Learning-Based Multi-Channel Spectrum Access in Full-duplex Cognitive Radio Networks with Unknown Primary User Activities

LEARNING-BASED MULTI-CHANNEL SPECTRUM ACCESS IN FULL-DUPLEX COGNITIVE RADIO NETWORKS WITH UNKNOWN PRIMARY USER ACTIVITIES

BY

MOHAMED HAMMOUDA, M.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2017) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	Learning-Based Multi-Channel Spectrum Access in Full-duplex Cognitive Radio Networks with Unknown Primary User Activities
AUTHOR:	Mohamed Hammouda M.Sc., (Electronics and Communications Engineering) Egypt-Japan University of Science and Technology (E-JUST), New Borg El-Arab, Egypt
SUPERVISOR:	Dr. Rong Zheng
CO-SUPERVISOR:	Dr. Tim Davidson
NUMBER OF PAGES:	xiii, 73

To my beloved parents.

Abstract

Cognitive radio had been proposed as a methodology for over-coming the inefficiency of the conventional static allocation of the available spectrum in wireless communication networks. The majority of opportunistic spectrum access schemes in cognitive radio networks (CRNs) rely on the Listen-Before-Talk (LBT) model due to the half-duplex nature of conventional wireless radios. However, LBT suffers from the problem of high collision rates and low secondary user throughput if time is misaligned among the secondary users (SUs) and the primary users (PUs). This problem can be mitigated by leveraging full-duplex (FD) communications that facilitate concurrent sensing and transmission. This thesis considers the problem of optimal opportunistic multi-channel spectrum sensing and access using FD radios in the presence of uncertain primary user (PU) activity statistics. A joint learning and spectrum access scheme is proposed. To optimize its throughput, the SU sensing period has to be carefully tuned. However, in absence of exact knowledge of the PU activity statistics, the PU's performance may be adversely affected. To address this problem, a robust optimization problem is formulated. Analysis shows that under some non-restrictive simplifying assumptions, the robust optimization problem is convex. The impact of sensing periods on the PU collision probability and the SU throughput are analyzed, and the optimal sensing period is found via convex optimization. An ε-greedy

algorithm is proposed for use by the SU to learn the PUs' activity statistics in multi-channel networks. It is shown that sublinear regrets can be attained by the proposed estimation and robust optimization strategy. Simulation studies demonstrate that the resulting robust solution achieves a good trade-off between optimizing the SU's throughput and protecting the PU.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Rong Zheng for her patient guidance, encouragement and continuous advice. She was always there when I needed help with my research. Words can not express how grateful I am for her support.

I would like also to thank Prof. Tim Davidson for his time and deep discussions from which I learnt a lot.

I am sincerely grateful to my parents, who have provided me through moral and emotional support in my life. I would like to thank my siblings and specially my brother Motaz for taking the time to read my thesis and provide me with feedback.

Special thanks go to my friends who always support and encourage me. I also thank the Cheryl Gies for her patience and assistance. And finally, also to everyone in the impact hub it was great sharing laboratory with all of you during last four years.

Thanks for all your encouragement!

Acronyms

AASIC	Analog Active Self Interference Cancellation
ADC	Analog-to-Digital Converter
ASA	Authorized Shared Access
ASIC	Active Self Interference Cancellation
AWGN	Additive White Gaussian Noise
CR	Cognitive Radio
CRN	Cognitive Radio Networks
CSCG	Circularly Symmetric Complex Gaussian
CUS	Collective Use of Spectrum
DSA	Dynamic Spectrum Access
EC	European Commission
ED	Energy Detector
FCC	Federal Communications Commission
FD	Full-Duplex
HD	Half-Duplex
LAT	Listen-And-Talk
LBT	Listen-Before-Talk
LNA	Low Noise Amplifier

LSA	Licensed Shared Access
MAB	Multi-Arm Bandit
MAC	Medium Access Control
PA	Power Amplifier
PDA	Personal Digital Assistant
PSIS	Passive Self Interference Suppression
PSK	Phase Shift Keying
PU	Primary Users
QoS	Quality of Service
RA	Receive Antenna
RF	Radio Frequency
RHS	Right Hand Side
RO	Receiving-Only
SaST	Sensing-And-Selectively-Transmit
SI	Self Interference
SIC	Self Interference Cancellation
SIS	Self Interference Suppression
SNR	Signal-to-Noise Ratio
SNRIR	Signal-to-Noise-plus-Residual-Interference Ratio
SO	Sensing-Only
ST	Sensing-And-Transmit
SU	Secondary User
TA	Transmit Antenna
WRAN	Wireless Regional Area Network

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

With the explosive growth in wireless data demand, spectrum scarcity has emerged as a primary problem for wireless service providers. Meanwhile, numerous studies conducted by agencies such as the Federal Communications Commission (FCC) in the United States have shown that much of the licensed spectrum remains unoccupied for long periods of time [2]. This has motivated the development of cognitive radio (CR) systems to efficiently exploit the under-utilized spectrum. CRs or secondary users (SUs) are wireless devices that can intelligently monitor and adapt to their environment and, hence, they are able to share the spectrum with the licensed primary users (PUs), operating whenever the PUs are idle. The key components of CR systems are mechanisms for *spectrum sensing* and *spectrum access*. Spectrum sensing allows the SUs to learn their environment prior to spectrum access, which is when the SUs actually transmit their data. Many existing opportunistic spectrum access solutions have been developed for radios that are half-duplex. To avoid excessive interference to

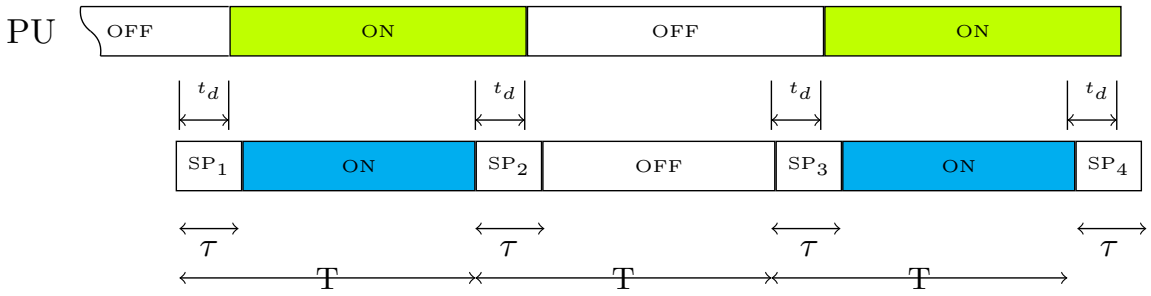


Figure 1.1: The misalignment problem

the PUs, an SU needs to first sense the availability of a selected channel for some time before commencing with transmission. This is known as the Listen-Before-Talk (LBT) model. However, LBT may suffer from the problem of time misalignment resulting in high collision rates and low PU throughput [3]. Consider the illustrative example in Figure 1.1, where "ON" and "OFF" indicate when the PU or the SU is active and inactive, respectively. In the example, the SU's first sensing period partially overlaps with the inactive period of the PU. As a result, the SU will start its transmission once it has sensed that the PU is idle. In reality, however, its transmission collides with that of the PU. The second sensing period partially overlaps with the active period of the PU. Consequently, the SU decides to refrain from transmitting the next frame while in reality it could have taken advantage of the idle spectrum. To resolve this problem, the best option is for the SU to sense and transmit at the same time. Recently, there has been significant progress in the development of full-duplex (FD) radios [4] that allow a device to transmit and receive at the same time over the same frequency by combining different self-interference cancellation techniques. In addition to potentially doubling the wireless capacity [5], we see that FD can also be leveraged to resolve the afore-mentioned misalignment problem as the SU can continue to sense the spectrum during its transmission. We find both analytically and

through simulation studies that the length of sensing/transmission periods in FD CR remains a critical parameter to be tuned in order to optimize the SU's throughput, and that the optimal sensing period depends on the PU statistics. These statistics are unknown to the SU and the SU needs to estimate them online.

In scenarios where there are multiple accessible channels with unknown PU activity statistics, a sequential learning approach that estimates the PU channel statistics while performing spectrum access is warranted. In sequential learning, an agent needs to make decisions in face of incomplete knowledge. Necessarily, there is a tension between operating on the best known channel and learning the availability of other PU channels, resulting in an instance of the so-called *exploration* and *exploitation* trade-off.

We will address this tradeoff in the development of a multi-channel opportunistic spectrum access scheme using FD radios under uncertainty in the PU statistics. We limit this study to a basic scenario, in which the SU transmitter-receiver pair communicate with one another by opportunistically accessing one of many PU channels. Multiple access schemes where multiple SUs access the same channel can be developed on top of the proposed approach and will be considered in our future work.

In our proposed spectrum access scheme, the SU operates at two time scales. At the micro time scale, the SU switches between sensing-only (SO), receiving-only (RO) and sensing-and-transmit (ST) states as dictated by the Sensing-and-Selectively-Transmit (SaST) protocol that we propose herein. At the macro-scale, which is of the same order of the PU state changes, the transition of the SU among PU channels is governed by a sequential learning policy. The efficiency of different sequential learning policies is measured in terms of their associated regret. Regret is defined

as the difference between the expected payoff gained by a "genie" (an unattainable ideal) who always uses the optimal scheme, and that obtained by the given policy.

1.2 Thesis Contributions

A closed-form analytic model is derived to characterize the impact of the sensing period and the PU idle probability on both the collision probability with the PU and the SU's throughput. To maximize the SU throughput while ensuring that the PU's performance is not adversely affected, we formulate a constrained robust optimization problem that accounts for the uncertainty in the SU's knowledge of the PU's statistics in the single channel case. We also develop a methodology based on sufficient statistics that enables the SU to learn the PU's statistics over time. We prove that under mild relaxation, such an optimization problem is convex, and hence can be efficiently solved. We also prove that its solution asymptotically converges to the optimal one.

In the multi-channel scenarios, we devise a joint learning and spectrum access policy using the solution to the robust optimization problem as a building block. In contrast to the analysis in [6–8], which is based on an idealized model of no sensing errors, the incorporation of spectrum sensing errors into our model results in sub-optimal throughput (and thus regret) even when the best channel is being accessed. This complicates the analysis of regrets. We show through rigorous analysis that the proposed scheme converges to the optimal solution incurring sublinear regrets. Extensive evaluation studies show that the theoretical model agrees with the simulation results, and the proposed algorithm provides a good trade-off between optimizing the SU's throughput and protecting the PUs. To the best of our knowledge, this is the first work that considers joint learning PU statistics and optimal control of the SU's

protocol parameters in a constrained setting. Our work makes the following new contributions.

- First, we formulate the single-channel robust optimal spectrum access problem under SaST with uncertainty in the SU's knowledge of the PU statistics and prove its convexity under mild assumptions.
- Second, we propose a joint learning and spectrum access policy for FD radios that takes into account unknown PU activities and spectrum sensing errors and operates on multiple channels.

1.3 Thesis Organization

The rest of the thesis is organized as follows. In Chapter 2, we first present definitions for CR, challenges that CR faces, spectrum sharing models, classifications of CR and the main functions of CR. We also introduce spectrum sensing and briefly discuss different spectrum sensing techniques. Then we discuss FD wireless communication in more detail. Research works related to FD communication and spectrum access schemes in cognitive radio networks (CRNs) are summarized. In Chapter 3, we discuss the proposed system model, assumptions, and the proposed learning and spectrum access strategy using FD radios. The performance of the SaST protocol is analyzed and the optimal spectrum access is formulated as a convex optimization problem for a single channel CRN in Chapter 4. In Chapter 5, we extend the problem formulation to the case of multi-channel CRNs. We also introduce the proposed strategy used to learn the PUs's activity statistics and analyze its performance. Finally, we conclude and discuss some future work in Chapter 6.

Chapter 2

Background and Related Work

In this chapter, we present background information on CRs and FD radios, and closely related work in literature.

2.1 Cognitive Radio

With the exponential growth of wireless devices and the prevalence of various radio technologies, e.g. cellular communication, WiFi, Bluetooth, the scarcity of wireless spectrum has arisen as a dominant problem in wireless communication. Underutilization of licensed spectrum motivated the development of CR systems. In CR, PUs are licensed users and can use the spectrum at any time. SUs transmit when the PUs are not active. FCC mandates that the negative impact of SUs on PU operations should be minimized. This is often accomplished by spectrum sensing and dynamic spectrum access (DSA).

2.1.1 Main Functions of Cognitive Radio

The cognitive cycle for DSA is shown in Figure 2.1 and has three main parts [1].

1. Spectrum analysis: For this task, a CR analyzes the spectrum to collect information about which channels are used and which are vacant. It also tries to identify each channel's capacity and maximum transmission rate.

2. Channel selection: After identifying vacant channels and their capacities, a CR chooses the channel that is best fit to its desired Quality of Service (QoS).

3. Controlling the transmission power and spectrum management: Having identified the best channel and its characteristics, the CR tunes its transmission settings to avoid interfering PUs or other SUs. It also keeps monitoring the channel in case it becomes unavailable.

2.1.2 Cognitive Radio Challenges

From the discussion above, one can see that an SU needs to be able to detect spectrum holes (unused frequency channels), use them when vacant, detect when a PU tries to use the bandwidth, and either hop to another spectrum or stop transmitting altogether when that happens. The fact that an SU is an unlicensed user means that PU's quality of service (QoS) may be negatively affected. To reduce this effect, SUs are required to detect the presence of a licensed user even at low signal-to-noise ratios (SNR), which is proven to be a challenging task to achieve [9], [10]. Furthermore, since SUs try to access the same channel at the same time, contention will ensure. This would degrade the overall efficiency of spectrum usage.

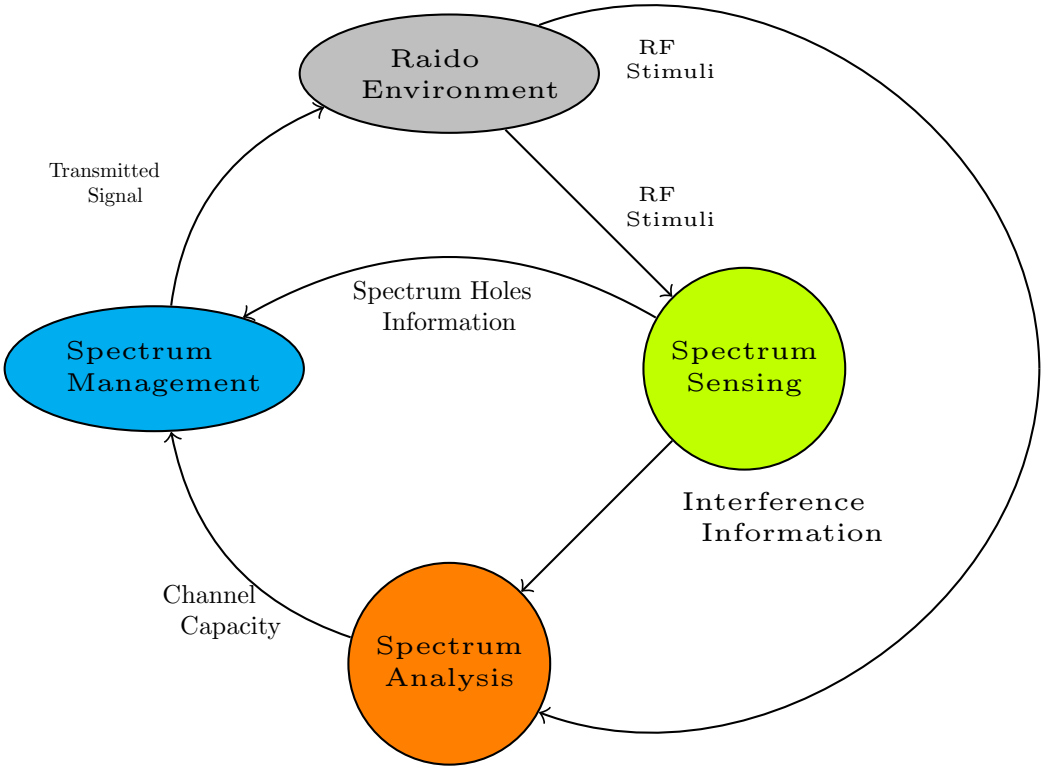


Figure 2.1: Cognitive cycle for dynamic spectrum access (DSA) [1]

2.1.3 Spectrum Sensing

According to the IEEE 802.22 standard an SU needs to achieve 0.9 probability of detection, and must be able to vacate the channel within two seconds in case a PU starts using it. This implies that continuous sensing is needed to determine if a PU is using the channel. Therefore, spectrum sensing is at the heart of CR systems. There are three main approaches: energy detection, feature detection and matched filter based detection.

2.1.3.1 Energy Detector

An energy detector (ED) samples the received signal from the PU channel, and takes a decision based on the energy of the signal. ED is usually used when little or no information is available regarding the waveform of PU's signal. It is easy to implement. However, if the information about the noise is not available, the performance of energy detection based spectrum sensing becomes unpredictable. Moreover, ED can not differentiate whether the signal comes from legitimate PUs and other SUs making it susceptible to emulation attacks [11].

2.1.3.2 Feature Detector

For many PUs, there are often specific features associated with the PUs' transmitted waveforms. These features can help the SU differentiate between noise or other SUs', and PUs' signals. For instance, in a cyclostationary feature detector, cyclostationary features such as a cyclic prefix, suffix, or the carrier frequency can be utilized. This type of detector has the advantage that it can easily differentiate between noise and PU signals because the former does not include any cyclic information [12], [13].

2.1.3.3 Matched Filter Detector

When the PU's transmitted signal and/or modulation scheme is/are known, the matched filter detector is the best detector [14]. This is due the fact that matched filters have faster response and are easy to implement. However, such information is not always available. One work around is to implement a matched filter for each modulation type (resulting in increase in system complexity), or to provide the SUs with limited information about the PUs under a certain regulating body.

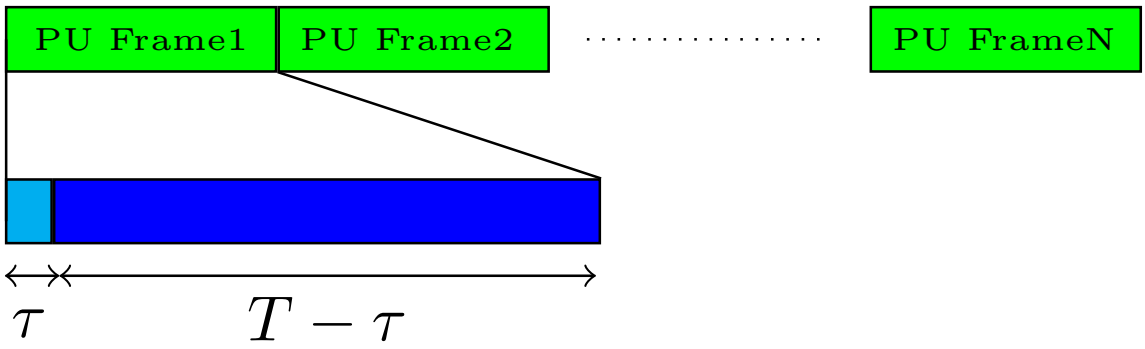


Figure 2.2: LBT Protocol

2.1.4 The Listen Before Talk Protocol

Many existing CR systems employ the so-called listen before talk (LBT) protocol for DSA as illustrated in Figure 2.2. In this protocol, an SU listens to (senses) the spectrum for a specific periodic interval of time (τ), and then talks (transmits) or remains silent (stays idle) for another interval of time ($T - \tau$) depending on the outcomes of spectrum sensing. The larger the sensing period is, the more the SU is certain of its decision about the presence of PUs and hence less interference to the PUs. However, this means that much time is wasted just for the sensing. To maximize the SU's throughput, a smaller value for τ is desirable as it increases its data transmission time. Therefore, in LBT, there exists a trade-off between protecting PUs and maximizing spectrum efficiency [15], [16].

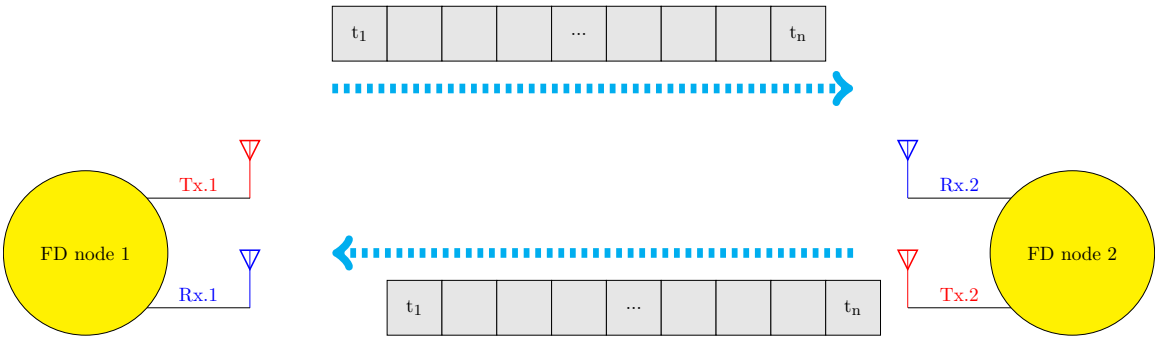


Figure 2.3: Full-duplex wireless communications

2.2 Full-Fuplex Wireless Communications

All commercial radios can transmit and receive simultaneously over two different frquency bands or transmit and receive over the same frequency but at different times. Recently, there has been significant progress in the development of FD radios [4], [17] that allow a device to transmit and receive at the same time by combining passive and active SIC techniques. Due to its potential to double the spectrum efficiency, FD communication is considered one of the candidate technologies for 5G and 802.11ax networks [5].

2.2.1 Self Interference Cancellation (SIC)

The main challenge faced in the implementation of FD wireless communication is the huge difference between the self-interference (SI) power resulting from the wireless device's own transmission and the power of the signal of interest received from another wireless device. The SI power as measured in some experiments was 50-100 dB higher than the power of the received signal of interest. The FD receiver receives the combination of the SI and the signal of interest. Intuitively speaking, if the receiver subtracts its own transmitted signal from the received combination, the resulting

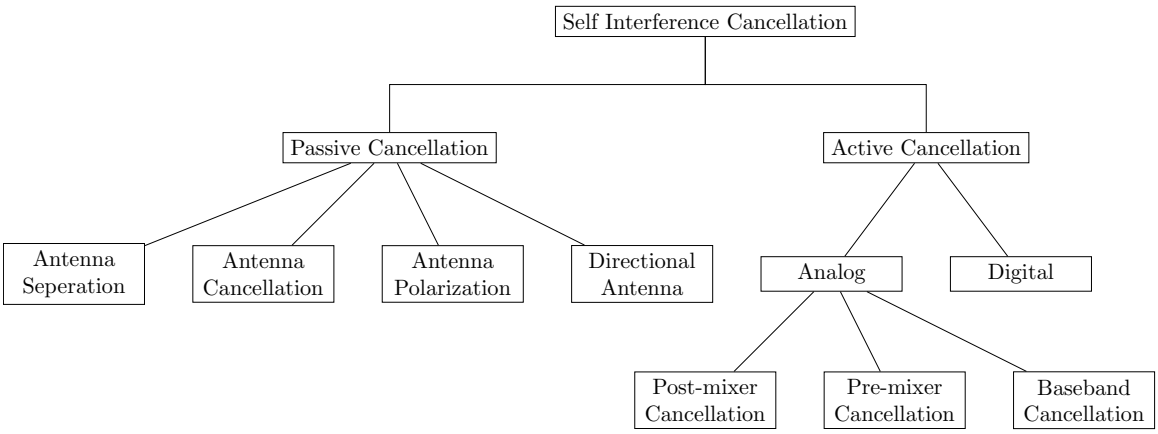


Figure 2.4: Different SIC techniques

signal should be the signal of interest. However this is difficult to implement in practice for to two reasons. First, the magnitude of the received signal could be too large due to the high SI power, causing saturation to the receiver's amplifier. Second, the high power of the SI signal compared to the received signal of interest, the automatic gain control (AGC) will be driven by the SI signal. In conventional digital receiver implementations, this can lead to high quantization noise on the signal of interest [18].

Therefore, it is important for an FD receiver to reduce SI prior to decoding the signal of interest. This process is called self-interference suppression (SIS). SIC aims to model the distortions of the received SI signal from the its original transmitted signal. Using those models, the FD receiver can compensate for those distortion. SIC techniques can be classified into passive SI suppression (PSIS) and active SI cancellation (ASIC), as shown in Figure 2.4. In the rest of this section, we give a brief introduction to different SIS schemes.

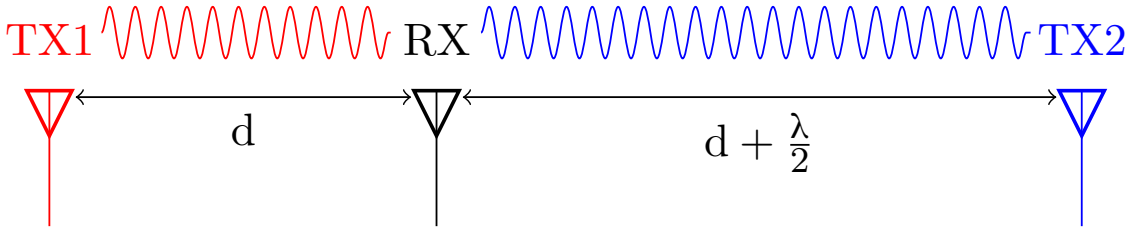


Figure 2.5: $\lambda/2$ based antenna cancellation

2.2.1.1 Passive Self Interference Suppression (PSIS)

PSIS aims to decrease the power of the propagated electromagnetic signal by electromagnetically isolating the transmit antenna(s) (TA(s)) from the receive antenna(s) (RA(s)) of the same node. This reduces the received power of the SI signal at the RAs [19]. There are various PSIS techniques including antenna separation, antenna cancellation, antenna polarization and directional antenna based suppression.

Antenna Separation In an FD node that has separate TAs and RAs, the path-loss between TAs and RAs can be simply increased by increasing the distances between those antennas. Around 40 dB of SI suppression can be achieved if antennas were separated by 30 cm [20]. Despite the simplicity of this approach, it is limited by the size of the device. Smaller form factor will limit the separation distance more and hence the achievable suppression.

Antenna Cancellation With antenna cancellation, multiple TAs are used to make the transmitted images of the signal add destructively at the RA. This can be achieved by two ways:

$\lambda/2$ **based antenna cancellation** In this technique, two TAs and one RA are used. The TAs are placed such that the distance from the two TAs to the RA are

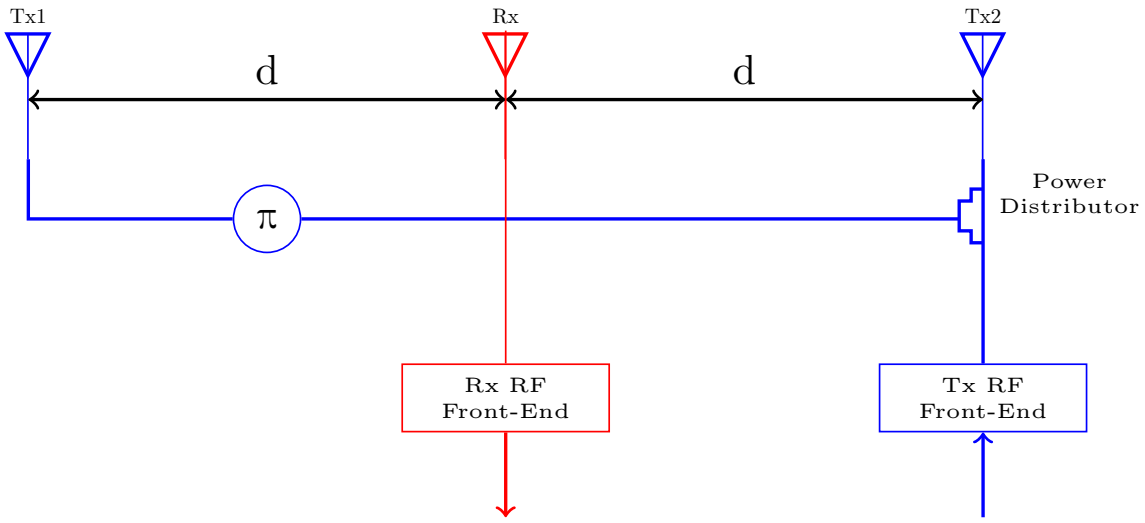


Figure 2.6: Symmetry based antenna cancellation

d and $d + \lambda/2$, where λ is the wavelength of the center carrier frequency. When the singals transmitted from both antennas are in phase, this result in the two signals arriving at the RA 180° out-of-phase. Hence adding destructively as illustrated in Figure 2.5. In [21], the authors showed that this approach can achieve up to 30 dB of SIS. Nonetheless, due to the dependency of this approach on the value of λ, it can be only applied to narrowband signals.

Symmetry based antenna cancellation This technique can be applied to both narrowband and wideband signals as it does not depend on λ. Proposed by the author of [22], the phase shift is done internally such that two copies of the singal feeding the two TAs are 180° out-of-phase as illustrated in Figure 2.6. It was shown in [23], that this approach can achieve up to 30 dB of SIS.

Antenna Polarization In this technique, the antennas of an FD node can be designed so that it only transmits vertically polarized signals while receiving only horizontally polarized ones. Based on the results shown in [24], this can help achieve

up to 15 dB of SIC.

Directional Antenna More passive SIS can be achieved using beamforming and/or directional antennas. This can be done through aligning the null direction of the TA towards the RA, as discussed in [25].

2.2.1.2 Active Self Interference Cancellation (ASIC)

PSIS techniques alone, are not sufficient for bringing the SI power down to the noise floor level. Hence, more powerful techniques are required. In ASIC, one exploits the knowledge of the transmitted signal to suppress it. To do so, one first estimates the SI channel and uses the knowledge of the transmitted signal to subtract the estimate from the received signal at the receiver. ASIC can be implemented in both the analog and digital domains as will be discussed below.

Analog Active Self Interference Cancellation Analog active self interference cancellation (AASIC) is performed in the analog domain, i.e. before the analog-to-digital converter (ADC) at the receiver. Based on where the reference signal is selected we can classify AASIC techniques into:

Post-mixer Analog Cancellation In this case, the cancellation signal is generated by by applying the RF transmitted signal after the mixer to a tapped delay line that emulates the SI channel. An example of post-mixer cancellation was proposed in [17]. The configuration of the canceller proposed in [17] is illustrated in Figure 2.7, where the blocks marked d_i are constant delays and the parameters a_i are tuned by a control algorithm that aims to minimize the undesired signal. It was shown in [4]

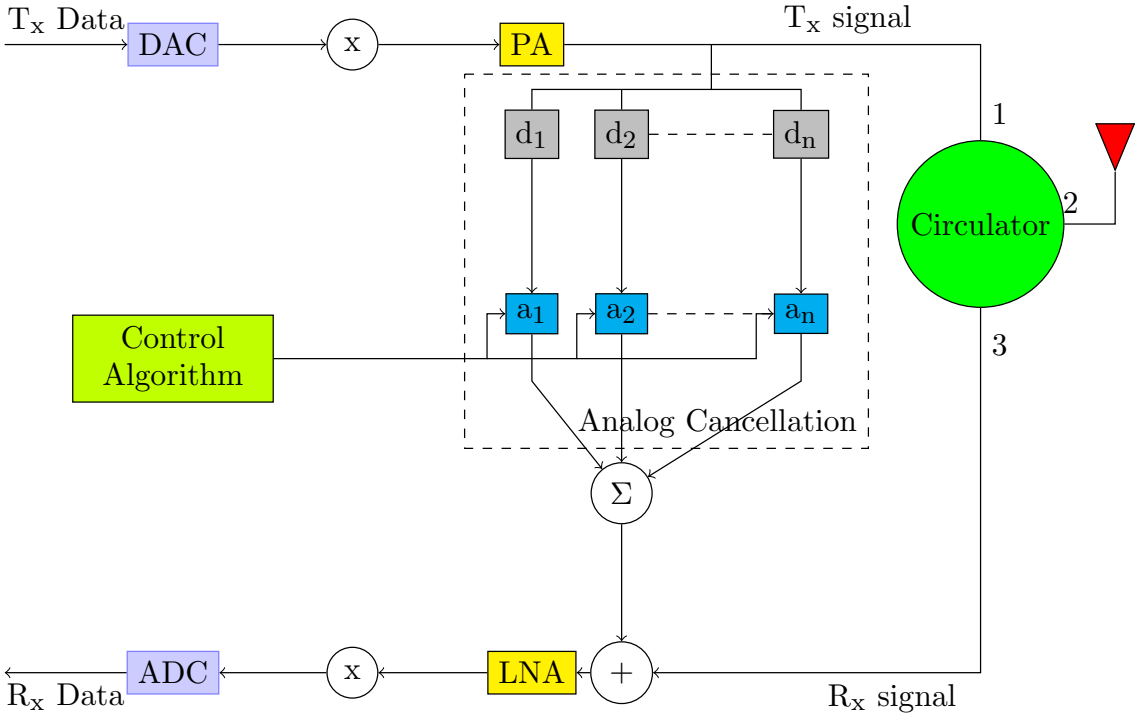


Figure 2.7: The post-mixer canceller, where PA stands for Power Amplifier and LNA is the Low Noise Amplifier

that this SIC scheme can achieve up to 45 dB of SI for a 10 MHz WiFi signal.

Pre-mixer Analog Cancellation In analog pre-mixer cancellation, the transmitted data is used to generate the analog cancellation signal [20]. Figure 2.8 illustrates the scheme, where c_1 is the cancellation signal that will be sent through an additional transmitter radio. It has been shown in [19] that this scheme can achieve up to 31 dB of SIC.

Baseband Analog Cancellation In baseband analog cancellation, the cancellation signal is generated in baseband and the cancellation is done in baseband [26], [27]. A block diagram of a baseband analog canceller is illustrated in Figure 2.9. The authors of [26] showed experimentally that this scheme can achieve more 10 dB of SIC.

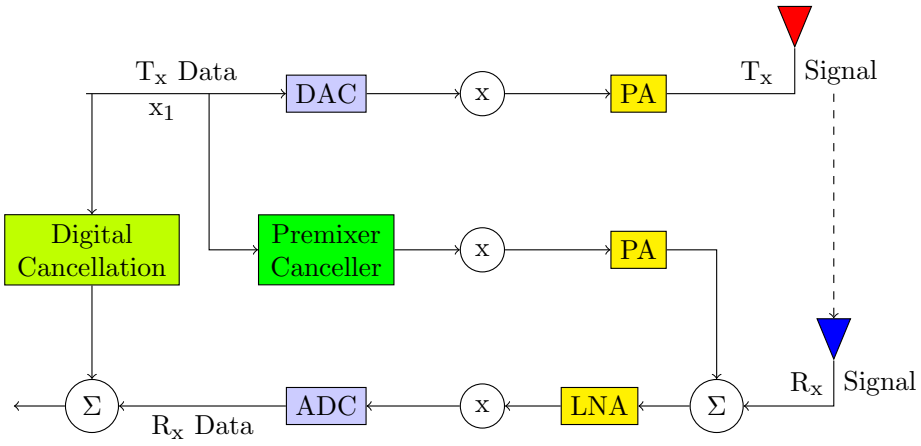


Figure 2.8: The pre-mixer canceller

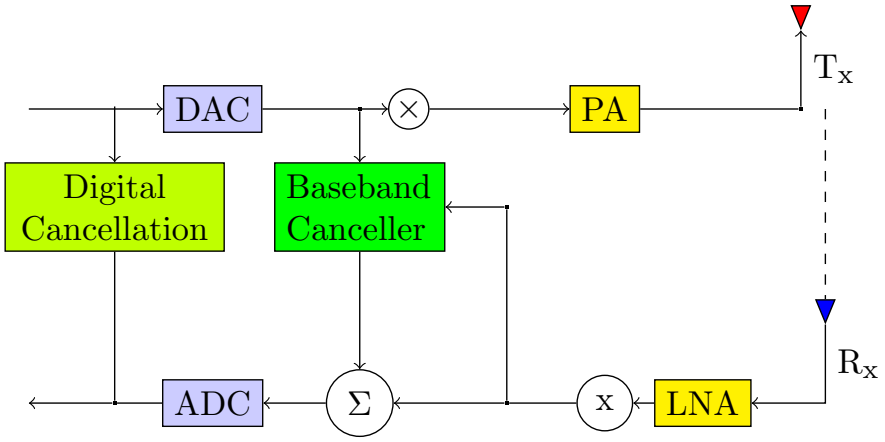


Figure 2.9: Baseband analog canceller

Digital Active Self Interference Cancellation This SIC scheme operates in the digital domain after the ADC at the receiver using the transmitted data as a reference signal, as shown show in Figure 2.8 and Figure 2.9. Operating in digital domain is advantageous as it requires less complex circuits than in analog domain. To perform digital cancellation, the cancellation circuit needs first to estimate the SI channel. The transmitted data are then processed using the channel estimate to generate the

samples to be subtracted from the received data. In [4], digital cancellation was implemented using the least square algorithm for channel estimation, due to its low complexity. In this case, the channel is the combination of the wireless channel and the analog cancellation circuit effects. The authors of [4] showed that their implementation could achieve 25-30 dB of digital cancellation on a 10 MHz WiFi signal.

2.3 Related Work

In this section, we discuss related work in CRNs and FD wireless communications.

2.3.1 Full-Duplex for Cognitive Radio

The idea of using FD techniques in CRNs has attracted a lot of interest recently [28]. In [29], [30], a Listen-and-talk (LAT) protocol was proposed with the help of the FD technique that allows SUs to simultaneously sense and access the vacant spectrum. In LAT, an SU switches between a sensing-only state, and a sensing-and-transmit states in close synchronization with the PU's activities. In [31], Cheng *et al.* considered the use of FD to allow non-time-slotted transmission of SUs, as an FD SU can abort the transmission when it senses the presence of a PU.

In [32], Afifi and Krunz considered an SU that can operate in either simultaneous transmit-and-sense mode or the simultaneous transmit-and-receive mode following a brief sensing period, and proposed an optimal mode-selection strategy that maximizes the SU's utility function subject to a constraint on the PU collision probability. In [33], it is assumed that the active/idle periods of different PU channels are identically distributed and are known beforehand through measurements. An SU switches to a different PU channel when the PU on the current channel is very likely to be busy. In [34], a MAC protocol was proposed for multichannel non-time-slotted CRNs. It is assumed that the PU also senses the spectrum before transmission to avoid collision. The SU alternates between FD and HD modes and uses FD in the absence of synchronization between the PU's and SU's frames. The proposed MAC protocol in [34] maximizes the channel utilization of the SUs in multi-channel settings without compromising PU throughputs. In our work, we aim to protect the PU without the

need for the PU to sense the spectrum.

2.3.2 Sequential Learning for Multi-channel CR

The problem of opportunistic spectrum access is similar to the multi-arm bandit (MAB) problem [35]. The MAB problem has been extensively investigated in the sequential learning literature. In the MAB problem, a gambler needs to decide the sequence of plays among a row of slot machines. The gambler's objective is to maximize the total reward earned through the sequence of played machines. In our case, the gambler is the SU and the slot machines are the PUs' channels. Several authors have investigated learning strategies to maximize the payoff to the SUs by choosing the order of PU channels to access.

Zheng *et al.* illustrated the theoretical foundation of the MAB problem and discussed its applications in the context of radio resource management [36]. Lai *et al.* applied the upper confidence bound 1 (UCB1) algorithm [37] to single SU channel selection and later extended it to consider Markovian payoffs for the case of multiple SUs in [7]. Liu and Zhao [8] formulated the problem of SU channel selection as a decentralized MAB problem, and presented a policy that achieves asymptotically logarithmic regret (defined as the difference between the throughput attainable over the best channel and that attained by the proposed policy) in time. Anandkumar [6] proposed two policies for distributed learning and access which achieve order optimality in terms of regret. In addition to learning the channel availability, the SUs also learn the other users' strategies and the total number of users in the system through channel feedback.

Kalathil *et al.* proposed in [38] an online learning algorithm for multiple SUs in

CRNs. The potential for collision among SUs picking the same channel is considered. The learning scheme in [38] is decentralized and the SUs do not coordinate in accessing a channel. The proposed learning algorithm was found to achieve a regret that is logarithmic in time. Existing work applying the principles of the MAB problem to full duplex cognitive radios assumes perfect sensing where an SU can always correctly sense whether a PU channel is available or not. Furthermore, only HD communication is allowed. In our work, we will use an ε-greedy algorithm [39] under the assumption of imperfect sensing and the use of FD.

2.3.3 Uncertainty in the knowledge of PU parameters

To be able to operate in a CRN, an SU needs to estimate a lot of parameters. These parameters include noise level, PUs' activity rates and channel gains between the SU and the PUs. In [40], [41] and [42], the effect of noise uncertainty on the performance of energy detectors in PU detection was discussed. Statistical modelling of the uncertainty and performance bounds for the probability of detection were investigated. The authors of [43], examined the generalized energy detector, where the squaring of amplitude of received samples in conventional energy detector is replaced by an arbitrary positive operator. It was shown that the conventional energy detector is the best energy detector under noise uncertainty. In [43], the SNR wall, defined as the SNR threshold below which it is impossible to satisfy a given detection requirement, was derived for the generalized ED in case of noise uncertainty. In [44], a robust optimization problem for SU power control was formulated to maximize the SU's throughput and limit the interference to PUs under the uncertainty in the PUs' locations. However, to the extent of our knowledge, we are the first to consider the

effect of uncertainty in the SU's knowledge of the PU's activity statistics.

Chapter 3

System Model and SU Strategy

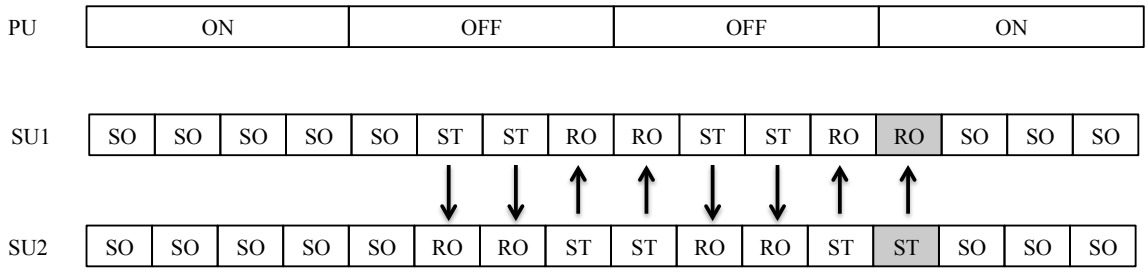


Figure 3.1: The SU to SU communication

To establish a model for the system of interest, consider a setting in which multiple PUs operate on K different channels $\mathcal{K} = \{1, 2, \ldots, K\}$. The occupancy of channel $k \in \mathcal{K}$ is assumed to follow a slotted Bernoulli process with an (unknown) parameter θ_k (called the *idle probability*) and a frame duration T . While the parameter θ_k is not known to the SU, the SU can estimate θ_k using past observations. The SU can utilize an FD radio to sense and transmit at the same time. The SU can be in one of three states: SO (sensing-only) state, RO (receive-only) and ST (sensing and transmitting) state. As the names suggest, in the SO state, the SU only senses the channel; in the RO state, the SU only receives the data transmitted from another SU; and in the

ST state, the SU senses the PU's channel and transmits to another SU on the same channel at the same time. The transmitting SU is the one responsible for sensing the PU's channel it is currently using.

The communication between two SUs is illustrated in Figure 3.1. In the RO and SO states, the SU uses HD. The SU uses FD only when it is in the ST state. As the ST and RO states can not exist in the same SU at the same time, we only consider SO and ST states in our modelling. Hence, we will not mention the RO state for the rest of this thesis.

3.1 SU Strategy

In CR, the SU should operate transparently to the PU. Hence, the SU should not assume any sort of coordination with the PU. In addition, in order to ensure that the PU's communication quality is not significantly affected, we restrict the collision probability with the PU to be within acceptable levels. We assume that the SU can estimate the boundaries of the PU's frames and exactly know their duration T. However, it does not have to be perfectly synchronized with it. At the beginning of a PU frame, the SU needs to make two decisions.

- Which PU channel in $\mathcal{K}$ to operate on, and
- What is the sensing and transmission slot duration τ.

Clearly, the SU should try to operate on the channel that will maximize its throughput. However, in absence of the exact knowledge of PU channel statistics, a sequential learning strategy needs to be adopted that updates the estimates of θ_k

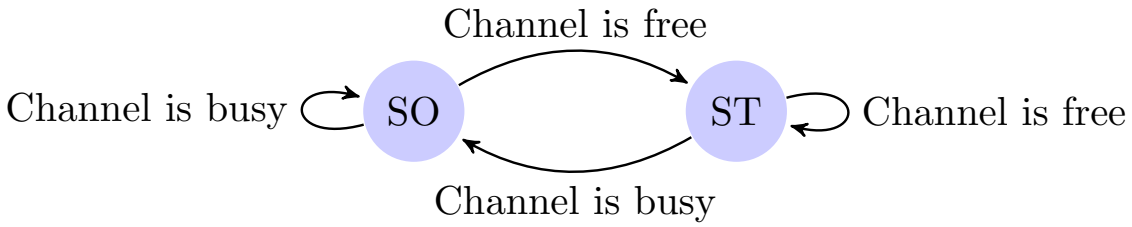


Figure 3.2: The SU activity modelled as a two-state machine

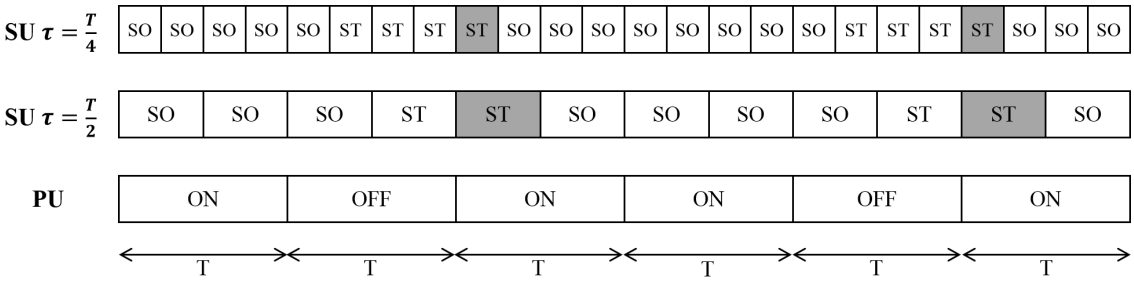


Figure 3.3: PU Frames vs. SU slots

and selects a channel in each frame. In Chapter 5, we adopt the ε-greedy policy for MABs [37] as our sequential learning strategy.

Upon selecting a PU channel, the SU starts from the SO state. After sensing the channel for a duration τ, depending on its detection of the currently accessed PU channel state (busy or idle), the SU either stays in the SO state or transits to the ST state for the next slot. It stays in this state for a further duration τ. In the ST state, it both senses and transmits simultaneously. The process is then repeated. The state transitions of the SU within one frame are illustrated Figure 3.2. We call this scheme the *Sensing-and-Selectively-Transmitting (SaST)* protocol. We will start by developing the SaST protocol for a single PU channel. Then, in Chapter 5, we will generalize the SaST protocol for the multi-channel case. In summary, the SU operates on two time scales: frames of size T, which coincide with that of PUs for channel selection; and slots of size $\tau < T$ for transitions between the SO and ST

states (Figure 3.3). The rationale behind this protocol is that the switching delay between transmitting and receiving states in typical wireless devices is much faster (tens of nano-seconds or less) than switching between channels, which may take up to several milliseconds. The key notations used in the rest of the thesis are summarized in 3.1.

Table 3.1: Key notations

Symbol	**Definition**
T	PU frame duration
τ	SU slot duration
γ	Ratio of the PU's signal power and the noise power at the SU
f_s	Sampling frequency
P_d	Probability of detection
P_f	Probability of false alarm
P_c	Probability of collision
η	Decision threshold
θ_k	The probability that PU k is idle
λ	SU throughput
$\bar{P}_c$	Maximum allowable collision probability
$C_0(C_1)$	SU throughput in absence of (with) PU transmissions

3.2 PU Detection

The SU uses energy detection to sense the presence of PUs. We refer to the situation when the PU is idle as hypothesis $\mathcal{H}_0$, and the situation when the PU is active as hypothesis $\mathcal{H}_1$. Under hypothesis $\mathcal{H}_1$, the discrete signal received at the SU can be represented as

$$y(m) = s(m) + u(m), \tag{3.1}$$

where $y(m)$, $s(m)$ and $u(m)$ are the amplitudes of the m^{th} sample of the total received signal, the (scaled version of the) PU's signal and the noise signal received at the SU receiver, respectively. Under hypothesis $\mathcal{H}_0$, the discrete signal received at the SU can be represented as

$$y(m) = u(m). \tag{3.2}$$

In hypothesis testing, a test statistic is a qauntity derived from the samples used to decide to support either $\mathcal{H}_0$ or $\mathcal{H}_1$ [45]. For energy detection, the test statistic is [46],

$$M = \frac{1}{\lceil \tau f_s \rceil} \sum_{m=1}^{\lceil \tau f_s \rceil} |y(m)|^2, \tag{3.3}$$

where τ is the sensing duration and f_s is the sampling frequency. Typically, many samples would be taken in each sensing duration, and hence, $f_s \gg \frac{1}{\tau}$. As energy detection is a threshold-based detection, the probability of false alarm P_f and the probability of detection P_d can be written as,

$$P_f = Pr(M > \eta | \mathcal{H}_0), \tag{3.4}$$

and

$$P_d = Pr(M > \eta | \mathcal{H}_1), \tag{3.5}$$

where $\eta \geq 0$, Pr indicates the probability and η is the detection threshold. Table 3.2 summarizes the various probabilities of interest. In frequency-flat additive white Gaussian noise (AWGN) channels with zero-mean circularly symmetric complex Gaussian (CSCG) noise and complex-valued PSK signals, the probability of detection is a function of the sensing duration τ and the decision threshold η [46],

Table 3.2: Probabilities of Interest

$P(\mathcal{H}_0\|\mathcal{H}_0) = 1 - P_f$	The probability that the SU observes that the PU is idle given the PU is idle.
$P(\mathcal{H}_0\|\mathcal{H}_1) = 1 - P_d$	The probability that the SU observes that the PU is idle given the PU is active (also called *the probability of misdetection* P_m).
$P(\mathcal{H}_1\|\mathcal{H}_0) = P_f$	The probability that the SU observes that the PU is active given the PU is idle (also called *the probability of false alarm* P_f).
$P(\mathcal{H}_1\|\mathcal{H}_1) = P_d$	The probability that the SU observes that the PU is active given the PU is active (also called *the probability of detection* P_d).

namely,

$$P_d(\eta, \tau) = \mathcal{Q}\left(\left(\frac{\eta}{\sigma_u^2 + \sigma_i^2} - \gamma - 1\right)\sqrt{\frac{\tau f_s}{2\gamma + 1}}\right), \tag{3.6}$$

where σ_p^2, σ_i^2 and σ_u^2 are the received PU's signal power, the residual self-interference power and the noise power at the SU's receiver, respectively. The term $\gamma = \frac{\sigma_p^2}{\sigma_u^2 + \sigma_i^2}$ is the signal-to-noise-plus-residual-interference ratio (SNRIR) of the received signal under hypothesis $\mathcal{H}_1$. The notion $\mathcal{Q}$ denotes the standard Q-function [47]. The probability of false alarm is given by

$$P_f(\eta, \tau) = \mathcal{Q}\left(\left(\frac{\eta}{\sigma_u^2 + \sigma_i^2} - 1\right)\sqrt{\tau f_s}\right). \tag{3.7}$$

When the SU operates in the SO state, $\sigma_i^2 = 0$. Therefore, when the SU operates in the ST mode, the SI power depends on the used self-interference cancellation (SIC) techniques used. Recent research shows that the self interference power can be reduced to or below the noise floor [48], [49] and [50]. However. in this thesis we assume perfect SIC, i.e. $\sigma_i^2 = 0$ and $\gamma = \frac{\sigma_p^2}{\sigma_u^2}$ in both ST and SO states. The effects of the SI residual power will be considered in our future work.

As $\gamma \geq 0$, the argument of the Q-function in (3.6) is always less than the corresponding one in (3.7). As the Q-function is monotonically decreasing function with its argument, we have that $P_d(\eta, \tau) \geq P_f(\eta, \tau)$.

To guarantee a specific level of protection for the PU, we can set the probability of detection to a fixed value $P_d = \bar{P}_d$. For a given τ, we can compute η as,

$$\eta = \sigma_u^2 \left(1 + \gamma + \mathcal{Q}^{-1}(\bar{P}_d)\sqrt{\frac{2\gamma + 1}{\tau f_s}}\right). \tag{3.8}$$

The resulting false alarm probability is

$$P_f(\tau, \bar{P}_d) = \mathcal{Q}\left(\sqrt{2\gamma + 1}\ \mathcal{Q}^{-1}(\bar{P}_d) + \gamma\sqrt{\tau f_s}\right). \tag{3.9}$$

3.3 The SaST protocol in a single channel network with known PU idle probability

The duration τ for which the SU senses (and transmits) is critical to the performance of both the SU and the PU. A smaller τ incurs larger false alarm and misdetection probabilities. Misdetecting an active PU causes collisions with the PU, while false alarms reduce the attainable throughput of the SU. Thus, τ needs to be chosen to maximize the secondary user throughput subject to the constraint on collision probability.

To find the optimal τ, in this section, we first provide an analytical characterization of the SU throughput and the PU collision probability constraint given the PU idle probability under the SaST protocol. First we will derive the formula for the simplest case of one SU and one PU. Then, we will generalize the formula for multiple PUs in

Chapter 5. For simplicity, we will remove the subscript k from θ_k for the single PU channel case in the rest of this chapter and in Chapter 4.

3.3.1 Two-state Markovian Approximation of PU Activities

First, we approximate the SU's view of the PU's activities as a two-state Markov chain. This allows a closed-form representation of the aforementioned quantities. It will be verified through simulations that such an approximation does not introduce too much error.

Consider discretized time points with separation τ. The PU can be in one of the two states: active ("1") or idle ("0"). Let P_{ij}^p be the probability that the primary user transitions from state i to j where $i, j \in \{0, 1\}$. Recall that the frame duration of the PU is T. The number of SU slots in T is T/τ. When the SU senses the last slot of an active frame of the PU, it sees the transition from the active to idle states if the PU next goes to an idle frame. Thus, $P_{10}^p = \frac{\tau\theta}{T}$. Likewise, when the SU senses the last slot of an idle frame of the PU, it sees the transition from the idle to active states if the PU goes to an active frame. We have $P_{01}^p = \frac{\tau(1-\theta)}{T}$. Therefore, the PU state transition matrix as observed by the SU is given by,

$$\mathbb{P}_{PU} = \begin{bmatrix} 1 - \frac{\tau(1-\theta)}{T} & \frac{\tau(1-\theta)}{T} \\ \frac{\tau\theta}{T} & 1 - \frac{\tau\theta}{T} \end{bmatrix} \tag{3.10}$$

Table 3.3: The joint system states

State	PU	SU
S_1	1	1
S_2	0	1
S_3	1	0
S_4	0	0

3.3.2 System Markov chain

To obtain a joint Markov model for both the SU and the PU, we identify the four possible states for the system shown in Table 3.3. In the table, "1" represents the PU/SU being in the active/ST state and "0" represents the PU/SU being in the idle/SO state. The state diagram of the combined Markov chain is shown in Figure 3.4, where P_{jk} is the transition probability from state S_j to state S_k.

Given the false alarm and detection probabilities of the spectrum sensing system, the state transition matrix is given by,

$$\mathbb{P} = \begin{bmatrix} (1-P_d)P_{11}^p & (1-P_d)P_{10}^p & P_d P_{11}^p & P_d P_{10}^p \\ (1-P_f)P_{01}^p & (1-P_f)P_{00}^p & P_f P_{01}^p & P_f P_{00}^p \\ (1-P_d)P_{11}^p & (1-P_d)P_{10}^p & P_d P_{11}^p & P_d P_{10}^p \\ (1-P_f)P_{01}^p & (1-P_f)P_{00}^p & P_f P_{01}^p & P_f P_{00}^p \end{bmatrix}. \tag{3.11}$$

To understand how $\mathbb{P}$ is derived, we note first that the transition of PU states is independent of those of the SU. On the other hand, transitions of the SU's state

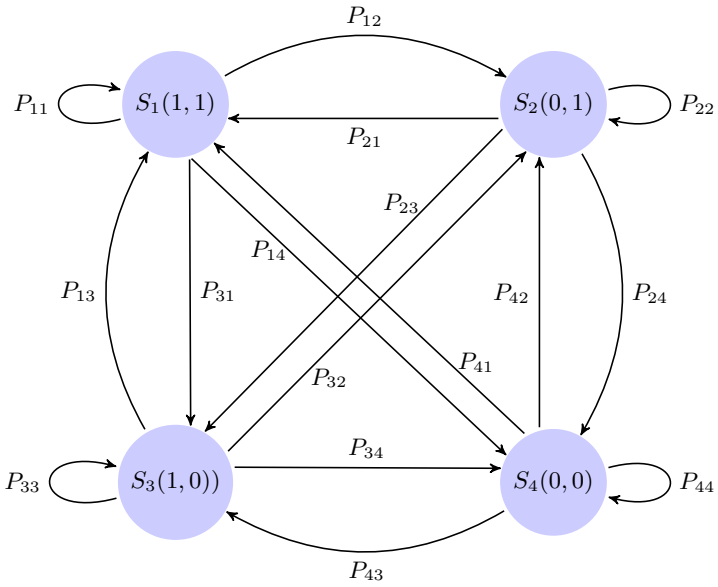


Figure 3.4: Combined Markov chain state diagram

depends on the current SU and PU states. As a concrete example,

$$Pr(S_2|S_1) =$$

$$Pr(\text{the PU transitions from active to idle}) \times Pr(\text{SU stays in the active state})$$

$$= (1 - P_d)P_{10}^p,$$

as the PU goes from the active state to the idle state with probability P_{10}^p, and SU stays in the active state when it misdetects an active PU as idle in the current slot with the probability $(1 - P_d)$.

3.3.3 Steady state probabilities

Let the steady state probabilities of the system be denoted by $\boldsymbol{\pi} = [\pi_{1,1}, \pi_{0,1}, \pi_{1,0}, \pi_{0,0}]$, where the first and the second subscripts indicate the state of the PU and the SU,

respectively. We have

$$\boldsymbol{\pi}\mathbb{P} = \boldsymbol{\pi}, \tag{3.12}$$

and

$$\sum_i \pi_i = 1. \tag{3.13}$$

Solving (3.12) and (3.13) yields,

$$\pi_{1,1} = (1-P_d)(1-\theta)P_{11}^p + (1-P_f)\theta P_{01}^p, \tag{3.14}$$

$$\pi_{0,1} = (1-P_f)\theta P_{00}^p + (1-P_d)(1-\theta)P_{10}^p, \tag{3.15}$$

$$\pi_{1,0} = P_f\theta P_{01}^p + P_d(1-\theta)P_{11}^p, \tag{3.16}$$

$$\pi_{0,0} = P_f\theta P_{00}^p + P_d(1-\theta)P_{10}^p. \tag{3.17}$$

In the special case when there is no sensing error, i.e., $P_d = 1$, and $P_f = 0$, $\mathbb{P}$ and $\boldsymbol{\pi}$ will be,

$$\mathbb{P} = \begin{bmatrix} 0 & 0 & P_{11}^p & P_{10}^p \\ P_{01}^p & P_{00}^p & 0 & 0 \\ 0 & 0 & P_{11}^p & P_{10}^p \\ P_{01}^p & P_{00}^p & 0 & 0 \end{bmatrix},$$

and

$$\boldsymbol{\pi} = \begin{bmatrix} \theta P_{01}^p & \theta P_{00}^p & (1-\theta)P_{11}^p & (1-\theta)P_{10}^p \end{bmatrix}. \tag{3.18}$$

Substituting (3.8), (3.9) and (3.10) in (3.11), (3.14)–(3.17), we can obtain the transition matrix and steady state probabilities for the joint PU and SU states as functions of τ, θ and $\bar{P}_d$.

3.3.3.1 Collision Probability and the SU's throughput

Now, we are in the position to analyze the performance of the SaST protocol in terms of the collision probability P_c and the SU's throughput λ. We define the collision probability as the probability that the SU is in the ST state while the PU is in the active state. Using Bayes' rule, this can be expressed as the conditional probability

$$P_c(\tau,\theta) = P(\text{SU is active} \mid \text{PU is active}) = \frac{P(\text{SU is active, PU is active})}{P(\text{PU is active})} = \frac{\pi_{1,1}}{1-\theta}, \quad (3.19)$$

where $\pi_{1,1}$ is the probability that the PU and SU are both active (collision with the PU). This is better than using $\pi_{1,1}$ as a measure for the effect on the PU because it protects the PU regardless of its activity rate. From (3.10) and (3.14), we have that

$$\begin{aligned} P_c(\tau,\theta) &= (1-P_d)(1-\frac{\tau\theta}{T}) + (1-P_f)\frac{\tau\theta}{T} \\ &= 1 - P_d + (P_d - P_f)\frac{\tau\theta}{T}, \end{aligned} \quad (3.20)$$

where we have left the dependence of P_d and P_f on τ implicit. The SU may transmit in two different cases. The first is when the PU is idle and the channel is free. In this case, the SU's channel capacity (SU throughput) is $C_0 = B\log_2(1+\gamma_s)$, where γ_s is the SNR of the transmitting SU's signal received at the receiving SU's receiver when only the SU is transmitting (no collision), and B is the bandwidth. On the other hand, when the SU collides with the PU, the SU throughput is $C_1 = B\log_2\left(1+\frac{\gamma_s}{1+\gamma}\right)$. Clearly, $C_0 \gg C_1$. The average SU throughput is thus,

$$\lambda(\tau,\theta) = \pi_{1,1}C_1 + \pi_{0,1}C_0, \quad (3.21)$$

where $\pi_{0,1}$ is the probability that the PU is idle and the SU is active. Because $\pi_{i,j}$ are functions of both τ and θ, the average SU throughput in (3.21) can thus be expressed as,

$$\lambda(\tau,\theta) = \pi_{1,1}C_1 + \pi_{0,1}C_0. \tag{3.22}$$

Using (3.17),(3.15) and (3.22) we find

$$\begin{aligned}\lambda(\tau,\theta) =& \frac{(C_0 - C_1)(P_d - P_f)\tau}{T}\theta^2 + \left(C_0(1-P_f) - C_1(1-P_d) + (2 - P_d - P_f)\frac{\tau(C_0 - C_1)}{T}\right)\theta \\ &+ (1-P_d)C_1 + (1-P_f)\left(C_0 - (C_0 - C_1)\frac{\tau}{T}\right),\end{aligned} \tag{3.23}$$

where, once again, we have left the dependence of P_d and P_f on τ implicit.For the rest of this thesis, we will drop the word "average" from the average SU throughput.

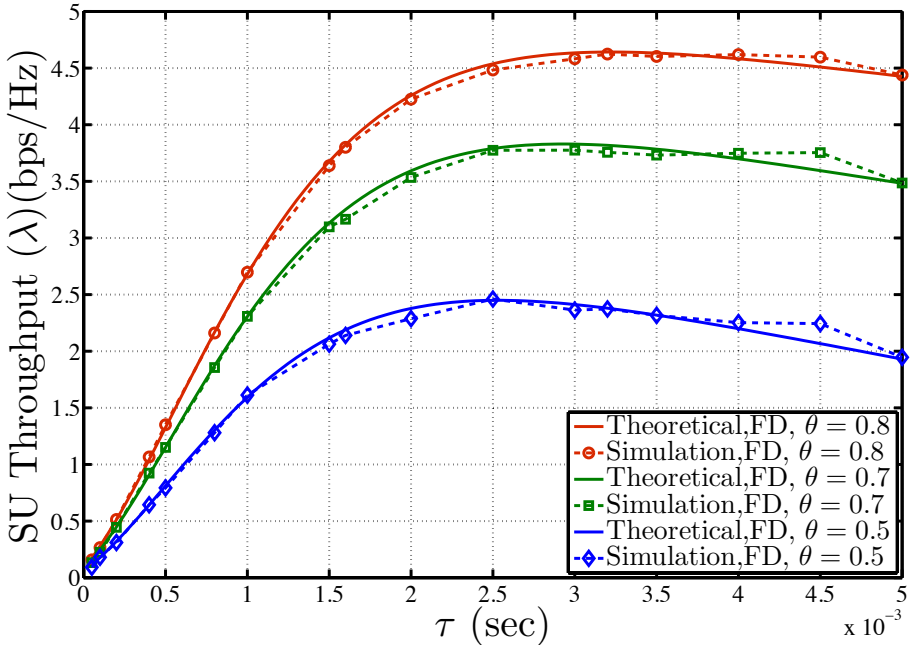
3.4 Model Validation

In order to validate our model, we simulated a single channel netwrok in which the PU transmits QPSK-modulated signals with passband bandwidth of 6MHz when active. The sampling frequency is at the Nyquist rate f_s = 6MHz. An AWGN channel model is assumed where the noise is characterized as a zero-mean CSCG process. In CRNs, we are interested in detecting PUs at very low SNR in order to protect the PU. Two different values are used for the SNR of the received signal from a PU at the SU receiver. To calculate the detection threshold for the energy detector, we set $\gamma = -15$dB. When examining collisions between the PU and SU, we are more interested in the high SNR regime, hence $\gamma = 20$dB is used. The target probability of detection is chosen to be $\bar{P}_d = 0.99$. The SNR of the SU transmissions is $\gamma_s = 20$dB,

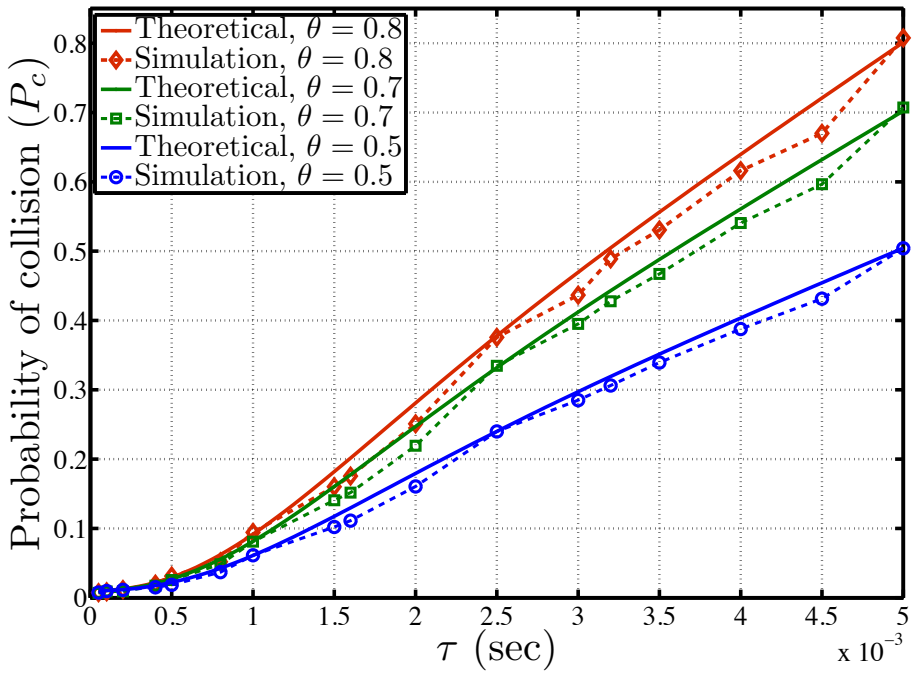
and thus the normalized SU throughputs in both cases are $\bar{C_0} = \log_2(1+\gamma_s) = 9.9582$ bps/Hz and $\bar{C_1} = \log_2\left(1+\frac{\gamma_s}{1+\gamma}\right) \approx 1$.

We validate the model by comparing both the SU throughput and the collision probability from the simulation and the theoretical results. The PU frame duration is set to be $T = 5$ msec. Figure 3.5a shows the theoretical and simulation results for the SU throughput as a function of the SU slot duration τ for different values of PU's idle probability θ. For example, for $\theta = 0.7$ the maximum achievable throughput is 3.829 bps/Hz at $\tau^* = 2.9$ msec. It is clear from the figure that simulation results almost coincide with the theoretical prediction. We also observe that the SU throughput is a concave function of the slot length τ. Figure 3.5b shows the collision probability and theoretical model. Again, we observe a good agreement between the two. Also, the monotonicity of P_c with respect to τ can be verified in the figure.

We notice also that both the analytical and simulation curves coincide when $T/\tau = q$, where q is a positive integer. In Figure 3.5, this is clear when τ= 0.5, 1, 2.5 and 5 msec. This is due the way our model was derived. However, even in that case we can that the collision probability analytic form acts as an upper-bound for the actual collision probability. This is a desirable property, as one of our goals is to protect the PU.



(a) SU throughput



(b) Collision probability

Figure 3.5: SU throughput (λ) and collision probability (P_c) vs slot duration (τ) for single channel FD spectrum sensing with $T = 5$msec and $\bar{P}_d = 0.99$

Chapter 4

Optimal Single-Channel SaST Protocol

The goal of this chapter is to design an optimal SaST protocol such that the SU throughput is maximized while the interference to the PU is kept acceptable. When the PU's idle probability is known, such a problem can be formulated in a straightforward manner. The problem can be formulated as:

$$\begin{aligned} \mathbf{P1}: \quad & \max_{\tau} \ \lambda(\tau, \theta) \\ \text{s.t.} \quad & P_c(\tau, \theta) \leq \bar{P}_c \\ & \tau \in [0, T], \end{aligned} \tag{4.24}$$

where $\lambda(\tau, \theta)$ was given in (3.23), $P_c(\tau, \theta)$ was give in (3.20) and $\bar{P}_c$ is the maximum allowable collision probability. Clearly, the solution to (4.24) guarantees that the collision probability would not exceed $\bar{P}_c$.

However, in practice, the PU's idle probability is not known a priori and needs to

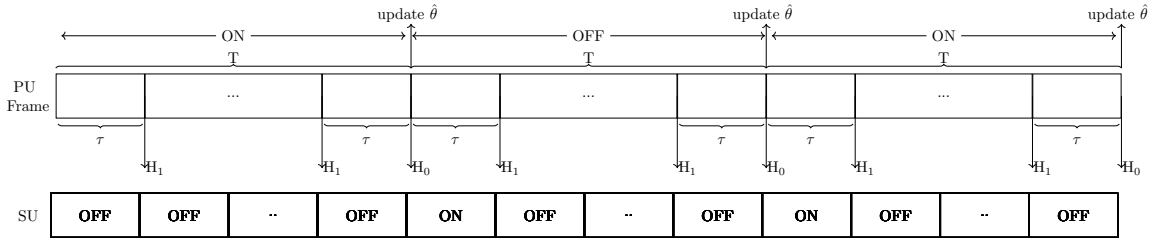


Figure 4.6: instants of updating the estimation of θ.

be estimated by the SU. A robust optimization is needed to account for the uncertainty in the PUs idle probability. Next, we first discuss the formulas for parameter estimation, and then present the robust optimization problem and its respective solution for single channel CRN.

4.1 Estimation of θ

The value of θ is unknown to the SU and it needs to be estimated. Let $\hat{\theta}(n)$ be the estimated idle probability of the PU at frame n. The value of $\hat{\theta}(n)$ can be updated every slot. However, since the PU state remains the same during its entire frame, we can update $\hat{\theta}(n)$ every PU frame duration T. To determine the actual state of the PU, we can take our decision about the channel to update the estimation of θ based on sensing outcomes through the whole PU frame duration T. This provides a more reliable decision compared to that at the slot level as shown in Figure4.6.

Denote by z_n the binary variable indicating whether the PU is sensed to be active in frame n. Using (3.3), z_n can be found by

$$z_n = \begin{cases} 0 & \text{if } M \leq \eta \\ 1 & \text{if } M > \eta \end{cases}$$

Lemma 1. *Let $P_d(T)$ and $P_f(T)$ denote the probabilities of detection and false alarm as the result of sensing outcomes over duration T, respectively. Then the estimator*

$$\hat{\theta}(n) = \frac{P_d(T) - \frac{\sum_{i=1}^{n} z_i}{n}}{P_d(T) - P_f(T)}, \tag{4.25}$$

is an unbiased estimator of θ at frame n.

Proof. $\mathbb{E}\left[\hat{\theta}(n)\right] = \mathbb{E}\left[\frac{P_d(T) - \frac{\sum_{i=1}^{n} z_i}{n}}{P_d(T) - P_f(T)}\right] = \frac{P_d(T) - \frac{\sum_{i=1}^{n} \mathbb{E}[z_i]}{n}}{P_d(T) - P_f(T)} = \frac{P_d(T) - \mathbb{E}[z_i]}{P_d(T) - P_f(T)}$.

Note that $\mathbb{E}[z_i] = \theta P_f(T) + (1-\theta) P_d(T)$. Then $\mathbb{E}\left[\hat{\theta}(n)\right] = \theta$. Hence, $\hat{\theta}(n)$ is an unbiased estimator of θ at frame n. □

4.2 Robust Optimization Problem

As the value of θ is estimated by the SU, a robust optimization is needed to account for the uncertainty in the PU's idle probability. In this case, what is known to the SU is the observations it makes regarding the PU's activity. The SU seeks to account for the uncertainty in its estimate of $\hat{\theta}$ by constraining its estimate of the collision probability with the PU, given its observations. The optimization problem **P1** can be rewritten as

$$\begin{aligned} \mathbf{P2}: \quad & \max_{\tau} \ \lambda(\tau, \hat{\theta}(n)) \\ \text{s.t.} \quad & P(col|z_1, z_2, \ldots, z_n) \leq \bar{P}_c \\ & \tau \in [0, T], \end{aligned} \tag{4.26}$$

where $P(col|z_1, z_2, \ldots, z_n)$ is the posterior estimated collision probability given the past observations of the PU activities at the SU receiver up to frame n. It was shown in in the proof of Lemma 1 that the unbiased estimator $\hat{\theta}(n)$ in (4.25) is a sufficient statistic for the observations. Hence $P(col|z_1, z_2, \ldots, z_n) = P(col|\hat{\theta}(n))$.

4.3 Solving the Optimization Problems

Equations (3.20) and (3.23) in Chapter 3 explicitly state the dependence of the collision probability P_c and the SU throughput λ on the slot length τ and the PU idle probability θ. Thus, to solve **P1** and **P2**, we need to first analyze the properties of λ and P_c with respect to τ and θ.

Lemma 2. *The SU throughput $\lambda(\tau,\theta)$ has the following properties:*

a) For fixed τ, $\lambda(\tau,\theta)$ is a continuous and monotonically increasing function of θ.

b) For fixed θ, $\lambda(\tau,\theta)$ is a concave function of τ for $\tau > \tau_0 = \left(-\frac{b}{2}+\sqrt{-\frac{1}{\gamma}+\frac{b^2}{4}}\right)^2/f_s$, where $b = \gamma\sqrt{2\gamma+1}\mathcal{Q}^{-1}(\bar{P}_d)$.

Proof. a) In (3.23), since $1-P_f(\tau) > 1-P_d(\tau)$ for all values of τ and since $C_0 \gg C_1$, the coefficients of the first and second order terms in right hand side of (3.23) are positive. In other words, with fixed τ, $\lambda(\tau,\theta)$ monotonically increases with $\theta \in (0,1)$.

b) To determine the concavity of the objective function, we take the 2nd order derivative of $\lambda(\tau,\theta)$ with respect to τ,

$$\frac{\partial^2}{\partial\tau^2}\lambda(\tau,\theta) = 2\frac{(1-\theta)\theta}{T}(C_0-C_1)\frac{d}{d\tau}P_f(\tau) - \theta\left[C_0-(C_0-C_1)\frac{(1-\theta)\tau}{T}\right]\frac{d^2}{d\tau^2}P_f(\tau). \tag{4.27}$$

From the definition of $P_f(\tau)$ in (3.9), we have that

$$\frac{d}{d\tau}P_f(\tau) = -\frac{\gamma}{2}\sqrt{\frac{f_s}{2\pi\tau}}e^{-\frac{x^2}{2}}, \tag{4.28}$$

where

$$x = \sqrt{2\gamma + 1}\ \mathcal{Q}^{-1}(\bar{P}_d) + \gamma\sqrt{\tau f_s}. \tag{4.29}$$

We observe that $\frac{d}{d\tau}P_f(\tau)$ is negative and that $C_0 > C_1$. Hence, the first term on the right hand side (RHS) of (4.27) is negative. Given that observation, since $\theta > 0$ and $\tau \leq T$, a sufficient condition for $\frac{\partial^2}{\partial\tau^2}\lambda(\tau,\theta)$ to be negative is that $\frac{d^2}{d\tau^2}P_f(\tau)$ is positive. Evaluating the second derivative of $P_f(\tau)$ we have

$$\frac{d^2}{d\tau^2}P_f(\tau) = \frac{\gamma}{4}\sqrt{\frac{f_s}{2\pi}}e^{-\frac{x^2}{2}}\left(\tau^{-\frac{3}{2}} + \tau^{-1}\gamma x\sqrt{f_s}\right), \tag{4.30}$$

where x was defined in (4.29). Hence, a sufficient condition for $\frac{d^2}{d\tau^2}P_f(\tau)$ to be positive is

$$\tau^{-\frac{1}{2}} + \gamma\sqrt{(2\gamma+1)f_s}\mathcal{Q}^{-1}(\bar{P}_d) + \gamma\sqrt{\tau}f_s > 0. \tag{4.31}$$

Therefore, a sufficient condition for $\lambda(\tau,\theta)$ to be concave in τ is for the sampling rate f_s to be high enough that the number of samples in the SU slot, τf_s, satisfies

$$\tau f_s > \left(-\frac{b}{2} + \sqrt{-\frac{1}{\gamma} + \frac{b^2}{4}}\right)^2, \tag{4.32}$$

where $b = \gamma\sqrt{2\gamma+1}\mathcal{Q}^{-1}(\bar{P}_d)$. When $\bar{P}_d = 0.99$ and $\gamma = 0$dB, the lower bound is less than 5. And as γ increases this lower bound approaches 0. Hence, in most situations of interest this lower bound is close to 0. The value for τ_0 in Lemma 2b) is obtained by rearranging the expression in (4.32).

□

Lemma 3. *The collision probability $P_c(\tau,\theta)$ has the following properties:*

a) For fixed τ, $P_c(\tau,\theta)$ is a continuous and monotonically increasing function of θ.

b) For fixed θ, $P_c(\tau, \theta)$ is a continuous and monotonically increasing function of τ.

Proof. a) From (3.20), since $P_d(\tau) - P_f(\tau) > 0$, it is clear that $P_c(\tau, \theta)$ increases linearly with θ for any fixed value of τ.

b) From (4.28), we see that $\frac{d}{d\tau}P_f(\tau)$ is strictly negative and thus $P_f(\tau)$ decreases with τ. Therefore, by (3.20), $P_c(\tau, \theta)$ monotonically increases with τ since $P_d(\tau) - P_f(\tau) > 0$.

□

When combined, the results in points b) of Lemmas 2 and 3 show that in almost all cases of practical interest problem **P1** is a convex optimization problem, and hence any locally optimal solution is globally optimal. In particular, if the value of τ for which $\partial\lambda(\tau, \theta)/\partial\tau$ is zero lies in $(\tau_0, T]$ then that value of τ is a globally optimal solution.

To develop a strategy for solving the robust optimization problem in **P2** we make the following observation:

Lemma 4. *Under specific conditions, the probability of collision $P(col|oz_1, z_2, \ldots, z_n)$ is upper-bounded by*

$$P_c^{UB}(\tau, \hat{\theta}(n)) = (1 - P_d)(\epsilon_n + 1) + (P_d - P_f(\tau)) \left[\hat{\theta}(n) + \epsilon_n + \sqrt{\frac{-1}{2n}\ln(\epsilon_n)}\right] \frac{\tau}{T}, \quad (4.33)$$

where $\hat{\theta}(n)$ is the SU's estimate of θ after n observations, and ϵ_n is any number such that $0 \leq \epsilon_n \leq 1$. A sufficient condition for this upper-bound to be convex in τ and ϵ_n is $e^{-2n} \leq \epsilon_n \leq \frac{e^{-0.5}}{n}$ and $\tau_0 < \tau \leq T$.

Proof. The estimate $\hat{\theta}(n)$ is an unbiased estimate for θ, i.e. $\mathbb{E}[\hat{\theta}(n)] = \theta$. Hence, using the Chernoff-Hoeffding bound [37], we have

$$P\left(\hat{\theta}(n) \leq \theta - \frac{a}{n}\right) \leq e^{\frac{-2a^2}{n}} \quad \forall a \geq 0, \tag{4.34}$$

or

$$P\left(\theta \geq \hat{\theta}(n) + \frac{a}{n}\right) \leq e^{\frac{-2a^2}{n}} \quad \forall a \geq 0. \tag{4.35}$$

Let $e^{\frac{-2a^2}{n}} = \epsilon_n$, then

$$P\left(\theta \geq \hat{\theta}(n) + \sqrt{-\frac{\ln(\epsilon_n)}{2n}}\right) \leq \epsilon_n. \tag{4.36}$$

Given the observation (estimation) $\hat{\theta}$, the probability of collision is found as

$$P(col|\hat{\theta}) = \int_{\theta} P(col|\theta)P(\theta|\hat{\theta})d\theta \tag{4.37}$$

$$P(col|\hat{\theta}) = \int\limits_{\theta \leq \tilde{\theta}} P(col|\theta)P(\theta|\hat{\theta})d\theta + \int\limits_{\theta > \tilde{\theta}} P(col|\theta)P(\theta|\hat{\theta})d\theta \tag{4.38}$$

As $P(col|\theta)$ and $P(\theta|\hat{\theta})$ are both positive, then given $\tilde{\theta}$ and that $\theta \in [0, 1]$,

$$\int_0^{\tilde{\theta}} P(col|\theta)P(\theta|\hat{\theta})d\theta \leq \max_{0 \leq \theta \leq \tilde{\theta}} \{P(col|\theta)\} \int_0^{\tilde{\theta}} P(\theta|\hat{\theta})d\theta. \tag{4.39}$$

Similarly

$$\int_{\tilde{\theta}}^1 P(col|\theta)P(\theta|\hat{\theta})d\theta \leq \max_{\tilde{\theta} \leq \theta \leq 1} \{P(col|\theta)\} \int_{\tilde{\theta}}^1 P(\theta|\hat{\theta})d\theta. \tag{4.40}$$

From the monotonic nature of $P_c(\tau, \theta)$ with θ (see Lemma 3),

$$\max_{0 \leq \theta \leq \tilde{\theta}} P(col|\theta) = P_c(\tau, \tilde{\theta}) = (1 - P_d) + (P_d - P_f(\tau)) \frac{\tau\tilde{\theta}}{T}, \tag{4.41}$$

and

$$\max_{\tilde{\theta}\leq\theta\leq 1} P(col|\theta) = P_c(\tau,1) = (1-P_d) + (P_d - P_f(\tau))\,\frac{\tau}{T}. \tag{4.42}$$

If we choose $\tilde{\theta} = \hat{\theta}(n) + \sqrt{-\frac{\ln(\epsilon_n)}{2n}}$, then using (4.36) we get

$$\int_{\tilde{\theta}}^{1} P(\theta|\hat{\theta})d\theta = P\left(\theta \geq \hat{\theta}_n + \sqrt{-\frac{\ln(\epsilon_n)}{2n}}\right) \leq \epsilon_n, \tag{4.43}$$

and

$$\int_{0}^{\tilde{\theta}} P(\theta|\hat{\theta})d\theta \leq 1 \tag{4.44}$$

Using (4.37) – (4.44), an upper bound for (4.37) can be expressed as

$$P(col|\hat{\theta}(n)) \leq P_c\left(\tau, \hat{\theta}(n) + \sqrt{-\frac{\ln(\epsilon_n)}{2n}}\right) + \epsilon_n P_c(\tau,1). \tag{4.45}$$

The upper bound on the right hand side can then be expressed as

$$P_c^{UB}(\tau,\epsilon_n;\hat{\theta}(n)) = (1-P_d)(\epsilon_n+1) + (P_d - P_f(\tau))\left[\hat{\theta}_n + \epsilon_n + \sqrt{\frac{-1}{2n}\ln(\epsilon_n)}\right]\frac{\tau}{T}. \tag{4.46}$$

For ease of notation, in some parts of the remaining analysis we will drop the subscript n from ϵ_n. To determine the convexity of $P_c^{UB}(\tau,\epsilon_n;\hat{\theta}(n))$ in τ and ϵ_n, we will evaluate the Hessian matrix of the function $f(\tau,\epsilon) = P_c^{UB}(\tau,\epsilon;\hat{\theta})$, which we denote by $\left[\begin{smallmatrix} f_{\tau\tau} & f_{\tau\epsilon} \\ f_{\epsilon\tau} & f_{\epsilon\epsilon} \end{smallmatrix}\right]$, and determine whether it is positive semidefinite.

From (4.46) we have that

$$f_{\epsilon\epsilon} = -\frac{\tau(-\epsilon + 2(\epsilon+1)\log(\epsilon) + 1)(P_d - P_f(\tau))}{4n^2 T\epsilon^2 \left(-\frac{\log(\epsilon)}{2n}\right)^{3/2}}, \tag{4.47}$$

$$f_{\tau\tau} = \frac{\left(\epsilon\left(\sqrt{-\frac{\log(\epsilon)}{2n}} - 1\right) - \sqrt{-\frac{\log(\epsilon)}{2n}} - \hat{\theta}(\epsilon+1)\right)\left(\tau P_f''(\tau) + 2P_f'(\tau)\right)}{T}, \tag{4.48}$$

$$f_{\epsilon\tau} = f_{\tau\epsilon} = \frac{\left(2\epsilon\left(2n(\hat{\theta}+1)\sqrt{-\frac{\log(\epsilon)}{2n}} + \log(\epsilon)\right) + \epsilon - 1\right)\left(P_d - \tau P_f'(\tau) - P_f(\tau)\right)}{4nT\epsilon\sqrt{-\frac{\log(\epsilon)}{2n}}} \tag{4.49}$$

A condition for the Hessian matrix of f to be positive semidefinite is that $f_{\tau\tau} \geq 0$, $f_{\epsilon\epsilon} \geq 0$ and $f_{\tau\tau}f_{\epsilon\epsilon} - f_{\tau\epsilon}f_{\epsilon\tau} \geq 0$. A sufficient condition for this is when $e^{-2n} \leq \epsilon \leq \frac{e^{-0.5}}{n}$. Hence, when $e^{-2n} \leq \epsilon_n \leq \frac{e^{-0.5}}{n}$, the upper bound in (4.33) is convex.

□

Using insights from Lemma 4 we can safely approximate the robust optimization problem in (4.26) by the following convex optimization problem

$$\begin{aligned} \mathbf{P3}: \quad & \max_{\tau,\epsilon_n} \ \lambda(\tau, \hat{\theta}(n)) \\ \text{s.t.} \quad & P_c^{UB}(\tau, \epsilon_n; \hat{\theta}(n)) \leq \bar{P}_c \\ & \tau \in (\tau_0, T] \\ & e^{-2n} \leq \epsilon_n \leq \frac{e^{-0.5}}{n} \end{aligned} \tag{4.50}$$

Here safety means that every feasible solution to **P2** is feasible for **P1** [51]. As $n \to \infty$, $\hat{\theta}(n) \to \theta$. Therefore, if $\epsilon_n \to 0$ and $-\frac{1}{2n}\ln(\epsilon_n) \to 0$, then $P_c^{UB}(\tau, \epsilon_n; \hat{\theta}(n)) \to P_c(\tau, \theta)$. That means that as $n \to \infty$ the formulation in **P3** approaches the original formulation

in **P1**. In the robust optimization literature, **P1** is called the nominal problem while **P2** is called the robust optimization problem.

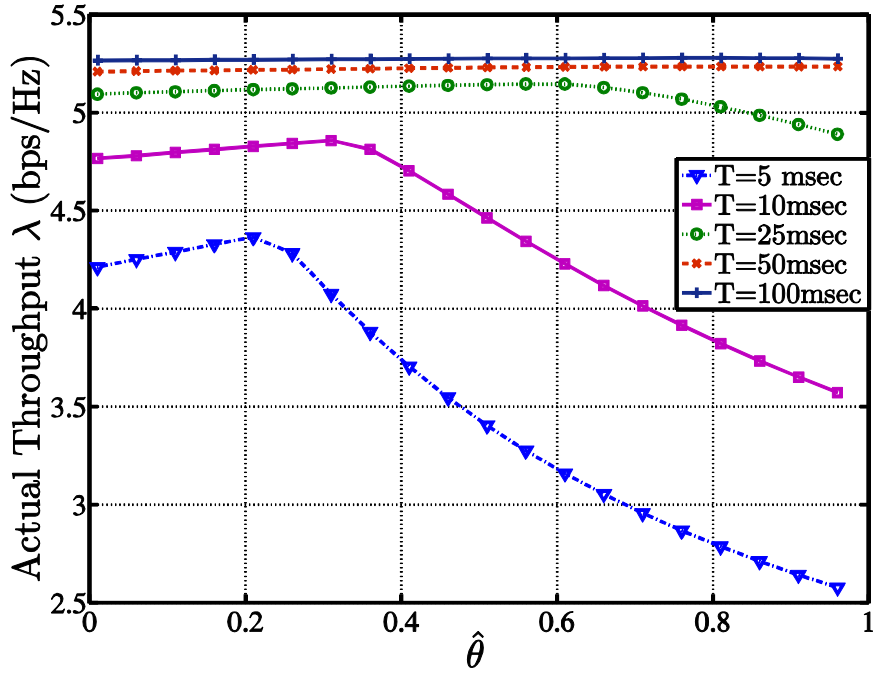
4.4 Performance Evaluation

Figures 4.7a and 4.7b together illustrate the motivation for finding a robust solution when $\hat{\theta}$ differs from θ. The results in both figures are the outcome of solving the nominal problem (**P1**) substituting $\hat{\theta}$ for θ. We show in Figures 4.7a and 4.7b the achievable throughput and the actual collision probability, respectively. In both figures, $\theta = 0.8$. We notice that for the smaller values of T, designs obtained when θ is under estimated appear to provide greater throughput , i.e., $\hat{\theta} < \theta$, $\lambda > \lambda^*$ where λ^* is the maximum throughput achieved using **P1** when $\hat{\theta} = \theta$. However, this is due to the fact that the collision probability constraint is violated, i.e. $P_c > \bar{P}_c$. For example, when $T = 5$msec, if $\hat{\theta} = 0.7$, the throughput (Figure 4.7a) is approximately 2.9 bps/Hz, which is greater than the optimal throughput for $\hat{\theta} = \theta = 0.8$, which is 2.6 bps/Hz. This is at the expense of higher collision probability (Figure 4.7b) at around 0.115, which violates the collision probability constrain of 0.1.

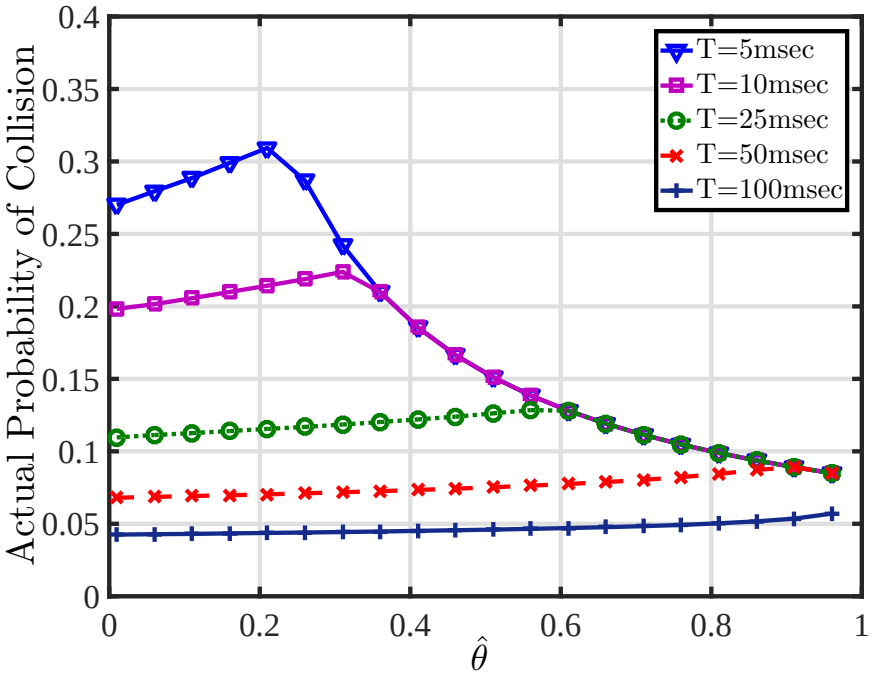
On the other hand, we observe that for larger T (e.g., $T = 50$msec and $T = 100$msec, the $\bar{P}_c$ constraint is always satisfied (Figure 4.7b) and the throughput is less sensitive to the estimation errors in $\hat{\theta}$ (Figure 4.7a). Thus, we conclude that it is desirable to operate the SaST protocol under large T's.

Figure 4.8b combines learning with optimization. It shows how using robust optimization prevents the actual probability of collision P_c from exceeding the limit of $\bar{P}_c$. However, if we use the nominal optimization problem, the actual probability of collision exceeds $\bar{P}_c$. This robustness comes at the expense of less SU throughput as

seen in Figure 4.8a. We also notice that as time increases the results obtained from solving **P3** approaches the results obtained from solving **P1**. This agrees with the discussion we made earlier in this chapter.

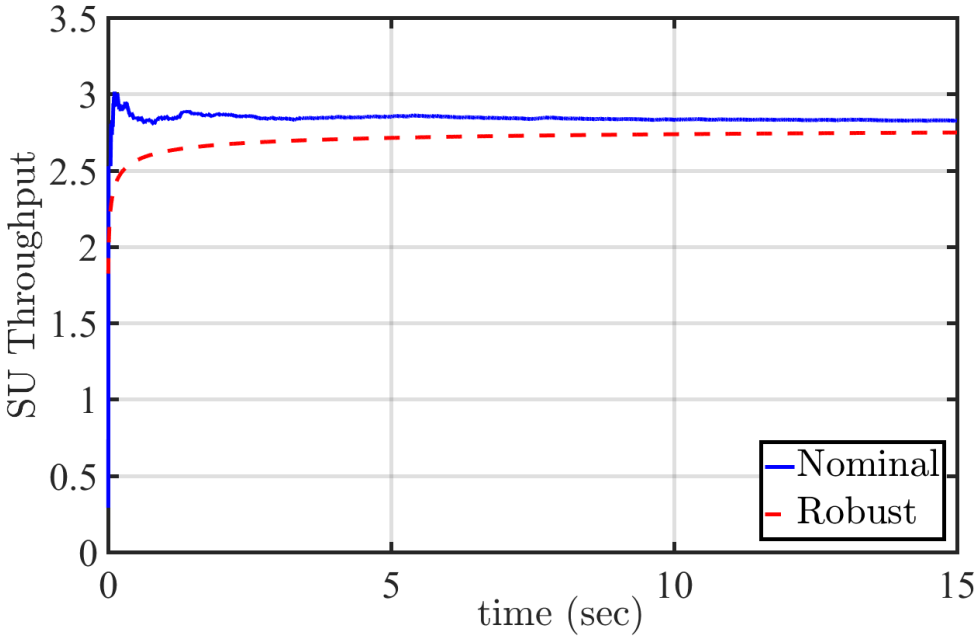


(a) The actual SU throughput

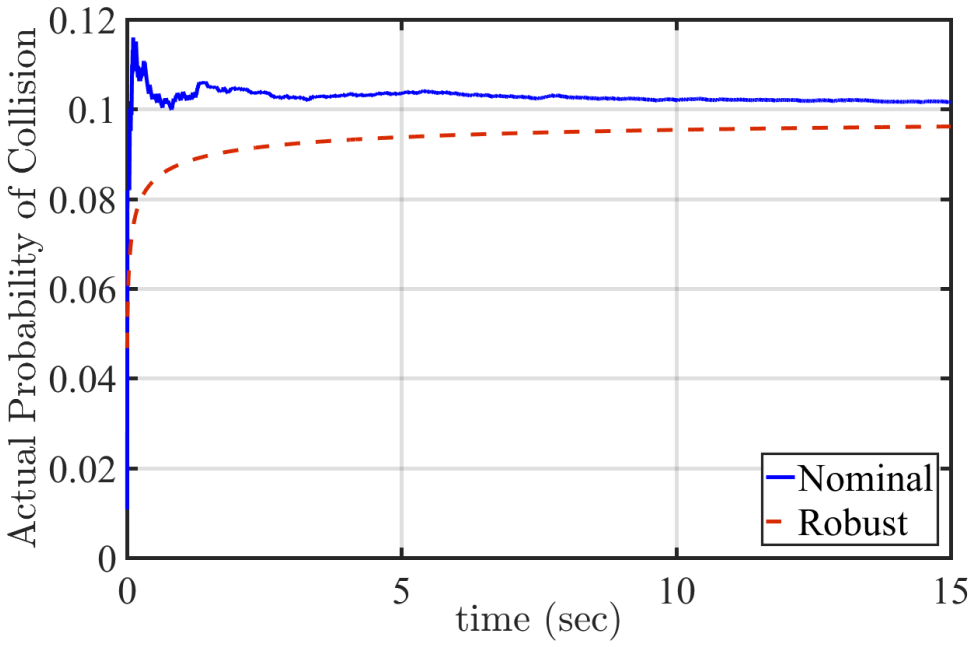


(b) The actual collision probability

Figure 4.7: The actual SU throughput and the actual collision probability vs the estimation of the PU idle probability ($\hat{\theta}$) for $\bar{P}_c = 0.1$, $\bar{P}_d = 0.99$ and different values of PU frame duration (T) in single channel FD spectrum sensing



(a) The actual SU throughput



(b) The actual collision probability

Figure 4.8: The actual probability of collision and the actual SU throughput vs time in case of solving nominal problem (**P1**) and robust problem (**P3**), $T = 5$msec, $\theta = 0.9$, $\bar{P}_d = 0.99$ and $\bar{P}_c = 0.1$

Chapter 5

Multi-channel SaST Protocol

In the previous two chapters, we investigated the problem of the SU trying to access the channel of one PU with unknown idle probability θ. However, in practice, multiple PU channels are available for access with idle probabilities are not known a priori. Therefore, an optimal SU strategy needs to learn the statistics on individual PU channel(s), decide which channel to operate on and the slot length to use. In this process, the SU faces the trade-off between exploiting the best channel known so far and exploring less known channels.

5.1 Estimation of PUs' Idle Probabilities

Before presenting the learning strategy, we first extend the notations from Chapter 4 to the multi-channel case. Let $\hat{\theta}_k(n), k \in \mathcal{K}$ denote the estimated idle probability of PU channel k, $L_k(n)$ denote the number of times the PU channel k has been accessed up until frame n and $k_n \leq n$ denote the channel accessed at frame n. Similar to 4.25

the unbiased estimator of the idle probability of PU channel k, θ_k is given by

$$\hat{\theta}_k(n) = \frac{P_d(T) - \frac{\sum_{i=1}^{n} z_i \cdot \mathbb{I}_{\{k_i=k\}}}{L_k(n)}}{P_d(T) - P_f(T)}, \tag{5.1}$$

where $\mathbb{I}_{\{.\}}$ is the Kronecker delta function.

5.2 Sequential Learning Algorithm

In multi-channel SaST, learning and channel switching occur at the time scale of PU frames, while spectrum sensing and access are performed on a slot by slot basis. The SU first selects each PU channel once. Then, at the beginning of each PU frame, the SU decides which PU channel to sense and access using the following modified ε-greedy policy [37]. Every T seconds, the SU will do the following:

1. With probability $\varepsilon_n = \min\left\{1, \frac{c}{n}\right\}, n > K$, the SU picks a random channel uniformly from $\mathcal{K} = \{1, 2, \ldots, K\}$, where c is a constant.

2. With probability $1 - \varepsilon_n$, the SU picks the channel k_n based on some criteria we will discuss later in this chapter.

In both cases, the SU updates the estimate $\hat{\theta}_{k_n}$ using (5.1) based on the observations in the $(n-1)^{th}$ PU frame. The SU uses the updated estimate $\hat{\theta}_{k_n}(L_{k_n}(n))$ and solves **P3** to find τ_{k_n} to use this value in sensing and accessing the PU channel for the rest of the nth PU frame. The SU can pick a PU channel based on one of two criteria:

- Maximum λ (denoted as λ_{max} scheme): In this case, the SU solves **P3** for all K channels using the current estimate $\hat{\theta}_k(n)$ and picks the channel $k_n = \arg\max_{k \in \mathcal{K}} \lambda\left(\tau_k, \hat{\theta}_k(n)\right)$ as the channel to sense.

- Maximum $\hat{\theta}$ (denoted as θ_{max} scheme): In this case, the SU picks the channel $k_n = \arg\max_{k \in \mathcal{K}} \hat{\theta}_k(n)$ as the channel to sense. This choice arises from insight into Lemma.2, which shows that if τ is fixed the SU throughput increases with θ. We note however, that this selection scheme overlooks the collision constraint and the uncertainty in $\hat{\theta}$.

In the λ_{max} selection scheme the SU has to solve **P3** K times whenever it picks a channel. On the other hand, the θ_{max} scheme requires the SU to solve **P3** only once. This shows that the θ_{max} is much more computationally efficient than the λ_{max}. In the rest of this chapter, we assume the SU is using the θ_{max} scheme whenever it is supposed to pick a channel on a non-random sense.

5.3 Regret and Performance Analysis

In the sequential learning literature, regret is typically used to characterize the performance of an online policy Φ with respect to an optimal offline solution. In our case the regret will be the difference between the throughput when selecting the channel using SaST and the throughput obtained if the best channel was selected. Lemma 2 proves that the best channel is the one that has the highest idle probability θ. In particular, in multi-channel CRNs, we can define the expected cumulative regret using the online policy Φ at frame n as,

$$R_n^{\Phi} = \mathbb{E} \sum_{t=1}^{n} \{(\lambda^* - \lambda_{k_t}(t))\, T\}, \tag{5.2}$$

where $\lambda^* = \max_{k \in K} \lambda_k$ and $\lambda_{k_t}(t)$ are the average throughput of the best PU channel and channel k_t at frame t, respectively. There are two sources of regret. First, when

the SU is accessing an inferior channel, it has lower throughput since the respective PU is more active and the transmission opportunity is less. This first type of regret is denoted by $R^{\Phi}_{\text{inf}}(n)$. Second, even when the SU is accessing the best channel, it incurs lower throughput than the optimal off-line scheme due to the use of the "safe" approximation of the collision constraint that led to the convex formulation of the robust slot length optimization problem in which leads to using a suboptimal τ by the SU. We denote this second type of regret by $R^{\Phi}_{\text{opt}}(n)$.

As the SU uses the ε-greedy policy described earlier, the percentage of time that an inferior channel is selected decreases over time. Furthermore, if j^* denotes the best channel and as the best channel is selected more often, $\hat{\theta}_{j^*}$ would approach θ_{j^*} and the instantaneous regret would approach zero. Therefore, we expect that the total expected regret would only grow sublinearly over time. This intuition is formalized in the following theorem.

Theorem 1. *The expected cumulative regret can be attributed to two sources:*

- *Regret incurred when selecting an inferior channel (R^{Φ}_{inf}), where $R^{\Phi}_{inf} = O(\log n)$.*
- *Regret incurred when selecting the best channel (R^{Φ}_{opt}), where $R^{\Phi}_{opt} = o(n)$.*

And hence the cumulative regret is $R^{\Phi}_{n} = o(n)$.

Proof. As discussed earlier, the expected regret can be attributed to two sources: the regret incurred when an inferior channel is accessed, and the regret incurred when the optimal throughput is not attained due to uncertainty in the PU activities in the best channel. In other words,

$$R^{\Phi}_{n} = R^{\Phi}_{inf}(n) + R^{\Phi}_{opt}(n). \tag{5.3}$$

Next, we will show that both regrets grow sublinearly over time under some weak conditions.

Regret incurred when selecting the best channel When selecting the best channel j^*, the SU may not be able to transmit at the optimal throughput due to the estimation error in $\hat{\theta}_{j^*}$ and relaxation in the robust optimization problem. However, intuitively, as the number of times selecting the best PU channel increases, **P3** approaches **P1** and $\hat{\theta}_{j^*}(n)$ approaches θ_{j^*}. This results an instantaneous regret that decreases in time. Next, we prove this intuition rigorously. Note that to show that the cumulative regret $R^{\Phi}_{opt}(n)$ is $o(n)$, it suffices to show that the instantaneous regret in the best channel converges to zero. To do so, we recall from (5.1) that

$$\hat{\theta}_j(n) = \frac{P_d(T) - \frac{\sum_{i=1}^{n} z_i \cdot \mathbb{I}_{\{k_i=j\}}}{L_j(n)}}{P_d(T) - P_f(T)},$$

and that as $n \to \infty$, $L_j(n) \to \infty$ and $\frac{\sum_{i=1}^{n} z_i \cdot \mathbb{I}_{\{k_i=j\}}}{L_j(n)} \to \mathbb{E}\left[z_i | k_i = j\right]$. Hence,

$$\lim_{n\to\infty} \hat{\theta}_j(n) = \frac{P_d(T) - \mathbb{E}\left[z_i | k_i = j\right]}{P_d(T) - P_f(T)}. \tag{5.4}$$

Therefore,

$$\lim_{n\to\infty} \hat{\theta}_j(n) = \frac{P_d(T) - (\theta_j P_f(T) + (1-\theta_j) P_d(T))}{P_d(T) - P_f(T)}, \tag{5.5}$$

and finally

$$\lim_{n\to\infty} \hat{\theta}_j(n) = \theta_j.$$

Regret incurred when selecting an inferior channel Let j^* be the index of the best channel. The probability that at time t channel $j \neq j^*$ will be chosen is,

$$\mathbb{P}(k_t = j) \leq \frac{\varepsilon_t}{K} + \left(1 - \frac{\varepsilon_t}{K}\right) \mathbb{P}\left(\hat{\theta}_j(t-1) \geq \hat{\theta}_{j^*}(t-1)\right) \tag{5.6}$$

We have

$$\mathbb{P}\left(\hat{\theta}_j(t-1) \geq \hat{\theta}_{j^*}(t-1)\right) \leq \mathbb{P}\left(\hat{\theta}_j(t-1) \geq \theta_j + \frac{\delta_j}{2}\right) + \mathbb{P}\left(\hat{\theta}_{j^*}(t-1) \leq \theta_{j^*} - \frac{\delta_j}{2}\right) \tag{5.7}$$

where $\delta_j = \theta_{j^*} - \theta_j$. We bound the first term as follows,

$$\mathbb{P}\left(\hat{\theta}_j(t-1) \geq \theta_j + \frac{\delta_j}{2}\right) = \mathbb{P}\left(\hat{\theta}_j(t-1) - \theta_j \geq \frac{\delta_j}{2}\right).$$

From (5.5) this can be written as

$$\mathbb{P}\left(\hat{\theta}_j(t-1) \geq \theta_j + \frac{\delta_j}{2}\right) = \mathbb{P}\left(\frac{1}{L_j(t)} \sum_{i=1}^{t} z_i \cdot \mathbb{I}_{\{k_i=j\}} - \mathbb{E}\left[z_i\right] \leq -\left(P_d(T) - P_f(T)\right) \frac{\delta_j}{2}\right) \tag{5.8}$$

Define $\breve{\delta}_j = \left(P_d(T) - P_f(T)\right) \delta_j$. From Chernoff-Hoeffding bound, we have

$$\mathbb{P}\left(\frac{1}{L_j(t)} \sum_{i=1}^{t} z_i \cdot \mathbb{I}_{\{k_i=j\}} - \mathbb{E}\left[z_i | k_i = j\right] \leq -\frac{\breve{\delta}_j}{2}\right) \leq e^{-\breve{\delta}_j^2 L_j(t)/2} \tag{5.9}$$

Denote $x_t = \frac{1}{2K} \sum_{i=1}^{t} \varepsilon_i$. Following the steps of the proof in [37] with minor modifications, we have

$$\mathbb{P}\left(\hat{\theta}_j(t-1) \geq \theta_j + \frac{\delta_j}{2}\right) \leq x_t e^{-\frac{x_t}{5}} + \frac{2}{\breve{\delta}_j^2} e^{-\breve{\delta}_j^2 \lfloor x_t \rfloor/2} \tag{5.10}$$

The same bound holds for $\mathbb{P}\left(\hat{\theta}_{j^*}(t-1) \leq \theta_{j^*} - \frac{\delta_j}{2}\right)$. For $t \geq t' = c$ we have

$$x_t \geq \frac{c}{K} \ln \frac{te^{1/2}}{c}$$

Using (5.6) – (5.10) and the lower bound on x_t, we get

$$\mathbb{P}(k_t = j) \leq \frac{c}{Kt} + \frac{2c}{K} \ln \left(\frac{te^{1/2}}{c}\right) \left(\frac{te^{1/2}}{c}\right)^{\frac{-c}{5K}} + \frac{4}{\breve{\delta}_j^2} \left(\frac{te^{1/2}}{c}\right)^{\frac{-\breve{\delta}_j^2 c}{2K}}, \tag{5.11}$$

for $j \neq j^*$.

With properly chosen c, the last two terms decrease faster than $1/t$. Thus,

$$\begin{aligned} R_{inf}^{\Phi}(n) &= \sum_{t=1}^{n} (\lambda^* - \lambda_{k_i}(t))T & (5.12) \\ &\leq \sum_{t=1}^{n} \sum_{j \neq j^*} P(k_t = j)\lambda^* T & (5.13) \\ &\leq 2c \log(n) \lambda^* T, & (5.14) \end{aligned}$$

where $\lambda^* = \max_{k \in \mathcal{K}} \lambda_k$. As a result, the cumulative regret $R_n^{\Phi} = o(n)$. □

A couple of comments are in order. First, since the expected cumulative regret is sublinear over time, the instantaneous regret actually converges to zero. Thus, the proposed policy is a non-regret policy. Second, in determining the regrets, we are comparing against an optimal policy that stays in the most available channel the whole time. a PU frame is out of the scope of the thesis.

5.4 Performance Evaluation

To evaluate the performance of the proposed multichannel SaST protocol, we have simulated a CRN that has $K = 10$ channels. The idle probabilities of the PUs are [0.2 0.8 0.7 0.1 0.6 0.4 0.25 0.9 0.45 0.65]. Therefore, the optimal channel is $j^* = 8$, which has the highest idle probability of $\theta_8 = 0.9$. Figure 5.1 shows $\hat{\theta}$ for 3 channels out of the 10 PU channels using ε-greedy algorithm and robust optimization. We see that $\hat{\theta}_8$ converges faster than other channels to $\theta_8 = 0.9$. We notice also that $\hat{\theta}$ corresponding to the channel with smallest θ converges much slower and after 15 seconds, the estimation of $\hat{\theta}_4$ is still far from the true value which is $\theta_4 = 0.1$. This is due to the fact that using ε-greedy algorithm, the best channel is accessed more often through time.

Using ε-greedy algorithm and robust optimization, Figure 5.2 shows the total cumulative regret with its two components: the regret from selecting an inferior channel $R^{\Phi}_{\text{inf}}(n)$, and the regret from inaccurate estimation in the optimal channel $R^{\Phi}_{\text{opt}}(n)$ for $T = 5$msec. We can see how the cumulative regret grows sublinearly, which illustrates the convergence of the multi-channel SaST protocols. We observe that $R^{\Phi}_{\text{inf}}(n)$ dominates $R^{\Phi}_{\text{opt}}(n)$. Figure 5.3 compares the cumulative regret when the SU chooses the PU channel to be sensed based on λ_{max} criterion and the regret when the SU chooses the channel based on θ_{max} in both robust and nominal formulations. As expected, the instantaneous regret of the the nominal and robust problem will be equal as $\hat{\theta}$ converges to θ. In the nominal formulation, θ_{max} and λ_{max} are the same. Figure 5.4 shows the actual probability of collision for the multi-channel case. It is clear in Figure 5.4a that using the solving the robust problem, the actual collision probability never exceeds $\bar{P}_c = 0.1$. However, in Figure 5.4b, the collision probability

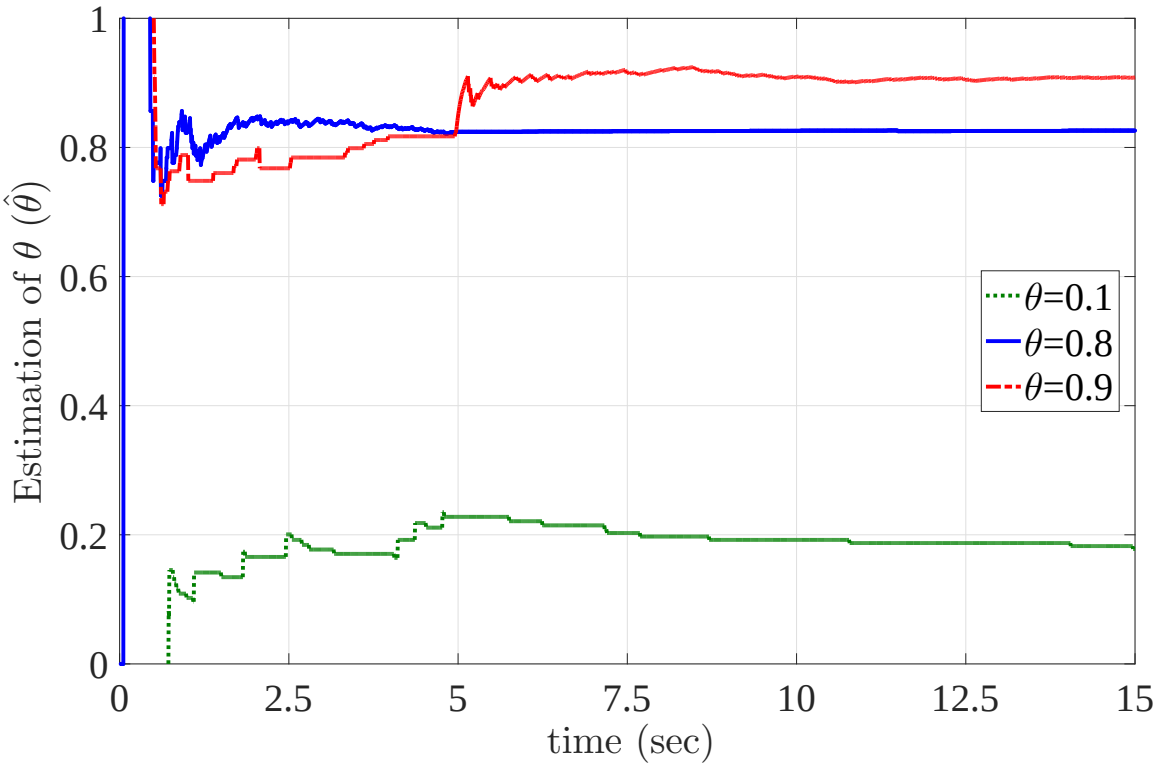


Figure 5.1: Convergence of the estimated idle probability ($\hat{\theta}$) on 3 out of 10 available channels using ε-greedy algorithm and robust optimization, $T = 5$msec

constraint is violated. It is also clear that the cumulative regret from solving the nominal problem is less than the ones from solving the robust problem. This agrees with the results in Figure 4.8a and which comes at the expense of violating the PU's constraint as illustrated in Figure 5.4 which is something the PU can not tolerate. Hence, robust optimization is a must in order to keep the PU protected.

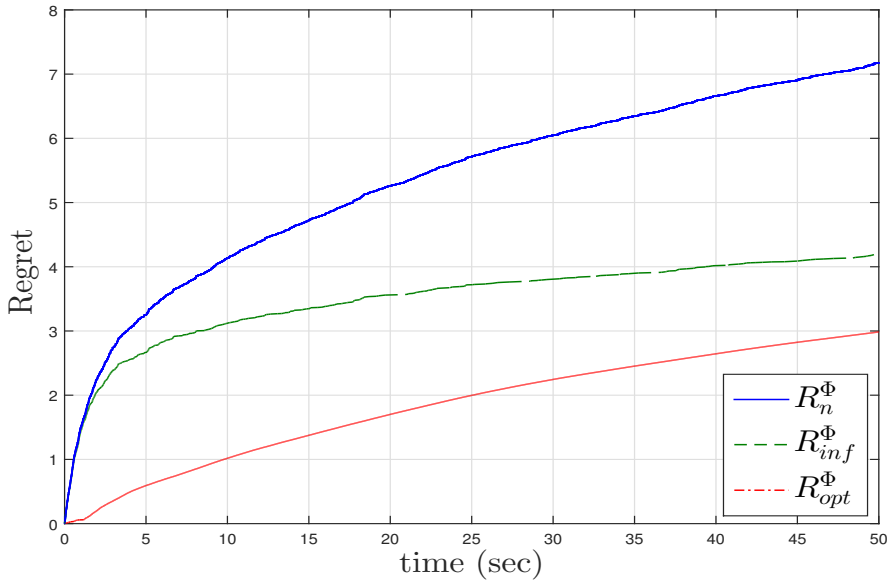


Figure 5.2: Regrets vs time using ε-greedy algorithm and robust optimization for $K = 10$, $T = 5$msec $c = 30$ and, $\bar{P}_c = 0.1$

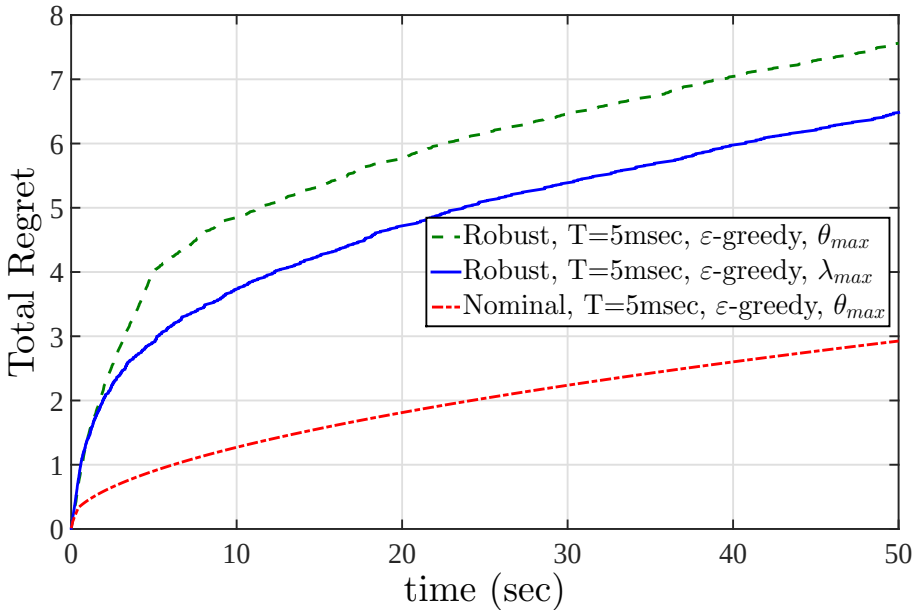
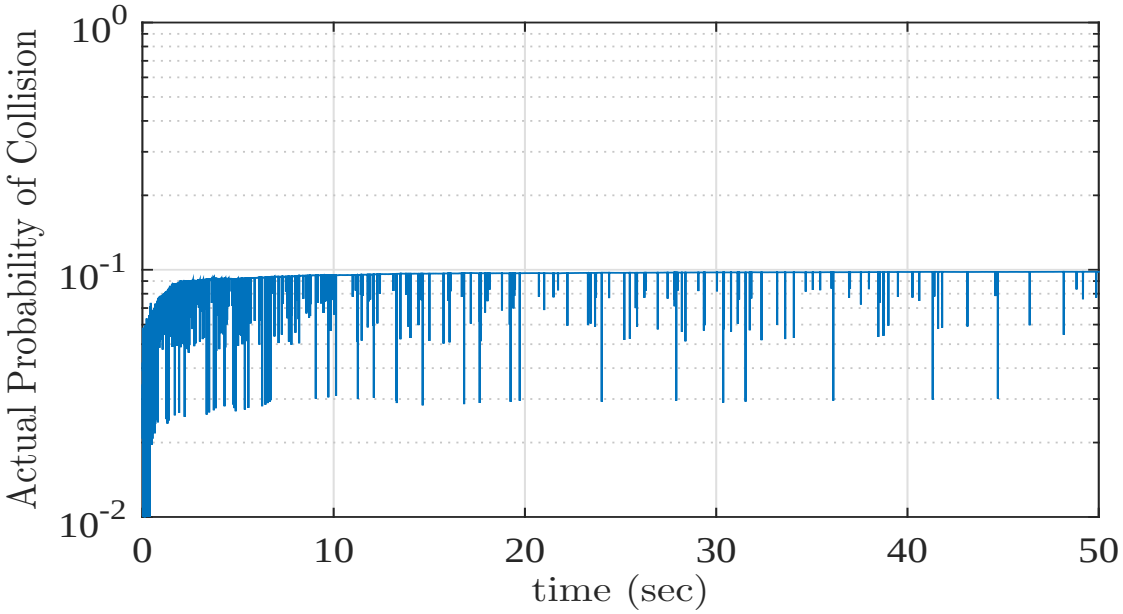
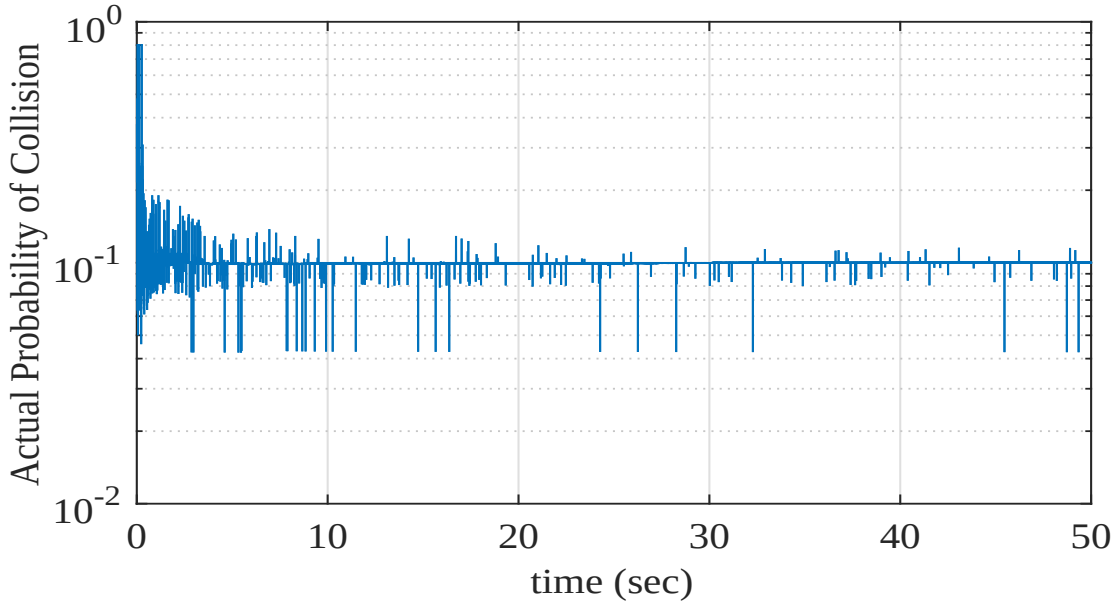


Figure 5.3: Regret vs time for maximum λ and maximum θ channel selection criteria in case of solving nominal problem (**P1**) and robust problem (**P3**)



(a) Solving the robust problem (**P3**)



(b) Solving the nominal problem (**P1**)

Figure 5.4: The actual probability of collision vs time for multichannel case using ε-greedy algorithm in case of solving nominal problem (**P1**) and robust problem (**P3**), $T = 5$msec, $\bar{P}_d = 0.99$ and $\bar{P}_c = 0.1$

Chapter 6

Conclusion and Future Work

FD wireless communication has is an enabling technology for CR. It can help an SU react as fast as possible to changes in PUs' activities, hence limiting the interference with the PUs to its minimum. Our research focused on the protecting the PUs during the stage at which the SUs are learning the wireless environment.

In Chapter 2, we introduced both CR and FD. We showed discussed the work done by other researchers on the track of applying FD in CR.

In Chapter 3, we derived our system model and explained the FD SU strategy. We introduced the different states which the SU may be in. We showed the how the SU's slot duration can affect the PU. We derived the state machine for both the SU and the combined PU-SU system. We derived the state transition probabilities of the Markov chain. We derived the formula for the SU's throughput and collision probability as functions of the SU slot duration and the PU idle probability. We showed through simulation, that our analytical model is a good approximation.

In Chapter 4, we introduced the trade-off between maximizing the SU's throughput and keeping the interference with the PU to an acceptable level. We formulated

both the nominal and robust optimization problems. We proved that the optimization problem is a convex problem. Our simulation results demonstrate that our proposed scheme protects the PU from excessive collision when the SU's estimation of the PU's statistics is not accurate.

In Chapter 5, we generalized the optimization problems to the case when there are multiple PU channels available. We used the ε-greedy algorithm for learning and exploiting the learned PUs activities. We proved that proposed scheme incurs sublinear cumulative regret over time and we illustrated that result through simulation.

In this thesis, the effect of the self-interference was not considered. As future work, the problem can be reformulated taking into consideration the residual self-interference. This will change the probability of detection and the probability of false alarm expressions. In consequence, the expressions of the SU's throughput and collision probability will also change. We assumed that the SU knows the limits of the PU's frame to update the PU's activity statistics at the end of each PU's frame. As future work, practical protocols that relax the synchronization requirement between PUs and SUs can be designed and implemented. That means to find an unbiased estimator for the PU's idle probability that can be updated every SU slot duration.

In this thesis, we assumed an SU uses FD to sense and transmit to another SU simultaneously over the same frequency channel. However, the communication between two SUs uses HD communication. The work in this thesis can be extended to investigate protocols that can use FD in the communication between two SUs as well. A comparison between the maximum achievable SU throughput using the proposed protocol in this thesis and the optimal SU throughput that can be achieved using HD techniques can be done as well. In addition, the proposed strategy can be extended

to take into account channel switching delay.

Bibliography

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, 2005.

[2] P. Kolodzy, "Spectrum policy task force," *Federal Commun. Comm., Washington, DC, Rep. ET Docket*, no. 02-135, 2002.

[3] M. Hammouda, M. El-Khamy, and M. El-Sharkawy, "Frequency Multiplexed Spectrum Sensing for Enhanced Primary User Protection," in *Proc. 8th Int. Conf. on Wireless Commun., Networking and Mobile Computing (WiCOM)*, Shanghai, Sept 2012, pp. 1–4.

[4] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th ACM Annu. Int. Conf. on Mobile Computing and Networking (MobiCom)*, Las Vegas, NV, 2011, pp. 301–312.

[5] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?" in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, Toronto, ON, April 2014, pp. 253–261.

[6] A. Anandkumar, N. Michael, and A. Tang, "Opportunistic spectrum access with multiple users: Learning under competition," in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, San Diego, CA, March 2010, pp. 1–9.

[7] L. Lai, H. Jiang, and H. Poor, "Medium access in cognitive radio networks: A competitive multi-armed bandit framework," in *Proc. 42nd Asilomar Conf. on Signals, Syst. and Comput.*, Pacific Grove, CA, Oct 2008, pp. 98–102.

[8] K. Liu and Q. Zhao, "Decentralized multi-armed bandit with multiple distributed players," in *Proc. Inform. Theory and Applicat. Workshop (ITA)*, San Diego, CA, Jan 2010, pp. 1–10.

[9] E. G. Larsson and M. Skoglund, "Cognitive radio in a frequency-planned environment: some basic limits," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4800–4806, 2008.

[10] N. Hoven and A. Sahai, "Power scaling for cognitive radio," in *Proc. Int. Conf. on Wireless Networks, Commun. and Mobile Computing*, vol. 1. IEEE, 2005, pp. 250–255.

[11] D. Pu, Y. Shi, A. V. Ilyashenko, and A. M. Wyglinski, "Detecting primary user emulation attack in cognitive radio networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*. IEEE, 2011, pp. 1–5.

[12] M. Matinmikko, M. Palola, H. Saarnisaari, M. Heikkilä, J. Prokkola, T. Kippola, T. Hänninen, M. Jokinen, and S. Yrjölä, "Cognitive radio trial environment: First live authorized shared access-based spectrum-sharing demonstration," *IEEE Veh. Tech. Mag.*, vol. 8, no. 3, pp. 30–37, 2013.

[13] Z. Ye, J. Grosspietsch, and G. Memik, "Spectrum sensing using cyclostationary spectrum density for cognitive radios," in *IEEE Workshop on Signal Process. Syst.* IEEE, 2007, pp. 1–6.

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

[15] W. Afifi, A. Sultan, and M. Nafie, "Adaptive sensing and transmission durations for cognitive radios," in *Proc. IEEE Int. Symp. on Dynamic Spectrum Access Networks (DySPAN)*, May 2011, pp. 380–388.

[16] Y. Pei, A. T. Hoang, and Y.-C. Liang, "Sensing-throughput tradeoff in cognitive radio networks: How frequently should spectrum sensing be carried out?" in *Proc. 18th IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Sept 2007, pp. 1–5.

[17] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," *ACM SIGCOMM Comput. Commun. Review*, vol. 43, no. 4, pp. 375–386, 2013.

[18] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4494–4510, Nov 2013.

[19] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, December 2012.

[20] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. 44th Asilomar Conf. on Signals, Syst. and Comput.*, Nov 2010, pp. 1558–1562.

[21] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. 16th Annu. Int. Conf. on Mobile Computing and Networking (MobiCom)*. ACM, 2010, pp. 1–12.

[22] A. K. Khandani, "Methods for spatial multiplexing of wireless two-way channels," Oct. 19 2010, uS Patent 7,817,641.

[23] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full-duplex wireless communications," in *Proc. of the 10th ACM Workshop on Hot Topics in Networks*. ACM, 2011, p. 17.

[24] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "Midu: Enabling mimo full duplex," in *Proc. 18th Annu. Int. Conf. on Mobile Computing and Networking*, ser. Mobicom '12. New York, NY, USA: ACM, 2012, pp. 257–268. [Online]. Available: http://doi.acm.org/10.1145/2348543.2348576

[25] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Proc. 45th Asilomar Conf. on Signals, Syst. and Comput. (ASILOMAR)*. IEEE, 2011, pp. 2002–2006.

[26] B. Kaufman, J. Lilleberg, and B. Aazhang, "An analog baseband approach for designing full-duplex radios," in *Proc. Asilomar Conf. on Signals, Syst. and Comput.* IEEE, 2013, pp. 987–991.

[27] ——, "Analog baseband cancellation for full-duplex: An experiment driven analysis," *arXiv preprint arXiv:1312.0522*, 2013.

[28] G. Liu, F. R. Yu, H. Ji, V. C. M. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 500–524, Secondquarter 2015.

[29] Y. Liao, T. Wang, L. Song, and Z. Han, "Listen-and-talk: Full-duplex cognitive radio networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, TX, Dec 2014, pp. 3068–3073.

[30] Y. Liao, L. Song, and Z. Han, "Full-duplex cognitive radio networks," in *Listen and Talk: Full-duplex Cognitive Radio Networks*. Springer International Publishing, 2016, pp. 19–50.

[31] W. Cheng, X. Zhang, and H. Zhang, "Full duplex spectrum sensing in non-time-slotted cognitive radio networks," in *Proc. Military Commun. Conf. (MILCOM)*, Baltimore, MD, Nov 2011, pp. 1029–1034.

[32] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, Turin, April 2013, pp. 1258–1266.

[33] ——, "Adaptive transmission-reception-sensing strategy for cognitive radios with full-duplex capabilities," in *Proc. IEEE Int. Symp. on Dynamic Spectrum Access Networks (DYSPAN)*, McLean, VA, April 2014, pp. 149–160.

[34] W. Cheng, X. Zhang, and H. Zhang, "Full-duplex spectrum-sensing and MAC-protocol for multichannel nontime-slotted cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 5, pp. 820–831, May 2015.

[35] M. N. Katehakis and A. F. Veinott, "The multi-armed bandit problem: Decomposition and computation," *Math. of Operations Research*, vol. 12, no. 2, pp. 262–268, 1987.

[36] R. Zheng and C. Hua, *Sequential Learning and Decision-Making in Wireless Resource Management*. Springer International Publishing, 2016.

[37] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2-3, pp. 235–256, May 2002. [Online]. Available: http://dx.doi.org/10.1023/A:1013689704352

[38] D. Kalathil, N. Nayyar, and R. Jain, "Decentralized learning for multiplayer multiarmed bandits," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2331–2345, Apr. 2014.

[39] T. H. Cormen, *Introduction to algorithms*. MIT press, 2009.

[40] M. Hammouda and J. Wallace, "Noise uncertainty in cognitive radio sensing: Analytical modeling and detection performance," in *Int. ITG Workshop on Smart Antennas (WSA)*, Dresden, March 2012, pp. 287–293.

[41] S. Gong, P. Wang, W. Liu, and W. Zhuang, "Performance bounds of energy detection with signal uncertainty in cognitive radio networks," in *Proc. IEEE Int. Conf. on Comput. Commun. (INFOCOM)*, Turin, April 2013, pp. 2238–2246.

[42] S. Gong, P. Wang, and W. Liu, "Spectrum sensing under distribution uncertainty in cognitive radio networks," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Ottawa, ON, June 2012, pp. 1512–1516.

[43] S. Kalamkar, A. Banerjee, and A. Gupta, "SNR wall for generalized energy detection under noise uncertainty in cognitive radio," in *Proc. 19th Asia-Pacific Conf. on Commun. (APCC)*, Denpasar, Aug 2013, pp. 375–380.

[44] J. Wang, J. Chen, Y. Lu, M. Gerla, and D. Cabric, "Robust power control under location and channel uncertainty in cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 4, no. 2, pp. 113–116, April 2015.

[45] G. Casella and R. L. Berger, *Statistical inference*. Duxbury Pacific Grove, CA, 2002, vol. 2.

[46] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[47] P. Flajolet, P. J. Grabner, P. Kirschenhofer, and H. Prodinger, "On Ramanujan's Q-function," *J. Comput. Appl. Math.*, vol. 58, no. 1, pp. 103–116, Mar. 1995. [Online]. Available: http://dx.doi.org/10.1016/0377-0427(93)E0258-N

[48] A. Balatsoukas-Stimming, A. C. Austin, P. Belanovic, and A. Burg, "Baseband and RF hardware impairments in full-duplex wireless syst.: experimental characterisation and suppression," *EURASIP J. on Wireless Commun. and Networking*, vol. 2015, no. 1, p. 142, 2015.

[49] D. Korpi, M. AghababaeeTafreshi, M. Piilila, L. Anttila, and M. Valkama, "Advanced architectures for self-interference cancellation in full-duplex radios: Algorithms and measurements," in *Proc. of the 50th Asilomar Conf. on Signals, Syst. and Comput.* IEEE, 2016, pp. 1553–1557.

[50] E. Antonio-Rodríguez, S. Werner, R. López-Valcarce, T. Riihonen, and R. Wichman, "Wideband full-duplex mimo relays with blind adaptive self-interference cancellation," *Signal Process.*, vol. 130, no. C, pp. 74–85, Jan. 2017.

[51] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton University Press, 2009.