Cognitive Communications in White Space: Opportunistic Scheduling, Spectrum Shaping and Delay Analysis

by

Shanshan Wang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2012 by the
Graduate Supervisory Committee:

Junshan Zhang, Chair
Guoliang Xue
Joseph Hui
Tolga Duman

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

A unique feature, yet a challenge, in cognitive radio (CR) networks is the user hierarchy: secondary users (SU) wishing for data transmission must defer in the presence of active primary users (PUs), whose priority to channel access is strictly higher. Under a common thread of characterizing and improving Quality of Service (QoS) for the SUs, this dissertation is progressively organized under two main thrusts: the first thrust focuses on SU's throughput by exploiting the underlying properties of the PU spectrum to perform effective scheduling algorithms; and the second thrust aims at another important QoS performance of the SUs, namely delay, subject to the impact of PUs' activities, and proposes enhancement and control mechanisms.

More specifically, in the first thrust, opportunistic spectrum scheduling for SU is first considered by jointly exploiting the memory in PU's occupancy and channel fading. In particular, the underexplored scenario where PU occupancy presents a long temporal memory is taken into consideration. By casting the problem as a partially observable Markov decision process, a set of multi-tier tradeoffs are quantified and illustrated.

Next, a spectrum shaping framework is proposed by leveraging network coding as a spectrum shaper on the PU's traffic. Such shaping effect brings in predictability of the primary spectrum, which is utilized by the SUs to carry out adaptive channel sensing by prioritizing channel access order, and hence significantly improve their throughput. On the other hand, such predictability can make wireless channels more susceptible to jamming attacks. As a result, caution must be taken in designing wireless systems to balance the throughput and the jamming-resistant capability.

The second thrust turns attention to an equally important performance metric, i.e., delay performance. Specifically, queueing delay analysis is conducted for SUs employing random access over the PU channels. Fluid approximation is taken and Poisson driven stochastic differential equations are applied to characterize the moments of the SUs' steady-state queueing delay. Then, dynamic packet generation control mechanisms are developed to meet the given delay requirements for SUs.

To my dear family.

ACKNOWLEDGEMENTS

First of all, I would like to express my deep gratitude and appreciation to my advisor, Professor Junshan Zhang, for his insightful guidance and tremendous support (professionally, financially and morally) throughput my stay as a Ph.D. student at ASU. He has been a wonderful mentor to me. From him, I learnt many things, and his unquenchable passion and amazing vision in pursuing significance has been a constant inspiration for me to pursue my potentials at all times. It is very fortunate for me to have him as my advisor.

I would like to express my sincere gratitude to my committee members, Professors Guoliang Xue, Joseph Hui and Tolga Duman, as well as my earlier committee members (for my qualifying exam), Professors Antonia Papandreou-Suppappola and Thomas Taylor. I thank them for their precious time in serving as my committee and all the valuable suggestions they have provided. Special thanks to my collaborators, Professor Lang Tong from Cornell University, Dr. Yalin E. Sagduyu from Intelligent Automation Inc., and Dr. Sugumar Murugesan, for their wonderful advices and generous help. I am very grateful to them. I would also like to thank Professor Douglas Cochran for the enlightening and delightful interactions during our meetings. In addition, many thanks to my former and current colleagues: Ashwin Kowdle, Chandra Theajaswi P.S., Dong Zheng, Qinghai Gao, Tuan Tran, Weiyan Ge, Ahmed Ewaisha, Brian Proulx, Dajun Qian, David Ganger, Kai Cai, Lei Yang, Ling Tang, Miao He, Mojgan Hedayati, Osman Yagan, Shibo He, Xiaowen Gong, and Yang Cao, for the pleasant and inspiring discussions.

Last but not least, I dedicate my thesis to my parents, who have always had such strong faith in me and are always there for me. I am also deeply

grateful to my husband, Feng Wang, who has been always so supportive and patient to me. It is such a bless to have them!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Spectrum Holes and Cognitive Radio

The traditional static spectrum allocation, where the frequency bands are designated to license holders, often called *primary users* (PUs), for exclusive use, is regarded as an obstacle for efficient utilization of the limited spectrum resources. Unlicensed users, often named as *secondary users* (SUs), are experiencing difficulty in finding communication opportunities, whereas the PUs, on the other hand, barely utilize their allocated spectrum. According to one spectrum measurement study conducted by the Shared Spectrum Company, the overall occupancy of the radio spectrum below 3GHz is less than 35%, even in the most crowded area near downtown Washington DC, where both government and commercial spectrum usage is intensive [1]. In another report by the FCC (Federal Communications Commission) Spectrum Policy Task Force [2], it is shown that a very limited portion of spectrum, namely $5\% \sim 15\%$, is utilized on average, leaving a large numbers of *"spectrum holes"* (or white space) unused, as Figure 1.1 demonstrates (whenever the PU is idle).

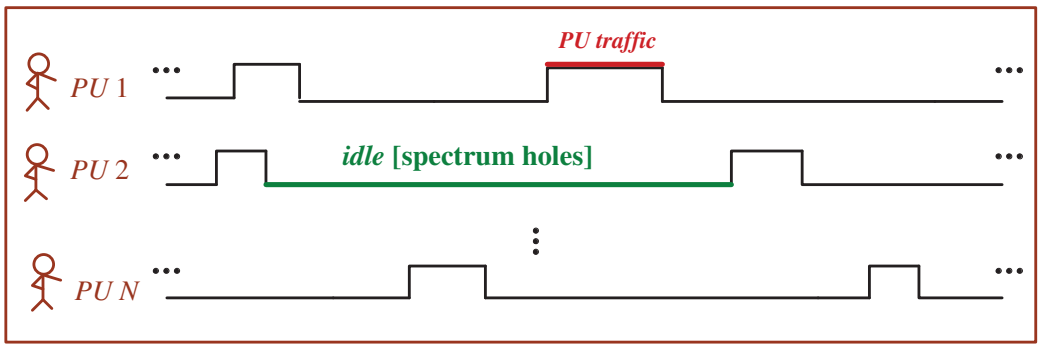

Figure 1.1: Flow level dynamics of primary user spectrum.

Dynamic spectrum access (DSA) has emerged as one promising solution to the aforementioned problem. The adaptive nature of DSA enables opportunistic access by the unlicensed users to the primary spectrum, which significantly mitigates the "virtual scarcity" problem in the existing spectrum management. There are three major models of DSA [3], as we briefly discuss below:

- **Dynamic Exclusive Use Model.** In the first model, the spectrum is assigned to licensed users for exclusive use only. Dynamic spectrum allocation among licensees is utilized, by making use of the spatial and temporal traffic characteristics of different services and dynamically allocating spectrum to coexisting (licensed) services accordingly. This approach is applicable to long-term commercial applications such as UMTS and DVB-T. Spectrum trading is permitted for licensed users to sell (partial) spectrum property rights to the SUs, where one main challenge lies in the precise definition of "property rights" of the spectrum, i.e., the "boundary" of frequency bands [4].

- **Open Sharing Model.** This model, also called "*license-exempt model*," indicates that no licenses exist or are required, and the spectrum is openly shared among peer users. The success of wireless services operating in the unlicensed industrial, scientific, and medical (ISM) radio band (for example, WiFi) gives much confidence behind this model. A key component of this model is "*spectrum etiquette*," a minimum set of restrictions placed by regulators to regulate the co-existence of peer users.

- **Hierarchical Access Model.** In the third model, "hierarchy" among users is introduced such that PUs have absolute priority over SUs and are entitled to use the assigned spectrum whenever needed. SUs, on the other hand, can utilize the spectrum only when the frequency bands are unoccupied, or the interference caused to PUs is limited under certain levels. Two approaches have been developed under this model, namely spectrum underlay and spectrum overlay. For spectrum underlay, SUs' data are carried by Ultra-Wide-Band (UWB) transmissions, where the signal power is spread over a wide frequency bandwidth, thus protecting PUs' transmissions from SUs' interference. Spectrum overlay, on the other hand, was first envisioned by Mitola under the term "spectrum pooling," and then investigated by the DARPA Next Generation (XG) program under the term "opportunistic spectrum access" [3]. Simply put, this is one approach where SUs opportunistically access the spectrum by detecting white space over time and space. In this approach, no strict constraints on the transmission power of SUs are imposed as in the spectrum underlay strategy. A typical application is the reuse of certain TV-bands that are not used for TV broadcast in a particular region.

Among the three models, the third one is perhaps of the most interest. Specifically, the hierarchical user structure therein gives rise to the challenge in designing a network that "balances" SUs' demands, while providing Quality of Service (QoS) performance guarantees to the PUs. Clearly, a key characteristic of such systems is their ability to exploit knowledge of the radio environment, and adapt their operations to it. In other words, to enable dynamic spectrum access in those systems, SUs are required to be equipped with cognitive capability to sense the dynamics in the communication environment, and utilize

it appropriately. In this context, *cognitive radio* (CR) emerges as the driving technology enabling such capabilities. Formally, as has been first introduced by Mitola [5]: "*A 'Cognitive Radio' is a radio that can change its transmitter parameters based on interaction with the environment in which it operates.*"

Since its appearance, cognitive radio systems have found potential in many areas, including rural communications, emergency applications, military scenarios and other delay-sensitive applications [6]. In the last decade, there has been a wealth of interest in studying cognitive radio, such as DARPA's NeXt Generation project and the IEEE standards 802.22 — to name a few. Needless to say, the optimal design of cognitive radio networks is very challenging. We highlight the major issues in the following.

- **Spectrum Sensing.** A first step towards opportunistic access is for SUs to identify the spectrum holes, and this is mainly performed at the PHY-layer. Three popular approaches for spectrum sensing include transmitter detection, cooperative detection and interference-based detection [7].

  In transmitter detection, SUs detect the signals from the PU's transmitter, based on their local observations. Matched filter, energy detector (radiometer) and cyclostationary feature detector are three main techniques applied for detection. One of the main focuses here is to develop detection algorithms that balance the tradeoffs between the probabilities of missing detection and false alarm [1, 3, 7–9].

  In transmitter detection, SUs perform detection independently without cooperation. While in a practical shadowing environment with fading, cognitive users may receive drastically different signal strengths and thus

different SNRs at different locations for a given transmitted primary signal. On many occasions, such signals can be very weak and thus difficult to detect for individual users. In the meantime, a secondary node seizing spectrum ranges without coordinating with others can lead to harmful interference with its surrounding neighbors, thus degrading the spectrum usage ("Tragedy of the Commons" [10]). These considerations, among other issues such as the hidden terminal problem, uncertainty in primary traffic pattern, and geographical locations of PUs and SUs, have motivated the development of cooperative detection, where SUs cooperate with each other in spectrum hole detection by exchanging information of their local observations on the spectrum. Though additional overhead may occur, the demands on the sensing sensitivity decreases, and thus a less sensitive detector can be employed [8].

Finally, in interference-based detection, the interference temperature model accounts for the cumulative RF energy from multiple SU transmissions and sets a maximum cap on their aggregate level. Then, the SUs can transmit over the primary spectrum as long as their overall signal power does not exceed this limit [7, 11].

- **Spectrum Tracking.** As the PUs' activities may vary rapidly in a random manner, to make a better decision on the usage of the detected spectrum holes, it is important for the SUs to come up with a smart channel selection strategy that "optimally" traces the availability across the whole spectrum. In this context, a tradeoff arises between applying the (usually) imperfect sensing results for immediate decisions on spectrum access, and storing them for future use. Algorithms applying sequential decision making have been proposed to characterize such

a tradeoff (see e.g., [12, 13]). In [12], the design of the optimal sensing order in a multi-channel cognitive radio network is addressed using the framework of *optimal stopping theory*. The authors considered both cases with known and unknown channel availabilities, and developed optimal strategies for channel selection accordingly. Similar approach was applied in, e.g., [14], to determine the optimal sensing policy. Besides, a partially observable Markov decision process (POMDP) was used in [13] to model SUs' sensing process and obtain the optimal sensing strategy.

It is noted that the accuracy in the modeling of spectrum occupancy is crucial to the efficiency of spectrum tracking [3]. Studies and testbed measurements (e.g., [15]) have shown that the primary spectrum exhibits a semi-Markovian property when transiting between the idle and busy states. Such observations have been utilized in many works (e.g., [13]) to model the underlying variation of the PU channels.

- **Spectrum Sharing.** In many scenarios, multiple SUs coexist in one system. It is important for individual users to achieve high utility, subject to fairness constraints. In this context, effective spectrum sharing becomes critical. Both centralized and distributed approaches have been explored to address the above issues. In centralized schemes, a central authority, such as a base station and a spectrum server, is assumed to have perfect knowledge of spectrum holes and allocates them to the requesting users [16, 17]. In contrast, in distributed algorithms, SUs sense and share the spectrum locally, without the assistance from the central authority. Interesting models based on graph coloring and game theory have been proposed under this category (see, e.g. [10, 18, 19]).

In the graph-based model, the network is represented as a bidirectional graph, where the vertices correspond to the users sharing the spectrum, and the undirected edges between vertices represent the interference among users. The color list of a vertex denotes the available spectrum to the user. Then, the graph coloring problem is defined as obtaining the optimal color (i.e., spectrum) assignment that maximizes the utility function for the vertices (i.e., SUs). The user utility can be constructed, e.g., based on the application layer requirements. Widely used criteria include Max-Sum-Bandwidth (MSB), Max-Min-Bandwidth (MMB) and Max-Proportional-Fair (MPF). While the optimal coloring is known to be NP-hard, suboptimal approaches have been proposed instead [10, 20, 21]. Models exploring local cooperation among SUs, such as local bargaining [22], have been developed as well.

On the other hand, the game theoretical approach models the PUs and SUs as players in a game, with varying payoff functions constructed based on the nature of the users (namely cooperative, selfish and malicious). The strategy space for the users varies according to their identity: For SUs, it includes the selection of the licensed band to use, the decision on transmission parameters (e.g., transmission power, frequency, bandwidth, modulation, channel coding, antenna pattern and time duration) to apply, the choice of price to pay for the spectrum usage and so on. For PUs, the strategy set may include which unused channel to lease to SUs, how much they charge, etc. Various models under this framework have been developed, such as the repeated spectrum sharing game model [18], auction-based spectrum sharing game and belief-assisted pricing model [23].

1.2 Main Scope of the Dissertation

In this dissertation, we study dynamic spectrum access by SUs under the hierarchical model with spectrum overlay, i.e., PUs have strictly higher priority over SUs in spectrum usage, and SUs opportunistically search for the spectrum holes for data transmission. Our main focus is on the QoS performance of the SUs, in terms of latency and throughput. While building upon the broader framework of spectrum tracking, our work goes beyond the existing work and contributes in the following main aspects:

- We consider opportunistic spectrum scheduling by exploiting the temporal correlation structure in *both* the channel fading and the PU occupancy states. In particular, the PU occupancy is assumed to possess a *long* temporal memory. We formulate the problem using the framework of partially observable Markov decision process and show that the optimal scheduling involves *multi-tier* "exploitation vs. exploration" tradeoffs. Such tradeoffs go beyond the classic "exploitation vs. exploration" tradeoff, as a result of the intricate interactions between the channel fading and PU occupancy states. For certain special cases, we establish the optimality of a simple greedy policy. To gain a better understanding of the tradeoffs for the general case, we introduce a full-observation genie-aided system, where the spectrum server collects channel fading states from all the PU channels. Using the genie-aided system, we decompose the multiple tiers of the tradeoffs, and examine them progressively.

- We propose a novel concept of "spectrum shaping" to investigate the fundamental issue of characterizing the properties of the actual spectrum structure, and identifying smart shaping approaches that lead to

a better tractability in the primary spectrum to the SUs. We develop smart sensing strategies for the SUs to effectively explore the primary spectrum under shaping, and propose cognitive transmission schemes where the SUs intelligently match their communication with the PUs' spectrum hole characteristics. On the other hand, we caution that the predictability induced by network coding makes wireless networks more vulnerable to jamming attacks, and a careful design is needed to balance the throughput performance and the jamming-resistant capability of the wireless systems.

- We study the delay performance in a cognitive radio network with random access. We take a nontraditional view via the stochastic fluid queue theory and model the queueing dynamics for the users using *Poisson Driven Stochastic Differential Equations* (PDSDE). Cognitive and distributed spectrum sensing mechanisms are developed accordingly that minimize the queueing delay of SUs under different scenarios. Further, dynamic packet generation control mechanisms are developed to achieve given delay requirements for SUs.

Throughout the dissertation, we assume that spectrum detection performed at the PHY-layer is perfect, and that the sharing of spectrum among SUs is carried out in a contention-based manner. In the following, we provide a brief overview of the main scope of the studies.

Consider a single SU opportunistically accessing one of multiple PU channels in the CR network. On the one hand, the SU intends to gain maximal immediate throughput by transmitting over the channel with the "best" conditions at the moment; while on the other hand, it is also necessary for

the SU to explore as many PU channels as possible at different time instants in order to obtain better understanding of the overall system dynamics, and therefore make informed optimal decisions for channel access in the near future.

In Chapter 2, we study such tradeoffs between "exploitation" and "exploration" with the objective of maximizing SU's throughput. In particular, we take further steps beyond the existing studies by: *1*) explore the utility of *both* the states (as opposed to one single state in the literature) - channel fading and PU occupancy, for opportunistic channel access; and *2*) incorporate a generic correlation structure in the PU occupancy state that features a *long* temporal memory (as opposed to memoryless or short-term memory model utilized in the literature) and develop a novel *"age"* model to capture such correlation . By casting the problem as a *partially observable Markov decision process* [24], we identify that a new set of *"multi-tier"* tradeoffs arise as a result of the intricate interactions between the PU occupancy and channel fading states. The optimality of a simple greedy policy is established under certain conditions. Further, by developing a genie-aided system with full observation of the channel fading feedbacks, we decompose and examine the multiple tiers of the tradeoffs in the original system, and recorded the impacts from channel fading and PU occupancy on such tradeoffs under the general setting.

The hierarchical channel access structure in a CR network determines that the SUs can transmit over PU channels only at idle instances of PUs. Therefore, to improve throughput, it is of great benefit for SUs to capture the spectrum access opportunities as quickly as possible. However, as illustrated in Figure 1.2, spectrum holes can take place in multiple dimensions, over space, time, and frequency. In particular, the traffic pattern of the PUs

directly determines the temporal spectrum availability and leaves room for SU transmissions whenever PUs do not have any packet traffic to transmit. The random nature of PUs' traffic, without smart processing, would lead to stochastic and sporadic transmission patterns, which hinder a quick discovery of spectrum holes.

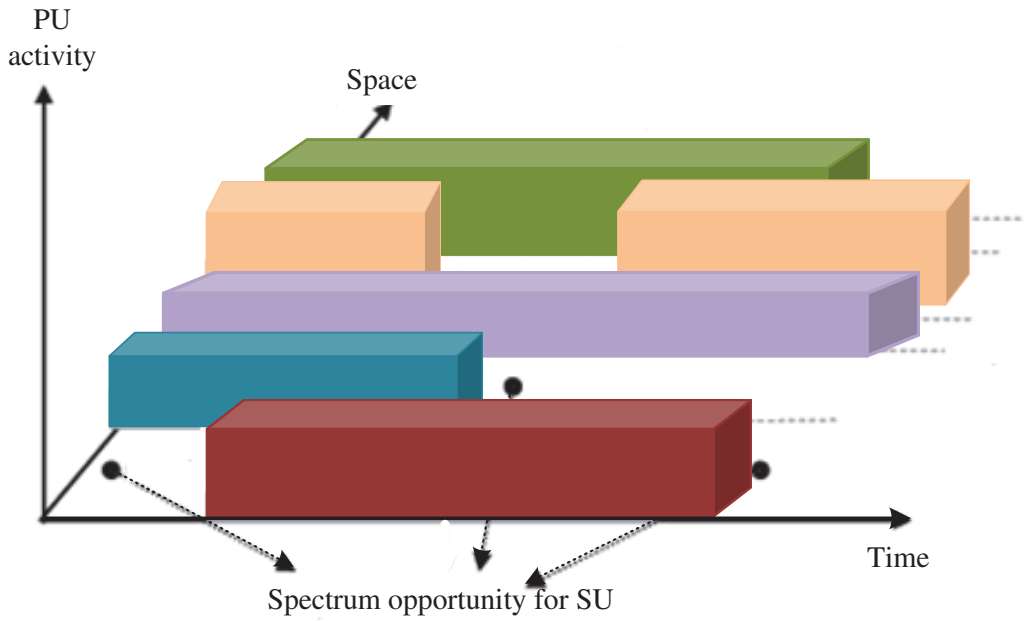

Figure 1.2: An illustration of random spatio-temporal structure of spectrum holes.

In Chapter 3, we formulate a spectrum shaping framework for a CR network with multiple PU channels, and propose to leverage *network coding* as a *"spectrum shaper"* to enhance the predictability on the primary spectrum to the SUs, thus boosting SUs' throughput. A key observation is that when PUs use network coding with batch operations, the busy periods on each of the PU channels are lower-bounded by the batch size, and the idle periods are

shaped based on the process of accumulating the packets. In other words, the primary spectrum is "shaped" by the coding process so that the distribution of the spectrum holes becomes more "regular" and "predictable," comparing to the case where classical retransmission-based retransmissions are applied. With this observation, we propose an adaptive channel sensing scheme with two-stage sensing, where SUs perform channel selection based on a "sensing list" that consists of channels predicted to be idle, and adaptively update it over time. We demonstrate that both PUs and SUs' throughput significantly improve, thanks to the shaping effect brought by network coding. Moreover, contention-based spectrum access of multiple SUs was considered and we developed a two-level backoff scheme based on the adaptive sensing to exploit the memory in network-coded transmissions and realize the throughput gains for the SUs.

Next, we show that the predictability induced by network coding makes wireless networks more vulnerable to jamming attacks targeting at the intended transmissions. Since the entropy of the busy period of network-coded transmission is less than that of the retransmission-based communication, a jammer can predict the spectrum's future status (busy or idle) more easily and increase the performance loss to the system by adaptive channel selection. Particularly, this jamming attack effect increases with the coding batch size, which, on the other hand, (desirably) improves the throughput of network-coded transmissions. This suggests that a careful design is needed to balance the throughput performance and the jamming-resistant capability of the wireless systems.

In addition to throughput, an equally important QoS metric for SUs is their delay. In Chapter 4, we study delay performance in cognitive radio

networks. It is noted that delay has always been one of the most important metrics in a wireless networks. It has important engineering implications, e.g., it can be used to characterize the number of users that can be supported under a given delay constraint. Unfortunately, the delay performance for cognitive radio networks is an under-explored area and not well understood in general, due to the analytical difficulties in characterizing queue interactions between primary and secondary networks and among SU networks. To this end, only a limited number of results have been reported in the literature [25, 26]. In Chapter 4, we take a step along this effort, and analyze the queueing delay for the SUs in a CR network with multiple PU channels. We apply the stochastic fluid queue theory and model the queueing dynamics of both PUs and SUs via *Poisson driven stochastic differential equations* (PDSDE). Based on the properties of PDSDE, we obtain the moments of SUs' queue lengths, which lead to the average queueing delay of the SUs. We focus on the light traffic regime that "decouples" the strong correlation across SUs due to the contention collisions, and characterize SUs' queue lengths accordingly under different system settings. Further, packet generation control under a given delay constraint is considered and analyzed using PDSDE.

Chapter 2

OPPORTUNISTIC SPECTRUM SCHEDULING BY JOINTLY EXPLOITING CHANNEL CORRELATION AND PU TRAFFIC MEMORY

2.1 Introduction

Being equipped with the cognitive capability, the SUs can be constantly aware of the environmental changes and adapt their operational parameters (such as power, frequency, etc.) accordingly. In a CR network, one unique environmental change is the PU's occupancy state. In particular, as a result of the hierarchical spectrum access structure, the transmissions of an SU are constrained by the stochastic nature of the PUs: An SU can access a channel only when the PU is inactive (i.e., the PU channel is idle), and must vacate the channel as soon as the PU is detected to return (i.e., the channel becomes busy). Clearly, an appropriate model for such idle-busy alternations is critical to determine the channel's accessability by an SU.

Another commonly existing variation is channel fading. Often times an *i.i.d.* flat fading model is used in abstracting fading channels, which fails to capture the temporal channel memory observed in realistic scenarios [27]. A basic model, namely the Gilbert-Elliot (GE) model [28], has been widely used in recent literature (see, e.g., [29–31]) as an alternative to capture the temporal correlation in the fading process. Specifically, the GE model uses a first-order Markov chain with two states: one representing a "good" channel where the user experiences error-free transmissions, and the other representing a "bad" channel with unsuccessful transmissions.

In most of the existing works (see, e.g., [29, 30, 32]), only one set of the system states – either the channel fading, or the PU occupancy – has been taken into consideration in developing spectrum access strategies by the SU. Furthermore, most studies rely on the assumption that the PU occupancy is *memoryless* across time (such as independent and identical distributed [33]), or has a *short-term* memory modeled by a first-order Markov chain (e.g., [29,32]). Recent works, however, have suggested that the PU occupancy may exhibit a *long* memory (see, e.g., [34]), for which a systematical study is lacking.

Motivated by these observations, in this work, we take steps forward, and: *1*) explore the utility of *both* the states - channel fading and PU occupancy, for opportunistic channel access; and *2*) incorporate a generic correlation structure in the PU occupancy state that features a *long* temporal memory. More specifically, we consider a CR network consisting of multiple PU channels and one SU opportunistically accessing one of the PU channels at a time.[1] A spectrum server is utilized to periodically schedule the SU to one of the channels for transmission. We formulate the problem as a *partially observable Markov decision process* [24], wherein the spectrum server makes scheduling decisions in terms of allocating a PU channel to the SU, based on *both* the channel's PU occupancy state and fading state. Going beyond the literature, we develop a novel *"age"* model to capture the *long* temporal correlation in the PU's occupancy state. In parallel, we model the channel fading state using a two-state first-order Markov chain, i.e., the GE model. Detailed formulation can be found in Section 2.2. Worth noting is that the usage of the spectrum server is consistent with the recent FCC ruling on the use of

[1] As can be seen from the following sections, even for the case with a single SU, there exist intricate tradeoffs and the study is highly nontrivial. It remains open to study the general model with multiple SUs.

a spectrum database in CR network operations [35]. Further, the spectrum server facilitates spectrum management, and enhances the scalability of the network [36].

Built upon the formulation, a main goal of this work is to understand the impact of both sets of system states, and their temporal correlation, on the SU's performance. We identify that *"multi-tier"* tradeoffs arise as a result of the intricate interactions between the PU occupancy and channel fading states. Such tradeoffs go beyond the classic "exploitation vs. exploration" tradeoff that was studied in the existing works (see, e.g., [29,30]). In the classic tradeoff, the answer to "whether to learn a channel or not" is purely dependent on the *one* system state under consideration. For instance, if channel fading is considered, then the channel with the "best" fading state will be scheduled to the user for "exploitation" (i.e., to obtain the maximal immediate gain), while the channel with the "least known knowledge on fading" will be favored for "exploration" (i.e., learning) for obtaining better understanding of the overall system and thus improving the total system gain in the long run. In contrast, important differences arise in our context due to the inclusion of *both* the system states. On the one hand, the stochastic nature of the PU traffic can result in termination of SU's transmission at any time. Therefore, the SU, hoping to maximize its immediate throughput (i.e., "exploitation"), would favor a channel that not only has the "best" channel fading (as in the classic setting), but is currently idle and can stay idle for as long as possible during the current scheduling period. On the other hand, due to the hierarchical spectrum access structure, the scheduling can only be performed among the idle PU channels. Accordingly, the PU channel favored for "exploration" in the classic sense (determined solely by channel fading) may no longer be preferred

if the channel is perceived to be occupied by the PU in the near future, since the learning results cannot be utilized for scheduling therein. Clearly, strong coupling between channel fading and PU occupancy exists and impacts the tradeoffs under the new setting. The scheduling process, in the presence of such coupling, can become even more complicated when temporal correlation in both sets of system states are included, particularly when such memory bears a long-term correlation structure. As a result, the answer to the same question: "whether to learn a channel or not" must now be determined by "balancing" the two sets of interacted system states, and by understanding the impact of the inherent memory on such a balancing. This is the main goal of our study.

The main contributions of this chapter can be summarized as follows:

- We study opportunistic spectrum scheduling by exploiting the temporal correlation structure in *both* the channel fading and the PU occupancy states. In particular, the PU occupancy is assumed to possess a *long* temporal memory. This, to the best of our knowledge, has not been addressed systematically in the literature so far.

- We formulate the problem using the framework of partially observable Markov decision process and show that the optimal scheduling involves *multi-tier* "exploitation vs. exploration" tradeoffs. Such tradeoffs go beyond the classic "exploitation vs. exploration" tradeoff, as a result of the intricate interactions between the channel fading and PU occupancy states.

- For certain special cases, we establish the optimality of a simple greedy policy, and examine the intricacy of the fundamental tradeoffs therein.

- To gain a better understanding of the tradeoffs for the general case, we introduce a full-observation genie-aided system, where the spectrum server collects channel fading states from all the PU channels. Using the genie-aided system, we decompose the multiple tiers of the tradeoffs, and examine them progressively.

The rest of the chapter is organized as follows. Section 2.2 introduces the basic setting and problem formulation in detail. In Section 2.3, we identify the fundamental tradeoffs and illustrate them by examining both the immediate reward and the total reward, and establish the optimality of a simple greedy policy for special cases. Section 2.4 further examines the tradeoffs by developing a genie-aided system that isolates the impact of channel fading and PU occupancy on the optimal reward. In Section 2.5, numerical results are presented with further discussions. Finally, concluding remarks are given in Section 2.6.

2.2 Problem Formulation

2.2.1 Basic Setting

We consider a CR network with one SU and N PUs.[2] Each PU is licensed to one of N independent channels, henceforth identified as PU channels. A PU generates packets according to a stationary process, transmits over its channel if there are backlogged packets, and leaves upon the completion of the transmissions. The PU traffic activity is assumed to be identical and independent across channels.[3] For the SU, we assume it is backlogged. Time is divided into two timescales: mini-slots and the control slots each constituting K mini-slots,

[2]Each user is assumed to be a pair of transmitter and receiver.

[3]Our study here can be readily extended to the case with nonidentical PU channels whilst the fundamental tradeoffs are retained therein.

as illustrated in Fig. 2.1. The length of each mini-slot is normalized to fit the transmission of one data packet of the PU or the SU. At the beginning of each control slot, the spectrum server schedules the SU to the "best" PU channel expected to yield the highest average throughput for the SU. The SU then transmits packets in the scheduled channel, until it detects the return of a PU.[4] Upon such an event, the SU suspends transmissions until the beginning of the next control slot, when the spectrum server re-schedules the SU to a PU channel based on most recent observations. At the end of each mini-slot when the SU transmitted a packet, it sends accurate feedback on the channel fading state (of the PU channel, as seen by the SU) corresponding to that mini-slot, to the spectrum server. The spectrum server uses this feedback (on channel fading), the observations on PU traffic, along with the memory inherent in these processes to perform informed scheduling decisions at the beginning of the next control slot. We discuss the system model and the scheduling problem formulation in more detail in the following.

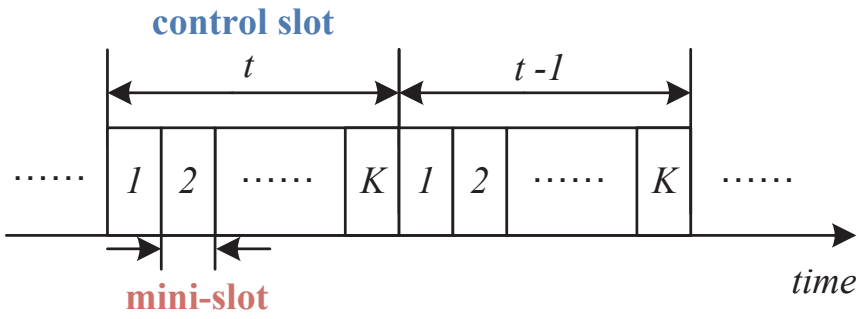

Figure 2.1: A sketch of the two timescale model.

[4]This can be accomplished by incorporating collision detection by the SU at the mini-slot timescale. We also assume that PU arrivals coincide with the mini-slot boundaries.

2.2.2 Problem Formulation

The opportunistic spectrum access at hand can be viewed as a sequential control problem, which we formulate as a partially observable Markov decision process. In the following, we introduce and elaborate the entities involved in the formulation.

Channel occupancy: The usage pattern on each of the PU channels can be modeled as an ON-OFF process at the mini-slot timescale, with ON denoting the *busy* state where the PU transmits data over the channel, and OFF the *idle* state where the PU is absent. Channel occupancy is the idle or busy state of the PU channels. Let $o_{t,k}(n)$ be a binary random variable, denoting whether PU channel n, for $n \in \{1, \ldots, N\}$, is idle ($o_{t,k}(n) = 0$), or not ($o_{t,k}(n) = 1$), in the kth mini-slot of control slot t. The corresponding idle probability is denoted by $\pi^o_{t,k}(n) \triangleq \Pr(o_{t,k}(n) = 0)$.

Idle/Busy age: The PU traffic is temporally correlated, i.e., the current occupancy state on each of the channel depends on the history of the channel occupancy. We introduce the notion of "age," defined as follows, to characterize the occupancy history:

*The **age** of a PU channel is the number of consecutive mini-slots immediately preceding the current mini-slot, during which the channel is in the same occupancy state as in the current mini-slot. The age is denoted as "idle age" if the channel is in idle state in the current mini-slot and "busy age" otherwise.* We use $x_t(n)$ to denote the age of channel n at the beginning of control slot t.

As noted earlier, we assume long memory in the PU occupancy. Specifically, with the definition of age in place, we adopt a family of functions monotonically decreasing in age, to denote the *conditional probability that a channel will be idle (or busy), given that it has been idle (or busy) for* $x \geq 1$ *mini-slots*:

$$\begin{aligned} P_I(x) &= \frac{1}{x^u + C_I}, \\ P_B(x) &= \frac{1}{x^u + C_B}, \quad u = 1, 2, ..., \end{aligned} \tag{2.1}$$

where C_I and C_B are normalizing constants taking positive values. Our occupancy model essentially imposes the following realistic correlation structures: *1*) the occupancy memory weakens with time, i.e., the impact of past occupancy events on the current occupancy state diminishes since the said event happened; *2*) the conditional probability that the PU channel is busy or idle now, is purely a function of the length of time the channel has been in the most recent state, and is independent of the channel occupancy history before the time of the latest transition to the most recent state. In sight of this, the quantities P_I and P_B defined in (2.1) are sufficient for capturing the temporal correlation in the channels' PU occupancy state.

Channel fading model: At the end of each mini-slot after transmitting a packet, the SU measures the channel fading between its transmitter and receiver on the scheduled channel, and feeds back this information to the spectrum server. Inspired by recent works [29,30,37], we capture the memory in the fading (of the PU channel) between the SU's transmitter and receiver using a two-state, first-order Markov chain, with state variations occurring at the mini-slot timescale. The Markov chain model is *i.i.d.* across the PU channels. Each state of the Markov chain corresponds to the degree of decodability of the data sent through the channel, where state 1 denotes full decodability and state 0 denotes zero decodability. Note that the states can

also be interpreted as a quantized representation of the underlying channel fading, in the sense that state 1 corresponds to "good" channel fading, while state 0 corresponds to "bad" fading. The probability transition matrix of this Markov chain is given as:

$$\mathbf{P} := \begin{bmatrix} 1-r & r \\ 1-p & p \end{bmatrix}, \tag{2.2}$$

where p is the conditional probability that the channel fading is good, given that it was good in the previous mini-slot; and r is the conditional probability that the channel fading is good, given that it was bad in the previous mini-slot. Throughout the paper, we will focus on the case when the fading channels are positively correlated[5], i.e., $p > r$.

Belief of channel fading state: Denote by $\pi^s_{t,k}(n)$ the belief of channel fading state in the kth mini-slot of control slot t on channel n. As is standard [24, 29], the fading state belief is a sufficient statistic that characterizes the current channel fading state as perceived by the SU. Further, let $f_{t,k}(a_t)$ be a binary random variable denoting the fading state feedback obtained at the end of the kth mini-slot in control slot t on the scheduled channel a_t. Also, define $\mathrm{T}^L(\cdot)$, for $L \in \{0, 1, \ldots\}$, as the Lth step belief evolution operator, taking the form: for $\gamma \in (0, 1)$,

$$\mathrm{T}^L(\gamma) = \mathrm{T}(\mathrm{T}^{L-1}(\gamma)), \tag{2.3}$$

with $\mathrm{T}^0(\gamma) = \gamma$ and $\mathrm{T}(\gamma) = \gamma p + (1-\gamma)r$. Now, the update of the fading state belief is governed by the underlying Markov chain model, and any new

[5]We focus on the positively correlated case, as it has been shown to capture more realistic scenarios than the negatively correlated model (see, e.g., [27, 30] and the references therein).

information obtained on the channel fading, i.e.:

$$\pi^s_{t,k+1}(n) = \begin{cases} p, & a_t = n, f_{t,k}(a_t) = 1, \\ r, & a_t = n, f_{t,k}(a_t) = 0, \\ \mathrm{T}(\pi^s_{t,k}(n)), & a_t \neq n. \end{cases} \tag{2.4}$$

Action space: This refers to the set of channels that the scheduling decision is made from. The spectrum server selects channels only from those that are currently idle[6], and the action space $\mathcal{A}_t$ in control slot t can thus be written as:

$$\mathcal{A}_t = \{n : o_{t,1}(n) = 0\}. \tag{2.5}$$

State: At the beginning of each control slot, the spectrum server makes the scheduling decision based on three factors: For each of the PU channels, *1*) the idle/busy state at the moment; *2*) the length of time the channel has been in the current occupancy state (i.e., age); and *3*) the fading state belief value. That is, the state of each PU channel n, is represented by a three-dimensional vector: $S_t(n) = [o_{t,1}(n), x_t(n), \pi^s_{t,1}(n)]$. Accordingly, the state of the system at the beginning of current control slot t is described by a $N \times 3$ matrix $\mathbf{S}_t$:

$$\mathbf{S}_t := [S_t(1); \ldots; S_t(N)] = \begin{bmatrix} o_{t,1}(1) & x_t(1) & \pi^s_{t,1}(1) \\ o_{t,1}(2) & x_t(2) & \pi^s_{t,1}(2) \\ \vdots & \vdots & \vdots \\ o_{t,1}(N) & x_t(N) & \pi^s_{t,1}(N) \end{bmatrix}. \tag{2.6}$$

Horizon: The horizon is the number of consecutive control slots over which scheduling is performed. We index the control slots in a decreasing order with control slot 1 being the end of the horizon.[7] Throughout the paper, we

[6]This is a policy level constraint to protect the PU's priority in spectrum access.

[7]For the mini-slots, we use the conventional increasing time indexing with the lower-case letter k.

denote the length of the horizon by m, i.e., the scheduling process begins at control slot m.

Stationary scheduling policy: A stationary scheduling policy $\mathcal{P}$ establishes a stationary mapping from the current state $\mathbf{S}_t$ to an action a_t in each control slot t.

Expected immediate reward: The expected immediate reward is the reward accrued by the SU within the current control slot. Specifically, the SU collects one unit of reward in each mini-slot, if the channel is idle and has good channel fading (i.e., conditions that indicate successful transmission by SU). Since the scheduled channel must be idle in the first mini-slot of the current control slot, the expected immediate reward can be calculated as:

$$R_t(\mathbf{S}_t, a_t) = \sum_{k=2}^{K} \pi^o_{t,k}(a_t)\pi^s_{t,k}(a_t) + \pi^s_{t,1}(a_t). \tag{2.7}$$

Total discounted reward: Given a scheduling policy $\mathcal{P}$, the total discounted reward, accumulated from the current control slot t, until the horizon, can be written as[8]

$$V_t(\mathbf{S}_t; \mathcal{P}) = R_t(\mathbf{S}_t, a_t) + \beta E_{\boldsymbol{\pi}^s_{t-1}} E_{\mathbf{O}_{t-1}} V_{t-1}\left(\mathbf{S}_{t-1}; \mathcal{P}\right), \tag{2.8}$$

where $\beta \in (0, 1)$ is the discount factor, facilitating relative weighing between the immediate and future rewards, and the expectation is taken with respect to fading state belief: $\boldsymbol{\pi}^s_{t-1} = \{\pi^s_{t-1,1}(1), \ldots, \pi^s_{t-1,1}(N)\}$, and PU occupancy: $\mathbf{O}_{t-1} = \{o_{t-1,1}(1), \ldots, o_{t-1,1}(N)\}$.

Objective function: The objective of the scheduling problem is to maximize the SU's throughput, i.e., SU's total discounted reward. A scheduling

[8]In the subsequent sections, we may drop $\mathcal{P}$ and the tiers of expectation to simplify the notation.

policy $\mathcal{P}^*$ is optimal if and only if the following optimality equation is satisfied:

$$V(\mathbf{S}_t;\mathcal{P}^*)=\max_{a_t\in\mathcal{A}_t}\left\{R_t(\mathbf{S}_t,a_t)+\beta E_{\boldsymbol{\pi}^s_{t-1}}E_{\mathbf{O}_{t-1}}V(\mathbf{S}_{t-1};\mathcal{P}^*)\right\}. \tag{2.9}$$

The function $V^*(\mathbf{S}_t) := V(\mathbf{S}_t;\mathcal{P}^*)$ is the objective function of the scheduling problem.

2.3 Fundamental Tradeoffs

The decision on opportunistic spectrum scheduling is made based on two sets of system states: the PU occupancy on the channel and the channel fading perceived by the SU. On the one hand, PUs may return in the middle of a control slot and hinder further transmissions of the SU, leading to a decreased reward for the SU. The temporal memory resident in the PU occupancy suggests that the past history of channel's occupancy, measured by the age, influences the occupancy state of the channel in the future. On the other hand, the PU channels may suffer from "bad" channel fading in the middle of a control slot, even if a PU does not return to hinder SU's transmissions.. Similar to the PU occupancy, the historic observations on the fading process would help determine the expected channel fading in the future. Note that by way of the channel feedback arrangement, an observation of a PU channel fading is made only when that channel is scheduled to the SU. Thus scheduling is inherently tied to channel fading learning. Roughly speaking, to maximize the SU's reward, the spectrum server must schedule a channel such that the combination of the perceived channel occupancy and channel fading strikes a "perfect" balance between the immediate gains and channel learning for future gains. We discuss this intricate tradeoff in the following.

2.3.1 Classic "Exploitation vs. Exploration" Tradeoff

In the existing literature (e.g., [29, 30, 32, 37, 38]), focus has been cast on considering only one of the factors: either channel fading or channel occupancy, along with the associated temporal correlation. The optimal decision is a mapping that best balances the tradeoff of "exploitation" and "exploration" on the single factor being considered. The exploitation side lets the scheduler choose the channel with the best perceived channel fading (or occupancy state) at the moment, corresponding to immediate gains; while the exploration side tends to favor the channel with the least learnt information so far, probing which can contribute to the overall understanding of the channel fading (or occupancy state) in the network, and thus better scheduling decisions in the future.

2.3.2 "Exploitation vs. Exploration" Tradeoff in Dynamics of Both Channel Fading and PU Occupancy States

In contrast to the existing works, we examine the tradeoffs when the temporal correlation in *both* the channel fading and PU occupancy are considered. While the classic tradeoff described above apparently exists, additional tradeoffs arise in our context due to the interactions between the two sets of system states. In particular, the long temporal memory in the PU occupancy state adds a new layer to the tradeoffs inherent in the problem. For instance, note that the SU can only transmit on *idle* channels. To carry out exploration, the channel being favored in the traditional sense, i.e., the one on which the least information is available, may no longer be the preferred choice if this channel is perceived to be unavailable (i.e., busy) for a prolonged duration in the near

future. In other words, it may not be worth learning the channel as the SU cannot utilize the learned knowledge in the near future.

Fig. 2.2 is a pictorial illustration of the impact from the occupancy state on the SU's expected reward. The history of occupancy, represented by the idle ages $x_t(1)$ and $x_2(t)$, affects both the immediate and future rewards of the SU. Specifically, as the idle age increases, the temporal memory in the PU's occupancy pushes the channel to transit to busy sooner (i.e., time point b comes earlier than a in the figure). Therefore, the average availability on the PU channel in the current control slot decreases, which leads to a smaller immediate reward for the SU.

Further, note that the latest mini-slot for which the spectrum server receives channel fading feedback is also the last mini-slot before the PU returns, i.e., time points a and b respectively for the two cases in Fig. 2.2. The duration d_1 (likewise, d_2) in Fig. 2.2 is an indication of how "fresh" the channel fading information is for the scheduling decision at the beginning of the next control slot, i.e., $t-1$. With $d_1 < d_2$, channel feedback is more fresh in the former case, with a lower idle age $x_t(1)$. Thus, age, through its effect on the *freshness* of feedback, and the availability of the PU channels in the future slots, adds another layer to the tradeoffs, thereby influencing the optimal scheduling decision.

To better perceive the intricate tradeoffs in the system, we proceed, in what follows, with a number of break-down results that aim at illustrating each tier of the tradeoffs progressively.

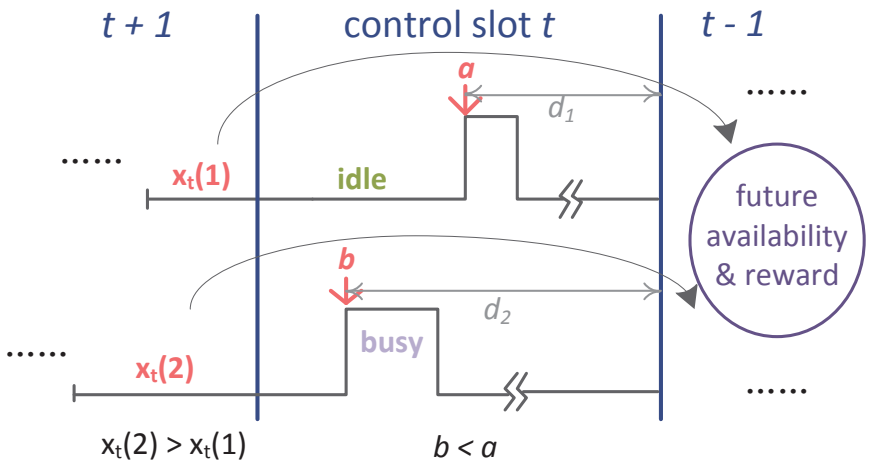

Figure 2.2: An illustration of the impact of age on the SU's reward.

2.3.3 Tradeoffs Inherent in Immediate Reward

Consider the PU channel scheduled in the current control slot t. Let k_0 denote the latest mini-slot before the PU of the scheduled channel returns in the current control slot. Clearly, k_0 is a random variable, taking values in $k_0 \in \{1, \ldots, K\}$. Let $p_z \triangleq \Pr(k_0 = z)$. With x denoting the idle age of the scheduled channel, k_0 is distributed as follows: For $K = 2$:

$$p_1 = 1 - P_I(x+1), \quad p_2 = P_I(x+1),$$

and for $K \geq 3$:

$$\begin{aligned} p_1 &= 1 - P_I(x+1), \\ p_z &= \prod_{j=1}^{z-1} P_I(x+j)(1 - P_I(x+z)), z = 2, \ldots, K-1, \\ p_K &= \prod_{j=1}^{K-1} P_I(x+j). \end{aligned} \tag{2.10}$$

In the following lemma, we establish the structure of the distribution of k_0.

Lemma 1. *The distribution of k_0 is monotonically decreasing in the idle age, x_t, for $z = 2, \ldots, K$, and monotonically increasing in x_t, for $z = 1$.*

Proof. For $K = 2$, it is straightforward to establish the conclusion. We next focus on $K \geq 3$. First, it is easy to show that for $z = 1$, p_1 is monotonically increasing in x since $P_I(x)$ is monotonically decreasing. Further, note that if two positively valued functions, $f_1(x) > 0$ and $f_2(x) > 0$, are both monotonically decreasing in x, i.e., for any positive integer $\Delta \geq 1$,

$$f_1(x) > f_1(x + \Delta), \quad f_2(x) > f_2(x + \Delta),$$

then the product of the two, $f_1(x)f_2(x)$, is also a decreasing function since

$$f_1(x)f_2(x) > f_1(x + \Delta)f_2(x + \Delta).$$

Therefore, p_K is monotonically decreasing in x.

Next, we show that $p_z, z = 2, \ldots, K - 1$ are monotonically decreasing in x. Based on the above argument, it is sufficient to show that for any $z = 2, \ldots, K - 1$, the following function,

$$g(x) \triangleq P_I(x + z - 1)(1 - P_I(x + z)),$$

is monotonically decreasing. We have the following simplifications: For $\Delta \geq 1$, a positive integer,

$$\begin{aligned}\frac{g(x)}{g(x+\Delta)} &= \frac{P_I(x+z-1)(1-P_I(x+z))}{P_I(x+\Delta+z-1)(1-P_I(x+\Delta+z))} \\ &= \frac{(x+\Delta+z-1)^u + C_I}{(x+z)^u + C_I} \cdot \frac{(x+\Delta+z)^u + C_I}{(x+\Delta+z)^u + C_I - 1} \cdot \frac{(x+z)^u + C_I - 1}{(x+z-1)^u + C_I} \\ &\triangleq B_1 B_2 B_3. \end{aligned} \tag{2.11}$$

Since $x \geq 0, C_I > 0, u \geq 1$ and $\Delta \geq 1$, we immediately have $B_1 \geq 1$, $B_2 > 1$. Further, using the binomial theorem, we obtain

$$(x + z)^u = \sum_{i=0}^{u} \binom{u}{i} (x + z - 1)^i \geq (x + z - 1)^u + 1, \tag{2.12}$$

and hence $B_3 \geq 1$. As a result, $\frac{g(x)}{g(x+\Delta)} > 1$ and $p_2, \ldots, p_{K-1}$ are monotonically decreasing in x. This concludes the proof. □

We next present a result that demonstrates the tradeoff inherent in the immediate reward with respect to age.

Proposition 1. *The immediate reward on the scheduled channel is monotonically decreasing in the idle age.*

Proof. As the system model implies, we can rewrite the immediate reward as the following weighted sum:

$$R_t(\mathbf{S}_t, a_t) = \sum_{z=1}^{K} \sum_{k=1}^{z} \pi_{t,k}^s(a_t) p_z, \tag{2.13}$$

where $p_z = \Pr(k_0 = z)$ is given by (2.10).

When $K = 2$, with idle age on a_t being x_t, we have $R_t(\mathbf{S}_t, a_t) = \pi_{t,1}^s(a_t) + P_I(x_t + 1)\pi_{t,2}^s(a_t)$. Apparently, it increases as x_t decreases.

For $K \geq 3$, denote by $\theta_z \triangleq \sum_{k=1}^{z} \pi_{t,k}^s(a_t), z = 1, \ldots, K$. It is clear that

$$\theta_1 < \theta_2 < \cdots < \theta_K, \tag{2.14}$$

and $\{\theta_z\}$'s are constants in the idle age x_t. To emphasize the role of the argument x_t, we rewrite $R_t(\mathbf{S}_t, a_t)$ as $R_t(x_t)$, and p_z as $p_z(x_t)$.

Utilizing the result of Lemma 1, and noting that for any positive integer $\Delta \geq 1$, $|p_1(x_t) - p_1(x_t + \Delta)| = \sum_{z=2}^{K} (p_z(x_t) - p_z(x_t + \Delta))$, we obtain:

$$\begin{aligned} R_t(x_t) - R_t(x_t + \Delta) &= \sum_{z=1}^{K} \theta_z \left(p_z(x_t) - p_z(x_t + \Delta)\right) \\ &> \theta_1 (p_1(x_t) - p_1(x_t + \Delta)) + \theta_2 \sum_{z=2}^{K} (p_z(x_t) - p_z(x_t + \Delta)) \\ &> 0, \end{aligned}$$

i.e., $R_t(x_t)$ is monotonically decreasing in the idle age x_t, and this establishes the proposition. □

The above result can be readily extended to the following corollary.

Corollary 1. *When all PU channels have equal fading state beliefs, the immediate reward is maximized by scheduling the SU to the channel with the lowest idle age.*

Next, recall that the channel fading is modeled by a positively-correlated Markov chain. Hence, if $\pi^{s}_{t,1}(a_t) > \pi^{s'}_{t,1}(a_t)$, then the inequality $\pi^{s}_{t,k}(a_t) > \pi^{s'}_{t,k}(a_t)$ holds, for all $k = 2, \ldots, K$. We present the following proposition without further proof.

Proposition 2. *The immediate reward on the scheduled channel is monotonically increasing in its fading state belief. Further, given equal idle ages across all PU channels, the immediate reward is maximized by scheduling the SU to the channel with the largest fading state belief value at the moment.*

2.3.4 Tradeoffs Inherent in Total Reward

In this subsection, we illustrate the tradeoffs inherent in the total reward by examining a special case with two channels $N = 2$ and number of mini-slots $K = 1$. In particular, we show that under these conditions, a simple greedy scheduling policy is optimal. The greedy policy is formally defined as follows: In any control slot, the greedy decision maximizes the immediate reward, ignoring the future rewards, i.e.,

$$\hat{a}_t = \max_{a_t \in \mathcal{A}_t} \{R_t(\mathbf{S}_t, a_t)\}. \tag{2.15}$$

We now formally record the result on greedy policy optimality in the following proposition.

Proposition 3. *The greedy policy is optimal with one mini-slot per control slot and two channels in the system.*

Proof. To prove the proposition, we begin with the following induction hypothesis:

Induction Hypothesis: With the length of the horizon denoted by $m, m \geq 2$*, assume that greedy policy is optimal in all the control slots* $t \in \{m-1, \ldots, 1\}$.

The proof proceeds in two steps: First, we fix a sequence of scheduling decisions $\vec{l} := \{a_m, \ldots, a_{t+1}\}$, and show that the expected immediate reward in control slot t, under the greedy policy, is not dependent on the scheduling decisions $\vec{l}$. Then, we provide induction based arguments to validate the induction hypothesis and hence establish that the greedy policy is optimal in all the control slots.

Let $U_t^{(\vec{l})}$ denote the expected immediate reward in slot $t \in \{m-1, \ldots, 1\}$, given the scheduling decisions $\vec{l}$. $U_t^{(\vec{l})}$ can be calculated as:

$$U_t^{(\vec{l})} = \sum_{\{o_t(1), o_t(2)\}} U_t^{(\vec{l})}(o_t(1), o_t(2)) \Pr(o_t(1), o_t(2)), \tag{2.16}$$

where $o_t(1)$ (likewise, $o_t(2)$) is the binary indicator of whether channel 1 (or 2) is idle ($o_t(1) = 0$) or not ($o_t(1) = 1$) in the tth control slot, and $\Pr(o_t(1), o_t(2))$ denotes the joint probability of both channels' availability status (idle or busy). There exist four realizations of the vector $(o_t(1), o_t(2))$, namely $\{(0,0), (0,1), (1,0), (1,1)\}$. In what follows, we show that the value of $U_t^{(\vec{l})}(o_t(1), o_t(2))$ calculated under each of the realizations is not dependent on

the scheduling decisions $\vec{l}$, i.e., for each realization, the reward calculated is identical when considering different scheduling decisions.

Case 1: When $(o_t(1), o_t(2)) = (0, 0)$. In this case, both channels are idle in control slot t. Let $\pi^s_{t+1}(n)$ be the fading state belief on channel n in control slot $t+1$. The expected immediate reward in control slot t, under the greedy policy, can be calculated as

$$U_t^{(\vec{l})}(0,0) = \pi^s_{t+1}(a_{t+1})p + (1 - \pi^s_{t+1}(a_{t+1}))\mathrm{T}(\pi^s_{t+1}(\tilde{a}_{t+1})), \tag{2.17}$$

where $\tilde{a}_{t+1} = \{1, 2\} \backslash a_{t+1}$. For notational convenience, we write $\alpha \triangleq \pi^s_{t+1}(a_{t+1})$ and $\tilde{\alpha} \triangleq \pi^s_{t+1}(\tilde{a}_{t+1})$. The reward $U_t^{(\vec{l})}(0,0)$ can now be further expressed as

$$\begin{aligned} U_t^{(\vec{l})}(0,0) &= p\alpha + (1-\alpha)\mathrm{T}(\tilde{\alpha}) \\ &= p\alpha + (1-\alpha)(\alpha p + (1-\alpha)r) \\ &= pP_1 + rP_2, \end{aligned} \tag{2.18}$$

where

$$P_1 \triangleq \alpha + \tilde{\alpha} - \alpha\tilde{\alpha}, \quad P_2 \triangleq (1-\alpha)(1-\tilde{\alpha}). \tag{2.19}$$

That is, P_1 is the probability that at least one of the channels experiences good channel fading in the previous control slot $t+1$, while P_2 is the probability that both channels see bad channel fading. It is noted that these probabilities are controlled by the underlying Markov dynamics only, and thus P_1 and P_2 are not dependent on the scheduling decisions $\vec{l}$. Therefore, $U_t^{(\vec{l})}(0,0)$ is unaffected by $\vec{l}$.

Case 2: When $(o_t(1), o_t(2)) = (0, 1)$. In this case, only channel 1 is idle and can be scheduled. The reward $U_t^{(\vec{l})}(0,1)$ is obtained as

$$U_t^{(\vec{l})}(0,1) = \begin{cases} p\pi^s_{t+1}(1) + r(1 - \pi^s_{t+1}(1)), & \text{if } a_{t+1} = 1, \\ \mathrm{T}(\pi^s_{t+1}(1)), & \text{if } a_{t+1} = 2. \end{cases} \tag{2.20}$$

It follows from (2.3) that

$$U_t^{(\vec{l})}(0,1)|_{a_{t+1}=1} = U_t^{(\vec{l})}(0,1)|_{a_{t+1}=2}, \tag{2.21}$$

i.e., $U_t^{(\vec{l})}(0,1)$ is not dependent on $\vec{l}$.

Case 3: When $(o_t(1), o_t(2)) = (1,0)$. Similar to *Case 2*, only channel 2 can be scheduled in this case, and we have:

$$U_t^{(\vec{l})}(1,0) = \begin{cases} \mathrm{T}(\pi_{t+1}^s(2)), & \text{if } a_{t+1} = 1, \\ p\pi_{t+1}^s(2) + r(1 - \pi_{t+1}^s(2)), & \text{if } a_{t+1} = 2. \end{cases} \tag{2.22}$$

Again, this indicates that

$$U_t^{(\vec{l})}(1,0)|_{a_{t+1}=1} = U_t^{(\vec{l})}(1,0)|_{a_{t+1}=2}, \tag{2.23}$$

i.e., $U_t^{(\vec{l})}(1,0)$ is not dependent on $\vec{l}$.

Case 4: When $(o_t(1), o_t(2)) = (1,1)$. In this case, both channels are busy, and it follows immediately that

$$U_t^{(\vec{l})}(1,1) = 0. \tag{2.24}$$

Clearly, $U_t^{(\vec{l})}(1,1)$ is not dependent on $\vec{l}$ as well.

Next, note that $\Pr(o_t(1), o_t(2))$ is a function of the ages $(x_t(1), x_t(2))$ only, which evolve independently from the scheduling decisions $\vec{l}$. Thereby, we conclude that $\Pr(o_t(1), o_t(2))$ is unaffected by the scheduling decisions $\vec{l}$, and so is the expected immediate reward in control slot t, i.e, $U_t^{\vec{l}} = U_t$.

Now, by extension, we have that the total reward collected from control slot t till the horizon is unaffected by $\vec{l}$, i.e.,

$$\sum_{t'=t}^{1} U_{t'}^{\vec{l}} = \sum_{t'=t}^{1} U_{t'}.$$

Thus, the greedy policy is optimal in control slot $t+1$ as well. Since $t \in \{m-1, \ldots\}$ is arbitrary, the greedy policy is optimal in every control slot $\{m, \ldots, 1\}$ under the induction hypothesis.

Finally, as the greedy policy is trivially optimal at the horizon, i.e., $t = 1$, using backward induction, we validate the induction hypothesis, and arrive at the conclusion that greedy is optimal in all control slots $t \in \{m, \ldots, 1\}$. This establishes the proposition. □

Remarks: Note that the tradeoffs inherent in the special case considered above, i.e., $K = 1, N = 2$, is more intricate than those observed in related recent works (e.g., [29,30]), where a control slot coincides with a mini-slot and only one of the states: channel fading or PU occupancy, is considered. This is because, despite $K = 1$, the question of "whether to learn a channel that may not be available for scheduling in the near future due to channel occupancy state" still exists. Thus the multiple layers of the tradeoffs discussed in the preceding subsections are retained in this special case, essentially adding value to our result on greedy optimality.

In the subsequent section, we proceed to further understand the tradeoffs in the original system by introducing a conceptual "genie-aided system."

2.4 Multi-tier Tradeoffs: A Closer Look via A Genie-Aided System

In the previous section, we examined the interaction between various state elements by quantifying the immediate reward, and also the total rewards in certain special cases. In order to obtain a more complete understanding of the inherent dynamics, we next introduce a full-observation genie-aided system that helps decompose and characterize the various tiers of the multi-

dimensional tradeoffs in the original system. Note that the results in the following hold for the general setting with arbitrary values of $N \geq 2$ and $K \geq 1$.

2.4.1 A Genie-Aided System

The genie-aided system is a variant of the original system with the following modification: The spectrum server receives channel fading feedbacks from *all* the channels and not only the scheduled channel. These feedbacks are collected at the same times as those of the feedbacks from the scheduled channel. Thus when the PU returns on the scheduled channel, the feedbacks from *all* the channels stop at once. Note that this is a conceptual system, without practical significance, which as we will see, is helpful in better understanding the complicated tradeoffs inherent in the original system.

2.4.2 Impact of Channel Fading on the Tradeoffs

We first illustrate the effect of channel fading on the optimal scheduling decisions in the following proposition.

Proposition 4. *When the idle ages are the same across all PU channels, it is optimal to schedule the SU to the channel with the highest fading state belief at the moment, i.e.,*

$$a_t^* = \arg\max_n \{\pi_{t,1}^s(n)\}. \tag{2.25}$$

Proof. First, from Proposition 2, the immediate reward is maximized when scheduling the channel with the highest fading state belief at the moment. Now, we focus on showing that the future reward does not change when the action in the current control slot varies and therefore establishing the proposition. Specifically, at the current control slot t, $t \geq 2$, consider an arbitrary

control slot in the future, $t_0 \in \{t-1, \ldots, 1\}$. In the following, we show that the expected immediate reward in this control slot, denoted by U_{t_0}, is not dependent on the current action a_t, and thus the future reward is not dependent on a_t.

Let k_0' denote the latest idle mini-slot in control slot t_0+1 before the PU returns. The reward U_{t_0} is then calculated as:

$$U_{t_0} = \sum_{k_0'=1}^{K} U_{t_0}(k_0') \Pr(k_0'), \tag{2.26}$$

where $\Pr(k_0')$ is the distribution of k_0', identical to that of k_0 as given in (2.10), and $U_{t_0}(k_0')$ is the expected reward in control slot t_0 for a given k_0'.

In the genie-aided system, the spectrum server obtains the feedbacks of the fading states from *all* N channels simultaneously, i.e., at the end of mini-slot k_0'. Place the channels on which good channel fading is observed at k_0' in the set $\mathcal{C}_1$, and the rest in another set $\mathcal{C}_0$. The characterization of the reward $U_{t_0}(k_0')$ can be further divided into the following cases.

- *Case 1:* $\mathcal{C}_0 = \emptyset$. This corresponds to the case where $f_{t_0+1,k_0'}(n) = 1, \forall n = 1, \ldots, N$. Let $W_{t_0}^{(\text{case1})}(n)$ be the expected reward in control slot t_0 on channel n in this case. We have

$$W_{t_0}^{(\text{case1})}(n) = \pi_{t_0,1}^{o}(n) \sum_{z=1}^{K} \sum_{k=1}^{z} \pi_{t_0,k}^{s}(n) p_z(x_{t_0}(n)), \tag{2.27}$$

  where $\pi_{t_0,1}^{s}(n) = \mathrm{T}^{K-k_0'}(p), \forall n \in \mathcal{C}_1$. It follows that $\pi_{t_0,k}^{s}(1) = \ldots = \pi_{t_0,k}^{s}(N), \forall k = 1, \ldots, K$. Further, since $x_t(1) = \ldots = x_t(N)$, we have $p_z(x_{t_0}(1)) = \ldots = p_z(x_{t_0}(N))$, and $\pi_{t_0,1}^{o}(1) = \ldots = \pi_{t_0,1}^{o}(N)$, and therefore, scheduling *any* of the idle channels in $\mathcal{C}_1$ achieves the same expected reward in control slot t_0, i.e.,

$$U_{t_0}(k_0')|_{\text{case1}} = W_{t_0}^{(\text{case1})}(n), \forall n \in \mathcal{C}_1, o_{t_0,1}(n) = 0. \tag{2.28}$$

Now, note that given equal idle ages on all N channels, the distribution of k_0' is identical across channels. Therefore, for all $n = 1, \ldots, N$, the values of $\pi^s_{t_0,k}(n), k = 1, \ldots, K$, $\pi^o_{t_0,1}(n)$ and $p_z(x_{t_0}(n))$ all stay unchanged if a different channel $a_t \neq a_t'$ is scheduled in the current control slot t. This implies that $U_{t_0}(k_0')|_{\text{case1}}$ is unchanged, and is thus not dependent on a_t.

- *Case 2:* $\mathcal{C}_1 = \emptyset$. In this case, $f_{t_0+1,k_0'}(n) = 0, \forall n = 1, \ldots, N$. The expected reward collected in control slot t_0 on channel n, denoted by $W^{(\text{case2})}_{t_0}(n)$, can be expressed the same as (2.27), where the channel strength belief $\pi^s_{t_0,1}(n) = \mathrm{T}^{K-k_0'}(r)$, for all $n \in \mathcal{C}_0$. Then, using the similar reasoning as in *Case 1*, we obtain that scheduling *any* of the idle channels in $\mathcal{C}_0$ achieves the same expected reward in control slot t_0, i.e.,

$$U_{t_0}(k_0')|_{\text{case2}} = W^{(\text{case2})}_{t_0}(n), \forall n \in \mathcal{C}_0, o_{t_0,1}(n) = 0. \tag{2.29}$$

  Further, the expected reward achieved in this case, $U_{t_0}(k_0')|_{\text{case2}}$, does not change when the action in current control slot t varies, i.e., $U_{t_0}(k_0')|_{\text{case2}}$ is not dependent on a_t.

- *Case 3:* $\mathcal{C}_0 \neq \emptyset, \mathcal{C}_1 \neq \emptyset$. In this case, we first show that the maximum expected reward in control slot t_0 is achieved by scheduling any of the idle channels in the set $\mathcal{C}_1$, which are perceived to have better channel fading state in the subsequent control slot t_0 than the channels in set $\mathcal{C}_0$. Specifically, picking any one of the channels from each of the set, $n_1 \in \mathcal{C}_1$ and $n_0 \in \mathcal{C}_0$, we have

$$\begin{aligned} W^{(\text{case3})}_{t_0}(n_1) &= \sum_{z=1}^{K} \sum_{k=1}^{z} \pi^s_{t_0,k}(n_1) p_z(x_{t_0}(n_1)), \\ W^{(\text{case3})}_{t_0}(n_0) &= \sum_{z=1}^{K} \sum_{k=1}^{z} \pi^s_{t_0,k}(n_0) p_z(x_{t_0}(n_0)), \end{aligned} \tag{2.30}$$

where

$$\begin{aligned}\pi^s_{t_0,k}(n_1) &= \mathrm{T}^{K-k_0'+k-1}(p),\\ \pi^s_{t_0,k}(n_0) &= \mathrm{T}^{K-k_0'+k-1}(r).\end{aligned} \tag{2.31}$$

Based on (2.3) and the property of the positively-correlated Markov chain, the following inequality holds: For all $k = 1, \ldots, K$,

$$\pi^s_{t_0',k}(n_1) = \mathrm{T}^{K-k_0'+k-1}(p) \geq \mathrm{T}^{K-k_0'+k-1}(r) = \pi^s_{t_0,k}(n_0),$$

with the equality achieved when $K \to \infty$. Further, applying similar argument as before, we obtain

$$W^{(\text{case3})}_{t_0}(n_1) > W^{(\text{case3})}_{t_0}(n_0).$$

Now, since $n_1 \in \mathcal{C}_1$ (likewise, $n_0 \in \mathcal{C}_0$) is arbitrary, from the conclusion drawn in *Case 1*, the maximal expected reward in control slot t under this case is achieved by scheduling the SU to *any* of the idle channels in the set $\mathcal{C}_1$, i.e.,

$$U_{t_0}(k_0')|_{\text{case3}} = W^{(\text{case3})}_{t_0}(n_1), \forall n_1 \in \mathcal{C}_1, o_{t_0,1}(n_1) = 0. \tag{2.32}$$

Next, based on the similar reasoning as in the previous cases, we know that $U_{t_0}(k_0')|_{\text{case3}}$ is not dependent on a_t.

Finally, note that $U_{t_0}(k_0')$ can be written as:

$$U_{t_0}(k_0') = \sum_{i=1}^{3} U_{t_0}(k_0')|_{\text{case } i}\Pr(\text{Case } i), \tag{2.33}$$

where

$$\begin{aligned}\Pr(\text{Case1}) &= \prod_{n\in\mathcal{C}_1} \pi^s_{t_0+1,k_0'}(n),\\ \Pr(\text{Case2}) &= \prod_{n\in\mathcal{C}_0} (1-\pi^s_{t_0+1,k_0'}(n)),\\ \Pr(\text{Case3}) &= \prod_{n\in\mathcal{C}_1} \pi^s_{t_0+1,k_0'}(n) \prod_{n'\in\mathcal{C}_0} (1-\pi^s_{t_0+1,k_0'}(n')),\end{aligned}$$

are quantities dependent on k_0' only. Based on the fact that the idle age at the moment are identical across the channels, the probabilities $\Pr(\text{Case } i), i = 1, 2, 3$, and the distribution $\Pr(k_0')$, are the same across the channels as well. Thus U_{t_0} is not dependent on a_t.

Since $t_0 \in \{t-1, \ldots, 1\}$ is arbitrary, by extension, we have that the total reward collected from control slot t_0 till the horizon, i.e., $\sum_{t'=t_0}^{1} U_{t'}$, which is the future reward of current control slot t, is unaffected by a_t. This concludes the proof and establishes the proposition. □

Remarks: Proposition 4 illustrates the effect of fading state belief on the optimal decisions in the genie-aided system. With ages equalized across the PU channels and the classic "exploitation vs. exploration" tradeoff neutralized (by definition of the genie-aided system), we saw that, higher fading belief favors the immediate reward and that the future reward is, in fact, unaffected by the current decision.

In the following, we study the effect of PU occupancy and age on the optimal decisions in the genie-aided system and, in turn, its impact on the original system.

2.4.3 Impact of PU Occupancy on the Tradeoffs

The following proposition identifies the effect of channel occupancy state on the optimal scheduling decisions.

Proposition 5. *When the fading state beliefs are the same across all PU channels, it is optimal to schedule the SU to the channel with the lowest idle age at the moment, i.e.,*

$$a_t^* = \arg\min_n \{x_t(n)\}. \tag{2.34}$$

Proof. We prove the proposition by showing that scheduling the channel with the lowest idle age favors: *1*) the immediate reward $R_t(\mathbf{S}_t, a_t)$; and *2*) the optimal future reward $V^*_{t-1}(\mathbf{S}_{t-1})$. The first part can be readily shown by appealing to Proposition 1 and Corollary 1. To show the second part, we adopt the following induction hypothesis:

Induction Hypothesis: The optimal future reward is convex in the fading state belief.

When channel a_t is scheduled in the current control slot t, the expected future reward can be evaluated as:

$$V^*_{t-1}(\mathbf{S}_{t-1})|_{a_t} = \pi^s_{t,k_0}(a_t)V^*_{t-1}(p) + (1-\pi^s_{t,k_0}(a_t))V^*_{t-1}(r),$$

where $V^*_{t-1}(p)$ and $V^*_{t-1}(r)$ represent the future reward calculated when the channel fading state observed in the k_0th mini-slot of control slot t is good or bad, respectively. More specifically,

$$V^*_{t-1}(p) \triangleq V^*_{t-1}\left(\pi^s_{t-1,1}(a_t) = \mathrm{T}^{K-k_0}(p)\right),$$
$$V^*_{t-1}(r) \triangleq V^*_{t-1}\left(\pi^s_{t-1,1}(a_t) = \mathrm{T}^{K-k_0}(r)\right).$$

Based on (2.3), we have, for $\gamma \in (0,1)$,

$$\mathrm{T}^L(\gamma) = (p-r)^L\gamma + r\frac{1-(p-r)^L}{1-(p-r)}, L = 0, 1, \ldots,$$

and $\lim_{L\to\infty}\mathrm{T}^L(\gamma) = \frac{r}{1-p+r} \triangleq \pi_{s|s}$, where $\pi_{s|s}$ denotes the steady-state probability of perceiving good channel fading on any of the PU channels. This indicates that, a smaller k_0, associated with a higher idle age at the moment (recall discussions in Section 2.3), results in a larger $K-k_0$ and thus a value of $\pi^s_{t-1,1}(a_t)$ closer to $\pi_{s|s}$, in which case,

$$V^*_{t-1}(\mathbf{S}_{t-1})|_{a_t} \to \pi^s_{t,k_0}(a_t)V^*_{t-1}(\pi_{s|s}) + (1-\pi^s_{t,k_0}(a_t))V^*_{t-1}(\pi_{s|s}) \triangleq V^*_{t-1}(a_t, E_{\pi_{s|s}}).$$

On the contrary, as k_0 becomes larger because of a lower idle age, the value of $\pi^s_{t-1,1}(a_t)$ deviates further away from $\pi_{s|s}$, but is closer to p (or r). Also, $\pi^s_{t,k_0}(a_t)$ gets closer to $\pi_{s|s}$. Therefore,

$$V^*_{t-1}(\mathbf{S}_{t-1})|_{a_t} \rightarrow \pi_{s|s} V^*_{t-1}(p) + (1-\pi_{s|s}) V^*_{t-1}(r) \triangleq E_{\pi_{s|s}} V^*_t(a_t).$$

Fig. 2.3 is a pictorial illustration of the above reasoning.

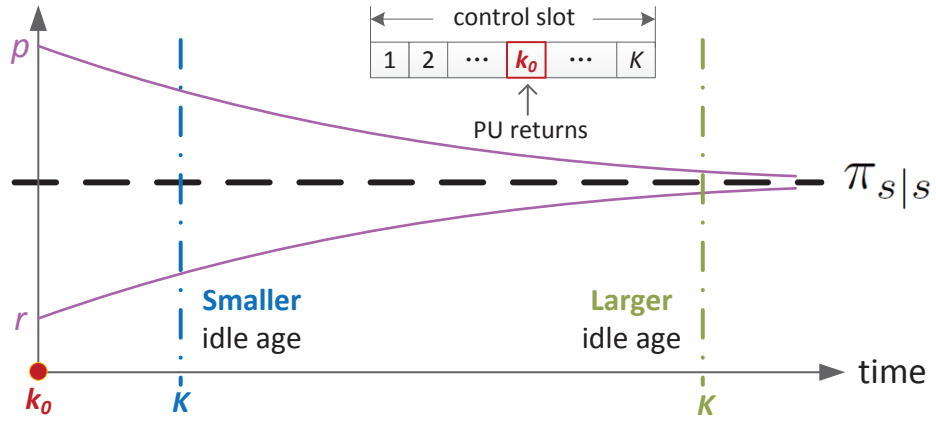

Figure 2.3: Illustration of the belief value variation and the effect of idle age.

Now, appealing to the induction hypothesis and Jensen's inequality [39], we have that $E_{\pi_{s|s}} V^*_t(a_t) > V^*_{t-1}(a_t, E_{\pi_{s|s}})$, and therefore the future reward is maximized by scheduling the channel with the lowest idle age, i.e.,

$$a^*_t = \min_{a_t}\{x_t(a_t)\} \quad s.t. \quad V^*_{t-1}(\mathbf{S}_{t-1})|_{a^*_t} = \max_{a_t}\{V^*_{t-1}(\mathbf{S}_{t-1})|_{a_t}\}.$$

Finally, we proceed to verify the induction hypothesis. At $t = 2$, the optimal future reward equals the optimal immediate reward at the horizon $t = 1$, i.e.,

$$V^*_1(\mathbf{S}_1) = R_1(\mathbf{S}_1, a^*_1) := \max_{a_1 \in \mathcal{A}_1}\{R_1(\mathbf{S}_1, a_1)\}. \tag{2.35}$$

Since $R_1(\mathbf{S}_1, a_1)$ is linear in the strength belief, using the property of convex function [39], we have that $V_1^*(\mathbf{S}_1)$ is convex in the strength beliefs, which validates the induction hypothesis. Then, using backward induction, we establish the proposition. □

Remarks: Proposition 5 explicitly illustrates the effect of the PU occupancy and age on the optimal decisions. With fading state beliefs equalized and the classic "exploitation vs. exploration" tradeoff neutralized (by definition of the genie-aided system), we saw that: *1*) lower idle age on the PU channel favors the immediate reward by allowing more idle time on the channel; and *2*) lower idle age also favors the future reward by way of better freshness of the channel fading feedback.

Thus by studying the full-observation genie-aided system, via the results in Propositions 4 and 5, we have decomposed the tradeoffs associated with the channel occupancy and the fading state beliefs in the original system. Indeed, the results in Propositions 4 and 5 rigorously support the understanding we developed earlier in Section 2.3 on the tradeoffs in the original system.

2.5 Numerical Results & Further Discussions

In this section, we evaluate and compare the optimal rewards of the original system (denoted as V_{ori}^*) and the genie-aided system (V_{genie}^*). Also, the optimal policy in the original system is compared to a *randomized* scheduling policy (with reward denoted as V_{random}^*), where the spectrum server chooses a channel, among all the idle ones, randomly and uniformly, and allocates it to the SU for data transmission. The numerical results are collected for a two channel system with $K = 2$, horizon length $m = 6$, and discounted fac-

tor $\beta = 0.9$. For notational convenience, denote $\Delta_{ga-ori} = V^*_{genie} - V^*_{ori}$ and $\Delta_{ori-rnd} = V^*_{ori} - V^*_{random}$.

Table 2.1 records the rewards obtained under various baseline cases, for various values of $\delta \triangleq p-r$, which broadly captures the temporal memory in the channel fading. In particular, as δ decreases, the channel fading memory fades and the difference between the baselines, which are primarily differentiated by the degree to which they exploit the memory in the system, tends to decrease.

Table 2.1: Comparison of rewards when the channel fading memory varies. System parameters used: $u = 1, C_I = 1, C_B = 2, x_t(1) = 10, x_t(2) = 5, \pi^s_{t,1}(1) = 0.4, \pi^s_{t,1}(2) = 0.7$

δ_{pr}	Δ_{ga-ori}	$\frac{\Delta_{ga-ori}}{V^*_{genie}}\%$	$\Delta_{ori-rnd}$	$\frac{\Delta_{ori-rnd}}{V^*_{ori}}\%$
0.8	0.0077	0.3%	0.372	14.5%
0.4	0.0068	0.29%	0.2827	9.42%
0.2	0.0023	9.8×10^{-4}%	0.1915	8.23%
0.1	0.0006	2.4×10^{-4}%	0.1836	7.93%

Next, in Table 2.2, we compare the baseline rewards under different values of the power exponent u, used in the definitions of P_I and P_B in (2.1). To build a better understanding of the trend reflected in these numerical results, we consider an arbitrary mini-slot, denoted as k', as the current mini-slot, and the following two exhaustive sets of PU occupancy histories: *1*) the set of histories $h^I_x, x = 1, 2, \ldots$, corresponds to the case when there are exactly x mini-slots between the current mini-slot k' and the most recent mini-slot (preceding k') when the channel was in busy state, i.e., the idle age of mini-slot k' is x; and *2*) the set of histories $h^B_x, x = 1, 2, \ldots$, corresponds to the case when there are exactly x mini-slots between the current mini-slot k' and the most recent mini-slot (preceding k') when the channel was in idle state, i.e.,

the busy age of mini-slot k' is x. In Fig. 2.4, we plot the two sets of occupancy histories. As has been pointed out in Section 2.2.2, the idle/busy age is a sufficient statistic for capturing the memory in the PU channels' occupancy states. Thereby, with the construction of h_x^I and h_x^B, we can examine the effect of the temporal correlation of PU occupancy on the system performance. Specifically, the conditional idle probability in mini-slot k', given occupancy history, can be obtained as:

$$\begin{aligned} \pi_{k'}^o|_{h_x^I} &= P_I(x) = \frac{1}{x^u + C_I}, \\ \pi_{k'}^o|_{h_x^B} &= 1 - P_B(x) = 1 - \frac{1}{x^u + C_B}. \end{aligned} \tag{2.36}$$

For $\pi_{k'}^o|_{h_x^I}$, it is clear, from Fig. 2.5, that as u increases, the conditional probability curves become steeper. Define the threshold point, x_0, such that for all $x > x_0$, the difference in $\pi_{k'}^o|_{h_x^I}$ is insignificant (below 10^{-2}). Now, note from the figure that the threshold x_0 decreases with increasing u, i.e., $x_0^{(u=5)} < x_0^{(u=3)} < x_0^{(u=1)}$. This indicates that the impact of different occupancy histories on the current PU occupancy state diminishes with increasing u and thus a decreased memory in the PU occupancy. Similar argument holds when considering h_x^B (see Fig. 2.6). In short, the exponent u broadly captures the PU occupancy memory, and as its value increases, the memory decreases and thus the rewards under various cases, as expected, come closer with increasing u. This is illustrated in Table 2.2.

In addition, as can be seen from both tables, the original system performs very close (within 1%) to the genie-aided system, while the cost of measuring and sending the channel fading feedback is only $\frac{1}{N}$ of the latter. Also, the optimal policy significantly outperforms the randomized policy, indicating that the temporal correlation structure in the channel fading and PU

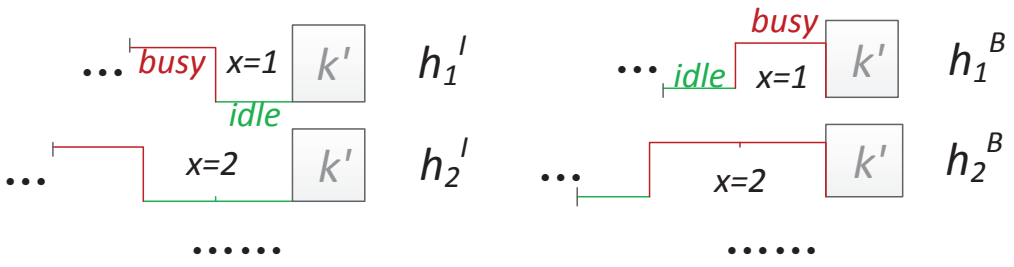

Figure 2.4: Illustration of occupancy histories h_x^I and h_x^B.

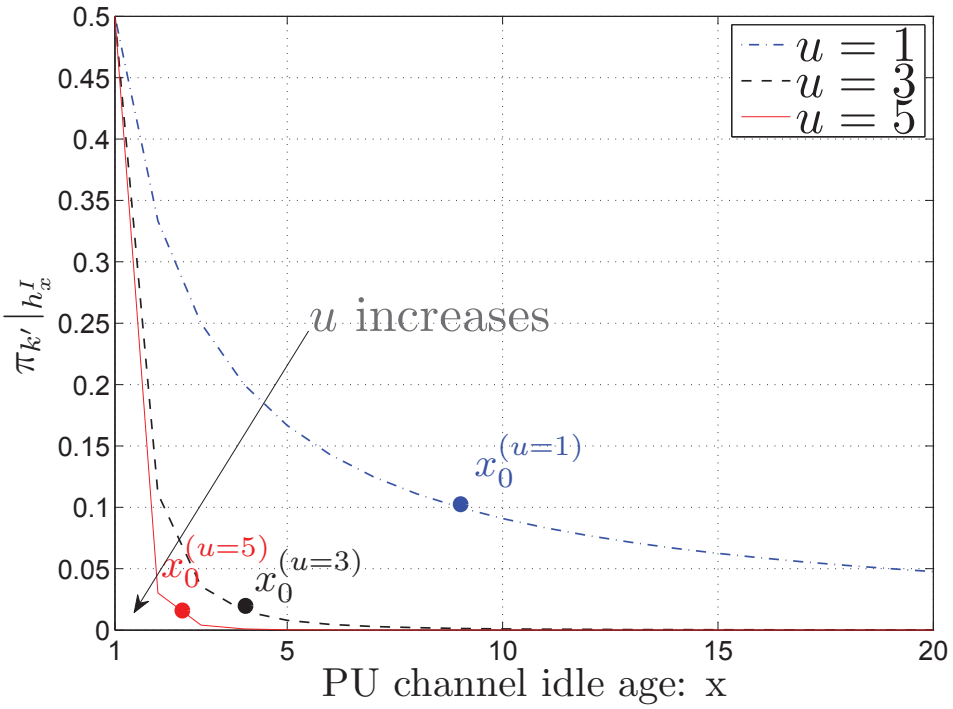

Figure 2.5: Conditional idle probability under h_x^I.

occupancy can greatly benefit the opportunistic spectrum scheduling, when appropriately exploited.

To evaluate the system performance further, Fig. 2.7 compares the optimal policy and a policy that explores memory in only the PU occupancy (with parameters $p = 0.9, r = 0.1, u = 1, C_I = 1, C_B = 2, x_t(1) = 0, x_t(2) =$

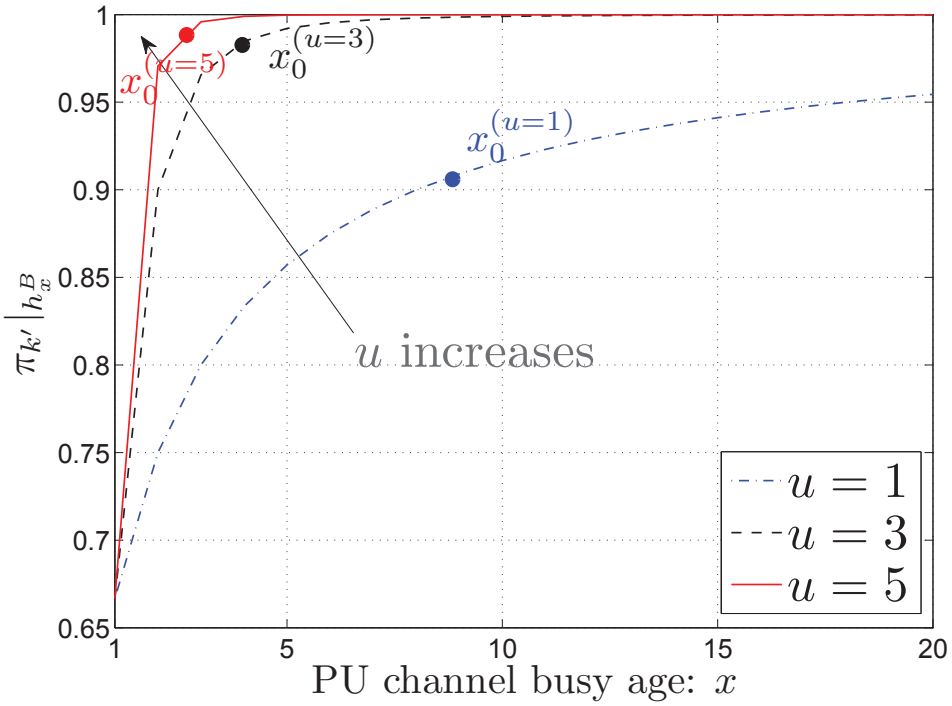

Figure 2.6: Conditional idle probability under h_x^B.

Table 2.2: Comparison of rewards when the occupancy memory varies. System parameters used: $p = 0.9, r = 0.1, C_I = 1, C_B = 2, x_t(1) = 0, x_t(2) = 1, \pi^s_{t,1}(1) = 0.4, \pi^s_{t,1}(2) = 0.7$

u	Δ_{ga-ori}	$\frac{\Delta_{ga-ori}}{V^*_{genie}}\%$	$\Delta_{ori-rnd}$	$\frac{\Delta_{ori-rnd}}{V^*_{ori}}\%$
1	0.0088	0.35%	0.3273	12.66%
3	0.006	0.26%	0.2201	9.59%
5	0.0043	0.2%	0.1839	8.28%

$5, \pi^s_{t,1}(1) = 0.4, \pi^s_{t,1}(2) = 0.7$). Clearly, a gain (of 14.3% at $m = 6$) exists when following the optimal scheduling policy, indicating the benefit of *jointly* exploring memory in both the channel fading and PU occupancy.[9]

[9]Note that it is also possible to compare the optimal policy with an algorithm obtained by exploring the memory in fading only. However, due to the long memory inherent in PU's occupancy, a fair comparison would require a large time horizon which is computationally expensive.

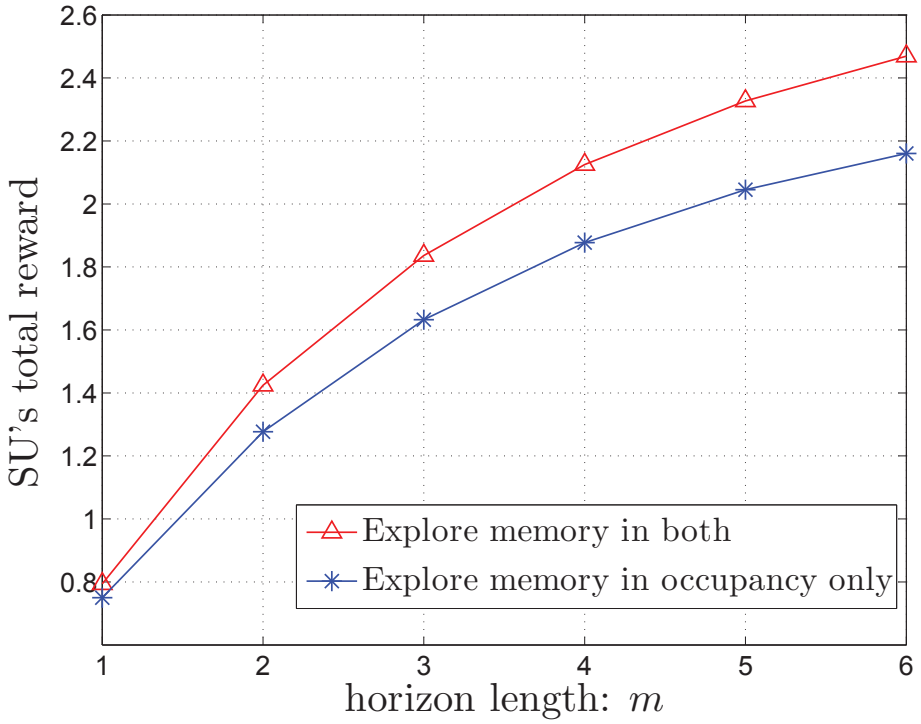

Figure 2.7: Gain of jointly exploring memory in channel fading and PU occupancy.

2.6 Conclusions

In this chapter, we studied opportunistic spectrum access for a single SU in a CR network with multiple PU channels. We formulated the problem as a partially observable Markov decision process, and examined the intricate tradeoffs in the optimal scheduling process, when incorporating the temporal correlation in *both* the channel fading and PU occupancy states. We modeled the channel fading variation with a two-state first-order Markov chain. The temporal correlation of PU traffic was modeled using "age" and a class of monotonically decreasing functions, which present a long memory. We identified a set of "multi-tier" tradeoffs arising from the intricate interactions between channel

fading and PU occupancy states, and examined such tradeoffs via quantifying both immediate and total rewards for the SU. The optimality of the simple greedy policy was established under certain conditions. Further, by developing a genie-aided system with full observation of the channel fading feedbacks, we decomposed and characterized the multiple tiers of the tradeoffs in the original system, and recorded the impacts from channel fading and PU occupancy on such tradeoffs under the general setting. Finally, we numerically studied the performance of various systems and/or policies and showed that the original system achieved an optimal total reward very close (within 1%) to that of the genie-aided system. Further, the optimal policy in the original system significantly outperformed randomized scheduling, as well as a policy that explores memory in only the PU occupancy, pointing to the merit of jointly exploiting the temporal correlation in both of the system states. The studies we have initiated here can be generalized to cases with multiple SUs and/or nonidentical PU channels, where the fundamental tradeoffs associated with the interactions among various state elements identified in this chapter will be retained. We believe that our formulation and the insights we have obtained open up new horizons in better understanding spectrum allocation in cognitive radio networks, with problem settings that go beyond the traditional ones.

Chapter 3

THE IMPACT OF INDUCED SPECTRUM PREDICTABILITY VIA WIRELESS NETWORK CODING

3.1 Introduction

Network coding is a novel networking paradigm that transforms the classical store-and-forward based routing and allows coding over packet traffic at the intermediate nodes. It has been shown that in general, network coding can improve the achievable throughput for multi-hop multicast communications with a single source to the min-cut capacity [40]. Such coding gain can be realized by linear network coding [41], and distributed implementation is possible through random linear network coding (RLNC) [42]. Network coding is not only beneficial in multi-hop communications [41–44], but also improves the throughput in single-hop broadcast channels under both backlogged and stochastic packet traffic [45–50].

Consider a source[1] multicasting to a set of receivers over lossy wireless channels. At the source, RLNC is applied so that the randomly arrived packets are put into batches, randomly coded and then transmitted to the receivers, as illustrated in Fig. 3.1. Specifically, the source buffers the packets into a batch $\mathbf{x} = [x_1, \ldots, x_m]$ of length m. Then, a vector of coding coefficients (with length m) is generated, randomly from the Galois field $GF(q)$ with field size q. Next, each coded packet y_i is formed as $y_i = \sum_{j=1}^{m} c_{ij} x_j$, $i = 1, 2, \ldots$, with the coding coefficient $c_{ij} \in GF(q)$. We assume a large value of q so that with high probability[2], each coded packet is *innovative*, and each receiver simply

[1] In this chapter, the source may be either a primary user (PU) base station in a cognitive radio (CR) network, or an arbitrary node with multicast traffic in a general wireless network.

[2] As specified in, e.g., [42, 45] and the references therein, $q \geq 64$ would suffice to make this probability arbitrarily close to 1.

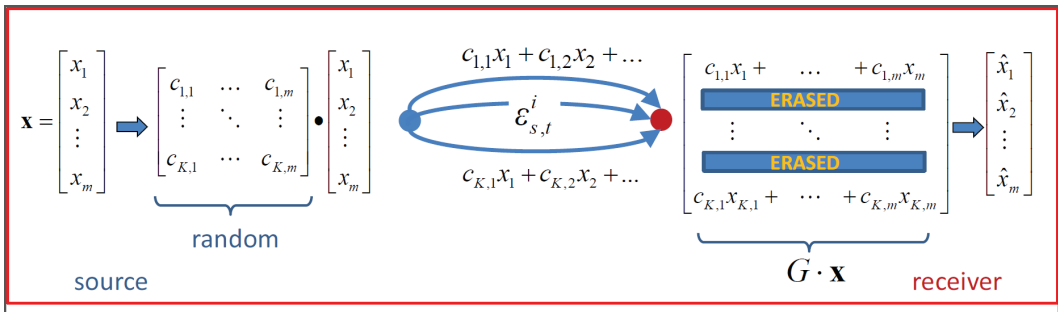

Figure 3.1: Network-coded transmissions over a lossy wireless channel.

needs to receive exactly m coded packets in order to decode the entire batch of original packets $\mathbf{x}$.

Observe that a *shaping* effect takes place in channel use under network coding[3] in the sense that the duration of each busy period is lower-bounded by the batch size m and that of the idle period depends on the process of accumulating randomly arrived packets. Compared to the traditional retransmission mechanism (e.g., ARQ) where individual packets are transmitted over the wireless channel independently, network-coded transmissions reduce the frequency of the transitions between the idle and busy periods and hence makes the channel usage more "predictable," as illustrated in Fig. 3.2.

In this study, we examine how to leverage this shaping effect in network-coded communications to infer the channel usage patterns. We first illustrate the importance of exploring this induced structure to discover spectrum holes in cognitive radio (CR) networks. Next, we also show that such a structure may cause a security threat when malicious jammers are present. Our findings lead to a new understanding of the inherent trade-offs between the throughput

[3]Such a shaping effect also applies to other block-based and rateless coding techniques, e.g., fountain codes.

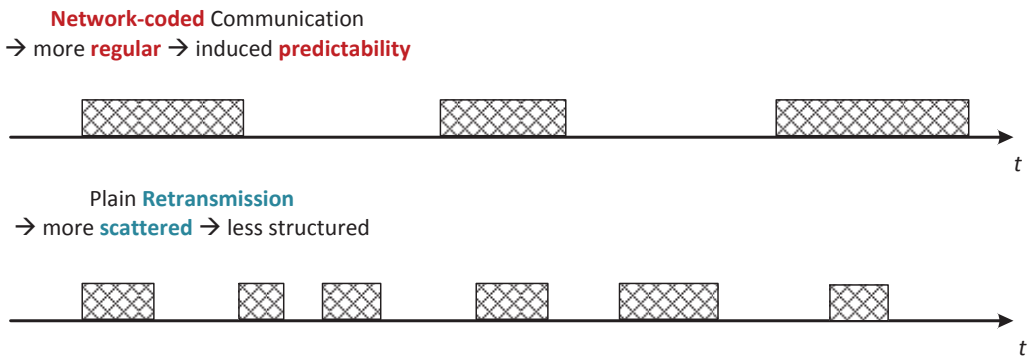

Figure 3.2: Induced structure on network-coded communication channel.

and security objectives resulting from the spectrum predictability induced by network coding.

Our first focus is on an overlay CR network[4] with multiple primary user (PU) channels, each representing a multicasting PU subnetwork consisting of one PU base station and a set of receiving PU nodes, and the packet traffic arrives randomly at the base station. Secondary users (SUs) opportunistically access these PU channels. It is clear that the PUs' throughput can be improved when network coding is applied for their data transmissions. This improved PU throughput indicates an extended availability of PU channels to the SUs since SUs are allowed to transmit over only the idle instances of PU channels. However, the problem of locating these spectrum holes remains challenging. Particularly, the random nature of packet arrivals at the primary subnetworks leads to stochastic and sporadic transmission patterns over the PU channels, and thereby "hides" the temporal spectrum holes from the SUs. Fortunately, network coding applied over the PU channels also provides potential benefits to the SUs due to the structure it induces on the primary spectrum (as Fig. 3.2

[4]Interested readers can refer to [5, 7, 51–53] and the references therein for more details regarding overlay CR networks.

indicates), i.e., this makes it easier for the SUs to "predict" the spectrum holes and find the transmission opportunities faster and more accurately. We note that this use of network coding here, as a *spectrum shaper* on the PU traffic, is similar in spirit to traffic shaping (e.g., leaky bucket [54]) where the outgoing traffic from the shaper is smoother and more regular as a result of the buffering process. In particular, network coding used here improves the throughput of both the PUs and SUs. We emphasize that the throughput gain for the SUs has two reasons: the spectrum opportunities for SUs increase (since PUs deliver their traffic faster) and SUs find these channel access opportunities easier (due to the batch transmission structure in channel use by the PUs).

With this insight, we propose an adaptive channel sensing scheme, in which each SU sets a distinct timer for every PU channel sensed to be busy and does not revisit it for a certain period of time (which depends on the batch size). This allows each SU to track a candidate list (it is also called "sensing list" in subsequent sections) of possibly idle PU channels and to perform a two-stage channel sensing: First, the SU chooses channels from the candidate list randomly and carries out sensing one by one searching for an idle channel, and in case there is no idle channel found in the candidate list, it continues to sense the rest of the PU channels. The timer evolution for the PU channels can be modeled as Markov chains that are coupled through the sensing list size. As expected, network coding applied at the PUs results in a higher throughput for SUs, and adaptive channel sensing outperforms random sensing. Next, we also consider joint channel sensing and channel contention for the case with multiple SUs. In particular, we develop an adaptive sensing scheme with two-level backoff, where detecting a PU transmission and colliding with peer SUs are treated separately.

Related work on spectrum sensing typically focuses on exploiting the geographic and temporal properties of the PU's signal power (e.g., energy, cyclostationary signal characteristics, and interference) such that the SUs can reliably transmit without interfering with the PU transmissions [7, 11, 55–61]. Different from the physical layer-centric sensing (either non-cooperative or cooperative [62,63]), our focus here is on the channel selection for spectrum sensing. Any physical layer spectrum sensing technique (e.g, energy detection [7], cyclostationary feature detection [7], etc.) can be applied here, and we assume that the sensing is perfect in the sense that the SU can accurately identify the PU's transmissions on a selected channel. Note that network coding applied at the PUs helps save the overall channel sensing effort of the SUs. However, different from the network coding-aided collaborative sensing [62], where the SUs disseminate the sensing information via network coding, in our work, each SU independently senses the PU channels and learns the network-coded traffic pattern, which is used for the future decisions on channel selection.

We should also caution that the spectrum predictability brought by network-coded communications may make the transmissions more susceptible to attacks from malicious users. Hereby, we examine the possible pitfalls of the predictability brought by network coding in a general wireless network with multiple channels. We consider the case where a jammer can observe one channel at a time and can jam the transmissions as long as it is active subject to its energy constraint. Recall that the busy periods are more structured under network-coded transmissions, which can be exploited by the jammer. We show that predictability-aware greedy attacks are indeed more detrimental compared to random attack under the retransmission-based communications. This points to the disadvantage of the spectrum predictability in face of a

malicious jammer and also points to the necessity of a careful design of system parameters to balance the throughput performance and security considerations under jamming attacks.

The rest of the chapter is organized as follows. In Section 3.2, we present a CR network model with multiple PU channels, and introduce the notion of spectrum shaping by network coding. We develop in Section 3.3 an adaptive sensing scheme for the SUs in the CR network, and evaluate its performance in Section 3.4. In Section 3.5, we extend the study to multiple SUs, where an adaptive sensing scheme with two-level backoff is proposed. Next, we show in Section 3.6 that the memory induced by network coding to the communication channel may pose a security threat when facing a malicious jammer. Finally we conclude the chapter in Section 3.7.

3.2 Induced Predictability via Network Coding in Cognitive Radio Networks

3.2.1 CR Network Model

We consider a CR network with N PU channels, and focus on the spectrum shaping effect of network coding, in terms of the structural properties of spectrum holes. We will then develop spectrum sensing methods for the SUs to discover and utilize the available spectrum resources. For ease of exposition, we first consider the case with one SU. Once the PU spectrum holes are identified, the next step of coordinating multiple SU transmissions can be handled via medium access control, and we will consider in Section 3.5 joint spectrum sensing and channel access for multiple SUs.

We assume that each PU channel is associated with a PU subnetwork consisting of one PU base station (BS) and L ($L \geq 2$) receivers (as shown in

Fig. 3.3). Within each PU subnetwork, the BS multicasts data packets to the receiving nodes. We assume that the time is slotted and synchronized among PU channels. In each slot, packets arrive at any PU channel, $j = 1, ..., N$, according to a stationary arrival process with a common rate λ. PU channels are lossy, with packet erasure probability ε. We consider independent erasures across channels and time slots.

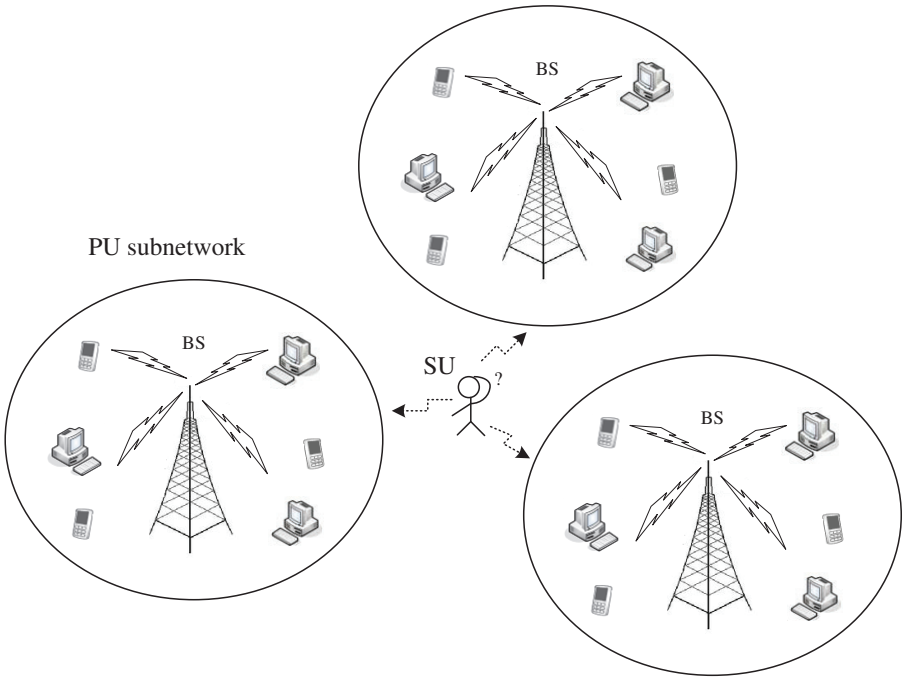

Figure 3.3: A CR network with multiple PU channels.

First, consider that the PUs' transmissions are carried out using network coding. Specifically, each BS accumulates a batch of m packets, encodes them via RLNC, and multicasts the coded packets to the receivers. Once all receivers of a PU subnetwork decode the entire batch of m packets, the BS proceeds with the next batch of m packets, provided that there are enough buffered packets. If not, the BS waits to accumulate the next batch of packets. In contrast, when the traditional retransmission mechanism is used, each

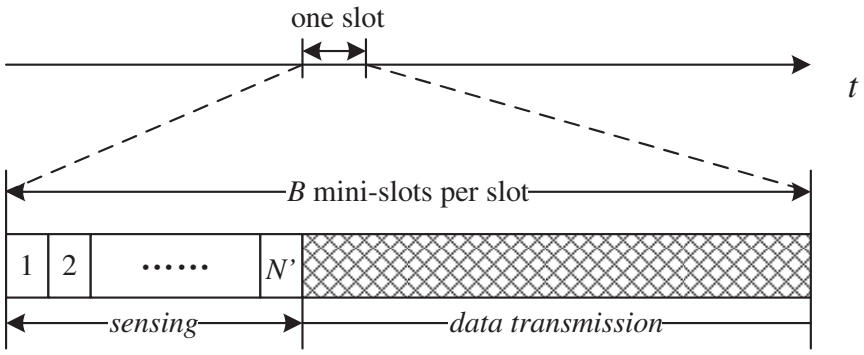

Figure 3.4: SU's slot structure for channel sensing and data transmission.

BS multicasts individual packets one by one independently. The BS keeps retransmitting each packet until all receivers successfully receive it, before transmitting the next packet.

In dynamic spectrum access, the SU opportunistically explores spectrum holes on PU channels for data transmissions. We assume synchronization across the SU and PUs. The packet transmission of a PU continues one slot (or time frame) and each slot of the SU amounts to B mini-slots, which are used either for sensing different PU channels or transmitting packets. Clearly, at most B channels can be sensed per slot, i.e., the SU would be able to sense all PU channels when needed only if $N \leq B$. Sensing is performed based on the "sensing list" $\mathcal{N}_t$, which consists of the channels that are considered by the SU as "possibly idle" in each slot. When sensing, the SU picks channels randomly and uniformly, one at a time without replacement. After channel sensing, the SU transmits data packets during the rest of the time slot, provided that an idle channel is detected; otherwise, the SU waits until the next slot and repeats the same steps. Fig. 3.4 illustrates the SU's slot structure, where N' indicates the end of the sensing phase when the SU decides to transmit.

3.2.2 Spectrum Shaping Effect via Network Coding

Let D_t represent the number of mini-slots used for sensing in slot t by the SU. Define a "good slot" to be the one in which the SU finds an idle channel and transmits. Correspondingly, a slot is called "bad" if the SU fails in obtaining any transmission opportunity therein. Let $\mathbf{1}_t$ be the indicator random variable indicating whether slot t is good ($\mathbf{1}_t = 1$) or bad ($\mathbf{1}_t = 0$). Since the SU can transmit data packets only over $B - D_t$ mini-slots (whenever SU finishes sensing) and only if the channel sensing is successful (i.e., it detects an idle PU channel), the SU's throughput can be obtained as

$$\begin{aligned} \eta_s &= \lim_{T_{tot}\to\infty} \frac{\sum_{t=1}^{T_{tot}} (B - D_t)\mathbf{1}_t}{T_{tot}} \\ &= E[(B - D_t)\mathbf{1}_t] = Bp_r - E[D_t\mathbf{1}_t], \end{aligned} \tag{3.1}$$

where T_{tot} is the time period (in number of slots) under observation, $p_r = \Pr(\mathbf{1}_t = 1)$ is the probability that the SU finds an idle PU channel in slot t, and equation (3.1) follows from the ergodicity of the channel sensing process. Clearly, the smaller D_t is, the larger the SU's throughput. In other words, the SU needs to find an idle channel (if any) as soon as possible in every slot. Under the assumption that only one channel can be observed at a time, the channel selection is critical to the SUs' throughput. Comparing different channel usage structures under network coding and retransmission schemes (illustrated in Fig. 3.2), we expect that the SU can reduce its sensing overhead by exploiting the underlying traffic structure on the PU channels. Specifically, when retransmission is used by the PUs, the PU channels experience fast idle-busy alternations. However, when network coding is used by the PUs, the transmission is performed on a batch basis with size m, and the transitions between the idle and busy states would become slower on the scale of m slots.

That is, network coding makes the spectrum structure more "regular" and hence easier for sensing.

With this insight, it is plausible for the SU to differentiate the PU channels and keep a shorter prioritized list for sensing (i.e., the "sensing list") when the PUs use network coding, i.e., we have $N_t \leq N$, for all t, where $N_t = |\mathcal{N}_t|$ is the size of the sensing list. In particular, we will show that the sensing list can be updated dynamically when network coding is used at the PUs, and this would yield a higher throughput for the SU.

3.2.3 PU's Throughput and Idle Probability

For completeness, we first summarize in the following the gain brought by network coding to the PUs. Denote by T_{NC} and $T_{0,RT}$ the completion time for one batch (when network coding is applied) and one packet (when retransmission is used), respectively.

Random network coding with large q ensures that each of the L PU receivers can decode the original m packets with high probability, as long as it receives exactly m coded packets [42]. Based on [45], the expected completion time is given by

$$E[T_{NC}] = m + \sum_{t=m}^{\infty} \left[1 - \left(\sum_{a=m}^{t} \binom{a-1}{m-1} (1-\varepsilon)^m \varepsilon^{a-m} \right)^L \right]. \tag{3.2}$$

The stability condition of the PU queues is given by $\lambda < \eta_p^{NC}$, where the maximum stable throughput η_p^{NC} is

$$\eta_p^{NC} = \frac{1}{E[T_{NC}]/m}. \tag{3.3}$$

From Little's theorem [64], the idle probability of each PU BS can be obtained as

$$P_{idle}^{NC} = 1 - \frac{\lambda E[T_{NC}]}{m}. \tag{3.4}$$

On the other hand, when retransmission is used by the PUs, the completion time of any individual packet is the number of time slots necessary for successful reception of it at all receivers. From (3.2) with $m = 1$, the expected service time is obtained as

$$E[T_{0,RT}] = \sum_{t=1}^{\infty} \left(1 - \left(\sum_{a=1}^{t} (1-\varepsilon)\varepsilon^{a-1} \right)^{L} \right) + 1. \tag{3.5}$$

The stability condition of the PU queues is given by $\lambda < \eta_p^{RT}$, where the maximum stable throughput η_p^{RT} is

$$\eta_p^{RT} = \frac{1}{E[T_{0,RT}]}. \tag{3.6}$$

It follows that the idle probability of the PU channels is

$$P_{idle}^{RT} = 1 - \lambda E[T_{0,RT}]. \tag{3.7}$$

As a result, we have $E[T_{NC}] < mE[T_{0,RT}]$ and therefore

$$\eta_p^{RT} < \eta_p^{NC}, \quad P_{idle}^{RT} < P_{idle}^{NC}, \tag{3.8}$$

As expected, network coding increases the stable throughput for PUs, which is the direct gain for the PUs by using network coding. Further, it provides the SU with extended availability of the primary spectrum due to the enlarged average idle probability at the PU channels. As we show in the subsequent sections, network coding in fact "shapes" the spectrum and increases the predictability of whether PU channels are idle or not. This in turn reduces the channel sensing time by the SU, and hence further improves the SU's throughput.

3.3 Exploiting Induced Predictability for Spectrum Sensing

3.3.1 Backoff-based Adaptive Sensing

When network coding is used by the PUs, the SU can predict "slower" transitions between the PUs' idle and busy states due to the additional buffering and batch transmissions of coded packets. Instead of keeping all N channels in the sensing list, the SU seeks for a shortened list to reduce the time it takes to find an idle channel. In each time slot, the SU carries out a two-stage sensing. First, the SU senses PU channels randomly from the sensing list $\mathcal{N}_t$ one by one, and stops whenever an idle channel is detected. In the meanwhile, the SU backs off on channels sensed to be busy for k slots (by setting up a countdown timer) and updates the list $\mathcal{N}_t$ every slot based on the channel sensing history.

If all channels in the first stage are found to be busy, the SU proceeds to the second stage and randomly searches for an idle channel in the backup list $\overline{\mathcal{N}}_t$, namely the list of channels excluding $\mathcal{N}_t$, with size $|\overline{\mathcal{N}}_t| = N - N_t$, until it finds one. Note that the sensing capability can be further improved by ordering channels in the back-up list according to their timer values and by giving priority to those channels with smaller timer values in the sensing order. However, this would increase the complexity significantly. Instead, we consider random channel selection (without replacement) from the backup list in the second stage.. In this stage, the SU moves the channel that it senses idle back to the sensing list, but does not change the timers of the channels that are detected to be busy. The detailed algorithm is summarized in **Algorithm 1**.

The intuition behind this two-stage sensing approach is to give higher sensing priority to the first stage with channels that are more likely to be idle,

Algorithm 1 Backoff-based Adaptive Sensing Under Network Coding

Initialization:
The SU keeps a sensing list $\mathcal{N}_t$ and a backup list $\overline{\mathcal{N}}_t$ for slot t.
Repeat in every slot:
$\Omega \rightarrow \emptyset$ (Ω is a local auxiliary set, consisting of channels sensed to be busy in the 1st stage).
for mini-slot $= 1 : B$ **do**
  if $\Omega \neq \mathcal{N}_t$ **then** {*1st stage sensing*}
    The SU chooses channel $j \in \mathcal{N}_t \backslash \Omega$ randomly and uniformly, and senses it.
    if channel j is idle **then**
      The SU transmits until the end of the slot.
    else
      The SU augments $\Omega \rightarrow \Omega \cup j$ and waits until the next mini-slot.
    end if
  else {*2nd stage sensing*}
    The SU chooses channel $j' \in \overline{\mathcal{N}}_t$ randomly and uniformly without replacement, and senses it.
    if channel j' is busy **then**
      The SU waits until the next mini-slot.
    else {Denote j' as j^*}
      The SU transmits until the end of the slot.
    end if
  end if
end for
At the end of the slot, the SU
a) Sets up countdown timers $\rightarrow k$ (slots), for all channels $j \in \Omega$. (The timers start clicking from the next slot.)
b) Clears timer of channel $j^* \rightarrow 0$, effective in the next slot and moves j^* back to the sensing list.
c) Collects channels $j'' \in \overline{\mathcal{N}}_t$ with timer value 1 (i.e., those who will have value 0 and be moved back to the sensing list in the next slot) and forms a set $\Theta = \cup j''$.
d) Updates the two channel lists:
  $\mathcal{N}_t \rightarrow \mathcal{N}_t \cup j^* \cup \Theta \backslash \Omega$,
  $\overline{\mathcal{N}}_t \rightarrow \overline{\mathcal{N}}_t \cup \Omega \backslash j^* \backslash \Theta$.

and then proceeds beyond the priority list, if needed. The backoff parameter k can be expressed as a function of the batch size m, namely $k = g(m), g : \mathbb{Z}^+ \to \mathbb{Z}^+$. Without prior information (i.e., when the SU visits a busy channel for the first time), on average the SU would find the PU transmissions in the middle of the batch service time. Therefore, we can intuitively let $k = \frac{E[T_{NC}]}{2}$. For any channel, the SU should choose as the backoff timer k the average remaining time it believes for the PU to complete the transmission of the current batch. We will evaluate the effects of k on the spectrum sensing and throughput performance in Section 3.4.

3.3.2 Throughput Analysis for Adaptive Sensing

In adaptive sensing, the SU dynamically updates the sensing list based on the sensing results from the current slot. It is clear that the sensing list size N_t is a random variable. For notational convenience, denote $p_n = \Pr(N_t = n)$. Note that the SU cannot sense all N channels when B is smaller than N, and accordingly, two cases shall be considered separately: $B \geq N$ and $B < N$.

3.3.3 The Case with $B \geq N$

The SU's throughput is intimately related to sensing at both of the two stages. Clearly, only if the SU fails in the first stage, i.e., all channels in the sensing list are found to be busy, the SU will proceed into the second stage. Accordingly, the probability $p_r = \Pr(\mathbf{1}_t = 1)$ can be obtained as

$$\begin{aligned} p_r &= \sum_{n=0}^{N} p_n \left\{ \left(1 - (1 - P_{idle}^{NC})^n\right) + \left((1 - P_{idle}^{NC})^n (1 - (1 - P_{idle}^{NC})^{N-n})\right) \right\} \\ &= 1 - (1 - P_{idle}^{NC})^N, \end{aligned} \tag{3.9}$$

where the first summand inside the braces denotes the probability that at least one channel in the sensing list (i.e., the first stage) is idle, and the second

summand denotes the probability that given the first stage fails, the SU can find an idle channel in the backup list (i.e., the second stage).

Further, the number of mini-slots used for finding an idle channel, provided that there exists at least one idle PU channel, is computed as

$$E[D_t\mathbf{1}_t] = \sum_{n=0}^{N} \frac{1-(1-P_{idle}^{NC})^n(1+nP_{idle}^{NC})}{P_{idle}^{NC}}\Big(1-(1-P_{idle}^{NC})^n\Big)p_n + \sum_{n=0}^{N}\sum_{d=n+1}^{N} dP_{idle}^{NC}(1-P_{idle}^{NC})^{(d-n)-1}\Big(1-(1-P_{idle}^{NC})^{N-n}\Big)(1-P_{idle}^{NC})^n p_n. \tag{3.10}$$

It can be seen that the SU's throughput depends on the distribution p_n, which is closely related to the probability that any given PU channel is in the sensing list. In what follows, we characterize this probability, denoted as π_0.

Recall that once a PU channel is sensed to be busy, the SU backs off on this channel by setting up a countdown timer with an initial value k. The evolution of such a timer follows a Markov chain, where states $0, ..., k$ correspond to the countdown values of the timer, or equivalently, the remaining time slots before the channel is moved back to the sensing list $\mathcal{N}_t$. The value of the timer decreases from k to 0 at the rate of 1 per slot. When it enters state 0, the corresponding channel is considered as a potential candidate in the sensing list again, and it remains in the list as long as it is sensed to be idle or it is simply not sensed. On the contrary, the channel is removed from the sensing list (to the backup list) in the next slot, provided that it is sensed to be busy. In the meantime, if the SU enters the second stage and senses the channels therein, the state of the idle channel it finally finds changes to 0 in the next slot. Thereby, the Markov chain for the timer should include transitions from states $i = 2, ..., k$ to state 0, as depicted in Fig. 3.5. Accordingly, the

transition probabilities can be expressed as

$$\begin{aligned} p_{i,i-1} &= 1 - p_b P_{idle}^{NC}, \quad i = 2, ..., k, \\ p_{i,0} &= p_b P_{idle}^{NC}, \quad i = 2, ..., k, \\ p_{1,0} &= 1, \\ p_{0,k} &= p_s(1 - P_{idle}^{NC}), \\ p_{0,0} &= p_s P_{idle}^{NC} + (1 - p_s), \end{aligned} \tag{3.11}$$

where p_s is the probability that a PU channel in the sensing list is sensed in the first stage, and p_b represents the probability that a PU channel in the backup list is sensed when the second stage sensing is needed.

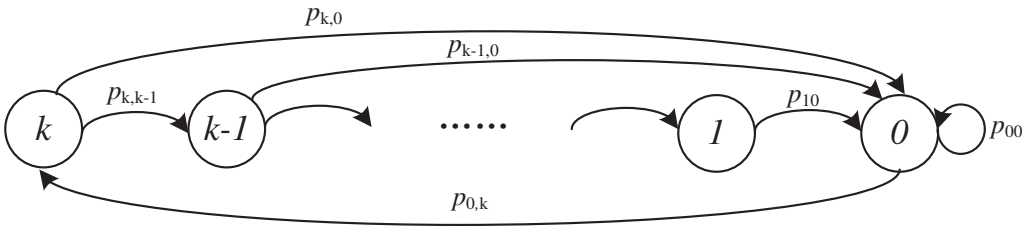

Figure 3.5: Markov chain for the timer of a given PU channel.

Recall that the PU channels in both lists are randomly chosen and sensed, until an idle channel is detected. The Markov chains of the timers for different PU channels are coupled with each other through the size of the sensing list N_t, which in turn depends on the states of the individual Markov chains, namely how many of them are in state 0. An approximation (which we validated via simulation, see Fig. 3.14 - Fig. 3.17 and related elaboration in Section 3.4.2) is necessary to integrate the effects of all PU channels (similar approach has been taken in e.g., [65], to analyze the IEEE 802.11 backoff mechanism). In particular, we average the sensing probability of the first

stage, p_s, and obtain that

$$p_s = \sum_{n=1}^{N} \sum_{x=0}^{n-1} \prod_{x'=0}^{x-1} \left(1 - \frac{1}{n-x'}\right) (1 - P_{idle}^{NC})^x \frac{1}{n-x} p_n. \tag{3.12}$$

Along the same line, the sensing probability of the second stage, p_b, can be characterized as

$$p_b = \sum_{n=0}^{N-1} (1 - P_{idle}^{NC})^n p_n \sum_{y=0}^{l'-1} \prod_{y'=0}^{y-1} \left(1 - \frac{1}{l'-y'}\right) (1 - P_{idle}^{NC})^y \frac{1}{l'-y}, \tag{3.13}$$

where $l' = N - n$.

The stationary distribution for the states $i = 0, 1, ..., k$ in the Markov chain is given by

$$\begin{aligned} \pi_0 &= \pi_0 p_{0,0} + \pi_1 + \sum_{i=2}^{k} \pi_i p_{i,0}, \\ \pi_{i-1} &= \pi_i (1 - p_b P_{idle}^{NC}), \quad i = 2, ..., k, \\ \pi_k &= \pi_0 p_{0,k}, \\ \sum_{i=0}^{k} \pi_i &= 1. \end{aligned} \tag{3.14}$$

Based on (3.11)-(3.14), π_0, the probability that any given PU channel is in the sensing list $\mathcal{N}_t$, can be computed as

$$\pi_0 = \left(1 + p_s(1 - P_{idle}^{NC}) \left(1 + \frac{(1 - p_b P_{idle}^{NC})(1 - (1 - p_b P_{idle}^{NC})^{k-1})}{p_b P_{idle}^{NC}}\right)\right)^{-1}. \tag{3.15}$$

It follows that the size of the sensing list has a binomial distribution with parameter π_0, i.e., $p_n = \Pr(N_t = n)$ is expressed as

$$p_n = \binom{N}{n} (\pi_0)^n (1 - \pi_0)^{N-n}. \tag{3.16}$$

Let $\mathbf{p} = [p_0, p_1, ..., p_N]$. We note that (3.16) consists of a fixed point equation of form $\mathbf{p} = T(\mathbf{p})$ for $\mathbf{p}$. Since $T : [0,1]^{N+1} \to [0,1]^{N+1}$ is a continuous mapping

on a compact space, there exists a solution to $\mathbf{p} = T(\mathbf{p})$. Furthermore, $\mathbf{p}$ consists of the stationary distributions of the ergodic irreducible Markov chain that represents the overall N PU channel states, it follows that there exists a unique solution to (3.16).

3.3.4 The Case with $B < N$

When $B < N$, the modeling of Markov chains and the corresponding analysis follow similarly from the case with $B \geq N$. It should be noted that only a subset of channels (B out of N) can be sensed at most in this case. Accordingly, we have that

$$\begin{aligned} p_r &= \sum_{n=0}^{B} \Pr(\mathbf{1}_t = 1 | N_t = n) p_n + \sum_{n=B+1}^{N} \Pr(\mathbf{1}_t = 1 | N_t = n) p_n \\ &= \sum_{n=0}^{B} \Bigg(\Big(1 - (1 - P_{idle}^{NC})^n \Big) + (1 - P_{idle}^{NC})^n \Big(1 - (1 - P_{idle}^{NC})^{B-n} \Big) \Bigg) p_n \\ &\quad + \sum_{n=B+1}^{N} \Big(1 - (1 - P_{idle}^{NC})^B \Big) p_n, \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} E[D_t \mathbf{1}_t] &= \sum_{n=0}^{B} E[D_t \mathbf{1}_t | N_t = n] p_n + \sum_{n=B+1}^{N} E[D_t \mathbf{1}_t | N_t = n] p_n \\ &= \sum_{n=0}^{B} \Bigg(\sum_{d=1}^{n} d P_{idle}^{NC} (1 - P_{idle}^{NC})^{d-1} \Big(1 - (1 - P_{idle}^{NC})^n \Big) \\ &\qquad + \sum_{d=n+1}^{B} d P_{idle}^{NC} (1 - P_{idle}^{NC})^{(d-n)-1} (1 - P_{idle}^{NC})^n (1 - (1 - P_{idle}^{NC})^{B-n}) \Bigg) p_n \\ &\quad + \sum_{n=B+1}^{N} \sum_{d=1}^{B} d P_{idle}^{NC} (1 - P_{idle}^{NC})^{d-1} \Big(1 - (1 - P_{idle}^{NC})^B \Big) p_n, \end{aligned} \tag{3.18}$$

where p_n is given by (3.16), with p_s and p_b changed to

$$p_s = \sum_{n=1}^{N} \sum_{x=0}^{x_0} \prod_{x'=0}^{x-1} \left(1 - \frac{1}{n-x'}\right) (1 - P_{idle}^{NC})^x \frac{1}{n-x} p_n,$$

$$p_b = \sum_{n=0}^{B-1} (1 - P_{idle}^{NC})^n p_n \sum_{y=0}^{(B-n)-1} \prod_{y'=0}^{y-1} \left(1 - \frac{1}{l'-y'}\right) (1 - P_{idle}^{NC})^y \frac{1}{l'-y} \quad (3.19)$$

for $x_0 = \min(B, n) - 1$ and $l' = N - n$. The throughput of SU can be then characterized based on (3.1).

3.3.5 Adaptive Sensing with Random Backoff

So far, we have studied the SU's throughput under adaptive sensing where the same backoff parameter k is chosen for all channels joining the backup list. In Section 3.4, we will evaluate the optimal k^* that achieves the maximum performance gain of adaptive sensing. Nevertheless, in practical systems, such k^* may be difficult to obtain offline for the SUs due to their limited access to the system parameters. In addition, the potential variations in those parameters over time may restrain the online calculation of k^* because of its high complexity and overhead. Under these circumstances, we may consider an alternative scheme where the backoff parameter k is a random variable, rather than fixed. The SU still keeps two lists (i.e., the sensing list and the backup list) and performs the two-stage sensing as before. One main difference from the previous scheme is that the SU now sets timers (for the PU channels sensed to be busy) with initial values chosen from a range $k \in [k_1, k_2], k_2 > k_1$, where $k_1, k_2 \in \mathbb{N}$, i.e., k are not necessarily identical across the backup channels. The Markov chain for the timer is updated accordingly and the probability π_0 can be then recalculated for both cases with $B \geq N$ and $B < N$.

3.4 Performance Evaluation for Adaptive Sensing

3.4.1 Prediction Accuracy of Spectrum Opportunities

In the backoff-based sensing strategy, the SU predicts the spectrum holes by adaptively updating the sensing list. The SU's throughput heavily hinges on the distribution of the sensing list size N_t, where the key parameter is the idle probability of the PU channels. In the actual underlying system, this probability is given by P_{idle}^{NC}, while in the *predicted* system built from the adaptive sensing strategy, the expected idle probability on the PU channels is represented by π_0. In order to examine the prediction accuracy, we define the following distance δ to quantify the difference between the two:

$$\delta = |\pi_0 - P_{idle}^{NC}|. \tag{3.20}$$

Intuitively, the prediction becomes more accurate with a smaller δ. To get a more concrete sense, we plot in Fig. 3.6 & Fig. 3.7 some examples on the comparison of π_0 and P_{idle}^{NC}, where we set $L = 20, m = 8, \lambda = 0.4, \varepsilon = 0.2$ and $k = 4$ as the default values. As can be seen from Fig. 3.6, the prediction π_0 closely tracks the idle probability P_{idle}^{NC} of the actual system for different erasure probabilities ε, indicating the robustness of channel tracking against channel variations. On the other hand, as Fig. 3.7 demonstrates, when the backoff parameter k increases, the difference δ first shrinks sharply and then increases slowly after k reaches a certain value. This points to an optimal backoff parameter:

$$k^* = \arg\min_k |\pi_0 - P_{idle}^{NC}|, \tag{3.21}$$

for capturing the spectrum holes. Intuitively, if k is chosen to be smaller than the optimal one, the sensing list would be longer than necessary with

redundant PU channels that are actually busy. On the other hand, if k is greater than (3.21), the SU tends to perform a conservative sensing policy with a shorter list of candidate channels to be sensed.

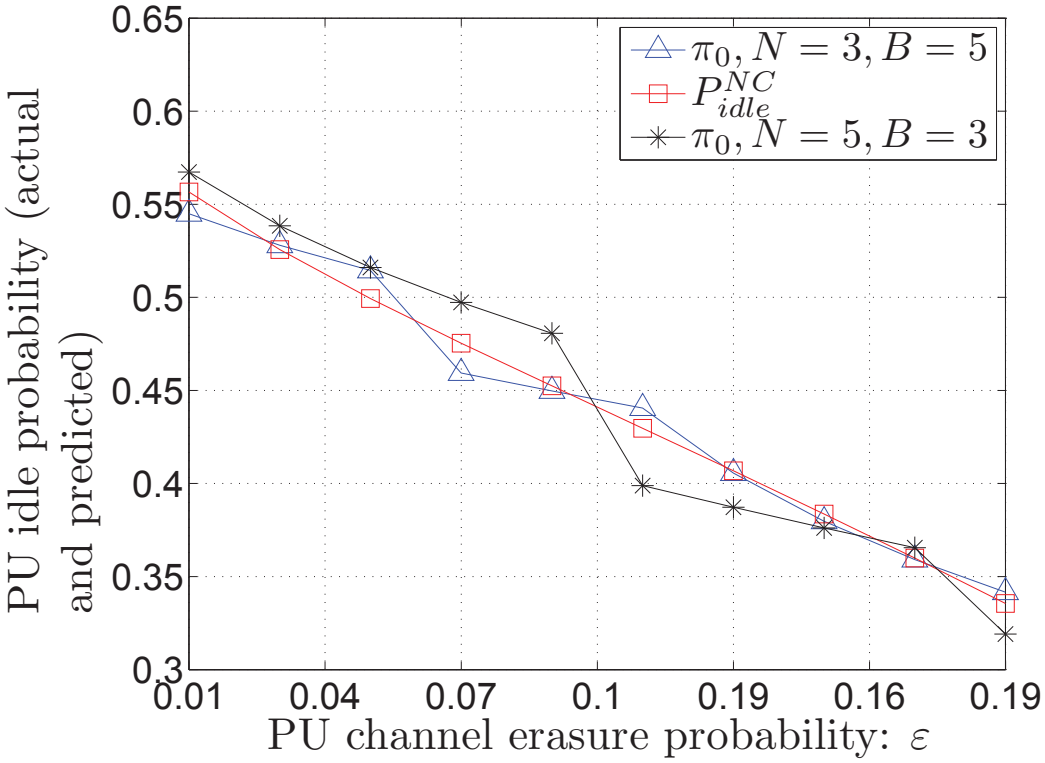

Figure 3.6: Predicted vs. actual PU idle probability.

3.4.2 SU's Throughput Gain under Network-coded PU Transmission

For the sake of comparison, we also study random channel sensing for the SU (under both cases when PUs use network coding or retransmission), where the SU senses one channel at a time, picked randomly and uniformly from all N PU channels without replacement, and continues this process, until it finds an idle channel and transmits on this channel using the remaining time of the slot. Intuitively, random sensing is more appropriate under the retransmission-based PU traffic where the channels exhibit fast alternations between the idle

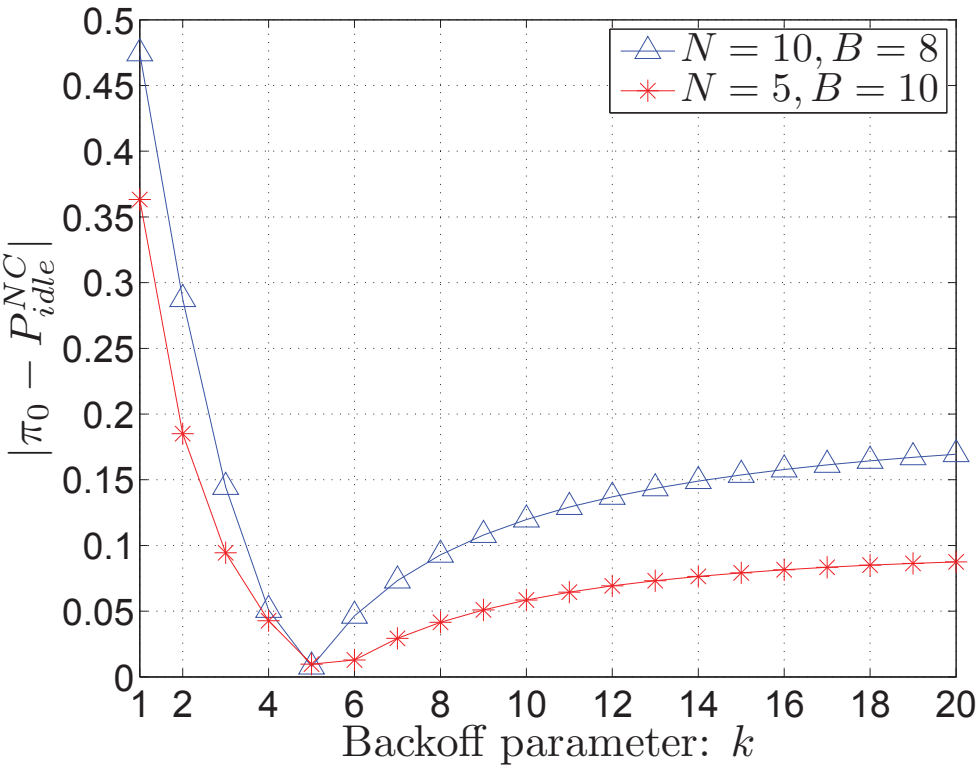

Figure 3.7: Prediction accuracy measured by $|\pi_0 - P_{idle}^{NC}|$.

and busy states. Nevertheless, as we show below, even with the same simple random sensing used by the SU, when PUs apply network coding, the SU's throughput can be significantly improved.

We illustrate, via numerical examples, the throughput gain of the SU when PUs apply network coding. First, Fig. 3.8 and Fig. 3.9 show that compared to the case when PUs use retransmission, the SU's throughput is significantly increased under network-coded PU traffic, where in either cases the SU simply adopts random sensing. Clearly, for any ε and λ, the SU's throughput improves with network coding, for both cases with $B \geq N$ and $B < N$. Further, as ε or λ increases, this throughput gain also increases. This set of results corroborate the first reason for SU's throughput improvement, namely

the spectrum opportunities for SUs increase on average since the PUs deliver their traffic faster with network coding.

The next example shows that by employing the adaptive sensing strategy, the SU's throughput can be *further* improved by almost 15%, as Fig. 3.10 - Fig.3.13 indicate, where the default values of parameters are $L = 20, m = 2, \lambda = 0.4, \varepsilon = 0.2$ and $k = 2$. The two-stage adaptive sensing strategy provides a further gain beyond random sensing, for all choices of parameters m, λ, ε, and k. Figures 3.10 and 3.11 show that this gain increases with λ and ε, and as illustrated in Fig. 3.12, the gain strongly depends on the network coding batch size as well. Moreover, Fig. 3.13 shows that the backoff parameter can be further adjusted by the SU to improve the gain of adaptive sensing. These observations here point to the second reason for SU's throughput gain, i.e., the SU can find the spectrum opportunities faster by exploiting the predictability structure on the PU channel brought by network coding.

Remarks: (1) As either λ or ε increases, the idle probability of the PU channels P_{idle}^{NC} decreases and the adaptive sensing scheme, in which the SU makes better use of the system information, reveals more gain over the random sensing scheme, where P_{idle}^{NC} dominates the performance.

(2) As m increases, the service rate of the PUs increases to the min-cut capacity given by $1 - \varepsilon$ [45]. This increases the idle probability of the PU channels for given arrival rates and thus shrinks the operational difference between the random and adaptive sensing strategies.

(3) Careful selection of the backoff parameter k improves the performance of the adaptive sensing strategy, as observed in Section 3.4.1. On one hand, we have $\pi_0 = 1$ for $k = 0$, i.e., the random sensing strategy serves

as a special case of the adaptive sensing scheme for $k = 0$. On the other hand, as k increases, almost all PU channels are included in the backup list and the adaptive sensing strategy performs again closely to the random sensing scheme. Therefore, we expect an optimal k to be an intermediate value yielding the highest performance gain (as Fig. 3.13 indicates).

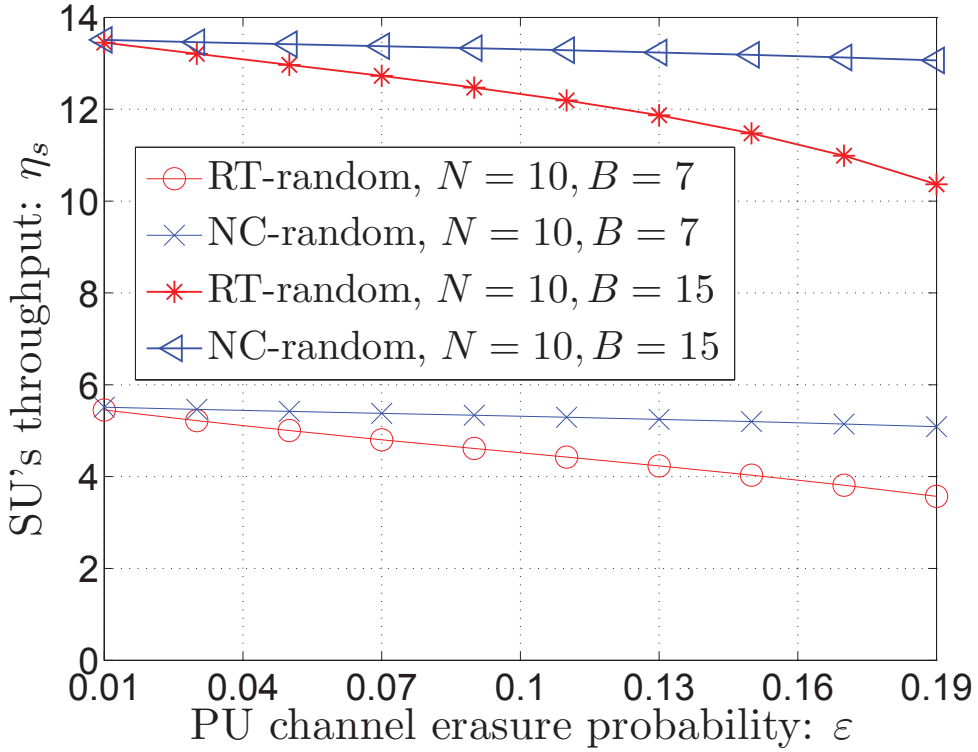

Figure 3.8: SU's throughput under random sensing: vs. ε.

Next, we take a closer look at the adaptive sensing scheme. As pointed out in Section 3.3, we approximate the sensing probabilities p_s and p_b in the Markov chain analysis and compute the SU's throughput accordingly. To validate this approximation, we perform Monte Carlo simulations (over 100 independent trials, each containing 10^4 time slots) and compare them with the numerical evaluation of the analysis, as shown in Fig. 3.14 - Fig. 3.17. The

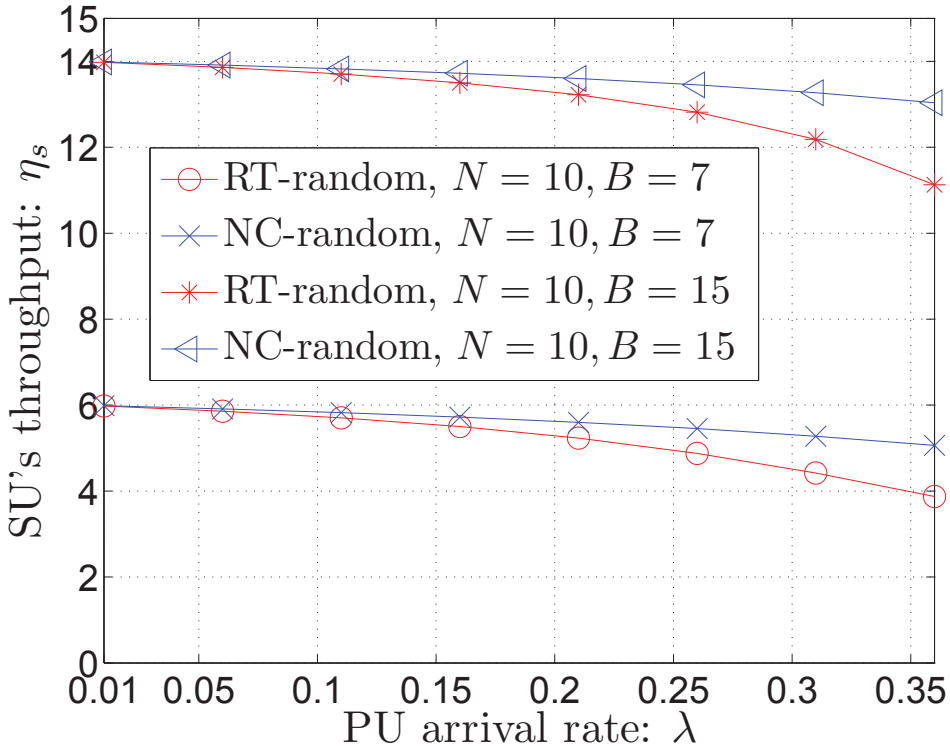

Figure 3.9: SU's throughput under random sensing: vs. λ.

default parameters are $L = 20, m = 5, \lambda = 0.4, \varepsilon = 0.1$ and $k = 2$. For better illustration, we also plot confidence intervals at the confidence level of 95% for the simulation results. We observe that the numerical results from the analysis match well with the simulations under variations in all parameters. On average, the maximum difference between the numerical and simulation results is less than 3%, indicating that the approximation in the analysis of adaptive sensing is appropriate.

Finally, Fig. 3.18 illustrates, via a few examples, the additional gain brought by the adaptive sensing with random backoff, compared to random sensing. It is indicated that adaptive sensing with random backoff achieves the same additional gain (over random sensing) as the fixed backoff case when

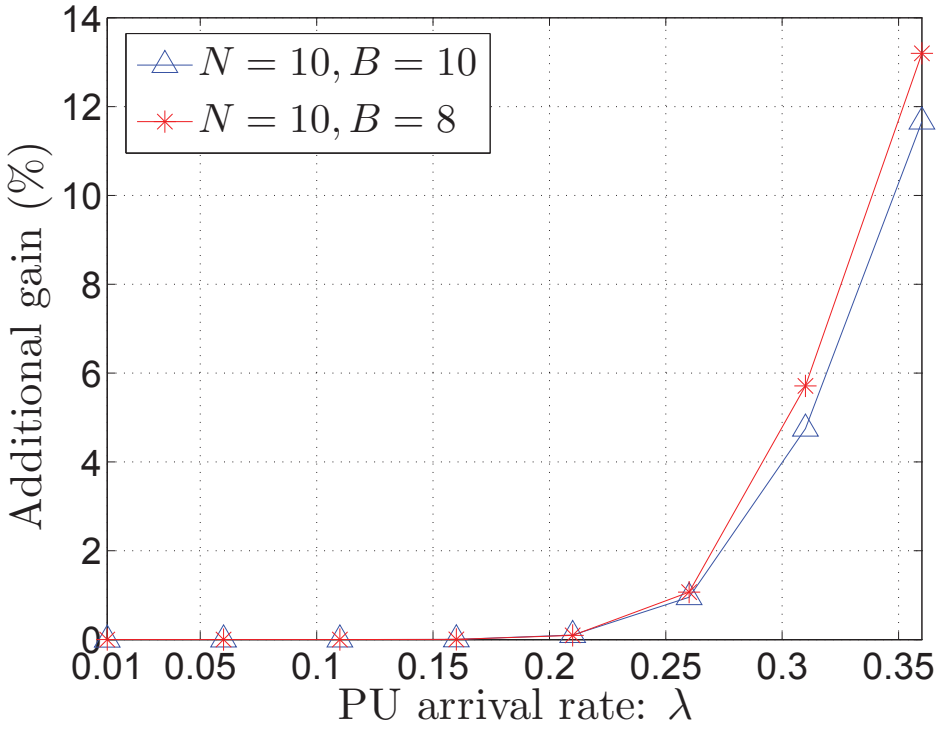

Figure 3.10: Additional gain (%) of adaptive sensing over random sensing: vs λ.

$k_2 = k_1 = 2$ (where the value 2 was the optimal k^* obtained in the previous evaluation for the fixed backoff case). As the range of k, i.e., $k_2 - k_1$ expands, the maximum gain decreases and this supports the intuition that the probability to pick the "best" k^* is now reduced to $\frac{1}{k_2 - k_1 + 1}$ due to the randomized choice of k.

3.5 Adaptive Sensing and Channel Access for the Case with Multiple SUs

In this section, we consider the case where multiple SUs (M of them) contend with each other to access the spectrum, and evaluate the aggregated system performance with the coupled SUs. In particular, building on the adaptive sensing studied in Section 3.3.1, we develop a two-level backoff scheme for

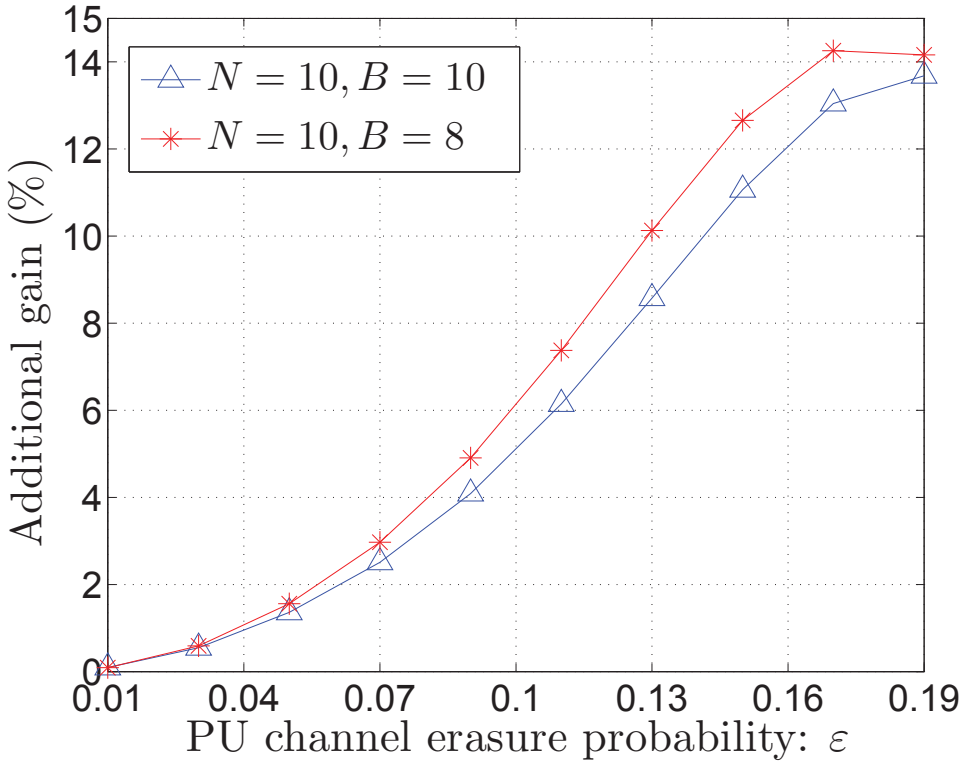

Figure 3.11: Additional gain (%) of adaptive sensing over random sensing: vs ε.

the SUs and study the average throughput of the SUs therein. For ease of exposition, we consider the special case where $B = 1$, i.e., only one channel is observed by each SU per time slot. We assume a collision channel, where a transmission is successful only if one node (PU or SU) is transmitting.

The two-level backoff strategy is outlined as follows. In every slot, each SU (with a nonempty sensing list, denoted as $\mathcal{N}_t^{(u)}$ for user $u, u = 1, \ldots, M$) chooses a channel randomly and uniformly, and senses it. If this channel is idle, the SU continues probing it with probability p_c; otherwise, it backs off and moves the channel to its backup list with probability 1 (which we

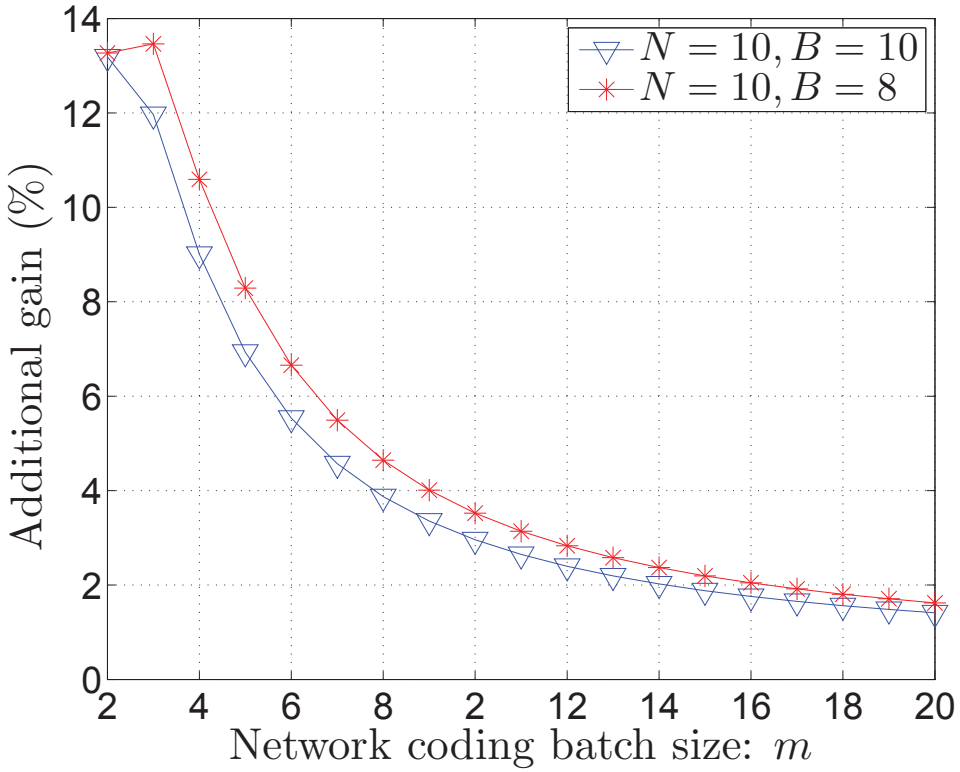

Figure 3.12: Additional gain (%) of adaptive sensing over random sensing: vs m.

call as the first level backoff). The SU can successfully seize the spectrum hole and transmit if no other SUs contend for the same channel at the same time; otherwise, each of the SUs who collided on the same channel moves this channel to its backup list with an identical probability q_b (this is the second level backoff). All channels that enter the backup list carry an initial timer value of k, counting down at the rate of 1 per slot. When the sensing list of one SU is empty, it randomly and uniformly chooses a channel from its backup list. The chosen channel will be moved back to the sensing list only if it is sensed to be idle and no other SUs contend for it simultaneously; otherwise, it stays in the backup list with the timer value unchanged.

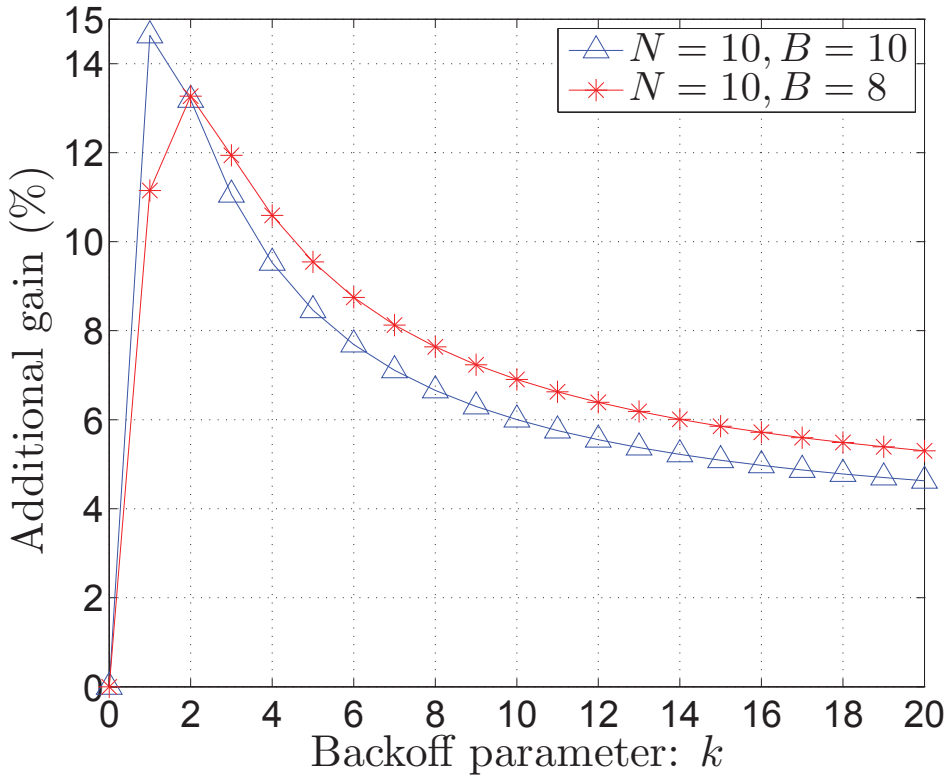

Figure 3.13: Additional gain (%) of adaptive sensing over random sensing: vs k.

It follows from the system model that the average throughput of the SUs can be obtained as

$$\begin{aligned}\eta_{s,avg}^{NC} &= \Pr(\mathbf{1}_t^{(u)} = 1) \\ &= P_{idle}^{NC} p_c \sum_{\substack{c_u \in \mathcal{N}_t^{(u)} \neq \emptyset, or \\ c_u \in \{1,\ldots,N\}, o.w.}} \frac{\prod_{l:c_l = c_u} (1 - p_c)}{N_t^u \mathbf{1}_{\mathcal{N}_t^{(u)} \neq \emptyset} + N \mathbf{1}_{\mathcal{N}_t^{(u)} = \emptyset}},\end{aligned} \tag{3.22}$$

where the event that $\mathbf{1}_t^{(u)} = 1$ is defined as the case when any SU can find a PU channel that is both idle and not contended by the other SUs; c_u represents the PU channel that user u chooses; and N_t^u is the sensing list size of user u. Clearly, based on (3.22), the characterization of the SUs' throughput requires both the distributions of $\{N_t^u\}$ for all $u = 1, \ldots, M$, and the channel identities (i.e., which channels) within each user's sensing list. However, these two

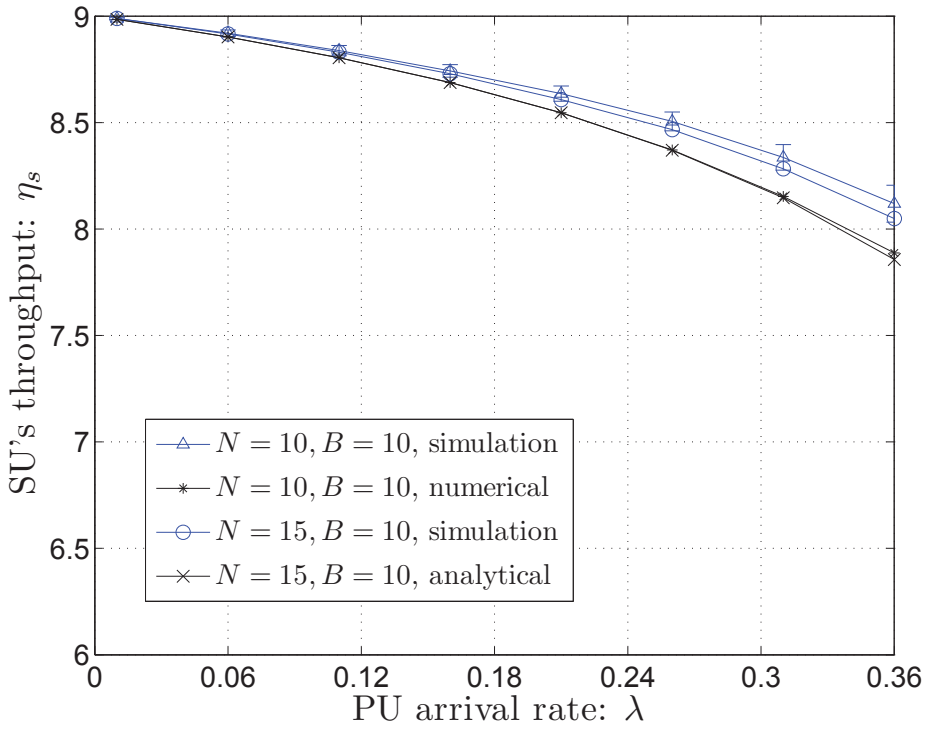

Figure 3.14: SU's throughput under adaptive sensing: vs λ.

quantities are time-varying and heterogeneous across the SUs in general, and more importantly, they evolve jointly across SUs due to the strong coupling that follows from the channel contention.

The exact analysis requires constructing a Markov chain with N^M states to model the coupled sensing lists across SUs, which is not tractable in this case. We next evaluate the throughput of the SUs under the proposed scheme via simulation, and compare that to the case when SUs apply random sensing under the retransmission-based PU communications. The av-

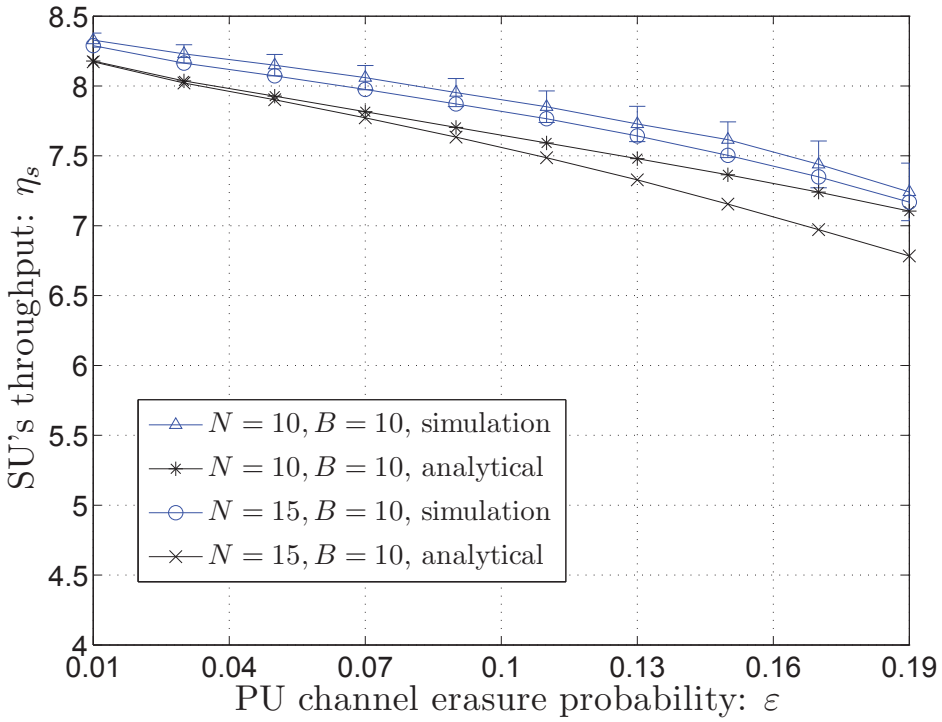

Figure 3.15: SU's throughput under adaptive sensing: vs ε.

erage throughput of the SUs in the latter case can be obtained as

$$\begin{aligned} \eta_{s,avg}^{RT} &= \sum_{j=1}^{N} \frac{1}{N} \sum_{u=1}^{M-1} \binom{M-1}{u} \left(\frac{1}{N}\right)^{u} \left(1-\frac{1}{N}\right)^{(M-1)-u} (1-p_c)^u p_c P_{idle}^{RT} \\ &= p_c P_{idle}^{RT} \left(1-\frac{p_c}{N}\right)^{M-1}. \end{aligned} \tag{3.23}$$

The optimal contention probability p_c that maximizes $\eta_{s,avg}^{RT}$ is given by $p_c^* = \min\left\{\frac{N}{M}, 1\right\}$ from the Karush-Kuhn-Tucker conditions for the optimality. In addition, it is clear that an upper-bound to (3.22) is given as $\eta_{s,avg}^{NC} < P_{idle}^{NC}$. Fig. 3.19 - Fig. 3.22 depict the average throughput of the SUs (obtained through Monte Carlo simulation over 10^5 time slots, with $M = 3$, $N = 10$ and $q_b = 0.5$). We set p_c to be the optimal value p_c^* and thus $\eta_{s,avg}^{RT}$ achieves the maximum value. As expected, by applying adaptive sensing (when the

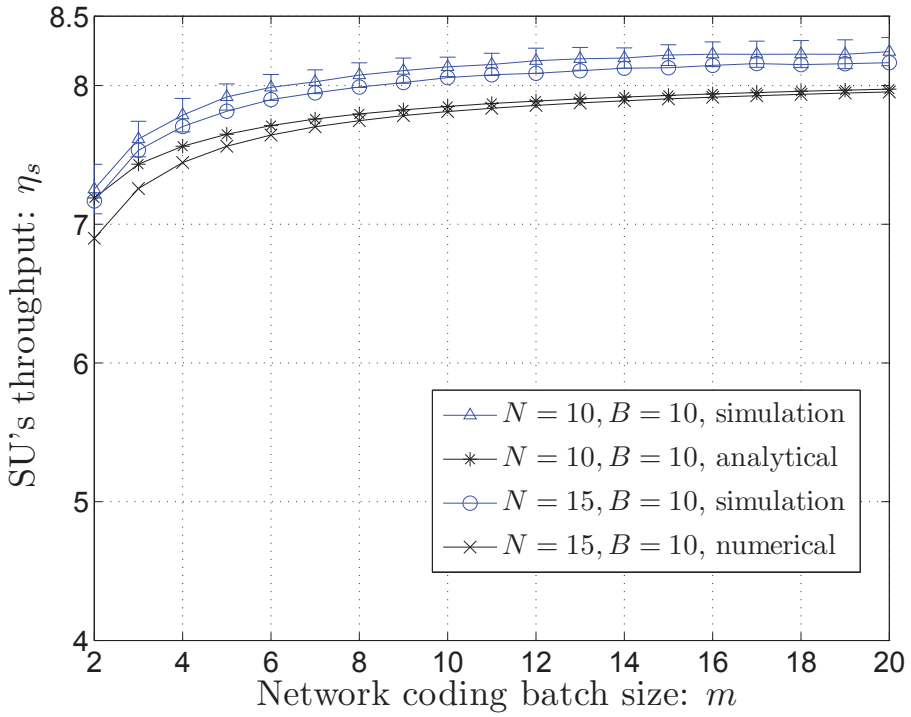

Figure 3.16: SU's throughput under adaptive sensing: vs m.

PUs use network coding), the SUs achieve a significant gain in throughput for the entire range of parameter variations (in λ, ε, m and L), over the case when random sensing is used by the SUs (with PUs applying retransmission). This again shows that network coding improves the throughput of the SUs, as the SUs exploit the memory on the PU spectrum induced by network coding.

3.6 Security Implications of Network Coding: The Case with Jamming Attacks

It can be seen from the previous sections that for dynamic spectrum access in a CR network, network coding is beneficial to the PU and SU transmissions due to its shaping effect on the PU channel and the resulting predictability. In this section, we study the potential drawbacks of network-coded transmissions

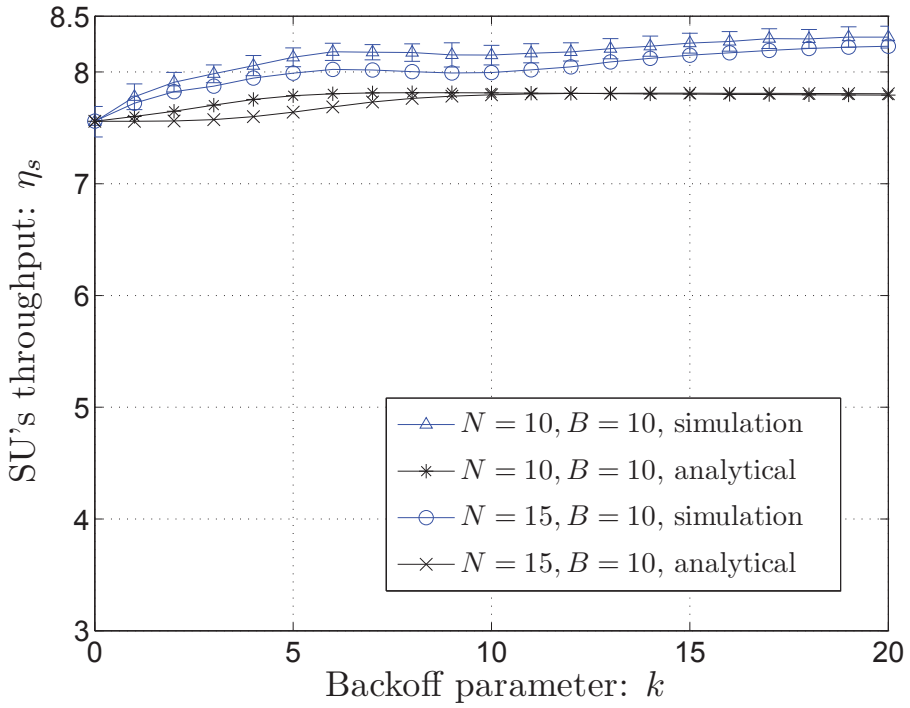

Figure 3.17: SU's throughput under adaptive sensing: vs k.

from the security perspective, when facing a malicious jammer. Particularly, instead of the specific CR network setting, we consider a general wireless network with multiple channels carrying out intended multicasting transmissions. First, we examine the predictability on the communication channel in terms of the entropy of the channel busy period. Next, we discuss different jamming strategies for the malicious user under network-coded or retransmission-based communications, and show that network coding may make the transmissions more susceptible to jamming because the channel occupancy patterns become more predictable.

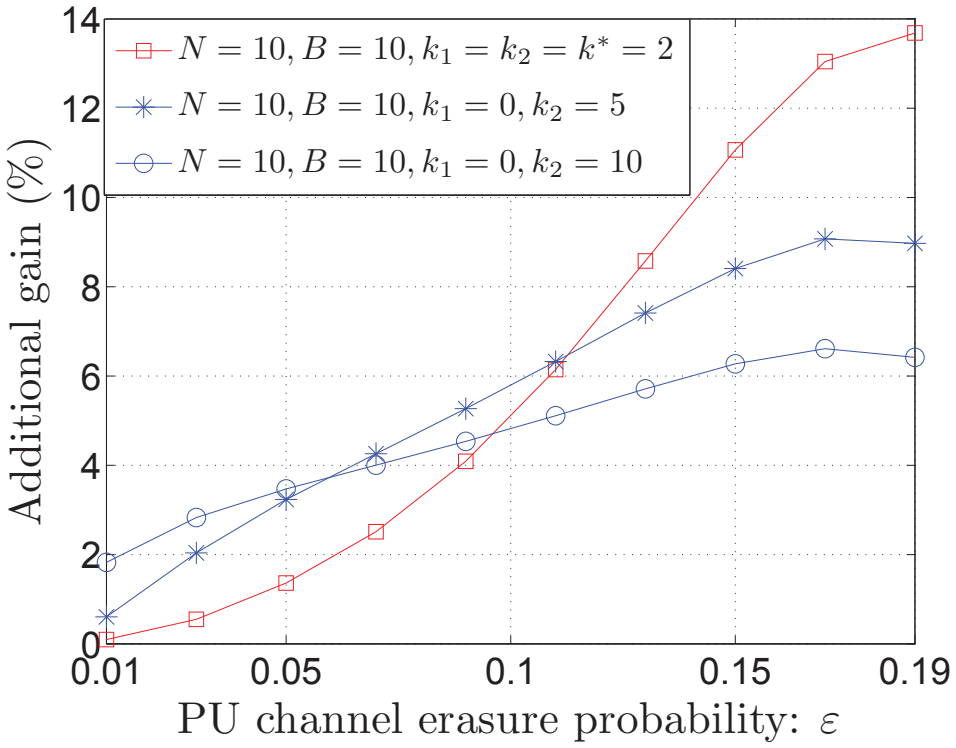

Figure 3.18: Additional gain (%) over random sensing by adaptive sensing with random backoff.

3.6.1 Entropy of Busy Period and Channel Predictability

As the adaptive sensing strategy implies, the characteristics of a channel's busy periods are closely related to its predictability. In this study, we use the entropy of the busy period to quantify channel predictability.

Denote T_b as the random variable for the length of a busy period on any channel. The entropy of the busy period is $H(\mathbf{Q}) = -\sum_t q_t \log(q_t)$, where $q_t \triangleq \Pr(T_b = t)$. To get a more concrete sense, we examine the entropy via numerical studies (collected over 10^6 time slots on one channel, with $L = 10$ and $\lambda = 0.35$). As Fig. 3.23 indicates, the entropy of the channel's busy period

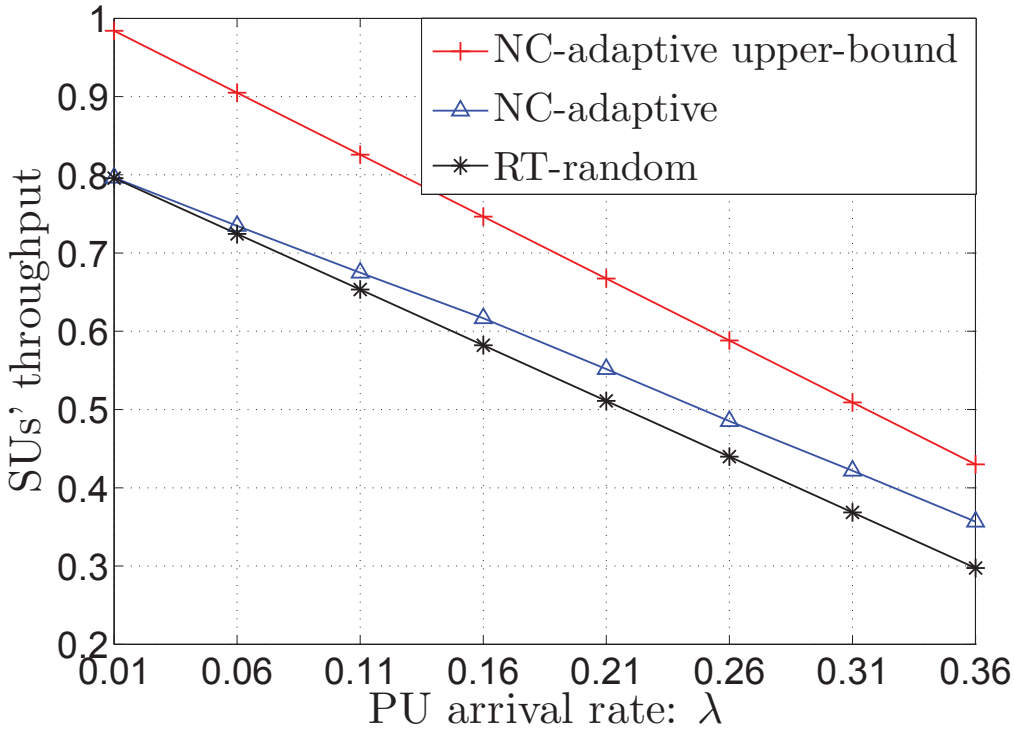

Figure 3.19: SUs' throughput under adaptive sensing with two-level backoff.: vs λ.

is smaller when network coding is applied, compared to retransmission-based communications. In addition, as m increases, the entropy decreases, indicating a more predictable transmission pattern and thus higher vulnerability to the potential attacks.

3.6.2 Attacker Model and Jamming Strategies

Energy is an important constraint for the activity of a wireless jammer [66], [67]. In the following, we consider an energy-constrained jammer that needs to recharge every T_a time slots, which we define as one attack period. In every slot, the jammer first senses a chosen channel, and then attacks it only if a transmission is detected. It is clear from the model that the average number

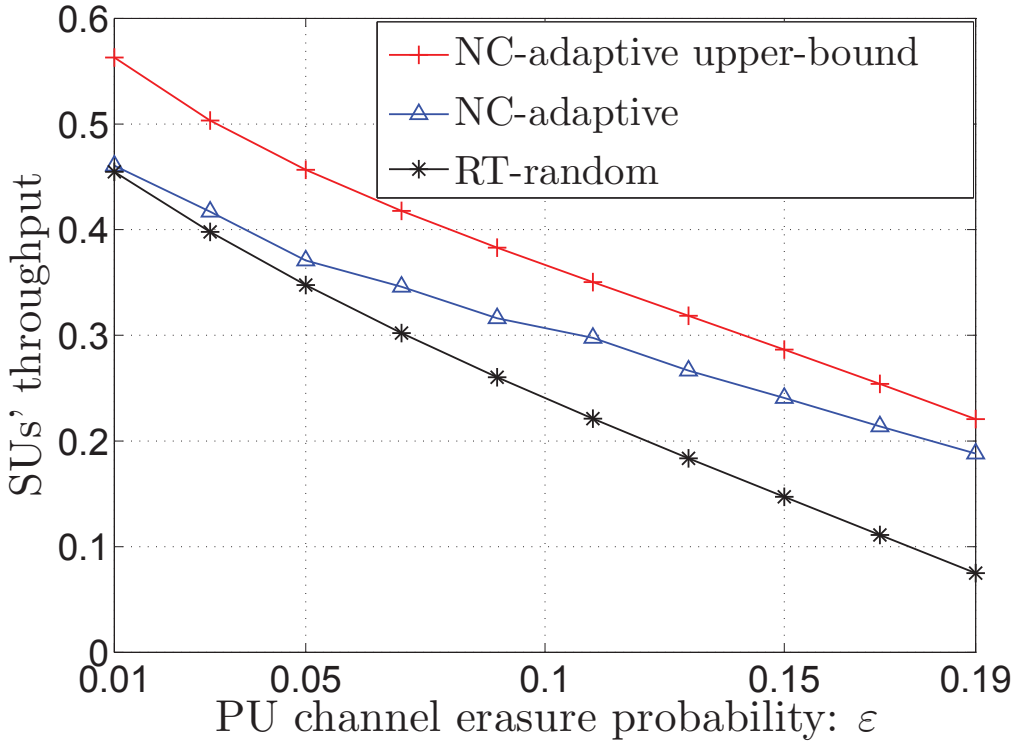

Figure 3.20: SUs' throughput under adaptive sensing with two-level backoff.: vs ε.

of jammed transmissions can be calculated as

$$\overline{U}_a = \sum_{t=1}^{T_a} \zeta_t, \tag{3.24}$$

where ζ_t is a Bernoulli random variable indicating whether the jamming attack is successfully launched in slot t ($\zeta_t = 1$), or not ($\zeta_t = 0$).

As expected, Section 3.6.1 confirms the intuition that the busy period of the network-coded transmissions is less random, compared to that under retransmission. Bearing the similar intuition as the SUs in a CR network, the jammer would also differentiate the two communication scenarios and apply different jamming strategies respectively. In particular, when network coding is applied by the intended transmission, the jammer would expect the channel

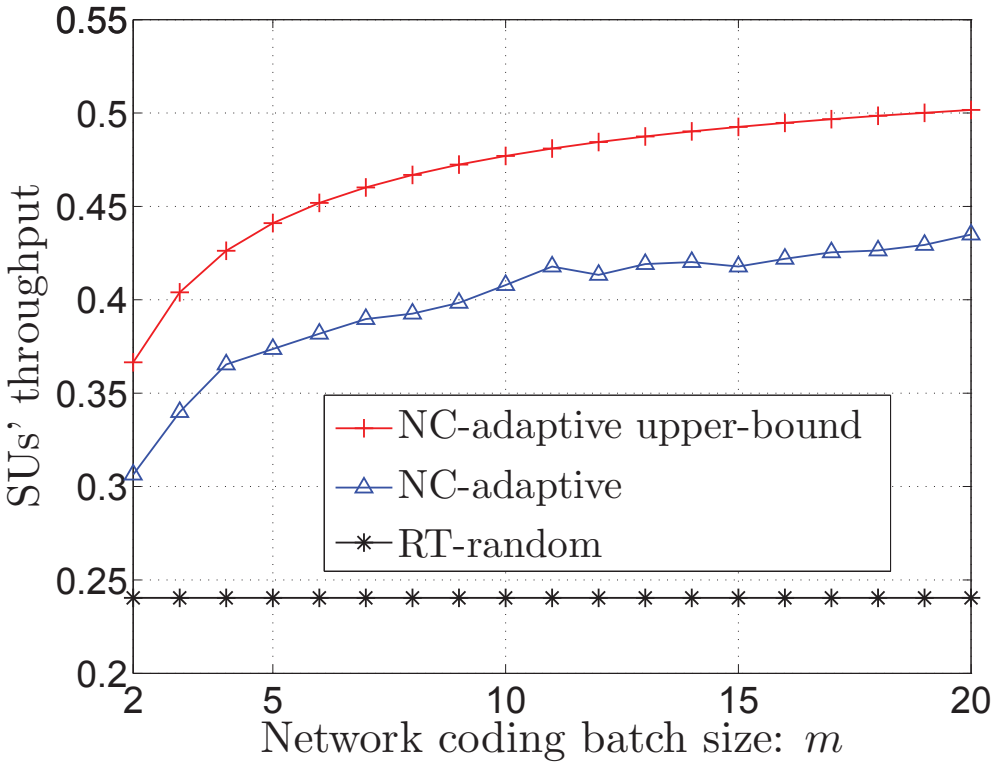

Figure 3.21: SUs' throughput under adaptive sensing with two-level backoff.: vs m.

to be busy for an additional amount of time once it is detected to be busy. We consider here a *greedy* jamming scheme, where the jammer stays on the chosen channel as long as it is busy, and switches to a different one in the next slot, if the current channel turns to be idle. Note that different strategies may be adopted for the jammer. We emphasize that the security concern we aim to illustrate is not due to the particular channel selection method used by the jammer, but follows from the memory that resides in any block-based network coding scheme. Thereby, we consider one representative scheme here (the greedy scheme) for illustration purposes.

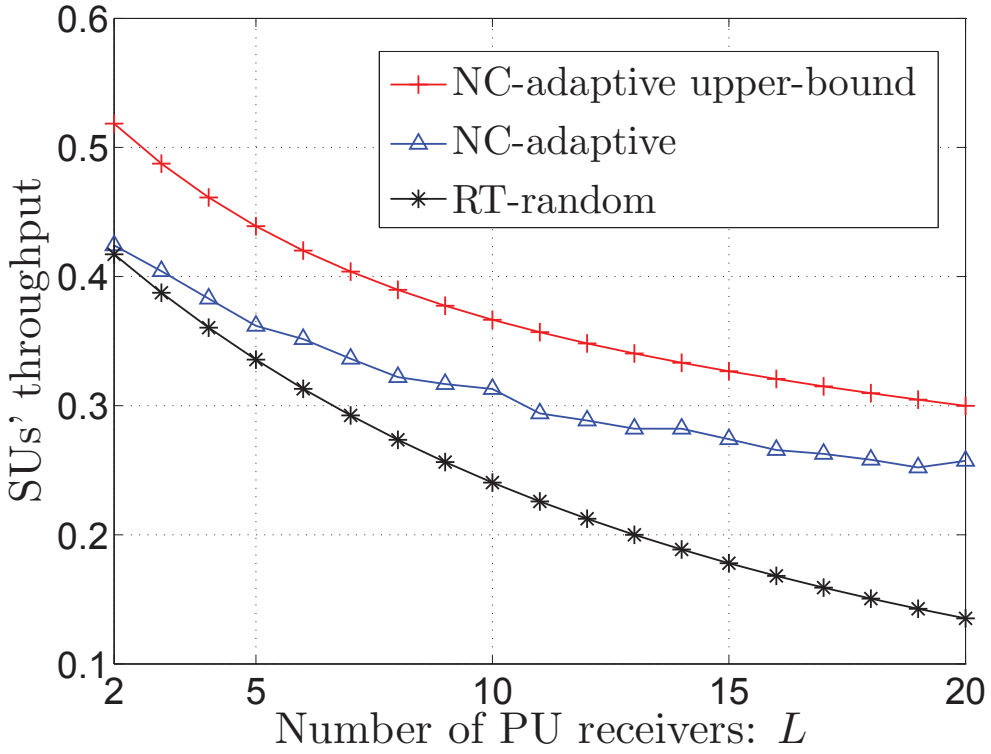

Figure 3.22: SUs' throughput under adaptive sensing with two-level backoff.: vs L.

3.6.3 Performance Evaluation

For comparison, we also consider that when retransmission is used for the intended transmission, the jammer applies a *random selection* scheme, in which it randomly chooses a channel from all N channels at a time, and attacks the transmission on this channel if it is busy or retries in the next slot otherwise. In this case, it is clear that

$$\overline{U}_a = (1 - P_{idle}^{RT})T_a. \tag{3.25}$$

On the other hand, for the case with network-coded transmissions, the jammer applies the greedy scheme as described above. It is noted that

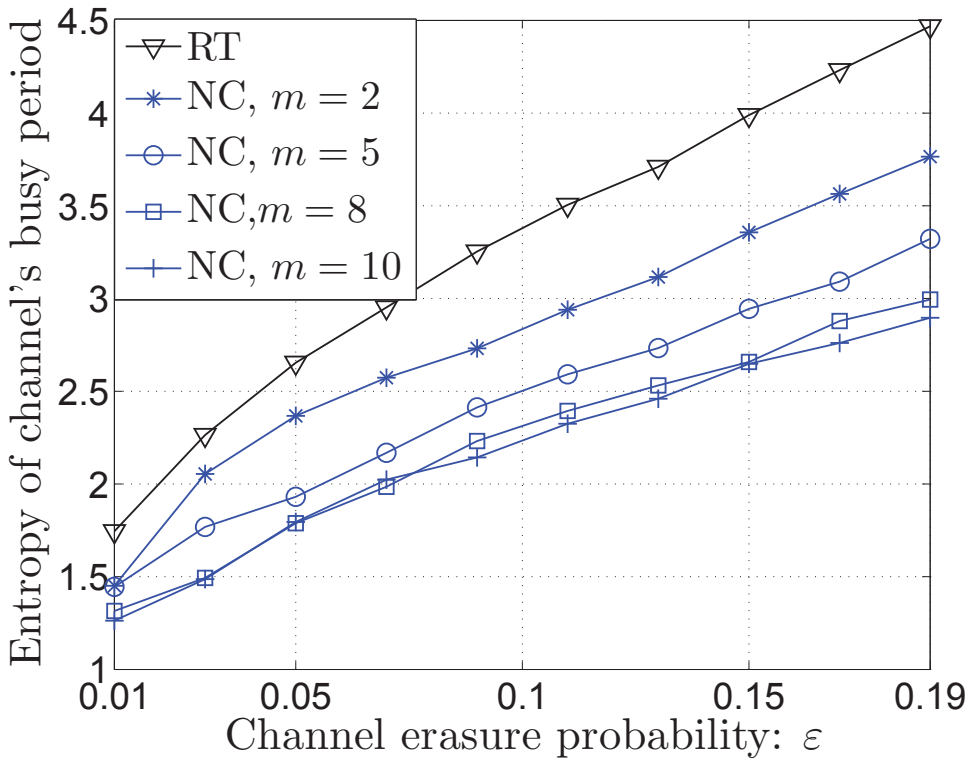

Figure 3.23: Channel predictability under network coding or retransmission.

the closed-form expression of $\overline{U}_a$ here is not attainable because a high-order memory (on the scale of m) resides in the network-coded transmissions, and the time instant the jammer accesses a particular channel is unknown and random with respect to the beginning of that busy period on the channel. Thereby, we carry out simulations to demonstrate the performance of the jammer in this case and compare it with the performance obtained from the retransmission-based traffic scenario.

The simulation results are collected over 10^6 time slots, with $T_a = 20$ and the default parameter values being $\lambda = 0.4, \varepsilon = 0.1, L = 10, N = 10$ and $m = 2$. As can be seen from Fig. 3.24 - Fig.3.27, the jammer's performance improves (undesirably) under network-coded channels, compared to

the case with retransmissions. In particular, Fig. 3.26 shows that the number of jammed transmissions increases with the batch size m, pointing to an increased security concern imposed by network coding, while the throughput of the intended transmission is expected to increase with m in the absence of a jammer [45]. Such observations imply a tradeoff between the throughput and security, and indicate that a careful design of the network-coded system is required to balance the system performance and security considerations.

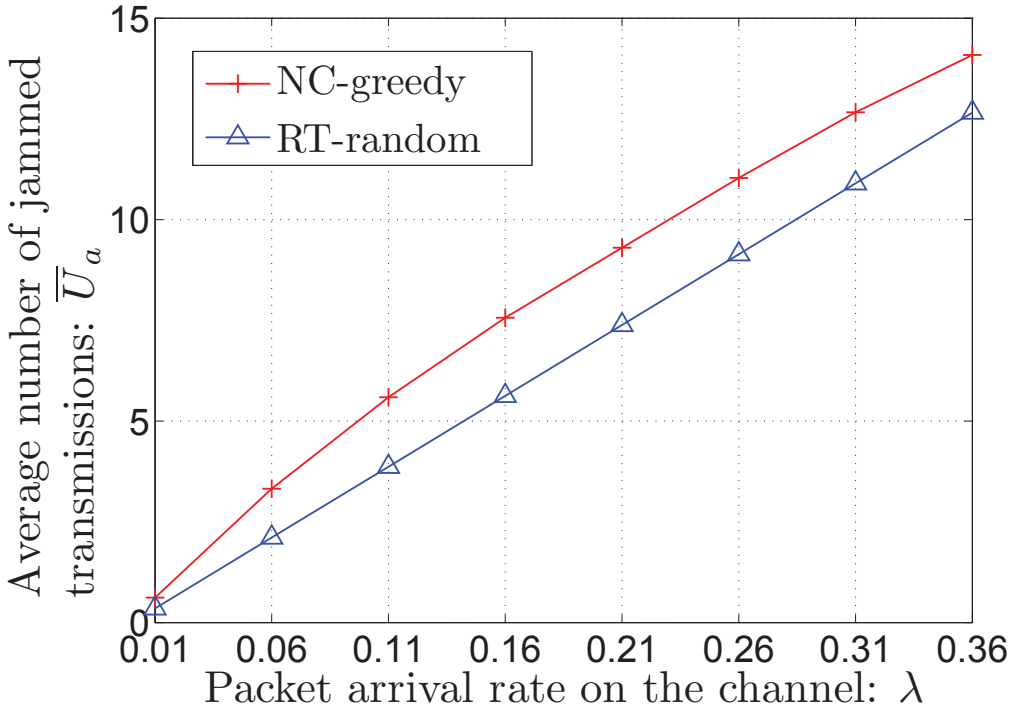

Figure 3.24: Average number of jammed transmissions under network coding (with greedy attack) or retransmission (with random attack): vs. λ.

3.7 Conclusion

With the observation that network coding induces a more "predictable" structure on the communication channel, we explored the possible advantages and

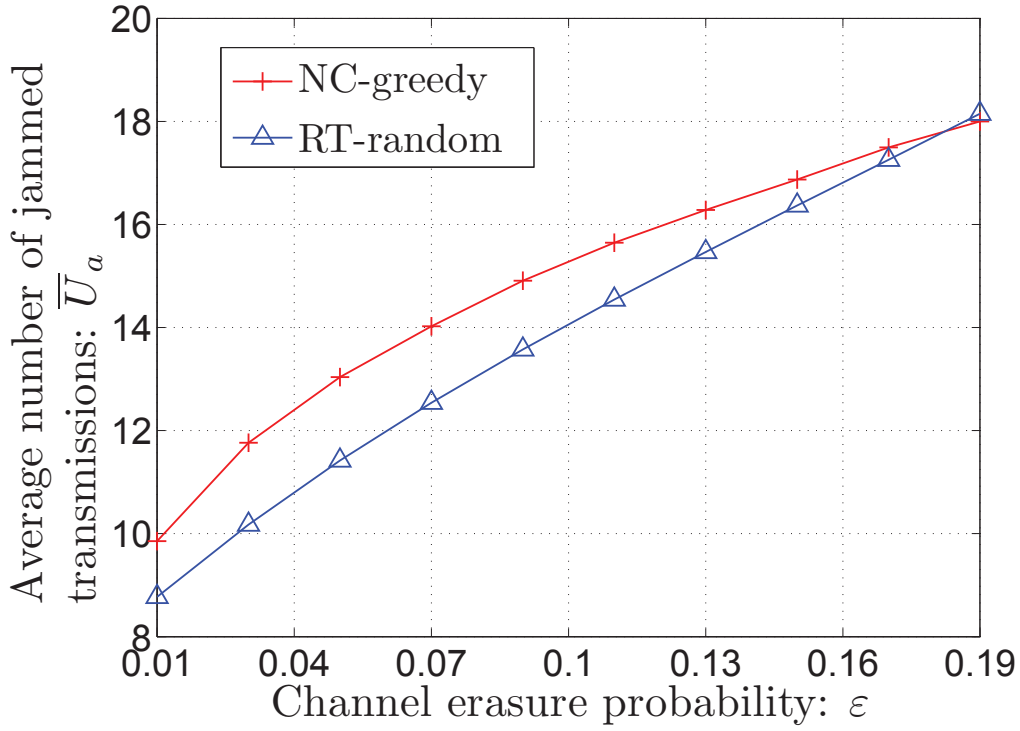

Figure 3.25: Average number of jammed transmissions under network coding (with greedy attack) or retransmission (with random attack): vs. ε.

disadvantages of such predictability in a wireless network. First, we considered a CR network with the SUs opportunistically accessing multiple PU channels. Each PU transmits packets to multiple receivers over lossy wireless channels via network coding or retransmission. Viewing network coding as a spectrum shaper, we showed that it increases the spectrum availability for the SUs and offers a more predictive structure to the PU spectrum, i.e., it improves the SU's prediction of spectrum holes on the PU channels. Based on the spectrum shaping effect of network coding, we developed adaptive channel sensing for the SUs, which is carried out by dynamically updating the list of the PU channels that are predicted by the SUs to be idle. Our analysis and numerical results showed that both PU and SU's throughput can be improved, when PUs

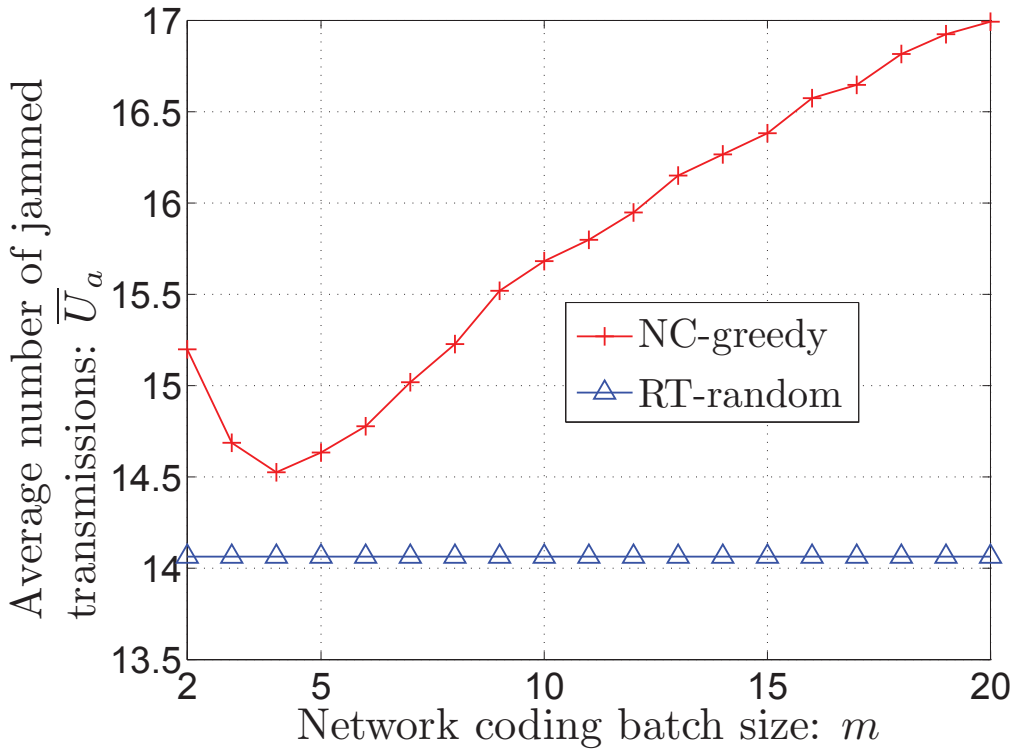

Figure 3.26: Average number of jammed transmissions under network coding (with greedy attack) or retransmission (with random attack): vs. m.

apply network coding, and the SUs can further improve this gain by applying adaptive channel sensing, with either fixed or randomized backoff parameter. Moreover, contention-based spectrum access of multiple SUs was considered and we developed a two-level backoff scheme based on the adaptive sensing to exploit the memory in network-coded transmissions and realize the throughput gains for the SUs.

Next, we showed that the predictability induced by network coding makes wireless networks more vulnerable to jamming attacks targeting at the intended transmissions. Since the entropy of the busy period of network-coded transmission is less than that of the retransmission-based communication, a

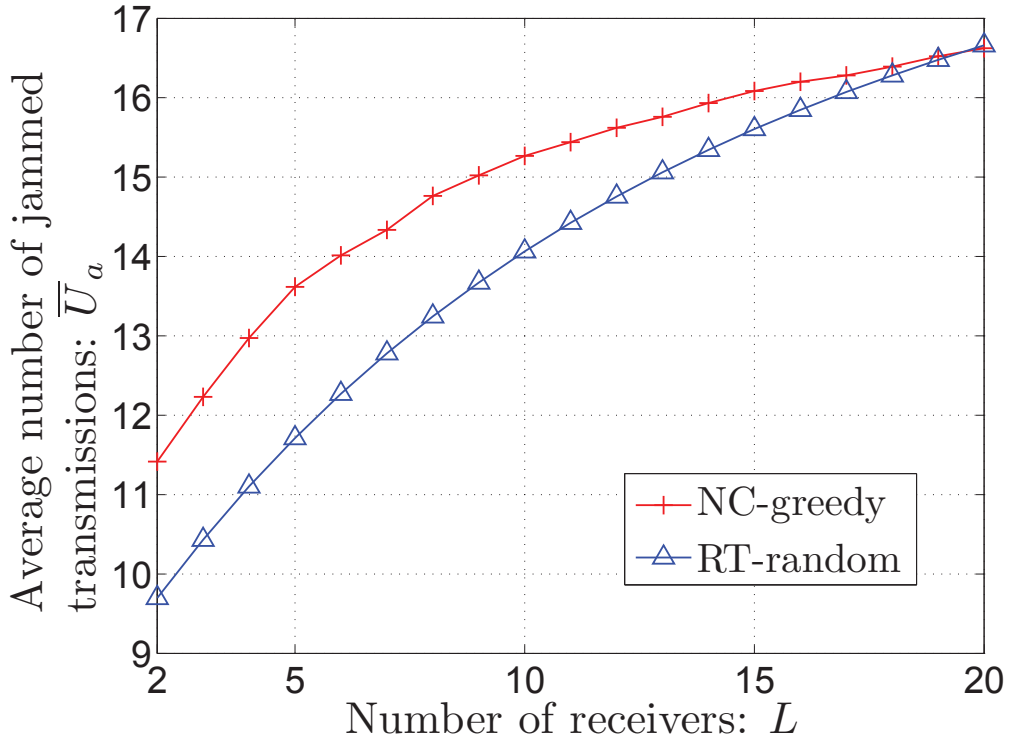

Figure 3.27: Average number of jammed transmissions under network coding (with greedy attack) or retransmission (with random attack): vs. L.

jammer can predict the spectrum's future status (busy or idle) more easily and increase the performance loss to the system by adaptive channel selection. Particularly, this jamming attack effect increases with the coding batch size, which, on the other hand, (desirably) improves the throughput of network-coded transmissions. This suggests that a careful design is needed to balance the throughput performance and the jamming-resistant capability of the wireless systems. We believe that our discussion here, on the "two sides of the coin" of network coding, can provide insight into the development of emerging technologies for wireless networks in general (including the CR networks), where network coding has been attracting an increasing practical interest (see, e.g., [68–70]).

Chapter 4

A CHARACTERIZATION OF DELAY PERFORMANCE OF COGNITIVE MEDIUM ACCESS

4.1 Introduction

In a hierarchical overlay cognitive network [3], an SU communicates opportunistically by exploiting spectrum "white space" left temporarily by PUs. As a result, transmissions of an SU is limited by the stochastic nature of PUs. An SU hoping to run certain applications (e.g. VoIP or streaming) would like to know what kind of rate and delay a secondary network can provide. In the same token, an owner of a secondary network would like to attract potential users by advertising a certain level of QoS assurance.

Characterizing the delay of a cognitive network is challenging. Specifically, the delay of an SU is affected by not only its own buffer and traffic properties, but also PUs' traffic characteristics, other competing SUs, and access policy of SUs. These interacting factors make delay analysis often analytically intractable, and only a limited number of results have been reported in the literature (see e.g., [25, 71, 72]).

We analyze in this chapter the delay performance in a cognitive radio network, where SUs contend for channels using an Aloha-based random access policy. In particular, an SU senses a channel owned by a PU and transmits only if the PU channel is idle. We model the PU's traffic generation as an ON-OFF process where the PU generates data only during the ON periods. For SUs, we assume that they generate data packets in each slot according to a Poisson distribution. Based on stochastic fluid queue theory, we model the system dynamics by using *Poisson driven stochastic differential equations* (PDSDE),

and analyze the steady state queue lengths of SUs accordingly. To facilitate tractability, we focus on the light traffic regime where the traffic intensity is low, as is often the case for delay analysis of buffered Aloha, e.g., [73, 74]. We consider the homogeneous case where the arrival rates of SUs are the same, and characterize the moments of the random queue lengths of SUs, for cases with a single PU channel (SCH) and multiple PU channels (MCH). Clearly, these moments provide critical statistical information about SUs' queueing length distribution. We also examine the impact of the PU traffic on SUs' queue lengths and the gain of using multiple PU channels. Adaptive algorithms, based on local information only, are developed to find the optimal contention probabilities that achieve the minimum mean queue lengths.

Next, we explore the gain of using two interfaces per SU, i.e., each SU is equipped with two interfaces (radios). Accordingly, each SU can sense two channels at a time and thus transmit on up to two channels, as long as the PU channels are idle and no contention collisions occur. Our analysis and numerical examples corroborate the intuition that the usage of two interfaces can greatly improve the delay performance by decreasing the mean queue lengths of SUs.

Furthermore, it is of equal importance to consider the scenario where stringent delay requirements are imposed on the SUs. In such a scenario there exists a maximum amount of traffic accommodable, necessitating traffic control. In this study, we consider packet generation control. As representative approaches, we develop two control mechanisms, one randomized and the other based on the queue lengths of the SUs.

The approach adopted in this chapter originates from the early work of Liu and Gong who studied the delay performance of priority queues using

fluid models [75]. Given the access structure of a hierarchical cognitive network, the problem of queueing analysis indeed resembles that of the priority queue problem. There are, however, nontrivial differences arising from cognitive radio specific applications. In particular, the problem considered in [75] arises from centralized scheduling of high and low priority queues, whereas, in this chapter, we consider multiple SUs competing for transmission opportunities by random access. This random access to the PU channels gives rise to the coupling across SUs' queue dynamics, which was not the case in [75] since only one low priority flow was considered there. In addition, in contrast to [75] where the single low priority flow receives a constant service rate whenever the buffer of the high priority flows is empty, in our study, SUs receive randomly arrived packets. As a result, the number of backlogged SUs is time-varying, and the service rate is random. Besides the work in [75], the delay performance of a multi-hop wireless ad hoc network was studied in [76], where *diffusion approximation* was used to characterize the average end-to-end delay. In [77], WLANs with access points connecting a fixed number of users in the presence of HTTP traffic was considered. A processor sharing queue with state-dependent service rate was used to model the system and analyze the mean session delay. In [78], queueing delay at nodes in an IEEE 802.11 MAC-based network was analyzed, where each node was modeled as a discrete time $G/G/1$ queue. Delay analysis for buffered Aloha was also studied (see [73, 74, 79–81] and references therein). In [79] and [81], the approach named "tagged user" was adopted. Specifically, the interfering users/nodes were modeled as "independent" queues in the sense that the analysis was conducted on one particular user, named the "tagged user," while the interference across users was incorporated into the characterization of the service time distribution of this tagged user. Another approach utilizing Markov chains with

reduced state space to approximate delay analysis can be found in [73, 74, 80]. Two Markov chains, one for the queueing dynamics at one user, and the other for the system status (i.e., the number of busy users, and/or the identities of the users (empty, busy or blocked)), were employed for characterizing the steady state distributions of the system as well as the delay. It is worth noting that the approximation worked well only for the light traffic regime, as has been pointed out in [73] and [74]. In [25] and [72], a large deviation approach was used to analyze delay characteristics of SUs. Inner and outer bounds on the large deviation rate region were obtained in [72] for a set of SUs with orthogonal sharing of spectrum opportunities.

We have a few more words on fluid models. Fluid approximation is a widely used tool for performance analysis in many fields, including communication networks and control techniques [82, 83]. It can provide a good approximation to the original systems by converting the discrete packets into a continuous fluid and offers greater tractability in analyzing the system performance. We should note that along a different avenue, the *deterministic* fluid model has been developed to analyze queueing systems, where microscopic fluctuations in the original systems are replaced by their mean values (see, e.g., [84, 85]). For a given random process $G(t)$, the resulted *fluid scale process*, obtained by using the *Functional Law of Large Numbers*, is defined as $\tilde{g}^{\beta}(t) = G(\beta t)/\beta$, i.e., the time and space are scaled by the same factor β for β being large. This deterministic model leads to the application of *ordinary differential equations* (ODE), which is in contrast to the stochastic differential equations we shall use in our context.

The rest of the chapter is organized as follows. In Section 4.2, we introduce the system model. Fluid flow approximation and PDSDE-based

analysis on the single PU channel case are given in Section 4.3. Section 4.4 studies the case with multiple channels, followed by the delay performance analysis when each SU is equipped with two interfaces in Section 4.5. Packet generation control for SUs under delay requirements is considered in Section 4.6. Finally, conclusions are drawn in Section 4.7.

4.2 System Model

Consider a time-slotted (with slot duration normalized to be 1) cognitive radio network with N PU channels and M SUs, where SUs contend for the channels using distributed random access policies when the PUs are inactive, as illustrated in Fig. 4.1. This model is of interest to many practical scenarios. For example, in a sensor network equipped with cognitive radios, sensors send out measurement data of the environment sporadically and opportunistically over "empty" PU channels.

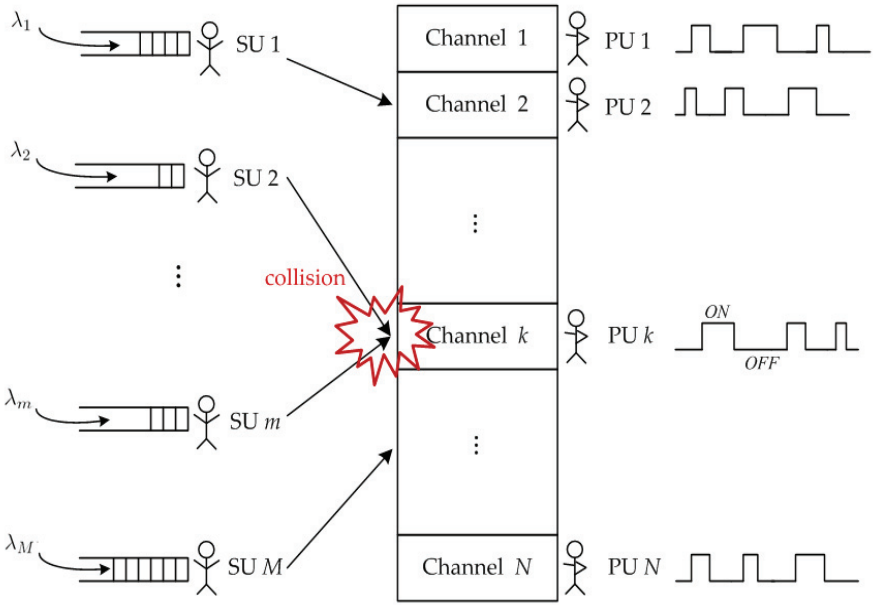

Figure 4.1: A cognitive radio network with multiple PUs and SUs.

Without loss of generality, we associate one PU with one channel (one can use a virtual PU to represent the PU activity). The data generation of the PU on channel j can be represented as a continuous-time ON-OFF process $x_j(t), j = 1, 2, \cdots, N$, i.e., when $x_j(t) = 1$ (ON periods), the PU generates data traffic at rate r_j; otherwise, no data is generated. The transmission rate on each channel is normalized to be 1. We are interested in the case where $r_j > 1$ during the ON periods (the case with $r_j \leq 1$ is trivial since the PUs' buffers are always empty). Let $A_{l,j}$ and $S_{l,j}$ denote the lth active and silent period of $x_j(t)$ respectively. We assume that[1] $\{A_{l,j}\}$ are $i.i.d.$ and follow an exponential distribution with $E[A_{l,j}] = 1/\mu_{H_j}$, and that $\{S_{l,j}\}$ are independent from $\{A_{l,j}\}$ and follow an exponential distribution with $E[S_{l,j}] = 1/\lambda_{H_j}$. It is worth noting that since PU's ON/OFF periods are typically much larger than the duration of one slot, we here neglect the edge effect where collisions between PUs and SUs occur when PUs transit from OFF to ON. That is, the probability that PUs generate new data during the middle of a slot and therefore preempt the transmission of SUs is negligible.

We assume that in each slot, each SU generates data packets according to a Poisson distribution with rate λ. In an overlay cognitive radio network, PUs have strict priority over SUs; SUs can transmit only if the channels are unoccupied by PUs. The channel access process is outlined as follows: each SU with backlogged data may choose a channel independently and uniformly at a time to probe. If the channel is sensed to be unoccupied, it contends for the channel with probability p. If the contention is successful (i.e., no other SUs are contending on the channel at the same time), the user then transmits its backlogged data. In fact, this simple random access policy turns to be

[1]This continuous-time Markovian model is widely used in the literature to model the PU's traffic (see, e.g., [25, 86]).

throughput optimal for small p and when there is only one SU [86]. Note that in practical scenarios, an SU would not have the knowledge of how many backlogged SUs there are, and accordingly we set the contention probability p to be oblivious of backlogged SUs.

For notational convenience, let $H_j(t)$ and $L_i(t)$ denote the queue lengths corresponding to PU j and SU i at time t, respectively, and P_{I_j} be the probability that PU j is idle, i.e., $P_{I_j} = \Pr(H_j(t) = 0)$. In the following, we shall focus on characterizing the queue lengths of SUs.

4.3 Multiple SUs Meet Single PU

4.3.1 Sample Path Description Using Poisson Driven Stochastic Differential Equations

We first consider the case with a single PU channel. For notational convenience, we drop the subscript j related to the PU parameters. In order to guarantee system stability, we enforce that

$$\lambda < \min\left\{\frac{1}{M}\left(1 - \frac{r\lambda_{_H}/\mu_{_H}}{\lambda_{_H}/\mu_{_H}+1}\right), \frac{1}{eM}P_I\right\}. \tag{4.1}$$

It is worth mentioning that the second term in (4.1) was established using the idea of "dominant systems," which has been used in characterizing the stability region of interacting queues in random access systems (e.g., [87, 88]). In our context, the "dominant system" is a system where an SU continues to probe the PU channel regardless of its buffer state (empty or backlogged). Accordingly, the stability region for this system is given by $\lambda < P_I p(1-p)^{M-1}$. Based on [87] and [88], the original system is stable if the dominant system is stable. In other words, the stability region obtained through the dominant system serves as an inner bound to that of the original system.

The queue dynamics of SU i can be written as

$$L_i(d+1) = [L_i(d) + U_i(d) - V_i(d)]^+,$$

where $U_i(d)$ and $V_i(d)$ stand for the arrivals and departures to/from SU i's queue during slot d.

To facilitate analysis, in the following, we take a macroscopic view on the queue evolution of SUs across multiple slots and use continuous approximation to characterize the dynamics in SUs' activities (as illustrated in Fig. 4.2). Let $\zeta_i(t)$ be the indicator random variable for the contention of SU i at time t (i.e., when it contends, $\zeta_i(t) = 1$; otherwise $\zeta_i(t) = 0$). The following stochastic differential equation is thus obtained:

$$dL_i(t) = dN_i(t) - (1 - \mathcal{I}_{H(t)})\zeta_i(t)\mathcal{I}_{L_i(t)} \prod_{k\in\{1,\ldots,M\}\setminus\{i\}} \left[1 - \mathcal{I}_{L_k(t)}\zeta_k(t)\right] dt, \quad (4.2)$$

where $\{N_i(t)\}$ are a set of Poisson counters with rate λ; and $\mathcal{I}_{f(t)}$ stands for the indicator function $\mathbf{1}(f(t) > 0)$.

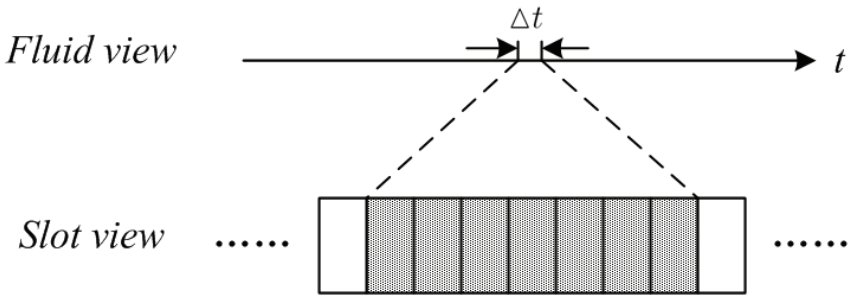

Figure 4.2: Fluid approximation of a slotted system.

Furthermore, it is clear that for the PU, its dynamics can be characterized as follows:

$$dH(t) = rx(t)dt - \mathcal{I}_{H(t)}dt. \quad (4.3)$$

Observe that (4.2) forms a set of Poisson driven stochastic differential equations (PDSDE) [89, 90]. Simply put, in a PDSDE, Poisson processes are the driving sources capturing the system dynamics, and this is in contrast to the conventional SDE where the Brownian Motion is used to describe the dynamics in the trajectory of a stochastic differential equation. In general, a PDSDE can be given as

$$z(t) = z(0) + \int_0^t f(z(\sigma), \sigma) d\sigma + \int_0^t g(z(\sigma), \sigma) dN_\sigma, \tag{4.4}$$

where N_σ is a Poisson counter. For the sake of completeness, we restate the definition of the solution to the above PDSDE [89].

Definition 1. *A function $z(\cdot)$ is a solution to* (4.4)*, in the Itô's sense, if on an interval where N_σ is constant, z satisfies $\dot{z} = f(z, t)$ and if N_σ jumps at t_1, z behaves in a neighborhood of t according to the rule*

$$\lim_{\substack{t \to t_1 \\ t > t_1}} z(t) = g(\lim_{\substack{t \to t_1 \\ t < t_1}} z(t), t_1) + \lim_{\substack{t \to t_1 \\ t < t_1}} z(t),$$

and $z(\cdot)$ is taken to be continuous from the left. When this definition is adopted, we can rewrite (4.4) *as*

$$dz(t) = f(z, t)dt + g(z, t)dN_\sigma(t).$$

Based on the properties of PDSDE [89], it can be shown that for $n \geq 2$,

$$dL_i^n(t) = nL_i^{n-1}(t)dL_i(t) + \sum_{k=2}^{n} \binom{n}{k} L_i^{n-k}(t)dN_i(t).$$

It follows that the moments of $L_i(t)$ in the steady state satisfy the following recursive equation[2]:

$$nE[L_i^{n-1}F] - \sum_{k=1}^{n} \binom{n}{k} E[L_i^{n-k}]\lambda = 0, \tag{4.5}$$

[2]We drop the time index t as the meaning is clear.

where

$$F = (1 - \mathcal{I}_{H(t)})\zeta_i(t)\mathcal{I}_{L_i(t)} \prod_{k\in\{1,\ldots,M\}\setminus\{i\}} \big[1 - \mathcal{I}_{L_k(t)}\zeta_k(t)\big].$$

4.3.2 Moments of SU Queue Lengths

We now start to study in more detail the moments of the queue lengths of SUs based on the above PDSDEs. Recall that SUs can access the channel only when the buffer of the PU is empty. With this observation, we first examine the idle period of the PU P_I. Note that the PU generates data at rate r only during an ON-period, and that the buffer is depleted at rate 1 as long as the queue is nonempty. The sample path description of the PU traffic then satisfies the following PDSDE:

$$\begin{aligned} dx(t) &= (1 - x(t))dN_{H_1}(t) - x(t)dN_{H_2}(t), \\ dH(t) &= rx(t)dt - \mathcal{I}_{H(t)}dt, \end{aligned} \tag{4.6}$$

where $N_{H_1}(t)$ and $N_{H_2}(t)$ are a pair of Poisson counters driving $x(t)$, with rate $\lambda_{_H}$ and $\mu_{_H}$ respectively. It is not difficult to show that in the steady state

$$P_I = 1 - \frac{r\lambda_{_H}}{\lambda_{_H} + \mu_{_H}}. \tag{4.7}$$

We then start characterizing the moments of the SU queue lengths. Based on (4.2), we observe that the M SU queues interact with each other through channel contention. In other words, besides the impact from PU activities, the service time of one SU also depends on other SUs' activities, and it turns out to be a quantity that follows a general distribution which is difficult to determine.

For ease of exposition, we shall focus on the light traffic regime and approximate the SU activities as if they were "weakly coupled" in the sense that

the event that one SU is idle (i.e., with no backlogged data) is independent from other SUs being idle. Similar approximations to "decouple" the interacting queues have been made in [79] and [81], among other works. According to the homogeneity assumption, this *idle probability* would be the same across all SUs. Let p_0 be this probability. It is clear that the number of backlogged SUs follows a Binomial distribution with its probability mass function given by

$$P_m = \binom{M}{m}(1-p_0)^m p_0^{M-m}, \tag{4.8}$$

where p_0 can be shown to satisfy [81] $p_0 = 1-\rho$, with $\rho = \frac{\lambda}{\mu}$ and μ being the mean service rate. In the case with a single PU channel, μ can be calculated as

$$\mu = \frac{1}{M}\sum_{m=1}^{M} mp(1-p)^{m-1}P_I P_m = pP_I(1-p_0)(1-p+pp_0)^{M-1},$$

where the characterization is done under the homogeneity assumption and is conditional on the number of backlogged SUs in the system. It follows that

$$p_0 = 1 - \frac{\lambda}{pP_I(1-p_0)(1-p+pp_0)^{M-1}}. \tag{4.9}$$

Now with all related parameters being characterized, we are in a position to calculate the moments of the queue lengths for SUs. Based on (4.2) and (4.5), the first two moments of SU i's queue length can be derived as

$$E[L_i] = \frac{\lambda}{-2\lambda + 2\alpha_s}, \quad E[L_i^2] = \frac{\lambda(\lambda+2\alpha_s)}{6(\lambda-\alpha_s)^2}, \tag{4.10}$$

where α_s is given by

$$\alpha_s = \sum_{m=1}^{M} p(1-p)^{m-1}P_m P_I. \tag{4.11}$$

4.3.3 Adaptive Algorithm for Optimal Contention Probability

The analysis above indicates that the contention probability, p, and the idle probability of one SU, p_0, are two key parameters to the characterization of

the mean queue lengths of SUs, and thus the delay performance. Intuitively speaking, when p is very small (approaching 0), SUs contend for the channel sporadically, and p_0 is small. On the other hand, when p is very large (approaching 1), all SUs with backlogs contend for the channel almost always, leading to a high contention collision among SUs, which makes the queue lengths increase. It is thus indicated that there exists an optimal value of p, which minimizes the mean queue lengths.

We note that (4.9) formulates a fixed point equation for the idle probability p_0, and p_0 is in itself an implicit function of the contention probability p, i.e., p is the argument of p_0. The following lemma proves the uniqueness on the solution to the fixed-point equation:

Lemma 2. *The fixed point equation (4.9) has a unique solution p_0.*

Proof. Let $\Gamma(\gamma) = (1-\gamma)(1-p+p\gamma)^{M-1}$, $\gamma \in [0,1)$. The first-order derivative of Γ w.r.t. γ is given by

$$\frac{d\Gamma(\gamma)}{d\gamma} = (1-p(1-\gamma))^{M-2}(Mp(1-\gamma)-1). \tag{4.12}$$

Recall from (4.13), we have $p \leq \frac{1}{M(1-p_0)}$, indicating that $\frac{d\Gamma(p_0)}{dp_0} \leq 0$ and $\Gamma(p_0)$ is nonincreasing in p_0. It follows that $1-\frac{\lambda}{\Gamma(p_0)}$ is nonincreasing in p_0 as well. Based on this monotonicity property (cf. [91]), we conclude that there is one unique solution to the the fixed point equation given by (4.9). □

Therefore, we obtain the optimal value of p by taking derivative with respect to p on both sides of (4.9) and setting $dp_0/dp = 0$. After some straightforward calculation, we obtain

$$p = \min\left\{\frac{1}{M(1-p_0)}, 1\right\}. \tag{4.13}$$

Intuitively speaking, $M(1-p_0)$ corresponds to the average number of backlogged users who would contend for channel access. Recall that p_0 is the probability that one SU's queue is empty. Accordingly, stochastic approximation algorithms, based on *local* information only, can be readily developed to find the optimal contention probability[3]. For simplicity, rewrite $\Phi_i(t) \triangleq 1 - \mathcal{I}_{L_i(t)}$. We note that the adaptation of p is based on the update of p_0. It follows that we can devise the following adaptive algorithms to obtain the optimal p. First, we use stochastic approximation to update p_0 as:

$$p_0(t+1) = \left(1 - \frac{1}{t+1}\right) p_0(t) + \frac{1}{t+1}\Phi_i(t+1). \tag{4.14}$$

Based on this adaptation, we next derive the adaptive algorithm for achieving the optimal p as[4]:

$$p(t+1) = p(t) + \frac{1 - Mp(t)(1-\Phi_i(t+1))}{t + Mp(t)(1-\Phi_i(t+1))}p(t). \tag{4.15}$$

The convergence of (4.14) and (4.15) can be shown by using the standard arguments from stochastic approximation [92].

We now illustrate by numerical examples the above results where the contention probability is set to be the optimal value given by (4.13), and SU's average queueing delay is transformed from $E[L_i]$ via Little's Law [64]. We also compare the analytical results obtained from the above fluid approximation with the Monte Carlo simulation studies of the underlying system. As shown in Fig. 4.3 (with $M = 20$ and $r = 1.2$), the queueing delay increases with the arrival rate of SUs, and as the *duty cycle* of $x(t)$, defined as $\tau \equiv \frac{\lambda_H}{\lambda_H + \mu_H}$, increases, the average queueing delay of SUs increases, indicating the impact

[3]Note that the SUs are statistically identical and will adopt the same update procedure.

[4]We assume that the number of SUs, M, is given and known to all SUs as a system parameter, and so is the number of PU channels, N, which will be used in the subsequent sections.

of the PU traffic on SUs' delay performance. In addition, the simulation and analytical results are shown to match with each other closely.

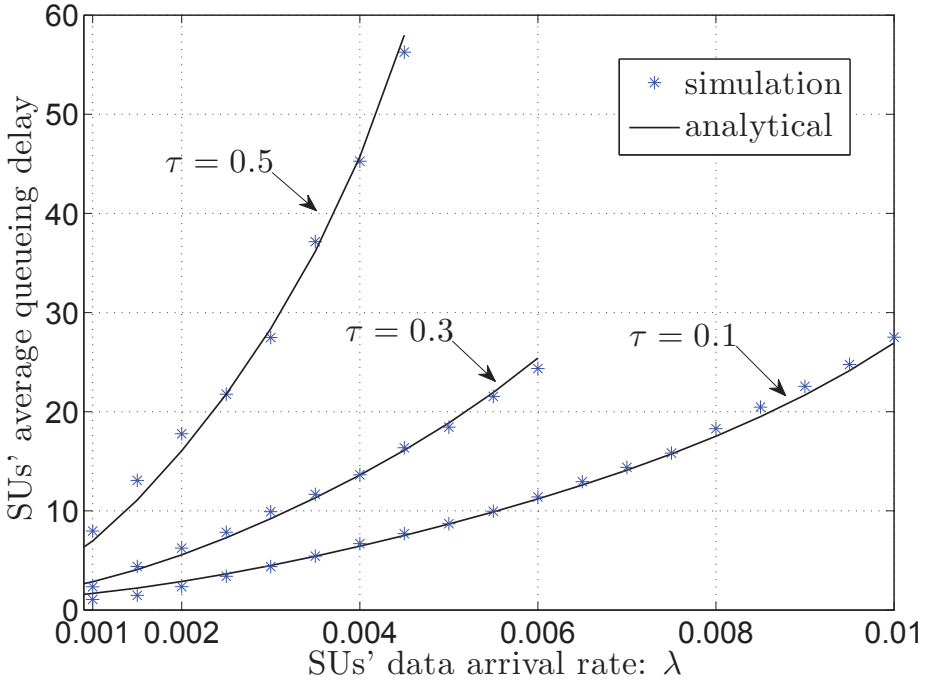

Figure 4.3: Average queueing delay of SUs for the case with a single PU channel.

4.4 Multiple SUs Meet Multiple PUs

We next consider the case where there are multiple PU channels, and examine the performance gain therein.

4.4.1 Sample Path Description Using Poisson Driven Stochastic Differential Equations

In this case, to keep the system stable, we enforce that

$$\lambda < \min\left\{\frac{1}{M}\left(N - \sum_{j=1}^{N}\frac{r_j\lambda_{H_j}/\mu_{H_j}}{\lambda_{H_j}/\mu_{H_j}+1}\right), \frac{N}{eM}P_I\right\}, \tag{4.16}$$

where, again, the second term was obtained along the same line as in the previous section using the idea of "dominant system."

Recall that each SU with backlogged data independently chooses one of the N PU channels uniformly at random. Let $\xi_{ij}(t)$ be an indicator random variable denoting that SU i chooses channel j at time t. As in the single PU channel case, we do continuous approximation when characterizing the dynamics of SU activities. The system dynamics can then be written as: for $j = 1, 2, \cdots, N$,

$$dH_j(t) = r_j x_j(t)dt - \mathcal{I}_{H_j(t)}dt, \tag{4.17}$$

and for $i = 1, 2, \cdots, M$,

$$dL_i(t) = dN_i(t) - \sum_{j=1}^{N}(1 - \mathcal{I}_{H_j(t)})\mathcal{I}_{L_i(t)}\xi_{ij}(t)\zeta_i(t) \prod_{k\in\{1,\ldots,M\}\setminus\{i\}} \Big[1 - \mathcal{I}_{L_k(t)}\xi_{kj}\zeta_k(t)\Big]dt. \tag{4.18}$$

Again, the coupling across SUs is observed in (4.18). We next carry out analysis on the moments of SUs' queue lengths by focusing on the light traffic regime as before.

4.4.2 Moments of SU Queue Lengths

Along the same line as in the single PU channel case, we first characterize the idle period of PUs. The sample path description for PU j is given by the following PDSDE:

$$\begin{aligned} dx_j(t) &= (1 - x_j(t))dN_{H_1}^{(j)}(t) - x_j(t)dN_{H_2}^{(j)}(t), \\ dH_j(t) &= r_j x_j(t)dt - \mathcal{I}_{H_j(t)}dt, \end{aligned} \tag{4.19}$$

where $N_{H_1}^{(j)}(t)$ (with rate λ_{H_j}) and $N_{H_2}^{(j)}(t)$ (with rate μ_{H_j}) are a pair of Poisson counters driving $x_j(t)$. For better tractability, we consider the case where the

PU channels are *i.i.d.* It follows that $P_{I_j} = P_{I_{j'}}, \forall\ j \neq j'$. Denote by $P_I = P_{I_j}$ for simplicity. It is easy to show that in the steady state, P_I can be calculated as is given in (4.7).

Next, we turn our attention to study the moments of SUs' queue lengths. Applying the PDSDE tools, we obtain for SU i,

$$E[L_i] = \frac{\lambda}{-2\lambda + 2\alpha_{\mathcal{M}}}, \quad E[L_i^2] = \frac{\lambda(\lambda + 2\alpha_{\mathcal{M}})}{6(\lambda - \alpha_{\mathcal{M}})^2}, \tag{4.20}$$

where

$$\begin{aligned} \alpha_{\mathcal{M}} &= \sum_{l=1}^{N}\sum_{m=1}^{M}\sum_{k=0}^{m-1}\binom{m-1}{k}p(1-p)^k\left(\frac{1}{N}\right)^{k+1}\left(1-\frac{1}{N}\right)^{m-(k+1)}P_m P_I \\ &= \sum_{m=1}^{M} pP_I\Big(1-\frac{p}{N}\Big)^{m-1}P_m, \end{aligned} \tag{4.21}$$

with P_m being given by (4.8). Furthermore, the mean service rate μ in this case can be calculated as:

$$\begin{aligned} \mu &= \frac{1}{M}\sum_{m=1}^{M} mp \sum_{j=1}^{N}\frac{1}{N}\sum_{k=0}^{m-1}\binom{m-1}{k}\Big(\frac{1}{N}\Big)^k\Big(1-\frac{1}{N}\Big)^{(m-1)-k}(1-p)^k P_I P_m \\ &= pP_I(1-p_0)\left(1-\frac{p}{N}+\frac{pp_0}{N}\right)^{M-1}. \end{aligned}$$

It follows that

$$p_0 = 1 - \frac{\lambda}{pP_I(1-p_0)\left(1-\frac{p}{N}+\frac{pp_0}{N}\right)^{M-1}}. \tag{4.22}$$

The characterization of P_m and $E[L_i]$ then follows.

4.4.3 Adaptive Algorithm for Optimal Contention Probability

Similar to the single PU channel case, taking derivative with respect to p and setting $dp_0/dp = 0$ yields that

$$p = \min\left\{\frac{N}{M(1-p_0)}, 1\right\}. \tag{4.23}$$

It is not difficult to see that $M(1-p_0)/N$ is the average number of backlogged SUs per PU channel. Based on (4.23), similar adaptive algorithms for obtaining optimal p can be developed as in the single PU channel case.

Meanwhile, we note that $p = \frac{N}{M(1-p_0)}$ holds when $N < M(1-p_0)$. In fact, this is the regime of interest when we characterize the gain of using multiple PU channels. Here we present numerical examples to illustrate the above analysis. The contention probabilities are set to be their optimal values. As illustrated in Fig. 4.4 (with $N = 5, M = 20$ and $r = 1.2$), the mean queue lengths of SUs decrease significantly when multiple PU channels are present, pointing to a *multi-channel gain* therein. An illustration of such a gain was depicted in Fig. 4.5, where the gain was defined as the ratio $E[L_i]^{(\mathcal{S})}/E[L_i]^{(\mathcal{M})}$, with the superscripts $\mathcal{S}$ and $\mathcal{M}$ denoting the cases with a single PU channel and multiple PU channels, respectively. It can be seen that as the arrival rate of SUs increases, or the duty cycle of PUs increases, the multi-channel gain increases as well.

4.5 Power of Two Interfaces

Intrigued by the celebrated results in [93] and [94], in this section, we explore the impact of using two interfaces (radios) by each SU on the delay performance in a cognitive radio network.

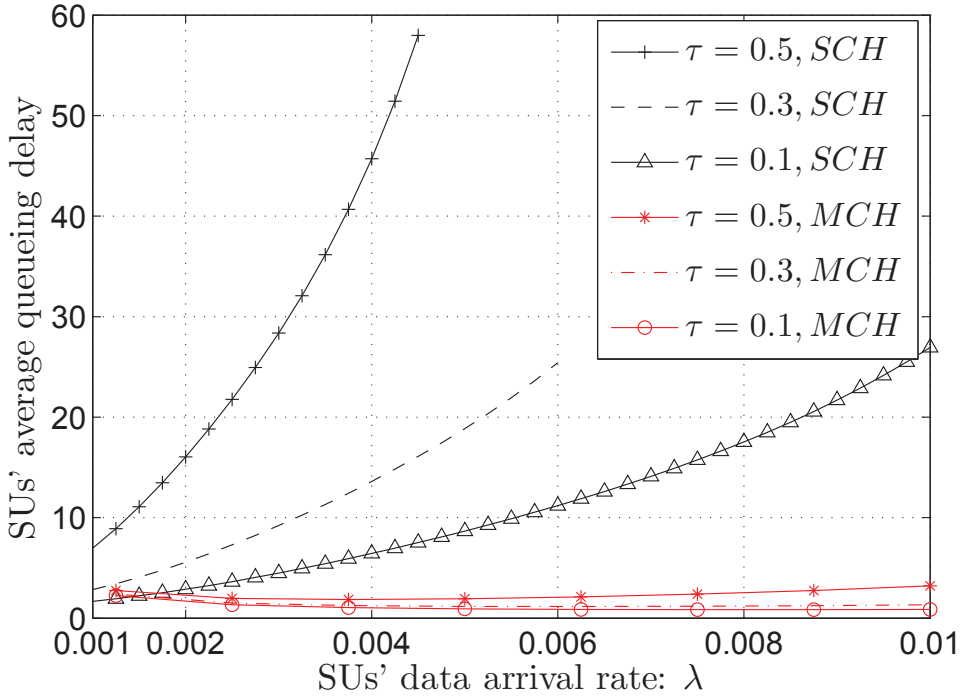

Figure 4.4: Comparison of average queueing delays, for cases with a single PU channel and with multiple PU channels.

4.5.1 System Model

In this new setting, each SU is equipped with two interfaces (this can be readily generalized to cases with more radios), and randomly chooses two channels independently and uniformly at a time. If the chosen PU channels (denoted as $c_1(t)$ and $c_2(t)$ for SU i) are unoccupied, the SU contends for each of them with probability p. If no collisions occur, it starts transmission of different packets on the channels. Clearly, in this case, each SU can access up to two channels for transmission at a time, thus decreasing the delay.

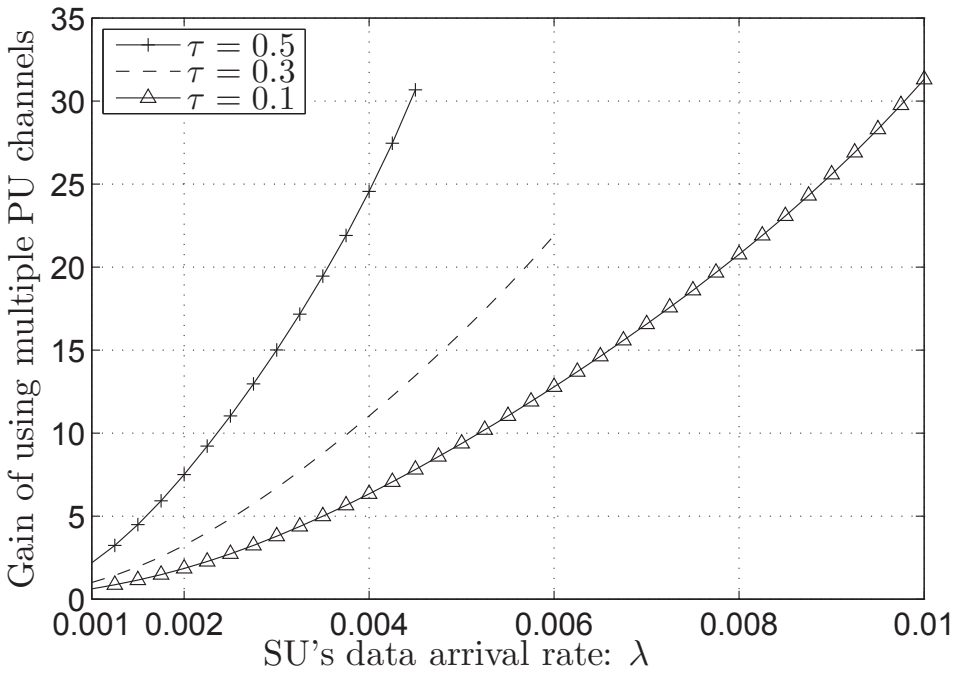

Figure 4.5: Gain of using multiple PU channels.

The queueing dynamics for the SUs, after fluid approximation, are updated as: for $i = 1, 2, \cdots, M$,

$$dL_i(t) \;=\; dN_i(t) - \sum_{\{c_1,c_2\}} \xi_{ic_1}(t)\xi_{ic_2}(t)\mathcal{I}_{L_i(t)}\zeta_i(t)\Big(D_1 + D_2\Big)dt, \qquad (4.24)$$

where

$$D_1 = \big(1 - \mathcal{I}_{H_{c_1}(t)}\big) \prod_{k\in\{1,\ldots,M\}\setminus\{i\}} \big[1 - \mathcal{I}_{L_k(t)}\mathcal{I}(\xi_{kc_1}(t)=1)\mathcal{I}(\zeta_k(t)=1)\big]$$

and, respectively,

$$D_2 = \big(1 - \mathcal{I}_{H_{c_2}(t)}\big) \prod_{k\in\{1,\ldots,M\}\setminus\{i\}} \big[1 - \mathcal{I}_{L_k(t)}\mathcal{I}(\xi_{kc_2}(t)=1)\mathcal{I}(\zeta_k(t)=1)\big]$$

denotes the event that channel c_1, and respectively, c_2, is available.

4.5.2 Moments of SU Queue Lengths

With the setting of *i.i.d.* PU channels unchanged, the idle probability of PUs, P_I, would be still given by (4.7) for the case where the ON-periods of $x_j(t)$'s are exponentially distributed.

As in the case with a single interface, we obtain that for SU i,

$$E[L_i] = \frac{\lambda}{-2\lambda + 2\alpha_c}, \quad E[L_i^2] = \frac{\lambda(\lambda + 2\alpha_c)}{6(\lambda - \alpha_c)^2}, \tag{4.25}$$

where

$$\alpha_c = \sum_{\{c_1,c_2\}} \sum_{m=1}^{M} \frac{1}{\binom{N}{2}}\, p \left\{2Z_2 + Z_1\right\} P_m,$$

with Z_1 and Z_2 being the conditional probabilities that one out of the two chosen channels is available, and both of them are available, given that the total number of backlogged SUs is m, respectively. To get a more concrete characterization of α_c, we denote by $\mathcal{A}_l$ and $\mathcal{B}_l$ the events that channel $c_l, l = 1, 2$, is unoccupied by PUs and un-contended by other SUs, given that the total number of backlogged SUs equals m. It is clear that $\{\mathcal{A}_l\}$ are independent from $\{\mathcal{B}_l\}$. It follows that

$$\begin{aligned} Z_1 &= \Pr(\mathcal{A}_1\mathcal{A}_2\bar{\mathcal{B}}_1\mathcal{B}_2) + \Pr(\mathcal{A}_1\mathcal{A}_2\mathcal{B}_1\bar{\mathcal{B}}_2) + \Pr(\mathcal{A}_1\bar{\mathcal{A}}_2\mathcal{B}_1) + \Pr(\bar{\mathcal{A}}_1\mathcal{A}_2\mathcal{B}_2) \\ &= 2P_I\Pr(\mathcal{B}_1) - 2P_I^2\Pr(\mathcal{B}_1\mathcal{B}_2), \\ Z_2 &= \Pr(\mathcal{A}_1\mathcal{A}_2\mathcal{B}_1\mathcal{B}_2) = P_I^2\Pr(\mathcal{B}_1\mathcal{B}_2), \end{aligned}$$

where $\Pr(\mathcal{B}_1)$ and $\Pr(\mathcal{B}_1\mathcal{B}_2)$ can be derived as

$$\Pr(\mathcal{B}_1) = \sum_{k=0}^{m-1} \binom{m-1}{k} \left(\frac{\binom{N-1}{1}}{\binom{N}{2}}\right)^k \left(1 - \frac{\binom{N-1}{1}}{\binom{N}{2}}\right)^{(m-1)-k} (1-p)^k,$$

and

$$\Pr(\mathcal{B}_1\mathcal{B}_2) = \sum_{k=0}^{m-1} \binom{m-1}{k} \left(1 - \frac{\binom{N-2}{2}}{\binom{N}{2}}\right)^k \left(\frac{\binom{N-2}{2}}{\binom{N}{2}}\right)^{(m-1)-k} (1-p)^k.$$

After some algebra, we obtain that

$$\alpha_c = \sum_{m=1}^{M} 2P_I p\Big(1 - \frac{2p}{N}\Big)^{m-1} P_m. \tag{4.26}$$

Therefore, the idle probability of one SU is given by

$$p_0 = 1 - \frac{\lambda}{2pP_I(1-p_0)\left(1 - \frac{2p}{N} + \frac{2pp_0}{N}\right)^{M-1}}. \tag{4.27}$$

We note that the calculation of μ here provides an upper bound on the per user throughput for SUs, since we have neglected the situation where the SUs have only one packet in the queue and can therefore transmit on only one of the two channels even if both of them are available. However, such cases would be of little interest and are not considered here.

4.5.3 Adaptive Algorithm for Optimal Contention Probability

Following similar steps as in the previous cases, we obtain the optimal contention probability to be

$$p = \min\left\{\frac{N}{2M(1-p_0)}, 1\right\}, \tag{4.28}$$

from where we note that $\frac{M(1-p_0)}{N/2}$ is intuitively the average number of backlogged SUs per PU channel. Adaptive algorithms similar to that described by (4.14) and (4.15) can be devised to find the optimal p.

Different from the multi-channel gain, the power of two choices is typically analyzed in the regime where $N >> M$. With this insight, we next focus on the case where $N \geq 2M$ and characterize the gain of using two interfaces per SU. It is clear that when $N \geq 2M$, the optimal contention probability is given by $p = 1$. It follows that $\alpha_{\mathcal{M}}$ and α_c can be rewritten as

$$\alpha_{\mathcal{M}} = \sum_{m=1}^{M} P_I\Big(1 - \frac{1}{N}\Big)^{m-1} P_m, \tag{4.29}$$

and

$$\begin{aligned}\alpha_c &= \sum_{\{c_1,c_2\}}\sum_{m=1}^{M}\frac{1}{\binom{N}{2}}\Big\{2\Pr(\mathcal{A}_1\mathcal{A}_2\mathcal{C}_1\mathcal{C}_2)+\big[\Pr(\mathcal{A}_1\mathcal{A}_2\bar{\mathcal{C}}_1\mathcal{C}_2)\\ &\qquad +\Pr(\mathcal{A}_1\mathcal{A}_2\mathcal{C}_1\bar{\mathcal{C}}_2)+\Pr(\mathcal{A}_1\bar{\mathcal{A}}_2\mathcal{C}_1)+\Pr(\bar{\mathcal{A}}_1\mathcal{A}_2\mathcal{C}_2)\big]\Big\}P_m\\ &= \sum_{m=1}^{M}2P_I\Big(1-\frac{2}{N}\Big)^{m-1}P_m, \end{aligned} \tag{4.30}$$

where $\mathcal{C}_l$ stands for the event that channel $c_l, l = 1, 2$, is not chosen by any other SUs, given that the total number of backlogged SUs is m.

Also, the mean service rate (denoted as $\mu_d, d = 1, 2$) and the idle probability (denoted as $p_{0,d}, d = 1, 2$) of SUs can be recalculated as follows: for the case with a single interface,

$$\begin{aligned}\mu_1 &= \frac{1}{M}\sum_{m=1}^{M}mP_I\left(1-\frac{1}{N}\right)^{m-1}P_m = P_I(1-p_{0,1})\left(1-\frac{1}{N}+\frac{p_{0,1}}{N}\right)^{M-1},\\ p_{0,1} &= 1-\frac{\lambda}{P_I(1-p_{0,1})\left(1-\frac{1}{N}+\frac{p_{0,1}}{N}\right)^{M-1}}, \end{aligned} \tag{4.31}$$

and for the case with two interfaces,

$$\begin{aligned}\mu_2 &= \frac{2}{M}\sum_{m=1}^{M}mP_I\left(1-\frac{2}{N}\right)^{m-1}P_m 2P_I(1-p_{0,2})\left(1-\frac{2}{N}+\frac{2p_{0,2}}{N}\right)^{M-1},\\ p_{0,2} &= 1-\frac{\lambda}{2P_I(1-p_{0,2})\left(1-\frac{2}{N}+\frac{2p_{0,2}}{N}\right)^{M-1}}. \end{aligned} \tag{4.32}$$

By solving the fixed-point equations for $p_{0,d}, d = 1, 2$, as given above, the mean queue lengths of SUs for the two cases can be readily derived.

We next characterize the gain of using two interfaces under this regime. Let $\psi \triangleq \frac{E[L_i]^{(\mathcal{M})}}{E[L_i]^{(\mathcal{C})}}$, where $E[L_i]^{(\mathcal{M})}$ and $E[L_i]^{(\mathcal{C})}$ denote the mean queue lengths of SUs for the cases with a single interface and two interfaces respectively.

When M is fixed and $N \to \infty$, we have that

$$\begin{aligned}
\lim_{N\to\infty} \psi &= \lim_{N\to\infty} \frac{\alpha_{\mathcal{C}} - \lambda}{\alpha_{\mathcal{M}} - \lambda} \\
&= \lim_{N\to\infty} \frac{2P_I(1 - p_{0,2}^M) - \lambda}{P_I(1 - p_{0,1}^M) - \lambda} \\
&= \frac{2\big(1 - (1 - \sqrt{\rho/2})^M\big) - \rho}{\big(1 - (1 - \sqrt{\rho})^M\big) - \rho},
\end{aligned} \tag{4.33}$$

where $\rho = \lambda/P_I$ is the traffic intensity. Fig. 4.6 depicts the gain as a function of the traffic intensity. As expected, the application of two interfaces provides significant gain by decreasing the mean queue lengths, and as the traffic intensity grows larger, the gain increases as well.

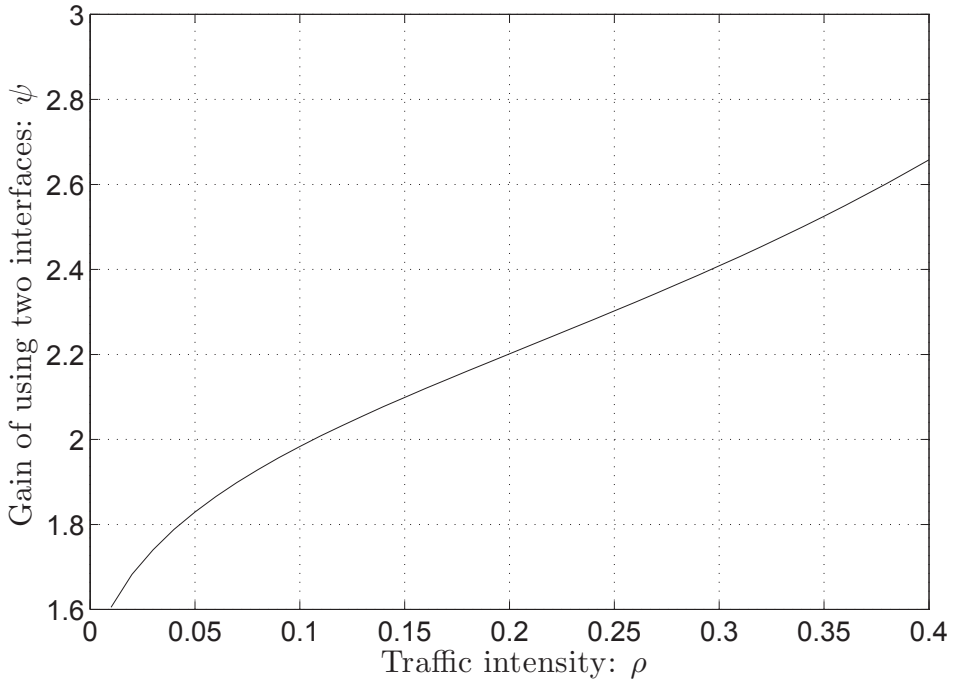

Figure 4.6: Gain of using two interfaces.

4.6 Adaptive Packet Generation Control Under Delay Constraints

In previous sections, we analyzed SUs' delay performance for different scenarios in the light traffic regime. As expected, larger delay can occur with increased arrival rate or decreased spectrum opportunities. Accordingly, when a stringent delay requirement is imposed on the SUs, effective control mechanisms (e.g., rate-limiting) are called for to regulate SUs' traffic in order to meet the requirement. In this section, we turn our attention to study such scenarios and design traffic control strategies that regulate SUs' packet generation to satisfy the delay constraint. In particular, we are interested in packet generation control, where the SUs either use a randomized strategy, or a queue-length-based control mechanism. In the following, we focus on the case with a single PU channel. The analysis readily extends to the case with multiple PU channels.

For notational convenience, let $y_i(t)$ be the control process that regulates SU i's packet generation, i.e., $y_i(t)$ is a Bernoulli random variable taking two values: 0 or 1. When $y_i(t) = 1$, SU i generates new packets, according to the Poisson distribution with rate λ, at time t; otherwise, no new packets are produced, as illustrated in Fig. 4.7.

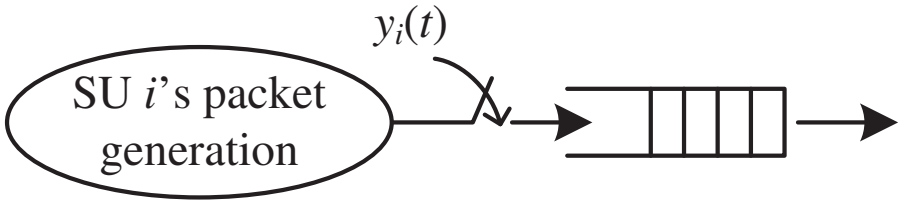

Figure 4.7: An illustration of packet generation control.

Applying fluid approximation, the PDSDE of SU i is written as

$$dL_i(t) = y_i(t)dN_i(t) - F(t)dt. \tag{4.34}$$

Based on the properties of PDSDE, we obtain for $n \geq 2$,

$$nE[L_i^{n-1}F] - \sum_{k=1}^{n} \binom{n}{k} E[L_i^{n-k}y^k]\lambda = 0. \tag{4.35}$$

Suppose that the delay requirement on the SUs is given as

$$\Pr(D \geq D_0) \leq \delta, \tag{4.36}$$

where D denotes the queueing delay of one SU; $D_0 \in \mathbb{N}$ and $\delta \in (0,1)$ are positive constants and known to all users *a priori*. Appealing to Markov's Inequality and Little's Law, a sufficient condition in meeting the delay constraint (4.36) can be expressed in terms of the SUs' mean queue length as follows:

$$E[L_i] = \lambda_0 E[D] \leq \lambda_0 \delta D_0, \tag{4.37}$$

where λ_0 is the average packet arrival rate of each SU, under the delay constraint.

4.6.1 Randomized Packet Generation Control by SUs

In the randomized packet generation control, SUs generate new packets with probability q, independently across users and time, i.e.,

$$y_i(t) = \begin{cases} 1, & \text{w.p. } q, \\ 0, & \text{w.p. } 1-q, \end{cases} \tag{4.38}$$

Based on (4.35), we obtain

$$E[L_i] = \frac{\lambda E[y_i^2]}{-2\lambda E[y_i] + 2\alpha_{\mathcal{S}}} = \frac{\lambda q}{-2\lambda q + 2\alpha_{\mathcal{S}}}, \tag{4.39}$$

where $\alpha_{\mathcal{S}}$ is given by (4.11) and P_m by (4.8), with $p_0 = 1 - \frac{\lambda_0}{\mu} = 1 - \frac{\lambda_0}{pP_I(1-p_0)(1-p+pp_0)^{M-1}}$ and $\lambda_0 = \lambda q$.

Using a similar approach, it can be shown that the optimal contention probability p is the same as given in (4.13), and the corresponding stochastic algorithm given by (4.14) and (4.15) can be applied to update p_0 and p by the SUs.

Intuitively, the larger the control parameter q, the higher the buffer occupancy and SUs' queueing delay. In particular, we are interested in finding out the maximum q satisfying (4.37), i.e.,

$$q_{max} = \max\{q \in [0,1] : E[L_i] \leq q\lambda\delta D_0\}. \tag{4.40}$$

Since a closed-form expression for q_{max} is not attainable, we next conduct numerical study to find q_{max} under different delay requirements and SUs' data arrival rates. As shown in Fig. 4.8 (with $M = 20$), when the arrival rate increases, or the delay requirement becomes more strict (i.e., with a smaller value of the product δD_0), the maximum traffic admission probability decreases.

4.6.2 Threshold-based Packet Generation Control by SUs

Different from the randomized control strategy outlined above, in the threshold-based control scheme, each SU decides whether to generate new packets by comparing its current queue length with a threshold L_0: if the queue length is smaller than or equal to L_0, the SU generates data at rate λ; otherwise, no traffic is generated, i.e.,

$$y_i(t) = \begin{cases} 1, & L_i(t) \leq L_0, \\ 0, & L_i(t) > L_0. \end{cases} \tag{4.41}$$

Correspondingly, we have

$$p_0 = 1 - \frac{\lambda \Pr(L_i \leq L_0)}{\mu}, \tag{4.42}$$

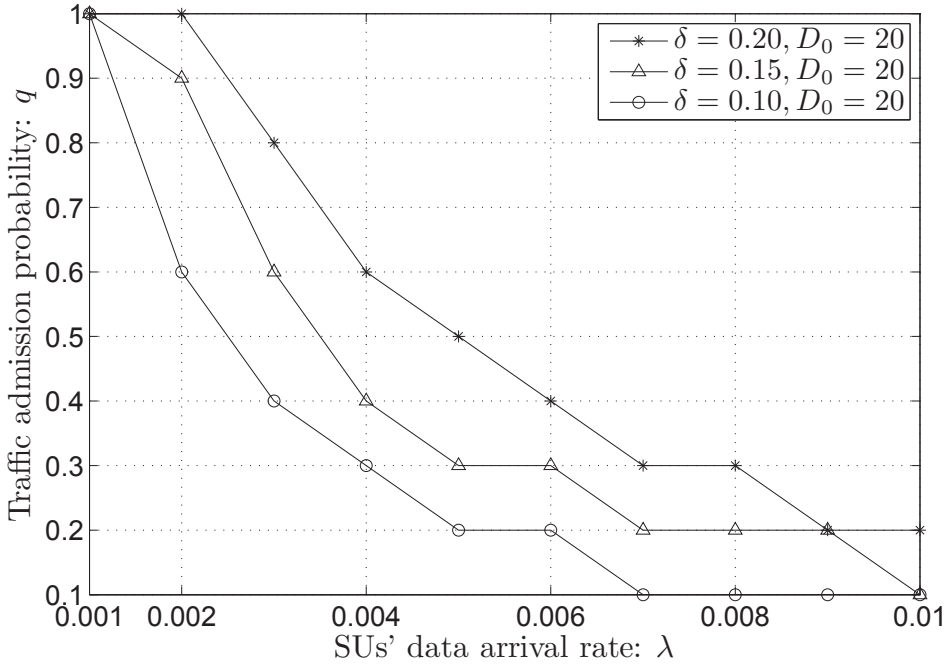

Figure 4.8: q_{max} under different SU arrival rates and delay requirements.

and the mean queue length of the SUs can be derived as (we omit details here for brevity)

$$E[L_i] = \frac{\lambda}{2\alpha_s}\left(E[Ly] + E[y^2]\right)$$
$$= \frac{\lambda}{2\alpha_s}\Big(2\sum_{k=1}^{L_0}\Pr(k \le L_i \le L_0) + \Pr(L_i \le L_0)\Big). \tag{4.43}$$

As in (4.37), a sufficient condition for meeting the delay constraint is

$$E[L_i] \le \lambda \Pr(L_i \le L_0)\delta D_0. \tag{4.44}$$

Clearly, based on (4.43), a closed-form expression on the average queue length is not attainable. However, distributed adaptive learning, similar to (4.14) and (4.15), can be performed by the SUs to dynamically adjusting the threshold L_0 and control the traffic accordingly. We next carry out simula-

tions (over 4×10^4 trials) to study the performance of the threshold-based control mechanism. Again, we are interested in obtaining the best L_0 that leads to the maximum traffic admission probability $\Pr(L_i \leq L_0)$ with which the sufficient condition can still be satisfied. Figs. 4.9 and 4.10 depict a few simulation results on the maximum probability $\Pr(L_i \leq L_0)$ and the corresponding threshold L_0, respectively. It can be seen that when λ increases, or the delay requirement becomes more stringent, the traffic admission probability $\Pr(L_i \leq L_0)$ decreases, and so does the threshold L_0.

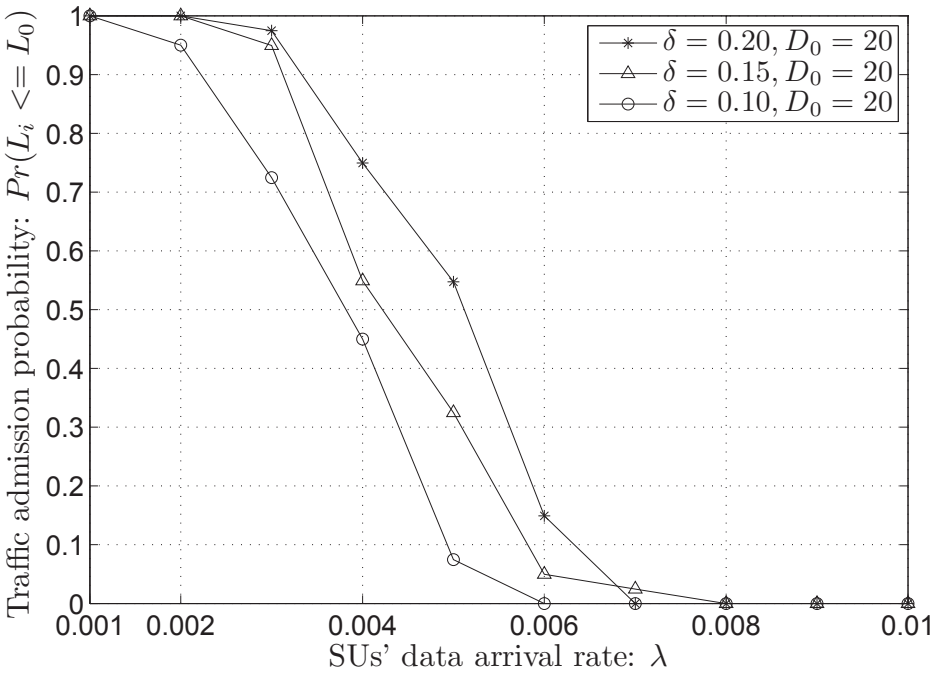

Figure 4.9: $\Pr(L_i \leq L_0)$ under different SU arrival rates and delay requirements.

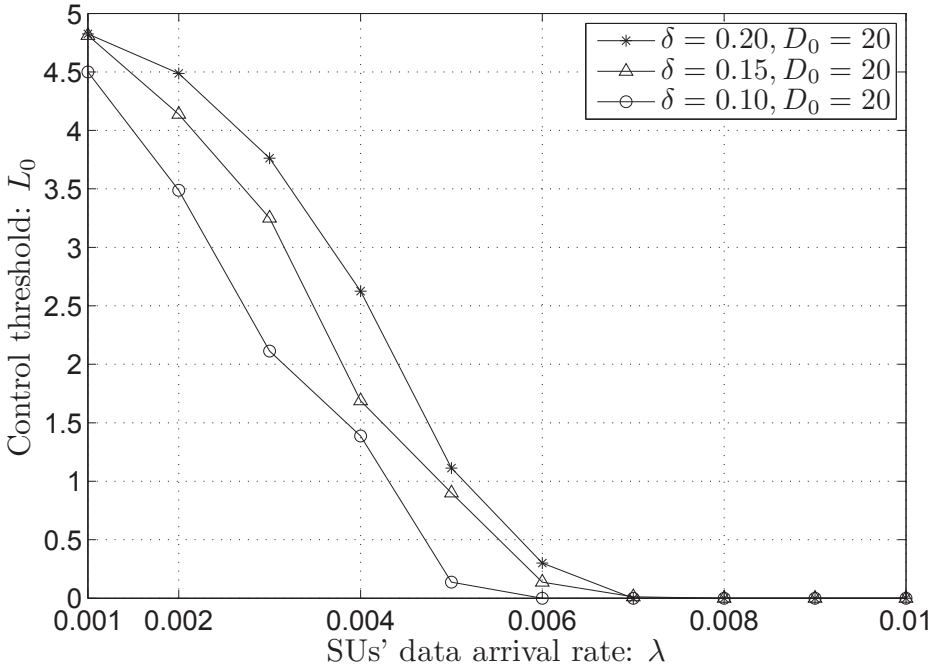

Figure 4.10: Control threshold L_0 under different SU arrival rates and delay requirements.

4.7 Conclusions

In this chapter, we have carried out delay analysis for a cognitive radio network. We took a stochastic fluid queue approach and modeled the system using Poisson driven stochastic differential equations. We characterized the moments of the queue lengths of SUs, for cases with a single PU channel and multiple PU channels. The impact of the PU traffic on SUs' queue lengths and the gain of using multiple PU channels were examined. Also, we explored the gain of using two interfaces per SU. Adaptive algorithms, using local information only, have been developed to find the optimal contention probabilities that achieve the minimum mean queue lengths and thus the minimum queueing delays of SUs.

Our analysis and numerical examples revealed that the mean queueing delay of SUs increases as the duty cycle of the PUs' traffic increases, pointing to the impact of PU activity on the delay performance of SUs. Also, when multiple PU channels were employed, we observed a decrease in the mean queueing delay, indicating a multi-channel gain. Moreover, if each SU is equipped with two interfaces, there is a decrease in the mean queueing delay because of the gain of using two choices.

Finally, we also studied packet generation control on the SUs, when delay constraints were imposed. We developed two control mechanisms, one randomized and the other utilizing SUs' queue lengths, and evaluated their performance.

Chapter 5

CONCLUSIONS AND DISCUSSIONS

5.1 Conclusions

In this disseration, we studied the opportunistic spectrum access in a cognitive radio network with an overlay model, where PUs have strictly higher priority over SUs in the spectrum usage. In particular, this dissertation consists of two main thrusts: In the first thrust, we investigated effective scheduling algorithms by either exploiting or inducing certain structure in the PU spectrum, and proposed optimal scheduling and smart spectrum tracking schemes to efficiently exploit these structures. Then, the second thrust was targeted at delay analysis in a cognitive radio network with random access, and the related issue of adaptive access control. In what follows, we provide concluding remarks for each of the works.

In Chapter 2, we studied opportunistic spectrum access for a single SU in a CR network with multiple PU channels. We formulated the problem as a partially observable Markov decision process, and examined the intricate tradeoffs in the optimal scheduling process, when incorporating the temporal correlation in *both* the channel fading and PU occupancy states. We modeled the channel fading variation with a two-state first-order Markov chain. The temporal correlation of PU traffic was modeled using "age" and a class of monotonically decreasing functions, which present a long memory. We identified a set of "multi-tier" tradeoffs arising from the intricate interactions between channel fading and PU occupancy states, and examined such tradeoffs via quantifying both immediate and total rewards for the SU. The optimality of the simple greedy policy was established under certain conditions. Further,

by developing a genie-aided system with full observation of the channel fading feedbacks, we decomposed and characterized the multiple tiers of the tradeoffs in the original system, and recorded the impacts from channel fading and PU occupancy on such tradeoffs under the general setting. Finally, we numerically studied the performance of various systems and/or policies and showed that the original system achieved an optimal total reward very close (within 1%) to that of the genie-aided system. Further, the optimal policy in the original system significantly outperformed randomized scheduling, as well as a policy that explores the memory in only the PU occupancy, pointing to the merit of jointly exploiting the temporal correlation in both of the system states. The studies we have initiated here can be generalized to cases with multiple SUs and/or nonidentical PU channels, where the fundamental tradeoffs associated with the interactions among various state elements identified in this work will be retained. We believe that our formulation and the insights obtained open up new horizons in better understanding spectrum allocation in cognitive radio networks.

Next, in Chapter 3, we considered a CR network with N PU channels and one SU, where each PU transmits packets to multiple receivers over lossy wireless channels via retransmission or network coding. Viewing network coding as a spectrum shaper, we showed that it increases the spectrum availability for the SU and offers a more predictive structure to the PU spectrum, i.e., it improves the SU's prediction of spectrum holes on PU channels. Based on the spectrum shaping effect of network coding, we developed different sensing strategies for the SU, where adaptive channel sensing is carried out by dynamically updating the list of the PU channels that are predicted by the SU to be idle. Our analysis and numerical results showed that compared to re-

transmission, both PU and SU's throughput can be improved when PUs apply network coding instead of retransmission, and the SU can further improve this gain by applying adaptive channel sensing (based on sensing history to reflect the PU traffic), with either fixed or randomized backoff parameter. Moreover, contention-based spectrum access of multiple SUs was considered and we developed a two-level backoff scheme based on the adaptive sensing to exploit the memory in network-coded transmissions and realize the throughput gains for the SUs. Further, we also showed that the predictability induced by network coding makes wireless networks more vulnerable to jamming attacks, and the performance degradation increases with the coding batch size. This suggests that a careful design is needed to balance the throughput performance and the jamming-resistant capability of the wireless systems.

Finally, in chapter 4, we carried out delay analysis for a cognitive radio network model. We took a stochastic fluid queue approach and modeled the system using Poisson driven stochastic differential equations. We characterized the moments of the queue lengths of SUs, for cases with a single PU channel and multiple PU channels. The impact of the PU traffic on SUs' queue lengths and the gain of using multiple PU channels were examined. Also, we explored the gain of using two interfaces per SU. Adaptive algorithms, using local information only, have been developed to find the optimal contention probabilities that achieve the minimum mean queue lengths and thus the minimum queueing delays of SUs. Our analysis and numerical examples revealed that the mean queueing delay of SUs increases as the duty cycle of the PUs' traffic increases, pointing to the impact of PU activity on the delay performance of SUs. Also, when multiple PU channels were employed, we observed a decrease in the mean queueing delay, indicating a multi-channel gain. Moreover, if each

SU is equipped with two interfaces, there is a decrease in the mean queueing delay because of the gain of using two choices. Finally, we also studied packet generation control on the SUs, when delay constraints were imposed. We developed two control mechanisms, one randomized and the other utilizing SUs' queue lengths, and evaluated their performances.

5.2 Discussions on Future Work

5.2.1 A Game-Theoretic View on Combating Jamming Attacks on Network-Coded Transmissions

As noted in Chapter 3, the predictability brought by network coding can make the wireless transmissions more vulnerable to malicious attacks. In particular, as previous analysis revealed, an increased network coding size can lead to more jammed transmissions, i.e., more severe damage. On the other hand, the throughput would be (desirably) improved if the coding size increases. Naturally, in the presence of the jamming, the legitimate user (e.g., PU in our previous context) must come up with corresponding strategies to combat the attacks and protect its throughput performance.

This interactive behavior between the user and the attacker can be modeled using game theory [95]. Consider a simple scenario with one legitimate user choosing from N independent channels for packet transmission using network coding, and a jammer aiming to attack the user's transmissions over time. Assume that the time is slotted. The key elements of the game are summarized as follows:

- *Players:* This game consists of two players: one legitimate user and one jammer.

- *Strategies:* The user and the jammer can choose various strategies given different system settings and constraints. We will give an example of strategy sets shortly in the following.

- *Payoffs:* The payoff for the user or the jammer is the utility they can achieve within a certain time frame. Payoffs should be functions of user strategies. The objective of each type of player is to maximize its payoff function.

One possible set of strategies for the players can be as follows: *1*) the user chooses which channel to transmit in each slot, where the channel conditions (say, SNR) is a random variable; and *2*) the jammer adjusts the probability $p_i(t)$ of attacking the transmission on channel i in each slot.

Based on the above strategies, the utility functions can be written out as: For the user:

$$U_s = E[T_{NC}] - C_s,$$

and for the jammer:

$$U_J = \sum_{t=1}^{E[T_{NC}]} \sum_{i=1}^{N} p_i(t)\mathbf{1}_t - C_J,$$

where T_{NC} denotes the completion time of transmitting (and successfully decoding) one batch of network-coded packets, which is now a function of the jammer's strategy as the packet erasure probability ε varies when an attack is launched on the channels; $\mathbf{1}_t$ is the indicator that the user is transmitting on the channel that jammer chooses in slot t; C_s is the cost for the user over the completion time, which can be obtained as:

$$C_s = \sum_{t=1}^{E[T_{NC}]} \sum_{i=1}^{N} p_i(t)\mathbf{1}_t,$$

and C_J is the energy cost for the jammer consumed during each successful attack, which can be written as:

$$C_J = \sum_{t=1}^{E[T_{NC}]} \mathbf{1}_t.$$

Note that the vector $\mathbf{p} \triangleq \{p_1(t), \ldots, p_N(t)\}$ can be simplified according to the specific system settings. For instance, if the user independently chooses *one* channel uniformly over time, we have $p_i(t) = \frac{1}{N}$. In some other cases, if the channel to be chosen in the next control slot is dependent on the previous history (and of course the memory in the channel condition variations), the expression for $p_i(t)$ will be a function of the varying parameters of the channel condition, subject to the constraint that $\sum_{i=1}^{N} p_i(t) = 1$. As a different example, if the user can access more than one channels at a time, the probabilities $\mathbf{p}$ would adapt accordingly as well.

5.2.2 Practical Considerations for Security Vulnerabilities in 802.22

Along with the research heat of CR networks, standardization effort has been made to pave the way for cognitive radios to be integrated into practical applications in the near future. IEEE 802.22 is the first wireless standard that incorporates detailed protocols for CR operations [96, 97].

Many practical security concerns emerge along with the standardization process. In what follows, we record some of these vulnerabilities, which deserve further investigation in future research. Note that the typical setting under the 802.22 standard is to have a base station (BS), serving multiple Consumer Premises Equipments (CPEs) which are essentially the same as the SUs studied in the dissertation.

- A key functionality for a CR network to successfully operate is to acquire spectrum availability information. As the standard specifies, three approaches are utilized to achieve this: spectrum sensing, geolocation (with the help of GPS to locate CPEs) and incumbent database. A BS relies on all these information to make informed decisions over time as to whether a frequency band is vacant or not. Clearly, attacks against any of the above functions can result in falsified information, thus potential interference from SU's transmission to that of the PU's. For instance, if the location information obtained via geolocation function is tampered, the BS will mistakenly consider a CPE to be far enough away from a PU and thus can carry out transmissions which could actually interfere the latter's communication severely.

- Superframe structure is adopted by 802.22. In particular, the BS sends special preamble and superframe control header (SCH) at the beginning of each superframe, utilizing which, the CPEs can associate with a particular BS during this frame. Any attacks on these preamble and SCH can lead to malfunction, or even worse, no service for the CPE in this superframe.

- 802.22 networks within radio range must synchronize their superframes to form self-coexistence. This is achieved by using beacons with time stamps (coexistence beacons). Falsifying these beacons (e.g., the time stamps) can destroy synchronization completely, and confuse CPEs who are trying to establish association with one of its nearby BS's.

- Quiet periods are used for spectrum sensing, and overlapping BS's must synchronize their quiet periods so that accurate sensing can be possible. Clearly, as soon as the synchronization of these quiet periods are

ruined, chaos arises in determining incumbent signals an thus the CPE's potential access to the PU channels is deprived of.

As many possible security vulnerabilities may appear, it is of great importance to come up with combating strategies while perfecting the standards. A detailed description of the IEEE 802.22 standard can be found in short surveys such as [96, 97].

5.2.3 Opportunistic Scheduling and Adaptive Learning

5.2.3.1 Opportunistic Spectrum Scheduling

In Chapter 2, opportunistic spectrum scheduling for a single SU has been studied. There are several extensions can be made building on this work:

- *Different Temporal Correlation Structure*: In our work, a novel concept of "age" was developed to capture the long memory in the PU's traffic. Based on this concept, a family of decreasing functions are defined that characterize the conditional probability of PU being idle or busy. A possible direction for future work is to consider a different set of functions, or an alternative formulation (i.e., use other quantities then the conditional probability functions) to capture this memory structure.

- *The Case with Multiple SUs – Fairness Considerations*: Another natural extension would be to examine the case when multiple SUs (say, M of them) are present. One interesting scenario is when the number of SUs is less than that of the PU channels, i.e., $M < N$, and the fairness among SUs is taken into consideration. Different fairness constraint can be included, such as max-min fairness, proportional fairness and so on.

- *Fully Distributed Scheduling*: In the current study, we have assumed the existence of a spectrum server that helps, to some extend, the scheduling process. It is also of interest to consider the case without spectrum server and the scheduling is fully distributed. However, caution must be taken in modeling the PU occupancy and channel fading states. Without careful design of the models, these two states may end up being combined into one "mega" state and therefore, the resulting study has the same flavor as those in the existing works.

5.2.3.2 Adaptive Learning of PU Traffic and Spectrum Environment

In addition, a further extension of this dissertation is to consider adaptive learning of PU traffic and spectrum environment. Throughout the dissertation, the SU's QoS performance is analyzed given certain knowledge of the PU traffic (e.g., the distribution of the packet generating process at PUs' queues) and/or the spectrum environment (e.g., the Markovian model of the channel fading state). Instead of such model-based approach, an interesting alternative is to take *learning-based* mechanisms where the SUs (and perhaps some intelligent attacker) learn the spectrum and traffic over time. To this end, different learning algorithms can be adopted and applied to all the topics that we have studied in the previous chapters as well as the above extensions:

- *Maximum-Likelihood Learning*: When the underlying channel model is known (such as the two-state Markov chain), while the statistics describing the model (such as the transition probabilities in the Markov chain) may not be accurate, the maximum-likelihood-based learning provides a way to learn and refine the model parameters, and therefore the model itself. For instance, a likelihood function of the transition parameter p,

$\Gamma(p)$, can be constructed by counting the number of transitions between the states of the Markov chain (see, e.g., [98]), and the optimal p can be obtained by setting $\frac{\partial \Gamma(p)}{\partial p} = 0$.

- *Q-Learning*: Q-learning [99] is a reinforcement learning approach applicable to scenarios when the underlying channel model is unknown, and the update is directly performed on the value function. A Q-learning algorithm establishes a mapping from the product space $\mathbf{S}_t \times \mathcal{A}_t$ to a real number, $Q(\mathbf{S}_t, a_t)$, denoting the "quality" of an action in a given state. At its core, the learning process carries along with an iterative value update as follows:

$$Q(\mathbf{S}_t, a_t) \leftarrow (1-\gamma_t) Q(\mathbf{S}_t, a_t) + \gamma_t \Big(R_t(\mathbf{S}_t, a_t) + \beta \max_{a_{t+1} \in \mathcal{A}_{t+1}} Q(\mathbf{S}_{t+1}, a_{t+1}) \Big), \tag{5.1}$$

where $\gamma_t \triangleq \gamma(\mathbf{S}_t, a_t)$ is the learning rate satisfying $0 < \gamma_t \leq 1$, and the value function at each iteration stage is chosen to be the maximum Q value, i.e., $V^*(\mathbf{S}_t) = \max_{a_t \in \mathcal{A}_t} Q(\mathbf{S}_t, a_t)$.

REFERENCES

[1] N. S. Shankar, C. Cordeiro, and K. Chanllapali, "Spectrum agile radios: utilization and sensing architectures," in *Proc. IEEE DySPAN'05*, Baltimore, MD, November 2005, pp. 160–169.

[2] "Report of the spectrum efficiency working group," FCC Spectrum Policy Task Force, Tech. Rep., November 2002.

[3] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access: signal processing, networking, and regulatory policy," *IEEE Signal Processing Magazine*, pp. 79–89, May 2008.

[4] Q. Zhao and A. Swami, "A survey of dynamic spectrum access: signal processing and networking perspectives," in *Proc. IEEE ICASSP'07*, Honolulu, HI, April 2007, pp. IV–1349–IV–1352.

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

[6] A. Jamalipour, "Cognitive heterogeneous mobile networks," *IEEE Wireless Communications Magazine*, vol. 15, no. 3, pp. 2–3, June 2008.

[7] I. F. Akyildiz, W.-Y. Leea, M. C. Vuran, and S. Mohantya, "Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, September 2006.

[8] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio newtorks: requirements, challenges and design trade-offs," *IEEE Communications Magazine*, pp. 32–39, April 2008.

[9] C.-X. Wang, H.-H. Chen, X. Hong, and M. Guizani, "Cognitive radio network management," *IEEE Vehicular Technology Magazine*, vol. 3, no. 1, pp. 28–35, March 2008.

[10] H. Zheng and C. Peng, "/collaboration and fairness in opportunistic spectrum access," in *Proc. IEEE ICC'05*, May 2005, pp. 3132–3136.

[11] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications, Special Issue on Cognitive Networks (Invited)*, vol. 23, pp. 201–220, 2005.

[12] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE transactions on wireless communications*, vol. 8, no. 1, pp. 297–307, 2009.

[13] S. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multi-channel opportunistic access," *IEEE transactions on Information Theory*, vol. 55, no. 9, pp. 4040–4050, 2009.

[14] N. B. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *Proc. ACM MOBIHOC'07*, 2007.

[15] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in WLAN channels: empirical model and its stochastic analysis," in *Proc. First International Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, August 2006.

[16] V. Brik, E. Rozner, S. Banerjee, and P. Bahl, "DSAP: a protocol for coordinated spectrum access," in *Proc. IEEE DySPAN'05*, Baltimore, MD, November 2005, pp. 611–614.

[17] M. M. Buddhikot, P. Kolodzy, S. Miller, kevin Ryan, and J. Evans, "DIMSUMNet: new directions in wireless networking using coordinated dynamic spectrum access," in *Proc. IEEE WoWMoM'05*, pp. 78–85.

[18] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 517–528, April 2007.

[19] L. Grokop and D. N. C. Tse, "Spectrum sharing between wireless networks," in *Proc. IEEE INFOCOM'08*, Arizona, USA, April 2008, pp. 201–205.

[20] J. Zhao, H. Zheng, and G.-H. Yang, "Distributed coordination in dynamic spectrum allocation networks," in *Proc. IEEE DySPAN'05*, Baltimore, MD, November 2005, pp. 259–268.

[21] L. Cao and H. Zheng, "Understanding the power of distributed coordination for dynamic spectrum management," *ACM Mobile Networks and Applications, Special Issue on Cognitive Radio Oriented Wireless Networks and Communications*, vol. 13, no. 5, pp. 477–497, October 2008.

[22] ——, "Distributed spectrum allocation via local bargaining," in *Proc. IEEE SECON'05*, September 2005, pp. 475–486.

[23] Z. Ji and K. J. R. Liu, "Dynamic spectrum sharing: a game theoretical overview," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 88–94, May 2007.

[24] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable markov processes over a finite horizon," *Operations Research*, vol. 21, no. 5, pp. 1071–1088, September 1973.

[25] A. Laourine, S. Chen, and L. Tong, "Queueing analysis in multichannel cognitive spectrum access: a large deviation approach," in *Proc. IEEE INFOCOM'10*, San Diego, CA, 2010.

[26] S. Chen and L. Tong, "Multiuser cognitive access of continuous time markov channels: maximum throughput and effective bandwidth regions," in *Proc. Workshop on Information Theory and its Applications*, San Diego, CA, February 2010.

[27] Q. Zhang and S. Kassam, "Finite-state markov model for rayleigh fading channels," *IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1688 – 1692, November 1999.

[28] E. N. Gilbert, "Capacity of burst-noise channels," *The Bell System Technical Journal*, pp. 1253–1265, September 1960.

[29] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multichannel opportunistic access: structure, optimality, and performance," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5431–5440, 2008.

[30] S. Murugesan, P. Schniter, and N. B. Shroff, "Multiuser scheduling in a markov-modeled downlink using randomly delayed arq feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1025 – 1042, February 2012.

[31] K. Kar, X. Luo, and S. Sarkar, "Throughput-optimal scheduling in multichannel access point networks under infrequent channel measurements," in *Proc. INFOCOM'07*, May 2007.

[32] Y. Chen, Q. Zhao, and A. Swami, "Distributed spectrum sensing and access in cognitive radio networks with energy constraint," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 783–797, February 2009.

[33] K. Liu and Q. Zhao, "Distributed learning in multi-armed bandit with multiple players," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5667 – 5681, November 2010.

[34] S. Geirhofer, L. Tong, and B. M. Sadler, "A measurement-based model for dynamic spectrum access in WLAN channels," in *Proc. IEEE Military Communications Conference (MILCOM)*, October 2006.

[35] F. C. Commission, "FCC frees up vacant TV airwaves for 'super wi-fi' technologies," FCC Spectrum Policy Task Force, Tech. Rep., September 2010.

[36] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40–48, April 2008.

[37] L. Johnston and V. Krishnamurthy, "Opportunistic file transfer over a fading channel: A pomdp search theory formulation with optimal threshold policies," *IEEE Transactions on Wireless Communications*, vol. 5, no. 2, pp. 394–405, 2006.

[38] X. Xiao, K. Liu, and Q. Zhao, "Opportunistic spectrum access in self similar primary traffic," *EURASIP Journal on Advances in Signal Processing: Special Issue on Dynamic Spectrum Access for Wireless Networking (invited)*, March 2009.

[39] B. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[40] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[41] S. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, February 2003.

[42] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger, "A random linear network coding approach to multicast," *IEEE transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, October 2006.

[43] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "Xors in the air: Practical wireless network coding," *IEEE/ACM Transaction on Networking*, vol. 16, no. 3, pp. 497–510, 2008.

[44] Y. Wu, P. Chou, Q. Zhang, K. Jian, W. Zhu, and S. Kung, "Network planning in wireless ad hoc networks: A cross-layer approach," *IEEE JSAC special issue on wireless ad hoc networks*, vol. 23, no. 1, pp. 136–150, January 2005.

[45] A. Eryilmaz, A. Ozdaglar, and M. Médard, "On the delay and throughput gains of coding in unreliable networks," *IEEE transactions on Information Theory*, vol. 54, no. 12, pp. 5511–5524, 2008.

[46] J. Kumar Sundararajan, D. Shah, and M. Médard, "ARQ for network coding," in *Proc. ISIT'08*, 2008, pp. 1651 – 1655.

[47] P. Larsson and N. Johansson, "Multi-user ARQ," in *Proc. VTC'06*, 2006, pp. 2052 – 2057.

[48] D. Lucani, M. Médard, and M. Stojanovic, "Random linear network coding for time-division duplexing: queueing analysis," in *Proc. IEEE ISIT'09*, July 2009, pp. 1423 – 1427.

[49] D. Nguyen, T. Tran, T. Nguyen, and B. Bose, "Wireless broadcast using network coding," *IEEE transactions on Information Theory*, vol. 58, no. 2, pp. 914–925, 2009.

[50] Y. E. Sagduyu and A. Ephremides, "On broadcast stability of queue-based dynamic network coding over erasure channels," *IEEE transactions on Information Theory*, vol. 55, no. 12, pp. 5463–5478, 2009.

[51] B. Fette, *Cognitive Radio Technology*. Academic Press, 2006.

[52] R. Thomas, D. Friend, L. DaSilva, and A. MacKenzie, *Cognitive Radio, Software defined radio, and adaptive wireless systems – Cognitive Networks*. Springer, 2007, pp. 17–41.

[53] A. M. Wyglinski, M. Nekovee, and T. Hou, Eds., *Cognitive Radio Communications and Networks: Principles and Practices*. Academic Press, 2010.

[54] R. L. Cruz, "A calculus for network delay, part I: network elements in isolation," *IEEE transactions on Information Theory*, vol. 37, no. 1, pp. 114–131, 1991.

[55] L. Cao and H. Zheng, "Stable and efficient spectrum access in next generation dynamic spectrum networks," in *Proc. INFOCOM'08*, 2008.

[56] N. Chang and M. Liu, "Competitive analysis of opportunistic spectrum access strategies," in *Proc. INFOCOM'08*, 2008.

[57] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE transactions on Information Theory*, vol. 54, no. 5, pp. 2053–2071, 2008.

[58] Y. Hou, Y. Shi, and H. Sherali, "Optimal spectrum sharing for multi-hop software defined radio networks," in *Proc. INFOCOM'07*, 2007.

[59] S. Huang, X. Liu, and Z. Ding, "Opportunistic spectrum access in cognitive radio networks," in *Proc. INFOCOM'08*, 2008.

[60] R. Tandra, M. Mishra, and A. Sahai, "What is a spectrum hole and what does it take to recognize one?" *IEEE special issue on Cognitive Radio*, pp. 824–848, May 2009.

[61] R. Urgaonkar and M. Neely, "Opportunistic scheduling with reliability guarantees in cognitive radio networks," in *Proc. INFOCOM'08*, 2008.

[62] N. Baldo, A. Asterjadhi, and M. Zorzi, "Cooperative detection and spectrum reuse using a network coded cognitive control channel," in *Proc. SECON Workshop'09*, 2009, pp. 1 – 7.

[63] M. Mishra, A. Sahai, and B. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. ICC'06*, 2006, pp. 1658 – 1663.

[64] D. Bertsekas and R. Gallager, *Data Networks, 2nd Edition.* Prentice Hall, 1992.

[65] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[66] P. Codenotti, A. Sprintson, and J. Bruck, "Anti-jamming schedules for wireless data broadcast systems," in *Proc. ISIT'06*, 2006, pp. 1856–1860.

[67] G. Lin and G. Noubir, "On link layer denial of service in data wireless LANs," *Wireless Communications and Mobile Computing*, vol. 5, no. 3, pp. 273–284, 2005.

[68] C. Fragouli, D. Katabi, A. Markopoulou, M. Médard, and H. Rahul, "Wireless network coding: Opportunities & challenges," in *Proc. MILCOM'07*, 2007.

[69] L. Geng, Y.-C. Liang, and F. Chin, "Network coding for wireless ad hoc cognitive radio networks," in *Proc. PIMRC'07*, 2007.

[70] P. Steenkiste, D. Sicker, G. Minden, and E. Dipankar Raychaudhuri, "Future directions in cognitive radio network research," NSF Workshop, Tech. Rep., 2009.

[71] S. Wang, J. Zhang, and L. Tong, "Delay analysis for cognitive radio networks with random access: a fluid queue view," in *Proc. IEEE INFOCOM'10*, San Diego, CA, 2010.

[72] S. Chen and L. Tong, "Multiuser cognitive access of continuous time markov channels: maximum throughput and effective bandwidth regions," in *Proc. UCSD Workshop on Information Theory and its Applications*, San Diego, CA, 2010.

[73] A. Ephremides and R.-Z. Zhu, "Delay analysis of interacting queues with an approximate model," *IEEE Transactions on Communications*, vol. 35, no. 2, pp. 194–201, 1987.

[74] E. Modiano and A. Ephremides, "A method for delay analysis of interacting queues in multiple access systems," in *Proc. IEEE INFOCOM'93*, vol. 2, San Francisco, CA, 1993, pp. 447–454.

[75] Y. Liu and W. Gong, "On fluid queueing systems with strict priority," *IEEE transactions on Automatic Control*, vol. 48, no. 12, pp. 2079–2088, December 2003.

[76] N. Bisnik and A. Abouzeid, "Queueing network models for delay analysis of multihop wireless ad hoc networks," in *Proc. International Conference on Wireless Communications and Mobile Computing'06*, Vancouver, Canada, 2006, pp. 773–778.

[77] D. Miorandi, A. A. Kherani, and E. Altman, "A queueing model for HTTP traffic over IEEE 802.11 WLANs," in *Proc. 16th ITC Specialist Seminar*, vol. 50, no. 1, 2004, pp. 63–79.

[78] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in *Proc. IEEE INFOCOM'04*, vol. 2, March 2004, pp. 1404–1413.

[79] T. K. Apostolopoulos and E. N. Protonotarios, "Queueing analysis of buffered slotted multiple access protocols," *Computer Communications*, vol. 8, no. 1, pp. 9–21, 1985.

[80] E. D. Sykas, D. E. Karvelas, and E. N. Protonotarios, "Queueing analysis of some buffered random multiple access schemes," *IEEE Transactions on Communications*, vol. 34, no. 8, pp. 790–798, 1986.

[81] T. Wan and A. U. Sheikh, "Performance and stability analysis of buffered slotted aloha protocols using tagged user approach," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 2, pp. 582–593, 2000.

[82] S. Meyn, *Control Techniques for Complex Networks*. Cambridge University Press, 2008.

[83] D. Anick, D. Mitra, and M. M. Sondhi, "Stochastic theory of a data-handling system with multiple sources," *The Bell System Technical Journal*, vol. 61, no. 8, pp. 1871–1894, 1982.

[84] A. Eryilmaz, P. Marbach, and A. Ozdaglar, "A fluid-flow model for backlog-based CSMA policies," in *Proc. IEEE WICON'08*.

[85] A. Eryilmaz and R. Srikant, "Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control," in *Proc. IEEE INFOCOM'05*, vol. 3, Miami, USA, March 2005, pp. 1794–1803.

[86] X. Li, Q. Zhao, X. Guan, and L. Tong, "Optimal cognitive access of markovian channels under tight collision constraints," *to appear in IEEE J. Select. Areas in Communications, Special Issue on Advances in Cognitive Radio Networks and Communications, 2011.*

[87] W. Luo and A. Ephremides, "Stability of N interacting queues in random access systems," *IEEE transactions on Information Theory*, vol. 45, no. 5, pp. 1579–1587, July 1999.

[88] R. R. Rao and A. Ephremides, "On the stability of interacting queues in a multiple access system," *IEEE transactions on Information Theory*, vol. 34, no. 5, pp. 918–930, September 1988.

[89] R. W. Brockett, "Lecture notes: stochastic control," Harvard University, Tech. Rep., 1983.

[90] R. W. Brockett, W. Gong, and Y. Guo, "Stochastic analysis for fluid queueing systems," in *Proc. IEEE Decision and Control'99*, Phoenix, AZ, 1999, pp. 3077–3082.

[91] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, "New insights from a fixed point analysis of single cell ieee 802.11 WLANs," *IEEE/ACM Transactions on Networking*, pp. 588 – 601, June 2007.

[92] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications, 2nd ed.* Springer, 2003.

[93] Y. Azar, A. Broder, A. Karlin, and E. Upfal, "Balanced allocations," in *Proc. 26th ACM Symposium on the Theory of Computing*, vol. 2, 1994, pp. 593–602.

[94] M. D. Mitzenmacher, "The power of two choices in randomized load balancing," Ph.D. dissertation, University of California at Berkeley, 1996.

[95] M. J. Osborne and A. Rubinstein, *A Course in Game Theory.* The MIT Press, 1994.

[96] C. Cordeiro, K. Challapali, D. Birru, and S. S. N, "Ieee 802.22: An introduction to the first wirless standard based on cognitive radios," *Journal of Communications*, vol. 1, no. 1, pp. 38–47, 2006.

[97] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shelhamme, and W. Caldwell, "Ieee 802.22: The first cognitive radio wirless regional area network standard," *IEEE Communications Magazine*, pp. 130–138, 2009.

[98] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Anti-jamming games in multi-channel cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 4–15, 2012.

[99] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, pp. 279–292, 1992.