative stage (or training stage), the feature vectors extracted at each SU are reported to the FC and KCCA is applied to retrieve non-linear projections. This method permits mapping the data into a high-dimensional *feature space* [Bishop, 2006], after which standard CCA is performed in the new space. Each i-th SU produces N feature vectors, $\{\mathbf{x}_{i1}, \mathbf{x}_{i2}, ...\mathbf{x}_{iN}\}$, and we aim to maximize the correlation among these feature vectors. The problem is solved by the method of Lagrange multipliers, where the solution contains the canonical weights corresponding to the i-th SU. After this training stage, we obtain the non-linear local detectors and a global test statistic as the sum of these non-linear local detectors which are evaluated during a testing stage.

Conclusions

First, we study the proposed approach by means of simulations. In this regard, we consider different scenarios in which noise or noise plus interference are present, and for which different features are extracted during the sensing period. Moreover, both operation modes are evaluated i.e. a distributed configuration (autonomous testing at each SU) and a centralized configuration (cooperative testing at the FC), and its performance is compared to an energy detector.

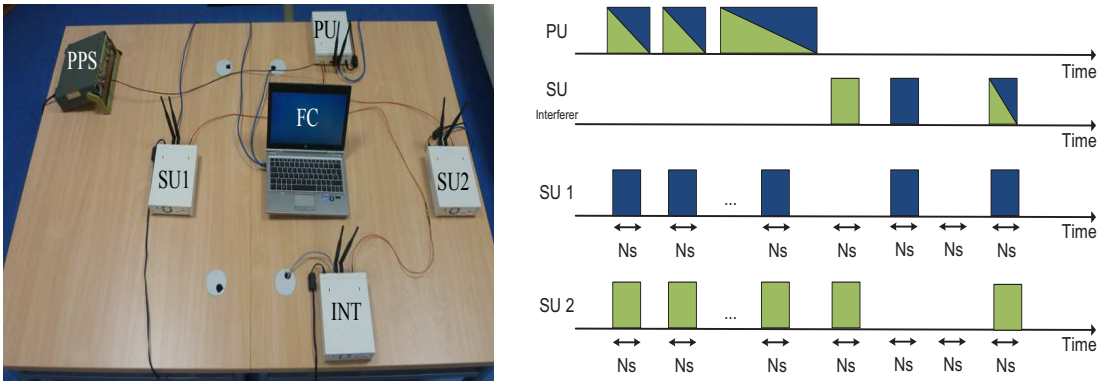
To evaluate the learning ability in the separation of both hypothesis, we focus on the obtained local detectors T_i (or decision functions) that should be able to assign negative values to the primary signal, and positive values to the noise signal. This is carried out in two scenarios and using both simulations and experimental measurements. Our observations considering our simulation results can be summarized as follows:

- For scenarios with only noise and using only energy measurements as input to our KCCA detector, we observe that both the KCCA and the energy detector attain the same performance, since the obtained tests are close to the optimal NP detector.

- In scenarios with noise plus interference, our KCCA detector obtains a significant gain over an energy detector. For both operations modes, our approach assigns negative values to the primary signal, whereas the noise and the interference are assigned positive values. These results show that our KCCA-based detector is able to learn to distinguish between the noise or noise plus interference and the primary signal. For these scenarios, a cooperative testing at the FC provides an improved detection performance over an autonomous testing at each SU, since in this mode all the feature vectors reported at the FC are employed to extract the global test. In addition, other extracted features such as the kurtosis or the cyclic statistics can be included to further boost the detection performance.

Our approach operates in blind manner, and can be applied to time-changing environments, since it adapts itself by retraining from time to time. After evaluating the detection performance in a simulation environment, we address its implementation in our platform to test the feasibility of this approach.

The scenario for validating our proposal requires several nodes, some of them as SUs that aim at detecting a PU, whereas other nodes are devoted to act as interferers. Nevertheless, the implementation of this scenario involves a higher complexity because of the time required to obtain the local statistics $T_i(x)$. Instead, the feasibility of our proposal is evaluated in an scenario recreated with a PU, two SU nodes, and an interfering node (INT). Two SUs aim at detecting a PU while an interfering node transmits randomly during the time that the PU stops transmitting. The testbed is prototyped with single-antenna N210 USRPs and synchronized in time by a



(a) (b)

Figure 2.11: (a) Implemented Testbed: a PU, two SU and an Interfering (INT) node, where the SUs report their measured energy levels to a fusion center, (b) Measurement procedure: a PU transmits using two bands of frequency channels represented by two blocks, each SU only senses in one of this bands and the interfering node transmits on any of these channels

PPS signal for simultaneous transmission and reception as it is shown in Fig. 2.11(a).

During each experiment, the PU transmits an OFDM signal using two bands of frequency channels (2-4 MHz & 4-6 MHz) represented by two blocks (i.e. 4 MHz), [1] each SU only senses one of these frequency bands, and the interfering node transmits on any of these channels following a Bernoulli distribution (with a probability of channel occupancy, $p = 0.5$). With 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. To clarify this configuration, Fig. 2.11(b) shows the implemented measurement procedure. In this procedure, each SU senses periodically acquiring each time N_s samples to compute the received energy level. Recall that our KCCA-based detector requires two stages: a training and testing stage. During a first training stage, the two SUs collect, each of them, a set of data composed of measured energy levels corresponding to each sensing period. These sets of data are stored at the central PC or fusion center (FC), where the canonical weights α_i are calculated, and consequently the local statistics $T_i(x)$ as well as a global statistic $T(x)$, are obtained. Next and during a testing stage, the performance of these detectors is evaluated using a new set of acquired data and compared against an energy detector.

Having implemented and configured our testbed, we evaluate the feasibility of our KCCA cooperative spectrum sensing detector. The values of SNR and SINR at the SU side are controlled by attenuating the maximum transmission power of a N210 USRP (i.e. at the PU and the INT node), and this allows us to define two possible

[1] This waveform follows an IEEE 802.11a format which is up-sampled so as to accomplish with the required configuration

cases: one in which the interference power lies below the received power of the PU signal and another one where the interference power levels are higher than the power of the PU signal. These cases are summarized here:

- In the first case where the interference power lies below the received power of the PU signal. The KCCA-based detector is able to separate the interference and noise from the primary signal. This is attained by mapping these signals to different values of the test statistics, which is observed by plotting the probability density function (PDF) of the measured energy levels under both hypotheses, and the obtained local statistics. The measurement PDFs follow Gaussian distributions which do not change abruptly as the detector operates under an almost stationary scenario. With these observations, we corroborate the learning ability to detect the PU signal by exploiting the correlation among the received signals, and in spite of the different noise variance at each SU as well as the interference power received at each SUs (because of the different channel frequencies).

- A more interesting case is addressed with a second setup where the interference power level is above that of the PU signal. In this case, the received interference varies more notoriously at each SU, and our KCCA framework is able to learn from the reported energy levels by each SU, the local statistics to detect the PU. Like the aforementioned case, this is corroborated by showing the PDF of received measurements as well as the local statistics. Moreover, the results (quantified in ROC curves) show a significant gain over an energy detector which is unable to distinguish between the noise plus interference and the primary signal. Moreover, our technique exhibits a much better performance than that of the energy detector as the interference level increases, since our KCCA framework exploits better the correlation of the received PU signal when more uncorrelated external interference is present.

  Within this second setup, a more challenging case is also addressed. In this scenario, the received power corresponding to the noise and the PU signal in one of the SU have similar energy values, for which the SU is unable to distinguish the data from the noise (e.g. with an energy detector). Nonetheless, the KCCA-based detector exploits the measurements from the other SU to extract the local and global statistics whose performance largely overcomes the performance of an energy detector, as it is confirmed with the obtained results in the corresponding ROC curves.

To sum up, our simulation results have been corroborated by the experimental results that show its robustness under the presence of interference, while obtaining a considerable advantage with respect to the use of an energy detector. It worth mentioning that although our KCCA framework may exploit additional knowledge about the PU signal to improve its performance detection (e.g. cyclic statistics), this feature was not verified here as it was explored after this measurement campaign in which we address its feasibility in a more realistic and challenging scenario.

Chapter 3

A Bayesian Approach for Detection in Cognitive Radio Networks

Title: A Bayesian Approach for Adaptive Multiantenna Sensing in Cognitive Radio Networks

Authors: Julio Manco Vásquez [1], Miguel Lázaro Gredilla [2], David Ramírez [3], Javier Vía [1], and Ignacio Santamaría [1]

Affiliations:

[1] Dept. of Communications Engineering, Univ. of Cantabria, Spain.
[2] Dept. Signal Theory and Commun., Univ. Carlos III de Madrid, Spain.
[3] Signal and System Theory Group, University of Paderborn, Germany.

Journal: Signal Processing, vol. 96, pp. 240, March, 2014.

A Bayesian Approach for Adaptive Multiantenna Sensing in Cognitive Radio Networks

Abstract
Recent work on multiantenna spectrum sensing in cognitive radio (CR) networks has been based on generalized likelihood ratio test (GLRT) detectors, which lack the ability to learn from past decisions and to adapt to the continuously changing environment. To overcome this limitation, in this paper we propose a Bayesian detector capable of learning in an efficient way the posterior distributions under both hypotheses. Our Bayesian model places priors directly on the spatial covariance matrices under both hypotheses, as well as on the probability of channel occupancy. Specifically, we use inverse-gamma and complex inverse-Wishart distributions as conjugate priors for the null and alternative hypothesis, respectively; and a binomial distribution as the prior for channel occupancy. At each sensing period, Bayesian inference is applied and the posterior for the channel occupancy is thresholded for detection. After a suitable approximation, the posteriors are employed as priors for the next sensing frame, which forms the basis of the proposed Bayesian learning procedure. The performance of the Bayesian detector is evaluated by simulations and by means of a CR testbed composed of universal radio peripheral (USRP) nodes. Both the simulations and experimental measurements show that the Bayesian detector outperforms the GLRT in a variety of scenarios.

Keywords Bayesian Inference Bayesian Forgetting Cognitive Radio Generalized Likelihood Ratio Test (GLRT) Multiantenna Spectrum Sensing.

3.1 Introduction

Cognitive Radio (CR) networks [Mitola and Maguire, 1999, Haykin, 2005, Akyildiz et al., 2006] rely on spectrum sensing as a key operation that secondary users (SU) must perform in order to identify whether a wireless communication channel is in use by a licensed primary user (PU) or not [Yucek and Arslan, 2009]. A reliable spectrum sensing stage is crucial to detect spectrum holes that can be subsequently filled with transmissions from SU [Nagaraj, 2009]. To this end, detectors employing multiple antennas have received increased attention recently because they do not require prior knowledge about the PU signalling scheme and are able to work with asynchronously sampled signals [Wang et al., 2010, Sala-Alvarez et al., 2012, Tugnait, 2012, Ramirez et al., 2011, Zhang et al., 2010, Taherpour et al., 2010, Ramirez et al., 2010]. These multiantenna detectors exploit the fact that under the null hypothesis (only noise) the signals received at the different antennas are spatially uncorrelated, whereas the presence of a PU induces some correlation and/or additional structure in the spatial covariance matrix.

Since the binary hypothesis testing problem involves some unknown parameters (e.g., noise variance and channel), the generalized likelihood ratio test (GLRT) approach has been typically followed to find one-shot detectors in several scenarios [Wang et al., 2010, Sala-Alvarez et al., 2012, Ramirez et al., 2011, Ramirez et al., 2010]. In [Ramirez et al., 2010], frequency and time-domain GLRTs have been derived that only exploit the spatial correlation induced by the presence of a PU, whereas the problem of detecting a rank-P primary user signal is addressed in [Ramirez et al., 2011][Sala-Alvarez et al., 2012]. However, these detectors do not take into account the smooth changes in the characteristics of the channel or the noise that can be expected between consecutive sensing frames. More precisely, it is reasonable to assume that the time scale of variation of the statistical parameters involved in the detection problem (for instance, noise variance or space-time PU activity pattern) are much longer than the sensing period. For instance, channel access patterns for primary users have been characterized as slowly time-varying in [Clancy and Walker, 2006] and more recently in [Lopez, 2011]. It is clear that detectors able to learn from past decisions would provide improved performance in these slowly time-varying scenarios. With this goal in mind, in this paper we propose an adaptive Bayesian framework for multiantenna sensing and evaluate its performance both by simulations and by means of a CR testbed.

Adaptive Bayesian detectors for radar applications have been proposed in [Bidon et al., 2008, Maio and Conte, 2010, Sohn et al., 2007], where a training set of data is available for the estimation of noise statistics. For cognitive radio applications, however, noise-only data is not always available; therefore, these adaptive Bayesian techniques cannot be applied to the scenarios considered in this paper. Bayesian detectors specific for cognitive radios have been previously proposed in [Font-Segura and Wang, 2010, Couillet and Debbah, 2010, Axel and Larsson, 2009, Couillet and Debbah, 2009, Jin, 2012]. Typically, these works assume a prior distribution for the unknown parameters and apply Bayesian inference to come up with

improved parameter estimates and, consequently, more reliable detectors. In comparison to these Bayesian approaches, our work presents two main novelties: first, our Bayesian detector places priors directly on the spatial covariance matrices under both hypotheses; and second, it includes learning and forgetting steps that allow to track the variations of the channel and noise characteristics from frame to frame. Our Bayesian approach is able to learn from past sensing frames when the coherence time of the propagation channel [Hlawatsch and Matz, 2011] is longer than the time elapsed between consecutive sensing periods. We refer to this situation as "smooth channel variations". Let us also remark that the proposed Bayesian detector is specifically tailored for multiantenna cognitive receivers and, consequently, this approach is not directly applicable to single-antenna SUs.

Specifically, our multiantenna Bayesian model uses inverse-gamma and complex inverse-Wishart distributions as conjugate priors for the null and alternative hypothesis, respectively; and a binomial distribution as the prior for channel occupancy. The reason for choosing these priors being that under Gaussian noise they are the conjugate priors for this problem and, therefore, the posteriors can be calculated in closed form. More precisely, the posterior conditioned on the channel state occupancy (idle or busy) adopts the same form as the prior. However, the unconditional posterior (marginalized over the channel state) becomes a convex combination of the priors. Since the marginalized or unconditional posteriors summarize the information gathered so far about the actual CR scenario, they are used as priors for the next sensing period: this represents the learning stage. To keep the learning process simple and scalable, the unconditional posterior (which is a linear combination of complex inverse-Wishart distributions when the PU is present) must be approximated within the family of the prior. Furthermore, the procedure is equipped with a forgetting mechanism based on [Kulhavý and Zarrop, 1993] that allows to work on non-stationary environments.

In this paper, we extend some initial results about this Bayesian approach which were presented in [Manco-Vasquez et al., 2012b] and [Manco-Vasquez et al., 2012a]. In particular, we consider the optimal approximation of the unconditional posterior according to the Kullback-Leibler (KL) distance, and compare its performance and computational cost with the simple approximation based on thresholding, which was discussed in [Manco-Vasquez et al., 2012b]. Also, we present a more in depth study of the proposed CR detector performance for different number of receiver antennas, observations and channel conditions. Finally, we also evaluate experimentally the performance of the proposed detector, in comparison to conventional GLRT-based detectors, using to this end a cognitive radio hardware platform based on Universal Software Radio Peripheral (USRP) devices [Ett, 2014].

The rest of the paper is organized as follows. The detection problem for the CR network is formulated in Section II. The Bayesian inference procedure for a single sensing frame is presented in Section III, where two different approximations of the posterior are derived. The learning and forgetting procedure for dealing with multiple sensing frames will be discussed in Section IV. In Section V, we analyze the simulation results for different settings including both stationary and non-stationary

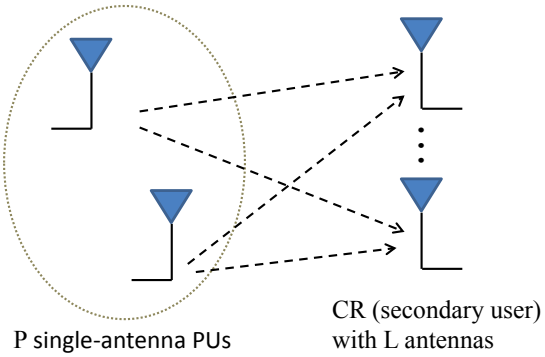


Figure 3.1: CR detection model: a cognitive user with L antennas tries to detect the presence of $P \geq 1$ single-antenna PUs or, equivalently, a single P-antenna PU.

environments, whereas the results obtained with the CR testbed are presented in Section VI. Finally, our main conclusions are summarized in Section VII.

3.2 Preliminaries

Notation

In this paper, we use bold-face lower case and bold-face upper case letters for column vectors and matrices, respectively; and light-face lower case letters for scalar quantities. The superscripts $(\hat{\cdot})$ and $(\check{\cdot})$ refer to the parameters of the posterior and prior distributions, respectively; and $(\tilde{\cdot})$ is used for estimated matrices and scalars. The determinant of a matrix $\mathbf{A}$ is denoted as $|\mathbf{A}|$, its trace as $\mathrm{trace}(\mathbf{A})$, the operator $\mathrm{diag}(\mathbf{A})$ refers to a diagonal matrix formed with the elements along the main diagonal of $\mathbf{A}$, $[\mathbf{A}]_{ij}$ denotes the ij element of the matrix, and the superscript $(\cdot)^H$ denotes Hermitian transpose. Finally, $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ indicates that $\mathbf{x}$ is a complex circular Gaussian random vector of mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$.

3.2.1 Problem Statement and GLRT detectors

We consider a cognitive receiver equipped with L antennas that wants to detect whether the channel is occupied by a primary user or not. During the t-th sensing frame, the cognitive receiver acquires $n = 0, \ldots, N-1$ snapshots denoted by $\mathbf{x}_t[n] \in \mathbb{C}^L$. The signal received during the t-th sensing period is stacked in a matrix: $\mathbf{X}_t = [\mathbf{x}_t[0], \ldots, \mathbf{x}_t[N-1]]$. The spectrum sensing problem can be formulated as a binary hypothesis test as follows

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}_t[n] = \mathbf{H}_t \mathbf{s}_t[n] + \mathbf{v}_t[n], \\ \mathcal{H}_0 &: \mathbf{x}_t[n] = \mathbf{v}_t[n], \end{aligned} \tag{3.1}$$

where $\mathbf{x}_t[n]$ is the acquired snapshot at time n, $\mathbf{s}_t[n] \in \mathbb{C}^P$ is the primary signal vector, which might represent the signal emitted by a single PU with P antennas or the signals emitted concurrently by P single-antenna PUs (see Fig. 3.1), $\mathbf{H}_t \in \mathbb{C}^{L\times P}$ describes the multiple-input multiple-output (MIMO) channel between the PU and the cognitive receiver, and $\mathbf{v}_t[n]$ is modeled as zero-mean additive white Gaussian circular noise. In our model, both the channel $\mathbf{H}_t$ and the transmitted signal $\mathbf{s}_t[n]$ are assumed to be random quantities. More specifically, taking into account that any spatial correlation and scaling of the primary signal can be absorbed in the channel matrix, we model $\mathbf{s}_t$ as a zero-mean circular complex Gaussian, spatially white and power-normalized. Under these assumptions, the distributions of the vector-valued observations under each hypothesis $\mathcal{H}_1$ and $\mathcal{H}_0$ can be modeled as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_t)$ and $\mathcal{CN}(\mathbf{0}, \mathbf{D}_t)$, respectively. Therefore, without any additional prior knowledge about the modulation format or signalling scheme used by the PU, the spectrum sensing problem amounts to testing between two different structures for the covariance matrix of $\mathbf{x}_t[n]$:

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}_t[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_t),\ n = 0, \ldots, N-1, \\ \mathcal{H}_0 &: \mathbf{x}_t[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{D}_t),\ n = 0, \ldots, N-1. \end{aligned} \tag{3.2}$$

Under $\mathcal{H}_1$, the $L \times L$ covariance matrix $\mathbf{R}_t$ can be written as $\mathbf{H}\mathbf{H}^H + \mathbf{D}$, i.e., a rank-$P$ matrix plus a scaled diagonal matrix. In this paper, we assume that the rank P, the channel $\mathbf{H}$, and the noise variance are all unknown parameters; therefore, all the structure of $\mathbf{R}_t$ under the alternative hypothesis reduces to being a positive semidefinite matrix. On the other hand, under $\mathcal{H}_0$, $\mathbf{D}_t$ is an arbitrary diagonal covariance matrix. In this way, our proposed Bayesian scheme is able to work in the most general setting.

Notice also that the likelihood under each hypothesis depends on unknown parameters and therefore the hypotheses are composite. The most typical approach to solve this kind of testing problems is the generalized likelihood ratio test (GLRT) [Kay, 1998]. When the noise is independent and identically distributed (iid) at each antenna ($\mathbf{D}_t = \sigma^2\mathbf{I}$) and $P \geq L-1$, the GLRT is the well-known sphericity test [Mauchly, 1940],[1] which is given by

$$\mathscr{L}_S = \frac{|\mathbf{S}_t|^{1/L}}{(1/L)\ \mathrm{trace}(\mathbf{S}_t)} \tag{3.3}$$

where $\mathbf{S}_t = \mathbf{X}_t\mathbf{X}_t^H/N$ is the sample covariance matrix.

A more general testing problem that can accommodate calibration uncertainties in the different antenna front-ends, takes into account a generic diagonal noise covariance matrix under $\mathcal{H}_0$. The GLRT in this case is the Hadamard ratio [Wilks, 1935] and is given by

$$\mathscr{L}_H = \frac{|\mathbf{S}_t|}{\prod_{i=1}^{L}[\mathbf{S}_t]_{ii}}. \tag{3.4}$$

[1] For rank-deficient signal covariance matrices the GLRT is, in general, more complicated, as it was shown in [Ramirez et al., 2011].

3.3 Bayesian inference on a single sensing frame

The Bayesian approach proposed in this paper assigns prior distributions to the covariance matrices under both hypotheses, as well as to the probability of channel occupancy. After discussing which priors should be used for this problem, in this section we perform exact Bayesian inference over a single sensing frame to derive the posteriors for the unknown parameters. Specifically, the posterior for the channel occupancy is the statistic used to decide whether the SU should transmit or not.

3.3.1 Prior distributions

Let us first introduce z_t as a binary hidden random variable that indicates whether a transmitter is present ($z_t = 1$) or not ($z_t = 0$). Let us also remind that all our information about $\mathbf{R}_t$ and $\mathbf{D}_t$ is that they are some unknown covariance matrices, respectively full Hermitian and diagonal. Following a proper Bayesian treatment, prior distributions on all the unknown parameters of the model (z_t, $\mathbf{R}_t$ and $\mathbf{D}_t$) must be placed. We will use the following:

$$p(z_t) = \text{Bernoulli}(z_t|\breve{\pi}_t) = \breve{\pi}_t^{z_t}(1-\breve{\pi}_t)^{1-z_t} \tag{3.5a}$$

$$p(\mathbf{R}_t) = \mathcal{CW}^{-1}(\mathbf{R}_t|\breve{n}_t, \breve{\mathbf{R}}_t) = \frac{|\breve{\mathbf{R}}_t|^{\frac{\breve{n}_t}{2}}|\mathbf{R}_t|^{-\frac{\breve{n}_t+L+1}{2}}\exp(-\frac{1}{2}\operatorname{trace}(\mathbf{R}_t^{-1}\breve{\mathbf{R}}_t))}{2^{\frac{\breve{n}_t L}{2}}\Gamma_L(\frac{\breve{n}_t}{2})} \tag{3.5b}$$

$$p(\mathbf{D}_t) = \mathcal{G}_L^{-1}(\mathbf{D}_t|\breve{m}_t, \breve{\mathbf{D}}_t) = \prod_{l=1}^{L}\mathcal{G}^{-1}([\mathbf{D}_t]_{ll}|\breve{m}_t/2, [\breve{\mathbf{D}}_t]_{ll}/2) \tag{3.5c}$$

$$= \frac{|\breve{\mathbf{D}}_t|^{\frac{\breve{m}_t}{2}}|\mathbf{D}_t|^{-\frac{\breve{m}_t+L+1}{2}}\exp(-\frac{1}{2}\operatorname{trace}(\mathbf{D}_t^{-1}\breve{\mathbf{D}}_t))}{2^{\frac{\breve{m}_t L}{2}}\Gamma^L(\frac{\breve{m}_t}{2})} \tag{3.5d}$$

where we have included the definitions of the Bernoulli distribution, the complex inverse-Wishart ($\mathcal{CW}^{-1}$) and the product of L independent inverse-gamma ($\mathcal{G}_L^{-1}$). Note the difference between $\Gamma_L(\cdot)$ (used to denote the multivariate gamma function) and $\Gamma^L(\cdot)$ (the standard gamma function raised to the L-th power). We denote the parameters of the prior distributions as $\breve{\pi}_t, \breve{n}_t, \breve{\mathbf{R}}_t, \breve{m}_t$ and $\breve{\mathbf{D}}_t$. When the SU starts sensing the environment (i.e., at $t = 0$), the priors should reflect our lack of knowledge about the sensed environment and, in this sense, they should be as uninformative as possible. In the Bayesian literature, typically Jeffreys uninformative priors [Jeffreys, 2012] are used because they are invariant under reparametrizations (unlike a uniform prior). For our problem, and adopting a more practical point of view, starting with uninformative priors at $t = 0$ amounts to choosing small initial parameters for the prior distributions (and 0.5 for the probability of channel occupancy). For instance, it can be proved the if $\breve{m}$ and $\breve{\mathbf{D}}$ tend to zero the product of L independent inverse-gamma becomes Jeffreys' prior. Therefore, we simply chose small values for the prior parameters when the learning procedure starts. After that, the prior parameters are adapted and learnt over time as new sensing frames are acquired according to the mechanism that will be described in Section 3.4. For a more

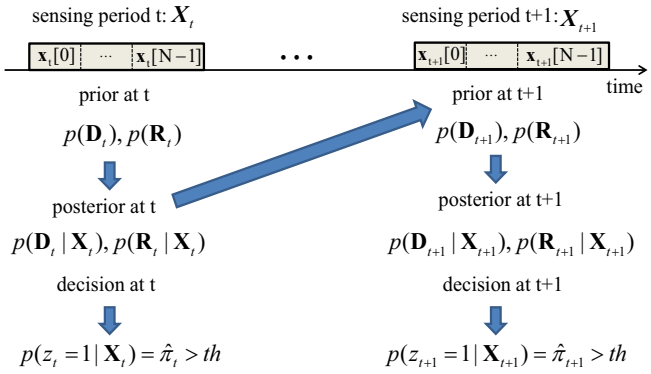


Figure 3.2: A Bayesian framework for spectrum sensing: the posteriors obtained after processing a sensing frame are employed as priors for the next sensing frame.

detailed discussion on the priors to be used for the multivariate Gaussian model the reader is referred to [Sun and Berger, 2006].

The main argument for the choice of these priors is analytical tractability: The complex inverse-Wishart distribution placed on $\mathbf{R}_t$, and the product of univariate inverse-gamma distributions placed on $\mathbf{D}_t$ are the conjugate priors for the distribution of full-rank covariance matrices and diagonal covariance matrices, respectively, when the observations follow a complex multivariate Gaussian distribution. As we will see in the next subsection, these conjugate priors allow us to exactly perform a Bayesian inference which is very convenient to avoid resorting to numerical integration methods. Nevertheless, let us also mention that several wireless standards use orthogonal frequency division multiplexing (OFDM) signals, which are typically modeled as zero-mean circular complex Gaussian signals, as well as MIMO technologies. For these practical scenarios, the signal model given in Eq. (4.1) and a flat prior for the covariance matrix reflecting our initial lack of knowledge about the main statistical parameters seems to be suitable.

3.3.2 Exact posterior distribution of z_t, $\mathbf{R}_t$ and $\mathbf{D}_t$

Since the noise is assumed to be Gaussian, the likelihoods of $p(\mathbf{X}_t|z_t = 0, \mathbf{D}_t)$ and $p(\mathbf{X}_t|z_t = 1, \mathbf{R}_t)$ can be written as

$$p(\mathbf{X}_t|z_t = 0, \mathbf{D}_t) = \prod_{n=1}^{N} \mathcal{CN}(\mathbf{x}[n]|z_t = 0, \mathbf{D}_t), \tag{3.6a}$$

$$p(\mathbf{X}_t|z_t = 1, \mathbf{R}_t) = \prod_{n=1}^{N} \mathcal{CN}(\mathbf{x}[n]|z_t = 1, \mathbf{R}_t). \tag{3.6b}$$

Given the hidden variable, z_t, priors are conjugate and therefore the posterior distributions (conditioned on the channel state) have the same form as the prior, but

with different parameters. For instance, we have

$$p(\mathbf{R}_t|\mathbf{X}_t, z_t = 1) = \mathcal{CW}^{-1}(\mathbf{R}_t|\hat{n}_t, \hat{\mathbf{R}}_t) \tag{3.7a}$$
$$p(\mathbf{D}_t|\mathbf{X}_t, z_t = 0) = \mathcal{G}_L^{-1}(\mathbf{D}_t|\hat{m}_t, \hat{\mathbf{D}}_t). \tag{3.7b}$$

where the posterior parameters, which clearly depend on the observed data $\mathbf{X}_t$, are given by

$$\hat{n}_t = \check{n}_t + N \tag{3.8a}$$
$$\hat{\mathbf{R}}_t = \check{\mathbf{R}}_t + \mathbf{S}_t \tag{3.8b}$$
$$\hat{m}_t = \check{m}_t + N \tag{3.8c}$$
$$\hat{\mathbf{D}}_t = \check{\mathbf{D}}_t + \mathrm{diag}(\mathbf{S}_t). \tag{3.8d}$$

When z_t is marginalized (by a direct application of the total probability theorem), each unconditional posterior becomes a convex combination of the posteriors for each hypothesis, yielding

$$p(\mathbf{R}_t|\mathbf{X}_t) = \hat{\pi}_t \mathcal{CW}^{-1}(\mathbf{R}_t|\hat{n}_t, \hat{\mathbf{R}}_t) + (1 - \hat{\pi}_t)\mathcal{CW}^{-1}(\mathbf{R}_t|\check{n}_t, \check{\mathbf{R}}_t) \tag{3.9a}$$
$$p(\mathbf{D}_t|\mathbf{X}_t) = \hat{\pi}_t \mathcal{G}_L^{-1}(\mathbf{D}_t|\check{m}_t, \check{\mathbf{D}}_t) + (1 - \hat{\pi}_t)\mathcal{G}_L^{-1}(\mathbf{D}_t|\hat{m}_t, \hat{\mathbf{D}}_t), \tag{3.9b}$$
$$p(z_t|\mathbf{X}_t) = \mathrm{Bernoulli}(z_t|\hat{\pi}_t) \tag{3.9c}$$

where $\hat{\pi}_t$ is given by

$$\hat{\pi}_t = \frac{p(\mathbf{X}_t|z_t = 1)p(z_t = 1)}{p(\mathbf{X}_t|z_t = 1)p(z_t = 1) + p(\mathbf{X}_t|z_t = 0)p(z_t = 0)}. \tag{3.10}$$

Recall that in (3.9a) and (3.9b) we use a breve ($\breve{}$) to denote the parameters of the *prior* distribution, whereas we use a hat ($\hat{}$) to denote the parameters of the *posterior* distribution. Finally, the marginal likelihood $p(\mathbf{X}_t|z_t)$ can be obtained analytically as

$$p(\mathbf{X}_t|z_t = 1) = \int p(\mathbf{X}_t|z_t = 1, \mathbf{R}_t)p(\mathbf{R}_t)\mathrm{d}\mathbf{R} = \frac{|\check{\mathbf{R}}_t|^{\frac{\check{n}_t}{2}}\Gamma_L(\frac{\hat{n}_t}{2})}{\pi^{\frac{NL}{2}}|\hat{\mathbf{R}}_t|^{\frac{\hat{n}_t}{2}}\Gamma_L(\frac{\check{n}_t}{2})} \tag{3.11a}$$

$$p(\mathbf{X}_t|z_t = 0) = \int p(\mathbf{X}_t|z_t = 0, \mathbf{D}_t)p(\mathbf{D}_t)\mathrm{d}\mathbf{D} = \frac{|\check{\mathbf{D}}_t|^{\frac{\check{m}_t}{2}}\Gamma(\frac{\hat{m}_t}{2})^L}{\pi^{\frac{NL}{2}}|\hat{\mathbf{D}}_t|^{\frac{\hat{m}_t}{2}}\Gamma(\frac{\check{m}_t}{2})^L}. \tag{3.11b}$$

After the posterior has been computed, the probability of a transmitter being present given observations $\mathbf{X}_t$ is simply $p(z_t = 1|\mathbf{X}_t) = \hat{\pi}_t$. Thus, we can occupy the channel when the collision probability $\hat{\pi}_t$ is below some desired threshold.

3.4 Bayesian inference over multiple frames

3.4.1 Learning from past sensing frames

The (unconditional) posteriors after processing the t-th frame summarize all statistical information observed so far. Therefore, a natural learning mechanism is to use

them as priors for the next sensing frame, as depicted in Fig. 4.1. More specifically, the proposed learning procedure is as follows: at each sensing frame the cognitive receiver updates the posterior distribution for $\mathbf{R}_t$ and $\mathbf{D}_t$ from priors existing at t and the likelihood obtained from $\mathbf{X}_t$; then, these posteriors become the priors to be used at the sensing period $t+1$. The procedure is started with uninformative priors at $t=0$.

A problem with a direct application of this idea is that, after applying Bayesian inference, the posterior distributions for $\mathbf{R}_t$ and $\mathbf{D}_t$ are convex combinations of the posteriors under each hypotheses, see Eqs. (3.9a) and (3.9b); and therefore the posterior does not belong to the same family distribution of the prior. For instance, the prior for $\mathbf{R}_t$ is a complex inverse-Wishart and the posterior is a linear combination of two complex inverse-Wisharts. To keep the process simple and scalable, it would be convenient to find an approximation of the posteriors within the family of each respective prior. In the next subsections we describe two possible approximations that can be applied to this end.

Thresholding-based approximation

A simple approximation to the posterior that falls within the same family as the prior can be obtained by truncating $\hat{\pi}_t$ to either 0 or 1, whichever it is closer. When this is done, Eq. (3.9a) and (3.9b) directly yields a posterior in the same family as the prior. In that case, when $\mathcal{H}_1$ is more probable, the posterior is obtained by performing only updates (4.6a) and (4.6b), whereas in the opposite case, only updates (4.6c) and (4.6d) are needed.

Kullback-Leibler approximation

A more rigorous approach is to find the approximation of the posteriors within the family of the priors that minimize the Kullback-Leibler distance. More precisely, the exact posteriors (reproduced here for convenience) are given by

$$p(\mathbf{R}_t|\mathbf{X}_t) = \hat{\pi}_t \mathcal{CW}^{-1}(\mathbf{R}_t|\hat{n}_t, \hat{\mathbf{R}}_t) + (1-\hat{\pi}_t)\mathcal{CW}^{-1}(\mathbf{R}_t|\check{n}_t, \check{\mathbf{R}}_t),$$
$$p(\mathbf{D}_t|\mathbf{X}_t) = \hat{\pi}_t \mathcal{G}_L^{-1}(\mathbf{D}_t|\check{m}_t, \check{\mathbf{D}}_t) + (1-\hat{\pi}_t)\mathcal{G}_L^{-1}(\mathbf{D}_t|\hat{m}_t, \hat{\mathbf{D}}_t).$$

Our problem consists in finding the approximations of these posteriors

$$q(\mathbf{R}_t|\mathbf{X}_t) = \mathcal{CW}^{-1}(\mathbf{R}_t|\tilde{n}_t, \tilde{\mathbf{R}}_t), \tag{3.13a}$$
$$q(\mathbf{D}_t|\mathbf{X}_t) = \mathcal{G}_L^{-1}(\mathbf{D}_t|\tilde{m}_t, \tilde{\mathbf{D}}_t), \tag{3.13b}$$

that minimize the Kullback-Leibler (KL) divergence. Therefore, we have to solve the following optimization problems

$$\{\tilde{m}_t, \tilde{\mathbf{D}}_t\} = \underset{\tilde{m}_t, \tilde{\mathbf{D}}_t}{\operatorname{argmin}} \ \mathrm{KL}(p(\mathbf{D}_t|\mathbf{X}_t)||q(\mathbf{D}_t|\mathbf{X}_t)) \tag{3.14a}$$
$$\{\tilde{n}_t, \tilde{\mathbf{R}}_t\} = \underset{\tilde{n}_t, \tilde{\mathbf{R}}_t}{\operatorname{argmin}} \ \mathrm{KL}(p(\mathbf{R}_t|\mathbf{X}_t)||q(\mathbf{R}_t|\mathbf{X}_t)). \tag{3.14b}$$

Fortunately, each of these minimization problems can be solved analytically except for a line search. The details of the derivation are relegated to 6.2, in the following we summarize the solution. In order to find $\{\tilde{m}_t, \tilde{\mathbf{D}}_t\}$ and $\{\tilde{n}_t, \tilde{\mathbf{R}}_t\}$, we first compute the following auxiliary quantities

$$\mathbf{K}_D = \hat{\pi}_t \check{m}_t \check{\mathbf{D}}_t^{-1} + (1-\hat{\pi}_t)\hat{m}_t \hat{\mathbf{D}}_t^{-1} \tag{3.15a}$$

$$\mathbf{K}_R = \hat{\pi}_t \hat{n}_t \hat{\mathbf{R}}_t^{-1} + (1-\hat{\pi}_t)\check{n}_t \check{\mathbf{R}}_t^{-1} \tag{3.15b}$$

$$k_D = -\ln|\mathbf{K}_D| + \hat{\pi}_t\left(L\psi\left(\frac{\check{m}_t}{2}\right) - \ln|\check{\mathbf{D}}_t|\right) + (1-\hat{\pi})\left(L\psi\left(\frac{\hat{m}_t}{2}\right) - \ln|\hat{\mathbf{D}}_t|\right) \tag{3.15c}$$

$$k_R = -\ln|\mathbf{K}_R| + \hat{\pi}_t\left(\psi_L\left(\frac{\hat{n}_t}{2}\right) - \ln|\hat{\mathbf{R}}_t|\right) + (1-\hat{\pi})\left(\psi_L\left(\frac{\check{n}_t}{2}\right) - \ln|\check{\mathbf{R}}_t|\right) \tag{3.15d}$$

where $\psi(\cdot)$ is the digamma function, and $\psi_L(\cdot) = \sum_{l=1}^{L}\psi(\cdot + (1-l)/2)$ defines the multivariate digamma function.

We then have to solve the following non-linear equations using, for instance, a few iterations of the Newton-Raphson method

$$k_D + L\ln(\tilde{m}_t) - L\psi\left(\frac{\tilde{m}_t}{2}\right) = 0, \tag{3.16a}$$

$$k_R + L\ln(\tilde{n}_t) - \psi_L\left(\frac{\tilde{n}_t}{2}\right) = 0. \tag{3.16b}$$

Finally, the covariance matrices $\tilde{\mathbf{D}}_t$ and $\tilde{\mathbf{R}}_t$ for the best approximation according to the KL distance are given by $\tilde{m}_t\mathbf{K}_\mathbf{D}^{-1}$ and $\tilde{n}_t\mathbf{K}_\mathbf{R}^{-1}$, respectively. These values are taken as the new parameters of the posterior distributions, that is

$$\tilde{\mathbf{D}}_t \to \hat{\mathbf{D}}_t, \tag{3.17a}$$

$$\tilde{\mathbf{R}}_t \to \hat{\mathbf{R}}_t, \tag{3.17b}$$

$$\tilde{n}_t \to \hat{n}_t, \tag{3.17c}$$

$$\tilde{n}_t \to \hat{m}_t. \tag{3.17d}$$

3.4.2 Forgetting in non-stationary environments

Since the channel may vary between consecutive frames, it is interesting to introduce a mechanism within the Bayesian framework to forget past data and hence be able to operate in a non-stationary environment. We assume here that no additional knowledge about the dynamical evolution of the channel, PU spectrum usage pattern or noise statistics is available. Therefore, we resort to the idea of Bayesian λ-forgetting [Kulhavý and Zarrop, 1993] that allows to forget in a principled manner with minimal assumptions. The basic idea of Bayesian forgetting is to use as prior distributions for frame $t+1$ a "smoothed" version of the posterior distributions obtained after processing frame t and the original prior distributions for $\mathbf{R}_t$ and $\mathbf{D}_t$ given by (3.5), i.e.,

$$p(\mathbf{D}_{t+1}|\mathbf{X}_t) \propto p(\mathbf{D}_t|\mathbf{X}_t)^{\lambda} p(\mathbf{D}_0)^{1-\lambda}, \tag{3.18a}$$

$$p(\mathbf{R}_{t+1}|\mathbf{X}_t) \propto p(\mathbf{R}_t|\mathbf{X}_t)^{\lambda} p(\mathbf{R}_0)^{1-\lambda}. \tag{3.18b}$$

Observe that according to this definition, when $\lambda = 0$, all the information obtained from previous data is forgotten and the process considers each frame independently (as the GLRT does), which is reasonable if abrupt changes occur in $\mathbf{R}_t$ and $\mathbf{D}_t$ between frames. When $\lambda = 1$, no forgetting occurs and the new posterior corresponds to the standard Bayesian posterior when $\mathbf{D}_t$ and $\mathbf{R}_t$ are constant across frames $\mathbf{D}_t = \mathbf{D}$, $\mathbf{R}_t = \mathbf{R}$ $\forall t$, which is reasonable under stationary conditions. Values of $\lambda \in [0,1]$ are therefore appropriate to model different evolution speeds in the channel, without having to define a concrete dynamical model. In another perspective, Eqs. (3.18) represent a change of the posterior in the direction of the prior: this has also been named as "back-to-the-prior" forgetting in [Van Vaerenbergh et al., 2012, Perez-Cruz et al., 2013] and a method for the selection of the forgetting value can be found in [Vaerenbergh et al., 2012].

With this forgetting step, the parameters of the prior distributions to be used for Bayesian inference at $t+1$ are given by

$$\check{n}_{t+1} = \lambda \hat{n}_t + (1-\lambda)\check{n}_0 \tag{3.19a}$$

$$\check{\mathbf{R}}_{t+1} = \lambda \hat{\mathbf{R}}_t + (1-\lambda)\check{\mathbf{R}}_0 \tag{3.19b}$$

$$\check{m}_{t+1} = \lambda \hat{m}_t + (1-\lambda)\check{m}_0 \tag{3.19c}$$

$$\check{\mathbf{D}}_{t+1} = \lambda \hat{\mathbf{D}}_t + (1-\lambda)\check{\mathbf{D}}_0. \tag{3.19d}$$

3.4.3 The proposed algorithm

The whole process is summarized in Algorithm 3.1. Since the algorithm only requires updating and storing $\hat{\mathbf{R}}_t$, $\hat{n}_t$, $\hat{\mathbf{D}}_t$, $\hat{m}_t$ from one frame to the next, it requires a fixed amount of memory and computation per sensing frame, which is $\mathcal{O}(L^2)$.

Algorithm 3.1 Online Bayesian Multiantenna Sensing

1: Initialize Parameters: λ, $\check{\mathbf{R}}_0$, $\check{n}_0$, $\check{\mathbf{D}}_0$, $\check{m}_0$
2: **for** Frame $t = 1, 2, \ldots$ **do**
3: Sense the medium N times through L antennas to get $\mathbf{X}_t$
4: Exact posterior: Compute $\hat{\mathbf{R}}_t$, $\hat{n}_t$, $\hat{\mathbf{D}}_t$, $\hat{m}_t$ and $\hat{\pi}_t$ using (4.6) and (3.11)
5: Take a decision about the channel occupancy based on $\hat{\pi}_t$, which is the probability of a PU being present at t.
6: Compute the approximated posterior parameters using KL minimization or thresholding
7: Forget: Compute $\check{\mathbf{R}}_t$, $\check{n}_t$, $\check{\mathbf{D}}_t$, $\check{m}_t$ using Eqs. (3.19)
8: **end for**

Let us notice that our Bayesian detector minimizes the probability of error just by deciding that the channel is busy when $p(z_t = 1|\mathbf{X}_t)$ is greater than $p(z_t = 0|\mathbf{X}_t)$, or,

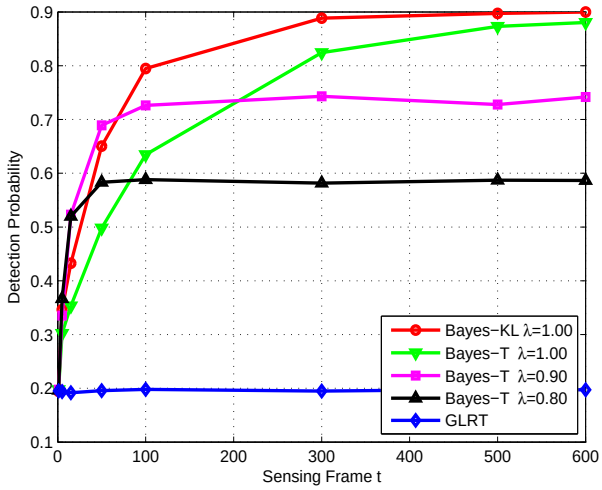


Figure 3.3: P_D for the Bayesian detector (Bayes-KL and Bayes-T) and the GLRT vs. the number of sensing frames in a time-invariant channel, $L = 5$, $P = 5$, $N = 50$, SNR $= -8$ dB and $P_{FA} = 0.1$.

in other words, when $p(z_t = 1|\mathbf{X}_t) > 0.5$. Notice also that, when the SU decides to transmit, the posterior probability for the channel occupancy can be translated into an "instantaneous" estimate of the false alarm probability. For instance, if $p(z_t = 1|\mathbf{X}_t) = 0.2$ and the SU decides to occupy the channel, the estimated probability of collision with the PU (i.e., a false alarm) would also be 0.2. Then, if we set the threshold $\eta = 0.2$ we ensure that the instantaneous[2] $P_{FA} < 0.2$. We consider this as an additional advantage of the proposed Bayesian procedure, for which we can easily identify the desired operation point (i.e, the threshold for the posterior probability of channel occupancy) just by assigning costs to the different decisions and setting the threshold accordingly.

Finally, we would like to stress that, for each new sensing frame, the proposed Bayesian scheme always updates the parameters of the posterior following the steps in Algorithm 3.1. Since we intend to operate in a probably changing environment, the sensing procedure must be continuously applied and we do not use any stopping rule.

3.5 Simulation results

In this section, we compare the performance of the proposed Bayesian detector with that of a GLRT-based detector (given by (3.4)) in different environments by means of Monte Carlo simulations. Unless otherwise stated, we assume a probability of channel occupancy given by $\breve{\pi}_t = 0.5$, a primary transmitter with $P = 5$ antennas and a secondary cognitive receiver with $L = 5$ antennas. The MIMO channel matrix

[2]This can be seen as a worst-case value, since $p(z_t = 1|\mathbf{X}_t) = 0.2$ does not necessarily mean that 20 % of the times the PU is active in this scenario, it is just our estimate of the posterior probability.

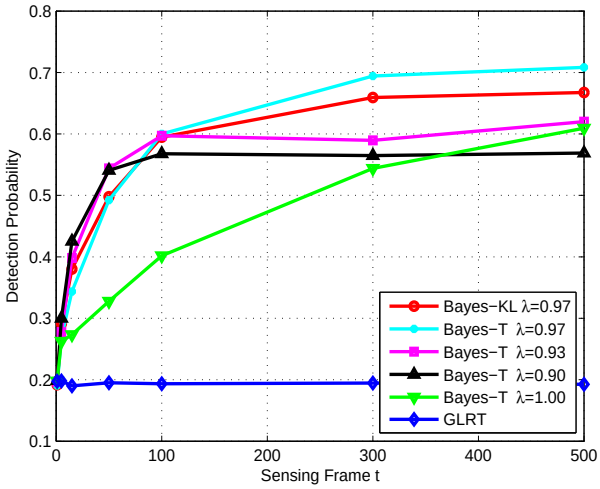


Figure 3.4: P_D for the Bayesian detector (Bayes-KL and Bayes-T) and the GLRT vs. the number of sensing frames in a slowly time-varying channel. $L = 5$, $P = 5$, $N = 50$, SNR $= -8$ dB, $P_{FA} = 0.1$ and $\lambda_{ch} = 0.9$.

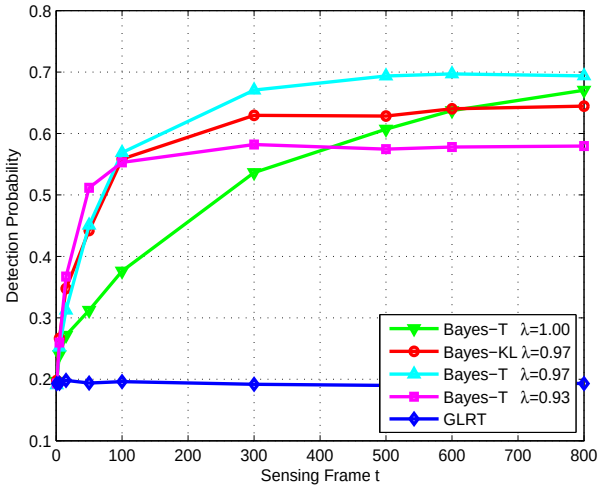


Figure 3.5: P_D for the Bayesian detector (Bayes-KL and Bayes-T) and the GLRT vs. the number of sensing frames in a fast time-varying channel. $L = 5$, $P = 5$, $N = 50$, SNR $= -8$ dB, $P_{FA} = 0.1$, and $\lambda_{ch} = 0.1$.

is assumed to be constant during the t-th sensing frame with i.i.d. entries distributed as $\mathcal{CN}(0,1)$. On the other hand, the channel evolves from frame to frame as $\mathbf{H}_{t+1} = \lambda_{ch}\mathbf{H}_t + (1-\lambda_{ch})\mathbf{P}_{t+1}$ [Baddour and Beaulieu, 2005], with $0 \leq \lambda_{ch} \leq 1$, and $\mathbf{P}_{t+1}$ a complex Gaussian noise matrix also with i.i.d. entries distributed as $\mathcal{CN}(0,1)$. For $\lambda_{ch} = 1$ we have a stationary channel, whereas for $\lambda_{ch} = 0$ it changes independently from frame to frame as a block fading channel [Goldsmith, 2005] [Chapter 4, Section 4.2].

3.5.1 P_D versus number of sensing frames

In this subsection, we study how the performance of the Bayesian detector evolves over time in stationary, slowly time-varying and fast time-varying environments. We start at $t = 0$ with an uninformative prior (we use small values for the parameters of the prior distributions) and then after each sensing frame we update the posterior (learning step), approximate the posterior using either truncation (denoted as Bayes-T in the plots) or KL minimization (denoted as Bayes-KL) and finally forget moving the approximated posterior towards the original uninformative prior with a forgetting factor λ. As a figure of merit we plot the detection probability $P_D = P(\hat{\pi}_t > \eta|\mathcal{H}_1)$ versus the number of sensing frames, where η is the threshold. We consider a fixed false alarm probability of $P_{FA} = 0.1$, and in each sensing frame the number of observations is $N = 50$[3]. For comparison we include the results obtained with the GLRT. In all examples we use a a signal-to-noise ratio SNR $= -8$ dB.

Stationary channel

We first consider a static scenario for which the channel remains constant over all sensing frames (i.e., $\lambda_{ch} = 1$). The results in Fig. 3.3 show that in this scenario, after just a few sensing frames, the Bayesian multiantenna detector provides a much higher P_D than the GLRT for different values of the forgetting parameter λ. After observing a sufficient number of frames the best results are obtained when using $\lambda = 1$ (which means no forgetting at all), as could be expected for this static environment. Interestingly, however, to forget a little ($\lambda = 0.9$) can be beneficial during the first sensing frames. This is explained because during the first sensing frames detection errors are more likely to occur and, consequently, the parameters of the posterior are not updated correctly. In this situation, it would be better not to trust so much on the

[3]The parameters in our simulations were not chosen to target any particular application or standard, since existing standards are mainly focused on static scenarios. In static scenarios, the sensing frames are typically much longer and the SNRs are also lower than those considered in our paper. As an example, the spectrum sensing requirements of the IEEE 802.22 wireless regional network (WRAN) standard 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 [Chen et al., 2007],[Cordeiro et al., 2007]. In our paper, we aimed at non-stationary environments and consequently considered much shorter (and frequent) sensing periods ($N = 50$ samples) in which the block-fading model is assumed to be valid.

observed data and apply the forgetting step. Finally, we also compare in the figure the performance of the two approximations of the posterior proposed in the paper. As expected, the KL-based approximation provides a better performance at the cost of a higher computational complexity.

Slowly time-varying channel

We now consider a non-stationary environment created by a slowly time-varying channel with $\lambda_{ch} = 0.9$. The results in Fig. 3.4 show again that the Bayesian detector outperforms the GLRT after just a few sensing frames. The optimal value of the forgetting factor, obtained by numerical results, for this scenario seems to be close to $\lambda = 0.97$, [4] and using a value of $\lambda = 1$ (no forgetting) strongly affects the performance. It is also clear that the convergence now is slower, since it takes more sensing frames to effectively learn and track the covariance matrices under both hypotheses. Finally, regarding the impact of the posterior approximation on the performance of the detector, we observe that in non-stationary environments it is better to use a thresholding-based approximation. In non-stationary environments, the importance of obtaining at each step an accurate approximation diminishes since the performance is limited by the variations observed from frame to frame.

Fast time-varying channel

In Fig. 3.5, we finally consider the case of a fast time-varying environment with $\lambda_{ch} = 0.1$. Remarkably, even in this highly non-stationary environment, the Bayesian detector outperforms the GLRT detector. This improvement can be attributed to the fact that the covariance matrix under $\mathcal{H}_0$ remains almost constant from frame to frame (only the channel changes) and, therefore, it can be learnt by the Bayesian detector. This improved estimate of the noise-only covariance matrix translates into a better P_D in comparison to the GLRT. For the reasons explained before, the simple truncation of the posterior performs better than the most accurate KL-based approximation in this rapidly varying scenario.

In order to provide a more complete understanding of the proposed method, we have repeated the experiments for $P_{FA} = 0.01$. The new results for stationary, slowly time-varying and fast time-varying channels are depicted in Fig. 3.6, Fig. 3.7, and Fig. 3.8, respectively.

3.5.2 Receiver Operating Characteristic

In this subsection, we obtain the receiver operating characteristic (ROC) curve of the detector after convergence, i.e., after processing a sufficient number of frames to reach the steady state. We study the ROC curve for different number of observations

[4] Let us point out that this value has no direct relationship with λ_{ch}, since λ is a design parameter and λ_{ch} is a characteristic of the channel.

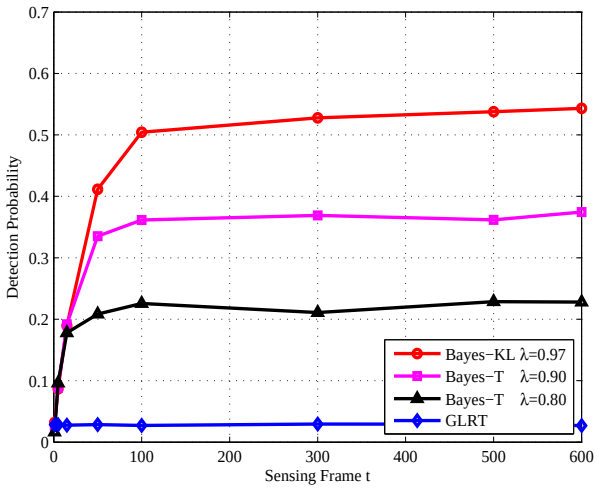


Figure 3.6: P_D for the Bayesian detector (Bayes-KL and Bayes-T) and the GLRT vs. the number of sensing frames in a time-invariant channel. $L = 5$, $P = 5$, $N = 50$, $SNR = -8$ dB and $P_{FA} = 0.01$.

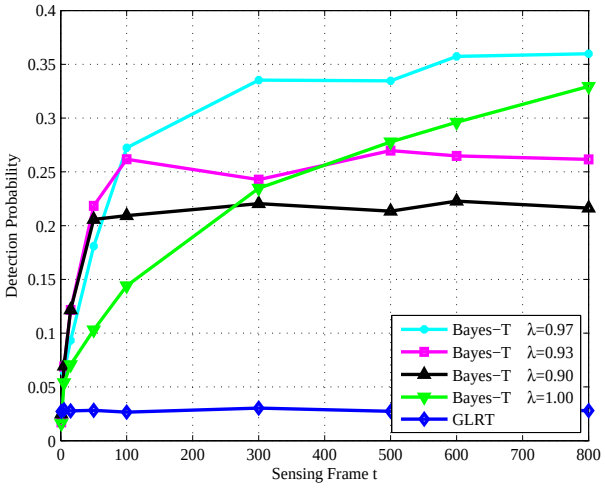


Figure 3.7: P_D for the Bayesian detector and the GLRT vs. the number of sensing frames in a slowly time-variant channel. $L = 5$, $P = 5$, $N = 50$, $SNR = -8$ dB, $P_{FA} = 0.01$ and $\lambda_{ch} = 0.9$.

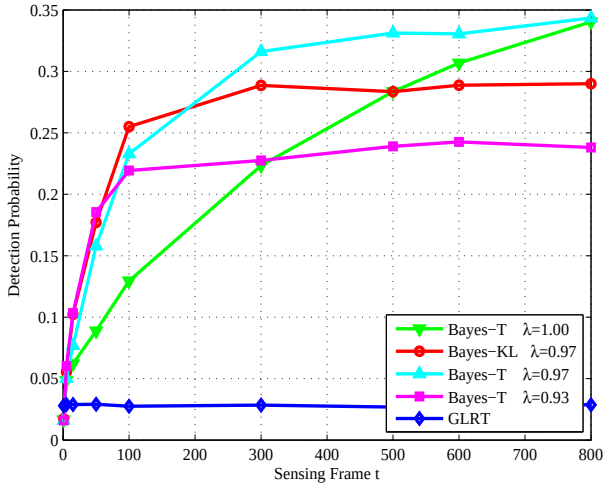


Figure 3.8: P_D for the Bayesian detector (Bayes-KL and Bayes-T) and the GLRT vs. the number of sensing frames in a fast time-variant channel. $L = 5$, $P = 5$, $N = 50$, $SNR = -8$ dB, $P_{FA} = 0.01$ and $\lambda_{ch} = 0.1$.

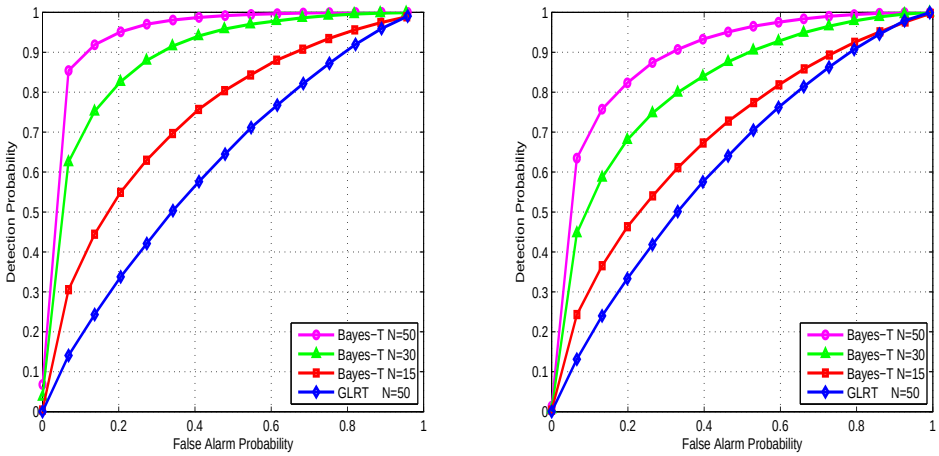
per sensing frame ($N = \{50, 30, 15\}$). As we haven seen previously, the approximation based on the KL distance is computationally more costly and does not provide any significant improvement in time-varying environments. Therefore, we concentrate on the results provided by the thresholding-based approximation referred to as Bayes-T, which seems to be better under non-stationary environments.

Fig. A.3.9(a) shows the results for the stationary channel, with $\lambda = 1$ and SNR $= -8$ dB. As we see, in steady-state the proposed Bayesian detector with only 15 snapshots per sensing frame outperforms the GLRT with 50 snapshots, which means a reduction of more than three times in the sensing time per frame. Fig. A.3.9(b) and Fig. A.3.9(c) show the results for slowly (using $\lambda_{ch} = 0.90$) and fast (using $\lambda_{ch} = 0.10$) time-varying environments, respectively; from which similar conclusions can be drawn.

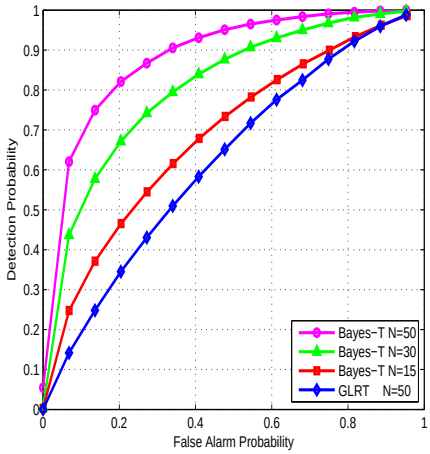
Finally, we fix the detection threshold and evaluate the probability of detection, P_D, and false alarm, P_{FA}, for different SNRs and in different scenarios. The number of samples per sensing frame is fixed to $N = 50$ and the rest of parameters is the same as in the previous section. The results are shown in Fig. A.3.10(a), Fig. A.3.10(b) and Fig. A.3.10(c) for stationary, slowly and fast time-varying scenarios.

3.5.3 Detection performance for a rank-P PU

In Fig. 3.11, we compare the probability of missed detection, P_M, for the Bayesian and GLRT detectors when the spatial rank of the PU signal varies. For the GLRT detector we have used the results in [Ramirez et al., 2011]. We consider an scenario with $L = 6$, $N = 50$, $P_{FA} = 0.1$, $P = \{1, ..., 6\}$ and SNR $= -8$ dB. In general, the performance of both detectors degrade for an increasing P, since as P increases



(a) (b)



(c)

Figure 3.9: ROC curves for the Bayesian and GLRT detector with $L = 5$, $P = 5$ and different number of snapshots per sensing frame. (a) Stationary channel with SNR $= -8$ dB and $\lambda = 1.0$. (b) Slowly time-varying channel with SNR $= -8$ dB, $\lambda = 0.97$ and $\lambda_{ch} = 0.90$. (c) Fast time-varying channel with SNR $= -8$ dB, $\lambda = 0.97$ and $\lambda_{ch} = 0.10$.

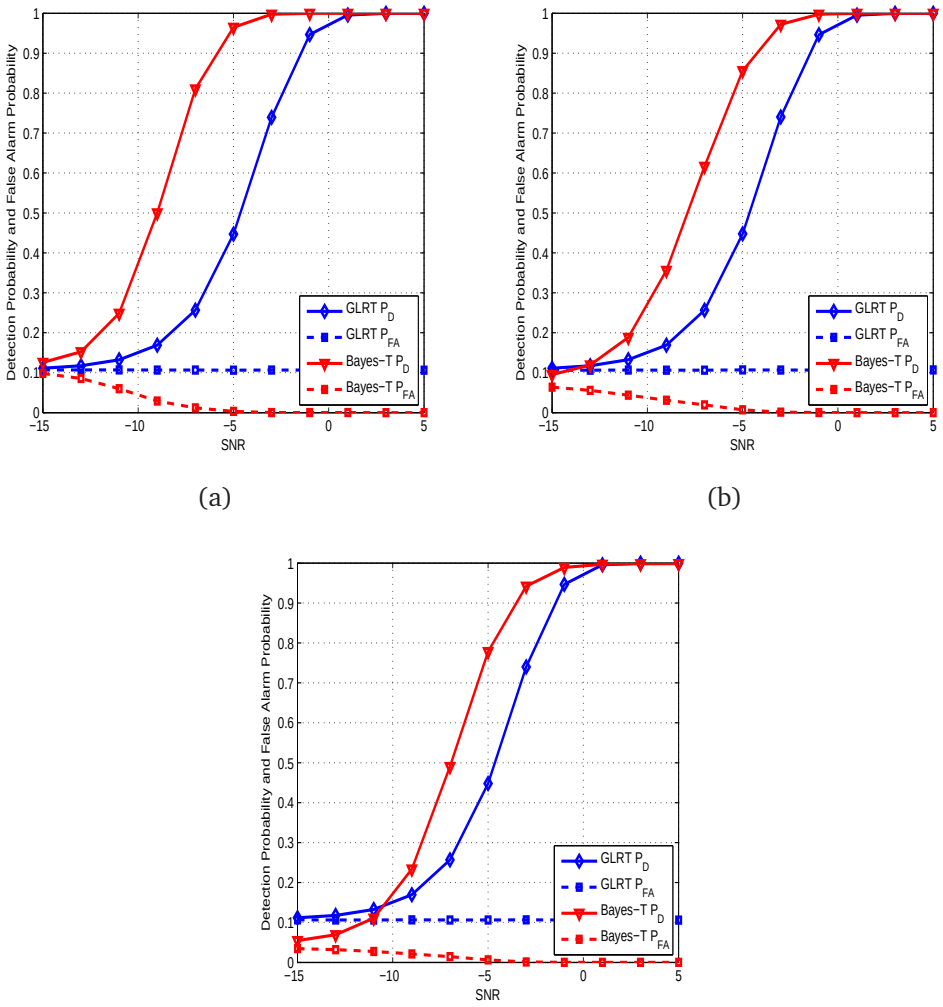


Figure 3.10: P_D and P_{FA} vs SNR with $L = 5$ antennas, and $N = 50$. (a) In a time-invariant channel ($\lambda = 1$ and $\lambda_{ch} = 1$). (b) In a slowly time-varying channel, ($\lambda = 0.97$ and $\lambda_{ch} = 0.95$). (c) In a fast time-varying channel ($\lambda = 0.95$ and $\lambda_{ch} = 0.10$).

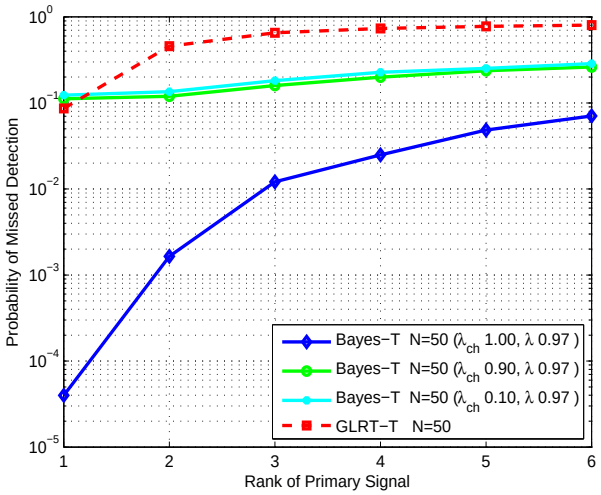


Figure 3.11: Probability of missed detection versus the rank of the primary signal P for both detectors with $L = 6$ antennas, $N = 50$, SNR $= -8$ dB and $P_{FA} = 0.1$. A static time-invariant channel with $\lambda = 1.0$ using a KL posterior approximation, a slowly time-varying channel with $\lambda = 0.97$ and $\lambda_{ch} = 0.90$, and a fast time-varying channel with $\lambda = 0.97$ and $\lambda_{ch} = 0.10$.

the covariance matrix under $\mathcal{H}_1$ has less structure to be exploited. Nevertheless, the Bayesian approach consistently provides better results than the GLRT, which validates again its ability to learn from the environment even when the actual model does not match exactly the presumed one.

3.6 Experimental Evaluation

In this section, we further validate the simulations by means of experimental measurements on a low-cost hardware cognitive platform. Specifically, our platform is composed of several N210 Universal Software Radio Peripheral (USRP) devices [Ett, 2014], each of them consists of a USRP motherboard and a Radio Frequency (RF) daughterboard (the XCVR2450 daughterboard based on a MAX2829 IC is able to cover ISM bands of 2.4GHz to 2.5GHz, and 4.9GHz to 5.8GHz). Basically, the motherboard consists of dual analog-to-digital converters (ADC) and digital-to-analog (DAC) converters connected to a Field Programmable Gain Array (FPGA). On the other hand, the daughterboard is a modular front-end used for analog operations such as up/down conversion.

In order to implement a multiantenna cognitive node, the N210 USRP includes a specific expansion port that allows coherent synchronization of two USRP2 units, as it is depicted in Fig. 4.2. Since the same clock (oscillators) and time reference are shared, both USRP nodes can start transmitting/receiving at the same time, thus avoiding any synchronization problem.

We have considered a simple scenario where a single-antenna PU access the channel according to a predefined pattern and a cognitive receiver with two antennas

senses periodically the medium and applies different detection procedures. The platform is controlled from a central PC, which allows us to define a pattern of spectrum occupancy as well as the sensing periods (see Fig. 4.4). Therefore, at each sensing period we know exactly the true hypothesis and hence we can estimate P_D and P_{FA} for a given threshold.

The experiments were conducted in the laboratory of the Signal Processing Group at the University of Cantabria, with a clear line of sight (LOS) between the PU and the cognitive receiver in a rather static environment. The PU transmits an orthogonal frequency division multiplexing (OFDM) 802.11a signal with a rate of 9 Mbps using BPSK symbols at a carrier frequency of 5.6 GHz, although during the detection stage the modulation format is assumed to be unknown by the SU.

Previous works describing measurements carried out at 5 GHz band in the same scenario (see Gutierrez et al [Gutiérrez et al., 2011]) showed that the measured indoor channel presents long coherence times in comparison to time devoted to each experiment. Therefore, the measured scenario closely represents a stationary environment. For each experiment, the 2-antenna USRP node senses the spectrum over a period of several seconds and then a large amount of data is stored in a $L \times M$ matrix, where M is the total number of samples and L is the number of antennas. On the other hand, the activity of the PU user, which is emulated by another USRP node, is controlled and recorded by a central PC. The activity pattern of the PU is recorded simultaneously while the SU is sensing the spectrum.

Using these data, the ROC curve is computed after processing 80 sensing frames so that our Bayesian detector reaches its steady-state performance. Since we know the channel status under which each sensing frame was acquired, the P_{FA} and the P_D can be easily estimated for different thresholds (to this end we used 5000 sensing frames) and the results depicted in Fig. 3.14 were obtained. The noise variance at each receiver antenna was measured and found to be very similar for the 2 RF branches, thus the Sphericity and Hadamard detectors [Ramirez et al., 2011] provided almost indistinguishable results which were both labeled in Fig. 3.14 as GLRT detector.

For each experiment, the SNR is controlled by the transmitter power and measured from the received signal at baseband. For the example shown here the measured SNR was -7.3 dB, the number of samples acquired by the SU during each sensing frame was $N = 50$, the number of SU antennas is $L = 2$ and the number of PU antennas is $P = 1$. In Fig. 3.14, we compare the ROC obtained by the proposed Bayesian detector working with a forgetting factor of $\lambda = 0.99$ and the GLRT detector for this setup. This figures corroborates the validity of the simulations carried out in Section 5. For a full detailed description of the experimental evaluation, the reader is referred to [Manco-Vasquez et al., 2012a] and [Gutiérrez et al., 2012], where a procedure to emulate time-varying scenarios is also described.

3.7 Conclusion

We have derived a new Bayesian framework for the problem of multiantenna spectrum sensing. We assume that the observations follow a Gaussian distribution under

Figure 3.12: N210 Ettus devices with the XCVR2450 daugtherboard installed. A two-antenna cognitive receiver is composed of two N210 boards connected through a MIMO cable.

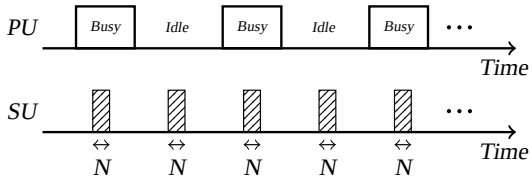


Figure 3.13: A PU transmits according to a preestablished sequence of states, and the SU senses periodically the wireless channel.

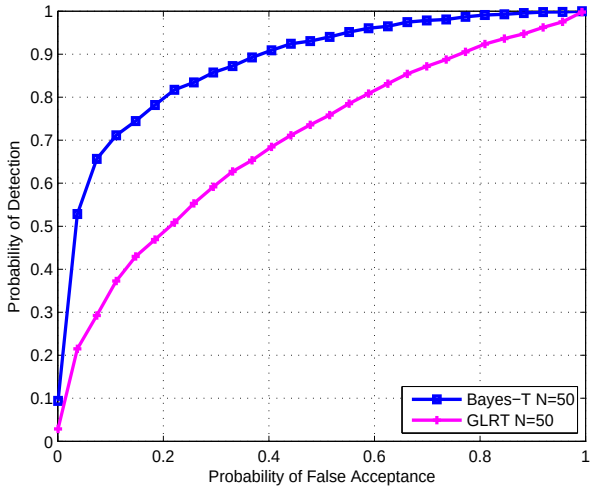


Figure 3.14: ROC curves for the Bayesian and GLRT detectors using the CR platform in a realistic indoor channel at 5,6 GHz.

both hypotheses, which allows us to choose conjugate priors and thereby to perform the exact Bayesian inference with closed-form expressions. Moreover, our Bayesian framework is able to exploit previous statistical information obtained from past sensing frames. To that end, we propose a forgetting mechanism where the posterior densities on the covariance matrices summarize this past information and the next Bayesian inference takes these posteriors as suitable priors. We evaluate the derived Bayesian framework in different scenarios, that is, stationary and non-stationary environments. The comparison between the Bayesian framework and GLRT detector under these scenarios as well as experimental evaluations shows that the Bayesian detector outperforms the GLRT. The most interesting findings are provided under a time-varying environment, where we showed that the Bayesian detector is able to efficiently learn the posterior.

3.8 Appendix: Derivation of KL approximation

In this appendix we find the pdf approximation that is closest in terms of the KL distance to the exact posterior. For notational simplicity, we omit the subindex t which refers to the sensing frame. We will only consider the approximation under $\mathcal{H}_1$, since the derivations under the null hypothesis are exactly the same. More precisely, under $\mathcal{H}_1$ the exact posterior is given by

$$p(\mathbf{R}|\mathbf{X}) = \hat{\pi}\mathcal{CW}^{-1}(\mathbf{R}|\hat{n},\hat{\mathbf{R}}) + (1-\hat{\pi})\mathcal{CW}^{-1}(\mathbf{R}|\check{n},\check{\mathbf{R}}), \tag{3.20}$$

and we want to find the approximation

$$q(\mathbf{R}|\mathbf{X}) = \mathcal{CW}^{-1}(\mathbf{R}|\tilde{n},\tilde{\mathbf{R}}), \tag{3.21}$$

that minimizes the KL divergence. That is, we want to solve

$$\{\tilde{n},\tilde{\mathbf{R}}\} = \underset{\tilde{n},\tilde{\mathbf{R}}}{\operatorname{argmin}}\ \mathrm{KL}(p(\mathbf{R}|\mathbf{X})||q(\mathbf{R}|\mathbf{X})) \tag{3.22}$$

where,

$$\mathrm{KL}(p(\mathbf{R}|\mathbf{X})||q(\mathbf{R}|\mathbf{X})) = \int p(\mathbf{R}|\mathbf{X})\ln(p(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R} - \int p(\mathbf{R}|\mathbf{X})\ln(q(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R}. \tag{3.23}$$

The first term in the right hand side of (3.23) is the negative of the differential entropy of the exact posterior and hence does not depend on $\{\tilde{n},\tilde{\mathbf{R}}\}$. Therefore, minimizing the KL is equivalent to solving the following maximization problem

$$\{\tilde{n},\tilde{\mathbf{R}}\} = \underset{\tilde{n},\tilde{\mathbf{R}}}{\operatorname{argmax}} \int p(\mathbf{R}|\mathbf{X})\ln(q(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R}. \tag{3.24}$$

Now, by substituting (3.20) and (3.21) into (3.24), we have

$$\int p(\mathbf{R}|\mathbf{X})\ln(q(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R} = \int \hat{\pi}\ln(\mathcal{CW}^{-1}(\mathbf{R}|\tilde{n},\tilde{\mathbf{R}}))\mathcal{CW}^{-1}(\mathbf{R}|\hat{n},\hat{\mathbf{R}})\mathrm{d}\mathbf{R}. + \int(1-\hat{\pi})\ln(\mathcal{CW}^{-1}(\mathbf{R}|\tilde{n},\tilde{\mathbf{R}}))\mathcal{CW}^{-1}(\mathbf{R}|\check{n},\check{\mathbf{R}})\mathrm{d}\mathbf{R}. \tag{3.25}$$

Integrating the first term on the right-hand side of Eq. (3.25) we get

$$\begin{aligned}
&\int \hat{\pi}\ln(\mathcal{CW}^{-1}(\mathbf{R}|\tilde{n},\tilde{\mathbf{R}}))\mathcal{CW}^{-1}(\mathbf{R}|\hat{n},\hat{\mathbf{R}})\mathrm{d}\mathbf{R}\\
&= \hat{\pi}\left[\int \ln\left(\frac{|\tilde{\mathbf{R}}|^{\frac{\tilde{n}}{2}}}{2^{-\frac{\tilde{n}L}{2}}\Gamma_L(\frac{\tilde{n}}{2})}\right)\mathcal{CW}^{-1}(\mathbf{R}|\hat{n},\hat{\mathbf{R}})\mathrm{d}\mathbf{R}\right] + \hat{\pi}\left[-\frac{\tilde{n}+L+1}{2}\int \ln|\mathbf{R}|\ \mathcal{CW}^{-1}(\mathbf{R}|\hat{n},\hat{\mathbf{R}})\mathrm{d}\mathbf{R}\right]\\
&\quad + \hat{\pi}\left[-\frac{1}{2}\int \mathrm{trace}(\mathbf{R}^{-1}\tilde{\mathbf{R}})\ \mathcal{CW}^{-1}(\mathbf{R}|\hat{n},\hat{\mathbf{R}})\mathrm{d}\mathbf{R}\right]\\
&= \hat{\pi}\left[\frac{\tilde{n}}{2}\ln|\tilde{\mathbf{R}}| - \frac{\tilde{n}}{2}L\ln(2) - \ln\Gamma_L\left(\frac{\tilde{n}}{2}\right)\right] + \hat{\pi}\left[-\frac{\tilde{n}+L+1}{2}\left(\ln|\hat{\mathbf{R}}| - L\ln(2) - \psi_L\left(\frac{\hat{n}}{2}\right)\right)\right]\\
&\quad + \hat{\pi}\left[-\frac{1}{2}\mathrm{trace}(\hat{n}\hat{\mathbf{R}}^{-1}\tilde{\mathbf{R}})\right]
\end{aligned} \tag{3.26}$$

where in the last step, we have used the fact that[5] $\mathbb{E}[\ln|\mathbf{R}|] = \ln|\hat{\mathbf{R}}| - L\ln(2) - \psi_L\left(\frac{\hat{n}}{2}\right)$ and $\mathbb{E}[\mathbf{R}^{-1}] = \hat{n}\hat{\mathbf{R}}^{-1}$. By the same procedure, the second term on the right-hand side of Eq. (3.25) is given by

$$\begin{aligned}
&\int(1-\hat{\pi})\ln(\mathcal{CW}^{-1}(\mathbf{R}|\tilde{n},\tilde{\mathbf{R}}))\mathcal{CW}^{-1}(\mathbf{R}|\check{n},\check{\mathbf{R}})\mathrm{d}\mathbf{R} \ =\\
&(1-\hat{\pi})\left[\frac{\tilde{n}}{2}\ln|\tilde{\mathbf{R}}| - \frac{\tilde{n}}{2}L\ln 2 - \ln\left(\Gamma_L\left(\frac{\tilde{n}}{2}\right)\right)\right] \ +\\
&(1-\hat{\pi})\left[-\frac{\tilde{n}+L+1}{2}\left(\ln|\check{\mathbf{R}}| - L\ln(2) - \psi_L\left(\frac{\check{n}}{2}\right)\right)\right] \ +\\
&(1-\hat{\pi})\left[-\frac{1}{2}\mathrm{trace}(\check{n}\check{\mathbf{R}}^{-1}\tilde{\mathbf{R}})\right]
\end{aligned}$$

Combining the two terms yields $\int p(\mathbf{R}|\mathbf{X})\ln(q(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R}$. In order to obtain the parameters that maximize this function, we have to take derivatives with respect to $\tilde{n}$ and $\tilde{\mathbf{R}}$ and equate them to zero. We first derive $\int p(\mathbf{R}|\mathbf{X})\ln(q(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R}$ with respect to $\tilde{\mathbf{R}}$, that is

$$\hat{\pi}\left[\frac{\tilde{n}}{2}\tilde{\mathbf{R}}^{-1} - \frac{\hat{n}}{2}\hat{\mathbf{R}}^{-1}\right] + (1-\hat{\pi})\left[\frac{\tilde{n}}{2}\tilde{\mathbf{R}}^{-1} - \frac{\check{n}}{2}\check{\mathbf{R}}^{-1}\right] = 0$$

[5] Recall that $\psi_L(a) = \frac{\partial}{\partial a}\ln(\Gamma_L(a)) = \sum_{l=1}^{L}\psi(a+(1-l)/2)$.

where we have applied the identities $\partial \ln|\mathbf{\Sigma_1}|/\partial\mathbf{\Sigma_1} = (\mathbf{\Sigma_1}^T)^{-1}$ and $\partial\,\mathrm{trace}(\mathbf{\Sigma_2\Sigma_1})/\partial\mathbf{\Sigma_1} = \mathbf{\Sigma_2}^T$. By defining $\mathbf{K}_R = \hat{\pi}\hat{n}\hat{\mathbf{R}}^{-1} + (1-\hat{\pi})\check{n}\check{\mathbf{R}}^{-1}$, it readily follows that $\tilde{\mathbf{R}} = \tilde{n}\mathbf{K}_R^{-1}$.

Now we take the derivative of $\int p(\mathbf{R}|\mathbf{X})\ln(q(\mathbf{R}|\mathbf{X}))\mathrm{d}\mathbf{R}$ with respect to $\tilde{n}$, which is given by

$$\hat{\pi}\left[\frac{1}{2}\ln|\tilde{\mathbf{R}}| - \frac{L}{2}\ln(2) - \frac{1}{2}\psi_L\left(\frac{\tilde{n}}{2}\right)\right] - \hat{\pi}\left[\frac{1}{2}\ln|\hat{\mathbf{R}}| - \frac{L}{2}\ln(2) - \frac{1}{2}\psi_L\left(\frac{\hat{n}}{2}\right)\right] +$$
$$(1-\hat{\pi})\left[\frac{1}{2}\ln|\tilde{\mathbf{R}}| - \frac{L}{2}\ln(2) - \frac{1}{2}\psi_L\left(\frac{\tilde{n}}{2}\right)\right] - (1-\hat{\pi})\left[\frac{1}{2}\ln|\check{\mathbf{R}}| - \frac{L}{2}\ln(2) - \frac{1}{2}\psi_L\left(\frac{\check{n}}{2}\right)\right]$$
$$= 0$$

Finally using $\tilde{\mathbf{R}} = \tilde{n}\mathbf{K}_R^{-1}$, and defining $k_R = -\ln|\mathbf{K}_R| + \hat{\pi}\left(\psi_L\left(\frac{\hat{n}}{2}\right) - \ln|\hat{\mathbf{R}}|\right) + (1-\hat{\pi})\left(\psi_L\left(\frac{\check{n}}{2}\right) - \ln|\check{\mathbf{R}}|\right)$, the non-linear equation $k_R + L\ln\tilde{n} - \psi_L\left(\frac{\tilde{n}}{2}\right) = 0$ is obtained. The same approach is followed to compute its counterpart under the hypothesis $\mathcal{H}_0$.

Chapter 4

Experimental Evaluation of the Proposed Bayesian Approach

Title: Experimental Evaluation of Multiantenna Spectrum Sensing Detectors using a Cognitive Radio Testbed

Authors: Julio Manco Vásquez, Jesus Gutierrez Teran, Jesus Perez Arriaga, Jesus Ibáñez, and Ignacio Santamaría

Affiliations: Dept. of Communications Engineering, Univ. of Cantabria, Spain.

Conference: IEEE International Symposium on Signals Systems and Electronics (ISSSE).

Experimental Evaluation of Multiantenna Spectrum Sensing Detectors using a Cognitive Radio Testbed

Abstract
Cognitive radio (CR) is a promising approach to improve the efficiency use of the wireless spectrum. One key element of this technology is spectrum sensing, which allows secondary users to detect the presence of licensed (primary) users. To this end, multiantenna spectrum sensing techniques have been proposed to detect the presence of a primary user based solely on the correlation structure of the signal received by a cognitive secondary receiver equipped with multiple antennas. Despite the numerous theoretical studies in the area of spectrum sensing, there exists a lack of experimental work evaluating the performance of these techniques in practice. In this paper, we test the performance of multiantenna Bayesian and generalized likelihood ratio test (GLRT) detectors on a cognitive radio platform. In comparison to one-shot GLRT detectors, the Bayesian detector is able to exploit past information from previous sensing periods, thus learning from the environment and improving its performance. Our cognitive platform is composed of Universal Software Radio Peripheral (USRP) nodes, that emulate the behavior of a single-antenna primary and a multiantenna cognitive receiver. Our measurements show that the Bayesian detector outperforms the GLRT detectors, in both stationary and non-stationary environments.

Keywords Cognitive Radio, Bayesian Detection, Multiantenna Spectrum Sensing, Cognitive Testbed, USRP.

4.1 Introduction

The availability of spectral resources is being limited by a growing demand of wireless services [Van de Beek et al., 2012]. Cognitive radio (CR) systems have been proposed to alleviate this limitation by achieving a more efficient use of the spectrum. In this context, a reliable spectrum sensing stage plays an important role, since it allows sharing the spectrum between legacy or primary users (PU) and non-legacy or secondary users (SU). By detecting holes in the spectrum, these gaps are filled with transmissions from SU, thus improving the usage of the frequency bands while avoiding interferences to licensed users.

Regarding the detection problem, some well-known approaches have been extensively studied, such as the energy-based detector, matched-filter detectors or the use of pilots [Yucek and Arslan, 2009, Cabric et al., 2006]. However, these well-know schemes also have some disadvantages, such as the requirement of a precise receiver calibration (knowledge of the noise variance), or the requirement of perfect synchronization. Recently, detectors employing multiple antennas have received increased attention because they do not require prior knowledge about the PU modulation format or the noise power, and are able to work with asynchronously sampled signals [Wang et al., 2010, Ramirez et al., 2011]. The vast majority of these approaches are one-shot detectors which are based on the GLRT. As an alternative to the GLRT approach, Bayesian detectors that exploit prior information about the environment have also been studied in [Manco-Vasquez et al., 2012b, Couillet and Debbah, 2010, Font-Segura and Wang, 2010].

In most of these works, the performance evaluation of the proposed spectrum sensing strategies is carried out by means of computer simulations. Recently, some experimental evaluations of spectrum sensing algorithms have been described in the literature. For instance, in [Grimm et al., 2011] an energy-based detector is evaluated in multiple bands by employing multiple antennas at the cognitive receiver, thus enhancing its performance by exploiting the spatial dimension. Also, in [Mate et al., 2011] a comparison performance was conducted between detection algorithms based on the eigenvalues of the received covariance matrix. Despite these few works, there still exist a need of experimental evaluations of spectrum sensing techniques in real scenarios using CR testbeds and platforms [Cabric and Brodersen, 2005, Mubaraq et al., 2004].

In an attempt to fill this gap, in this paper we experimentally evaluate the performance of a Bayesian detector, recently proposed in [Manco-Vasquez et al., 2012b], by means of experimental measurements on a low-cost hardware platform. More specifically, the platform is composed of USRP nodes using 5 GHz radios and allow us to emulate a scenario in which a single-antenna PU accesses the channel with a given probability of occupancy, and a secondary (cognitive) receiver, equipped with two antennas, senses the channel and performs Bayesian inference to learn the environment. The measurements were performed in an indoor scenario, for which long-coherence times are typically observed. To be able to assess the performance of the method in channels with higher mobility, we use a beamforming-based procedure at the PU to induce some pre-defined channel variability. Our experimental

results confirm the advantages of the Bayesian approach with respect to GLRT-based detectors.

The rest of the paper is organized as follows: In Section II, we give an overview of the proposed Bayesian spectrum sensing algorithm to be evaluated in our cognitive platform. The implemented testbed and the measurement setup are described in Section III and IV, respectively. We present the experimental results in Section V and finally, the paper concludes with a discussion of the results in Section VI.

4.2 Multiantenna Spectrum Sensing

By employing multiple antennas at the cognitive receiver, it is possible to fully exploit both the spatial and temporal structure of the received signals, and detect the presence of a PU without any knowledge about its signalling scheme or the noise variance [Manco-Vasquez et al., 2012b, Couillet and Debbah, 2010, Font-Segura and Wang, 2010]. In this work we consider a scenario with a single-antenna PU and a cognitive receiver that senses the wireless medium N times through L independent antennas. The received signal at the baseband is stacked in an observation matrix $\mathbf{X}_t = [\mathbf{x}_t[1], \ldots, \mathbf{x}_t[N]]$, where $\mathbf{x}_t[n] \in \mathbb{C}^L$ denotes the snapshot acquired at the n-th time instant within the t-th sensing period or frame (see Fig. 4.1).

The spectrum sensing problem can be formulated as the following binary hypothesis test

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}_t[n] = \mathbf{h}_t s_t[n] + \mathbf{v}_t[n], \\ \mathcal{H}_0 &: \mathbf{x}_t[n] = \mathbf{v}_t[n], \end{aligned} \tag{4.1}$$

where $\mathbf{x}_t[n]$ denotes the $L \times 1$ vector of the n-th sample of the receiver signal, $s_t[n] \in \mathbb{C}$ is the primary signal, $\mathbf{h}_t \in \mathbb{C}^{L\times 1}$ describes the (unknown) single-input multiple-output (SIMO) channel between the transmitter and the cognitive receiver, and $\mathbf{v}_t[n] \in \mathbb{C}^{L\times 1}$ is the additive white Gaussian noise. We assume that the vector-valued observations under each hypothesis follow complex Gaussian distributions

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}_t[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_t),\ n = 0, \ldots, N-1. \\ \mathcal{H}_0 &: \mathbf{x}_t[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{D}_t),\ n = 0, \ldots, N-1. \end{aligned} \tag{4.2}$$

where $\mathbf{R}_t$ is an arbitrary positive definite covariance matrix and $\mathbf{D}_t$ is a diagonal covariance matrix. The resulting hypothesis testing problem is in short a test on the structure of a covariance matrix under Gaussian data. This topic has been well researched in the statistics and signal processing literature.

Notice also that the likelihood under each hypothesis depends on unknown parameters and therefore the hypotheses are composite. The most typical approach to solve this kind of testing problems is the GLRT. When the noise is independent and identically distributed (iid) at each antenna, the GLRT is the well-known Sphericity

test [1], which is given by

$$\mathscr{L}_{spher} = \frac{|\mathbf{S}_t|^{1/L}}{(1/L)\, trace(\mathbf{S}_t)} \tag{4.3}$$

where $|\cdot|$ and $trace(\cdot)$ refer to the determinant and trace of a matrix respectively, and $\mathbf{S}_t = \mathbf{X}_t\mathbf{X}_t^H/N$ is the sample covariance matrix.

The iid case assumes a perfectly calibrated multiantenna cognitive receiver, which is in general hard to achieve. A more general testing problem that can accommodate calibration uncertainties in the different antenna front-ends, considers a generic diagonal noise covariance matrix under $\mathcal{H}_0$. The GLRT for this problem is the Hadamard ratio [Wilks, 1935] and is given by

$$\mathscr{L}_{Hadam} = \frac{|\mathbf{S}_t|}{\prod_{i=1}^{L}[\mathbf{S}_t]_{ii}} \tag{4.4}$$

where $[\mathbf{S}_t]_{ii}$ refers to the (i,i)th element of the sample covariance matrix $\mathbf{S}_t$.

4.2.1 Bayesian detector

The main goal of this paper is the practical evaluation in realistic environments of a multiantenna Bayesian detector for CR that has been recently proposed in [Manco-Vasquez et al., 2012b]. In comparison to GLRT (one-shot) based approaches, the Bayesian detector exploits statistical information obtained from past sensing periods. In the following, we provide a short description of this technique; for a full treatment of this topic the reader is referred to [Manco-Vasquez et al., 2012b].

In the proposed Bayesian framework, prior distributions are placed directly on the spatial covariance matrices under $\mathcal{H}_0$ and $\mathcal{H}_1$, as well as on the probability of the channel occupancy.

The main idea of the proposed procedure is depicted in Fig. 4.1. At each sensing frame, the cognitive receiver updates the posterior distributions for $\mathbf{R}_t$ and $\mathbf{D}_t$ from the priors existing at t and the likelihood obtained from $\mathbf{X}_t$. Conjugate priors are chosen in order to obtain closed-form expressions for the posteriors as indicated in [Manco-Vasquez et al., 2012b].

The posterior of the channel occupancy $p(z_t = 1|X_t) = \hat{\pi}_t$ is then thresholded and used for detection, where the hidden variable z_t indicates whether the PU is active $(z_t = 1)$ or not $(z_t = 0)$. Specifically, as it was shown in [Manco-Vasquez et al., 2012b], $\hat{\pi}_t$ can be obtained by blending the posterior and prior parameters according to

$$\hat{\pi}_t = \frac{1}{1 + \frac{|\check{\mathbf{D}}_t|^{\frac{\check{m}}{2}}|\hat{\mathbf{R}}_t|^{\frac{\hat{n}}{2}}\Gamma_L(\frac{\check{n}}{2})\Gamma(\frac{\hat{m}}{2})^L}{|\check{\mathbf{R}}_t|^{\frac{\check{n}}{2}}|\hat{\mathbf{D}}_t|^{\frac{\hat{m}}{2}}\Gamma_L(\frac{\hat{n}}{2})\Gamma(\frac{\check{m}}{2})^L}} \tag{4.5}$$

[1] Actually, for rank-deficient covariance matrices the GLRT is, in general, more complicated, as it was shown in [Ramirez et al., 2011]. However, when $L = 2$ (which is the scenario considered in the experimental setup of this paper) the GLRT coincides with the Sphericity test [Mauchly, 1940].

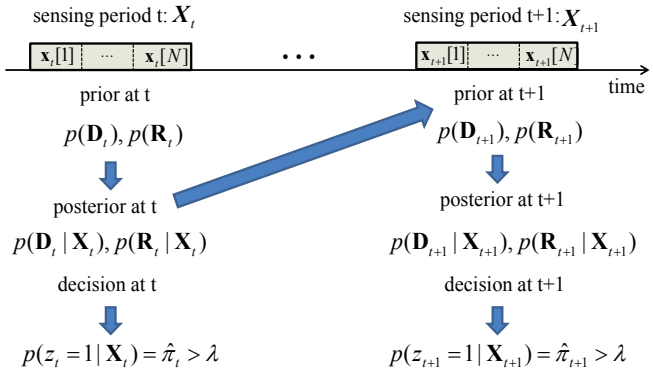


Figure 4.1: A Bayesian framework for spectrum sensing: the posteriors obtained after processing a sensing frame are employed as priors for the next sensing frame.

where $\Gamma_L(\cdot)$ and $\Gamma(\cdot)$ denote a multivariate gamma and a gamma function, respectively. Also, we have used $\check{}$ and $\hat{}$ to distinguish the prior and posterior parameters. The rest of the parameters are given by the following expressions:

$$\hat{n}_t = \check{n}_t + N \tag{4.6a}$$
$$\hat{\mathbf{R}}_t = \check{\mathbf{R}}_t + \mathbf{X}_t\mathbf{X}_t^H/N \tag{4.6b}$$
$$\hat{m}_t = \check{m}_t + N \tag{4.6c}$$
$$\hat{\mathbf{D}}_t = \check{\mathbf{D}}_t + \mathrm{diag}(\mathbf{X}_t\mathbf{X}_t^H/N). \tag{4.6d}$$

where $\mathrm{diag}(\cdot)$ refers to a matrix with their diagonal entries equal to the diagonal elements of the argument.

To handle non-stationary environments, a forgetting mechanism is introduced in the Bayesian detector. In particular, the prior parameters at the sensing frame $t+1$ are a smoothed version of the posterior parameters at t and those used in the (uninformative) prior at $t = 0$

$$\check{n}_{t+1} = \lambda\hat{n}_t + (1-\lambda)\check{n}_0 \tag{4.7a}$$
$$\check{\mathbf{R}}_{t+1} = \lambda\hat{\mathbf{R}}_t + (1-\lambda)\check{\mathbf{R}}_0 \tag{4.7b}$$
$$\check{m}_{t+1} = \lambda\hat{m}_t + (1-\lambda)\check{m}_0 \tag{4.7c}$$
$$\check{\mathbf{D}}_{t+1} = \lambda\hat{\mathbf{D}}_t + (1-\lambda)\check{\mathbf{D}}_0, \tag{4.7d}$$

where $\lambda \in [0,1]$ is a forgetting parameter. With this scheme, the cognitive receiver is able to both learn and forget information from past sensing frames, as needed. When $\lambda = 0$, all the information obtained from previous sensing frames is forgotten and the detector considers each frame independently (as the GLRT does). This would be adequate for very high-mobility scenarios and channels with high Doppler spreads. On the other hand, when $\lambda = 1$ no forgetting occurs and the posterior at t summarizes in an effective way all sensing data observed so far. This would be adequate in low-mobility scenarios.

4.3 Testbed Description

For our cognitive radio platform, we have chosen USRP devices [Ett, 2014] developed by Ettus Research. The testbed also includes a universal hardware drive (UHD) as a host driver and a set of Application Programming Interface (API) functions. This UHD driver can be utilized as a standalone application or with third-party applications (e.g. Matlab). For the present work, we have developed our own application called Universal Software Architecture for Software Defined Radio (USASDR). This software architecture allows working simultaneously with several USRP nodes, by means of a unique controller identified by an IP address that receives instructions from a remote PC running Matlab.

The cognitive radio platform is composed of N210 USRP nodes. Each of them consists of a USRP motherboard and a Radio Frequency (RF) daughterboard. Basically, the motherboard consists of dual analog-to-digital converters (ADC) and digital-to-analog (DAC) converters connected to a Field Programmable Gain Array (FPGA). The FPGA downsamples the signal in the digital domain and transfers it to a Gigabit Ethernet controller. On the other hand, the daughterboard is a modular front-end used for analog operations such as up/down conversion.

The flow of signal in the receive path starts by downconverting the frequency of the received signal from RF to Intermediate Frequency (IF), around DC. Specifically, the XCVR2450 daughterboard based on a MAX2829 IC is able to cover industrial, scientific and medical (ISM) bands of 2.4GHz to 2.5GHz, and 4.9GHz to 5.8GHz. At this stage we obtain I/Q analog signals that are subsequently digitized by the ADC. The USRP motherboard contains 14-bit ADCs with a sample rate of 100 MSamples/s. After sampling, the ADC transfers the data to an FPGA where data rate conversion is performed. Then, the FPGA transfers the results to Gigabit Ethernet controller which passes it over to the host computer where the rest of the signal processing tasks are performed.

Regarding the transmitter path, the same processing chain is repeated in reverse order. Firstly, the Gigabit Ethernet controller of the host computer sends the input data to the USRP. After receiving the complex signal, a digital up converter converts the signal to IF before passing it to a DAC with 16 bits of resolution and 400 MSamples/s of sample rate. The DAC finally sends the IF signal to the RF transceiver, where it is upconverted to RF and transmitted over the air.

In order to implement a multiantenna cognitive node, the N210 USRP node includes an specific expansion port that allows coherent synchronization of two USRP units, as it is depicted in Fig. 4.2. Since the same clock (oscillators) and time reference are shared, both USRP nodes can start transmitting/receiving at the same time, thus avoiding any synchronization problem.

4.4 Measurement Setup

We have considered a scenario where a single-antenna PU accesses the channel according to a predefined probability of occupancy and a cognitive receiver with two

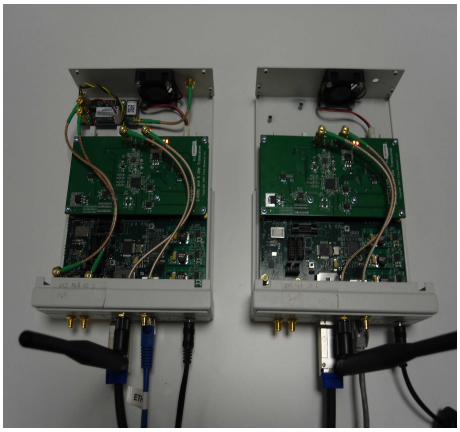


Figure 4.2: N210 Ettus devices with the XCVR2450 daugtherboard installed. A two-antenna cognitive receiver is composed of two N210 boards connected through a MIMO cable.

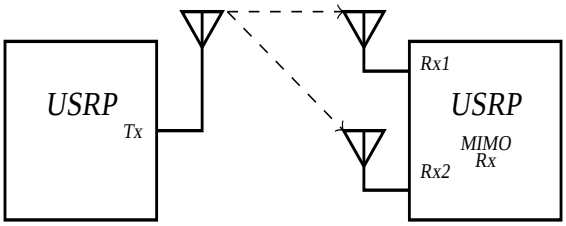


Figure 4.3: Scheme of the proposed scenario: A single-antenna PU and a secondary user with two antennas.

antennas senses the medium and performs the Bayesian detection procedure described in Section 4.2.1. The experiments were conducted in the laboratory of the Signal Processing Group at the University of Cantabria, with a clear line of sight (LOS) between the PU and the cognitive receiver. When active, the PU transmits an orthogonal frequency division multiplexing (OFDM) 802.11a signal[2]. Specifically, an OFDM waveform is generated with a rate of 9 Mbps using BPSK symbols, the resulting analog I/Q signals are sent to the RF front-end and transmitted at a carrier frequency of 5.6GHz. The power transmitted by the PU can be modified, thus allowing us to control the received signal-to-noise ratio (SNR). We compare the performance of the proposed Bayesian detector, the Sphericity (iid case) and the Hadamard (non-iid) GLRT detectors.

4.4.1 Measurement Procedure

The PU starts transmitting according to a predefined periodic sequence of states $\mathcal{H}_1$ and $\mathcal{H}_0$ (See Fig. 4.4), which are obtained from the probability of channel occupancy that we have taken as 0.5. The PU stops transmitting after the secondary user (SU) has sensed the wireless channel and N samples have been acquired by the $L = 2$

[2] Remember, however, that the modulation format is assumed to be unknown and therefore is not exploited by the spectrum sensing techniques considered in this work.

antennas. The data acquired by the SU are transferred to the PC and the three detectors are tested. Then, a new transmitting/sensing cycle (either under $\mathcal{H}_1$ or $\mathcal{H}_0$) is performed as shown in Fig. 4.4.

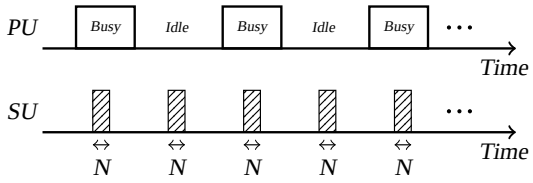


Figure 4.4: A PU transmits according to a periodic sequence of states, and the SU senses the wireless channel in each state.

At each sensing period, the signals acquired by the two USRP nodes are used to estimate the sample covariance matrix. As it was described in Section 4.2.1, the SU gathers statistical information summarized by the posterior densities on the covariance matrices. Therefore, we start computing the ROC curves after a given number of sensing frames has been sensed, so that the SU has obtained suitable prior information.

We carried out different measurement campaigns both in stationary and non-stationary environments. Typically, the indoor channel at 5 GHz under a controlled setup presents very long coherence times. In consequence, if no further action is taken the measurements are representative of a stationary environment. In order to test the performance under non-stationary environments, we should provide some mobility to either the PU or the SU. Since this mobility is very complicated to accomplish with our testbed, instead, we emulate a time-varying channel by transmitting the PU 802.11a signal using a time-varying beamformer. Basically, the idea consists on multiplying the PU OFDM signal by a 1×2 time-varying beamformer obtained according to an autoregressive model for fading channel, thus providing correlated Rayleigh processes [Baddour and Beaulieu, 2005] as it is shown in Fig.4.5.

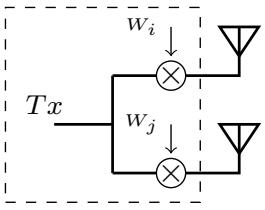


Figure 4.5: Emulation of a time-varying channel by transmit beamforming.

4.5 Experimental Results

We study the performance of the mentioned detectors by comparing their Receiving Operating Characteristic (ROC) curves obtained by measurements performed with our cognitive platform. For a given threshold we obtain the false alarm, P_{FA}, and

detection probabilities, P_D, by averaging the results obtained over multiple transmitting/sensing cycles. We conducted experiments under static, slowly time-varying and fast time-varying environments. The mobility speed of the channel is controlled by the transmit beamforming procedure depicted in Fig.4.5. In a static scenario, we would expect to obtain covariance matrices (under each hypothesis) that remain almost constant over consecutive sensing periods, thus boosting the performance of the Bayesian detector in comparison to GLRT counterparts. The number of samples per each sensing frame was set to $N = 50$, and the forgetting factor λ was selected depending on the scenario mobility.

Static scenario

In Fig. 4.6 we compare the ROC obtained by the Bayesian detector using $\lambda = 1$ (curve marked with squares), the Sphericity (circles) and the Hadamard (crosses) detectors. For this example the SNR measured at the receiver was -7.3 dB. The improvement of the Bayesian detector over the GLRTs is evident. Notice, also, that the ROCs for the Sphericity and the Hadamard detectors are almost identical: this is due to the fact that in this example the noise variance in the two receiving branches is very similar.

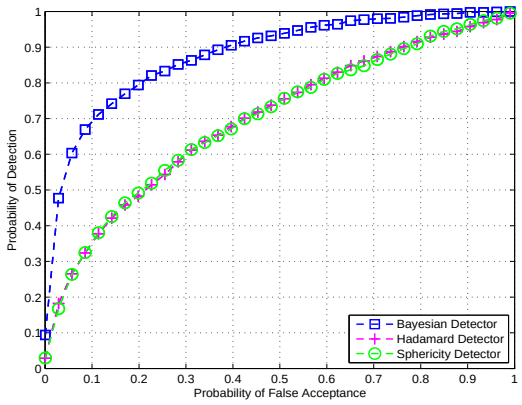


Figure 4.6: ROC curves for the Bayesian (squares), Sphericity (circles) and Hadamard (crosses) detectors, in a static environment. $N = 50$ and a SNR $= -7.3$ dB was measured.

Slowly time-varying scenario

In this scenario, we emulate a slowly time-varying channel by using the transmit beamforming procedure described previously. In this scenario, under a given hypothesis, the sample covariance matrix is correlated from frame to frame one as it is shown in Fig. 4.7. Nevertheless, the improvement of the Bayesian scheme is still evident.

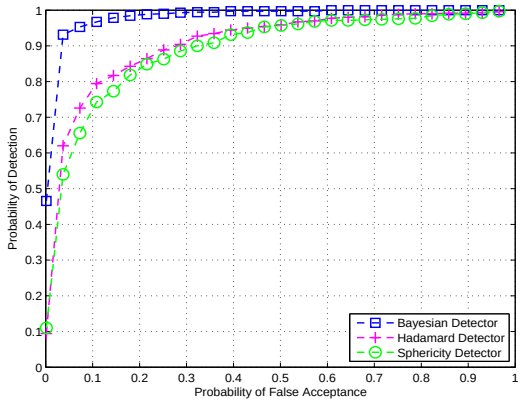


Figure 4.7: ROC curves for the Bayesian (squares), Sphericity (circles) and Hadamard (crosses) detectors, in a slowly time-varying environment. $N = 50$ and a SNR $= -1.18$ dB was measured.

Fast time-varying scenario

In our final experiment, we have studied the performance when the channel realizations in each sensing frame are iid (see Fig. 4.8). We note that our Bayesian detector still outperforms the other two detectors in this scenario. This can be explained to the fact that, under $\mathcal{H}_0$, the diagonal covariance matrix remains constant and can be effectively learnt by the Bayesian detector.

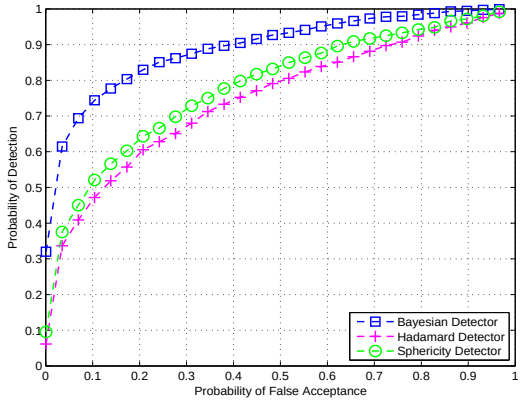


Figure 4.8: ROC curves for the Bayesian (squares), Sphericity (circles) and Hadamard (crosses) detectors, in a fast time-varying environment. $N = 50$ and a SNR $= -2.5$ dB was measured.

4.6 Conclusions

Using a cognitive radio testbed, in this paper we have evaluated the performance of three different multiantenna spectrum sensing algorithms in realistic static and time-varying environments. The experimental study shows that the Bayesian detector outperforms the GLRT detectors under the same conditions in different environments. A

significant gain is obtained by the Bayesian detector since it exploits past information and is able to learn from the environment.

Chapter 5

Kernel Canonical Correlation Analysis for Cooperative Spectrum Sensing

---

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

Authors: Julio Manco Vásquez, Steven Van Vaerenbergh, Javier Vía, and Ignacio Santamaría

Affiliations: Dept. of Communications Engineering, Univ. of Cantabria, Spain.

Journal: Transactions on Emerging Telecommunications Technologies, 2014.

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

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 (USRP) 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.

Keywords Cooperative Spectrum Sensing, Kernel Canonical Correlation Analysis, Hardware Testbed, USRP.

5.1 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 [Van de Beek et al., 2012, Elshafie et al., 2014]. 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 [Akyildiz et al., 2011].

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 [Parsa et al., 2008]. As shown in [Biglieri, 2011, Makarfi and Hamdi, 2013], 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 [Makarfi and Hamdi, 2013]. In [Guimarães et al., 2013], 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 [Nieminen et al., 2010, Song et al., 2010].

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 [Font-Bach et al., 2014], 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 [Zhao and Sasaki, 2013], 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 [Ho-Van, 2013], [Ding et al., 2014]. 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 [Thilina et al., 2013] 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 [Tan et al., 2013]. In [Ma et al., 2008, Yan et al., 2009], 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 [Harchaoui et al., 2013], which have been used previously in the context of cognitive radio networks, for instance in [Ding et al., 2013]. 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 [Van Vaerenbergh et al., 2013]. 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 [Ett, 2014], 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 5.2, we give an overview of the CSS problem. A detailed description of the proposed KCCA-based detector and its operation is presented in Section 5.3. In Section 5.4, 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 5.5. Finally, the paper concludes with a discussion of the obtained results in Section 5.6.

5.2 Cooperative Spectrum Sensing

Let us consider a cooperative spectrum sensing scenario where M SUs and a PU coexist in the same area [Akyildiz et al., 2011]. 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 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. 5.1, where a small cell (shadowed) within a heterogeneous network (HetNet) receives interference from neighboring cells during the time that the channel is considered vacant.

5.3 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 correla-

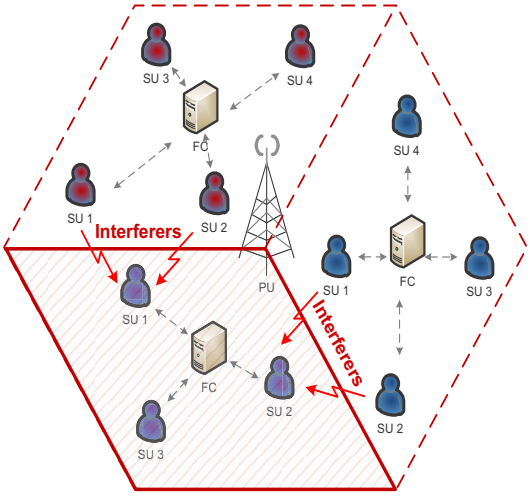


Figure 5.1: A spectrum sensing problem in a HetNet. Three SUs in a small cell cooperate to detect the presence of a PU, while two of them receive interference from other small cells. The interferences are independent of each other.

tion 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. 5.2. 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. 5.2):

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.

5.3.1 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

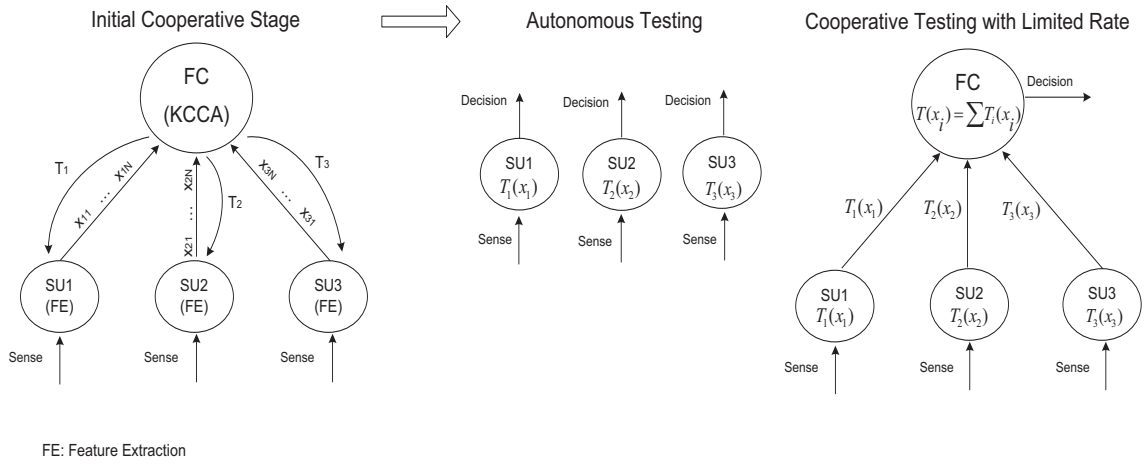


Figure 5.2: Operation of the proposed KCCA scheme: In an initial cooperative stage (left side of the picture), the measurements are reported to the FC to extract the local statistical tests T_i, then it starts operating either following a distributed (Autonomous Testing) or centralized configuration (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.

received signal r_i are sensed. We denote the feature vector extracted by the i-th SU during the n-th sensing period as,

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

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.

5.3.2 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 [Hardoon et al., 2004, Van Vaerenbergh et al., 2013]. This procedure consists in mapping the data into a high-dimensional space first, after which standard CCA is performed in the new space.

Kernel-Based Learning

In kernel-based learning (KBL), the data is transformed into a high-dimensional *feature space* [Bishop, 2006],

$$\Phi : \mathbf{x}_{in} \rightarrow \Phi(\mathbf{x}_{in}). \tag{5.1}$$

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,

$$\kappa(\mathbf{x}_{ij}, \mathbf{x}_{ik}) = \langle \Phi(\mathbf{x}_{ij}), \Phi(\mathbf{x}_{ik}) \rangle. \tag{5.2}$$

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

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

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 [Bishop, 2006]. 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,

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

Kernel Canonical Correlation Analysis for CSS

Consider a scenario in which M SUs are present, and each SU produces N feature vectors, $\{\mathbf{x}_{i1}, \mathbf{x}_{i2}, ...\mathbf{x}_{iN}\}$. 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[1].

The pairwise canonical correlations between the data sets, ρ_{ij}, are obtained in the context of KCCA as $\rho_{ij} = \mathbf{z}_i^\top \mathbf{z}_j = \alpha_i^\top \mathbf{K}_i \mathbf{K}_j \alpha_j$ [Bach and Jordan, 2003], where $\mathbf{z}_i = \mathbf{K}_i \alpha_i$ is a canonical variate obtained as the projection of the i-th set of data by means of the canonical vector α_i. A measure of the correlation associated to the i-th data set, ρ_i, can be subsequently obtained as,

$$\rho_i = \frac{1}{M-1} \sum_{\substack{j=1 \\ j \neq i}}^{M} \rho_{ij}, \tag{5.3}$$

and a generalized canonical correlation can be obtained as

$$\rho = \frac{1}{M} \sum_{i=1}^{M} \rho_i. \tag{5.4}$$

The maximization of ρ with respect to the canonical vectors α_i admits a trivial solution, which can be easily avoided by means of the following constraint on the energy of the canonical variates

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

Analogously, overfitting problems can be avoided by adding a regularization factor, c, to the norm of the projectors in the previous constraint [Hardoon et al., 2004].

$$\frac{1}{M} \sum_{i=1}^{M} \alpha_i^\top \mathbf{K}_i \mathbf{K}_i \alpha_i + c\ \alpha_i^\top \mathbf{K}_i \alpha_i = 1 \tag{5.5}$$

The canonical weights α_i are obtained by maximizing ρ subject to the restriction given in equation (5.5). This can be solved by the method of Lagrange multipliers, yielding the following generalized eigenvalue problem (GEV)

$$\frac{1}{M} \mathbf{R} \alpha = \beta \mathbf{D} \alpha, \tag{5.6}$$

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

$$\mathbf{R} = \begin{bmatrix} \mathbf{K}_1\mathbf{K}_1 & \cdots & \mathbf{K}_1\mathbf{K}_M \\ \vdots & \ddots & \vdots \\ \mathbf{K}_M\mathbf{K}_1 & \cdots & \mathbf{K}_M\mathbf{K}_M \end{bmatrix}, \tag{5.7}$$

[1] Several generalizations of CCA to more than two sets of variables can be found in [Kettenring, 1971, Vía et al., 2007].

and $\mathbf{D}$ is given by

$$\mathbf{D} = \begin{bmatrix} \mathbf{K}_1(\mathbf{K}_1 + c\mathbf{I}) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{K}_M(\mathbf{K}_M + c\mathbf{I}) \end{bmatrix}. \tag{5.8}$$

The solution α contains the different canonical weights as stacked vectors, $\alpha = [\alpha_1^\top, \alpha_2^\top, ..., \alpha_M^\top]^\top$, and it is retrieved as the eigenvector corresponding to the largest eigenvalue of the GEV problem (5.6) [Van Vaerenbergh et al., 2013]. The eigenvalue β 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.

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 [Hardoon et al., 2004],

$$\sum_{n=1}^{N} \Phi(\mathbf{x}_{in}) = 0, \quad i = 1, ..., M. \tag{5.9}$$

Since the transformations Φ 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

$$\tilde{\mathbf{K}}_i = \mathbf{N}_o \mathbf{K}_i \mathbf{N}_o^\top, \tag{5.10}$$

where $\mathbf{N}_o = \left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)$, $\mathbf{1}$ 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

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

a similar procedure is followed [Van Vaerenbergh, 2010], leading to

$$\tilde{\mathbf{k}}_i = \left(\mathbf{k}_i - \mathbf{1}^T \tilde{\mathbf{K}}_i / N\right) \mathbf{N}_o. \tag{5.11}$$

5.3.3 Local and Global Tests

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

$$T_i(\mathbf{x}_i) = \sum_{j=1}^{N} \alpha_{ij} \tilde{\kappa}(\mathbf{x}_i, \mathbf{x}_{ij}) \tag{5.12}$$

where α_{ij} refers to the j-th element of the canonical vector α_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 (5.12) 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

$$T_i(\mathbf{x}) = \sum_{i=1}^{M} T_i(\mathbf{x}_i), \tag{5.13}$$

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.

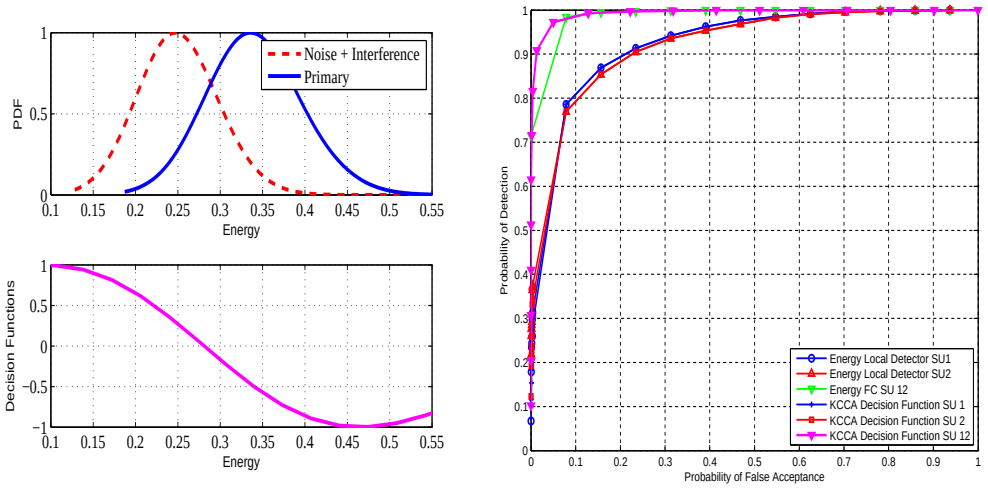
5.4 Simulation Results

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 [2]. Again, this value has been chosen mainly for computational reasons, as well as to avoid abrupt changes in the scenario statistics.

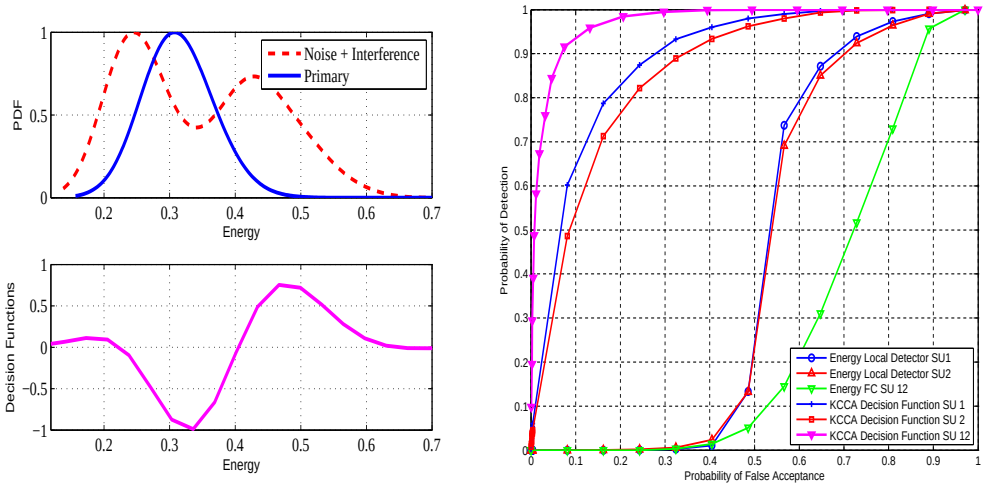
For the KCCA-based detector a Gaussian kernel is employed, the kernel width w_i for each set of data $\mathbf{x}_i$ is chosen by applying the Silverman's rule [Silverman, 1986, Van Vaerenbergh, 2010], and the regularization parameter is set to $c = 10$.

[2] 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 [Cordeiro et al., 2007].



(a) (b)

Figure 5.3: (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.



(a) (b)

Figure 5.4: (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.

5.4.1 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[3], 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. 5.3(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. 5.3(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. 5.4(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. 5.4(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. 5.5(a) and 5.5(b) we have considered a scenario with $M = 3$ SUs, where the advantages of including more information in the feature

[3]The PDFs for the results shown in this paper are obtained using a Parzen density estimator with a Gaussian kernel [Duda et al., 2000].

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. 5.5(a), where we observe that the energy of the primary signal almost overlaps with that of 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 [Shalvi and Weinstein, 1990], which is defined as the normalized fourth-order cumulant,

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

The PDF of the kurtosis is shown in Fig. 5.5(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. 5.5(c).

5.5 Experimental Results

5.5.1 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 [Ett, 2014].

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:

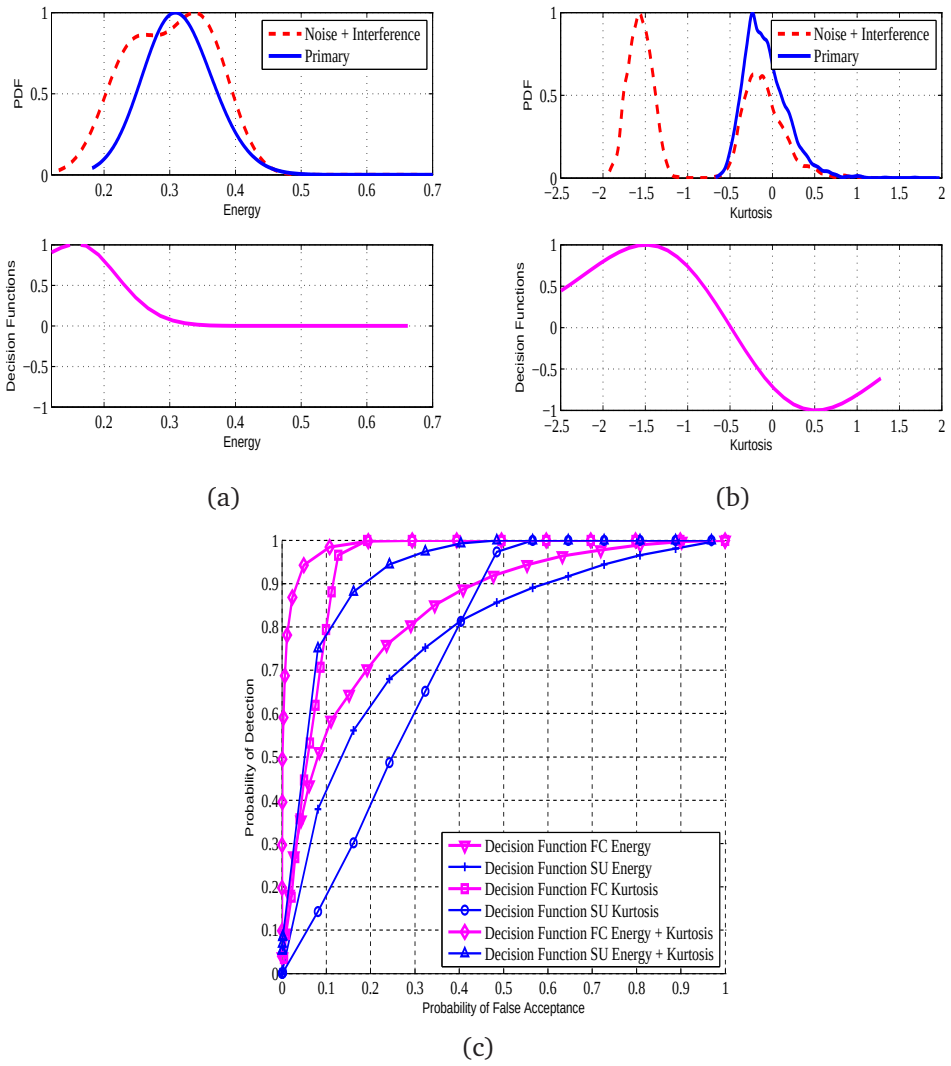


Figure 5.5: PDF of the received signal under both hypothesis, and the KCCA decision function for a SINR $\approx$ -7.45 dB at the SU 1 (a) Energy (b) Kurtosis. The corresponding ROC curves for local decisions (at each SU) and centralized decisions (at the FC) with a KCCA-based detector using only one or both features (c).

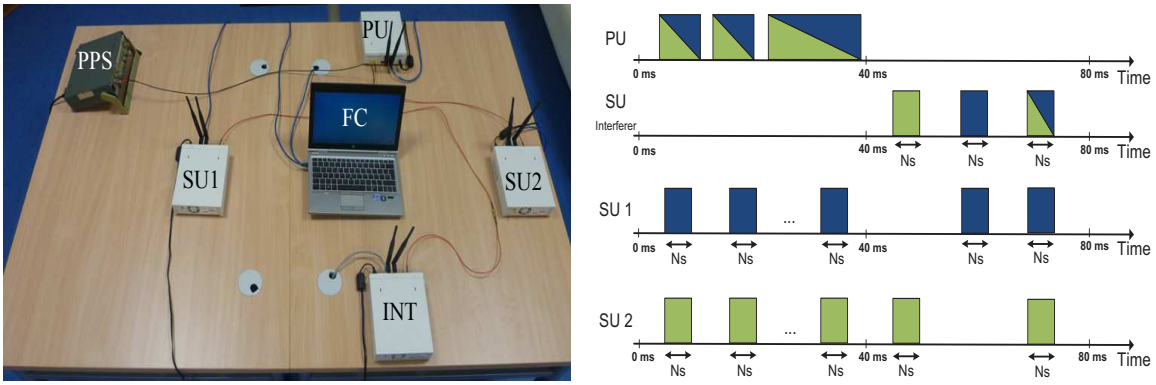
- The flow of the signal at the receiver side is similar to its counterpart, but in a reverse order. After capturing the data, a Gigabit Ethernet controller is in charge of transferring it to the host computer where the rest of the signal processing tasks are performed. A detailed description of the flow of data with our custom implementation can be found in [Manco-Vasquez et al., 2012a].

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. 5.6(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.

5.5.2 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. 5.6(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 α_i are calculated and used to form the statistic $T_i(x)$, whose performance is evaluated during an off-line process.



(a) (b)

Figure 5.6: (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.

5.5.3 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 [Gutiérrez et al., 2011] 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. 5.7(a) and 5.7(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. 5.8(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.

Example 2. A more interesting case is depicted in Figs. 5.7(c) and 5.7(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. 5.8(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. 5.7(e) and 5.7(f), and its corresponding ROC curve in Fig. 5.8(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. 5.4(a), 5.4(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.

5.6 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

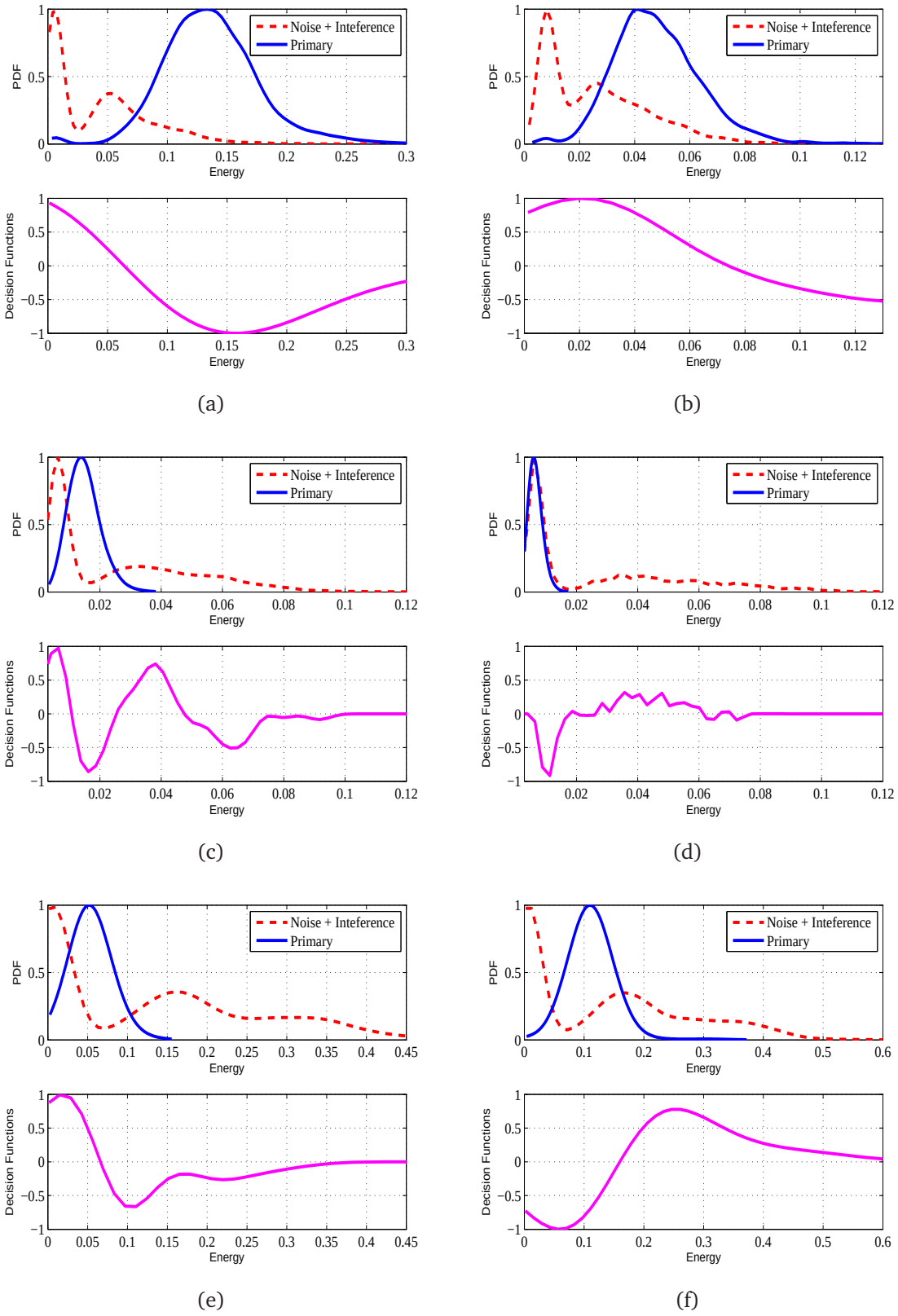


Figure 5.7: Three considered cases (rows): KCCA decision function and probability density function for the primary, the interfering and noise signal at SU 1 (left) and SU 2 (right). (a) and (b) with an approx. SNR 0.63 dB, (c) and (d) with an approx. SINR -11.4 dB and -9.2 dB respectively, and finally (e) and (f) with an approx. SINR -6.3 dB and -5.1 dB.

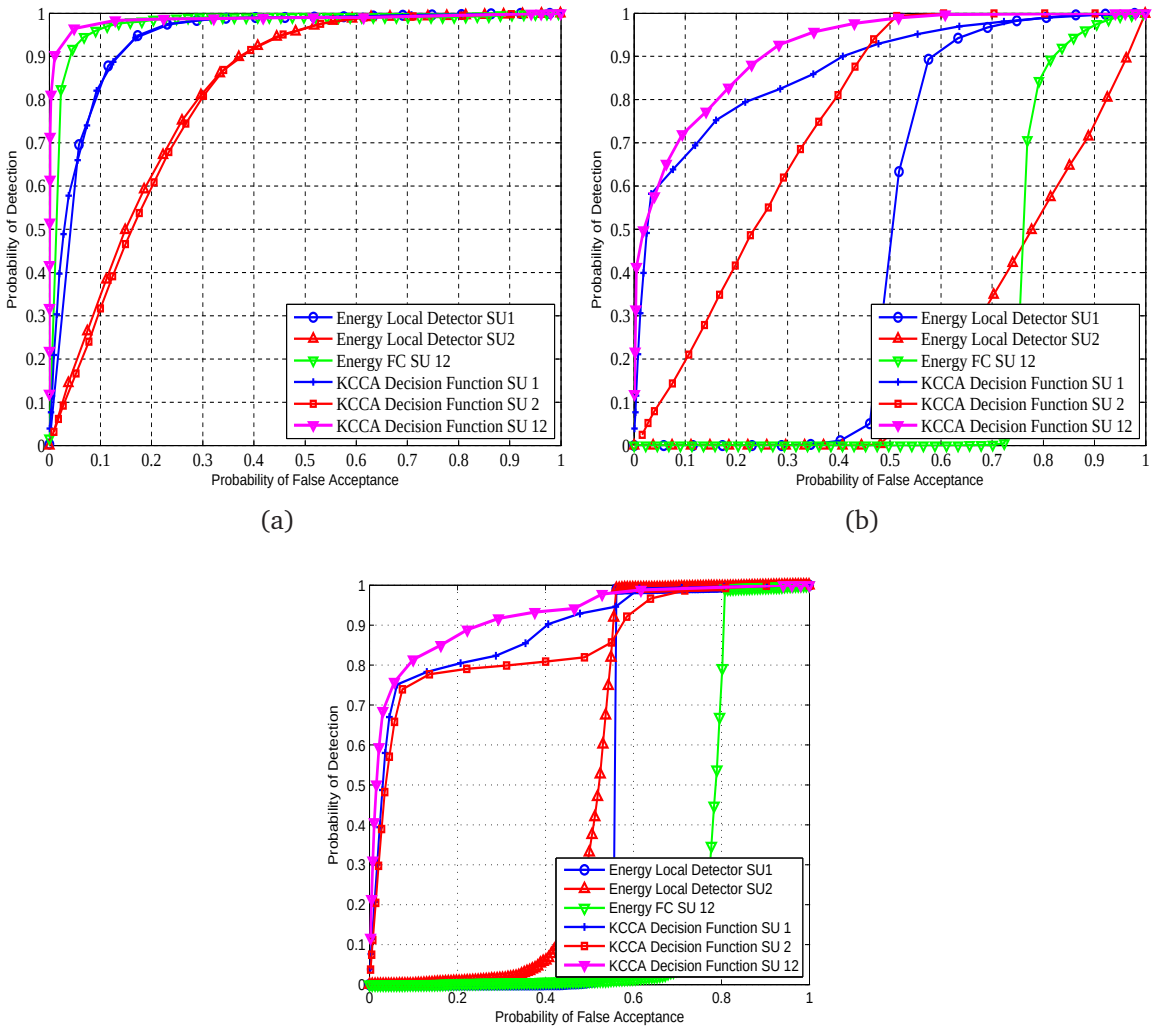


Figure 5.8: ROC Curves for the KCCA and energy detector and for three considered cases.

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.

Chapter 6
Conclusions and Future Lines

6.1 Conclusions

This dissertation tackles the detection problem in CR scenarios. In these scenarios, the requirements imposed by this technology demand the introduction of autonomous learning mechanisms that allow the radio devices to operate in unknown RF environments where prior knowledge is not available. In this context, we have proposed and evaluated novel detection schemes that meet the aforementioned requirements. For that end, several experimental evaluations are carried out to test the feasibility of these novel solutions. These approaches, namely, a Bayesian spectrum sensing framework, and a robust KCCA CSS scheme, are examined under more realistic conditions in a CR testbed utilizing software-based PHY-layer implementations, where the experimental results corroborates the obtained numerical results.

To summarize, this thesis proposes and evaluates spectrum sensing schemes tailored for CR networks employing, for that end, learning techniques as well as experimental evaluations.

6.2 Future Works

The conception of a CR technology entails a great effort toward its development by means of the research and the experimental evaluation of the developed technologies. Nevertheless, there are still many open issues that deserve further investigation. Some of these challenging problems related to the CR paradigm have been addressed in this thesis. However these approaches also motivate further research in the explored directions. Here, we list some research lines that could complement the presented work, and other ones, close to the studied subjects, that still remain unanswered.

- We have formulated probabilistically our knowledge about a CR scenario within a Bayesian framework, where however other relevant features could also be explored. For instance, the channel access patterns of the primary users charac-

terized as slowly time-varying, could be learnt to improve the detection performance. In addition, different forgetting factors are required to attain the best performance depending on the channel, for which it is interesting to derive explicit relations between the forgetting factor λ and the channel variability (e.g. λ_{ch} in our simulations). For instance, new estimation procedures taking into account other parameters involved such as the SNR, the number of antennas, or the number of acquired samples in each sensing frame could be considered.

- With the KCCA detector, the detection of a primary user is carried out exploiting the non-linear correlations among the received SU signals. In this scheme, the availability of more received signals yield better decision functions and a better detection performance. This improvement of the performance can be attained with more sensors and/or more acquired data at each SU, which requires additional research. For instance, multiantenna SUs provide additional spatial dimensions to be exploited in the search of non-linear correlations.

  Another step ahead would be the incorporation of a mechanism in the selection of some SUs to attain the same or even a better performance while alleviating the complexity of the algorithm. Eventually, the corresponding experimental evaluations taking into account a large number of sensors and/or antennas could also be addressed to test the feasibility of the approach.

- The acquired experience in the development of a CR platform allows the rapid implementation and experimental evaluation of other related aspects not considered in this thesis. In this regard, it is worth mentioning future works such as the evaluation of detection schemes in terms of achievable rate, as well as the rate adaptation approaches for CR links in time-varying channels where reinforcement learning could be employed.

Appendix A
Publications

This dissertation has given place to the following publications:

A.1 International Journal

1.1 A Bayesian Approach for Adaptive Multiantenna Sensing in Cognitive Radio Networks. J. Manco-Vásquez, Miguel Lázaro, David Ramírez, J. Vía, I. Santamaría. Signal Processing Elsevier, Volume 96, Pages 228-240. 2014.

1.2 Kernel Canonical Correlation Analysis for Robust Cooperative Spectrum Sensing in Cognitive Radio Networks. J. Manco-Vásquez, Jesus Ibáñez, J. Vía, I. Santamaría. Transactions on Emerging Telecommunications Technologies, 30 Oct. 2014.

A.2 International Conferences

2.1 Experimental Evaluation of Multiantenna Spectrum Sensing Detectors using a Cognitive Radio Testbed. J.Manco Vásquez, J. Gutiérrez, J. Pérez, J. Ibáñez, and I. Santamaría. IEEE International Symposium on Signals, Systems and Electronics (ISSSE). Potsdam, Germany October 2012.

2.2 Bayesian Multiantenna Sensing for Cognitive Radio. J. Manco-Vásquez, Miguel Lázaro, David Ramírez, J. Vía, I. Santamaría. IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM). Hoboken NJ, USA June 2012.

2.3 Adaptive Kernel Canonical Correlation Analysis Algorithms for Maximum and Minimum Variance. S. Van Vaerenbergh, J. Vía, J. Manco-Vásquez, and I. Santamaría. IEEE International Conference on Acoustics, Speech, and Signal

Processing (ICASSP). Vancouver, Canada May 2013.

2.4 Experimental Evaluation of a Cooperative Kernel-Based Approach for Robust Spectrum Sensing. J. Manco-Vásquez, S. Van Vaerenbergh, J. Vía, I. Santamaría. IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM). A Coruña, Spain June 2014.

Bibliography

[Boo, 2013] "Boost C++ libraries". `http://www.boost.org/`, 2013.

[Int, 2013] "Intel threading building blocks: (Intel TBB)". `http://threadingbuildingblocks.org/.Intel`, 2013.

[Poc, 2013] "Poco C++". 2013.

[Pyt, 2013] "Python". `http://www.python.org/`, 2013.

[Swi, 2013] "The Simple Wrapper and Interface Generator SWIG". `http://www.swig.org/`, 2013.

[Ett, 2014] "Ettus research LLC. Universal Software Radio". *http://www.ettus.com/*, 2014.

[Asg, 2015] "Asgard platform". A software framework for the development of SDR and cognitive test beds: http://blog.asgard.lab.es.aau.dk/, 2015.

[Pla, 2015a] "Communication research centre canada". www.crc.ca/coral, 2015.

[Pla, 2015b] "Cornet platform". http://cornet.wireless.vt.edu/trac/wiki/CORNET, 2015.

[Too, 2015] "The GNU radio software radio toolkit". gnuradio.org/, 2015.

[Akyildiz et al., 2011] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan. "Cooperative spectrum sensing in cognitive radio networks: A survey". *Phys. Commun.*, volume 4, pages 40–62, 2011.

[Akyildiz et al., 2006] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty. "Next generation dynamic spectrum access cognitive radio wireless networks: A survey". *Computer Networks (Elsevier)*, volume 50, no. 13, pages 2127–2159, 2006.

[Axel and Larsson, 2009] E. Axel and E. G. Larsson. "A Bayesian approach to spectrum sensing, denoising and anomaly detection". In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pages 2333–2336. 2009.

[Bach and Jordan, 2003] F. R. Bach and M. I. Jordan. "Kernel independent component analysis". *J. Mach. Learn. Res.*, volume 3, pages 1–48, 2003.

[Baddour and Beaulieu, 2005] K. Baddour and N. Beaulieu. "Autoregressive modeling for fading channel simulation". *IEEE Transactions on Wireless Communications*, volume 4, no. 4, pages 1650–1662, 2005.

[Bidon et al., 2008] S. Bidon, O. Besson, and J. Tourneret. "A bayesian approach to adaptive detection in nonhomogeneous environments". *IEEE Transactions on Signal Processing*, volume 56, no. 1, pages 205–217, 2008.

[Biglieri, 2011] E. Biglieri. "Effect of uncertainties in modeling interferences in coherent and energy detectors for spectrum sensing". In *Proc. of IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 2418–2421. 2011.

[Bishop, 2006] C. M. Bishop. *Pattern recognition and machine learning*. Springer New York, 2006.

[Cabric, 2008] D. Cabric. "Addressing feasibility of cognitive radios". *IEEE Transactions on Signal Processing*, volume 25, no. 6, pages 85–93, 2008.

[Cabric and Brodersen, 2005] D. Cabric and R. Brodersen. "Physical layer design issues unique to cognitive radio systems". In *Proc. of IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, volume 2, pages 759–763. 2005.

[Cabric et al., 2006] D. Cabric, A. Tkachenko, and R. Brodersen. "Spectrum sensing measurements of pilot, energy, and collaborative detection". In *Proc. of IEEE Military Communications Conference, MILCOM.*, pages 1–7. 2006.

[Chaudhari et al., 2014] S. Chaudhari, M. Kosunen, S. Mäkinen, M. Laatta, V. Koivunen, J. Ryynänen, and M. Valkama. "Measurement campaign for collaborative sensing using cyclostationary based mobile sensors". In *the 7th IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks DySPAN*, pages 283–290. 2014.

[Chen et al., 2007] H.-S. Chen, W. Gao, and D. G. Daut. "Spectrum sensing using cyclostationary properties and application to ieee 802.22 wran". In *Global Communications Conference, GLOBECOM, Washington, USA*. 2007.

[Chen and Qiu, 2011] Z. Chen and R. Qiu. "Cooperative spectrum sensing using Q-learning with experimental validation". In *Proc. of IEEE Southeastcon*, pages 405–408. 2011.

[Chowdhury et al., 2011] K. Chowdhury, R. Doost-Mohammady, W. Meleis, M. Di Felice, and L. Bononi. "Cooperation and communication in cognitive radio networks based on TV spectrum experiments". In *Proc. of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–9. 2011.

[Clancy and Walker, 2006] T. Clancy and B. Walker. "Predictive dynamic spectrum access". In *Proc. SDR Forum Technical Conference*. 2006.

[Cordeiro et al., 2007] C. Cordeiro, M. Ghosh, D. Cavalcanti, and K. Challapali. "Spectrum sensing for dynamic spectrum access of TV bands". In *2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*. 2007.

[Couillet and Debbah, 2009] R. Couillet and M. Debbah. "Bayesian inference for multiple antenna cognitive receivers". In *Proc. IEEE Wireless Communications and Networking Conference, WCNC 2009*, 6, pages 1–6. 2009.

[Couillet and Debbah, 2010] R. Couillet and M. Debbah. "A Bayesian framework for collaborative multi-source signal sensing". *IEEE Transactions on Signal Processing*, volume 58, no. 10, pages 5186–5195, 2010.

[Dandawate and B. Giannakis, 1994] A. V. Dandawate and G. B. Giannakis. "Statistical tests for presence of cyclostationarity". *IEEE Transactions on Signal Processing*, volume 42, no. 9, pages 2355–2369, 1994.

[Ding et al., 2013] G. Ding, Q. Wu, Y.-D. Yao, J. Wang, and Y. Chen. "Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions". *IEEE Signal Processing Magazine*, volume 30, no. 4, pages 126–136, 2013.

[Ding et al., 2014] G. Ding, Q. Wu, Y. Zou, J. Wang, and Z. Gao. "Joint spectrum sensing and transmit power adaptation in interference-aware cognitive radio networks". *Transactions on Emerging Telecommunications Technologies*, volume 25, no. 2, pages 231–238, 2014.

[Doyle et al., 2010] L. Doyle, P. Sutton, K. Nolan, J. Lotze, B. Ozgul, T. Rondeau, S. Fahmy, H. Lahlou, and L. DaSilva. "Experiences from the iris testbed in dynamic spectrum access and cognitive radio experimentation". In *IEEE Symposium on New Frontiers in Dynamic Spectrum*, pages 1–8. 2010.

[Duda et al., 2000] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley, 2000.

[Elshafie et al., 2014] H. Elshafie, N. Fisal, M. Abbas, W. A. Hassan, H. Mohamad, N. Ramli, S. Jayavalan, and S. Zubair. "A survey of cognitive radio and TV white spaces in Malaysia". *Transactions on Emerging Telecommunications Technologies*, pages 652–656, doi: 10.1002/ett.2778, 2014.

[Font-Bach et al., 2014] O. Font-Bach, N. Bartzoudis, A. Pascual-Iserte, M. Payaro, L. Blanco, D. López, and M. Molina. "Interference management in LTE-based HetNets: a practical approach". *Transactions on Emerging Telecommunications Technologies*, volume 26, pages 195–215, 2014.

[Font-Segura and Wang, 2010] J. Font-Segura and X. Wang. "GLRT-based spectrum sensing for cognitive radio with prior information". *IEEE Transactions on Communications*, volume 58, no. 7, pages 2137–2146, 2010.

[Goldsmith, 2005] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[Grimm et al., 2011] M. Grimm, A. Krah, N. Murtaza, R. K. Sharma, M. Landmann, R. Thom, A. Heuberger, and M. Hein. "Performance evaluation of directional spectrum sensing using an over-the-air testbed". In *Proc. of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management*, page 5 pp. ACM, 2011.

[Guimarães et al., 2013] D. A. Guimarães, R. A. A. de Souza, and A. N. Barreto. "Performance of cooperative eigenvalue spectrum sensing with a realistic receiver model under impulsive noise". *Journal of Sensor and Actuator Networks*, volume 2, no. 1, pages 46–69, 2013.

[Gutiérrez et al., 2011] J. Gutiérrez, Ó. González, J. Pérez, D. Ramírez, L. Vielva, J. Ibáñez, and I. Santamaría. "Frequency-domain methodology for measuring MIMO channels using a generic test bed". *IEEE Transactions on Instrumentation and Measurement*, volume 60, no. 3, pages 827–838, 2011.

[Gutiérrez et al., 2012] J. Gutiérrez, J. Ibañez, and J. Perez. "Beamforming-based emulation of spatial and temporal correlated miso channels". In *Proc. of IEEE International Symposium on Signals Systems and Electronics (ISSSE), Potsdam, Germany*. 2012.

[Harchaoui et al., 2013] Z. Harchaoui, F. Bach, O. Cappe, and E. Moulines. "Kernel-based methods for hypothesis testing: A unified view". *IEEE Signal Processing Magazine*, volume 30, no. 4, pages 87–97, 2013.

[Hardoon et al., 2004] D. R. Hardoon, S. R. Szedmak, and J. R. Shawe-taylor. "Canonical correlation analysis: An overview with application to learning methods". *Neural Comput.*, volume 16, no. 12, pages 2639–2664, 2004.

[Haykin, 2005] S. Haykin. "Cognitive radio: Brain-empowered wireless communications". *IEEE Journal on Selected Areas in Communications*, volume 23, no. 2, pages 201–220, 2005.

[Hlawatsch and Matz, 2011] F. Hlawatsch and G. Matz. *Wireless Communications over Rapidly Time-Varying Channels*. Elsevier Academic Press, 2011.

[Ho-Van, 2013] K. Ho-Van. "Exact outage analysis of underlay cooperative cognitive networks with maximum transmit-and-interference power constraints and erroneous channel information". *Transactions on Emerging Telecommunications Technologies*, volume 26, pages 772–788, 2013.

[Jeffreys, 2012] H. Jeffreys. "An invariant form for the prior probability in estimation problems". In *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences*, volume 186, pages 453–461. 2012.

[Jin, 2012] Y. Jin. "Cognitive multi-antenna radar detection using bayesian inference". In *Proc. of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. 2012.

[Kay, 1998] S. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Sadle River, NJ: Prentice Hall, 1998.

[Kettenring, 1971] J. Kettenring. "Canonical analysis of several sets of variables". *Biometrika*, volume 58, no. 3, pages 433–451, 1971.

[Kremo et al., 2014] H. Kremo, O. Altintas, H. Tanaka, M. Kitamura, K. Inage, and T. Fujii. "Cooperative spectrum sensing in the vehicular environment: An experimental evaluation". In *Proc. of IEEE 79th Vehicular Technology Conference (VTC Spring)*, pages 1–5. 2014.

[Kulhavý and Zarrop, 1993] R. Kulhavý and M. B. Zarrop. "On general concept of forgetting". *International Journal of Control*, volume 58, no. 4, pages 905–924, 1993.

[Lopez, 2011] M. Lopez. *Spectrum usage models for the analysis, design and simulation of cognitive radio networks*. PhD Thesis, Universitat Politecnica de Catalunya, 2011.

[Lu et al., 2012] L. Lu, X. Zhou, U. Onunkwo, and G. Li. "Ten years of research in spectrum sensing and sharing in cognitive radio". *EURASIP Journal on Wireless Communications and Networking*, volume 2012, no. 1, page 28 pp., 2012.

[Ma et al., 2008] J. Ma, G. Zhao, and Y. Li. "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks". *IEEE Transactions on Wireless Communications*, volume 7, no. 11, pages 4502–4507, 2008.

[Maio and Conte, 2010] A. D. Maio and E. Conte. "Adaptive detection in gaussian interference with unknown covariance after reduction by invariance". *IEEE Transactions on Signal Processing*, volume 58, no. 6, pages 2925–2934, 2010.

[Makarfi and Hamdi, 2013] A. Makarfi and K. Hamdi. "Interference analysis of energy detection for spectrum sensing". *IEEE Transactions on Vehicular Technology*, volume 62, no. 6, pages 2570–2578, 2013.

[Manco-Vasquez et al., 2012a] J. Manco-Vasquez, J. Gutierrez-Teran, J. Perez-Arriaga, J. Ibañez, and I. Santamaria. "Experimental evaluation of multi-antenna spectrum sensing detectors using a cognitive radio testbed". In *Proc. of IEEE International Symposium on Signals Systems and Electronics (ISSSE), Potsdam, Germany*. 2012.

[Manco-Vasquez et al., 2012b] J. Manco-Vasquez, M. Lazaro-Gredilla, D. Ramirez, J. Via, and I. Santamaria. "Bayesian multiantenna sensing for cognitive radio". In *Proc. of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. 2012.

[Manco-Vásquez et al., 2014a] J. Manco-Vásquez, M. Lázaro-Gredilla, D. Ramírez, J. Vía, and I. Santamaría. "A Bayesian approach for adaptive multiantenna sensing in cognitive radio networks". *Signal Processing*, volume 96, page 240, 2014.

[Manco-Vásquez et al., 2014b] J. Manco-Vásquez, S. Van Vaerenbergh, J. Vía, and I. Santamaría. "Experimental evaluation of a cooperative kernel-based approach for robust spectrum sensing". In *8th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. A Coruña, Spain, 2014.

[Manco-Vásquez et al., 2014c] J. Manco-Vásquez, S. Van Vaerenbergh, J. Vía, and I. Santamaría. "Kernel canonical correlation analysis for robust cooperative spectrum sensing in cognitive radio networks". *Transactions on Emerging Telecommunications Technologies*, 2014.

[Mate et al., 2011] A. Mate, K.-H. Lee, and I.-T. Lu. "Spectrum sensing based on time covariance matrix using GNU radio and USRP for cognitive radio". In *Systems, Applications and Technology Conference (LISAT), 2011 IEEE Long Island*, pages 1–6. 2011.

[Mauchly, 1940] J. W. Mauchly. "Significance test for sphericity of a normal n-variate distribution". *The Annals of Mathematical Statistics*, volume 11, no. 2, pages 204–209, 1940.

[Mishra et al., 2005] S. Mishra, D. Cabric, C. Chang, D. Willkomm, B. van Schewick, A. Wolisz, and R. Brodersen. "A real time cognitive radio testbed for physical and link layer experiments". In *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN*, pages 562–567. 2005.

[Mitola and Maguire, 1999] J. Mitola and G. Maguire. "Cognitive radio: Making software radios more personal". *IEEE Personal Communications*, volume 6, no. 4, pages 13–18, 1999.

[Mubaraq et al., 2004] S. Mubaraq, D. Cabric, C. Chang, D. Willkomm, B. van Schewick, A. Wolisz, and R. W. Brodersen. "Implementation issues in spectrum sensing for cognitive radios". In *Proc. of IEEE Signals, Systems and Computers, 2004*, volume 1, pages 772–776. 2004.

[Nagaraj, 2009] S. V. Nagaraj. "Entropy-based spectrum sensing in cognitive radio". *Signal Processing*, volume 89, no. 2, pages 174–180, 2009.

[Nieminen et al., 2010] J. Nieminen, R. Jantti, and L. Qian. "Primary user detection in distributed cognitive radio networks under timing inaccuracy". In *Proc. of IEEE Symposium on New Frontiers in Dynamic Spectrum*, pages 1–8. 2010.

[Parsa et al., 2008] A. Parsa, A. Gohari, and A. Sahai. "Exploiting interference diversity for event-based spectrum sensing". In *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN*, pages 1–12. 2008.

[Perez-Cruz et al., 2013] F. Perez-Cruz, S. Van Vaerenbergh, J. Murillo-Fuentes, M. Lazaro-Gredilla, and I. Santamaria. "Gaussian processes for nonlinear signal processing". *IEEE Signal Processing Magazine*, volume 30, no. 4, pages 40–50, 2013.

[Ramirez et al., 2011] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria. "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas". *IEEE Transactions on Signal Processing*, volume 59, no. 8, pages 3764–3774, 2011.

[Ramirez et al., 2010] D. Ramirez, J. Via, I. Santamaria, and L. L. Scharf. "Detection of spatially correlated Gaussian times series". *IEEE Transactions on Signal Processing*, volume 58, no. 10, pages 5006–5015, 2010.

[Safatly et al., 2014] L. Safatly, B. Aziz, A. Nafkha, Y. Louet, Y. Nasser, A. El-Hajj, and K. Kabalan. "Blind spectrum sensing using symmetry property of cyclic autocorrelation function: from theory to practice". *EURASIP J. Wireless Comm. and Networking.*, volume 12, no. 11, pages 5244–5453, 2014.

[Sala-Alvarez et al., 2012] J. Sala-Alvarez, G. Vazquez-Vilar, and R. Lopez-Valcarce. "Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation". *IEEE Transactions on Signal Processing*, volume 60, no. 6, pages 3065–3078, 2012.

[Shalvi and Weinstein, 1990] O. Shalvi and E. Weinstein. "New criteria for blind deconvolution of nonminimum phase systems (channels)". *IEEE Transactions on Information Theory*, volume 36, no. 2, pages 312–321, 1990.

[Silverman, 1986] B. W. Silverman. *Density estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, London, UK, 1986.

[Sohn et al., 2007] K. J. Sohn, H. Li, and B. Himed. "Parametric GLRT for multichannel adaptive signal detection". *IEEE Transactions on Signal Processing*, volume 55, no. 11, pages 5351–5360, 2007.

[Song et al., 2010] S. H. Song, K. Hamdi, and K. Letaief. "Spectrum sensing with active cognitive systems". *IEEE Transactions on Wireless Communications*, volume 9, no. 6, pages 1849–1854, 2010.

[Sun and Berger, 2006] D. Sun and J. O. Berger. "Objective priors for the multivariate normal model". In *8th World Meeting on Bayesian Statistics, Alicante, Spain*, pages 1–26. 2006.

[Sutton et al., 2010] P. Sutton, J. Lotze, H. Lahlou, S. Fahmy, K. Nolan, B. Ozgul, T. Rondeau, J. Noguera, and L. Doyle. "Iris: an architecture for cognitive radio networking testbeds". *IEEE Communications Magazine*, volume 48, no. 9, pages 114–122, 2010.

[Taherpour et al., 2010] A. Taherpour, M. Nasiri-Kenari, and S. Gazor. "Multiple antenna spectrum sensing in cognitive radios". *IEEE Transactions on Wireless Communications*, volume 9, no. 2, pages 814–823, 2010.

[Tan et al., 2013] X. Tan, H. Zhang, Q. Chen, and J. Hu. "Opportunistic channel selection based on time series prediction in cognitive radio networks". *Transactions on Emerging Telecommunications Technologies*, volume 25, pages 1126–1136, 2013.

[Tandra and Sahai, 2008] R. Tandra and A. Sahai. "SNR walls for signal detection". *IEEE Journal of Selected Topics in Signal Processing*, volume 2, no. 1, pages 4–17, 2008.

[Thilina et al., 2013] K. 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*, volume 31, no. 11, pages 2209–2221, 2013.

[Tonelli and Buthler, 2013] O. Tonelli and J. Buthler. *The Asgard Platform*. A book (ed) of platforms, 1st edn. Wiley, New York, 2013.

[Tugnait, 2012] J. Tugnait. "On multiple antenna spectrum sensing under noise variance uncertainty and flat fading". *IEEE Transactions on Signal Processing*, volume 60, no. 4, pages 1823–1832, 2012.

[Vaerenbergh et al., 2012] S. V. Vaerenbergh, I. Santamaria, and M. Lazaro-Gredilla. "Estimation of the forgetting factor in kernel recursive least squares". In *Proc. of IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. 2012.

[Van de Beek et al., 2012] J. Van de Beek, J. Riihijarvi, A. Achtzehn, and P. Mahonen. "TV white space in Europe". *IEEE Transactions on Mobile Computing*, volume 11, no. 2, pages 178–188, 2012.

[Van Vaerenbergh, 2010] S. Van Vaerenbergh. *Kernel methods for nonlinear identification, equalization and separation of signals*. Ph.D. thesis, University of Cantabria, 2010.

[Van Vaerenbergh et al., 2012] S. Van Vaerenbergh, M. Lazaro-Gredilla, and I. Santamaria. "Kernel recursive least-squares tracker for time-varying regression". *IEEE Transactions on Neural Networks and Learning Systems*, volume 23, no. 8, pages 1313–1326, 2012.

[Van Vaerenbergh et al., 2013] S. Van Vaerenbergh, J. Vía, J. Manco-Vásquez, and I. Santamaría. "Adaptive kernel canonical correlation analysis algorithms for maximum and minimum variance". In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*. 2013.

[Vía et al., 2007] J. Vía, I. Santamaría, and J. Pérez. "A learning algorithm for adaptive canonical correlation analysis of several data sets". *Neural Networks*, volume 20, no. 1, page 139–152, 2007.

[Wang et al., 2010] P. Wang, J. Fang, N. Han, and H. Li. "Multiantenna-assisted spectrum sensing for cognitive radio". *IEEE Transactions on Vehicular Technology*, volume 59, no. 4, pages 1791–1800, 2010.

[Wilks, 1935] S. S. Wilks. "On the independence of k sets of normally distributed statistical variables". *The Annals of Mathematical Statistics*, volume 3, no. 3, pages 309–326, 1935.

[Yagi et al., 2010] Y. Yagi, T. Ohno, H. Murata, K. Yamamoto, and S. Yoshida. "Experimental study of cooperative spectrum sensing for cognitive radio". In *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–5. 2010.

[Yan et al., 2009] Y. Yan, A. Li, and H. Kayama. "Study on soft decision based cooperative sensing in cognitive radio network". In *IEEE 70th Vehicular Technology Conference Fall*, pages 1–5. 2009.

[Yucek and Arslan, 2009] T. Yucek and H. Arslan. "A survey of spectrum sensing algorithms for cognitive radio applications". *IEEE Communications Surveys & Tutorials*, volume 11, no. 1, pages 116–130, 2009.

[Zhang et al., 2010] R. Zhang, T. Lim, Y.-C. Liang, and Y. Zeng. "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach". *IEEE Transactions on Communications*, volume 58, no. 1, pages 84–88, 2010.

[Zhao and Sasaki, 2013] B. Zhao and S. Sasaki. "Active spectrum sensing for cognitive radio networks". *Transactions on Emerging Telecommunications Technologies*, volume 26, pages 789–799, 2013.