AN ABSTRACT OF THE DISSERTATION OF

MohammadJavad NoroozOliaee for the degree of Doctor of Philosophy in Computer Science presented on August 10, 2015.

Title: Design and Analysis of Distributed Resource Allocation Techniques for Large-Scale Dynamic Spectrum Access Networks

Abstract approved: _____

Bechir Hamdaoui

The widespread use of wireless devices that we have recently been witnessing, such as smartphones, tablets, laptops, and wirelessly accessible devices in general, is causing an unprecedented growth in the required amount of the wireless radio spectrum. On the other hand, the spectrum resource has, for the last several decades, been allocated to spectrum users in static manners. These static allocation methods were found to be very inefficient. This inefficient use of the spectrum resource, coupled with its limited availability nature, has led to a shortage in the spectrum supply. Consequently, the concept of dynamic spectrum access (DSA) has emerged as an alternative allocation approach with great potential for solving this shortage problem. In the DSA context, there are two types of users: primary users (PUs) and secondary users (SUs). While PUs have exclusive access rights to use their licensed spectrum bands at all time, SUs are allowed to use these bands only opportunistically. That is, prior to using any licensed band, SUs must first sense the band to make sure that it is vacant. When a PU returns while

SUs are using its band, SUs must vacate immediately. In the first part of this dissertation, we address the problem of resource management in DSA networks. Specifically, we develop resource and service management techniques to support SUs with certain QoS (Quality of Service) requirements in large-scale DSA networks. The proposed techniques empower SUs to seek and exploit spectrum opportunities dynamically and effectively, thereby maximizing the SUs' long-term received service satisfaction. Our proposed techniques are efficient in terms of optimality, scalability, distributivity, and fairness. In the second part of this dissertation, we model, characterize, and analyze the key performance metrics of these DSA networks. Specifically, we use a continuous-time Markov process to derive and analyze the forced termination and blocking probabilities of SUs under two realistic limitations. First, we investigate the impact of the channel handoff agility limitation on the performance. Here, due to this agility limitation, which is imposed essentially by hardware restrictions, SUs can only switch to their immediate neighboring channels whenever they have to due to, for e.g., the return of a PU. We show that such a limitation has great impact on the achievable performances of DSA networks. Second, we study the impact of the adjacent channel interference on the performance of DSA schemes, which often rely on the use of guard bands to handle such an interference. We model and analyze the impact of guard band deployment methods on the performance of DSA schemes. Our study serves as design guidelines for choosing appropriate guard band deployment methods when designing DSA schemes.

Design and Analysis of Distributed Resource Allocation Techniques for Large-Scale Dynamic Spectrum Access Networks

by

MohammadJavad NoroozOliaee

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented August 10, 2015
Commencement June 2016

Doctor of Philosophy dissertation of MohammadJavad NoroozOliaee presented on August 10, 2015.

APPROVED:

_____

Major Professor, representing Computer Science

_____

Director of the School of Electrical Engineering and Computer Science

_____

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

MohammadJavad NoroozOliaee, Author

To Fatemeh and Hasan,

ACKNOWLEDGEMENTS

No word conveys how much I am grateful to the people who helped me to reach this point in my life. Words are never powerful enough to show my gratitude and feelings, although they are the only means I have. Nevertheless, I wish I would be able to express a small portion of my appreciation to the people who helped me.

I would love to express my sincere gratitude to my advisor Bechir Hamdaoui, for his unconditional support and valuable guidance during my PhD. It is an honor for me to be working with him for six years which is undoubtedly the best part of my life. I am also extremely grateful to the EECS Associate Head, Bella Bose, for his invaluable support. I will never forget how kind he is to me and how much he helped me during these years.

I would also like to thank Kagan Tumer, Mohsen Guizani, Taeib Znati, Ammar Rayes and Xiuzhen Cheng for all of their valuable assistance in this work. My committee members, Alan Fern, Sinisa Todorovic, Scott Sanner, Huaping Liu and Hector Vergara cannot be thanked enough for being helpful and patient during the PhD process . I wish to express my gratitude to Abdollah Tavakoli Farsoni for serving on my masters and doctorate committee as well. I know thanking him is not enough for his tremendous support and help. I want to thank the School of EECS staff members, particularly, Ferne Simendinger, Nicole Thompson and Colisse Franklin for their availability and invaluable guidance and assistance during my studies at OSU.

Friendship is the most precious thing we have in our lives and my friends play an important role in my life. I am very thankful to each and every one of them.

Last and most, I would like to thank my family. My wife, Hanieh has been extremely patient and supportive of me throughout this entire process and has made countless

sacrifices to help me get to this point. I dedicate this dissertation to my parents without them this dissertation would not exist. My dear sisters and brother, Shirin, Farzaneh and Jalal who are always encouraging me whenever I feel exhausted.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

Chapter 1: Introduction

1.1 Dynamic Spectrum Access

The unprecedented growth in the use of wireless devices, applications and services has created high demand for the wireless radio spectrum resource. This growth has also led to the development of new communication techniques, such as multiple-input multiple-output (MIMO) [50, 52–54] and cooperative communication [45], and to the emergence of new wireless technologies, such as WiFi [46, 49, 51], LTE [30], and femtocell networks [22, 23] to be able to meet this high demand. The spectrum supply, on the other hand, has not been keeping up with this demand, thereby creating an unexpected supply shortage. The reason for this shortage is reported to be due not to the scarcity of the spectrum, but rather to the inefficiency of current spectrum allocation methods, thus leaving plenty of unused spectrum opportunities along both the time and frequency dimensions. In effect, recent studies conducted by FCC (US federal communication commission) and other regulatory bodies revealed vast temporal and geographical variations in the utilization of the licensed spectrum, ranging from 15% to 85% [1, 10, 69]. The results of these studies persuaded regulatory bodies to urge for more efficient and adaptive spectrum allocation policies. Consequently, FCC has since been revising and improving its regulations to allow and promote opportunistic and dynamic access to the spectrum. Cognitive radio, proposed by Joseph Mitola III in 1999 [98], emerges as the potential technology that can enable such an opportunistic and dynamic access capability.

This dissertation has two major parts. In the first part, we begin by addressing the

problem of resource management in Dynamic Spectrum Access (DSA) networks, and in the second part, we use a continuous-time Markov process to model, characterize and analyze key performance metrics of these DSA networks.

1.2 Design of Efficient Resource Allocation Schemes

During this past decade, DSA has attracted significant research attention from academia, industry and government agencies, resulting in numerous works ranging from protocol design and optimization [3, 21, 27, 35, 37, 42, 57, 62, 71, 85, 91–93, 114, 118, 122, 140] to market-oriented resource management strategies and architectural paradigms [13, 17, 25, 32, 63, 70, 74, 76, 100, 133, 138]. Research attention has also been paid to the development of prediction model-based channel selection techniques that can adapt themselves to enable effective DSA [8, 26, 84, 88, 90, 127, 130, 131]. One key challenge with relying on prediction models, however, is that the unique characteristics of DSA environment make it too difficult to construct accurate models that can capture DSA dynamics, unless some unrealistic assumptions are made, leading to inaccurate prediction of the environment's behaviors. As a result, there have also been some efforts on developing learning-based techniques that do not require such models, yet can still perform well by learning directly from interaction with the environment [12, 19, 44, 64, 86, 89, 128, 129]. Instead of using models, these techniques rely on learning algorithms [119, 121] to learn from the past and present interaction experience to decide what to do best in the future.

In the first part of this dissertation, we address the problem of resource management in DSA networks. Specifically, we develop resource and service management techniques to support secondary users (SUs) with certain QoS (Quality of Service) requirements in large-scale DSA networks. The proposed techniques empower SUs to seek and ex-

ploit spectrum opportunities dynamically and effectively, thereby maximizing the SUs' long-term received service satisfaction levels. Our techniques are efficient in terms of optimality, scalability, distributivity, and fairness.

1.3 Performance Analysis of Cognitive Radio Networks

Research attention has also been given to deriving models and studying behaviors of the cognitive radio and dynamic spectrum access performance [37, 55, 58, 91, 135]. Generally, most of these performance studies model cognitive radio access by means of Markov chains, and use these models to derive and analyze network performances. For example, in [9, 117, 126, 132, 137], Markov chains are used to model and study the forced termination and blocking probabilities of SUs in a cognitive multichannel access system consisting of primary and cognitive users. However, one common unrealistic assumption made in these existing works that we address in this dissertation is that SUs, when accessing the multichannel system opportunistically, are allowed to switch/jump to any available channel in the system, regardless of the frequency gap between the target and the current channels [65]. However, due to hardware limitations, SUs can actually jump only so far from where the operating frequency of their current channel is, given an acceptable switching delay that users are typically constrained by [66]. Therefore, in Chapter 4, we study the performance of cognitive radio networks, but while considering realistic channel switching (or handoff) agility, where SUs can only switch to channels that are immediate neighbors of their current operating channels.

Additionally, many works have been done in the literature, proposing models for performance evaluation of CRNs (e.g., [16, 72, 114, 136]). In most of these works, adjacent-channel interference (ACI) is often not considered, requiring thus ideal transmission

filters. The lack of ideal filters, as in the case of real systems, causes, however, spectrum spill-over. Therefore, in order to protect adjacent PU and CU transmissions, there should be a frequency separation. Such a separation is referred to as a guard band. However, using guard bands restricts effective spectrum utilization. CUs should account for the impact of guard bands when choosing channels for their transmission. The good news is that when two contiguous channels belong to the same CU, there is no need for a guard band between them. Throughout, a set of contiguous channels assigned to the same CU is called a frequency block.

The ACI impact on CRN performances has previously been studied in [136], where a centralized solution for adaptive guard-band setting was proposed. The proposed solution uses a dynamic guard-band configuration to minimize ACI, thus requiring a central server for frequency planning. In [136], the authors did not consider channel aggregation. In Chapter 5, we consider CU transmissions to be carried over multiple contiguous (i.e., bonded) or noncontiguous (i.e., aggregated) available bands. Experimental studies were conducted to show the benefits of channel bonding and aggregation in the context of CRNs (e.g., [24, 75, 81]). In [24], the authors show the impact of using channel bonding on the performance of an IEEE 802.11 network. They experimentally proved that channel bonding can significantly improve the network performance in terms of throughput, transmission range, and power consumption. However, they did not consider channel aggregation in their work. In their work, they simply considered single-hop scenarios. The authors in [75] model an ad-hoc opportunistic spectrum access network with channel bonding as a Markov process in order to study the throughput performance. Their results show that channel bonding generally enhances network performance under certain circumstances. Network size and the number of available channels determine the level of enhancement. In their model, nevertheless, the authors did not consider the guard band

constraint. Another study investigates the possibility of opportunistic spectrum access under strict limit on PUs' service rate [81]. The study shows that reliable communications cannot be guaranteed by traditional spectrum access policies for CUs. Therefore, their results proposed channel aggregation as a means to provide reliable communications for CUs. The papers mentioned above do not incorporate the guard band constraints in their design. Therefore, in Chapter 5, we model a multichannel CRN access as a continuous-time Markov process under the realistic assumption of non-ideal filters (i.e., guard bands are used). We then use our model to derive system performance metrics under various channel assignment schemes. The second part of this dissertation, which includes Chapter 4 and Chapter 5, uses a continuous-time Markov process analysis to model and characterize the performance of DSA networks while accounting for the two aforementioned physical limitations.

1.4   Contributions and Dissertation Organization

1.4.1   Research Contributions

We summarize the contributions of this dissertation in this section.

- In Chapter 2 and 3, we derive and propose efficient objective functions for SUs that are aligned with system objective in that when SUs aim to maximize them, their collective behaviors also lead to good system-level performance, thereby resulting in increasing each SU's long-term received rewards. More specifically, we propose objective functions that are (*i*) near-*optimal*, in that they allow SUs to achieve rewards close to the maximal achievable rewards, (*ii*) *scalable*, in that they perform well in systems with a small as well a large number of users, (*iii*) *learnable*,

in that they allow SUs to reach up near-optimal rewards very quickly, and (*iv*) *distributive*, in that they are implementable in a decentralized manner by relying on local information only. We want to emphasize that the focus of this Chapter is not on learning, but rather on the design of objective functions that promote coordination, so we used existing learning techniques to achieve our goals. We also emphasize that, in this work, we benefit from the concepts and techniques that were previously used in multi-agent coordination but has not been used in this context.

- In Chapter 4, we study the performance of cognitive radio networks, but while considering realistic channel switching (or handoff) agility, where SUs can only switch to channels that are immediate neighbors of their current operating channels. We modelled the system under the aforementioned constraint as a continuous Markov process since it has been popular in the literature [9, 117, 126, 132, 137]. This makes the accuracy of the results in our work comparable to the results of previous works. Results show that for a fixed primary user load, the forced access termination probability of cognitive users decreases significantly as the number of target handoff channels increases. The results also show that the gap between the forced termination probabilities for different numbers of target handoff channel set sizes increases with the primary user arrival rate. Unlike the forced termination probability, the blocking probability is shown not to depend on the level of spectrum agility. Also, we observe that as the primary user arrival rate increases, the cognitive spectrum access efficiency reduces, regardless of the level of handoff agility. The spectrum efficiency depends, however, on spectrum agility level, and the higher the agility, the higher the spectrum efficiency. To summarize, Chapter 4

makes the following contributions: 1) performance modelling of cognitive radio network access with limited spectrum handoff agility, 2) validation of the derived analytic results via computer simulations, and 3) study and analysis of the impact of spectrum handoff agility on the performance behaviors of dynamic spectrum access networks.

- In Chapter 5, we again model a multichannel CRN access as a continuous-time Markov process under the realistic assumption of non-ideal filters (i.e., guard bands are needed). We then use our model to derive system performance metrics, such as blocking probability, forced termination probability, degradation probability, CU utilization, service degradation, spectrum efficiency, under five different channel assignment schemes. These channel selection schemes are designed to improve some of the system performance metrics, such as spectrum efficiency and service degradation. Furthermore, we consider two scenarios of CRNs: FDM and D-OFDM. While in a FDM-based CRN, two neighboring CU transmissions are not allowed to share the same guard band, in the D-OFDM-based CRN, neighboring CU transmissions are allowed to do so.

1.4.2   Dissertation Organization

This dissertation is composed of six chapters. Chapter 1 provides an introduction to DSA networks along with the challenges that emerge in the design of resource allocation schemes in such networks. It also highlights our major contributions gained in this research. In Chapter 2 and Chapter 3, we derive and evaluate private objective functions for large-scale, distributed opportunistic spectrum access (OSA) systems. By

means of any learning algorithms, the derived objective functions enable OSA users to assess, locate, and exploit unused spectrum opportunities effectively by maximizing the users' average received rewards. Chapter 4 studies the performance of cognitive radio networks while considering realistic channel handoff agility, where cognitive users can only switch to their neighboring channels, whereas in Chapter 5, we use continuous-time Markov analysis to model and analyze the adjacent channel interference and its impact on CRN performance under different spectrum sharing schemes. Finally, in Chapter 6, we conclude by summarizing and highlighting the main contributions and results of the work presented in this dissertation.

Chapter 2: Efficient Objective Functions for Coordinated Learning in Large-Scale Distributed OSA Systems

In this Chapter, we derive and evaluate private objective functions for large-scale, distributed opportunistic spectrum access (OSA) systems. By means of any learning algorithms, these derived objective functions enable OSA users to assess, locate, and exploit unused spectrum opportunities effectively by maximizing the users' average received rewards. We show that the proposed objective functions are: near-*optimal*, as they achieve high performances in terms of average received rewards; highly *scalable*, as they perform well for small- as well as large-scale systems; highly *learnable*, as they reach up near-optimal values very quickly; and *distributive*, as they require information sharing only among OSA users belonging to the same band.

2.1 INTRODUCTION

Federal Communications Commission (FCC) foresees opportunistic spectrum access (OSA) as a potential solution to the spectrum shortage problem [94, 95]. Essentially, OSA improves spectrum efficiency by allowing unlicensed or secondary users (SUs) to exploit unused licensed spectrum, but in a manner that limits interference to licensed or primary users (PUs). That is, prior to using a licensed band, SUs must first sense the band to check its vacancy, and if it is vacant, they can then use it for as long as no PUs are present. When any PUs return to their band, SUs must vacate immediately.

During the past few years, due to its apparent promises, OSA has created significant research efforts, resulting in numerous works ranging from protocol design and performance optimization [27, 37, 42, 71, 85, 91, 114, 116, 118, 122, 140] to market-oriented management strategies and architectural paradigms [25, 32, 70, 74, 76, 100, 133, 138]. More recently, research efforts have also been given to the development of optimal channel selection techniques that rely on spectrum availability prediction models to adapt themselves to the environment so as to promote effective OSA [8, 26, 78, 84, 88, 90, 127]. The challenge, however, is that the OSA environment has very unique characteristics that make it too difficult to construct models that can capture its dynamics without making assumptions about the environment itself. Such assumptions are often unrealistic, leading to inaccurate prediction of the environment's behaviors.

As a result, there have also been some efforts on developing learning-based techniques that do not require such models, yet can still perform well by learning directly from interaction with the environment [11, 19, 64, 86, 89, 102, 128, 129]. Instead of using models, these techniques rely on learning algorithms (e.g., reinforcement learners [119, 121] and evolving neuro-controllers [4, 20]) to learn from past and present interaction experience to decide what to do best in the future. In essence, learning algorithms allow SUs to learn by interacting with the environment, and use their acquired knowledge to select the proper actions that maximize their own (often selfish) objective functions, thereby "hopefully" maximizing their long-term cumulative received rewards.

The key challenge that we address in this work is that when SUs' objective functions are not carefully coordinated, learning algorithms can lead to poor performances in terms of SUs' long-term received rewards. In other words, when SUs aim at maximizing their intrinsic (not carefully designed) objective functions, their collective behavior often leads to worsening each other's long-term cumulative rewards, a phenomenon known as the

"tragedy of the commons" [68]. It is, therefore, imperative that objective functions be designed carefully so that when SUs maximize them, their collective behavior does not result in worsening each other's performance.

With this in mind, in this work, we derive efficient objective functions for SUs that are aligned with system objective in that when SUs aim to maximize them, their collective behaviors also lead to good system-level performance, thereby resulting in increasing each SU's long-term received rewards [105, 107]. More specifically, we propose objective functions that are (*i*) near-*optimal*, in that they allow SUs to achieve rewards close to the maximal achievable rewards, (*ii*) *scalable*, in that they perform well in systems with a small as well as a large number of users, (*iii*) *learnable*, in that they allow SUs to reach up near-optimal rewards very quickly, and (*iv*) *distributive*, in that they are implementable in a decentralized manner by relying on local information only.

We want to emphasize that the focus of this Chapter is not on learning, but rather on the design of objective functions that promote coordination, and that can be used by any learning algorithms.

The rest of the Chapter is organized as follows. In Section 2.2, we present the spectrum and traffic models. Section 2.3 states the motivation as well objective of this work. In Section 2.4, we propose the objective functions. In Section 2.5, we derive upper bounds on the maximal achievable Section 2.6, we evaluate the proposed functions. In Section 2.7, we discuss the applicability of the proposed techniques to other models. conclude the paper in Section 2.8.

2.2 SYSTEM MODEL

We assume that spectrum is divided into m non-overlapping bands, and that each band is associated with many PUs. We consider a time-slotted system, where PUs are assumed to arrive and leave at the beginning and at the end of time slots. An *OSA agent*, or simply *an agent*, is a group of two or more SUs who want to communicate together. In order to communicate with each other, all SUs in the group must be tuned to the same band. At the end of each time step, by means of a reinforcement learning algorithm [119], each agent selects the "best" available spectrum band, and uses it during the next time step. This process repeats at each step.

At each time step, each agent receives a service that is passed to it from the environment/system. One possible service metric is the amount of throughput that the visited spectrum band offers the agent. Another possible metric is the reliability of the communication carried on the spectrum band, which can be measured through, for example, SNR (signal to noise ratio), PSR (packet success rate), etc. What service metric to use and how to quantify it are beyond the scope of this work. Here, we assume that once the agent switches to a particular band, the received service level can immediately be quantified by monitoring the metric in question. Hereafter, we then assume that each band j is characterized by a value V_j that represents the maximum/total service level that the band can offer.

In this Chapter, we consider the elastic traffic model, where the agent's received reward (i.e., satisfaction) increases proportionally to the service it receives from using the spectrum band so long as the received quality-of-service (QoS) level is higher than a certain (typically low) threshold R. But when the received QoS level is below the threshold R, the agent's reward decreases rapidly (e.g., exponentially) with the received

QoS level; i.e., the reward/satisfaction goes almost immediately to zero when the received QoS level is below R. This traffic model is suitable for elastic applications, such as file transfer and web browsing, where the higher the received service quality level, the better the quality/reward perceived by these applications. But when the received QoS level is below a certain low threshold (i.e., R), the quality of these applications becomes unacceptable. Formally, the reward $r_j[n_j[t]]$ (also often referred to simply as $r_j[t]$ for simplicity of notation) contributed by band j at time step t can be written as:

$$r_j[n_j[t]] = \begin{cases} V_j/n_j[t] & \text{if} \quad n_j[t] \leq V_j/R \\ Re^{-\beta \frac{n_j[t]R-V_j}{V_j}} & \text{otherwise} \end{cases} \quad (2.1)$$

where $n_j[t]$ denotes the number of agents that choose band j at time step t, and β is a reward decaying factor. Note that here we assume that the total service level V_j offered by any band j is split equally among all the $n_j[t]$ agents that use band j at time t. In Fig. 2.1, we show for the sake of illustration the agent reward $r_j[n_j[t]]$ as a function of the number of agents $n_j[t]$ when $\beta = 50$ and $V_j/R = 4$.

From the system's perspective, the system or global reward can be regarded as the sum of all agents' received rewards. Formally, at any time step t, the global reward $G[t]$ is

$$G[t] = \sum_{j=1}^{m} n_j[t]r_j[n_j[t]] \quad (2.2)$$

where m is the number of spectrum bands. If each agent i selecting band j receives the reward function contribution $r_j[t]$ for that band (a "local reward"), then the per-agent average reward $\bar{r}[t]$ at time step t is

$$\bar{r}[t] = \frac{\sum_{j=1}^{m} n_j[t]r_j[t]}{\sum_{j=1}^{m} n_j[t]} = \frac{G[t]}{\sum_{j=1}^{m} n_j[t]} \quad (2.3)$$
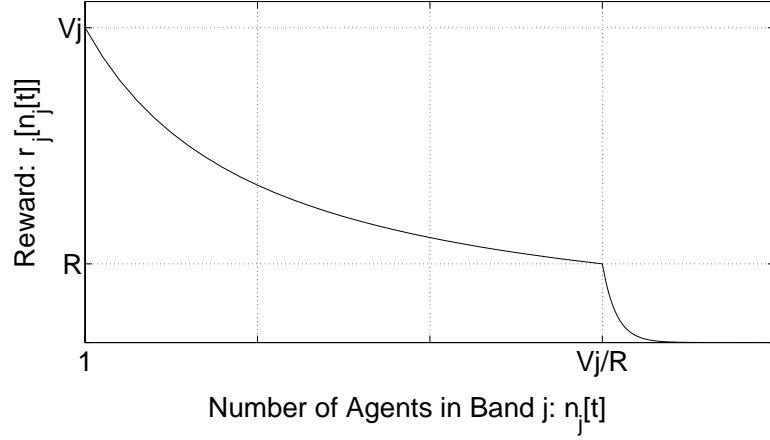
Figure 2.1: Elastic reward function: $\beta = 50$ and $V_j/R = 4$.

2.3 MOTIVATION AND OBJECTIVE

The goal of this work is to design efficient objective functions for OSA agents, so that when agents aim to maximize them, their collective behaviors lead to good system-level performance, thereby resulting in increasing each agent's long-term received rewards. Hereafter, let g_i denote agent i's objective function. Although the objective functions (g_i for agent i) that we derive in this Chapter are designed to be used by any learning algorithm, throughout this work, we choose to use the ϵ-greedy Q-learner [119] (with a discount rate of 0 and an ϵ value of 0.05) for the purpose of evaluating the effectiveness of our developed functions only. At each episode (or time step) t, each agent i aims then at maximizing its own private objective function $g_i[t]$ using its own Q-learner.

At the end of every episode, each agent selects and takes the action with the highest entry value with probability $1 - \epsilon$, and selects and takes a random action among all possible actions with probability ϵ. After taking an action, the agent then computes the reward that it receives as a result of taking such an action (i.e., as a result of using

the selected band), and uses it to update its Q-table. A table entry $Q(a)$ corresponding to action a is updated via $Q(a) \leftarrow (1 - \alpha)Q(a) + \alpha u$, where α (here, the value of α is set to 0.5) is the learning rate, and u is the received reward from taking action a. All the results presented in this Chapter are based on this Q-learner. Readers are referred to [119] for more details on the Q-learner.

Again, we want to emphasize that the focus of this work is not on learning, but rather on designing optimal objective functions that OSA agents can aim to maximize, and that can be used by any learners.

2.3.1   Motivation

The key question that arises naturally is which objective function g_i should each OSA agent i aim to maximize so that its received reward is maximized? There are two intuitive choices that one can think of. One possible objective function choice is for each agent i using band j to selfishly go after the intrinsic reward r_j contributed by the band j as defined in Eq. (2.1); i.e., $g_i = r_j$ for each agent i using band j. A second also intuitive choice is for each agent to maximize the global (i.e., total) rewards received by all agents; i.e., $g_i = G$ for each agent i as defined in Eq. (2.2), hoping that maximizing the overall received rewards will eventually lead to maximizing every agent's long-term average received rewards.

For illustration purposes, we plot in Fig. 2.2 the per-agent average received reward $\bar{r}[t]$ (measured and calculated via Eq. (2.3)) under each of these two private objective function choices. In this experiment, we consider an OSA system with a total number of agents equal to 500 and a total number of bands m equal to 10. There are two important observations that we want to make regarding the performance behaviors of these two
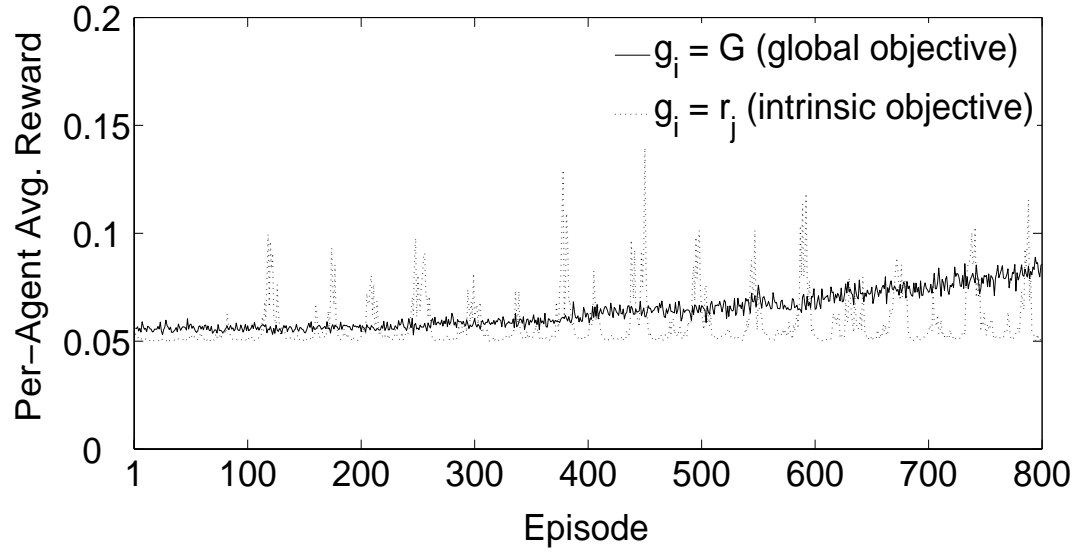
Figure 2.2: Per-agent average achieved reward $\bar{r}[t]$ as a function of episode t under the two private objective functions: intrinsic choice ($g_i = r_j$) and global choice ($g_i = G$) for $R = 2$, $\beta = 2$, $V_j = 20$ for $j = 1, 2, \ldots, 10$.

objective functions, and that constitute the main motivation of this work. First, note that when agents aim to maximize their own intrinsic rewards (i.e., $g_i = r_j$ for each agent i using band j), the per-agent average received reward presents an oscillating behavior: it ramps up quickly at first but then drops down rapidly too, and then starts to ramp up quickly and drop down rapidly again, and so on, which explains as follows. With the intrinsic objective function, an agent's reward, by design, is sensitive to its own actions, which enables it to quickly determine the proper actions to select by limiting the impact of other agents' actions, thus learning about good spectrum opportunities fast enough. However, agents' intrinsic objectives are likely not to be aligned with one another, which explains the sudden drop in their received reward right after learning about good opportunities; i.e., right after their received reward becomes high.

The second observation is regarding the second objective function choice, G. Observe that, unlike the intrinsic function, when each agent i sets its objective function g_i to the global reward function G, this results in a steadier performance behavior where the per-agent average received reward increases continuously, but slowly. With this function choice, agents' rewards are aligned with one another by accounting for each other's actions, and thus are less (or not likely to be) sensitive to the actions of any particular agents. The alignedness feature of this function is the reason behind the observed monotonic increase in the average received reward. However, the increase in the received reward is relatively slow due to the function's insensitivity to one's actions, leading to slow learning rates.

Therefore, it is imperative that private objective functions be designed with two (usually conflicting) requirements in mind: (*i*) *alignedness*; when agents maximize their own private objectives, they should not end up working against one another; instead, their collective behaviors should result in increasing each agent's long-term received

rewards, and (*ii*) *sensitivity*; objective functions should be sensitive to agents' own actions so that proper action selections allow agents to learn about good opportunities fast enough.

With this in mind, we now state our objective in this work.

2.3.2 Objective

The objective of this work is to design efficient private objective functions for large-scale, distributed OSA systems. More specifically, we aim to devise private objective functions with the following design requirements and objectives. First, they should be optimal in that they should enable agents to achieve high rewards. Second, they should be scalable in that they should perform well in OSA systems with a small as well as a large number of agents. Third, they should be learnable in that they should enable OSA agents to find and locate spectrum opportunities quickly. Fourth, they should be distributive in that they should be implementable in a decentralized manner. The objective functions that we derive in this work meet all of these design requirements.

2.4 OBJECTIVE FUNCTION DESIGN

We first begin by summarizing the concepts of factoredness and learnability, both of which are essential for capturing as well as ensuring the two required design properties: alignedness and sensitivity. Then, we derive and present efficient objective functions that meet the above design requirements by striking a good balance between alignedness and sensitivity.

2.4.1 Factoredness and Learnability

Again, let g_i denote the function that OSA agent i aims to maximize as its objective, and that we want to derive. Let z characterize the joint move of all OSA agents in the system. Here, the global (total) reward, G, is a function of z, which specifies the full system state (G can then precisely be written as $G(z)$). Hereafter, we use the notation $-i$ to specify all agents other than agent i, and z_i and z_{-i} to specify the parts of the system state controlled respectively by agent i and agents $-i$. The system state z can then be written as $z = z_i + z_{-i}$.

For the joint actions of multiple OSA agents to lead to good overall average reward, two requirements must be met. First, we must ensure that an OSA agent aiming to maximize its own private objective function also leads to maximizing the global (total achievable) rewards, so that its long-term average received rewards are indeed maximized. This means that the agents' private objective functions ($g_i(z)$ for agent i) need to be "aligned" or "factored" with the global reward function ($G(z)$) for a given system state z. Formally, for systems with discrete states, the degree of *factoredness* of a given private objective function g_i is defined as [6]:

$$\mathcal{F}_{g_i} = \frac{\sum_z \sum_{z'} h[(g_i(z) - g_i(z'))\,(G(z) - G(z'))]}{\sum_z \sum_{z'} 1} \tag{2.4}$$

for all z' such that $z_{-i} = z'_{-i}$, where $h[x]$ is the unit step function, equal to 1 if $x > 0$, and zero otherwise. Intuitively, the higher the degree of factoredness of an agent's private objective function g_i, the more likely it is that a change of state will have the same impact on both the agent's (i.e., local) and the total (i.e., global) received rewards. A system is fully factored when $\mathcal{F}_{g_i} = 1$.

Second, we must ensure that each OSA agent can discern the impact of its own actions on its private objective function, so that a proper action selection allows the agent to quickly learn about good spectrum opportunities. This means that the agent's private objective function should be more sensitive to its own actions than the actions of other agents. Formally, the level of sensitivity or *learnability* of a private objective function g_i, for agent i at z, can be quantified as [6]:

$$\mathcal{L}_{i,g_i}(z) = \frac{E_{z_i'}[|g_i(z) - g_i(z_{-i} + z_i')|]}{E_{z_{-i}'}[|g_i(z) - g_i(z_{-i}' + z_i)|]} \tag{2.5}$$

where $E[\cdot]$ is the expectation operator, z_i''s are parts of the system states, controlled only by agent i, that are resulting from agent i's alternative actions at z, and z_{-i}''s are parts of the system states, controlled by agent $-i$, that are resulting from agent $-i$'s alternative joint actions. So, at a given state z, the higher the learnability, the more $g_i(z)$ depends on the move of agent i. Intuitively, higher learnability means that it is easier for an agent to achieve higher rewards.

Unfortunately, these two requirements are often in conflict with one another [6]. Therefore, the challenge in designing objective functions for large-scale OSA systems is to find the best tradeoff between factoredness and learnability. Doing so will ensure that agents can learn to maximize their own objectives while doing so will also lead to good overall system performance; i.e., their collective behaviors will not result in worsening each other's received rewards.

2.4.2   Efficient Objective Functions

The selection of an agent's private objective function that provides the best performance hinges on balancing the degree of factoredness and the level of learnability. In general and as discussed in the previous section, a highly factored private objective function will experience low learnability, and a highly learnable function will have low factoredness [6].

To provide some intuition on how we designed our proposed functions, we will again revisit to the behaviors of the global reward function, illustrated earlier in Section 2.3.1. Recall that when agents set the global reward G as their objective functions (i.e., $g_i = G$ for each agent i), their collective behaviors did indeed result in increasing the total system achievable rewards (i.e., did result in a fully factored system), because agents' private objectives are aligned with system objective. The issue, however, is that because G depends on all the components of the system (i.e., all agents), it is too difficult for agents (using G as their objective functions) to discern the effects of their own actions on their objectives, resulting then in low learnability rates.

The key observation leading to the design of our functions is that by removing the effects of all agents other than agent i from the function G, the resulting agent i's private objective function will have higher learnability than G, yet without compromising its alignedness quality. Formally, these functions can be written as

$$D_i(z) \quad \equiv \quad G(z) - G(z_{-i}) \tag{2.6}$$

where z_{-i} again represents the parts of the state on which agent i has no effect. These difference functions have also been shown to lead to good system performance in other domains, such as multi-robot control [7] and air traffic flow regulation [125]. First, note

that these proposed functions (D_i for agent i) are fully factored, because the second term of Eq. 2.6 does not depend on agent i's actions. On the other hand, they also have higher learnability than G, because subtracting this second term from G removes most of other agents' effects from agent i's objective function. Intuitively, since the second term evaluates the value of the system without agent i, subtracting it from G provides an objective function (i.e., D_i) that essentially measures agent i's contribution to the total system received rewards, making it more learnable without compromising its factoredness quality.

By substituting Eq. (2.2) into Eq. (2.6), explicitly noting the time dependence t, and for clarity, removing the implicit dependence on the full state z, the objective function D_i for agent i selecting band j at time t can then be written as:

$$
\begin{aligned}
D_i[t] &= \sum_{k=1}^{m} n_k[t] r_k[n_k[t]] - \left(\sum_{k=1,k\neq j}^{m} n_k[t] r_k[n_k[t]] + (n_j[t] - 1) r_j[n_j[t] - 1] \right) \\
&= n_j[t] r_j[n_j[t]] - (n_j[t] - 1) r_j[n_j[t] - 1]
\end{aligned} \tag{2.7}
$$

It is important to note that, by taking away agent i from the second term of the function D_i, the terms corresponding to all spectrum bands k, except the band j that agent i is using, cancel out. This explains why D_i (as shown in Eq. (2.7)) depends on band j only. Therefore, the proposed function D_i is simpler to compute than the global function G. More specifically and importantly, it is fully decentralized as agents implementing/using it as their objectives need to gather and share information only with the agents that belong to the same band. This constitutes one important property among few others (to be described later) that this proposed function has. Next, we formally prove the claims that we made regarding the performances of the proposed function.

Proposition 2.4.1 *The function D_i is fully factored.*

Proof Differentiating both sides of Eq. (2.6) w.r.t. agent i's state z_i yields $\frac{\partial}{\partial z_i} D_i(z) = \frac{\partial}{\partial z_i} G(z) - \frac{\partial}{\partial z_i} G(z_{-i})$, which in turn yields $\frac{\partial}{\partial z_i} D_i(z) = \frac{\partial}{\partial z_i} G(z)$ since $\frac{\partial}{\partial z_i} G(z_{-i}) = 0$.

Proposition 2.4.2 *The expected value of the learnability of D_i is higher than the expected value of the learnability of G.*

Proof We now sketch this proof. From Eq (2.5), the inner term of the numerator of D_i's learnability is equal to $D_i(z) - D_i(z_{-i} + z_i')$, which, Eq. (2.6), can also be written as $G(z) - G(z_{-i}) - (G(z_{-i} + z_i') - G(z_{-i}))$ or equivalently as $G(z) - G(z_{-i} + z_i')$. Hence, the numerator of the learnability is the same for D_i and G. Therefore, any gains in learnability must come from the denominator. Now, for a state z where agent i picked band j and a state z' where it did not, the inner term of the denominator of D_i's learnability is:

$$
\begin{aligned}
DEN_{\mathcal{L},D} &= D_i(z) - D_i(z'_{-i} + z_i) \\
&= n_j g_j[n_j] - (n_j - 1)g_j[n_j - 1] - \left((n_j' + 1)g_j[n_j' + 1] - n_j' g_j[n_j']\right)
\end{aligned}
$$

where we dropped the t terms for clarity and where n_k' is the number of agents that choose band k in the alternate state z'. That is the denominator consists of two terms, representing two bands *that differ by only one user.* Now, let us focus on the denominator for the learnability of G for a state z where agent i picked band j and a state z' where

it picked band k:

$$
\begin{aligned}
DEN_{\mathcal{L},G} &= G(z) - G(z'_{-i} + z_i) \\
&= \sum_{l=1, l\neq j, l\neq k}^{m} n_l g_l[n_l] - n'_l g_l[n'_l] \\
&+ \quad n_k g_k[n_k] - (n'_k - 1)g_k[n'_k - 1] \\
&+ \quad n_j g_j[n_j] + (n'_j + 1)g_j[n'_j + 1]
\end{aligned}
$$

Now, here, there are also two terms, representing two bands (j and k) *that differ by only one user.* The expected magnitude of these values will be the same as those for the *only two terms for $DEN_{\mathcal{L},D}$*. However, there are $m-2$ terms that differ by as many as the total number of agents minus 1. As a consequence, we have $E[DEN_{\mathcal{L},G}] >> E[DEN_{\mathcal{L},D}]$ leading to D having much higher learnability on average than does G.

2.5   Optimal Achievable Rewards

In this section, we derive a theoretical upper bound on the maximum/optimal achievable rewards. This upper bound will serve as a means of assessing how well the developed objection functions perform when compared not only with intuitive objective functions; i.e., intrinsic ($g_i = r_j$ for each agent i using band j) and global ($g_i = G$ for each agent i), but also with the optimal achievable performances.

Without loss of generality and for simplicity, let us assume that $V_j = V$ for $j = 1, 2, \cdots, m$. Let n denote the total number of agents in the system at any time. First, note that when $n \leq m\frac{V}{R}$, the maximum global achievable reward is simply equal to mV (assume $n \geq m$), which corresponds to having each band contain no more than $\frac{V}{R}$ agents. Therefore, in what follows, we assume that $n > m\frac{V}{R}$, and let $c = \frac{V}{R}$, which denotes the

capacity (in terms of number of supported OSA agents) of each spectrum band.

Now, we start by proving the following lemma, which will later be used for proving our main result.

Lemma 2.5.1 *The global received reward of an OSA system reduces less when a new OSA agent joins a more crowded spectrum band than when it joins a less crowded band.*

Proof Recall that when a band j has $n' > c$ agents, its reward is $G_j(n') = n'Re^{-\beta(\frac{n'}{c}-1)}$. If a new agent joins this band, the new reward becomes $G_j(n'+1) = (n'+1)Re^{-\beta(\frac{n'+1}{c}-1)}$. First, it can easily be shown that when $n' > c \geq 1$, $G_j(n') > G_j(n'+1)$; i.e., the reward when joining band j decreases by $\Delta_j(n') \equiv G_j(n') - G_j(n'+1)$. Now we can easily see that $\Delta_j(n')$ decreases when n' increases. Hence, the greater the number n' (i.e., the more crowded the band), the smaller the decrease in reward.

Theorem 2.5.2 *When there are n agents in the system, the global reward reaches its maximal only when $m-1$ bands (out of the total m bands) each has exactly c agents, and the m-th band has the remaining $n - c(m-1)$ agents.*

Proof Let $k = n - mc$, and let us refer to the agent distribution stated in the theorem as C. Note that C corresponds to when $m-1$ bands each has exactly c agents and the other m-th band has the remaining $c+k$ agents (since $n - c(m-1) = c+k$). We proceed with the proof by comparing C with any possible distribution C' among all possible distributions. Let $c+k_1$ be the number of agents in the most crowded band in C', $c+k_2$ be the number of agents in the second most crowded band in C', and so forth. We just need to deal with the bands that each contains more than c agents. If there are p bands each containing more than c agents, then we know that $\sum_{i=1}^{p} k_i \geq k$.

For each band having $c + k'$ agents, let ϵ_i be the amount by which the global reward is reduced when agent i joins the band for $i = 1, 2, \cdots, k'$. From Lemma 2.5.1, it follows that $\epsilon_i > \epsilon_{i+1} > 0$, for all $i = 1, 2, \cdots, k' - 1$.

Note that for the distribution C, the global reward is reduced by $t = \sum_{i=1}^{k} \epsilon_i$, and for C', it is reduced by $t' = \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i$. It remains to show that $t' - t > 0$ for any $C' \neq C$. We consider three different scenarios:

- $k_1 > k$: Here, we have

$$
\begin{aligned}
t' - t &= \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \\
&= \sum_{i=k}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i
\end{aligned}
$$

which is greater than zero.

- $k_1 = k$: In this scenario, we have

$$
\begin{aligned}
t' - t &= \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \\
&= \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i
\end{aligned}
$$

which is also greater than zero.

- $k_1 < k$: In this scenario, we have

$$
\begin{aligned}
t' - t &= \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \\
&= \underbrace{\sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i}_{part\ a} - \underbrace{\sum_{i=k_1}^{k} \epsilon_i}_{part\ b}
\end{aligned}
$$

Since $k_1 + k_2 + \cdots + k_p \geq k$, the number of ϵ_i terms in *part a* is greater than the number of terms in *part b*. From Lemma 2.5.1, we know that the largest term in *part b* is ϵ_{k_1}, which is smaller than the smallest term ϵ_{k_2} in *part a*. Hence, *part a* is greater than *part b*, and thus $t' - t$ is greater than zero.

In all scenarios, we showed that $t' - t > 0$. Therefore, the global reward for any distribution C' is smaller than that for the distribution C; i.e., C is the distribution that corresponds to the maximal global achievable reward.

Corollary 2.5.3 *The per-agent average achievable reward is at most* $(m-1)V/n + (R - (m-1)V/n)e^{-\beta(\frac{nR}{V} - m)}$.

Proof The proof follows straightforwardly from Theorem 2.5.2 by calculating the global achievable reward for the derived optimal agent distribution.

Note that this upper bound (that we derived and stated in Corollary 2.5.3) is the maximum/optimal average reward that an agent can achieve (it is a theoretical upper bound). In the next section, we will evaluate the performances of the proposed objective functions in terms of their achievable rewards, and compare them against these optimal achievable performances.

2.6 PERFORMANCE EVALUATION

In this section, we evaluate and compare the performances of the proposed objective functions in terms of the per-agent average achievable rewards with the optimal achievable rewards calculated through Corollary 2.5.3 as well as with those achievable under each of the two intuitive functions r_j and G.

2.6.1 Static OSA Systems

We first consider evaluating the proposed objective functions under the same experiment, conducted in Section 2.3.1, where again the total number of agents considered in this experiment is equal to 500, and that of bands is equal to $m = 10$. Here, the simulated system is static, in that all agents are assumed to enter and leave the OSA system time. In this section, we ignore the PUs' activities. Dynamic OSA systems as well as PUs' activities will be considered later in Section 2.6.2.

Fig. 2.3 shows the per-agent average achievable reward normalized w.r.t. the optimal achievable reward under each of the three functions: intrinsic ($g_i = r_j$), global ($g_i = G$), and proposed ($g_i = D_i$). The figure clearly shows that the proposed function D_i achieves substantially much better performances than the other two. In fact, when using D_i, an agent can achieve up to about 90% of the total possible, achievable reward, whereas it only can achieve up to about 20% when using any of the other two functions. Another distinguishing feature of the proposed D_i function lies in its learnability; that is, not only does D_i achieve good rewards, but also it does so quite fast, as the received rewards ramp up rapidly, quickly reaching near-optimal performance.

We also study the proposed function with regard to another performance metric:
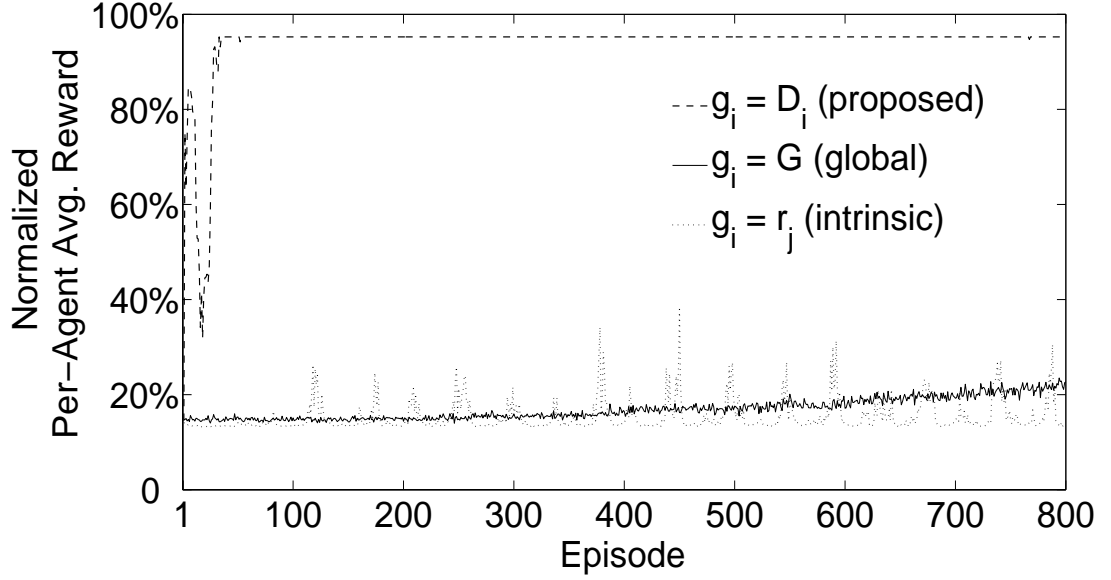
Figure 2.3: Per-agent average achieved reward normalized w.r.t. maximum achievable reward under intrinsic function ($g_i = r_j$), global function ($g_i = G$), and proposed function ($g_i = D_i$): $R = 2$, $\beta = 2$, $V = 20$.

scalability. For this, we plot in Fig. 2.4 the per-agent average achievable reward under each of the three studied objective functions when varying the number of OSA agents, n, from 100 to 800 while keeping the number of bands $m = 10$ the same. Observe that D_i outperforms the other two functions substantially when it also comes to scalability. Note that D_i achieves high rewards, even for large numbers of agents, whereas the achievable reward under either of the other two functions drops dramatically with the number of agents. We therefore conclude that the proposed function D_i is very scalable, and works well in systems with small as well as large numbers of agents.

To summarize, the obtained results show that the proposed objective function (i) achieves near-*optimal* performances in terms of per-agent average rewards, (ii) is highly *scalable* as it performs well for small- as well as large-scale systems, (iii) is highly *learnable*
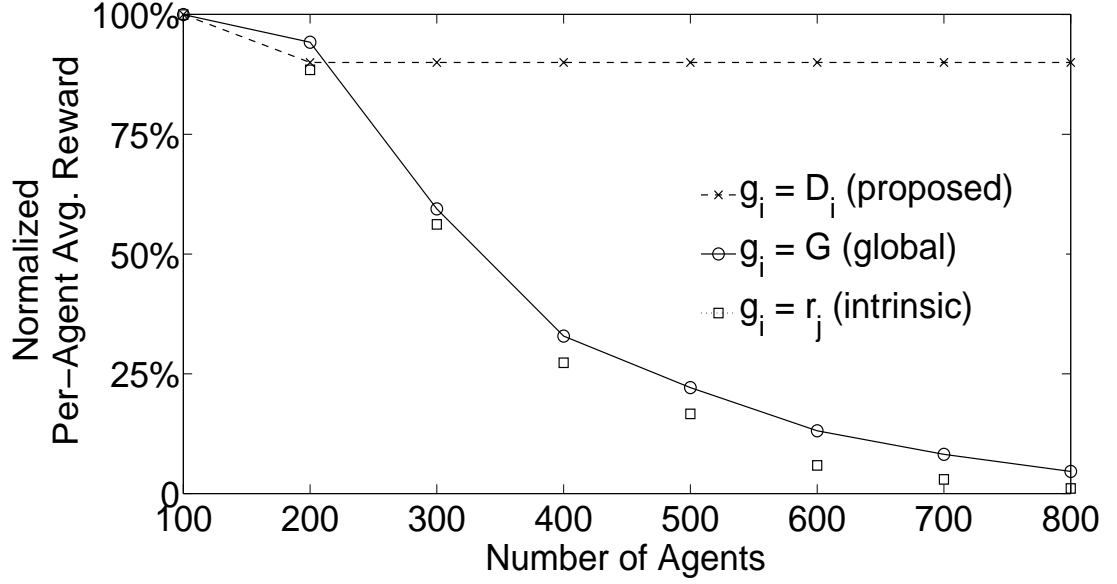
Figure 2.4: Per-agent average achieved reward normalized w.r.t. maximum achievable reward under intrinsic ($g_i = r_j$), global ($g_i = G$), and proposed ($g_i = D_i$) functions for various numbers of agents: $R = 2$, $\beta = 2$, $V = 20$.

as rewards reach up near-optimal values very quickly, and (iv) is *distributive*, as it requires information sharing only among agents belonging to the same band.

2.6.2 Dynamic OSA Systems

In this section, we further assess how well these objective functions perform (i) by considering dynamic OSA networks (both with and without PUs' activities), and (ii) by investigating another important performance metric, fairness, in addition to the optimality and scalability metrics.

Recall that in the previous section, we considered an OSA system in which all agents enter the system, use the available spectrum bands, and then leave the system all at

the same time. That is, the number of agents does not change over time, and remains the same during the course of the entire system's lifetime. In this section, we want to investigate how well these obtained results hold when considering dynamic systems, in which agents enter and leave at different independent times. OSA agents can then be viewed as data sessions/flows that start and finish at different times, and independently from one another. We will study these dynamic systems (i) without as well as (ii) with the presence of PUs' activities, and see how well these functions behave under such systems.

2.6.2.1   Without PUs' Activities

To mimic the dynamic behaviors of OSA agents, we assume that agents (e.g., data sessions) arrive according to a Poisson process with arrival rate λ, and stay in the system for an exponentially distributed duration of mean μ. Let $\kappa = \lambda\mu$ represent the average number of agents that are using the system at any time. In this section, we first begin by studying dynamic OSA systems without considering the presence of PUs. The impact of PUs will be investigated in the next section.

Fig. 2.5 shows the normalized per-agent average received reward for $\frac{\lambda}{\mu} = 1$ when κ equals 500 (Fig. 2.5(a)), 750 (Fig. 2.5(b)), and 1000 (Fig. 2.5(c)). The figure shows that the proposed function D_i still possesses its distinguishing performance features/trends even under dynamic behaviors. These trends are as follows. First, note that the function D_i receives near-optimal rewards, which are more than 90% of the maximal achievable rewards. Second, it is highly learnable as it reaches up near-optimal behaviors quite fast. Third, it is highly scalable as the achievable rewards do not drop below $\approx$ 90% of the maximal achievable reward even when the average number of agents in the system is

increased from 500 to 1000. Fourth, it outperforms the other two functions significantly, and this is regardless of the number of agents.

For completeness, we also study these same performances under different values of the ratio $\frac{\lambda}{\mu}$. Recall that for a given number of agents, the higher the ratio $\frac{\lambda}{\mu}$, the shorter the sessions' durations. For example, when $\kappa = 500$, $\frac{\lambda}{\mu} = 1$ implies that the sessions' average duration μ and arrival rate λ are both equal to ≈ 22.3, whereas $\frac{\lambda}{\mu} = 5$ implies that $\mu = 10$ and $\lambda = 50$. Figs. 2.6 and 2.7 show the normalized per-agent average received reward for two more ratios: $\frac{\lambda}{\mu} = 5$ and $\frac{\lambda}{\mu} = 20$. Observe that the performances of D_i are still close to optimal, and are much higher than those obtained under r_j and G regardless of the ratio $\frac{\lambda}{\mu}$; i.e., whether $\frac{\lambda}{\mu}$ equals 1, 5, or 20.

Fairness is also another important performance metric to evaluate. We want to assess how fair D_i is when compared with the other two functions. Fig. 2.8 depicts the coefficient of variations (CoV)[1] of the per-agent average received rewards under the three studied functions for various combinations of arrival rates λ and durations μ of OSA agents. Observe that D_i achieves CoV values similar to those achievable under any of the other two functions, and this is regardless of sessions' durations and arrival rates. In other words, these results show that the proposed function, when used in practical, dynamic network settings, not only does it achieve good performances in terms of optimality, scalability, and learnability, but it does so while reaching a fairness quality as good as those reached through the two other functions. Next, we show that this is also true in the presence of primary users.

---

[1]CoV is the ratio of the standard deviation to the mean of the agents' received rewards; we use this metric as a means of assessing the fairness, which reflects how close agents' received rewards are to one another.

(a) $\kappa = 500$

(b) $\kappa = 750$

(c) $\kappa = 1000$

Figure 2.5: Normalized per-agent average received reward under Poisson arrival traffic model: $\lambda/\mu = 1$

(a) $\kappa = 500$
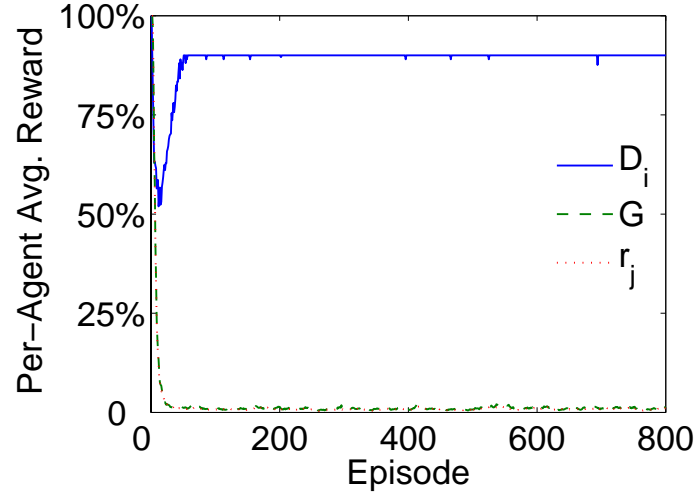
(b) $\kappa = 750$
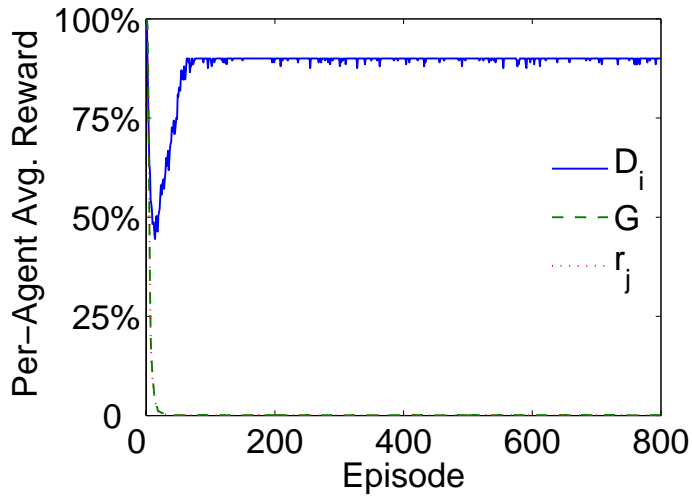
(c) $\kappa = 1000$

Figure 2.6: Normalized per-agent average received reward under Poisson arrival traffic model: $\lambda/\mu = 5$
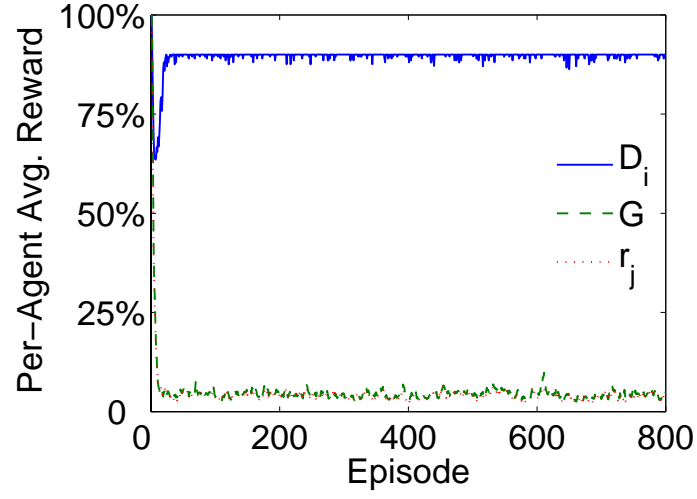
(a) $\kappa = 500$
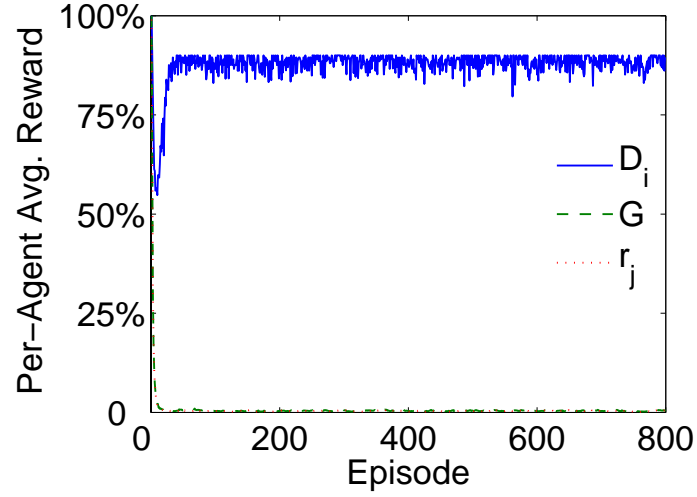
(b) $\kappa = 750$
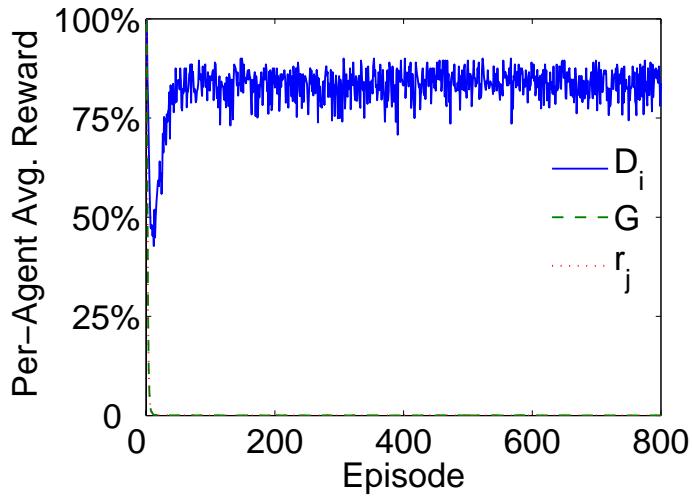
(c) $\kappa = 1000$

Figure 2.7: Normalized per-agent average received reward under Poisson arrival traffic model: $\lambda/\mu = 20$
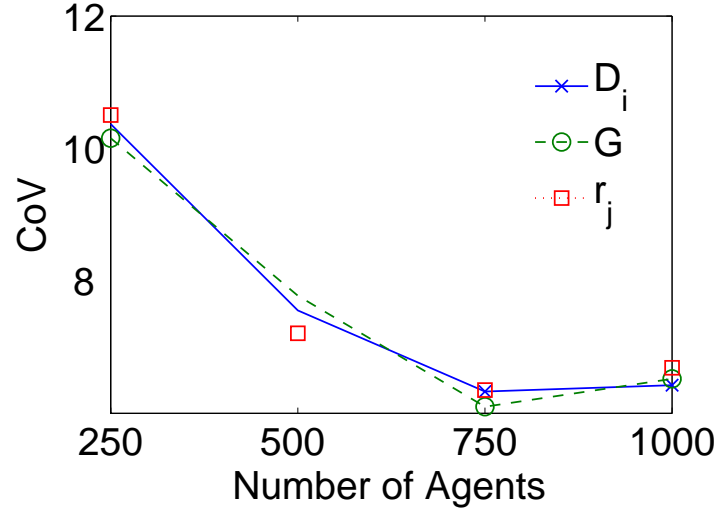
(a) $\lambda/\mu = 1$
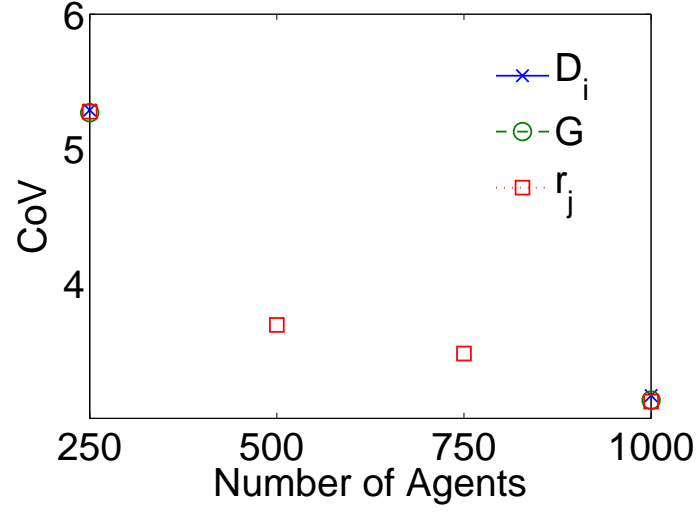
(b) $\lambda/\mu = 5$
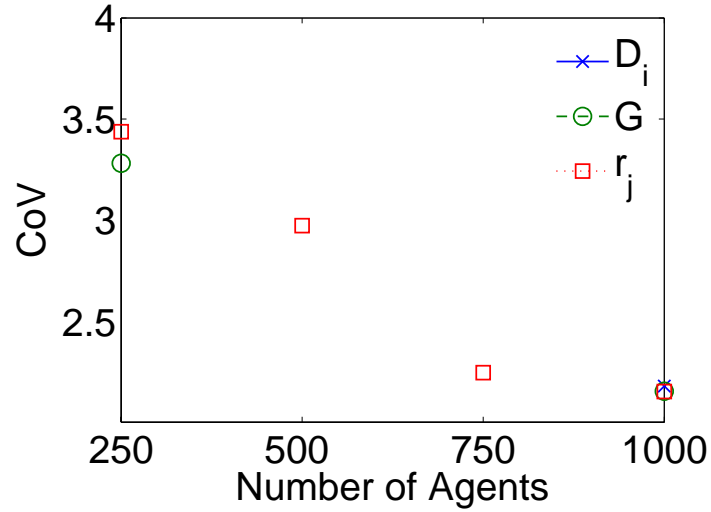
(c) $\lambda/\mu = 20$

Figure 2.8: Coefficient of variation (CoV) of normalized per-agent average reward under r_j, G, and D_i: dynamic OSA system without PUs' presence

2.6.2.2   With PUs' Activities

We now study the impact of the presence of PUs on the performance of the studied objective functions. In this experiment, we consider the same dynamic model used in Section 2.6.2.1 but while assuming and accounting for PUs' activities. We assume that each band is associated with a set of PUs that enter and leave the band at random times. We model PUs' activities on each band as a renewal process alternating between ON and OFF periods, which represent the time during which PUs are respectively present (ON) and absent (OFF). For each spectrum band j, we assume that ON and OFF durations are exponentially distributed with means ν_j^{ON} and ν_j^{OFF}, respectively[2]. We use $\eta_j \equiv \nu_j^{ON}/(\nu_j^{OFF} + \nu_j^{ON})$ to denote the *PU traffic load* on band j.

Figs. 2.9, 2.10, and 2.11 show the normalized per-agent average reward for PU traffic loads η of 10%, 30%, and 50%, respectively ($\eta = \eta_j \ \forall j$). The subfigures in each figure each corresponds to a different average number of agents (i.e., for $\kappa = 500$, $\kappa = 750$, and $\kappa = 1000$). There are three observations that we can make out of these results. First, observe that the proposed function D_i achieves rewards higher than those achieved under the other two functions, and for all combinations of number of agents κ and primary user loads η. Second, unlike in the case of dynamic OSA without primary users, as expected, the achievable rewards reach zero when primary users are present, but quickly reach up high values as soon as PUs leave their bands. Third, also as expected, when the PU traffic load increases (i.e., η is increased), the total achievable average reward decreases, since rewards will also be taken away by PUs themselves. For example, Fig. 2.11 shows that when the PU load $\eta = 50\%$, D_i reaches up to only about 50% of the maximal achievable reward. This is because 50% of the total reward has already been received by the

---

[2]Recall that learners do not actually need prior knowledge of primary users' traffic behavior. Here, the exponential distributions will be used to generate samples so as to be able to mimic the OSA environment.
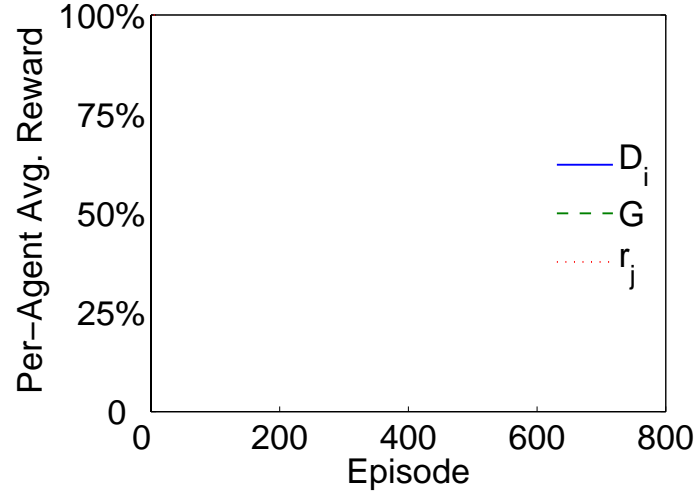
(a) $\kappa = 500$
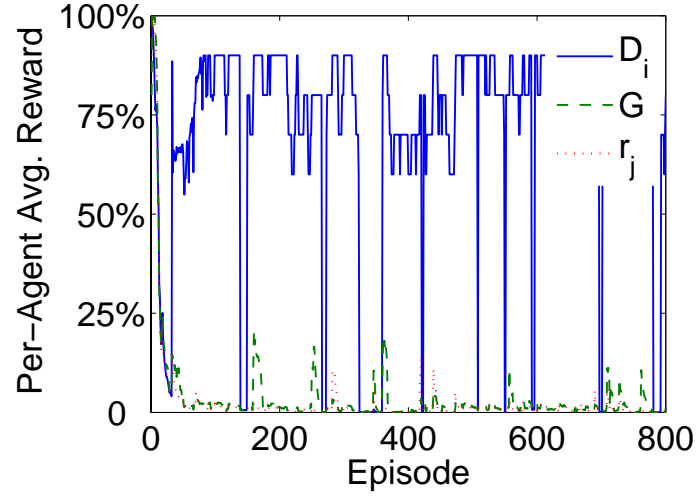
(b) $\kappa = 750$
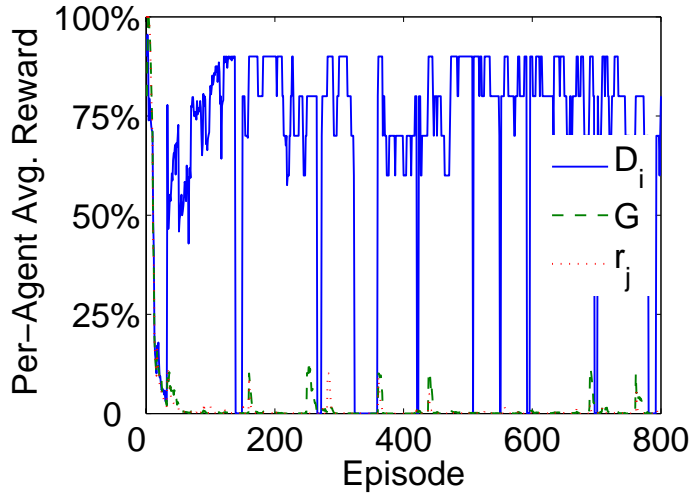
(c) $\kappa = 1000$

Figure 2.9: Normalized per-agent average reward under OSA agent traffic with Poisson arrival of $\frac{\lambda}{\mu} = 1$ and with PUs traffic load of $\eta = 10\%$
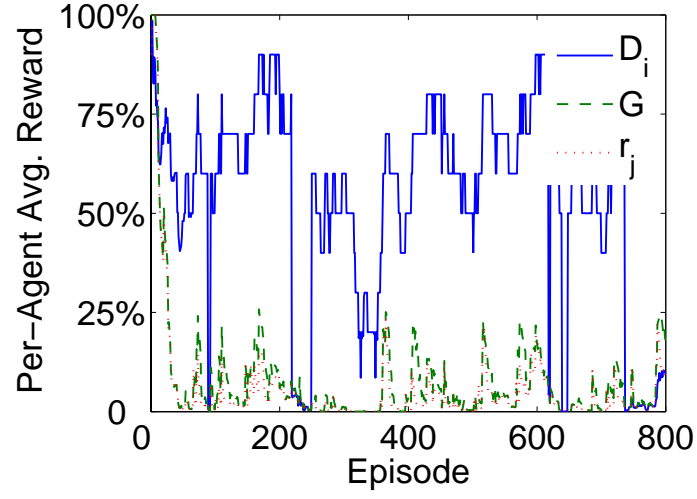
(a) $\kappa = 500$
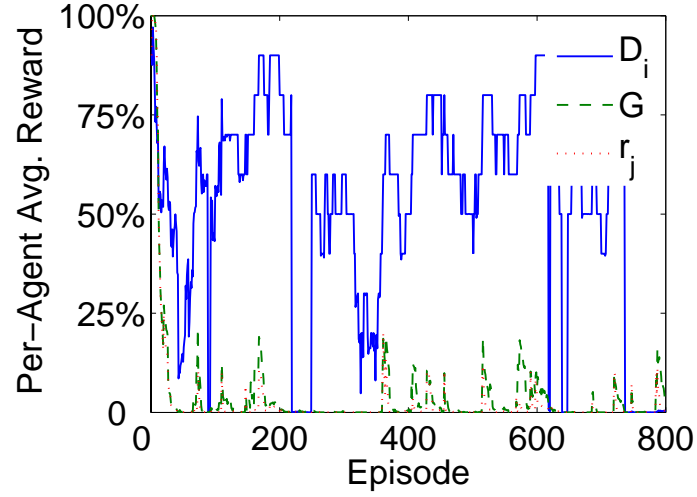
(b) $\kappa = 750$
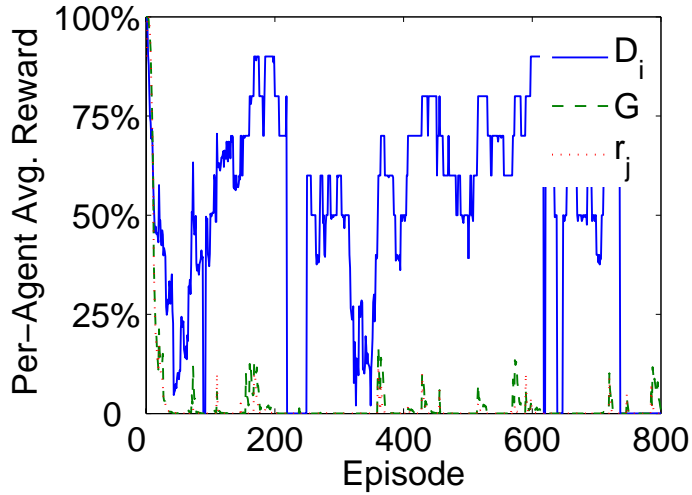
(c) $\kappa = 1000$

Figure 2.10: Normalized per-agent average reward under OSA agent traffic with Poisson arrival of $\frac{\lambda}{\mu} = 1$ and with PUs traffic load of $\eta = 30\%$
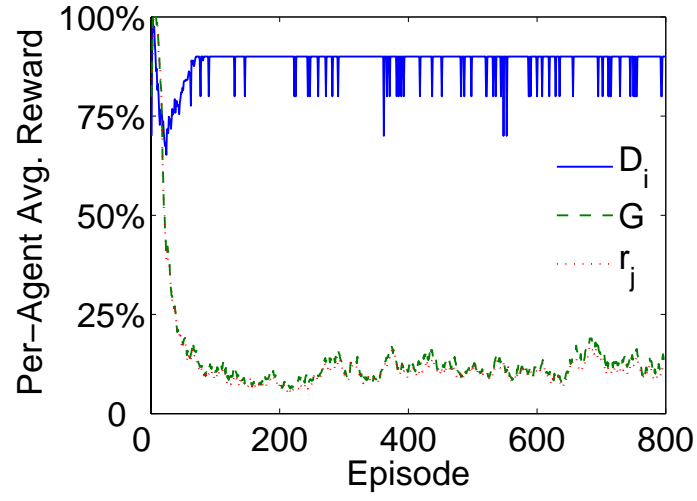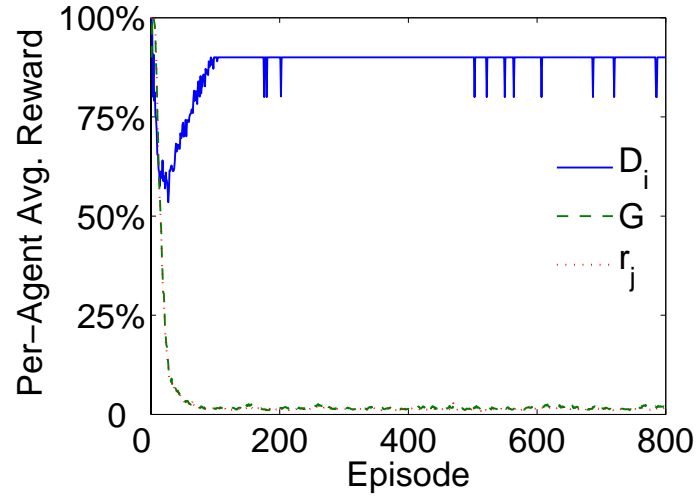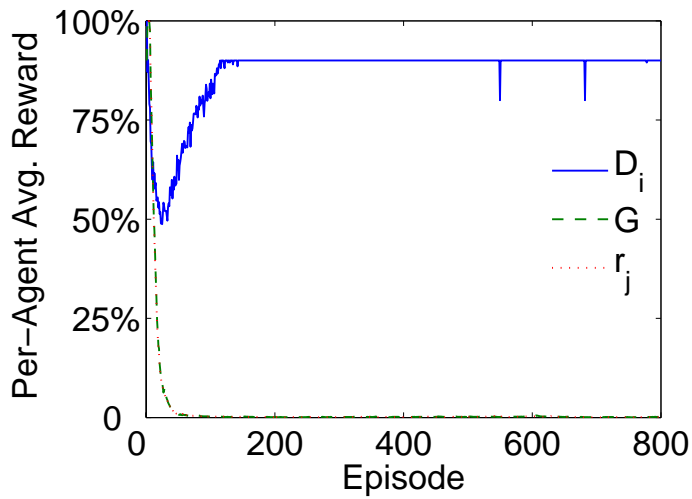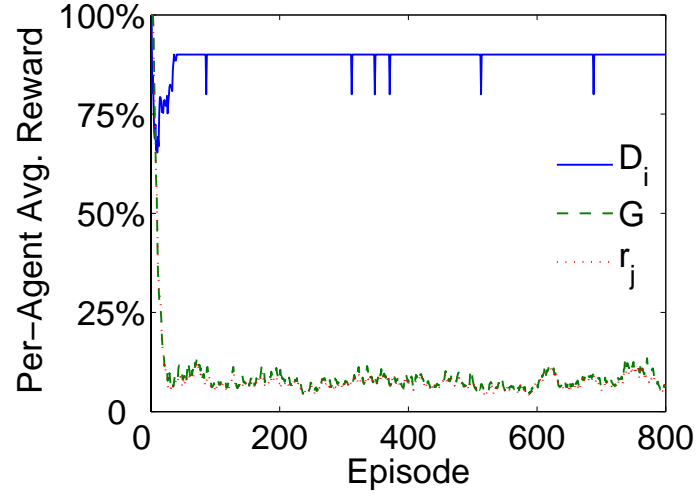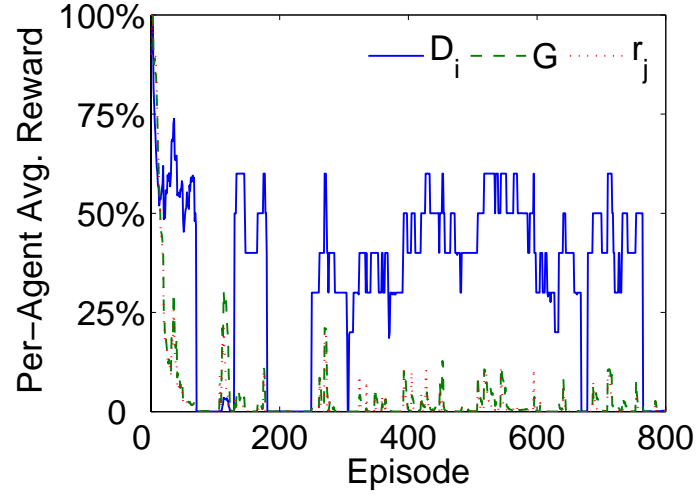
(a) $\kappa = 500$
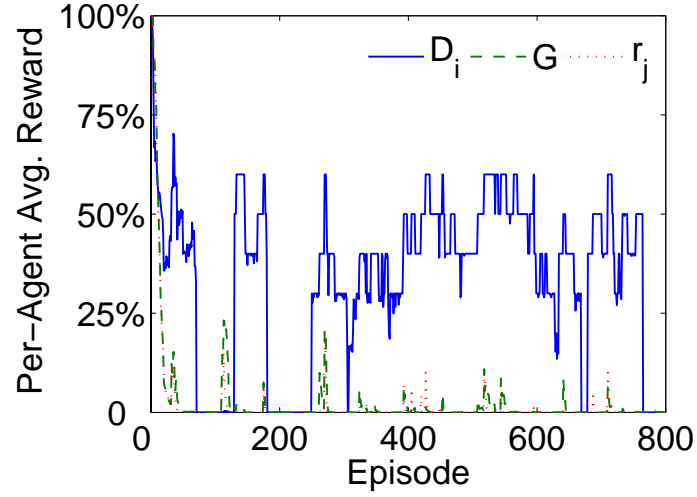
(b) $\kappa = 750$
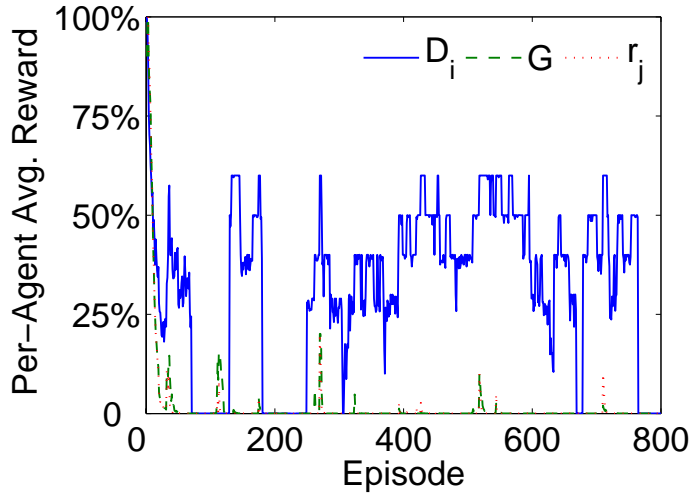
(c) $\kappa = 1000$

Figure 2.11: Normalized per-agent average reward under OSA agent traffic with Poisson arrival of $\frac{\lambda}{\mu} = 1$ and with PUs traffic load of $\eta = 50\%$

primary users (i.e., $\eta = 50\%$). However, this achieved reward under D_i is still considered as about 90% of the total available/achievable rewards. Thus, we conclude that the proposed function yields performances that are close to optimal even in the presence of PUs, and this is true regardless of the number of OSA agents and/or the PU traffic load.

We now show in Fig. 2.12 the coefficient of variations (CoV) of the per-agent average received rewards under the three studied functions for various PU traffic loads. Like when PUs are absent, when PUs are present, we also observe that the proposed objective function achieves CoVs similar to those achievable under any of the other two functions, independently of PU traffic loads. We also observe that CoV increases with the number of agents.

To summarize, the proposed objective functions, when used in practical, dynamic network settings, are shown to achieve good performances in terms of optimality, scalability, and learnability while also reaching a fairness quality as good as those reached through the two other functions.

2.7 DISCUSSION

In this work, we assumed that agents (i.e., SUs) all compete for all available spectrum bands. In other words, we implicitly assumed that all agents interfere with one another, and hence, must share the spectrum resources. However, we want to emphasize that the techniques that we derived in this Chapter apply also when considering models in which agents may not necessarily interfere with all other agents, and hence, some of them may coexist without needing to share the spectrum resources. This could be, for example, because agents are located fart apart from one another. When considering this non-interfering model, the number of agents that compete for the bands will then vary
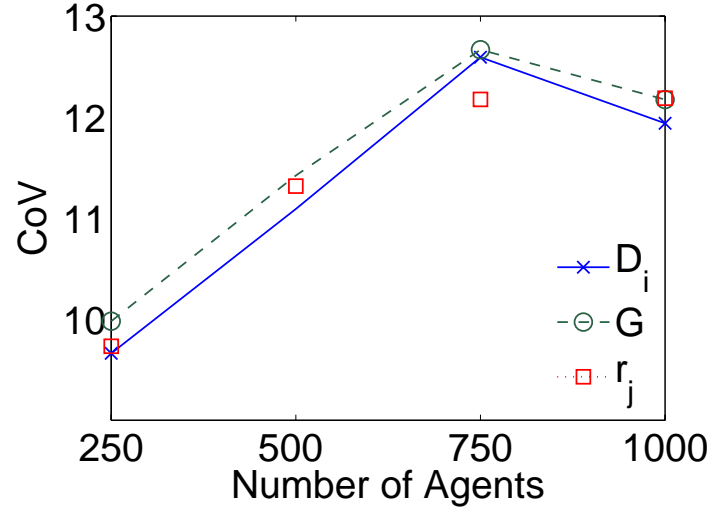
(a) $\eta = 10\%$
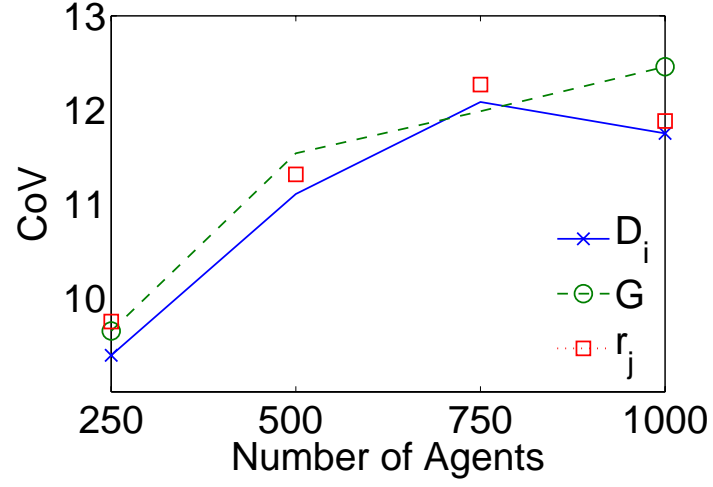
(b) $\eta = 30\%$
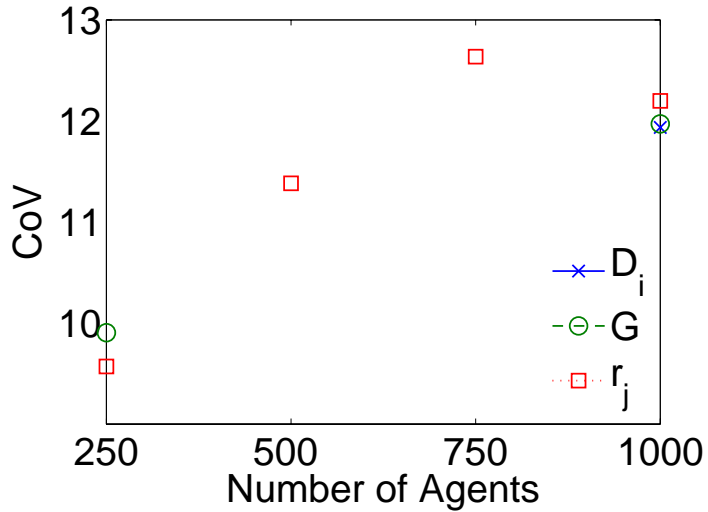
(c) $\eta = 50\%$

Figure 2.12: Coefficient of variation (CoV) of normalized per-agent average achieved reward under r_j, G, and D_i, under various PUs traffic loads

from one location to another, and hence, different locations may experience different numbers of competing agents. We argue that this does not impact the performance of our techniques, since as shown in Section 2.6.1, our techniques scalable; i.e., they perform well regardless of the number of agents in the system. In addition, the dynamic OSA systems simulated in Section 2.6.2 are composed of a number of agents that indeed changes from one time step to another, yet the techniques are still shown to perform well. Therefore, our techniques perform well in systems whose number of agents changes (whether in time as in the case of systems with dynamic behaviors, or in space as in the case of systems with non-interfering agents).

2.8 CONCLUSION

In this Chapter, we propose and evaluate efficient private objective functions that OSA users can use to locate the best spectrum opportunities. OSA users can rely on any learning algorithms to maximize these proposed objective functions, thereby ensuring near-optimal performances in terms of the long-term average received rewards. We showed that these proposed functions (i) receives near-*optimal* rewards, (ii) are highly *scalable* as they perform well for small- as well as large-scale systems, (iii) are highly *learnable* as rewards reach up near-optimal values very quickly, and (iv) are *distributive* as they require information sharing only among users belonging to the same spectrum band.

Chapter 3: Maximizing Secondary-User Satisfaction in Large-Scale

DSA Systems Through Distributed Team Cooperation

We develop resource and service management techniques to support secondary users (SUs) with QoS requirements in large-scale distributed dynamic spectrum access (OSA) systems. The proposed techniques empower SUs to seek and exploit spectrum opportunities dynamically and effectively, thereby maximizing the SUs' long-term received service satisfaction levels. Our techniques are efficient in terms of optimality, scalability, distributivity, and fairness. First, they enable SUs to achieve high service satisfaction levels by quickly locating and accessing available spectrum opportunities. Second, they are scalable by performing well in systems with small as well as large numbers of SUs. Third, they can be implemented in a decentralized manner by relying on local information only. Finally, they ensure fairness among SUs by allowing them to receive equal amounts of service.

3.1 Introduction

Dynamic spectrum access (OSA) has been recognized as a key networking solution for solving the recently observed shortage problem in spectrum supply [2, 94, 95]. In the OSA context, there are two types of users: primary users (PUs) and secondary users (SUs). PUs are the users who have exclusive access rights to use the licensed spectrum bands. SUs, on the other hand, are allowed to use spectrum bands only opportunistically.

Therefore, prior to using any licensed band, SUs must first sense the band to make sure that it is vacancy, and when a PU returns while SUs are using its band, SUs must vacate immediately. Spectrum sensing and PU detection techniques are beyond the scope of this work, and we assume that SUs use existing sensing [39, 79, 80, 112] and signal classification [109, 113] techniques for detecting and coping with PU activities.

OSA has great potentials for improving spectrum efficiency through distributed access and management of spectrum resources [18, 33, 40, 42, 71, 114, 118, 122, 140]. As a result, it has generated a lots of research interests in developing channel selection techniques that SUs can use to adapt themselves to the OSA environment [26, 38, 61, 86, 139]. Zhao et al. [139] propose a prediction model that captures the OSA environment's dynamics under periodic channel sensing. The authors use a simple two-state Markovian model to mimic PUs' activities on each channel, and use this model to derive an optimal access policy that leads to the maximization of spectrum utilization. Similarly, Liu et al. [86] model PUs' activities as a discrete-time Markov chain, which is then used to develop channel decision policies for two SUs in a two-channel OSA system. Chen et al. [26] propose OSA access methods that integrate physical-layer's with MAC-layer's sensing and access policy. They also assume that PUs' activities follow a discrete-time ON/OFF Markov process.

Most of the proposed models developed for deriving optimal spectrum selection make a Markovian process model assumption about PUs' activities, which may not be accurate. Unlike traditional communication environments, the OSA environment gives rise to some unique characteristics, making it too difficult to model its dynamics and behaviors. This fact has created research interests to develop new distributed techniques that promote effective OSA [19, 34, 64, 86, 89, 110, 128, 129]. For instance, game-theoretic approaches have been the focus of many researchers who used game-theory to develop

distributed dynamic access methods [89, 110]. The authors in [89] study a OSA system with multiple, non-cooperative SUs with restricted information exchange. In this work, ON/OFF PUs' activities are modeled as an i.i.d. Bernoulli process, and OSA is formulated as a multi-armed bandit problem with multiple, non-cooperating agents. In [110], the authors investigated distributed OSA networks with non-cooperative, selfish users by studying, through game-theoretic approaches, the impact of incomplete information on system performance. They show that the lack of information can degrade the performance substantially. Learning-based techniques are also of a particular interest to OSA because they can easily be implemented in a decentralized manner without requiring any prior knowledge of the OSA environment's dynamics. Instead, these learning algorithms allow SUs to use their knowledge acquired from past and present interactions with the environment to take the proper actions that lead to maximizing the long-term amount of service that the SUs receive from accessing the OSA system. In other words, SUs first define and choose their objectives, then rely on a learning algorithm as a means to maximize these objectives. However, when these objectives are not designed carefully, learning algorithms can lead to poor overall system performance. This is because the collective behavior of the SUs aiming to maximize poorly designed objective functions is likely to yield a low overall received system service, thereby worsening the amount of service each SU receives. It is, therefore, essential that SUs' objective functions be carefully designed so that when the SUs go after maximizing them, their behavior as a whole leads to the maximization of the amount of service that each SU receives from accessing the OSA system.

In this work, we propose efficient management techniques for OSA networks that allow SUs to maximize their received service satisfaction through efficient spectrum resource allocation [47, 48, 103, 106]. We consider a distributed OSA system with multiple,

non-overlapping spectrum bands. We also assume that each SU implements a learning algorithm (e.g., a reinforcement learner [119]) so it can use to maximize its objective function, thus enabling it to locate and select the best available spectrum opportunities. We want to emphasize that the focus of this work is not on learning algorithms, but rather on the design of efficient techniques that can be used by any learning algorithm to promote effective resource utilization.

We show that the proposed techniques enable SUs to achieve high service satisfaction levels by allowing them to quickly locate and exploit available spectrum opportunities; are very scalable by performing well in systems with a small as well as a large number of SUs; can be implemented in a decentralized manner by relying on local information sharing only; and ensure fairness among SUs by allowing them to receive approximately equal amounts of service.

The rest of the Chapter is organized as follows. In Section 3.2, we present the model and describe the motivation of this work. In Section 3.3, we present our proposed resource and service management techniques. In Section 3.4, we derive the optimal performance behaviors. We evaluate the performances of the proposed techniques, and compare them with those achievable under existing approaches in Section 3.5. Finally, we conclude the Chapter in Section 3.6.

3.2 PROBLEM STATEMENT

When a group of two or more SUs want to communicate, all members of the group must first select and switch to the same band; in the remainder of the Chapter, we will refer to these groups as *agents*. At each time step, each agent using a band receives a service that is passed to it from that band. The amount of service that the band offers an

agent can be measured in terms of, for example, amount of throughput, reliability of the communication, SNR, packet success rate, etc. We assume that once the agent switches to a particular band, it can immediately quantify and measure the amount of service that it receives from using such a band. The methods that agents use to measure the service received as a result of using a band are beyond the scope of this work. Throughout, let V_j be the total amount of service that spectrum band j offers.

Although the proposed resource and service management techniques can be used by all learning algorithms, we choose to use in this work the ϵ-greedy Q-learner [119] with a discount rate of 0 and an ϵ value of 0.05 for evaluation purposes. More details on the Q-learner can be found in [119]. We want to reiterate that this work in not on learning, but rather on the development of management techniques for OSA that can be used by any learning algorithms.

3.2.1   Traffic Model

In this Chapter, we study the inelastic traffic model, in which an agent receives a *constant* service satisfaction level when the band it uses offers an amount of service that is greater than a certain required threshold, Q, and receives an almost zero service satisfaction level when the amount of service offered by the band is below the threshold. Under this inelastic traffic model, receiving an amount of service that is less than what is required (i.e., Q) is not acceptable (which explains why the service satisfaction level drops immediately to zero), while receiving an amount that is higher than what is required is not beneficial either (which explains why the service satisfaction level remains constant). This inelastic model suits well applications with QoS requirements, such as video and audio applications, where receiving a QoS level higher than what the appli-

cation requires does not typically improve the quality, whereas if the received level is lower than the required one, the application experiences a significant degradation in its quality. Formally, the service satisfaction level, $r_j[t]$, any agent using band j receives at time step t can be written as:

$$r_j[t] = \begin{cases} 1 & \text{if} \quad n_j[t] \leq V_j/Q \\ e^{-\beta \frac{n_j[t]Q - V_j}{V_j}} & \text{otherwise} \end{cases} \tag{3.1}$$

where $n_j[t]$ is the number of agents using band j at episode t, and β is a decaying factor. Note that when $n_j[t]$ is greater than $c_j \equiv V_j/Q$, the service satisfaction level decreases exponentially. This means that none of the agents will be satisfied with the service they receive from band j if the band has more agents than c_j (c_j here represents band j's capacity; i.e., the maximum number of agents that the band can support while satisfying their required service levels). From the system's perspective, the global or system service satisfaction level can be regarded as the sum of all agents' service satisfaction levels. Formally, by letting m denote the number of available spectrum bands, the global service satisfaction level, $G[t]$, at time step t can be expressed as

$$G[t] = \sum_{j=1}^{m} n_j[t] r_j[t] \tag{3.2}$$

3.2.2   Motivation

The goal of this work is to develop efficient resource and service management techniques for large-scale, distributed OSA systems. Specifically, we aim to derive scalable and distributed objective functions for SUs that are aligned with system objective, so that when SUs (i.e., agents) aim to maximize them, they indeed lead to the maximization of

their long-term received service satisfaction levels. By means of any learning algorithm, these functions will enable SUs to efficiently find and locate spectrum opportunities, thus increasing the long-term service satisfaction level that each SU can receive from accessing the OSA system. With this in mind, the question that arises now is which objective function g_i should each agent i maximize so that its received service satisfaction level is maximized?

Intuitively, one can think of two function choices. One possible objective function choice is to have each agent i using band j maximize its inherent service satisfaction level r_j received from band j as defined in Eq. (3.1); i.e., $g_i = r_j$ for each agent i using band j. A second also intuitive choice is for each agent to maximize the global/total service satisfaction levels that all agents receive; i.e., $g_i = G$ for each agent i as defined in Eq. (3.2), hoping that maximizing the global received service satisfaction levels eventually leads to maximizing every agent's long-term average received service satisfaction level.

For illustration purposes, we measure and show in Fig. 3.1 the system/global service satisfaction levels received by all agents under each of these two private objective function choices. We consider a OSA system with $n = 1600$ agents and $m = 10$ spectrum bands.

Now we make the following two key observations. First, note that when agents aim to maximize their own inherent received service satisfaction level (i.e., $g_i = r_j$ for each agent i using band j), the global/system service satisfaction level received by all agents presents an oscillating behavior: it ramps up quickly at first but then drops down rapidly too, and then starts to ramp up quickly and drop down rapidly again, and so on, which explains as follows. With the inherent objective function, an agent's received service satisfaction level, by design, is sensitive to its own actions, which enables it to quickly determine the proper actions to select by limiting the impact of other agents' actions, thus learning about good spectrum opportunities fast enough. However, agents'
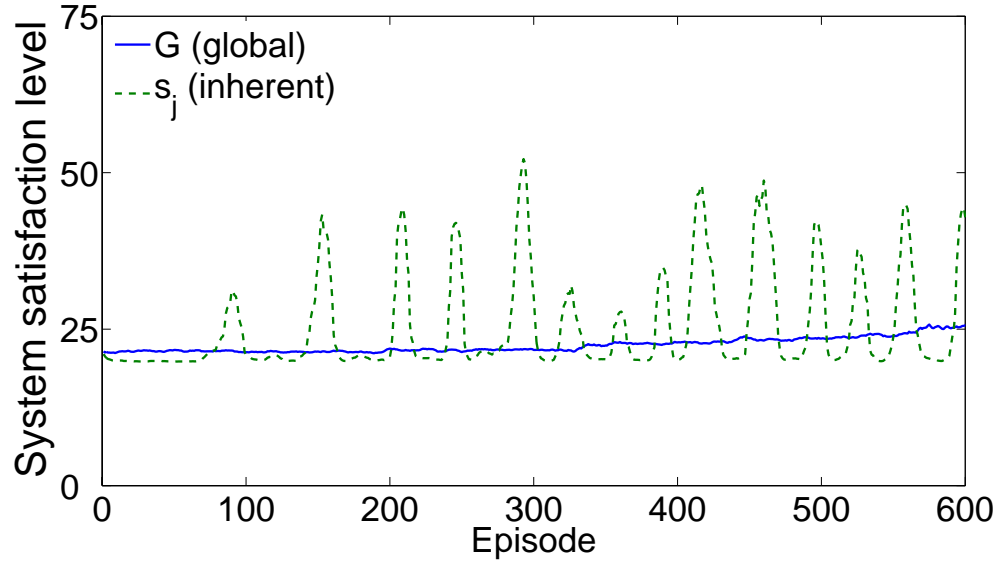
Figure 3.1: System service satisfaction level under the two private objective functions: inherent choice ($g_i = r_j$) and global choice ($g_i = G$) for $m = 10$, $\beta = 2$, and $V_j/Q = 50$ for $j = 1, 2, \ldots, 10$.

inherent objectives are not aligned with one another, which explains the sudden drop in their received service satisfaction level right after learning about good opportunities.

Second, observe that, unlike the inherent function, the global function results in a steadier performance behavior where the system received service satisfaction increases continuously, but slowly. With this function choice, agents' objectives are aligned with one another by accounting for each other's actions, and thus are less sensitive to the actions of any particular agents. The alignedness feature of this function is the reason behind the observed monotonic increase in the overall system performance. However, the increase in the performance is relatively slow due to the function's insensitivity to one's actions, leading to slow learning rates.

Therefore, objective functions must be designed with two conflicting requirements in mind: (*i*) *alignedness*; when agents maximize their own private objectives, their collective behavior should indeed result in increasing each agent's long-term received service satisfaction level, and not in worsening it, and (*ii*) *sensitivity*; objective functions should be sensitive to agents' own actions so that proper action selections allow agents to learn about good opportunities fast enough.

3.2.3   Work Objective

Our goal is to derive efficient objective functions for large-scale OSA systems that maximize SUs' total received service satisfaction levels. Specifically, we aim to derive objective functions that *i*) enable SUs to achieve high service satisfaction levels by allowing them to quickly locate and exploit available spectrum opportunities; *ii*) are scalable by performing well in systems with a small as well as a large number of SUs/agents; *iii*) are implementable in a decentralized manner by relying on local information only; and *iv*)

are fair by allowing agents to receive approximately equal amounts of service.

3.3 Resource and Service Management

The challenge in designing objective functions for OSA systems is to find the best balance between alignedness and sensitivity. Doing so will ensure that agents can learn to maximize their own objectives while also achieving good overall system performance; i.e., their collective behavior will not worsen each other's received service satisfaction level. Throughout, we use g_i to denote the objective function of agent i that we aim to derive in this work.

In this section, we will first present the difference objective function, proposed in [5] and shown to perform well in various domains, such as multi-robot coordination [7] and air traffic control [125]. This difference function will be used here as the basis for comparing the performance of our proposed function. Then, we present our proposed objective functions, whose performances, shown in Section 3.5, are compared against those achievable under the two intuitive functions (r_j and G), against those achievable under the existing difference objective function, and against a theoretical upper bound that we derive and state in Section 3.4.

3.3.1 Difference Objective Functions

Recall that, as illustrated in Section 3.2.2, when agents set the global service satisfaction level, G, as their objectives (i.e., $g_i = G$ for each agent i), their collective behaviors did indeed result in increasing the total (system) service satisfaction levels, because agents' private objectives are aligned in this case with that of the system. However, because G

depends on all agents, it is too difficult for agents (using G as their objective functions) to discern the effects of their own actions on their objectives, resulting then in low learnability rates. The authors in [5] address the above issue by proposing the *difference objective functions*, which provide a good balance between alignedness and sensitivity, leading to good system performance. The basic idea is that by removing the effects of all agents other than agent i from the function G, the resulting difference objective function will have higher learnability (or sensitivity) than G. These difference functions can formally be written as

$$D_i(t) \equiv G(t) - G_{-i}(t) \tag{3.3}$$

where $G_{-i}(t)$ is the system service satisfaction level at time step t when agent i is absent from the system. ($G(t)$ is given in Eq. 3.2.) Intuitively, since the second term evaluates the system satisfaction level without agent i, subtracting it from G provides an objective function that essentially measures agent i's contribution to the total received system service satisfaction level, making it more learnable. The difference function D_i can be thought of as the *individual or agent contribution* to the system. Now by substituting Eq. (3.2) into Eq. (3.3) and after some algebraic manipulation, D_i for agent i selecting band j at time t can then be written as:

$$D_i[t] = n_j[t]r_j[n_j[t]] - (n_j[t]-1)r_j[n_j[t]-1] \tag{3.4}$$

3.3.2 Team Contribution Objective Functions

We now present our proposed functions. Our key idea is that instead of removing the impact of all agents other than agent i from the global service satisfaction level G (which led to the difference function design), we consider removing the impact of only those agents that may not be aligned with the agent itself. That is, in terms of contribution, we propose that an agent's objective function accounts for not only its contribution, but also for the contributions of all the agents that are aligned with it; i.e., those agents that share the resource with. More specifically, we propose that when the agents sharing a particular band/resource make, as a team, a positive contribution to the overall system performance, each agent in the team gets rewarded the team contribution; i.e., the sum of all agents' contributions. But when the team contribution is negative (i.e., the resource is overcrowded, and hence none of the agents sharing it meet their required service), each agent in the team gets rewarded its own (negative) contribution only. The intuition is that when a group of agents (sharing a particular resource) succeed, they should celebrate their success as a team, but when they fail, each individual is only responsible for its own failure.

The proposed functions can then be thought of as the team or resource contribution to the entire system, and hence, they will be termed as *team (or resource) contribution objective functions*. Formally, when agent i chooses band j, its team contribution function can be written as

$$
T_i(t) = \begin{cases} \sum_{k=1}^{n_j[t]} D_k[t] & \text{if} \quad n_j[t] \leq V_j/Q \\ D_i[t] & \text{otherwise} \end{cases} \tag{3.5}
$$

where again $n_j[t]$ is the number of agents using band j at episode t and $D_i[t]$ is the individual contribution function of agent i using band j, given in Eq. 3.4. Note that

because D_i is the same for all agents sharing spectrum band j, Eq. 3.5 can be rewritten as

$$T_i(t) = \begin{cases} n_j[t]D_i[t] & \text{if} \quad n_j[t] \leq V_j/Q \\ D_i[t] & \text{otherwise} \end{cases} \tag{3.6}$$

With the proposed function, SUs are capable of effectively distributing themselves across the bands in a way that benefits all of them by increasing the amounts of service they receive in the long-term. Thus, the proposed technique can be thought of as a resource allocation method that enables SUs to quickly locate best spectrum opportunities, and distribute themselves among the available bands effectively without needing any cooperation from one another.

3.3.3 Distributed Computation of Team Contribution Function T_i

Before proceeding with the performance evaluation of the proposed objective functions in terms of optimality, scalability, learnability, and fairness, we want to shed some light on their implementation aspects. Specifically, we want to discuss methods that agents can use to compute them in a distributed manner in spite of the large number of interacting agents, the restricted information sharing, and the limited communication and coordination capability among agents. Note that the design of computation methods for the proposed functions is beyond the scope of this work, and is in itself a different challenging problem. But here we only want to give some insights and reiterate on the distributed feature of these proposed functions.

Note that, by taking away agent i from the second term of the function D_i (as shown in Eq. 3.4), the terms corresponding to all spectrum bands k except the band agent i is using cancel out, thus making the proposed functions implementable in a decentralized

manner; i.e., each agent can implement them by relying on local information that can be observed locally by the agent itself. Let us now elaborate further on this. From the expression of $D_i[t]$ given in Eq. 3.4, note that $D_i[t]$ depends only on $n_j[t]$, the number of agents that happen to be contending with agent i for band j. Hence, in order to compute/estimate $D_i[t]$, one needs to estimate $n_j[t]$ given the information that agent i observes locally. Now an agent i using band j can easily/locally quantify the service, $a_i(t)$, it receives once it uses the OSA system, which can for e.g. be measured in terms of the amount of throughput the agent receives. Thus, assuming that all agents sharing a band will roughly receive the same amount of throughput, and that V_j is known to all agents, the number of agents, $n_j[t]$, using band j can be estimated to $V_j/a_i(t)$, which can then be used to estimate/compute $D_i[t]$. Hence, the function $D_i[t]$ can be computed by using information that an agent can observe/measure locally, and so can the function $T_i(t)$.

3.3.4   Performance Comparison: T_i versus D_i

We now want to compare the performance of T_i with that of D_i in terms of their ability to increase the overall achieved service satisfaction level. For this, we first introduce the concept of "factoredness", which basically captures how aligned the agents' objectives are. Intuitively, the higher the degree of factoredness, the more likely it is that a change of state will have the same impact on the value of the objective function and on the achieved system satisfaction level. In other words, the more factored the objective function is, the more likely the system satisfaction level increases as agents maximize their objective functions, which eventually results in a higher long-term per-agent achieved service satisfaction level.

Let $z(t)$ characterize the joint move of all OSA agents in the system at time t. The global service satisfaction level, G, is then a function of $z(t)$, which can precisely be written as $G(z(t))$. The system state $z(t)$ basically captures the agent-channel assignment information and depends on the actions taken by the agents. For simplicity of notation, we often omit throughout the Chapter the dependency of these states on time t. With this, for systems with discrete states, the degree of *factoredness* for a given objective function g_i can formally be defined as [6]:

$$\mathcal{F}_{g_i} = \frac{\sum_z \sum_{z'} h[(g_i(z) - g_i(z'))\,(G(z) - G(z'))]}{\sum_z \sum_{z'} 1} \tag{3.7}$$

for all system states z and z' such that $z_{-i} = z'_{-i}$, where z_{-i} (or z'_{-i}) represents the system state that does not depend on the state of agent i (i.e., the parts of the system state controlled by all agents other than agent i), and $h[x]$ is the unit step function, equal to 1 if $x > 0$, and zero otherwise. A system is said to be fully factored when $\mathcal{F}_{g_i} = 1$.

Proposition 3.3.1 *The degree of factoredness of the proposed objective function T_i is higher than that of the difference objective function D_i, i.e., $\mathcal{F}_{T_i} \geq \mathcal{F}_{D_i}$.*

Proof Note that the only term in $\mathcal{F}_{D_i}$ that is different from that in $\mathcal{F}_{T_i}$ is $g_i(z) - g_i(z')$; everything else is the same. Let us then compute and compare this term for T_i and D_i. Let n_j and n'_j be the number of users in spectrum band j for system state z and z', respectively. Again, let us denote band j's capacity by c_j and define $c_j = V_j/Q$ (here, we assume $c_j = c$ for $j = 1, ..., m$). We consider the following four cases for n_j and n'_j that cover all possible cases:

- $n'_j > c > n_j$:

    In this case, from Eqs. 3.1 and 3.4, we can see that $D_i[z] - D_i[z']$ is positive.

Similarly, from Eqs. 3.4 and 3.6, it follows that $T_i(z) - T_i(z') = n_j D_i[z] - D_i[z']$ is positive. Thus, the term $g_i(z) - g_i(z')$ is positive for both objective functions, and hence, there is no difference between $\mathcal{F}_{D_i}$ and $\mathcal{F}_{T_i}$ since each depends on the sign of the term, and not on its value.

- $n_j > c > n'_j$:

  Eqs. 3.1 and 3.4 imply that $D_i[z] - D_i[z']$ is negative, and similarly, from Eqs. 3.4 and 3.6, we easily see that the value of $T_i(z) - T_i(z') = D_i[z] - n'_j D_i[z']$ is also negative. Thus, the term $g_i(z) - g_i(z')$ is negative for both objective functions. Hence, $\mathcal{F}_{D_i} = \mathcal{F}_{T_i}$.

- $n'_j, n_j > c$:

  Eq. 3.6 implies that $T_i(z) - T_i(z') = D_i[z] - D_i[z']$. Thus, the term $g_i(z) - g_i(z')$ has the same sign for both objective functions, and hence, there is no difference between $\mathcal{F}_{D_i}$ and $\mathcal{F}_{T_i}$ in this case either.

- $n'_j, n_j < c$:

  From Eqs. 3.1 and 3.4, it follows that $D_i[z] - D_i[z']$ is zero, but from Eq. 3.6, it follows that the value of $T_i(z) - T_i(z')$ depends on n_j and n'_j and is not zero unless $n_j = n'_j$. This is the only case where the term $g_i(z) - g_i(z')$ in $\mathcal{F}_{D_i}$ is different from that in $\mathcal{F}_{T_i}$. So for these terms, the numerator in Eq. 3.7 is greater when $g_i = T_i$ than when $g_i = D_i$. This is because $D_i[z] - D_i[z']$ is equal to zero, and hence so is the step function value, whereas $T_i(z) - T_i(z')$ is not always equal to zero (when $n'_j, n_j < c$) and the step function value is equal to 1 for some values of n_j and n'_j. Thus, $\mathcal{F}_{T_i} \geq \mathcal{F}_{D_i}$. This completes the proof.

3.4 Optimal Service Satisfaction

We now derive the optimal achievable service satisfaction level. This derivation will serve as a means of assessing how well the developed objective functions perform when compared not only with existing objective functions, but also against the optimal achievable performances.

Without loss of generality and for simplicity, let us assume that $V_j = V$ for $j = 1, 2, \cdots, m$. Let n denote the total number of agents in the system at any time. In what follows, we assume that $n > m\frac{V}{Q}$ (when $n \leq m\frac{V}{Q}$, the problem is trivial), and let $c = \frac{V}{Q}$, which denotes the capacity (in terms of the number of supported agents) of each spectrum band. Now, we start by proving the following lemma, which will later be used for proving our main result.

Lemma 3.4.1 *The system/global service satisfaction level reduces less when a new agent joins a more crowded spectrum band than when it joins a less crowded band.*

Proof Recall that when a band j has $n' > c$ agents, the total service satisfaction level offered by the band is $G_j(n') = n'e^{-\beta(\frac{n'}{c}-1)}$. If a new agent joins this band, the new total service satisfaction level offered by the band becomes $G_j(n'+1) = (n'+1)e^{-\beta(\frac{n'+1}{c}-1)}$. First, it can easily be shown that when $n' > c \geq 1$, $G_j(n') > G_j(n'+1)$. Hence, the total service satisfaction level offered by a band j decreases by $\Delta_j(n') \equiv G_j(n') - G_j(n'+1)$ when a new agent joins the band. Now we can easily see that $\Delta_j(n')$ decreases when n' increases. Hence, the greater the number n' (i.e., the more crowded the band), the smaller the decrease in the total service satisfaction level when a new agent joins the band.

Theorem 3.4.2 *When there are n agents in the system, the global service satisfaction*

level reaches its maximal only when $m - 1$ bands (out of the total m bands) each has exactly c agents, and the m-th band has the remaining $n - c(m - 1)$ agents.

Proof Let $k = n - mc$, and let us refer to the agent distribution stated in the theorem as C. Note that C corresponds to when $m - 1$ bands each has exactly c agents and the other m-th band has the remaining $c + k$ agents (since $n - c(m-1) = c + k$). We proceed with the proof by comparing C with any possible distribution C' among all possible distributions. Let $c + k_1$ be the number of agents in the most crowded band in C', $c + k_2$ be the number of agents in the second most crowded band in C', and so forth. We just need to deal with the bands that each contains more than c agents. If there are p bands each containing more than c agents, then we know that $\sum_{i=1}^{p} k_i \geq k$.

For each band having $c + k'$ agents, let ϵ_i be the amount by which the global service satisfaction level is reduced when agent i joins the band for $i = 1, 2, \cdots, k'$. From Lemma 3.4.1, it follows that $\epsilon_i > \epsilon_{i+1} > 0$, for all $i = 1, 2, \cdots, k' - 1$.

Note that for the distribution C, the global service satisfaction level is reduced by $t = \sum_{i=1}^{k} \epsilon_i$, and for C', it is reduced by $t' = \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i$. It remains to show that $t' - t > 0$ for any $C' \neq C$. We consider three different scenarios:

- $k_1 > k$: Here, we have

$$
\begin{aligned}
t' - t &= \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \\
&= \sum_{i=k}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i
\end{aligned}
$$

  which is greater than zero.

- $k_1 = k$: In this scenario, we have

$$
\begin{aligned}
t' - t &= \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \\
&= \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i
\end{aligned}
$$

  which is also greater than zero.

- $k_1 < k$: In this scenario, we have

$$
\begin{aligned}
t' - t &= \sum_{i=1}^{k_1} \epsilon_i + \sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \\
&= \underbrace{\sum_{i=1}^{k_2} \epsilon_i + \cdots + \sum_{i=1}^{k_p} \epsilon_i}_{part\ a} - \underbrace{\sum_{i=k_1}^{k} \epsilon_i}_{part\ b}
\end{aligned}
$$

  Since $k_1 + k_2 + \cdots + k_p \geq k$, the number of ϵ_i terms in *part a* is greater than the number of terms in *part b*. From Lemma 3.4.1, we know that the largest term in *part b* is ϵ_{k_1}, which is smaller than the smallest term ϵ_{k_2} in *part a*. Hence, *part a* is greater than *part b*, and thus $t' - t$ is greater than zero.

In all scenarios, we showed that $t' - t > 0$. Therefore, the global service satisfaction level for any distribution C' is smaller than that for the distribution C; i.e., C is the distribution that corresponds to the maximal achievable global service satisfaction level.

Corollary 3.4.3 *The system service satisfaction level that a OSA system can achieve is at most $(m - 1)V/Q + (n - (m - 1)V/Q)e^{-\beta(\frac{nQ}{V} - m)}$.*

Proof The proof follows from Theorem 3.4.2 by calculating the achievable global service satisfaction level for the derived optimal agent distribution.

Note that the optimal achievable system satisfaction level (derived and stated in Corollary 3.4.3) is a theoretical upper bound on the sum of all agents' possible achievable service satisfaction levels. In the next section, we will evaluate the performances of the proposed objective functions as well as those of the existing difference function, and compare them against this upper bound.

3.5   Performance Evaluation and Analysis

In this section, we evaluate the effectiveness of the proposed resource management techniques in terms of their achievable system service satisfaction levels. To do so, we measure and compare the achievable satisfaction levels under each of the four functions: inherent ($g_i = r_j$), global ($g_i = G$), difference ($g_i = D_i$), and proposed ($g_i = T_i$). In addition to comparing the performance of the proposed function with those achievable under existing ones, we compare it against the optimal achievable performance, derived in Corollary 3.4.3.

3.5.1   Simulation Method and Setting

We consider a OSA system consisting of m non-overlapping spectrum bands and a large number of SUs agents all using the system opportunistically. We assume that each agent uses the Q-learning algorithm to implement the proposed objective function. Each agent does so independently from all other agents, and as long as it needs to access the OSA system. At each episode, each agent receives an amount of service (i.e., throughput) that is passed to it from the system. The learning algorithm utilizes this amount of service to compute and maximize its objective function so as to help the agent make the

best spectrum decision/choice. All simulation scenarios are run (using MATLAB) until the measured achievable satisfaction level reaches its maximum peak. Each simulation point plotted in each figure is averaged over all runs.

Unless stated otherwise, throughout this performance evaluation section, the decaying factor β is set to 2, the number of agents is set to 1600, the number of bands is set to 10, and the capacity V_j/Q is set to $c = 50$ for all j.

3.5.2   Static OSA without Primary Users

We begin by considering a static OSA system, in which all SUs enter and leave the system at the same time. We also ignore PU activities for the moment. Fig. 3.2 shows the system service satisfaction level normalized w.r.t. the optimal service satisfaction level (derived and stated in Corollary 3.4.3) achieved under each of the four functions: inherent, global, difference, and proposed.

The figure shows that the proposed function, T_i, outperforms substantially the two intuitive functions, r_j and G, and outperforms the difference function, D_i, by about 25% in terms of the overall system service satisfaction levels. When compared to the optimal achievable performances, the proposed function T_i is shown to achieve about 85 to 90% of the maximal achievable service satisfaction levels. Also, observe that our proposed function is very learnable as it enables agents to reach up their achievable service satisfaction levels quite quickly.

Figure 3.2: Normalized system service satisfaction levels under the four studied functions: inherent ($g_i = r_j$), global ($g_i = G$), difference ($g_i = D_i$), and proposed ($g_i = T_i$) at various time steps.

3.5.3 Static OSA with Primary Users

We again consider a static system where SUs enter and leave at the same time, but with the presence of PU activities. We model PU activities on each channel as a renewal process alternating between ON and OFF periods [73, 83, 134], which represent the time during which primary users are respectively present (ON) and absent (OFF). For each channel j, we assume that ON and OFF durations are exponentially distributed with means ν_j^{ON} and ν_j^{OFF}, respectively. We use $\eta_j \equiv \nu_j^{ON}/(\nu_j^{OFF} + \nu_j^{ON})$ to denote the *PU traffic load* on channel j.

Fig. 3.3 shows the service satisfaction levels under various PU traffic loads. As it can be seen, even when considering PU activity, T_i still outperforms the other objective functions. However, the performance difference gap decreases as the PU traffic load increases. This is expected because the system satisfaction level, under any of the function, decreases as the PU load increases, since PU's presence makes the resources less available and hence, the overall system capacity decreases. Also, note that the achievable satisfaction levels, under any of the studied function, drop to zero whenever PUs come back, as it forces SUs to leave that channel, resulting then in an immediate decrease of the system service satisfaction levels.

3.5.4 Dynamic OSA without Primary Users

Now, we consider a dynamic OSA system, in which SUs (i.e., the agents) can independently enter and leave the system at various different times. To model the dynamic behaviors of SUs, we assume that agents arrive according to a Poisson process with arrival rate λ. Each agent is characterized with an exponentially distributed duration of

(a) $\eta = 10\%$

(b) $\eta = 50\%$

Figure 3.3: Normalized system satisfaction level under the four studied functions: inherent ($g_i = r_j$), global ($g_i = G$), difference ($g_i = D_i$), and proposed ($g_i = T_i$) at various time steps with PUs traffic load of $\eta = 10\%, 50\%$

mean μ, during which the agent seeks and exploits available spectrum opportunities. We use $\kappa = \lambda\mu$ to designate the *OSA agent load*, which essentially represents the average number of agents that are using the system at any time. In this section, PU activities are ignored; they will be considered in the next section.

In Fig. 3.4, we show the achieved performances under each of the four studied functions when considering dynamic behaviors of SUs: (Fig 3.4(a) for $\frac{\lambda}{\mu} = 1$; Fig 3.4(b) for $\frac{\lambda}{\mu} = 20$). Observe that the proposed objective function T_i outperforms all the other functions even when considering dynamic behaviors. Note that as the ratio $\frac{\lambda}{\mu}$ increases, the system satisfaction levels (under any of the function) decrease. This is because the higher the ratio $\frac{\lambda}{\mu}$, the lesser time (on average) SUs spend in the system (provided that κ is kept constant), and hence, the shorter the exploration time; i.e., SUs do not have enough time to explore better spectrum opportunities. This explains why when $\frac{\lambda}{\mu}$ increases, the system satisfaction level decreases.

3.5.5   Dynamic OSA with Primary Users

We again consider a dynamic OSA system, but while also accounting for the activities of PUs. As in the previous scenario, we assume that agents arrive according to a Poisson process with arrival rate λ. Each agent is characterized with an exponentially distributed duration of mean μ. Figs. 3.5, and 3.6 show the system service satisfaction levels normalized w.r.t. the optimal service satisfaction level in a dynamic OSA system with PU activity for various values of the SU traffic ratio $\frac{\lambda}{\mu}$ and the PU traffic load η. In all cases, the proposed objective function T_i outperforms the other objective functions, but the performance gain depends on how loaded the system is. When the PU traffic load is relatively low as in Figs. 3.5(a), and 3.6(a) when $\eta = 10\%$, both the difference
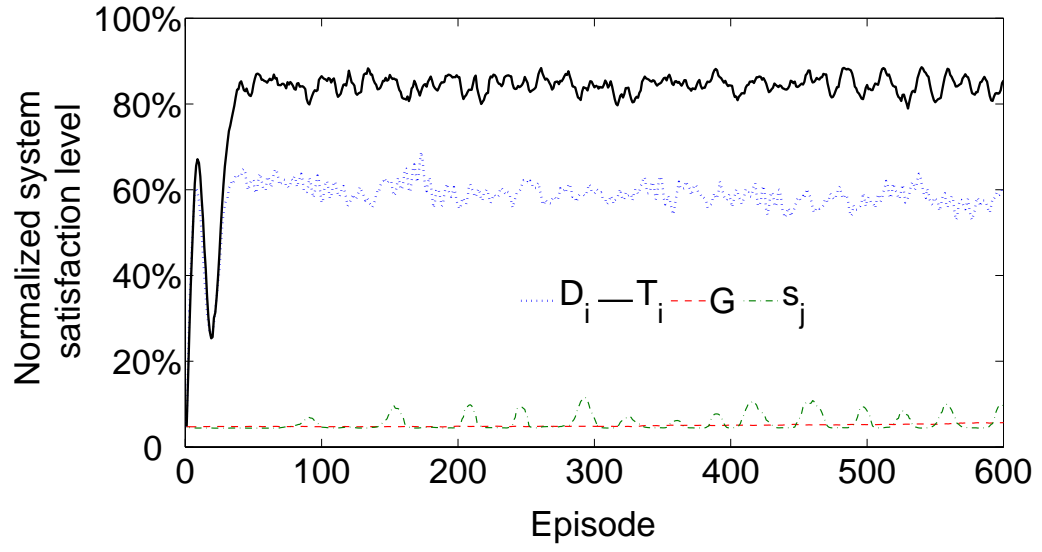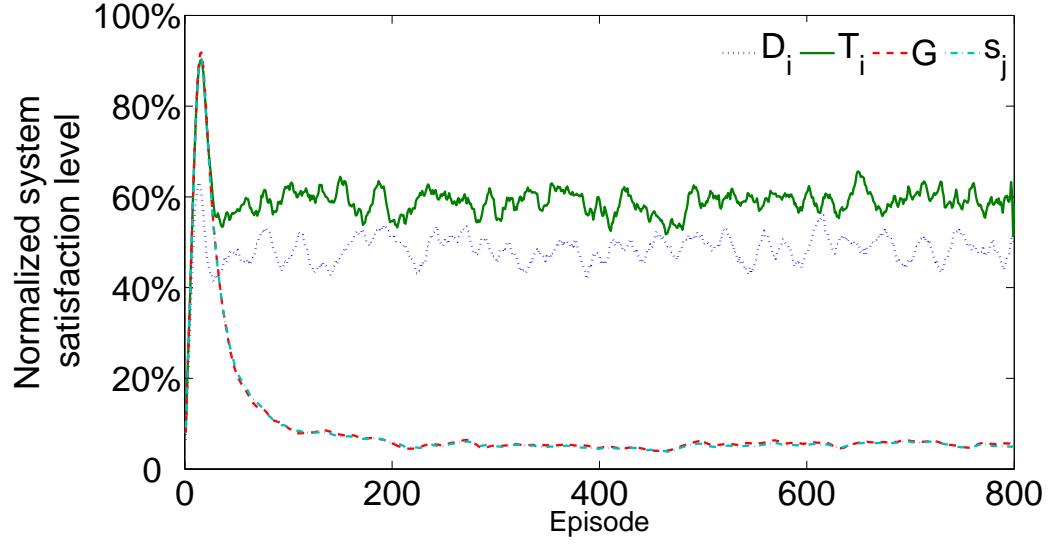
(a) $\frac{\lambda}{\mu} = 1$
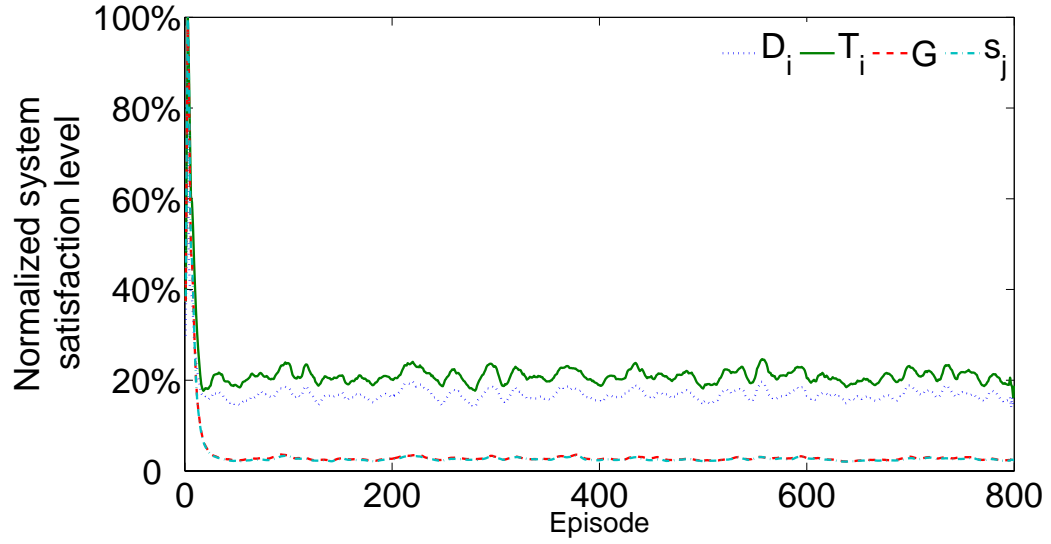
(b) $\frac{\lambda}{\mu} = 20$

Figure 3.4: Normalized system satisfaction level under the four studied functions: inherent ($g_i = r_j$), global ($g_i = G$), difference ($g_i = D_i$), and proposed ($g_i = T_i$) at various time steps under OSA agent traffic with $\frac{\lambda}{\mu} = 1, 20$ and without PU activities (total number of agents $\kappa = 1600$).

and the team contribution functions outperform the other two functions substantially. But as expected, when the PU load η increases, all functions achieve small service satisfaction levels, because the system is already loaded by primary users and hence there is no available spectrum for SUs to exploit. Likewise, as the ratio $\frac{\lambda}{\mu}$ grows (i.e., as the time each SU spends in the system decreases), the system satisfaction levels decrease, because when SUs spend lesser times in the system, it may not be enough for them to find good spectrum opportunities.

3.5.6   Scalability Performance

To also study scalability performance, we plot in Fig. 3.7 the normalized system service satisfaction level while varying number of agents, n, from 800 to 1600 while keeping m equal to 10. Since it takes some time for the technique to converge (to reach its maximum performance level), the performance values presented in this and the next subsections are measured after 600 episodes, which gave enough time for the performance to reach its best.

Observe that T_i is highly scalable. Note that as the number of agents increases, T_i maintains high achievable system service satisfaction level, whereas the satisfaction level under r_j or G drops dramatically with the number of agents. When compared with the difference function D_i, our proposed function T_i still achieves satisfaction levels that are about 30% higher than those achievable under D_i.
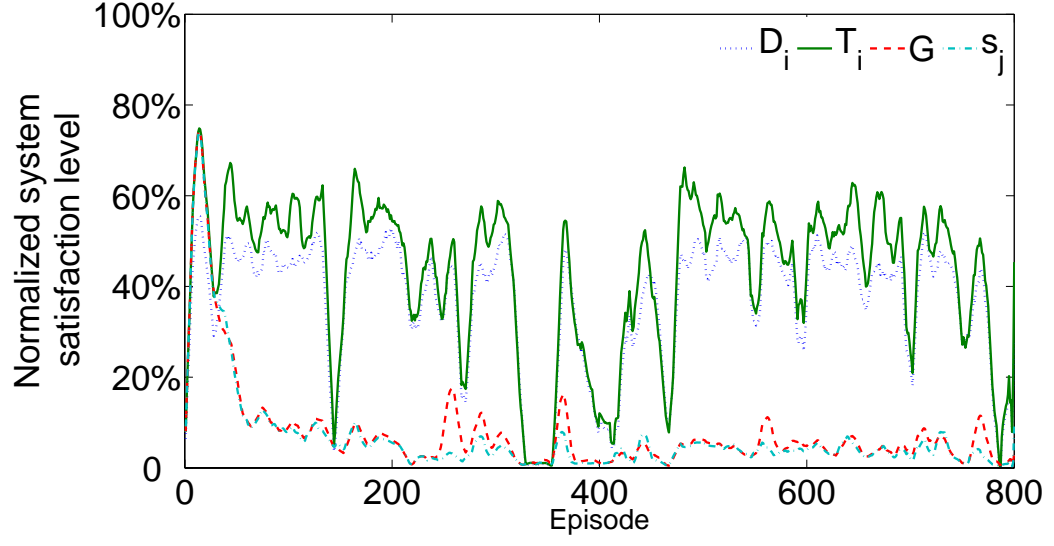
(a) $\eta = 10\%$
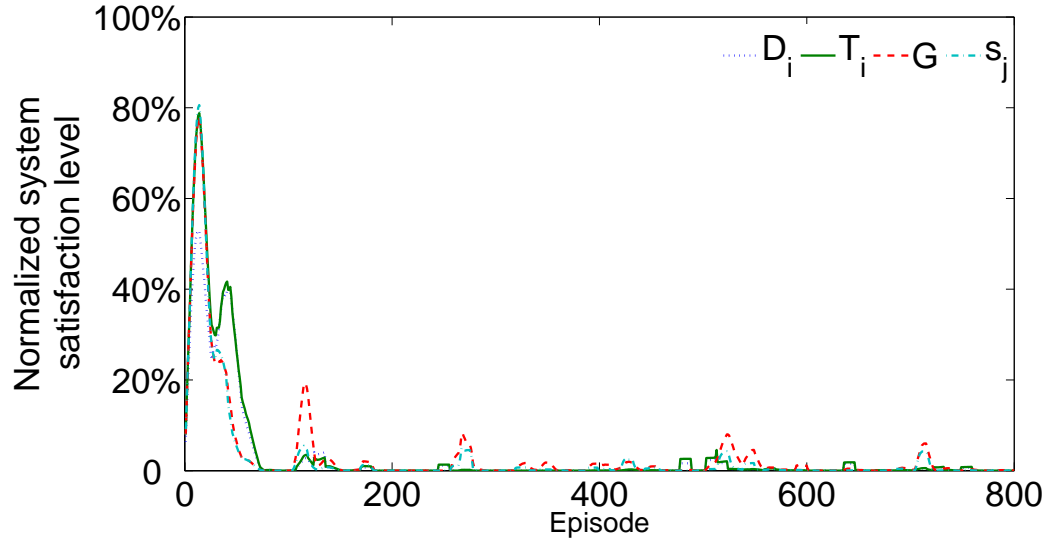
(b) $\eta = 50\%$

Figure 3.5: Normalized system satisfaction level under the four studied functions: inherent ($g_i = r_j$), global ($g_i = G$), difference ($g_i = D_i$), and proposed ($g_i = T_i$) at various time steps under OSA agent traffic of $\frac{\lambda}{\mu} = 1$ and with PU activities of $\eta = 10\%, 50\%$
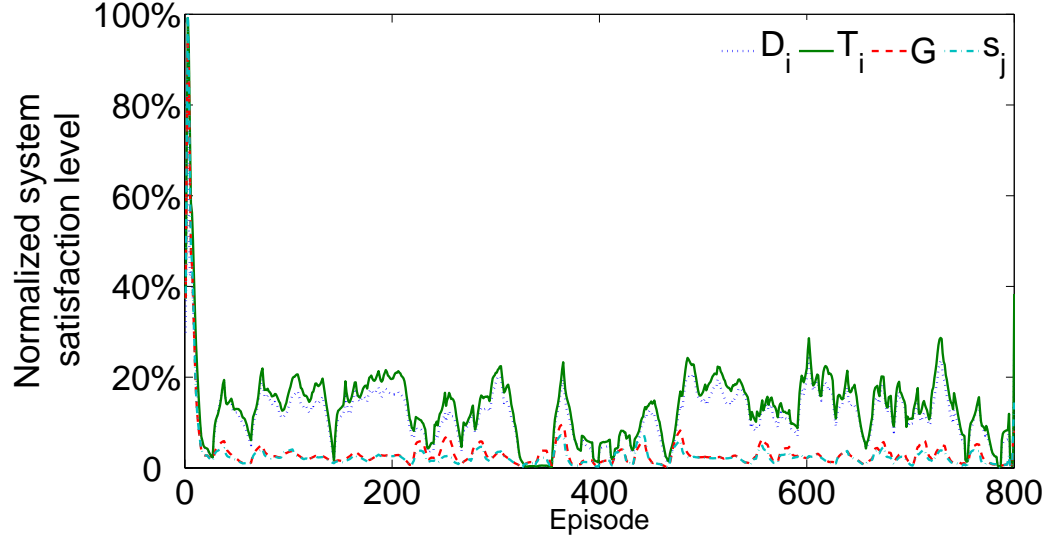
(a) $\eta = 10\%$
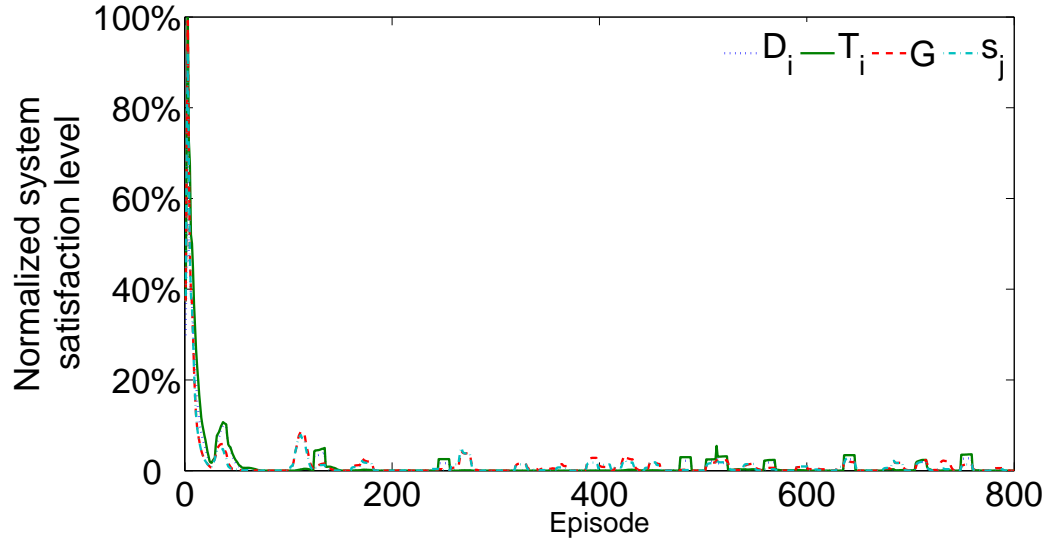
(b) $\eta = 50\%$
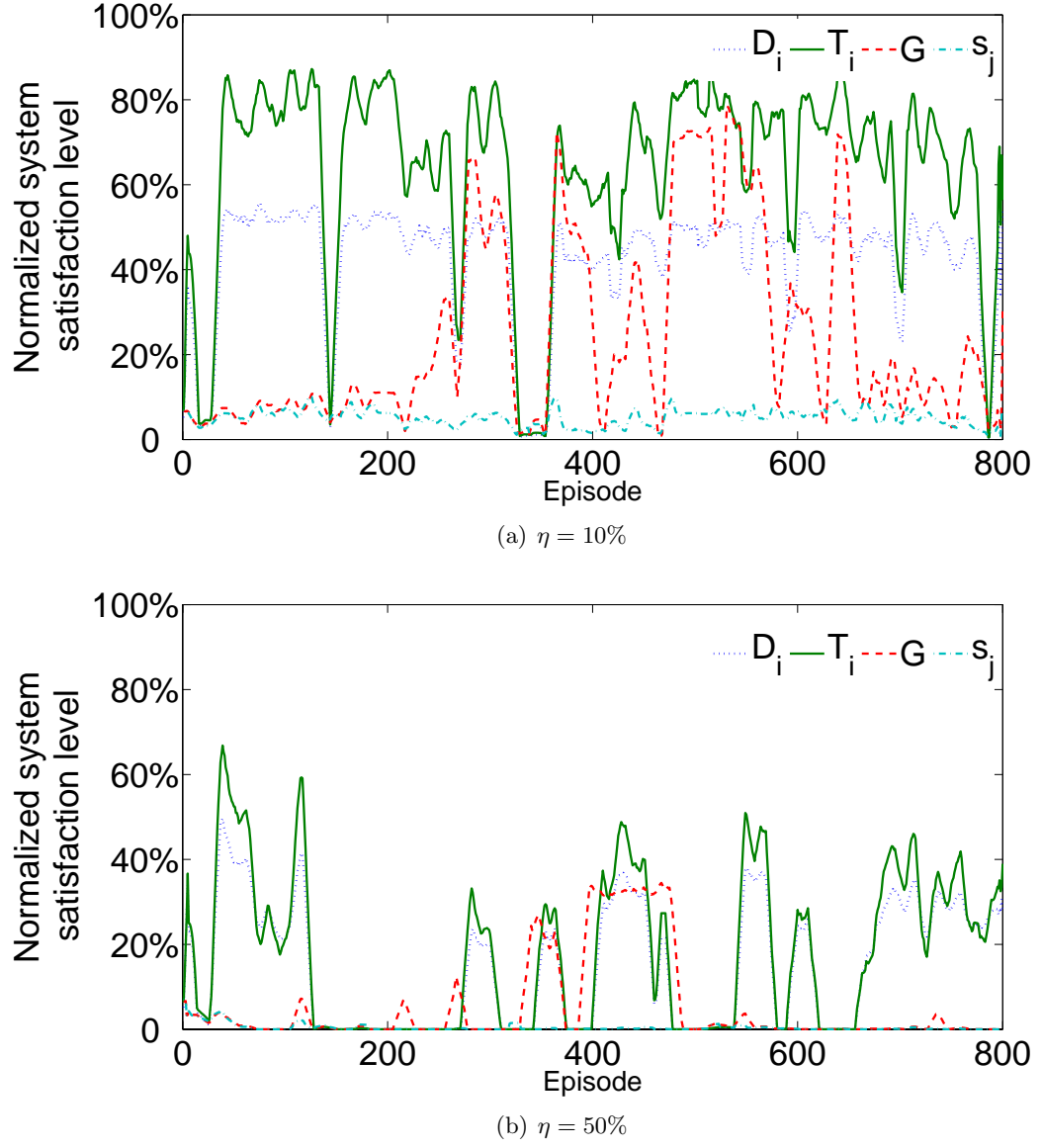
Figure 3.6: Normalized system satisfaction level under the four studied functions: inherent $(g_i = r_j)$, global $(g_i = G)$, difference $(g_i = D_i)$, and proposed $(g_i = T_i)$ at various time steps under OSA agent traffic of $\frac{\lambda}{\mu} = 20$ and with PU activities of $\eta = 10\%, 50\%$
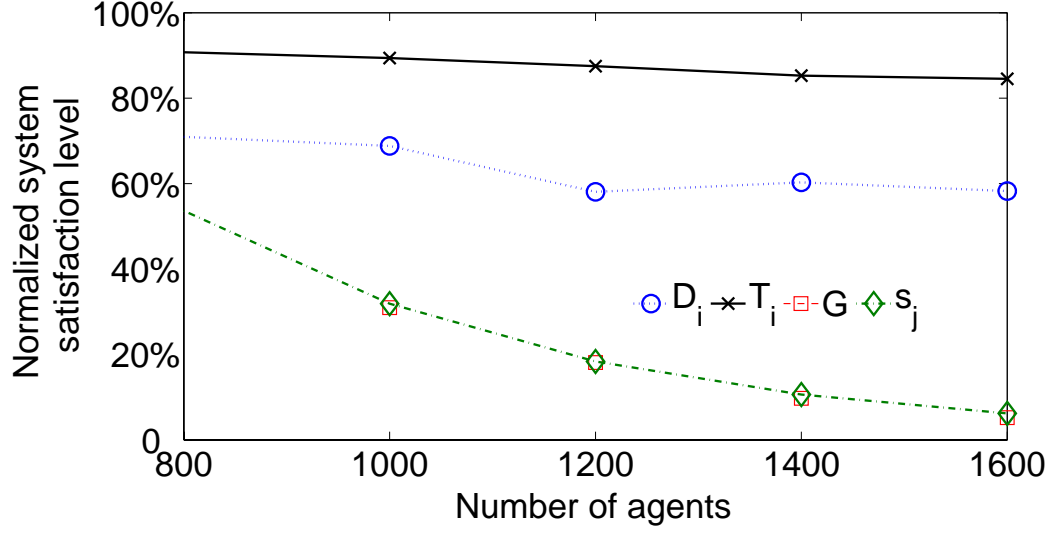
Figure 3.7: Normalized system service satisfaction levels under inherent $(g_i = r_j)$, global $(g_i = G)$, difference $(g_i = D_i)$, and proposed $(g_i = T_i)$ functions for various numbers of agents.

3.5.7   Fairness Performance

To also see how well the proposed functions do when it comes to fairness, we plot in Fig. 3.8 the coefficient of variations (CoV)[1] of the received system service satisfaction levels for various numbers of agents.

Observe that the proposed function achieves CoV values approximately similar to those achievable under any of the other three studied functions. These results show that not only the proposed function achieve good performance in terms of optimality, scalability, and learnability, but also does so while ensuring a fairness quality as good as those achieved via the other approaches.

---

[1]CoV is the ratio of the standard deviation to the mean of the agents' received service satisfaction levels; we use this metric as a means of assessing the fairness, which reflects how close agents' received satisfaction levels are to one another.
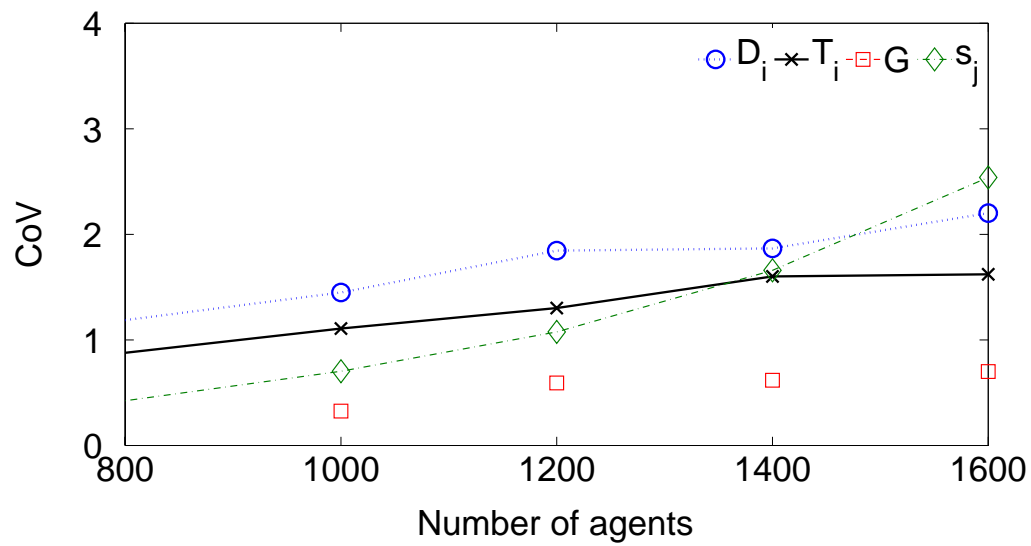
Figure 3.8: Coefficient of variation (CoV) of satisfaction levels under inherent ($g_i = r_j$), global ($g_i = G$), difference ($g_i = D_i$), and proposed ($g_i = T_i$) functions for various numbers of agents.

To summarize, we showed that the proposed objective functions achieve high satisfaction levels of agents' received service, are highly scalable as they perform well regardless of the number of agents and/or the capacity of spectrum bands, are highly learnable by enabling agents to reach up high values very quickly, are distributive as they require information sharing only among agents belonging to the same spectrum band, and are fair by allowing agents to receive similar amounts of service.

3.5.8   Discussion

There are three points that are worth mentioning and clarifying. Firstly, we want to reiterate that the reason for why our proposed objective functions are capable of achieving high service satisfaction levels is mainly because they lead to a distribution of agents across the available bands that is very close to the optimal agent distribution stated through Theorem 3.4.2, thus yielding near-optimal achievable performances and very scalable results regardless of the number of agents. That is, under the proposed techniques, $m-1$ bands will each have about c agents, whereas the rest of the agents will all go to the mth band. In other words, unlike the other two functions which tend to jam all bands by distributing the agents uniformly across all bands, the proposed functions avoid band jamming by distributing the agents across the bands in a way that benefits all agents by increasing their long-term average achievable service. Now, one may think that this may be unfair to those agents that happen to be in the most crowded band (i.e., the mth band), in that they will receive very low satisfaction levels when compared with those that happen to be in one of the other $m-1$ bands. Fortunately, this is not the case. Our experiments (not included in this Chapter due to limited number of figures) indicate that the most crowded band does not always contain the same set of agents.

That is, agents belonging to this crowded band (which offers the least per-agent service) change over time, since agents move across bands at different time steps. The fact that the same agents do not get stuck in the most crowded channel is what ensures fairness among agents by allowing different agents to receive approximately equal amounts of service. This is justified via the fairness results shown in Fig. 3.8.

Secondly, the OSA systems that we are considering in this work are very large-scale complex systems by nature, where a large number of SUs enter and leave the system independently and at various different times. Hence, it is envisioned that different components of such systems will be managed and controlled by different network owners. This nature makes it too difficult to incorporate admission control. Therefore, it is anticipated that SUs in such future systems will themselves decide whether to join the system or not based on the application they run and the service they are receiving. The techniques that we propose in this work allow SUs to achieve the best possible service satisfaction from using the OSA system through distributive and coordinated learning, and it is left up to the user to judge whether to stick around or not given the service it receives.

Thirdly, in order for a group of SUs (e.g., a SU transmitter-receiver pair or an agent) to communicate on a given data channel, the SUs must first agree on the data channel before switching to it. SUs must rely on MAC protocols to do so. Most MAC designs for cognitive networks typically designate one channel, referred to as control channel, where all control messages needed for selecting data channels take place. Numerous MAC protocols have already been proposed in the literature to enable and coordinate multiple access in cognitive radio/dynamic spectrum access networks; [29, 31] present two surveys on MAC protocols. In this work, we assume that SUs use one the existing MAC protocols to negotiate data channels.

3.6 CONCLUSION

In this Chapter, we proposed efficient resource and service management techniques to effectively support SUs in large-scale OSA systems. We showed that the proposed techniques achieve high service satisfaction levels, are very scalable by performing well in small- as well as large-scale systems, are highly learnable by reaching up high values fast, are distributive by requiring information sharing only among agents belonging to the same band, and ensure fairness among SUs by allowing them to receive equal amounts of service.

Chapter 4: Analyzing Cognitive Network Access Efficiency Under Limited Spectrum Handoff Agility

Most existing studies on cognitive radio networks assume that cognitive users can switch to any available channel, regardless of the frequency gap between the target channel and the current channel. However, due to hardware limitations, cognitive users can actually jump only so far from where the operating frequency of their current channel is. This Chapter studies the performance of cognitive radio networks while considering realistic channel handoff agility, where cognitive users can only switch to their neighboring channels. We use continuous-time Markov process to derive and analyze the forced termination and blocking probabilities of cognitive users. Using these derived probabilities, we then study and analyze the impact of limited spectrum handoff agility on cognitive spectrum access efficiency. We show that accounting for realistic spectrum handoff agility reduces performance of cognitive radio networks in terms of spectrum access capability and efficiency.

4.1 INTRODUCTION

Cognitive radio access network paradigm allows cognitive users (SUs) to exploit unused licensed spectrum on an instant-by-instant basis, so long as it causes no harmful interference to primary users (PUs). For this, SUs must ensure that licensed bands are vacant before using them, and they must vacate them immediately upon the return of any PUs

to their licensed bands.

Cognitive radio access, also referred to as dynamic or opportunistic spectrum access, has great potential for improving spectrum efficiency and increasing achievable network throughput of wireless communication systems. The research issues and topics that have been addressed in these recent years are numerous, ranging from fundamental networking issues to practical and implementation ones. A few examples of such issues and topics are spectrum access management [14, 25, 32, 41, 77, 120], adaptive and learning technique development [43, 60, 87, 129] and spectrum prediction models [8, 26, 36, 79, 84, 88]. Research efforts have also been given to deriving models and studying behaviors of the cognitive radio access performance [37, 91, 135]. Generally, most of these performance studies model cognitive radio access by means of Markov chains, and use these models to derive and analyze network performances. For example, in [9, 117, 126, 132, 137], Markov chains are used to model and study the forced termination and blocking probabilities of SUs in a cognitive multichannel access system consisting of primary and cognitive users. However, one common unrealistic assumption made in these existing works that we address in this Chapter is that SUs, when accessing the multichannel system opportunistically, are allowed to switch/jump to any available channel in the system, regardless of the frequency gap between the target and the current channels [65]. But due to hardware limitations, SUs can actually jump only so far from where the operating frequency of their current channel is, given an acceptable switching delay that users are typically constrained by [66]. Therefore, in this Chapter, we study the performance of cognitive radio networks, but while considering realistic channel switching (or handoff) agility, where SUs can only switch to channels that are immediate neighbors of their current operating channels [104].

The rest of the Chapter is organized as follows. In Section 4.2, we state the system

model. In Section 4.3, we model and derive analytically the forced termination and blocking probabilities. Section 4.4 validates the derived results, and analyzes the performance behaviors. Section 4.5 investigates the impact of spectrum handoff agility on cognitive spectrum access efficiency. Finally, in Section 4.6, we conclude our work.

4.2 Multichannel Access System Model

We consider a cognitive radio multichannel access system with m primary bands, $B_1,...,B_m$, where each band is composed of n sub-bands, giving a total of mn sub-bands, termed $A_1,...,A_{mn}$. Two types of users are present in the system. Primary users (PUs) who have exclusive access rights to B_1 to B_m, and cognitive users (SUs) who are allowed to use the A_1 to A_{mn} sub-bands, but in an opportunistic manner; i.e., so long as they do not cause any harmful interference to PUs. Throughout this work, we assume that SUs are equipped with single-radio devices.

While PUs have strict priority to use the spectrum bands, SUs are allowed to use a sub-band only when the sub-band's associated primary band is vacant; i.e., not being used by any PUs. Here, we ignore the spectrum handoff and spectrum sensing delays, simply because both of them are bounded [66, 97] and hence do not impact the blocking and the forced termination probabilities; i.e., the performances that we study in this work. These delays, however, need to be accounted for when analyzing system throughput and spectrum utilization performances. Therefore, we assume throughout that SUs are always aware (with some bounded delay) of the presence of PUs, and that as soon as any PUs reclaim their band, SUs are capable of immediately (or with some bounded delay) vacating the band and switch to another idle sub-band, if any exists. In our model, we assume that, during spectrum handoff [132], SUs can jump to any channel/band situ-

ated at no more than k bands away from its current operating band; the set of possible channels to which a SU is able to jump to is referred to as the *target handoff channel set*. Specifically, if a SU is currently using a sub-band belonging to primary band B_i, the SU can only jump to any sub-band from B_{i-k} to B_{i+k} when handoff is initiated.

4.3   MODELLING AND CHARACTERIZATION

We model the channel selection process as a continuous-time Markov process, defined by its states and transition rates. In this section, we define the states and calculate the state transition rates. As stated previously, mn sub-bands are shared by both PUs and SUs. Thus, we define each state as an m-tuple $(i_1, ..., i_m)$ in which i_j, for $j = 1, 2, \ldots, m$, indicates the number of SUs in band j if $i_j > -1$, otherwise i_j is equal to -1, indicating that band j is occupied by a PU. Note that i_j takes on values between -1 and n (i.e., $-1 \leq i_j \leq n$). Thus, the total number of states is $(n+2)^m$ and all these states are valid. We assume that arrivals of SUs and PUs both follow Poisson processes with arrival rates λ_c and λ_p, respectively, and the service times are exponentially distributed with rates μ_c and μ_p, respectively. There are four cases/events under which a state changes, and thus we only have to consider these four cases to compute the transition rate matrix $\mathbf{Q}$. In what follows, let $(i_1, ..., i_m)$ be the current state.

Case 1: First, consider that a SU arrives to the system and selects spectrum band j. The next possible states are then $(i_1, ..., i_j + 1, ..., i_m)$ for all $-1 < i_j < n$. Assuming that the number of these states is α, the transition rate from $(i_1, ..., i_m)$ to $(i_1, ..., i_j + 1, ..., i_m)$ is then λ_c/α. The states whose i_j value is either -1 or n do not change, because the SU will be blocked and denied access to the system in this case. Note that α might be

different for different states and it can be calculated via $\alpha = \sum_{l=1,-1<i_l<n}^{m} 1$.

Case 2: Second, consider that a SU leaves spectrum band j. In this case, the next possible states are $(i_1, ..., i_j - 1, ..., i_m)$ for all $i_j > 0$. Thus, the transition rates from $(i_1, ..., i_m)$ to $(i_1, ..., i_j - 1, ..., i_m)$ is $i_j \mu_c$.

Case 3: Third, when a PU leaves band j, the next states are $(i_1, ..., i'_j, ..., i_m)$ where $i'_j = 0$ and $i_j = -1$. Assuming that the number of occupied bands by PUs is β which means that the number of next states is also β, the transition rate from $(i_1, ..., i_m)$ to $(i_1, ..., i'_j, ..., i_m)$ is then μ_p/β, where as stated earlier $i'_j = 0$ and $i_j = -1$. Note that β might be different for different states and it can be calculated via $\beta = \sum_{l=1,i_l=-1}^{m} 1$.

Case 4: Fourth, consider that a PU arrives to spectrum band j. Note that PUs do not select any band upon their arrivals, since they already have their predefined bands to operate on. In this case, the next states are $(i_1, ..., i'_{j-k}, ..., i'_j, ..., i'_{j+k}, ..., i_m)$ where $i'_j = -1$ and $\sum_{l=j-k,l\neq j}^{j+k} (i'_l - i_l) = i_j$ if user is not forced to terminate. User access termination occurs when none of the adjacent bands can accommodate the cognitive user that is required to vacate band j. Thus, the next states are $(i_1, ..., i'_{j-k}, ..., i'_j, ..., i'_{j+k}, ..., i_m)$ where $i'_j = -1$ and $(i'_l = n$ or $i'_l = -1)$ for $j - k \leq l \leq j + k$. When the user is forced to terminate, the transition rate is λ_p, and when there is no termination, the transition rate is as follows

$$\gamma_{s'}^s = \lambda_p \left(\frac{1}{2k - \sum_{l=j-k,i'_l=-1}^{j+k} 1} \right)^{i_j} \prod_{l=0,l\neq k}^{2k} \left(\frac{i_j}{i'_{j-k+l} - i_{j-k+l}} \right) \tag{4.1}$$

where $\gamma_{s'}^s$ denotes the transition rate from state s to state s', where $s = (i_1, ..., i_m)$ and $s' = (i'_1, ..., i'_m)$. Thus far, we computed the transition rates, and we were able to

determine the transition rate matrix $\mathbf{Q}$. One can solve the system of equations

$$\boldsymbol{\pi}.\mathbf{Q} = 0 \;\; \text{and} \;\; \sum_{i=1}^{(n+2)^m} \pi_i = 1 \tag{4.2}$$

where π_i is the stationary probability of state i and $\boldsymbol{\pi}$ is the stationary probability matrix.

Note that only transition rates that are calculated in **Case 4** depend on the value k since spectrum handoff is expected to occur only in this case. In the other three cases, no spectrum handoff occurs, and hence, next states do not depend on the value of k. Thus, in terms of model complexity, we can say that the construction of the transition matrix $\mathbf{Q}$ depends on the value k, because as k increases, the number of possible transitions increases and hence so is the number of non-zero entries of $\mathbf{Q}$. The size of $\mathbf{Q}$ which is $(n+2)^m \times (n+2)^m$ does not, however, depend on k. Therefore, the complexity of solving Eq. 4.2 may not depend on k if general algorithms are used. But if customized algorithms (those that take advantage of matrix sparsity) are used instead, such a complexity may be reduced depending on the value of k.

Now, the forced termination probability P_f of a cognitive user can be defined as

$$P_f = \frac{\sum\limits_{(s,s')\in T} \pi_s \gamma_{s'}^s}{(1 - P_b)\lambda_c} \tag{4.3}$$

where T is the set that contains all state pairs (s, s') in which a user is forced to terminate when transitions from s to s', and P_b is the blocking probability to be defined later. Formally, T can be defined as $T = \{(s, s') = ((i_1, ..., i_m), (i'_1, ..., i'_m))|N_c(s) > N_c(s') \;and\; N_p(s) < N_p(s')\}$ where $N_c(s)$ and $N_c(s')$ are the numbers of SUs in state s and s', respectively, and $N_p(s)$ and $N_p(s')$ are the numbers of PUs in state s and

s', respectively. The number of SUs in state $s = (i_1, ..., i_m)$, $N_c(s)$, can be written as $N_c(s) = \sum_{j=1, i_j \neq -1}^{m} i_j$. Similarly, the number of PUs in state $s = (i_1, ..., i_m)$, $N_p(s)$, can be written as $N_p(s) = \sum_{j=1, i_j = -1}^{m} 1$. When a new SU arrives to the system and cannot find any empty sub-band, because the bands are occupied by either PUs or any other SUs, the user is denied access to the system. In this case, we say that the SU is blocked, and the blocking probability P_b can then be written as

$$P_b = \sum_{s \in B} \frac{\pi_s \lambda_c}{\sum_{s \in S, s \neq s'} \gamma_{s'}^s} \tag{4.4}$$

where B is the set of all the states in which blocking occurs when a new SU arrives to the system, and is defined as $B = \{s = (i_1, ..., i_m) | \forall j \ 1 \leq j \leq m, \ -1 < i_j < n\}$.

It is worth mentioning that although the focus of this work is on analyzing how the value k affects the forced termination and blocking probabilities and not so much on how one chooses k, one can first set (decide on) acceptable/predefined blocking and forced termination probabilities, and then use Eqs. 4.3 and 4.4 to find the value of k that meets such probabilities. Basically, setting upper (lower) bounds on blocking and forced termination probabilities results in a set of values for k that satisfy these upper (lower) bounds.

4.4 ANALYTIC RESULT VALIDATION AND ANALYSIS

In this section, we validate our derived analytic results via MATLAB simulations, and analyze the performance of cognitive radio spectrum access systems with limited channel handoff by studying the impact of the target handoff channel set size on cognitive users' forced termination and blocking probabilities.
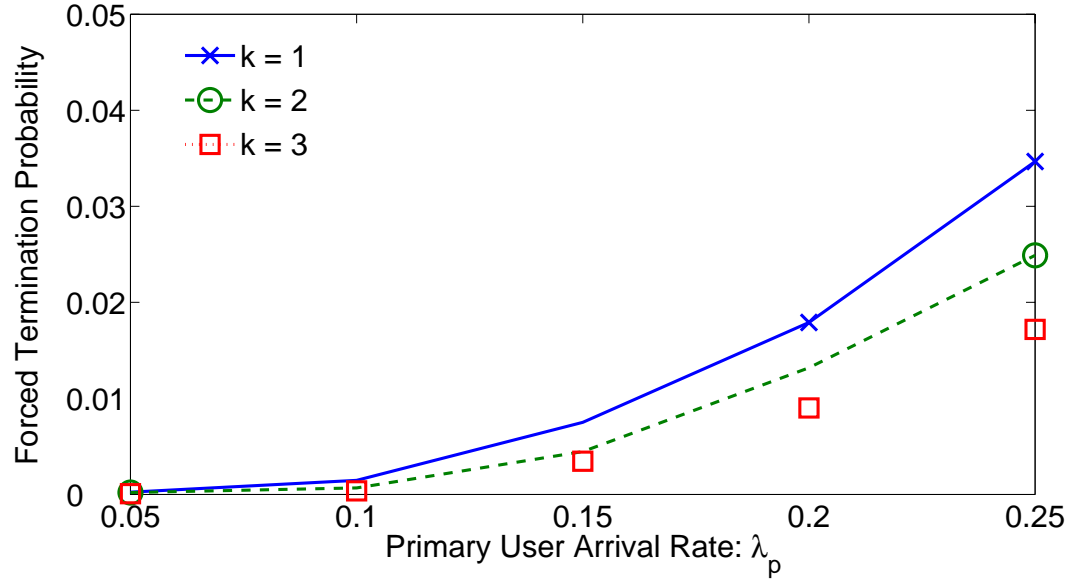
(a) Analytic results
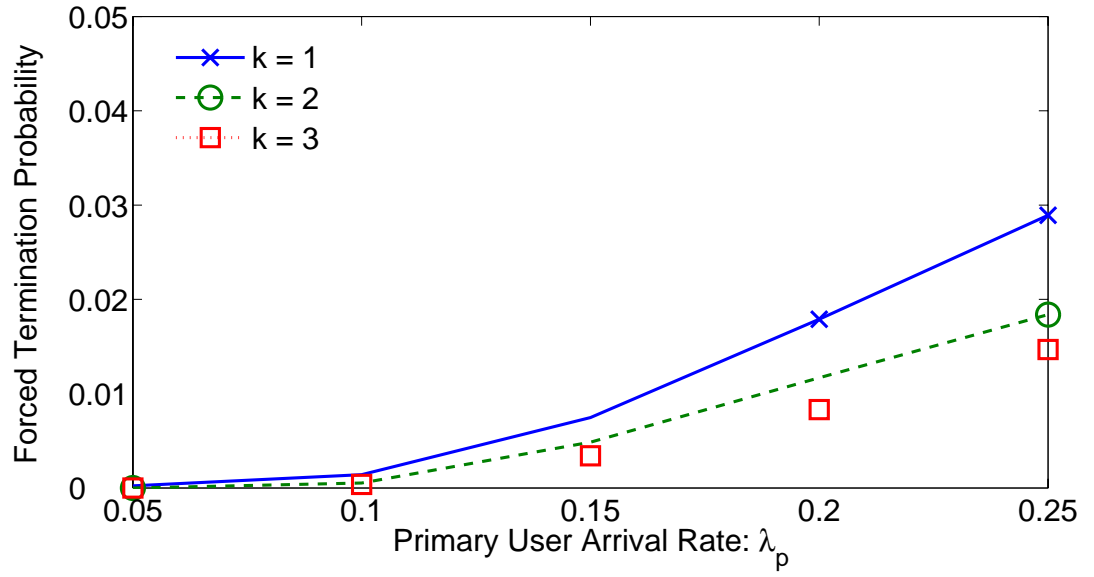
(b) Simulation results

Figure 4.1: Forced termination probability as a function of the primary arrival rate λ_p for $k = 1, 2, 3$: $m = 7$ and $n = 2$ ($\mu_p = 0.06$, $\lambda_c = 0.68$, $\mu_c = 0.82$)

4.4.1 Impact of Handoff Agility on Termination Probability

Fig. 4.1(a) plots the derived forced termination probability of cognitive users as a function of the primary user arrival rate λ_p for three different values of the number of target handoff channels, k. The termination probability is defined as the probability that a cognitive user, already accessing and using a channel whose PU has returned, is forced to cease communication as a result of none of the channels in its target handoff channel set is vacant. First and as expected, observe from the figure that as the primary user arrival rate (i.e., PU load) increases, the probability that cognitive users (already using the system) are forced to leave the system due to not finding an available band in their target handoff channel set increases. Second, for a given primary user arrival rate λ_p, the greater the number of target handoff channels, the lower the forced termination probability. Again, this trend of performance behavior is expected, as having more channels to switch to, increases the chances of cognitive users finding available bands, which explains the decrease in the forced termination probability of cognitive users. Third, the gap between the forced termination probabilities for different numbers of target handoff channel set sizes increases with the primary user arrival rate. To validate the derived analytic results, we use MATLAB to simulate a multichannel access system with primary and cognitive users arriving to the system according to Poisson process with arrival rates λ_p and λ_c, respectively. In these simulations, we compute the actual forced termination probability of cognitive users, measured as the ratio of the number of terminated users to the total number of accepted users. Fig. 4.1(b) shows the values of forced termination probabilities of the simulated cognitive radio spectrum access network again for three values of k. Observe that the simulated performance behaviors of cognitive systems in terms of the forced termination probability match well those obtained via our analytic

results. This validates our derived models.

4.4.2 Impact of Handoff Agility on Blocking Probability

We now study the impact of channel handoff agility on the blocking probability of cognitive users, defined as the probability that a cognitive user, attempting to access the multichannel system, is denied access to the system due to not finding any available channels. Fig. 4.2(a) depicts the derived blocking probability of cognitive users as a function of the primary user arrival rate λ_p for three different values of the number of target handoff channels, $k = 1, 2, 3$.

First, observe that, as expected, the blocking probability of cognitive users increases with the primary user arrival rate. That is, as the rate of primary users increases, the network becomes more and more loaded, resulting in higher blocking probability. Second, observe that unlike the case of forced termination probability, the blocking probability does not depend on the level of spectrum agility; i.e., the value of k. This is because any new cognitive user wanting to access the system does so by selecting any available channel, which leads to the same chances of being able to find an available channel, regardless of how agile spectrum handoff is for existing users (users that are already using the system). Like the case of termination probability studied in the previous section, we now validate our analytic results of blocking probability using MATLAB. We simulate a multichannel system with primary and cognitive users arriving to the system according to Poisson process with arrival rates λ_p and λ_c, respectively. In these simulations, we compute the actual blocking probability of cognitive users, measured as the ratio of the number of blocked users to the total number of arrived users. We show in Fig. 4.2(b) the blocking probabilities of the simulated cognitive network for three values of channel handoff set

size, $k = 1, 2, 3$. Observe that the simulated blocking probability performance behaviors of cognitive systems, shown in Fig. 4.2(b), match well those obtained via our analytically derived results, shown in Fig. 4.2(a). This validates the analytic blocking probability performance that we derived in this work.

4.5 SPECTRUM EFFICIENCY EVALUATION

In this section, we study the impact of spectrum handoff agility on cognitive spectrum access efficiency. We use the blocking and forced termination probabilities derived in the previous sections to evaluate the efficiency of cognitive spectrum access while considering different levels of spectrum handoff agility, k.

We first begin by defining and deriving the cognitive spectrum access efficiency while assuming no limited spectrum handoff agility. This will serve as an upper bound on the maximal achievable spectrum efficiency when considering limited handoff agility. Let us assume that there is no limited spectrum handoff agility, meaning that cognitive users are allowed to switch to any available band without any handoff restriction. Using classic Markovian analysis [96], one can write the probability p_j that j bands (out of m bands) are occupied by PUs as $p_j = \frac{\rho^j}{j! \sum_{i=0}^{m} \rho^i / i!}$ (here only PUs are considered). where $\rho = \lambda_p / \mu$ and $\mu = \mu_p = \mu_c$. It follows that the average number E_p of spectrum bands occupied by PUs can be written as $E_p = \rho - \frac{\rho^{m+1}}{m! \sum_{j=1}^{m} \rho^j / j!}$.

Similarly, the average number E_c of bands occupied by either PUs or SUs can be written as (now both PUs and SUs are considered) $E_c = \rho_c - \frac{\rho_c^{m+1}}{m! \sum_{j=1}^{m} \rho_c^j / j!}$ where $\rho_c = (\lambda_p + \lambda_c)/\mu$.

Note that here both PUs and SUs are treated the same, in that both types are allowed to use the spectrum with equal access rights and opportunities. Again, this simplified
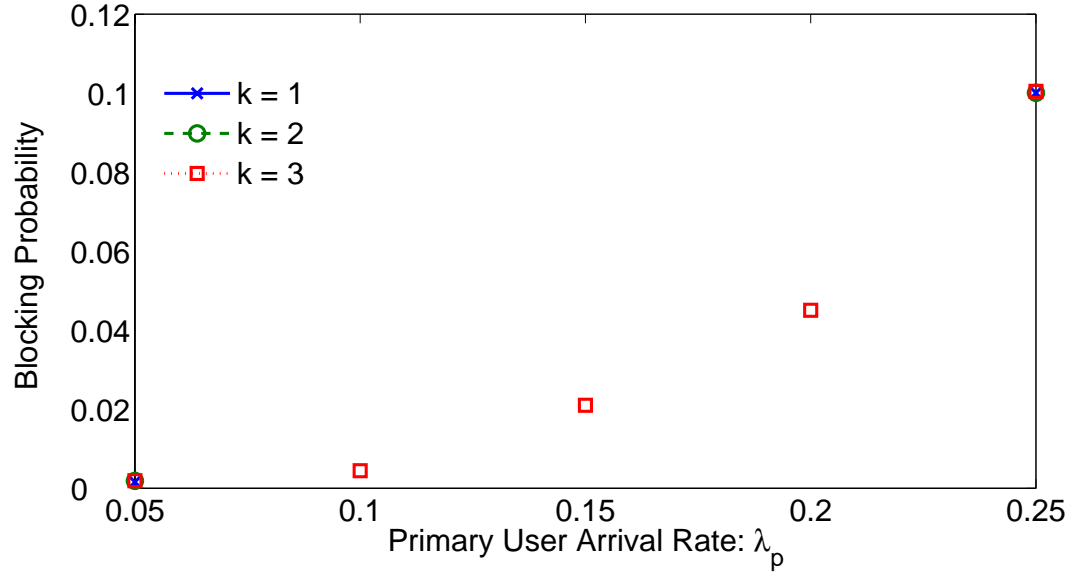
(a) Analytic results
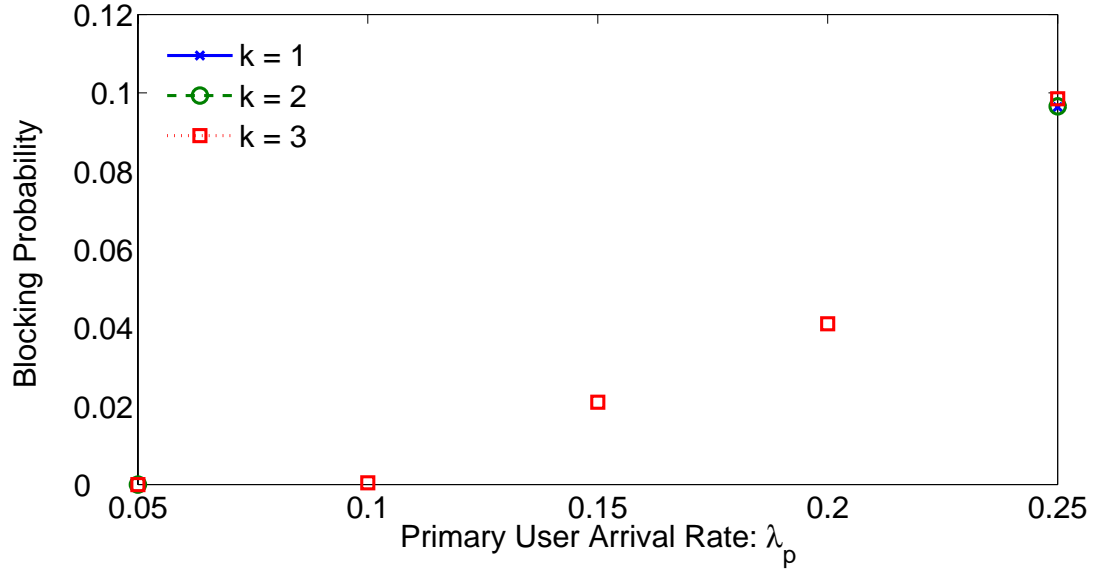
(b) Simulation results

Figure 4.2: Blocking probability as a function of the primary arrival rate λ_p for $k = 1, 2, 3$: $m = 7$ and $n = 2$ ($\mu_p = 0.06$, $\lambda_c = 0.68$, $\mu_c = 0.82$)

cognitive spectrum access is introduced so that it can be used as an upper bound on the maximum achievable spectrum efficiency when considering realistic cognitive network access, where PUs have spectrum access priority over SUs, and when SUs have limited spectrum handoff agility.

We now define the *ideal cognitive spectrum access efficiency, η,* (or cognitive spectrum access efficiency with no limited spectrum handoff agility) as

$$\eta = \frac{E_c - E_p}{m - E_p} \tag{4.5}$$

In Fig 4.3, we measure the cognitive spectrum access efficiency with limited handoff agility, normalize it with respect to the ideal spectrum access efficiency (given in Eq. 4.5), and show it for various values of primary user arrival rates when $k = 1$, $k = 2$, and $k = 3$. The values of blocking and termination probabilities used in this study are extracted from the results shown in the previous section. First, observe that as the primary user arrival rate increases, the spectrum efficiency of cognitive radio network access with limited channel handoff capability reduces, and this is regardless of the value of the parameter k of handoff agility. Second, note that the efficiency of cognitive spectrum access depends on how agile spectrum handoff is, and the higher the agility, the higher the spectrum efficiency. For example, when $\lambda_p = 0.8$, having a spectrum handoff agility of value $k = 3$ yields a cognitive spectrum efficiency of about 60% of the ideal efficiency (obtained when channel handoff is not limited to k channels), whereas, when handoff is limited to $k = 1$, the efficiency reaches about 18% of the ideal efficiency only. Third, the figure also shows that the difference between achievable spectrum efficiencies under different numbers of target handoff channels increases with the primary user arrival rate. That is, the higher the arrival rate, the higher the gap between the spectrum efficiency under $k = 3$ and that

under $k = 1$. Note that $k = 3$ is the case that there is no limited spectrum handoff agility since $m = 7$. To summarize our findings, in this work, we demonstrate the impact of the commonly made assumption of considering that cognitive users can handoff/switch to any available band, regardless of how far the target band is from the current band, on the performance behaviors of cognitive radio spectrum access systems. Our results show the importance of considering realistic spectrum handoff agility (i.e., with restricted/limited target handoff channel set) when assessing the achievable network performances and the spectrum access efficiency of cognitive radio networks. We found (both analytically and numerically) that the achievable cognitive radio performance in terms of system access capability and spectrum access efficiency can be significantly lesser than what is usually claimed in existing works, due to the limited nature of spectrum handoff agility that most works ignore and do not take into account. We therefore conclude that making unrealistic spectrum handoff assumption may lead to very inaccurate and misleading results, and it is then imperative that performance studies of cognitive radio networks do account for the restricted agility of channel switching.

4.6 SUMMARY OF FINDINGS AND CONCLUDING REMARKS

This Chapter models and analyzes the performance behaviors of cognitive radio networks enabled with dynamic multichannel access capability, but while considering realistic channel handoff agility assumptions, where cognitive users can only switch to vacant channels that are immediate neighbors of their current channels. Using Markov chain analysis, we model cognitive access networks with restricted channel handoff agility as a continuous-time Markov process, and analytically derive the forced access termination and blocking probabilities of cognitive users, and evaluate the spectrum access efficiency
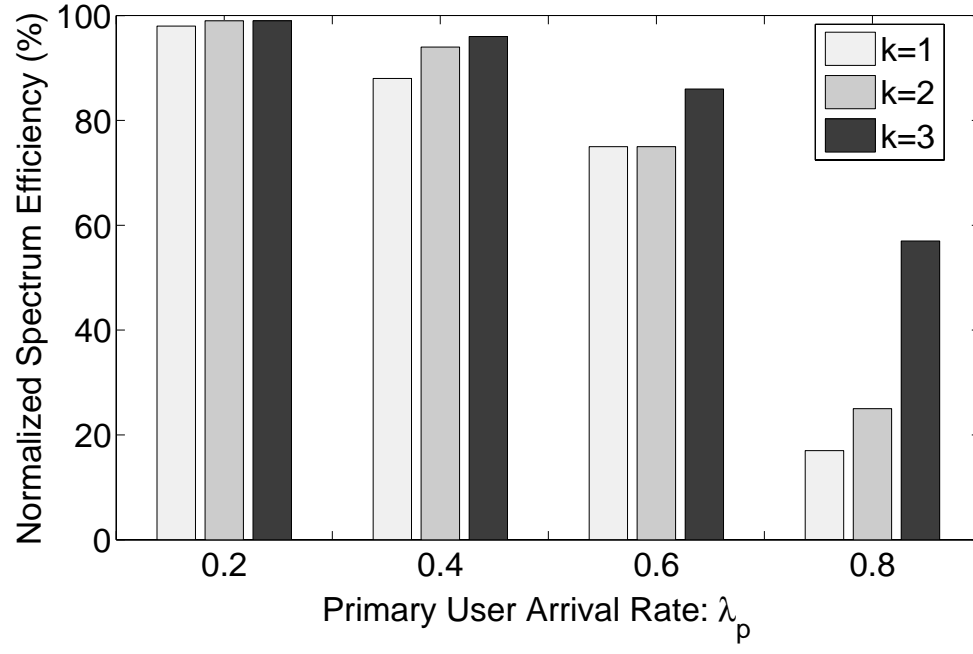
Figure 4.3: Spectrum efficiency as a function of the primary arrival rate λ_p for $k = 1, 2, 3$.

of cognitive networks. Using MATLAB simulations, we also validate our analytically derived performance results.

Our obtained results demonstrate the impact and importance of considering realistic channel handoff agility in cognitive radio access on the cognitive radio network performances in terms of users' blocking and termination probabilities, as well as cognitive spectrum access efficiency. This work demonstrates the cognitive radio performance implications of the commonly made spectrum-handoff agility assumption of allowing cognitive users to switch to any available band, regardless of how far the target band is from the current band.

Our findings in this work show that the achievable performance of cognitive radio networks in terms of spectrum access capability and efficiency can be significantly lesser than what existing works usually claim, due to the limited nature of spectrum handoff agility that most works do not account for. We conclude that making unrealistic assumption regarding the spectrum handoff agility may lead to very inaccurate and misleading results, and it is then imperative that performance studies of cognitive radio networks do account for the restricted agility of channel switching when modeling and assessing the performance of such networks.

Chapter 5: On the Impact of Guard Bands on Spectrum Bonding and Aggregation in Multi-Channel Cognitive Radio Network Access

Adjacent channel interference (ACI) is often not considered when designing spectrum sharing schemes for cognitive radio networks (CRNs). In practice, it is necessary to avoid interference by deploying guard bands between two distinct receptions. However, using guard bands typically reduces spectrum efficiency. In this work, we use continuous-time Markov analysis to study the impact of guard bands on CRN performance under different spectrum sharing schemes. We specifically derive and study several CRN performance metrics, such as blocking probability, forced termination probability, and spectrum efficiency. Our study shows the importance of taking guard bands into consideration when characterizing and analyzing CRN performances.

5.1 INTRODUCTION

FCC and other regulatory bodies conducted studies on the causes of spectrum scarcity at a given time in any location. The results of the studies indicate that less than 10 percent of the spectrum is utilized [1, 15]. Therefore, FCC was convinced that the use of cognitive radio, which provides the capability of opportunistic spectrum access, is necessary in order to improve spectrum utilization. Consequently, a pervasive research has been done on cognitive radio networks recently. Opportunistic spectrum access should guarantee that cognitive radio users do not affect primary radio users. Cognitive

radio users (CU) may coexist with primary radio users (PU) without making harmful interference with them. CUs have to vacate a channel if it is reclaimed by a PU. An important objective in this domain is how to model, characterize and analyze the system performance considering practical constraints.

5.1.1 Motivation

Many works have been done in the literature, proposing models for performance evaluation of CRNs (e.g., [16, 56, 59, 72, 108, 114, 136]). In most of these works, adjacent-channel interference (ACI) is often not considered, hence they require ideal transmission filters while, in practice, signal filtering causes spectrum spill-over due to not having ideal filters. In order to protect adjacent PU and CU transmissions, there should be a frequency separation. Such separation is referred to as a guard band. However, using guard bands restricts effective spectrum utilization. CUs should consider guard band issue when they choose channels for their transmission. However, if two contiguous channels belong to the same CU, there is no need for a guard band between them. A set of contiguous channels assigned to the same CU is called a frequency block.

We note that ACI impact in CRNs was previously studied in [136]. Specifically, a centralized solution for adaptive guard-band setting was proposed. Their solution uses a dynamic guard-band configuration to minimize ACI, requiring a central server for frequency planning. In [136], the authors did not consider channel aggregation nor did they deal with the channel assignment problem.

In this work, we consider CU transmissions to be carried over multiple contiguous (i.e., bonded) or noncontiguous (i.e., aggregated) available bands. Experimental studies were conducted on the benefits of channel bonding and aggregation in the context of

CRNs (e.g.,[24, 75, 81]). Experiments done in [24] showed the impact of using channel bonding on the performance of an IEEE 802.11 network. They experimentally proved that channel bonding can significantly improve the network performance in terms of throughput, range, and power consumption. However, they did not consider channel aggregation in their work. In their work, they simply considered single-hop scenarios. Also, authors of [75] modeled an adhoc opportunistic spectrum access network with channel bonding as a Markov process in order to study the throughput. Their results show that channel bonding generally improves network performance under certain circumstances. Network size and the number of available channels determine the level of improvement. In their model, nevertheless, the authors did not consider the guard band constraint. Another study investigated the possibility of opportunistic spectrum access under strict limit on PUs service rate [81]. The study showed that reliable communications cannot be guaranteed by traditional spectrum access policies for CUs. Therefore, their results proposed channel aggregation as a means to provide reliable communications for CUs. The papers mentioned above do not incorporate the guard band constraints in their design.

5.1.2 Contributions

In this work, we model a multichannel CRN access as a continuous-time Markov process under the realistic assumption of non-ideal filters (i.e., guard bands are needed). We then use our model to derive system performance metrics, such as blocking probability, forced termination probability, degradation probability, CU utilization, service degradation, spectrum efficiency, under five different channel assignment schemes. These channel selection schemes are designed to improve some of the system performance metrics such

as spectrum efficiency or service degradation.

Furthermore, we consider two scenarios of CRNs: FDM and D-OFDM. While in a FDM-based CRN, two neighboring CU transmissions are not allowed to share the same guard band, in the D-OFDM-based CRN, neighboring CU transmissions are allowed to do so.

5.1.3 Organization

The rest of the Chapter is organized as follows. In Section 5.2, we describe the system model considering constraints imposed by adjacent channel interference. Section 5.3 introduces five guard-band-aware channel assignment schemes which will be used for analysis. In Section 5.4, we describe our continuous time Markov process model and derive network performance metrics. Analytic results and analysis are presented in Section 5.5. Finally, Section 5.6 includes the summary and concluding remarks.

5.2 SYSTEM MODEL

We consider a cognitive radio multi-channel network which contains a set of n contiguous non-overlapping spectrum bands/channels. These bands are used by two types of geographically coexisting users; Primary Users (PUs) who have exclusive rights to access the spectrum bands, and Cognitive Users (CUs) who are allowed to use the bands as long as they do not cause interference to PUs.

CUs are only allowed to use a band only when there is no PU transmission operating over that band because PUs have a strict priority to use the spectrum bands. We assume that CUs are always aware of the presence of PUs and that, as soon as a PU reclaims

its band, CUs immediately vacate the band and switch to another idle band, if any exists. However, we know that there is a delay associated with spectrum handoff and spectrum sensing but since these delays are bounded [67, 97] and they do not impact the performance metrics that we study in this Chapter, we can ignore these delays.

The multichannel access capability can be implemented using frequency division multiplexing (FDM), or discontinuous orthogonal frequency division multiplexing (D-OFDM) [72, 101, 111]. First, we consider FDM-based and D-OFDM-based CRNs in which each CU transmission can be fulfilled over, at most, m bands simultaneously. In both models, A CU transmission may be carried over multiple contiguous (i.e., bonded) or noncontiguous (i.e., aggregated) available bands, depending on the spectrum opportunities.

5.2.1 FDM-based CRNs

Under FDM, we assume that each CU is equipped with m half-duplex radio transceivers and uses tunable raised-cosine pulse filters. The number of bands belonging to a frequency block and roll-off factor of the filter β determines the number of guard bands required for that frequency block. β is a measure for the excess bandwidth of the filter due to a spill-over. Formally, a CU transmission that uses a frequency block of k adjacent channels has an excess bandwidth on each side of the frequency block of $\Delta f = kW\frac{\beta}{2}$. In this case, ACI can be alleviated using only one guard band of bandwidth W on each side of the frequency block. That is, $\Delta f \leq W$, implying $k \leq \frac{2}{\beta}$. Considering practical values for k, β, and W, the above condition often holds. For example, with $\beta = 0.1$ and $W = 3$MHz, $k \leq 20$ channels (i.e., a data rate of up to 60 Mbps). Accordingly, it is reasonable to assume that a guard band of bandwidth W on each side of a frequency

block is sufficient to protect the reception over that block and avoid harmful interference to neighboring transmissions. Therefore, two guard bands separate two frequency blocks assigned to neighboring distinct CU transmissions. This means that if a guard band is reserved for a CU transmission, it cannot be reused (shared) by another CU transmission [115]. However, we consider only one guard band between a CU transmission and a PU transmission.

5.2.2   D-OFDM-based CRNs

In this case, a single half-duplex radio can be used for a CU transmission over multiple contiguous or non-contiguous channels. For a given CU transmission and a set of assigned channels, sub-carriers assigned to the selected channels are used for the transmission and the rest of sub-carriers are assigned zero power [72, 111]. In [123, 124], authors have shown that the main source of interference to any demodulated sub-carrier is the nearest sub-carriers of an immediate neighbor frequency block that is assigned to a different transmissions. As a result, it is sufficient to consider one guard band between two frequency blocks that are allocated to two different CU transmissions. Note that the size of the frequency blocks is disregarded [123, 124].

The difference discussed above regarding guard band reservation makes system analysis under each model different. Therefore, in this paper, we analyze the system under both reservation paradigms. In general, the difference in the transmission powers of two adjacent frequency bands determines the required amount of guard band. For conventional multi-channel networks, such difference must be taken considered during the channel assignment process. On the other hand, for opportunistic CRNs, transmission power is already limited strictly to ensure that a CU transmission does not create

harmful interference to adjacent-channel PUs. For example, the FCC demands stringent power limits on opportunistic communications over specific white spaces (the spectral mask used in an idle TV channel must be at least 55 dB less than the highest average power used by PUs [28]). Accordingly, in our work, we assume that the power difference between two distinct and adjacent CU transmissions is small as required by the FCC.

5.3 CHANNEL SELECTION SCHEMES

The required number of guard bands depends on how bands are allocated to CUs transmissions. Allocation of the bands is determined by the channel assignment scheme used by each CU. Therefore, we need to consider various channel selection algorithms in order to study and analyze the impact of guard-bands in CRNs. Below we describe five different channel selection schemes used in this work.

- **Greedy Algorithm**

  Simplicity and low overhead of this algorithm makes it attractive for use in multi-channel systems [16, 82, 99]. Upon its arrival, the CU chooses the first n available bands as it scans through the bands. Note that the CUs choose the frequency blocks considering the required guard bands. This approach is simple and quick as it does not require to sense all channels.

- **Minimizing the Number of Frequency Blocks - Min FB**

  In [115], the authors propose a Binary Linear Programming algorithm that minimizes the number of frequency blocks when assigning CU channels. Their scheme tries to reduce the number of required guard bands for a given transmission by exploiting the benefit of considering already existing guard bands with channel

bonding and aggregation. Refer to [115] for more details.

- **Maximizing the Number of Guard Bands - Max FB**

  This scheme is proposed to maximize the number of frequency blocks as opposed to minimizing the number of frequency blocks. Maximizing the number of frequency blocks can be useful since when a PU arrives in the system, it is less probable that the affected CUs are forced to terminate all of their transmissions.

- **Minimizing the Number of Guard Bands - Min GB (Local)**

  We propose an assignment scheme which aims to minimize the number of guard bands. Reducing the number of required guard bands results in higher spectrum efficiency as more bands can be used for CU transmissions. When a CU arrives, it chooses a set of channels that causes the least number of guard bands to be added to the existing ones. Min FB algorithm tries to reduce the number of guard bands indirectly by reducing the number of frequency blocks, however, Min GB (Local) tries to reduce the number of guard bands directly.

- **Minimizing the Number of Guard Bands - Min GB (Global)**

  Another approach to channel assignment is to make all the CUs minimize the total number of frequency blocks when a new CU arrives. This algorithm is different from the previous one as in this algorithm, all the CUs enter the process of choosing a set of channels and some type of coordination is needed to fulfill this. In the previous algorithm, Min GB(Local), when a CU arrives, other CUs continue their transmission on their current channels. Thus, it is possible that the algorithm does not always provide optimal solution while MinGB(Global) always chooses an optimal solution. One may argue that why we need other algorithms if MinGB(Global) is optimal. The reason is that the problem of finding an optimal

solution to minimize the number of guard bands is NP-hard, hence MinGB(Global) is not an efficient algorithm. Moreover, MinGB(Global) requires a central coordinator which makes even solving the problem more sophisticated. We want to emphasize that we only propose this algorithm for comparison purposes.

Remark: As noted earlier, the focus of this work is on analysis of guard-band aware CRNs rather than designing a channel selection scheme. Also, efficient implementation of the channel selection schemes is out of the scope of this work. It is important to note that considering fixed guard bands is not possible in this work since CUs access the frequency bands opportunistically and enforcing fixed guard bands limits the capability of CRNs, which defeats the purpose of CRN access.

5.4 Modelling and Characterization

We model the channel selection process as a continuous-time Markov process, which is defined by its states and transition rates. Here, we need to define the states and state transition rates. Bands are used by both PUs and CUs. Therefore, we define each state as an n-tuple, $(a_1, \cdots, a_n)$ in which a_i, for $i = 1, \cdots, n$, indicates that band i is assigned to CU numbered a_i, if $a_i > 0$; or, if a_j is equal to -1, it indicates that band j is occupied by a PU. If a_j is equal to 0, it means that the corresponding band is not assigned for any transmissions (i.e., the band is either idle or used as guard). Note that if $a_i = a_j > 0$, bands i and j are assigned to the same CU numbered a_i. It is important to keep track of the bands used by each CU since we need to know which bands become idle when a CU transmission is over or which CUs are affected when a PU reclaims a band.

We try to reduce the number of states in our model since the complexity of solving Markov process balance equations depends directly on the number of states. Hence, in

order to reduce the number of states in our model, we add other constraints without loss of generality. We require that, in any state $(a_1, \cdots, a_n)$, if

$$0 < a_i < a_j$$

for some $i, j = 1, \cdots, n$, then

$$\min\{k|a_k = a_i, k = 1, \cdots, n\} < \min\{k|a_k = a_j, k = 1, \cdots, n\}$$

and

$$\min\{a_k|a_k > 0, k = 1, \cdots, n\} = 1$$

We know that a_i takes only values between -1 and $n/3 + 1$ (i.e., $-1 \geq a_i \geq n/3 + 1$). Thus, the number of states is at most $(n/3 + 3)^n$. Note that, some of the states we count are invalid state as they do not satisfy the above constraints.

We model arrivals and departure of CUs and PUs both as Poisson processes with arrival rates λ_c and λ_p, respectively, and the service times are exponentially distributed with rates μ_c and μ_p, respectively. Arrival or departure of a PU or a CU create a possible state transition. In order to compute transition rates we need to look at four cases/events under which a state transition occurs; thus, we only have to consider these four cases to compute the transition rate matrix $\mathbf{Q}$. Let $s = (a_1, \cdots, a_n)$ denote the current system state in all the following cases.

- First, consider that a CU arrives to the system and selects at most n spectrum bands. The next state depends on the bands selected for CU transmission and is determined by the channel selection algorithm used by the CU. If there are at least three contiguous idle bands the transition rate from current state s to the

new state s' is λ_c. Note that the algorithms we use in this work are deterministic and result in one possible new state. Moreover, incoming CU will be blocked and denied access to the spectrum bands if the current state s does not contain at least three contiguous idle bands.

- Second, consider that a CU leaves spectrum band i. In this case, the bands used for that CU transmission become idle and the transition rate from current state s to the new state s' is μ_c.

- Third, when a PU leaves band i, the next state is $s' = (a_1, \cdots, a'_i, \cdots, a_n)$, where $a'_i = 0$ and $a_i = -1$. Assuming that the number of occupied bands by PUs α, which means that the number of succeeding states is also α, the transition rate from s to s' is then μ_p/α, where α can be different for different states, and it can be calculated via $\alpha = \sum_{l=1, a_l=-1}^{n} 1$.

- Fourth, consider that a PU arrives to spectrum band i. Note that PUs operate on a predefined band hence they do not select any band upon their arrivals. In this case, affected CU transmission has to find new idle bands to proceed. Thus, the next state is determined by the channel selection algorithm used and since, as mentioned earlier, the algorithms used in this work are deterministic, there is only one possible next state. Hence, the transition rate to the new state s' is λ_p. Note that if the all transmissions belonging to same CU are affected and that CU does not find any bands to proceed its transmission then that CU transmission is terminated.

Note, only in case 1, a CU might be totally blocked since CU arrival only takes place under that case. Also, a CU might be forced to terminate its transmission under case 4

since an ongoing CU transmission can only be affected when a PU reclaims its band.

Thus far, we computed the transition rates, and we were able to determine the transition rate matrix $\mathbf{Q}$. One can solve the following system of equations:

$$\boldsymbol{\pi}.\mathbf{Q} = 0 \;\; and \;\; \sum_{s \in S} \pi_s = 1 \tag{5.1}$$

where S is the set of possible states, $\boldsymbol{\pi}$ is the stationary probability matrix and π_s is the stationary probability of state s. Now, we can define our performance metrics based on our model and the derived stationary probabilities.

1. **Blocking Probability (P_b):**

   The blocking probability is defined as the probability that a cognitive user, attempting to access the multichannel system, is rejected accessing to the system due to not finding any available band.

   The blocking probability P_b of a CU can be formally defined as

   $$P_b = \sum_{s \in B} \frac{\lambda_c \pi_s}{\sum_{s \in S, s \neq s'} \gamma_{s'}^{s}} \tag{5.2}$$

   where $\gamma_{s}^{s'}$ is the transition rate from state s to s' and B is the set of states in which blocking occurs when a new CU arrives to the system, and is defined as

   $$B = \{s \in S | N_a(s) = 0\}$$

   where $N_a(s)$ is the number of available bands (i.e., number of available bands excluding guard bands) in state s. The number of available bands in state $s =$

$(a_1, \cdots, a_n)$ can be written as

$$N_a(s) = \sum_{i=1}^{n} \Pi_{j=i-1}^{i+1} I(a_j) \tag{5.3}$$

where $I(a)$ is an indicator function defined as

$$I(a) = \begin{cases} 1 & \text{if } a = 0 \\ 0 & \text{if } a \neq 0 \end{cases}$$

2. **Forced Termination Probability (P_f):**

   The probability that a cognitive user, already accessing and using a set of channels, is forced to stop its transmission as a result of not finding an available band upon a PU arrival is called the forced termination probability.

   The forced termination probability P_f of a CU can be defined as

   $$P_f = \frac{\sum_{(s,s') \in T} \pi_s \gamma_s^{s'}}{(1 - P_b)\lambda_c} \tag{5.4}$$

   where $\gamma_s^{s'}$ is the transition rate from state s to s', P_b is the blocking probability as defined in Eq.5.2 and T is the set of pairs of states in which a CU is forced to terminate all of its transmissions when transitioning from s to s', and is defined as

   $$T = \{(s, s') | N_a(s') = 0, N_{CU}(s) = N_{CU}(s') + 1, N_{PU}(s) = N_{PU}(s') - 1\}$$

   where $N_a(s')$, as defined in Eq.5.3, denotes the number of available bands to CUs

in state s', $N_{CU}(s)$ and $N_{PU}(s)$ denote the number of cognitive users and primary users in state s, respectively. The number of primary users in state $s = (a_1, \cdots, a_n)$ can be defined as

$$N_{PU}(s) = \sum_{i=1, a_i=-1}^{n} 1. \tag{5.5}$$

And the number of cognitive users in state $s = (a_1, \cdots, a_n)$ can be defined as

$$N_{CU}(s) = \max_i \{a_i\}, \quad for \ \ 1 \leq i \leq n. \tag{5.6}$$

3. **Service Degradation Probability (P_d):**

   The degradation probability that at least one CU transmission is forced to stop as a result of not finding an available band when a PU reclaims its channel.

   Similarly, the degradation probability P_d of a CU can be defined as

   $$P_d = \frac{\sum_{(s,s') \in D} \pi_s \gamma_s^{s'}}{(1 - P_b)\lambda_c} \tag{5.7}$$

   where again $\gamma_s^{s'}$ is the transition rate from state s to s', P_b is the blocking probability as defined above and D is the set of pairs of states in which a CU is forced to terminate some but not all of its transmissions when transitioning from s to s', and is defined as

   $$D = \{(s, s') | N_a(s') = 0, N_{CU}(s) = N_{CU}(s'), N_{PU}(s) = N_{PU}(s') - 1\}$$

   where $N_a(s')$, as defined in Eq.5.3, denotes the number of available bands to CU

in state s', $N_{CU}(s)$ and $N_{PU}(s)$ denote the number of cognitive users and primary users in state s, as defined in Eq.5.6 and Eq.5.5, respectively.

4. **Spectrum Efficiency (Ξ):**

   In addition to forced termination, blocking and degradation probabilities, we derive spectrum efficiency, average number of guard bands, service degradation and cognitive user utilization. Spectrum efficiency, Ξ, is the ratio of number of bands used for either a PU or a CU transmission to total number of bands. We formally define spectrum efficiency as

$$\Xi = \sum_{s \in S} \pi_s \xi_s \tag{5.8}$$

   where ξ_s is the number of bands used for a PU or CU transmission in state $s = (a_1, \cdots, a_m)$ and is defined as $\xi_s = \sum_{i=1, a_i \neq 0} 1$.

5. **Average Number of Guard Bands (χ):**

   We formally define the average number of guard bands, χ,as

$$\chi = \sum_{s \in S} \pi_s \kappa_s \tag{5.9}$$

   where κ_s is the number of guard bands in state $s = (a_1, \cdots, a_n)$ and is defined as

$$\kappa_s = \sum_{\substack{i=1 \\ (a_{i-1} \neq 0 \ \bigvee \ a_{i+1} \neq 0) \\ a_i = 0}}^{n} 1$$

   We also define a_0 and a_{n+1} to be equal to zero in order to avoid out of range indices.

6. **CU Utilization (H):**

   The fraction of channels used by cognitive users is defined as CU utilization.

   We define the CU Utilization, H,as

$$H = \sum_{s \in S} \pi_s \eta_s \tag{5.10}$$

   where η_s is the number of bands used for CU transmission in state $s = (a_1, \cdots, a_n)$ and is defined as

$$\eta_s = n - N_{PU}(s) - N_a(s) \tag{5.11}$$

   where n is total number of channels in the system. Also, $N_{CU}(s)$ and $N_{PU}(s)$ denote the number of cognitive users and primary users in state s, as defined in Eq.5.6 and Eq.5.5, respectively.

7. **Amount of Service Degradation (Δ):** Service degradation is defined as the number of transmission losses by CU when a PU arrival forces the CU to stop some or all of its transmissions.

   We define the amount of service degradation, χ,as

$$\Delta = \sum_{s \in S} \pi_s \delta_s \tag{5.12}$$

   where δ_s is the number of bands needed for CU transmission in addition to the allocated bands so that each CU can use all of its capacity. Formally, δ_s, for state

$s = (a_1, \cdots, a_n)$, is defined as

$$\delta_s = m \times N_{CU} - \eta_s$$

where m is the maximum number of channels a CU uses, $N_{CU}(s)$ denotes the number of cognitive users in state s, as defined in Eq.5.6 and η_s denotes the number if bands used for CU transmission in state s, as defined in Eq.5.11.

5.5 ANALYTIC RESULTS AND ANALYSIS

In this section, we analyze the performance of multichannel cognitive radio spectrum access system. We study the impact of guard bands on the performance of the system under five various channel selection schemes. Our performance metrics include (1) blocking probability (P_B), (2) forced termination probability (P_f), (3) degradation probability (P_d), (4) CU Utilization (H), (5) service degradation (Δ), (6) spectrum efficiency (Ξ) and (7) the number of guard bands (χ).

We use MATLAB to generate transition rate matrix $\mathbf{Q}$ first, as explained in Section 5.4, then solve Eq. 5.1, and finally calculate aforementioned performance metrics in a multichannel access system where primary and cognitive users arrive to the system according to Poisson process with arrival rates λ_p and λ_c and service times μ_p and μ_c, respectively.

In all of our analytical results we considered a system with $\mu_p = 10$, $\lambda_c = 1$, $\mu_c = 10$, $n = 10$. Also, $m = 3$ when the performance metric is in terms of PU arrival rate λ_p and $\lambda_p = 0.1$ where the performance metric is in terms of the number of channels accessed by each CU.

5.5.1 Analytic Results with no Guard Band Reuse

In this section, we analyze the FDM-based CRN performance where guard band reuse is not allowed.

5.5.1.1 Impact of Guard Bands on Blocking Probability

Fig.5.1(a) depicts the derived blocking probability of CUs as a function of the primary user arrival rate, λ_p, and Fig.5.1(b) depicts blocking probability as a function of the maximum number of channels used by each cognitive user, m, under five different channel assignment schemes.

First, observe that the blocking probability, defined in Eq. 5.2, of cognitive users increases with the primary user arrival rate. That is, as the rate of primary users increases, the network becomes more and more loaded, resulting in higher blocking probability. Second, as the number of channels per CU increases, the number of occupied bands by CUs increases, resulting in higher blocking probability as expected. This trend of performance behavior is expected, as having less channels to choose from, decreases the chances of cognitive users finding available bands, which explains the increase in the blocking probability of cognitive users. Third, as expected, the blocking probability is smaller when the channel assignment scheme tries to minimize the number of guard bands, resulting in leaving more available channels for the newly incoming CUs, thus having smaller blocking probabilities.
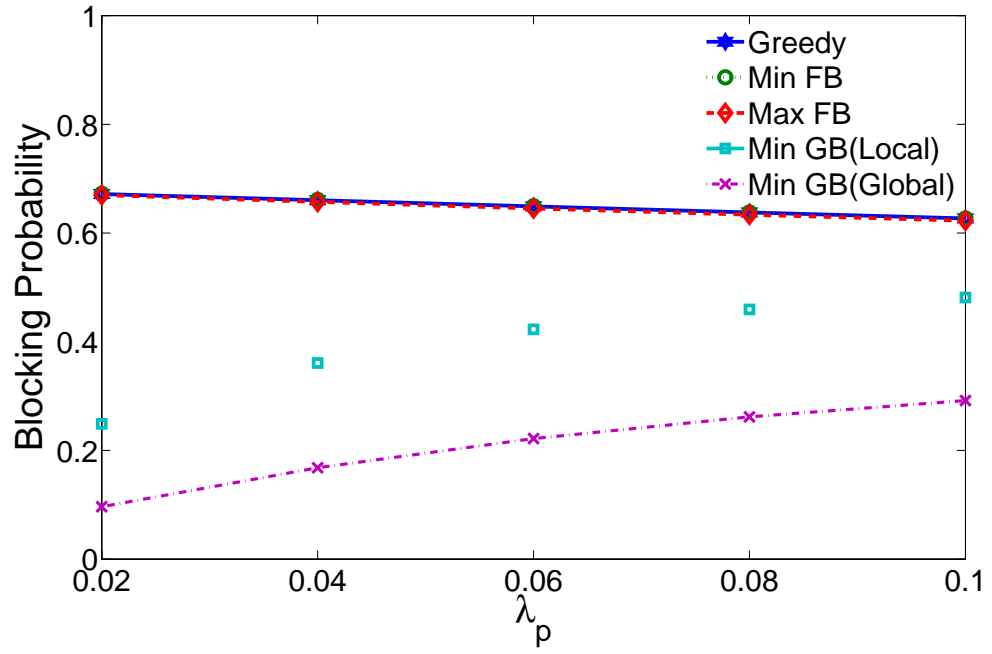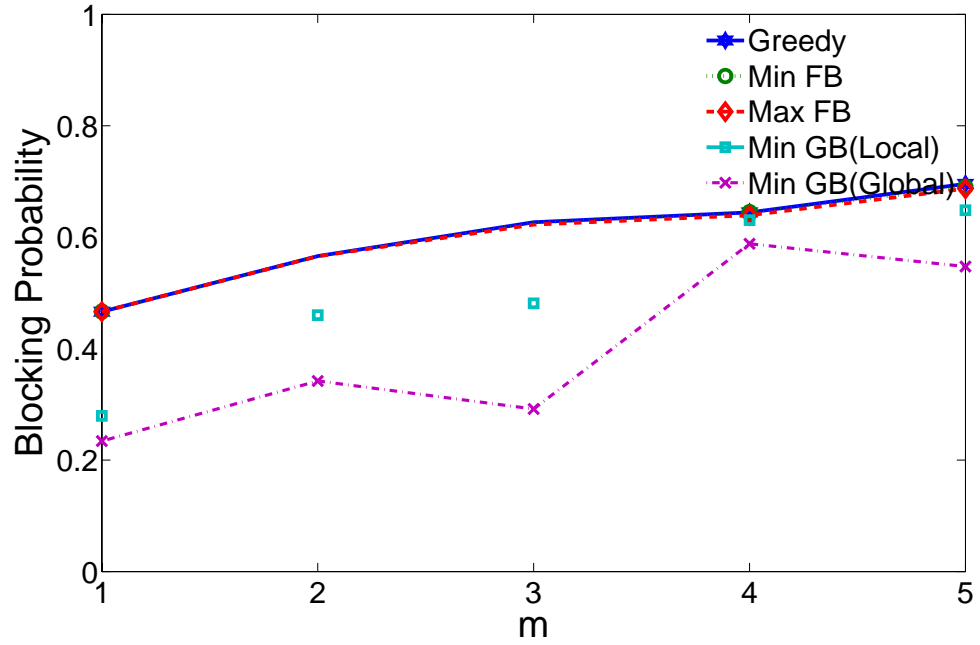
(a) P_b vs. λ_p

(b) P_b vs. m

Figure 5.1: Blocking probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band reuse, for $\lambda_p = 0.1$

5.5.1.2   Impact of Guard Bands on Termination Probability

We study the impact of guard bands on forced termination probability of cognitive users. Fig.5.2(a) plots the derived forced termination probability of cognitive users as a function of the primary user arrival rate, λ_p, and Fig.5.2(b) depicts the forced termination probability as a function of the number of radios for each cognitive user, m, under five different channel assignment schemes.

First, observe that the termination probability increases slowly as the primary user arrival rate increases. That is, as the rate of primary users increases, the chance of a CU being affected by a PU activity increases, resulting in slightly higher forced termination probability. Second, observe that when the CUs use three radios, the forced termination probability is smallest. This is because as the number of radios increases, the probability that a CU is affected by a PU activity decreases. However, at some point the CUs might not be able to use all their radios due to band unavailability, thus we observe an increase in the forced termination probability as the number of radios in use is decreased. Unlike the blocking probability, the forced termination probability does not significantly depend on the channel assignment schemes we considered in this work since they all try to minimize the number of guard bands which results in having fewer frequency blocks, hence the same level of influence by PU activity.

5.5.1.3   Impact of Guard Bands on Degradation Probability

Now, we study the degradation probability when a PU reclaims its channel and at least a CU transmission is forced to stop transmission as a result of not finding an available band. Fig.5.3 plots the derived degradation probability of cognitive users as a function
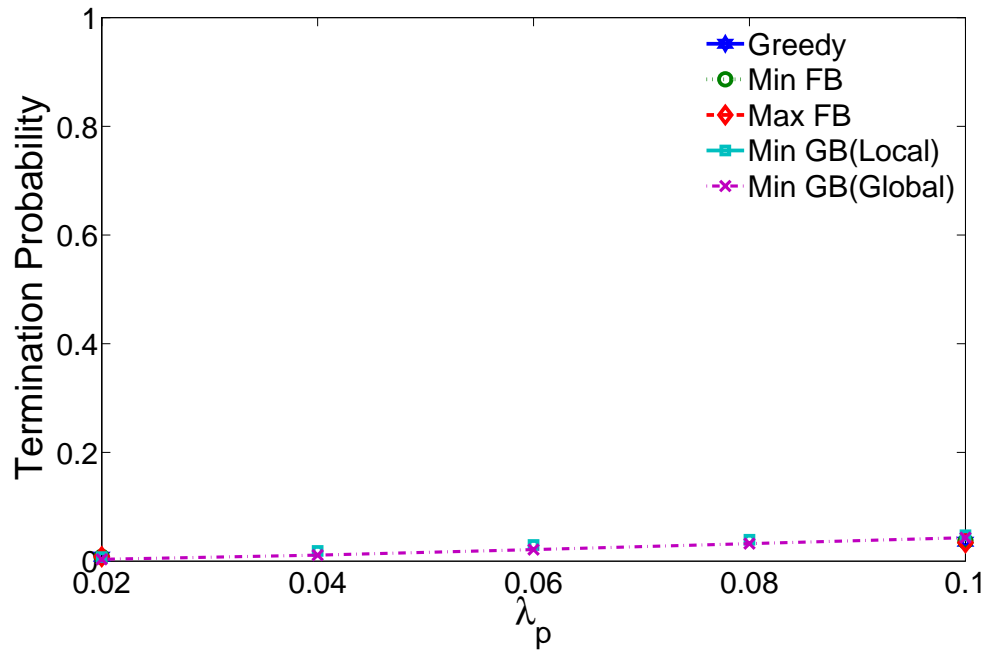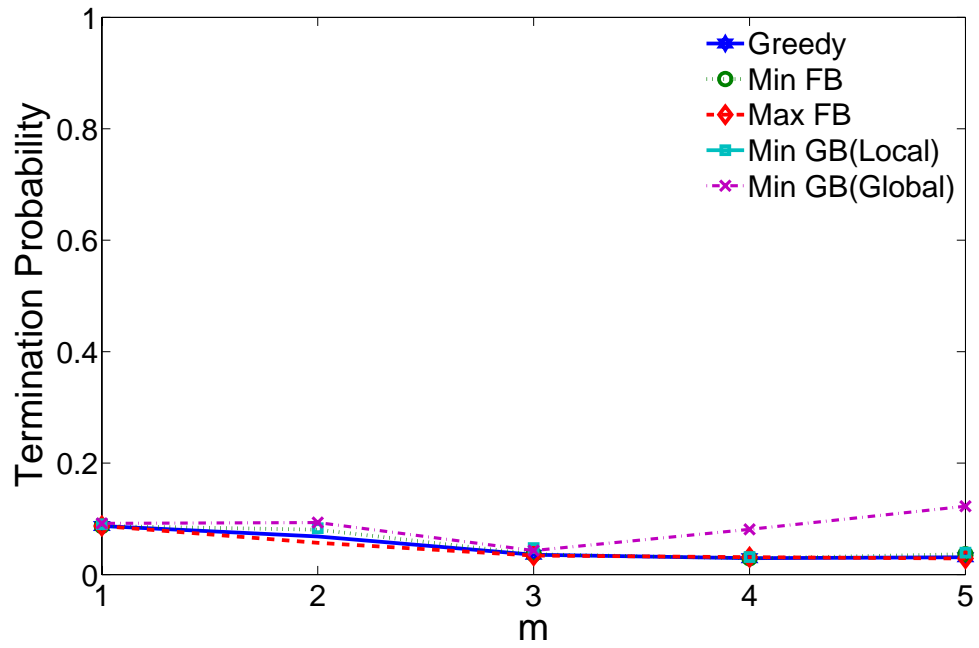
(a) P_f vs. λ_p

(b) P_f vs. m

Figure 5.2: Forced termination probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band, for $\lambda_p = 0.1$

of the primary user arrival rate, λ_p, and the number of channels used by each cognitive user, m, under five different channel assignment schemes.

First, observe that, similar to forced termination probability, the degradation probability increases slowly as the primary user arrival rate increases. That is, as the rate of primary users increases, it is more likely that a CU be affected by a PU arrival, resulting in higher degradation probability. Second, if each CU has more ongoing transmissions, the probability of forced terminating reduces, as shown in Fig.5.2(b), however, the degradation probability increases when the number of channels used by each CU increases, as shown in Fig.5.3(b), since the probability of a CU being affected by a PU activity increases. Third, we observe that degradation probability is relatively larger than forced termination probability, depicted in Fig.5.2. This shows that, upon a PU arrival, CUs are less likely to terminate all of their transmissions, rather CUs have to stop some of their transmissions due to not finding available channels.

5.5.1.4   Impact of Guard Bands on CU Utilization

Now we study the impact of considering guard bands on utilization of cognitive users. Fig.5.4(a) plots the normalized CU utilization as function of the primary user arrival rate, λ_p, and Fig.5.4(b) depicts the normalized CU utilization as a function of the number of channels used by each cognitive user, m, under five different channel assignment schemes.

First, we observe that CU utilization increases slowly under Min GB algorithms as the primary user arrival rate increases. However, CU utilization decreases slowly under other three selection schemes. Although CU utilization is lower under Min GB algorithms but it also increases as the PU arrival rate increases. Second, observe that as the number of channels used by each CU increases, CU utilization increases under all
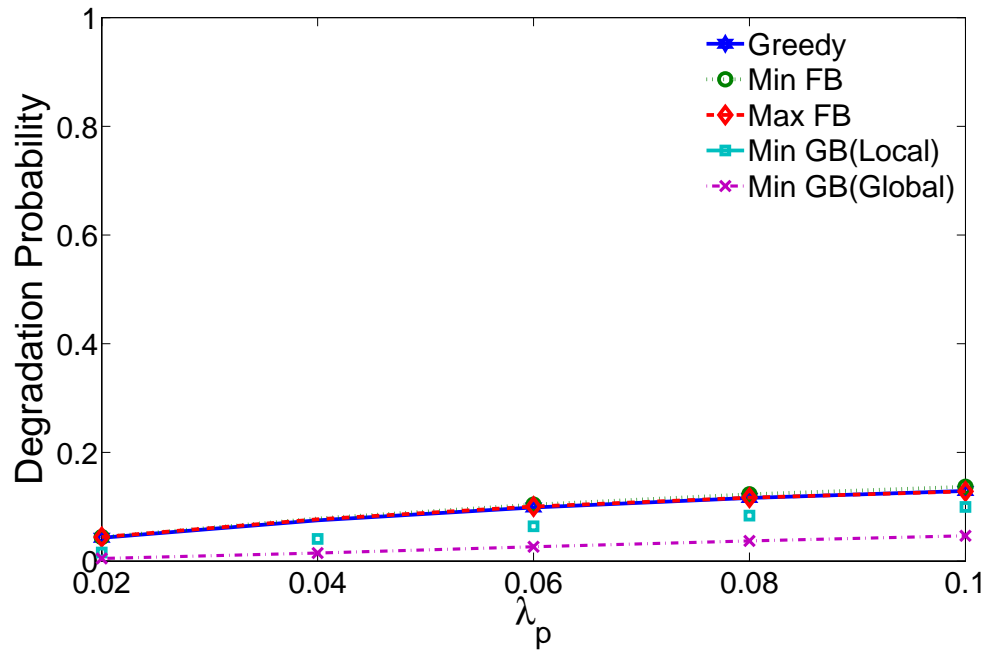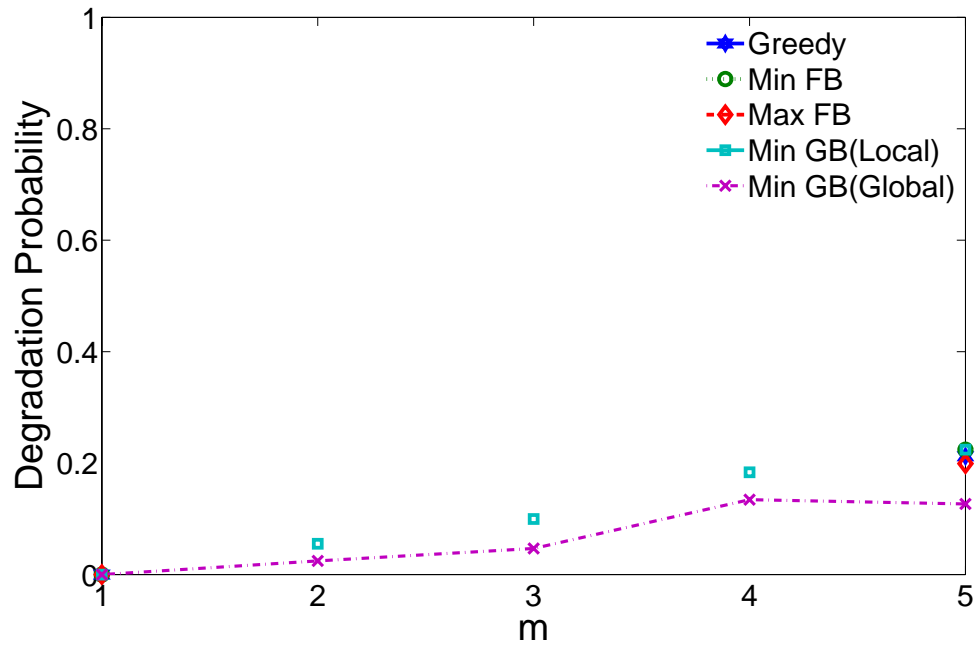
(a) P_d vs. λ_p

(b) P_d vs. m

Figure 5.3: Forced termination probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band reuse, for $\lambda_p = 0.1$

five selection schemes. The reason is that as the number of channels increases, the need for guard bands per CU decreases, thus, increasing the number of available channels and CU utilization.

5.5.1.5 Impact of Guard Bands on Service Degradation

We are interested in the impact of guard bands on the average amount of service degradation when a PU arrival forces a CU to stop some or all of its transmissions. Fig.5.5(a) shows the average service degradation of cognitive users as a function of the primary user arrival rate, λ_p, and Fig.5.5(b) depicts the forced termination probability as a function of the number of channels used by each cognitive user, m, under five different channel assignment schemes.

We observe that MinGB(Local) scheme incurs slightly more service degradation upon a PU arrival. On the other hand, we observe that MinGB(Global) scheme incurs slightly less service degradation upon a PU arrival. We also observe that the amount of service degradation increases as the primary user arrival rate increases since as the number of PUs increases CUs finds fewer channels for their transmission. Finally, as expected, the amount of service degradation increases as the number of channels used by each CU increases since the probability that a PU arrival affects more channels used by CU transmissions increases.

5.5.1.6 Impact of Guard Bands on Spectrum Efficiency

We study the impact of considering the guard bands on spectrum efficiency. Fig.5.6 depicts the spectrum efficiency as a function of the primary user arrival rate, λ_p, and as
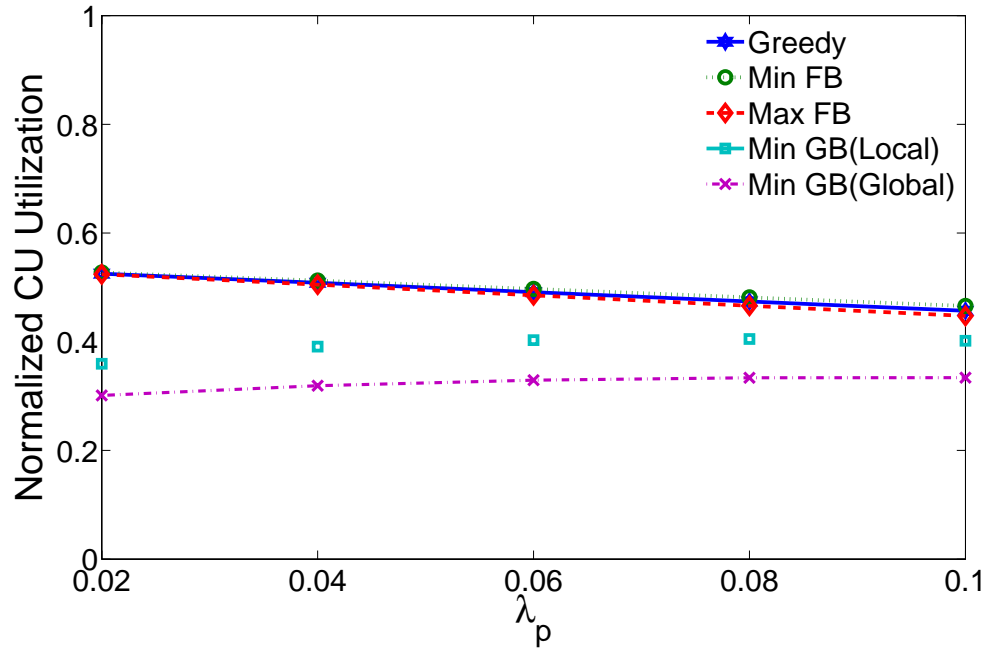
(a) H vs. λ_p
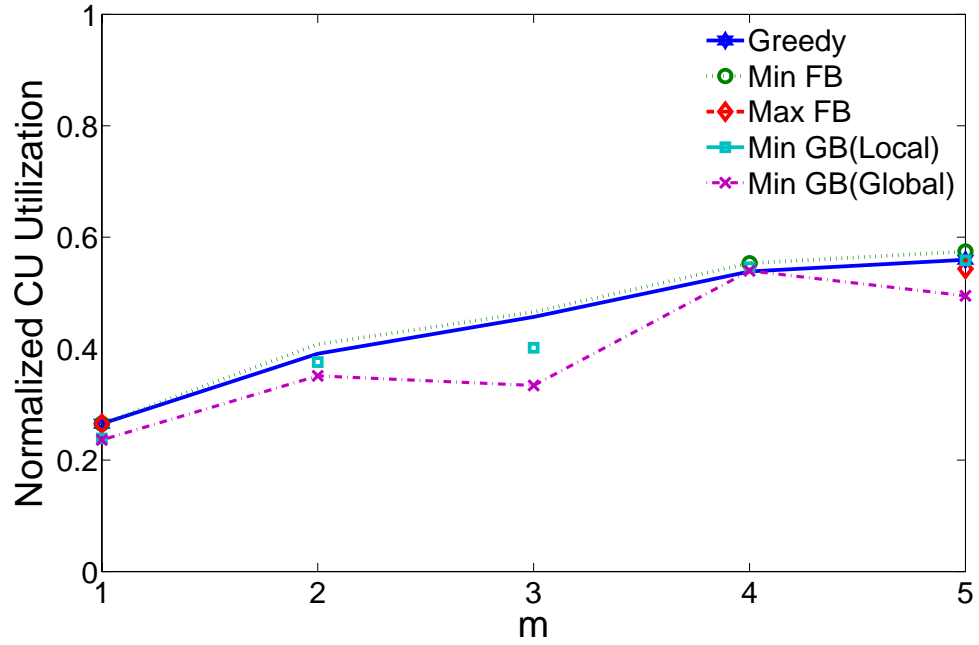
(b) H vs. m

Figure 5.4: Normalized CU utilization as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band reuse, for $\lambda_p = 0.1$

(a) Δ vs. λ_p
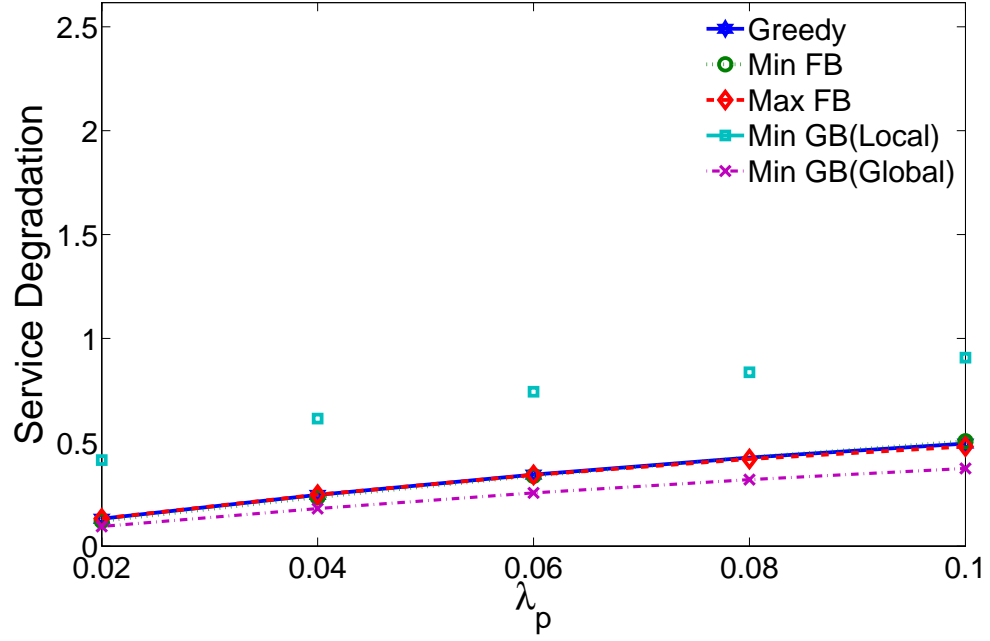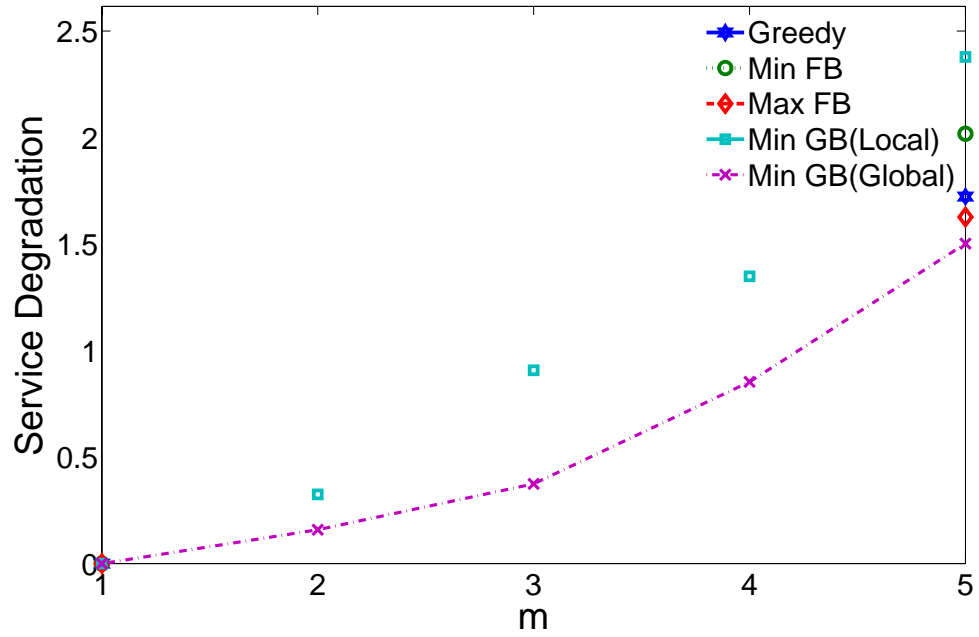
(b) Δ vs. m

Figure 5.5: Forced termination probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band reuse, for $\lambda_p = 0.1$

a function of the number of channels used by each cognitive user, m, under five different channel assignment schemes.

We make three observations. First, in Fig.5.6(a), observe that the spectrum efficiency does not depend on primary user arrival rate as expected. That is, when a the primary user arrival rate increases, the spectrum is used by the primary users instead of cognitive users, hence the number of bands used for transmission does not change rather the type of users change. Second, in Fig.5.6(b), observe that the spectrum efficiency increases as the number of radios used by each cognitive user increases. The reason is that fewer CUs may access the spectrum, hence the number of frequency blocks that are used by CU decreases, thus requiring fewer guard bands. Therefore, user may access the spectrum bands more efficiently. Third, Fig.5.6(b) also shows that the spectrum efficiency increases as the number of channels used by each CU increases, this is simply due to channel bonding done by the CUs in order to reduce the number of guard bands and increasing efficiency.

5.5.1.7   Average Number of Guard Bands under Various Channel Assignment Schemes

We are also interested in the average number of guard bands incurred under each channel assignment scheme described in Section 5.3. The average number of guard bands is shown in Fig.5.7 as a function of primary user arrival rate, λ_p, (shown in Fig. 5.7(a)) and the number of channels used by each cognitive user, m,(shown in Fig. 5.7(a)) under five different channel assignment schemes.

First, we observe that, in Fig.5.7(a), as the primary user arrival rate, λ_p, increases,

(a) Ξ vs. λ_p

(b) Ξ vs. m

Figure 5.6: Spectrum Efficiency as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band reuse, for $\lambda_p = 0.1$
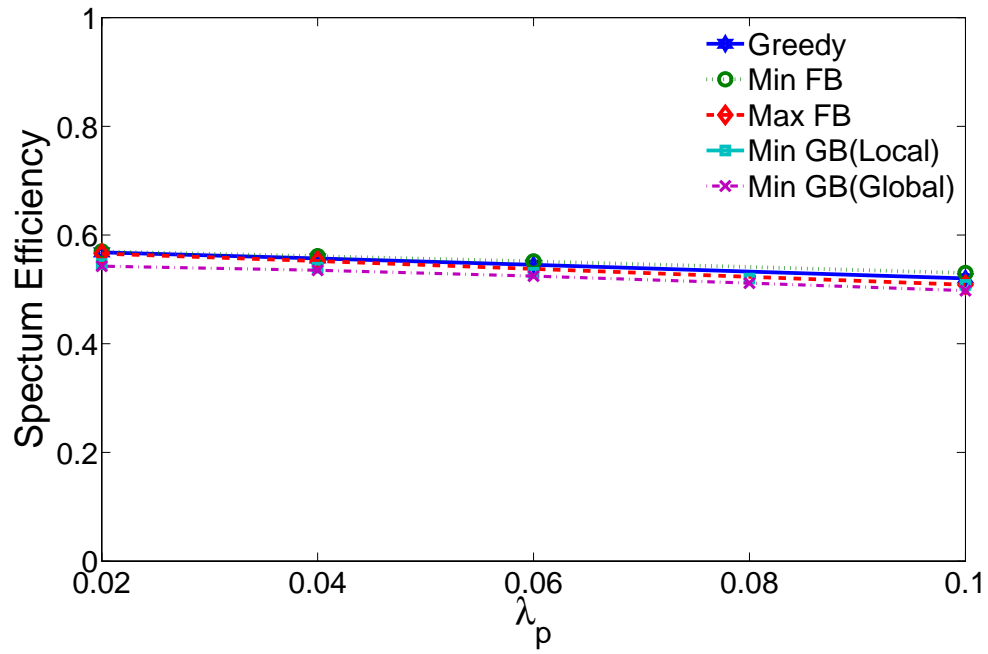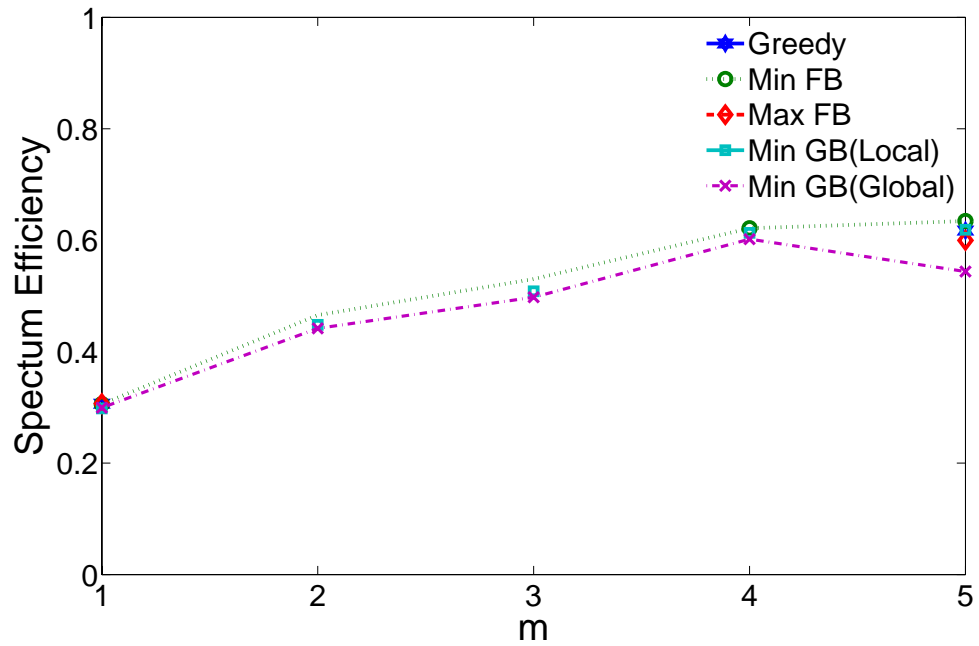
the average number of guard bands increases due to not finding contiguous bands to use as a frequency block; hence CUs need to use more frequency blocks, demanding more guard bands. Second, in Fig.5.7(b), we observe that average number of guard bands decreases as the number of radios used by each CU increases. That is, increasing the number of radios helps the CUs to bond contiguous bands forming larger frequency blocks which requires fewer number of guard bands. Third, as expected, the channel assignment scheme, which intuitively reduces the number of guard bands, imposes fewer guard bands.

5.5.2 Analytic Results with Guard Band Reuse

Next, we analyze the D-OFDM-based CRN performance, where guard band reuse is allowed, under various channel selection schemes.

5.5.2.1 Impact of Guard Bands on Blocking Probability

Fig.5.8(a) and Fig.5.8(b) depict the derived blocking probability of CUs as a function of the primary user arrival rate, λ_p, and the maximum number of channels used by each cognitive user m, respectively, under five different channel assignment schemes.

First, observe that the blocking probability of cognitive users decreases with the primary user arrival rate. That is, as the rate of primary users increases, the network becomes more and more loaded, thus expecting higher blocking probability. However, a PU arrival does not only affect the blocking probability, it also affects degradation probability which will be explained later. On the other hand, as the number of PUs increases, the probability of a PU leaving the system increases, therefore, there are more
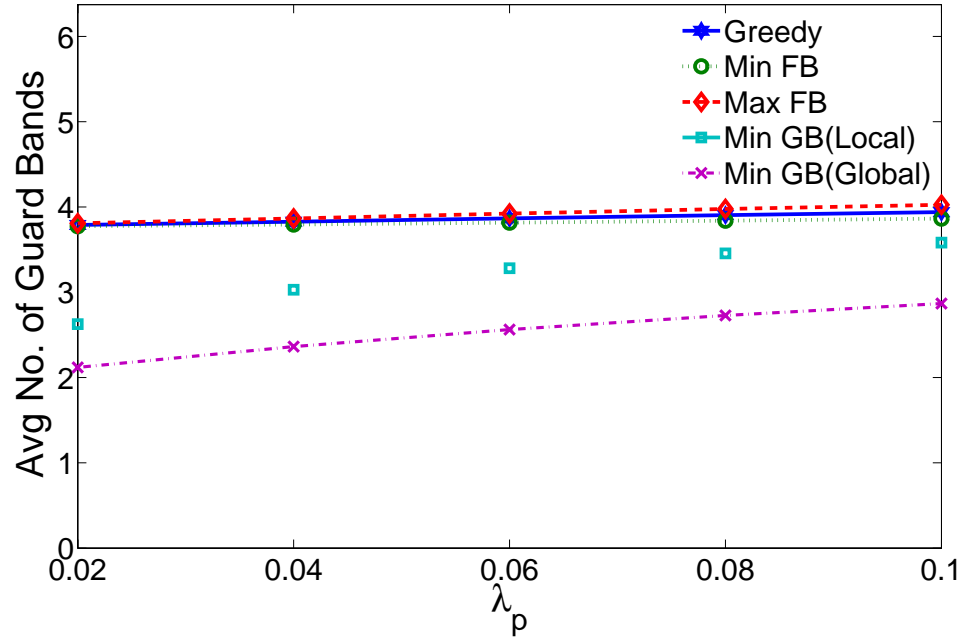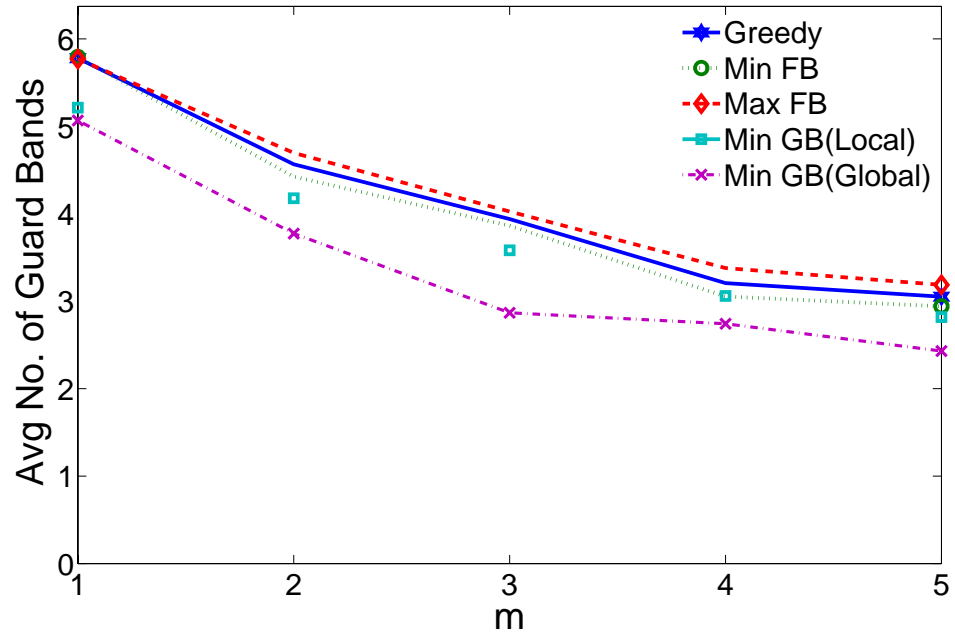
(a) χ vs. λ_p

(b) χ vs. m

Figure 5.7: Average number of guard bands as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes without guard band reuse, for $\lambda_p = 0.1$

idle channels to be used for CUs transmissions. This explains the decrease in blocking probability. Second, as the number of channels per CU increases, the number of occupied bands by CUs increases which generally results in higher blocking probability. This trend of performance behavior is expected, as having less channels to choose from, decreases the chances of cognitive users finding available bands, which explains the increase in the blocking probability of cognitive users. Third, MinGB(Global) results in significantly smaller blocking probabilities when the number of channels per CU increases from 3 to 4. When CUs access more channels, there will be also more idle channels upon CU departure and this is why the blocking probability is reduced.

5.5.2.2   Impact of Guard Bands on Termination Probability

We study the impact of guard bands on forced termination probability of cognitive users. Fig.5.9(a) and Fig.5.9(b) plots the derived forced termination probability of cognitive users as a function of the primary user arrival rate the number of channels accessed by each cognitive user m under five different channel assignment schemes.

We observe that the termination probability increases very slowly as the primary user arrival rate. That is, as the rate of primary users increases, the chance of a CU affected by a PU activity increases, however, the termination probability does not depend on PU arrival rate significantly, rather it depends on total network traffic. Next, observe that when the CUs access more channels, the forced termination probability becomes generally smaller. As the number of channels used by each CU increases, the probability that a CU is affected by PU arrivals decreases.
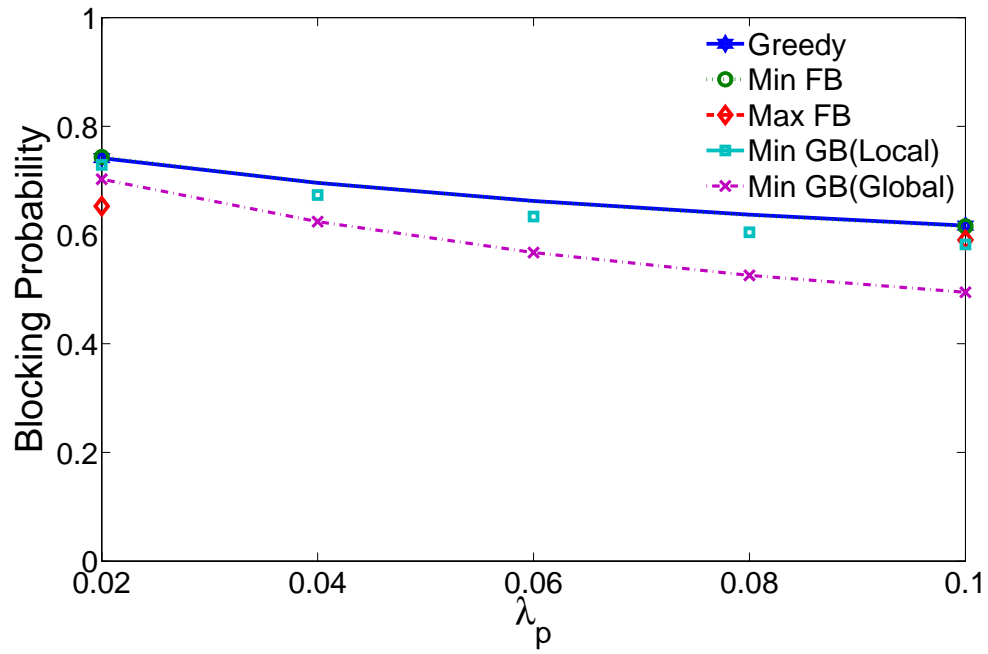
(a) P_b vs. λ_p
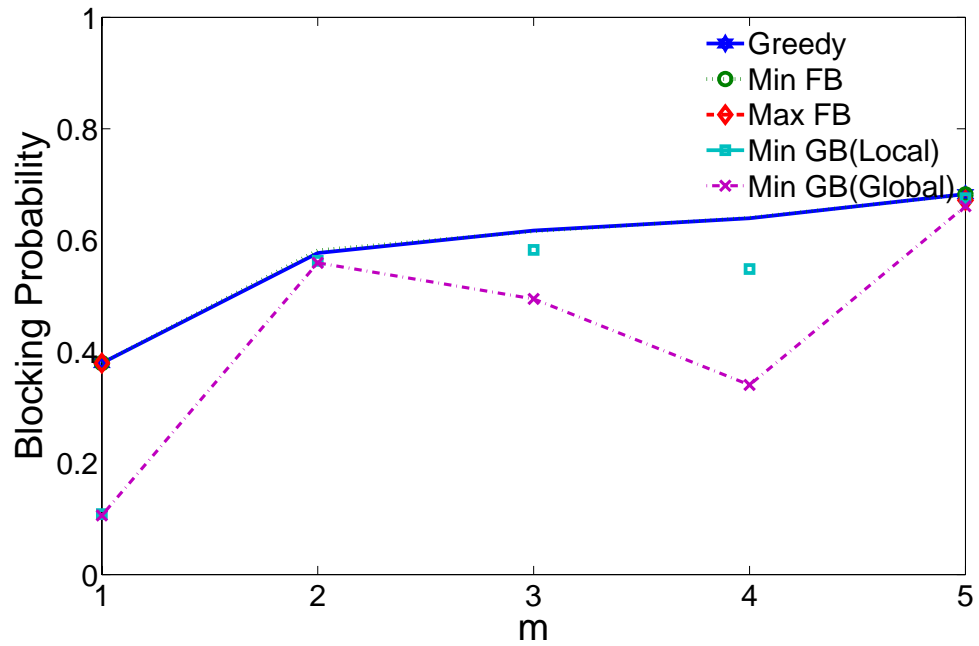
(b) P_b vs. m

Figure 5.8: Blocking probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$
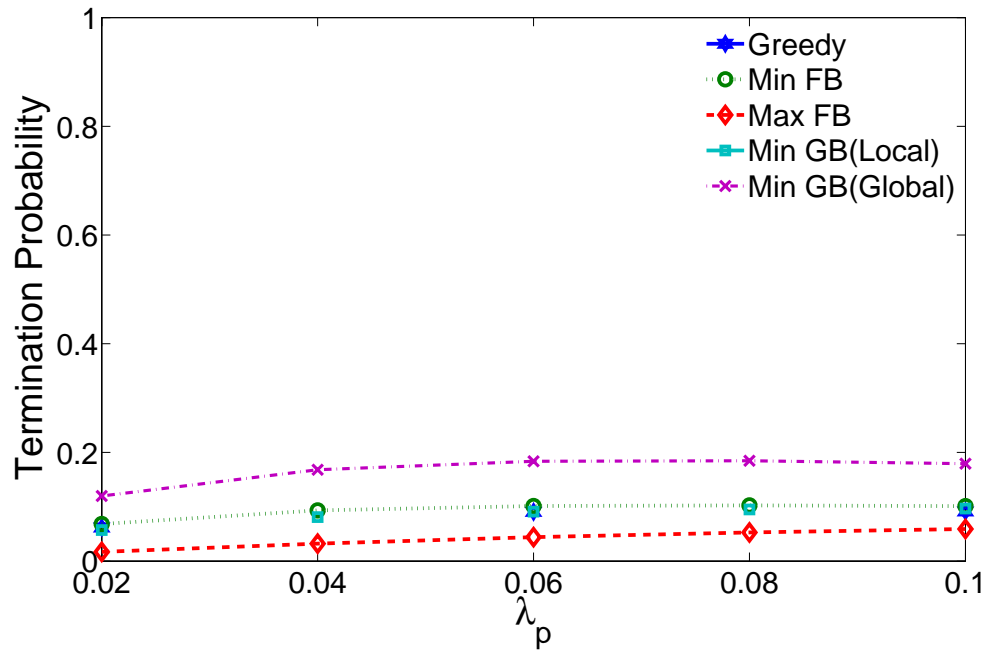
(a) P_f vs. λ_p
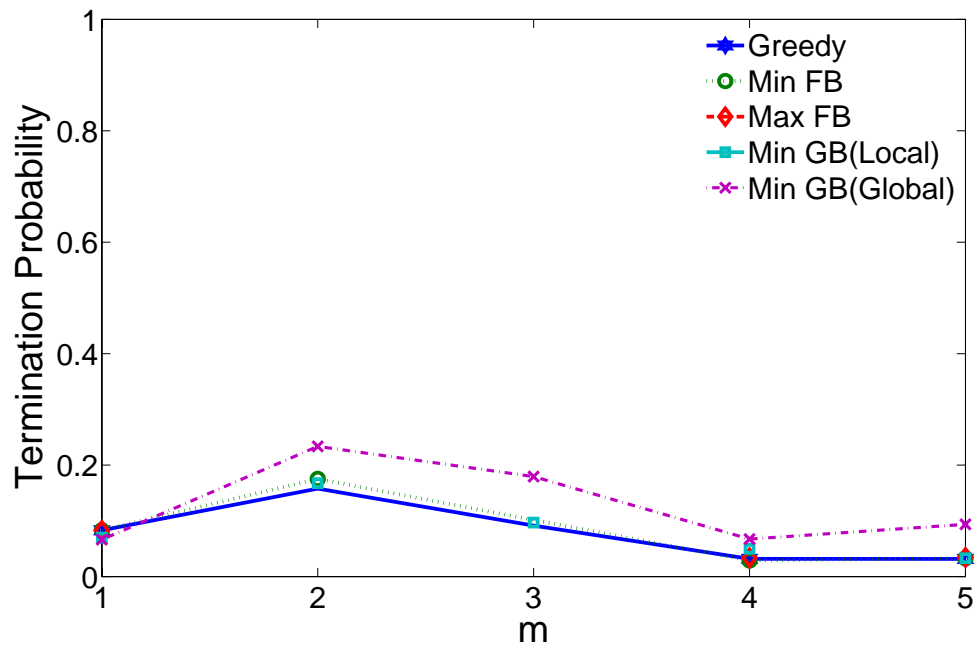
(b) P_f vs. m

Figure 5.9: Forced termination probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$

5.5.2.3   Impact of Guard Bands on Degradation Probability

Now, we study the degradation probability when a PU reclaims its channel and at least a CU transmission is forced to stop transmission as a result of not finding an available band. Fig.5.10 plots the derived degradation probability of cognitive users as a function of the primary user arrival rate, λ_p, and the number of channels used by each cognitive user, m, under five different channel assignment schemes.

First, observe that, similar to forced termination probability, the degradation probability increases slowly as the primary user arrival rate increases. That is, as the rate of primary users increases, it is more likely that a CU be affected by a PU arrival, resulting in higher degradation probability. Second, if each CU access more channels for its transmissions, the probability of forced terminating reduces, as shown in Fig.5.9(b), in contrast, the degradation probability increases when the number of channels used by each CU increases, as shown in Fig.5.10(b), since the probability of a CU being affected by a PU activity increases. Third, we observe that degradation probability is relatively larger than forced termination probabilities, depicted in Fig.5.9. This demonstrates that CUs are less likely to terminate all of their transmissions upon a PU arrival, rather CUs are forced to cease some of their transmissions due to not finding available channels.

5.5.2.4   Impact of Guard Bands on CU Utilization

Next, we study the impact of considering guard bands on utilization of cognitive users. Fig.5.11 plots the normalized CU utilization as function of the primary user arrival rate, λ_p, and the number of channels used by each cognitive user, m, under five different channel assignment schemes.
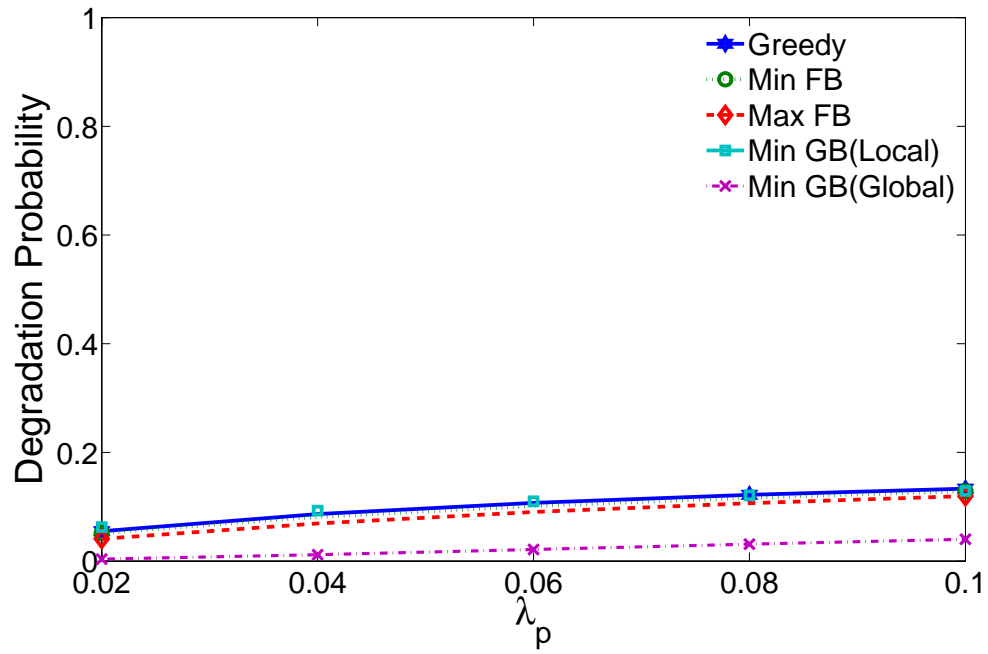
(a) P_d vs. λ_p
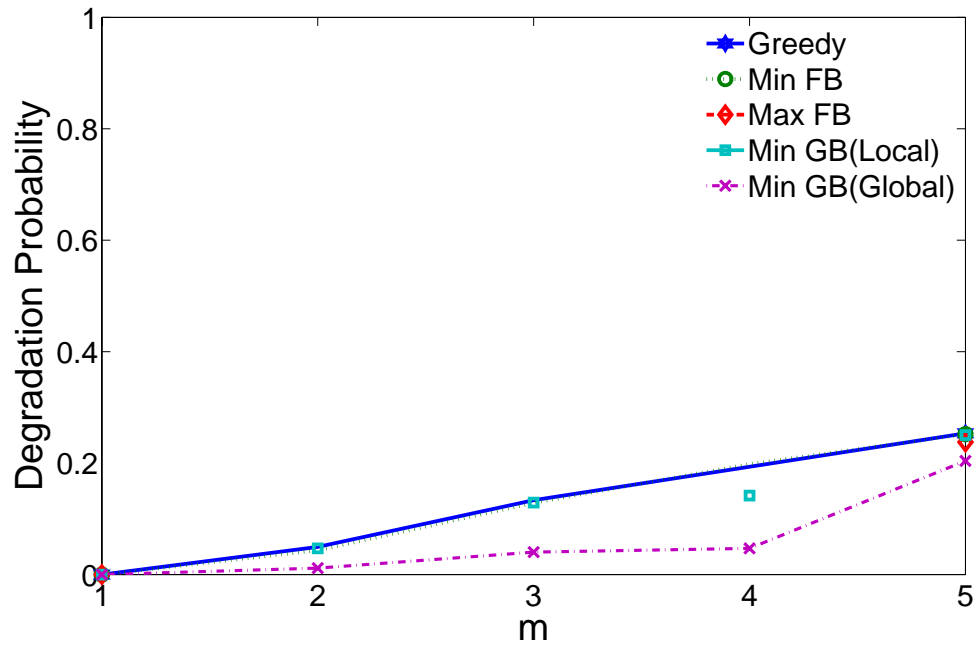
(b) P_d vs. m

Figure 5.10: Forced termination probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$

We observe that CU utilization decreases slowly as the primary user arrival rate increases. As the number of PUs increases the number of channels assigned to CUs decrease, hence the CU utilization decreases. Also, we observe that as the number of channels used by each CU increases, CU utilization increases under five selection schemes. The reason is that as the number of channels increases, the need for guard bands per CU decreases generally, thus, increasing the number of available channels and CU utilization.

5.5.2.5   Impact of Guard Bands on Service Degradation

We are also interested in the impact of deploying guard bands on the average amount of service degradation when a PU arrival forces a CU to stop some or all of its transmissions. Fig.5.12 shows the average service degradation of cognitive users as a function of the primary user arrival rate, λ_p, and the number of channels used by each cognitive user, m, under five different channel assignment schemes.

We observe that MinGB(Local) scheme incurs more service degradation upon a PU arrival and it increases as the PU arrival rate increases. On the other hand, we observe that MinGB(Global) scheme incurs less service degradation upon a PU arrival and the degradation decreases as the PU arrival rate increases. We also observe that the amount of service degradation increases generally as the primary user arrival rate increases. As PU activity increases CUs finds fewer channels for transmission. Finally, as expected, the amount of service degradation increases as the number of channels used by each CU increases since the probability that a PU arrival affects more channels used by CU transmissions increases.
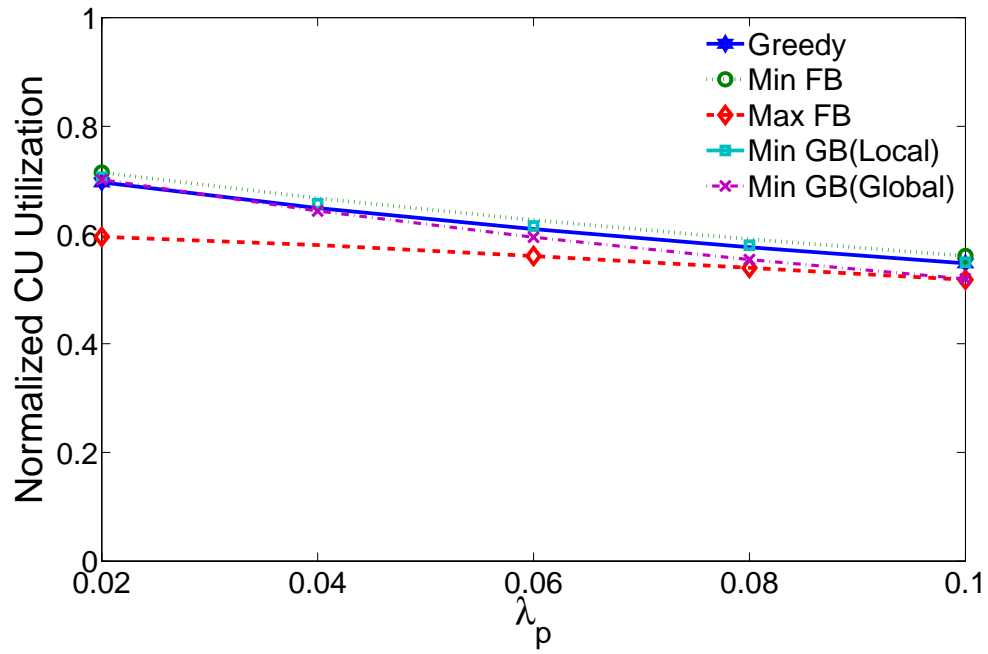
(a) H vs. λ_p
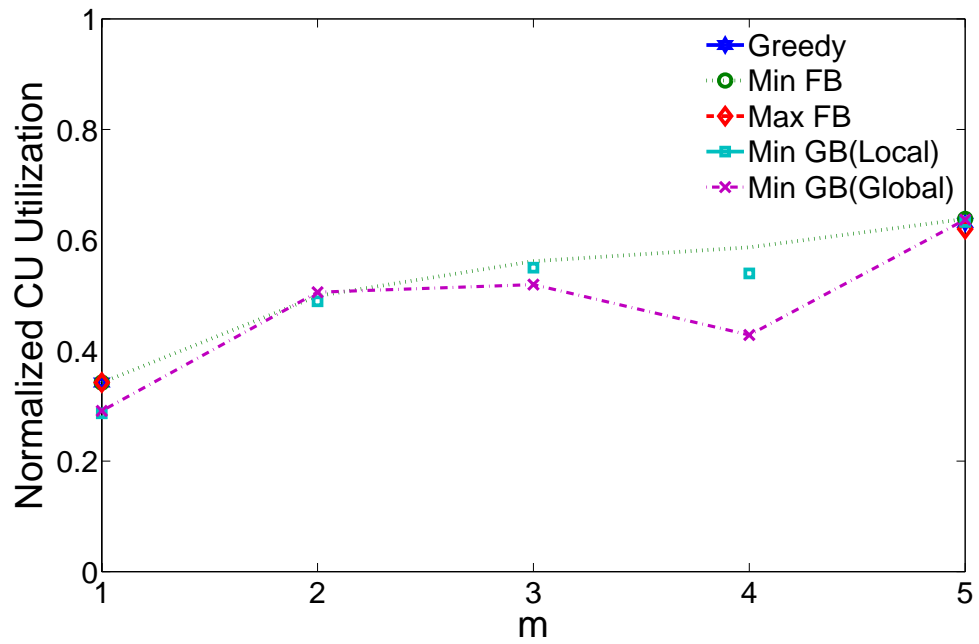
(b) H vs. m

Figure 5.11: Normalized CU utilization as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$

(a) Δ vs. λ_p
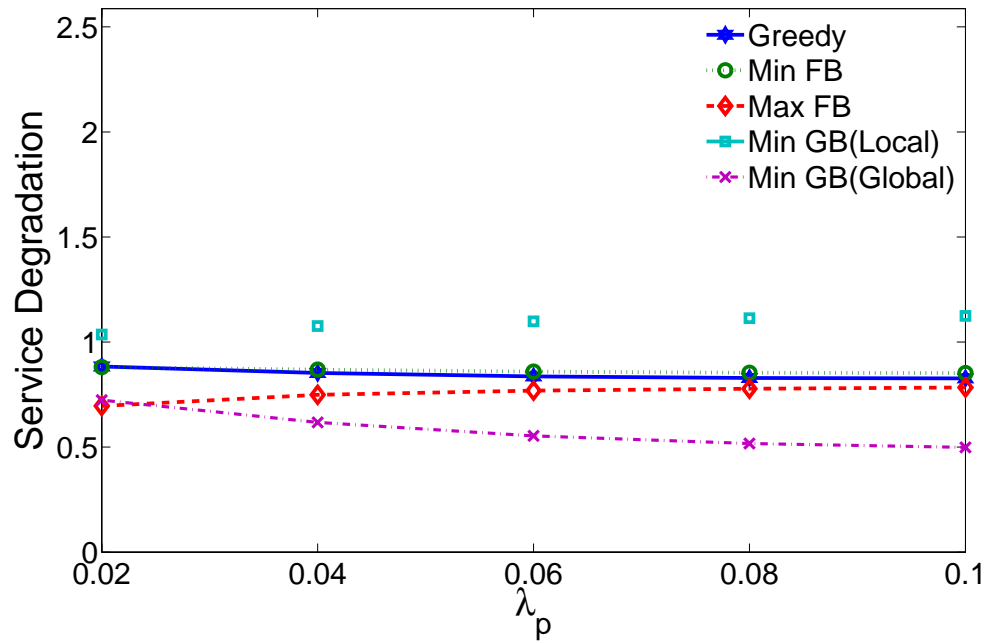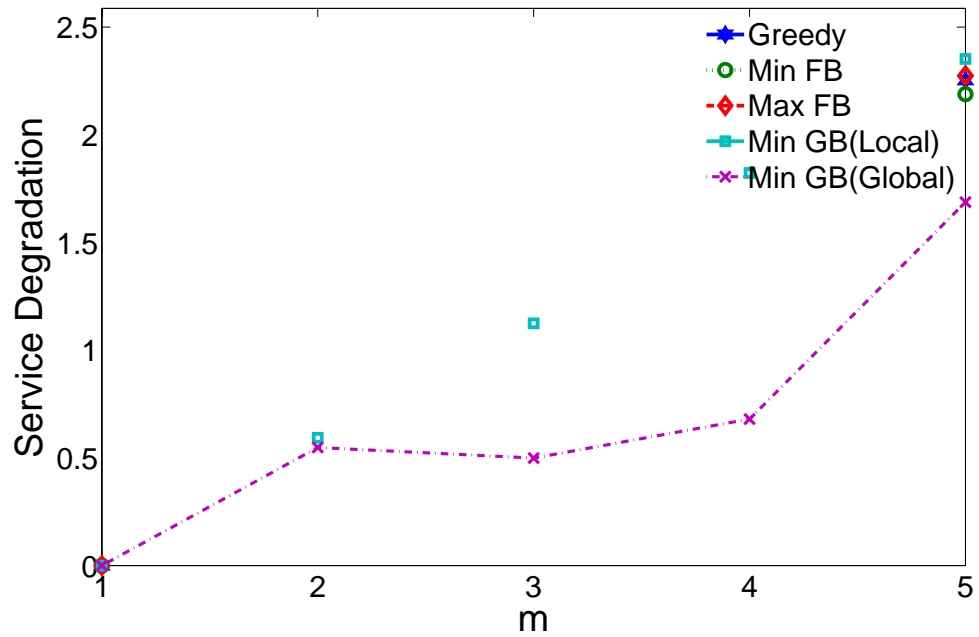
(b) Δ vs. m

Figure 5.12: Forced termination probability as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$

5.5.2.6   Impact of Guard Bands on Spectrum Efficiency

We study the impact of considering the guard bands on spectrum efficiency. Fig.5.13 depicts the spectrum efficiency as a function of the primary user arrival rate, λ_p, and as a function of the number of channels used by each cognitive user, m, under five different channel assignment schemes.

We make two key observations. First, in Fig.5.13(a), observe that the spectrum efficiency decreases as the PU arrival rate increases since a channel used by a PU needs one guards at each side and as the number of PUs increases the number of required guard bands increases, thus reducing spectrum efficiency. Second, in Fig.5.13(b), observe that the spectrum efficiency increases as the number of radios used by each cognitive user increases. The reason is that the length of each CU frequency blocks increases since each CU can bond more channels together, thus requiring fewer guard bands. Therefore, spectrum efficiency increases.

5.5.2.7   Average Number of Guard Bands under Various Channel Assignment Schemes

Last, we are interested in the average number of guard bands incurred by each channel assignment scheme described in Section 5.3. Average number of guard bands as a function of the primary user arrival rate, λ_p, and the number of channels used by each cognitive user, m, are shown in Fig.5.14.

First, we observe that, in Fig.5.14(a), as the primary user arrival rate, λ_p, increases, the average number of guard bands increases due to not finding contiguous bands to use as a frequency block; hence CUs need to use more frequency blocks, demanding more

(a) Ξ vs. λ_p
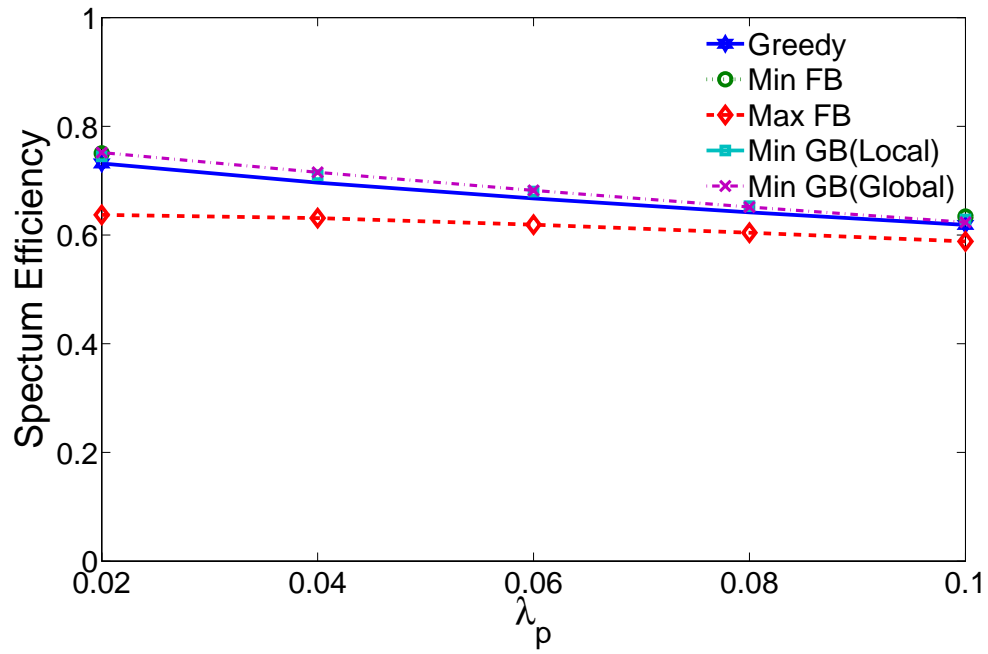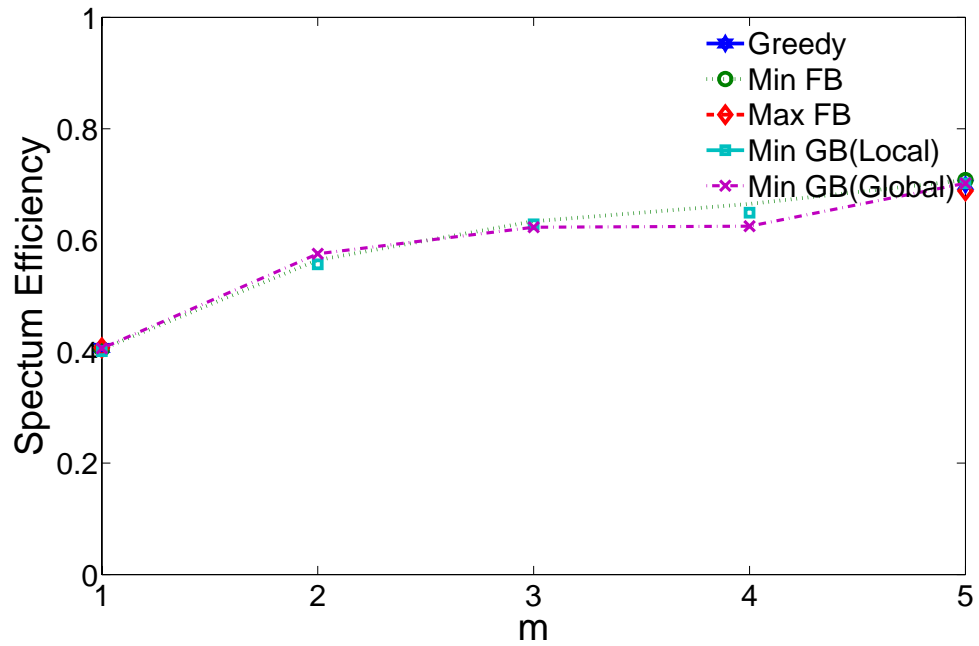
(b) Ξ vs. m

Figure 5.13: Spectrum Efficiency as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$

guard bands. Second, in Fig.5.14(b), we observe that average number of guard bands decreases as the number of channels used by each CU increases. That is, increasing the number of channels helps the CUs to bond contiguous bands forming wider frequency blocks which requires fewer number of guard bands. Third, as expected, the channel assignment scheme, MinGB(Global) , which intuitively must reduce the number of guard bands, imposes fewer guard bands and MaxFB requires more guard bands as expected.

5.6 Summary and Concluding Remarks

In this Chapter, we propose a continuous time Markov model for cognitive radio networks. Our model consider using non-ideal filters, thus requiring guard bands to prevent adjacent channel interference. In our model, we consider CUs that have the capability of bonding and aggregating channels to access multiple channels simultaneously. Two types of cognitive radio model are considered in this work (1) FDM-based CRNs and (2) D-OFDM-based CRNs. In FDM-based CRNs, each frequency block requires to have one guard band at each end. In this case, it is not allowed to reuse a guard band for two adjacent frequency blocks, however, in D-OFDM-based CRNs, guard band reuse is allowed.

Exploiting the proposed model, we derive and analyze various network performance metrics including (1) blocking probability, (2) forced termination probability, (3) degradation probability, (4) CU Utilization, (5) service degradation, (6) spectrum efficiency and (7) the number of guard bands under five various channel selection schemes. Our analysis shows that that disregarding ACI results in inaccurate and misleading outcomes. Therefore, it is imperative to consider guard bands when spectrum sharing schemes are designed for cognitive radio networks.

(a) χ vs. λ_p

(b) χ vs. m

Figure 5.14: Average number of guard bands as a function of (a) the primary arrival rate λ_p, for $m = 3$ (b) the maximum number of channels m, under five different channel assignment schemes with guard band reuse, for $\lambda_p = 0.1$

Chapter 6: Conclusions

In Chapter 2, we propose and evaluate efficient private objective functions that OSA users can use to locate the best spectrum opportunities. OSA users can rely on any learning algorithms to maximize these proposed objective functions, thereby ensuring near-optimal performances in terms of the long-term average received rewards. We showed that these proposed functions (i) receives near-*optimal* rewards, (ii) are highly *scalable* as they perform well for small- as well as large-scale systems, (iii) are highly *learnable* as rewards reach up near-optimal values very quickly, and (iv) are *distributive* as they require information sharing only among users belonging to the same spectrum band.

In Chapter 3, we proposed efficient resource and service management techniques to effectively support SUs in large-scale OSA systems. We showed that the proposed techniques achieve high service satisfaction levels, are very scalable by performing well in small- as well as large-scale systems, are highly learnable by reaching up high values fast, are distributive by requiring information sharing only among agents belonging to the same band, and ensure fairness among SUs by allowing them to receive equal amounts of service.

Chapter 4 models and analyzes the performance behaviors of cognitive radio networks enabled with dynamic multichannel access capability, but while considering realistic channel handoff agility assumptions, where cognitive users can only switch to vacant channels that are immediate neighbors of their current channels. Using Markov chain analysis, we model cognitive access networks with restricted channel handoff agility as a continuous-time Markov process, and analytically derive the forced access termination

and blocking probabilities of cognitive users, and evaluate the spectrum access efficiency of cognitive networks. Using MATLAB simulations, we also validate our analytically derived performance results. Our obtained results demonstrate the impact and importance of considering realistic channel handoff agility in cognitive radio access on the cognitive radio network performances in terms of users' blocking and termination probabilities, as well as cognitive spectrum access efficiency. This chapter demonstrates the cognitive radio performance implications of the commonly made spectrum-handoff agility assumption of allowing cognitive users to switch to any available band, regardless of how far the target band is from the current band. Our findings in Chapter 4 show that the achievable performance of cognitive radio networks in terms of spectrum access capability and efficiency can be significantly lesser than what existing works usually claim, due to the limited nature of spectrum handoff agility that most works do not account for. We conclude that making unrealistic assumption regarding the spectrum handoff agility may lead to very inaccurate and misleading results, and it is then imperative that performance studies of cognitive radio networks do account for the restricted agility of channel switching when modeling and assessing the performance of such networks.

In Chapter 5, we propose a continuous time Markov model for cognitive radio networks. Our model consider using non-ideal filters, thus requiring guard bands to prevent adjacent channel interference. In our model, we consider CUs that have the capability of bonding and aggregating channels to access multiple channels simultaneously. Two types of cognitive radio model are considered in this chapter (1) FDM-based CRNs and (2) D-OFDM-based CRNs. In FDM-based CRNs, each frequency block requires to have one guard band at each end. In this case, it is not allowed to reuse a guard band for two adjacent frequency blocks, however, in D-OFDM-based CRNs, guard band reuse is allowed. Exploiting the proposed model, we derive and analyze various network perfor-

mance metrics including (1) blocking probability, (2) forced termination probability, (3) degradation probability, (4) CU Utilization, (5) service degradation, (6) spectrum efficiency and (7) the number of guard bands under five various channel selection schemes. Our analysis shows that that disregarding ACI results in inaccurate and misleading outcomes. Therefore, it is imperative to consider guard bands when spectrum sharing schemes are designed for cognitive radio networks.
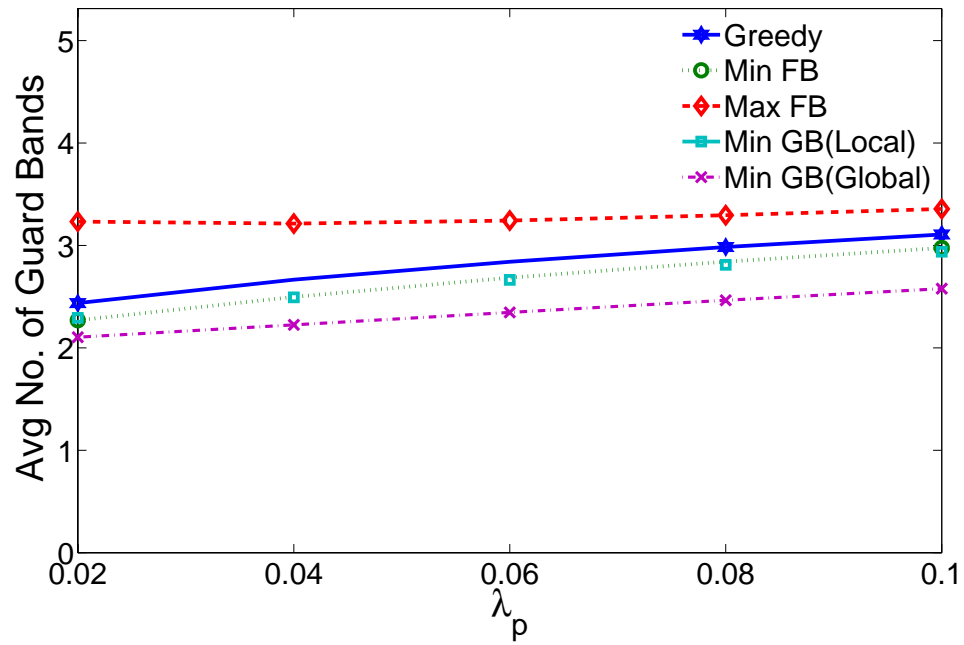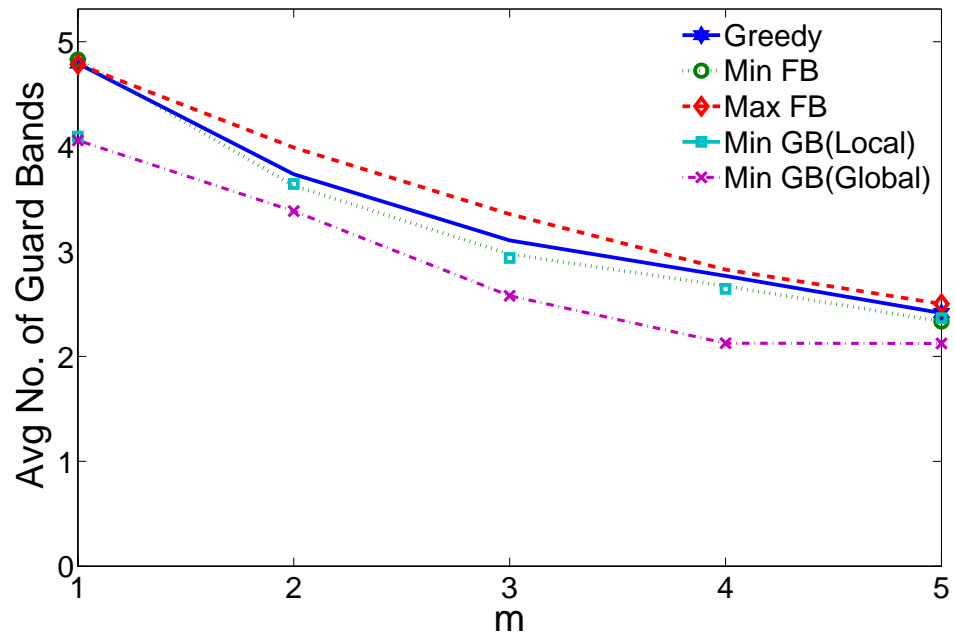
Bibliography

[1] Fcc, spectrum policy task force report, et docket no. 02-155. Nov. 2002.

[2] *FCC, Spectrum Policy Task Force (SPTF), Report of the Spectrum Efficiency WG, Report ET Docet no. 02-135, November*, 2002.

[3] Nadia Adem and Bechir Hamdaoui. Delay performance modeling and analysis in clustered cognitive radio networks. In *Global Communications Conference (GLOBECOM), 2014 IEEE*, pages 193–198. IEEE, 2014.

[4] A. Agogino and K. Tumer. Unifying temporal and structural credit assignment problems. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, New York, NY, July 2004.

[5] A. Agogino and K. Tumer. Multi agent reward analysis for learning in noisy domains. In *Proc. of the Fourth Int'l Joint Conf. on Autonomous Agents and Multi-Agent Systems*, Utrecht, Netherlands, July 2005.

[6] A. K. Agogino and K. Tumer. Analyzing and visualizing multiagent rewards in dynamic and stochastic environments. *Journal of Autonomous Agents and Multi Agent Systems*, 17(2):320–338, 2008.

[7] A. K. Agogino and K. Tumer. Efficient evaluation functions for evolving coordination. *Evolutionary Computation*, 16(2):257–288, 2008.

[8] S. H. A. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari. Optimality of myopic sensing in multi-channel opportunistic access. *IEEE Transactions on Information Theory*, 2009.

[9] W. Ahmed, J. Gao, H. A. Suraweera, and M. Faulkner. Comments on analysis of cognitive radio spectrum access with optimal channel reservation. *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*, 8(9):4488–4491, September 2009.

[10] Ian F Akyildiz, Won-Yeol Lee, Mehmet C Vuran, and Shantidev Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Computer Networks*, 50(13):2127–2159, 2006.

[11] Omar Alsaleh, Bechir Hamdaoui, and Alan Fern. Q-learning for opportunistic spectrum access. In *Proceedings of the 6th International Wireless Communications and Mobile Computing Conference*, pages 220–224. ACM, 2010.

[12] Omar Alsaleh, Pavithra Venkatraman, Bechir Hamdaoui, and Alan Fern. Enabling opportunistic and dynamic spectrum access through learning techniques. *Wireless Communications and Mobile Computing*, 11(12):1497–1506, 2011.

[13] Tamara Alshammari, Bechir Hamdaoui, Mohsen Guizani, and Ammar Rayes. Overcoming user selfishness in dsa systems through credit-based resource allocation. In *Communications (ICC), 2014 IEEE International Conference on*, pages 318–323. IEEE, 2014.

[14] Tamara Alshammari, Bechir Hamdaoui, Mohsen Guizani, and Ammar Rayes. Malicious-proof and fair credit-based resource allocation techniques for dsa systems. *Wireless Communications, IEEE Transactions on*, 14(2):606–615, 2015.

[15] Haythem A Bany Salameh and Marwan Krunz. Channel access protocols for multi-hop opportunistic networks: challenges and recent developments. *Network, IEEE*, 23(4):14–19, 2009.

[16] Haythem A Bany Salameh, Marwan M Krunz, and Ossama Younis. Mac protocol for opportunistic cognitive radio networks with soft guarantees. *Mobile Computing, IEEE Transactions on*, 8(10):1339–1352, 2009.

[17] Mahdi Ben Ghorbel, Bechir Hamdaoui, Rami Hamdi, Mohsen Guizani, and MohammadJavad NoroozOliaee. Distributed dynamic spectrum access with adaptive power allocation: Energy efficiency and cross-layer awareness. In *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*, pages 694–699. IEEE, 2014.

[18] Mahdi Ben Ghorbel, Bassem Khalfi, Bechir Hamdaoui, and Mohsen Guizani. Resources allocation for large-scale dynamic spectrum access system using particle filtering. In *Globecom Workshops (GC Wkshps), 2014*, pages 219–224. IEEE, 2014.

[19] U. Berthold, M. Van Der Schaar, and F. K. Jondral. Detection of spectral resources in cognitive radios using reinforcement learning. In *Proceedings of IEEE DySPAN*, pages 1–5, 2008.

[20] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[21] Nessrine Chakchouk and Bechir Hamdaoui. Traffic and interference aware scheduling for multiradio multichannel wireless mesh networks. *Vehicular Technology, IEEE Transactions on*, 60(2):555–565, 2011.

[22] Nessrine Chakchouk and Bechir Hamdaoui. Statistical characterization of uplink interference in two-tier co-channel femtocell networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2012 8th International*, pages 968–973. IEEE, 2012.

[23] Nessrine Chakchouk and Bechir Hamdaoui. Uplink performance characterization and analysis of two-tier femtocell networks. *Vehicular Technology, IEEE Transactions on*, 61(9):4057–4068, 2012.

[24] Ranveer Chandra, Ratul Mahajan, Thomas Moscibroda, Ramya Raghavendra, and Paramvir Bahl. A case for adapting channel width in wireless networks. *ACM SIGCOMM computer communication review*, 38(4):135–146, 2008.

[25] L. Chen, S. Iellamo, M. Coupechoux, and P. Godlewski. An auction framework for spectrum allocation with interference constraint in cognitive radio networks. In *Proceedings of IEEE INFOCOM*, 2010.

[26] Y. Chen, Q. Zhao, and A. Swami. Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors. *IEEE Trans. on Inf. Theory*, May 2008.

[27] G. Cheng, W. Liu, Y. Li, and W. Cheng. Joint on-demand routing and spectrum assignment in cognitive radio networks. In *Proceedings of IEEE ICC*, 2007.

[28] Federal Communications Commission et al. Second report and order and memorandum opinion and order, in the matter of unlicensed operation in the tv broadcast bands (et docket no. 04-186) and additional spectrum for unlicensed devices below 900 mhz and in the 3 ghz band (et docket no. 02-380), fcc 08-260, 2008.

[29] C. Cormiob and K. R. Chowdhurya. A survey on MAC protocols for cognitive radio networks. *Ad Hoc Networks*, 7(7), 2009.

[30] D Dechene and Abdallah Shami. Energy-aware resource allocation strategies for lte uplink with synchronous harq constraints. *IEEE Transactions on Mobile Computing*, (99):1–1, 2014.

[31] A. De Domenico, E. Calvanese Strinati, and M.-G. Di Benedetto. A survey on MAC strategies for cognitive radio networks. *IEEE COMMUNICATIONS SURVEYS & TUTORIALS*, 14(1), 2012.

[32] L. Duan, J. Huang, and B. Shou. Competition with dynamic spectrum leasing. In *Proceedings of IEEE DySPAN*, 2010.

[33] Christopher Edmonds, Daniel Joiner, Steven Springer, Kris Stephen, and Bechir Hamdaoui. Cognitive wireless mesh network testbed. In *Wireless Communications and Mobile Computing Conference, 2008. IWCMC'08. International*, pages 373–376. IEEE, 2008.

[34] Samina Ehsan, Bechir Hamdaoui, and Mohsen Guizani. Feasibility conditions for rate-constrained routing in power-limited multichannel wsns. In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pages 1–6. IEEE, 2011.

[35] Samina Ehsan, Bechir Hamdaoui, and Mohsen Guizani. Radio and medium access contention aware routing for lifetime maximization in multichannel sensor networks. *Wireless Communications, IEEE Transactions on*, 11(9):3058–3067, 2012.

[36] Chong Feng, Wenbo Wang, and Xiao Jiang. Cognitive learning-based spectrum handoff for cognitive radio network. *International Journal of Computer and Communication Engineering*, 1, November 2012.

[37] Z. Feng and Y. Yang. Throughput analysis of secondary networks in dynamic spectrum access networks. In *Proceedings of IEEE INFOCOM*, 2010.

[38] Y. Gai, B. Krishnamachari, and R. Jain. Learning multiuser channel allocations in cognitive radio networks: A combinatorial multi-armed bandit formulation. In *Proceddings of IEEE DySPAN*, 2010.

[39] M. Gandetto and C. Regazzoni. Spectrum sensing: a distributed approach for cognitive terminals. *IEEE Journal of Selected Areas in Communications*, April 2007.

[40] Mohsen Guizani, Bassem Khalfi, Mahdi Ben Ghorbel, and Bechir Hamdaoui. Large-scale cognitive cellular systems: resource management overview. *Communications Magazine, IEEE*, 53(5):44–51, 2015.

[41] G. Gur, S. Bayhan, and F. Alagoz. Cognitive femtocell networks: an overlay architecture for localized dynamic spectrum access. *IEEE Wireless Communications*, 17(4), 2010.

[42] B. Hamdaoui and K. G. Shin. OS-MAC: An efficient MAC protocol for spectrum-agile wireless networks. *IEEE Transactions on Mobile Computing*, August 2008.

[43] Bechir Hamdaoui. Adaptive spectrum assessment for opportunistic access in cognitive radio networks. *Wireless Communications, IEEE Transactions on*, 8(2):922–930.

[44] Bechir Hamdaoui. Optimal discovery of bandwidth opportunities in spectrum agile networks. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pages 1–5. IEEE, 2008.

[45] Bechir Hamdaoui, Tamara Alshammari, and Mohsen Guizani. Exploiting 4g mobile user cooperation for energy conservation: challenges and opportunities. *Wireless Communications, IEEE*, 20(5):62–67, 2013.

[46] Bechir Hamdaoui, Moncef Elaoud, and Parameswaran Ramanathan. A delay-based admission control mechanism for multimedia support in ieee 802.11 e wireless lans. *Wireless Networks*, 15(7):875–886, 2009.

[47] Bechir Hamdaoui, M NoroozOliaee, Kagan Tumer, and Ammar Rayes. Coordinating secondary-user behaviors for inelastic traffic reward maximization in large-scale osa networks. *Network and Service Management, IEEE Transactions on*, 9(4):501–513, 2012.

[48] Bechir Hamdaoui, MohammadJavad NoroozOliaee, Kagan Tumer, and Ammar Rayes. Aligning spectrum-user objectives for maximum inelastic-traffic reward. In *Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on*, pages 1–6. IEEE, 2011.

[49] Bechir Hamdaoui and Parameswaran Ramanathan. Lifetime-throughput tradeoff for elastic traffic in multi-hop hotspot networks. In *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, volume 3, pages 1565–1569. IEEE, 2004.

[50] Bechir Hamdaoui and Parameswaran Ramanathan. Link-bandwidth calculation for qos routing in wireless ad-hoc networks using directional communications. In *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, volume 1, pages 91–94. IEEE, 2005.

[51] Bechir Hamdaoui and Parameswaran Ramanathan. Sufficient conditions for flow admission control in wireless ad-hoc networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9(4):15–24, 2005.

[52] Bechir Hamdaoui and Parameswaran Ramanathan. A cross-layer admission control framework for wireless ad-hoc networks using multiple antennas. *Wireless Communications, IEEE Transactions on*, 6(11):4014–4024, 2007.

[53] Bechir Hamdaoui and Parameswaran Ramanathan. Cross-layer optimized conditions for qos support in multi-hop wireless networks with mimo links. *Selected Areas in Communications, IEEE Journal on*, 25(4):667–677, 2007.

[54] Bechir Hamdaoui and Kang G Shin. Characterization and analysis of multi-hop wireless mimo network throughput. In *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, pages 120–129. ACM, 2007.

[55] Bechir Hamdaoui and Kang G Shin. Constraint design and modelling of multi-hop multi-band wireless mimo networks. In *Wireless Communications and Mobile Computing Conference, 2008. IWCMC'08. International*, pages 916–919. IEEE, 2008.

[56] Bechir Hamdaoui and Kang G Shin. On the achievable throughput of multi-band multi-antenna wireless mesh networks. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pages 1–5. IEEE, 2008.

[57] Bechir Hamdaoui and Kang G Shin. Throughput behavior in multihop multi-antenna wireless networks. *Mobile Computing, IEEE Transactions on*, 8(11):1480–1494, 2009.

[58] Bechir Hamdaoui and Kang G Shin. Maximum achievable throughput in multiband multiantenna wireless mesh networks. *Mobile Computing, IEEE Transactions on*, 9(6):838–849, 2010.

[59] Bechir Hamdaoui, Kang G Shin, and Megha Maiya. Constraint design and throughput evaluation in multi-band wireless networks using multiple-input multiple-output links. *IETE Technical Review*, 26(2):101–107, 2009.

[60] Bechir Hamdaoui, Pavithra Venkatraman, and Mohsen Guizani. Opportunistic exploitation of bandwidth resources through reinforcement learning. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6. IEEE, 2009.

[61] Rami Hamdi, Mahdi Ben Ghorbel, Bechir Hamdaoui, and Mohsen Guizani. Design and implementation of distributed dynamic spectrum allocation protocol. In *Communications Workshops (ICC), 2014 IEEE International Conference on*, pages 274–278. IEEE, 2014.

[62] Soumaya Hamouda and Bechir Hamdaoui. Dynamic spectrum access in heterogeneous networks: Hsdpa and wimax. In *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, pages 1253–1257. ACM, 2009.

[63] Soumaya Hamouda and Bechir Hamdaoui. Dynamic allocation of spectrum resources in heterogeneous networks under interference constraints. *International*

Journal of Autonomous and Adaptive Communications Systems, 5(4):329–341, 2012.

[64] Z. Han, C. Pandana, and K. J. R. Liu. Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning. In *Proceedings of IEEE WCNC*, 2007.

[65] H. Harada. A software defined cognitive radio prototype. In *Proc. of IEEE PIMRC*, 2007.

[66] H. Harada. A feasibility study on software defined cognitive radio equipment. In *Proc. of IEEE DySPAN*, 2008.

[67] Hiroshi Harada. A feasibility study on software defined cognitive radio equipment. In *New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on*, pages 1–12. IEEE, 2008.

[68] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

[69] Samer Hijazi, Bala Natarajan, Marco Michelini, Zhiqiang Wu, and Carl R Nassar. Flexible spectrum use and better coexistence at the physical layer of future wireless systems via a multicarrier platform. *Wireless Communications, IEEE*, 11(2):64–71, 2004.

[70] J. Jia, Q. Zhang, and M. Liu. Revenue generation for truthful spectrum auction in dynamic spectrum access. In *Proceedings of ACM MobiHoc*, 2009.

[71] J. Jia, Q. Zhang, and X. Shen. HC-MAC: a hardware-constrainted cognitive MAC for efficient spectrum management. *IEEE Journal on Selected Areas in Communications*, January 2008.

[72] Juncheng Jia, Jin Zhang, and Qian Zhang. Cooperative relay for cognitive radio networks. In *INFOCOM 2009, IEEE*, pages 2304–2312. IEEE, 2009.

[73] Chunxiao Jiang, Yan Chen, KJ Ray Liu, and Yong Ren. Renewal-theoretical dynamic spectrum access in cognitive radio network with unknown primary behavior. *Selected Areas in Communications, IEEE Journal on*, 31(3):406–416, 2013.

[74] X. Jing and D. Raychaudhuri. Global control plane architecture for cognitive radio networks. In *Proceedings of IEEE ICC*, 2007.

[75] Shaunak Joshi, Przemysław Pawełczak, Danijela Čabrić, and John Villasenor. When channel bonding is beneficial for opportunistic spectrum access networks. *Wireless Communications, IEEE Transactions on*, 11(11):3942–3956, 2012.

[76] G. S. Kasbekar and S. Sarkar. Spectrum auction framework for access allocation in cognitive radio networks. In *Proceedings of ACM MobiHoc*, 2009.

[77] Bassem Khalfi, Mahdi Ben Ghorbel, Bechir Hamdaoui, and Mohsen Guizani. Optimal power allocation for smart-grid powered point-to-point cognitive radio system. In *Computing, Communications and IT Applications Conference (ComComAp), 2014 IEEE*, pages 316–320. IEEE, 2014.

[78] Bassem Khalfi, Mahdi Ben Ghorbel, Bechir Hamdaoui, and Mohsen Guizani. Dynamic power pricing using distributed resource allocation for large-scale dsa systems. In *Wireless Communications and Networking Conference (WCNC), 2015 IEEE*, pages 1090–1094. IEEE, 2015.

[79] H. Kim and K. G. Shin. Efficient discovery of spectrum oppotunities with MAC-layer sensing in cognitive radio networks. *IEEE Transactions on Mobile Computing*, May 2008.

[80] K. Kim, I. A. Akbar, K. K. Bae, J.-S. Um, C. M. Spooner, and J. H. Reed. Cyclostationary approaches to signal detection and classification in cognitive radio. In *Proceedings of IEEE DySPAN*, 2007.

[81] Vinod Kone, Lei Yang, Xue Yang, Ben Y Zhao, and Haitao Zheng. On the feasibility of effective opportunistic spectrum access. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 151–164. ACM, 2010.

[82] Marwan Krunz and David Manzi. Channel access and traffic control for dynamic-spectrum networks with single-transmit, dual-receive radios. *Computer Communications*, 34(8):935–947, 2011.

[83] M. Levorato, L. Badia, , U. Mitra, and M. Zorzi. An analysis of cognitive networks for unslotted time and reactive users. In *Proceedings of IEEE MASS*, 2010.

[84] Xin Li, Qianchuan Zhao, Xiaohong Guan, and Lang Tong. Optimal cognitive access of markovian channels under tight collision constraints. *Selected Areas in Communications, IEEE Journal on*, 29(4):746–756, 2011.

[85] X. Lin and S. Rasool. A distributed joint channel-assignment, scheduling and routing algorithm for multi-channel ad hoc wireless networks. In *Proceedings of IEEE INFOCOM*, 2007.

[86] H. Liu, B. Krishnamachari, and Q. Zhao. Cooperation and learning in multiuser opportunistic spectrum access. In *Proceedings of IEEE ICC*, 2008.

[87] K. Liu and Q. Zhao. Distributed learning in multi-armed bandit with multiple players. *IEEE Trans. on Signal Processing*, 58(11), November 2010.

[88] K. Liu, Q. Zhao, and B. Krishnamachari. Dynamic multichannel access with imperfect channel state detection. *IEEE Trans. on Signal Processing*, 58(5), May 2010.

[89] Keqin Liu and Qing Zhao. Distributed learning in cognitive radio networks: Multi-armed bandit with distributed multiple players. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 3010–3013. IEEE, 2010.

[90] X. Liu, B. Krishnamachari, and H. Liu. Channel selection in multi-channel opportunistic spectrum access networks with perfect sensing. In *Proceedings of IEEE DySPAN*, 2010.

[91] Y. Ma, D. I. Kim, and Z. Wu. Optimization of OFDMA-based cellular cognitive radio networks. *IEEE Transactions on Communications*, 58(8), 2010.

[92] Megha Maiya and Bechir Hamdaoui. An improved ieee 802.11 mac protocol for wireless ad-hoc networks with multi-channel access capabilities. In *High Performance Computing and Simulation (HPCS), 2011 International Conference on*, pages 162–168. IEEE, 2011.

[93] Megha Maiya and Bechir Hamdaoui. imac: improved medium access control for multi-channel multi-hop wireless networks. *Wireless Communications and Mobile Computing*, 13(11):1060–1071, 2013.

[94] M. McHenry. Reports on spectrum occupancy measurements, shared spectrum company. In *www.sharedspectrum.com/?section=nsf_summary*.

[95] M. McHenry and D. McCloskey. New York city spectrum occupancy measurements. *Shared Spectrum Conf.*, Sept. 2004.

[96] N. F. Mir. *Computer Communication Networks*. Prentice Hall.

[97] Jelena Misic and Vojislav B Misic. Performance of cooperative sensing at the mac level: Error minimization through differential sensing. *Vehicular Technology, IEEE Transactions on*, 58(5):2457–2470, 2009.

[98] Joseph Mitola III and Gerald Q Maguire Jr. Cognitive radio: making software radios more personal. *Personal Communications, IEEE*, 6(4):13–18, 1999.

[99] Asis Nasipuri and Samir R Das. Multichannel csma with signal power-based channel selection for multihop wireless networks. In *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, volume 1, pages 211–218. IEEE, 2000.

[100] T. R. Newman, S. M. S. Hasan, D. Depoy, T. Bose, and J. H. Reed. Designing and deploying a building-wide cognitive radio network testbed. *IEEE Communications Magazine*, 48(9), 2010.

[101] KE Nolan, PD Sutton, LE Doyle, TW Rondeau, B Le, and CW Bostian. Dynamic spectrum access and coexistence experiences involving two independently developed cognitive radio testbeds. In *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pages 270–275. IEEE, 2007.

[102] Mohammad Javad Norooz Oliaee, Bechir Hamdaoui, and Mohsen Guizani. Adaptive service function for system reward maximization under elastic traffic model. In *Global Communications Conference (GLOBECOM), 2013 IEEE*, pages 4781–4785. IEEE, 2013.

[103] M NoroozOliaee and Bechir Hamdaoui. Distributed resource and service management for large-scale dynamic spectrum access systems through coordinated learning. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*, pages 522–527. IEEE, 2011.

[104] M NoroozOliaee, Bechir Hamdaoui, Xiuzhen Cheng, Taieb Znati, and Mohsen Guizani. Analyzing cognitive network access efficiency under limited spectrum handoff agility. *Vehicular Technology, IEEE Transactions on*, 63(3):1402–1407, 2014.

[105] M NoroozOliaee, Bechir Hamdaoui, and Kagan Tumer. Achieving optimal elastic traffic rewards in dynamic multichannel access. In *High Performance Computing and Simulation (HPCS), 2011 International Conference on*, pages 155–161. IEEE, 2011.

[106] MohammadJavad NoroozOliaee, Bechir Hamdaoui, and Mohsen Guizani. Maximizing secondary-user satisfaction in large-scale dsa systems through distributed team cooperation. *Wireless Communications, IEEE Transactions on*, 11(10):3588–3597, 2012.

[107] MohammadJavad NoroozOliaee, Bechir Hamdaoui, and Kagan Tumer. Efficient objective functions for coordinated learning in large-scale distributed osa systems. *Mobile Computing, IEEE Transactions on*, 12(5):931–944, 2013.

[108] MohammadJavad NoroozOliaee, Bechir Hamdaoui, Taieb Znati, and Mohsen Guizani. Forced spectrum access termination probability analysis under restricted channel handoff. In *Wireless Algorithms, Systems, and Applications*, pages 358–365. Springer Berlin Heidelberg, 2012.

[109] T. J. O'Shea, T. C. Clancy, and H. J. Ebeid. Practical signal detection and classification in gnu radio. In *Proceedings of the SDR 07 Technical Conference and Product Exposition*, 2007.

[110] A. Ozyagci and J. Zander. Distributed dynamic spectrum access in multichannel random access networks with selfish users. In *Proceedings of International Wireless Communications and Networking Conference*, 2010.

[111] Jeffrey D Poston and William D Horne. Discontiguous ofdm considerations for dynamic spectrum access in idle tv channels. In *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pages 607–610. IEEE, 2005.

[112] Z. Quan, S. Cui, and A. H. Sayed. Optimal linear cooperation for spectrum sensing in cognitive radio networks. *IEEE Journal of Selected Topics in Signal Processing*, Februray 2008.

[113] J. Riihijarvi and P. Mohonen. Exploiting spatial statistics of primary and secondary users towards improved cognitive radio networks. In *Proceedings of International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, 2008.

[114] Haythem A Bany Salameh, Marwan Krunz, and Ossama Younis. Cooperative adaptive spectrum sharing in cognitive radio networks. *IEEE/ACM Transactions on Networking (TON)*, 18(4):1181–1194, 2010.

[115] Haythem Ahmad Bany Salameh, Marwan Krunz, and David Manzi. Spectrum bonding and aggregation with guard-band awareness in cognitive radio networks. *Mobile Computing, IEEE Transactions on*, 13(3):569–581, 2014.

[116] Akhil Sivanantha, Bechir Hamdaoui, Mohsen Guizani, Xiuzhen Cheng, and Taieb Znati. EM-MAC: an energy-aware multi-channel MAC protocol for multi-hop wireless networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2012 8th International*, pages 1159–1164. IEEE, 2012.

[117] Yi Song and Jiang Xie. Performance analysis of spectrum handoff for cognitive radio ad hoc networks without common control channel under homogeneous primary traffic. *INFOCOM*, 2011.

[118] H. Su and X. Zhang. Cross-layer based oppotrunitic MAC protocols for QoS provisionings over cognitive radio wireless networks. *IEEE Journal on Selected Areas in Communications*, January 2008.

[119] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[120] M. Alkaee Teleghan and B. Hamdaoui. Efficiency-revenue optimality tradeoffs in dynamic spectrum allocation. In *Proc. of IEEE GLOBECOM*, 2010.

[121] G. Tesauro. Practical issues in temporal difference learning. *MLJ*, 8:257–277, 1992.

[122] M. Timmers, S. Pollin, A. Dejonghe, L. Van der Perre, and F. Catthoor. A distributed multichannel MAC protocol for multihop cognitive radio networks. *IEEE Transactions on Vehicular Technology*, 59(1), 2010.

[123] Andrea M Tonello, Nicola Laurenti, and Silvano Pupolin. Analysis of the uplink of an asynchronous multi-user dmt ofdma system impaired by time offsets, frequency offsets, and multi-path fading. In *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, volume 3, pages 1094–1099. IEEE, 2000.

[124] Imene Trigui, Mohamed Siala, and Hatem Boujemâa. Optimized pulse shaping for ofdm multi-user communications over doubly dispersive channels. In *Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on*, pages 1–4. IEEE, 2007.

[125] K. Tumer and A. Agogino. Distributed agent-based air traffic flow management. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 330–337, Honolulu, HI, May 2007.

[126] Vamsi Krishna Tumuluru, Ping Wang, Dusit Niyato, and Wei Song. Performance analysis of cognitive radio spectrum access with prioritized traffic. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, 2012.

[127] J. Unnikrishnan and V. V. Veeravalli. Cooperative sensing for primary detection in cognitive radio. *IEEE Journal of Selected Topics in Signal Processing*, 2(1), 2008.

[128] J. Unnikrishnan and V. V. Veeravalli. Dynamic spectrum access with learning for cognitive radio. In *Proc. of Asilomar Conference on Signals Systems and Computers*, Oct. 2008.

[129] J. Unnikrishnan and V. V. Veeravalli. Algorithms for dynamic spectrum access with learning for cognitive radio. *IEEE Transactions on Signal Processing*, 58(2), August 2010.

[130] Pavithra Venkatraman and Bechir Hamdaoui. Cooperative q-learning for multiple secondary users in dynamic spectrum access. In *2011 7th International Wireless Communications and Mobile Computing Conference*. 2011.

[131] Pavithra Venkatraman, Bechir Hamdaoui, and Mohsen Guizani. Opportunistic bandwidth sharing through reinforcement learning. *Vehicular Technology, IEEE Transactions on*, 59(6):3148–3153, 2010.

[132] Lianfeng Shen Xiaorong Zhu and Tak-Shing Peter Yum. Analysis of cognitive radio spectrum access with optimal channel reservation. *IEEE COMMUNICATIONS LETTERS*, 11(4):304–306, April 2007.

[133] H. Xu, J. Jin, and B. Li. A secondary market for spectrum. In *Proceedings of IEEE INFOCOM*, 2010.

[134] D. Xue, E. Ekici, and X. Wang. Opportunistic periodic mac protocol for cognitive radio networks. In *Proceedings of IEEE Globecom*, 2010.

[135] S. Mangold Y. Xing, R. Chanddramouli and S. Shankar. Dynamic spectrum access in open spectrum wireless networks. *IEEE J. Sel. Areas Commun.*, 24:626–636, March 2006.

[136] Lei Yang, Ben Y Zhao, and Haitao Zheng. The spaces between us: Setting and maintaining boundaries in wireless spectrum access. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, pages 37–48. ACM, 2010.

[137] Yong Yao, Said Rutabayiro Ngoga, David Erman, and Adrian Popescu. Performance of cognitive radio spectrum access with intra- and inter-handoff. *IEEE ICC*, 2012.

[138] S. Yarkan and H. Arslan. Exploiting location awareness toward improved wireless system design in cognitive radio. *IEEE Communications Magazine*, January 2008.

[139] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler. Opportunistic spectrum access via periodic channel sensing. *IEEE Transactions on Signal Processing*, February 2008.

[140] C. Zou and C. Chigan. On game theoretic DSA-driven MAC for cognitive radio networks. *Computer Communications*, 32(18), 2009.