Cognitive Radios: System Design Perspective

Danijela Branislav Cabric
Robert W. Brodersen

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2007-156
http://www.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-156.html

December 17, 2007

Cognitive Radios: System Design Perspective

by

Danijela Branislav Čabrić

Diploma (University of Belgrade, Yugoslavia) 1998
M.S. (University of California, Los Angeles) 2001

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering-Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Robert W. Brodersen, Chair
Professor Anant Sahai
Professor Bozidar Stojadinović

Fall 2007

The dissertation of Danijela Branislav Čabrić is approved:

Chair Date

Date

Date

University of California, Berkeley

Fall 2007

Cognitive Radios: System Design Perspective

Abstract

Cognitive Radios: System Design Perspective

by

Danijela Branislav Čabrić

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Robert W. Brodersen, Chair

A major shift in wireless communications is now emerging with the development of cognitive radios, which attempt to share spectrum in a fundamentally new way. Cognitive radios address the problem of poor spectrum utilization exhibited in many frequency bands. On a conceptual level, cognitive radio networks sense the spectral environment and adapt transmission parameters to dynamically reuse available spectrum. The novelty of this approach requires us to re-architect the mechanisms for using radio frequencies and find a way for multiple systems to co-exist through sharing rather than fixed allocations.

This dissertation explores some fundamental questions in cognitive radio system design by bridging the theoretical and practical aspects of the physical and network layers. A system level design involved a closed loop research approach connecting the-

oretical analysis and development of new algorithms with their implementation and experimental verification. A wireless testbed platform with capabilities for real-time signal processing and protocols, networking and multiple antenna communication was developed to support this research approach.

Spectrum sensing has been identified as a key enabling functionality for cognitive radios, therefore the goal of this research was to address its feasibility, performance limits and implementation issues. A major challenge in the spectrum sensing design is the requirement to detect very weak signals of different types in a minimum time with high reliability. To solve this problem, a cross-layer design approach was applied involving sensing radio front-end, digital signal processing and networking solutions. By exploiting spatial filtering for interference suppression, statistical signal processing to combat channels uncertainties and network cooperations to improve detection reliability, we show the practically achievable limits for sensing weak signals in wideband cognitive radio channels. As a result, an architecture of a spectrum sensing function is proposed and its performance and implementation complexity are characterized using the developed testbed platform.

Professor Robert W. Brodersen
Dissertation Committee Chair

To my family.

Contents

List of Figures

List of Tables

Acknowledgments

I would first like to thank my advisor Professor Robert Brodersen, whose vision, technical guidance and support throughout this project have been invaluable. I deeply admire his research approach and genuine interest in understanding the true nature of things. His advising style with always positive attitude gave me just enough freedom to endeavor into a new research area, thought me what a real research impact means, and most importantly helped me build confidence. It has been a true fortune and privilege working with him.

I would also like to thank Professor Anant Sahai for sharing his knowledge and experience, and for serving on my dissertation committee. Professor Sahai's expertise in theoretical aspects of communications and signal processing was of great help during both formulation and solution of many of the research problems investigated in this thesis.

I would like to acknowledge Professor Bozidar Stojadinovic for his participation in my qualifying exam and for serving on my dissertation committee.

I am also grateful for the opportunity to be part of an exceptional research environment at the BWRC. One of most beneficial aspects of working at BWRC was the opportunity to interact with incredibly smart and talented students. Due to the collaborative aspects of the research project I was involved in, I have many fellow students to thank: Artem Tkachenko, Mubaraq Mishra, Jing Yang, Rahul Tandra, Dejan Markovic, Hayden So, Chen Chang, David Sobel, and Tina Smilkstein. I also

enjoyed stimulating discussions and friendship with Jana van Gruenen, Nate Pletcher, Simone Gambini, Luis Alarcon, Farhana Sheikh and Kevin Camera.

Thanks to Pam Atkinson and Isabel Blanco of CalView Program at UC Berkeley who gave me the opportunity to master my teaching skills by facilitating distance learning program to working professionals and students.

I also wish to thank Center for Circuits, Systems and Software (C2S2) program and BWRC member companies for the financial support. Special gratitude goes to BWRC staff Gary Kelson, Tom Boot, Brian Richards, Sue Mellers, Fred Burghardt, Kevin Zimmermann and Deirdre McAuliffe Bauer for doing a wonderful job in organizing and administering many activities of the center.

My warmest thanks go to Lara and Zeljka for their friendship and support over the years.

I owe a special thanks to my husband Srdjan for his patience, love and encouragement throughout this journey. He definitely made me a better and stronger person. I am truly grateful for his commitment and unconditional support during writing this thesis.

Finally, and most importantly, I would like to express gratitude to my parents and my sister. Even thousands miles apart, they have been present through every step of my life, providing support in difficult times, and enjoying every one of my accomplishments. Their endless love and belief in me have been a constant source of inspiration, and this thesis is dedicated to them.

Chapter 1

Introduction

1.1 Motivation

The demand for wireless connectivity and crowding of unlicensed spectra have pushed the regulatory agencies and technology developers to be ever more aggressive in providing new ways to use spectra. In order to enable future wireless systems for commercial use or public services, new technologies that can provide an order of magnitude increase in system capacity are needed, to either support more users or higher data rates. However, it is commonly believed that there is a spectrum scarcity at frequencies that can be economically used for wireless communications. The National Telecommunications and Information Administration's (NTIA) frequency allocation chart (Figure 1.1) indicates overlapping allocations over all of the frequency bands, which reinforces the scarcity mindset.

Under this static frequency allocation wireless systems are regulated through fixed spectrum assignments, operating frequencies and bandwidths, with constraints on power emission that limits their range. Due to these constraints, most communications systems are designed so that they achieve the best possible spectrum efficiency within the assigned bandwidth using sophisticated modulation, coding, multiple antennas and other techniques. The most advanced systems are approaching Shannon's channel capacity limit but cannot provide desired system capacity increase.

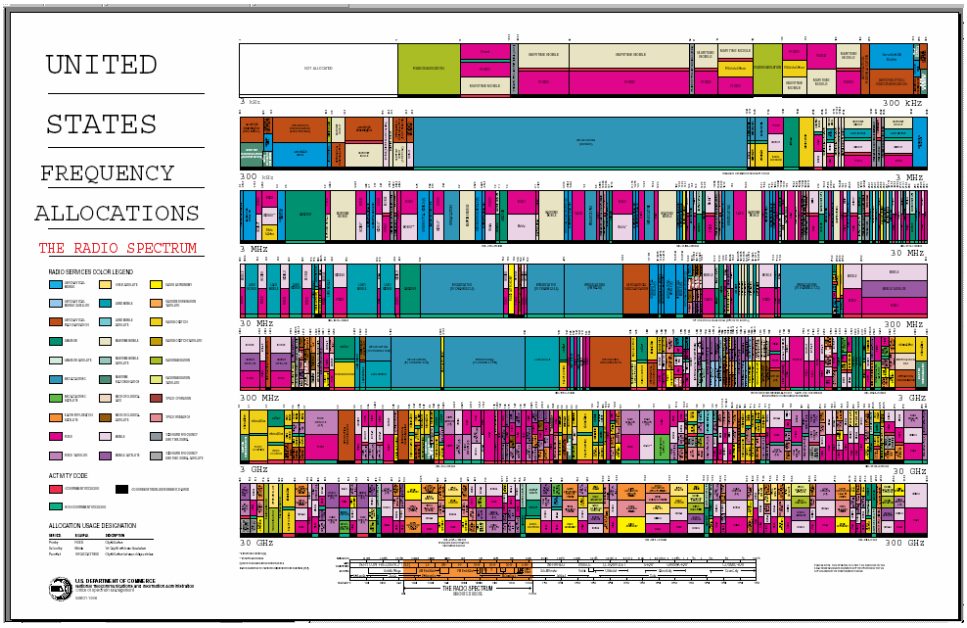
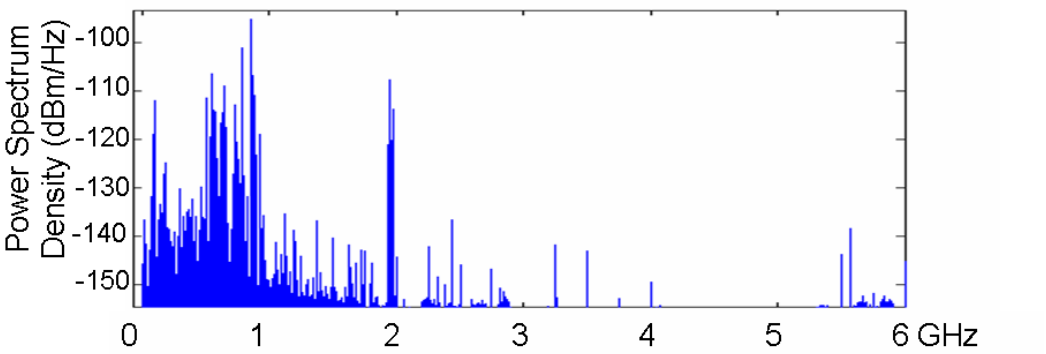


Figure 1.1: The NTIA's frequency allocation chart

While the current spectrum allocation leaves no available bandwidth for future wireless systems, actual measurements of spectrum utilization show that many assigned frequency bands are not being used at every location and time. Figure 1.2 shows measurements taken in downtown Berkeley which reveal a typical utilization

of roughly 30% below 3 GHz, and 0.5% in the 3 to 6 GHz frequency band. The graph shows usage over a very short period of time and represent the power spectral density (PSD) of the received 6 GHz wide signal collected for a span of $50\mu s$ sampled at 20 GS/s.



Freq (GHz)	0~1	1~2	2~3	3~4	4~5	5~6
Utilization(%)	54.4	35.1	7.6	0.25	0.128	4.6

Figure 1.2: Measurement of spectrum utilization (0-6 GHz) in downtown Berkeley [1]

Measurements taken over 10 minutes in the same Berkeley location (Figure 1.3) show that there are also temporal gaps in the spectrum usage even in the 0 to 2.5 GHz, which is considered to be very crowded. This view is supported by various recent measurements in the US and elsewhere. Even studies done by the FCC's Spectrum Policy Task Force reported vast temporal and geographic variations in the usage of allocated spectrum with utilization ranging from 15% to 85% [2] in major US metropolitan areas. These measurements seriously question the suitability of the current regulatory regime and possibly provide the opportunity to solve the spectrum

bottleneck.

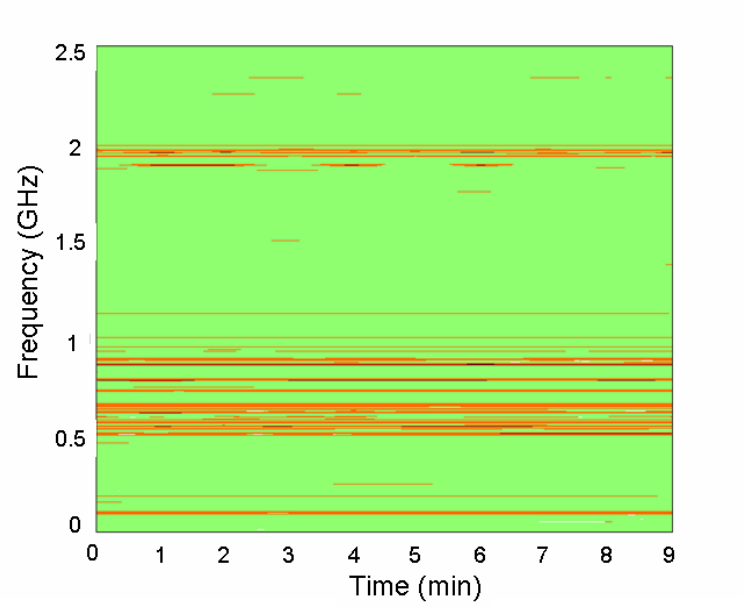


Figure 1.3: Temporal variation of the spectrum utilization (0-2.5 GHz) in downtown Berkeley

Unfortunately, creating a new spectrum allocation chart based on the usage distribution is not only impractical but also inefficient, because it is not possible to predict and optimize allocation that would suit all current and future wireless systems. Furthermore, any change in the spectrum allocation could create an opposition from the current users/owners of the spectrum. Therefore, the solution to this problem should preserve rights and access priorities of "primary" users.

One approach to improve the spectrum utilization is to allow a secondary use of the spectrum, that would be controlled by primary users. This approach would require enhancement of a primary user system infrastructure to provide either temporary licenses or implement markets for selling the unused spectrum. Clearly, there is a

substantial cost involved in its implementation that primary users might not be willing to invest at this stage, when competition for spectrum usage is not very high. A more flexible approach, that suits the current regime of spectrum under-utilization, is to allow secondary users to get access to frequency bands already allocated to primary users while these are not using it, but without direct primary user's control over access. One of the mechanisms to accomplish this is called "opportunistic spectrum sharing". Under the opportunistic sharing regime, secondary users are allowed to operate in certain frequency bands without the consent of the primary users of these bands, as long as they do not interfere with the primary user.

The simplest approach to share the spectrum without interfering with primary users is based on severe restrictions on transmitted power levels and operation over "ultra" wide bandwidths (Ultra Wideband or UWB transmission). This approach has been approved by the FCC [3] in 2002, and presents the first step forward to improve spectrum utilization through sharing. However, the UWB approach did not properly address the existance of "white space" in the spectrum utilization. After realizing UWB limitations and showing that some interference is acceptable, the FCC has issued a Notice of Proposed Rule Making [4] advancing "Cognitive Radio" technology as a candidate to implement more aggressive opportunistic spectrum sharing. The intention was to allow higher transmission power if the radio can reliably identify white spaces and prevent the interference to primary users.

Given the large amount of under-utilized spectrum, a cognitive radio technology

could be a solution that provides orders of magnitude improvement in system capacity. For example, if cognitive radios operate in a gigahertz wide bandwidths where the utilization is below 10%, the achievable throughput can be improved at least 10 times and the number of users can be 10 times larger with respect to narrowband systems that share a single 10 MHz channel. Taking into account temporal under-utilization, the number of cognitive radios users sharing a single band can be further increased. As a result, many communications services like cellular phones or wireless local area networks could support advanced applications such as real-time video or peer-to-peer networking. Also, this additional user capacity could enable integration of different technologies into ubiquitous wireless infrastructure needed in multimedia or home gateway networks.

More importantly, flexibility in usage of available spectrum bands through opportunistic sharing would allow deployment of heterogenous services based on location and application demands. This opportunity is particularly attractive for socially important services like public safety and emergency relief that require large throughput and large number of wireless users during relatively short period of time but with low latency and high priority. These high demands could not be accommodated using fixed frequency allocation since the cost and efficiency of spectrum resource usage would not be justified. Cognitive radio technology perfectly addresses these emergency situation scenarios given the ability to temporarily occupy large bandwidths with high priority access at virtually no cost per spectrum resource usage.

1.2 Overview of Related Work

Historically, the term cognitive radio was first introduced by Mitola [5] as "the point in which wireless personal digital assistants (PDAs) and the related networks are sufficiently computationally *intelligent* about radio resources and related computer-to-computer communications to: (a) detect user communications needs as a function of use context, and (b) to provide radio resources and wireless services most appropriate to those needs." Thus a cognitive radio is able to automatically select the best and cheapest service for a radio transmission and is even able to delay or bring forward certain transmissions depending on the currently or soon to be available resources. The learning and reasoning capabilities are key to cognitive radio operation simply implemented in software as a high layer functionality [5], [6]. The emphasis on software implementation of cognitive radios is also supported by Software Defined Radio (SDR) developers. However, this approach lacks description of specific radio architecture for physical and network layers needed to advance cognitive radio technology.

While at this time there are no cognitive radio networks in commercial deployment, the basic ideas of spectrum sharing with interference avoidance have already been demonstrated. In the middle of the 5 GHz band of 802.11a, there are two approximately 200 MHz wide bands (5150-5350 MHz and/or 5470-5725 MHz) which are shared with sensitive aeronautical navigation radars. While these radars have limited geographic distribution, it was deemed necessary to protect them from 802.11a trans-

missions. A technique called Dynamic Frequency Selection (DFS) is employed, which involves sensing if a radar signal is present and then avoiding those frequency channels to prevent interference. While the relatively simplistic strategy of abandoning the impacted 802.11a frequency channels is far from optimal in terms exploitation of the spectrum, it does indicate that the regulators are willing to accept that spectrum can be shared through sensing and avoidance.

There is also an evidence that regulators are willing to go even further and actually experiment with the cognitive radio operation by allowing shared use in the digital TV bands from 400-800 MHz [7], [8]. The primary application being promoted is to use the long range capability of the TV carrier frequencies to provide internet access in rural areas. The technology is currently being developed by IEEE 802.22 standard group [9], addressing the critical issues of interference protection set by broadcasters and designing a network architecture for mutual co-existence. Given the static TV channel allocations, this technology would be just a first step toward truly dynamic spectrum access. However, the technical proof of its feasibility would clear the way for regulatory approval and opening of other licensed spectrum for spectrum sharing.

Another natural band to allow cognitive radio operation in is the same 3 to 10 GHz band that is already in use by the UWB radios, given that utilization is very low. This band is sufficiently wide that the true advantages of cognitive radio operation could be demonstrated and the simultaneous availability of two approaches to sharing could actually be mutually supportive. The primary users of this spectrum are WiMAX

systems that are much more advanced than TV broadcast system. The strategy to avoid interference between UWB and WiMAX systems is called "detect and avoid" (DAA), which also incorporates basic cognitive functions like sensing and transmit power adaptation [10].

From the military perspective, DARPA has followed a much more general approach to cognitive radios and tried to define an architectural framework to implement policy based intelligent (neXt Generation) XG radios [11]. The XG framework proposes a cognitive radio system design which is quite different from the one presented in this thesis. Here, we briefly outline its main characteristics. XG radios are based on the Software Defined Radio solutions and do not address any specific frequency band or application at this time. As a result, the XG framework does not deal with the feasibility and implementation issues involved in physical and network layers of cognitive radios, which are main focus of this thesis. XG framework follows a high level abstraction principles for radio specific functions, assuming that radio and network behavior can be described through a set of policies and reasoning algorithms. XG radios have regulable kernels, which can be controlled via policy rules which would allow radios to transcend regulatory borders with simple policy changes. This so called "policy engine" approach would enable the world-wide operation of cognitive radios given different frequency allocations in different countries, and many military applications. Main research efforts are spent on developing the policy description language and reasoning algorithms for an implementation of policy rules.

The growing interest in cognitive radio research from signal processing and communication communities in academia and industry has spurred an increasing number of papers in the recent years [12], [13]. There are a large number of proposals for all communication layers, but the unified system architecture has not been clearly defined. In addition, most of these research results rely on theoretical analysis or computer simulations. However, these approaches are not sufficient or satisfactory to: i) influence and convince regulators to approve cognitive radios, ii) remove concerns of primary users or provide feasible technical requirements for co-existence, and iii) promote cognitive radio technology for wide range of applications.

Unfortunately, there has not been a clear definition of what actually constitutes a cognitive radio. A relatively conservative definition would be that a cognitive radio is network of radios that co-exists with higher priority primary users, by sensing their presence and modifying its own transmission characteristics in such a way that they do not yield any harmful interference. It is this sensing function and ability to rapidly modify their transmitted waveform that is the unique characteristic and challenge of cognitive radio implementation. More generically speaking, a cognitive radio is an environment-aware radio, i.e. it has awareness of the RF channel, location, user profiles etc. to which it is capable of adapting. However, in this thesis we shall constrain our attention to the spectrum awareness and adaptation aspects (spectrum agility), that impact the physical and network layers of a cognitive radio.

1.3 Thesis Organization

This thesis attempts to develop a system design framework for cognitive radio technology by covering all design aspects including theoretical analysis, implementation and experimental evaluation in realistic scenarios. To demonstrate this approach we focus on a novel spectrum sensing functionality that is tightly integrated in both physical and network layers.

Chapter 2 introduces a cognitive radio spectrum sharing technology. We define cognitive radios in a very broad sense with capability to share all three dimensions in signal space (frequency, time and space), and give examples of technologies being developed using these concepts. In this chapter, we also propose a cognitive radio system architecture with physical and network layer functionalities. Although in this thesis we focus in depth only on the spectrum sensing part of the system, this architecture allowed us to address problems in the system context rather than separate building blocks.

Chapter 3 presents a testbed platform that has been used for the design, implementation and experimental studies reported in this thesis. We provide a technical description of the testbed hardware and software architecture while putting an emphasis on its suitability for implementation of real-time signal processing and protocols. These capabilities in addition to support for networking and multiple antenna communication can be used for exploration of a wide range cognitive radios systems with the unified design flow.

In Chapter 4, we describe the problem of spectrum sensing that will be discussed throughout the rest of the thesis. We start by introducing its requirements and performance metrics, and give insights on the trade-offs between them in the practical implementations. Based on this problem formulation, we define the design space of possible techniques that can be used on a radio level and physical and network layers.

Chapter 5 provides an in depth analysis of physical layer techniques using digital signal processing for sensing primary user signals. The techniques involve detection of general signal parameters such as energy, pilots or features of primary user modulation types. While these techniques have been studied in the literature using theoretical models, we complement the theory with experimental data that shows fundamental limits and practical gains achievable in the system implementation. In addition, we explain why these sensitivity limits exist by identifying radio impairments that cause them, and develop techniques for robust physical layer sensing.

Chapter 6 explores the possibility of using network cooperation to improve the spectrum sensing reliability. Similar to the physical layer sensing analysis, we start from the theoretical predictions of achievable gains. Through the development and implementation of network sensing using the testbed, we identify different strategies for network cooperation, and experimentally characterize their performance and limitations in actual wireless environments. This study also reveals implementation constraints and required interfaces with physical layer sensing that should be used in a protocol design for spectrum sensing.

Chapter 7 addresses the design of cognitive radio sensing receiver front-ends to support spectrum sensing over a wide frequency bandwidth without sensitivity degradation. The main design issue is the need to accommodate large dynamic range signals. Conventional CMOS radio architecture cannot support wide frequency bands and large dynamic range signals, thus new approaches must be developed. We propose and discuss spatial filtering approach using multiple antennas as a way to selectively suppress strong blockers.

Chapter 8 concludes this work and proposes some steps for future research.

Chapter 2

Cognitive Radio Spectrum Sharing

Cognitive radios could provide a paradigm shift in the way that spectra is regulated and used. However, the novelty of this approach makes it difficult to leverage the experience of present wireless systems. There are many challenges across all layers of a cognitive radio system design, from its application to its implementation. A systematic framework for a cognitive radio system design needs to be addressed at the very early stage, so that system functions, models, and requirements can have corresponding metrics and that key questions could be addressed across a larger research community.

2.1 Spectrum Sharing

In spectrum sharing, choosing the radio transmit power that causes minimal interference to primary users presents the crucial challenge. In principle, the transmit

power constraint cannot be globally set so that it meets interference requirements at any location and time for arbitrary primary users. The extreme and quite conservative strategy would be to use to the UWB transmission [3] with severe transmit power constraints. UWB communication relies on the fact that if the bandwidth is increased, then reliable data transmission can occur even at power levels so low that primary radios in the same spectral bands are not affected. Figure 2.1 illustrates the transmitter power spectrum density profiles in UWB underlay spectrum sharing approach.

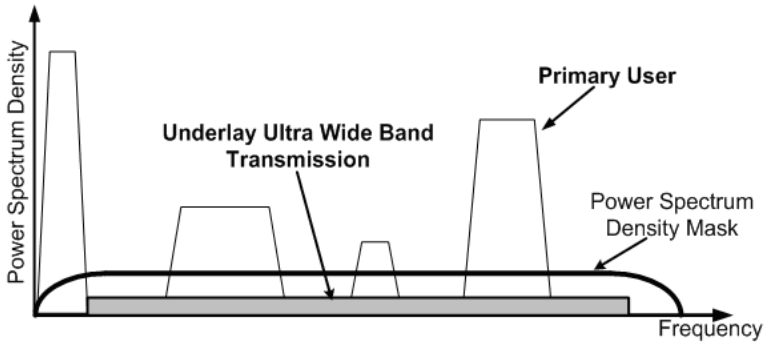


Figure 2.1: Transmit power allocation in UWB spectrum sharing

In order to overcome the transmit power limitations, systems are allowed to use very large bandwidth so that they can trade off the data rate for robustness. This strategy is suitable for noise-like channels where signal-to-noise ratio (SNR) can be improved by spreading or coding. However, spreading transmission power equally across a wide bandwidth could be largely sub-optimal in case of strong in-band interference.

Opposite to UWB, a cognitive radio approach does not necessarily limit the trans-

mission power, but rather attempts to share the spectra through a dynamic avoidance strategy and would use higher transmit power in "white spaces" to maximize capacity. Unique to cognitive radio operation is the requirement to sense the environment over huge swaths of spectrum and adapt to it since the radio does not have primary rights to any pre-assigned frequencies. Note that cognitive radios are allowed to transmit increased transmit power levels but must not cause co-channel and adjacent channel interference to primary users in the vicinity. Figure 2.2 illustrates the transmitter power spectrum density profiles in cognitive radio overlay spectrum sharing approach.

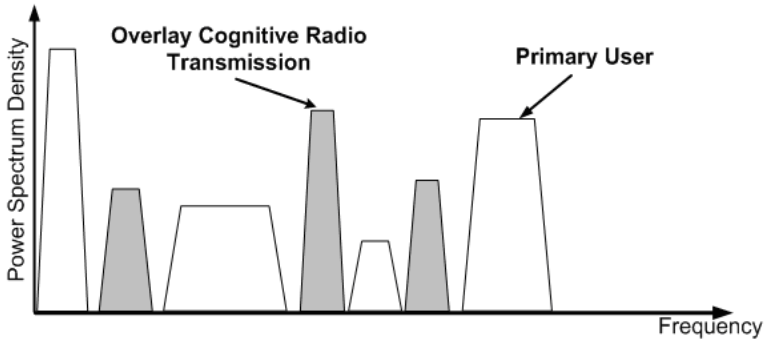


Figure 2.2: Transmit power allocation in cognitive radio spectrum sharing

Figures 2.1 and 2.2 present only spectrum sharing in frequency domain. The advantages of cognitive radio approach can be better understood if we consider spatial and temporal dimensions too. The total increase in system capacity using cognitive radio approach comes from the fact that cognitive radios are exploiting all three dimensions for sharing.

In order to illustrate cognitive radio spectrum sharing using spatial dimension we chose a scenario where two cognitive radio networks reuse two TV channel frequencies

(Figure 2.3). This scenario is currently being addressed by the IEEE 802.22 Working Group [9]. Spatial domain sharing deals with the frequency reuse which is virtually applied in any wireless system design. Frequency reuse is typically done through frequency planning, resulting in an inefficient utilization of frequencies across different geographical regions. On the other hand, cognitive radios can implement adaptive opportunistic frequency reuse on a much smaller spatial scale.

To guarantee interference requirements set for the frequency reuse, cognitive radios must sense primary user signals and ensure operation far enough beyond the service contour or protection region of a primary system, in this case TV broadcast. The service contour is defined by the minimum signal strength needed for reliable reception of a TV signal. If located further away from the service contour, cognitive radios can transmit at higher power levels and establish longer range communication links. This capability allows cognitive radio networks to scale with size of a geographical region with under-utilized spectrum.

Spectrum sharing in time domain is applicable when primary user systems use their assigned frequency band infrequently, in bursts, or during regular time intervals. Figure 2.4 illustrates the time and frequency domain usage of primary user spectrum when shared with cognitive radios. Clearly, exploiting time domain spectrum opportunities requires very dynamic spectrum access and presents the most challenging implementation of cognitive radio spectrum sharing. While exploiting the spatial dimension is the first step towards cognitive radio implementation in TV

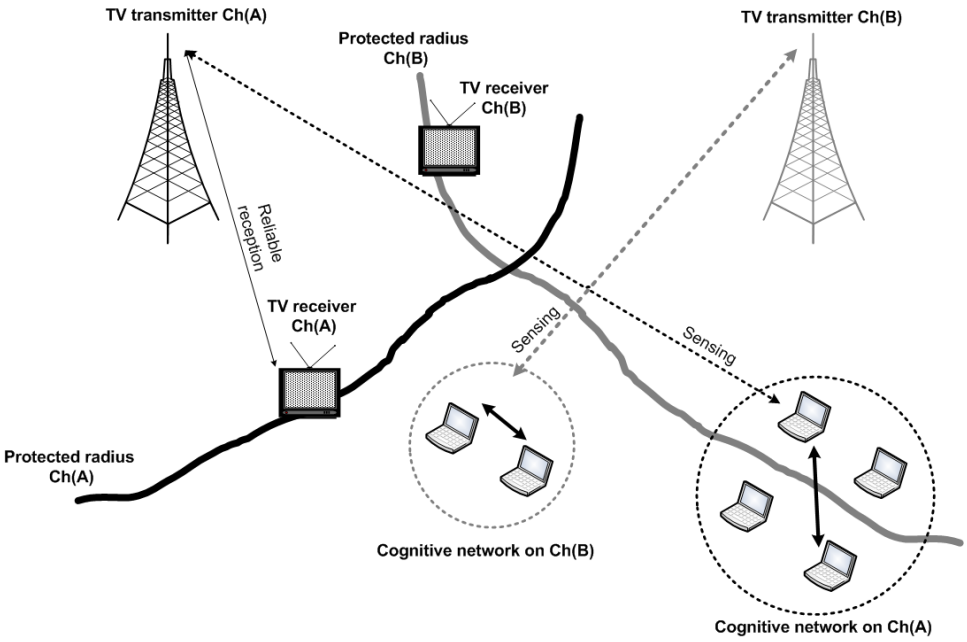


Figure 2.3: Spatial domain view of cognitive radio spectrum sharing

bands, future cognitive radio systems targeting other primary user bands will need to incorporate spectrum sharing in all three domains.

2.2 Cognitive Radio System

The next steps in developing and implementing cognitive radio technology require definition of a system architecture. Conventional communications systems architectures are defined and standardized using seven ISO/OSI layers. Even though cognitive radios are quite different from traditional wireless systems, it is reasonable to assume that a cognitive radio framework would be based on ISO/OSI layering methodology. A further advantage of layering approach is the possibility to enhance existing layers

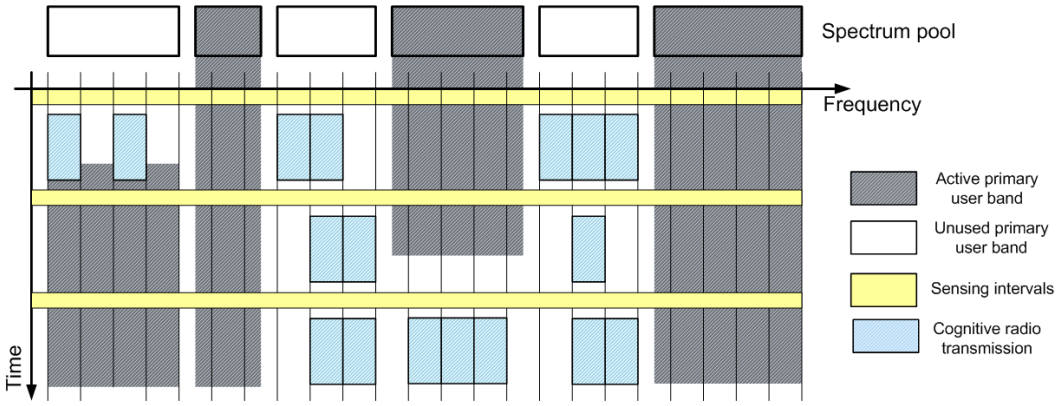


Figure 2.4: Time domain view of cognitive radio spectrum sharing

of conventional radios with new cognitive functionalities.

In [14] we proposed physical (PHY), link, and network layer functions of an envisioned cognitive radio system for opportunistic virtually unlicensed spectrum sharing. Here, we overview the proposed architecture and discuss the role of cognitive radio specific functions (Figure 2.5). Higher layers are not addressed since they are dependent on the specific application. This architecture provided basis and underlining framework for the research and development of a cognitive radio system presented in this thesis.

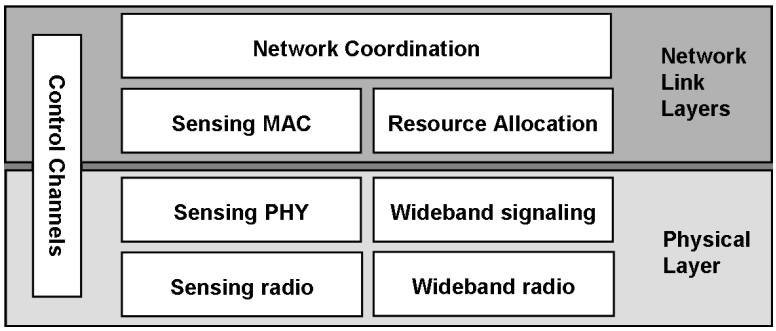


Figure 2.5: Physical and network layer functions of a proposed cognitive radio system architecture

2.2.1 Physical Layer Functions

Sensing radio and Sensing PHY

Since cognitive radios are secondary users of unoccupied spectrum they do not have *a priori* right to any frequency band. Their communication is strictly conditional on the reliable detection of primary user transmissions in their vicinity. As a result, cognitive radios must operate in a much wider frequency bandwidth than conventional radios which spans multiple primary user bands and perform frequent measurements of primary users' activity through spectrum sensing.

There are certainly a number of approaches that could be used to check the presence of primary user signals, like databases or beacons [4]. However, the only autonomous and flexible approach is based on measurements of the actual occupancy at a given location and time. Even in database and beacon approaches, spectrum sensing could add robustness and responsiveness to changes in the environment because it provides a real-time feedback of spectrum occupancy. Therefore, we argue that spectrum sensing should be considered as an essential part of any cognitive radios system. On the physical layer, spectrum sensing requires a wideband radio front-end for real-time spectrum measurements (Sensing radio) and a baseband signal processor of these measurements (Sensing PHY). Regardless of the application, spectrum sensing requirements are based on primary user modulation type, power, frequency and temporal parameters.

This new radio functionality involves the design of various analog and digital

circuits as well as signal processing techniques in order to meet challenging radio sensitivity requirements and wideband frequency agility. Given the novelty of this function and many issues involved in its design, architecture, and implementation, the main focus of these thesis will be on spectrum sensing.

Wideband radio and signaling

After identifying available spectrum opportunities using spectrum sensing, cognitive radios should establish transmission links for communication that provide the best spectrum utilization and capacity while avoiding interference to primary users. Properties of these communication links are quite different from those found in conventional narrow-band communication system that use a pre-defined modulation scheme. First, spectrum bands available for transmission could be spread over a wide frequency range, with variable bandwidths and band separations. Second, for optimal spectrum and power efficiency every cognitive radio estimates the quality of unoccupied frequencies in order to provide higher layers with signal-to-noise ratio and interference measurements needed for power and bit allocation. These measurements can be obtained from spectrum sensing. Lastly, different applications might require different selections of frequency bands based on propagation characteristics or power.

Clearly, cognitive radio approach requires a dynamic spectrum allocation supported by frequency agile modulation schemes tightly coupled with spectrum sensing function. In terms of implementation, there is a need for the development of trans-

mission scheme flexible enough to allow assignments of any band to any user and also scalable with the number of users and bands. In the ideal case, this flexible wideband transmission would be realized by digital domain waveform synthesis, where a set of parameters specifies transmission bands and power control [15].

2.2.2 Network and Link Layer Functions

Sensing MAC

While each radio performs spectrum sensing locally, the utilization of available spectrum resources must be controlled from the higher layers. Therefore, all spectrum sensing measurements must be shared among cognitive radios to establish a global view of the spectrum opportunities. The sharing mechanisms depend on the network architecture. In the centralized networks, an access point or a base station plays a role of a coordinator for exchange of spectrum sensing measurements. On the other hand, ad-hoc network architectures require robust distributed approaches for sharing. In both cases, a special protocol running in parallel with data medium access control (MAC) is needed to support sensing on the network layer (Sensing MAC). These two protocols must be tightly synchronized and responsive to changes in spectrum utilization caused by either primary user activities or other cognitive radio networks in the vicinity. On one hand network layer spectrum sensing adds an overhead and system complexity, but on the other it provides mechanisms to improve sensing reliability and possibly simplify the designs of Sensing radio and Sensing PHY.

Resource allocation

As stated earlier, cognitive radio approach shows promise to achieve orders of magnitude improvement in system capacity with respect to systems with fixed frequency allocations. The basic strategy to meet this goal is through adaptation and optimization of radio links across all available degrees of freedom (frequency, space, and time) identified through spectrum sensing. This function is typically achieved using the resource allocation implemented at the network layer. While resource allocation algorithms have been studied for a long period of time in different communication systems and networks, it is evident that cognitive radios approach bring new challenges.

In cognitive radio scenarios, available resources change on a variable time scale and are continuously updated through spectrum sensing. Therefore, algorithms for resource allocation and their physical implementation need to meet very strict timing constraints. In addition, the objective of resource allocation is not only to maximize a system capacity but also to meet interference constraints of primary systems. As discussed earlier, these interference constraints are met through adaptive transmit power allocation based on spectrum sensing measurements. In principle, if a radio can meet specified sensitivity levels, it should be allowed to transmit higher power levels as it is located far away from protected radius. Based on this rationale, there is an explicit trade-off between the sensing sensitivity and allowable transmission power [16] which requires integration of both sensing MAC and PHY with resource

allocation in cognitive radio systems.

Network Coordination and Control Channels

Clearly, the coordination of network layer functions for dynamic spectrum access is a very challenging design problem which has not been addressed in wireless communications systems designs so far. For the purposes of exchange of sensing information, adaptive resource allocation, and co-existence of multiple cognitive systems, we argue that network layer must be supported by standardized control channels. We proposed a hierarchical approach with two different kinds of control channels: a universal control channel (UCC) and group control channels (GCCs).

The UCC is globally unique, known to every cognitive radio *a priori*, whose main purpose is to allow co-existence of multiple cognitive networks. The need for the UCC arises from the requirement that while sensing the primary user band, noise or interference added by nearby cognitive users should be minimized, in order to achieve good sensitivity. In addition, when multiple cognitive systems are operating in the same spectrum pool then mutual interference and resource allocation should be coordinated through agreements on sharing strategies (e.g. time-division multiple access or carrier sense mechanisms). As a result, the UCC channel must cover large distances but its throughput requirements are fairly low. On the other hand the GCC is set up to exchange sensing information and support Sensing MAC, perform resource allocation and link maintenance. It is local within one group and thus has a

shorter range but possibly higher throughput than the UCC. There could be several active GCCs in the same geographic region.

Note that UCC and GCC are logical concepts, which need to be mapped to a physical communication channel.In the case of wideband cognitive radios systems, an implementation of control channels using UWB transmission is especially attractive. Their main advantage is that they are unlicensed with low impact on other kinds of communication thus could be continuously transmitted. Furthermore, multiple GCCs can be implemented independently using different spreading codes. However, there are severe limitations on transmit power of UWB emissions which limits their range. Since control channels require fairly low data rates, spreading gain can be used to increase the range. This example shows that UWB and cognitive radios can be complementary technologies.

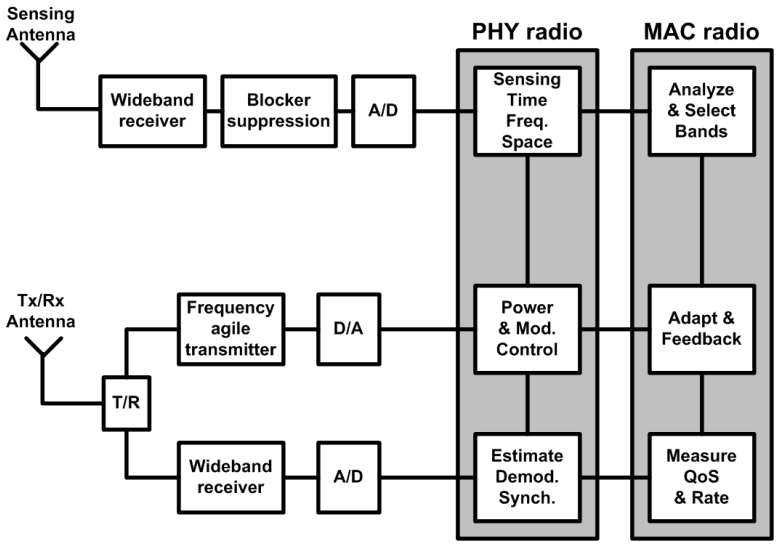


Figure 2.6: Block level diagram of a cognitive radio receiver

From this high level abstraction of system functions, we would like to develop a radio architecture that can support them with a feasible implementation. Taking further steps in this direction, we propose a functional level block diagram of a radio receiver (Figure 2.6). Note that due to its vital functionality for cognitive operation, spectrum sensing requires a separate receiver. As a result, it requires an independent antenna and radio front-end with sampling wide frequency band and processing the signals to find 'white spaces' in frequency, time, and space. Our goal is to understand the performance limits and implementation issues involved in the design of spectrum sensing while taking into account integration with appropriate radio blocks. For that purpose, we developed a testbed platform for implementation of radio blocks from Figure 2.6. Next chapter presents the testbed and its capabilities.

Chapter 3

Testbed for Cognitive Radio System Research

In this chapter we present an experimental testbed platform developed for exploration and demonstration of cognitive radio systems. The testbed is particularly suited for the development of physical and network layer functionalities and their experimental characterization in realistic wireless scenarios. Advanced capabilities such as real-time high-speed signal processing and protocol implementation, support for multiple network interactions and multiple antenna operation are described in detail this chapter. In the later chapters, this testbed is used for the design, implementation, and experimental studies of a spectrum sensing functionality.

3.1 Motivation and Requirements

The idea of cognitive radios has created a great interest in academic and industrial research. As a result, there is a large number of proposals for their physical and network layer functionalities. However, most of research studies and results that evaluate cognitive radios rely on a theoretical analysis or computer simulations. Given the novelty of functionalities and lack of proper models of shared spectrum environments, these analytical approaches cannot fully characterize cognitive radio performance. Furthermore, these studies often neglect practical system limitations and dependencies between different functionalities. It is evident that in order to enable this technology and fully understand system design issues in its implementation, proposed cognitive radios systems should be verified and demonstrated in realistic scenarios through physical implementation and experimental studies.

There is a considerable debate whether it is possible to build a cognitive radio system that does not disturb primary users. Unfortunately, this debate cannot be resolved on a theoretical basis. Clearly, primary user concerns can only be addressed by working systems that demonstrate that the amount of interference is sufficiently low to justify allowing opportunistic spectrum sharing. Furthermore, this demonstration is important to regulators whose confidence in cognitive radios is not sufficient enough to go forward with the regulatory framework that would allow this technology.

From the design perspective, it is far from clear what mechanisms are best suited to implement cognitive system architecture presented in Chapter 2. Conventional

design methodologies rely on layering and clearly defined interfaces between different functionalities, so that they can be independently developed. While there are a plethora of techniques in the literature proposing individual functions, none of them have been demonstrated as a part of the system in a real-world scenario. In addition to demonstration, it is desired to characterize their performance through an extensive experimentation in a predefined set of test cases and establish metrics for comparison of different implementations. The need for experiments is also stressed by the inability to realistically model all random sources encountered in a wireless channel, receiver circuitry, and interference environment.

Our approach to cognitive radio system design exploration is to use a testbed platform that provides flexibility and rapid reconfiguration for implementing different approaches, and allows experiments under controlled but realistic environments. A use of real-time testbed operation allows us to perform a large set of experiments for various primary user types, receiver settings, and network configurations. This particularly important for a comprehensive evaluation of a statistical behavior that computer simulation address with extensive Monte Carlo simulations. Besides addressing their performance limitations, a physical implementation of complex signal processing algorithms and protocols provides estimates of their hardware complexity.

In communications system research, testbeds have been predominantly used to test single point-to-point links and measure established metrics such as BER vs. SNR, and effective throughput under different wireless channel propagation environ-

ments. Certainly, cognitive radio research requires new testbed capabilities not only because of new functionalities but also due to complex interaction between physical and network layers that have to be addressed jointly. Based on the cognitive system architecture presented in Chapter 2 we map a set of requirements that a cognitive radio testbed should satisfy:

- Real-time baseband processing for spectrum sensing and agile transmission with high computational throughput and low latency
- Integration of physical and network layers through real-time protocol implementation on embedded processors
- Sufficiently wide bandwidth radio front-end with agility over multiple subchannels and scalable number of antennas for spatial processing
- Central processing of information exchange between multiple radios for controlled physical and network layer development and analysis (e.g. control channel implementation)
- Ability to perform controlled experiments in different environments, (e.g. shadowing and multipath, indoor and outdoor environments)
- Support for multiple radios which can be emulated as either primary users or cognitive radios

- Reconfigurability and fast prototyping through a software design flow for algorithm and protocol description

Several testbeds have been developed and used for cognitive radio experiments. Some of them use existing technologies for wireless access in order to modify and evaluate new protocols. For example, Open Access Research Testbed for Next-Generation Wireless Networks (ORBIT) [17] provides an experimental environment to evaluate protocols in different applications under real-world settings utilizing a radio-grid emulator that consists of radio nodes such as 802.11 a/b/g devices. Also, researchers at Virginia Tech use commercial-off-the-shelf (COTS) equipment to experiment with game theory approaches for distributed spectrum sharing protocols and implementation of a cognitive policy engine [18]. Unfortunately, without the access to the physical layer implementations new functionalities like spectrum sensing cannot be supported with these testbeds.

There is also an initiative to create a software defined implementation of cognitive radios and use open source programming model called GNU radio [19] to support various independent cognitive radio developments. While this approach provides great flexibility through software, it cannot support high computational throughput for real-time processing and controlled physical and network layer integration.

The cognitive radio testbed that fulfils all requirements outlined above was developed at Berkeley Wireless Research Center (BWRC) [20]. In the next section we describe the testbed architecture and capabilities.

3.2 Testbed Architecture

The BWRC cognitive radio testbed hardware architecture consists of Berkeley Emulation Engine (BEE2) [21], reconfigurable 2.4 GHz radio modems, and fiber link interface for connection between BEE2 and radios. The software architecture consists of Simulink-based design flow and BEE2 specific operating system that provide an integrated environment for implementation and simple data acquisition during experiments.

3.2.1 Real-time baseband processing

Real-time baseband processing implemented using BEE2 (Figure 3.1) consists of 5 Xilinx Vertex-2 Pro 70 FPGAs [22] in a single compute module with integrated 500 giga-operations per second (GOPS). These 5 FPGAs provide plenty of parallelism, which can be used to implement computationally intense signal processing algorithms even for multiple radios. In addition, FPGAs offer rapid reconfigurability which can speed up architectural exploration for the final ASIC implementation. Furthermore, the design complexity of FPGA designs (in terms of multipliers, logic slices, and memory) could be used for area estimation of an ASIC with the same functionality. Due to limited on-chip memory resources, each FPGA is connected to 4GB DDR memory. These memory capabilities are very useful for logging experimental statistics over a long period time for trace processing.

In addition to dedicated logic resources, each FPGA embeds a PowerPC 405 core

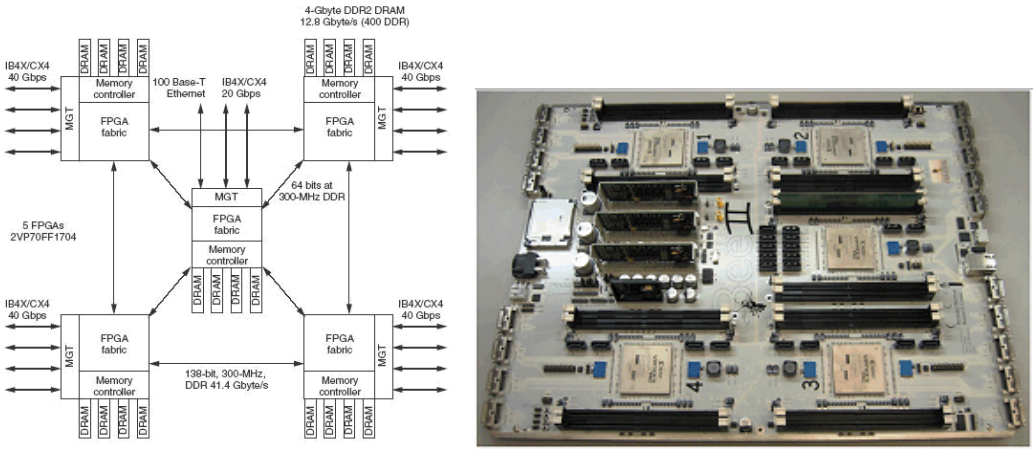


Figure 3.1: BEE2 system architecture (left) and BEE2 implementation (right)

for minimized latency between and maximized data throughput between microprocessor and reconfigurable logic. To support protocol development and interface between other networked devices, the PowerPC runs modified version of Linux and a full IP protocol stack. Since FPGAs run at clock rates similar to that of the processor cores, system memory, and communication subsystems, all data transfers within the system have tightly bounded latency and are well suited for real-time applications.

In order to interface this real-time processing engine with radios and other high throughput devices, multi-gigabit transceivers (MGTs) on each FPGA are utilized together with physical XAUI 4X (IB4X) electrical connection to form a 10 Gbps full-duplex links. There are a total of 18 such interfaces per BEE2 board allowing independent connections of 18 radios. Each individual MGT channel is software-configurable to communicate and exchange data at any rate below 10 Gbps.

3.2.2 Reconfigurable wireless modem

Support for wireless networking and flexibility to implement various sensing algorithms and adaptive transmission techniques require a highly reconfigurable radio modems. Such radio modem is designed to operate in an unlicensed 2.4 GHz ISM band where it can be adaptively tuned over 85 MHz of bandwidth with programmable center frequency and several gain control stages. Top level block diagram and implementation are presented in Figure 3.2. Both received signal strength (RSSI) and automatic gain control (AGC) are measured in real-time to support optimal signal conditioning on the receiver end. This is particularly important in spectrum sensing where any non-linearity or noise amplification greatly degrades receiver sensitivity. It also features dual antenna configuration for switched antenna diversity. This feature allows calibration of receiver noise and gain through signal processing rather than physical connection of a noise meter. The analog/baseband processing is implemented

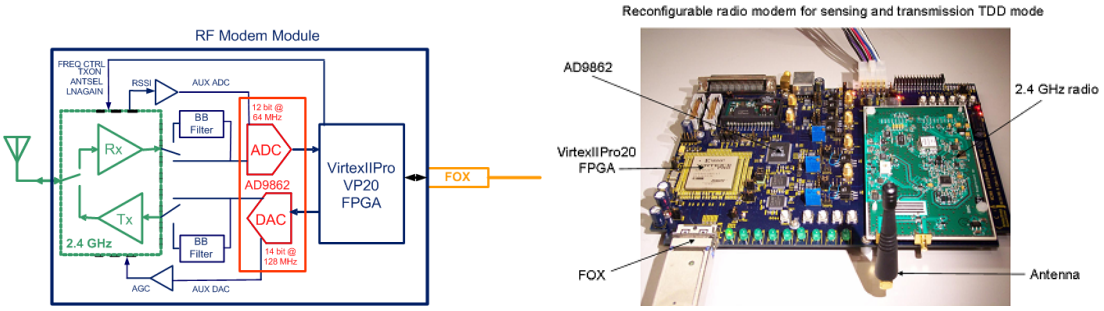


Figure 3.2: Reconfigurable 2.4 GHz radio modem: architecture (left) and implementation (right)

with a 14-bit 128 MHz D/A converters, 12-bit 64 MHz A/D converters, and 32 MHz

wide baseband filters. The high resolution of signal converters is chosen so that there is insignificant degradation of receiver sensitivity when digital signal processing algorithms are evaluated. Digital signal processing algorithms such as FFTs, matched filters, and others can be implemented with on-board Xilinx Virtex-IIPro20. In addition, this FPGA is used to implement radio control functions, calibration of analog impairments and real time access to programmable radio registers. In order to make A/D and D/A samples available for further processing on BEE2 side XAUI Infiniband connectors are also integrated on the radio board.

3.2.3 Networking Capabilities

While XAUI interface can provide a transfer of raw A/D bits (2 channels of 12 bits @ 64 MHz) and D/A bits (14 bits @ 128 MHz) at 2.5 Gbps, it can only extend up to couple of meters distance. For experimentation in network scenarios, it is necessary that radios can be spatially distributed while the information is centrally processed on BEE2. For that purpose, a fiber link between BEE2 and radio modem is deployed using optical transceivers compatible with XAUI Infiniband connectors. An optical cable can connect radio modems at distances up to 100m away from BEE2. In addition, the optical link provides good analog signal isolation on the front-end side from the digital noise sources created by BEE2. However, this optical interface introduces asynchronous operation between BEE2 and radio modems. The XAUI interface and protocol is implemented on both sides of the link for the synchronization

of data packets. In addition, Ethernet interface to BEE2 allows that other networked devices such as laptop computers and 802.11 equipment can be directly connected to BEE2. Therefore, a scenario that involves both cognitive radios and primary users can be emulated with this testbed, as illustrated in Figure 3.3.

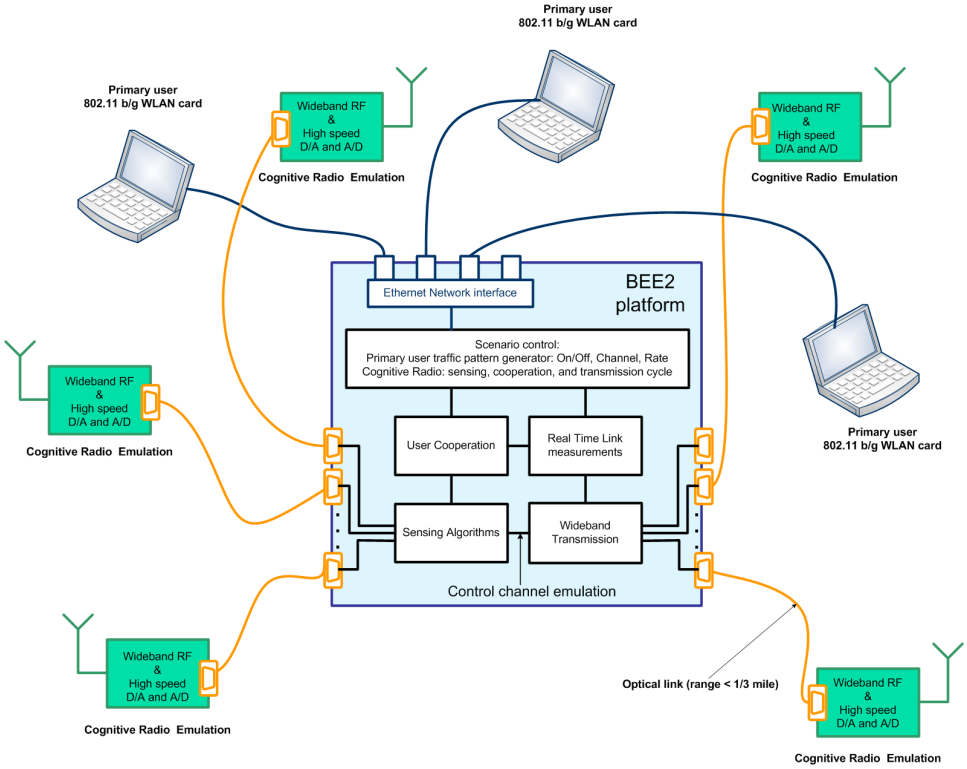


Figure 3.3: Emulation of cognitive radios and primary users using BWRC testbed

3.2.4 Multiple Antenna Capabilities

Advanced physical layer capabilities such as multiple antennas are highly desirable for exploration of techniques like antenna diversity, spatial filtering, or space-time coding. Given the large number of parallel interfaces multiple antenna front-end can

be created using existing hardware. However, a simple parallel connection of multiple radios does not create a multiple antenna transceiver. In order to ensure synchronous operation, every radio must have the center frequency derived from the common local oscillator (LO) and the same digital clock (CLK). Otherwise, uncorrelated noise and phase noise sources degrade the performance of multiple antenna system. The system for LO, CLK and power distribution is designed for scalability up to 16 antenna. Figure 3.4 presents a 4 antenna subsystem integrated with BEE2 platform.

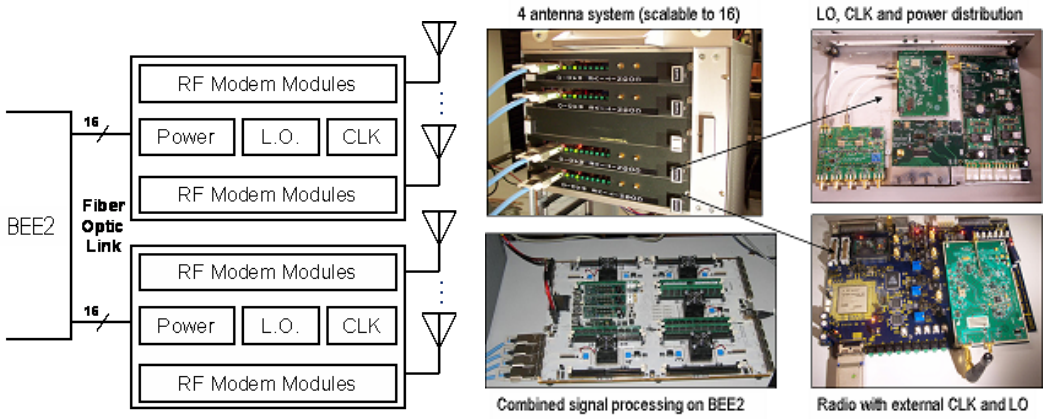


Figure 3.4: Multiple antenna radio front-end: architecture (left) and implementation (right)

3.3 Software Design Flow

In order to make effective use of the developed hardware, software design flow for rapid implementation and experimentation of algorithms and protocols is needed. Commonly used algorithm and protocol description languages are standard Matlab and C. Therefore, the software tool chain is built around Matlab/Simulink from Math-

works [23] coupled with the Xilinx System Generator [22] for mapping high-level block diagrams and state machine specifications to FPGA configurations. This environment supports simultaneous development of signal processing algorithms and digital design description for their hardware realization. Therefore no translation is required and allows signal processing researchers to realize hardware implementation of developed algorithms.

The original Xilinx System Generator library is enhanced by a set of parametrizable blocks to support interfaces with hardware components such as A/D converters, radio configuration registers and DDR memory. Simulink design is translated directly to FPGA configuration through BEE2 enhanced Xilinx Platform Studio (BPS) [24]. Furthermore, the tool provides the developer with hardware estimates of the design in terms of number of multiplications, logic slices, and memory. This is extremely important in the design optimization process.

One of the key features in the design flow is the ability to communicate and control hardware registers, block RAMs, DRAMs, and software running on control FPGA in real-time. This feature allows rapid post-processing of acquired signals in MATLAB or access to radio configuration registers during the experiment via automated scripts. Furthermore, it allows an implementation of protocols in C programming language and its direct integration with underlying hardware. This was enabled by enhancing Linux operating system through abstraction of hardware registers and memory on the user FPGA using file mapping [25]. BEE2 can be connected to the local area network

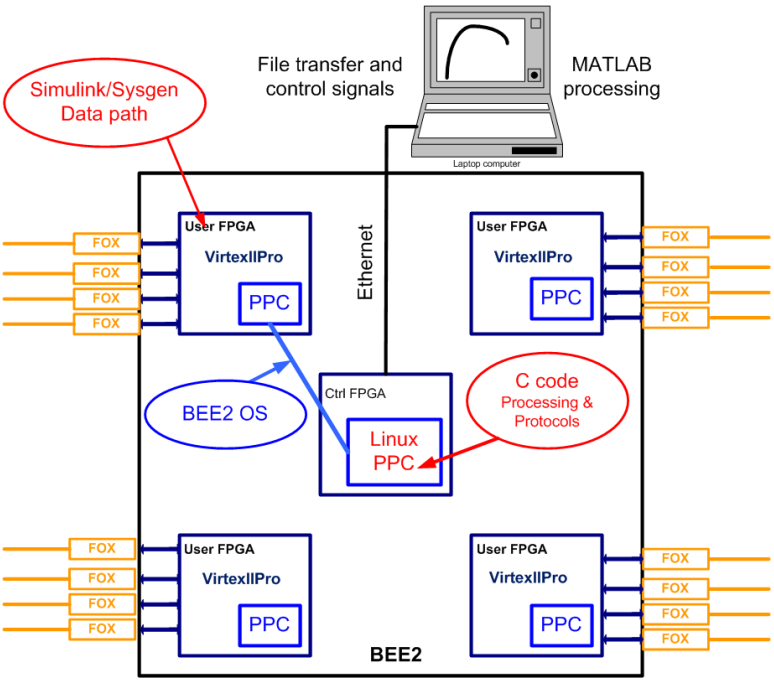


Figure 3.5: Software design flow for mapping of algorithms and protocols on BEE2 and experimental control

so that registers and memory can be accessed and transferred to laptops or PCs via Ethernet. Figure 3.5 illustrates the mapping process of algorithms and protocols on BEE2 as well as experiment control via Ethernet.

With defined cognitive system architecture and developed testbed platform for its implementation, we proceed with the development of spectrum sensing functionality.

Chapter 4

Spectrum Sensing Design

Spectrum sensing has been identified as a key enabling functionality to ensure that cognitive radios would not interfere with primary users, by reliably detecting primary user signals. In addition, reliable sensing plays a critical role on communication links of cognitive radios since it creates spectrum opportunities for them. In order to efficiently utilize the available opportunities, cognitive radios must sense frequently all degrees of freedom (time, frequency, space) while minimizing the time spent in sensing. Given the novelty and critical role of this function in a cognitive radio system, in the rest of this thesis we focus in depth on its design, implementation, and experimental characterization. The goal of this chapter is to define requirements and metrics for spectrum sensing functionality. We propose a cross-layer design approach in order to meet challenging requirements and solve critical design issues in its implementation.

4.1 Problem Formulation

In general, primary users have not been very receptive at the idea of cognitive radios and opportunistic spectrum sharing. In particular, they are concerned that cognitive radios will harmfully interfere with their operation. However in this argument, it is not very well understood what is considered harmful interference.

First example of harmful interference is when a cognitive radio may not be able to reliably detect a primary user signal and therefore may start sending although the primary user is using that frequency band. This is the classic "hidden terminal problem" in wireless networks where a receiver is unable to "hear" the transmitter and starts its own transmission thereby interfering with the intended receiver of the transmission. Second example is when a cognitive radio is using a frequency band that was deemed free by the sensing process but may not be able to reliably detect that a primary user has reappeared. Therefore, it may not vacate the frequency band quickly enough and therefore continue to send creating harmful interference to the primary user transmission.

From these two example, we can see that there are clearly two requirements on a cognitive radio sensor that influence the amount of harmful interference. In the first example, it is evident that sensor sensitivity plays a key role in reliable detection. In the second example, the sensing time needed to meet the required sensitivity is another requirement for sensing performance. We now discuss what primary and cognitive radio system parameters influence these two requirements.

4.1.1 Sensitivity requirements

Typical spectrum sensing scenario involving primary user system (transmitter and receiver) and cognitive radio sensor is illustrated in Figure 4.1. Since cognitive radios are secondary users of a primary user spectrum, their fundamental requirement is to avoid interference to primary user receivers in their vicinity. Note that primary user networks have no requirement to change their infrastructure for spectrum sharing with cognitive networks. Therefore, cognitive radio sensor must base its decision about primary user activity through reception of primary transmitter signal.

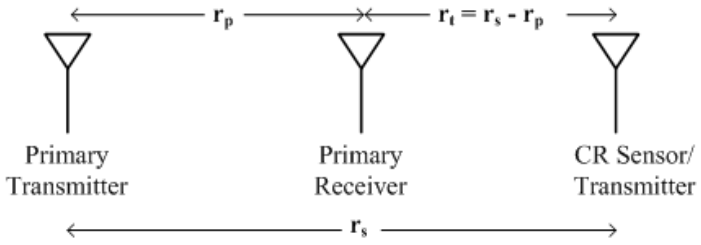


Figure 4.1: Spectrum sensing scenario

The difficulty in sensing transmitter signal strength is that neither distance nor channel between cognitive radio sensor and potential victim primary user receiver can be reliably estimated. Certainly, primary user signal decays with distance, thus spectrum sensing for far away radios becomes increasingly difficult. However, other wireless channel effects like multipath and shadowing can significantly degrade receiver signal strength independently of distance. Unfortunately, these two effects are main contributors to a hidden terminal problem. In order to overcome a hidden ter-

minal problem, cognitive radio sensor must be able to detect the signal affected by the worst case of these channel conditions.

Lets estimate the required sensitivity of a cognitive radio sensor needed to avoid a hidden terminal problem. A power level diagram in Figure 4.2 presents relationship between primary user and cognitive radio signals of interest. In the design of primary user systems, transmitter signal power is set so that receiver within certain distance have a minimum required signal to noise ratio (SNR) to successfully decode at the target rate R. For most communication systems decoding SNRs are in the range from 5 to 20 dB [26]. Signals received by a cognitive radio sensor are further attenuated by additional path loss, multipath fading and shadowing. Note that path loss exponents α's and fading margins Δ's in Figure 4.2 are dependent on the environment. Typical scenarios encounter 20 to 40 dB of fading margins [27], which is in effect sets the required sensitivity gain ρ. Typical path loss exponents are between 2 and 5 [28]. Taking these empirical numbers as a worst case condition, it can be estimated that sensing SNR falls into the negative range approximately from 0 to -35 dB. For example, for a cognitive radio operation in licensed TV bands IEEE 802.22 working group and FCC have defined required SNR sensitivities for primary user signals to be: -22 dB for DTV signals and -10 dB for wireless microphones [29].

Clearly, spectrum sensing is a very challenging receiver design problem where negative SNR detection reliability become key design issues. For complete characterization of sensing requirements, we now address sensing reliability through statistical

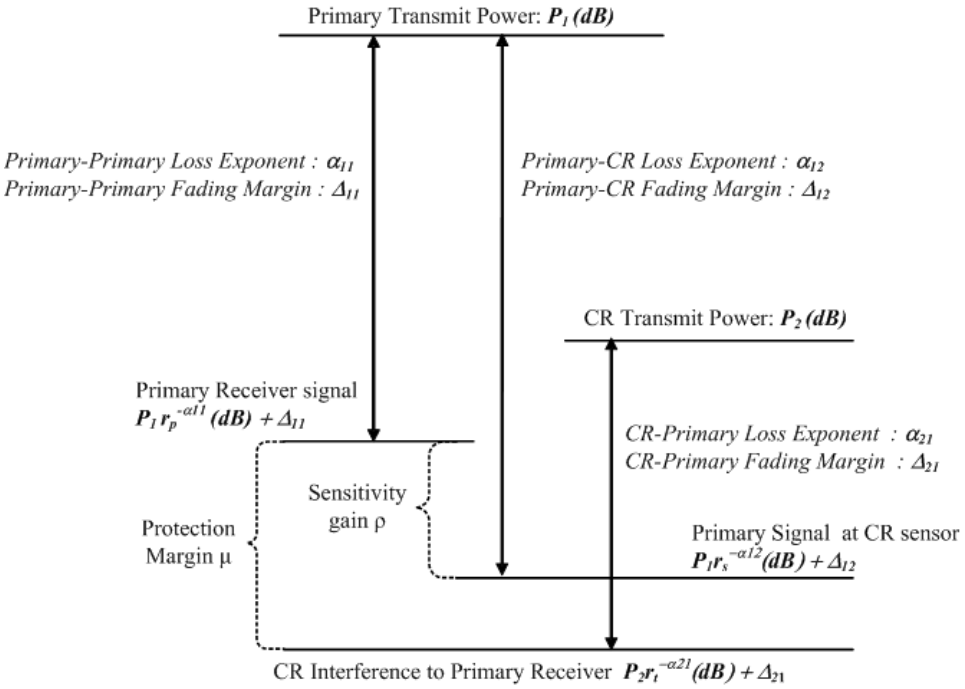


Figure 4.2: Power level diagram of relevant signals for spectrum sensing

parameters of sensing detection process. Unreliable sensing is defined by the total probability that sensor misses the detection of an active primary signal and results in harmful interference by a cognitive radio transmitter. First note that unreliable sensing can be caused by two different scenarios. First scenarios is when the received signal at the sensor is below the margins set for the worst case hidden terminal problem. Second scenario is related to the sensor's internal decision process about primary signal presence, when signal is above the required sensitivity level [30].

Given the conservative measures for the worst case fading margins, the probability of the first scenario is considered to be very small. Therefore, the sensor reliability is dominated by the detector performance in negative SNRs. This implies that probabil-

ity of harmful interference to a primary user is approximately equal to probability of missed detection P_{md} in the decision process, given that the received signal strength is above required sensitivity. Clearly, this P_{md} number is a subject of negotiation with primary users, and it depends on their QoS degradation tolerances. For example, practical numbers for P_{md} proposed by TV broadcasters and considered by IEEE 802.22 are in the range from 1% to 10%.

In summary, spectrum sensing sensitivity requirements are set by two parameters: minimum detectable signal strength or sensing SNR, and probability of missed detection under this sensing SNR.

4.1.2 Sensing time requirements

In order to constrain interference, spectrum sensing design should achieve orthogonality between primary and cognitive transmission across different degrees of freedom. Our discussion, so far, has focused on the spatial domain meaning that if the absence of a primary system can be detected reliably within a certain area, and, moreover, if secondary transmissions can be confined to this area, acceptable harmful interference is generated towards the primary system. Reusing TV bands that are not broadcast across the entire country are a prominent example of this paradigm.

In addition to spatial domain, cognitive radio need to consider the time domain to achieve orthogonality. This means that primary user frequency bands must be sensed periodically within some predetermined time interval. The sensing interval is set by

the primary user system QoS tolerances. For example, in the case of reuse of TV spectrum each DTV channel must be sensed every 2 seconds [9]. Given the broadcast nature of the TV system and steady channel assignments, this timing requirement is quite relaxed. On the other hand, many of the multi-access communications systems employed today have a bursty transmission behavior. If they are not operating under full traffic, their under-utilized spectrum can be shared with cognitive radios. By finding a means to predict the bursty medium access of such standards it is possible to take advantage of these transmission opportunities in the time domain [31]. For these packet oriented multi-access channels, sensing intervals are tightly constrained by the packet length and are in the order of milliseconds or less.

The sensing interval requirement presents the maximum time a cognitive radio sensor could spend for primary user detection. On the other hand, cognitive radio system objective is to utilize the available spectrum resources as efficiently as possible. From the cognitive radio system design perspective, an effective time used for sensing should be minimized since cognitive radios cannot communicate during this time. While minimization of sensing time is not a matter of regulation or primary user requirements (as long as the sensing time is less than sensing interval), it clearly impacts the feasibility and usefulness of a cognitive radio approach.

Our interest is to characterize the design space of spectrum sensing approaches in terms of achievable sensitivities and sensing times. Figure 4.3 illustrates the trade-off involved in these two requirements. Intuitively, better sensitivity is achieved at the

expense of longer sensing time. The minimum sensing time presents the fundamental limit, and identifies which primary user requirements cannot be met. The achievable region consists of various techniques that will be explored in details in later chapters.

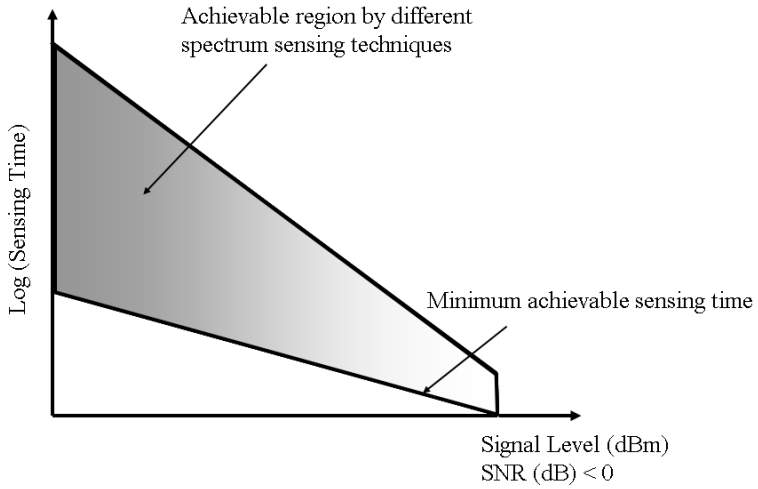


Figure 4.3: Trade-off between sensing time requirements and SNR

4.2 Cross-layer Design Approach

Spectrum sensing architecture designs depend on spectrum utilization regimes and specific primary user system. The most flexible architecture scalable to regimes with higher spectrum utilization should allow a cognitive radio to simultaneously sense the spectral environment over a wide frequency band [32]. Within a wide band of interest there could be different types of primary user systems with various ranges of operation and traffic patterns. We have seen that even though there are a large variety of primary user systems, cognitive radio's knowledge of their characteristics

and requirements for interference protection can be abstracted by a few generally applicable parameters.

Now, we discuss issues in the implementation of wideband spectrum sensing function. We argue that its challenging requirements must be addressed across multiple layers. Figure 4.4 presents the layers involved in the design and outlines the key challenges in their implementation. Sensing radio includes an antenna and RF/analog subsystems with mixing, amplification and conversion stages for reception of adequate signal strengths in primary user frequency bands. Physical layer sensing involves digital signal processing techniques for negative SNR detection, cognition, and adaptation to noise and interference environments. In addition, signal processing techniques can relax challenging requirements for analog circuits and A/D conversion, and enhance overall radio sensitivity. Higher layers involve a protocol for sensing and network management to coordinate spectrum sensing within and across cognitive radio networks. Through cooperation in the signal detection process, network sensing provides means to improve reliability for a given sensitivity requirement. In the subsequent sections, we elaborate the reasoning behind our approach.

4.2.1 Sensing radio

In conventional radios, receiver sensitivity is a quantitative measure that specifies the minimum signal level that can be detected for targeted modulation scheme and bit error rate (BER). However, the cognitive radio front-end is used to sense wide range

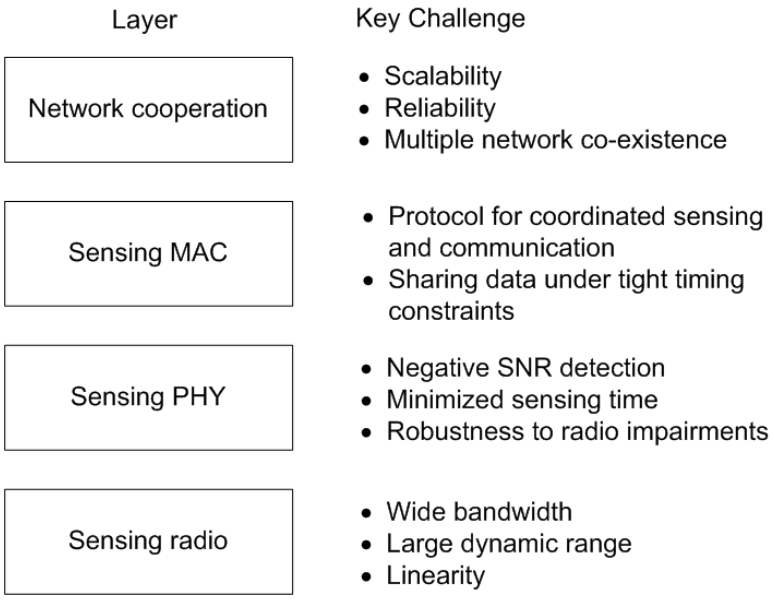


Figure 4.4: An overview of design challenges in spectrum sensing across different layers

of signals and modulations, thus such sensitivity metric does not apply. Furthermore, spectrum sensing requires detection of weak signals whose power levels are so low that they fall below the receiver noise floor where effective BER is 0.5. Under these circumstances, it is a paramount requirement to condition a received signal from the cognitive radio antenna to detector circuits with minimal signal-to-noise degradation and signal distortion. This degradation is determined by linearity, sampling rate, accuracy and power of the front-end circuitry.

The architecture of a sensing radio and its circuits' requirements depend on the regime where cognitive radio system operate. We distinguish three different regimes presented in Figure 4.5:

- **Low spectrum utilization ($N << K$):** In case bandwidth utilization by

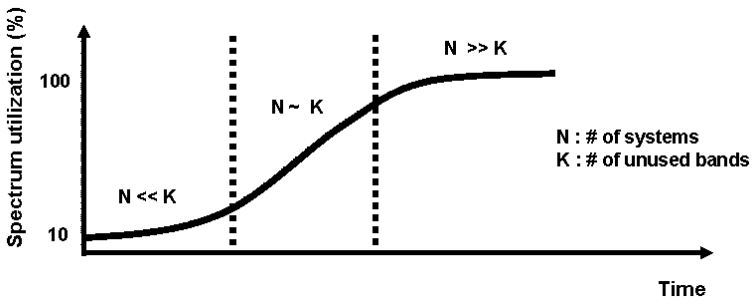


Figure 4.5: Different regimes of cognitive radio development

primary systems is low (e.g. below 10%) and there are limited temporal and spatial variations in the spectrum usage by the primary system, we can say that there is no spectrum scarcity for the cognitive radio network.

- **Medium spectrum utilization ($N \sim K$):** Over a period of time, the progression of cognitive radio technology will increase the spectrum utilization. The temporal and spatial variations will also increase and the number of unused bands will become comparable to the number of cognitive networks sharing the same spectrum pool.

- **High spectrum utilization ($N >> K$):** Eventually, there will be multiple cognitive networks competing for the same spectrum, so that spectrum will become truly scarce resource. In addition, there could be significant temporal and spatial variations due to dynamic allocation.

Low spectrum utilization regime

This regime corresponds to an early stage of a cognitive radio network deploy-

ment, which is a situation today with the abundance of spectrum available in some frequency bands. An example of such regime would be the TV bands, where current IEEE 802.22 standard is proposed. It has been shown that there are approximately from 5 to 20 (6 MHz wide) unoccupied TV channels available for cognitive operation in this band, especially in the rural areas. Due to continuous transmission and static geographic TV transmitter allocations, these bands show no temporal variation of primary user signals. Under these conditions, spectrum sensing can be done sequentially by searching one frequency band at the time through frequency sweeping using a tunable local oscillator (LO), as illustrated in Figure 4.6. The main challenge for this architecture is to design a wideband voltage controlled oscillator (VCO) with the tuning range over a band of interest (in the order of 1 GHz). A phase-locked loop (PLL) controlling the VCO should guarantee short settling time and small phase noise. The biggest advantage of this architecture is that the baseband portion of the radio is still narrowband which can provide a good selectivity and sensitivity. As a result, a fixed narrowband baseband filter for channel selection and low resolution, low speed A/D converter can be optimized for reduced cost and power.

Medium spectrum utilization regime

Once the number of unoccupied bands and competing systems become comparable cognitive radios cannot afford sequential sensing, as the probability of sensing the occupied band is increased. In addition, some bands (for example public safety or radar bands) are very infrequently used, but still cognitive radios must detect primary

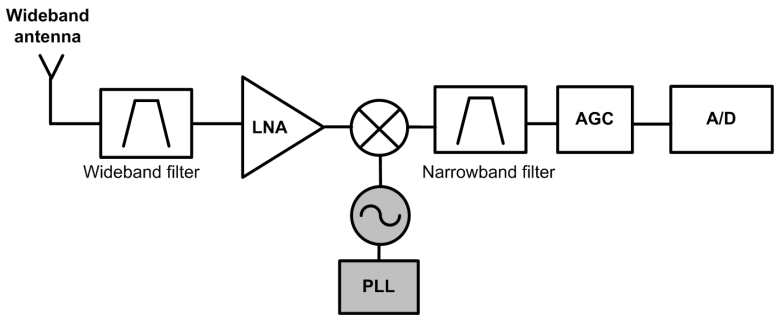


Figure 4.6: Sensing radio architecture for regimes with low spectral utilization

user reappearance and leave those bands. Therefore, the front-end circuitry should support simultaneous parallel sensing over several frequency bands, as illustrated in Figure 4.7.

From the circuits perspective, the number of radio components increases. For most practical solutions, the number of parallel paths should be bounded to four or five. Even though there are multiple LOs, which consume large amount of power for precision control, they are operating at fixed frequencies and therefore have relaxed requirements. In order to justify parallelization, the frequency planning should be sufficiently wide. As a result, the baseband portion of parallel radio channels has increased bandwidth and would require A/D converters with higher speed (100-300 MHz) and resolution (5-8 bits) than in the case of low spectrum utilization regime.

High spectrum utilization regime

If the radio spectrum environment is highly variable, both because of different types of primary user systems, propagation losses, and competing cognitive networks, then the implementation of the spectrum sensing function requires the highest degree

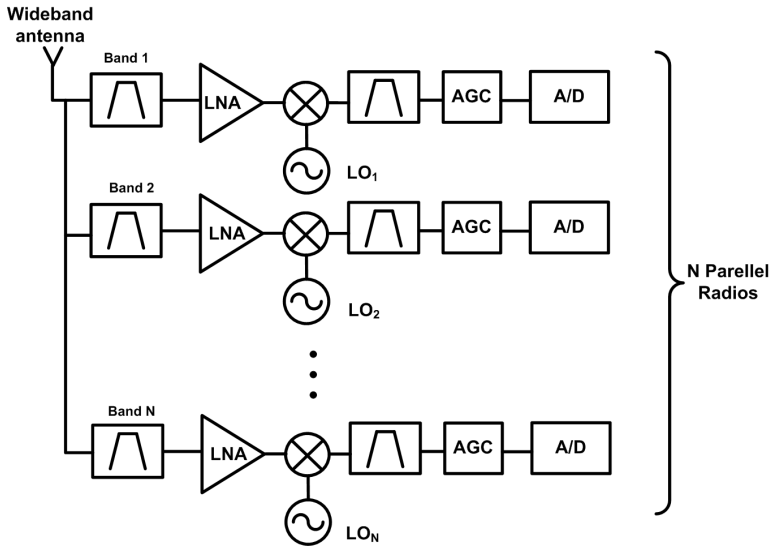


Figure 4.7: Sensing radio architecture for regimes with medium spectral utilization

of flexibility. Figure 7.1 depicts an architecture of a wideband RF front-end capable of simultaneous sensing of several GHz wide spectrum, for example. This architecture is commonly proposed for software-defined radios except for significantly narrower bandwidths. For cognitive radios, all front-end circuitry should be wideband, which further imposes high speed sampling requirement for the A/D converter. In addition, the RF signal presented at the antenna of such a front-end includes signals from close and widely separated transmitters, and from transmitters operating at widely different power levels and channel bandwidths. As a result, the large dynamic range becomes the main challenge since it requires high linearity circuits and high resolution A/D converter. In this regime, reducing the strong in-band primary user signals is necessary to receive and process weak signals.

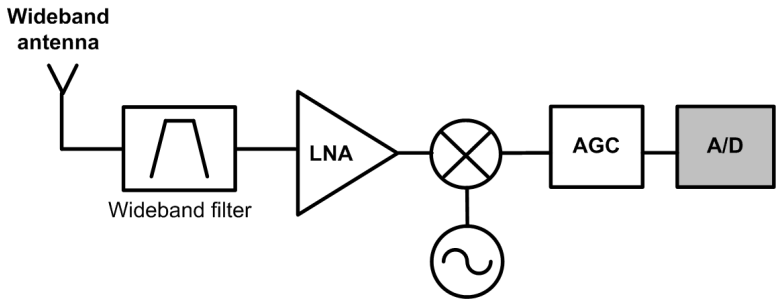


Figure 4.8: Sensing radio architecture for regimes with high spectral utilization

4.2.2 Physical layer spectrum sensing

After analog reception and sampling of a wideband signal, digital signal processing techniques need to be utilized to further increase radio sensitivity by processing gain. As discussed in Section 4.1.1 primary user systems require protection even in the worst case scenarios when the received signal could be far below noise floor. As a result, the new problem that cognitive radios face is the reliable detection in negative SNR regimes, where classical detection and demodulation methods used in positive SNR regimes might not be adequate. While in theory negative SNR regimes often do not need to be treated separately, actual implementation of signal detection might pose some limitations on the feasible SNRs [33]. Our goal is to verify these theoretical results through implementation and experiments. In case of disagreements, we aim to provide adequate signal models to support further theoretical studies of sensing in negative SNR regimes.

In contrast to classical digital communication receivers, a relaxed requirement for spectrum sensor is that a primary signal does not need to be demodulated and de-

coded but only detected. Therefore, there is no need for synchronization, decoding or protocol compatibility, which can significantly simplify sensing PHY implementation. This setting opens the room for new signal processing techniques that are generally applicable to different classes of primary signals, provide desired processing gains, and have flexible implementation so that wide bandwidths can be sensed.

The general approach for physical layer sensing, that we will follow, relies on the estimation of a specific primary user property or parameter (e.g. energy or a pilot signal). This estimated parameter is then used in the detection process, implemented as a simple hypothesis testing. This approach also allows identification and classification of primary user signals based on the knowledge of their unique characteristics. Chapter 5 provides detailed analysis and experimental results obtained by cognitive radio testbed of different signal processing techniques for physical layer spectrum sensing.

4.2.3 Network layer spectrum sensing

Physical layer sensing allows each radio to identify available spectrum opportunities locally. From the system design perspective, multiple radios need to simultaneously perform spectrum sensing and jointly decide on available spectrum opportunities that will be used for spectrum allocation. Therefore, spectrum sensing needs to be incorporated at the network layer. Clearly, the design of sensing functionality becomes more complex involving many dependencies and requiring robust cooperation

among radios.

While cooperation has become a common strategy to improve network performance in various settings (e.g. sensor networks), in the context of spectrum sensing for cognitive radios cooperation brings unique challenges. Implicitly, cooperation requires means of sharing sensing information among cognitive radios. Therefore, a common control channel is needed to coordinate sensing of the group of cognitive radios. Using a control channel, a cognitive radio network should implement a designated sensing protocol in parallel with the regular medium access protocol. Both latency and capacity of the control channel need to be minimized and scalable to large networks. In addition to scalability, control channel should be reliable and maintained by multiple cognitive radio networks.

With support of network infrastructure for cooperative sensing, cognitive radios can exploit inherent diversity in individual measurements and improve detection by relying on users with favorable channel conditions. The achievable gains are function of many different parameters involving underlying physical layer sensing, wireless channel characteristics and a size of the network. In order to fully understand benefits from cooperation in spectrum sensing, we use our testbed to quantify achievable gains and establish design guidelines. For example, one of the questions of interest is: Given the network topology what is the sufficient number of cooperative users that meets the required sensing reliability? Chapter 6 addresses a number of questions related to cooperation gains through analysis and experimental results.

Chapter 5

Physical Layer Spectrum Sensing

In this chapter we explore system design approaches for spectrum sensing performed on the physical layer of a single radio. The design space is diverse as it involves various types of primary user signals, traffic patterns, and interference requirements. In order to provide flexibility and take advantage of digital CMOS implementation, we focus on digital signal processing techniques. In case of spectrum sensing the need for signal processing is two-fold: improvement of radio front-end sensitivity by processing gain and primary user identification based on knowledge of the signal characteristics.

We start with the in-depth analysis of conventional signal processing techniques and identify regimes where these techniques are applicable. Our study is complemented by experimental data that shows fundamental limits and practical gains achievable through signal processing. The goal of this chapter is to present a unified theoretical and practical system design view of spectrum sensing functionality

together with guidelines for its implementation in wideband cognitive radio baseband processors.

5.1 Signal Processing Design Space

As describer in Chapter 4, the key challenge of spectrum sensing is the detection of weak signals in noise with a very small probability of miss detection. Since spectrum sensing is considered as a class of a detection problem, we first review existing approaches in the literature. First we focus on two classical types of signal detection problems: i) detecting known signals in additive noise (*coherent detection*) and ii) detection of random signals in additive noise (*non-coherent detection*).

These two cases give lower and upper bounds on performance of all other derived detectors. Coherent processing means that signal can be optimally process through a matched filter [34], which achieves the highest processing gain and reduction of noise. It also allows for recognition of specific signals like pilots, for example. Among the non-coherent methods, energy detector or radiometer has been used in wide range of applications, mostly because of its simplicity. Early work by [35], [36] showed that the radiometer is the optimal detector when the receiver knows only the power of the signal and noise. However, in cognitive radio applications this assumption is often not valid and recent theoretical work by [33] showed that the presence of uncertainties in noise power severely limits radiometer performance.

In addition to these two conventional approaches, we investigate the possibility

to use statistical signal processing techniques for recognition of features present in modulated signals [32] for spectrum sensing in negative SNR regimes. The theoretical framework for modeling and analysis of modulated signal properties is based on the property of cyclostationarity [37]. Many techniques for signal interception and classification were developed based on this framework [38], and so far have been mostly used in military applications with positive SNR regimes. However, there are a number of questions, that are missing in the literature, related to their performance in negative SNR, comparison with coherent/non-coherent approaches, robustness to noise and interference, and implementation issues. While this approach holds promise to solve the sensing problem for a wide class of primary user signals, the answers to these questions are needed to further pursue research on feature sensing techniques for specific primary user bands.

In the next three sections we develop, evaluate, and compare sensing approaches based on energy, pilot and feature detection, respectively. Based on these studies, we make design recommendations for physical layer implementation of spectrum sensing.

5.2 Energy Detection

The most basic approach for detecting signals in the presence of noise is based on energy measurement. Also, it is the most general technique since it applies to any signal type. It requires minimum information about the signal, including only signal bandwidth and carrier frequency. In communications and signal processing literature,

energy detection is studied as a hypothesis testing problem and performance is measured by a resulting pair of detection and false alarm probabilities (P_d, P_{fa}). While the positive SNR regimes have been very well understood, it is unclear if the same analysis and performance apply in the highly negative SNR regime.

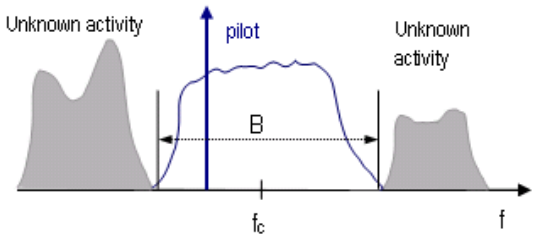


Figure 5.1: Spectrum picture at the input of sensing radio for primary user signal detection

5.2.1 Theoretical performance

Let us revisit the theoretical analysis first. The detection is the test of the following two hypotheses: under H_0 primary user signal is not present and there is only noise at input of the receiver, and under H_1 primary user signal and noise are present at the input of the receiver. Given that the center frequency f_c and bandwidth B of the primary user signal are known, input signal is downconverted and sampled at the Nyquist rate, $f_s = 2B$. Discrete time model of this hypothesis test is:

$$H_0 : y[n] = w[n] \quad n = 1, ..., N \tag{5.1}$$

$$H_1 : y[n] = s[n] + w[n] \quad n = 1, ..., N \tag{5.2}$$

where N samples represent the observation interval equivalent to the sensing time.

Both signal $s[n]$ and noise $w[n]$ samples are modeled as independent Gaussian random variables with zero mean and variance σ_s^2 and σ_w^2, respectively. A decision statistic for energy detector is measured energy over N samples:

$$\varepsilon(y) = \sum_{n=1}^{N} y[n]^2 \tag{5.3}$$

In order to compute the P_d and P_{fa} we need to determine the pdf of decision statistic under both hypothesis. Because the test statistic is a sum of N $i.i.d$ gaussian random variables then its pdf is chi-square χ_N^2 [39].

The detection is performed by a threshold test on the measured energy. There are number of ways that a threshold can be set. In the spectrum sensing situations, threshold γ is set to meet the fixed P_{fa}.

Then,

$$P_{fa} = P_r(\varepsilon(y) > \gamma | H_0) = Q_{\chi_N^2}(\frac{\gamma}{\sigma_w^2}) \tag{5.4}$$

$$P_d = P_r(\varepsilon(y) > \gamma | H_1) = Q_{\chi_N^2}(\frac{\gamma}{\sigma_w^2 + \sigma_s^2}) \tag{5.5}$$

Note that P_{fa} depends only on the noise variance, thus the threshold can be set regardless of the primary user signal level.

For a large number of samples N, (larger than 250 [36]), χ_N^2 can be approximated

with a Gaussian random variable, i.e. $\chi^2_N \sim \mathcal{N}(N, 2N)$. Then, we can re-express P_d and P_{fa} as:

$$P_{fa} = Q(\frac{\frac{\gamma}{\sigma_w^2} - N}{\sqrt{2N}}) \tag{5.6}$$

$$P_d = Q(\frac{\frac{\gamma}{\sigma_w^2+\sigma_s^2} - N}{\sqrt{2N}}) \tag{5.7}$$

If the number of samples used in sensing is not limited, an energy detector can meet any desired P_d and P_{fa} simultaneously. The minimum required number of samples is a function of the signal to noise ratio $SNR = \sigma_s^2/\sigma_w^2$:

$$N = 2[(Q^{-1}(P_{fa}) - Q^{-1}(P_d))SNR^{-1} - Q^{-1}(P_d)]^2 \tag{5.8}$$

In the low $SNR << 1$ regime, number of samples required for the detection, that meets specified P_d and P_{fa}, asymptotically scales as $1/SNR^2$. This scaling law is characteristic for non-coherent detection, i.e. detectors whose sensing time scale as $1/SNR^2$ will be termed as non-coherent.

5.2.2 Implementation and experimental results

In terms of implementation, there are a number of choices for energy detection based sensors. The main design goals are to optimally filter the signal bandwidth, and minimize the contribution of the out-of-band noise and interfering signals for increased sensitivity. Analog implementations (Figure 5.2 (a)) require analog pre-

filter with fixed bandwidth which becomes quite inflexible for simultaneous sensing of narrowband and wideband signals. Digital implementations offer more flexibility by using FFT based spectral estimates. FFT based architecture inherently supports various bandwidth types and allows sensing of multiple signals simultaneously. The size of the FFT is the critical parameter because larger FFT size improves the bandwidth resolution and detection of narrowband signals but at the same time increases the sensing time. Figure 5.2 (b) presents the architecture for wideband energy detector spectrum sensor.

In practice, it is common to choose a fixed FFT size to meet the desired resolution with a moderate complexity and low latency. Then, the number of spectral averages becomes the parameter used to meet the detector performance goal. We consider this approach in our experiments. The energy detector is implemented on the FPGA of the reconfigurable wireless modem. It is designed using 1024 point FFT with a fully parallel pipelined architecture for the fastest speed. Therefore the measured sensing times do not include any computational latency. Due to A/D sampling at 64 MHz, this implementation has 62.5 kHz FFT bin resolution. Each block of FFT outputs is averaged using an accumulator with programmable number of averages.

The goal of our experimental study was to evaluate and verify the theoretical results on the performance and possible limitations of the wideband energy detector. The measurements were performed for a 4MHz wide QPSK signal at 2.485 GHz carrier, in the SNR regime from -7 to -23 dB. For each signal level, we collected two

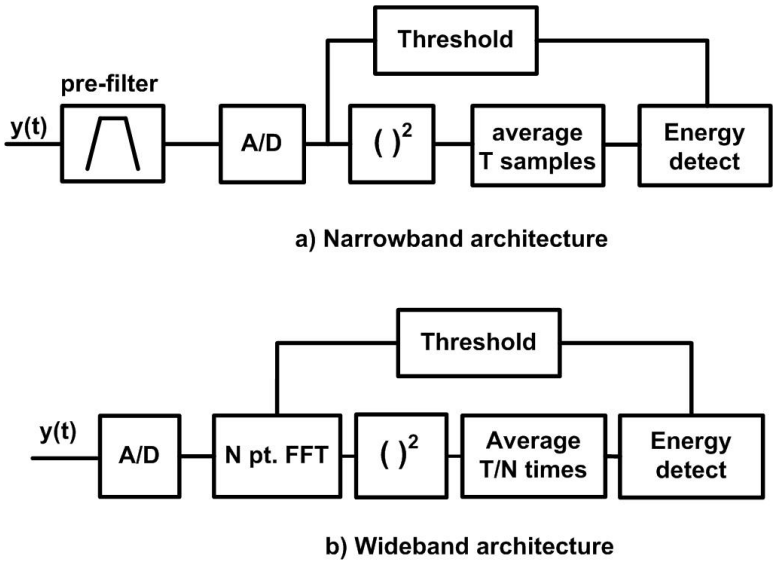


Figure 5.2: Energy detector implementations: a) narrowband architecture, b) wideband architecture

sets of energy detector outputs: one in the presence, and the other in the absence of the QPSK signal. From "no input signal" data, we estimated the detection threshold to meet the specified $P_{fa} = 5\%$. Then, we applied the threshold to the data where signal was present and computed the P_d. In order to accurately estimate the P_d and P_{fa} each detection measurement is repeated 1000 times. From these measurements, sensing time vs. SNR relationship is derived for $P_{fa} = 5\%$ and $P_d = 90\%$. Figure 5.3 presents the experimental results.

Results show that the theoretically predicted performance holds for SNRs above -20 dB. However, below -20 dB SNR the detection becomes progressively harder and at -23dB signal could not be detected regardless of the sensing time duration. Unfortunately, this deviation has two major consequences. First, energy detection is not a robust method for sensing of very weak signals. Second, this deviation

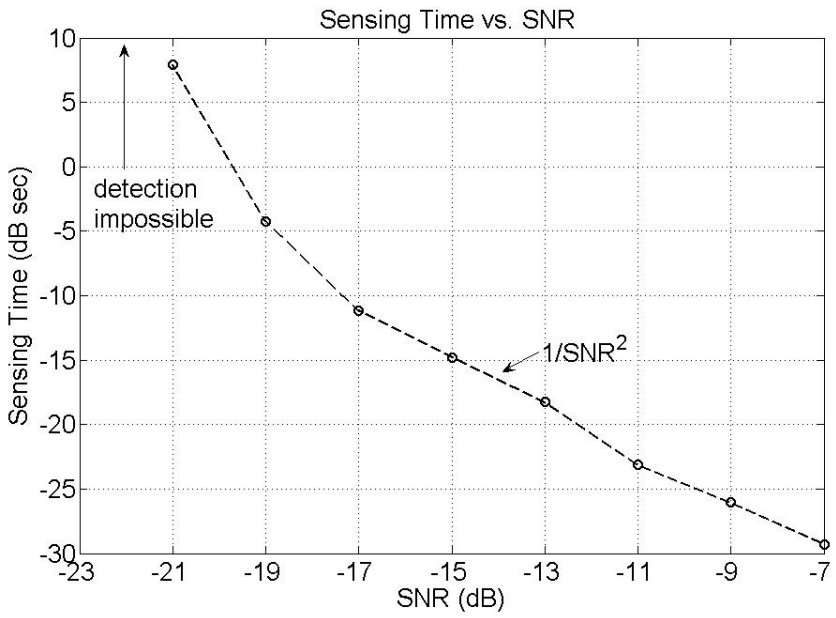


Figure 5.3: Measured sensing time vs. SNR for energy detector

from theoretical results shows that modeling of highly negative SNRs regimes is not adequate.

5.2.3 Limitation due to noise uncertainty

In order to understand why the detection suddenly becomes impossible we need to revisit our signal model. There, we have made two very strong assumptions that are typically made in communications system analysis. First, we assumed that noise is white, additive and Gaussian wide-sense stationary process, with zero mean and known variance. However, noise is an aggregation of various sources including not only thermal noise at the receiver and underlined circuits, but also interference due to nearby unintended emissions, weak signals from transmitters very far away, etc.

Second, we assumed that noise variance is precisely known to the receiver, so that the threshold can be set accordingly. However, this is practically impossible as noise could vary over time due to temperature change, ambient interference, filtering, etc. Figure 5.4 shows the measurement of the noise power level in the receiver used for testing of energy detection. Therefore, for signal detection in highly negative SNRs the noise model needs to incorporate this temporal changes.

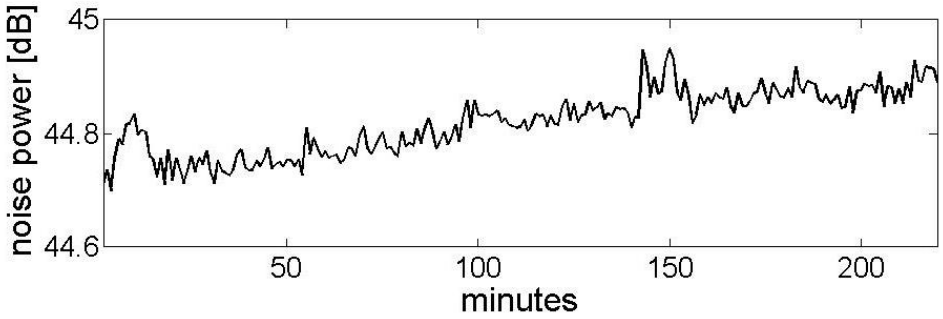


Figure 5.4: Noise power variation over time

This deviation from theoretical model for the time varying noise process become particularly important when the signal strength is below the estimation error of the noise variance. In that case, the detection threshold, which is set based on estimated noise variance, is set too high and weak signals could never be detected. Although a sensing receiver periodically estimates noise power and compensates for slowly varying temperature change, there is a resulting estimation error due leakage of the adjacent band interference. Quantitatively, if there is a residual x dB noise estimation error, then the detection is impossible below $SNR_{wall} = 10\log_{10}[10^{(x/10)} - 1]dB$ [40].

In our experimental study SNR_{wall} occurs at -23 dB which corresponds to approx-

imately $x = 0.03$ dB of residual noise uncertainty. Note that the maximum noise level variation was 0.2 dB and noise estimator was using the same amount of time to sense the signal and estimate the receiver noise variance. In typical cognitive radio applications, an adjacent band interference due to primary blockers could be much higher, and also the amount of time to estimate noise variance could be much shorter than a sensing time. In that case, we can expect that SNR_{wall} will be significantly higher than -20 dB. Thus, for primary user bands with stringent sensitivity requirements energy detector is not a viable sensing method.

5.3 Pilot Detection

While energy detection is a fairly general approach, it neglects the presence of deterministic signals like pilots that primary users embed in their transmission in order to perform synchronization and acquisition. The power of the pilot tone is typically 1% to 10% of the total transmitted power. One special case of a pilot signal, frequently present in primary user broadcast systems, is a sinewave tone used for receiver synchronization (Figure 5.1). On the other hand, packet based communications systems use short packet preambles as pilots. The conventional methods for synchronization of pilots and preambles are known for positive SNR regimes [41],[42]. Here, we will examine techniques of sensing weak pilots in negative SNR regimes and address their practical feasibility through implementation and experiments.

5.3.1 Theoretical performance

The benefit of pilot signals, if they are perfectly known to cognitive radio sensor, is that it can be processed coherently. Coherent processing achieves the best possible robustness with respect to noise. The detection problem is similar to energy detector case, where the receiver needs to decide between two hypothesis with targeted P_d and P_{fa}. Since the pilot is simply added to the data bearing signal, the optimal detector is the matched filter [43] that projects the received signal in the direction of the pilot:

$$T(y) = \sum_{n=1}^{N} y[n] x_p[n]^* \tag{5.9}$$

Under either hypothesis $T(y)$ is Gaussian.

$$T(y) \quad = \quad \begin{cases} \mathcal{N}(0, {\sigma_w}^2 \varepsilon_p) & \text{under } H_0 \\ \mathcal{N}(\varepsilon_p, {\sigma_w}^2 \varepsilon_p) & \text{under } H_1 \end{cases} \tag{5.10}$$

where $\varepsilon_p = \sum_{n=1}^{N} x_p[n]^2$. Now, probability of false alarm and detection are simply evaluated from the Gaussian distribution.

$$P_{fa} = P_r(T(y) > \gamma | H_0) = Q(\frac{\gamma}{\sqrt{\sigma_w^2 \varepsilon_p}}) \tag{5.11}$$

$$P_d = P_r(T(y) > \gamma | H_1) = Q(\frac{\gamma - \varepsilon_p}{\sqrt{\sigma_w^2 \varepsilon_p}}) \tag{5.12}$$

where again threshold γ is set to meet desired P_{fa}.

If the sensing time is not limited then both desired P_d and P_{fa} can be simultaneously met. The required number of samples is a function of the pilot $SNR_p = \varepsilon_p^2/\sigma_w^2$:

$$N = [Q^{-1}(P_{fa}) - Q^{-1}(P_d)]^2 SNR_p^{-1} \tag{5.13}$$

Since matched filter pilot detection uses the optimal processing, it could achieve detection with the minimum possible number of samples. Therefore, the scaling law of $N \sim 1/SNR_p$ gives a lower bound on the sensing time performance for any possible sensing detector type. However, if pilot is weak with respect to the energy of the signal then the required number of samples can be higher than for the energy detector.

5.3.2 Energy vs. pilot detection: sensing packets with preambles

The trade-off between energy and pilot detection is best demonstrated in the case of sensing primary user systems with packet communication. Note that in such cases sensing time is limited by the packet length. As a result, the energy in the packet is constrained, and the energy in the pilot is determined by the length of the preamble. Clearly, there are several approaches to perform the detection. One approach is to perform only energy detection, second would be to coherently process a preamble, while the third approach would be to jointly detect preamble and remaining energy in the packet.

A preamble data is typically a pseudo-noise sequence which is known to have good autocorrelation properties. Therefore, if a cognitive radio sensor knows the packet structure then it can perform coherent processing of a preamble through a set of matched filter correlators. This method can also facilitate the optimal energy detection in the case of asynchronous packet communication. In that case the hybrid approach of combined preamble and energy detection can be applied, which could improve overall sensing performance. Since there is a dependence between these two detection stages, there is an optimal trade-off between the probabilities of detection and false alarm per stage so that total probability of detection is maximized.

We compare these three approaches with respect to preamble length in two negative SNR regimes. Figure 5.5 shows that energy detector is a good candidate in SNRs above -10dB and short packet lengths (1000 symbols). Preamble detectors are superior in lower SNRs but their performance strongly depends on the preamble size. As expected, combined detection provides additional performance gain needed for robust detection in wide range of SNRs at the cost of more complex sensing receiver.

5.3.3 Sensitivity to frequency offset

Until now we have considered theoretical performance of a sensor that has a perfect synchronization with a primary user pilot. In all practical receivers with noisy oscillators and circuitry, it is almost impossible to achieve perfect synchronization even in

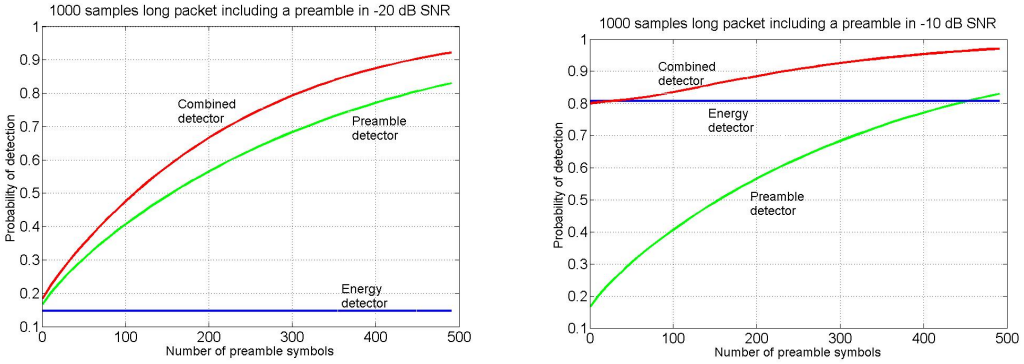


Figure 5.5: Comparison of different approaches for packet detection in negative SNR

positive SNR. Typically, synchronization loops are able to estimate and correct frequency offsets, but their operation is conditioned on positive SNRs where estimation process can be done reliably and fast. However, in negative SNRs these loops are driven by noise and cannot perform robust synchronization. Since sensing receiver does not demodulate the signal, can we expect that synchronization requirements are not very tight?

First lets examine this question theoretically. Consider the sinewave pilot: $x_p[n] = Ae^{j(w_0 n+\varphi)}$ with the nominal carrier frequency w_0. Suppose there is a frequency offset w between the primary transmitter and sensing receiver:

$$\widetilde{x_p}[n] = x_p[n]e^{jwn} \tag{5.14}$$

Then, the decision statistic of an unsynchronized pilot detector becomes:

$$\tilde{T} = \sum_{n=1}^{N} y[n] x_p[n]^* e^{-jwn} \sim A^* e^{-j\varphi} \sum_{n=1}^{N} e^{-jwn} \tag{5.15}$$

Unfortunately, if the sensing interval N becomes comparable or larger than the period of the frequency offset ($2\pi/w$), then coherent processing gain decreases and eventually becomes equal to zero. Therefore, in the presence of frequency offsets the pilot detection has limits on sensing time and detectable signal levels. Now, what are the achievable sensing levels and corresponding frequency offsets?

We answer this question through an experimental study performed on the sinewave pilot at 2.485 GHz. The sinewave signal levels were from -110 dB to -136 dB. This range was chosen based on the requirements from IEEE 802.22 standard for DTV pilot sensing [9]. The frequency offset between the input signal and sensing radio are controlled within sub-Hz accuracy. Figure 5.6 presents experimental results. First, it was confirmed that under perfect synchronization very weak signals can be detected and the scaling of sensing time with respect to SNR is as theoretically predicted $N \sim 1/SNR_p$. However, under very small frequency offsets 10-100 Hz detection becomes impossible for very weak signals. In most practical receivers, tolerances of the frequency oscillators are in the order of 1KHz-50KHz, which would make detection of even stronger pilot signals extremely difficult.

Although pilot sensing performance is severely affected by frequency offset, it does

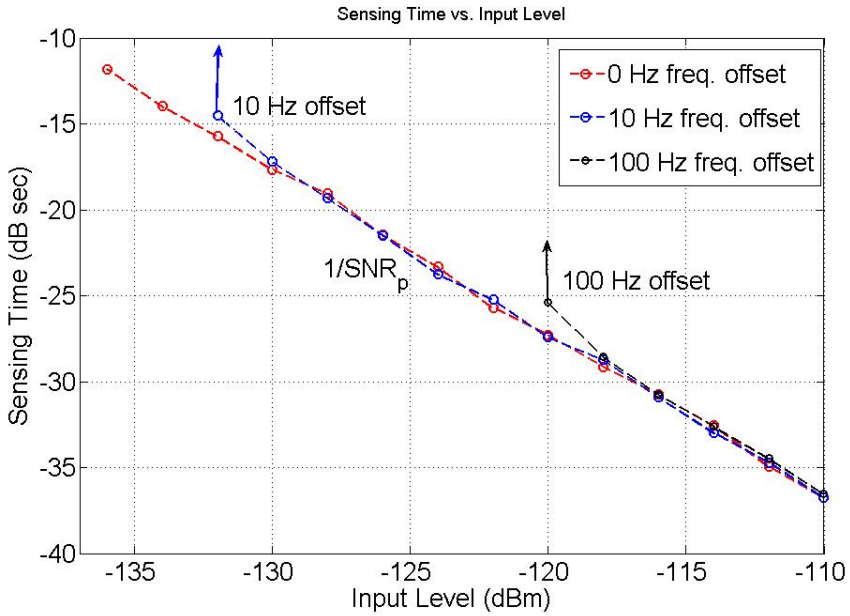


Figure 5.6: Performance of sinewave pilot sensing in negative SNR

not present fundamental limitation in achieving required sensitivity. At the cost of increased receiver complexity, i.e. adding a set of parallel correlators that address every possible frequency offset within desired accuracy, pilot sensing could approach theoretically expected limits. In the wideband implementation where sensing receiver has to sense multiple bands and pilots total complexity makes this approach too expensive in terms of hardware and consequently processing power.

5.3.4 Partial coherent approach for sinewave pilot detection

In this section we discuss approaches for pilot sensing that can benefit from coherent processing without tight synchronization requirements. Also, the emphasis is on simple implementation for wideband sensing.

First, observe that frequency offset cancels coherent processing gain only if the sensing time is in the order of frequency offset time period. If sensing is performed over much shorter time intervals, the coherent processing gain is preserved with the small degradation with respect to the ideal case. By exploiting this property, a suboptimal pilot detection can be implemented with two stage processing. First stage correlates the received signal with a known pilot for a period of time M, and then the second stage non-coherently averages these coherent pieces K times. Theoretically, the new decision statistic for pilot sensing becomes:

$$T = \sum_{k=1}^{K} [\sum_{m=1}^{M} y[m]\widetilde{x_p}[kM+m]^*]^2 \tag{5.16}$$

Note that the total sensing time is $N = KM$. In this mixed (coherent and non-coherent) approach P_d depends more on the choices for M and K than on N. It can be shown that:

$$P_d = Q(\frac{Q^{-1}(P_{fa})}{\sqrt{1+2MSNR_p}} - \sqrt{\frac{K}{2(1+2MSNR_p)}} MSNR_p) \tag{5.17}$$

Under assumption that K is sufficiently large we can identify 3 limiting regimes based on the length of coherent processing M.

Regime 1 if $MSNR_p << 1$ then

$$P_d = Q(Q^{-1}(P_{fa}) - \sqrt{\frac{K}{2}} MSNR_p) \tag{5.18}$$

which scales with K same as in energy detector case. The impact of M is the increase of the sensitivity by $logM$ dB. More importantly, it effectively moves the SNR_{wall} by $logM$ dB.

Regime 2 if $MSNR_p >> 1$ then

$$P_d = Q(\frac{Q^{-1}(P_{fa})}{\sqrt{2MSNR_p}} - \sqrt{\frac{KMSNR_p}{4}}) \tag{5.19}$$

In this regime detector behaves similarly to the optimal coherent detector.

Regime 3 if $MSNRp \approx 1$ the detector is in a transient regime, between coherent and non-coherent processing.

The special case where this approach is particularly convenient is partial coherent sinewave pilot sensing. First note that the optimum matched filter for sinewave detection can be realized using the FFT with the length equal to N. However, in practice it is expensive to implement large size FFTs. Furthermore, all efficient implementations have length that is power of 2, so that arbitrary lengths and sensing times could not be supported. It is desirable to have a fixed and short FFT that acts as a first stage of the partial coherent processor and then use the non-coherent averaging. In this case, the length of the FFT determines M.

In order to verify performance and robustness of the proposed technique, we implement two different partial coherent detectors for sinewave pilot using conventional 256 point and 1024 point FFTs. In the experimental study, we use the input sinewave

signal that is within 10 kHz offset with respect to 2.485 GHz carrier. We also compare it with the coherent approach under frequency offset of 100 Hz. For the considered input signal levels and frequency offset range, chosen FFT sizes correspond to the Regime 1 where performance is improved by $\log M$ dB. Experimental results presented in Figure 5.7 confirm that there is a 6 dB improvement in the SNR_{wall} using $\log 6 = 4$ times longer FFT. Also, there is an addition 6 dB sensitivity gain compared to the matched filter detection under very tight frequency offset (100Hz). Although proposed approach is suboptimal and increases sensing time, the benefits of increased sensitivity and simple but robust implementation for practical frequency offsets makes it a favorable candidate for sensing multiple sinewave pilots in wideband regimes.

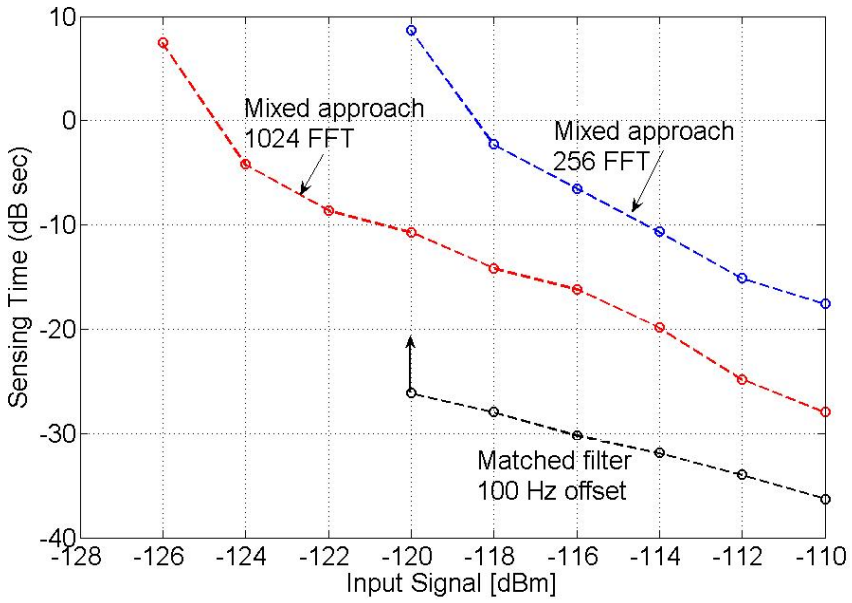


Figure 5.7: Performance of partial coherent detector for weak sinewave sensing

5.4 Feature Detection

Previous discussion shows that coherent processing always provides robust detection with sensing time proportional to the energy in the pilot. However, we have seen that obtaining coherent processing gains comes at the price of synchronization. In addition, some primary user signals do not have easily detectable pilot or the energy in the pilot could be very low. A natural question arises: Is there any other information that cognitive radio sensor can detect in primary signals in order to exploit coherent processing?

Some intuition can be gain by understanding what information is thrown away by a non-coherent energy detector. In essence, the energy detector threats noise and signal in the same way, as wide-sense stationary white processes. However, modulated signals are in general coupled with sine wave carriers, pulse trains, repeating spreading, hoping sequences, or cyclic prefixes which result in built-in periodicity. Even though the data is a wide-sense stationary random process, these modulated signals are characterized as cyclostationary, since their statistics, mean and autocorrelation, exhibit periodicity. This periodicity is introduced intentionally in the signal format so that a receiver can exploit it for: parameter estimation such as carrier phase, pulse timing, or direction of arrival [37], [38]. This information can then be used for detection of a random signal with a particular modulation type in a background of noise and other modulated signals.

5.4.1 Theoretical background

In the previous theoretical analysis we assumed that signals of interest are either known or wide-sense stationary random processes. Therefore, it was sufficient to represent them with time independent first and second order moments, i.e. mean and autocorrelation function. Now, we make a distinction between primary user modulated signals and noise. First, lets revisit a definition of cyclostationary signals, which are characterized by mean and autocorrelation function that are periodic in time. Mathematically:

$$m_X(t) = E[X(t)] = E[X(t + T_0)] \tag{5.20}$$

$$R_X(t+\tau/2, t-\tau/2) = E[X(t+\tau/2)X^*(t-\tau/2)] = E[X(t+T_0+\tau/2)X^*(t+T_0-\tau/2)] \tag{5.21}$$

Period T_0 corresponds to embedded periodicity in the modulated signal, e.g. symbol rate or carrier. From the Fourier series analysis we know that every periodic signal has a Discrete Fourier series representation [44]:

$$R_X(t + \tau/2, t - \tau/2) = \sum_{\alpha} R_X^{\alpha}(\tau) e^{j2\pi\alpha\tau} \tag{5.22}$$

where $\alpha = \frac{k}{T_0}$ for $k = 0, 1, 2, ...$ The Fourier series coefficients $R_X^{\alpha}(\tau)$, which depend on the *lag* parameter τ, are called *cyclic autocorrelation functions*. Each cyclic autocorrelation function has its frequency domain conterpart related by Wienner re-

lationship:

$$S_X^\alpha(f) = \mathcal{F}\{R_X^\alpha(\tau)\} = \int_{-\infty}^{\infty} R_X^\alpha(\tau) e^{-j2\pi f\tau} d\tau \tag{5.23}$$

$S_X^\alpha(f)$ is called *spectral correlation function* (SCF), because it can be thought of as the density of correlation between the two spectral components at frequencies $f + \alpha/2$ and $f - \alpha/2$. This can be observed from the analogy with the definition of conventional autocorrelation since one can express (SCF) as [37]:

$$S_X^\alpha(f) = \lim_{T\to\infty} \lim_{\Delta t\to\infty} \int_{-\Delta t/2}^{\Delta t/2} \frac{1}{T} X_T(t, f + \alpha/2) X_T^*(t, f - \alpha/2) dt \tag{5.24}$$

where finite time Fourier transform is given by:

$$X_T(t, f) = \int_{t-T/2}^{t+t/2} x(u) e^{-j2\pi f u} du \tag{5.25}$$

Spectral correlation function is also termed as *cyclic spectrum*. Unlike power spectrum density, which is real-valued one dimensional transform, the spectral correlation function is two dimensional transform, in general complex-valued and the parameter α is called cycle frequency. Power spectral density is a special case of a spectral correlation function for $\alpha = 0$. The distinctive character of spectral redundancy makes signal selectivity possible. Signal analysis in cyclic spectrum domain preserves phase and frequency information related to timing parameters in modulated signals. As a

result, overlapping features in the power spectrum density are non-overlapping feature in the cyclic spectrum. Different types of modulated signals (such as BPSK, QPSK, SQPSK) that have identical power spectral density functions can have highly distinct spectral correlation functions [45], [46]. More importantly, stationary noise and interference exhibit no spectral correlation. This property has been confirmed by measurement of SCF of 2.4 GHz receiver noise presented in Figure 5.8 which presents the basis for robust feature detection.

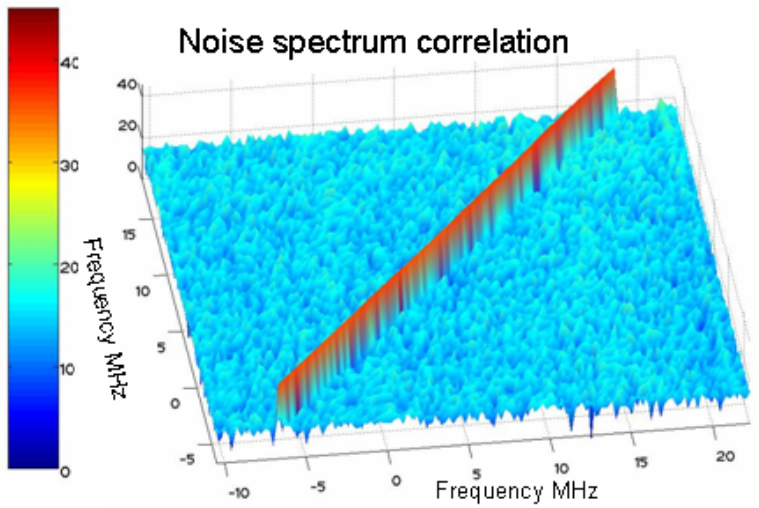


Figure 5.8: Spectral correlation function of the noise and background interference at the 2.4 GHz receiver input

5.4.2 Feature estimation and detection

The properties presented in the previous section describe statistics of features that a specific class of primary user signals have. Note that different primary user signals have different feature, thus these parameters can be used for signal classi-

fication. There is an existing body of literature addressing signal interception in positive SNRs [47]. Also, there are number of methods derived to test the presence of cyclostationarity in the signal [48].

In the spectrum sensing situation, the problem is more difficult as the features need to be estimated and detected with the minimum amount of time in negative SNR. Features need to be estimated in the given limited amount of time. Certainly, the quality of estimation depends on SNR.

First lets discuss the estimation step. Spectral correlation function is simply estimated from the frequency domain processing of the received signal. Here, we outline the commonly used method exploiting the FFT. Given NN_{FFT} samples divided in blocks of N_{FFT} samples, spectral correlation function is estimated as:

$$\widetilde{S_x^\alpha}(f) = \frac{1}{NN_{FFT}} \sum_{n=0}^{N} X_{N_{FFT}}(n, f + \alpha/2) X^*_{N_{FFT}}(n, f - \alpha/2) \tag{5.26}$$

where $X_{N_{FFT}}(n, f)$ is the N_{FFT} point FFT around sample n.

Note that ideal spectral correlation function assumes infinite number of samples (Eq. 5.24). When the observation interval is there is finite, there is a residual estimation error that prevents signal separation from noise. Figure 5.9 illustrates how the feature estimation improves with increased number of spectral correlation averages N.

Before we define a feature detector, we first comment on the required sampling rate. Till now, we have stated that sampling the received signal at Nyquist is sufficient

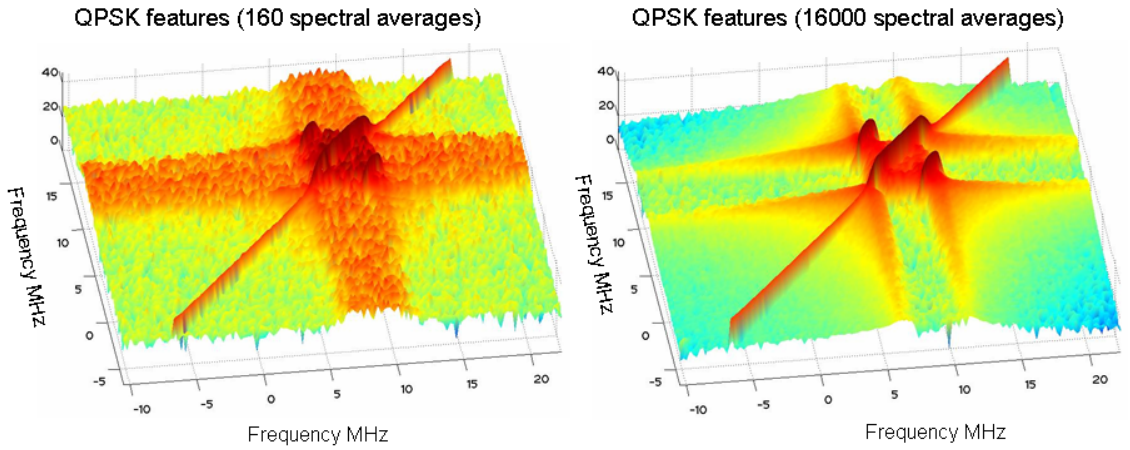


Figure 5.9: Feature estimation accuracy with respect to sensing time

for detection purposes. However, in order to capture the features from the signal samples we need to oversample the received signal. The reason for this requirement is that features are consequence of the spectral redundancy in the signal. One source of redundancy is a pulse shaping filter used to limit the bandwidth of any modulated signals.

To better understand this phenomenon, we will take a common example of a class of amplitude modulated signals and analyze their feature associated with the symbol rate. General representation of a pulse shape amplitude modulated signal is:

$$x(t) = \sum_{n=-\infty}^{\infty} a_n q(t - nT_s) \tag{5.27}$$

where a_n is data sequence, $q(t)$ is a pulse shaping filter and T_s is a symbol rate. The cyclic spectrum of $x(t)$ that has a maximum feature energy is given by:

$$S_X^{\alpha}(f) = \frac{1}{2T_s} Q(f + \alpha/2) Q^*(f - \alpha/2) \quad \text{for} \quad \alpha = 1/T_s \tag{5.28}$$

where $Q(f)$ is a pulse shaping filter in Fourier domain. In case of most commonly used square root raised cosine filter, the feature spans over:

$$\frac{1-\beta}{2T_s} < f < \frac{1+\beta}{2T_s} \tag{5.29}$$

where β is a roll-off factor of the square root raised cosine filter. Figure 5.10 illustrates how different βs influence the energy of the feature corresponding to the symbol rate.

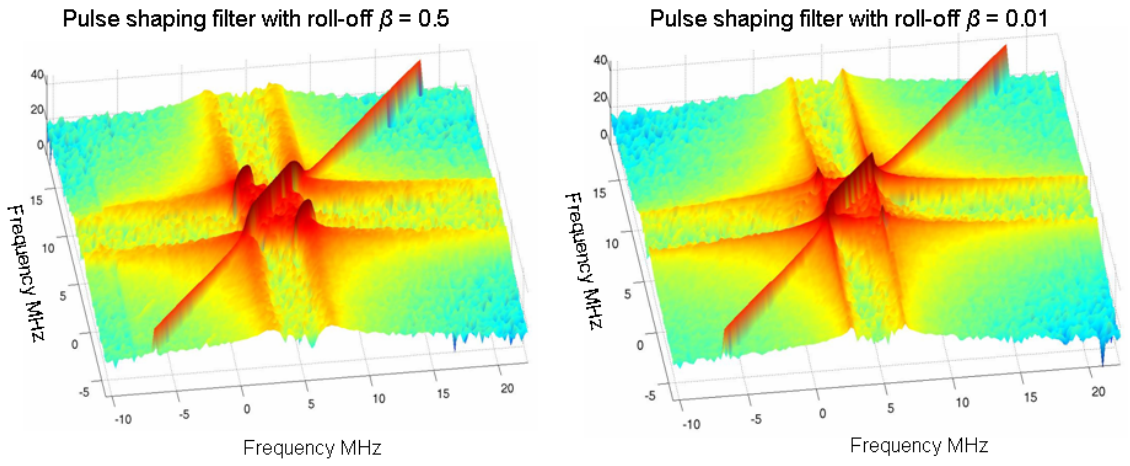


Figure 5.10: Comparison of SCFs for signals with different pulse shaping filters

Since filters with $\beta = 0$ are not realizable, every practical system will have some energy in the feature. Larger β contributes to a larger spectrum redundancy and would provide larger energy in the feature. From the sampling frequency point of

view, larger β requires larger oversampling ratio.

Now, assuming that the signal is sampled at a high enough rate, we can convert the detection problem into a discrete time binary hypothesis testing problem with the threshold test, similar to the energy and pilot detection case. Note that we assume that primary user signal format and its ideal SCF $S_S^\alpha(f)$ are perfectly known at the sensing receiver. The optimal test statistic is the projection of the estimated SCF of the received signal $y[n]$ onto the ideal primary user SCF at the particular feature α chosen to have maximum energy.

$$T_\alpha(N) = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} S_S^\alpha(f)\widetilde{S_Y^\alpha}(f)df \tag{5.30}$$

where f_s is the sampling frequency and N is the number of collected N_{FFT} block samples. Note that since the signal and noise are not correlated, resulting SCF is equal to $S_Y^\alpha(f) = S_X^\alpha(f) + S_W^\alpha(f)$. Since for $\alpha \neq 0$ and infinite estimation interval $S_W^\alpha(f) = 0$, signal detection is robust to noise regardless of the SNR. This theoretical result holds promise for improved sensing sensitivity using feature detectors. However, this optimal detection method cannot be implemented in practice due to additional uncertainty embedded in the received signal caused by the wireless channel. Wireless channel changes frequency domain response of the received signal $Y(f) = H(f)S(f) + W(f)$. The problem is that $H(f)$ is a complex function which also alters spectral correlation function of y.

$$S_Y^{\alpha}(f) = H(f + \frac{\alpha}{2})H(f - \frac{\alpha}{2})^* S_S^{\alpha}(f) + S_W^{\alpha}(f) \tag{5.31}$$

As a result of fading, optimal detection from Eq. 5.30 can result in cancellation or attenuation of decision statistic. Therefore, we must apply suboptimal detection approach that is robust to complex channel effects. A suboptimal detector is often called *cyclic spectrum analyzer* and its implementation is simply integration of the SCF in frequency domain for a specific α:

$$T_{\alpha}(N) = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mid \widetilde{S_Y^{\alpha}}(f) \mid^2 df \tag{5.32}$$

Note that this suboptimal sufficient statistics is obtained through a non-linear squaring operation. Therefore, it is expected that this detection approach will fall into the category of non-coherent detectors in terms of sensing time requirements, similar to energy detector performance. However, due to operation in the spectral correlation domain space ($\alpha \neq 0$), we rely on uncorrelated properties of the noise to improve robustness to noise variations and overcome SNR_{wall}. We will further evaluate this hypothesis through implementation and experimental study.

5.4.3 Sensitivity to clock offset

Given that feature detectors exploit additional coherency in the primary user signal that is not present in noise, the natural question arises if there are any robustness

issues involved in the detector performance. Remember that in case of pilot detection coherent processing was highly susceptible to carrier frequency offsets and resulted in limits on sensing sensitivity. In case of cyclostationary processing of the features related to the symbol period, it is evident that sampling of the incoming signal becomes critical. Intuitively, a random sampling of a periodic signal would not resemble its periodicity. In the previous analysis, we inherently assumed that the sampling clock is an integer multiple of symbol rate. Now, we examine what happens if there is a sampling clock offset. As a result of sampling clock offset there is a drift in sampling point instances within a symbol time. First, note that a constant time offset introduces a phase offset in both frequency and cyclic spectrum domain.

$$\overline{x}(t) = x(t - t_0) \longleftrightarrow \overline{X}(f) = X(f)e^{-j2\pi f t_0} \tag{5.33}$$

$$S_{\overline{X}}^{\alpha} = S_X^{\alpha} e^{-j2\pi\alpha t_0} \tag{5.34}$$

Since cyclic spectrum is a complex function, this time offset has a different effect on it than on a power spectrum density being a real function. Now, lets assume that there is a residual sampling clock offset Δ from the ideal sampling clock T_0, i.e. $T - T_0 = \Delta$, where T is the sampling clock interval. Then, in the process of estimation of spectral correlation function every estimate obtained by correlation of FFT bins will have a phase offset. Lets assume that N FFTs are used to estimate $\widetilde{S}_{\overline{X}}^{\alpha}(f)$.

$$\widetilde{S}_X^{\alpha}(f) = \frac{1}{N}\sum_{i=1}^{N}\overline{S_X^{\alpha}}(i,f) \tag{5.35}$$

where each FFT serves to estimate instant spectrum correlation:

$$\overline{S_X^{\alpha}}(f) = \frac{1}{N_{FFT}}\overline{X}_{N_{FFT}}(n, f+\alpha/2)\overline{X}^{*}_{N_{FFT}}(n, f-\alpha/2) \tag{5.36}$$

In the presence of a clock offset, each FFT block will have a phase offset with respect to previous due to shift in the sampling point by t_i:

$$\overline{S_X^{\alpha}}(f) = S_X^{\alpha}(i,f)e^{-j2\pi\alpha t_i} \tag{5.37}$$

where $t_i = iN_{FFT}\delta$, and $S_X^{\alpha}(i,f)$ is the cyclic spectrum under perfect sampling.

In the process of averaging, these estimates with different phase offsets are non-coherently added, and resulting cyclic spectrum has attenuated features:

$$\widetilde{S_X^{\alpha}}(f) = S_X^{\alpha}(f)\sum_{i=1}^{N}e^{-j2\pi\alpha t_i} = S_X^{\alpha}(f)\frac{\sin(2\pi\alpha\Delta N_{FFT}N/2)}{\sin(2\pi\alpha\Delta N_{FFT}/2)}e^{-j2\pi\alpha\Delta N_{FFT}(N+1)/2} \tag{5.38}$$

The feature at α is smeared if $\Delta \approx 1/\alpha N N_{FFT}$. Figure 5.11 illustrates SCF with perfect sampling with stong features corresponding to symbol rates and SCF with sampling offset where these dominant features are smeared as predicted by Eq. 5.38.

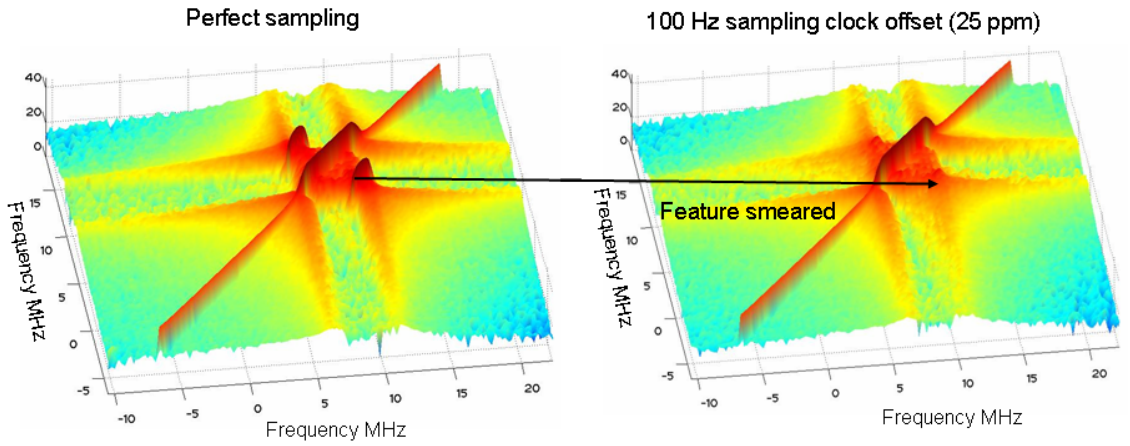


Figure 5.11: Spectral correlation function of 4MHz QPSK signal with perfect sampling (left) and with 100Hz sampling offset (right)

5.4.4 Robust approach for feature detection

Because of this sensitivity to sampling clock offset, we need to modify our original method for spectrum correlation function estimation (Eq 5.26). First note that, similar to the coherent processing in the presence of frequency offset, the number of samples that can be used for sensing is limited by the following ratio:

$$NN_{FFT} < \frac{\alpha}{\delta} \tag{5.39}$$

where sampling clock frequency offset δ and sampling time offset Δ are related by:

$$\Delta = \frac{\delta}{\alpha(\alpha - \delta)} \tag{5.40}$$

Taking this into account, we propose partially coherent feature processing that prevents cancellation of the features by performing two stage processing. First stage

performs coherent feature estimation based on Eq. 5.26 over a sufficiently short time interval, and then the second stage non-coherently averages these partial estimates. Based on the maximum expected sampling offset δ_{max}, number of coherent averages M_1 is chosen to be smaller than $\alpha/\delta_{max}N_{FFT}$. The SCF estimate then becomes:

$$S_x^{\alpha}(f)' = \frac{1}{T_{FFT}} \sum_{m=1}^{M_2} | \sum_{k=1}^{M_1} X(k + mM_2, f + \alpha/2)X^*(k + mM_2, f - \alpha/2)| \qquad (5.41)$$

where the total number of averages is $N = M_1M_2$.

5.4.5 Implementation and experimental results

Due to highly non-linear processing and issues with modeling of noise in cyclostationary domain, we analyze the performance of ideal and partial coherent feature processor through implementation and experiments.

Feature detectors are always implemented in digital domain [49]. Direct algorithms first compute the spectral components of the data through FFT, and then perform the spectral correlation directly on the spectral components. The computational complexity of a spectral correlation function estimator is easily estimated. For a stream of N_{FFT} samples it requires a computation of N_{FFT} point FFT, which requires $N_{FFT} * logN_{FFT}$ multiplications, and N_{FFT}^2 multiplication for cross multiplications. Note that this algorithm is extremely parallel so that the computation of the spectral correlation function can be organized across frequency or across cycle plane

independently.

Figure 5.12 presents the cyclostationary feature detector implementation that is robust to sampling offsets based on 5.41. The first stage averages SCF in the complex domain so that features are coherently added and the noise is cancelled. The second stage averages magnitude of the output of the first stage, and therefore it changes the processing from coherent to non-coherent. The output of the detector is obtained through integration of the energy in the SCF that lies in (α, f) space where signal of interest has theoretically predicted features. The implementation has a flexibility of dynamically changing the region of interest in the two-dimensional SCF domain. Therefore, it is suitable for use in wideband scenarios for parallel search of multiple features.

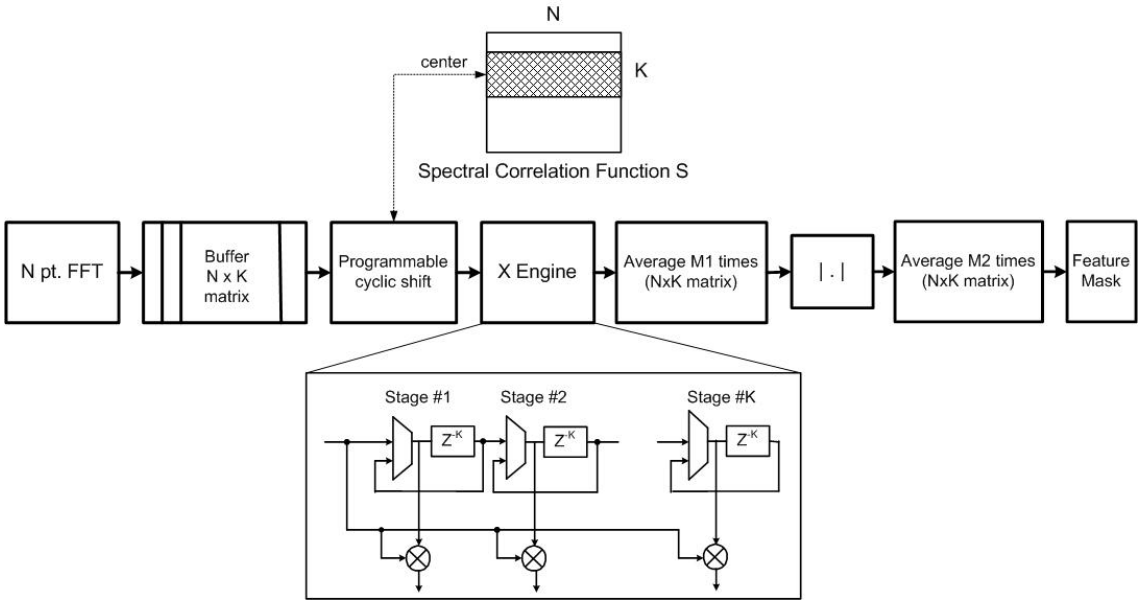


Figure 5.12: Implementation of a cyclostationary feature detector for robustness under sampling clock offsets

Figure 5.13 presents the experimental results that compare feature detectors with energy detector in the presence of sampling clock offsets. Under stationary white noise, feature detectors (even with perfect sampling) have a performance loss with respect to energy detector. This is due to the fact that energy contained in the feature is related to the pulse shaping filter roll-off. In case of $\beta = 0.5$ the loss is approximately 6 dB. On the other hand, a sampling clock offset of 100 Hz at 64 MHz makes the detection of signals below -15 dB SNR impossible. However, once the proposed 2 stage averaging is deployed the detector achieves the desired probability of detection and false alarm at the penalty of increased detection time. The number of averages in the first stage is chosen from the Eq.5.41. For SNRs below -15 dB, the number of averages in the second stage has to be increased with respect to perfect sampling feature detection. The proposed scheme performs comparable even for sampling offsets of 1 KHz, with approximately 0.5 dB loss with respect to 100 Hz.

5.4.6 Comparison with energy detection

Results in Figure 5.13 show that even under perfect sampling feature detector is inferior to energy detector. This result is expected as energy of the feature is always less or equal than energy of the signal. On the other hand, we have already seen that energy detector is not robust to noise variance uncertainty caused by temperature variations and out-of-band interference. What is the performance of feature detector in this case?

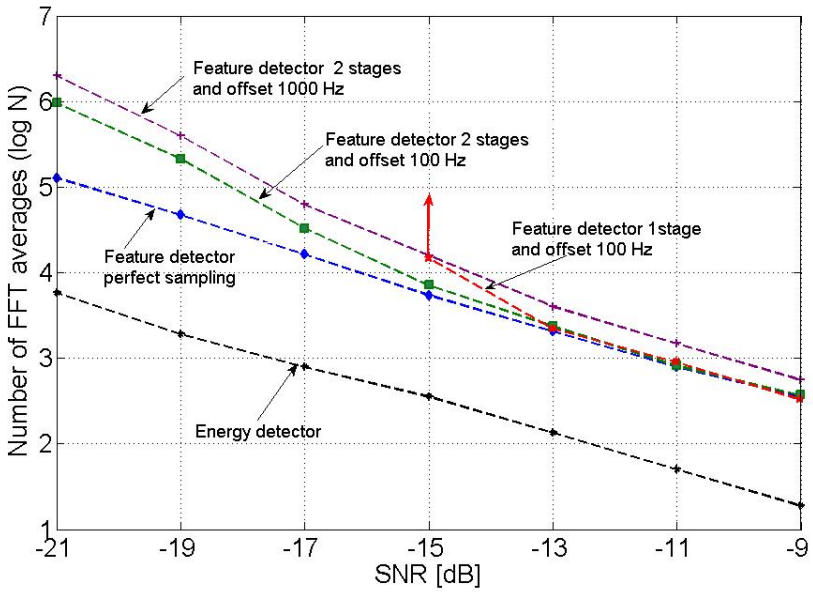


Figure 5.13: Performance of cyclostationary feature detectors in negative SNR

This question was in part answered theoretically by [50], [51]. Here, we review their result. A performance of the detector is measured by a *deflection coefficient* d defined by:

$$d = \frac{|E[T_\alpha(N)|H_1] - E[T_\alpha(N)|H_0]}{\sqrt{var[T_\alpha(N)|H_0]}} \tag{5.42}$$

Intuitively, deflection coefficient corresponds to SNR at the output of the detector. Under noise variance uncertainty x and input SNR, deflection coefficients for energy and feature detectors scale with the number of samples as:

$$d_{energy} \sim \frac{\sqrt{N}SNR}{\sqrt{1 + x(1+N)}} \tag{5.43}$$

$$d_{feature} \sim \frac{K_\alpha \sqrt{N} SNR}{\sqrt{1+x}} \tag{5.44}$$

where K_α represents ratio between signal energy and feature energy. Note that under $x = 0$, both coefficients scale as $\sqrt{N}$ which confirms the non-coherent behavior demonstrated in experimental results. However, for any $x > 0$ deflection coefficient for feature detector saturates for large N, while for feature detector it monotonically improves. The interpretation of saturation for energy detector case is another view of SNR_{wall} defined earlier. Eq. 5.44 theoretically predicts that feature detectors do not have SNR_{wall} due to noise uncertainty. We verify this result experimentally and measure performance in terms sensing time (measured by number of FFT averages) vs. SNR.

To test the robustness of feature detectors under varying noise and interference, we experimented with the adjacent channel interference coming from the commercial 802.11g WLAN with a continuous traffic generated by video camera data transfer between two laptops. Figure 5.14 shows the performance of both energy and feature detectors in the presence of this strong adjacent band interference.

Due to spectral leakage of the FFT, energy detector suffers from the large variation in the noise-plus-interference level. This variation progressively degrades the energy detector performance and at $SNR = -18$ dB detection becomes impossible. On the other hand, feature detector robustly detects the weak signals and outperforms the energy detector. Note that there is a slight degradation in performance of feature

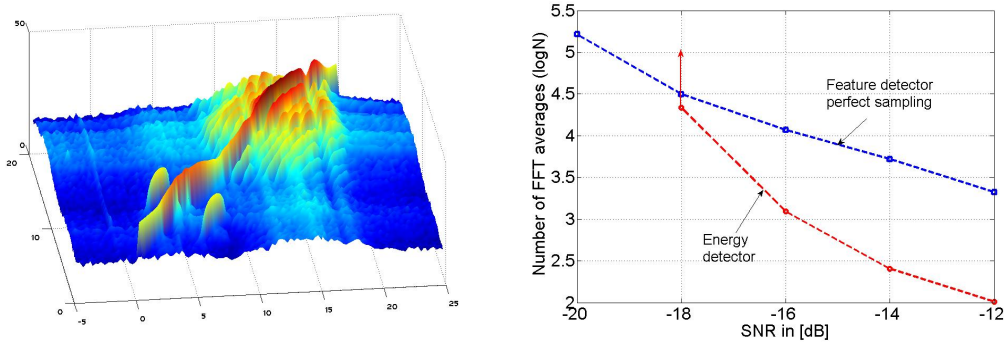


Figure 5.14: Features of QPSK signal and the adjacent 802.11g signal (left) and detector performance in non-stationary noise due to adjacent band interference (right)

detector due to leakage of the interference signal in SCF domain.

5.5 Summary

In this chapter we systematically explored the design space of signal processing algorithms and fundamental trade-offs in their performance with respect to sensing time and sensing sensitivity. We found that the primary user signal characteristics like active signal time, percentage of deterministic signal energy, and redundancy of modulation present generalized parameters for optimal selection of spectrum sensing approach for the required SNR sensitivity. The presented framework provided a classification of sensing schemes using coherent, non-coherent or mixed signal processing and their scaling of sensing time with respect to SNR. Furthermore, we showed that there are physical limits on achievable SNR and sensing time in the presence practical receiver design constraints and realistic noise and interference sources. We also provided realistic models of receivers and their parameters that need to be incorporated

Parameter	Energy	Pilot	Feature
Sensing time	$1/SNR^2$	$1/SNR_p$	K_α/SNR^2
Limitation	SNR_{wall}	Freq. offset (0.01 ppm)	Clk. offset (10 ppm)
Remedy	N/A	2 stage processing	2 stage processing
Complexity	18 mult, 320Kb	40 mult, 64Kb	82 mult, 3200Kb
Proposed usage	Band selection	Robust sensing if pilots are available	Robust sensing and signal classification

Table 5.1: Summary of physical layer spectrum sensing methods

in further theoretical studies of sensing in negative SNRs.

Table 5.1 presents the summary of our study together with general recommendation for use of certain detector types.

Energy detector can quickly and reliably identify presence of strong primary user signals in fairly high SNR regimes, therefore it is suitable for selection of candidate bands where more elaborate processing should be done. Unfortunately, energy detector can not be used for reliable sensing in highly negative SNR regimes due to SNR_{wall} caused by estimation errors or variations in noise and interference. In addition, when the time available for sensing is limited, energy detector sensitivity level could be insufficient for the primary user requirement.

For faster detection or higher sensitivity, energy detection should be combined with more sensitive coherent or mixed approaches. Exploiting coherent processing gain for deterministic pilots or preambles is beneficial if they contain the significant percentage of signal energy. However, achieving high sensitivities is often not possible due to sensitivity to very small frequency offsets. Mixed approaches provide a robust but yet simple way to exploit partial coherent processing gain and overcome the

synchronization requirements. As a side effect, mixed approaches result in an increase of the required sensing time.

In the absence of deterministic signals, sensing receiver can exploit modulation inefficiency and detect signal features. Although in principle non-coherent methods, feature detectors allow signal recognition and provide robustness to non-stationary noise and interference channel conditions. The sensing times of feature detectors are comparable to energy detectors sensing time within a constant factor. Their main advantage is that there is no limit in the achievable sensitivity, i.e. no SNR_{wall}. Their weakness is the sensitivity to sampling clock offsets, but it could be overcome by simple mixed processing approach.

Chapter 6

Network Layer Spectrum Sensing

Up to this point we have considered spectrum sensing performed by a single radio in noise dominated channels. From the system design perspective, multiple cognitive radios need to simultaneously perform spectrum sensing and jointly decide on available spectrum opportunities. Therefore, there is inherent requirements for the node cooperation and exchange of sensing information. Through cooperation cognitive radio network can also exploit inherent diversity in individual measurements and improve detection by relying on users with favorable channel conditions. In this chapter we explore the strategies for network cooperation, achievable gains in actual wireless environments and implementation constraints.

6.1 Motivation and Theoretical Background

In Chapter 5 we have seen that meeting primary user requirements for high sensitivity and short sensing time with a single radio physical layer sensing is not always possible. Remember that the primary user protection requirements are set by the worst case channel condition resulting in highly negative SNR regimes. In wireless network scenarios, these worst case channel conditions are determined by three dominant effects: multipath fading, shadowing, and local interference [28]. Although they are considered as negative effects, on their positive side is an inherent variability at various location and at different times. As a result, it is reasonable to expect that not all cognitive radios in the network will simultaneously experience the worst channel conditions. Intuitively, this channel diversity increases the probability that there are radios in the network that have good channel conditions thus can provide reliable sensing for the whole network.

Recently, use for cooperation in wireless have been studied extensively, especially in respect of achieving diversity gains and lower outage probabilities. However, the gains obtained through cooperation in spectrum sensing need to be quantified through improvement of overall probability of the detection or through a decrease in required sensing time. Next, we present theoretical framework to address these gains. First, we will look into improvement of sensing reliability.

Lets assume that there are M radios in one cognitive radio network, and each radio has a physical layer spectrum sensor. For simplicity, we consider only energy

based sensing. Recall from Chapter 5 that each sensor performs hypothesis test with a particular threshold. We now modify our model to account for fading, shadowing and local interference, and take the assumption that each radio experiences independent fading h^i, $i = 1, ..., M$.

$$H_{0,i} : y_i[n] = w_i[n] \quad n = 1, ..., N \tag{6.1}$$

$$H_{1,i} : y_i[n] = h_i s[n] + w_i[n] \quad n = 1, ..., N \tag{6.2}$$

where h_i is a wireless channel (path loss, shadowing and multipath) from a primary user transmitter to a cognitive radio sensor i. Probability distribution function of h_i's depends on the wireless environment and placement of radios. Effects of local interference are merged with noise, thus each radio has noise with variance σ_i^2. Note that we model interference as a Gaussian process independent of the receiver noise.

If all radios perform sensing over the same period time (synchronized sensing) and use the same threshold γ for detection, then their $P_{fa,i}$ and $P_{d,i}$ are:

$$P_{fa,i} = Q(\frac{\frac{\gamma}{\sigma_{w,i}^2} - N}{\sqrt{2N}}) \tag{6.3}$$

$$P_{d,i} = Q(\frac{\frac{\gamma}{\sigma_{w,i}^2 + |h_i|^2 \sigma_s^2} - N}{\sqrt{2N}}) \tag{6.4}$$

In the simplest case, we assume that all radios experience the same noise and local interference $\sigma_{w,i}^2 = \sigma_w^2$ for all i. In addition, channel coefficients h_is are independent

and identically distributed as $f_{h_i}(x)$. For a given pdf of h_i, we can compute the average probability of detection for each radio as:

$$\overline{P_d} = \int_x P_{d,i}|_{h_i=x} f_{h_i}(x) dx \tag{6.5}$$

Since both P_{fa} and $\overline{P_d}$ now depend only on a detection threshold, we can characterize the sensor performance with a pair of $P_{fa}, \overline{P_d}$ for a given threshold γ. Sweeping through all possible thresholds, a receiver operating curve (ROC) can be obtained. ROCs are commonly used to measure the detector performance.

Now, we are ready for analyze cooperation gains. Assume that M radios combine their measurements by "OR" operation, i.e. if any radio detects the primary user signal the whole network will take his decision. If all measurements are independent, then average probability of detection for the network Q_D monotonically increases with number of radios M:

$$Q_D = 1 - (1 - \overline{P_d})^M \tag{6.6}$$

Unfortunately, the probability of false alarm for the network Q_F also monotonically increases with M as:

$$Q_F = 1 - (1 - P_{fa})^M \tag{6.7}$$

Earlier, we said that cooperation gain can be measured by effective decrease in

the required sensing time. First note that as the number of radios increases, the probability that every radio experiences a deep fade decreases. For the "OR" combining rule, the user with the best SNR will detect the signal first. All radios that operate in SNRs bellow the best user's SNR will not be able to detect a primary user signal. For a given joint channel distribution of M radios, we can measure the probability of exceeding certain $SNR(M)$. The probability of detection for the whole network will be governed by this $SNR(M)$. Note that $SNR(M)$ is monotonically increasing with M. From the signal processing analysis of energy detector, we know that a higher SNR corresponds to a shorter sensing time for a fixed P_d. Given that $SNR(M)$ monotonically increases with M, then sensing time of the best user as well as overall network monotonically decreases with M. In case of energy detector, this reduction in the sensing time can be significant, as for a 10 dB improvement in SNR, nominal sensing time can be decreased 100 times.

6.2 Limitations on Cooperative Gains

In the previous theoretical analysis of cooperation gains in terms of Q_D and Q_F, we made two strong assumptions on channel and noise statistics in order to simplify the derivation and find upper bounds on cooperation gains. Now, we will revise our model and discuss the implications on practically achievable cooperation gains.

In modeling of statistics of channel coefficients h_i's we assumed that they are both independent and identically distributed. As a result, the cooperation gains in Q_D and

Q_F were maximized. However, channel coefficients are a result of superposition of three components (path loss, shadowing, and multipath) that do not necessarily need to be independent for all radios. While path loss for small to medium networks can be assumed equal for all radios, the other two effects could have quite different characteristics. For example, shadowing can exhibits high correlation if two radios are blocked by the same obstacle. There are a number of measurements reported in the literature [52] showing that different wireless environments and carrier frequencies have largely different shadowing statistics. A measure of shadowing correlation is described by the coefficient ρ and modeled as an exponential function of distance: $\rho = e^{-ad}$, where parameter a depends on the properties of shadowing obstacles in the environment.

Measurements of shadowing in indoor environments show that the correlation coefficient is independent of wavelength over a frequency octave, but it is dependent on the topography. It was estimated that 90% correlation distance is typically 1m, 50% is around 2m, and slowly decays to 30% over 8m [52]. Therefore, if the cognitive radio network is deployed in a limited area, like typical indoor environment, increasing the number of radios introduces the correlation which in effect limits the cooperation gain. Simulation results in [53], [54] have shown such decrease in cooperation gain based on theoretical models of correlation. In the next section, we will present the experimental results that show effects of spatial distribution of radios in the network on cooperation gains in indoor environments.

Our second modeling assumption was made for equal distribution of noise and local interference across all radios. As a result, every sensor could apply the same detection threshold and achieve equal P_{fa}. However, in all practical situations these two assumptions do not hold. First, due to circuits variability, temperature difference and aging, each radio has different aggregate local noise, typically characterized as a noise figure. Even after manufacturing, the tolerances on receiver noise figure are in the order of 1 dB. Furthermore, we have seen in Chapter 5 that strong primary signals leakage from adjacent bands causes additional variability in energy levels and decreases robustness of energy based sensors. While this variability in noise and interference sources exists in practically all receivers and does not cause significant performance loss in positive SNR, the question is if it affects the network sensing performance in negative SNR. The theoretical treatment of this problem in positive SNR with perfect knowledge of noise and interference distributions can be found in [55]. The optimal solution involves high computational complexity needed to solve a set of coupled nonlinear equations that scales exponentially with number of users. We investigate this question in our experimental study by testing two practical types of threshold rules in the local decision process: 1) a predetermined (fixed) threshold set by the centralized processor or 2) an adaptive threshold independently estimated at each radio based on real-time measurements of noise and ambient interference in the band of interest.

6.3 Experimental Results

In this section we present the experimental study of cooperative gains achievable in an energy detection based spectrum sensing network. The performance is measured in terms of ROC (Q_F vs. Q_D).

6.3.1 Experimental setup

The experiments were conducted inside the Berkeley Wireless Research Center. The floor plan of the center is shown in Figure 6.1. The figure also shows 54 locations on a 2m by 2m grid, that covers a cubicle area, library and conference room, where all wireless measurements were taken. In all experiments, the transmitter was located inside the lab. Therefore, the signal path between the transmitter and all receiver positions included propagation through either concrete or wooden walls, supporting beams, medium and large size metal cabinets, and general office furniture. The area covers a balanced variation of obstacles which are typical for indoor non-line-of-sight environments.

Due to operation in the unlicensed ISM band, interference sources had to be considered. All 802.11 b/g, Bluetooth,and ZigBee equipment was shut down during the experimentation, in order to minimize potential interference. Also, experiments were performed overnight to eliminate influence of people in the building.

For each sensing location, data was collected for three different transmitter configurations: idle spectrum i.e. no signal, sinewave signal and QPSK signal. The idle

Figure 6.1: BWRC floor plan with transmitter and receiver locations

spectrum was sensed in order to to compare two different threshold rules described in the previous sections. For each location and data type, spectrum was sensed 200 consecutive times using 3200 averages (51.2 ms) in energy detector.

For the sinewave, the signal generator transmitted a -40 dBm signal at 2.485GHz. For the 4 MHz wide QPSK signal, the signal generator transmitted a -30 dBm signal in 2.483GHz - 2.487GHz band, centered at 2.485GHz. The average SNR measured across 54 location was -10 dB, with standard deviation of 3 dB. The bandwidth ratio of QPSK signal to sinewave is approximately 10*1og10(4 MHz/62.5kHz)=18 dB, thus a 10 dB difference in transmit power favors the sinewave case in terms of the receiver SNR. It was expected that sinewave performance would be more affected by multipath, thus 8 dB power gain was added.

6.3.2 Impact of distance

First, we explore the possibility to use multiple antennas at sensing receiver as a proxy for cooperative communication. Although multiple antennas add complexity and cost per radio, they would effectively minimize the network overhead in implementing cooperative sensing. In conventional wireless systems, increasing the number of antennas in one radio can significantly improve received SNR via beamforming. In general, beamforming algorithms exploit wireless channel spatial diversity [56], due to multipath scattering, to increase the antenna gain in a desired signal direction.

Maximization of a signal processing gain in multiple antenna systems requires that antennas are separated by distances larger than $\lambda/2$, where λ is a wavelength of a carrier frequency. This is a commonly used rule that is based on the correlation coefficient measurements and modeling in multipath dominated environments. In addition, an optimal combining requires knowledge of multipath gains and phases for coherent processing. This knowledge is commonly acquired through channel estimation based on pilots. In principle, reliable estimation of a channel phase requires positive SNR and takes time, which makes it infeasible for spectrum sensing applications. In our system implementation we used the sub-optimal approach where received signals from all antennas were non-coherently added through energy combining. This technique is equivalent to an equal gain combining in multiple antenna receive diversity applications [28].

Next, we measure how this cooperative gain scales with number of antennas in

terms of Q_D and Q_F. Results in Figure 6.2 show that 2 antennas improve Q_D by 10% and 4 antennas maximize it by 25%. However, the gain is also saturated with more than 4 antennas, i.e. there is no benefit in adding more. This saturation is explained by limited number of degrees of freedom, i.e. spatial clusters, in indoor wireless channels. Multiple antenna channel measurements [57] show that in most indoor environments a received signal is distributed in 1 to 3 spatial clusters. Note that number of resolvable clusters correspond to number of antennas. In addition to gain saturation, we can also observe that the gains are less than theoretically predicted by Eq. 6.6. Although antennas were separated by $\lambda/2 = 12cm$, this result suggests that there is significant correlation between antennas. Intuitively, this can be explained by considering a shadowing component in wireless channel, which cannot be combated by multiple antennas.

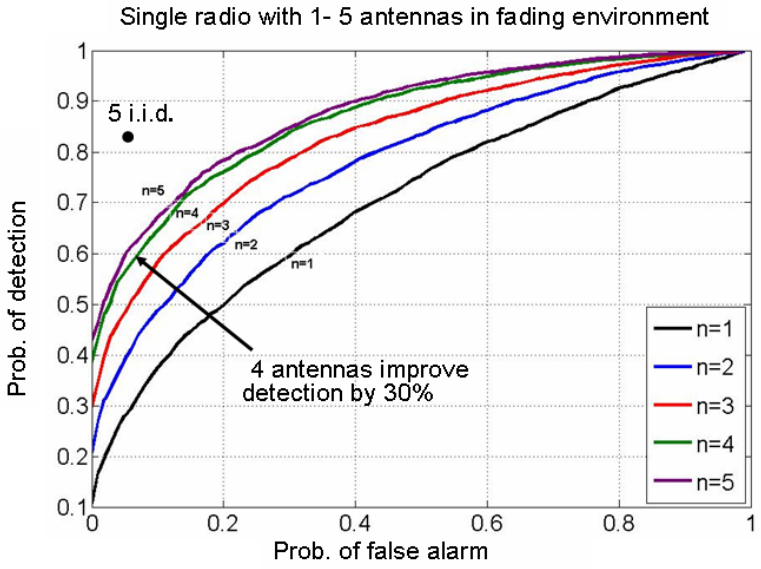


Figure 6.2: Small scale cooperation gain and benefits of multiple antennas

Shadowing is a large scale phenomenon that depends on objects in the wireless environment of interest. The key characteristics of an object are its size and material. In order for signal to be shadowed by an object, its size has to have a dimension that is much larger than a signal wavelength λ. Typical object sizes in our environment are from 2m (cubicle dividers) to 8m (metal cabinets). Therefore, we placed radios on 2m by 2m grid and measured the cooperation gain as a function of distance between two cooperative radios. Figure 6.3 presents the measurement results. First note that resulting Q_D for the fixed Q_F monotonically increases with radio separation, which suggests that correlation between two radios decreases with distance monotonically. This confirms earlier theoretical model of exponential decay of correlation with distance. The most significant gains were obtained for separation of 2m (approximate size of a cubicle divider), which turns out to be the dominant shadowing source in our environment.

Now, we are ready to answer our original question about relationship between cooperation gain and number of radios. By setting the separation to be at minimum 2m, we measure the gain obtained with 2 to 5 cooperative radios. Figure 6.4 presents the measurements results. Results show that in typical indoor environments, a few cooperative radios can achieve very reliable sensing. With 5 radios, $Q_D = 95\%$ was achieved by using cooperation among radios with poor detection $P_d = 60\%$. This result shows great potential for use of cooperation in spectrum sensing, especially in scenarios like indoor WLAN networks where the typical number of cooperative radios

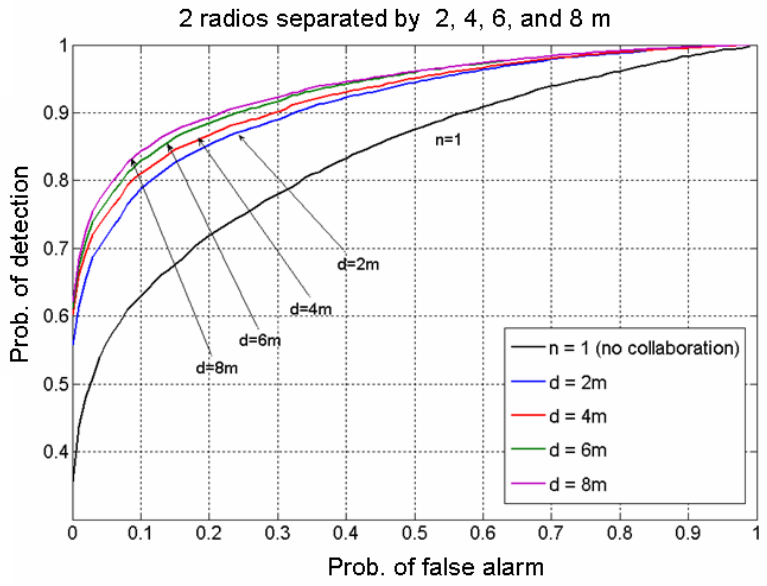


Figure 6.3: Large scale cooperation gain

and their spatial distribution closely correspond to our measurement setup.

6.3.3 Impact of signal bandwidth

In addition to spatial diversity, wireless channels are typically characterized by frequency diversity. Frequency diversity means that different radio frequencies sufficiently wide apart will experience independent fading characteristics. As a result, signals with wider bandwidths are not affected by the worst case channel conditions caused by multipath. The separation between uncorrelated frequencies is called coherence bandwidth and it is inversely proportional to a delay spread of a wireless channel [28]. In effect, frequency diversity, similar to spatial diversity, also depends on multipath profile in a given wireless environment. In typical indoor environments

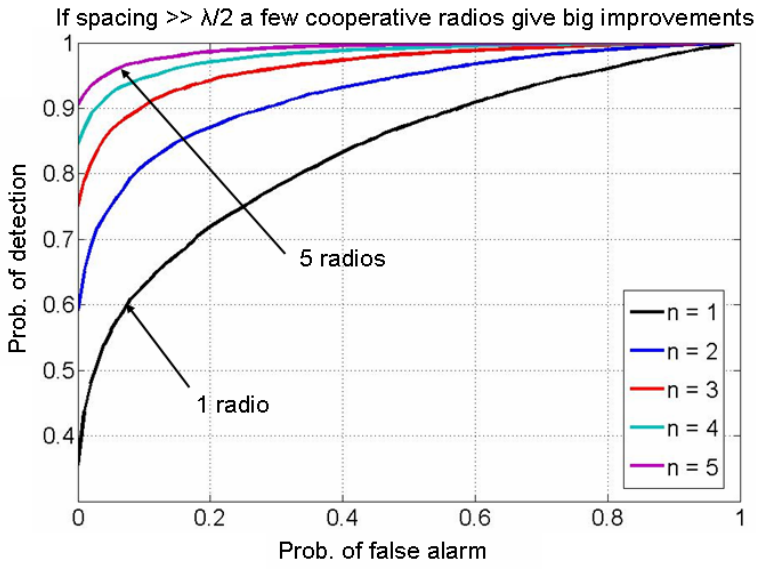


Figure 6.4: Cooperation gain vs. number of radios

coherence bandwidths are in the order of 1MHz.

In order to estimate combined gains from cooperation and frequency diversity, we used two distinct types of signals in our experimental study: narrowband sinewave pilot and wideband 4 MHz QPSK signal. Cooperation gains in terms of Q_D for the fixed $Q_F = 10\%$ are compared in Figure 6.5. Since 4 MHz bandwidth occupies approximately 4 coherence bandwidths, the frequency diversity gains can be exploited in case of wideband QPSK signal. On the other hand, sinewave signals cannot exploit any frequency diversity and suffer from the worst case channels caused by deep multipath fades. Note that SNR for sinevawe signal was 8 dB higher than for QPSK. Even with this advantage, cooperation gains for sinewave sensing were inferior to wideband signals, where with 5 cooperative radios Q_D reaches 99%. As a result, network cooperation exploited together with frequency diversity presents a robust spectrum

sensing technique simply implemented with an equal gain energy combining.

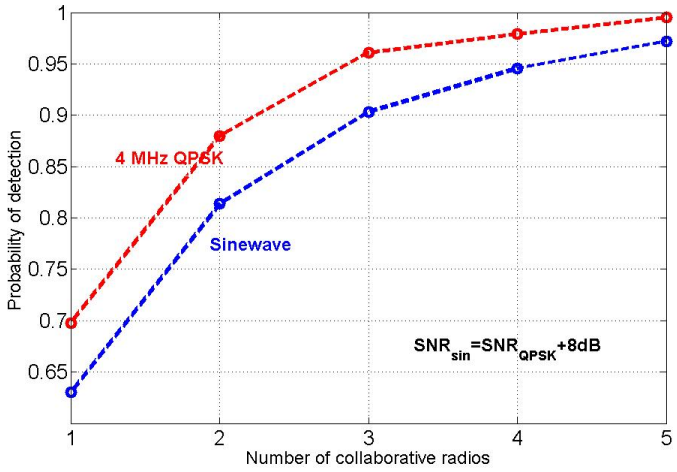


Figure 6.5: Comparison of cooperation gains between narrowband (sinewave) and wideband (4 MHz QPSK) signals

6.3.4 Impact of detection threshold

In our theoretical analysis, we pointed that a modeling assumption about noise and interference sources across different radios being i.i.d does not hold in general. As a result, the threshold applied in energy detection across different radios need to be adaptively estimated rather than kept fixed. Using real radio implementation and actual wireless channel environment, we were able to test the effects of different threshold rules.

Figure 6.6 compares experimentally obtained ROCs for two different threshold rules applied to the same set of measurement data. Note that due to time varying characteristics of both noise and interference sources, it was essential that two algo-

rithms are tested simultaneously. For the fixed threshold rule, the same threshold was applied for all possible combinations of cooperating radios. For the estimated threshold, each radio measures its aggregate noise and interference and computes the threshold to meet P_{fa}. While thresholds across radios are different, each one has the same P_{fa}.

Results shows that the performance of the fixed threshold rule is significantly suboptimal with respect to that of an estimated threshold rule. Not surprisingly, even in the case of single radio energy detection under varying noise and interference, it is essential to apply the location and time relevant threshold obtained via estimation. If the network cooperation is implemented using same threshold for all radios, gains are still present, but are significantly reduced with respect to optimal adaptive threshold rule. The gap between the two rules varies form 15% to 25%. As a result, even a moderate $Q_D = 90\%$ and $Q_F = 10\%$ can never be met using the fixed threshold rule. While adaptive threshold rule guarantees more reliable sensing, it requires frequent receiver noise calibration and accurate interference estimation. This requirement introduces additional circuit complexity and sensing time.

6.4 Implementation Issues in Cooperation

In the previous two sections we analyzed theoretical and experimental performance of spectrum sensing using network cooperation. Both discussions presented performance under the assumption that information from all radios is gathered at

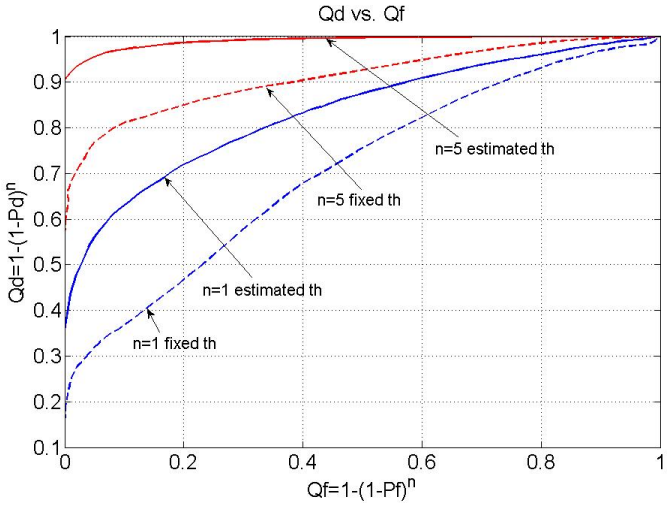


Figure 6.6: Impact of threshold rule on cooperation gains

one location (centralized processor), where different sensing strategies can be implemented. However, in practical network scenarios cooperation requires means of sharing sensing information among cognitive radios. In addition to sharing, timing coordination is necessary in order to efficiently utilize sensed spectrum opportunities. In Chapter 2 we proposed common control channels to implement network cooperation and satisfy these two requirements. With the better understanding of algorithms and their performance in a real system we now discuss implementation issues in the system design of network cooperation.

A common control channel essentially means that a cognitive radio network has a designated sensing protocol in parallel with a regular medium access protocol. While most of MAC protocols have embedded control information for network management, their requirements are not as stringent and critical as for the spectrum sensing functionality. Recall that primary user protection requirements are based on sensing time

and sensing reliability. While our experimental study showed that cooperation does improve reliability with decreased sensing time per radio. However, we did not incorporate the additional latency due to information exchange among radios. This latency depends on the load of the control channel. In the most general case, the load scales linearly with the number of radios and number of sensed frequency bands. Note that control channel requires a physical communication channel which occupies spectrum resources, thus its load should be minimized.

In large networks, scalability of the control channel protocol becomes an issue. The underlying signal processing on the physical layer determines what kind of information radios can share through cooperation. In the simplest case, when all radios use energy detection, reported information could be either hard decision or soft decision. In the case of hard decision reporting, the overall traffic is minimized as radios send only one bit per sensed frequency band. Sharing soft decision enables optimum detection, and allows identification of radios with poor channel and noise conditions. Furthermore, soft decisions allow estimation of the correlation between radios. Given that radios operate in a limited area, there is a high likelihood of many correlated measurements. A solution to the scalability problem lies into algorithms that identify uncorrelated radios and minimize number of participants.

In this study we explored cooperation strategies based on energy detection. While the cooperation gains are achievable even with this simple approach, there are some disadvantages that complicate protocol implementation. The main drawback of en-

ergy detector is that it cannot distinguish primary user transmission from other cognitive radios. Therefore, when energy sensing takes place all cognitive radios in the area should resume transmission and perform sensing within the same sensing interval. This imposes additional silent time, loss in effective system throughput, and coordination with medium access protocol. The other two techniques we studied in Chapter 5, pilot and feature detection, can distinguish transmission from different sources in a robust and reliable way. Furthermore, these techniques do not require quite periods. However, optimal combining strategies in terms of soft and hard combining need to be investigated. Also, further research is needed for flexible and scalable cooperation protocols that can optimally exploit versatile physical layer spectrum sensing capabilities.

6.5 Summary

This chapter explored the improvements in spectrum sensing performance achievable through network cooperation. Consistent with theoretical probabilistic analysis, the experimental study showed that sensing reliability improves monotonically with number of cooperative radios. Due to this gain, sensing time of individual radios can be reduced. However, there are a number of critical parameters that influence these improvements. Monotonic increase in Q_D is achieved only if radios provide independent sensing measurements. Therefore, network topology determines the required number of users and their optimal spatial distribution. Multiple antennas also provide

sensitivity gain but the gain is limited by spatial diversity in a wireless channel.

While cooperation is beneficial for any type of primary signal sensing, higher gains are achievable for wideband signals, since frequency diversity can be exploited. In terms of implementation, the important design parameter is the detection threshold used by each cooperating radio. Due to variability of noise and interference sources, thresholds need to be adapted per each radio independently rather than kept fixed for all radios in the network. From the system design perspective, network cooperation requires dedicated protocol to share and coordinate individual sensing measurements. Physical implementation would require a common control channel with minimized load, and low latency that is scalable with number of cooperative users. Further research is needed for common control channel cooperation protocols that can support different physical layer spectrum sensing techniques like pilot or feature detection.

Chapter 7

Sensing Radio

After evaluating approaches for spectrum sensing at the physical and network layers, we return to the design of a sensing radio with better understanding of its design specifications. In Chapter 5 we have seen that physical layer spectrum sensing was highly affected by the time varying interference sources from adjacent bands. This implies that the sensing radio signal reception and conditioning should minimize the effect of strong primary signals leakage caused by saturation, clipping or non-linearities. From the experimental study of network layer sensing in Chapter 6, we learnt that wideband sensing and multiple antennas provide additional degrees of freedom that can be exploited by network cooperation. In this chapter, we introduce new radio architectures for spectrum sensing that address requirements driven by higher layers, and also result in a feasible implementation.

7.1 Wideband Sensing Radio Requirements

In order to enable maximum spectrum utilization, future cognitive radios must be able to sense and communicate over wide bandwidths. From the implementation point of view, all front-end circuitry should be wideband which introduces many challenges in RF radio design. In typical cognitive radios sensing scenario, the RF signal presented at the antenna includes signals from close and widely separated transmitters, and from transmitters operating at widely different power levels and channel bandwidths. As a result, the large dynamic range becomes the main challenge as it sets the stringent requirements on circuits linearity and resolution of A/D converters.

The scaling of CMOS technology improves the transistor operating frequency, which could be used for increased bandwidth and sampling rate, but adversely affects the design of precision analog circuitry. With the reduction in supply voltage which scales with the transistor features, the inherent transistor gain decreases, the voltage headroom decreases making it difficult to accommodate large dynamic range signals. Therefore, reducing the strong in-band primary user signals is necessary to receive and process weak signals for spectrum sensing.

Lets review which circuit blocks from Figure 7.1 present major limitations for spectrum sensing. First, cognitive radio sensing antenna should provide wideband reception with minimal signal loss. Recently, ultra-wideband antennas in 0-1 GHz and 3-10 GHz bands have been developed, and can be used for cognitive radios too. In contrast with low noise amplifiers (LNA) for UWB systems, where noise figure

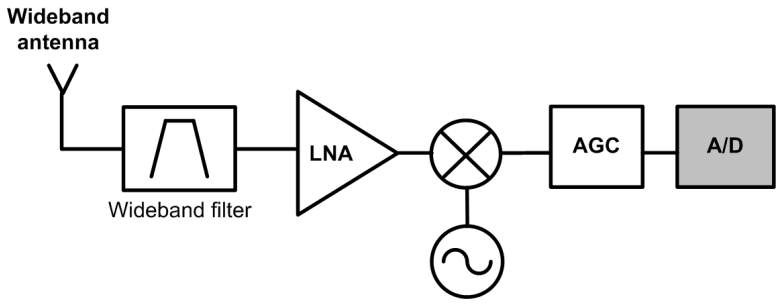


Figure 7.1: Sensing radio architecture for regimes with high spectral utilization

was not an issue and was traded for power dissipation, sensing radios require minimal noise figure (e.g 2-3 dB) for improved sensitivity with low power. In addition, the critical design specification for amplifiers and mixers is to maintain the linearity across entire dynamic range and bandwidth. However, these are also conflicting requirement with respect to power consumption. Large dynamic range and sampling of wideband signals further require high precision and high speed A/D converters. Unfortunately, the design of high speed A/D converters has fundamental limits in terms of achievable resolution. For example, designing a 1 GHz A/D converter with 12 bit resolution is infeasible due to comparator ambiguity and clock jitter impairments [58]. These constraints in the circuits implementation requires us to rethink the sensing radio architecture. In the next section, we discuss new design approaches for dynamic range reduction in wideband radios and compare several architectures that attempt to overcome this limitation.

7.2 Addressing the Dynamic Range Problem

Measurements of spectral utilization from 0 to 6 GHz ref show that there are typically a few strong narrowband primary signals that cause large dynamic range problem. These strong signals are of no interest to detect, since spectrum sensing focuses on processing of weak signals. In order to solve the dynamic range problem, we recognize that strong signals can be resolved and removed in frequency, time, or space domain. Next, we outline three different approaches for reducing the dynamic range by "filtering" strong blocking signals in these three domains.

7.2.1 Frequency domain filtering

In conventional radios, fixed filters are used to provide frequency discrimination, pre-condition input signal, and relax the requirements on the radio circuit components. However, cognitive radios encounter scenarios where, at every location and time, different strong primaries fall in-band. Therefore, cognitive radios require bandpass and bandstop filters with challenging specifications: high center frequency, narrowband with large out-of-band rejection, and tuning ability. Tunable bandpass filters are used for channel selection and reduction of out-of-band interference, while a tunable bandstop filters are needed for isolation of transmit and receive paths, and in-band blocker rejection.

Filters are external components and not favorable for integrated single chip radios. In general, CMOS implementation and integration lead low cost and power, so the

number of external filters should be minimized. Furthermore, sharp roll-off RF filters need high Q, which leads to high power consumption and large circuit area to accommodate passive elements, inductors, and capacitors. Non-ideal filters cause leakage of the strong signal across bands and can degrade weak signal sensing performance.

Tunable RF filters require exploration of non-conventional filter architectures. Novel techniques for filtering like RF-MEMs show some promise. However, MEMS filters suffer from insertion loss, and are also harder to design for high frequencies. In addition, they require a long time to tune to the desired band. The majority of lumped-element MEMS-tuned filter designs use fixed inductors and achieves tuning via adjustable MEMS capacitors, so their frequency is in general limited to 3 GHz [59]. Tuning bandwidth and center frequency is possible via switching distributed resonators, but again resolution is limited by the placement density and electrical sizes of resonators, since the design is discrete in nature. In order to improve resolution, number of resonator elements must increase. It is required that on-going research efforts in RF-MEMs filters address these requirements and provide better filtering flexibility.

7.2.2 Time domain cancellation

Filtering in time domain essentially means selective cancellation or subtraction of large signals. Cancellation techniques are motivated by the result from a multiuser detection theory [60] developed for an interference channel. This result establishes

that it is possible to decode a strong interfering signal first, reconstruct it and subtract from the incoming signal so that weak signals can be decoded too. It is important to note that it is sufficient to attenuate signal, not perfectly cancel, thus estimation error can be tolerated. Now, lets see how this result applies to a wideband receive signal that includes a few strong primary narrowband signals.

The architecture that utilizes time domain cancellation [61] is presented in Figure 7.2. The main idea for this approach is to use two low-resolution A/D converters with N and M bits in order to achieve a high resolution A/D converter of N+M bits. For example, under strong interference conditions with up to 50 dB dynamic range increase with respect to desired weak signals, it is possible to realize an equivalent 12 bit A/D converter using two 6 bits converters. This architecture effectively behaves as a two stage pipelined A/D converters, where the first stage deploys adaptive signal processing to provide reduced dynamic range to the second stage. The incoming signal goes through two paths, one branch deals with interference estimation and reconstruction, while the other is a programmable analog delay line used to align the signals in two branches so that proper cancellation timing is achieved. In general, the trade-off between the number of bits in two stages of A/D conversion depends on the interference strength, i.e. strong interference situations require more bits in the first stage A/D converter. The key challenge in this approach is matching of the latency through the two paths using an analog delay line.

The main advantage of digital signal processing is a capability to adaptively pre-

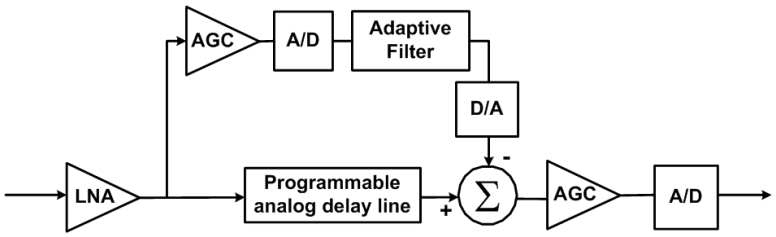


Figure 7.2: Feedforward architecture for time domain digitally assisted analog interference cancellation system

dict strong blocking signals at different frequencies and bandwidths. Note that this adds flexibility with no additional hardware cost as opposed to frequency filtering approach. Similar idea, also called active cancellation, is developed for collocated transmitter and receiver interference cancellation present in multiple standard radio systems. The feasibility of active cancellation and the first circuit demonstration is reported in [62]. While this approach could relax the requirements on A/D conversion, it does not address the problems of linearity in the LNA and mixing stages. Essentially, the linearity problem has to be solved very early in the reception process, ideally at the antenna. This brings us to the third method for signal filtering using spatial dimension.

7.2.3 Spatial filtering using multiple antennas

Lastly, we explore an alternative approach for dynamic range reduction that filters the received signals in the spatial domain by using multiple antennas. This idea is inspired by recent theoretical work on multiple antenna channels identifying that spatially received signals occupy a limited number of directions or spatial clusters

[56]. Measurements across ultra-wide frequency band show that in typical indoor environments there are 2 to 4 spatial clusters [57]. In addition, the angular spread of these clusters is in the order of 10 to 30 degrees. Figure 7.3 illustrates the power distribution across frequency and space at the input of sensing receiver. The narrowband character of strong blocking signals and their confined distribution across spatial dimension offer new possibilities for design of radio architectures with analog signal processing. Next, we propose an architecture and implementation of sensing radio with multiple antennas for dynamic range reduction using spatial filtering.

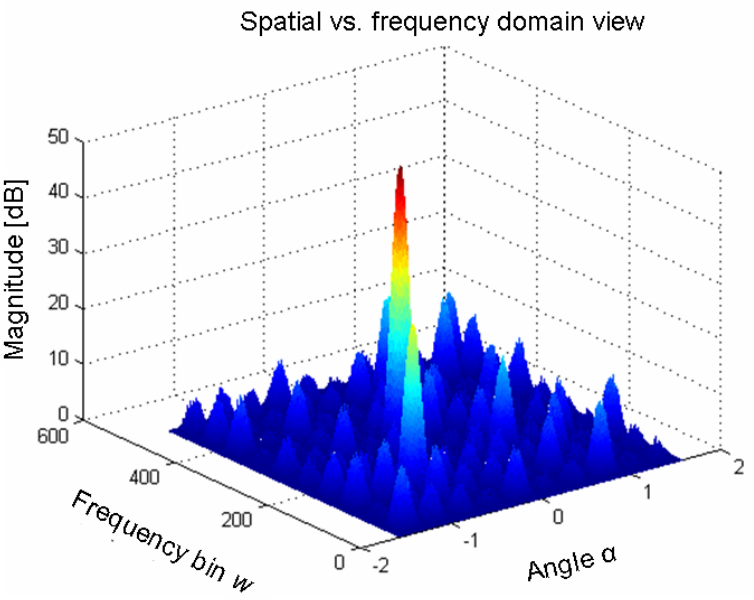


Figure 7.3: Received energy distribution across frequency and spatial dimensions

7.3 Multiple Antenna Sensing Radio Architecture

Enhancing radio receiver with multiple antennas allows improvement of received signal strength, throughput, or link robustness. For example, in Chapter 6 we have seen that multiple antennas can exploit channel diversity and improve reliability of spectrum sensing. Here, we are after a different problem. Our goal is to actively suppress blocking signals and selectively receive frequency bands of interest where spectrum sensing will be preformed by a physical layer sensing. Commonly, selective and adaptive signal reception is achieved through beamforming techniques. There is a large existing body of literature addressing digital beamforming algorithms [63]. Here, in contrast to digital beamforming, multiple antenna processing must be done in the analog domain, before the LNAs, mixers and automatic gain control circuits, so that wideband signal can be properly amplified with minimum distortion and for the best utilization of the number of bits in the A/D converter.

The proposed architecture of a wideband RF front-end which is enhanced with an antenna array for spatial filtering is shown in Figure 7.4. Note that digital beamforming techniques can be realized using multiple antenna receiver with parallel receiver chains and A/D converters, while here after the combining stage there is only one receiver branch and one A/D converter. There is a similarity between this approach and electronically steerable passive array radiator (ESPAR), where also only single output port is observable [64]. However, ESPARs are typically designed for a narrow frequency band and would not be suitable for wideband sensing.

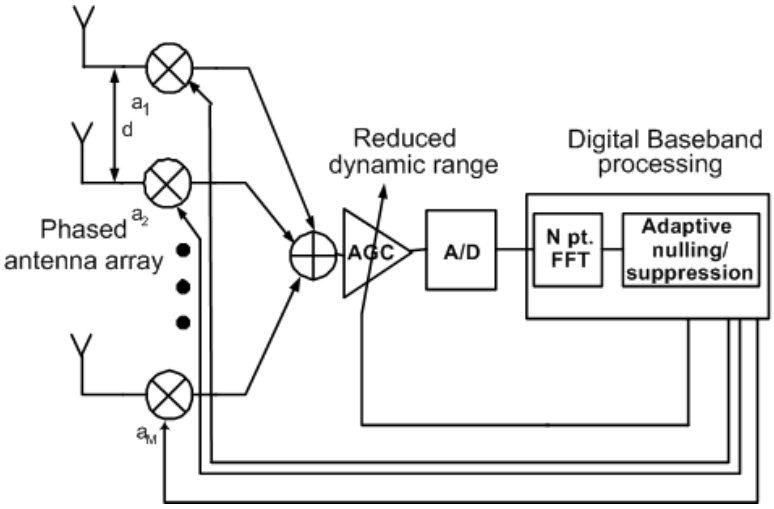


Figure 7.4: Wideband RF front-end with phased antenna array for spatial filtering

In terms of implementation, the proposed architecture could be designed as a phased antenna array where the antenna array coefficients are computed in the digital back-end and fed back to analog phase shifters which then adjust the gains and phases of the antenna elements. The use of simple phase shifters is particularly attractive due to their very low latency needed for fast convergence of the desired array response. Figure 7.5 presents two possible approaches for an implementation of complex RF phase shifters.

Clearly, there is an additional hardware complexity and cost involved in the implementation of the proposed architecture. They are both increasing with number of antennas and precision of array coefficients. Therefore, in practice, both number of antennas and precision of array coefficients need to be minimized. Furthermore, the realization of phase shifters with high precision is not feasible. The precision limitation is more stringent on the phase than on the gain [65]. Next, we analyze impact

of number of antennas and coefficient quantization on the sensing radio design and set the requirements for a given performance goal.

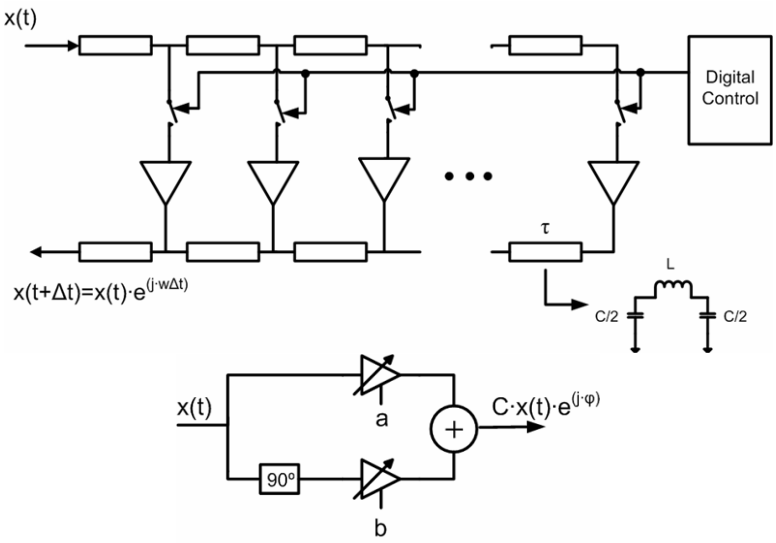


Figure 7.5: Implementation of RF phase shifters: switched delay lines (top) and vector modulators (bottom)

7.3.1 Impact of number of antennas

Phased antenna arrays are typically implemented as an uniform linearly distributed array of antennas, where all antenna elements are identical and equally spaced by distance d. Each antenna element is weighted by an array coefficient $a(n)$, where $n = 0, ..., N-1$. Coefficients are typically complex numbers.

Array response in the particular angular direction ϕ is described by Nth order polynomial:

$$A(e^{j\theta}) = \sum_{n=0}^{N-1} a(n) e^{-jn\theta} \tag{7.1}$$

where $\theta = 2\pi k \cos(\phi)$ and $k = d/\lambda$.

Essentially, a signal arriving from the direction ϕ will be weighted by a complex number $A(e^{j\theta)}$. Thus, based on the properties of an array response polynomial different directions will be weighted differently.

Since we are interested in suppressing directions where blocking signals occur, next we study how many blockers can be suppressed and how to synthesize array that satisfy desired constraints. To answer this questions, we represent an array response as a polynomial in a complex z-plane:

$$A(z) = a(N-1)z^{N-1} + a(n-2)z^{n-2} + ... + a(1)z + a(0) \tag{7.2}$$

The array response can be factored in terms of $N-1$ nulls $\{z_1, .., , z_{N-1}\}$:

$$A(z) = K(z - z_1)(z - z_2) \cdots (z - z_{N-1}) \tag{7.3}$$

Positions of nulls and their correspondence to angular directions are illustrated in Figure 7.6 using a complex plane representation. Note that if the null lies on the unit circle than the corresponding direction can be perfectly nulled. If the null lies outside the unit circle than this direction can only be attenuated. The amount of attenuation depends also on the position of other nulls.

Based on this analysis, we can establish that with N element array a maximum of $N-1$ interferers can be suppressed. This fact implies that in typical scenarios with

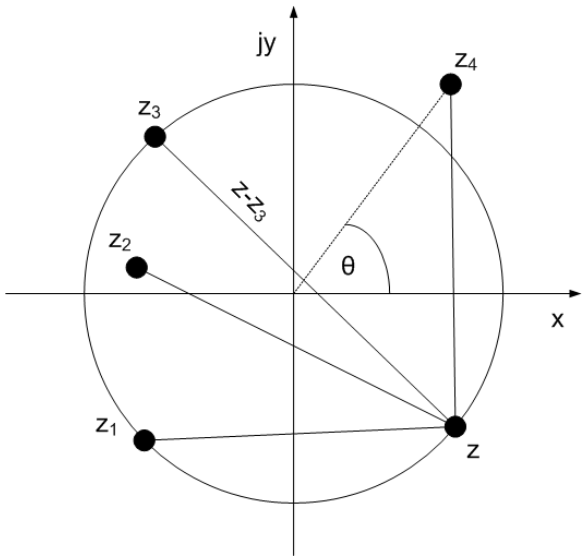


Figure 7.6: Determining an array response and nulls using complex plane

2 to 4 spatial clusters, the required number of antennas for suppression is from 3 to 5. Note that positioning nulls perfectly on a unit circle, i.e. complete nulling of a particular interferer, requires infinite precision of array coefficients. As we pointed out earlier, in practice coefficients are quantized and the achievable resolution is limited to 5 to 6 bit precision. Next, we evaluate how much suppression can be achieved under this coefficient precision constraint.

7.3.2 Impact of array coefficients precision

Because of inherently hard to predict nature of quantization errors, a statistical analysis of the effect of coefficient quantization on array response is appropriate. The aim of such analysis is to predict how much accuracy for the coefficients is required without knowing the values of the array coefficients. The statistical model used in

the analysis assumes that the errors due to quantization of different coefficients are statistically independent, and that each error is uniformly distributed between $-Q/2$ and $Q/2$, it has zero mean and variance $Q^2/12$, where Q is the quantization step size [66], [67].

Lets denote $e(n) = a(n) - a^*(n)$ to be the error in the coefficient $a(n)$ due to quantization. Lets define the error function by:

$$\epsilon(e^{j\theta}) = A^*(e^{j\theta}) - A(e^{j\theta}) \tag{7.4}$$

where $A(e^{j\theta}) = \sum_{n=0}^{N-1} a(n)e^{-jn\theta}$ is the ideal array response and $A^*(e^{j\theta}) = \sum_{n=0}^{N-1} a^*(n)e^{-jn\theta}$ is the array response with quantized coefficients, $\theta = 2\pi k\cos(\phi)$ and $k = d/\lambda$. Now,

$$|\epsilon(e^{j\theta})|^2 = (\sum_{n=0}^{N-1} e(n)e^{-jn\theta})(\sum_{n=0}^{N-1} e(n)^* e^{jn\theta}) \tag{7.5}$$

$$= \sum_{n=0}^{N-1}\sum_{k=0}^{N-1} e(n)e^*(k)e^{-j(n-k)\theta} \tag{7.6}$$

Therefore average error can be computed as,

$$\overline{|\epsilon(e^{j\theta})|^2} = \sum_{n=0}^{N-1} \overline{e(n)^2} = N\frac{Q^2}{12} \tag{7.7}$$

using the independence assumption. Thus a measure of the deviation of $|\epsilon(e^{j\theta})|$ from zero can be defined as

$$\sigma_\epsilon = \frac{Q}{2}\sqrt{\frac{N}{3}} \tag{7.8}$$

With high probability $|\epsilon(e^{j\theta})|$ ought to be bounded by two to three times its standard deviation:

$$|\epsilon(e^{j\theta})| \widetilde{\leq} 2\sigma_\epsilon \tag{7.9}$$

Now, suppose that we were interested in designing an array with the ideal response $L(\theta)$. If we use N element antenna array then we can meet the specification within a margin δ, i.e.

$$|A(e^{j\theta}) - L(\theta)| \leq \delta \tag{7.10}$$

Then, if we quantize the array coefficients with b number of bits, the the resulting array response $A^*(e^{j\theta})$ will deviate from the desired response by:

$$|A^*(e^{j\theta}) - L(\theta)| \leq 2^{-b}\sqrt{\frac{N}{3}} + \delta \tag{7.11}$$

Now lets define the inband rejection, equal to effective dynamic range reduction, of the unquantized and quantized array response:

$$DR = -20\log_{10}(max|A(e^{j\theta}) - L(\theta)|) \tag{7.12}$$

and

$$DR^* = -20\log_{10}(max|A^*(e^{j\theta}) - L(\theta)|) \tag{7.13}$$

Note that

$$|A^*(e^{j\theta}) - L(\theta)| \leq |A^*(e^{j\theta}) - A(e^{j\theta})| + |A(e^{j\theta}) - L(\theta)| \tag{7.14}$$

Thus,

$$DR^* \widetilde{\geq} -20\log_{10}(10^{-\frac{DR}{10}} + 2^{-b}\sqrt{\frac{N}{3}}) \tag{7.15}$$

It can be seen that for each number of bits b there is a maximum value of inband rejection. For example, if the desired dynamic range reduction is 25 dB, then the array coefficients must be quantized with minimum 5 bits. The loss in the dynamic range reduction is a function of desired rejection and number of bits.

Note that that sensitivity to coefficient quantization becomes larger if we are trying to achieve high inband rejection. However, if we target inband rejection below 25 dB than the loss is tolerable.

Now, lets see what happens if we increase the number of antennas. Then, to achieve the desired inband rejection DR we need approximately:

$$b \approx \frac{DL^*}{6} + \frac{5}{3}\log_{10}\frac{N}{3} \tag{7.16}$$

Figure 7.7 shows the dependence.

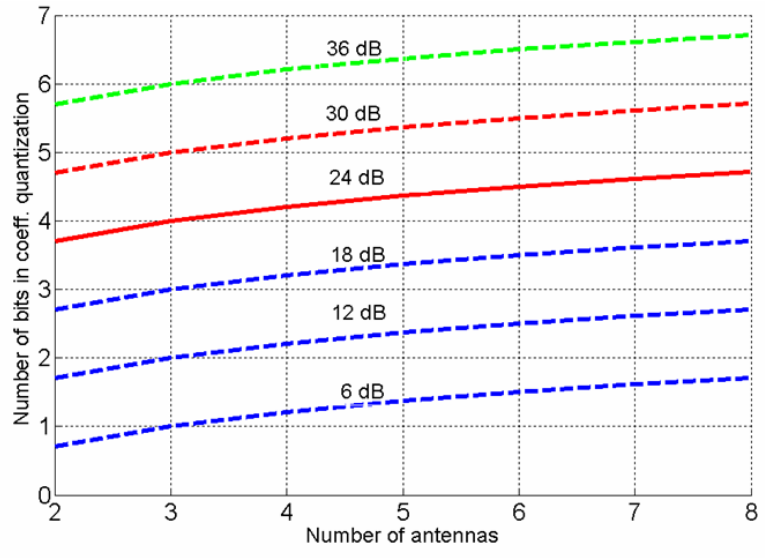


Figure 7.7: Relationship between number of antennas and coefficient quantization for a given inband rejection

Based on the presented analysis, for typical scenarios with 4 strong blockers and required dynamic range reduction of 25 to 30 dB, it is feasible to design a 4 element phased antenna array with 5 bit precision of array coefficients. Now, lets consider the signal processing algorithms for computation of coefficients that achieve dynamic range reduction.

7.4 Spatial Filtering Approach

There is a number of signal processing algorithms developed for derivation of the optimum set of antenna element weights. These algorithms fall into two major categories: i) Reference based algorithms and ii) Direction-of-arrival (DOA) based algorithms.

The reference based algorithms use knowledge of a desired reference signal together with measurement of the received signal to adaptively change the antenna weights according to a predefined criterion or constraint. Some examples of the design criteria are [68]: maximizing the output power in the look direction (conventional beamforming), minimizing the noise and interference come from all directions other than desired signal (Capon beamforming), or maintaining unit gain in the look direction while zero gain in the interfering directions (zeros-forcing beamformer).

In contrast, the DOA based algorithms use the mathematical method based on eigenvalue and eigenvector analysis [69] to extract the DOA of all signals in the same frequency band including the DOAs of desired user as well as co-channel interferers and the DOAs of the multipath reflections of these original sources of signals. Once the nature of signal sources is determined, the antenna coefficients can be set form an appropriate beam pattern to boost the desired signal and suppress the interferers.

The problem of a dynamic range reduction considered here is different from both categories. First, our design criteria is to suppress a number of strong signals by a certain margin, rather than maximize, minimize, or null a particular direction. Second, there is no desired signal, while there are a few interfering signals. Third and most important difference is that both reference based and DOA algorithms require vector observations. They are commonly used in digital beamforming architectures. In the proposed architecture, there is only one signal observation which contains combined signals from all direction thus neither of these techniques can be used.

Therefore, new algorithms need to be developed for dynamic range reduction criterion given the analog antenna array architecture.

7.4.1 Dynamic range reduction algorithm

While we cannot directly apply any of the algorithms from the literature, we can still partially use or modify some of them. For example, a beamforming technique can be used to detect directions of arrival of blocking signal sources. The idea is to "steer" the antenna beam in one direction at a time and measure the output power. The steering location which results in maximum power yields the DOA estimate of a blocking signal. The calculation of array coefficients can be based on a modified zero-forcing beamforming method. There, the optimum weight vector is basically a minimum solution of a linear system, which is constrained by predetermined array responses at the desired and the co-channel users directions. The number of total interfering sources is assumed to be less than or equal the number of array elements. Here, the only difference is that optimum weight vector satisfies a constraint rather than being a minimum of a linear system.

A simple algorithm that we propose can be derived by exploiting the fact that strong primary users occupy distinct frequency bands and spatial directions of arrival. First, our goal is to resolve a 2 dimensional (spectral and spatial) map of the received signal energy (Figure 7.8). By applying an FFT on a wideband signal at the output of the A/D, a power profile in frequency domain is measured. In order to obtain

the estimate of angles of arrivals, the antenna array coefficients must sweep through many directions. Given M antenna elements, any set of $K \geq M$ independent array coefficients is sufficient to obtain the estimation of spatial distribution.

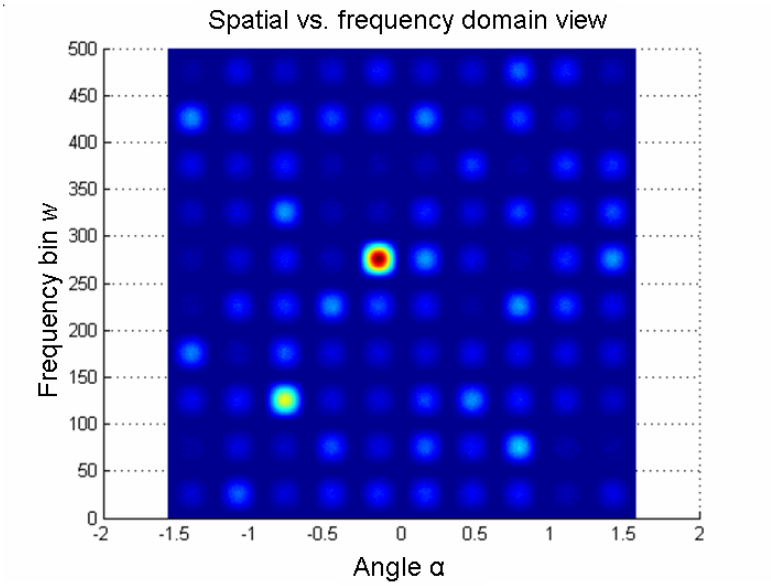


Figure 7.8: A spatial and spectral distribution of received signal energy

Let matrix A denote a set of K array coefficients used for sequential sweeping:

$$A = [\underline{a}(1)^T \underline{a}(2)^T ... \underline{a}(M)^T]_{M \text{x} K} \tag{7.17}$$

where $\underline{a}(k) = [a_1(k) a_2(k) ... a_M(k)]$ is an array weighting vector for the direction k. These K sweeping directions, i.e. array vectors, need to form an orthogonal set of basis in order to uniquely resolve the spatial dimension.

For each antenna sweep the output of the FFT for frequency f is:

$$Y(f,k) = \underline{a}(k)^T \underline{X}(f) + w(f,k) \tag{7.18}$$

where $\underline{X}(f)$ is the vector denoting input signal at frequency f on all M antennas, and $w(f,k)$ is the noise at frequency f during kth sweep. After sweeping through K different directions, vector of received signals $\underline{Y}(f)$, corresponding to a primary user at frequency f, is expressed as:

$$\underline{Y}(f) = A\underline{X}(f) + \underline{w}(f) \tag{7.19}$$

Note that vector $\underline{Y}(f)$ measures the received signal only in K pre-determined directions. Given that the strong blocking signal directions are not known in advance, it is very likely that these K directions will not coincide with blocking signal directions. However, all other directions can be estimated from these K measurements by applying the Least Squares (LS) estimation [70]:

$$\widehat{\underline{X}}(f) \;\; \text{s.t. } \min\|\underline{Y}(f) - A\underline{X}(f)\|^2 \tag{7.20}$$

$$\widehat{\underline{X}}(f) = (A^*A)^{-1}A^*\underline{Y}(f) \tag{7.21}$$

Given the measurement of received signal both in frequency and space, the strong blocking signals can be resolved. Our goal now is to find a set of coefficients that

suppresses the dynamic range. First, the algorithm selects M strongest signals in frequency domain $\{\underline{\widehat{X}}(f_1), ..., \underline{\widehat{X}}(f_M)\}$ that will be used to set coefficients. The simplest approach to meet the dynamic range requirement is to find set of coefficients that satisfy linear constraint:

$$\underline{a}_{opt}^{T}[\underline{\widehat{X}}(f_1)\underline{\widehat{X}}(f_2)...\underline{\widehat{X}}(f_M)] \leq \underline{C} \tag{7.22}$$

where $\underline{C}$ is a vector of constraints on the received power set by the desired dynamic range reduction quantity.

7.4.2 Design example

The proposed algorithm was tested with several examples of wideband signals containing strong blockers. Based on our design specifications derived earlier, we used a 4 element antenna array and 5 bit coefficient quantization. The target dynamic range reduction was set to 25 dB. Figure 7.9 presents one scenario of interest, with two blockers which are 30 dB and 40 dB stronger than the other signals in band. Both signals are narrowband, and come from different directions, 45° and 70° respectively. Note that in this example the spatial separation between blocking signals is very small, making it difficult for selective suppression.

A 128 point FFT was used for the frequency domain processing, and 8 sweeps were performed for the spatial distribution estimation. The array coefficients were computed based on the proposed algorithm. Results in Figure 7.9 show that these

blocking signals are suppressed by approximately 25 dB. These could potentially save 3-4 bits in A/D converter resolution. Note that the array response suppresses blocking signals but also changes the gain in other receive directions. However, the gain is limited by the number of antennas. Essentially with 4 antenna array, the maximum gain in one direction could be $\log 4 = 6$ dB, given that coefficients are set to create a narrow beam in that direction. Therefore, the proposed algorithm compute an array response that filters the inputs signal with the required dynamic range reduction without significant degradation of other signals in band.

Note that this algorithm uses sequential sweeping, which inherently adds latency in the computation of array coefficients. The convergence time increases linearly with number of antennas and the size of FFT. However, typical channel variations have Doppler frequencies in the order of 1 to 100 Hz, which results in long time constants (1ms to 1s) for the frequency bands of interest. For commonly used sizes of FFT (124 to 1024), number of antennas (< 10) and sampling speeds in the order of GHz the convergence times are approximately from 10 to 100 μs. One approach to speed up the convergence time and address scenarios with sudden appearance of strong blockers is to design an adaptive algorithm for the coefficient adaption.

The challenge in designing adaptive algorithms is again due to the analog array architecture. Adaptive least mean squares (LMS) and other gradient based methods require vector observations of the received signal, thus are not applicable in this context. An alternative approach could be to use a linearized random search (LRS)

developed in [71] that adapts coefficients in a linear time with number of antennas using random perturbations. Our future work will address the design and performance analysis of adaptive algorithms based on LRS for the dynamic range reduction.

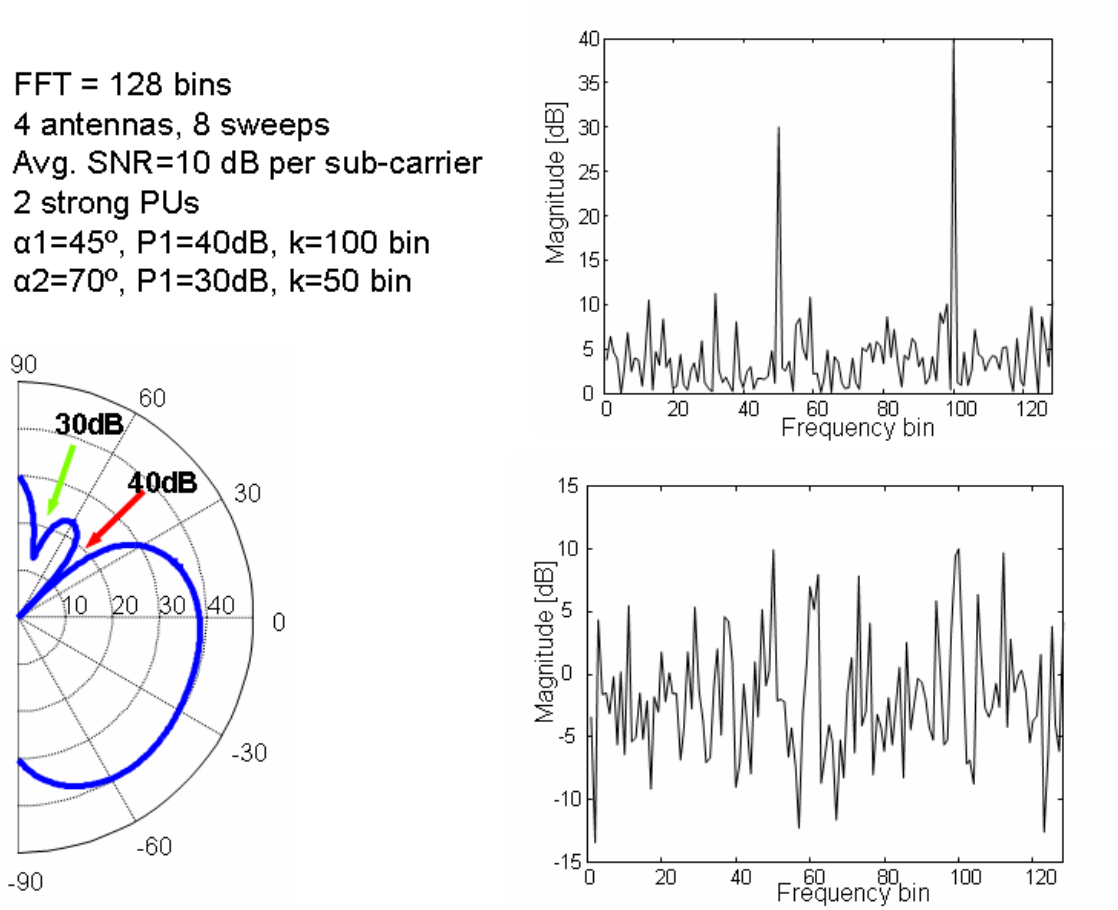


Figure 7.9: Performance of a spatial filtering approach for 2 blocking signals using 4 antenna array sensing receiver: parameters (top left), input signal (top right), resulting array response (bottom left), spatially filtered output (bottom right)

Chapter 8

Conclusions

This dissertation explores some fundamental questions in cognitive radio system design by bridging the theoretical and practical aspects of the physical and network layers. A system level design involved a closed loop research approach connecting theoretical analysis and development of new algorithms with their implementation and experimental verification. A wireless testbed platform with capabilities for real-time signal processing and protocols, networking and multiple antenna communication was developed to support this research approach.

Spectrum sensing has been identified as a key enabling functionality for cognitive radios, therefore the goal of this research was to address its feasibility, performance limits and implementation issues. A major challenge in the spectrum sensing design is the requirement to detect very weak signals of different types in a minimum time with high reliability. To solve this problem, a cross-layer design approach was applied

involving sensing radio front-end, digital signal processing and networking. As a result, an architecture of a spectrum sensing function is proposed (Figure 8.1) and its performance and implementation complexity are characterized using the developed testbed platform.

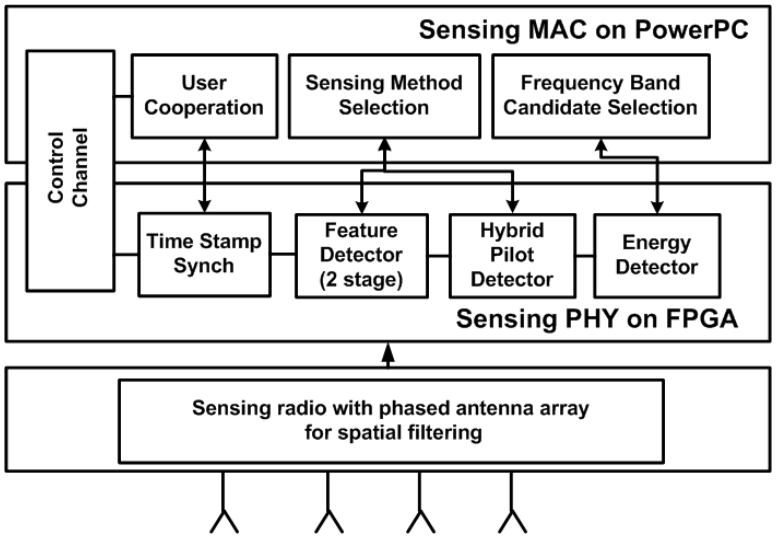


Figure 8.1: Proposed architecture for spectrum sensing implementation

8.1 Research Contributions

In this thesis, we developed a framework for the design of cognitive radio systems that use spectrum sensing for opportunistic spectrum sharing. This framework involves a systematic exploration of functions across different layers needed for spectrum sensing implementation and evaluation of their performance and feasibility. The main research contributions of this research are:

- Proposed a cognitive radio system architecture and specified physical and network layer functionalities. This architecture presents the basis for integration of core cognitive radio functions, spectrum sensing and adaptive transmission over wide bandwidths, with protocols and control channels for cognitive network management.

- Developed a wireless testbed platform for exploration of a wide range cognitive radio systems with the unified design flow. Key features involve real-time high speed signal processing and protocols, multiple antenna and multiple radio networking configurations, and simple programming interface using standard Matlab and Linux/C environment. The testbed is used for: i) experimental characterization of new functionalities under real wireless and interference channels and radio impairments; ii) design optimization to satisfy real-time constraints and practical implementation complexity.

- Characterized physical layer signal processing techniques for spectrum sensing based on energy, pilot and feature detection in terms of practically achievable sensitivities, sensing times, implementation complexity and robustness to radio and channel impairments. Developed new techniques for pilot and feature detection to overcome limitations due to radio parameter uncertainties. Results are supported by theoretical analysis, physical implementation and experimental verification.

- Characterized network layer techniques for spectrum sensing based on network cooperation in terms of improvements in sensing reliability with respect to number of radios, signal parameters and wireless channel properties. Experimental data provided insights for optimal cooperation strategies and development of protocols for spectrum sensing.

- Developed a novel radio architecture exploiting antenna array processing to spatially suppress blocking signals, and improve sensitivity of wideband sensing radios. Analyzed the design requirements and proposed algorithms for its feasible circuits implementation.

8.2 Future Work

In this dissertation, we focused mainly on the spectrum sensing portion of a cognitive radio system architecture proposed in Chapter 2. Further research is required for the development of other functionalities that interact with spectrum sensing and their joint evaluation using the presented design framework.

It has been shown that network cooperation will be essential for the robust and reliable operation of cognitive radios. However, there are a number of challenges involving implementation of cooperative protocols including scalability, network synchronization and coordinated access. As a potential solution to this problem, we proposed two kinds of control channels: universal and one per group of cognitive

radios. However, control channels impose hard requirements on network architecture and physical layer functions for communication on control channels. An existing body of ad-hoc protocols could provide directions for the development of lightweight distributed approaches to control channels. Of more fundamental importance is the understanding the impact of cooperation on the overall system capacity and spectrum utilization.

While spectrum sensing enables cognitive radio systems, the adaptive transmission over white spaces is critical for capacity increase and effectiveness of spectrum sharing. On the physical layer, it would be interesting to explore new signaling strategies for different regime of spectrum availability. In addition, a design of modulation schemes that can meet co-channel and adjacent channel interference without external filtering is desirable for a flexible digital implementation.

Use of spatial dimension and multiple antennas presents many new opportunities. As discussed in Chapter 7, resolving signals in angular domain allows selective processing of interference sources. In contrast to commonly used techniques for beamforming in a desired signal direction, in spectrum sharing scenarios multiple antenna arrays are used for suppression of signals and minimization of interference. Further theoretical work as well as design of adaptive algorithms and their implementation is required to better understand how to optimally use this additional degree of freedom. Better understanding of multiple antenna channels in spectrum sharing also requires measurements for channel model development using testbed platforms.

Bibliography

[1] J. Yang, "Spatial channel characterization for cognitive radios," Master's thesis, University of California, Berkeley, 2004.

[2] FCC, "Spectrum policy task force report," in *ET. Docket No. 02-155*, Nov. 2002.

[3] ——, "First report and order," in *FCC Docket No. 02-48*, Feb. 2002.

[4] ——, "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," in *Notice of Proposed Rule Making (NPRM) FCC Docket No. 03-322*, Dec. 2003.

[5] J. Mitola, "Cognitive radio an integrated agent architecture for software defined radio," Ph.D. dissertation, KTH Royal Institute of Technology, Stockholm, 2000.

[6] ——, "Cognitive radio for flexible mobile multimedia communications," in *Sixth International Workshop on Mobile Multimedia Communications (MoMuC99)*, June 1999.

[7] FCC, "Unlicensed operation in the TV broadcast bands," in *Notice of Proposed Rule Making (NPRM) FCC Docket No. 04-113*, May 2004.

[8] C.Cordeiro and et. al., "IEEE 802.22: The first worldwide wireless standard based on cognitive radios," in *First IEEE Int'l Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[9] I. P802.22/D.01. Draft standard for wireless regional area networks part 22: Cognitive wireless RAN medium access control (MAC) and physical layer (PHY) specifications: Policies and procedures for operation in the TV bands. [Online]. Available: http://www.ieee802.org/22

[10] S.M.Mishra, S. ten Brink, R. Mahadevappa, and R. W. Brodersen, "Detect and avoid: An Ultra-Wideband/WiMAX coexistence mechanism," in *IEEE Communications Magazine*, June 2007.

[11] D. X. W. Group, "XG Policy language framework," in *BBN Technologies Request for Comments*, Apr. 2004.

[12] 1st IEEE Symposium on New Frontiers on Dynamic Spectrum Acsess, "DySPAN05," Baltimore, Nov. 2005.

[13] 2nd IEEE Symposium on New Frontiers on Dynamic Spectrum Acsess, "DySPAN07," Dublin, Apr. 2007.

[14] D. Cabric, S. M. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz, "A cognitive radio approach for usage of virtual unlicensed spectrum," in *IST Mobile Summit 2005*, June 2005.

[15] D. Cabric and R. W. Brodersen, "Physical layer design issues unique to cognitive radio systems," in *IEEE Symposium on Personal, Indoor, Mobile Radio Communication (PIMRC)*, Sept. 2005.

[16] N. Hoven, "On the feasibility of cognitive radio," Master's thesis, University of California, Berkeley, 2005.

[17] D. Ryachaudhuri, I. Seskar, M. Ott, S. Ganu, K. Ramachandran, H. Kremo, R. Siracusa, H. Liu, and M. Singh, "Overview of the orbit radio grid testbed for evaluation of next-generation wireless network protocols," in *IEEE Wireless Communications and Networking WCNC*, Mar. 2005.

[18] C. J. Rieser, T. W. Rondeau, C. W. Bostian, and T. M. Gallagher, "Cognitive radio testbed: Further details and testing of a distributed genetic algorithm based cognitive engine for programmable radios," in *IEEE Military Communications Conference MILCOM*, Nov. 2004.

[19] [Online]. Available: http://www.gnu.org/software/gnuradio/

[20] S. M. Mishra, D. Cabric, C. Cheng, D. Willkomm, B. V. Schewick, A. Wolisz, and R. W. Brodersen, "A real time cognitive radio testbed for physical and network level experiments," in *First IEEE Int'l Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[21] C. Chang, J. Wawrzynek, and R. W. Brodersen, "BEE2: A high-end reconfigurable computing system," *IEEE Design and Test of Computers*, vol. 22, June 2005.

[22] [Online]. Available: http://www.xilinx.com/

[23] [Online]. Available: http://www.mathworks.com/products/simulink/

[24] [Online]. Available: http://www.bwrc.eecs.berkeley.edu/bee2/

[25] H. So, A. Tkachenko, and R. W. Brodersen, "A unified hardware/software runtime environment for FPGA based reconfigurable computers using BORPH," in *Intl Conference on Hardware-Software Codesign and System Synthesis*, Oct. 2006.

[26] J. Proakis, *Digital Communications*, 4th ed. Mc-Graw Hill, 2001.

[27] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, 1st ed. Cambridge University Press, 2005.

[28] T. S. Rappaport, *Wireless communications: principles and practice*, 1st ed. Prentice Hall, 1996.

[29] G. Chuinard, D. Cabric, and M. Ghosh, "Sensing thresholds," in *IEEE 802.22-06/005/r3*, May 2006.

[30] A. Sahai and D. Cabric, "Spectrum sensing: fundamental limits and practical challenges," in *First IEEE Int'l Symposium on Dynamic Spectrum Access Networks (DySPAN), Tutorial*, Nov. 2005.

[31] S. Geirhofer, L. Tong, and B. M. Sadler, "Cognitive radios for dynamic spectrum access - dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, May.

[32] D. Cabric, S. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *38th Asilomar Conference on Signals, Systems and Computers*, Nov. 2004.

[33] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Fourty-Second Allerton Conference on Communication, Control and Computing*, Sept. 2004.

[34] D. Middleton, "On the detection of stochastic signals in additive normal noise - Part I," *IEEE Transaction on Information Theory*, vol. 3, 1957.

[35] D. Slepian, "Some comments on the detection of gaussian signals in gaussian noise," *IEEE Transaction on Information Theory*, vol. 4, 1958.

[36] H. Urkowitz, "Energy detection of unknown deterministic signals," in *Proc. IEEE*, vol. 55, 1967, pp. 523–531.

[37] W. A. Gardner, *Statistical spectral analysis: a non-probabilistic theory*, 1st ed. Prentice Hall, 1987.

[38] ——, *Cyclostationarity in communications and signal processing*, 1st ed. IEEE Press, 1994.

[39] R. Durret, *Probability: Theory and Examples*, 3rd ed. Belmont:Duxbury Press, 2004.

[40] R. Tandra, "Fundamental limits on detection in low SNR," Master's thesis, University of California, Berkeley, 2005.

[41] U. Mengali and A. D'Andrea, *Synchronization techniques for digital receivers*, 1st ed. Springer, 1997.

[42] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital communication receivers: synchronization, channel estimation, and signal processing*, 1st ed. Wiley, 1997.

[43] S. M. Kay, *Fundamentals of statistical signal processing: Detection theory.* Prentice Hall, 1998, vol. 2.

[44] A. V. Oppenheim, A. S. Willsky, and I. T. Young, *Signals and Systems*, 1st ed. Prentice Hall, 1983.

[45] W. A. Gardner, "Spectral correlation of modulated signals: Part I - analog modulation," *IEEE Transaction on Communications*, vol. 35, June 1987.

[46] W. A. Gardner, W. A. Brown, and C. K. Chen, "Spectral correlation of modulated signals: Part II - digital modulation," *IEEE Transaction on Communications*, vol. 35, June 1987.

[47] W. A. Gardner, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Transaction on Communications*, vol. 36, Aug. 1986.

[48] A. V. Dandawate and G. B. Giannakis, "Statistical test for presence of cyclostationarity," *IEEE Transaction on Signal Processing*, vol. 42, Sept. 1994.

[49] G. K. Yeung and W. A. Gardner, "Search-efficient methods for detection of cyclostationary signals," *IEEE Transaction on Signal Processing*, vol. 44, May 1996.

[50] C. M. Spooner, "Performance evaluation of detectors for cyclostationary signals," Master's thesis, University of California, Davis, 1988.

[51] W. A. Gardner and C. M. Spooner, "Signal interception: Performance advantages of cyclic-feature detectors," *IEEE Transaction on Communications*, vol. 40, Jan. 1992.

[52] J. C. Liberti and T. S. Rappaport, "Statistics of shadowing in indoor radio channels at 900 and 1900 MHz," in *IEEE Military Communications Conference MILCOM*, Oct. 1992.

[53] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," in *IEEE International Conference on Communications (ICC)*, 2006.

[54] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environment," in *First IEEE Int'l Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[55] I. Y. Hoballah and P. K. Varshney, "Distributed Bayesian signal detection," *IEEE Transaction on Information Theory*, vol. 35, Sept. 1989.

[56] A.Poon, R.W.Brodersen, and D.Tse, "Degrees of freedom in spatial channels," *IEEE Transaction on Information Theory*, vol. 51, Feb. 2005.

[57] A. Poon and M. Ho, "Indoor multiple antenna characterization from 2 to 8 GHz," in *IEEE International Conference on Communications (ICC)*, June 2004.

[58] R. H. Walden, "Analog-to-digital converters survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, Apr. 1999.

[59] B. E. Carey-Smith, P. A. Warr, P. R. Rogers, M. A. Beach, and G. S. Hilton, "Flexible frequency discrimination subsystems for reconfigurable radio front ends," *EURASIP Journal on Wireless Communications and Networking, Special Issue on Reconfigurable Radio for Future Generation Wireless Systems*, vol. 2005, Mar. 2005.

[60] S. Verdu, *Multiuser Detection*, 1st ed. Cambridge University Press, 1998.

[61] D. Cabric, M. Chen, D. Sobel, S. Wang, Y. Yang, and R. W. Brodersen, "Novel radio architectures for UWB, 60 GHz, and cognitive wireless systems," *EURASIP Journal on Wireless Communications and Networking, Special Issue on CMOS RF Circuits for Wireless Applications*, vol. 2006, Apr. 2006.

[62] A. Raghavan, E. Gebara, E. M. Tentzeris, and J. Laskar, "Analysis and design of an interference canceller for collocated radios," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, Nov. 2005.

[63] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing*, 1st ed. John Wiley and Sons, 2002.

[64] J. Cheng, Y. Kamiya, and T. Ohira, "Adaptive beamforming of ESPAR antenna using sequential perturbation," in *IEEE Microwave Theory and Techniques (MTT)*, June 2001.

[65] S.T.Smith, "Optimum phase-only adaptive nulling," *IEEE Transactions on Signal Processing*, vol. 53, Nov. 2005.

[66] D.S.K.Chan and L.R.Rabiner, "Analysis of quantization errors in the direct form for finite impulse response digital filters," *IEEE Transactions on Audio and Electroacustics*, vol. 21, Aug. 1973.

[67] L. C. Godara, "The effect of phase-shifter errors on the performance of an antenna-array beamformer," *IEEE Journal of Oceanic Engineering*, vol. 10, 1985.

[68] S. Haykin, *Adaptive Radar Signal Processing*, 1st ed. John Wiley and Sons, 2007.

[69] R. O. Smidth, "Multiple emitter location and signal parameter estimation," *IEEE Transcations on Antennas and Propagation*, vol. 34, Mar. 1986.

[70] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 1996.

[71] B. Widrow and J. McCool, "A comparison of adaptive algorithms based on the methods of steepest descent and random search," *IEEE Transactions on Antennas and Propagations*, vol. 24, Sept. 1976.