Universidad
Politécnica
de Cartagena

Coexistence Policies in Cognitive Radio

Ph.D. dissertation (draft)

Supervised by:
Juan José Alcaraz Espín
`juan.alcaraz@upct.es`

Department of Information and
Communication Techonlogies

Mario López-Martínez
`mario.lopez@upct.es`

January 2015

Abstract

Public regulatory agencies have traditionally granted access to radio-electric spectrum through fixed, long term licenses for large geographical areas. The increasing demand of wireless communication has led to an almost fully assigned spectrum. However, it is sparsely and unevenly used. A dramatic improvement on spectrum usage efficiency is possible through the operation of cognitive radios, capable of gathering information about their surrounding spectrum environment and adapting their transmission parameters accordingly. Cognitive radios allow unlicensed users to dynamically exploit unused fragments of spectrum. Our main concern in this work is protecting licensed users from interferences caused by unlicensed users' transmissions, and creating incentives for the licensed networks to implement mechanisms easing the coexistence with unlicensed users. Another important guideline of our work is the focus on low-complexity algorithms that can operate at the small time-scale of the spectrum opportunities. This thesis emphasizes the use of mathematical tools such as stochastic modeling, dynamic programming, reinforcement learning, among others, in the development of mechanisms for dynamic spectrum access. The first part of this thesis focuses on Opportunistic Spectrum Access (OSA), proposing schemes that simultaneously increase spectrum efficiency and improve the protection of primary users. The developed mechanisms do not require direct signaling between the primary and the secondary devices. In the second part, we focus on incentives, by developing two frameworks for spectrum trading, in which the licensed users can lease unused bandwidth to unlicensed ones, in exchange of money and relay services, respectively. We also include an extensive survey on the topic, unique in its scope.

Contents

Introduction

1.1 Why Coexistence Policies?

Wireless communications need, among other resources, radio-electric spectrum, which is finite. Government agencies, such as the FCC in the United States or the Electronic Communications Committee (ECC) through national administrations in Europe, grant access to the spectrum through fixed, long-term licenses on large geographical areas, in what is called the "command-and-control" management scheme. Licenses were given to operators in a first-come-first-served basis initially. As the demand of spectrum grew, the procedure changed to national lotteries and comparative hearings. Nowadays, licenses are granted through spectrum auctions generating large revenues to the public administration. Nevertheless, the ever-growing demand of spectrum due to the popularization of mobile services has left these traditional policies obsolete since almost all the spectrum has already been assigned.

However, several reports have been pointing out that most of the licensed spectrum show very little use. In 2002, the FCC Spectrum Policy Task Force highlighted that the typical channel occupancy of a public safety community, for example, could be less than 15% [1]. The European Commission, in 2007, also stated that "some users hold large amounts of valuable spectrum that they do not use to its full capacity" [2].

From early on, experts suggested it would be necessary to switch to a more flexible management to improve spectrum efficiency [3, 4]. In the US, the FCC started allowing license holders to lease their licenses under different constraints in 2003 [5], whereas in Europe is taking some more time: by 2011, only 9 out of 22 european countries surveyed by the ECC allowed spectrum leasing [6].

Cognitive radios [7] could elasticize spectrum management even further by enabling the automatic exploitation of unused spectrum opportunities anywhere in time, space, frequency, power and/or in CDMA codes. A *cognitive radio* has two main characteristics [8]: cognitive capability and reconfigurability. Cognitive capability refers to the ability of the radio to become aware of its surrounding radio

environment. Reconfigurability refers to the device's ability of changing its trans-mission parameters dynamically.

In the most typical scenario, secondary users (SUs, unlicensed), equipped with cognitive radios, would look for unused fragments of spectrum in licensed bands and would use them for their transmissions subject to interference constraints for the protection of primary users (PUs, licensed). That is, PUs and SUs will coexist in some areas. This situation corresponds to the "Hierarchical Access Model", also known as *Opportunistic Spectrum Access (OSA)*, one of the possible coexistence approaches under *Dynamic Spectrum Access (DSA)* [9]. Along the lines of cur-rent regulation models is the Dynamic Exclusive Use Model, also called *spectrum trading*, which considers leasing or selling licenses but at a smaller time-scale, even in a real-time basis and with a free choice of technology. Lastly, the "Spectrum Commons Model" assumes open sharing with no categories of users.

DSA has recently gained popularity because of the TV digitalization, in order to take advantage of the spectrum opportunities it has generated. Therefore, the last few years have seen some DSA standards such as IEEE 802.22 [9], ECMA-392 [10], 802.11af [12], 802.19.1 [11] and IEEE SCC41 [13].

No matter which model is considered, DSA is still annoying licensed operators at present, because it implies additional costs to them such as changes in their infrastructures, QoS impairments due to interferences, or profit reduction due to increased competition, while they have already paid for their licenses. Protecting and creating incentives for PUs instead of forcing them avoids that actual spectrum owners, as well as prospective ones, feel discouraged from investing in new spectrum technologies and services: "why am I going to spend money in spectrum if the government is going to give it for free to others?".

It is needed that these operators believe in an automated mechanism that con-trols their profits. But this idea alone is scary: "How big is the risk of experiencing a great economic loss?". Especially if they take into account incidents like the one experienced in 2012 by the Knight Capital Group in the related field of High Frequency Trading of stocks [10] where the firm's algorithms caused a loss of $440 million in 30 minutes (four times its net income from 2011).

It is hard to convince the telecommunications industry to adopt automated dynamic spectrum management mechanisms if there is no strong evidence that these algorithms are safe against wrong decisions (due to disadaptation to real time, for example). We believe that optimality, then, should be considered as a secondary objective.

1.2   Motivation and Objective

As discussed in the previous section, while DSA has been widely accepted as the most promising approach to address the spectrum inefficiency problem, there are still some concerns that hinder its fully adoption by the telecom industry [11].

Spectrum holders and regulators alike demand a higher level of protection for PU communications, and better incentives to accept coexistence with other users

in their spectrum. In consequence, our[1] *motivation* in this thesis is, *to develop dynamic spectrum access mechanisms focusing on the protection and creation of incentives to the PUs.* We directed our attention to the Opportunistic Spectrum Access framework, as it is more conservative regarding PU's protection compared to other DSA approaches (such as the spectrum overlay and underlay models, consisting of concurrent transmissions of PUs and SUs), and to spectrum trading in our search for incentives. Improvements in the Spectrum Commons Model are out of the scope of this thesis, since it removes property rights over spectrum and its challenges are significantly different.[2]

More specifically, we set the following sub-objectives regarding unresolved issues in OSA:

- Improving existing MAC protocols for hardware-constrained radios, specifically by considering the effect of the scanning delay on the information uncertainty (addressed in chapters 2 and 4).

- Allowing the secondary devices to better assess the impact of their activity on primary users' performance. The previous literature relied on OSA protocols design to keep PU performance above certain bounds that are heterogeneous in nature and difficult to assess in practice (chapters 2 and 5).

- Exploring the options available to the primary operators for easing the coexistence with secondary devices (chapters 3 and 4).

- Proposing an application of OSA with practical interest to the telecommunications industry, and therefore motivating its adoption (chapter 5).

With respect to the open challenges in spectrum trading, we pursued the following sub-goals:

- Providing an unified, structured and critic view of the existing solutions on this field, its specific challenges and open issues. An study with such a broad scope had not been done before (chapter 6).

- Balancing PUs' QoS with the economic revenue obtained by selling spectrum opportunities to SUs (chapter 7).

- Exploring mechanisms not based on monetary values. Most previous works are focused on economic transactions. However, other types of trades could add non-substitutable value such as trading spectrum opportunities in exchange for relay services (chapter 8).

We also have two common sub-objectives for OSA and spectrum trading:

---

[1]I will be using the first person plural by default throughout my thesis. Most of the ideas and work is a team effort involving my supervisor and other authors. This should not be considered as something that works against showing my autonomy, but even my personal decisions where strongly inspired by others' opinions. Meaningful research is not done (or should not be done) by a single person in a vacuum [12]. Thus, I consider artificial the use of singular forms here. In other parts of this thesis, such first person plural also includes the reader, primarily when describing mathematical derivations, as a natural way of walking him through the process.

[2]There has been a long ongoing debate about whether such a drastic change of paradigm is convenient [13], given the success witnessed in unlicensed bands (e.g. in the ISM band).

- Dealing with imperfect information scenarios. That is, those for which some of the parameters of their stochastic models are not known or are not stationary (chapters 5 and 8).

- Keeping complexity low. The main tendency of previous (and contemporaneous) solutions in the area has been to propose increasingly complex models. Nonetheless, the resources, especially time, required to reach such elaborated solutions have not been considered enough.

For that reason, it is more practical to design models that may not fully exploit spectrum in a particular moment (for example, by not taking into account spectrum geographical reuse) but that are capable of reducing the uncertainty due to the changes over time, and thus, providing higher guarantees on the agents' benefit.

1.3   Contributions and Structure

Our contribution can be divided into two main blocks, coincident with the structure of the thesis. Part I features our contributions in the Opportunistic Spectrum Access framework. This part is focused on protecting the PUs from interference caused by SUs access to spectrum, while trying to increase the spectrum usage efficiency. The most common approach in previous literature is to study the sensing and access from the SUs perspective, that is, focusing exclusively on the design of SU access protocols (independently if the works are addressing the impact on PUs or on SUs). The original concept of cognitive radios, especially under an OSA approach, was to use such cognitive capabilities to make SUs access transparent (or as transparent as possible) to the PUs. However, we introduce the novel approach of allowing the primary network to adapt its resource allocation strategy in order to better coexist with cognitive devices.

In particular, the contributions described in every chapter and publications are as follows.

Chapter 2: Background Detection of Primary User Activity in Opportunistic Spectrum Access [14]. We propose a collision detection mechanism for the SU, simultaneous to packet reception. Our proposal exploits local characterization of the PU activity, power levels and packet errors at each time slot. Based on these data, we develop a Maximum *A Posteriori* (MAP) estimator as the basis for the collision detection system. We describe the general methodology to build this estimator, and study the system's robustness against parameter misestimations. The proposed scheme can be incorporated to existing OSA protocols. Our results show a notable throughput improvement for the SUs while assuring a lower collision probability with PU transmissions in most cases, even under severe estimation inaccuracies.

Chapter 3: Bandwidth Reservation as a Coexistence Strategy in Opportunistic Spectrum Access Environments [15]. We study a bandwidth reservation (BR) scheme by which the primary network (PN) keeps a set of adjacent channels free of PU transmissions. These reserved channels only accommodate PU traffic when all the non-reserved channels are used, and the SUs only occupy available

channels within the reserved spectrum. From the point of view of the PN, BR entails a tradeoff between the benefits of an improved coexistence with SUs, and the capacity reduction associated to having fewer options for PU channel allocation. The SUs are characterized by a hardware limited radio, imperfect spectrum sensing, bayesian estimation of PU activity and multichannel access. By means of a Markov-reward model, we compute the expected PU capacity with and without BR, considering propagation effects, interference and random locations. The results show that, in a non-congested PN with SU activity, the interference reduction capability of BR increases the overall capacity of the PN compared to not using BR.

Chapter 4: Uncertainty-Aware Opportunistic Spectrum Access in Coexistence - Friendly Systems [16]. This chapter comprises two main contributions: first we develop an optimal SU access strategy that considers the effect of delay in the uncertainty of the information available at the SUs. This effect has been largely ignored in the design of OSA protocols for hardware limited radios. Second, this OSA scheme is designed for coexistence-oriented scenarios (such as the bandwidth reservation of previous chapter), and evaluated under these conditions. The formulation combines SU throughput and PU performance degradation, characterized by the overlapping time with SU activity. As most OSA protocols, our approach requires the SU to characterize the traffic profile of the PUs. The results of this chapter also illustrate the robustness of our approach to misestimation of PU traffic intensity.

Chapter 5. Response Surface Methodology for Efficient Spectrum Reuse in Cellular Networks [17, 18]. This chapter addresses one promising application of the OSA strategy: spectrum reuse of legacy access networks. As cellular network technology evolves, the operators deploy new generation networks while maintaining their legacy networks, since not all users upgrade their terminals at the same pace. Therefore, the spectrum associated to these legacy networks becomes gradually underused. By means of cognitive radio techniques, the operator can allow its new generation terminals to reuse this spectrum without any modification (and therefore any investment) on their legacy infrastructure. We propose a semi-decentralized scheme in which a cellular operator guides the secondary access by broadcasting some operational parameters of the access strategy. The mechanism dynamically learns the optimal parameters by means of a response surface methodology (RSM), implying a very small signalling overhead. Our results show a notable capacity improvement compared to the classical approaches of exploiting spatial or temporal opportunities.

Part II of the thesis contains our contributions in the field of automated spectrum trading. It is especially in this part where our low-complexity-intention has a greater significance, since the mainstream trend mostly neglects the importance of real-time adaptation, as chapter 6 shows. Our contributions in this topic can be summarized as follows:

Chapter 6. Automated spectrum trading mechanisms: understanding the big picture [19]. This work presents a general view of automated spectrum trading.

Our work is unique in its scope, focusing on automated spectrum trading as the most promising mechanism to solve spectrum scarcity. We identify the specific aspects that make spectrum different from conventional goods and the impact that these features have on automatic trading. We discuss past, current approaches and future research lines, highlighting their advantages and disadvantages. Each subtopic is presented and developed explicitly with references to previous works where each aspect has been considered or addressed. Finally, as a conclusion of the surveying effort, we highlight overlooked issues of automated spectrum trading, with a special focus on practical implementation.

Chapter 7 An MDP Framework for Centralized Dynamic Spectrum Leasing [20]. This work addresses the design of centralized DSA MAC protocols comprising dynamic spectrum auction. We explore the possibilities of a formal design based on a Markov decision process (MDP) formulation. We propose a design framework to balance the grade-of-service (given by the blocking probability for PUs) and the expected economic revenue provided by the SUs' accepted bids. This trade-off can be managed in two ways. One consists of computing a single objective value given by a combination of the blocking probability and the expected revenue. The weights assigned to each objective determine the point in the Pareto front where the obtained policy lies. The other approach, consists of solving for one of the objectives and setting constraints on the remaining ones. This chapter focuses on centralized trading in which we explore the use of MDP and CMDP formulations to balance benefit and grade of service for PUs. The main advantages of this approach is that it assures operating at global optimum and reduces the computational effort at SUs.

Chapter 8. A superprocess with Upper Confidence Bounds for Cooperative Spectrum Sharing [21, 22]. In this last chapter we focus on Cooperative Spectrum Sharing (CSS). CSS is is an appealing approach for primary networks to share spectrum with SUs because it increases the transmission range or rate of the PUs. Most previous works are focused on developing complex algorithms which may not be fast enough for real-time variations such as channel availability and/or assume perfect information about the network. Instead, we develop a learning mechanism for a primary network to enable CSS in a strongly incomplete information scenario with low computational overhead. Our mechanism is based on a Markovian variant of multi-armed bandits (MABs) called superprocess, enhanced with the concept of Upper Confidence Bound (UCB) from stochastic MABs. By means of Monte-Carlo simulations we show that, despite its low computational overhead, it converges to the optimal solution and outperforms baseline approaches such as ϵ-greedy. This algorithm can be extended to include more sophisticated features while maintaining its desirable properties such as low computational overhead and fast speed of convergence.

1.4   Methodology

This thesis is part of the project "ChAracterization, EvaLuation, Planning and IMprovement of Key Technologies for the Future Internet: Knowledge and Transfer

Chapters	Tools used by proposed mechanisms	Numerical Evaluation
2	Statistical inference, MRM, discrete optimization	MRM, Monte-Carlo sim.
3	MRM	Stochastic model, MRM, Monte-Carlo sim.
4	POMDP	MRM
5	Stochastic optimization (RSM)	Monte-Carlo sim.
7	CMDP	MRM
8	Reinforcement learning (MABs), MDP	Monte-Carlo sim.

Table 1.1: Relation of mathematical tools and thesis chapters

(CALM)", funded by the Spanish Ministry of Economy and Competitiveness. The CALM project is organized as a set of objectives interrelating key technologies for the Future Internet, *with a special focus on mathematical tools* for the development, analysis and improvement of networks and protocols.

Most of our proposed coexistence mechanisms and policies in this thesis are grounded in Markov Decision Processes (MDPs) [23]. The suitability of this tool for our scenario is straightforward to see: Markov models "are applicable to any dynamic system whose evolution over time involves uncertainty" [24], as is the case of spectrum occupancy over time. The constrained-MDP (CMDP) formulated in chapter 7, and the Partially Observable MDP (POMDP) in chapter 4, are especially remarkable because of their far-from-common presence in previous DSA works.

Trying to characterize incomplete information scenarios, by taking into account all the possible values of the (often continuous) unknown parameters, is computationally infeasible with MDPs. This is known as "the curse of dimensionality" [25]. In the chapters devoted to these scenarios, we resort to reinforcement learning and stochastic optimization approaches. Some of the tools used in this thesis like Response Surface Methodology (RSM) (chapter 5) or Multi-Armed bandits (MABs) (chapter 8) are of new applicability to coexistence issues in cognitive radio and are becoming hot topics, not only in the research community [26] but also in highly profitable businesses like online advertisement [27].

Regarding the analysis and numerical evaluation of our mechanisms, we used Markov Reward Models (MRM). Either accompanying the previous tool or as the main performance evaluation resource, we resorted to the simple but powerful concept of Monte-Carlo simulation.

We rely on well-known stochastic models of the different real world parameters of interest, such as the Poisson traffic model, the log-normal distribution to describe shadow fading, etc. Such models have been widely used and validated in previous literature [28]

Because of our focus in algorithms and the link layer of cognitive networks, a characterization involving a full-protocol stack or a deep modeling of physical effects is not needed. Therefore, regarding the software support, we mainly worked with MATLAB [29], Mathematica [30] for some mathematical derivations, and the OMNeT++ simulator [31] in the particular case of chapter 3 (used for numerical evaluation). We have also employed other development platforms such as Python

with SciPy [32] and R [33], because of their suitability for machine learning appli-
cations and data analysis.

Part I

Opportunistic Spectrum Access

Background Detection of Primary User Activity in Opportunistic Spectrum Access

2.1 Introduction

Motivation

One of the main problems of OSA is that current radio frequency front-ends cannot perform sensing and transmission in the same channels at the same time. In consequence, in OSA protocols, spectrum sensing and SU data transmission are performed in two separated phases. The scanning phase allows the SUs to find free channels to transmit in. However, during the transmission phase, a channel occupied by an SU can be eventually used by a PU transmission causing a potentially harmful interference to the PU receiver. In consequence, in OSA protocols, an SU is only allowed to transmit during a limited amount of time after which it performs a new sensing procedure (*periodic spectrum sensing*). According to a given PU activity model, the length of the SU transmission period is set to the value that maximizes the SU throughput under a given collision probability constraint. The generic OSA model described is referred to as *optimized OSA* in this chapter, and is used as a starting point and as a benchmark for our proposal. We propose a collision detection mechanism that performs PU activity estimation during SU packet reception, and uses it to improve the classic optimized OSA mechanism.

Related Work

Most of the works about collision detection for single-radio transceivers only focus on the detection carried out at the transmitter side, such as periodic spectrum sensing in OSA protocols [34], [35], [36]. To the best of our knowledge, there are only two exceptions to this [37, 38]. Paper [37] theoretically studies the impact of different receiver-side detectors, but it does not detail their implementations. The

work in [38] implements an energy-detector [39, Ch. 4] at the SU Rx. However, as opposed to ours, they assume that the secondary transmission is always correctly decoded and can be substracted from the received signal.

We adopt a worst case PU QoS criteria used in previous works [36], [40], [41], by which any collision is considered to cause harmful interference to PU receivers. As in [36], our model also includes the effect of transmission overlap in SU performance. Additionally, we use the statistical description of the fading in the direct and interfering links to evaluate this performance.

Perfect collision detection is implicitly assumed in [42, 43]. In this work, we relax that assumption, taking into consideration the multiple factors that hinder perfect detection: path loss, fading effects, and decoding probability as a function of the modulation used and the signal-to-interference-and-noise ratio. To overcome these difficulties and perform an optimal collision detection, the SU receiver can make use of diverse estimations performed during the successive spectrum scanning and sensing periods. As an example, an accurate PU traffic estimator is described in [44].

Our Contribution

We refer to our proposed mechanism as *background detection* (BD). In contrast to classic spectrum sensing, in which licensed channels are sensed while the SU transmitter is idle, the BD mechanism operates at the SU receiver during ongoing SU transmissions. In consequence, BD uses information available at the SU receiver during packet reception to decide, by means of a Maximum *A Posteriori* (MAP) estimator, if there has been an overlap in each transmission slot. The MAP estimator exploits existing knowledge at the SU, such as the average power of the received signals, the presence or absence of decoding errors, the PU traffic pattern, and the statistical description of the fading processes. Our mechanism is used simultaneously with optimized OSA, and we refer to the combined scheme as OSA-BD. We also evaluate the robustness of the system under inaccuracies in these data. In particular, it is shown that, while optimized OSA is very reliant on an accurate PU traffic characterization, OSA-BD achieves higher robustness against PU traffic misestimations.

The rest of this chapter is organized as follows. Section 2.2 describes the PU and SU networks and states the technical foundations of our proposal. Section 2.3 discloses the background detection mechanism, formulating the MAP rule and then computing the detection probabilities. We describe the incorporation of our idea to optimized OSA in Section 2.4, and Section 2.5, contains the numerical evaluation results and comparison to benchmack schemes. Concluding remarks are written in Section 2.6.

Figure 2.1: SU pair communicating over a downlink channel of a PU BS. The SU Tx is causing interference to two PUs using the same channel. The SU Rx, equipped with our BD mechanism, is capable of detecting the signal from the PU BS while receiving the SU Tx's packets and thus, sends the SU Tx a signal to abort transmission.

2.2   System Description

General Considerations

The mechanism proposed in this paper is very generic, and is applicable whenever the primary and the secondary wireless networks present some general characteristics described in this section. The spectrum of the primary network is assumed to be divided into orthogonal channels. The SUs can scan these channels to detect spectrum opportunities in them, as well as monitor PU activity and estimate the PU traffic pattern. One type of PU access networks that fits well on this description is cellular networks, but other types may be feasible for our proposal as well.

As an example, let us consider the system in Fig. 2.1, representing the coverage area (cell) of a PU base station (PU BS), where a cognitive pair, consisting of an SU transmitter (SU Tx) and an SU receiver (SU Rx), tries to perform an SU transmission on a spectrum opportunity within this PU cell. It is assumed that the SU transmits in PU downlink channels. The BD scheme could be used for PU uplink channels as well but it would be less effective, due to possible hidden terminal problems (*i.e.* a PU Tx out of range of the SU Rx detection). Before transmitting, both the SU Tx and the SU Rx scan the licensed spectrum looking for available channels and exchange control messages containing their sensing outcomes over a dedicated control channel. Once the cognitive pair makes the access decision, the SU Tx starts the SU transmission over one or more available channels.

The SU transmission is divided into time-slots. At the end of each time-slot, the SU Rx submits a control message to the SU Tx over the control channel. In general, this message would contain an ACK (or NACK) of the data transmitted on the time-slot.

It is assumed that the signal power received at the SU Rx remains constant during a symbol time T_s of the SU Tx transmission. In this case, the probability of correct signal detection at the SU Rx is defined in terms of the probability

that the signal to interference ratio (SINR) is above a given threshold, which is determined by the minimum acceptable bit error rate (BER) for correct decoding, the transmission rate and the modulation used. Let us consider that, given the SU data packet length, and the information and redundancy bits in it, SU packets are correctly received if the BER during the packet reception is above certain value (e.g. 10^{-3}). This BER objective is attained when the ratio E_b/N_0 is above a threshold γ_b, where E_b is the energy per bit and N_0 is the noise spectral density (including both thermal and interference noise). The threshold depends on the modulation used (e.g. for BPSK, BER $= 10^{-3}$ for $\gamma_b = 7$ dB). Therefore, the threshold in terms of SINR is given by $\gamma = \gamma_b/(BT_b)$ where B is the channel bandwidth and T_b the bit duration.

Principles of the Background Detection Scheme

Each SU data packet may be received at the SU Rx with or without PU interference. Let P denote the signal power at the SU Rx during the reception of an SU data packet. The SU signal is received with power Y at the SU Rx and, in case of transmission overlap, the signal transmitted by the PU BS is received with power X at the SU Rx. Let N denote the average white noise power in the channel bandwidth at the SU Rx during the packet reception. Note that P, X, and Y are random variables, while N is a constant determined by the noise factor of the receiver, which is a known design parameter. Therefore, we can subtract N from the power observation. Then, under PU-SU transmission overlap, $P = X + Y$ and, in absence of overlap, $P = Y$. Similarly, $\texttt{SINR} = \frac{Y}{N+X}$, with overlap, and $\texttt{SINR} = \frac{Y}{N}$ without overlap. If $\texttt{SINR} > \gamma$ the packet is correctly received and, if $\texttt{SINR} \leq \gamma$ the packet is impossible to decode (error event).

After the reception of an SU data packet, the SU Rx can obtain two observations: P, and the occurrence (or absence) of an error event. The idea of the background detection (BD) mechanism is to use these observations to allow the SUs estimate the presence of PU activity during SU transmissions. If PU activity is detected, the OSA algorithm may decide to end the ongoing SU transmission to protect from interference any PU terminal in range of the SU Tx (in Fig. 2.1, the PU terminals pointed by a dashed arrow) in case any of them is receiving PU BS transmissions on the channel where the overlap was detected.

The BD scheme is essentially a binary hypothesis testing mechanism conducted during SU transmission to decide on the most feasible value of a random variable Θ, which takes two possible values $\{0, 1\}$ at each SU packet reception. The event $\Theta = 1$ indicates that there has been PU interference during the packet reception, and $\Theta = 0$ indicates that the reception was free of PU interference. The probability density functions (PDF) of Y and X are $f_Y(y)$ and $f_X(x)$ respectively. In case that both the SU Tx and the PU BS signals experience Rayleigh fading over their respective links to the SU Rx, we have $f_Y(y) = \frac{1}{p_{ss}} e^{-\frac{y}{p_{ss}}}$, and $f_X(x) = \frac{1}{p_{ps}} e^{-\frac{x}{p_{ps}}}$, where p_{ss} is the average power received from the SU Tx and p_{ps} is the average power received from the PU BS. Both distributions, as well as the average powers can be obtained from the previous sensing history of the SUs. Similarly, this sensing history provides

information about the PU traffic profile and intensity on the scanned channels. One basic metric is the probability of PU activity during the transmission time of an SU packet ($\mathbf{P}\left(\Theta = \theta\right)$), with $\theta \in \{0, 1\}$. Therefore, assuming $f_Y(y)$, $f_X(x)$ and $\mathbf{P}\left(\Theta = \theta\right)$ are known, the hypothesis testing is based on Bayesian estimation, resulting in a maximum a posteriori probability (MAP) decision rule. In Section 2.5 we evaluate the sensitivity of the BD mechanism to the inaccuracies in the estimation of the PDFs and traffic parameters.

Summarizing, once the SU transmission has started, the SU Rx obtains, upon each packet reception, the following observation:

$$
(P, E) = \begin{cases} (p_r, 0) & \text{, if } \texttt{SINR} > \gamma \\ (p_r, 1) & \text{, if } \texttt{SINR} \leq \gamma \end{cases}
\tag{2.1}
$$

And, upon this observation, the MAP decision rule selects the hypothesis having the maximum posterior distribution over the two possible values of θ. Next section provides a detailed derivation for the BD MAP rule, and obtains the probabilities of correct and incorrect detection of the MAP estimator.

2.3 PU Activity Detection in Background

MAP Rule

Given the observation $(P = p_r, E = e)$, the MAP rule selects the hypothesis H_i for which the *a posteriori* probability $\mathbf{P}\left(\Theta = \theta | P = p_r, E = e\right)$ is largest. By Bayes' rule this is equivalent to selecting the hypothesis maximizing $p_\Theta\left(\theta\right) f_{P,E|\Theta}\left(p_r, e | \theta\right)$. Because E is a discrete random variable, we have to consider separately each of its possible values (E = 0 if the SU packet is received without errors and E = 1 otherwise) for the computation of the *a posteriori* probability. Let E_0 and E_1 refer to the events E = 0 and E = 1, respectively. Similarly, we use H_0 and H_1 to refer to the events $\Theta = 0$ and $\Theta = 1$. Let us derive the expression of $f_{P,E|\Theta}\left(p_r, e | \theta\right)$ for each possible combination of E_0 and E_1 with H_0 and H_1.

Case E_1, H_1

Applying the definitions we have

$$
\begin{aligned}
f_{P,E_1|H_1}\left(p_r, E_1 | H_1\right) &= \frac{\partial}{\partial p_r} \mathbf{P}\left(P \leq p_r, \texttt{SINR} \leq \gamma | H_1\right) \\
&= \frac{\partial}{\partial p_r} \mathbf{P}\left(X + Y \leq p_r, Y \leq N\gamma + X\gamma\right)
\end{aligned}
\tag{2.2}
$$

We will first find the joint PDF of P and X and then integrate to find the PDF of P.

$$
\begin{aligned}
\mathbf{P}\left(X+Y \le p_r, Y \le N\gamma + X\gamma | X = x\right) &= \mathbf{P}\left(Y \le p_r - x, Y \le N\gamma + x\gamma | X = x\right) \\
&= \mathbf{P}\left(Y \le \min\left(p_r - x, N\gamma + x\gamma\right) | X = x\right) \\
&= \mathbf{P}\left(Y \le \min\left(p_r - x, N\gamma + x\gamma\right)\right) \\
&= \int_{-\infty}^{\min(p_r-x,N\gamma+x\gamma)} f_Y(y)\,dy \\
&= \begin{cases} \int_{-\infty}^{p_r-x} f_Y(y)\,dy & \text{, if } p_r \le N\gamma + x\left(\gamma + 1\right) \\ \int_{-\infty}^{N\gamma+x\gamma} f_Y(y)\,dy & \text{, otherwise} \end{cases}
\end{aligned}
$$
(2.3)

where the third equality comes from the independence of the PU and SU signals, X and Y. By differentiating both sides with respect to p_r we obtain

$$
f_{P,E_1|H_1,X}\left(p_r, E_1 | H_1, x\right) = \begin{cases} f_Y\left(p_r - x\right) & \text{, if } p_r \le N\gamma + x\left(\gamma + 1\right) \\ 0 & \text{, otherwise} \end{cases}
$$
(2.4)

Then, the joint PDF of P and X is

$$
\begin{aligned}
f_{P,X,E_1|\Theta}\left(p_r, x, E_1 | H_1\right) &= f_{X|\Theta}\left(x|H_1\right) f_{P,E_1|\Theta}\left(p_r, E_1|H_1\right) \\
&= f_X\left(x\right) f_{P,E_1|\Theta}\left(p_r, E_1|H_1\right) \\
&= \begin{cases} f_X\left(x\right) f_Y\left(p_r - x\right) & \text{, if } p_r \le N\gamma + x\left(\gamma + 1\right) \\ 0 & \text{, otherwise} \end{cases}
\end{aligned}
$$
(2.5)

from which we finally obtain

$$
\begin{aligned}
f_{P,E_1|H_1}\left(p_r\right) &= \int_{-\infty}^{\infty} f_{P,X,E_1|\Theta}\left(p_r, x, E_1|H_1\right)\,dx \\
&= \int_{\frac{p_r-N\gamma}{(1+\gamma)}}^{\infty} f_X\left(x\right) f_Y\left(p_r - x\right)\,dx
\end{aligned}
$$
(2.6)

where $f_{P,E_1|H_1}\left(p_r\right)$ denotes $f_{P,E_1|H_1}\left(p_r, E_1|H_1\right)$ in a more compact form. When both links are characterized by Rayleigh fading, $f_X\left(x\right)$ and $f_Y\left(y\right)$ are exponential distributions and we have

$$
f_{P,E_1|H_1}\left(p_r\right) = \begin{cases} \dfrac{e^{-\frac{p_r}{p_{ps}}} - e^{\left(-\frac{p_r}{p_{ss}} + \frac{(p_{ps}-p_{ss})(p_r-N\gamma)}{p_{ps}p_{ss}(1+\gamma)}\right)}}{p_{ps} - p_{ss}} & \text{, if } p_r \ge N\gamma \\[2em] \dfrac{e^{-\frac{p_r}{p_{ps}}} - e^{-\frac{p_r}{p_{ss}}}}{p_{ps} - p_{ss}} & \text{, otherwise} \end{cases}
$$
(2.7)

Case E_0, H_1

Proceeding similarly to the case above we have

$$
\begin{aligned}
f_{P,E_0|H_1}\left(p_r, E_0|H_1\right) &= \frac{\partial}{\partial p_r}\mathbf{P}\left(P \le p_r, \mathtt{SINR} > \gamma | H_1\right) \\
&= \frac{\partial}{\partial p_r}\mathbf{P}\left(X + Y \le p_r, Y > N\gamma + X\gamma\right)
\end{aligned}
$$
(2.8)

To obtain the joint PDF of X and P we need the following probability

$$
\begin{aligned}
\mathbf{P}\left(X+Y\le p_r, Y>N\gamma+X\gamma | X=x\right) &= \mathbf{P}\left(X+Y\le p_r, Y>N\gamma+X\gamma\right) \\
&= \mathbf{P}\left(N\gamma+X\gamma<Y\le p_r-x\right) \\
&= \begin{cases} \displaystyle\int_{N\gamma+X\gamma}^{p_r-x} f_Y(y)\,dy & ,\text{if } x<\frac{p_r-N\gamma}{1+\gamma} \\ 0 & ,\text{otherwise} \end{cases}
\end{aligned}
\tag{2.9}
$$

By differentiating with respect to p_r we obtain

$$
f_{P,E_0|H_1,X}\left(p_r, E_0|H_1, x\right) = \begin{cases} \displaystyle\frac{\partial}{\partial p_r}\int_{N\gamma+X\gamma}^{p_r-x} f_Y(y)\,dy & ,\text{if } x<\frac{p_r-N\gamma}{1+\gamma} \\ 0 & ,\text{otherwise} \end{cases}
\tag{2.10}
$$

Then, after obtaining the joint PDF of X and P, we integrate over x to get

$$
f_{P,E_0|H_1}\left(p_r\right) = \int_{-\infty}^{\frac{p_r-N\gamma}{1+\gamma}} \left(\frac{\partial}{\partial p_r}\int_{N\gamma+X\gamma}^{p_r-x} f_Y(y)\,dy\right) f_X(x)\,dx
\tag{2.11}
$$

which, for Rayleigh fading has the following form

$$
f_{P,E_0|H_1}\left(p_r\right) = \begin{cases} \displaystyle\frac{1}{p_{ps}-p_{ss}}e^{-\frac{p_r}{p_{ss}}}\left(e^{\frac{(p_{ps}-p_{ss})(p_r-N\gamma)}{p_{ps}p_{ss}(1+\gamma)}}-1\right) & ,\text{if } p_r>N\gamma \\ 0 & ,\text{otherwise} \end{cases}
\tag{2.12}
$$

Case E_1, H_0

In this case there is no signal received from the PU and the conditional PDF of P is defined as follows

$$
\begin{aligned}
f_{P,E_1|H_0}\left(p_r, E_1|H_1\right) &= \frac{\partial}{\partial p_r}\mathbf{P}\left(P\le p_r, \mathtt{SINR}\le\gamma|H_1\right) \\
&= \frac{\partial}{\partial p_r}\mathbf{P}\left(Y\le p_r, Y\le N\gamma\right)
\end{aligned}
\tag{2.13}
$$

The probability in the right hand side is given by

$$
\begin{aligned}
\mathbf{P}\left(Y\le p_r, Y\le N\gamma\right) &= \mathbf{P}\left(Y\le\min\left(p_r, N\gamma\right)\right) \\
&= \begin{cases} \displaystyle\int_{-\infty}^{p_r} f_Y(y)\,dy & ,\text{if } p_r\le N\gamma \\ \displaystyle\int_{-\infty}^{N\gamma} f_Y(y)\,dy & ,\text{otherwise} \end{cases}
\end{aligned}
\tag{2.14}
$$

By differentiating with respect to p_r we obtain

$$
f_{P,E_1|H_0}\left(p_r\right) = \begin{cases} f_Y(p_r) & ,\text{if } p_r\le N\gamma \\ 0 & ,\text{otherwise} \end{cases}
\tag{2.15}
$$

The expression for Rayleigh fading is straightforward.

Case E_0, H_0

The PDF of P is now given by

$$
\begin{aligned}
f_{P,E_0|H_0}\left(p_r, E_0|H_1\right) &= \frac{\partial}{\partial p_r} \mathbf{P}\left(P \le p_r, \mathrm{SINR} > \gamma | H_1\right) \\
&= \frac{\partial}{\partial p_r} \mathbf{P}\left(Y \le p_r, Y > N\gamma\right)
\end{aligned}
\tag{2.16}
$$

where $\mathbf{P}\left(Y \le p_r, Y > N\gamma\right)$ equals $\mathbf{P}\left(N\gamma < Y \le p_r\right)$ when $p_r > N\gamma$ and is 0 otherwise. Therefore we have

$$
f_{P,E_0|H_0}\left(p_r\right) = \begin{cases} \dfrac{\partial}{\partial p_r} \displaystyle\int_{N\gamma}^{p_r} f_Y(y)\, dy & ,\text{if } p_r > N\gamma \\ 0 & ,\text{otherwise} \end{cases}
\tag{2.17}
$$

The particularization for Rayleigh fading is straightforward.

Given the above expressions, the MAP rule is specified by partitioning the observation space into disjoint sets in which each of the two hypothesis is chosen. In case a reception error has occurred (E_1), the following equation provides the threshold value(s) for the received power

$$
\mathbf{P}\left(H_0\right) f_{P,E_1|H_0}\left(p_r\right) = \mathbf{P}\left(H_1\right) f_{P,E_1|H_1}\left(p_r\right)
\tag{2.18}
$$

And in the E_0 event, the corresponding equation is

$$
\mathbf{P}\left(H_0\right) f_{P,E_0|H_0}\left(p_r\right) = \mathbf{P}\left(H_1\right) f_{P,E_0|H_1}\left(p_r\right)
\tag{2.19}
$$

We denote the power threshold for the E_0 event by $p^*_{E_0}$, and the threshold for the E_1 event by $p^*_{E_1}$.

Let us determine the power thresholds for the case of Rayleigh fading. Considering exponential distributions for X and Y in (2.18), we obtain the following two equations for E_1

$$
\mathbf{P}\left(H_0\right) \cdot 0 = \mathbf{P}\left(H_1\right) \frac{e^{-\frac{p_r}{p_{ps}}} - e^{\left(-\frac{p_r}{p_{ss}} + \frac{(p_{ps} - p_{ss})(p_r - N\gamma)}{p_{ps}p_{ss}(1+\gamma)}\right)}}{p_{ps} - p_{ss}} \quad ,\text{if } p_r > N\gamma \tag{2.20}
$$

$$
\mathbf{P}\left(H_0\right) \frac{e^{-\frac{p_r}{p_{ss}}}}{p_{ss}} = \mathbf{P}\left(H_1\right) \frac{e^{-\frac{p_r}{p_{ps}}} - e^{-\frac{p_r}{p_{ss}}}}{p_{ps} - p_{ss}} \quad ,\text{if } p_r \le N\gamma \tag{2.21}
$$

The first equation (2.20) only holds for $p_r = \infty$. Therefore, whenever $p_r > N\gamma$, the MAP rule selects the hypothesis H_1, which is rational because if $p_r > N\gamma$, the hypothesis H_0 implies that $\mathrm{SINR} > \gamma$ which is inconsistent with E_1. In other words, if $p_r > N\gamma$, E_1 is caused by PU interference. It can be checked that the solution of second equation (2.21) is

$$
p^* = \frac{p_{ss}p_{ps}}{p_{ps} - p_{ss}} \log\left(\frac{p_{ps} - p_{ss}}{p_{ss}} \frac{\mathbf{P}\left(H_0\right)}{\mathbf{P}\left(H_1\right)} + 1\right)
\tag{2.22}
$$

If p^* is a real number and $p^* \le N\gamma$, then $p^*_{E_1} = p^*$ is the threshold below which the MAP rule selects the hypothesis H_0 in the event E_1. In case (2.21) has no solution

in $(0, N\gamma]$, then $p_{E_1}^* = N\gamma$ if the left hand side of (2.21) is larger than the right hand side, and $p_{E_1}^* = 0$ otherwise.

Considering Rayleigh fading, the equation for E_0, (2.19) results in the following two equations

$$\mathbf{P}(H_0) \frac{e^{-\frac{p_r}{p_{ss}}}}{p_{ss}} = \mathbf{P}(H_1) \frac{e^{-\frac{p_r}{p_{ss}}} \left(e^{\frac{(p_{ps}-p_{ss})(p_r-N\gamma)}{p_{ps}p_{ss}(1+\gamma)}} - 1 \right)}{p_{ps} - p_{ss}} \qquad \text{, if } p_r > N\gamma \qquad (2.23)$$

$$\mathbf{P}(H_0) \cdot 0 = \mathbf{P}(H_1) \cdot 0 \qquad \text{, if } p_r \le N\gamma \qquad (2.24)$$

The second equation is trivially held for every $p_r \le N\gamma$ because it corresponds to an unfeasible event: the absence of error when $p_r \le N\gamma$ i.e. when SINR $\le \gamma$. For $p_r > N\gamma$ the first equation has the following solution

$$p^* = \frac{p_{ss}p_{ps}(1+\gamma)}{p_{ps}-p_{ss}} \left(\log\left(\frac{p_{ps}-p_{ss}}{p_{ss}} \frac{\mathbf{P}(H_0)}{\mathbf{P}(H_1)} + 1 \right) + \frac{N\gamma(p_{ps}-p_{ss})}{p_{ss}p_{ps}(1+\gamma)} \right) \qquad (2.25)$$

If p^* is a real number and $p^* > N\gamma$, then $p_{E_0}^* = p^*$ is the threshold above which the MAP rule selects the hypothesis H_1 in the event E_0. Otherwise one of the two hypothesis always holds. The threshold is defined as $p_{E_0}^* = N\gamma$ if H_1 always holds or as $p_{E_0}^* = \infty$ if H_0 always holds.

We have therefore determined four disjoint sets in the observation space:

$$\begin{aligned}
R_{H_0 E_0} &= \left\{ (p_r, e) : p_r \le p_{E_0}^*, e = 0 \right\} \\
R_{H_1 E_0} &= \left\{ (p_r, e) : p_r > p_{E_0}^*, e = 0 \right\} \\
R_{H_0 E_1} &= \left\{ (p_r, e) : p_r \le p_{E_1}^*, e = 1 \right\} \\
R_{H_1 E_1} &= \left\{ (p_r, e) : p_r > p_{E_1}^*, e = 1 \right\}
\end{aligned} \qquad (2.26)$$

When X and Y are not described by exponential distributions, the procedure to determine the $R_{H_i E_j}$ sets is analogous to the one described. Section 2.5 shows the impact of fading distribution mismatch. These sets are not only useful to establish the MAP rule, but also to estimate the probabilities of correct or incorrect detection.

Detection Probability

For each possible combination of H_0, H_1 and E_0, E_1, two outcomes of the MAP estimation are possible, $\hat{H}_0$, $\hat{H}_1$, corresponding to the hypothesis $\hat{\theta} = 0$, and $\hat{\theta} = 1$, respectively. Selecting $\hat{H}_1$ when H_0 is true is generally referred to as a Type I error, and selecting $\hat{H}_0$ when H_1 is true is a Type II error.

Let us first consider the analysis of the probabilities for H_1, i.e. when transmission overlap is present. The probability of a Type II error when the SU packet is received correctly is defined as

$$\mathbf{P}(\hat{H}_0, E_0 | H_1) = \mathbf{P}\left((P, E) \in R_{H_0 E_0} | H_1 \right), \qquad (2.27)$$

the probability of correct PU activity detection when the SU packet is received correctly is

$$\mathbf{P}(\hat{H}_1, E_0 | H_1) = \mathbf{P}\left((P, E) \in R_{H_1 E_0} | H_1 \right), \qquad (2.28)$$

the Type II error probability when the SU packet received is erroneous is

$$\mathbf{P}(\hat{H}_0, E_1 | H_1) = \mathbf{P}\left((P, E) \in R_{H_0 E_1} | H_1\right), \tag{2.29}$$

and the probability of correct PU activity detection when the SU packet received is erroneous is

$$\mathbf{P}(\hat{H}_1, E_1 | H_1) = \mathbf{P}\left((P, E) \in R_{H_1 E_1} | H_1\right), \tag{2.30}$$

Let us obtain these probabilities for sets $R_{H_i E_j}$ defined with two thresholds ($p_{E_0}^*$ and $p_{E_1}^*$) and then particularize for Rayleigh fading. The probability $\mathbf{P}(\hat{H}_0, E_0 | H_1)$ is given by

$$
\begin{aligned}
\mathbf{P}(\hat{H}_0, E_0 | H_1) &= \mathbf{P}\left(P \le p_{E_0}^*, \text{SINR} > \gamma | H_1\right) \\
&= \mathbf{P}\left(X + Y \le p_{E_0}^*, Y > N\gamma + X\gamma\right) \\
&= \mathbf{P}\left(N\gamma + X\gamma < Y \le p_{E_0}^* - X\right) \\
&= \int_0^{x^*} \int_{N\gamma + x\gamma}^{p_{E_0}^* - x} f_Y(y) f_X(x)\, dy\, dx
\end{aligned}
\tag{2.31}
$$

where $x^* = \frac{p_{E_0}^* - N\gamma}{\gamma + 1}$. Solving for X and Y exponentially distributed we obtain the following probability

$$\mathbf{P}(\hat{H}_0, E_0 | H_1) = e^{-\frac{x^*}{p_{ps}}} p_{ss} \frac{e^{\frac{p_{E_0}^*}{p_{ss}}} \left(e^{-\frac{x^*}{p_{ss}}} - e^{\frac{x^*}{p_{ps}}}\right)}{p_{ss} - p_{ps}} + \frac{e^{-\frac{((N+x^*)\gamma)}{p_{ss}}} \left(e^{x^*\left(\frac{1}{p_{ps}} + \frac{\gamma}{p_{ss}}\right)} - 1\right)}{p_{ss} - p_{ps}\gamma} \tag{2.32}$$

The probability of correct PU detection at the E_0 event is obtained as follows:

$$
\begin{aligned}
\mathbf{P}(\hat{H}_1, E_0 | H_1) &= \mathbf{P}\left(P > p_{E_0}^*, \text{SINR} > \gamma | H_1\right) \\
&= \mathbf{P}\left(X + Y > p_{E_0}^*, Y > N\gamma + X\gamma\right) \\
&= \mathbf{P}\left(Y > \max\left(p_{E_0}^* - X, N\gamma + X\gamma\right)\right) \\
&= \int_0^{x^*} \int_{p_{E_0}^* - x\gamma}^{\infty} f_Y(y) f_X(x)\, dy\, dx \\
&\quad + \int_{x^*}^{\infty} \int_{N\gamma + x\gamma}^{\infty} f_Y(y) f_X(x)\, dy\, dx
\end{aligned}
\tag{2.33}
$$

The expression for Rayleigh fading is

$$\mathbf{P}(\hat{H}_1, E_0 | H_1) = \frac{e^{-\frac{p_{E_0}^*}{p_{ss}}} \left(-1 + e^{\left(\frac{1}{p_{ss}} - \frac{1}{p_{ps}}\right)x^*}\right) p_{ss}}{p_{ps} - p_{ss}} + \frac{e^{-\frac{x^*}{p_{ss}} - \frac{\gamma(N+x^*)}{p_{ss}}} p_{ss}}{p_{ss} + p_{ps}\gamma} \tag{2.34}$$

The probability of a Type II error at the E_1 event is given by:

$$
\begin{aligned}
\mathbf{P}(\hat{H}_0, E_1 | H_1) &= \mathbf{P}\left(P \le p_{E_1}^*, \text{SINR} \le \gamma | H_1\right) \\
&= \mathbf{P}\left(X + Y \le p_{E_1}^*, Y \le N\gamma + X\gamma\right) \\
&= \mathbf{P}\left(Y \le \min\left(N\gamma + X\gamma, p_{E_1}^* - X\right)\right) \\
&= \mathbf{P}\left(Y \le p_{E_1}^* - X\right) \\
&= \int_0^{p_{E_1}^*} \int_0^{p_{E_1}^* - x} f_Y(y) f_X(x)\, dy\, dx
\end{aligned}
\tag{2.35}
$$

where the last equality comes from the fact that $p_{E_1}^* \leq N\gamma$.

$$\mathbf{P}(\hat{H}_0, E_1 | H_1) = \frac{p_{ps}\left(1 - e^{-\frac{p_{E_1}^*}{p_{ps}}}\right) + \left(-1 + e^{-\frac{p_{E_1}^*}{p_{ss}}}\right)p_{ss}}{p_{ps} - p_{ss}} \tag{2.36}$$

The probability of correct PU activity detection at E_1 is:

$$\begin{aligned}
\mathbf{P}(\hat{H}_1, E_1 | H_1) &= \mathbf{P}\left(P > p_{E_1}^*, \mathtt{SINR} \leq \gamma | H_1\right) \\
&= \mathbf{P}\left(X + Y > p_{E_1}^*, Y \leq N\gamma + X\gamma\right) \\
&= \mathbf{P}\left(p_{E_1}^* - X < Y \leq N\gamma + X\gamma\right) \\
&= \int_0^{p_{E_1}^*} \int_{p_{E_1}^* - x}^{N\gamma + x\gamma} f_Y(y) f_X(x) \, dy \, dx \\
&\quad + \int_{p_{E_1}^*}^{\infty} \int_0^{N\gamma + x\gamma} f_Y(y) f_X(x) \, dy \, dx
\end{aligned} \tag{2.37}$$

For Rayleigh fading the expression of this probability is

$$\mathbf{P}(\hat{H}_1, E_1 | H_1) = \frac{e^{-\frac{p_{E_1}^*}{p_{ps}}} p_{ps}}{p_{ps} - p_{ss}} + \frac{e^{-\frac{p_{E_1}^*}{p_{ss}}} p_{ss}}{-p_{ps} + p_{ss}} - \frac{e^{-\frac{N\gamma}{p_{ss}}} p_{ss}}{p_{ss} + p_{ps}\gamma} \tag{2.38}$$

.

It can be checked that for each H_k event (k=0,1), $\sum_{i:0,1} \sum_{j:0,1} \mathbf{P}(\hat{H}_i, E_j | H_k) = 1$. Fig. 2.2 illustrates the integration regions for the detection probabilities in case of transmission overlap (H_1).

Figure 2.2: *Integration regions to compute the probabilities of correct detection, type I and type II errors under the H_1 event (transmission overlap).*

Let us now obtain the probabilities for H_0, i.e. in absence of transmission overlap. The probabilities are defined in the following way

$$\begin{aligned}
\mathbf{P}(\hat{H}_0, E_0 | H_0) &= \mathbf{P}\left((P, E) \in R_{H_0 E_0} | H_0\right) \\
\mathbf{P}(\hat{H}_1, E_0 | H_0) &= \mathbf{P}\left((P, E) \in R_{H_1 E_0} | H_0\right) \\
\mathbf{P}(\hat{H}_0, E_1 | H_0) &= \mathbf{P}\left((P, E) \in R_{H_0 E_1} | H_0\right) \\
\mathbf{P}(\hat{H}_1, E_1 | H_0) &= \mathbf{P}\left((P, E) \in R_{H_0 E_1} | H_0\right)
\end{aligned} \tag{2.39}$$

We now obtain these probabilities as in the H_1 case. The correct estimation probability $\mathbf{P}(\hat{H}_0, E_0|H_0)$ is given by

$$
\begin{aligned}
\mathbf{P}(\hat{H}_0, E_0|H_0) &= \mathbf{P}\left(P \leq p^*_{E_0}, \texttt{SINR} > \gamma|H_0\right) \\
&= \mathbf{P}\left(Y \leq p^*_{E_0}, Y > N\gamma\right) \\
&= \mathbf{P}\left(N\gamma < Y \leq p^*_{E_0}\right) \\
&= \int_{N\gamma}^{p^*_{E_0}} f_Y(y)\,dy
\end{aligned}
\tag{2.40}
$$

which, in the case of Rayleigh fading equals $e^{-N\gamma/p_{ss}} - e^{-p^*_{E_0}/p_{ss}}$. The Type I error $\mathbf{P}(\hat{H}_1, E_0|H_0)$ is

$$
\begin{aligned}
\mathbf{P}(\hat{H}_1, E_0|H_0) &= \mathbf{P}\left(P > p^*_{E_0}, \texttt{SINR} > \gamma|H_0\right) \\
&= \mathbf{P}\left(Y > p^*_{E_0}, Y > N\gamma\right) \\
&= \mathbf{P}\left(Y > \max\left(p^*_{E_0}, N\gamma\right)\right) \\
&= \mathbf{P}\left(Y > p^*_{E_0}\right) \\
&= \int_{p^*_{E_0}}^{\infty} f_Y(y)\,dy
\end{aligned}
\tag{2.41}
$$

For Rayleigh fading $\mathbf{P}(\hat{H}_1, E_0|H_0) = e^{-p^*_{E_0}/p_{ss}}$. The correct estimation probability at the E_1 event is

$$
\begin{aligned}
\mathbf{P}(\hat{H}_0, E_1|H_0) &= \mathbf{P}\left(P \leq p^*_{E_1}, \texttt{SINR} \leq \gamma|H_0\right) \\
&= \mathbf{P}\left(Y \leq p^*_{E_1}, Y \leq N\gamma\right) \\
&= \mathbf{P}\left(Y \leq \min\left(N\gamma + X\gamma, p^*_{E_1}\right)\right) \\
&= \mathbf{P}\left(Y \leq p^*_{E_1}\right) \\
&= \int_0^{p^*_{E_1}} f_Y(y)\,dy
\end{aligned}
\tag{2.42}
$$

which for Rayleigh fading is $1 - e^{-p^*_{E_1}/p_{ss}}$. Finally the Type I error probability at the E_1 event is

$$
\begin{aligned}
\mathbf{P}(\hat{H}_1, E_1|H_0) &= \mathbf{P}\left(P > p^*_{E_1}, \texttt{SINR} \leq \gamma|H_0\right) \\
&= \mathbf{P}\left(Y > p^*_{E_1}, Y \leq N\gamma\right) \\
&= \mathbf{P}\left(p^*_{E_1} < Y \leq N\gamma\right) \\
&= \int_{p^*_{E_1}}^{N\gamma} f_Y(y)\,dy
\end{aligned}
\tag{2.43}
$$

which equals $e^{-p^*_{E_1}/p_{ss}} - e^{-N\gamma/p_{ss}}$ for Rayleigh fading.

2.4   Application to OSA

In the following sections, we analyze the use of the proposed background detection (BD) mechanism in an OSA system, studying its effect on the SU throughput and collision probability with PU transmissions, compared to common practical approaches in OSA.
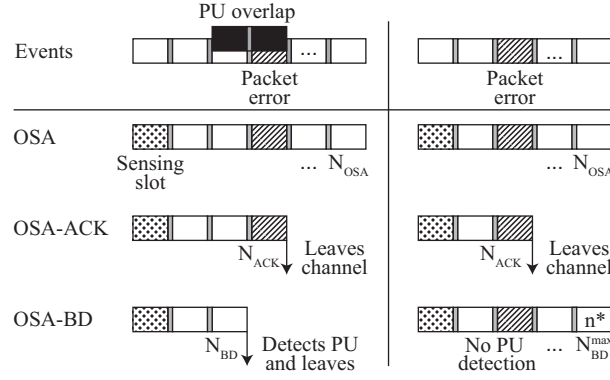
Figure 2.3: Transmission periods with an OSA system, an enhanced OSA with an ACK mechanism and our proposal. On the left, how these mechanisms react to a PU overlap that causes a packet error on its second transmission slot. On the right, how these systems react to packet errors without PU presence. Inter-slot periods are required for signaling.

We assume single-channel transmission capabilities for the SU, leaving for future research the generalization to the multiple-channel case. We also assume that it takes certain number of slots n_s for the SU to detect at least one available channel. For the sake of simplicity, we will consider n_s fixed to 5, although, as we explain later, the resulting expressions can be easily generalized for the case of a randomly distributed n_s. An SU transmission period consists of these scanning slots followed by a number of consecutive transmission time-slots. The SU transmits one data packet in each one of these consecutive time-slots. The PU inter arrival time and the PU channel holding time are characterized by geometric distributions, i.e. the equivalent of a Poisson traffic model for discrete-time. In every time-slot, a channel free of PU traffic is occupied by a PU transmission with probability p, and an ongoing PU transmission ends with probability q. We can therefore characterize the PU activity in a channel as a two-state discrete time Markov chain (DTMC), whose transition matrix is $M = \left[\begin{smallmatrix} 1-p & p \\ q & 1-q \end{smallmatrix}\right]$. Considering the n-th time-slot of an SU transmission period, ϕ_n denotes the probability distribution of the DTMC state on this time-slot. In the single-channel traffic model considered, ϕ_n contains two elements: $\phi_n(1)$ and $\phi_n(2)$, corresponding to no PU activity and PU activity in the channel respectively. The probability distribution in the scanning slot is $\phi_0 = (1, 0)$ if the detection is perfect. With imperfect sensing, $\phi_0 = (1 - p_{\bar{n}}, p_{\bar{n}})$, where $p_{\bar{n}}$ denotes the probability of false negative (the sensing outcome indicates that the channel is free of PU activity when it is not). Given the transition matrix M, the probability distribution for the n-th time-slot is given by $\phi'_n = M^n \phi'_n$, where "′" stands for the transpose operation.

Two parameters characterize the *system performance*: the SU *throughput* (T), defined as the expected number of correctly received SU packets per time-slot in an SU transmission period, and the *collision probability* P_c, defined as the PU-SU transmission overlap probability per time-slot in an SU transmission period. These parameters can be computed by means of a Markov reward process characterizing

the number of consecutive SU transmission slots.

We compare 3 approaches: A generic optimized OSA, an optimized OSA incorporating ACKs, and our proposal OSA-BD. Let us summarize them.

OSA. In (optimized) OSA, the SU transmits in the channel during a fixed number of slots N_{OSA}. This number is set to the integer solving

$$
\begin{aligned}
\max_n\ &T(n) \\
\text{subject to}\quad &P_c(n) \le \alpha \\
&0 \le n \le n_{OSA}^{max}
\end{aligned}
\tag{2.44}
$$

where n_{OSA}^{max} stands for avoiding unbounded n in case $T(n)$ attains its maximum at $n \to \infty$ and $P_c(n)$ remains under α for every n. Note that, provided that $\phi_0 = (1,0)$, $P_c(n)$ is a monotonically increasing sequence bounded by $p/(p+q)$. If several values of n attain the maximum, the OSA algorithm picks the one with the lowest P_c. We refer to this value as N_{OSA}.

The expression for $P_c(n)$ is given by

$$
P_c(n) = \frac{1}{n+n_s} \sum_{i=1}^{n} \phi_i(2)
\tag{2.45}
$$

Defining $\mathbf{P}(E_0|H_0)$ as the probability of correctly receiving an SU packet in the absence of collision, and $\mathbf{P}(E_0|H_1)$ as the probability of E_0 in a time-slot with collision, $T(n)$ is given by the following equation

$$
T(n) = \frac{1}{n+n_s} \left[\sum_{i=1}^{n} \phi_i(1)\mathbf{P}(E_0|H_0) + \sum_{i=1}^{n} \phi_i(2)\mathbf{P}(E_0|H_1) \right]
\tag{2.46}
$$

The probability $\mathbf{P}(E_0|H_0)$ is obtained as follows

$$
\begin{aligned}
\mathbf{P}(E_0|H_0) &= \mathbf{P}\left(\text{SINR} > \gamma | H_0\right) \\
&= \mathbf{P}\left(Y > N\gamma\right) \\
&= \int_{N\gamma}^{\infty} f_Y(y)\,dy
\end{aligned}
\tag{2.47}
$$

and $\mathbf{P}(E_0|H_1)$ is given by

$$
\begin{aligned}
\mathbf{P}(E_0|H_1) &= \mathbf{P}\left(\text{SINR} > \gamma | H_1\right) \\
&= \mathbf{P}\left(Y > N\gamma + X\gamma\right) \\
&= \int_0^{\infty} \int_{N\gamma+x\gamma}^{\infty} f_Y(y) f_X(x)\,dy\,dx
\end{aligned}
\tag{2.48}
$$

For Rayleigh fading we have

$$
\begin{aligned}
\mathbf{P}(E_0|H_0) &= e^{-\frac{N\gamma}{p_{ss}}} \\
\mathbf{P}(E_0|H_1) &= \frac{p_{ss} e^{-\frac{N\gamma}{p_{ss}}}}{p_{ss} + p_{ps}\gamma}
\end{aligned}
\tag{2.49}
$$

The SU uses estimations of p and q to optimize the duration of the transmission period under collision probability constraints. In our numerical results we evaluate the impact of the estimation inaccuracies on the performance.

OSA-ACK. With this version of the protocol, when an SU Tx packet is not decoded correctly by the SU Rx, the SU Rx sends a NACK to the SU Tx and vacates the channel, aborting transmission. This is the same as assuming that every decoding error is due to collision with a PU. Thus, the number of consecutive SU transmission slots, N_{ACK}, is a random variable. We set a maximum number of consecutive transmission slots also for this case, denoted by N_{ACK}^{max}.

OSA-BD. Because BD can detect PU activity simultaneously to SU packet reception, the SU transmission can be aborted at any time-slot. Thus, the number of consecutive transmission slots, N_{BD}, denotes a random variable. We also determine a maximum value for consecutive transmissions in OSA-BD, denoted by N_{BD}^{max}.

Both OSA-ACK and OSA-BD are used on top of an optimized OSA. This implies that N_{ACK}^{max} and N_{BD}^{max} are set to the optimal integer solving (2.44) for each mechanism. Note that the underlying Markov process is different for each access strategy. Fig. 2.3 illustrates the behavior of each scheme. Let us describe the Markov-reward model of the SU transmission process for OSA-BD. It consists in a discrete-time Markov chain (DTMC) in which each state is associated to a reward defined in terms of T of P_c, depending on what is being computed. The DTMC, denoted by Z_k, characterizes a process of consecutive SU transmissions, where k enumerates time-slots. The integer value i taken by Z_k. i.e. the state of the DTMC at the k-th time-slot, has the following interpretation:

- If i is odd, the channel is free of PU activity (H_0) at the $\lceil i/2 \rceil$-th time-slot.

- If i is even, the channel presents PU activity (H_1) at the $i/2$-th time-slot.

After the scanning time-slot ($i = 0$), the SU transmits with probability 1, and enters state $i = 1$ with probability $1 - p$ and state $i = 2$ with probability p. The DTMC advances to a higher state if the BD does not detect PU activity ($\hat{H}_0$), and returns to state 0 otherwise, which means the end of the ongoing SU transmission. Therefore, the DTMC transition probabilities are defined as

$$
\begin{aligned}
p_{0,1} &= 1 - p \\
p_{0,2} &= p \\
p_{2k-1,2k+1} &= (1-p)\mathbf{P}(\hat{H}_0|H_0) &&\text{, for } 0 < k < n \\
p_{2k-1,2k+2} &= p\mathbf{P}(\hat{H}_0|H_0) &&\text{, for } 0 < k < n \\
p_{2k-1,0} &= 1 - \mathbf{P}(\hat{H}_0|H_0) &&\text{, for } 0 < k < n \\
p_{2k,2k+1} &= q\mathbf{P}(\hat{H}_0|H_1) &&\text{, for } 0 < k < n \\
p_{2k,2k+2} &= (1-q)\mathbf{P}(\hat{H}_0|H_1) &&\text{, for } 0 < k < n \\
p_{2k,0} &= 1 - \mathbf{P}(\hat{H}_0|H_1) &&\text{, for } 0 < k < n \\
p_{n,0} &= 1
\end{aligned}
\tag{2.50}
$$

where $\mathbf{P}(\hat{H}_0|H_0)$ and $\mathbf{P}(\hat{H}_0|H_1)$ can be obtained as follows

$$
\begin{aligned}
\mathbf{P}(\hat{H}_0|H_0) &= \mathbf{P}(\hat{H}_0, E_0|H_0) + \mathbf{P}(\hat{H}_0, E_1|H_0) \\
\mathbf{P}(\hat{H}_0|H_1) &= \mathbf{P}(\hat{H}_0, E_0|H_1) + \mathbf{P}(\hat{H}_0, E_1|H_1)
\end{aligned}
\tag{2.51}
$$

where $\mathbf{P}(\hat{H}_0, E_0|H_1)$, $\mathbf{P}(\hat{H}_0, E_1|H_1)$, $\mathbf{P}(\hat{H}_0, E_0|H_0)$, and $\mathbf{P}(\hat{H}_0, E_1|H_0)$, are obtained by (2.31), (2.35), (2.40), and (2.42)respectively. Fig. 2.4 depicts the graph of the DTMC.
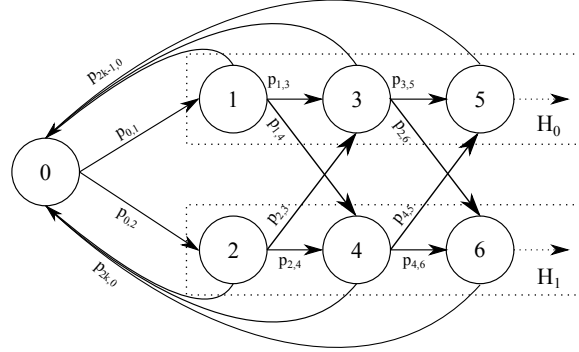

Figure 2.4: Graph of the DTMC characterizing the SU transmission process of OSA-BD.

For a generic reward $r(i)$, the expected average reward is defined as $\bar{r} = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K} r(Z_k)$. For an ergodic, finite-state Markov chain, and provided that $|r(i)|$ is bounded for every i, the total expected average reward $\bar{r}$ is $\bar{r} = \sum_{i=0}^{n} r(i) \pi_i$, where π_i are the steady state probabilities of the DTMC obtained by solving the equilibrium equations $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{T}$, $\|\boldsymbol{\pi}\| = \sum_{i=0}^{n} \pi_i = 1$, where $\mathbf{T}$ is the transition matrix containing the probabilities defined in (2.50). Let $r_T(i)$ and $r_{P_c}(i)$ denote the per-state reward for computing T and P_c respectively. According to the definitions of T and P_c, and because the DTMC models consecutive SU transmissions, we have that $\bar{r}_T = T$ and $\bar{r}_{P_c} = P_c$. When using BD, we will use the terms $\bar{r}_T(n)$ and $\bar{r}_{P_c}(n)$ to refer to the expected averages for the SU throughput and collision probability given a maximum length, n, of SU transmissions. The optimal n is obtained by solving the following optimization problem:

$$\max_n \bar{r}_T(n) \text{ subject to } \bar{r}_{P_c}(n) \leq \alpha \tag{2.52}$$

which is solved for integers $n \in \{0, 1, \ldots, n_m\}$. Note that the trivial values for $n = 0$ are $\bar{r}_T(0) = 0$ and $\bar{r}_{P_c}(0) = 0$. If there is more than one solution, the optimal transmission limit, n^*, is the lowest one attaining the maximum throughput.

Let us see how to obtain the per-stage reward functions $r_T(i)$ and $r_{P_c}(i)$, for every i. The per-stage throughput $r_T(i)$ is defined as the expected number of SU packets correctly received at the i-th SU transmission time-slot. By definition, for every $i > 0$, the SU transmits one packet, therefore $r_T(i) = \mathbf{P}(E_0|i) \cdot 1$, where $\mathbf{P}(E_0|i)$ is the probability of the E_0 event at the i-th time-slot, and is given by

$$\mathbf{P}(E_0|i) = \begin{cases} \mathbf{P}(E_0|H_0) & \text{, if } i \text{ odd} \\ \mathbf{P}(E_0|H_1) & \text{, if } i \text{ even} \end{cases} \tag{2.53}$$

for $i > 0$, where $\mathbf{P}(E_0|H_0)$ and $\mathbf{P}(E_0|H_1)$ are given by (2.47) and (2.48) respectively. Similarly, the per-stage collision probability is $r_{P_c}(i) = 0$ when i is odd, and $r_{P_c}(i) = 1$ when i is even. For $i = 0$ we have $r_T(0) = 0$ and $r_{P_c}(0) = 0$.

The specific structure of matrix $\mathbf{T}$ allows an efficient computation of the steady-state probability vector $\boldsymbol{\pi}$. Let us express $\mathbf{T}$ in a block-matrix form

$$\mathbf{T} = \begin{pmatrix} \mathbf{B}_{0,0} & \mathbf{B}_{0,1} & \mathbf{0} & \mathbf{0} & \cdots \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \mathbf{A} & \mathbf{0} & \cdots \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{A} & \cdots \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \mathbf{A} \end{pmatrix} \tag{2.54}$$

where $\mathbf{B}_{0,0} = (0,0)$, $\mathbf{B}_{0,1} = ((1-p), p)$, and

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix} (1-p)\mathbf{P}(\hat{H}_0|H_0) & p\mathbf{P}(\hat{H}_0|H_0) \\ q\mathbf{P}(\hat{H}_0|H_1) & (1-q)\mathbf{P}(\hat{H}_0|H_1) \end{pmatrix} \\
\mathbf{B} &= \begin{pmatrix} 1-\mathbf{P}(\hat{H}_0|H_0) \\ 1-\mathbf{P}(\hat{H}_0|H_1) \end{pmatrix} \\
\mathbf{1} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}
\end{aligned}
\tag{2.55}
$$

Let us define $\boldsymbol{\pi}_k = (\pi_{2k-1}, \pi_{2k})$, denoting the probability of the events H_0 and H_1 at the k-th transmission slot. Applying the equilibrium equations we have $\boldsymbol{\pi}_0 = \sum_{k=1}^{n-1} \boldsymbol{\pi}_j \mathbf{B} + \boldsymbol{\pi}_n \mathbf{1}$, $\boldsymbol{\pi}_1 = \boldsymbol{\pi}_0 \mathbf{B}_{0,1}$, and $\boldsymbol{\pi}_k$ can be expressed in the matrix-geometric form $\boldsymbol{\pi}_k = \boldsymbol{\pi}_1 \mathbf{A}^{k-1} = \boldsymbol{\pi}_0 \mathbf{B}_{0,1} \mathbf{A}^{k-1}$, for $k > 1$. The normalization condition is $\boldsymbol{\pi}_0 + \sum_{k=1}^{n-1} \boldsymbol{\pi}_j \mathbf{1} + \boldsymbol{\pi}_n \mathbf{1} = 1$, which combined with the equilibrium equations results in the following expression for $\boldsymbol{\pi}_0$

$$\pi_0 = \frac{1}{2 + \sum_{k=1}^{n-1} \mathbf{B}_{0,1} \mathbf{A}^{k-1}(\mathbf{1} - \mathbf{B})} \tag{2.56}$$

Because $\mathbf{A}$ is strictly positive we know, by the Perron-Frobenius theorem, that it has one positive eigenvalue, with multiplicity 1, which is strictly higher than all other eigenvalues. Since the dimension of $\mathbf{A}$ is 2×2, it has only two eigenvalues and thus we conclude that $\mathbf{A}$ has two distinct eigenvalues, λ_1 and λ_2 (with their corresponding eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$). This implies that $\mathbf{A}$ can be expressed as $\mathbf{A} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^{-1}$ where $\mathbf{M} = [\mathbf{e}_1\mathbf{e}_2]$ and $\boldsymbol{\Lambda}$ is a diagonal matrix containing (λ_1, λ_2) in its diagonal. Then

$$\mathbf{A}^{k-1} = \mathbf{M} \begin{pmatrix} \lambda_1^{k-1} & 0 \\ 0 & \lambda_2^{k-1} \end{pmatrix} \mathbf{M}^{-1} \tag{2.57}$$

which allows us to express (2.56) in the following closed form

$$\pi_0 = \left(2 + \mathbf{B}_{0,1}\mathbf{M} \begin{pmatrix} \frac{1-\lambda_1^{n-1}}{1-\lambda_1} & 0 \\ 0 & \frac{1-\lambda_2^{n-1}}{1-\lambda_2} \end{pmatrix} \mathbf{M}^{-1}(\mathbf{1} - \mathbf{B}) \right)^{-1} \tag{2.58}$$

which enables an efficient computation of $\boldsymbol{\pi}_k = \boldsymbol{\pi}_0 \mathbf{B}_{0,1} \mathbf{A}^{k-1}$, for $k = 1, \ldots, n$.

Parameter	Assigned value	Parameter	Assigned value
p_{ss}	-72.9 dBm	$\mathbf{P}(H_1)$	1/3
p_{ps}	-70.5 dBm	$p_{E_0}^*$	-61.8 dBm
N	-103 dBm	$p_{E_1}^*$	-95.9 dBm
γ	7 dBm	$\mathbf{P}(\hat{H}_1\vert H_0)$	$2.65 \cdot 10^{-6}$
p, q	$10^{-3}, 2 \cdot 10^{-3}$	$\mathbf{P}(\hat{H}_0\vert H_1)$	0.1026

Table 2.1: Parameter setting of the reference scenario used in numerical evaluations

The Markov model for OSA-ACK would be similar, and only requires changing the value of correct detection probability in case of no PU activity and missed detection in case of PU activity to

$$
\begin{aligned}
\mathbf{P}(\hat{H}_0\vert H_0) &= \mathbf{P}(E_0\vert H_0) \\
\mathbf{P}(\hat{H}_0\vert H_1) &= \mathbf{P}(E_0\vert H_1)
\end{aligned}
\tag{2.59}
$$

respectively. That is to say, as we introduced previously, OSA-ACK assumes that every decoding error is due to collision with a PU, thus, a SU with OSA-ACK goes back to sensing whenever there is a packet error: with probability $1 - \mathbf{P}(E_0\vert H_0)$ in case there is no PU activity, and with probability $1 - \mathbf{P}(E_0\vert H_1)$ in case of PU activity.

2.5   Numerical Results

We consider a cognitive pair located within a PU BS cell. The distance between the PU BS and the SU Rx is $d_{ps} = 1000$ m. The PU BS transmission power on each channel, p_p, is 20 dBm. The transmission antenna height is $h_t = 10$ m and the receiver is at $h_r = 1.5$ m. The transmission and reception antenna gains are $g_t = 4$ dB and $g_r = 2$ dB, respectively. With these parameters, the average interference power at the SU receiver due to PU BS transmission is given by the pathloss equation $p_{ps} = \dfrac{p_p(h_t h_r)^2 g_t g_r}{d_{ps}^{-4}}$, which equals -70.5 dBm. Assuming that the SU Tx antenna height and gain are 1.5 m and 4 dB, respectively, and that the distance between SU Tx and the SU Rx is $d_{ss} = 250$ m, we obtain $p_{ss} = -72.9$ dBm. Considering a channel bandwidth B = 2 MHz, and an SU Rx noise figure equal to 18 dB, the total noise power is $N = -103$ dBm. The cognitive pair transmits at a rate of 2 Mbit/s using BPSK modulation and, as explained in Section 2.2, each packet is assumed to arrive correctly if BER$< 10^{-3}$, therefore the SINR threshold is $\gamma = 7$ dB. The main parameters are summarized in Table 2.1. We will consider the SU Rx accurately knows these parameters, except when stated otherwise.

It is interesting to assess how the performance figures vary with the optimization parameter n (in time-slots), as shown in Fig. 2.5. Setting the collision probability threshold to $\alpha = 0.025$, the optimal n value for optimized OSA is $N_{OSA} = 49$. For OSA-ACK and OSA-BD, $T(n)$ increases monotonically with n, while $P_c(n)$ remains

below 0.025, thus, n is unbounded in eq. (2.44). We select $N_{ACK}^{max} = N_{BD}^{max} = 100$, which assures $P_c < 0.05$ in case of malfunction of these mechanisms.
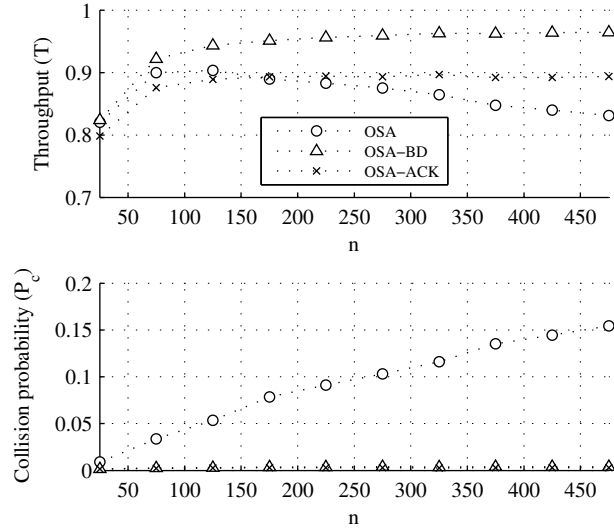

Figure 2.5: Throughput and collision performance with optimized OSA, OSA-ACK and OSA-BD vs. the number of time-slots (n) of the SU transmission. As OSA-ACK and OSA-BD can abort transmissions, their P_c remain low.

The advantage of the OSA-BD mechanism over OSA-ACK is shown in Fig. 2.6. BD is especially useful compared to the ACK mechanism when p_{ss} is small in relation to p_{ps}. In that case, SU packet errors are caused by propagation issues more often than by PU interference, therefore, the ACK mechanism aborts transmission earlier and attains less throughput per transmission period.

Sensitivity to Estimation Inaccuracies

In this section we evaluate and compare the robustness of OSA, OSA-ACK and OSA-BD. Optimized OSA and the BD mechanism are heavily reliant on the characterization of the PU traffic in the channel. To evaluate the impact of traffic estimation inaccuracies, let us consider that the arrival traffic intensity p differs from the estimated one, which is used for the computation of N_{OSA} in optimized OSA and for the threshold computation in BD. In Fig. 2.7 we can see that, in all three systems, T decreases and P_c increases but OSA-BD does not suffer a severe degradation and still outpeforms the other two mechanisms.

Another critical aspect is the channel characterization. Let us consider that the BD design is based on an errored estimation of the average primary power at the SU Rx, p_{ps}. Fig. 2.8 shows the performance of each mechanism under different p_{ps} estimation errors, and two types of fading: Rayleigh and Nakagami with $m = 2$ (while BD is designed for Rayleigh fading in all cases). We see that OSA-BD maintains its advantage with respect to optimized OSA in almost every case, and with respect to OSA-ACK regarding throughput. The statistical
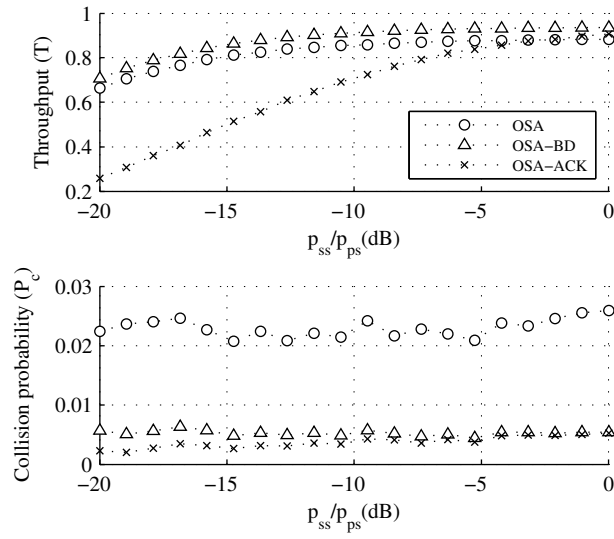
Figure 2.6: Throughput and collision performance with optimized OSA, OSA-ACK and OSA-BD vs. different ratios of secondary to primary power at the SU receiver. $\alpha = 0.025, \mathrm{N_{OSA}} = 49, \mathrm{N_{ACK}^{max}} = \mathrm{N_{BD}^{max}} = 100$.
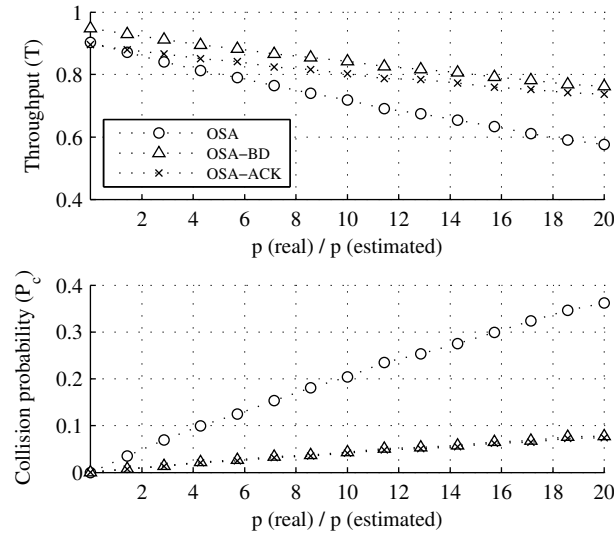

Figure 2.7: Throughput and collision performance with optimized OSA, OSA-ACK and OSA-BD vs. the inaccuracy on the estimation of the traffic intensity. $\alpha = 0.025, \mathrm{N_{OSA}} = 49, \mathrm{N_{ACK}^{max}} = \mathrm{N_{BD}^{max}} = 100$.

characterization of the fading shows a moderate impact on BD performance. In contrast, overestimating p_{ps} shows a more harmful effect on the collision probability. The reason is that, as the real p_{ps} decreases respect the estimated value, the probability of Type II error approaches 1 and, in consequence the OSA-BD tends to operate as an OSA without BD. In Fig. 2.8, $\mathrm{N_{ACK}^{max}}$ and $\mathrm{N_{BD}^{max}}$ are set to 100 which, as explained previously, results in a $\mathrm{P_c}$ near 0.05 in case of BD or ACK malfunction. We can guarantee that the worst-case collision probability of OSA-

BD or OSA-ACK is no worse than OSA's, by simply setting $N_{BD}^{max} \leftarrow N_{OSA}$ and $N_{ACK}^{max} \leftarrow N_{OSA}$, respectively.
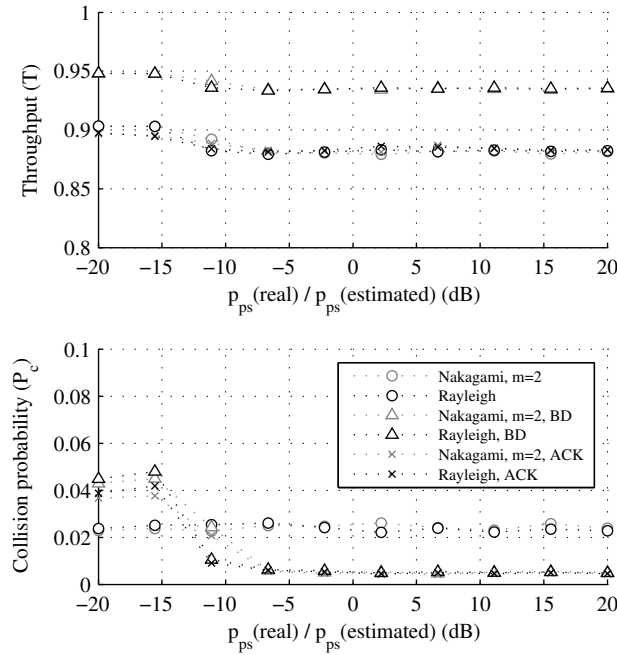

Figure 2.8: Throughput and collision performance with optimized OSA, OSA-ACK and OSA-BD vs. the inaccuracy on the estimation of the primary power received at the SU Rx, for different fading channel characterizations. $\alpha = 0.025, N_{OSA} = 49, N_{ACK}^{max} = N_{BD}^{max} = 100$.

2.6 Conclusions

We have considered the problem of interference caused by transmission overlap in OSA. On top of the existing approach of optimized periodic sensing, we have introduced a novel mechanism called Background Detection. This scheme operates at the SU receiver and uses a Maximum *A Posteriori* estimator to decide, upon each SU packet reception, if there has been collision with a PU packet, aborting transmission if necessary. Our results showed that BD allows for longer SU transmission periods, improving SU throughput while reducing collision probability. In addition, although BD is also reliant on traffic estimations, it enhances the robustness of the OSA protocol against estimation inaccuracies, especially in terms of collision probability. BD shows promising results and it is a simple mechanism to implement (not requiring additional hardware or signaling).

Bandwidth Reservation as a Coexistence Strategy in Opportunistic Spectrum Access Environments

3.1 Introduction

Because of the sensing and transmission problem described in the previous chapter, consequence of the hardware limitations of cognitive devices, there is always some probability of PU and SU transmission overlap and therefore some level of interference at the PUs. In fact, OSA relies on the idea that it makes sense to minimally reduce the QoS of the primary network (PN) if it would result in a much higher spectrum efficiency [45]. Previous research efforts have been mainly focused on designing OSA for the best SU performance at a desired level of PU QoS protection [46]. In contrast, the PN point of view has not received enough attention, let alone the consideration of the PN as an active agent collaborating in OSA. And, as we expalined in the introduction, the consent of the spectrum owners is vital to the implementation of DSA mechanisms.

Motivation

The question addressed in this chapter is: is it worth for the PN to reserve part of its spectrum for easing OSA to a secondary network (SN)? The reserved channels will only be used by the PUs if the non-reserved ones are occupied by other PU transmissions. Intuitively, it seems to be beneficial for the SUs, since it eases finding spectrum opportunities and assures that the free channels are adjacent, simplifying the SU's transmission hardware. The incentives for the PN are not so clear. On the one hand it might avoid collisions with SU transmissions, resulting in less interference at PU receivers, but on the other hand it would reduce the amount of channels available to the PN resource allocation algorithm. In a PN

exploiting the channels with best instantaneous propagation conditions, the reservation scheme would reduce the achievable rate. Determining when the positive aspects compensate the negative ones is not a trivial question and depends on several characteristics of the system: the PU traffic, the SU traffic, the configuration of OSA parameters, the transmission parameters, etc.

The situation considered is not only theoretical. There are scenarios where it would be a feasible coexistence framework. For example, in areas where the regulators force OSA in a licensed spectrum band, it would be an affordable way for the licensed operator to minimize the harmful effects of SU activity. Or if a spectrum owner decides to grant access to certain spectrum bands in exchange for a pre-established flat-rate fee for SUs. This solution could be an alternative to the more complex automatic spectrum trading mechanisms in the literature [47], [48], and more feasible in the early stages of cognitive radio deployment, not requiring the introduction of new protocols, interfaces, etc, as spectrum trading does [49]. With the scheme described in this chapter, the licensed operator maintains the reserved band free of PU activity with a high probability and the cognitive users assure, by means of their sensing capabilities and short transmissions periods, a small collision probability and a low interference level with PU communications.

Related Work

In many networking technologies we can find examples of mechanisms for exploiting spare network resources. For example, in circuit switched networks, it is usual to use protection provisioned capacity to accommodate low-priority traffic in failure-free conditions [50]. In wireless networks, channel reservation schemes were applied to reduce the dropping probability of handover calls [51], [52]. Although the bandwidth reservation idea studied in this paper is conceptually similar to those classic schemes, its application in opportunistic spectrum access environments is completely different in many technical aspects: interference and propagation issues, the sensing of PU activity, the variability of the channel gain, the effect of BR on PU Shannon capacity, among others. Therefore, new research challenges have to be addressed.

Despite the intense research effort conducted on cognitive radio so far, the issue of bandwidth reservation for OSA and its impact on SU and PN performance has not been yet addressed in detail. The term channel reservation in cognitive radio environments is used in [53] and later works like [54], [55] and [56] to refer to a mechanism by which some of the channels not used by the PN are kept free of SU traffic by the secondary network (SN) itself. These "reserved channels" are used by the SN to accommodate ongoing SU transmissions that, forced by PUs, have to leave their initially occupied channels. In that case, the reservation is performed by the SN at the cost of an increased SU blocking probability. That idea is indeed an adaptation of the classic handoff channel reservation scheme to the cognitive environment.

To the best of our knowledge, only one work [57] has considered reserving part of the spectrum so that PUs only occupy the second part when the first one is

congested. While [57] assumed that the PN was completely unaffected by the introduction of a SN and by the reservation scheme, we base our study on the fact that, in real systems, both SU activity and bandwidth reservation do actually have an impact on PU performance. Therefore, we present the idea not as a way to minimize the SN blocking probability, but as a strategy to simultaneously improve PU and SU capacities in a coexistence scenario. Moreover we develop a detailed, realistic model comprising hardware constraints, multichannel OSA, propagation effects and interference.

In the framework of spectrum trading, a recent work [58] proposed a system in which the PN reserves a set of channels for dynamic spectrum auction. This work follows the concept of spectrum pooling, where secondary users temporarily rent from a pool of spectrum resources from primary users [39]. This approach requires the implementation of new protocols to support the required PU-SU signaling and makes the spectrum sensing no longer needed. It is therefore essentially different to the opportunistic access framework of our proposal.

There have been extensive research efforts involving OSA for low-complexity, hardware constrained devices [59], [60] as well as multichannel access [61], [62]. Several works like [63], [64], [65] and [66] consider both characteristics simultaneously. As in [63] we consider that, because of the SU's hardware limitations, sensing can only be conducted within a small portion of the spectrum (channels) and requires a non-negligible time (scanning delay). Our model also includes imperfect sensing, bayesian estimation and the effects of PU interference on SU achievable bit rate.

Because we are interested in evaluating the capacity of the licensed network, the resource allocation scheme of the PN has to be capable of exploiting all the available bandwidth, implying full channel awareness. This is the goal of OFDM scheduling mechanisms [67]. Our model considers one simple but effective mechanism described in [67] and [68]. In works centered on cognitive MAC, the PN is usually characterized by collision probability [65], [43], [69] and [40], and less frequently by overlapping time [41]. However, for more general evaluations of PN and SN coexistence, interference level at PU receivers or PN Shannon capacity is preferred [70]. In OFDM related works, performance is usually evaluated in terms of capacity or achievable rate [67], [68], as we do in this work. The novelty in our analysis is obtaining closed-form analytical expressions for the capacity under OSA interference, with and without BR.

Our Contribution

The main contribution of this chapter is to develop a detailed framework for evaluating the impact of the bandwidth reservation scheme on PU and SU performance. Based on the numerical results obtained, we discuss under which conditions and why it is recommendable or not to implement a PN channel reservation scheme for OSA and how to configure it. As we explain in Section 3.2, the SUs are characterized by hardware-limited radio, imperfect spectrum sensing, multichannel access, and a Maximum A Posteriory (MAP) estimator of the PU activity in the scanned

channels. Section 3.2 also describes the PN resource allocation model. In Section 3.3 we derive the transition probabilities of the Markov model for the PU-SU channel occupation process. By means of the Markov-reward model detailed in Section 3.4, we evaluate the joint PU and SU performance in terms of expected total capacity, interference power and collision probability.

3.2 System Description

Licensed Network

The system under study consists of two overlay wireless networks. The licensed one operates with centralized access coordinated by an access node (AN). As in previous works (e.g. [69], [64],[71], [72], [73]), PU traffic is assumed to follow a Poisson model: PU inter-arrival time and channel holding time are exponential random variables with rates λ_p and μ_p respectively.

We are evaluating the impact of the spectrum reservation scheme on the capacity of the licensed network. Therefore we consider that the AN's resource allocation mechanism exploits all the available bandwidth independently of the number of PUs connected. We only consider resource allocation in the downlink bandwidth, were the OSA operates.

Radio Resource Allocation of the Licensed Network

The downlink spectrum consists of N channels or subcarriers, each one with a bandwidth of W Hz. At each time-slot, each one of the n_p PUs in the system is assigned one of the N channels and part of the total transmission power p_{PU} of the AN. At time-slot t, the wireless link between the j-th PU and the AN is associated to N channel gains, one for each channel. The channels are assumed to present a constant gain over the bandwidth W and over the duration τ of the time-slot, and change over consecutive time-slots due to fading effects (block fading model, [28]). A perfect channel state information is assumed, so that the AN knows all the channel gains at each time-slot. The objective of the resource management algorithm is to assign one channel to each PU and to allocate the power p_{PU} among the n_p assigned channels, at each time-slot, such that the total PU capacity is maximized. A well known strategy to attain this objective is to use a flat transmit power spectral density (PSD) [68], i.e. all the assigned subcarriers are allocated the same power (p_{PU}/n_p). This simplifies the problem to a subcarrier allocation problem which is solved by allocating the best available subcarrier to the PU with the least achieved rate, and then allocating the best subcarrier among the $N-1$ remaining ones to the PU with the second least rate, and so on. This strategy is shown to assure fairness among users and to achieve 96% of the optimum capacity ([68], [67]).

Bandwidth Reservation for OSA

The bandwidth reservation (BR) mechanism considered implies that the AN reserves m adjacent downlink channels for OSA, and the remaining $N - m$ channels are assigned to the n_p PUs following the subcarrier allocation strategy described, whenever $n_p \leq N - m$. In case $n_p > N - m$, part of the m reserved channels (concretely $n_p + m - N$) are assigned to PU transmissions, such that the n_p channels used for PUs remain adjacent among them. Summarizing, the spectrum available for PU downlink transmission contains $\max(n_p, N - m)$ adjacent channels. Figure 3.1 shows an example of the BR mechanism studied.
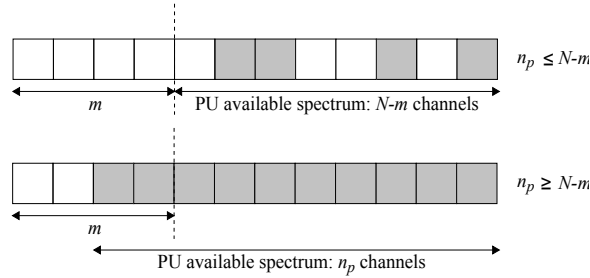
Figure 3.1: Channel reservation scheme in the licensed spectrum. Greyed-out channels are occupied by PUs.

Secondary Network

The unlicensed or secondary network operates in a decentralized, ad-hoc fashion. Communication is always performed between pairs of SUs (*cognitive pairs*) consisting of one sender and one receiver. Every SU is assumed to be under the coverage area of the same licensed access node, thus the objective is to transmit data using one or several of the N licensed channels causing the less possible degradation to PU communications. The SU access follows the interweave paradigm [39], avoiding simultaneous transmissions with PUs.

In line with the low complexity, hardware limited approach of previous works (e.g. [45] as [63], [60]), the MAC protocol of the SUs is similar to the hardware-constrained MAC described in [63]. Summarizing, HC-MAC comprises three consecutive phases: *contention*, *sensing* and *transmission*. The contention procedure allows a pair of SUs (sender and receiver) to reserve the use of the spectrum in a certain area avoiding collisions with other SU transmissions. Then, the cognitive pair starts to sense the spectrum in fixed-duration sensing slots to detect PU activity on each channel. As in [63], the hardware limitations of the SUs are: (1) Each SU is equipped with a single radio that can either transmit or receive, but not at the same time. (2) When scanning the spectrum to detect PU activity, an SU can only sense one narrowband channel during each sensing slot. (3) Once a cognitive pair decides to start a transmission, it can use up to $n_{s,max}$ noncontiguous channels simultaneously.

For the sake of completeness, we consider imperfect sensing, which can result in false positive or false negative PU activity detection. The false positive and false

negative probabilities, p_f and p_n, are computed for an energy level detector, using the formulas presented in [45] and the references therein. The parameters involved are the sensing slot duration s, the channel bandwidth W, the observed signal to noise ratio, and the detection threshold θ.
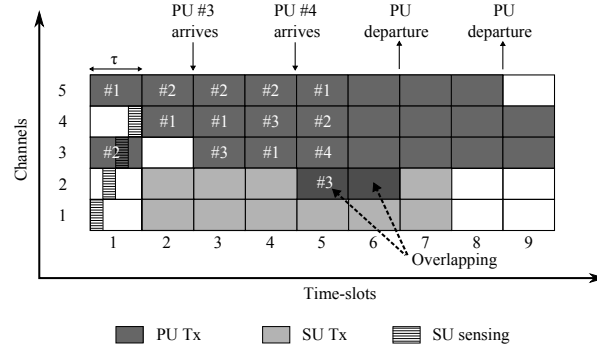

Figure 3.2: Sensing and access phases in OSA. The SU scans $\Delta = 4$ channels, and detects $a = 3$ available channels. In this case, the SU is leaving 1 free channel. This happens when the safety margin is $k = 1$ or when the SU channel occupation limit is $n_{s,max} = 2$. During the SU transmission, two PUs arrive, causing transmission overlap in channel 2 during time-slots 5 and 6. The numbers $\#n$ identify the PU to which the channel is assigned at each time-slot.

During the sensing phase, the SU scans Δ channels. Because this phase lasts one time-slot, the sensing slot duration is $s = \tau/\Delta$. If BR is used, the SU only transmits in the available channels found in the m reserved channels. If not, the SU can transmit in the free channels found in Δ. In both cases the SU can use up to $n_{s,max}$ and, depending on the configuration, it may have to leave k free channels, as a safety margin. Figure 3.2 shows an example of the sensing and access phases. Table 3.1 summarizes the most relevant parameters of the model. Some of them have been already presented in this section and others are explained in Sections 3.3 and 3.4. We also include a brief list of the abbreviations used.

3.3   Spectrum Occupation Model

Spectrum Occupation Process

Upon a PU arrival, the incoming PU is admitted in if there are available channels, and remains in the system for a random time, exponentially distributed with rate μ_p. If $n_p = N$, incoming PUs are rejected (blocking event). Because the system operates in a time-slot basis, we translate the PU channel occupation process into a set of probabilities $P_p(j|i)$, for $i, j \in \{0, 1, \ldots, N\}$, defined as the probability that $n_p = j$ in time-slot $t + 1$, given that $n_p = i$ in time-slot t. These $P_p(j|i)$ are the elements of the $(N + 1) \times (N + 1)$ transition probability matrix $\mathbf{T}$. The stationary distribution $\bar{\pi} = (\pi_0, \pi_1, \ldots, \pi_N)$ is obtained from the equilibrium equations $\bar{\pi} =$

Abbreviation	Definition
PU	Primary User
PN	Primary Network
AN	Access Node
SU	Secondary User
OSA	Opportunistic Spectrum Access
BR	Bandwidth Reservation
MAC	Medium Access Control
MAP	Maximum A Posteriori
PSD	Power Spectral Density

Notation	Definition
W	channel bandwidth
τ	duration of the time-slots
Δ	number of scanned channes
k	safety margin
$n_{s,max}$	maximum channels for SU Tx
m	reserved channels in BR
λ_p	PU's arrival rate
μ_p	PU's service rate
π_i	probability of PU activity in channel i
p_s	probability of SU arrival
q_s	probability of SU departure
ρ_s	SU traffic intensity
T_s	average duration of an SU transmission
n_s	number of SU's in the system
n_p	number of PU's in the system
$C_{j_N}(n_p)$	j-th PU capacity without interference
$C_N(n_p)$	PU capacity without interference
$C_{j,N}^I(n_p, n_s)$	j-th PU capacity with interference
$C_N^{nbr}(n_p, n_s)$	per state PN capacity without BR
$C_N^{br}(n_p, n_s)$	per state PN capacity with BR
$\bar{R}_{PU}^{nbr}$	normalized PN capacity without BR
$\bar{R}_{PU}^{br}$	normalized PN capacity with BR
$C_N'^{nbr}(n_p, n_s)$	per state SU capacity without BR
$C_N'^{br}(n_p, n_s)$	per state SU capacity with BR
$\bar{R}_{SU}^{nbr}$	normalized SU capacity without BR
$\bar{R}_{SU}^{br}$	normalized SU capacity with BR

Table 3.1: Summary of the most relevant parameters of the model

$\bar{\pi}\mathbf{T}$, $\|\bar{\pi}\|_1 = \sum_{i=1}^m \pi_i = 1$. This process is essentially a discrete-time M/M/N/N queue (in Kendall's notation), with blocking probability equal to π_N.

Let us now consider the PU process in conjunction with OSA. At each time slot, the probability that an SU attempts to access the spectrum is p_s. An incoming SU scans the spectrum and then starts an SU transmission period using the detected free channels up to a maximum $n_{s,max}$. Considering $n_{s,max}$ as a safety limit, the OSA MAC assures that no other SU attempts to access the spectrum if another SU is using it. The model could be generalized to accommodate several concurrent SU transmissions, although, for evaluation purposes, it would be equivalent to increase $n_{s,max}$ in the presented model. For an incoming SU, we define the opportunistic access probability, $O(n_s', n_p)$, as the probability that the SU occupies n_s' channels in time-slot $t+1$ given that there are n_p PUs using the spectrum in time-slot t. The duration of an SU transmission period is, in average, T_s, after which the SU leaves the spectrum. The termination probability of an SU transmission is $q_s = 1/T_s$. As a result, considering the SU access process in absence of PU activity, the probability that the spectrum contains an SU transmission is given by $\rho_s = p_s/(p_s + q_s)$, which we will refer to as the SU traffic intensity. Note that the actual probability of SU activity in the system will be lower than ρ_s, in general, because, the spectrum sensing may not always find spectrum opportunities.

This model describes the spectrum occupation process as a Markov chain, Z_t with a state space consisting of a set of pairs, (n_s, n_p), such that $n_s \in \{0, 1, \ldots, n_{s,max}\}$ and $n_p \in \{0, 1, \ldots, N\}$. The transition probabilities from state $Z_t = (n_s, n_p)$ to $Z_{t+1} = (n_s', n_p')$ are given by

$$P\left(n_s', n_p'|n_s, n_p\right) = \begin{cases} p_s O_s\left(n_s', n_p\right) P_p\left(n_p'|n_p\right) & \text{if } n_s = 0,\ n_s' \neq 0 \\ q_s\left(n_s\right) P_p\left(n_p'|n_p\right) & \text{if } n_s \neq 0,\ n_s' = 0 \\ \left((1-p_s) + p_s O_s\left(0, n_p\right)\right) P_p\left(n_p'|n_p\right) & \text{if } n_s = n_s' = 0 \\ (1-q_s) P_p\left(n_p'|n_p\right) & \text{if } n_s = n_s' \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$(3.1)$$

The following subsections develop the derivation of $O_s\left(n_s', n_p\right)$ for each spectrum reservation policy.

SU Spectrum Sensing

After the sensing phase, the SU generates an estimate $\hat{h}$ of the number h of occupied channels in the scanned spectrum. The estimation is based upon the sensing outcome, $\mathbf{x} = (x_1, \ldots, x_\Delta)$, where $x_i = 1$ if the SU detected PU activity in channel i and $x_i = 0$ otherwise. We consider that the SU obtains $\hat{h}$ by a Maximum A Posteriori (MAP) estimation, $\hat{h} = \text{argmax}_h P\left(h|\mathbf{x}\right)$, which, applying Bayes' rule is equivalent to $\hat{h} = \text{argmax}_h P\left(h\right) P\left(\mathbf{x}|h\right)$ (see [74]), where $P\left(h\right)$ and $P\left(\mathbf{x}|h\right)$ can be obtained from the PU traffic and sensing error models, and depend on the spectrum reservation policy.

Without Spectrum Reservation

The conditional distribution $P(\mathbf{x}|h)$ of each outcome can be obtained as follows. For each $\mathbf{x}$ the number of positive detections is $\|\mathbf{x}\|_1 = \sum_{i=1}^{\Delta} x_i$. For a given h, the number of false positives f_p in $\mathbf{x}$ ranges from $f_{p,\min} = (\|\mathbf{x}\|_1 - h)^+$ to $f_{p,\max} = \min(\|\mathbf{x}\|_1, \Delta - h)$. For each f_p, the following equations provide the number of correct positives c_p, false negatives f_n and correct negatives c_n in $\mathbf{x}$:

$$
\begin{aligned}
c_p &= \|\mathbf{x}\|_1 - f_p \\
f_n &= h - (\|\mathbf{x}\|_1 - f_p) \\
c_n &= \Delta - \|\mathbf{x}\|_1 - f_n
\end{aligned}
\tag{3.2}
$$

Then, adding the probabilities of all the possible f_p values yields

$$
P(\mathbf{x}|h) = \sum_{f_p = f_{p,\min}}^{f_{p,\max}} \binom{\Delta - h}{f_p} \binom{h}{c_p} p_n^{f_n} (1 - p_n)^{c_p} p_f^{f_p} (1 - p_f)^{c_n}
\tag{3.3}
$$

When j PUs are randomly allocated over N channels, the number h of PUs in the Δ scanned channels is a hypergeometrical random variable denoted by $P_H(h; N, j, \Delta)$. To obtain the distribution $P(h)$ we apply the law of total probability on every possible value of h:

$$
P(h) = \begin{cases} \sum_{j=h}^{N - \Delta + h} \pi_j P_H(h; N, j, \Delta) & \text{, if } h > 0 \\ \sum_{j=0}^{N - \Delta} \pi_j P_H(h; N, j, \Delta) & \text{, if } h = 0 \end{cases}
\tag{3.4}
$$

With Spectrum Reservation

The SU is assumed to scan, at least, the m reserved channels. Therefore, $\mathbf{x} = (\mathbf{x}_m, \mathbf{x}_{\Delta-m})$, where $\mathbf{x}_m$ is the observation obtained in the first m scanned channels and $\mathbf{x}_{\Delta-m}$ is the observation in the $\Delta - m$ channels scanned in the non-reserved spectrum. Similarly, $h = h_m + h_{\Delta-m}$, where h_m and $h_{\Delta-m}$ are the PUs in the reserved and non-reserved channels respectively. The probabilities $P(\mathbf{x}_{\Delta-m}|h_{\Delta-m})$ and $P(h_{\Delta-m})$ are computed using (3.3) and (3.4) respectively, for $\Delta - m$ scanned channels. In the reserved spectrum, the PUs occupy contiguous channels, therefore, for a given h_m, we can obtain f_p, c_n, f_n, and c_p in $\mathbf{x}_m$ as follows

$$
\begin{aligned}
f_p &= \sum_{j=h_m+1}^{m} x_j \\
c_n &= m - h_m - \sum_{j=h_m+1}^{m} x_j \\
f_n &= h_m - \sum_{j=1}^{h_m} x_j \\
c_p &= \sum_{j=1}^{h_m} x_j
\end{aligned}
\tag{3.5}
$$

and the conditional probability $P(\mathbf{x}_m|h_m)$ is given by

$$
P(\mathbf{x}_m|h_m) = p_n^{f_n} (1 - p_n)^{c_p} p_f^{f_p} (1 - p_f)^{c_n}
\tag{3.6}
$$

The distribution $P(h_m)$ is directly obtained from $\bar{\pi}$:

$$
P(h_m) = \begin{cases} \pi_{N-m+h} & \text{, if } h_m > 0 \\ \sum_{j=0}^{N-m} \pi_j & \text{, if } h_m = 0 \end{cases}
\tag{3.7}
$$

SU Spectrum Access

Let $D_\Delta(a, h)$ denote the probability of finding a number a of available channels given that, during the sensing phase, h out of the Δ scanned channels are occupied by PU transmissions. In order to compute $D_\Delta(a, h)$ let us define $X_{\hat{h}}$ as the set of possible values of $\mathbf{x}$ for which the outcome of the MAP estimator equals $\hat{h}$:

$$X_{\hat{h}} = \left\{ \mathbf{x} | \hat{h} = \underset{h}{\operatorname{argmax}} \; P(h) \, P(\mathbf{x}|h) \right\}.$$

Finding a free channels implies detecting PU activity in $\hat{h} = \Delta - a$, therefore $D_\Delta(a, h)$ is given by

$$
\begin{aligned}
D_\Delta(a, h) &= P(\hat{h}|h) \\
&= \sum_{\mathbf{x} \in X_{\hat{h}=\Delta-a}} P(\mathbf{x}|h)
\end{aligned}
\tag{3.8}
$$

Let us consider the no BR and the BR cases separately for the computation of $O_s(n_s', n_p)$.

Without Spectrum Reservation

The probability that, in any time slot, h channels out of Δ scanned channels are used by PUs, when there are n_p PUs in the spectrum, is hypergeometrical. Therefore

$$
O_s(n_s', n_p) = \begin{cases}
\displaystyle\sum_{h=h_{min}}^{h_{max}} D_\Delta(n_s' + k, h) \, P_H(h; N, n_p, \Delta) & \text{if } n_s' < n_{s,max} \\
\displaystyle\sum_{a=n_s'+k}^{\Delta} \sum_{h=h_{min}}^{h_{max}} D_\Delta(a, h) \, P_H(h; N, n_p, \Delta) & \text{if } n_s' = n_{s,max}
\end{cases}
\tag{3.9}
$$

where $h_{min} = (\Delta + n_p - N)^+$ and $h_{max} = \min(\Delta, n_p)$ are the minimum and maximum number of PUs within Δ.

With Spectrum Reservation

In this case, an incoming SU occupies n_s' channels, always within the m reserved ones. Therefore, the SU detects $a_m = \min(m, n_s' + k)$ in the m reserved channels and $a_{\Delta-m} = (n_s' + k - m)^+$ in the no reserved ones. For $n_p > N - m$, $h_m = n_p + m - N$ and $h_{\Delta-m} = \Delta - m$, therefore

$$O_s(n_s', n_p) = D_m(a_m, h_m) \, D_{\Delta-m}(a_{\Delta-m}, h_{\Delta-m}) \tag{3.10}$$

For $n_p \leq N - m$, $h_m = 0$ and each value of $h_{\Delta-m}$ has a hypergeometric probability $P_H(h_{\Delta-m}; N, n_p, \Delta - m)$. Multiplying (3.9) by the $D_m(a_m, h_m)$ term, we can obtain $O_s(n_s', n_p)$. In case $n_s' = n_{s,max}$, the outer summation is done over $a_{\Delta-m} = \Delta - m, \dots, n_s' + k - m$.

3.4 Markov Reward Model for PU and SU Capacities

In this section we construct a Markov reward model to estimate the achievable transmission rate for PUs and SUs. Such a model requires the definition of a reward function $R(n_s, n_p)$ providing the Shannon capacity at each state. The objective is to obtain the expected average reward defined as

$$\bar{R} = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} R(Z_t) \tag{3.11}$$

For an ergodic, finite-state Markov chain, and provided that $|R(n_s, n_p)|$ is bounded for every (n_s, n_p), the total expected average reward $\bar{R}$ is

$$\bar{R} = \sum_{n_p=0}^{N} \sum_{n_s=0}^{m-k} R(n_s, n_p) P(n_s, n_p) \tag{3.12}$$

where $P(n_s, n_p)$ are the steady state probabilities of the Markov process defined by (3.1).

Propagation model

Before deriving the expressions for PU and SU capacities with and without BR, we have to establish the models for the pathloss and fading effects on the transmitted signals. Let p_{PU} denote the total available transmission power at the PU Tx (AN) for the whole spectrum. The power transmitted in each channel, when there are n_p PUs, is $p_p = p_{PU}/n_p$ (flat transmit PSD). The average received power at a PU terminal, after considering the path loss, is referred to as p_{pp}. Similarly, p_{SU}, p_s, and p_{ss}, are the total power at the SU Tx, the SU power in one channel, and the average SU received power in one channel, respectively. For interference computations, we use p_{sp} to refer to the average SU power received at a PU Rx location and p_{ps} for the PU power at an SU Rx location. Figure 3.3 shows a diagram with the transmission and interference powers between PUs and SUs.
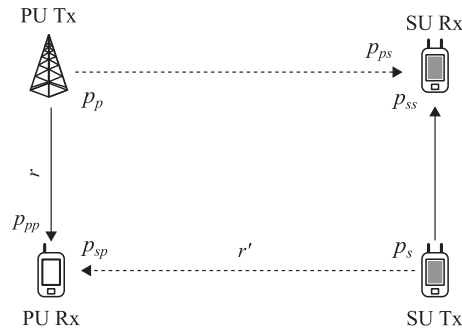

Figure 3.3: *Diagram of the interference between PUs and SUs.*

For path loss estimation, we consider a common empirical method known as piecewise linear model (see [28]). In this model, for a given distance r between the

Tx and the Rx, the average received power is approximately equal to $p_{pp} = p_p K g(r)$, where K is a constant depending on the antenna gains, and $g(r)$ is given by

$$g(r) = \begin{cases} 1 & \text{if } r \leq 1 \\ r^{-n_1} & \text{if } 1 < r \leq r_c \\ r_c^{-n_1} \left(\frac{r}{r_c}\right)^{-n_2} & \text{if } r > r_c \end{cases} \qquad (3.13)$$

The parameters n_1, n_2 and r_c are empirically determined depending on the propagation environment. Typical values are around $n_1 = 2$, $n_2 = 4$, and $r_c = 100$ (that we use in our model). The same $g(r)$ is also used for SU signals.

The instantaneous received power changes in consecutive time-slots due to multipath fading effects, and is characterized by a probability density function (*pdf*). In particular, assuming that the fading amplitude follows a Rayleigh distribution, the received power has the following *pdf*, $f(p) = e^{-p/\bar{p}}/\bar{p}$, where $\bar{p}$ is the average power at the Rx location.

The capacity expressions of the following subsections are obtained for the $g(r)$ and $f(p)$ functions described. Note, however that for different fading models (e.g. Nakagami or Rice) or different path loss models (two ray, Okumura-Hata, etc) the procedure is similar, although the equations may be more mathematically involved or require numerical evaluation.

PU downlink capacity without interference

Let us consider a PU that, in a given time-slot, is the j-th one to be assigned a channel by the AN. Because the Tx power on each channel remains constant for a fixed n_p, the maximum achievable rate per channel is determined by the additive white Gaussian noise (AWGN) capacity model (see [28]). The capacity for the j-th PU is given by $C_{j,N} = W\log_2(1 + \text{SNR}_p)$, where SNR_p is the average signal to noise ratio at the j-th PU Rx. The N subscript remarks that there are N available channels to allocate PU transmissions. Each one of the N channels has a different gain for the j-th PU Rx, but, because $j-1$ channels are already assigned, the AN selects, for the j-th PU, the channel with the highest gain among the $N-j+1$ remaining ones. In consequence the received power will be the highest among $N-j+1$ possible values.

The cumulative distribution function (CDF) for Rayleigh fading is $F(p) = (1 - e^{-p/p_{pp}})$. Therefore, the CDF of the highest value among $N-j+1$, is given by $P(P \leq p) = (1 - e^{-p/p_{pp}})^{N-j+1}$, resulting in the following *pdf*

$$f_j(p) = \frac{dP(P \leq p)}{dp} = \frac{N-j+1}{p_{pp}} e^{-\frac{p}{p_{pp}}} \left(1 - e^{-\frac{p}{p_{pp}}}\right)^{N-j} \qquad (3.14)$$

We can now obtain SNR_p

$$\text{SNR}_p = \int_{p=0}^{\infty} \frac{p}{N_0 W} f_j(p) dp = \frac{p_{pp} H_{N-j+1}}{N_0 W} \qquad (3.15)$$

where $H_n = \sum_{i=0}^{n} i^{-1}$ is the n-th harmonic number. In the rest of this section, we will refer to H_{N-j+1} as $H^{(j)}$. Applying $p_{pp} = p_{PU}Kg(r)/n_p$, we define

$$\text{SNR}_p(r) = \frac{p_{PU}H^{(j)}Kg(r)}{n_p N_0 W} \tag{3.16}$$

where r denotes the distance between the PU Tx and Rx. We could obtain the expectation of $C_{j,N}$ over all possible r values in the cell but, for mathematical tractability, we will consider upper bounds for the average capacity on the coverage area. Because of the concavity of the $\log(1+x)$ function, we know, by Jensen's inequality, that $\mathbf{E}\left[\log_2\left(1+\text{SNR}_p(r)\right)\right] \le \log_2\left(1+\mathbf{E}\left[\text{SNR}_p(r)\right]\right)$. Considering that all PU Rx locations are equally probable and the PU Tx is at the center of the cell, the *pdf* of the distance r over a cell of radius R is given by $f_r(r) = 2r/R^2$. Therefore

$$\underset{r}{\mathbf{E}}\left[\text{SNR}_p(r)\right] = \frac{p_{PU}H^{(j)}K}{n_p N_0 W}\underset{r}{\mathbf{E}}\left[g(r)\right] \tag{3.17}$$

where

$$\underset{r}{\mathbf{E}}\left[g(r)\right] = \int_{r=0}^{R} g(r)f_r(r)dr = \frac{\left(2R^2\left(1+\log(r_c)\right)-r_c^2\right)}{R^4} \tag{3.18}$$

The resulting $C_{j,N}(n_p)$ expression is

$$C_{j,N}(n_p) = W\log_2\left(1 + \frac{p_{PU}H^{(j)}K}{n_p N_0 W}\underset{r}{\mathbf{E}}\left[g(r)\right]\right) \tag{3.19}$$

And finally, in absence of OSA interference, the total expected downlink capacity for an N-channel system containing n_p PU transmissions is obtained by summing the average capacities $C_{j,N}(n_p)$ over the n_p PUs

$$C_N(n_p) = \sum_{j=1}^{n_p} C_{j,N}(n_p) \tag{3.20}$$

PU downlink capacity with interference

Let us now consider a time-slot where the channel assigned to the j-th PU is occupied by an SU transmission, causing interference at the PU Rx. By the AWGN model, the capacity for the j-th PU in this time-slot is now given by $C_{j,N}^{I} = W\log_2\left(1+\text{SINR}_p\right)$, where SINR_p is the average signal to interference and noise ratio at the j-th PU Rx. For the general case of the PU Tx and SU Tx located at different places, the PU signal and the interference signal follow independent fading processes. With the interference power *pdf* given by $f^I(p) = e^{-p/p_{sp}}/p_{sp}$, the SINR_p is obtained as follows

$$\begin{aligned}
\text{SINR}_p &= \int_{y=0}^{\infty}\int_{x=0}^{\infty} \frac{x}{y+N_0 W}f_j(x)f^I(y)dxdy \\
&= \int_{y=0}^{\infty} \frac{p_{pp}H^{(j)}}{y+N_0 W}f^I(y)dy \\
&= \frac{p_{pp}H^{(j)}}{p_{sp}}\exp\left(\frac{N_0 W}{p_{sp}}\right)E_1\left(\frac{N_0 W}{p_{sp}}\right)
\end{aligned} \tag{3.21}$$

where $E_1(a) = \int_a^\infty t^{-1} e^{-t} dt$, is the exponential integral function (equivalent to the upper incomplete gamma function $\Gamma(0, a)$).

Considering that $p_{sp} = p_{su} K' g(r') / n_p$, where r' is the distance between the interfering SU Tx and the PU Rx, the SINR_p (3.21) depends now on r and r'. To obtain an upper bound of $C_{j,N}^I$ over the possible r, r' values in the cell, we make the following considerations

$$\underset{r,r'}{E} \left[\log_2 \left(1 + \text{SINR}_p(r, r') \right) \right] \leq \log_2 \left(1 + \underset{r,r'}{E} \left[\text{SINR}_p(r, r') \right] \right) \tag{3.22}$$

which follows from Jensen's inequality. Assuming independence between r and r', we have that

$$
\begin{aligned}
\underset{r,r'}{E} \left[\text{SINR}_p(r, r') \right] &= \underset{r'}{E} \left[\underset{r}{E} \left[\text{SINR}_p(r, r') \right] \right] \\
&= \underset{r'}{E} \left[n_s H^{(j)} p_{PU} K \underset{r}{E} \left[g(r) \right] \frac{\exp\left(\frac{n_s N_0 W}{p_{su} K' g(r')} \right) E_1 \left(\frac{n_s N_0 W}{p_{su} K' g(r')} \right)}{n_p p_{su} K' g(r')} \right]
\end{aligned}
\tag{3.23}
$$

The function involving the expectation over r' in the last term, that is, $\phi_{\text{SINR}_p}(r') = \underset{r}{E} \left[\text{SINR}_p(r, r') \right]$, is strictly increasing in r' (SINR_p increases as the interfering Tx moves away) and concave for $r' > \delta$. In particular, δ is the smallest r' such that $r' \geq r_c$ and $\frac{\partial^2 \phi_{\text{SINR}}(r')}{\partial r'^2} \leq 0$. Let $f_{r'}(r')$ and $\bar{r}'$ denote the *pdf* and the average value of r' respectively. If $f_{r'}(r') > 0$ for every r', and $\delta < \bar{r}'$ then

$$\underset{r'}{E} \left[\phi_{\text{SINR}_p}(r') \right] \leq \phi_{\text{SINR}_p}(\bar{r}') \tag{3.24}$$

which can be easily checked by considering an alternative distribution $f_{r''}(r'')$, such that $f_{r''}(r'') = 0$ for $r'' < \delta$, $f_{r''}(r) \geq f_{r'}(r)$ for $r \geq \delta$, and having an expected value $\bar{r}'' = \bar{r}'$. We have that

$$\underset{r'}{E} \left[\phi_{\text{SINR}_p}(r') \right] < \underset{r''}{E} \left[\phi_{\text{SINR}_p}(r'') \right] \leq \phi_{\text{SINR}_p}(\bar{r}') \tag{3.25}$$

where the first inequality follows from the strictly increasing property of $\phi_{\text{SINR}_p}(r')$ in r', and the second one from the concavity of $\phi_{\text{SINR}_p}(r'')$ in the $f_{r''}(r'')$ domain ($r'' \geq \delta$) allowing us to apply Jensen's inequality in this domain. Numerical evaluations showed that, for typical configurations, $\delta = r_c << \bar{r}'$ and previous expression can be used to obtain an upper bound for the average PU capacity over the cell. Note that this approach is equivalent to consider that the interfering SU is located at a distance $\bar{r}'$ from the PU Rx. When every SU and PU locations within the cell are equally probable, the average distance is $\bar{r}' = \frac{128 R}{45 \pi}$ (see [75]).

For a pessimistic estimation of the capacity under interference, using a more straightforward reasoning we can obtain a lower bound for $\underset{r'}{E} \left[\phi_{\text{SINR}_p}(r') \right]$ by simply setting $g(r') = \underset{r'}{E} \left[g(r') \right]$ in (3.23).

Finally, the capacity for the j-th PU under interference, $C_{j,N}^I(n_s, n_p)$, is given by

$$C_{j,N}^I(n_s, n_p) = W \log_2 \left(1 + \frac{n_s H^{(j)} p_{PU} K \underset{r}{E} \left[g(r) \right] \exp\left(\text{SNR}_s(\bar{r}')^{-1} \right) E_1 \left(\text{SNR}_s(\bar{r}')^{-1} \right)}{n_p p_{su} K' g(\bar{r}')} \right) \tag{3.26}$$

Where $\mathrm{SNR}_s(r') = p_{SU}K'g(r')/n_s N_0 W$. The total capacity if all the n_p channels experience interference is $C_N^I(n_s, n_p) = \sum_{j=1}^{n_p} C_{j,N}^I(n_s, n_p)$. Next we derive the expressions for the total downlink capacity with and without BR.

No bandwidth reservation for OSA

Given n_p and n_s, the number i of collisions may take a value ranging from $c_{min} = (n_p + n_s - N)^+$ to $c_{max} = \min(n_p, n_s)$. If there are i collisions, each one of the n_p PUs may be colliding probability $p_c = \frac{i}{n_p}$, or may not, with probability $1 - p_c$. The expected capacity for the j-th PU is given by $C_{j,N}^{(i)}(n_p, n_s) = C_{j,N}^I(n_s, n_p)p_c + C_{j,N}(n_p)(1 - p_c)$. Summing the n_p average capacities, we obtain the total capacity for i collisions

$$
\begin{aligned}
C_N^{(i)}(n_p, n_s) &= \sum_{j=1}^{n_p} C_{j,N}^{(i)}(n_p, n_s) \\
&= \sum_{j=1}^{n_p} C_{j,N}^I(n_s, n_p)\frac{i}{n_p} \\
&\quad + \sum_{j=1}^{n_p} C_{j,N}(n_p, n_s)\left(1 - \frac{i}{n_p}\right) \\
&= C_N^I(n_s, n_p)\frac{i}{n_p} + C_N(n_p, n_s)\left(1 - \frac{i}{n_p}\right)
\end{aligned}
\tag{3.27}
$$

The probability of i collisions equals $P_H(i; N, n_s, n_p)$, therefore, the total capacity without BR, $C_N^{nbr}(n_p, n_s)$, is given by

$$
C_N^{nbr}(n_p, n_s) = \sum_{i=c_{min}}^{c_{max}} C_N^{(i)}(n_p, n_s) P_H(i; N, n_s, n_p)
\tag{3.28}
$$

The associate reward function $R_{PU}^{nbr}(n_p, n_s)$ accounts for the PU capacity only when there is PU activity, normalizing respect $C_N(n_p)$, as follows

$$
R_{PU}^{nbr}(n_p, n_s) = \frac{C_N^{nbr}(n_p, n_s)}{(1 - \pi_0) C_N(n_p)}
\tag{3.29}
$$

where $(1 - \pi_0)$ is the probability that there is at least one PU transmission. It can be checked that, in expression (3.12), the factors $P(n_s, n_p)/(1 - \pi_0)$ correspond to the probability distribution over the states where $n_p > 0$.

With bandwidth reservation for OSA

According to the BR mechanism definition, when $n_p \le N - m$, the system is collision free and the AN has $N - m$ channels to dynamically allocate PU transmissions, and the obtained downlink capacity with BR, $C_N^{br}(n_p, n_s)$, equals the no-interference capacity for a spectrum with $N - m$ channels, $C_{N-m}(n_p)$, (3.20). When $n_p > N - m$, the PUs occupy part of the reserved spectrum, and therefore collisions are possible. Given n_s, the number of channels with collisions is $c = (n_s - N + n_p)^+$, and $C_N^{br}(n_p, n_s) = C_{n_p}^{(c)}(n_p, n_s)$, (3.27). Summarizing, we have

$$
C_N^{br}(n_p, n_s) = \begin{cases} C_{N-m}(n_p) & \text{, if } n_p \le N - m \\ C_{n_p}^{(c)}(n_p, n_s) & \text{, if } n_p > N - m \end{cases}
\tag{3.30}
$$

The reward function $R_{PU}^{br}(n_p, n_s)$ is obtained as in (3.29).

SU capacity without interference

During an SU transmission period, the SU uses the same power $p_s = p_{SU}/n_s$ in each of the n_s occupied channels. Therefore, the SU AWGN capacity per channel is given by $C' = W\log_2(1 + SNR_s)$, where SNR_s is the average SU signal to noise ratio in one channel. With Rayleigh fading and an average received power $p_{ss} = p_{SU}K'g(r')/n_s$, we have $SNR_s(r') = p_{SU}K'g(r')/n_sN_0W$. Taking the expectation over r' we have that $\underset{r'}{\mathbf{E}}[\log_2(1 + SNR_s(r'))] \leq \log_2\left(1 + \underset{r'}{\mathbf{E}}[SNR_s(r')]\right)$. This upper bound defines the SU capacity for n_s channels as

$$
\begin{aligned}
C'(n_s) &= n_sW\log_2\left(1 + \underset{r'}{\mathbf{E}}[SNR_s(r')]\right) \\
&= n_sW\log_2\left(1 + \frac{p_{SU}K'\underset{r'}{\mathbf{E}}[g(r')]}{n_sN_0W}\right)
\end{aligned}
\tag{3.31}
$$

The term $\underset{r'}{\mathbf{E}}[g(r')]$ can be computed exactly for the distribution of two points chosen at random within a disk of radius R

$$
f_{r'}(r') = \frac{4r'}{\pi R^2}\cos^{-1}\left(\frac{r'}{2R}\right) - \frac{2r'^2}{\pi R^3}\sqrt{1 - \frac{r'^2}{4R^2}}
\tag{3.32}
$$

However, for the sake of simplicity, we present an expression for $\underset{r'}{\mathbf{E}}[g(r')]$ obtained with a close approximation of (3.32): $\tilde{f}_{r'}(r') = c_1x + c_2x^2 + c_3x^3$. The coefficients $c_1 = -(64 - 27\pi)/3\pi R^2$, $c_2 - 4(3\pi - 8)/\pi R^2 =$ and $c_3 = -(128 - 45\pi)/12\pi R^4$ are obtained from the equations $\tilde{f}_{r'}(0) = 0$, $\tilde{f}_{r'}(2R) = 0$, $\int_0^{2R}\tilde{f}_{r'}(r')dr' = 1$, and $\int_0^{2R}r'\tilde{f}_{r'}(r')dr' = 128\pi/45\pi$. The resulting expression is

$$
\begin{aligned}
\underset{r'}{\mathbf{E}}[g(r)] &\approx \int_{r'=0}^{2R}g(r)\tilde{f}_{r'}(r')dr' \\
&= c_1 + \frac{-c_3 + 2c_3r_c^2}{4} + \frac{c_2}{12}\left(-8 + 24r_c - \frac{6r_c^2}{R}\right) - \frac{c_1c_3^2}{8R^2} + c_1Ln[r_c] + c_3r_c^2Ln\left[\frac{2R}{r_c}\right]
\end{aligned}
\tag{3.33}
$$

SU capacity with interference

During a time-slot, the SU per-channel capacity under PU interference is $C'^I = W\log_2(1 + SINR_s)$, where $SINR_s$ denotes the SU signal to interference ratio. When both signals experience Rayleigh fading, the $SINR_s$ is given by

$$
\begin{aligned}
SINR_s &= \int_{y=0}^{\infty}\int_{x=0}^{\infty}\frac{x}{y + N_0W}f'(x)f^I(y)dxdy \\
&= \int_{y=0}^{\infty}\frac{p_{ss}}{y + N_0W}f^I(y)dy \\
&= \frac{p_{pp}}{p_{ps}}e^{\frac{N_0W}{p_{ps}}}E_1\left(\frac{N_0W}{p_{ps}}\right)
\end{aligned}
\tag{3.34}
$$

where $p_{ss} = p_{SU}K'g(r')/n_s$ and $p_{ps} = p_{PU}Kg(r)/n_p$. For the average capacity over all r and r' values, we can apply the same reasoning used for $C^I_{j,N}(n_p, n_s)$, to conclude that

$$
\underset{r',r}{\mathbf{E}}[\log_2(1 + SINR_s(r', r))] \leq \log_2\left(1 + \underset{r'}{\mathbf{E}}[SINR_s(r', \bar{r})]\right)
\tag{3.35}
$$

The corresponding SU per-channel capacity with interference, for an (n_p, n_s) pair is

$$C'^I(n_s, n_p) = W\log_2 \left(1 + \frac{n_p p_{SU} K' \underset{r'}{\mathbf{E}} [g(r')] \exp\left(\mathrm{SNR}'_p(\bar{r})^{-1}\right) E_1\left(\mathrm{SNR}'_p(\bar{r})^{-1}\right)}{n_s p_{PU} K g(\bar{r})}\right)$$

(3.36)

where $\mathrm{SNR}'_p(r) = p_{PU} K g(r)/n_p N_0 W$.

without Bandwidth Reservation

Let z denote the number of channels with overlapping transmissions. Given n_s and n_p, z takes random values ranging from $z_{\min} = (n_s + n_p - N)^+$ to $z_{\max} = \min(n_s, n_p)$, with hypergeometric distribution. The SU capacity, without BR, in the n_s used channels, $C'^{nbr}(n_s, n_p)$ is obtained by taking expectation over the z values:

$$C'^{nbr}(n_s, n_p) = \sum_{z=z_{\min}}^{z_{\max}} \left((n_s - z) C'(n_s) + z C'^I(n_s, n_p)\right) P_H(z; N, n_p, n_s) \quad (3.37)$$

To define the reward function $R_{SU}^{nbr}(n_s, n_p)$, we consider the maximum SU capacity, $\rho_{SU} C'(n_{s,\max})$, corresponding to an ideal case where the SUs access the spectrum on every attempt, and transmit in $n_{s,\max}$ channels without interference. Therefore

$$R_{SU}^{nbr}(n_s, n_p) = \frac{C'^{nbr}(n_s, n_p)}{\rho_s C'(n_{s,\max})} \quad (3.38)$$

with Bandwidth Reservation

In this case, given n_s and n_p, the number of channels with overlapped transmissions is $z_{\min} = (n_s + n_p - N)^+$, and therefore

$$C'^{br}(n_s, n_p) = (n_s - z_{\min}) C'(n_s) + z_{\min} C'^I(n_s, n_p) \quad (3.39)$$

The reward $R_{SU}^{br}(n_s, n_p)$, is obtained as in (3.38).

3.5 Numerical Results

System Configuration

The model developed in previous section allows us to obtain numerical results upon which we address the main issues presented in the introduction of the chapter: Is it justified the use of spectrum reservation for OSA? Under which conditions? Can we jointly improve the performance of both PUs and SUs?

To jointly evaluate PU-SU configurations, we have considered several combinations of PU and SU traffic intensities. To illustrate it, we show the results for $\rho_s = 0.5$ and two PU traffic intensities: $\lambda_p = 0.1$ (low PU traffic) and $\lambda_p = 0.5$ (high PU traffic). Considering that each PU remains in the spectrum for $1/\mu_p = 60$

Parameter	Assigned value
number of channels, N	15
channel bandwidth, W	200 KHz
time slot duration, τ	200 µs
PU average Tx time, $1/\mu_p$	60 s
SU average Tx time, T_s	$10 \times \tau$
sensing detection threshold, θ	17.8 dB
PU Tx power, p_p	10 W
SU Tx power, p_s	4 W
cell radius, R	1000 m
$g(r)$ threshold distance, r_c	100 m
$g(r)$ propagation factor, $K = K'$	4 dB
receivers' noise figure, F	7 dB

Table 3.2: Parameter setting of the reference scenario used in numerical evaluations

s, the PU blocking probabilities are 0.2% for $\lambda_p = 0.1$, and 50% for $\lambda_p = 0.5$. In all cases, the licensed spectrum band is divided into $N = 15$ narrowband channels of $W = 200$ KHz, the time-slot duration is $\tau = 200$ µs, and the average duration of an SU transmission is $T_s = 10$ time-slots. The remaining parameters of the model, such as transmission powers and cell radius, are provided in Table 3.2. Regarding the sensing error probabilities, the detection threshold is $\theta = 17.8$ dB as in [45], resulting in $p_f = 8.3 \times 10^{-3}$ and $p_n = 1.4 \times 10^{-16}$ for $\Delta = 6$.

At each traffic scenario, we show the configurations of the Δ and k parameters for two $n_{s,max}$ values: 3 and 6. With BR, the amount of reserved channels is $m = n_{s,max}$. The values of the safety parameter k range from 0 to 3 and the scanned channels Δ range from $n_{s,max} + k$ to 10. We obtain, for each ($n_{s,max}$, k, Δ) tuple, the expected PU and SU normalized capacities with and without BR ($\overline{R}_{PU}^{nbr}$, $\overline{R}_{SU}^{nbr}$, $\overline{R}_{PU}^{br}$, $\overline{R}_{SU}^{br}$) defined in previous section. We are also interested in evaluating PU performance in terms of collision probability and average interference power at PU receivers, two usual metrics in previous works. In particular, the average interference, $\bar{R}_I^{nbr}$, $\bar{R}_I^{br}$, can be computed using the following per-state rewards in (3.12):

$$
\begin{aligned}
R_I^{nbr}(n_p, n_s) &= \frac{1}{n_p} \sum_{i=c_{min}}^{c_{max}} i \frac{p_{su} K' \mathbf{E}[g(r')]}{r'} \frac{}{n_s} P_H(i; N, n_s, n_p) \\
R_I^{br}(n_p, n_s) &= \frac{1}{n_p} \frac{p_{su} K' \mathbf{E}[g(r')]}{r'} \frac{}{n_s} (n_s + n_p - N)^+
\end{aligned}
\tag{3.40}
$$

The average collision probability ($\bar{R}_c^{nbr}$, $\bar{R}_c^{br}$) can be obtained from (3.40) by removing the signal power factor.

Figure 3.4 shows the joint SU-PU performance for the low PU traffic scenario. Squares correspond to no BR (NBR), and circles to BR performance values. Red and green points correspond to $n_{s,max} = 3$ and 6, respectively. Larger point sizes indicate larger k. Increasing Δ allows the discovery of more spectrum opportunities, implying more SU capacity, but less PU capacity (especially without BR) due to a
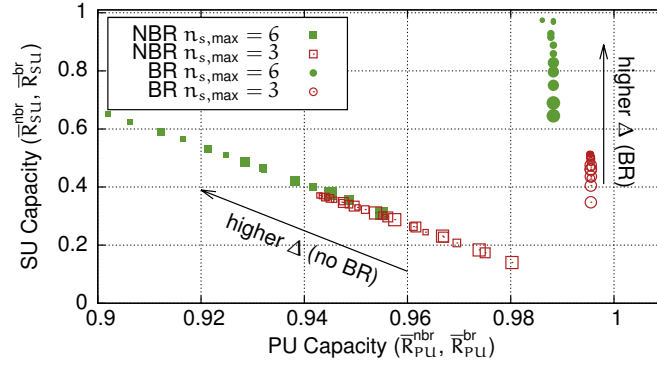
Figure 3.4: Normalized PU and SU capacities for a low PU traffic scenario $(\lambda_p = 0.1)$. Larger points denote higher k values.
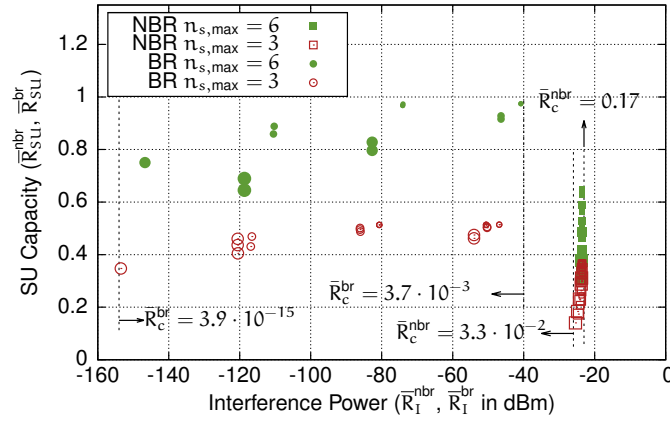

Figure 3.5: SU capacity versus interference caused to the PUs for low PU traffic. Larger points denote higher k values. Collision probabilities at the extreme points are indicated.

higher collision probability. The protection parameter k causes the opposite effect: higher k increases the protection level to PUs (higher $\overline{R}_{PU}^{nbr}$) and reduces $\overline{R}_{SU}^{nbr}$. In consequence, the best PU performance, when no BR is used, is attained when $\Delta = n_{s,max} + k$. When BR is used, it can be seen that, for each $n_{s,max}$, all the $\overline{R}_{PU}^{br}$ values are higher than $\overline{R}_{PU}^{nbr}$. Figure 3.5, showing the average interference at the PUs vs. SU capacity, illustrates the reason of the differences between using BR or not: the superior performance of BR in terms of collision/interference, which compensates the negative effect of having fewer channels to allocate at the AN. However, although BR always manages to reduce collision/interference, it is not always enough to compensate the drawbacks of BR in terms of PU achievable rate. In particular, when the PU traffic is high, there are values of $\overline{R}_{PU}^{nbr}$ than cannot be achieved with BR, as can be seen in Figure 3.6. Note, however, that the PU capacity varies in a very short range. Nevertheless, Figure 3.7 shows that BR also reduces the interference in the high traffic regime.
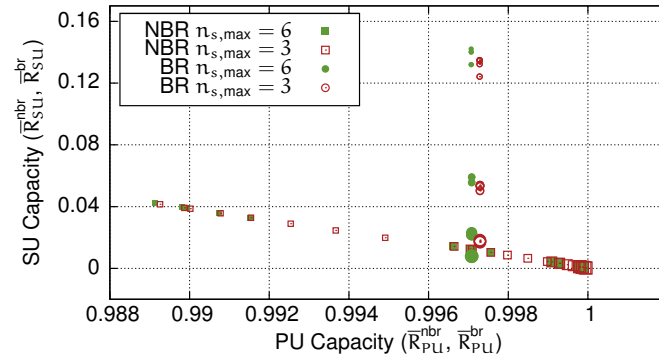
Figure 3.6: SU and PU capacities for a high PU traffic scenario ($\lambda_p = 0.5$). Larger points denote higher k values.
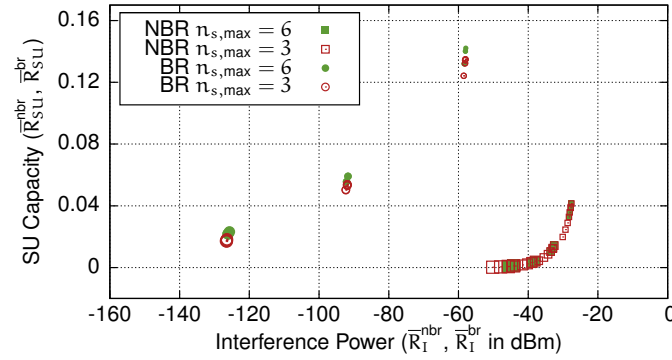

Figure 3.7: SU capacity versus interference caused to the PUs for high PU traffic.

Studying PU motivation to use BR

We have seen that the PU traffic intensity is closely related to the performance limits on $\overline{R}_{PU}^{br}$. This suggests that an AN implementing BR could decide to switch this feature on or off, depending on its own traffic load. To see where the traffic threshold could be located, we represent, in Figure 3.8, the highest values of $\overline{R}_{PU}^{br}$ and $\overline{R}_{PU}^{nbr}$ for $n_{s,max} = 4$. To assess the effect of SU traffic as well, several ρ_s values are used $(0.1, 0.5, 0.8)$. It can be seen that, without BR, the PU capacity is higher for larger λ_p, because the spectrum sensing detects more PU activity and incoming SUs either do not transmit or transmit using very few channels, reducing the effects of interference/collision. We can obtain an estimation of the traffic thresholds without resorting to the Markov model by means of approximate expressions for $\overline{R}_{PU}^{nbr}$ and $\overline{R}_{PU}^{br}$. Assuming that the spectrum contains an SU transmission with probability ρ_s, we can compute the following average:

$$\tilde{R}_{PU}^* = C_N^* \left(\left\lfloor \tilde{n}_p \right\rceil, \left\lfloor \tilde{n}_s \right\rceil \right) \rho_s + C_N^* \left(\left\lfloor \tilde{n}_p \right\rceil \right) (1 - \rho_s) \tag{3.41}$$
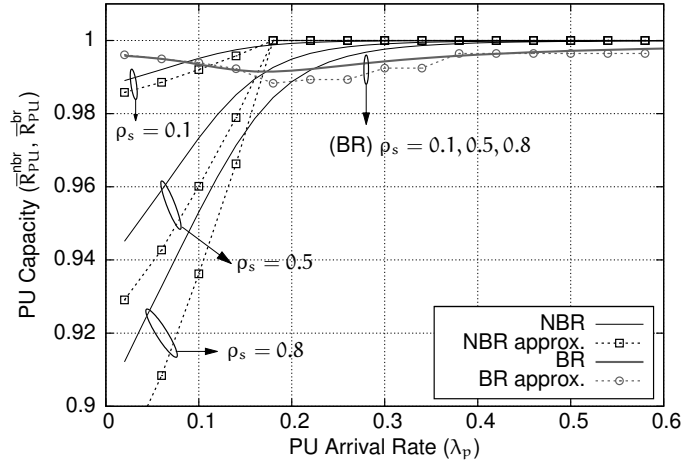
Figure 3.8: Effect of PU and SU traffic intensities on PU capacity. Approximate values $\tilde{R}_{PU}^{nbr}$, $\tilde{R}_{PU}^{br}$ are represented with points and dotted lines. For each SU traffic intensity there is a PU traffic threshold value below which the PU would prefer to activate BR.

where $* =$ nbr or br, $\lfloor x \rceil$ denotes the closest integer to x, and $\tilde{n}_p$, $\tilde{n}_s$ are estimates of the average values of n_p and n_s. We have that $\tilde{n}_p = \sum_{i=0}^{N} i\pi_i$, and $\tilde{n}_s = \sum_{i=0}^{N} \theta_*(i)\pi_i$, where

$$
\begin{aligned}
\theta_{nbr}(n_p) &= \min\left(\frac{N-n_p}{N}\Delta - k, n_{s,max}\right) \\
\theta_{br}(n_p) &= \min\left(N - n_p - k, n_{s,max}\right)
\end{aligned}
\tag{3.42}
$$

and $\frac{N-n_p}{N}\Delta$ is the expected number of available channels in Δ. The approximate capacity values obtained with this approximation are also shown in Figure 3.8. Finally, Figure 3.9 depicts the interference power at the PU receivers for different PU and SU traffic regimes. When the traffic increases, the BR average interference power converges to the no BR case, because the reserved channels are more frequently used by PU transmissions resulting in higher collision probability.

3.6   Conclusions

Even with the sensing capabilities of cognitive radios, OSA is always associated to some degree of transmission overlapping with PU transmissions and therefore to harmful interference at PU receivers. This chapter studies the idea of a primary network reserving some part of its spectrum for OSA to reduce harmful interference from SUs. The reserved channels are only used for PU transmissions when the rest of the channels overflow. We have developed a Markov-reward model for evaluating the PU and SU capacities in coexistence scenarios with and without bandwidth reservation. The numerical results showed that using BR implies moderate PU capacity reduction, e.g. reserving 4 out of 15 channels, resulted in a PU capacity reduction of 1% in the worst case. If the PU traffic intensity is below a certain value,
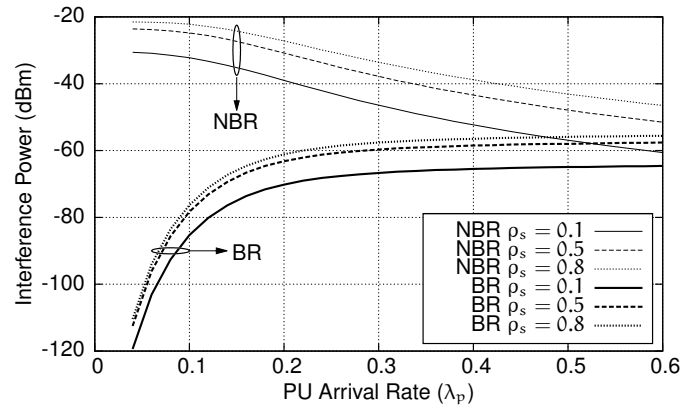
Figure 3.9: Effect of PU and SU traffic intensities on the interference at PUs.

the PU capacity obtained with BR is higher than the capacity achieved without BR, because of the better interference (and collision probability) performance of BR. The AN is therefore motivated to activate BR under low traffic regimes. For the SUs there is a remarkable performance improvement when using the reserved spectrum, resulting in a higher spectrum efficiency. Our model is applicable to the design of BR activation policies, using either the Markov framework or a simpler approximate formulation.

Uncertainty-Aware Opportunistic Spectrum Access in Coexistence-Friendly Systems

4.1 Introduction

In our work on PU bandwidth reservation, described in the previous chapter, we have shown that the benefits of reduced PU-SU collisions outweigh the negative consequences of a more limited selection of channels available to the PUs. In this chapter we develop an OSA mechanism further exploiting that finding. We also continue improving MAC protocols for SUs with hardware-constrained cognitive radios, a sub-objective we began to pursue in chapter 2. More specifically, we address an issue that has not received much attention in multi-channel OSA protocols: the increase in the uncertainty of channel occupancies with the elapsed time from the observations.

Motivation

Let us consider a system where a hardware-constrained cognitive transmitter intends to transfer some data to a cognitive receiver using one or several narrowband channels within certain licensed spectrum band (*multi-channel* access). Specifically, the hardware constraints imply that: (1) fine sensing can only be conducted within a small portion of spectrum (sensing constraint) at a time; (2) the spectrum used by an SU has a limited bandwidth (transmission constraint) and (3) SUs are equipped with a single cognitive radio which cannot sense and transmit simultaneously. After a contention phase, these SUs (cognitive pair) start to scan the spectrum in consecutive periods (scanning slots) in order to detect which channels are free of PU activity and thus available for transmission. During each scanning slot, the cognitive pair (CP) senses only one channel and the SUs exchange messages informing each other about the observation outcome. Therefore, each scanning slot introduces a time overhead that needs to be considered in the OSA protocol de-

sign. After each scanning slot, the cognitive pair has to decide whether to continue scanning or to transmit (and how many channels to transmit in).

Optimizing this process is challenging because of the trade-offs involved. On the one hand, sensing more channels allows the SU to find more available channels to transmit in, potentially achieving higher throughput. But on the other hand, the scanning time overhead increases as more channels are sensed, which is contrary to the throughput goal. In addition, *using many scanning slots implies more uncertainty about the information obtained in earlier slots*: the channel occupancy of earlier sensed channels may have changed before the SU transmission (*scanning delay*). In a real system, where sensing results are not fully reliable, this uncertainty is even higher. To the best of our knowledge, this last aspect has never been included in a sequential sensing mechanism with multichannel access, and is the first key issue addressed in this chapter.

The second key issue is *to evaluate the impact of the channel allocation scheme for PUs on OSA performance*. If each channel is randomly occupied by a PU, both the collision probability and the post-scanning uncertainty variation over time is the same for every channel. However, the licensed operator could allocate its channels to the PUs with a different pattern, more friendly to the coexistence with OSA nodes. Specifically, we show that a simple sequential channel assignment for PUs reduces collisions with SUs. In addition, if the licensed network has the ability to aggregate channels with ongoing PU sessions into a compacted spectrum block (*spectrum merging*), the reduction is even greater. It is easy to see that SUs' throughput will be dramatically improved. On the other hand, for primary systems that can dynamically change the channels assigned to ongoing PU transmissions, using these policies reduces the ability of the base station to exploit the channels with best instantaneous propagation conditions. In those cases, the negative impact in PUs' throughput is compensated by the effect of the reduced collisions, under certain traffic intensities, as shown in the previous chapter and is also highlighted in [76]. For other scenarios, the reduction in PUs' throughput due to spectrum merging is minimal.

Obviously, this OSA framework requires the SU nodes to have a previous knowledge of PU traffic descriptors. This is a typical requirement for OSA protocols, and multiple previous works offer solutions to this problem. However, traffic estimations are generally imperfect. Therefore, we also evaluate the impact of the difference between estimated and real PU traffic parameters. Our findings show that this difference has a surprisingly low impact on PU performance.

Related Work

MAC designs for hardware-constrained radios were introduced in [63]. However its authors focused only on optimizing SU's throughput and did not provide a method to compute the actual overlapping time which is one of the objectives of our model. In addition, it uses a worst case approach, setting a fixed duration of T seconds for SU transmissions, where T is the maximum overlap time with SU activity that a PU receiver could tolerate.

Several works consider multichannel access with sensing constraint [64, 65, 66]. The proposal in [64] optimizes the discovery of spectrum opportunities. In contrast to our model, the system in [64] is assumed to be collision-free. In [65], the focus is on optimizing the sensing phase considering, among other aspects, the trade-off between its duration (and reliability) and collision probability. As we will explain later, our model incorporates similar considerations, but in the design of access policies that include not only the sensing phase but also the number of channels to transmit in. Finally, the approach in [66] is perhaps the closest to ours, although there are two significant differences: first, [66] aimed to optimize SU's throughput without quantifying the effect on PU's performance. Second, our formulation incorporates the increase in the uncertainty of past channel observations over time, which is crucial for evaluating the effect of PU channel allocation schemes.

Other works consider the uncertainty increase over time, but only across different sensing episodes, not during the ongoing one. Those works propose models with either single channel sensing and immediate transmission [77, 78], or with a fixed *a priori* set of channels to sense [69, 79, 80]. In contrast, we aim at an optimal multichannel access strategy, which implies that the number of channels scanned on any sensing episode must depend on the sequence of observations received during this episode.

The previous chapter and this work are in line with the recent increase of interest in non-transparent access of SUs to PUs' spectrum [81, 82]. This trend is in contrast to the traditional approach maintained since the origin of cognitive radio [83], specially under the spectrum interweave OSA model [69]. Such change of paradigm is probably motivated by the application of OSA techniques to heterogeneous cellular networks (HetNets). HetNets also imply overlay networks (small cells, often with cognitive abilities, over the macro-cell network), but a substantial difference with classical interweave OSA: that all the HetNet users are PUs and thus, QoS must be guaranteed for both the macro and the small cell users. Because of this, in an HetNet scenario, it is intuitively convenient for the operator to be coexistence-friendly with its associated small-cells, even though they implement cognitive abilities. While some works [76, 84] have studied channel assignment policies for HetNets, the design of optimal multichannel OSA strategies in coexistence-friendly environments is still an open issue.

Our Contribution

This chapter comprises the following main contributions:

1. We address an issue that has been largely ignored in the design of multichannel OSA protocols for hardware constrained radios: considering the effect of the scanning delay in the uncertainty of the information available at the SUs. Finding an optimal OSA strategy implies solving a Partially Observable Markov Decision Process (POMDP) with finite horizon. This is a challenging task because of the inherent complexity of POMDPs, which is increased by the inclusion of delay-dependent uncertainty.

2. The OSA scheme is designed and evaluated for different levels of *coexistence friendliness* at the spectrum owner, reflected on its spectrum management policies. Specifically: random allocation, sequential allocation and sequential allocation with spectrum merging. We also show how the selected spectrum management policy relates to the delay effects considered in the OSA design.

4.2   System Description

Licensed Network

The system under study consists of two overlay wireless networks. The licensed one operates with centralized access coordinated by an access node (AN). The spectrum used by this network comprises N adjacent narrowband channels. Upon each PU arrival, the AN assigns an available channel to the incoming PU. SU activity is not detected by either the PUs or the AN, therefore a PU may occupy a channel with an ongoing SU transmission. Similarly to many related works (*e.g.* [69], [64], [73]), PU traffic is assumed to follow a Poisson model: PU inter-arrival time characterized by an exponential random variable with rate λ and channel holding time described by an exponential random variable with rate μ.

The criteria for PU channel allocation depend on the system, *e.g.* the AN may assign them according to channel quality measurements reported by the PUs. So far, in OSA research it has been usually assumed that channels are occupied by PUs according to independent identically distributed (i.i.d.) processes. In that scheme, that we refer to as *random allocation*, every channel can be assigned to an incoming PU with the same probability. However, when the same spectrum band is accessed by opportunistic nodes, random allocation is not optimal in terms of collision probability. In this work we investigate the benefits of using different spectrum management strategies for allocating channels to PUs. Therefore, we consider two additional schemes:

1. *Sequential allocation*: Considering the N channels numbered from 1 to N, each incoming PU is always assigned the highest numbered available channel.

2. *Sequential allocation with spectrum aggregation*: channels are assigned to PUs sequentially, es explained before, but when one PU releases its channel, all the spectrum currently in use by ongoing PU communications is rearranged so that all the active PUs occupy the adjacent channels with the highest numbers (merged spectrum).

Fig. 4.1 illustrates the arrival and departures of PUs under these allocation schemes. Many wireless access systems, *e.g.* OFDMA-based systems, allow both sequential allocation of channels (subcarriers) and channel rearrangement for ongoing communications. However, in some cases, sequential allocation may imply assigning channels whose propagation conditions are not optimal for the incoming PUs. For those cases, we assume that the AN uses the bandwidth reservation scheme of the previous chapter, which would be equivalent to spectrum merging. In that chapter it was concluded that the benefits of easing secondary access overcome its draw-
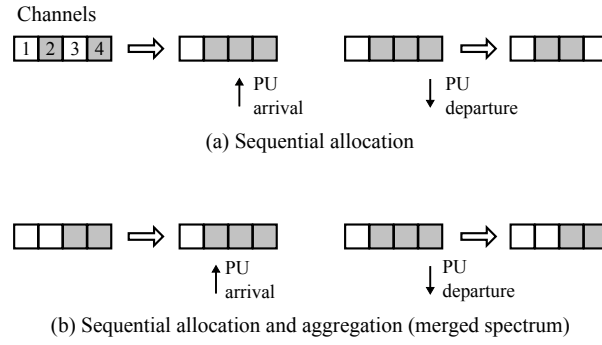
Channels

(a) Sequential allocation

(b) Sequential allocation and aggregation (merged spectrum)

Figure 4.1: Channel allocation schemes for PUs.

backs when the primary traffic intensity is below certain threshold. Therefore, the use of a coexistence-friendly spectrum management scheme implies the implicit assumption that the PU traffic intensity is below that threshold, which is reasonable since OSA is conceived to increase the efficiency of underused spectrum of, e.g. legacy networks.

Secondary Network

The unlicensed or secondary network operates in a decentralized, ad-hoc fashion. Communication is always performed between pairs of SUs (*cognitive pairs*) consisting of one sender and one receiver. Every SU is assumed to be under the coverage area of the same licensed access node, thus the SU's objective is to transmit data using one or several of the N licensed channels causing the less possible interference to PU communications. Secondary nodes are constrained to some hardware limitations: (1) Each SU is equipped with a single radio that can either transmit or receive, but not at the same time. (2) When sensing the spectrum to detect PU activity, an SU can only sense one of the N channels. (3) Once a cognitive pair decides to start a transmission, it can use up to M noncontiguous channels simultaneously.

The SUs implement the MAC protocol described in [63] (HC-MAC) with a different sensing and accessing decision policy. Summarizing, HC-MAC comprises three consecutive phases: *contention*, *sensing* and *transmission*. The contention procedure allows a pair of SUs (sender and receiver) to reserve the use of the spectrum in a certain area avoiding collisions with other SU transmissions. After winning a contention, the cognitive pair starts to sense the spectrum in fixed-duration sensing slots. During the τ seconds of each slot both the sender and the receiver sense the same channel and exchange the sensing result. At the end of every slot, a decision is made whether to sense the next channel or to transmit using one or more of the available channels detected. Because of the information exchange, the cognitive pair acts as a single entity when making the transmission decision provided that both nodes operate under the same policy. Similarly to [63] and [59], in case of overlapping with PU activity, the SU transmission is considered to be successful, *i.e.,* it is assmmed that the received power at the SU receiver from
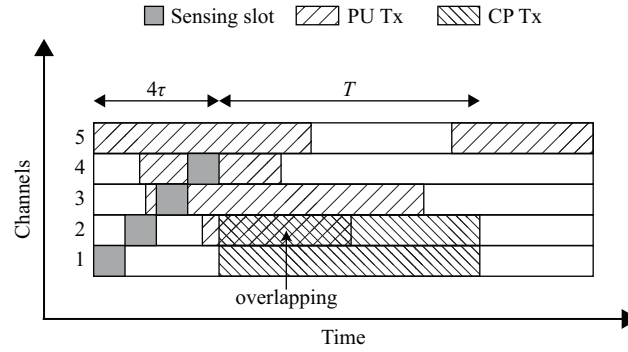
Figure 4.2: OSA example. After sensing channels 1 to 4, the CP decides to transmit in channels 1 and 2, resulting in an overlapping event on channel 2.

the corresponding SU transmitter is much higher than the interference from PUs. Nevertheless, since the model also estimates the expected overlapping time at each state, it could be readily extended to the case in which the SU transmissions can be harmed by PU interference.

The OSA policy considered in this work differs from the one in HC-MAC in its decision space and its computation procedure. We use an extended decision space such that instead of using the total number of available channels, the sender may use any smaller number, including quitting the sensing process without transmitting. The differences in policy computation result from a more general formulation incorporating aforementioned aspects, namely: observation uncertainty due to imperfect sensing and delay overhead; modeling PU activity on the whole spectrum band with diverse channel allocation strategies; and PU's QoS characterization in terms of overlapping time. Fig. 4.2 depicts a simplified example of SU access incurring in an overlapping event.

4.3   POMDP Formulation

In this section we formulate the opportunistic access problem as a POMDP with finite horizon and obtain an equivalent MDP by means of the *state augmentation* approach. The formulation in this section uses a generic objective, and therefore it is applicable to any optimization criterion, e.g. increasing SU throughput, capacity, etc. Section 4.4 describes the computation of a specific cost function balancing SU throughput and overlapping time.

The Spectrum Process

Considering only PU activity, each of the N channels can be in two states: 1, when the channel is assigned to one PU, and 0 when the channel is unassigned. Let $Y(t)$ denote the aggregate state of the N channels at time t. $Y(t)$ is a continuous-time Markov chain (CTMC) that we refer to as the *spectrum process*. Each channel allocation scheme results in a different stochastic behavior for $Y(t)$. The set of

	$\mathcal{S}$	q_{ij}	
Random	$\{0, 1, \ldots, 2^N\}$	$\begin{cases} \frac{\lambda}{\left\|(i)_2\right\|_1}, & \text{if} \quad \sum_{n=1}^N (j_n - i_n) = 1 \\ \mu, & \text{if} \quad \sum_{n=1}^N (i_n - j_n) = 1 \\ -\lambda - \left\|(i)_2\right\|_1 \mu, & \text{if} \quad i = j \\ 0, & \text{otherwise} \end{cases}$	
Sequential	$\{0, 1, \ldots, 2^N\}$	$\begin{cases} \lambda, & \text{if} \quad j = i + 1 \\ \mu, & \text{if} \quad \sum_{n=1}^N (i_n - j_n) = 1 \\ -\lambda - \left\|(i)_2\right\|_1 \mu, & \text{if} \quad i = j \\ 0, & \text{otherwise} \end{cases}$	
Merged	$\{0, 1, \ldots, N\}$	$\begin{cases} \lambda, & \text{if} \quad j = i + 1 \\ i\mu, & \text{if} \quad j = i - 1 \\ -\lambda - i\mu, & \text{if} \quad i = j \\ 0, & \text{otherwise} \end{cases}$	

Table 4.1: State space $\mathcal{S}$ and components of the infinitesimal generator $\mathbf{Q}$ characterizing $Y(t)$ for random channel allocation, sequential channel allocation, and sequential allocation with spectrum merging.

(countable) values that $Y(t)$ can take, is denoted by $\mathcal{S}$. Let $i, j \in \mathcal{S}$ be two states of $Y(t)$ and let $(i)_2$ denote the binary representation of i as an N-tuple, *i.e.* $(i)_2 = (i_1, \ldots, i_N)$, where $i_k = 1$ if the k-th channel is occupied by a primary user and $i_k = 0$ otherwise. We will also refer to i_k as the state of channel k. The complementary of $(i)_2$ is denoted by $\overline{(i)_2}$. The 1-norm of $(i)_2$ is defined as

$$\left\|(i)_2\right\|_1 = \sum_{k=1}^N |i_k| \tag{4.1}$$

Note that previous expression provides the number of occupied channels in the system while $\left\|\overline{(i)_2}\right\|_1$ yields the number of free channels.

Table 4.1 summarizes, for each channel allocation scheme, the set $\mathcal{S}$, and the elements q_{ij} of infinitesimal generator, $\mathbf{Q}$, of $Y(t)$.

Partial Observations and the Information Vector

The cognitive pair *partially observes* $Y(t)$ in consecutive periods (scanning slots) of τ seconds. In particular, during the k-th scanning slot, only the k-th channel of the spectrum is sensed. At the end of the sensing slot, the CP obtains the sensing result z_k: $z_k = 1$ if the k-th channel is estimated to be occupied, or $z_k = 0$ otherwise. Therefore, z_k is a partial (and possibly unreliable) observation of $Y(t)$. Hence, we face a POMDP problem, which implies that decisions should be based

Notation	Definition
N	number of channels in the spectrum band
M	maximum usable channels by the CP
k	scanning slot (decision stage)
τ	duration of a scanning slot
T	duration of a CP transmission
R	data rate of a single channel
$Y(t)$	state of the spectrum process at time t
Y_k	value of $Y(t)$ at the end of k-th scanning slot
$\mathcal{S}$	state space of $Y(t)$ (and Y_k)
$\mathbf{Q}$	infinitesimal generator of $Y(t)$
q_{ij}	elements of $\mathbf{Q}$
λ	arrival rate for PU population
μ	PU departure rate
$\mathbf{P}$	transition probability matrix for Y_k
p_{ij}	elements of $\mathbf{P}$
Z_k	sensing outcome on channel (and slot) k
X_k	state (information vector) of the CP system
$\mathcal{S}(k, s_k)$	states in $\mathcal{S}$ such that channel k is in state s_k
$\mathcal{X}_k$	state space of the information vector
$p_{\bar{p}}$	probability of false positive detection
$p_{\bar{n}}$	probability of false negative detection
u_k	scanning decision at stage k
m_k	selected channels at stage k
$d_k = (u_k, m_k)$	CP decision at stage k
$\mathcal{D}_k(x_k)$	control space for d_k
$\mu_k(x_k)$	function mapping CP states to decisions
$g_k(x_k, d_k)$	per-stage cost of the CP system
η	admissible policy for the CP system
J_η	cost function for policy η
J^*	optimal cost for the CP system
$p_k(*\vert*)$	transition probability of the CP system
$\boldsymbol{\pi}_{x_k}$	*a posteriori* estimation of Y_k given x_k
$\hat{\boldsymbol{\pi}}_{x_k}$	*a priori* estimation of Y_{k+1} given x_k
b_{max}	maximum achievable throughput by the CP
$b_k(m_k)$	throughput attained with decision m_k
c_{max}	maximum overlapping time
$c_k(x_k, m_k)$	expected overlapping time given x_k and m_k
$l_{i,x_k}(t)$	expected time of $Y(t)$ in i given x_k

Table 4.2: Summary of notation

on *imperfect state information*. Moreover, note that the observations z_k obtained in consecutive scanning slots $k = 1, 2, \ldots$, correspond to samples of $Y(t)$ taken in different instants. Let us formalize these aspects and their implications.

Let us define the *observation process* Z_k taking values from the observation space $\{0, 1\}$. We assume that the channel activity detection depends on the values of the spectrum process at the end of each slot. Therefore we define the discrete-time Markov chain (DTMC)

$$Y_k = Y(k\tau), \quad \text{for } k = 1, \ldots, N. \tag{4.2}$$

The observations are originated, in general as

$$Z_k = h_k(Y_k), \quad \text{for } k = 1, \ldots, N. \tag{4.3}$$

Assuming perfect detection and $Y_k = i$, with binary representation $(i)_2 = (i_1, \ldots, i_N)$, the function on the right hand side has the following form:

$$h_k(i) = i_k, \quad \text{for } k = 1, \ldots, N. \tag{4.4}$$

In case the detection process is subject to errors such as false positive and false negative detections characterized by the probabilities $p_{\bar{p}}$ and $p_{\bar{n}}$ respectively, the function $h_k(i)$ provides random values in the observation space according to the following conditional distribution:

$$\begin{aligned}
P(h_k(i) = 1 | i_k = 1) &= 1 - p_{\bar{n}} \\
P(h_k(i) = 0 | i_k = 1) &= p_{\bar{n}} \\
P(h_k(i) = 1 | i_k = 0) &= p_{\bar{p}} \\
P(h_k(i) = 0 | i_k = 0) &= 1 - p_{\bar{p}}
\end{aligned} \tag{4.5}$$

for $k = 1, \ldots, N$.

The observation process Z_k is a DTMC. To show this fact, let us define the set $\mathcal{S}(k, s_k)$ containing the states in S such that channel k is in state s_k:

$$\mathcal{S}(k, s_k) = \{i \in \mathcal{S} | i_k = s_k\} \tag{4.6}$$

For the case of perfect detection, the probability distribution of Z_k is given by

$$P(Z_k = z_k) = P(Y_k \in \mathcal{S}(k, z_k)) \tag{4.7}$$

with z_k representing a particular sensing result on channel k. For the case of imperfect detection, Z_k is given by

$$\begin{aligned}
P(Z_k = 1) &= P(Y_k \in \mathcal{S}(k, 1))(1 - p_{\bar{n}}) + P(Y_k \in \mathcal{S}(k, 0))\, p_{\bar{p}} \\
P(Z_k = 0) &= P(Y_k \in \mathcal{S}(k, 0))(1 - p_{\bar{p}}) + P(Y_k \in \mathcal{S}(k, 1))\, p_{\bar{n}}
\end{aligned} \tag{4.8}$$

which can be written more compactly as

$$P(Z_k = z_k) = P(Y_k \in \mathcal{S}(k, z_k))\, p_c(z_k) + P(Y_k \in \mathcal{S}^c(k, z_k))\, p_i(z_k) \tag{4.9}$$

where $\mathcal{S}^c(k, z_k)$ denotes the complementary of $\mathcal{S}(k, z_k)$, and the functions $p_c(z_k)$ and $p_i(z_k)$ represent the probabilities of correct and incorrect detection respectively, and have the following expressions

$$\begin{aligned} p_c(z_k) &= 1 - z_k p_{\bar{n}} - (1 - z_k) \, p_{\bar{p}} \\ p_i(z_k) &= 1 - z_k p_{\bar{p}} - (1 - z_k) \, p_{\bar{n}} \end{aligned} \tag{4.10}$$

Then, since Y_k is a DTMC, Z_k is also a DTMC.

Let X_k denote the amount of information available to the cognitive pair at the end of slot k and call it the *information vector*. Assuming an initial information X_0 (*e.g.* $X_0 = 0$ in case of no initial information at the beginning of the sensing process) the information vector is generated iteratively as

$$X_k = I(Z_k, X_{k-1}), \quad \text{for } k = 1, \ldots, N, \tag{4.11}$$

where $I(Z_k, X_{k-1})$ is defined as

$$I(Z_k, X_{k-1}) = \begin{cases} X_{k-1}, & \text{if } Z_k = 1 \\ (k, X_{k-1}), & \text{if } Z_k = 0 \end{cases} \tag{4.12}$$

Therefore, the information vector X_k contains the position of the channels that have been observed idle up to stage k. Let $\mathcal{X}_k$ be the set of possible values for X_k. Because the cognitive pair must make the k-th decision with all the information available and not only with the k-th observation, Z_k, it can use the information vector X_k as the state for the system describing the cognitive pair (*CP system*). Following this approach, known as *state augmentation*, the imperfect state information problem casts into a fully observable state information one, as explained in [24]. The state space of the CP system is then $\mathcal{X}_k$, in which we must include a special *termination* state, denoted by F, corresponding to the data transmission from the cognitive transmitter, or to the transmitter quitting the medium without transmitting. In Markov modeling terms, the termination state F is an absorbing state.

Decisions and Policies

Based on $x_k \in \mathcal{X}_k$, the access algorithm selects a decision $d_k = (u_k, m_k)$,[1] where $u_k = 1$ if the decision is to continue scanning and $u_k = 0$ otherwise. The second component, m_k, contains the channels selected to transmit in. The amount of selected channels is denoted by $n \leq M$. Note that, if $u_k = 0$ and $m_k = 0$, the decision made is to quit the medium access process after scanning the k-th channel without transmitting. The decision variables u_k and m_k take values from the sets $\mathcal{U}_k$ and $\mathcal{M}_k(x_k)$, where

$$\mathcal{U}_k = \begin{cases} \{0, 1\}, & \text{if } 0 \leq k \leq N - 1 \\ \{0\}, & \text{if } k = N \end{cases} \tag{4.13}$$

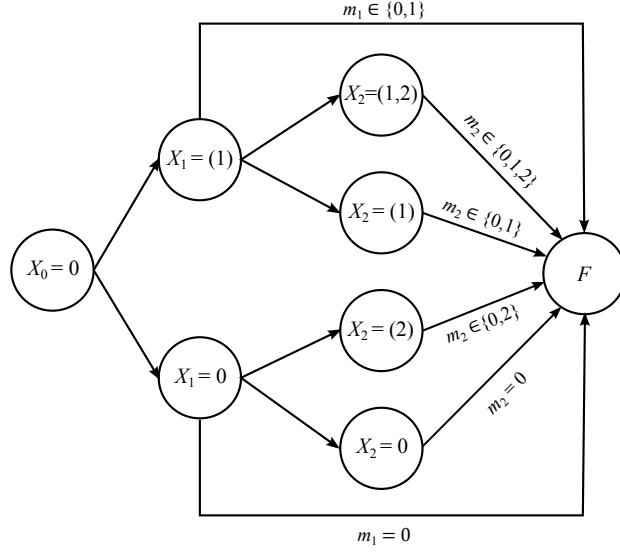[1] For $x_k = F$ there is no decision to be made.

Figure 4.3: States and transitions of the MDP for a two-channel spectrum. Arrows ending at the termination state F denote transitions associated to $u_k = 0$. The sets $M_k(x_k)$ for these transitions are explicitly stated.

and $\mathcal{M}_k(x_k)$ is the set containing all the partitions of x_k. For compactness we define $\mathcal{D}_k(x_k) = \mathcal{U}_k \times \mathcal{M}_k(x_k)$.

A *policy* η is a mapping providing the decision $d_k \in \mathcal{D}_k(x_k)$ to be applied given the information $x_k \in \mathcal{X}_k$ available at time k. That is, $\eta = \{\mu_0, \mu_1, \ldots, \mu_N\}$, where each function $\mu_k(x_k) : \mathcal{X}_k \to \mathcal{D}_k(x_k)$ for $k = 0, \ldots, N$. As an example, Fig. 4.3 illustrates the diagram of a small system with only two channels.

The Equivalent MDP Problem

At the k-th stage each pair state-decision is associated to a *per-stage cost*, denoted by $g_k(x_k, d_k)$ which accumulates over time. For the special case that $x_k = F$ (termination state) the per-stage cost equals 0, therefore we can define an auxiliary function $\tilde{g}_k(x_k, d_k)$, such that

$$g_k(x_k, d_k) = \begin{cases} 0, & \text{if} \quad x_k = F \\ \tilde{g}_k(x_k, d_k), & \text{if} \quad x_k \in \mathcal{X}_k \end{cases} \tag{4.14}$$

Because the termination state is absorbing, no additional cost is accumulated when the CP system reaches it. In our context, $\tilde{g}_k(x_k, d_k)$ may estimate the overlapping time, the throughput or a combination of both. The formulation of the objectives and therefore the computation of the per-stage cost is addressed in Section 4.4.

The objective of the finite-horizon MDP is to find a policy $\eta = \{\mu_0, \mu_1, \ldots, \mu_N\}$ that minimizes the *cost function*

$$J_\eta = \mathop{\mathbb{E}}_{X_0^*, \ldots, X_N^*} \left\{ \sum_{k=0}^{N} g_k(X_k^*, \mu_k(X_k^*)) \right\} \tag{4.15}$$

In compact notation the MDP problem consists of finding the optimal cost J^*

$$J^* = \min_\eta J_\eta \tag{4.16}$$

The computation of the expectation on the right hand side of (4.15) relies on the transition probabilities $p_k(x_{k+1}|x_k, d_k)$, providing the probability of making a transition from state x_k in time k to state x_{k+1} in time $k+1$ when the decision is $d_k = (u_k, \cdot)$.[2]

Computing the Transition Probabilities

The transition probabilities $p_k(x_{k+1}|x_k, u_k)$ for every $x_k, x_{k+1} \in \mathcal{X}_k$, $u_k \in \mathcal{U}_k$, $k = 0, \ldots, N-1$, are:

$$p_k(x_{k+1}|x_k, u_k) = \begin{cases} 1, & \text{if } x_k = F \text{ and } x_{k+1} = F \\ 0, & \text{if } x_k = F \text{ and } x_{k+1} \neq F \\ (1 - u_k), & \text{if } x_k \neq F, x_{k+1} = F \\ u_k P(X_{k+1} = x_{k+1}|x_k), & \text{otherwise} \end{cases} \tag{4.17}$$

where $P(X_{k+1} = x_{k+1}|x_k)$ denotes the probability that the information vector equals x_{k+1} at stage $k+1$ conditioned on its value x_k at previous stage. Let us describe how to compute $P(X_{k+1} = x_{k+1}|x_k)$.

Recall that each state x_k depends on the observed process Y_k. At the initial state x_0, where no observation is still available, Y_0 is a random variable, characterized by a probability distribution that we will denote by $\boldsymbol{\pi}_0$. Because the spectrum process $Y(t)$ is active long before secondary users attempt to access the spectrum, the probability distribution $\boldsymbol{\pi}_0$ can be assumed to be the steady-state probability distribution (*probability vector*) of Y_k. The transition probability matrix of Y_k, $\mathbf{P}(t)$, is obtained from the Kolmogorov *forward* equation

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{P}(t)\mathbf{Q} \tag{4.18}$$

The elements $p_{ij}(t)$ of $\mathbf{P}(t)$ provide the transition probabilities between any pair of states, $i, j \in \mathcal{S}$ of $Y(t)$ when the process runs for t units of time. Since $Y_k = Y(k\tau)$, the transition matrix for Y_k is $\mathbf{P} = \mathbf{P}(\tau)$. The solution of (4.18) is $\mathbf{P}(t) = e^{t\mathbf{Q}}$, therefore:

$$\mathbf{P} = \mathbf{P}(\tau) = e^{\tau\mathbf{Q}} \tag{4.19}$$

The probability vector $\boldsymbol{\pi}_0$ is the unique solution to the equations below:

$$\boldsymbol{\pi}_0 = \boldsymbol{\pi}_0 \mathbf{P}, \quad \|\boldsymbol{\pi}_0\|_1 = 1 \tag{4.20}$$

Let us define two additional probability distributions, $\boldsymbol{\pi}_{x_k}$, and $\hat{\boldsymbol{\pi}}_{x_k}$, required to obtain $P(X_{k+1} = x_{k+1}|x_k)$. First, let $\boldsymbol{\pi}_{x_k}$ denote the probability distribution of Y_k given the information vector x_k:

$$\boldsymbol{\pi}_{x_k}(i) = P(Y_k = i|x_k), \quad \text{for } i \in \mathcal{S} \tag{4.21}$$

---

[2]Note that m_k does not affect the transition probabilities of x_k.

Second, let $\widehat{\boldsymbol{\pi}}_{x_k}$ denote the probability distribution of Y_{k+1} estimated when the available information is x_k:

$$\widehat{\boldsymbol{\pi}}_{x_k}(i) = P\left(Y_{k+1} = i | x_k\right), \quad \text{for } i \in \mathcal{S} \tag{4.22}$$

Essentially, $\widehat{\boldsymbol{\pi}}_{x_k}$ is an *a priori* estimate of the distribution of Y_{k+1}, *i.e.* before receiving the observation z_{k+1} at stage $k+1$. Once z_{k+1} is known, it is possible to compute the *a posteriori* estimate of Y_{k+1}, $\boldsymbol{\pi}_{x_{k+1}}$.

For the perfect detection case, the conditional probability $P\left(X_{k+1} = x_{k+1} | x_k\right)$ is related to $\widehat{\boldsymbol{\pi}}_{x_k}$ by the following equation:

$$\begin{aligned}
P\left(X_{k+1} = x_{k+1} | x_k\right) &= P\left(Z_{k+1} = z_{k+1} | x_k\right) \\
&= P\left(Y_{k+1} \in \mathcal{S}(k+1, z_{k+1}) | x_k\right) \\
&= \sum_{j \in \mathcal{S}(k+1, z_{k+1})} P\left(Y_{k+1} = j | x_k\right) \\
&= \sum_{j \in \mathcal{S}(k+1, z_{k+1})} \widehat{\boldsymbol{\pi}}_{x_k}(j)
\end{aligned} \tag{4.23}$$

For the imperfect detection case, the conditional probability $P\left(\boldsymbol{X}_k = x_k | x_{k-1}\right)$ is given by

$$\begin{aligned}
P\left(\boldsymbol{X}_k = x_k | x_{k-1}\right) &= P\left(Z_k = z_k | x_{k-1}\right) \\
&= P\left(Y_k \in \mathcal{S}(k, z_k) | x_{k-1}\right) p_c(z_k) + \\
&\quad P\left(Y_k \in \mathcal{S}^c(k, z_k) | x_{k-1}\right) p_i(z_k) \\
&= \sum_{j \in \mathcal{S}(k, z_k)} P\left(Y_k = j | x_{k-1}\right) p_c(z_k) + \\
&\quad \sum_{j \in \mathcal{S}^c(k, z_k)} P\left(Y_k = j | x_{k-1}\right) p_i(z_k) \\
&= \sum_{j \in \mathcal{S}(k, z_k)} \widehat{\boldsymbol{\pi}}_{x_{k-1}}(j) p_c(z_k) + \\
&\quad \sum_{j \in \mathcal{S}^c(k, z_k)} \widehat{\boldsymbol{\pi}}_{x_{k-1}}(j) p_i(z_k)
\end{aligned} \tag{4.24}$$

Now we specify how to compute the conditional distributions $\widehat{\boldsymbol{\pi}}_{x_k}$ and $\boldsymbol{\pi}_{x_k}$. The conditional distribution $\widehat{\boldsymbol{\pi}}_{x_k}$ can be obtained from the *a posteriori* estimate of previous stage $\boldsymbol{\pi}_{x_{k-1}}$

$$\begin{aligned}
\widehat{\boldsymbol{\pi}}_{x_k}(i) &= P\left(Y_{k+1} = i | x_k\right) \\
&= \sum_{j \in \mathcal{S}} P\left(Y_{k+1} = i | Y_k = j\right) P\left(Y_k = j | x_{k-1}\right) \\
&= \sum_{j \in \mathcal{S}} p_{ji} \boldsymbol{\pi}_{x_{k-1}}(j)
\end{aligned} \tag{4.25}$$

for $i \in \mathcal{S}$, where p_{ji} are the elements of the transition matrix P obtained in (4.19). Using vector notation we have $\widehat{\boldsymbol{\pi}}_{x_k} = \boldsymbol{\pi}_{x_k} P$.

Regarding $\boldsymbol{\pi}_{x_k}$, for the particular case of x_0 we have, by definition

$$\boldsymbol{\pi}_{x_0}(i) = P\left(Y_0 = i\right) = \boldsymbol{\pi}_0(i), \quad \text{for } i \in \mathcal{S} \tag{4.26}$$

and, for $k = 1, \ldots, N$ we have, for the perfect detection case:

$$
\begin{aligned}
\boldsymbol{\pi}_{x_k}(i) &= P\left(Y_k = i | x_k\right) \\
&= P\left(Y_k = i | z_k, x_{k-1}\right) \\
&= \frac{P\left(Y_k = i, Z_k = z_k | x_{k-1}\right)}{P\left(Z_k = z_k | x_{k-1}\right)} \\
&= \begin{cases} \dfrac{P\left(Y_k = i | x_{k-1}\right)}{P\left(Z_k = z_k | x_{k-1}\right)}, & \text{if } i \in \mathcal{S}(k, z_k) \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
\tag{4.27}
$$

where $P\left(Y_k = i | x_{k-1}\right) = \hat{\boldsymbol{\pi}}_{x_{k-1}}(i)$, and $P\left(Z_k = z_k | x_{k-1}\right)$ is given by (4.23). And for the imperfect detection case:

$$
\begin{aligned}
\boldsymbol{\pi}_{x_k}(i) &= P\left(Y_k = i | x_k\right) \\
&= P\left(Y_k = i | z_k, x_{k-1}\right) \\
&= \frac{P\left(Y_k = i, Z_k = z_k | x_{k-1}\right)}{P\left(Z_k = z_k | x_{k-1}\right)} \\
&= \begin{cases} \dfrac{P\left(Y_k = i | x_{k-1}\right) p_c(z_k)}{P\left(Z_k = z_k | x_{k-1}\right)}, & \text{if } i \in \mathcal{S}(k, z_k) \\ \dfrac{P\left(Y_k = i | x_{k-1}\right) p_i(z_k)}{P\left(Z_k = z_k | x_{k-1}\right)}, & \text{if } i \in \mathcal{S}^c(k, z_k) \end{cases}
\end{aligned}
\tag{4.28}
$$

where $P\left(Y_k = i | x_{k-1}\right) = \hat{\boldsymbol{\pi}}_{x_{k-1}}(i)$, and $P\left(Z_k = z_k | x_{k-1}\right)$ is given by (4.24). Summarizing, the conditional distributions $\hat{\boldsymbol{\pi}}_{x_k}$ and $\boldsymbol{\pi}_{x_k}$ can be recursively generated starting with

$$
\boldsymbol{\pi}_{x_0} = \boldsymbol{\pi}_0,
\tag{4.29}
$$

and using the following equations. For the perfect detection case:

$$
\begin{aligned}
\hat{\boldsymbol{\pi}}_{x_{k-1}} &= \boldsymbol{\pi}_{x_{k-1}} \mathbf{P}, \\
\boldsymbol{\pi}_{x_k}(i) &= \begin{cases} \dfrac{\hat{\boldsymbol{\pi}}_{x_{k-1}}(i)}{\sum_{j \in \mathcal{S}(k, z_k)} \hat{\boldsymbol{\pi}}_{x_{k-1}}(j)}, & \text{if } i \in \mathcal{S}(k, z_k) \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
\tag{4.30}
$$

for $k = 1, \ldots, N$. And for the imperfect detection case:

$$
\begin{aligned}
\hat{\boldsymbol{\pi}}_{x_{k-1}} &= \boldsymbol{\pi}_{x_{k-1}} \mathbf{P}, \\
\boldsymbol{\pi}_{x_k}(i) &= \begin{cases} \dfrac{\hat{\boldsymbol{\pi}}_{x_{k-1}}(i) p_c(z_k)}{\displaystyle\sum_{j \in \mathcal{S}(k, z_k)} \hat{\boldsymbol{\pi}}_{x_{k-1}}(j) p_c(z_k) + \sum_{j \in \mathcal{S}^c(k, z_k)} \hat{\boldsymbol{\pi}}_{x_{k-1}}(j) p_i(z_k)}, & \text{if } i \in \mathcal{S}(k, z_k) \\ \dfrac{\hat{\boldsymbol{\pi}}_{x_{k-1}}(i) p_i(z_k)}{\displaystyle\sum_{j \in \mathcal{S}(k, z_k)} \hat{\boldsymbol{\pi}}_{x_{k-1}}(j) p_c(z_k) + \sum_{j \in \mathcal{S}^c(k, z_k)} \hat{\boldsymbol{\pi}}_{x_{k-1}}(j) p_i(z_k)}, & \text{if } i \in \mathcal{S}^c(k, z_k) \end{cases}
\end{aligned}
\tag{4.31}
$$

for $k = 1, \ldots, N$.

4.4   Obtaining Pareto-Optimal Policies

Cost Function

The cost function $\tilde{g}_k(x_k, d_k)$ represents a convex combination of two opposite objectives: the throughput $b_k(m_k)$ and the average overlapping time $c_k(x_k, m_k)$:

$$\tilde{g}_k(x_k, d_k) = \left[\alpha \frac{c_k(x_k, m_k)}{c_{max}} + (\alpha - 1) \frac{b_k(m_k)}{b_{max}} \right] (1 - u_k) \qquad (4.32)$$

where $0 \leq \alpha \leq 1$. The objectives are divided by their maximum values c_{max} and b_{max} in order to obtain dimensionless (normalized) values. Note that $b_k(x_k, m_k)$ is added with negative sign in the convex combination because the cost function is minimized in problem (4.16). According to multi-objective optimization theory, given a value of α, the solution of the finite-horizon MDP with the cost defined in (4.32) yields a Pareto-optimal solution regarding these two opposite objectives, *i.e.* a solution for which there does not exists any policy such that one objective can be improved without worsening the other. The set of Pareto-optimal solutions is referred to as the Pareto-front of the problem and we can estimate it by iteratively solving the MDP for a representative discrete range of α values.

Throughput

The maximum throughput is attained when M channels are found idle in the first M scanning slots and the decision made is to transmit using these M channels. Given a transmission rate R and scanning and transmission times equal to τ and T respectively, the maximum throughput is

$$b_{max} = \frac{MRT}{M\tau + T}. \qquad (4.33)$$

When the access algorithm decides to transmit in n channels at stage k the throughput is

$$b_k(m_k) = \frac{nRT}{k\tau + T}, \ n \leq k \qquad (4.34)$$

Average Overlapping Time

The maximum average overlapping time occurs when, after selecting M channels to transmit in, these M channels turn out to be occupied by PUs when the SU starts to transmit and remain occupied for at least T seconds. Therefore, in the worst case of having M collisions lasting for T seconds, the total overlapping time is:

$$c_{max} = MT. \qquad (4.35)$$

Accordingly, given a vector m_k selecting n channels to transmit in after the k-th scanning slot, the average overlapping time is given by

$$c_k(x_k, m_k) = \sum_{j=1}^{n} \mathop{\mathrm{E}}_{\substack{Y(t) \\ t \in [k\tau, k\tau + T)}} \left[t \text{ in } \mathcal{S}(m_k(j), 1) \Big| x_k \right] \qquad (4.36)$$

where the expectation operation calculates the expected total time that the spectrum process $Y(t)$ spends in the set of states $\mathcal{S}(m_k(j), 1)$ during the period $[k\tau, k\tau + T)$. Recall that, according to its definition in (4.6), $\mathcal{S}(m_k(j), 1)$ corresponds to the set of states in $\mathcal{S}$, such that channel $m_k(j)$ is occupied by a PU. Therefore, (4.36) is the sum of the expected overlapping times on each of the n channels selected. Let us denote by $l_{i,x_k}(t')$ the expected total time that $Y(t)$ spends in state $i \in \mathcal{S}$ during the interval $[k\tau, k\tau + t']$ given the information x_k, after k scanning slot. Then,

$$c_k(x_k, m_k) = \sum_{j=1}^{n} \sum_{i \in \mathcal{S}(m_k(j), 1)} l_{i,x_k}(T - k\tau) \qquad (4.37)$$

The expected total time spent in a state, $l_{i,x_k}(t')$, is a known performance parameter of Markov models (see [85]). Let $\boldsymbol{L}_{x_k}(t')$ indicate the vector whose elements are $l_{i,x_k}(t')$, and let $\boldsymbol{\pi}_{x_k}^{(t)}$ denote the transient state probability vector of $Y(t)$ conditioned on x_k. By definition, $\boldsymbol{L}_{x_k}(t')$ can be expressed as:

$$\boldsymbol{L}_{x_k}(t') = \int_{k\tau}^{k\tau + t'} \boldsymbol{\pi}_{x_k}^{(t)} dt, \qquad (4.38)$$

At time $k\tau$, the value of $Y(t)$ is given by $Y(k\tau) = Y_k$, whose probability distribution conditioned on x_k is $\boldsymbol{\pi}_{x_k}$. Therefore, $\boldsymbol{\pi}_{x_k}^{(k\tau)} = \boldsymbol{\pi}_{x_k}$, given by (4.30) and (4.31), allowing us to compute $\boldsymbol{L}_{x_k}(t')$ as the solution of the following differential equation

$$\frac{d\boldsymbol{L}_{x_k}(t')}{dt'} = \boldsymbol{L}_{x_k}(t')Q + \boldsymbol{\pi}_{x_k}, \quad \boldsymbol{L}_{x_k}(0) = \boldsymbol{o} \qquad (4.39)$$

Backward Induction Algorithm

Once the POMDP has been reformulated as a fully observable state estimation one, and given that the horizon is finite, the MDP problem can be solved by means of the backward induction algorithm, also known as dynamic programming (DP) algorithm. In addition, because both the state and the control spaces are finite, there exists at least one deterministic Markovian policy which is optimal [23]. This is the type of policies defined in Subsection 4.3, where each function μ_k maps states (not past histories) into decisions, *i.e.* μ_k is Markovian. Similarly, $\mu_k(x_k)$ always gives the same outcome with probability 1, *i.e.* μ_k is deterministic.

The algorithm to find optimal Markovian deterministic policies for the MDP problem (4.16) is the following:

1. Set $k = N$ and

$$J_N^*(x_N) = \min_{d \in \mathcal{D}_N(x_N)} g_N(x_N, d) \text{ for all } x_N \in \mathcal{X}_N^* \qquad (4.40)$$

Set

$$\mu_N^*(x_N) = \arg \max_{d \in \mathcal{D}_N(x_N)} g_N(x_N, d). \qquad (4.41)$$

2. Substitute $k - 1$ for k and compute $J_k^*(x_k)$ for each $x_k \in \mathcal{X}_k^*$ by

$$J_k^*(x_k) = \min_{d \in \mathcal{D}_k(x_k)} \left\{ g_k(x_k, d) + \sum_{j \in \mathcal{X}_k^*} p_k(j|x_k, d) J_{k+1}^*(j) \right\} \qquad (4.42)$$

channels	memory (bytes)	multiplications	time (hours)
4	16	85296	4.7×10^{-10}
8	256	3.1×10^8	1.7×10^{-6}
12	4096	1.3×10^{12}	0.007
16	65536	5.2×10^{15}	28.67
20	1048576	2.1×10^{19}	117427
24	1.7×10^7	8.7×10^{22}	4.8×10^8

Table 4.3: Off-line computational effort and SU memory requirements
for OSA under random or sequential PU channel assignment

Set

$$\mu_k^*(x_k) = \arg \min_{d \in \mathcal{D}_k(x_k)} \left\{ g_k(x_k, d) + \sum_{j \in \mathcal{X}_k^*} p_k(j|x_k, d) J_{k+1}^*(j) \right\} \qquad (4.43)$$

3. If $k = 0$ stop. Otherwise return to step 2.

At the end of previous algorithm, (4.42) provides the optimal cost function for $k = 0$, $J_0^*(x_0)$. Because the scanning process starts with null initial information, *i.e.* $x_0 = 0$, the scalar value $J_0^*(0)$ is the optimal cost J^* defined in (4.16). Note that both (4.41) and (4.43) may result in more than one optimal decision per state, meaning that there may be more than one optimal policy η^* for the problem.

Computational Cost

Given the dimension of the problem N, and the spectrum management policy, it is straightforward to evaluate the computational cost of obtaining the optimal OSA policy. Table 4.3 shows the computational costs, including computation time in a 50 Gflops processor, for the case of random or sequential channel allocations, considering different number of channels N. Even if the policy is computed off-line, the computational effort becomes cumbersome for a spectrum size of around 20 channels. Nevertheless, a pre-computed policy implies memory requirements too. Considering the decision tree structure of the MDP (see Fig. 4.3) an SU only needs to store the information states where a transmission decision has to be made. Table 4.3 shows the worst case, in which 2^N states should are stored.

Therefore, for larger N values, approximation methods could be applied. For example, an approximate evaluation of matrix P, or state aggregation *e.g.* by modeling PU channel occupation in blocks of channels instead of individual ones. On the other hand, with the merged spectrum strategy, the implementation requirements of OSA reduce dramatically as shown in Table 4.4. This advantage adds to the ones shown in following section.

channels	memory (bytes)	multiplications	time (seconds)
4	4	1804	3.6×10^{-8}
8	8	11928	2.4×10^{-7}
12	12	37412	7.5×10^{-7}
16	16	85296	1.7×10^{-6}
20	20	162620	3.3×10^{-6}
24	24	276424	5.5×10^{-6}

Table 4.4: Off-line computational effort and SU memory requirements for OSA under spectrum merging for PUs

4.5   Numerical Results

Performance Evaluation

In this subsection we compare the performance of Pareto-optimal OSA policies for the three PU channel allocation strategies previously explained: random, sequential and sequential with spectrum merging for PUs. We use a Poisson traffic model, with 8 new PU arrivals per minute, and and average channel holding time of 1 minute, i.e. an offered traffic of 8 Erlangs over a 10-channels band. The parameter settings of the evaluation scenario are shown in Table 4.5.

Parameter	Value
N	10 channels
M	3 channels
λ	$\frac{8}{60}$ s^{-1}
$1/\mu$	60 s
T	10 s
τ	500 ms

Table 4.5: Parameter settings

As explained in previous section, the per-stage cost function (4.32) consists of a convex combination of two opposite parameters: throughput and overlapping time. For a given weight α in (4.32), a policy η^* solving the MDP problem yields an expected normalized throughput, denoted by J_{th} and an expected total overlapping time denoted by J_c. For any policy $\eta = \{\mu_0, \ldots, \mu_N\}$, J_{th} is obtained by the following recursive equations:

$$J_{th,k}(x_k) = g_{th,k}(x_k, \mu_k(x_k)) + \sum_{j \in \mathcal{X}_k^*} p_k(j|x_k, \mu_k(x_k)) J_{th,k+1}(j) \qquad (4.44)$$

for $k = N-1, \ldots, 0$, starting with

$$J_{th,N}(x_N) = g_{th,N}(x_N, \mu_N(x_N)) \text{ for all } x_N \in \mathcal{X}_N^*. \qquad (4.45)$$

In these equations, $g_{th,k}$ denotes the per stage throughput and is given by (4.14) and (4.32) with $\alpha = 1$. The total throughput is $J_{th} = J_{th,0}(x_0)$. The expected total overlapping time J_c is obtained similarly.

Fig. 4.4 shows the Pareto fronts obtained under the three PU channel allocation patterns considered. As expected, OSA performs better when channels are assigned sequentially to PUs, and much better when PUs always occupy adjacent channels into a compacted spectrum block. To give an idea of this improvement, let us set the constraint $J_c \leq 0.1$. Depending on how PU channels are managed, the maximum achievable J_{th} can be 0.4, 0.58 or 0.7. Conversely, if the goal is to keep J_{th} close to 0.6, the minimum achievable J_c values are 0.35, 0.14 or 0.03 for each PU channel management strategy.
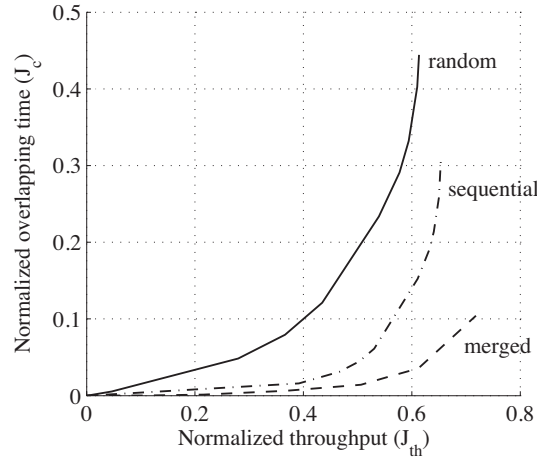
Figure 4.4: Pareto fronts showing the optimal combinations of normalized optimal throughput and normalized overlapping time for three different PU channel allocation patterns.

To show the benefit of MDP-based policies let us consider the following family of *baseline policies*, η^r: A CP following η^r decides to transmit over M channels whenever the number of available channels detected is $n_f = M + r$. Otherwise, the CP pair keeps on scanning the spectrum. If the total number of available channels in the spectrum is $n_f < M + r$, then the SU node transmits in $n = \max(n_f - r, 0)$ channels. The integer parameter $r > 0$ represents a bandwidth safety margin to reduce collisions. Fig. 4.5 compares the performance of baseline policies η^r for $r = 0, \ldots, N - 1$, to MDP-based policies under each PU channel management scheme. Note that the lower the safety margin r, the greater the throughput, at the cost of an increased overlapping time. Although MDP-based policies outperform the reference ones in every scenario, it is interesting how close the reference performance is to the Pareto front in the merged spectrum scenario, suggesting that, in this case, it may be acceptable to use a simple set of rules similar to η^r, alleviating the need to solve the MDP problem off-line. This also illustrates the utility of Pareto front computations for evaluating OSA policies. The performance of both sets of policies is also similar under every PU strategy for the case in which the overlapping time is of no importance ($\alpha = 0$), since both decide to transmit over the maximum M channels (if M or more channels are available).
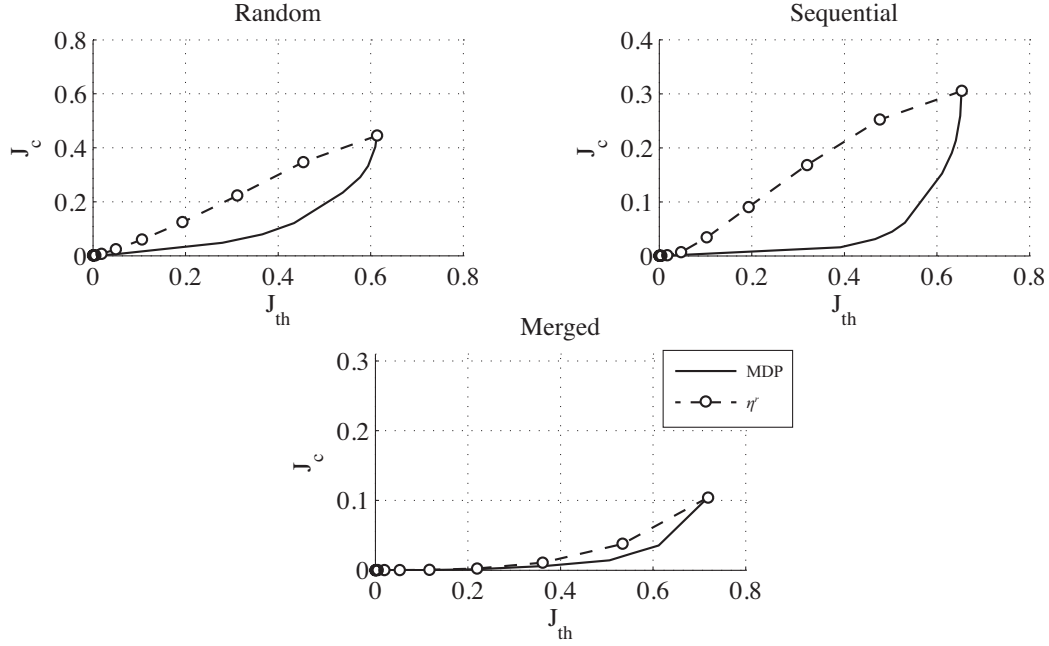
Figure 4.5:   MDP-based policies vs.   baseline policies η^r with $r =$ $0, \ldots, N-1$.

One reasonable concern about the proposed procedure for obtaining OSA policies is that it may be too reliant on PU traffic estimation accuracy. Even if the Poisson model is suitable for PU traffic, the λ and μ values used for off-line policy computation may be very different from realistic ones, which may also change over time. The question is how much does the accuracy of traffic descriptors affect the performance figures, especially overlapping time which determines the QoS of PUs. Next subsection addresses this issue.

Sensitivity to Traffic Estimation Accuracy

This subsection evaluates the sensitivity of each performance figure to drifts on the values of PU traffic descriptors λ and μ. By drift we refer to the difference between the value used for off-line policy computation (estimated value) and the value that actually describes PU traffic (nominal value). Drifts are positive when PU traffic is larger than the estimated one, and negative otherwise.

For the three PU channel management scenarios considered so far, the OSA policy is computed using the nominal λ and μ values specified in Table 4.5. Then, at each scenario, we evaluated the performance of the OSA policy in terms of throughput and overlapping time for drifts ranging from negative to positive values. In OSA, the main concern is to keep PU's QoS degradation at low values. Therefore, selected OSA policies have low overlapping times at nominal traffic descriptors, as shown in Table 4.6.

Fig. 4.6 shows the throughput and overlapping time values attained by the OSA policy in a system with random PU channel assignment, for different combinations of the estimated λ and μ values. The nominal traffic descriptors are $\lambda = \frac{8}{60}$ s^{-1}

PU channel allocation	random	sequential	merged
normalized throughput	0.43	0.53	0.72
normalized overlapping time	0.12	0.06	0.10

Table 4.6: Performance figures in the reference scenario with zero drift.

(8 arrivals per minute) and $\mu = \frac{1}{60}$ s^{-1} (average holding time: 1 minute), thus the axis in Fig. 4.6 contains negative and positive drifts on both parameters. It was expectable that positive and negative drifts resulted in reduced and improved throughput respectively. Nevertheless, the overlapping time remains at relatively low values for every combination of λ and μ. This shows a very desirable feature of the OSA strategy proposed: contrary to the usual belief, PU traffic estimation inaccuracies have little impact on PU's QoS. This effect is also observed in the scenarios of sequential PU channel assignment (see Fig. 4.7) and spectrum merging for PUs (see Fig. 4.8).

Clearly, if real PU traffic is heavier than predicted, the SU sensing process will provide, more frequently, observations corresponding to congested spectrum states. Therefore, independently of the accuracy of the traffic estimation, if the policy is intended to operate with low overlapping time, it will make the conservative decisions associated to these spectrum states. The only difference is that these decisions will be made more frequently. The consequence is that, in many cases, the overlapping time is even reduced under positive drifts. This is not always the case in systems with spectrum merging for PUs, which is expectable because in this scenario channels are more intensively used than in the others.
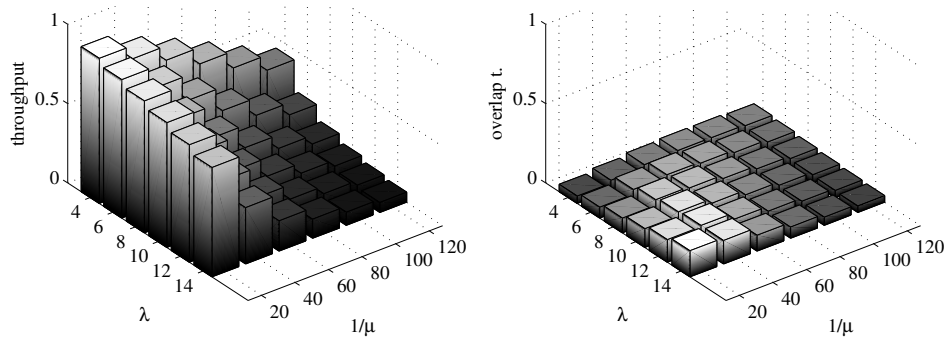
Figure 4.6: Normalized SU throughput and expected overlapping time with random channel assignment for PUs (λ is given in arrivals per minute).

4.6 Conclusions

We have addressed the design of OSA policies for hardware constrained cognitive networks, and the impact of different levels of coexistence friendliness of the PU. The problem was formulated as a partially observable Markov decision process with
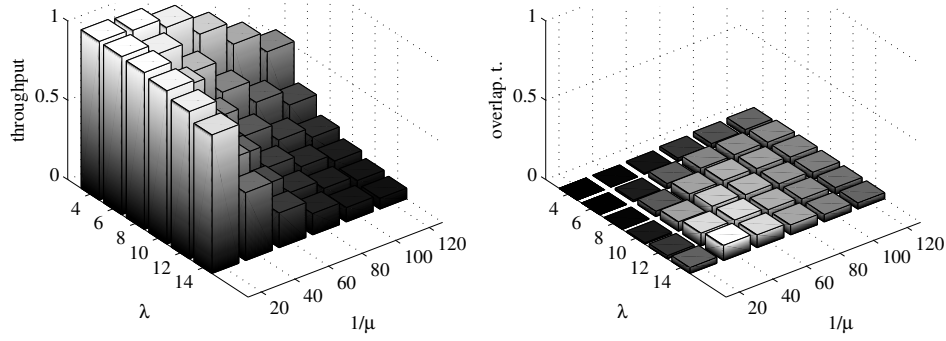
Figure 4.7: Normalized SU throughput and expected overlapping time with sequential channel assignment for PUs (λ is given in arrivals per minute).
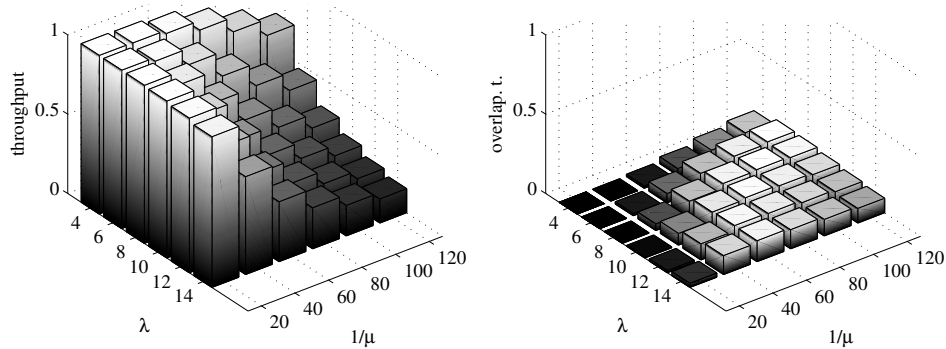

Figure 4.8: Normalized SU throughput and expected overlapping time with spectrum merging for PUs (λ is given in arrivals per minute).

finite horizon in which, at the end of each scanning slot, the spectrum state estimation is updated considering not only the new observation but also the effect of the delay on older observations. OSA policies need to balance SU's throughput and performance degradation of PUs, characterized by the overlapping time, which is a more general metric than collision probability. Regarding the PU allocation schemes, the overall performance improves when channels are assigned to PUs sequentially instead of randomly and improves even more if PUs always occupy adjacent channels. We also showed the robustness of Pareto-optimal policies against inaccuracies in PU traffic estimation, supporting the use of off-line precomputed policies. The results described in this chapter also reinforce the idea that, in situations of low or moderate PU traffic in which the operation of cognitive networks is justified, it is convenient for the licensed network to adjust its own channel allocation polices to facilitate spectrum sharing with SUs.

Response Surface Methodology for Efficient Spectrum Reuse in Cellular Networks

5.1 Introduction

As we pointed out in the introduction, the wireless industry incumbents are reluctant to allow DSA mechanisms because of interference concerns and because they do not get anything in return. Finally, in this last chapter devoted to OSA, we propose an application scenario that does not involve sharing with other entities, but sharing an operator's resources among its own networks. Our OSA proposal is specifically suited for cellular networks. In this way, we hope to go one step further in increasing the industry's interest in OSA, directly in line with its requests (as the industry reports in [86]).

Motivation

Approximately every decade, a new cellular access technology is developed and introduced in the market. However, because not all users upgrade their terminals at the same pace, the operators have to keep their legacy networks working, and the spectrum assigned to these old networks becomes gradually more and more underused. One way to increase the spectrum efficiency is to allow the users of the newest networks to opportunistically access the spectrum of the legacy networks. Applying the concepts of cognitive radio [39], the terminals of the newest network would be the secondary users (SUs), and the terminals of the legacy network would be the primary users (PUs). The secondary access to the spectrum of the primary (legacy) network should not cause noticeable degradation to PU transmissions. In this scenario, the operator owns and control both the primary and the secondary networks, and therefore can establish the rules for secondary access and monitor the impact of these rules on the performance of both SU and PU transmissions.

However, we assume *no modifications on the legacy network*, implying that the operation of the secondary network is *transparent* to the primary network.

We consider that the SUs establish point-to-point connections among them (cognitive pairs) in an ad-hoc fashion. We propose a semi-decentralized secondary access scheme allowing each SU to access a set of primary frequency bands using either temporal or spatial spectrum holes (opportunities). By giving more access options to each SU, this approach can outperform previous, more limited, mechanisms. To fully exploit its potential, we propose an on-line algorithm that, with very small computational and signaling overhead, allows the system to learn the optimal SU access policy in terms of capacity.

Related Work

Other works have proposed the combination of temporal and spatial sensing by the SUs [87], [88], [89], [90]. Works like [87] and [88] consider a single secondary transmitter and exploit information from spatial sensing to improve the performance of temporal sensing. The single transmitter model is also applied in [89] for vehicular networks. In [90] the capacity is optimized by randomizing the access strategies, but for a single PU, single SU scenario. However, the motivational scenario of our work comprises multiple secondary transmitters with multiple primary base stations. Thus, we must consider relevant features that are not captured by the single transmitter model, such as the inter SU-interference and the frequency reuse in a cellular structure.

In cellular networks, previous works as [91] and [92] have addressed spatial spectrum reuse by a secondary network. However, these works assume that the secondary users only scan one frequency. As pointed out by their authors, scanning all the frequencies would improve the performance.

Our Contribution

Our contribution in this work is a semi-distributed mechanism for opportunistic spectrum access in which *all the spectrum of the system can be made available* to each SU. Each SU decides locally which frequency band to access in, according to a probability distribution. *A centralized entity learns the optimal probability distribution over the frequency bands of the cellular network*, which is periodically broadcasted among the SUs. The learning algorithm applies the response surface methodology (RSM) [93], whose application is novel in interference management problems.

In the following section we describe the system. Section 5.3 formulates the design problem and Section 5.4 presents the RSM algorithm to solve it. Finally, Sections 5.5 and 5.6 presents the numerical results and the conclusions of this work.

5.2   System description

The system considered comprises a legacy cellular network (primary network), a secondary network establishing ad-hoc point-to-point links, and a secondary access controller (SAC) which monitors the system performance and broadcasts the operation parameters of the SU access strategy (with small signaling overhead, as we discuss in Section 5.4). The primary network contains base stations (PBSs) and primary users (PUs). Each PBS covers a certain geographical area (primary cell) and is assigned a frequency band different from its adjacent cells. We assume a frequency reuse scheme of 7 frequency bands (reuse factor 7), denoted by $f_1, \ldots, f_7$. Nevertheless, the proposed method can be applied to other reuse factors as well. Each frequency band is divided into $2N$ orthogonal channels (N downlink and N uplink channels). Secondary access is constrained to downlink channels. Time is divided into equal duration time-slots, which is usual in most cellular systems. The data transmitted over a single PU channel in a time-slot is referred to as *packet*.

The secondary network consists of pairs of secondary users (SUs) entering and leaving the system in an ad-hoc fashion. This network model can characterize femtocells, terminals acting as relays, or any short-range transmission using available spectral resources of the legacy cellular network. The SAC is associated to the network to which the SUs belong. In particular, we assume that the SUs are the users of a new generation cellular network. Because both networks (the newest and the legacy one) belong to the same operator, we consider that the SAC can retrieve some information from the legacy network.

The SUs can detect the power of the pilot tones of the neighboring PBSs. With this information (and possibly with the aid of the SAC) each SU can infer its position with respect to the surrounding cells and therefore be aware of the PBS power levels from each frequency band. Indeed, the SU does not need to estimate exactly these PBS power levels, it just needs to establish an ordering of the frequencies $f_1, \ldots, f_7$ from lower to higher PBS power level. For a generic SU, let $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_7)$ be the vector of frequencies ordered in increasing received power, which clearly depends on the location of the SU. Note that ϕ_7 corresponds, in general, to the frequency band of the primary cell where the SU is located. Each SU builds its own $\boldsymbol{\phi}$ vector by scanning the PBS pilot tones on each frequency band. This vector only needs to be updated when the SU changes its location.

Example. In Figure 5.1, any SU located in area A has $\phi_1 = f_1$ (similarly, any SU in area B has $\phi_1 = f_2$). But because each SU is closer to a different PBS, $\phi_7 = f_5$ for SU 1, and $\phi_7 = f_6$ for SU 2. These SUs have also different values of ϕ_2 and ϕ_4.

When an SU transmits over any band $\phi_i = f_j$, it may cause some interference to the closest PUs using the same channel. Obviously, these PUs are located at cells using the same band f_j. Given how vector $\boldsymbol{\phi}$ is built, the average level of SU interference caused at PU receivers is proportional to the ordering index i of the selected band ϕ_i, i.e. a single SU transmission over ϕ_1 causes interference on the most distant cell, a transmission over ϕ_2 interferes on the second most distant cell and so forth. In our system, an OSA strategy exploiting only *spatial opportunities*,
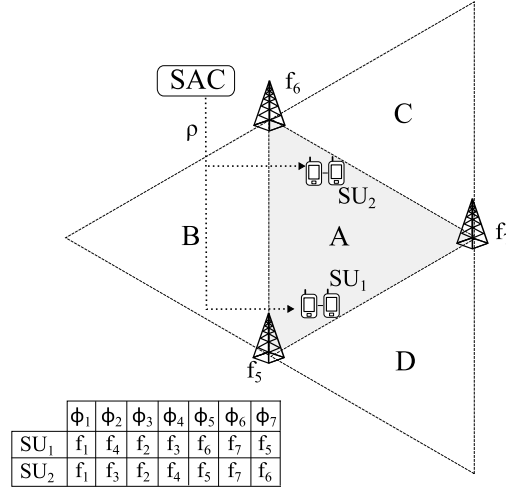
Figure 5.1: *Example of the system with 2 SU pairs. Every SU in area A associates f_1 to ϕ_1. The remaining f_j-ϕ_i associations for each SU depend on how close the SU is to areas B, C or D. The SAC broadcasts a vector of operation parameters ρ to the SUs.*

implies that each SU only tries to access over ϕ_1. However, what really matters is the aggregate interference at the receivers. Note that, in accordance to the ad-hoc nature of the secondary network (small SU transmission range, low antenna heights, indoor locations in many cases) the PU receivers could still decode packets even if some potential interference sources are closer, i.e. if some SUs select other bands different from ϕ_1. In this case, the benefit of giving more frequency band options to the SUs is to decrease the inter-SU interference, increasing the capacity of the secondary network.

Besides, when an SU tries to access the band of a close PBS, i.e. a band associated to a higher interference power such as ϕ_7, the SU can perform *PU activity detection* on the channels within the band [39]. This activity detection allows the SU to detect free channels before starting transmission, and to stop SU transmission when detecting overlap with PU transmissions. One way to implement this latter capability is the Background Detection mechanism described in Chapter 2. The closer the SU is to the PBS, the more reliable is the PU activity detection and therefore the less impact on PU communication. The free channels in ϕ_7 are identified as *temporal opportunities* [17] in previous works.

As shown in [17] the typical strategies of using only spatial opportunities (ϕ_1 channels) or only temporal ones (ϕ_7 free channels) are not optimal, and an hybrid scheme combining both types of access increases the capacity of the system. The challenge is to find the optimal balance between the fraction of the SU population exploiting each type of spectrum opportunities. In this work we incorporate a learning strategy to address this issue, and we extend the hybrid access idea by allowing the SU to use 3 or more bands from its ϕ vector.

Now let us describe the proposed **secondary access scheme** over a spectrum comprising R frequency bands (R denotes the frequency reuse factor of the primary cellular network). Let us define the probability distribution $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_R)$, where

$0 \le \rho_i \le 1$ for $i = 1, \ldots, R$, and $\sum_{i=1}^{R} \rho_i = 1$. Let $\mathcal{P}$ denote the set of possible values of $\boldsymbol{\rho}$. Before transmitting, each SU randomly selects one band, ϕ_i from its ϕ vector, according to the probability distribution $\boldsymbol{\rho}$, i.e. $P(\text{select band } \phi_i) = \rho_i$. Once the SU has selected a particular band, ϕ_i, if the SU is capable of detecting enough PBS power in ϕ_i to perform PU activity detection, then the SU tries to access over temporal opportunities, occupying only channels that are free of PU activity, that is, using an OSA scheme like the ones described in previous chapters. Otherwise, the SU will access any channel with a sufficiently low SINR (considering also the interference from other SUs).

The value of $\boldsymbol{\rho}$ should be determined by the SAC, which broadcasts it to the SUs. The objective pursued in $\boldsymbol{\rho}$ computation is to maximize the system performance. In the envisioned scenario, the performance metric is the capacity achieved by the SUs, subject to a given constraint on the probability of correct packet reception at the PU receivers. Next section formalizes the problem in these terms. As we show in Section 5.4, the SAC aims to *learn* the optimal $\boldsymbol{\rho}$[1] based on the performance measurements retrieved from the primary and the secondary devices.

Note that the outlined mechanism performs a semi-decentralized resource allocation. The $\boldsymbol{\rho}$ learning process is centralized at the SAC, while each SU decides autonomously which channel to occupy, based on $\boldsymbol{\rho}$ and its own spectrum sensing over the selected band. The signaling overhead consists of periodic broadcast messages from the SAC to the SUs *announcing* the updates of $\boldsymbol{\rho}$, and one feedback packet from each terminal to the SAC after each session, containing the performance metric measured during the session.

5.3 Problem Formulation

In this section we define the performance metrics for both the PUs and SUs and formulate the problem that the SAC needs to solve to determine the ρ vector.

Let us consider a secondary network with a given traffic intensity characterized by its arrival rate per area unit, and a random transmission time. As stated above, the vector ρ determines how the SUs distribute themselves over the available spectrum and the proportion of SUs using PU activity detection. In consequence, the SINR at each PU or SU receiver depends on ρ. Let N_t^{SU} and N_t^{PU} denote the number of SUs and PUs in the system, respectively, at time-slot $t = 1, 2, \ldots$. The SINR over time at an SU receiver $s \in \{1 \ldots N_t^{SU}\}$, is a discrete time stochastic process induced by ρ, and denoted by $\Gamma_t^s(\rho)$. The expected normalized capacity per active SU pair is given by

$$C_s(\boldsymbol{\rho}) = E \left\{ \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{s=1}^{N_t^{SU}} \frac{\log_2 (1 + \Gamma_t^s(\boldsymbol{\rho}))}{N_t^{SU} C_{max}^S} \right\} \tag{5.1}$$

---

[1]It is advisable to divide the system into relatively homogeneous regions in terms of PBS density, traffic intensity and type of terrain, so that a suitable ρ can be found for each region.

where C_{max}^S is the maximum achievable SU capacity per Hertz. The expectation is taken over Γ_t^s and N_t^{SU}.[2] At a PU receiver, it is assumed that a data packet transmitted on time-slot t is correctly decoded if its SINR, Γ_t^{PU}, is greater than a given detection threshold, γ^{PU}. Then, the probability of correct detection at a PU is defined as

$$P_c(\boldsymbol{\rho}) = E\left\{\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{j=1}^{N_t^{PU}} \frac{\mathbb{I}_{\{\Gamma_t^j(\boldsymbol{\rho}) > \gamma^{PU}\}}}{N_t^{PU}}\right\} \tag{5.2}$$

where Γ_t^j refers to the SINR process at the j-th active PU, and $\mathbb{I}_{\{z\}}$ is an indicator function which equals 1 if condition z holds, and equals 0 otherwise.

Therefore, the objective of the SAC is to find $\boldsymbol{\rho}$ solving the following problem:

$$\begin{aligned} \max_{\boldsymbol{\rho}} \quad & C_s(\boldsymbol{\rho}) \\ \text{subject to} \quad & P_c(\boldsymbol{\rho}) \geq P_{c,min} \\ & \boldsymbol{\rho} \in \mathcal{P} \end{aligned} \tag{5.3}$$

where $P_{c,min}$ denotes the minimum acceptable P_c.

Finding an optimal $\boldsymbol{\rho}$ is a challenging task because the (multiple) SINR stochastic processes capture the interaction of a random number of randomly located terminals with fading effects among each pair of them. Moreover, these processes may have very different characterizations. For example, the interference caused by the SUs using ϕ_7 or ϕ_6 can be very intense in general but, thanks to the PU activity detection in closer cells, it would be infrequent and last short periods of time (depending on the detection accuracy). On the other hand, the interference from the SUs using bands that usually do not allow PU activity detection (e.g. ϕ_1 or ϕ_2) will be sustained over time, but will be less intense (similar to the co-channel interference from other PBSs).

The problem (5.3) is formally a *stochastic optimization problem with stochastic constraints* [93], that must be solved online, i.e. while the system is operating. Therefore, the SAC has to dynamically *learn* an optimal $\boldsymbol{\rho}$. One feasible way to address it is by means of the Response Surface Methodology (RSM) [94]. However, the inclusion of stochastic constraints introduces high complexity in the formulation. In the following section we develop the particular case in which $P_{c,min}$ is sufficiently low for condition $P_c(\rho) \geq P_{c,min}$ to hold at the optimal $\boldsymbol{\rho}$ (in numerical results, $P_c(\boldsymbol{\rho})$ at the optimal $\boldsymbol{\rho}$ is never less than 5% of the P_c obtained in absence of SU access). This allows us to remove the stochastic constraint regarding $P_c(\boldsymbol{\rho})$. We leave as future work the inclusion of constraints in the RSM formulation.

5.4   Response Surface Method (RSM) Algorithm

The aim of the algorithm is the maximization of the expected value of the average capacity function (5.1) on a closed convex feasible domain $\mathcal{P} \subset \mathbb{R}^R$ for the input vector $\boldsymbol{\rho}$, that is, to find a $\boldsymbol{\rho}$ solving $\max_{\boldsymbol{\rho} \in \mathcal{P}} C_s(\boldsymbol{\rho})$.

---

[2]Note that the traffic intensity in a cellular communication network varies during a day, but if it is observed during a smaller time window, e.g. 1 hour, the traffic arrival process can be considered stationary, with constant intensity.

RSM allows us to find an approximate solution to this problem by successively estimating the gradient of the objective function and using these estimations in stochastic gradient ascent steps. At each one of these steps, numbered by $n = 1, 2, \ldots$, the system generates one update of the input vector $\boldsymbol{\rho}_{(n)}$ according to the following expression

$$\boldsymbol{\rho}_{(n+1)} = \boldsymbol{\rho}_{(n)} + \alpha_{(n)} \hat{\nabla} C_s(\boldsymbol{\rho}_{(n)}) \tag{5.4}$$

where $\hat{\nabla} C_s$ denotes the estimation of the gradient ∇C_s, and $\alpha_{(n)}$ is the step-size weighting factor. A standard condition for the selection of $\alpha_{(n)}$, assuring the convergence of $\boldsymbol{\rho}_{(n)}$ [93], is $\sum_{n=1}^{\infty} \alpha_{(n)} = \infty$, $\sum_{n=1}^{\infty} \alpha_{(n)}^2 < \infty$.

Let us discuss the computation of $\hat{\nabla} C_s$. Given $\boldsymbol{\rho}_{(n)}$ at step n, we consider a subdomain $\mathcal{S}_{(n)}$ of the feasible domain $\mathcal{P}$ such that $\boldsymbol{\rho}_{(n)} \in \mathcal{S}_{(n)} \subset \mathcal{P}$. Note that $\boldsymbol{\rho}_{(n)}$ is the point at which the estimate $\hat{\nabla} C_s$ must be computed. Therefore, we need to estimate the objective function $C_s(\boldsymbol{\rho}_{(n)})$ on $\mathcal{S}_{(n)}$, by taking samples $y^{(i)}, i = 1, \ldots, p$, of the function. For this, we need a finite set of points $\boldsymbol{\rho}_{(n,i)}, i = 1, \ldots, p$, generally called *design points*, belonging to $\mathcal{S}_{(n)}$. These points are chosen by the decision maker (the SAC in our case), and can be, for example, random perturbations of $\boldsymbol{\rho}_{(n)}$, falling within $\mathcal{S}_{(n)}$.

Let $t_{(n)}$ denote the time-slot in which the update $\boldsymbol{\rho}_{(n)}$ is obtained. Given the set of p decision points, the SAC can obtain samples $y^{(i)}$ by this simple procedure:

1. Determine p sampling instants $t_{(n,i)} = t_{(n)} + iT$ for $i = 1 \ldots p$, where T is a sufficiently long time period for measuring performance at the SUs.

2. At $t_{(n,i-1)}$ (where $t_{(n,0)} = t_{(n)}$), the SAC signals the design point $\boldsymbol{\rho}_{(n,i)}$ to the SUs, for $i = 1 \ldots p$.

3. At $t_{(n,i)}$, the SAC obtains the capacity samples from each active SU and averages them to obtain $y^{(i)}$.

Thus, the estimates can be expressed as

$$y^{(i)} = \frac{1}{T} \sum_{k=t_{(n,i-1)}}^{t_{(n,i)}} \sum_{s=1}^{N_k^{su}} \frac{\log_2 \left(1 + \Gamma_k^s(\boldsymbol{\rho}_{(n,i)}) \right)}{N_k^{su} C_{max}^S} \quad i = 1, 2, \ldots, p, \tag{5.5}$$

Note that the stochastic ascent algorithm needs p periods of length T to upgrade $\boldsymbol{\rho}_{(n+1)}$.

The objective function C_s is then approximated on $S_{(n)}$ by a polynomial response surface model $\hat{C}_s(\boldsymbol{\rho}) = \hat{C}_s(\boldsymbol{\rho}|\beta_0, \beta_1 \ldots \beta_R)$. The coefficients β_j, are determined by least squares estimation. Therefore, the RSM-gradient estimator $\hat{\nabla} C_s(\boldsymbol{\rho}_{(n)})$ at $\rho_{(n)}$ is defined by the gradient (with respect to ρ)

$$\hat{\nabla} C_s(\boldsymbol{\rho}_{(n)}) = \nabla \hat{C}_s(\rho_{(n)}). \tag{5.6}$$

Thus, C_s is approximated on $S_{(n)}$ by the linear empirical model

$$\hat{C}_s(\rho_n) = \beta_0 + \boldsymbol{\beta}_I'(\boldsymbol{\rho} - \boldsymbol{\rho}_{(n)}) \tag{5.7}$$

where

$$\boldsymbol{\beta}' = (\beta_0, \boldsymbol{\beta}_I') = (\beta_0, \beta_1, \ldots, \beta_R) \tag{5.8}$$

is the $(R+1)$-vector of unknown coefficients of the linear model (and $\boldsymbol{\beta}'$ denotes the transpose of the columns vector $\boldsymbol{\beta}$). Having samples $y^{(i)}$ of the function values $C_s(\boldsymbol{\rho}_{(n,i)})$ at the design points $\boldsymbol{\rho}_{(n,i)}, i = 1, \ldots, p$, in $S_{(n)}$, we can obtain, by least squares, the following estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}} = (\mathbf{W}^\mathsf{T}\mathbf{W})^{-1}\mathbf{W}'\boldsymbol{y}. \tag{5.9}$$

Here, the $p \times (R+1)$ matrix $\mathbf{W}$ and the p dimensional vector $\boldsymbol{y}$ are defined by

$$\mathbf{W} = \begin{pmatrix} 1 & \boldsymbol{\delta}^{(1)} \\ 1 & \boldsymbol{\delta}^{(2)} \\ \vdots & \vdots \\ 1 & \boldsymbol{\delta}^{(p)} \end{pmatrix}, \quad y = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(p)} \end{pmatrix} \tag{5.10}$$

with $\boldsymbol{\delta}^{(i)} = \boldsymbol{\rho}_{(n,i)} - \boldsymbol{\rho}_{(n)}$, for $i = 1, 2, \ldots, p$. Note that $(\mathbf{W}^\mathsf{T}\mathbf{W})$ in (5.9) is invertible whenever the columns of $\mathbf{W}$ are linearly independent, which can be easily guaranteed by a proper selection of the design points.

In case $C_s(\boldsymbol{\rho})$ is concave on $\mathcal{P}$, the RSM algorithm will approach the global optimum following the stochastic gradient ascent iterations (5.4). The concavity of $C_s(\boldsymbol{\rho})$ has been observed in numerical simulations in the scenarios considered in the following Section.

5.5   Numerical Results

Evaluation Framework

This subsection describes the Monte-Carlo methodology and the scenario used to evaluate C_s. The primary network uses a 7-band frequency planning, as previously stated, with irregular shaped cells having an average radius $r = 700$ m. We consider pathloss and multipath fading. To compute the pathloss attenuation over distance, $A(d)$ (dB), we use the following piecewise dual-slope model [28]:

$$A(d) = \begin{cases} K + 10\gamma_1 \log_{10}(d/d_0) & d_0 \leq d \leq d_c \\ K + 10\gamma_1 \log_{10}(d_c/d) + 10\gamma_2 \log_{10}(d/d_c) & d > d_c \end{cases} \tag{5.11}$$

The critical distance d_c is notably smaller for SU transmission than for PBS transmissions, because the PBSs are located at high outdoor locations while SUs are, in general, located either indoors or at ground level. All the signals are assumed to experience Rayleigh fading.

The area used to generate random terminal locations is a 4.2 Km $\times$ 4.8 Km rectangle. We focus on the downlink channels of one frequency band. We consider an scenario in which the primary traffic intensity is low 0.16 Erlangs/Km^2, so that the average occupation of the spectrum by the primary terminals is only 5%. In contrast, the SU traffic intensity, 27.5 Erlangs/Km^2, is high in comparison. In this case, the SU capacity is mostly determined by the inter-SU interference and

Parameter	Assigned value
Primary transmitters	
number of downlink channels, N	5
cell radius, r	700 m
average received power at PU	-78 dBm
SINR detection threshold at PU,γ^{PU}	-17 dB
baseline noise at PU ($N_0 + I_{PU}$)	-110 dBm
Secondary transmitters	
average SU Tx power per channel	0.5 W
SU link distance	90 m
probability of PU activity detection	0.9
probability of overlap detection	0.8
Propagation	
pathloss exponents, γ_1, γ_2	2.4, 4.2
propagation factor K	46.7 dB
critical distance for PBS transmission	$1.2r$ m
critical distance for SU transmission	100 m
RSM	
dimensions of ρ,R	$1\ldots5$
measuring period T (in time-slots)	300
number of samples per step p	$3(2^{R-1}+1)$

Table 5.1: Parameter setting of the reference scenario used in numerical evaluations

the RSM algorithm is essentially performing *interference management* in the SU network. We also discuss the effect of higher PU spectrum occupation and a different SU traffic intensity. According to the ad-hoc nature of the secondary network, the average SU link distance considered is 90 m. Table 5.1 summarizes the simulation parameters considered.

Convergence and Usage of the Frequency Bands

Let us consider the case of 3-dimensional $\boldsymbol{\rho}$ vectors, R = 3. In this case, the $\boldsymbol{\rho}$ vector of a particular SU is $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3)$, being ρ_3 the probability of trying to access the band ϕ_7 of the primary cell where the SU is in, and ρ_1 and ρ_2 the probabilities of accessing ϕ_1 and ϕ_2, the bands where the SU receives less PBS power. The SU can only perform PU activity detection on ϕ_7 channels, while the channels of the other bands are considered spatial opportunities. Fig. 5.2 shows the values of ρ_1 and ρ_2 (ρ_3 is simply $1 - \rho_1 - \rho_2$) over consecutive update steps $n = 1, 2, \ldots$. The initial vector is $\rho_0 = (0, 0, 1)$. Although ρ_0 is a rather poor initial guess, we see that the RSM algorithm stabilizes after 20 iterations. The final value is $\boldsymbol{\rho} = (0.29, 0.42, 0.29)$. Fig. 5.3 shows the estimated value of C_s as a function of ρ_1 and ρ_2. Consistently with the result obtained by RSM, the maximum values of C_s lie on the line $\rho_1 + \rho_2 = 0.7$.
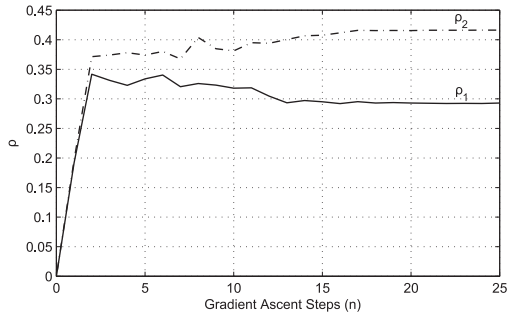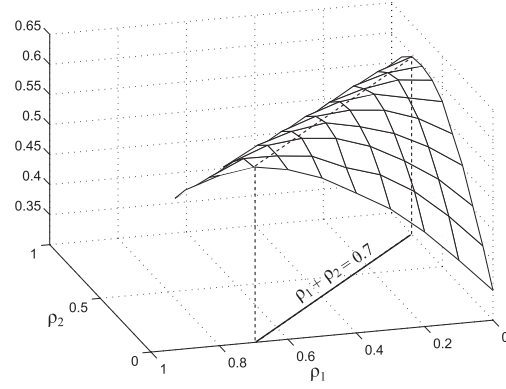
Figure 5.2: Consecutive RSM updates of ρ_1 and ρ_2.

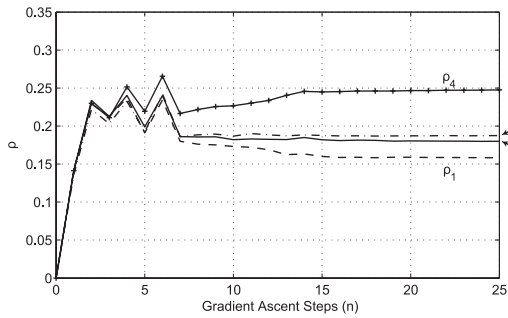Figure 5.3: Estimation of C_s over ρ_1 and ρ_2.

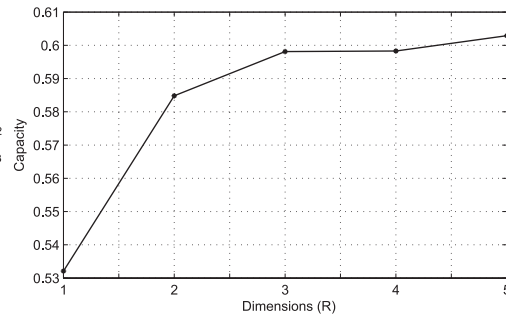Figure 5.4: Consecutive RSM updates of $\rho_1 \dots \rho_4$.

Figure 5.5: Maximum SU capacity obtained with RSM versus the number of dimensions of ρ, R.

Let us now consider the results for $R = 5$. The probability ρ_5 is now associated to ϕ_7, allowing PU activity detection, while $\rho_1 \dots \rho_4$ are associated to $\phi_1 \dots \phi_4$, where PU activity detection is assumed to be unfeasible. The initial $\boldsymbol{\rho}$ vector is $(0, 0, 0, 0, 1)$. Figure 5.4 shows that, in this case, the convergence of RSM is as fast as with $R = 3$. Note that, as in the case of $R = 3$, the algorithm is not assigning ρ_1 the highest probability. The reason is that the algorithm aims to minimize inter SU interference by separating the SU transmitters using the same band. Recall that in Fig. 5.1 all the SUs in triangular region A shared the same ϕ_1 but had different values for ϕ_2. The aggregated probability of using spatial opportunities is $\sum_{i=1}^{4} \rho_i = 0.77$, which is higher than in the $R = 3$ case. This result confirms the idea that, the more spectrum is available to each SU, the more spectrum the SU exploits. Having more spectrum options per SU (higher R) also allows the SU network to achieve a higher capacity, as illustrated by Fig. 5.5.

Effect of the Traffic Intensity

Fig. 5.6 shows the average SU capacity attained by the RSM algorithm with $R = 2$, versus the average spectrum usage by the PUs, for different SU traffic intensities. As expected, the more spectrum occupied by PU traffic, the smaller the achieved capacity. Similarly, more SU traffic implies less SU capacity. The reason of the
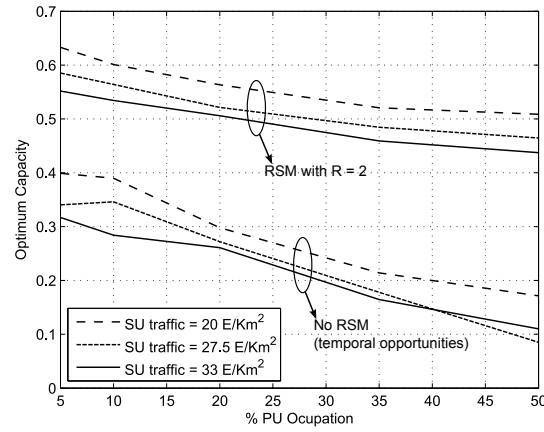
Figure 5.6: Maximum SU capacity obtained with RSM under different traffic intensities. For comparision, the figure also shows the capacity for a SU network exploiting only temporal opportunities (no RSM).

reduction on the maximum achieved capacity is the increment of the aggregate interference power at the SU receivers, from both the PBSs and other SU pairs. Compared to the typical approach of exploiting only temporal opportunities, the benefit of using RSM is noticeable for every traffic situation. In all cases, the reduction of the probability of correct detection at the PU receivers (P_c) was less than 5%.

5.6 Conclusions

Motivated by the problem of spectrum reuse in cellular networks, we presented a semi-distributed mechanism allowing the secondary network to learn the most efficient spectrum access strategy. This mechanism exploits both spatial and temporal opportunities, and is especially effective for interference management in highly dense secondary networks. The learning approach is based on response surface methodology (RSM) which, according to our numerical results, improves notably the system capacity compared to the usual strategies of exploiting only spacial or temporal opportunities, and shows a fast convergence rate even when it is poorly initialized. Because of that, the system can adapt its control vector ρ to variations on the traffic intensity or user distribution. Surprisingly, to the best of our knowledge, the use of RSM on this framework had not been previously reported. Our future work is focused on incorporating performance constraints for the primary network in the RSM formulation.

Part II

Automated Spectrum Trading

Automated spectrum trading mechanisms: understanding the big picture

6.1 Introduction

In this second part of the thesis, we focus on automated spectrum trading mechanisms. Compared to other DSA approaches, automated spectrum trading has the advantage of providing economic (and/or other) incentives to the entities involved, especially important to licensed operators, encouraging the adoption of DSA and promoting motivation for future investments in primary spectrum acquisitions. Nevertheless, deploying a secondary spectrum trading market could also cause the opposite result if is not properly crafted, that is, undermining the competition structure, by giving rise to collusive behaviors and/or service providers becoming monopolists (see the introduction of Sect. 6.5 of this chapter and [95], especially its Sect. 5.2 "Comparison of competition: tipping effects").

This chapter contains a comprehensive survey on automated spectrum trading which, to the best of our knowledge, is novel in its treatment of the subject in the following key aspects: its scope (a general view of spectrum trading), the proposed classification of the different issues and approaches in spectrum trading, its tutorial nature, its coverage of the most up-to-date works, and the identification of the main open problems.

Motivation of automated spectrum trading

Previous works in economic theory are not easily translated to spectrum trading because of the specific features of the traded good. A spectrum market poses the following additional challenges:

- Rapid variations with time
- Imperfect information

- Complex resource allocation considering re-utilization and heterogeneity of the good

Both spectrum supply and demand changes are related to data traffic intensity which, in general, experiences *rapid variations with time.* In addition, spectrum characteristics are also highly variable, in particular, availability and channel quality parameters. If trading decisions could be made in such short time-scales (possibly in the order of seconds), the spectrum exploitation efficiency would increase. This implies that the agents making the decisions in the spectrum markets have to be computerized. Automatic transactions are present in financial trading ([96]), but that market does not show the specific issues of spectrum trading.

Imperfect information not only refers to the competitive behavior of agents in a traditional market, by which they are not willing to reveal their true valuation of a good. It also refers to the fact that agents may not have complete or reliable information about the market due to the communication complexity that it would involve, specially in ad-hoc networks. It entails unfeasible requirements such as, for example, all entities knowing all channel gains between each other.

Resource allocation becomes more complicated with spectrum as a traded good because, depending on the mutual interference ranges of the entities, it can be *spatially reused.* It may also be considered that spectrum is an *heterogeneous* good, in the sense that the same spectrum portion can be valuated differently by each user, depending on its position, technology (spectrum efficiency use), etc.

Rapid variations with time and resource allocation in complex models are tightly interrelated. Optimal resource allocation in complex models requires a computational time that makes it hard to keep up with the rapid changes, which, in turn, may render a solution inefficient because of a time dis-adaptation. Dealing with imperfect information makes optimal resource allocation more complex, even intractable. On the contrary, local estimations could be considered, sacrificing model optimality for the sake of speed.

Our contribution

Our contribution in this chapter is summarized as follows:

- Our work is unique in its scope, focusing on automated spectrum trading as the most promising mechanism to solve spectrum scarcity.
- We identify the specific aspects that make spectrum different from conventional goods and the impact that these features have on automatic trading.
- We discuss past and current approaches, as well as future research lines, highlighting their advantages and disadvantages.
- Each sub-topic is presented and developed explicitly with references to previous works where each aspect has been considered or addressed.
- This work can serve as an introduction to the field for novel researchers, and is useful to experienced ones for its critical discussion of the most recent trends.
- Finally, as a conclusion of the surveying effort, we highlight overlooked issues of automated spectrum trading, with a special focus on practical implemen-

tation.

Compared to previous works in this area such as [97, 98, 99, 100, 101, 102, 103], our work covers a broader range of aspects, as summarized in Table 6.1. Although the scope of [97] claims to be "economic approaches", it devotes part of the survey to dynamic spectrum access, which has its own particularities. In addition, it is structured around a classification of related works where the items are a mix of mathematical tools and market principles, specially focusing on game theory, while our work extends the classification including further aspects in spectrum trading such as its objectives, mechanisms, possible market forms, etc. We also perform a deeper identification of key issues in the entities' decision, especially real-time adaptation, which could be considered as one of the most critical aspects for a spectrum trading mechanism to achieve an efficient solution. There is a small essay on spectrum sharing and trading in [98], and [99] presents a general study of spectrum trading with a notable tutorial approach. Compared to these works, our survey offers an extensive coverage of the advances in this field during the last five years, showing an upgraded taxonomy comprising the most novel proposals. From a different perspective other works cover only part of the issues addressed in this survey: [100] is a survey of spectrum trading for Cooperative Secondary Spectrum Access (CSSA) under imperfect information, [101] is concerned with self-organization paradigms in cognitive radio and [102, 103] develop a summary of game theoretic and auction approaches in dynamic spectrum access respectively.

The rest of this chapter is organized as follows. In section 6.2 we show key points in the spectrum trading mechanism design. Section 6.3 reviews the challenges arising as a consequence of the special properties of spectrum as a trading good. The mathematical tools used to study these trading models are discussed in section 6.4. How those techniques are used in different market forms is analyzed in section 6.5. Finally, we provide some guidelines and feedback for future research in section 6.6 and summarize the main conclusions of this survey in section 6.7.

Table 6.1: Summary of related works."E" means the topic is developed explicitly, "I" means the topic is developed implicitly, "M" means the topic is mentioned but not developed, "X" means the topic is not present.

	Our work	[97]	[98]	[99]	[100]	[101]	[102]	[103]
Scope	Spec. trading	DSA and spec. trading	DSA and spec. trading	DSA	CSSA in imperfect info.	Self-org.	Game Th. DSA	Spec. auctions
Type of work	Survey	Survey	Model	Book	Survey	Survey	Survey	Survey
Year	2013	2010	2008	2008	2013	2013	2011	2012
Number of references	98	36	14	<30	15	15	74	32
Economic objectives	**E**	X	X	I	X	X	**E**	**E**
CSSA	**E**	X	X	**E**	**E**	**E**	M	M
Real-time variations	**E**	X	X	X	X	I	I	M
Imperfect information	**E**	I	M	**E**	**E**	I	**E**	M
Spatial reuse	**E**	X	X	X	M	X	X	**E**
Heterogeneity	**E**	M	X	M	X	I	M	M
Formulation frameworks	**E**	**E**	**E**	**E**	**E**	M	**E**	M
Market forms	**E**	**E**	**E**	**E**	X	X	X	M
Implementation issues	**E**	M	M	**E**	**E**	**E**	I	**E**
Protection against malfunctions	**E**	**E**	X	**E**	X	X	**E**	**E**

6.2   Trading mechanism design

Basics

The most basic form of spectrum trading (and any trading in general) is found when the transaction takes place between just two entities. Usually one of these entities, the "seller", offers its unused spectrum opportunities to interested users, "buyers", in exchange for something (usually money). Although this situation may appear as simplistic, it is useful as more than a didactic example, since it corresponds to a real scenario in which there is only one primary and one secondary network, both with a centralized structure so the transaction is negotiated between their base stations as in Fig. 6.1, and, in fact, some authors have studied this market form as explained later in Section 6.4).
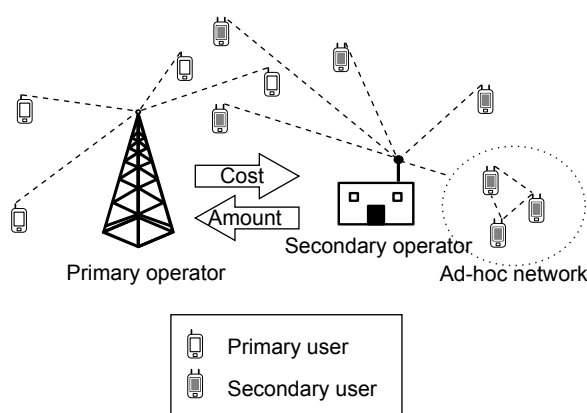
Figure 6.1: Basic spectrum trading

For the buyer, the trade-off between its interest in what the seller offers and its cost is expressed by the utility function $u_i(x)$, where x is an n-tuple comprising the conditions of the trade, *e.g.,* the amount of spectrum bought and the price paid. Similarly, the seller is also in a compromise between the revenue it obtains from the buyer and the costs associated with the sale: interferences, higher blocking probability (because it would have less available spectrum) and fixed costs because of the investment in infrastructure. This trade-off is expressed by the profit function $\pi_i(y)$, where y denotes the conditions of the trade. It can also be called utility function. Both functions return unidimensional values:

$$u_i, \pi_i : \mathbb{R}^N \to \mathbb{R} \tag{6.1}$$

An example of a simple utility function for buyer i, where bandwidth is the traded good, has the following form:

$$u_i(b_i, c_i) = k_i b_i - c_i b_i \tag{6.2}$$

where b_i is the amount of commodity to buy (bandwidth), c_i is its cost per unit for buyer i (it may not be the same for every buyer) and k_i characterizes the value

Table 6.2: Classification of different works by their economic objective

Main economic objective	Works
Benefit of the spectrum owner	[105, 48, 47, 49, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121]
Social optimum	[122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 98, 132, 133, 134, 135, 136, 137, 138]
Benefit of intermediaries	[139, 140, 141, 142, 143, 144, 145, 146, 147, 118, 119, 148, 146]

given by the SU to the spectrum units. The valuation of the bandwidth is then $k_i b_i$, while $c_i b_i$ represents the cost. It is far from trivial to select which metrics should be combined in a utility/profit function to quantify the impact of spectrum trading on each agent's outcome, and which constraints should be imposed. The metric used strongly influences the outcome of the trading models [104]. One usual approach is to use logarithmic functions to model the "law of diminishing returns": if the obtained resource is increased, at some point satisfaction will saturate because no more resources are really needed.

When there are multiple sellers and buyers in an area, the conflict between sellers and buyers gets more complex because of the added interactions brought by competition among sellers. The price per unit of a seller depends on the price per unit of the other sellers, as rational buyers will buy from the cheapest. There is also competition among buyers for the finite resource or in terms of mutual interferences and price charged depending on their total aggregate demand. The key of a spectrum trading mechanism design is to obtain policies for the entities involved so that they agree in the amount of commodity to be sold (supply) and bought (demand) as well as its price, in order to reach some particular objective.

Economic Objectives

In the introduction of this thesis, we explained that spectrum is scarce because it has been assigned inefficiently. Accordingly, every automated spectrum trading system tries to improve that efficiency, but who should benefit the most from that improvement? Politics aside, the first answer that may come to our minds would be "those who need it the most", that is to say, *maximizing the social welfare*. Even if we think it is the only possible answer, we still have to consider a fact about motivation: in order for the system to be feasible, all entities need to have an incentive to participate. In other words, no matter what the objective to maximize is, the utility functions of all the other entities need to be taken into account and be better than without the trading system (otherwise, they will not participate). And furthermore, there is one group of entities, the licensed spectrum owners, whose motivation to take part in the trading system is crucial. Then, trading mechanisms may focus on one of these other related economic objectives: *maximizing the profit of the licensed spectrum* and *maximizing the profit of a secondary spectrum*

provider.

- *Maximizing the social welfare.* It means maximizing an aggregate utility comprising the benefit perceived by both the spectrum owners and the unlicensed users. It is worth noting that if we do not say anything else, the concept does not include any sense of fairness. It just increases the sum of the utilities but that does not imply an equal contribution from each. That clearly has an effect on the motivation of the entities and it has been studied by multiple works like [149, 150, 125, 48]. Social welfare with a fairness consideration could be seen as optimizing the usage of spectrum for the benefit of the whole community and therefore, it would be the objective to be pursued by the regulator authorities. This objective should be carefully balanced with the next one.

- *Maximizing the benefit of the licensed spectrum owner.* "Benefit" in the sense of the utility function we shown in the previous subsection: as the combined positive and negative effects of their sale, in a variety of forms and constraints. As explained in the introduction, any form of spectrum sharing may affect licensed operators and even if that does not occur, they will be willing to receive a compensation in order to allow coexistence. Their cooperation will be boosted if they receive a proper incentive.

- *Maximizing the benefit of a secondary spectrum provider.* A secondary spectrum provider is an entity that buys spectrum from a primary licensed owner and sells it to unlicensed users (similar to a virtual operator renting a cellular access network). This objective could be considered a particular way to give incentives to primary operators (POs). As primary operators would be able to sell their unused spectrum to these entities, they would be less reluctant to make investments in spectrum. The risk they incur is lower. In addition, changes in spectrum ownership would be more dynamic and the ability to match demands of it would increase, as these secondary operators (SOs) could focus on making direct transactions and would not go through the process of asking for long-term licenses

There are few works whose objective is maximizing the benefit of the unlicensed users ([151, 139, 152] to cite some) and thus it has not been included as a main objective in the previous list. It would go against the idea of giving an extra incentive to primary operators. Again, we would like to remark that this does not mean that their benefit has not been taken into account: every model still aims to be incentive-compatible for every entity involved and thus, secondary utilities are also considered. There are also other approaches that do not fit exactly in any of the previous categories such as [139, 153], which leave the door open to use any optimization criteria from above. Table 6.2 provides a classification of different previous works by their economic objective.

What goods can be traded

On the side of the spectrum owner, the offered good consists of spectrum opportunities. These opportunities can be

- Frequency bandwidth (FDMA-like Medium Access Control [MAC]) [151, 116, 140, 143, 148, 137, 118, 119, 131, 132, 135].

- Power (CDMA) [122, 48, 47, 123, 111, 112, 113].

- Time[1] (TDMA) [126, 106, 107, 109, 110, 130].

- Transmission rate or capacity [139, 141, 114] .

- Admission to the system [121, 105, 49].

- A combination of previous resources (referred to as "channels" without specification) [138, 153, 128, 142, 120, 152, 129, 136, 133].

Although most of the models are flexible to use any of these ways of considering spectrum opportunities, they may influence the formulation of utility/profit functions (*e.g.,* if power is being traded, the impact of channel gains on the utility/profit functions may be higher)

On the side of the spectrum buyer, spectrum is typically traded for money, but not always. There is a current trend in the design of trading mechanisms that considers the possibility of SUs offering their cooperation to the spectrum owner, increasing the coverage, the battery life and/or the throughput of the licensed network. This increasingly popular approach is known as *Cooperative Secondary Spectrum Access* (CSSA). It also receives the names of *Cooperative Spectrum Sharing* (CSS) or *Cooperative Cognitive Radio Networks* (CCRNs) [106, 107, 108, 113, 109, 110, 154, 155, 156, 157].

The premise is simple: a primary user will lease some transmission opportunities if the SUs, in exchange, act as transmission relays for the primary nodes. This is different from other market-driven spectrum sharing approaches in that it fosters the creation of transmission opportunities in the PU spectrum in an exogenous way: by increasing the transmission rate of the PU, it reduces its spectrum usage. Because of that, CSSA is particularly useful when the PU own demands are so high that it rarely has spectrum to lease.

One of the earlier proposals [106], considers one primary transmitter and receiver, and the transmission time is divided into blocks, each block containing three stages. In the first stage, the PU communicates with its intended receiver via direct link and SUs also receive the information (broadcast). On the second stage, some SUs re-transmit the same information to the primary receiver. Finally, the secondary transmitters can communicate with their intended secondary receivers. Before each transmission block, the PU tries to optimize its profit function by selecting which SUs it is willing to use as relays and determining a time assignment for each stage based on knowledge about what SUs will do ("best response") depending on the system status (*e.g.,* channel gains). After the PU makes its decision public, the selected SUs try to maximize their utilities taking into account that they will use the same amount of power for cooperating with the PU as for their own transmission. That election includes the possibility of not transmitting at all

---

[1]Most papers outside this category also consider the time dimension, since channels are not leased forever. However, the difference is that they consider a fixed amount of time and do not charge a price per timeshare unit.

in case the bandwidth offered by the PU is not worth acting as relay. That decision is modeled with a non-cooperative power control game.

Given that framework, in [107] the authors propose the consideration that a PU has certain traffic demand and once it is satisfied, it will have no incentive in improving its transmission rate and therefore it will not be willing to allow SUs to access its spectrum. In order to avoid that situation, the authors propose that SUs should also pay some monetary value to the primary. They also consider a different MAC for SUs, TDMA, rather than power control, claiming that simultaneous transmission of all SUs on an interference channel, as in [106], is unrealistic because a SNR constraint cannot be met. In [107] SUs' interactions are modeled with a non cooperative payment selection game: they have to decide how much they are willing to pay and transmission time is assigned proportionally to the amount each secondary paid with respect to the total paid by all users. Two later works [108, 109], consider more than one primary and secondary links. In the first one, [108], the framework is extended to consider a primary and a secondary networks, both centralized. This implies that there are more than one primary transmitter-receiver pair and optimal selection of relays needs to be considered. On the other hand, SUs no longer compete as the central SUs' infrastructure takes into account aggregated utility, and thus, the payment mechanism changes. However, the proposal in [109], does the opposite, and assumes multiple selfish entities with no central authority. In addition, the authors use a different technique, coalitional games, where POs form cooperating coalitions with SUs, either involving a payment mechanism or no monetary exchange at all. A different mathematical approach is proposed in [110] and [154], which apply the idea of contract-based trading and study different cases of incomplete information: the authors consider different types of SUs, according to their willingness to pay and their link qualities with PU receivers. An optimal contract between the PU and each SU type comprises the SU transmission powers (for PU and SU data), and the time allocated for the SU own data transmission. The contracts are designed such that each type of user will choose the contract that is designed for its type while maximizing the PU's utility. CSSA under incomplete information is also addressed by the latest works [155, 156], applying bargaining and an auction-like mechanism, both with learning algorithms. A recent survey on the topic can be found in [100].

Other related works explore alternative flavors of CSSA. For example, paper [157] develops a model in which the collaborating SUs are rewarded in the time domain, with increased access opportunities to a base station (or access node), instead of bandwidth; [158] explores the possibility of both PUs and SUs acting as relays of each other assuming a cooperative attitude between them; and [159] shows the optimal strategy of a SU which is allowed to dynamically decide whether it enters a CSSA or a spectrum leasing market.

A close related topic to CSSA is the *overlay* cognitive radio network paradigm [160], by which a SU acts as relay for the PU and, at the same time, transmits its own message. The interference constraint on the SU transmission is enlarged thanks to the cooperative transmission and the use of interference cancellation techniques [161].

Market mechanisms

As we showed earlier, buyers and sellers have conflicting objectives. They have to agree on the conditions of the trade, mainly amount of the traded good and its price, but they can also negotiate its quality, whether it is a temporal or a permanent lease of its exploitation rights, etc. Their negotiation process follows one of the following general mechanisms: *pricing, auction* or *bargaining*.

Pricing

Initially, it was the most common mechanism in spectrum trading works. The sellers compute their offers and announce them. Buyers choose whether to connect with the one they are most interested in or not to establish any connection. It is almost like a "supermarket of spectrum" but with the associated peculiarities described when discussing spectrum as a traded good.

In order to find the optimal price that would maximize the seller's utility function (profit), the seller would need to know information such as the demand function of the buyers or the existence and strategy of other buyers. Early research in pricing assumed the sellers know that information or neglected the cost of obtaining it. That made pricing an interesting approach at first because of the simplicity of its mathematical models compared to other market mechanisms, while allowing optimal (even closed-form) solutions. Additionally, under some mathematical models, pricing could become completely decentralized, allowing more scalability and speed.

Nevertheless, obtaining that information implies, in practice, more communication between the entities. This overhead grows exponentially with their number of participants in the market and causes a delay that may harm the optimality of the trade (see section 6.3). Besides, it is very unlikely that such information would be revealed in a competitive market (see section 6.3 "Imperfect information").

The more interest was put in imperfect information markets, the less attention was given to pricing in its traditional form. Instead, the research effort is currently focused on different market mechanisms and, more recently, on a different pricing technique known as contract theory (see section 6.4).

Taking that into account, the essential advantage of pricing is how decentralized and fast it could still be: each seller could set its own price based on a local estimation (or we could even think of some regulator authority fixing some rules or public prices), and announce it. A buyer contrasts the offers it can access and simply connects to the most convenient with no negotiation, no additional message exchange. Local estimations avoid the incomplete information inconvenience and maintain pricing decentralization, albeit with a suboptimal solution, compared to an ideal complete information optimization.

Auction

Auctions have received much attention in spectrum trading and they are the most active research topic today. An auction is a centralized approach, where a regulator authority or a spectrum owner itself, gathers all the relevant information from buyers

and sellers and makes a decision that optimizes a predefined objective. It is worth highlighting that spectrum trading auctions have much lower geographical and time scales than traditional FCC-style auctions to allocate spectrum licenses.

Auctions have been used along History to sell items and specially to sell limited common goods [162, 163] because of their notion of giving an item to the one that valuates it the most and because the transaction can take place in public removing suspicions about the regulators benefiting any buyer for a bribe.

In its most basic form, buyers submit "bids" to the auctioneer according to their valuation of the item and thus, the buyer who most valuates the auctioned good, makes the highest bid and gets the item. By assigning the spectrum to the ones that valuate them most, social welfare could be maximized (although they can instead be designed to focus on optimizing the profit of the seller).

This is not as straightforward as it could be seen. For example, focusing on social welfare as a strict goal could lead to starvation of some buyers with less "bid power" (*i.e.,* smaller budgets), discouraging them from competition and finally, ending up with a lower social welfare [125]. The system could be degraded in some other way, for instance due to "vindictive bidding", in which users with no chance of winning or no real interest in the auctioned good increase their bids just to make winning users pay a higher price.

Another strong point of auctions for spectrum trading is that they could be designed to handle the imperfect information issue which strongly affected the pricing mechanism (see section 6.3 "Imperfect information").

There are several types of auctions and the challenge is to properly set their rules in order to achieve different sub-objectives. Even so, it is also unlikely that the seller would simply let the auction run by some rules and accept the result, without really getting involved in it (even if it is the auctioneer). In consequence, there are several mechanisms for the sellers to intervene, such as setting a reservation price (it will not sell unless bids reach some value) or double auctions, where multiple sellers also bid with the price they are willing to sell for, while multiple buyers are also bidding [134, 136, 137].

Under the same imperfect information scenario, auctions may achieve more efficient solutions than pricing but doing so under a strict time constraint remains unresolved.

Bargaining

Bargaining was considered since the beginning of spectrum trading, but only by a few works. It is getting more attention presently. Like an old fashioned market, buyer and seller agree (or not) on a price as a result of a negotiation. Bargaining in spectrum trading has only been studied in one-to-one trades. If there are multiple buyers and sellers, negotiation could be decomposed in pairs ("one-on-one bargaining"). The negotiating pairs are matched in different ways, e.g., by initiative of the PU as in [106] or by random matching as in [138].

One-on-one bargaining have the advantage of being the most decentralized trading mechanism of all, and has the least communication overhead. Because

of that, as we discuss in section 6.6, we believe it its one of the most promising approaches, at least in the short term. When there are more than a pair of entities involved, the one-on-one approach is suboptimal, but is an effective decomposition strategy for the more complex many-to-many bargaining problem [164], where each node would locally optimize its choice of bargaining partner (based on, for example, previous trades, geographical position, estimations of channel gains, etc.).

6.3   Spectrum as a traded good

As we sketched out in the introduction of this chapter, the radio-electric spectrum has some distinctive aspects with respect to conventional goods and commodities, that need to be taken into account in the design of spectrum trading mechanisms. An extensive survey of the literature allowed us to classify these characteristics in four main groups: *real-time variations, imperfect information, spatial reuse*, and *heterogeneity.*

Real-time variations

In a wireless network, the traffic generated at a particular region can vary rapidly over time. For example, as it would happen in a sport event, a conference, sensors upon an alert, or even day-to-day uses may have irregular patterns. Apart from that, channel conditions (*i.e,* variations on the channel gain because of multipath propagation, shadowing, interferences, etc.) could also vary quickly, having an impact on users preference for different spectrum bands (*e.g.,* a user may not need much bandwidth in a good channel condition because there is less probability of errors).

Adapting to these variations would lead to a more efficient use of resources, no matter the objective pursued by the trading mechanism. For that reason, speed is important in spectrum trading algorithms. On the other hand, the more aspects an algorithm takes into account, the better solution it will compute for a particular scenario, but the slower it will become. Therefore, there is a trade-off between speed and complexity but both properties are required to achieve an efficient solution. This is particularly remarked in works like [153, 165].

Spectrum trading mechanisms consume time on the execution of the algorithms (*computational overhead*) and on the message exchange (*communication overhead*), required to disseminate the information upon which the agents make their decisions and perform the trading.

To the best of our knowledge, there are no works measuring the impact of the computational time, which is critical in frameworks (especially auctions) that easily become NP-complete. Communication overhead is not usually considered [137, 131] and, regarding battery-constrained sensor networks, for example, it is not only important in the sense of time but also in the sense of power consumed: it would be interesting to include it in the trading algorithms so that it may not be a good strategy to even try to perform a spectrum trading if the probability of obtaining spectrum is low.

After a careful review of the existing works, it can be concluded that this issue should receive more attention. It is crucial for developing efficient algorithms and thus, efficient trading systems, which is the ultimate goal of the spectrum trading paradigm. See section 6.6 for a detailed explanation.

Imperfect information

Most spectrum trading algorithms assumed that the entities involved in them (either the spectrum owner or the auctioneer alone, in a centralized approach, or both sellers and buyers, in a decentralized one) have global knowledge of the network, such as number of competitors, channel gains and even private information like the valuation that each player gives to the traded good. It is unlikely that in a competitive environment these entities would reveal that information, especially taking into account that they may improve their utilities by hiding it or even lying, at the cost of worsening the whole system. This would also imply an additional computational overhead on all entities because they would need to find which would be their best response rather than simply telling their true valuation of the items.

Auctions could be designed in such a way that there are no incentives for cheating, which it is called the "truthfulness" property [153, 125, 128, 120, 133, 134, 136]. This is done by designing the auction so that reporting their true valuation is a dominant strategy, in other words, bidding their true valuation rewards them with a higher utility than with any other strategy. A very popular classical auction design for truthfulness is the Vickery-Clarke-Groves (VCG) auction (see [166]), that may be suitable for spectrum auctions because it is also multi-unit (packs of channels could be sold). Nevertheless, VCG and other classical mechanisms may not meet the requirements of a real-time market and become computationally intractable, especially when considering other aspects such as spatial reuse. In addition, suboptimal approaches of these algorithms lose their truthfulness property, and therefore efforts are oriented to combine the objectives of remaining truthful while dealing with the particularities of spectrum and being efficient.

Apart from truthfulness in auctions, there is also a technique to achieve truthfulness in pricing, which is called "contract theory" [110, 47], and it is based on offering different combinations of spectrum and price, computed through estimations.

The rest of the approaches are based on learning algorithms, using information from previous trades to perform estimations of the missing information. The most popular and currently active method to deal with the lack of complete information is dynamic games [151, 141, 116, 148, 117, 152, 127, 117] which are explained in Section 6.4. But there are others such as stochastic learning as in [142], or other learning algorithms in market-equilibrium approaches [129, 130, 131, 98, 132].

Spatial Reuse

This feature is observed in [153, 125, 128, 121, 126, 165]. Traditionally, licenses are granted for a wide geographical area, larger than the radio coverage area of

users. Therefore, the spectrum owner could divide spectrum usage rights into different non-interfering regions. That is to say, unlike most typical traded goods, the same spectrum band could be sold more than once to non-conflicting users to increase the income.

   Fig. 6.2 shows an illustrative example of one operator willing to sell one unused channel to three cognitive pairs with overlapping broadcast regions. Assuming that an auction is being used, the user in the middle, with the highest bid, would win the auction and thus the channel. However, the operator would get more revenue and the social welfare would be optimized if it assigned the channel to the non-interfering users of the side. Following this policy, however, there is an additional problem of starvation for the users in the middle, as they would need to bid higher than the combined bid of the ones of the side. Under these considerations, the problem of optimal allocation has been shown to be NP-hard [125] and suboptimal algorithms have to be used.
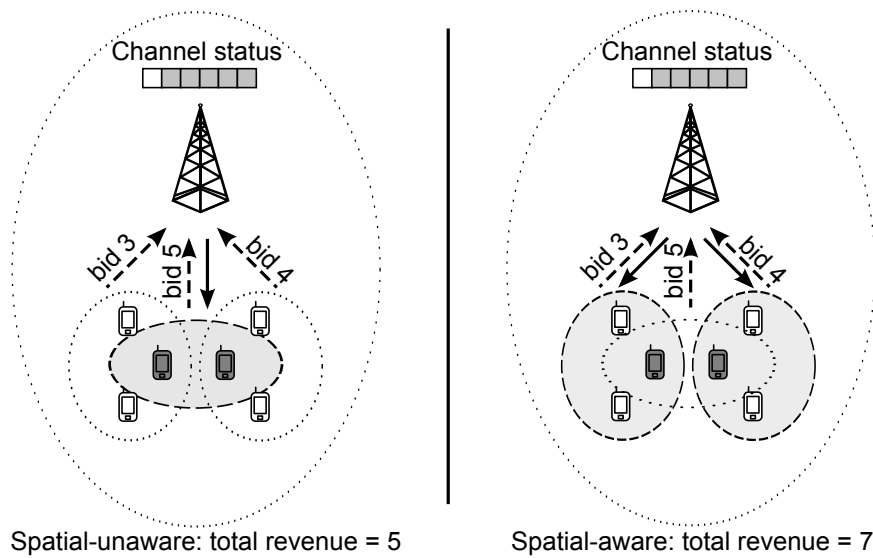

Spatial-unaware: total revenue = 5   Spatial-aware: total revenue = 7

Figure 6.2: Example comparison of a spatial unaware and aware spectrum owner

Heterogeneity

Spectrum heterogeneity [47, 110, 142, 141, 116, 146, 117, 130] refers to the consideration that different portions of the good have different characteristics. This issue is explained in detail in [142], but basically the idea behind it is that lower frequency signals travel better through obstacles and further at the cost of less amount of information transmitted and lower directivity for the transmission antennas. Another aspect considered to differentiate the spectrum is the continuity of the spectrum opportunity: it may consist of several, non-contiguous blocks, and the transmission hardware has a limit on the amount of discontinuous bands that can be used simultaneously. Finally, the spectrum QoS is perceived differently

depending on the geographical distribution of the nodes or the amount of traffic carried in it.

All this leads to different valuations by users depending on their preferences and thus, it has an impact on their utility functions, demands and service selection (even selection of more than one operator [141]) which enable more niche markets that should be taken into account in pricing. For example, [142, 146, 117] include a physical propagation model (reflecting the heterogeneity of spectrum) into the utility functions of users, *i.e,* utility depends on the received power at the SU receiver, and this power depends on parameters such as the central frequency of the channel assigned for that communication. Different channels sold by different spectrum owners would have different frequencies and thus, different valuations for the SUs.

Another example: [116, 130] include a parameter ("substitutability factor") that represents the ability of users to switch frequencies freely within the spectrum. In a duopoly, the effect on the demands of the SUs would be as follows:

$$D_i(c_i, c_j) = \frac{k_i - c_i - \delta(k_i - c_j)}{1 - \delta^2} \tag{6.3}$$

where D_i and c_i represent the demand (for example, bandwidth) and cost per unit to the i-th operator, k_i denotes the spectral efficiency or valuation of the SUs of the spectrum sold by the duopolists and δ is the substitutability factor. When δ is large, the demand to one operator is affected by the fluctuations in cost of the spectrum of the other operator, *i.e.,* both spectrum chunks are interchangeable and users are free to switch from one to the other. In contrast, when δ is small, users are not as free and demands to each operator are independent.

Another cause of heterogeneity comes from the side of users: most works consider users with different valuations of the good, mainly through the use of parameters in their utility functions such as the spectral efficiency or any other that indicates their different willingness to pay. Along this line, there are some works that take this idea further like [142], which considers users more sensitive to quality or to price, or [47, 110] that classify users according to their channel conditions (causing different preferences about how much quality to buy), and [49, 142], considering users with different and limited budgets.

6.4 Formulation frameworks: proposed solutions

As we have shown, the designer of a spectrum trading mechanism has to decide: what objective should be optimized, which market mechanism should be used, and what special features of the spectrum as a traded good should be considered. In addition, the trading mechanisms can be mathematically characterized to obtain some insights into how each trading algorithm would achieve the design goals (although in the end, the ultimate objective is to improve spectrum utilization).

When it comes to the mathematical model, there are several techniques available to formulate it. The election of one or the other depends mainly on which of the issues of spectrum we are more interested in addressing and the preferred

Table 6.3: Classification of papers by market mechanism and formulation framework

Market mechanism	Formulation framework	Works
Pricing	Game Theory	[167, 148, 109, 151, 116, 152, 141, 127, 146, 121, 144, 145, 143, 140, 147, 117, 168, 48, 114, 122, 142]
	Contract Theory	[110, 47, 169, 154]
	Optimization	[105, 49]
	Market-equilibrium	[129, 130, 131, 98, 132, 126]
Auction	Auction games	[123, 153, 124, 135, 128, 137, 118, 119, 139, 120, 133, 134, 136, 135]
Bargaining	Bargaining games	[138, 164, 149, 115, 170]

market mechanism for doing so. Table 6.3 shows what formulation frameworks are usually associated to each market mechanism, and the works applying these associations. In this section, we explain which and how the characteristics of the spectrum trading scheme have been addressed by each modeling approach, highlighting the benefits and drawbacks of each one.

Dominant trend: game theory

Since the early works on spectrum trading, game theory [171] has been the most used framework. *Game theory* is a mathematical discipline studying situations involving individual rational decision makers (*players*) pursuing their own (and often conflicting) objectives. In the framework of game theory, the interaction among these individuals is a *game*. The actions (*moves* or *decisions*) made by the players bring outcomes ("payoffs") to them. The mapping of an action to every possible situation a player can perceive (*information state*) is a player's *strategy*, which can be, in general, random distributions over the action space. Agents are assumed to be rational in the sense that they will choose strategies that maximize, in expectation, their individual payoffs.

Clearly, the game theoretic framework matches the spectrum trading scenario. For example, from the SUs' side, each user has the goal of maximizing its utility function by requesting a portion of a finite shared resource from an operator, at some cost. The fact of sharing a limited resource makes their actions to have an influence on each other. In this case, their interactions in spectrum trading can be seen as a game, where users are the players, their moves are their requests of spectrum opportunities and their payoffs are the value of their utility functions (*e.g.* the data rate they get minus the payment to the PU). This was indeed the origin of the works that used game theory in spectrum trading: it was used to study operators' resource allocation to secondary users in works that used pricing as the market mechanism. For example, in [151], the interaction between SUs when requesting spectrum to a primary operator comes from the price of spectrum,

which depends on the total demand of all SUs. That is to say, the cost of a spectrum unit is determined by the amount of spectrum requested by other SUs, and this influences the amount of spectrum a particular SU requests. In [48] and [122] the spectrum opportunities are traded in the form of uplink power and the interaction between SUs is their mutual interference. On the other side, the cost per unit of spectrum was fixed by the spectrum owners using a different technique, optimization.

That research trend evolved to consider the spectrum owners'side also as a game, modeling their competition [142, 114, 167] and, as we remarked in section 6.2, using other market mechanisms apart from pricing. Each spectrum owner has the goal of maximizing its profit function by exploiting its spectrum for its own use and/or selling part of it, obtaining some revenue for that. However, if there is more that one owner in an area willing to sell spectrum, they will compete for attracting buyers from a common pool of users. Then, for example under a pricing market mechanism, spectrum owners would be the players, their moves would be the price per unit of good and their payoffs would be the variation of their profit functions. Another parallel evolution was to use game theory in more complex markets, from the perspective of multiple secondary operators, to discover the best joint strategy for investing in spectrum from a primary operator and selling it to SUs [141, 147, 148].

Most frequent game theoretic approaches

The basic game theoretic formulation, considering a static one-shot game where agents have full information about other agents' actions and payoffs, is currently regarded as unfeasible to model most of the more complex situations in spectrum trading. Situations involving imperfect information, successive moves from the agents and the subsequent learning from past moves, and cooperation among agents. In this subsection we focus on the approaches that are receiving more attention, and the cases where each one is applicable. These game theoretic models are:

- Dynamic games.
- Auctions.
- Bargaining games.
- Contract theory.
- Cooperative games.

Dynamic games

Dynamic games have been used from the beginning to deal with the issue of imperfect information. Dynamic games are a type of games where players interact more than once and they adjust their strategies based on what they have observed in previous rounds, *e.g.*, from their own payoffs. The main advantage is that, in contrast to static games, players are not assumed to know information such as

the actions taken by other agents and their payoffs. The drawbacks are that market is not cleared in one shot, and optimal solutions are not achieved under some circumstances, so their stability needs to be addressed.

It is important to note that "dynamic" in most of these works means to go through several iterations in a fixed situation in order to converge to the solution of the equivalent static game with complete information. This is the case of the works using repeated games [151, 116, 152, 141, 127], where the same one-shot game is played several (or even infinite) rounds. Nevertheless, dynamic games offer more than this, such as the possibility of adapting the strategies to situations that change over time, such as variations in the supply and demand, or in the spectrum value (e.g. changes in channel gains) [172]. Another aspect that can be modeled by a dynamic game (in the sense of time) is the transient period required by the players to achieve a stable equilibrium in works that focus on investing in spectrum and pricing strategies for secondary operators. These works, that try to find the optimal joint strategies for investing in spectrum from a primary operator and selling it to SUs, consider multi-layered markets, as we will further explain in section 6.5. These markets are studied as interrelated multi-stage games. An example could be a two-stage game where a lower stage models the operator selection of SUs and an upper stage models the competition of secondary operators in investing in spectrum from the primary operators. Until [148], all works treated each stage independently, assuming the others had already reached equilibrium. And finally, players in a dynamic game could consider long term profits [127, 148, 116, 138]. This would bring results closer to optimality but at the cost of more complex formulations and computational overhead.

With an additional degree of difficulty, *stochastic games*, a generalization of dynamic games which deal with uncertainty at a deeper level, are starting to be applied to spectrum trading [167]. The main difficulties regarding stochastic games arise from the lack of a mature and complete theoretic framework, in contrast to other game theoretic techniques. Not even approximate solution concepts, such as those based on multi-agent reinforcement learning, are assured to converge [173].

Even so, researchers are not focusing much on game theory applied to pricing now, as we said in section 6.2. This is in favor of other techniques that seem to naturally deal better with imperfect information such as auctions, bargaining and contracts, although these are subsets of game theory and also went through the previously commented evolutions.

Auctions

Auctions can be described as games, played among the bidders, where imperfect information refers to the players true valuation of the item, and the strategies are the bids they are submitting to the auctioneer (see Section 6.3). Although they have been applied to automated spectrum trading since the origins of its study, it is the technique that is receiving more attention nowadays. As we pointed out in section 6.3, the main challenge of the works following this trend is to develop complex auction models under imperfect information, maintaining truthfulness and

being computationally tractable in real time. An extensive survey on the topic can be found in [103].

Bargaining

With respect to bargaining, as we said in section 6.2, it is only used in its one-on-one form in spectrum trading. *Bargaining games* are a subset of game theory on their own but they can also be casted to an equivalent problem of another subset of game theory which is cooperative games (see section "C.Operators bargaining - first phase" of [115]. In bargaining games, the strategies of a player are the offers to make to the other entity and whether to accept them or not. Again, recently the idea is to incorporate dynamics to the bargaining model, so that it takes into account multiple negotiations with the same entity, trying to improve its solutions [174, 164]. Bargaining research is also in the direction of using it in multi-stage market models to speed up the trading [115, 138].

Contract Theory

Contract theory is a technique directly related to pricing. In contract theory, the spectrum seller, rather than looking for an optimal price of spectrum, offers different spectrum items (for example, channels with different qualities) at different prices. If it offers the correct combinations of items and prices, the different type of users would respond according to their true valuations of the spectrum and the seller would optimize its revenue or social welfare. The major drawback is that the spectrum seller still needs to have market power (be a monopolist). This framework is currently making its first steps so the researchers' efforts are oriented to apply it to spectrum trading in diverse scenarios, specially in Cooperative Secondary Spectrum Access (CSSA, see section 6.2).

Cooperative games

Finally, it is worth saying that almost all works try to find sets of strategies that constitute an equilibrium of the game. There are different types of equilibria (depending on the type of game), being the Nash equilibrium (NE) the most common. In an NE, no player has motivation to unilaterally change its current action because it would not improve its payoff. In other words, in an NE, the decision of each user is its best response to the decisions of the others (which are also best responses). But a particular NE does not necessarily entail optimality, that is, some or even all the agents would obtain a higher payoff with a set of actions that is not a NE. Working in such an operation point (efficient outcome) would be desirable. To achieve this goal, it could be assumed that rational agents would be willing to reach an agreement to operate in an efficient outcome. When several or all the agents do that, we say that they cooperate, and the game is referred to as a *cooperative game*.

Most spectrum trading works assume non-cooperative behavior, since it is natural to think that each user is on a free-for-all share of resources trying to maximize

its satisfaction on its own, with no communication or help from other users, which are considered competitors. However, sometimes this selfish behavior leads to inefficient outcomes (*e.g.,* over-exploitation [175], or equilibria that are not optimal) and different techniques are used to drive players to better equilibrium points, mainly through some kind of communication between them. An example of these techniques are binding agreements "enforced" by a third party, such that it is in the best interests of all players to follow the recommendations of that party. The solution is called "correlated equilibrium". Despite it has been used in dynamic spectrum access [176, 177], as far as we know it has not yet been exploited in spectrum trading works. Bargaining games are also an example of improving the efficiency of the outcome through communication.[2] Another example are "commitments", where a player takes a binding action on him and informs the others so he can persuade them to take convenient actions to him. This would be the case of the leader in Stackelberg games.

In a Stackelberg game [106, 107, 108, 113, 140, 145, 148, 157] at least one of the agents, the "leader", makes its decision first and the rest of the players ("followers") react to it. The leader assumes rationality of the followers, so that they will choose their best possible responses after seeing the leader's decision. Then the leader can force a situation where that best response of the players is one that is in the leader's interest. The solution to this game is the Stackelberg equilibrium, where the leader receives a better payoff than in a simultaneous game. The leaders of Stackelberg games may not only be the spectrum owners and the followers be the users, but also it can be used to model market where some entities have priorities over others, such as a partnership of a spectrum owner and a secondary operator [148].

Cooperative games, also called "coalitional games" ([109] and a small snippet on [142]), go a step further in the relationship among players. Coalitional games focus on studying group utilities. They are closely related to dynamic games as shown in [116], specially infinite horizon games, where cooperation is encouraged due to the long term interaction that could be established among players and the possibility to punish in future actions those who do not respect agreements. Coalitional game theory is also used to analyze collusive behaviors in [127, 168, 118], when entities of the same type (spectrum owners on one side, or users on the other side) cooperate to affect market prices, harming the rest of the players and jeopardizing the efficiency of the system. Furthermore, [168] proposes a cooperation mechanism among small spectrum owners to fight against this price war. Unfortunately, we have not found many examples of cooperative games as we point out in Section 6.6.

---

[2] "Bargaining games" term is used here in a general sense and the improvement in efficiency they bring is compared to the efficiency of other games with the same number of players. That is to say, comparing multi-player games with no communication between entities to many-to-many bargaining games or comparing any other two player game with no communication to a one-on-one bargaining game. Multi-player markets as decomposition of multiple one-on-one bargaining games are less efficient than other multi-player games.

Alternative frameworks

In this subsection we discuss other approaches that have been considered an alternative to game theoretic formulation in the past but, for several reasons, do not currently receive the attention of the spectrum trading community.

Optimization

Whereas *optimization* has been widely used for dynamic spectrum access, only a few works in the past years have modeled a spectrum trading process entirely on optimization [105, 49]. Optimization, also called "mathematical programming", is the search of the best values of some parameters such that selected objective functions are maximized or minimized, subject to a number of constraints. The main difference between game theory and these works is that the latter does not consider the other agents as rational decision makers, *i.e.* as independent agents pursuing their own objectives.

Note that this idea does not imply a tyrannical attitude from the spectrum owners. In fact, it is still interesting for them to make the trading incentive-compatible for these users, either because nobody would buy its services otherwise or because a regulator may enforce this type of optimization.

In the context of spectrum trading it is developed in a single spectrum owner market form (monopoly).[3] The parameters to look for are usually price per commodity unit (often bandwidth or power), the constraints to respect are related to degradation caps on PUs' activity and the goal of the optimization could be as diverse as pointed out when describing the design of trading mechanisms (see section 6.2).

The main advantage of optimization is that, in theory, it could achieve the best solution among all techniques described, from a system perspective or in the case of only one decision maker with market power (a monopolist). The definition of "best" depends on the objective of the optimization, and thus, of that trading mechanism (see section 6.2): *e.g.* best in terms of the sum of the utility functions of all users, or in terms of fairness, best from the PU monopolist perspective in terms of its revenue, etc. Nevertheless, that is only possible in very simple models. It suffers from the "curse of dimensionality" and models become easily intractable, as explained in [172].

It is interesting to note that optimization is usually combined with other frameworks: it has been used for setting prices in monopoly works that used game theory for resource allocation among SUs (see section 6.5) and it is necessary for computing equilibria in games, for determining the winner of an auction, etc.

The efforts are mostly oriented to formulate the problem as a convex optimization problem (the objective function and the constraint sets are convex). Convexity has notable mathematical advantages and make very large optimization problems

---

[3]It could also be carried out by a third party, a regulator authority, although no examples have been found of that, probably because that would not be accepted by the entities involved as they would not have any way to influence it.

tractable. If the problem is non-convex, in general, simpler versions of the problem (sub-optimization) are solved.

When the decisions are made over stages, or time-slots, the optimization problem can be solved by means of a decomposition principle known as *dynamic programming*. This is the single-agent equivalent to a dynamic game. In spectrum trading this framework has been used in [105] to model a spectrum trading system as a Markov Decision Process (MDP). An MDP models uncertainty introducing system states and rewards that are not completely determined by the actions taken in each stage (partially random changes). A stochastic game is a generalization of an MDP to multiple decision makers. Compared to non-dynamic frameworks, the optimal strategy solving an MDP is farsighted, and therefore better in the long term than a strategy that only optimizes the payoff obtained in the current stage (myopic strategy).

Market-equilibrium

Borrowed from microeconomic theory, D. Niyato and E. Hossain develop in [129, 130, 131, 98, 132, 126] a solution to the conflict of interests between a single buyer and a single seller, later extended to the multiple buyers and sellers case. It is based on the demand and supply theory: differentiating the profit function of the seller with respect to the amount of good to trade, a supply function $S(\mathbf{c})$ can be obtained which expresses how much of the trading good should be sold given a price in order to maximize its satisfaction. Similarly, a demand function $D(\mathbf{c})$ can be obtained for the buyer, differentiating its utility functions with respect to the amount of good to buy, showing how much of the trading good should be bought given a price in order to maximize its satisfaction. The intersection of both functions is the *market-equilibrium*, where supply equals demand.

Solving that equation may imply having perfect information of the network (*e.g,* the demand function could depend on the channels gains) which is a hard constraint to meet in a real environment, so they proposed several distributed and iterative learning algorithms to reach that solution.

In [131], the authors compare the market equilibrium mechanism against two other approaches, competition and cooperation among entities. Their results show that market-equilibrium is the most stable approach and implies a low communication overhead, at the cost of a lower profit for the spectrum sellers and lower social welfare, as well as certain requirements on the available bandwidth for sale (otherwise equilibrium will not exist).

In [129] the authors extend the idea to create interrelated multi-level markets where there is a first market between primary and secondary services, another between secondary and tertiary (because the secondary can also sell spectrum opportunities) and so on. These multi-level markets may be different categories of services or multi-hop networks. A practical point of view is presented in [132], where the authors apply their model to an integrated WiFi-WiMAX network as in Fig. 6.3. In this practical example, a WiMAX base station (primary provider) offers delay-sensitive access as its primary service, but it can also sell secondary

access in the form of loss-sensitive connections. The bandwidth available to sell to secondary providers depends on the PUs' demand and that trade-off is expressed in the primary provider profit function. The secondary providers are able to re-sell that spectrum to a tertiary service provider at the same time, in a similar trade-off to the one that faces the primary service. This tertiary service provider is a best-effort WiFi access point. This work is specially interesting because it shows the application of their trading algorithm to a specific scenario and technologies.

This technique has not been completely abandoned since the work in [126] is from 2011 and extends the idea further by applying this market equilibrium model to a market with several entities that can buy and sell spectrum and other features as spatial reuse, budget constraints and multiple channels to buy/sell.
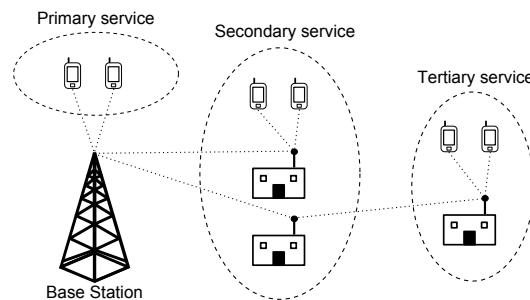
Figure 6.3: Application of [132] to a WiFi-WiMax network

We present in Fig. 6.4 a diagram summarizing the trade-off between how satisfactory the solution of a particular scenario is with each framework and how difficult it is to implement in a real scenario due to the communication and computation overheads (with regard to trades in real-time). This figure also features the research status of each approach. The color of each circle is related to the attention that the technique is currently receiving and the size of the circle is proportional to the amount of works using this technique. Our conclusion, as we develop in the final section of the chapter, is that attention should now be directed more to the practical issues of implementing spectrum trading algorithms rather than trying to push their analytical optimality forward: the key of automated spectrum trading should be adaptability to time-varying spectrum and self-organization (distributed, local based algorithms), not only to "make them happen" with the real phenomena they have to deal, but also to convince actual and potential spectrum owners, who are the ones that would have to actively adopt them in the end.

6.5 Market forms: possible situations of the market.

In this section we present some spectrum trading systems with an holistic view, focusing on the entities involved and the relationships they establish between each other. We classify them in *monopoly* and *oligopoly* markets, based on the number of spectrum sellers (see table 6.4).
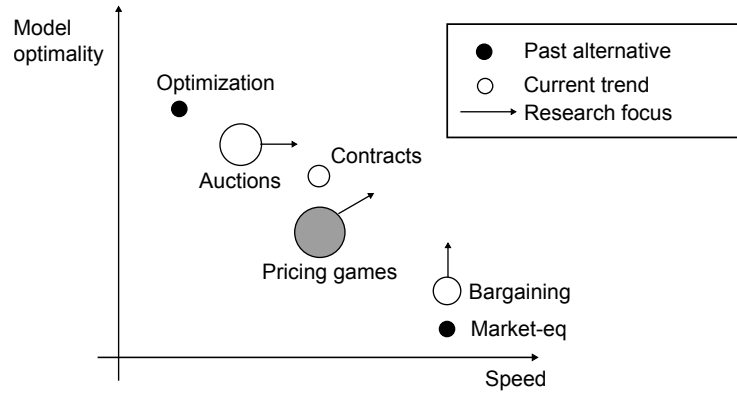
Figure 6.4: *Formulation frameworks: optimality of solutions versus implementability. The size of each circle represents the amount of works in the area. Game theory (pricing) is shaded because interest in it is decreasing*

Monopoly

In a monopoly situation only one entity operates as spectrum seller without any competition, and multiple SUs buy the spectrum opportunities offered by this single entity.

This situation is clearly less favorable for the spectrum buyers than the case of multiple sellers, where an increased competition would lower the prices and spectrum would be used in a more efficient way. However, it can be found in many practical situations, *e.g.,* if only one operator is willing to trade its spectrum; if there are several spectrum sellers and all but one are congested with primary traffic; if in a given area only one operator has infrastructure to provide coverage in the spectrum bands that are traded.

The most common market mechanism in this situation is pricing, where SUs make their own decisions about the amount of spectrum they want, although their actions may influence each other. In this case, SUs are modeled as agents in a game-theoretic formulation. On the other hand the monopolist optimizes the spectrum price, assuming it knows how SUs would react to the price. Most works consider the actions of the monopolist as part of the game [106, 111, 107, 112, 113, 108].

We have previously mentioned that the game theory pricing approach evolved to contract theory in order to handle the imperfect information issue. L. Gao *et al.* in [47] and later together with L. Duan *et al.* in [110] introduced and developed the concept of "quality discrimination spectrum trading": a spectrum trading market where SUs are classified into multiple (discrete) categories according to their preference for a given spectrum quality when buying a channel. In the specific case considered in these works, "quality" stands for the maximum allowable transmission power. All computational effort lays at the monopolist's base station or access point, which has to obtain the optimal set of qualities-prices and associate them to each SU consumer type, so that PU's revenue is maximized while making

Table 6.4: Classification of papers by market form.

Market forms	Works
Monopoly	[151, 105, 122, 48, 47, 49, 138, 123, 153, 124, 135, 125, 128, 106, 111, 107, 112, 113, 108, 110, 140, 145, 120]
Oligopoly	[109, 138, 142, 141, 116, 114, 143, 144, 148, 115, 146, 147, 168, 117, 137, 118, 119, 139, 134, 136]

these associations (called "contracts") compatible with the interests of the SUs, *i.e.,* each SU from a type should consider its assigned quality-price as its best option, even considering not to transfer at all.

Another popular approach in monopolies are auctions, [123, 153, 124, 135, 125, 128, 120] where the winner determination is formulated as an optimization problem for the monopolist. For the SUs, two approaches can be identified: in truthful mechanisms, they only have to submit their true valuations of the traded good [153, 125, 128, 120]; in other works, the bids submitted by the SUs correspond to the outcome of a game played among them [123, 124].

Other models rely entirely on optimization, such as [105]. In that work the monopolist uses dynamic programming to obtain policies on when to allow or block SUs access to its channels, considering that the price charged will vary as a function of the occupation of the system. That variation on the charged price will also influence the arrival rate of SUs (decreasing with price).

A fully decentralized approach is presented in [122], where SUs are modeled in a non-cooperative game to determine the amount of power they are using to transmit (power is the traded good), given a price, which is computed by each user in a distributed fashion, using only local information (e.g, channel gains of its neighbors), avoiding the need of global information. The monopoly part of [49] proposes a distributed optimization similar to [122] among the monopolist and the SUs, where the monopolist computes part of the parameters for the pricing and broadcasts them to the SUs, which infer the rest.

Oligopoly

A monopoly may not be the most usual situation, as it is highly likely that, in a region, more than one spectrum owner will be interested in selling, and furthermore, it is a desirable situation from the point of view of resource exploitation and users' welfare because, in general, sellers competition lowers prices.

However, setting up a competitive spectrum trading market is not easy. As licensed spectrum owners' profit diminishes, they would be willing to use aggressive market strategies to eliminate competition and end up as monopolists or to forbid entrance to newcomers. For example, licensed operators may acquire spectrum not to increase service quality or to provide a new service but to make the others fail to acquire it and thus lose their competitive power. Or, they may acquire extra spectrum to speculate with it: obtaining benefit from the trading itself. Collusions, as we introduced in section 6.4 when describing cooperative games, also belong to

this category.

To avoid these issues, a well designed and enforced regulation can help to encourage and protect competition while keeping incentives for spectrum owners. This issue is addressed in [95, 168]. In particular, in [95], this effect of competition can be seen in a comparative study of secondary-use frameworks (auction, pricing and brokerage). The authors also propose and analyze the effect of a maximum amount of spectrum trading on the different frameworks and show when each of them is more convenient to use depending on the state of maturity of the market. They argue that the final objective should be to use pricing, as it is the one that produced better results under their experiments, considering not only profits but fair competition issues, users' utilities, and spectrum efficiency. Aggressive strategies are studied in [168] from two perspectives: the short-term aggressive strategies to capture more users of the market and the long-term predatory pricing. It also shows a cooperative responding strategy for the smaller spectrum owners to avoid monopoly.

Three-layered market model

Competition among sellers is introduced in most of these works on a three-layered market model as in Fig. 6.5, composed of spectrum owners whose spectrum is licensed by a regulatory entity (such as the FCC) and provides service to a set of PUs on long term subscriptions, and secondary spectrum operators which buy unused spectrum to the primary providers and sell it to SUs in a typically shorter time basis (even real-time/on demand). The first layer of the market is the (generally auction) process where primary operators (POs) get spectrum licenses. In the second layer those operators sell portions of their spectrum to the secondary operators (SOs). In the third layer market, which operates with the smallest time-scale, the SOs sell their services to the end users over the leased spectrum.

The market layers should not be confused with the number of game stages modeling these markets. Considering that the focus is in the second and third market layers, the game usually comprises three stages: spectrum investment (SOs buying spectrum), spectrum pricing (SOs setting a price) and demand distribution (SUs selecting SO). However, a market model can have more than three layers, by allowing SUs to re-sell spectrum, establishing ternary, quaternary markets, and so forth. Most works on this section focus on SOs.

The reason of existence of these virtual[4] intermediaries to do the secondary trading is that they could focus on it, being near the end costumer and thus, crafting tailored plans and/or adding innovative or exclusive services [143, 178].

In addition, intermediaries provide a way to ease the entrance to the market to a new PO and therfore, to improve users' welfare and spectrum utilization as well as providing incentives for existing POs since it gives them the option of selling their bandwidth excess without having to worry about designing pricing and/or marketing plans for SUs to buy it.

---

[4] "Virtual" as they have no spectrum license and may not even have infrastructure as a Mobile Virtual Network Operator "MVNO" for cellular networks.
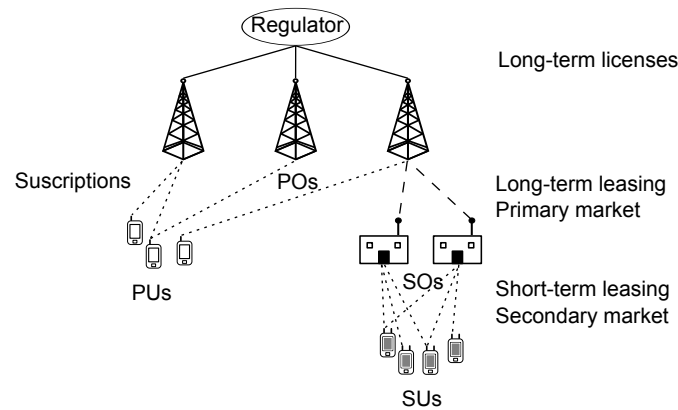
Figure 6.5: Three-layered market

Some works introduce competition into the pricing problem and focus on this middle layer by jointly studying spectrum investment and pricing [141, 140, 143, 144, 145, 147][5]. There are also works outside this three-layered market model [139, 142, 114, 117] even with a different relation between primary and secondary operators [115].

Focusing on SOs as intermediary entities, it is interesting to evaluate the impact of the price they pay to the spectrum owners on the final price charged to SUs. There are several issues of concern around spectrum investment: some works like [141] assume that, no matter if the spectrum is bought to the regulator or to PUs, these are long-term transactions (*i.e.,* for years), a legacy from the old command-and-control law framework, while SOs are looking to satisfy SUs' needs almost on demand. Therefore SOs try to adapt real time dynamic pricing strategies for SUs to those long-term investment decisions. Another point of interest is that competition can also take place at the investment stage among SOs either directly, where they are trying to get "best quality" spectrum, *e.g,* [147], and/or they are charged based on their aggregate demand [141]; or indirectly, as the bandwidth amount offered by each operator to the SUs and its aggregate total will influence prices and profits [143]. And while most of the works on spectrum trading only consider leasing as the way to obtain spectrum, L. Duan, J. Huang et. al. have studied the influence of spectrum sensing in [140, 145] as an additional way to obtain spectrum, proving that this possibility always increases SO's expected profit and users' payoffs.

Another alternative is sharing a spectrum band, proposed in [114], where two operators have a fixed (licensed) part of spectrum and an unlicensed common band which can be used under congestion and for free. Their results show that when the unlicensed to licensed band ratio increases, the profit of the providers decreases as well as the social welfare. On the other hand, users' welfare increases, so it is the role of the regulator to fix the amount of shared band in this trade-off.

---

[5]Works [140, 145] do not really study sellers' competition but they are closely related to other works featured in this section, showing the same structure and dealing with investment as if they were on a competitive environment.

Other market models

Several works show a different market structure. In [139] a variation of a private commons regime is modeled (the authors call it "mixed commons/property-rights") with a centralized Spectrum Policy Server (SPS) per geographic domain. When each user enters the system, it connects to this broker which obtains its location and a function called "acceptance probability of services" related to its utility function. Then, the SPS starts an iterative bidding process of service offers from operators where the winning bid is the one with the highest acceptance probability, which is shown to the user and accepted with that probability. An extension to a multi-user environment is also studied. The authors of [142] consider that a centralized controller does not match the reality in a system that is distributed by nature and proposed a similar model with operators and users interacting directly. In [117], the authors are also interested in a different market model, without brokers or any centralized entity, where the primary base stations should be the ones to also serve SUs by selling them primaries' unused spectrum directly. The authors of [115] are close to the three-layered market model (it is a three-layered market indeed) and study DSA but in a *Mobile Virtual Network Operator* fashion, *i.e.,* when the virtual operator buys spectrum to the licensed one, it becomes its competitor for a common pool of users and both play a non-cooperative game on the price per subscription charged to those users.

6.6   Open research lines and future trends

Based on the drawbacks of current proposals, which we have pointed out in the preceding sections, and considering the issues that are recently receiving more attention from the research community, we would like to highlight the following unresolved challenges:

Real-time adaptation versus optimality

In spectrum trading research, the main tendency has been building increasingly complex models, which deal with more issues of spectrum in more convoluted markets, and aiming to solve them optimally. Nonetheless, the resources required to reach such elaborated solutions have not been considered enough, specially time. Time consumption is a relevant issue because, as we have been pointing out throughout our survey, multiple important parameters in spectrum trading experience rapid variations over time: spectrum opportunities, demand (unplanned peaks of traffic), valuation of the spectrum (changes in channel gains, mobility of entities), etc.

In consequence, we think that time consumption should be regarded as a key feature in spectrum trading algorithms: if obtaining an optimal solution takes so long that the system's parameters vary significantly during the computation time, this solution will be not optimal when applied. In addition, the more time devoted to negotiation among the agents, the less time used in transmission. For that reason, it is more practical to design models that may not fully exploit spectrum in a

particular moment (for example, by not taking into account spectrum geographical reuse) but that are capable of reducing the uncertainty due to the changes over time, and thus, providing higher guarantees on the agents' satisfaction.

According to the computation and communication overhead, it would be fair to say that the most suitable frameworks to achieve real-time adaptation are those that are decentralized and compute solutions based on local observations (imperfect information), otherwise it would imply lots of entities communicating with a central node. Despite that idea, achieving real time operation in complex auction models is a hot topic in spectrum trading research [150, 179, 138]. This is due to the fact that optimality is still in the spotlight, as pointed out in [150]: "The decentralized approach has several advantages that make it attractive: lower complexity than the centralized approach, robustness and scalability. However, with decentralized approaches there is no guarantee that optimal solutions can be achieved. To optimize objectives, such as global efficiency and fairness, and some important parameters, such as price of anarchy and price of stability, the spectrum data for all SUs in the network should be considered"

Nevertheless, two approaches are, in our opinion, more promising to overcome the time-efficiency issue: bargaining and contract theory. One-on-one bargaining offers simplicity on its algorithms, while it can be extended to include more features such as dynamic behavior and adaptation: the bargain process can be made over several stages in order to reach better agreements, and it can take into account past trading history and far-sighted decision making [174]. When multiple agents interact, one-on-one bargaining is not optimal in general (many-to-many bargaining is indeed optimal, but is a much more complex problem, lacking the simplicity and efficiency of the one-to-one scheme), as it is not considering all available entities and relationships as a whole. But the negotiation among pairs can be easily done in real-time, so that there are no utility losses because of computational or communication delays. In order to compensate the sub-optimality of this approach, it is possible to study beforehand which pairs of entities are best matches to each other (rather than matching pairs randomly).

Many-to-many bargaining is studied as a cooperative game where all entities must be committed to the agreement, and therefore deviating from the *grand coalition* should not be in the interest of any subset of entities. It appears to be less tractable because it would require intense message exchange between the members of the coalition. Cooperative games in general, however, are interesting because they can look for higher and longer term utilities (allowing Pareto optimal outcomes).

Pricing was seen as an effective way to deal with time variations, as each entity could compute the optimal price on its own and then users would choose the most convenient for them. However, it its simplest form, its algorithms required much knowledge about the network: for example, price was established based on how much users value spectrum (their demand) and that information is unlikely to be shared. Users could play strategically with that in order to get lower prices by reporting lower valuations. The price menu offered by contract theory tackles that imperfect information and maintains decentralization and low communication,

although, again, it comes at the cost of lower utilities with respect to an ideal global optimization.

Implementation and applications

Almost all the efforts on spectrum trading research have been devoted to the analytical study of the economic aspects of the trading. It is true that other related technical challenges, such as spectrum sensing or MAC protocols [180, 46], to cite some, are studied in cognitive radio papers, but are hardly found in spectrum trading works. However, combining technical and economic aspects in trading mechanism research, even in prototype design, would throw light on relevant, non-trivial issues. For example, how to put together the stages of a MAC protocol with those of the spectrum trading algorithms, that is to say, how the trading algorithm translates into an exchange of control packets. In fact, these issues may have an impact on the performance of the trading algorithm such that it needs to be tuned (*e.g.* to handle or reduce time delays). Nevertheless, these implementation issues have received little attention (unlike, for example, the authors of [122], who designed and described a practical MAC protocol for their algorithm). In contrast, there have been numerous MAC proposals [46] for dynamic spectrum access with cognitive radio.

The same happens with regard to applications of these models in practical network environments with specific wireless technologies. There are only few works that consider that in spectrum trading (IEEE 802.22 standard-based wireless regional area networks in TV white spaces in [132], Wi -Fi and WiMax in [181]). Because of the issues and challenges discussed in previous sections, it would be fair to say that implementation and applications of spectrum trading would need more than a simple extension of the models for dynamic spectrum access without trading.

And all this is particularly important because it threatens the main reason of existence of spectrum trading: creating an incentive to incumbent operators.

Incentives to incumbent operators

As we stated in the introduction, there are some other ways to optimize the use of spectrum apart from automated spectrum trading, such as a forced dyamic spectrum access. What makes spectrum trading different from other proposals is that it creates incentives to spectrum owners so that they would be willing to facilitate the access to their unused spectrum to other secondary entities. Creating incentives instead of forcing the operators avoids that actual spectrum owners as well as prospective ones feel discouraged from investing in new spectrum technologies and services: "why am I going to spend money in spectrum if the government is going to give it for free to others?"

Therefore, it is needed that these operators believe in an automated mechanism that controls their profits. But this idea alone is scary. "Is there any security on obtaining benefit"? The short answer is "yes" but the long answer would be "only in

theory". It is hard to convince he telecommunications industry to adopt spectrum trading if there are no test-bed experiments and there is no strong evidence that these algorithms do not lead to economic loses (due to dis-adaptation to real time, for example). Optimality, then, remains in the background.

On the other hand, secondary users also need an incentive to request spectrum from the primary operator (*e.g.* inexpensive tariffs). Otherwise, the trading market will not work because they would all prefer to become primary users.

In a different and complementary approach, something can be added to money as an incentive to spectrum owners: services, specifically, secondary users acting as relays in Cooperative Secondary Spectrum Access (as we have pointed out in section 6.2). Albeit money is always an incentive and satisfaction due to it never saturates, this adds a new and non-substitutable reason for sharing: *e.g.* secondary users can act as relays for increasing the primary operator range. Furthermore, the most promising feature of CSSA is using it when the primary operator is congested, in order to increase its transmission rate. Due to that increase, new spectrum opportunities would be generated and there would be more spectrum sharing and efficiency.

However, CSSA provides no incentive to primary operators when they have low traffic volume and they are not interested in augmenting their transmission rate. For that reason, hybrid approaches of CSSA and economic exchanges are being studied [149, 107]. New approaches are also exploring the idea of offering other services to the primaries like secure transmissions [182] or offload services [183, 184]

Protection against malfunctions

Along the same line of implementation issues, it is not difficult to think that a system is likely to behave differently as planned due to intentional or unfortunate events.

Intentional failures: untruthfulness and collusions

Although, like every aspect of social interactions, spectrum trading would have a regulation framework protected by the law, it is by no means irrelevant to design mechanisms that prevent or discourage market manipulation. It is so because of the precautionary principle (i.e. "better to be safe than sorry"), it is easier to be proactive and avoid such situations; and most of the times market manipulation could be hard to prove: for example, if an algorithm depends on entities reporting their true valuation of the spectrum, how could someone prove that they are being untruthful?

As we have previously commented, truthfulness has been studied for a few years and it is still a hot topic. The essential problem here is that combining truthfulness with some other goals such as real time, optimality, spatial reuse, etc. soon makes the algorithm intractable and there is no accepted solution to that so far.

Collusions, cooperation of a set of players to influence market prices and harm the others, have received little attention, perhaps because they could be more easily

spotted by a regulator authority, as it has happened in other markets. They are, however, a serious threat to the correct operation of the system, not only because it may turn it inefficient, but also because it can end up ruining the non-colluding entities, erasing competition and destroying the market.

Unfortunate events: lack of rationality

Failures can also occur in a system when control information is not correctly received due to propagation issues or contains any error from the source. This event would not only cause immediate loses or suboptimal trades but it can also push the system to a permanent situation of inefficient operation. All approaches assume that entities are rational and do not study the impact of such failures. They also do not consider the implementation of reactive mechanisms to make decisions under the irrational behavior that would take place. To the best of our knowledge, there are no published works addressing this aspect of spectrum trading.

Furthermore, even without considering erroneous information, irrational behavior may take place. In section 6.4 we explained how optimization was discarded as a main tool to model spectrum trading because it was affected by an inherent "curse of dimensionality" . Then, game theory, showing a decentralized approach studying entities as individual rational decision makers, became the preferred method. Sadly, the applicability of game theory beyond very limited models and the whole concept of finding an equilibrium are being called into question. In fact, [185] suggests that failure to converge to equilibrium in some complex games is independent of the learning algorithm and the behavior of players ends up being essentially random.

Complex game models

This point could be considered as a contradiction with the previous one, where we cited how complex game theory models may not be tractable. Nevertheless, we still think that some complex formulations have not been explored enough yet and feature some promising properties so as to conclude they should not be abandoned this soon.

We showed that a method to deal with the uncertainties of spectrum is trying to reduce it with simple (not necessarily optimal) real-time algorithms. However, it is not the only way to do so. Another way to do it is thinking about long term rewards and relationships using cooperative games. Cooperation and considering the history of play gives the feel that it is a more secure and profitable system: cooperation allows to reach higher profits than competition, allows fair sharing of revenues, there are punishment strategies if someone deviate from coalitions, etc. But the drawback has already been stated: "coalitional games are inherently difficult to solve" [109]. In addition, they require more communication between the entities which may push them far away from real time and may creat a dis-adaptation that may not be compensated with the long term commitments.

A promising way to achieve the benefits of cooperative games but with smaller computation and communication overheads, could be to maintain some local measure of "reputation" about the other players as in [174]. Reputation is a way to

"compress" past histories, expressing how likely is to achieve a good transaction with those other players and thus, it can be used to prioritize one player over the others. That would encourage every involved entity to share its resources.

Another interesting research line is exploring the effect of transient times in multi-stage games. An example of a multi-stage game could be that of an oligopoly of spectrum providers serving to a common group of users: at one stage, the users perform the service selection according to the prices and perceived quality; at the next stage, service providers decide the spectrum to lease and its price. Both stages are interrelated but almost all works solve them by backward induction, assuming that one of the stages is in equilibrium before including it in the other stage game. In reality, both dynamic stages would be looking for equilibrium at the same time and considering the impact of such transient states could increase profits and reduce the convergence time, at the cost of a more complex formulation [148].

Lastly, stochastic games deal with uncertainty by trying to incorporate it to the model. However, as we pointed out before in the survey, they are still a subject of research for theoretical mathematicians so its use in spectrum trading requires more than trying to apply it and still lacks of a more complete theoretical foundation.

Complex market joint studies

We have shown in section 6.5 that the most common market model in spectrum trading works is the three-layered one of Fig.6.5; where the spotlight is put on the decisions of the secondary operators in the middle layer. Even if this structure is not followed, a very extended assumption is that a primary operator provides service to its subscribers (primary users) and sells unused spectrum to a secondary operator who, in turn, tries to convince other entities (secondary users) to pay for the services provided over its leased spectrum.

But in reality, users are not primary or secondary by default: an entity willing to use spectrum would consider to become a primary user paying a subscription to a primary operator or become a secondary use served by a secondary operator. Each service profile is different: the one of the primary operator is more reliable, possibly with a higher transmission rate and better QoS. On the contrary, the secondary operator could possibly be more aggressive in price. The point is that there are some unexplored interactions such as a primary owner considering that when it its selling spectrum to a secondary operator, it becomes its competitor as in [115], or a primary operator dealing at the same time with primary and secondary users [117]. And on the users side, for example, the idea of users selecting which network they prefer, primary or secondary [186].

These overlooked considerations are indeed relevant in a real scenario and could be used to boost the efficiency of a trading market, or, at least, keep it working: let's imagine a secondary operator offering the most suitable service to users such that these users would never decide to turn into primary users, starving the primary operator.

6.7   Conclusions

Automated spectrum trading is a promising strategy for increasing spectrum efficiency while providing incentives to spectrum owners and users to take part in it. It faces challenges because of its novelty in wireless networks and the fact that techniques from other markets are not applicable due to the peculiarities of spectrum as a traded good. In consequence, it raises some concerns among the wireless industry incumbents about its feasibility and value. Current and prospective spectrum operators may not support this functionality unless their profits are guaranteed. Of course they can be forced to implement such a market but they may feel discouraged to invest in spectrum and deploy new services. The biggest obstacles to implementation are the fast variations of the environment parameters and dealing with imperfect information, while trying to reach the best possible solution. "Best solution" refers here to maximizing the utilities of the agents, including fairness considerations (depending on the economic objectives). This survey showed that the research community tends to build increasingly complex models (such as sophisticated auctions) and is trying to improve their speed. However, researchers are devoting less attention to real-time operation capabilities, which is important not only for the optimality of the solution itself (the solution of a complex model is of no use if the real situation has changed and has nothing to do with it), but also to assure that the algorithm will not be economically harmful to the agents, at least on the long run. In consequence, it can be concluded that suboptimal but faster mechanisms should receive more attention in the near future, in particular the approaches using bargaining and contract theory. Also, in the line of providing incentives to spectrum operators, cooperative secondary spectrum access could be considered an appealing approach. It adds a non-substitutable value to these operators such as increased range, transmission rate or security.

An MDP Framework for Centralized Dynamic Spectrum Leasing

7.1 Introduction

In the next two chapters we propose two automated spectrum trading mechanisms, suitable for different scenarios. The one described in this chapter is appropiate for situations of low spectrum usage by PUs. In such scenario, the spectrum owner may be willing to obtain some benefit from its unused resource by selling it to SUs. Nevertheless, it still has to meet the demands of the PUs (to some extent), assuring that the blocking probability for the PUs remains under a desired threshold. We show how to compute a policy that balances those two objectives by means of an MDP and a constrained-MDP formulations.

Motivation

In situations of low spectrum usage, a licensed operator may decide to sell spectrum opportunities to SUs in the *secondary spectrum market* to obtain extra-revenue from its resources. A regulator may encourage this behavior, for example, by setting a fee proportional to the amount of spectrum held. [187].

The operator does not know the future demands from its users in advance (apart from a statistical characterization). What would happen if, at any given moment, all of its spectrum is in use and a new PU arrives to the system? As the licensed operator, it could reclaim the spectrum it leased to SUs if needed, but the process is not inmediate and the incurred delay could result in, for example, a PU missed call. We refer to the probability of such event as the *blocking probability*. In addition, reclaiming its spectrum may not be a convenient solution for the operator (*e.g.* because it may require a compensation to the SU).

Therefore, a licensed operator should carefully balance the revenue obtained from allowing the SUs to access its spectrum, with the blocking probability of the PUs. In order to do this in real-time, a protocol is required to support negotiations on access price, channel holding time, *etc.*, between the spectrum owner and the

SUs. In this work, we consider a centralized bid-auction model, in which SUs bid for the spectrum of a single spectrum owner. We consider a continuous bid process, in which each incoming SU makes its bid upon arrival. Because of the real-time constraint, the operator only knows one bid at a time, and has to make its decision for each incoming SU, before knowing future bids. We explore the use of MDP and CMDP formulations to balance benefit and grade of service for PUs. The main advantages of this approach is that it assures operating at global optimum and reduces the computational effort at SUs.

Related Work

Centralized dynamic spectrum trading has been addressed in previous works. In the proposal by Yu *et. al.* [188], the spectrum broker controls the access of SUs based on a threshold rule computed by means of an MDP formulation with the objective of minimizing the blocking probability of secondary users. In order to cope with the non-stationarity of traffic conditions, the authors propose a finite horizon MDP instead of an infinite horizon one. The drawback is that the policy cannot be computed off-line, imposing a high computational overhead on the system. Tang *et. al.* study in [82] several admission control schemes at a centralized spectrum manager. The objective is to meet the traffic demands of SUs, increasing spectrum utilization efficiency while assuring a grade of service in terms of blocking probability to primary users. Among the schemes analyzed, the best performing one is based on a constrained Markov decision process (CMDP).

When the problem is distributed, dynamic game theory is used instead of MDP. In [152], spectrum trading from PUs to SUs is modeled as a non-cooperative dynamic game, using a Markov chain to describe groups of SUs buying opportunities from PUs. Given its distributed implementation, the main goal of this approach is finding the system's equilibrium. The objective in [153] and [120] is maximizing the profit of PUs, with especial interest in assuring bidding truthfulness. Wang *et. al.* present in [135] a bandwidth auction mechanism between an PU and multiple SUs, where the objective is to maximize the SU's payoff and reach Nash equilibrium in a competitive and distributed manner, while the PU sets a minimum remaining bandwidth for its own use as a constraint.

Our contribution

This chapter addresses the design of centralized DSA MAC protocols for real-time dynamic spectrum auction. We explore the possibilities of a formal design based on a Markov decision process (MDP) formulation and propose, in Section 7.2, a design framework to balance the grade-of-service (given by the blocking probability for PUs) and the expected economic revenue provided by the SUs' accepted bids. This trade-off can be managed in two ways. One consists of computing a single objective value given by a combination of the blocking probability and the expected revenue. The weights assigned to each objective determine the point in the Pareto front where the obtained policy lies. The other approach, which is presented in

Section 7.3, consists of solving for one of the objectives and setting constraints on the remaining ones. This strategy results in a constrained MDP formulation (CMDP) and the policies obtained are not necessarily deterministic.

7.2   System model

The model includes a spectrum bidding procedure in which SUs send their bid offers, within a finite countable set of prices (for mathematical tractability), for the use of a channel. Each bid offer consists of the price that the SU is willing to pay for each minute of channel holding time. Taking into account the trade-off between the blocking probability of PUs and the expected benefit obtained from spectrum rental, SUs can be accepted or rejected. Accepted SUs are given access to the band and they are charged for it the price they selected. PUs are always accepted if the there are available channels.

Incoming traffic is characterized by a classic Poisson model. PUs arrive with a rate of λ_L arrivals per unit of time. The arrival rate for SUs is denoted by λ_U. The licensed spectrum managed by the central controller is assumed to be divided into channels (or bands) with equal bandwidth. Each user occupies a single channel. The average holding times for PUs and SUs are given by $1/\mu_L$ and $1/\mu_U$ respectively, where μ_L and μ_U denote the departure rate for each class. Because a Poisson traffic model is considered, both the inter-arrival time and the channel holding times are exponentially distributed random variables for both user classes. The model can be easily extended including more user classes, the probability that a user occupies two or more channels, and so on. Essentially the procedure is the same, but the Markov chain would comprise more states as more features are considered in the model. In this model, the state of the Markov chain is determined by the number of channels k occupied by PUs, and the number of channels s occupied by SUs. Because spectrum is a limited resource, there is a finite number N of channels. Figure 7.1 depicts a diagram of the model and its parameters. Note that we can map all the possible combinations of (k, s) for $0 \leq k \leq N$, $0 \leq s \leq N$ and $k+s \leq N$ into a single integer i such that

$$0 \leq i \leq \frac{N(N+1)}{2} + N + 1. \tag{7.1}$$

The number in the right hand side of previous equation is the total number of states that we will denote by N_T.

The bidding prices are classified into a finite set of values: $\mathbb{B} = \{b_1, b_2, \dots b_m\}$ given in money charged per unit of time. Each price has a probability p_i of being offered by an incoming user. Obviously $\sum_{i=1}^{m} p_i = 1$. Figure 7.1 illustrates this model.

The objective of the MDP is to obtain the maximum economic profit with the minimum impact on PUs. In the framework of MDPs we have to define the actions and the costs of these actions. Let $g(i, u)$ denote the instantaneous cost of taking action u at state i. The control u at each stage determines the admitted and rejected bidding prices. Logically, the control should be defined as a threshold, *i.e.*
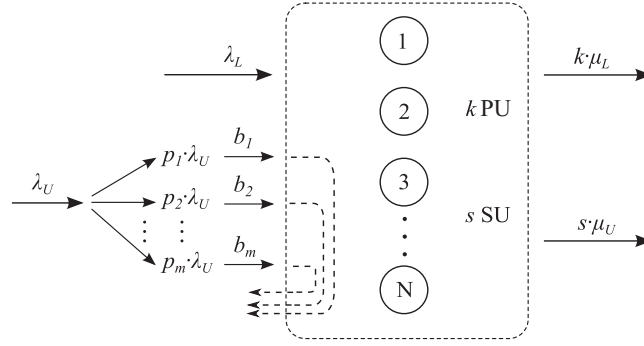
Figure 7.1: Diagram of the auction-based access model. SUs can offer up to m different bid prices. Each price is assigned a probability. The access policy decides upon each bid according to the price offered and the system's state.

when $u = i$ only bids equal or above p_i are admitted. For notation convenience, the control $u = m + 1$ indicates that no bid is accepted. The per-stage reward function $g(i, u)$ is given by the linear combination of the the blocking probability per unit of time, $g_L(i, u)$, and the expected benefit at stage i when decision u is made, $g_U(i, u)$. Therefore $g(i, u) = \alpha g_L(i, u) + \beta g_U(i, u)$ where the scalars α and β are weighting factors. Note that $\beta < 0$ since the objective is to minimize $g(i, u)$. By adjusting α and β we can compute a Pareto front for both objectives.

The function $g_L(i, u)$, is given by the probability that the system cannot provide a channel to an incoming PU:

$$g_L(i, u) = \begin{cases} 1, & \text{if } i \equiv (k, s) \text{ and } k + s = N \\ 0, & \text{otherwise} \end{cases} \tag{7.2}$$

where the symbol "$\equiv$" denotes equivalence, *i.e.* i corresponds to a state (k, s) such that $k + s = N$.

To determine $g_L(i, u)$, let B_i denote the expected income when an SU whose bidding price is b_i is accepted. Since the average channel holding time for PUs is $1/\mu_U$, then $B_i = b_i/\mu_U$. Given a control u, $P(r|u)$ denotes the conditional probability that the bidding price of the next accepted SU is b_r, and is given by

$$P(r|u) = \begin{cases} \dfrac{p_r}{\sum_{j=u}^{m} p_j}, & \text{if } r \geq u \\ 0, & \text{otherwise} \end{cases} \tag{7.3}$$

Let us define $\tilde{g}_U(i, u, j)$ as the average benefit associated to the transition from state i to state j. Its expression is

$$\tilde{g}_U(i, u, j) = \begin{cases} p_u \sum_{r=1}^{m} B_r P(r|u), & \text{if } j = i + 1 \\ 0, & \text{otherwise} \end{cases} \tag{7.4}$$

where $p_u = \lambda_U/(\lambda_U + \lambda_L)$ denotes the probability that the next arrival corresponds to an SU. Therefore, the per-stage benefit $g_U(i, u)$ is given by

$$\begin{aligned} g_U(i, u) &= \sum_{j=1}^{N_T} \tilde{g}_U(i, u, j) p_{ij}(u) \\ &= p_{i,i+1}(u) p_u \sum_{j=1}^{N_T} B_r P(r|u). \end{aligned} \tag{7.5}$$

The objective is to find a policy minimizing the expected value of the combined cost, given by

$$\lim_{K \to \infty} \frac{1}{E\{t_K\}} E \left\{ \int_0^{t_K} g\left(x(t), u(t)\right) \right\} \qquad (7.6)$$

where t_K is the completion time of the K-th transition. The problem can be solved by formulating its auxiliary discrete-time average cost problem. Let γ be a scalar greater than the transition rate out of any state of the chain, $i.e.$ $\gamma > \nu_i(u)$ for every i. We can compute the transition probabilities $\tilde{p}_{ij}(u)$ for the auxiliary discrete-time problem from the probabilities $p_{ij}(u)$ of the original problem as

$$\tilde{p}_{ij}(u) = \begin{cases} \frac{\nu_i(u)}{\gamma} p_{ij}(u) & , \text{ if } i \neq j \\ 1 - \frac{\nu_i(u)}{\gamma} & , \text{ if } i = j \end{cases} \qquad (7.7)$$

We can now formulate Bellman's equation (see [24]) to obtain the optimum average cost λ

$$\tilde{h}(i) = \min_{u \in \{0,1\}} \left[\alpha g_L(i, u) + \beta g_U(i, u) \nu_i(u) - \lambda + \sum_{j=1}^{N_T} \tilde{p}_{ij}(u) \tilde{h}(j) \right] \qquad (7.8)$$

for $i = 1, \ldots, n$. Multiplying $g_U(i, u)$ by $\nu_i(u)$ yields benefit per unit of time. It can be anticipated that the structure of this problem, essentially a connection admission control problem, requires a threshold type solution in which upcoming SUs of each class will only be admitted into the system if the number of occupied channels is below certain threshold. There will be one threshold per bidding price. By properly adjusting the weighting factors α and β we can compute a Pareto front allowing us to determine the maximum possible benefit for a given blocking objective for the PUs.

7.3 Constrained MDP

When several objectives concur in an MDP problem, another feasible approach is to optimize one of them subject to constraints on the other objectives. This strategy results in a CMDP formulation of the problem. Solving MDPs by iterative methods such as policy or value iteration allows us to find deterministic policies, $i.e.$ policies that associate each system's state $i \in S$ to a single control u. However, these policies do not, in general, solve CMDP problems. Instead, the solution of a CMDP is a randomized policy, $i.e.$ it associates each state to a probability distribution defined over the elements in $U(i)$.

There are mainly two approaches to solve CMDPs, linear programming (LP) and Lagrangian relaxation of the Bellman's equation. This proposal follows the former one. Each feasible LP formulation relies on the use of the $dual$ variables $\phi(i, u)$, defined as the stationary probability that the system is in state i and chooses action u under a given randomized stationary policy. The problem addressed in this chapter results, under every stationary policy, in a truncated birth-death process, since PUs are always accepted. In consequence, every resulting Markov chain

is *irreducible*, in other words, it is recurrent and there are not transient states. Moreover, the state and action spaces are finite. Under these circumstances, as shown in [23], every feasible solution of the LP problem corresponds to some randomized stationary policy. Therefore, if the constrained problem is feasible, then there exists an optimal randomized stationary policy.

The LP approach consists of expressing the objective and the constraints in terms of $\phi(i, u)$. Once the problem is discretized, the average cost is defined as

$$\lambda = \lim_{K \to \infty} \frac{1}{K} E \left\{ \sum_{k=0}^{K} g_U(x_k, u_k) \right\} \tag{7.9}$$

where k denotes the decision epoch of the process. The constraints are defined similarly to the main objective: each constraint imposes a bound on an average cost related to a different per-stage cost. In our case, it is given by

$$c = \lim_{K \to \infty} \frac{1}{K} E \left\{ \sum_{k=0}^{K} g_L(x_k, u_k) \right\} \leq \beta \tag{7.10}$$

where $g_L(x(t), u(t))$ is the real-valued function providing the per-stage cost associated to the constraint β. Therefore the average reward MDP with a single constraint is defined as the following minimization problem:

$$\min_{\mu} \lambda \text{ subject to } c \leq \beta \tag{7.11}$$

Given the characteristics of the problem (finite state and action spaces and recurrent Markov chain under every policy), the limits in (7.9) and (7.10) exist and are equal to

$$\lambda = \sum_{i \in S} \sum_{u \in U(i)} g_U(i, u) \phi(i, u) \tag{7.12}$$

and

$$c = \sum_{i \in S} \sum_{u \in U(i)} g_L(i, u) \phi(i, u) \tag{7.13}$$

respectively. In addition, the following conditions must be hold by the *dual* variables:

$$\sum_{u \in U(j)} \phi(j, u) = \sum_{i \in S} \sum_{u \in U(i)} p_{ij}(u) \phi(i, u) \tag{7.14}$$

for all $j \in S$, which is closely related to the balance equations of the Markov chain and

$$\sum_{i \in S} \sum_{u \in U(i)} \phi(i, u) = 1, \tag{7.15}$$

which, together with $\phi(j, u) \geq 1$ for $i \in S$ and $u \in U(i)$ correspond to the definition of $\phi(i, u)$ as a limiting average state action frequency. In consequence,

the LP for the CMDP has the following formulation:

$$
\min_{\phi} \quad \sum_{i \in S} \sum_{u \in U(i)} g_U(i, u)\, \phi(i, u)
$$

$$
\text{s.t.} \quad \sum_{i \in S} \sum_{u \in U(i)} g_L(i, u)\, \phi(i, u) \leq \beta
$$

$$
\sum_{u \in U(j)} \phi(j, u) - \sum_{i \in S} \sum_{u \in U(i)} p_{ij}(u)\, \phi(i, u) = 0 \qquad (7.16)
$$

$$
\sum_{i \in S} \sum_{u \in U(i)} \phi(i, u) = 1
$$

$$
\phi(j, u) \geq 1
$$

Assuming that the problem is feasible and ϕ^* is the optimal solution of the LP problem above, the stationary randomized optimal policy μ^* is generated by

$$
q_{\mu^*(i)}(u) = \frac{\phi^*(i, u)}{\sum_{u' \in U(i)} \phi^*(i, u')} \qquad (7.17)
$$

for cases where the sum in the denominator is nonzero. Otherwise, the state is transitory and the control is irrelevant. Note that $q_{\mu^*(i)}(u)$ denotes the probability of choosing action u at state i under policy μ^*.

7.4   Numerical Results

We will consider three scenarios characterized by the asymmetry between the traffic intensity of PUs and SUs. In every scenario, the average holding time is equal for every user, independently of their type. Therefore the service rate $\mu_L = \mu_U = 5$. Assuming that the time unit is an hour, this results in an average holding time of 12 minutes per connection. The total traffic ($\lambda = \lambda_L + \lambda_U$) is 40 calls/h, which yields a total incoming traffic of 8 Erlangs. In a wireless cell covering 2.5 km^2 of urban area (cell radius equal to 400 m), with 2000 people per km^2 and a 10% aggregate market penetration (PUs and SUs), the number of covered users is around 500, and the resulting traffic intensity is 0.016 Erlangs per user. The number of available channels is set to $N = 10$, in order to evaluate the system in a relatively congested situation. With the assumed traffic intensity we can estimate the blocking probability of the system for the aggregate traffic by means of the well-known Erlang's B formula (see [189]):

$$
E(n, \rho) = \frac{\frac{\rho^n}{n!}}{\sum_{j=0}^{j=n} \frac{\rho^j}{j!}} \qquad (7.18)
$$

where n is the number of channels and ρ denotes the utilization factor. In our case $\rho = \lambda/\mu_L = \lambda/\mu_U$. According to this formula, if the system accepted every incoming user, the total blocking probability would be $E(10, 8) = 0.12$. As we will see, this probability is an upper bound for the blocking probability of the PUs and a lower bound for the SUs.

Parameter	scenario 1	scenario 2	scenario 3
λ_L (calls/h)	30	20	10
λ_U (calls/h)	10	20	30
$\mu_L = \mu_U$ (calls/h)	5	5	5
N	10	10	10

Table 7.1: Parameters setting for the three scenarios of the priority based access problem.

The three scenarios are summarized in Table 7.1.

Additionally we define three classes of SUs, characterized by the price that they offer per minute of channel occupation. The bid offers per class are: class 1: 0.01 $/m, class 2: 0.02 $/m and class 3: 0.03 $/m. We define the probability of an SU incoming call being of each class. The SU class probability distribution is: class 1 probability: 0.5, class 2 probability: 0.3 and class 3 probability: 0.2. We summarize SU class definition in Table 7.2.

SU class	class 1	class 2	class 3
offered price ($/m)	0.01	0.02	0.03
probability	0.5	0.3	0.2

Table 7.2: Classification of SUs in terms of their bid offers and their corresponding probabilities.

Note that both the offered prices and their probability distributions are static, *i.e.* they do not change over time and are independent of the system occupation. It is not unrealistic taking into account typical tariff policies of wireless operators. In this environment the class structure and the probability distribution may be seen as types of contracts for SUs and market penetration of each type of contract respectively. However, for a more dynamical auction process, where bidders are able to change their bid offers adaptively, the model should be revised. One possibility would be to define one probability distribution for each state. More detailed modeling strategies would increase the complexity of the MDP solving algorithm or even make them intractable. This is a classic problem of MDPs, known as the *curse of dimensionality* and is typically addressed by means of the heuristic approach of approximate dynamic programming.

Figure 7.2 shows the Pareto fronts for the auction-based system in the three scenarios. It can be observed that, for the same traffic intensity (the three scenarios receive 40 calls per unit of time) when the traffic share of the SUs is higher, the income obtained from SUs increases at the time that the blocking probability of the PUs diminishes. It is interesting to check that, especially in scenarios 2 and 3, a very small increment of PU blocking probability can multiply the benefit obtained from spectrum leasing by a factor of 2 or 3. On the other hand, these figures also indicate that once the income surpasses certain threshold, Pareto-optimal policies can only produce small increments of the income by dramatically rising PU blocking probability.
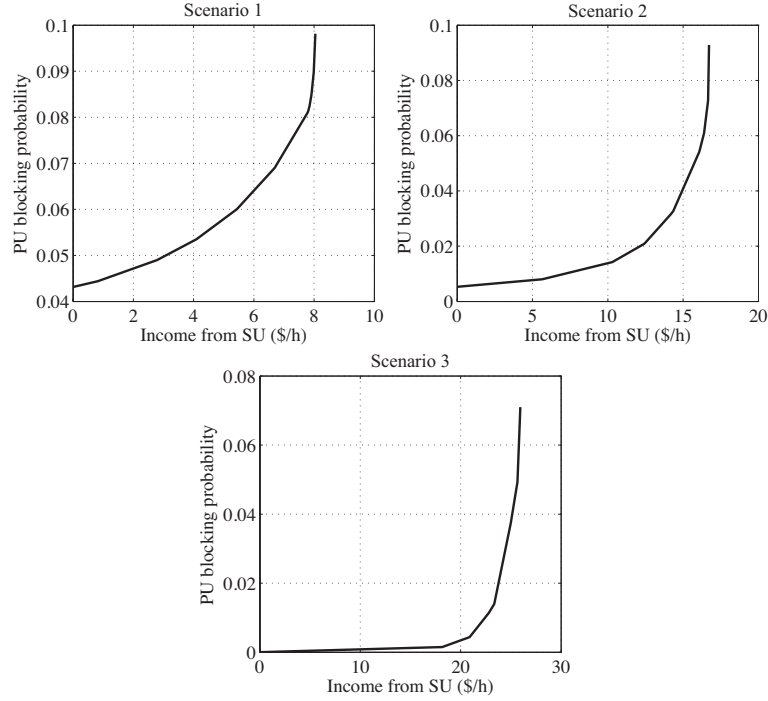
Figure 7.2: Pareto fronts obtained for the auction-based access

Fig. 7.3 shows the policies attaining a blocking probability for PUs of 0.04 for each scenario. The height of the bars corresponds to the lowest accepted class (prices) at each state. States where bars' height is 0 correspond to states where all SUs are rejected, independently of their bids. As expected, Pareto-optimal policies are threshold type. As the traffic intensity of PUs reduces respect to that of the SUs, there are more states where SUs are admitted in the system, and lower prices are accepted, *i.e.* the thresholds move toward smaller numbers. Interestingly, it is the total amount of occupied channels, and not the type of users occupying them, what determines the thresholds.

7.5   Conclusions

In this chapter we propose an MDP framework for real-time auction-based spectrum leasing. The SUs are classified according to the price they are willing to pay for the use of the spectrum. The main issue of the problem addressed is that two contrary objectives coexist: to reduce blocking probability for PUs and to increase the income received from spectrum leasing. For these type of problems there does not exist an *optimal* policy, but a set of *Pareto optimal* policies.

We have shown how to compute policies at the Pareto front by weighting the objectives in an MDP problem or by reformulating the problem as a constrained MDP. Numerical solutions of the proposed equations show the influence of the traffic share on system's performance and on the structure of Pareto-optimal policies.
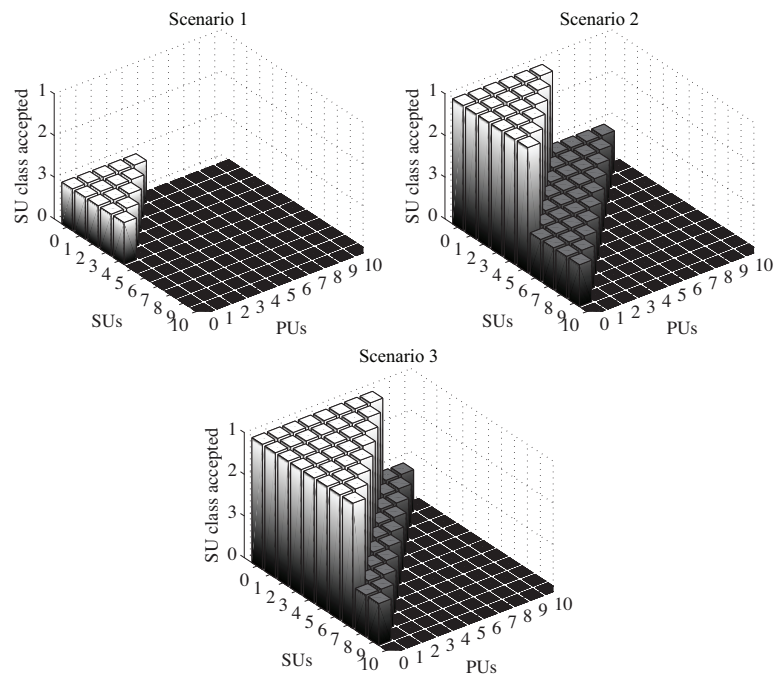
Figure 7.3: Graphical representation of policies attaining 0.04 PU block-
ing probability for the auction-based access

CHAPTER 8

A superprocess with Upper Confidence Bounds for Cooperative Spectrum Sharing

8.1 Introduction

The mechanism we proposed in the previous chapter is appropiate when the spectrum owner has unused spectrum to lease. What happens if the PU demands are higher and there are not so many opportunities to sell? Is automated spectrum trading possible in such context? In this chapter we give a positive answer to that question, in the framework of *Cooperative Spectrum Sharing* (CSS), also known as *Cooperative Secondary Spectrum Access* (CSSA). More specifically, we are interested in the trading scenario introduced by [106], whose basic premise is that SUs without license may act as transmission relays for a PU in exchange for transmission opportunities in the spectral resources of the PU.

CSS is different from other market-driven spectrum sharing approaches in that it fosters the creation of transmission opportunities in the PU spectrum in an exogenous way: by increasing the transmission rate of the PU, it reduces its spectrum usage. Because of that, CSS is particularly useful when the PU's own demands are so high that it would rarely have spectrum to lease.[1]

As we highlighted in chapter 6, most previous works are focused on developing complex algorithms which may not be fast enough for real-time decisions and/or assume perfect information about the network. Instead, we develop a learning mechanism for a PU to enable CSS in a strongly incomplete information scenario, with low computation overhead. Our mechanism is based on a Markovian variant of multi-armed bandits (MABs) called superprocess, enhanced with the concept of Upper Confidence Bound (UCB) from stochastic MABs. By means of Monte-Carlo

---

[1]It could be the case that the PU has a low traffic volume and it is not interested in increasing its transmission rate. For that reason, hybrid approaches combining CSS and other DSA ideas, with monetary transactions, have also been studied [107, 149].

simulations we show that, despite its low computational overhead, it outperforms baseline approaches such as ϵ-greedy.

Motivation

A CSS system poses the following *key challenges*: 1) spectrum resource allocation; 2) the SUs may belong to self-interested networks different from the PU's, and thus, the PU should not expect the SUs to collaborate in maximizing the PU's profit; 3) the PU has to undergo a negotiation process with the nearby SUs, having no previous information about them, in general; 4) spectrum opportunities may happen on a short timescale (of the order of seconds or less), thus, for CSS to be effective, this negotiation must be carried out in real-time; 5) the situation around the PU can change quickly (*e.g.,* SUs arriving and leaving the system).

Previous works in CSS have not addressed these issues simultaneously. As we discussed in chapter 6 and we explain in Section 8.1, they are focused on requirements 1-3, but the time needed to reach elaborated solutions (requirements 4 and 5) is not considered in detail. Multiple factors in spectrum trading (*e.g.,* supply, demand, channel gains) vary rapidly with time, and trying to reach a complex allocation solution for a particular spectrum opportunity could take so long that it may render the solution impractical.

We focus on meeting all the aforementioned requirements from the perspective of a PU willing to transmit (a primary transmitter or PT), requesting help from SUs to communicate with an intended primary receiver (PR). The PT has to gradually learn the optimal SU-offer combination based on its accumulated observations. Our proposal aims to achieve a balance between the time devoted by the PT to exploration of options and the time spent exploiting the best alternative known so far, to maximize the PT's payoff (in terms of net transmitted data) over time.

Related Work

Many previous works in CSS such as [106, 107, 108, 190] consider perfect information scenarios. They assume global knowledge of the network, actions and payoffs of other entities present in the system, and even private information like the value an entity gives to the traded good. It is unlikely that the SUs belonging to a different operator would reveal private information to the PT, taking into account that they may obtain higher payoffs by hiding it or even lying, at the cost of worsening the whole system. Some of this private information, in addition, is non-measurable. For example, if a PT suspects an SU is lying about its transmission power, it could check what the transmission power of an SU is, but there is no way to know how much battery it has left. The incompleteness of information in our scenario is stronger than in most works addressing this issue. In [154], the probability distribution of SU-PR average channel gains is fully known, and [156] assumes that the PT knows the set of such averages for the SUs present in its coverage area. Instead, we completely rely on learning to discover this information, although prior belief distributions can be easily included in our mechanism.

Other approaches [191, 161, 109] consider cooperation between networks belonging to different operators. Although it is possible to design mechanisms that provide incentives to selfish SUs to collaborate with the PT (cooperation in terms of game theory [149]), they require the exchange of several messages between these individuals, and therefore a loss in transmission efficiency, as pointed out in [131].

These selfish entities, however, do have incentives to make strategies and/or collude against the PT. In our proposal, the PT employs one-to-one bargaining instead of broadcast offers, as [155], but considering more than one SU. The benefits of one-to-one transactions versus broadcast mechanisms [156, 154, 192, 107, 106] are: the reduction of the strategic power of the SUs, as they cannot overhear public offers or other information about their competitors; their robustness against collusions of SUs [193]; and the reduction of the communication overhead on the control channels in comparison to widespread mechanisms such as auctions [156, 192] (*e.g.,* multiple rounds of bidding messages).

Auction mechanisms are also a popular approach to handle the incomplete information issues in CSS. Truthful auctions allow optimal solutions without assuming the knowledge of private information, but at the cost of complex winner-selection algorithms and/or increased communication overheads. This may not be suitable for spectrum trading in short timescales. There are other mechanisms such as the stochastic optimization based on contract theory proposed in [154], in which the authors do not assume perfect information and are concerned about the complexity of the algorithm. Nevertheless, in contrast to our proposal, they do not implement any learning process from successive interactions with SUs. Independent learning is a key feature of self-organizing networks and scalability, as indicated in the survey on market mechanisms for CSS with incomplete network information in [100].

Our work is inspired by [155], where the authors consider CSS between a PU and a cognitive SU pair in a repeated bargaining game. The bargaining partner selection problem in our work is not comparable to the one formulated in [194] because we do not assume knowledge about the impatience of other players, and we consider the PU has full bargaining power and learning abilities, among other differences. Bargaining partner selection is also different from traditional best-relay selection problems in multi-hop networks or cooperative networks [191, 158], which expect an internetwork collaborative behavior in the form of relays revealing their true link qualities, battery life, etc.

Our research in this chapter is built on the techniques for multi-armed bandit with dependent arms developed by D.B. Brown and J.E. Smith in [195] and S. Pandey, D. Chakrabarti and D. Agarwal in [196]. In [195], a MAB superprocess is used to develop an optimal policy for exploring oil and gas fields in the North Sea. The state space of their problem is finite, unlike ours. Our dependency model is also different enough from [196] to require substantial changes in their algorithm, such as the integration with an MDP. Multi-armed bandits have been previously used in spectrum sharing, mainly for sensing or resource allocation [197]. Surveys on bandits and other machine learning algorithms in cognitive radio can be found in [26] and [198]. To the best of our knowledge, however, this it the first work

exploiting MABs in CSS.

Other related works explore alternative flavors of CSS. For example [158] explores the possibility of both PUs and SUs acting as relays of each other assuming a cooperative attitude between them. The work in [159] shows the optimal strategy of an SU which is allowed to dynamically decide whether it enters a CSSs or a spectrum leasing market. A closely related topic to CSS is the overlay cognitive radio network paradigm [160], by which an SU acts as relay for the PU and, at the same time, transmits its own message. The interference constraint on the SU transmission is enlarged thanks to the cooperative transmission and the use of interference cancellation techniques [161].

Table 8.1 summarizes the most significant features of our work and compares it to previous works. Note that for those works analyzing different configurations, we select and compare to the one closest to our proposal.

Table 8.1: Summary of related works

	Our work	[156]	[154]	[155]	[149]	[108]	[191]
Information	**Strongly incomplete**	Incomplete	Incomplete	**Strongly incomplete**	Complete	Complete	Complete
Learning	**Exploration - exploitation**	Greedy search	No	**Repeated game**	N/A	N/A	N/A
Penalization for rejected offers	**Yes**	No	N/A	**Yes**	N/A	N/A	N/A
Communication	**Unicast**	Broadcast	Broadcast	**Unicast**	Broadcast	Broadcast	Broadcast
Technique	**MAB - superprocess**	Matching Theory	Contract Theory	Bargaining	Co-op. bargaining	Stackelberg game	MILP optimization
Market form	**1 PU - n SUs**	**1 PN - n SUs**	**1 PU - n SUs**	1 PU - 1 SU	**1 PU - n SUs**	1 PN - 1 SN	n PUs - n SUs
Objective	**Max. PU reward**	**Max. PU reward**	**Max. PU reward**	Equilibrium (bias towards SUs)	Max. SU rewards and fairness	Equilibrium (bias towards PU)	Configurable
Year	**2014**	**2014**	**2014**	2013	2012	2010	2013

N/A = Not applicable, PN = Primary Network, SN = Secondary Network, MILP = Mixed Integer Linear Programming
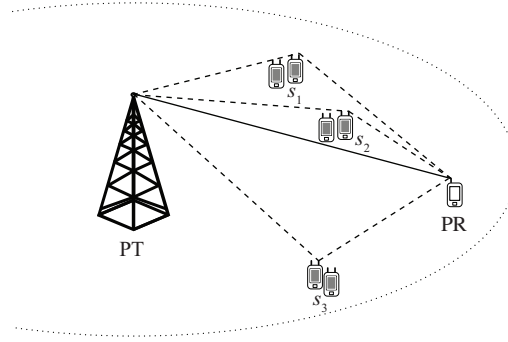
Figure 8.1: Cooperative Spectrum Sharing.

Our Contribution

We develop a new approach for the PT learning problem in the context of CSS, by means of a stochastic multi-armed bandit (MAB) problem with dependent arms. This approach allows the PT to group the information gathered from SUs instead of treating every possible action independently. Thus, the PT learns in a more efficient way from its observations compared to other reinforcement learning approaches. We formulate the main problem in Section 8.2. The reward function of our MAB comprises two factors: one is drawn from a probability distribution, and the other one is determined by a Markov chain. To the best of our knowledge, there are no algorithms in the previous literature for this type of rewards in MABs. In Section 8.3, *we describe our MAB-MDP algorithm*, in which we integrate a stochastic MAB and a Markov Decision Process (MDP). MAB-MDP allows us to introduce some key concepts used in Section 8.4 to build *a new algorithm for solving the stochastic MAB with dependent arms. We call this algorithm Superprocess with Upper Confidence Bounds (Super-UCB)*, because it combines the classic UCB indices used in stochastic MABs and an extension to Markovian MABs [199] called superprocess. *Super-UCB makes a more efficient use of the available information, achieving a better performance without increasing the computational overhead. We evaluate the performance, scalability and robustness of Super-UCB with Monte Carlo simulations* in Section 8.5. Our solution is not only directly implementable, without strong assumptions, but also extensible to more complex scenarios, as summarized in the conclusions in Section 8.6.

8.2   System Model

The protocol considers a PU transmitter (PT) - receiver (PR) pair and a set of SU cognitive pairs in the coverage area of the PT denoted by $\mathcal{S} \equiv \{s_1, s_2, \ldots, s_S\}$, as in Fig. 8.1. The system is under the "exclusive-use" coexistence model [9], by which the PUs are the only entities with the right to transmit in a certain band. When the PU pair's channel condition is not suitable for direct transmission, the PT would be willing to use the SUs as relays.

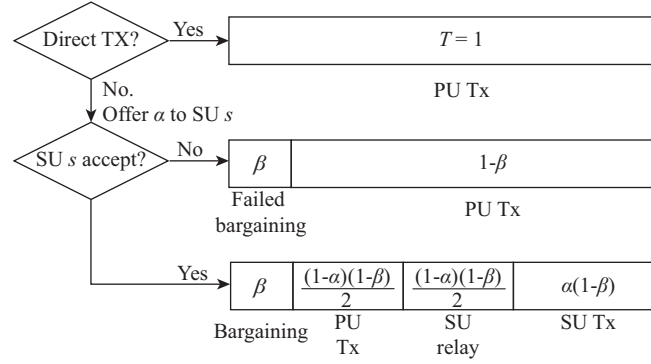In exchange for its services, the PT makes an offer to the SU, consisting of a

Figure 8.2: Frame structures.

certain amount of time for SU data transmission over the PT channels. The SUs transmit with fixed power, the same for relaying and their own transmissions. The SUs are assumed to have their own, but limited, spectral resources. Therefore, although it is not crucial for their communication purposes, the SUs may benefit from the additional spectrum resources obtained from the PT. As a consequence, it is the PT who contacts the SU and initiates the bargaining.

Time is divided into fixed transmission periods or *frames*, which we will consider of duration $T = 1$ units of time and numbered as $n = 1, 2, \ldots$. The offer that the PT makes, denoted by $\alpha \in [0, 1]$, is the fraction of the transmission period during which the SU is allowed to transmit its own data. For tractability, we discretize that interval in equal increments and $\mathcal{A} = \{\alpha_1, \ldots, \alpha_A\}$ denotes the set of possible offers to make to SU $s \in \mathcal{S}$.[2]

Each transmission period is composed of a decision phase, an optional bargaining phase, and a communication phase, as in [106, 155]. See also Fig. 8.2.

- Decision phase. During this phase, the PT has to choose which SU to bargain with and what offer to make to it: $(s, \alpha) \in \mathcal{S} \times \mathcal{A}$. The PT aims to maximize its accumulated payoff, and its decisions are based on previous interactions.

- Bargaining phase. We consider the PT has full bargaining power and makes a take-it-or-leave-it offer to an SU. This phase models the time spent in sending that message and waiting for the answer.

- Communication phases. If the offer is rejected by the SU, the PT directly transmits for the rest of the frame to its intended receiver. If the offer is accepted by the SU, then the communication phase is divided into three time periods:

  - Direct transmission of the PT. The PT transmits its information to its receiver and to the selected SU.

  - SU forwarding. The selected SU re-transmits the information to the PR.

  - SU's own transmission.

---

[2] In practice we will assume $\alpha_A = 0.9$, and $\alpha_1 > 0$, as neither the SU nor the PT may "work for free". Also note that our model could consider different offer sets $\mathcal{A}_s$ for each SU.

Payoff Functions

PT's payoff. For the cooperative transmission, the system employs the decode and forward relay communication scheme from [200]. For a given SU $s \in \mathcal{S}$, the links initially involved are: PT-PR, PT-s, and s-PR. Let $\Gamma_{P,P}^{(n)}$, $\Gamma_{P,s}^{(n)}$, and $\Gamma_s^{(n)}$ be their respective SNR values averaged over the duration of transmission frame n. The achievable data rate satisfies $R_P^{(n)} = K \min\{\log(1 + \Gamma_{P,s}^{(n)}), \log(1 + \Gamma_{P,P}^{(n)} + \Gamma_s^{(n)})\}$, where K is a constant. We assume that the SUs always decode the PT data correctly in the first phase, and thus we focus on the s-PR link and we have $R_P^{(n)} = K \log(1 + \Gamma_{P,P}^{(n)} + \Gamma_s^{(n)})$. Since the PT-PR and s-PR links are considered to be in bad and good propagation conditions respectively, *i.e.*, $\Gamma_{P,P}^{(n)} \ll \Gamma_s^{(n)}$, and $1 \ll \Gamma_s^{(n)} \; \forall n$, the achievable data rate can be approximated as $R_P^{(n)} \approx K \log(\Gamma_s^{(n)})$.

The time allocated to the s-PR transmission within a frame can be considered long compared to fast fading variations. That is, the effect of multipath is assumed to be negligible in terms of the average SNR, $\Gamma_s^{(n)}$. Therefore, $\Gamma_s^{(n)}$ is mainly determined by pathloss attenuation and shadowing, remaining constant during an s-PR transmission. Successive SU relay phases between s and the PR are sufficiently distant in time for the channel to decorrelate. Thus, the SNR samples $\Gamma_s^{(n)}$ are modeled as i.i.d. random variables, following a log-normal distribution which typically characterizes shadowing [28], *i.e.*, $\log(\Gamma_s^{(n)}) = \gamma_s^{(n)} \sim N(\mu_{\gamma_s}, \sigma_{\gamma_s})$.

The PR can observe the value of $\gamma_s^{(n)}$ and feed it back to the PT at the end of each transmission frame. Then, the reward for the PT in frame n is the transmission rate it gets, multiplied by the time that transmission lasts:

$$W^{(n)}(s, \alpha) = \begin{cases} (1 - \beta)(1 - \alpha) \log(\Gamma_s^{(n)}) & \text{if } \alpha \text{ is accepted} \\ (1 - \beta) \log(1 + \Gamma_{P,P}^{(n)}) & \text{otherwise} \end{cases} \qquad (8.1)$$

For the sake of clarity, from now on we set $\beta = 0$ and $\log(1 + \Gamma_{P,P}^{(n)}) = 0$ for all n. Nevertheless, our model is also applicable with different values of these parameters.

SU's payoff. The payoff obtained by an SU acting as relay is the difference between its net transmitted data during a frame when using the PT channel, αR_s^P, and the net data it would transmit when using its own spectral resources R_s. An SU accepts any offer that provides positive payoff, which results in a threshold behavior. SU s accepts an offer $\alpha_a \in \mathcal{A}$ whenever $\alpha_a \geq \alpha_s^*$, where α_s^* is the minimum offer SU s is willing to accept. The *type of an SU*, τ_s, is the index of the smallest offer that this SU accepts, i.e. $\tau_s = \arg\min_a\{\alpha_a : \alpha_a \geq \alpha_s^*, \alpha_a \in \mathcal{A}\}$. If the link between the SU pair is stable (*e.g.,* a close ad hoc connection) and the offered PT bandwidth is constant (only the time offered changes), the thresholds, and therefore the type of each SU, remain constant over multiple transmission frames.

Multi-armed bandit formulation

Mathematically, we model the sequential decisions of the PT as a multi-armed bandit (MAB) problem. The PT selects SU-offer pairs (s, α) from the action set $\mathcal{U} = \mathcal{S} \times \mathcal{A}$. In the MAB model each $u = (s, \alpha)$ is an arm, and $\mathcal{U}$ is the set of arms.

At round n, the arm pulled is $u^{(n)} = (s^{(n)}, \alpha^{(n)})$, and the reward received by the PT is $W^{(n)} = W^{(n)}(u^{(n)})$. The history of the system up to time n is defined as the sequence of decisions and observed rewards $W^{(0)}, u^{(1)}, W^{(1)}, \ldots u^{(n)}, W^{(n)}$ (where $W^{(0)}$ corresponds to the initial samples $\gamma_s^{(0)}$). A *policy* π is a function that, at each round n, prescribes a decision $u^{(n+1)}$ based on the history of the system. Therefore π induces a history $W^{(0)}, u_\pi^{(1)}, W_\pi^{(1)}, u_\pi^{(2)}, W_\pi^{(2)}, \ldots$. The usual performance metric in learning problems is the *regret* [201]. At decision stage n, we define the regret of a policy π by $r^{(n)} = \max_u \mathbb{E}\left[\sum_{k=1}^n \left(W^{(k)}(u) - W_\pi^{(k)}\right)\right]$. The regret quantifies the performance loss of a learning policy with respect to a policy that makes the best decision on average. In a realistic setting, the PT does not know such best decision, and faces the challenge of learning it while trying to maximize the reward. This is known as the *exploration - exploitation* tradeoff, a term coined in [202] and extensively used afterwards in the related literature [203, 25]. It describes the problem of finding the optimal balance between getting immediate rewards and gathering additional information to make better decisions in future rounds. In MABs, this is the balance between pulling the arms that seem to be better in expectation and those that may seem to be worse initially but could potentially be the best.

Correlated arms. Learning about an arm of the MAB should not be treated independently of the other arms. The reward obtained from an SU s when offered a particular α provides information about the rewards of all the other offers (the other arms of the same SU). Before making an offer α to an SU s, the PT holds a prior belief about the acceptance probability of each offer by s. The acceptance or rejection of an offer implies an update of this belief. In addition, when the PT observes a sample of γ_s and updates its sample mean $\overline{\gamma}_s^{(n)}$, that affects the belief about the reward of all the arms $\{(s, \alpha), \ \forall \alpha \in \mathcal{A}\}$ of that SU s.

In the next sections we explain how, inspired by [196] and [195], respectively, we handle correlation by grouping the observations of an SU, exploiting mutual information. *We reduce the original MAB* with $|\mathcal{U}| = S \times A$ arms to a MAB with S arms, one for each SU, where each arm integrates the information the PT has observed about each SU up to round n, that is, the acceptance and rejection of offers, and the SNR samples $\gamma_s^{(n)}$. We must then answer two questions: *how do we represent an SU in this reduced MAB?* Once an SU has been chosen by means of the reduced MAB, *what offer should the PT make to the selected SU?*

There are two fundamental problems that prevent us from using previous MAB algorithms. The first problem is that the reward of our model does not match those of previous works. We encode the acceptance and rejection of offers as knowledge states. Thus, under a given policy, we can see the changes in the PT's knowledge as a Markov chain. Then, it is clear by (8.1) that the rewards of our reduced MAB come simultaneously from the probability distribution of $\gamma_s^{(n)}$ and from the Markov process of the knowledge states about the offers. Previous algorithms for MABs model the rewards when pulling an arm as drawn from a probability distribution (stochastic MABs), from a Markov chain (Markovian MABs), or chosen by an adversary (adversarial MABs). *There are no algorithms in the literature for mixed stochastic and Markovian bandits.* Our proposals combine techniques from both

types of bandits to solve the reduced MAB. The *MAB-MDP algorithm* relies on a classic algorithm for stochastic MABs, augmented with an MDP, while the *Super-UCB algorithm* is based on an algorithm for Markovian MABs augmented with elements from stochastic MABs. The main difference between our two algorithms is that Super-UCB integrates more information about SUs in the PT's policy than MAB-MDP as we will explain further in Section 8.3.

The second problem when using classic MAB algorithms in our reduced MAB is that those algorithms only dictate one decision in every round (which arm to pull), whereas the PT needs to make one more decision (the offer to the SU). The MDP of our MAB-MDP algorithm acts as an additional mechanism to the MAB to obtain such policy. In the case of Super-UCB, we make use of an extended model of Markovian MABs which handles MABs whose arms are MDPs instead of plain Markov chains, thus providing a policy for arm and action selection.

8.3   MAB - MDP Algorithm

In this proposal, we handle the reduced MAB as a stochastic MAB to obtain a policy for selecting which SU to bargain with, and integrate it with an MDP for the offer selection.

Offer selection policy. We are interested in finding a policy π_s^{MDP} that, given an SU $s \in \mathcal{S}$, maps the PT's knowledge about the type of s, to the next offer α to s. Given α, the reward ψ that the MDP observes is $(1-\alpha)$ if the offer is accepted, and 0 otherwise. Therefore, a given policy π_s^{MDP} induces a sequence of rewards $\psi^{(1)}, \psi^{(2)}, \ldots, \psi^{(n)}$. The uncertainty about the time horizon n is captured by a discount factor $\delta < 1$, characterizing the expected lifetime of the system, e.g., the probability that the s-PR pair remains active in each frame. The search problem of discovering the optimal offer to s consists of finding the policy maximizing $\mathbb{E}\left[\sum_{n=1}^{\infty} \delta^n \psi^{(n)}\right]$. This problem is formulated as an MDP as follows.

The set of *states* of the MDP for SU s, denoted by $\mathcal{X}_s$, are all the possible knowledge states about the type of s. Assuming uniform probability for the type of an SU, the state of the MDP is completely defined by a two dimensional vector x, whose elements, h and l, contain the index of the highest offer rejected by s, and the index of the lowest offer that s accepted, respectively. The PT knows that the SU accepts every offer α_a with index $a \geq l$, and rejects those with index $a \leq h$, $\alpha_a \in \mathcal{A}$. The initial state is $(0, A)$, since the PT does not know anything about rejections initially ($h = 0$) and α_A is known to be surely accepted ($l = A$).

Each *transition probability* $P_s(x, x', \alpha_a)$ from a state x to a state x', with $x, x' \in \mathcal{X}_s$, given an offer α_a, is determined by the PT's beliefs about the acceptance probability of the offer α_a at current knowledge state, x. This probability is defined as:

$$P(\alpha_s^* \leq \alpha_a | x) = \begin{cases} 0 & \text{for } 0 < a \leq h \\ \frac{a-h}{l-h} & \text{for } h < a < l \\ 1 & \text{for } l \leq a \leq A \end{cases} \qquad (8.2)$$
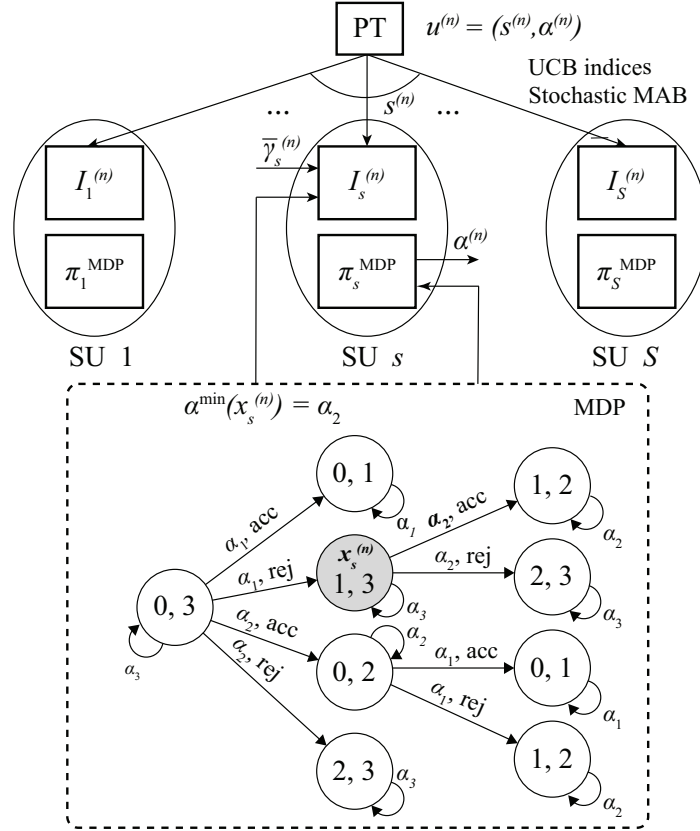
Figure 8.3: MAB-MDP algorithm for S SUs and A = 3 possible offers . Each SU is chosen according to a MAB and the offer according to an MDP. "acc" represents the event that the offer is accepted and "rej" that it is rejected.

Therefore, the transition probabilities between every pair of states are given by:

$$P_s(x, x', \alpha_a) = \begin{cases} P(\alpha_s^* \leq \alpha_a | x) & \text{for } a = l' \\ 1 - P(\alpha_s^* \leq \alpha_a | x) & \text{for } a = h' \\ 0 & \text{otherwise} \end{cases} \tag{8.3}$$

Finally, the reward associated to a transition of the MDP is:

$$\psi(x', \alpha_a) = \begin{cases} 1 - \alpha_a & \text{for } a = l' \\ 0 & \text{for } a = h' \end{cases} \tag{8.4}$$

We can now formulate the Bellman equation that allows us to obtain the value function V_s for each state:

$$V_s(x) = \max_{\alpha_a \in \mathcal{A}} \sum_{x' \in \mathcal{X}_s} P_s(x, x', \alpha_a)(\psi(x', \alpha_a) + \delta V_s(x')) \tag{8.5}$$

which can be readily solved offline by standard algorithms such as policy iteration [24].

SU representation in the reduced MAB. In a classic stochastic MAB, rewards when pulling an arm are drawn from a probability distribution associated to that arm. Proposed policies in the literature compute an index for each arm (dependent

on that arm only), $I_s^{(n)}$, and dictate to pull the arm with the highest index on each decision round. A commonly used family of index policies are the *Upper Confidence Bound (UCB)* policies proposed by [199]. The index of these policies consists of the sample average reward obtained from arm s up to round n, plus an additional term, the UCB, related to the uncertainty of that estimation.

In our case, the reward W when pulling arm s depends not only on s but also on α: $W^{(n)}(s, \alpha) = (1 - \alpha)\log(\Gamma_s^{(n)})$. When choosing arm s and offering α, the reward is drawn from the Gaussian distribution $\log(\Gamma_s^{(n)}) \sim N(\mu_{\gamma_s}, \sigma_{\gamma_s})$, multiplied by a constant $(1 - \alpha)$. Let us denote by $x_s^{(n)} = (h_s^{(n)}, l_s^{(n)}) \in \mathcal{X}_s$ the state of SU s at round n. We characterize the SU in the reduced MAB by its *presumed best achievable offer* $\alpha^{\min}(x_s^{(n)})$, denoting the minimum $\alpha \in \mathcal{A}$ in state $x_s^{(n)}$ with positive belief of being accepted and also being achievable by the offer selection policy π_s^{MDP}.[3]

Note that, in general, $\alpha^{\min}(x_s^{(n)}) \neq \alpha_a$, with $\alpha_a = \pi_s^{\text{MDP}}(x_s^{(n)})$. That is, the PT uses an optimistic representation of SU s type in a given state $x_s(n)$, but the PT may not directly jump into the corresponding optimal offer $\alpha^{\min}$. Instead, the PT may approach it in successive interaction rounds by reducing the offer every time, gradually increasing the risk of rejection, in order to balance exploration and exploitation as dictated by π_s^{MDP}.

By applying this SU characterization to the UCB index for Gaussian distributions shown in [199], we obtain:

$$I_s^{(n)} = \widehat{W}_s^{(n)} + 4\hat{\sigma}_{W_s}(x_s^{(n)})\sqrt{\frac{\ln(n_{\{W>0\}} + 2)}{n_s + 1}} \qquad (8.6)$$

where $\widehat{W}_s^{(n)}$ is the estimated average reward of arm s up to round n, and $\hat{\sigma}_{W_s}(x_s^{(n)})$ is the estimated standard deviation of the rewards of arm s, n_s is the number of times arm s has been pulled up to round n, and $n_{\{W>0\}}$ is the number of rounds in which the PT obtained a positive reward from any SU. Since s is characterized by its presumed best achievable offer in state $x_s^{(n)}$, we have that $\widehat{W}_s^{(n)} = (1 - \alpha^{\min}(x_s^{(n)}))\overline{\gamma}_s^{(n)}$ and $\hat{\sigma}_{W_s}(x_s^{(n)}) = (1 - \alpha^{\min}(x_s^{(n)}))\sigma_{\gamma_s}$, with $\overline{\gamma}_s^{(n)}$ and σ_{γ_s} denoting the sample mean and known standard deviation of $\gamma_s^{(n)}$, respectively.

In practice, removing the factor 4 from the UCB yields a significantly lower regret, while still achieving convergence. However, despite this empirical evidence, there are no theoretical bounds on the regret for this version of the UCB.

We proceed to give a detailed **description of the MAB - MDP algorithm**. Its diagram and pseudo-code can be found in Fig. 8.3 and Algorithm 8.4, respectively. Initially, for each SU $s \in \mathcal{S}$:

1. The PT computes the exploration-exploitation policy π_s^{MDP} for the offers $\alpha \in \mathcal{A}$.

---

[3] For a policy π_s^{MDP} that ends up exploring all the offers of an SU, $\alpha^{\min}(x_s^{(n)})$ is simply the lowest α not rejected for the current state x. For more conservative policies that may dictate not to explore all offers, $\alpha^{\min}(x_s^{(n)})$ is the lowest α that policy is willing to explore given the state $x_s^{(n)}$. The complete notation would be $\alpha^{\min}(x_s^{(n)}, \pi_s^{\text{MDP}})$ but we use the former to improve readability.

2. Given that policy, for the initial state $x_s^{(0)} = (0, A)$, the PT computes the presumed best achievable offer $\alpha^{\min}(x_s^{(0)})$ of each SU. The PT chooses that offer as representative of the SU.

3. With $\alpha^{\min}(x_s^{(0)})$ and the initial sample of the SNR $\gamma_s^{(0)}$, the PT builds the UCB index $I_s^{(0)}$ (8.6).

Then, on each round n:

4. The PT selects the SU z with the highest UCB index $I_z^{(n)}$ and the offer $\alpha_a \in \mathcal{A}$ is indicated by the MDP policy π_z^{MDP}.

5. If the SU *rejects* the offer, the PT updates its knowledge state (MDP) $x_z^{(n+1)}$ and checks if $\alpha^{\min}(x_z^{(n+1)}) \neq \alpha^{\min}(x_z^{(n)})$. If that is the case, the PT computes $I_z^{(n+1)}$ according to (8.6). Otherwise, $I_z^{(n+1)} = I_z^{(n)}$.

6. If the SU *accepts*, the PT updates $x_z^{(n+1)}$, and sets $n_z = n_z + 1$. The indices I_s, for all $s \in \mathcal{S}$, are also updated because $n_{\{W>0\}} = n_{\{W>0\}} + 1$ after the PT receives a positive reward.[4]

Note that it cannot be assured that this algorithm performs a fully optimal learning. An optimal learning algorithm would imply solving an MDP comprising all the information gathered by the PT. Our MAB-MDP algorithm decomposes the learning problem into two simpler sub-problems (the MAB and the MDP). Each sub-problem can be solved very efficiently because it uses only partial information about the global system. Despite its simplicity, MAB-MDP shows a remarkably low regret in numerical evaluation.

8.4 Super-UCB Algorithm

The MAB-MDP algorithm of the previous section is based on a stochastic MAB. In this section we follow the Markovian MAB approach. Our objective is to benefit from all the information gathered about the SU types, instead of considering only their presumed best achievable offer.

For the classic Markovian MAB with independent arms in which a decision-maker (the PT in our scenario) selects an arm to play and receives a random reward from a Markov chain, Gittins [204] developed an index policy proven to be optimal if the transition probabilities are known. The computation of the indices for each state of each arm is performed offline. The online sequential decision is as simple as in the stochastic MAB index policies: at each decision stage, the decision-maker has to pull the arm with the highest index for their current states. Then, the state of the pulled arm changes and so does the index of that arm, which is replaced by the one associated to the new state.

Superprocess. As we are grouping the arms of the MAB by SU, the rewards when choosing an SU are not drawn from Markov reward chains but from Markov Decision Processes (MDP). The PT does not only select which SU to interact with

---

[4]In a classic stochastic MAB, every arm pull increases $n_{\{W>0\}}$ because every arm pull gives a reward from a probability distribution. This is not our case due to the possibility that an offer is rejected.

Input: $\mathcal{X}_s, \mathcal{A}, P_s, \gamma_s^{(0)}; \; \forall s \in \mathcal{S}$
Output: $(s^{(n)}, \alpha^{(n)}); \; s^{(n)} \in \mathcal{S}, \alpha^{(n)} \in \mathcal{A}; \; \forall n$

1: Initialize. $\forall s \in \mathcal{S}$:
2:     $n_{\{W>0\}}, n_s^{(n)}, n \leftarrow 0$
3:     $\pi_s^{\mathsf{MDP}} \leftarrow$ Solve MDP $(\mathcal{X}_s, \mathcal{A}, P_s)$
4:     $x_s^{(0)} \leftarrow (0, A)$
5:     $\widehat{W}_s^{(0)} \leftarrow (1 - \alpha^{\min}(x_s^{(0)}))\gamma_s^{(0)}$
6:     $\hat{\sigma}_{W_s}(x_s^{(0)}) \leftarrow (1 - \alpha^{\min}(x_s^{(0)}))\sigma_{\gamma_s}$
7:     $I_s^{(0)} = \widehat{W}_s^{(0)} + 4\hat{\sigma}_{W_s}(x_s^{(0)})\sqrt{\frac{\ln(n_{\{W>0\}}+2)}{n_s+1}}$
8: **loop**
9:     $s^{(n)} \leftarrow z \leftarrow \max_s I_s^{(n)}$
10:     $\alpha^{(n)} \leftarrow \alpha_a \leftarrow \pi_z^{\mathsf{MDP}}(x_z^{(n)})$
11:     **if** $\alpha_a < \alpha_z^*$ **then** {SU rejects}
12:         collect reward 0
13:         $x_z^{(n+1)} \leftarrow (a, l_z^{(n)})$
14:         $\widehat{W}_z^{(n+1)} \leftarrow (1 - \alpha^{\min}(x_z^{(n+1)}))\overline{\gamma}_z^{(n)}$
15:         $\hat{\sigma}_{W_z}(x_z^{(n+1)}) \leftarrow (1 - \alpha^{\min}(x_z^{(n+1)}))\sigma_{\gamma_z}$
16:         $I_z^{(n+1)} \leftarrow \widehat{W}_z^{(n+1)} + 4\hat{\sigma}_{W_z}(x_z^{(n+1)})\sqrt{\frac{\ln(n_{\{W>0\}}+2)}{n_z+1}}$
17:     **else**
18:         collect reward $(1 - \alpha_a)\log(\Gamma_z^{(n)})$
19:         $x_z^{(n+1)} \leftarrow (h_z^{(n)}, a)$
20:         $\overline{\gamma}_z^{(n+1)} \leftarrow \frac{\overline{\gamma}_z^{(n)} n_z + \gamma_z^{(n)}}{n_z+1}$
21:         $\widehat{W}_z^{(n+1)} \leftarrow (1 - \alpha^{\min}(x_z^{(n)}))\overline{\gamma}_z^{(n+1)}$
22:         $n_z \leftarrow n_z + 1$
23:         $n_{\{W>0\}} \leftarrow n_{\{W>0\}} + 1$
24:         **for all** $s \in \mathcal{S}$ **do**
25:             $I_s^{(n+1)} \leftarrow \widehat{W}_s^{(n+1)} + 4\hat{\sigma}_{W_s}(x_s^{(n)})\sqrt{\frac{\ln(n_{\{W>0\}}+2)}{n_s+1}}$
26:         **end for**
27:     **end if**
28:     $n \leftarrow n + 1$
29: **end loop**

Figure 8.4: MAB-MDP algorithm in pseudo-code

(pulling an arm) but also which offer to make (selecting an action of the MDP). If we consider the SUs as independent, this problem can be formulated as an extension to MABs called *superprocess*. P. Whittle [205] extended the Gittins index policies to superprocesses. Regarding the offer selection policy, there are different choices that we discuss later.

Ideally, such MDP of an arm should encode all information or beliefs the PT has about an SU, that is, both the acceptance and rejection of offers, and the samples of $\gamma_s^{(n)}$. Unfortunately, including the information about $\gamma_s^{(n)}$ would lead to a continuous-state MDP which is intractable in practice. Combining the superprocess with Approximate Dynamic Programming techniques is not an option either: as shown in [196], the approximations obtained for Gittins indices are not tight enough (not even to be used as heuristics). In order to integrate the uncertainty about μ_{γ_s} in the superprocess, without making the state space explode, we propose an approximate method by including the PT's beliefs about μ_{γ_s} in the reward of the MDP in the form of a UCB index, instead of including them as state variables. Then, in every round in which the UCB indices get updated, that is, rounds in which the PT's beliefs about μ_{γ_s} change, the PT solves the superprocess using the new UCB indices in the rewards of the MDPs, obtaining a new set of Gittins indices for all states.

Solution to the superprocess. In order to obtain the Gittins index for state x of an arm s, Whittle introduces a retirement action into the Bellman equation. The retirement action leads to a ficticious final state with a unique reward M. Then, the value of state x at round n, $V_s^{(n)}(x, M)$, is obtained by solving:

$$V_s^{(n)}(x, M) = \max \left\{ M, \max_{\alpha_a \in \mathcal{A}} \sum_{x' \in \mathcal{X}_s} P_s(x, x', \alpha_a) \left(\phi_s^{(n)}(x', \alpha_a) + \delta V_s^{(n)}(x', M) \right) \right\}$$

(8.7)

where $\phi_s^{(n)}$ denotes the reward function at round n, associated to a transition of the MDP of SU s, and it is defined as follows:

$$\phi_s^{(n)}(x', \alpha_a) = \begin{cases} (1 - \alpha_a) J_s^{(n)} & \text{for } a = l' \\ 0 & \text{for } a = h' \end{cases}$$

(8.8)

and $J_s^{(n)}$ denotes a UCB index of the form $J_s^{(n)} = \left(\overline{\gamma}_s^{(n)} + 4\sigma_{\gamma_s} \sqrt{\frac{\ln(n_{\{W>0\}}+2)}{n_s+1}} \right)$. The *Gittins index* $M_s^{(n)}(x)$ for SU s in state x and round n is the smallest M that makes the PT indifferent between taking the retirement action or continuing in the sequential decision problem. The retirement reward symbolizes the reward that could be achieved by switching to the other arms.

In order to find the Gittins indices $M_s^{(n)}(x), \forall x \in \mathcal{X}_s$, (8.7) has to be repeatedly solved for different values of M. Following the approach in [195], we can formulate (8.7) as a linear program [24] and use the techniques of parametric linear programming with parameter M. Let $\mathbf{v}_s^{(n)}$ denote the vector of values representing the value function $V_s^{(n)}(x, M), \forall x \in \mathcal{X}_s$. The corresponding linear program to (8.7) for

SU s at round n is:

$$\underset{v_s^{(n)}}{\text{minimize}} \sum_{x \in \mathcal{X}_s} v_s^{(n)}(x)$$

$$\text{subject to } v_s^{(n)}(x) \geq \sum_{x' \in \mathcal{X}_s} P_s(x, x', \alpha_a)(\phi_s^{(n)}(x, x', \alpha_a)$$

$$+ \delta v_s^{(n)}(x')), \forall x \in \mathcal{X}_s, \alpha_a \in \mathcal{A} \tag{8.9}$$

$$v_s^{(n)}(x) \geq M, \forall x \in \mathcal{X}_s$$

$V_s^{(n)}(x, M)$ is known to be piecewise linear, increasing, and convex in M [195]. Obtaining the value of $V_s^{(n)}(x, M)$, then, consists of finding the slopes and breaking points of the pieces. Using the Gaas-Saaty algorithm for parametric linear programming [206], the authors of [195] proposed the *Frontier Algorithm* to compute the Gittins indices $M_s^{(n)}(x)$. The Gaas-Saaty algorithm is a variation of the traditional simplex method for linear programs.

Given (8.7) and the particular form of $\phi_s^{(n)}$, there is no need to solve the superprocess on each round in response to the changes in $J_s^{(n)}$. The next result shows that it is enough to solve Eq. (8.7) once per SU, using a reference index $J_s^{\prime(n)} = 1$.

Proposition 1. *Let $J_s^{(n)}$ be the UCB index of an SU s at round n, and let $\mathcal{X}_s$ be the state space of s. The Gittins indices for each state $x \in \mathcal{X}_s$ at round n are given by:*

$$M_s^{(n)}(x) = J_s^{(n)} M_s'(x) \tag{8.10}$$

where $M_s'(x)$ is the Gittins index for state x when the UCB index is $J_s^{\prime(n)} = 1$.

Proof. Consider Eq. (8.7) for round n of SU s. Dividing both sides of the expression by $J_s^{(n)}$, we obtain:

$$\frac{V_s^{(n)}(x, M)}{J_s^{(n)}} = \max\left\{\frac{M}{J_s^{(n)}}, \max_{\alpha_a \in \mathcal{A}} \sum_{x' \in \mathcal{X}_s} P_s(x, x', \alpha_a)\left(\frac{\phi_s^{(n)}(x', \alpha_a)}{J_s^{(n)}} + \delta\frac{V_s^{(n)}(x', M)}{J_s^{(n)}}\right)\right\} \tag{8.11}$$

Observe that $\frac{\phi_s^{(n)}(x', \alpha_a)}{J_s^{(n)}} = \phi_s^{\prime(n)}(x', \alpha_a)$, where:

$$\phi_s^{\prime(n)}(x', \alpha_a) = \begin{cases} (1 - \alpha_a)J_s^{\prime(n)} & \text{for } a = l' \\ 0 & \text{for } a = h' \end{cases} \tag{8.12}$$

Denoting $\frac{V_s^{(n)}(\cdot, M)}{J_s^{(n)}} = V_s^{(n)}(\cdot, \frac{M}{J_s^{(n)}})$, Eq. (8.11) becomes:

$$V_s^{(n)}\left(x, \frac{M}{J_s^{(n)}}\right) = \max\left\{\frac{M}{J_s^{(n)}},\right.$$

$$\left.\max_{\alpha_a \in \mathcal{A}} \sum_{x' \in \mathcal{X}_s} P_s(x, x', \alpha_a)\left(\phi_s^{\prime(n)}(x', \alpha_a) + \delta V_s^{(n)}\left(x', \frac{M}{J_s^{(n)}}\right)\right)\right\} \tag{8.13}$$

which is essentially equal to Eq. (8.7), with $J_s^{\prime(n)} = 1$ and M scaled by a factor $J_s^{(n)}$. It is then straightforward to see that $\frac{M_s^{(n)}(x)}{J_s^{(n)}} = M_s'(x)$. $\qquad\square$

This fact greatly reduces the computation overhead of the algorithm, because the solution to the superprocess only needs to be computed once, and can be done offline.

Offer selection policy. Regarding the action to take on a given state for an arm, there are different choices. The Whittle - Gittins index policy for superprocesses turns out to be optimal only if there is a dominant action $\alpha_s^{DOM}(x)$, for each arm s and state $x \in \mathcal{X}_s$, that achieves the maximum in (8.7) for all values of $M \in [0, M_s'(x)]$. If that is the case, the action indicated by the policy is the dominant action. In our case, however, the action indicated by the policy solving (8.7) changes with the retirement value. This means that the choice of the action for a given SU depends on what the PT knows about the others (*i.e.,* depending on the types of the other SUs, the PT may decide to explore more or less a particular SU). Thus, it is suboptimal to consider each SU in isolation when deciding on the offer to make. Nevertheless, even for this suboptimal case, the Gittins - Whittle policy can still be used as a heuristic, providing tight upper bounds (¡ 2%) on the values of $V_s^{(n)}(x)$ (Eq. (8.7)) as reported by [195]. For the suboptimal cases, [196] suggests to choose the actions that maximize $V_s^{(n)}(x, M_s'(x))$. However, such policy leads to under-exploration of an SU, while a policy selecting the actions that maximize $V_s^{(n)}(x, 0)$ in (8.7) attains a lower regret, as noticed in [195] and verified by our simulations.[5] We will denote such policy by π_s^{SPR}. The Frontier Algorithm provides all the previously mentioned policies.

Description of the Super-UCB algorithm. Its diagram and pseudo-code can be found in Fig. 8.5 and Algorithm 8.6, respectively. Initially, for each SU $s \in \mathcal{S}$:

1. We obtain the reference Gittins indices $M_s'(x)$ and the offer selection policy π_s^{SPR} for every state x of SU s, solving (8.7) with the Frontier Algorithm.

2. The PT computes the UCB index $J_s^{(0)}$ and the Gittins indices $M_s^{(0)}$ with the initial samples $\gamma_s^{(0)}$. [6]

Then, on each round n:

3. The PT selects the SU z with the highest Gittins index $M_z^{(n)}(x_z^{(n)})$, $z \in \mathcal{S}$, and the offer $\alpha_a \in \mathcal{A}$ indicated by π_z^{SPR}.

4. If the SU *rejects* the offer, the PT updates its current knowledge state $x_z^{(n+1)}$, and the SU is represented by the Gittins index $M_z^{(n+1)}(x_z^{(n)})$ of the new state.

5. If the SU *accepts*, the PT updates $x_z^{(n+1)}$, the SU is represented by the Gittins index $M_z^{(n+1)}(x_z^{(n)})$ of the new state, and the PT sets $n_z = n_z+1$. The indices $J_s^{(n+1)}$ for all $s \in \mathcal{S}$, are also updated because $n_{\{W>0\}} = n_{\{W>0\}} + 1$ after the PT receives a positive reward.

Note that it is still not assured that the algorithm performs fully optimal learning, which, as we have already seen, is intractable. Nevertheless, *Super-UCB makes a more effective use of the available information with respect to the MAB-MDP algorithm.* This is reflected in a regret gap as shown in Section 8.5. When choosing

---

[5]This does not mean that the PT never switches arms, as that is dictated by the Gittins indices. Rather, the offer selection policy explores an SU as if the PT never switched SU.

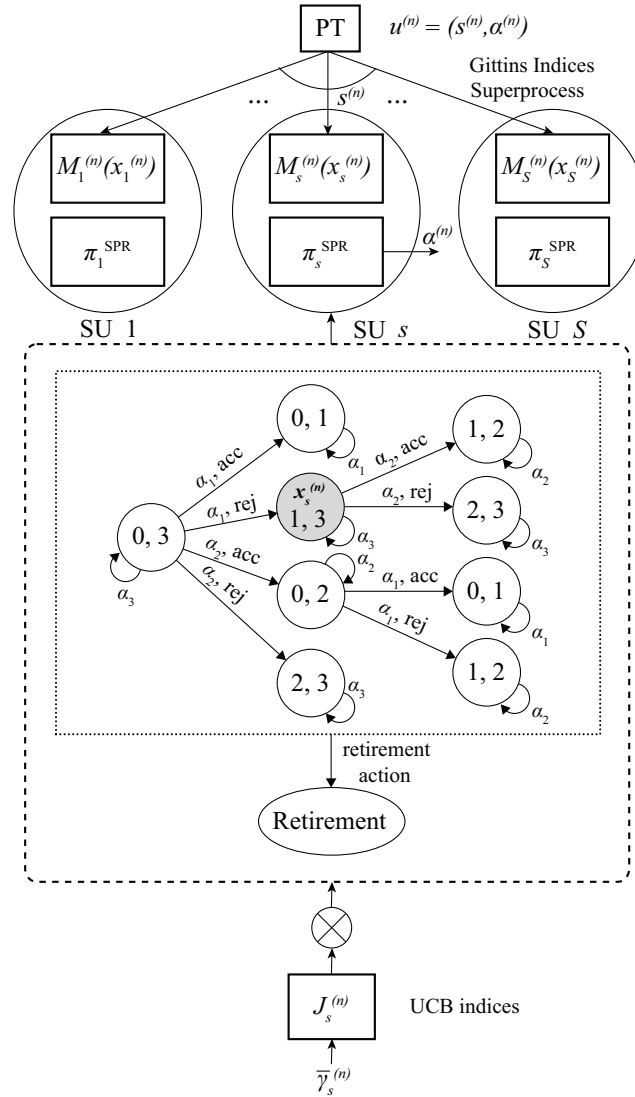[6]We use $M_s^{(0)}$ for $M_s^{(0)}(x)$, $\forall x \in \mathcal{X}_s$ to improve readability.

Figure 8.5: Super-UCB algorithm for S SUs and A = 3 possible offers. SUs and offers are chosen according to a MAB superprocess, whose reward includes a UCB index as a factor. "acc" represents the event that the offer is accepted and "rej" that it is rejected.

which SU s to bargain with, Super-UCB takes into account all the PT knowledge about the type of an SU, while MAB-MDP represents each SU by its presumed best achievable offer (or best possible type given the observations). This makes MAB-MDP more prone to over-exploration of different SUs. To illustrate this, let us consider the following example comprising two SUs, s_1 and s_2. At a given round, PT knows that s_1 accepts α_2 but has not tried α_1 on s_1 yet ($\alpha_1 < \alpha_2$). No offer has been made to s_2 so far. For MAB-MDP, both SUs are assumed to be type-1, since both of them can potentially accept α_1. Therefore, the choice made by MAB-MDP in this case would only be determined by PT's estimations of $\mu_{\gamma_{s_1}}$ and $\mu_{\gamma_{s_2}}$, that is, the sample means $\overline{\gamma}_{s_1}^{(n)}$ and $\overline{\gamma}_{s_2}^{(n)}$, and the number of samples n_{s_1} and n_{s_2} at round n, according to (8.6). Knowing that s_1 accepts at least α_2 would

Input: $\mathcal{X}_s, \mathcal{A}, P_s, \gamma_s^{(0)}; \ \forall s \in \mathcal{S}$

Output: $(s^{(n)}, \alpha^{(n)}); \ s^{(n)} \in \mathcal{S}, \alpha^{(n)} \in \mathcal{A}; \ \forall n$

1: Initialize. $\forall s \in \mathcal{S}$:
2:      $n_{\{W>0\}}, n_s^{(n)}, n \leftarrow 0$
3:      $[M_s', \pi_s^{\mathsf{SPR}}] \leftarrow$ Frontier Algorithm$(\mathcal{X}_s, \mathcal{A}, P_s)$
4:      $x_s^{(0)} \leftarrow (0, A)$
5:      $J_s^{(0)} = \gamma_s^{(0)} + 4\sigma_{\gamma_s}\sqrt{\frac{\ln\left(n_{\{W>0\}}+2\right)}{n_s^{(n)}+1}}$
6:      $M_s^{(0)} \leftarrow M_s' J_s^{(0)}$
7: **loop**
8:      $s^{(n)} \leftarrow z \leftarrow \max_s M_s^{(n)}(x_s^{(n)})$
9:      $\alpha^{(n)} \leftarrow \alpha_a \leftarrow \pi_s^{\mathsf{SPR}}(x_z^{(n)})$
10:      **if** $\alpha_a < \alpha_z^*$ **then** {SU rejects}
11:         collect reward 0
12:         $x_z^{(n+1)} \leftarrow (a, l_z^{(n)})$
13:      **else**
14:         collect reward $(1 - \alpha_a)\log(1 + \Gamma_z^{(n)})$
15:         $x_z^{(n+1)} \leftarrow (h_z^{(n)}, a)$
16:         $\overline{\gamma}_z^{(n+1)} \leftarrow \frac{\overline{\gamma}_z^{(n)} n_z + \gamma_z^{(n)}}{n_z + 1}$
17:         $n_z \leftarrow n_z + 1$
18:         $n_{\{W>0\}} \leftarrow n_{\{W>0\}} + 1$
19:         **for all** $s \in \mathcal{S}$ **do**
20:           $J_s^{(n+1)} \leftarrow \overline{\gamma}_s^{(n+1)} + 4\sigma_{\gamma_s}\sqrt{\frac{\ln\left(n_{\{W>0\}}+2\right)}{n_s+1}}$
21:           $M_s^{(n+1)} \leftarrow M_s' J_s^{(n+1)}$
22:         **end for**
23:      **end if**
24:      $n \leftarrow n + 1$
25: **end loop**

Figure 8.6: Super-UCB algorithm in pseudo-code

have no influence on the choice made.

Note that, when computing $\widehat{W}_s^{(n)}$ in MAB-MDP, if we represent the SUs by their average type instead of the presumed best possible type, the algorithm becomes greedier, making decisions mostly based on average rewards. Then, MAB-MDP would incur in over-exploitation, and its performance in terms of regret would be notably worse. The inconvenience of basing decisions on average rewards is well reported in reinforcement learning literature [203]. *Super-UCB, however, can use the information about SU types in a more balanced way thanks to the retirement action in the superprocess formulation.* The Gittins indices $M_s^{(n)}$ of an SU s are related to the expected reward of interacting with s, but that expectation takes into account the possibility of not exploring that SU infinitely and switching to a potentially more rewarding one.

Table 8.2: Summary of the information exploited by each algorithm

	$\overline{\gamma}_s$	σ_{γ_s}	Possible SU types for SU selection	Possible SU types for offer selection
ϵ-greedy	✓	×	/	/
Super-0CB	✓	×	✓	✓
MAB-MDP	✓	✓	/	✓
Super-UCB	✓	✓	✓	✓

✓ = exploited, / = partly exploited, × = not exploited

8.5   Numerical Results

In addition to the MAB-MDP, we use two more algorithms as benchmarks for Super-UCB: an ϵ-greedy algorithm and a standard superprocess which we call Super-0CB. We evaluate them over time, versus the standard deviation of the SNR of the channels σ_{γ_s}, the number of SUs S and the number of offers A. We also observe the robustness of MAB-MDP and Super-UCB against mis-estimations of σ_{γ_s}.

Comparison with ϵ-greedy. Given the main features of our proposed mechanism (real-time operation and exploration-exploitation tradeoff), the best candidates for comparison are the families of reinforcement learning heuristics [207, 203] known as ϵ-greedy policies. Specifically, we will compare our algorithm to the ϵ-linear-descending strategy. This strategy chooses the alternative $u = (s, \alpha)$ with best expected reward with probability $1 - \epsilon^{(n)}$, and performs random exploration with probability $\epsilon^{(n)}$, which, in each round n, takes the value $\epsilon^{(n)} = \min\left(\epsilon^{(0)}/n, 1\right)$, or $\epsilon^{(n)} = \min\left(\epsilon^{(0)} \ln(n)/n, 1\right)$ for a log-descending variant, where $\epsilon^{(0)}$ is a tunable parameter. The alternative u is chosen according to: $u = (s, \alpha) = \arg\max_{s,\alpha}(1 - \alpha)\overline{\gamma}_s^{(n)}P(\alpha_s^* \leq \alpha)$. As noted in the literature [199, 207], the ϵ-descending greedy strategy performs as well as (and most of the time, even better than) many other complex policies. Its main drawback is that the $\epsilon^{(0)}$ parameter has to be carefully chosen and there is little that can be said theoretically about its optimal value except for distributions with support on $[0, 1]$ (see [199]). Based on empirical tests of a wide range of values for $\epsilon^{(0)}$, we set it to 10 for all figures.

Comparison with a superprocess without upper confidence bounds (Super-0CB). We also compare to an algorithm solving the superprocess in (8.7) with $J_s^{(n)} = \overline{\gamma}_s^{(n)}$ at every round n, which we will call *Super-0CB*. This is equivalent to compute the Gittins indices at every n assuming that the true values of $\mu_{\gamma_s}^{(n+i)}$, $i = 0, 1, 2, ...,$ $\forall s \in \mathcal{S}$, are the sample means observed up to round n, that is, assuming zero upper confidence bound. We set the discount factor δ of MAB-MDP, Super-UCB, and Super-0CB to 0.98. The information taken into account by each algorithm is summarized in Table 8.2.

Performance over time

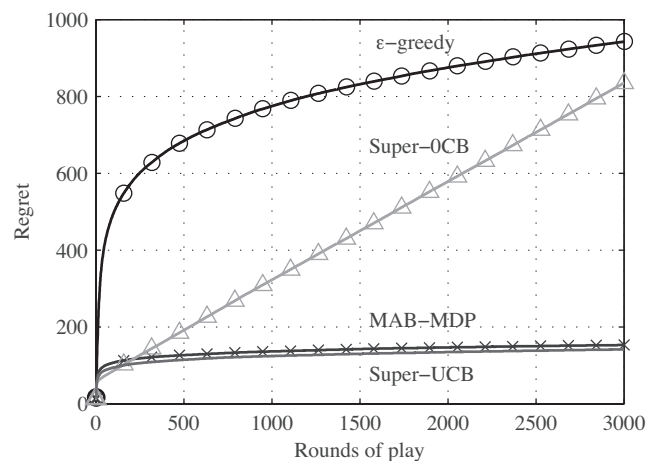Fig. 8.7 shows the performance of our two proposals MAB-MDP and Super-UCB over time compared to ϵ-greedy.

Figure 8.7: Regret over time for the MAB-MDP, Super-UCB, Super-0CB and the ε-greedy algorithms, for S = 7 SUs and A = 7 possible offers. The standard deviation of the SNR of each arm s is $\sigma_{\gamma_s} = 5$ dB and the average SNRs μ_{γ_s} are chosen uniformly from the range $[5, 40]$ dB. The results are averaged over 2000 independent experiments.

As expected, MAB-MDP and Super-UCB outperform ε-greedy because they make a more effective use of information in their decision processes. The shape of both curves of the MAB-MDP and Super-UCB is the same but, as previously anticipated, with a gap. The reason for this gap is the loss of information in the MAB-MDP algorithm.

No surprise is also what happens with Super-0CB over time: the regret increases linearly on average due to the cases where the best arm showed bad initial samples and thus, a suboptimal arm is selected. It does perform better than the other algorithms in the initial frames simply because it does more exploitation than MAB-MDP and Super-UCB, and still balances exploration and exploitation of the offers of SUs unlike the ε-greedy.

Performance versus standard deviation

Figure 8.8 shows the performance of the algorithms versus different values of σ_{γ_s}. The standard deviation for log-normal shadowing varies from 4 to 13 dB depending on the environment, and is related to the variations of the blockage that objects cause in the signal path [28]. The higher the deviations, the harder the problem, as the samples of each distribution may not preserve the rank of their mean values, that is to say, bad arms may appear as the best ones and viceversa. The MAB-MDP and Super-UCB standard deviation terms are adjusted accordingly.

ε-greedy and Super-0CB algorithms experience notable performance degradation under high variance. This is because they rely more on the estimates of the mean reward. As the variance increases, more rounds are needed for the sample mean to converge to the true mean. Nevertheless, due to its random exploration nature, ε-greedy manages to eventually find the optimal arm, which is something that cannot be said about Super-0CB and explains its poor results. The proposed
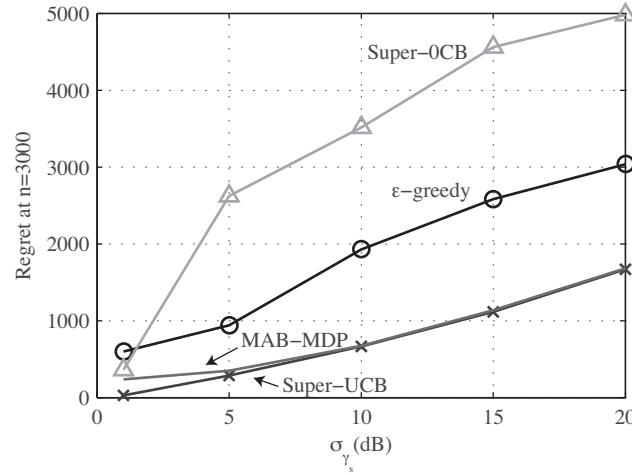
Figure 8.8: Regret at $n = 3000$ for the MAB-MDP and the ϵ-greedy algorithms for different values of σ_{γ_s}. Mean SNRs and SU types are fixed over experiments. $\mu_\gamma = \{35, 32.5, 32.5, 30, 30, 27, 27\}$ and SU types $= \{1, 1, 1, 2, 2, 3, 3\}$.

UCB-based algorithms show certain robustness under high variances, since they take the standard deviation into account when making their decisions.

Performance under mis-estimation of standard deviation

Fig. 8.9 shows the effect of mistestimating the standard deviation σ_{γ_s} in our algorithms. ϵ-greedy and Super-0CB perform a frequentist inference and therefore do not make use of the variance.

Underestimating or overestimating σ_{γ_s} in the UCB index implies being less or more optimistic, respectively, about the values of the true average rewards of the SUs, based on the sample means observed. Thus, an overestimation implies more exploration of each SU and an increased probability of finding the optimal arm. An underestimation implies more exploitation of SUs and less exploration. As shown in the figure, overestimation is safer than underestimation in terms of regret. Overestimation causes a slower convergence towards the optimal arm but underestimation could make the algorithm converge to a suboptimal solution, which is associated to a linear regret with the number of rounds. The difference between the rewards of the optimal and of that suboptimal arm in each round accumulates over time. Note that, under severe underestimation, both MAB-MDP and Super-UCB can even do worse than the ϵ-greedy algorithm, but no worse than Super-0CB Underestimation may be justified in a scenario where the SUs are expected to stay a short time in the coverage area of the PT. Then, the PT may be better off sacrificing the search of an optimal arm for exploiting suboptimal arms with "good enough" rewards.

Performance against number of SUs and discretization levels of the offers

Figs. 8.10 and 8.11 illustrate how regret grows with respect to the optimal policy for the different algorithms when increasing the number of SUs in the system
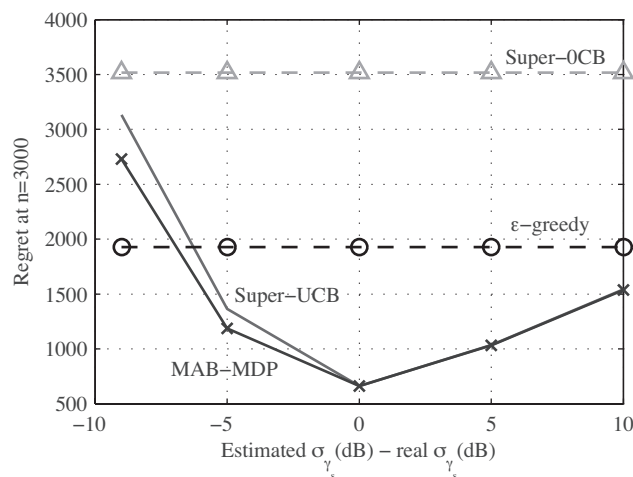
Figure 8.9: Regret at $n = 3000$ for different estimations of σ_{γ_s}, with the real value of $\sigma_{\gamma_s} = 10$ dB. Mean SNRs and SU types are fixed over experiments.

or the granularity of the offers the PT can make. Note that a higher regret represents worse performance with respect to the optimal policy. With a higher number of SUs or a higher number of possible offers, the ideal reference policy obtains better rewards. In short: finding the optimal arm is harder but that optimal arm provides a higher reward than the optimal arm of a scenario with fewer SUs or fewer possible offers. That said, both figures show that the proposed algorithms scale well when increasing the size of the problem, as their regret grows dramatically less than for the ϵ-greedy or the basic Super-0CB algorithm. This is because of the more effective use of the available information made by the UCB-based algorithms. Moreover, the computational cost or memory requirements do not experience a significant growth, even in the worst case, which is increasing the number of offers. The state space of the MDP of an SU, $\mathcal{X}_s$, grows with the number of offers, A, as $|\mathcal{X}_s| = A(A+1)/2$, which for a fine-grained range of 20 possible offers becomes a tractable problem with 210 states and 20 actions in each state.

8.6 Conclusion and Future Work

We have proposed a spectrum trading mechanism in Cooperative Spectrum Sharing (CSS) using multi-armed bandits (MABs), from the perspective of a primary transmitter (PT), modeling channels under shadowing effects. We have focused on a scenario where the PT has no knowledge of the performance of the SUs acting as relays, or about the offers they are willing to accept. Our two algorithms, MAB-MDP and Super-UCB, have been shown to be able to learn payoff-maximizing actions for the PT with little communication or computation overhead. Our numerical results indicate that, despite their simplicity, they significantly outperform the classical exploration-exploitation ϵ-greedy algorithm, with Super-UCB featuring a better overall performance. They are shown to be robust to inaccuracies
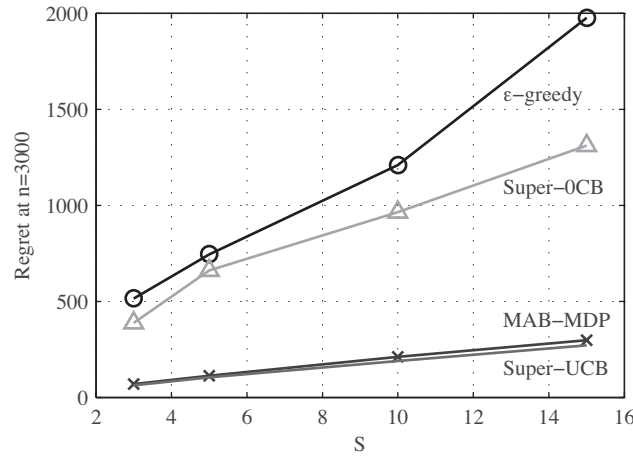
Figure 8.10: Regret at $n = 3000$ vs. the number of SUs, S. Number of possible offers $A = 7$. Again, for every arm s, $\sigma_{\gamma_s} = 5$ dB and μ_{γ_s} are chosen uniformly from the range $[5, 40]$ dB.
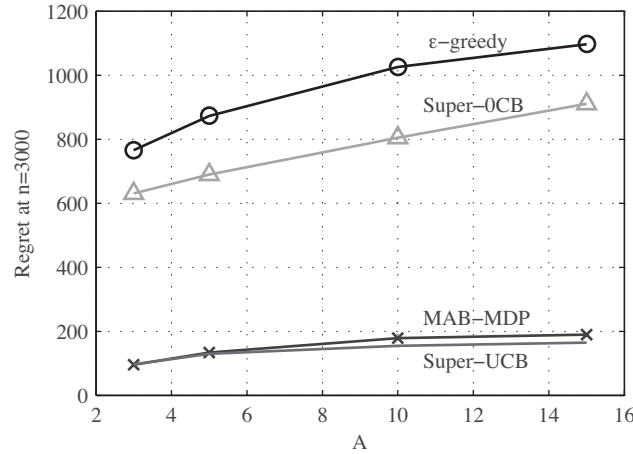

Figure 8.11: Regret at $n = 3000$ vs. the number of possible offers, A. Number of SUs $S = 7$.

in the little information they need and to scale well when the size of the problem increases, *i.e.,* for more SUs and available offers. This work can be the starting point to develop more complex scenarios. Considering the explosion of MAB variants in the recent literature, as the next steps it would be possible and interesting to study: 1) how to exploit the spatial fading correlation across different SUs, 2) extension of the algorithms to a multiple PT and multiple PR case, 3) inclusion of more dimensions to learning, such as learning the staying time of SUs in the PT coverage area or the distribution of the SU types, and 4) inclusion of SU and PU strategic behaviors.

Final remarks

Summary of conclusions

We have addressed the shortcomings of existing Dynamic Spectrum Access (DSA) mechanisms, especially in the protection and incentives for licensed users. Our proposals bring novel looks to DSA's open challenges.

OSA. Under the framework of Opportunistic Spectrum Access (OSA), we have specifically explored two areas: the sensing-throughput trade-off in hardware-constrained radios (chapters 2 and 4) and PU spectrum management for coexistence-friendly SU access (chapters 3 and 4).

We have improved the SU's periodic sensing approach by adding a novel background detection mechanism (chapter 2), which achieves better SU throughput while reducing collision probability with the PU. Our mechanism is simple to implement and robust against estimation inaccuracies. In chapter 4, we have proposed an MDP-based sequential sensing policy that takes into account the loss of reliability of a sensing result over time. Such policy improves SU throughput and reduces collision probability with the PUs. Many previous works have considered that channel states are static during the sensing phase, which is a strong assumption in many real scenarios.

We have shown that, through coexistence-friendly spectrum allocation, a primary operator could significantly improve SU's throughput while improving or not harming its own. More specifically, the primary operator could opt for either reserving part of its spectrum for SU access (chapter 3), or allocating channels to primary users in a sequential way (chapter 4). Therefore, a primary operator would use this simple coexistence measure in its self-interest, even though the initial motivation of cognitive radio (especially under OSA) was SU transparent access.

Finally, we showed an application of OSA beyond the abstract terms "PU" and "SU": a mechanism for an operator to reuse its legacy licensed and underused cellular network (PUs) by newer cognitive terminals (SUs) (chapter 5). This mechanism is based on a novel manner of simultaneously exploiting temporal and spatial spectrum opportunities, improving the system's capacity while protecting the legacy PUs. The proposed mechanism comprises an optimal learning algorithm to find the solution of an stochastic optimization problem. As additional benefits, this approach is highly adaptive and model-free.

Automated spectrum trading. Regarding automated spectrum trading, we have proposed two mechanisms providing different types of incentives to PUs: economic (chapter 7) and relay services (chapter 8). The mechanism described in

chapter 7, which is especially well-suited for lightly-loaded primary networks, aims to balance the blocking probability of PUs with the revenue obtained from leasing channels to SUs. The work developed in chapter 8 is more appropriate for heavy-loaded primary networks. In that scenario, the SUs help to increase the data rate of the PUs in exchange for spectrum opportunities and as a consequence, create more spectrum opportunities. This later mechanism does not assume any knowledge about the SUs and thus, incorporates a learning mechanism of payoff-maximizing actions. Both proposed mechanisms are simple to implement and to extend to more complex scenarios, do not make unrealistic assumptions and offer a dramatic improvement over baseline approaches.

In addition, we have elaborated a comprehensive taxonomy in the topic providing evidences of one of our main points in this thesis: that the research community was not addressing the incentives problem of PUs correctly. As we stated in the introduction of the thesis, researchers have been focused on maximizing the profit of PUs, and as a consequence, building increasingly complex models that are not suitable for real-time operation.

So what?

We believe this thesis represents a significant push in making DSA a more attractive mechanism for the wireless industry to meet the explosion in mobile data demand.

Such attractiveness is relevant given the reluctance to implement DSA of incumbent mobile operators and other spectrum owners like TV broadcasters, despite their claims of a "spectrum crisis" [208]. To date, both regulators and operators have only taken baby steps towards exploiting the full potential of DSA, under any of its possible frameworks [209, 11].

As we stated in the introduction, the wireless industry main concerns regarding DSA are interferences and the absence of incentives for them to adopt it. They do not only argue that the risk of interference is harmful to their activities (when they act as primary users), but also that it discourages investments (when they act as secondary users, for example, in the federal band between 1755-1850 Mhz), since they cannot predict how much they will be able to exploit a given shared band [86]. Another surprising objection is, as Chris Guttman-McCabe (vice president of regulatory affairs at CTIA - The Wireless Association) puts it, that "we are not aware that [cognitive] technology exists in a really truly commercially viable or scalable format [for advanced spectrum sharing]" [210]. The principal interest of the wireless industry is for the governments to clear more spectrum for their exclusive use [211, 210].

Most of the work during the last 10 years in DSA (1367 works listed in Thompson Reuters' Web of Science) including this thesis, as well as important figures such as the President's Council of Advisors on Science and Technology (PCAST) [187], or Martin Cooper, the inventor of the cellphone [208, 212], counterargument the industry's objections. Martin Cooper asserts, however, that the true reason behind the industry's opposition to sharing is that "licensing spectrum is a zero-sum game.

When a company gets the license for a band of radio waves, it has the exclusive rights to use it. Once a company owns it, competitors can't have it".

Regulators also have some objections to the widespread use of DSA, namely political pressures from the industry (*e.g.* the broadcasting lobby with regard to TV white spaces [209, p. 103]), and the loss of an important source of revenue: spectrum auctions [213].

In any case, the work developed in this thesis is also relevant to other approaches enjoying more support from the industry, such as DSA for coexistence with legacy networks, as our proposal in chapter 5, and for coexistence with small cells (HetNets) enriched with cognitive capabilities, as well as "voluntary" spectrum trading.

Future Work

The conflict of interest between the industry and social welfare has no trivial solution and may even depend on political views on the matter.

It is worth noting that the wireless industry does not roundly refuse DSA. In fact, they claim to be already using it within their networks and have sharing agreements *e.g.* for Amazon's Kindle or for Garmin's GPS navigator [86]. What they refuse is enforced, unpredictable sharing of spectrum. Thus, regarding the viability of adoption, it is a safer approach for the research community to further study those approaches approved by the incumbents:

- Authorized / Licensed Shared Access

- HetNets

- Automated spectrum trading

- Millimeter wave spectrum

Authorized or Licensed Shared Access (LSA) is a DSA initiative proposed by the industry. For SUs to transmit in a spectrum band operating under LSA, they must obtain a license from the PU and follow a set of negotiated conditions. LSA relies in geo-location databases, as an alternative to spectrum sensing, preferred by incumbents and regulators alike 3 We could also include in this point DSA within the operator's networks. As we modeled in chapter 5, an operator may be willing to access its legacy network spectrum resources with newer terminals, without incurring in the cost of upgrading its old equipment to support coexistence. The challenges to face in this scenario are similar to those in plain OSA, although protection to PUs is likely to lose importance in favor of overall performance, seeking Pareto optimal points of operation (as we did in chapter 4).

As we introduced in chapter 4, *HetNets* also imply two-layer networks, in this case an overlay of small cells over a macro-cell network, usually with no structured deployment. Due to their unplanned nature and density, it is convenient to equip these small cells with cognitive abilities to avoid interference with macro cell transmissions. As in the previous approach, a substantial difference with classical interweave OSA is that all the HetNet users are PUs and thus (users of the same

operator), and QoS must be guaranteed for both the macro and the small cell users.

With regard to *automated spectrum trading*, summarizing what we explained in the conclusions of chapter 6, the main focus should be on real-time adaptation instead of optimality, applications in practical network environments, alternatives to economic trades such as Cooperative Spectrum Sharing, protection against market failures (collusions, lack of rationality, etc.), and more complex analytic studies (cooperative game theory, stochastic games, learning mechanisms).

Millimeter wave spectrum is strongly in line with the industry's desires of more spectrum. The cost and power consumption of its required equipment has decreased and previous propagation issues are now seen as surmountable [214]. They key challenge it brings is "beamforming", networking through narrow beams. Nevertheless, it also unlocks a new dimension for spectrum sharing. Combined with small cells, millimeter wave spectrum may become the future "beachfront spectrum"[7]

Although a deep study on regulation issues is out of the scope of this thesis, it is impossible to neglect its influence in the successful implementation of DSA. Policymakers should be focused on:

- Spectrum sharing policy enforcement and monitoring

- Standardization and design of general interference guidelines

- Design of incentive mechanisms for spectrum owners

- Micro-economic market studies

Policymakers should *enforce sharing in some cases*, despite the warnings of the wireless industry. This point should not be seen as contrary to our motivation or to the previous discussion, but more like a complementary and simultaneous approach, as we indicated in chapter 3. There is enough evidence of the benefits of spectrum sharing for the social welfare, and reallocation is not always convenient, as in the federal band. Contrary to the claims of the industry, this may force them to invest in new technologies to improve coexistence. Along the same line, regulators could enforce build-out milestones and roll-out obligations. In addition, complex disputes over interferences are expected, thus an intense *effort in monitoring* will be needed. Even advocates of the free-market should take into account that the transition from the current situation may need the regulator intervention to avoid harm to social welfare [95]. The degree of interventionism should be carefully balanced with its administrative burden, so as not to lose dynamism.

Regulators could also apply incentive mechanisms for spectrum owners to give up part of their spectrum. This is a controversial point, since most of them got access to spectrum for free and giving them a compensation could be considered by the society as unfair. Nevertheless, taking decisions based on historical fairness considerations instead on the current social welfare interests is a known fallacy called the "sunk-cost fallacy" [209, p. 111].

---

[7]Beachfront spectrum is how the band between 225 Mhz and 3,7 Ghz is informally called, which is highly valuated because of its propagation characteristics [187].

With respect to *standardization and interference guidelines*, their positive impact is twofold: first, they will help streamline the administrative processes, and second, they will help remove uncertainty about how much spectrum each entity gets to use, which is one of the strongest objections of the wireless industry incumbents to DSA. Standardization is especially critical in Europe, where each country has pursued its own spectrum policies.

Finally, as pointed out in the Report on Radio Spectrum Competition issues, by the European Radio Spectrum Policy Group [215], there is a need for more complex *micro-economic market studies*, such as the long-term impact of DSA and possible anticompetitive outcomes. Also quoting from the same report, "to date there is limited practical experience": more insight in real-life implementation of DSA is necessary, although there have been important steps in that direction recently [216].

Bibliography

[1] Federal Communications Commission Spectrum Policy Task Force, "Report of the Spectrum Efficiency Working Group," Tech. Rep., 2002, ET Docket No. 02-135.

[2] Commission of the European Communities, "Impact assessment accompanying document to COM(2007)697, COM(2007)698, and COM(2007)699," Tech. Rep., 2007.

[3] T. M. Valletti, "Spectrum trading," *Telecommunications Policy*, vol. 25, no. 10-11, pp. 655–670, Oct. - Nov. 2001.

[4] E. A. Jorswieck, L. Badia, T. Fahldieck, E. Karipidis, and J. Luo, "Spectrum sharing improves the network efficiency for cellular operators," *IEEE Commun. Mag.*, vol. 52, no. 3, pp. 129–136, Dec. 2013.

[5] J. W. Mayo and S. Wallsten, "Enabling efficient wireless communications: The role of secondary spectrum markets," *Information Economics and Policy*, vol. 22, no. 1, pp. 61–72, Mar. 2010, Wireless Technologies.

[6] Electronic Communications Committee (ECC), "Description of practices relative to trading of spectrum rights of use," Tech. Rep., 2011, ECC Report 169.

[7] J. Mitola and J. Maguire, G.Q., "Cognitive radio: making software radios more personal," *IEEE Personal Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[8] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.

[9] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[10] J. Silver-Greenberg and B. Protess, "Trying to be nimble, Knight Capital stumbles," *Dealbook, The New York Times*, Aug. 2012. [Online]. Available: http://dealbook.nytimes.com/2012/08/02/trying-to-be-nimble-knight-capital-stumbles/?_r=0

[11] R. B. Kelly and A. Lafrance, "Spectrum trading in the EU and the US - shifting ends and means," in *The International Comparative Legal Guide to: Telecommunication Laws and Regulations*. Global Legal Group, 2012.

[12] M. Csikszentmihalyi, *Creativity: The Psychology of Discovery and Invention*, reprint ed. Harper Perennial, 2013.

[13] J. Peha, "Approaches to spectrum sharing," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[14] J. Alcaraz, M. López-Martínez, J. Vales-Alonso, and J. Garcia-Haro, "Background detection of primary user activity in opportunistic spectrum access," in *IEEE ICC*, 2015.

[15] ——, "Bandwidth reservation as a coexistence strategy in opportunistic spectrum access environments," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 478–488, Mar. 2014.

[16] J. Alcaraz, M. López-Martínez, J. Vales-Alonso, J. Ayala-Romero, and J. Garcia-Haro, "Uncertainty-aware opportunistic spectrum access in coexistence-friendly systems," *IEEE Transactions on Cognitive Communications and Networking*, 2015.

[17] J. Alcaraz, J. Ayala-Romero, M. López-Martínez, and J. Vales-Alonso, "Combining dual tessellation and temporal opportunities for spectrum reuse in cellular systems," in *11th International Symposium on Wireless Communications Systems (ISWCS)*, Aug. 2014, pp. 486–490.

[18] ——, "Response surface methodology for efficient spectrum reuse in cellular networks," in *IEEE ICC*, 2015.

[19] M. López-Martínez, J. J. Alcaraz, J. Vales-Alonso, and J. Garcia-Haro, "Automated spectrum trading mechanisms: Understanding the big picture," *Wireless Networks*, vol. 21, no. 2, pp. 685–708, Jan. 2015.

[20] J. Alcaraz, M. López-Martínez, J. Vales-Alonso, and J. Garcia-Haro, "An MDP framework for centralized dynamic spectrum auction," in *35th IEEE Sarnoff Symposium (SARNOFF)*, May 2012, pp. 1–5.

[21] J. Alcaraz, J. Ayala-Romero, M. López-Martínez, and J. Vales-Alonso, "Multi-armed bandits with dependent arms for cooperative spectrum sharing," in *IEEE ICC*, 2015.

[22] M. López-Martínez, J. J. Alcaraz, B. L., and M. Zorzi, "A superprocess with upper confidence bounds for cooperative spectrum sharing," *IEEE Trans. Mobile Comput.*, 2015.

[23] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley-Interscience, 2005.

[24] D. Bertsekas, *Dynamic Programming and Optimal Control*, 4th ed. Athena Scientific, 2012, vol. I.

[25] W. B. Powell and I. O. Ryzhov, *Optimal Learning*. Wiley, 2012.

[26] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136–1159, Oct. 2013.

[27] J. M. White, *Bandit Algorithms for Website Optimization*. O'Reilly, 2012.

[28] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press.

[29] MATLAB, *Release 2014b*. Natick, MA: The MathWorks Inc., 2014. [Online]. Available: http://www.mathworks.com/products/matlab/

[30] Mathematica, *Version 10.0*. Champaign, IL: Wolfram Research Inc., 2014. [Online]. Available: http://www.wolfram.com/mathematica/

[31] A. Varga, *OMNeT++, Version 4.6*, 2015. [Online]. Available: http://www.omnetpp.org

[32] E. Jones, T. Oliphant, P. Peterson, and et al., *SciPy: Open Source Scientific Tools for Python. Version 0.15.0*, 2015. [Online]. Available: http://www.scipy.org

[33] RStudio: Integrated development environment for R, *Version 0.98.1091*. Boston, MA: RStudio, 2014. [Online]. Available: http://www.wolfram.com/mathematica/

[34] X. Zhou, J. Ma, G. Li, Y. H. Kwon, and A. Soong, "Probability-based optimization of inter-sensing duration and power control in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4922–4927, Oct. 2009.

[35] W.-Y. Lee and I. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845–3857, Oct. 2008.

[36] J. Zhang, L. Qi, and H. Zhu, "Optimization of MAC frame structure for opportunistic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2036–2045, Jun. 2012.

[37] G. Chockler, M. Demirbas, S. Gilbert, C. Newport, and T. Nolte, "Consensus and collision detectors in wireless ad hoc networks," in *Proc. of the 24th Annual ACM Symposium on Principles of Distributed Computing*, 2005, pp. 197–206.

[38] S. Stotas and A. Nallanathan, "Overcoming the sensing-throughput tradeoff in cognitive radio networks," in *IEEE ICC*, May 2010, pp. 1–5.

[39] E. Biglieri, A. Goldsmith, L. Greenstein, N. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. Cambridge Univ. Press.

[40] S. Huang, X. Liu, and Z. Ding, "Optimal transmission strategies for dynamic spectrum access in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 8, no. 12, pp. 1636–1648, Dec. 2009.

[41] ——, "Opportunistic spectrum access in cognitive radio networks," in *IEEE INFOCOM*, Apr. 2008.

[42] C. Zhang and K. Shin, "What should secondary users do upon incumbents' return?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 417–428, Mar. 2013.

[43]   E. Jung and X. Liu, "Opportunistic spectrum access in multiple-primary-user environments under the packet collision constraint," *IEEE/ACM Trans. Netw.*, vol. 20, no. 2, pp. 501–514, Apr. 2012.

[44]   W. Gabran, C.-H. Liu, P. Pawelczak, and D. Cabric, "Primary user traffic estimation for dynamic spectrum access," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 544–558, Mar. 2013.

[45]   P. Pawelczak, S. Pollin, H.-S. So, A. Bahai, R. Venkatesha Prasad, and R. Hekmat, "Performance analysis of multichannel medium access control algorithms for opportunistic spectrum access," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 3014–3031, Jul. 2009.

[46]   A. De Domenico, E. Strinati, and M. Di Benedetto, "A survey on MAC strategies for cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 1, pp. 21–44, Feb. 2012.

[47]   L. Gao, X. Wang, Y. Xu, and Q. Zhang, "Spectrum trading in cognitive radio networks: A contract-theoretic modeling approach," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 843–855, Apr. 2011.

[48]   H. Yu, L. Gao, Z. Li, X. Wang, and E. Hossain, "Pricing for uplink power control in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1769–1778, May 2010.

[49]   L. Yang, H. Kim, J. Zhang, M. Chiang, and C. W. Tan, "Pricing-based spectrum access control in cognitive radio networks with random access," in *IEEE INFOCOM*, Apr. 2011, pp. 2228–2236.

[50]   M. P. J.P. Vasseur and P. Demeester, *Network recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS*.   Morgan Kaufmann, 2004.

[51]   R. Ramjee, R. Nagarajan, and D. Towsley, "On optimal call admission control in cellular networks," in *IEEE INFOCOM*, vol. 1, Mar. 1996, pp. 43–50 vol.1.

[52]   J. Vazquez-Avila, F. Cruz-Perez, and L. Ortigoza-Guerrero, "Performance analysis of fractional guard channel policies in mobile cellular networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 301–305, Feb. 2006.

[53]   X. Zhu, L. Shen, and T.-S. Yum, "Analysis of cognitive radio spectrum access with optimal channel reservation," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 304–306, Apr. 2007.

[54]   J. Martinez-Bauset, V. Pla, and D. Pacheco-Paramo, "Comments on "analysis of cognitive radio spectrum access with optimal channel reservation"," *IEEE Commun. Lett.*, vol. 13, no. 10, pp. 739–739, Oct. 2009.

[55]   W. Ahmed, J. Gao, H. Suraweera, and M. Faulkner, "Comments on "analysis of cognitive radio spectrum access with optimal channel reservation"," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4488–4491, 2009.

[56] J. Lai, R. P. Liu, E. Dutkiewicz, and R. Vesilo, "Optimal channel reservation in cooperative cognitive radio networks," in *IEEE VTC*, May 2011, pp. 1–6.

[57] P. K. Tang, Y. H. Chew, L. C. Ong, and M. Haldar, "Performance of secondary radios in spectrum sharing with prioritized primary access," in *IEEE MILCOM*, Oct. 2006, pp. 1–7.

[58] G. Wu, P. Ren, and Q. Du, "Recall-based dynamic spectrum auction with the protection of primary users," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 2070–2081, Nov. 2012.

[59] J. Park, P. Paweczak, and D. Cabric, "Performance of joint spectrum sensing and MAC algorithms for multichannel opportunistic spectrum access ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 7, pp. 1011–1027, Jul. 2011.

[60] Y. Li, S. Jayaweera, M. Bkassiny, and K. Avery, "Optimal myopic sensing and dynamic spectrum access in cognitive radio networks with low-complexity implementations," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2412–2423, Jul. 2012.

[61] W. S. Jeon, J. A. Han, and D. G. Jeong, "A novel MAC scheme for multichannel cognitive radio ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 6, pp. 922–934, Jun. 2012.

[62] D. Xu, E. Jung, and X. Liu, "Efficient and fair bandwidth allocation in multichannel cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 8, pp. 1372–1385, Aug. 2012.

[63] J. Jia, Q. Zhang, and X. Shen, "HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 106–117, Jan. 2008.

[64] H. Kim and K. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 533–545, May 2008.

[65] W. Gabran, P. Pawelczak, and D. Cabric, "Throughput and collision analysis of multichannel multistage spectrum sensing algorithms," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3309–3323, 2011.

[66] H. T. Cheng, H. Shan, and W. Zhuang, "Stopping rule-driven channel access in multi-channel cognitive radio networks," in *IEEE ICC*, Jun. 2011, pp. 1–6.

[67] S. Sadr, A. Anpalagan, and K. Raahemifar, "Radio resource allocation algorithms for the downlink of multiuser OFDM communication systems," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 3, pp. 92–106, Aug. 2009.

[68] W. Rhee and J. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *IEEE VTC*, vol. 2, May 2000, pp. 1085–1089.

[69] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[70] C. Sun, G. Villardi, Z. Lan, Y. Alemseged, H. Tran, and H. Harada, "Optimizing the coexistence performance of secondary-user networks under primary-user constraints for dynamic spectrum access," *IEEE Trans. Veh. Technol.*, vol. 61, no. 8, pp. 3665–3676, Oct. 2012.

[71] X. Gelabert, O. Sallent, J. Perez-Romero, and R. Agusti, "Flexible spectrum access for opportunistic secondary operation in cognitive radio networks," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2659–2664, Oct. 2011.

[72] L. Jiao, F. Li, and V. Pla, "Modeling and performance analysis of channel assembling in multichannel cognitive radio networks with spectrum adaptation," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2686–2697, Jul. 2012.

[73] S. Tang and B. Mark, "Modeling and analysis of opportunistic spectrum sharing with unreliable spectrum sensing," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1934–1943, Apr. 2009.

[74] D. Bertsekas and J. Tsitsiklis, *Introduction to Probability*, 2nd ed. Athena Scientific, 2008.

[75] H. Solomon, *Geometric Probability*, 2nd ed. SIAM, 2008.

[76] H. ElSawy and E. Hossain, "Channel assignment and opportunistic spectrum access in two-tier cellular networks with cognitive small cells," in *IEEE GLOBECOM*, Dec. 2013, pp. 4477–4482.

[77] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multi-channel opportunistic access: structure, optimality, and performance," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5431–5440, Dec. 2008.

[78] S. Filippi, O. Cappe, and A. Garivier, "Optimally sensing a single channel without prior information: The tiling algorithm and regret bounds," *IEEE Trans. Signal Process.*, vol. 5, no. 1, pp. 68–76, Feb. 2011.

[79] U. Berthold, F. Fu, M. van der Schaar, and F. Jondral, "Detection of spectral resources in cognitive radios using reinforcement learning," in *IEEE DySPAN*, Oct. 2008, pp. 1–5.

[80] J. Lundén, V. Koivunen, S. Kulkarni, and H. Poor, "Reinforcement learning based distributed multiagent sensing policy for cognitive radio networks," in *IEEE DySPAN*, May 2011, pp. 642–646.

[81] X. Mao, H. Ji, V. Leung, and M. Li, "Performance enhancement for unlicensed users in coordinated cognitive radio networks via channel reservation," in *IEEE GLOBECOM*, Dec. 2010, pp. 1–5.

[82] P. K. Tang, Y. H. Chew, W.-L. Yeow, and L. C. Ong, "Performance comparison of three spectrum admission control policies in coordinated dynamic

spectrum sharing systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3674–3683, 2009.

[83] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology (KTH), Kista, Sweden, 2000.

[84] C. de Lima, M. Bennis, and M. Latva-aho, "Coordination mechanisms for self-organizing femtocells in two-tier coexistence scenarios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2212–2223, Jun. 2012.

[85] D. Bolch, *Queueing Networks and Markov Chains : Modeling and Performance Evaluation With Computer Science Applications*, 2nd ed. Wiley-Interscience, 2006.

[86] CTIA - The Wireless Association, "In the matter of promoting more efficient use of spectrum through dynamic spectrum use technologies. comments of CTIA - the wireless association," Tech. Rep., 2011, eT Docket No. 10-237.

[87] T. Do and B. Mark, "Joint spatial-temporal spectrum sensing for cognitive radio networks," in *CISS*, Mar. 2009, pp. 124–129.

[88] Q. Wu, G. Ding, J. Wang, and Y.-D. Yao, "Spatial-temporal opportunity detection for spectrum-heterogeneous cognitive radio networks: Two-dimensional sensing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 516–526, Feb. 2013.

[89] G. Ding, Q. Wu, F. Song, X. Li, and J. Wang, "Joint exploration and exploitation of spatial-temporal spectrum hole for cognitive vehicle radios," in *IEEE ICSPCC*, 2011, pp. 1–4.

[90] M. Khoshkholgh, K. Navaie, and H. Yanikomeroglu, "Access strategies for spectrum sharing in fading environment: Overlay, underlay, and mixed," *IEEE Trans. Mobile Comput.*, vol. 9, no. 12, pp. 1780–1793, Dec. 2010.

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

[92] E. Axell, E. G. Larsson, and D. Danev, "Capacity considerations for unco-ordinated communication in geographical spectrum holes," *Phys. Commun.*, vol. 2, no. 1-2, pp. 3–9, Mar. 2009.

[93] K. Marti, *Stochastic Optimization Methods*. Springer, 2008.

[94] E. Angün, J. Kleijnen, D. den Hertog, and G. Gürkan, "Response surface methodology with stochastic constraints for expensive simulation," *Journal of the Operational Research Society*, vol. 60, no. 6, pp. 735–746, Jun. 2009.

[95] "An analytic research on secondary-spectrum trading mechanisms based on technical and market changes," *Computer Networks*, vol. 56, no. 1, pp. 3 – 19, Jan. 2012.

[96]   J. Adler, *Raging bulls: how Wall Street got addicted to light-speed trading*,
       Aug. 2012. [Online]. Available: http://www.wired.com/business/2012/08/
       ff_wallstreet_trading/all/

[97]   S. Maharjan, Y. Zhang, and S. Gjessing, "Economic approaches for cognitive
       radio networks: A survey," *Wireless Personal Communications*, vol. 57, no. 1,
       pp. 33–51, Mar. 2011.

[98]   D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A
       market-equilibrium-based approach," *IEEE Wireless Commun. Mag.*, vol. 15,
       no. 6, pp. 71–80, Dec. 2008.

[99]   E. Hossain, D. Niyato, and Z. Han, *Dynamic spectrum access and manage-
       ment in cognitive radio networks*.   Cambridge Univ. Press, 2012.

[100] J. Huang, "Market mechanisms for cooperative spectrum trading with in-
       complete network information," *IEEE Commun. Mag.*, vol. 51, no. 10, pp.
       201–207, Oct. 2013.

[101] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimiza-
       tion approaches for cognitive radio technologies: a survey," *IEEE Wireless
       Commun. Mag.*, vol. 20, no. 2, pp. 36–42, Apr. 2013.

[102] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic ap-
       proaches for multiple access in wireless networks: A survey," *IEEE Commun.
       Surveys Tuts.*, vol. 13, no. 3, pp. 372–395, 2011.

[103] Y. Zhang, D. Niyato, P. Wang, and E. Hossain, "Auction-based resource
       allocation in cognitive radio systems," *IEEE Commun. Mag.*, vol. 50, no. 11,
       pp. 108–120, Nov. 2012.

[104] Y. Zhao, S. Mao, J. O. Neel, and J. Reed, "Performance evaluation of
       cognitive radios: Metrics, utility functions, and methodology," *Proc. IEEE*,
       vol. 97, no. 4, pp. 642–659, Apr. 2009.

[105] H. Mutlu, M. Alanyali, and D. Starobinski, "Spot pricing of secondary spec-
       trum usage in wireless cellular networks," in *IEEE INFOCOM*, Apr. 2008.

[106] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pick-
       holtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE
       J. Sel. Areas Commun.*, Jan. 2008.

[107] J. Zhang and Q. Zhang, "Stackelberg game for utility-based cooperative
       cognitiveradio networks," in *ACM MobiHoc*.

[108] Y. Yi, J. Zhang, Q. Zhang, T. Jiang, and J. Zhang, "Cooperative
       communication-aware spectrum leasing in cognitive radio networks," in *IEEE
       DySPAN*, Apr. 2010, pp. 1–11.

[109] D. Li, Y. Xu, X. Wang, and M. Guizani, "Coalitional game theoretic approach
       for secondary spectrum access in cooperative cognitive radio networks," *IEEE
       Trans. Wireless Commun.*, vol. 10, no. 3, pp. 844–856, Mar. 2011.

[110] L. Duan, L. Gao, and J. Huang, "Contract-based cooperative spectrum sharing," in *IEEE DySPAN*, May 2011, pp. 399–407.

[111] S. Jayaweera and T. Li, "Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games," vol. 8, no. 6, pp. 3300–3310, Jun. 2009.

[112] S. Jayaweera, G. Vazquez-Vilar, and C. Mosquera, "Dynamic spectrum leasing: A new paradigm for spectrum sharing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2328–2339, Jun. 2010.

[113] G. Vazquez-Vilar, C. Mosquera, and S. Jayaweera, "Primary user enters the game: Performance of dynamic spectrum leasing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 1–5, Dec. 2010.

[114] P. Maille and B. Tuffin, "Price war with partial spectrum sharing for competitive wireless service providers," in *IEEE Global Telecommunications Conference, GLOBECOM*, Nov. 2009, pp. 1–6.

[115] L. Guijarro, V. Pla, B. Tuffin, P. Maille, and J. Vidal, "Competition and bargaining in wireless networks with spectrum leasing," in *IEEE GLOBECOM*, Dec. 2011, pp. 1–6.

[116] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE J. Sel. Areas Commun.*, Jan. 2008.

[117] S. Dixit, S. Periyalwar, and H. Yanikomeroglu, "A competitive and dynamic pricing model for secondary users in infrastructure based networks," in *IEEE VTC*, 2010, pp. 1–5.

[118] S. Sengupta, M. Chatterjee, and S. Ganguly, "An economic framework for spectrum allocation and service pricing with competitive wireless service providers," in *IEEE DySPAN*, Apr. 2007, pp. 89–98.

[119] S. Sengupta and M. Chatterjee, "An economic framework for dynamic spectrum access and service pricing," *IEEE/ACM Trans. Netw.*, vol. 17, no. 4, pp. 1200–1213, Aug. 2009.

[120] J. Jia, Q. Zhang, Q. Zhang, and M. Liu, "Revenue generation for truthful spectrum auction in dynamic spectrum access," in *ACM MobiHoc*, 2009, pp. 3–12.

[121] G. Kasbekar and S. Sarkar, "Spectrum auction framework for access allocation in cognitive radio networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 6, pp. 1841–1854, Dec. 2010.

[122] F. Wang, M. Krunz, and S. Cui, "Price-based spectrum management in cognitive radio networks," *IEEE Sel. Topics in Signal Process.*, vol. 2, no. 1, pp. 74–87, Feb. 2008.

[123] J. Huang, R. A. Berry, and M. L. Honig, "Auction-based spectrum sharing," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 405–408, Jun. 2006.

[124] J. Huang, Z. Han, M. Chiang, and H. Poor, "Auction-based resource allocation for cooperative communications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1226–1237, 2008.

[125] A. Gopinathan and Z. Li, "Strategyproof auctions for balancing social welfare and fairness in secondary spectrum markets," in *IEEE INFOCOM*, Apr. 2011, pp. 3020–3028.

[126] P. Xu, S. Kapoor, and X.-Y. Li, "Market equilibria in spectrum trading with multi-regions and multi-channels," in *IEEE GLOBECOM*, Dec. 2011, pp. 1–5.

[127] Z. Ji and K. Liu, "Multi-stage pricing game for collusion-resistant dynamic spectrum allocation," *IEEE J. Sel. Areas Commun.*, Jan. 2008.

[128] Y. Zhu, B. Li, and Z. Li, "Truthful spectrum auction design for secondary networks," in *IEEE INFOCOM*, Mar. 2012, pp. 873–881.

[129] D. Niyato and E. Hossain, "Hierarchical spectrum sharing in cognitive radio: A microeconomic approach," in *IEEE WCNC*, Mar. 2007, pp. 3822–3826.

[130] ——, "Equilibrium and disequilibrium pricing for spectrum trading in cognitive radio: A control-theoretic approach," in *IEEE GLOBECOM*, Nov. 2007, pp. 4852–4856.

[131] ——, "Market-equilibrium, competitive, and cooperative pricing for spectrum sharing in cognitive radio networks: Analysis and comparison," *IEEE Wireless Commun. Mag.*, vol. 7, no. 11, pp. 4273–4283, Nov. 2008.

[132] ——, "A microeconomic model for hierarchical bandwidth sharing in dynamic spectrum access networks," *IEEE Trans. Comput.*, vol. 59, no. 7, pp. 865–877, Jul. 2010.

[133] X. Zhou and H. Zheng, "TRUST: A general framework for truthful double spectrum auctions," in *IEEE INFOCOM*, Apr. 2009, pp. 999–1007.

[134] S. Wang, P. Xu, X. Xu, S. Tang, X. Li, and X. Liu, "TODA: Truthful online double auction for spectrum allocation in wireless networks," in *IEEE DySPAN*, Apr. 2010, pp. 1–10.

[135] X. Wang, Z. Li, P. Xu, Y. Xu, X. Gao, and H.-H. Chen, "Spectrum sharing in cognitive radio networks: An auction-based approach," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 40, no. 3, pp. 587–596, Jun. 2010.

[136] L. Gao, Y. Xu, and X. Wang, "Map: Multiauctioneer progressive auction for dynamic spectrum access," *IEEE Trans. Mobile Comput.*, vol. 10, no. 8, pp. 1144–1161, Aug. 2011.

[137] H. Xu, J. Jin, and B. Li, "A secondary market for spectrum," in *IEEE INFOCOM*, Mar. 2010, pp. 1–5.

[138] D. Xu, X. Liu, and Z. Han, "Decentralized bargain: A two-tier market for efficient and flexible dynamic spectrum access," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1697–1711, 2013.

[139] O. Ileri, D. Samardzija, and N. Mandayam, "Demand responsive pricing and competitive spectrum allocation via a spectrum server," in *IEEE DySPAN*, Nov. 2005, pp. 194–202.

[140] L. Duan, J. Huang, and B. Shou, "Cognitive mobile virtual network operator: Investment and pricing with supply uncertainty," in *IEEEINFOCOM*, Mar. 2010, pp. 1–9.

[141] J. Jia and Q. Zhang, "Competitions and dynamics of duopoly wireless service providers in dynamic spectrum market," in *ACM MobiHoc*, 2008, pp. 313–322.

[142] Y. Xing, R. Chandramouli, and C. Cordeiro, "Price dynamics in competitive agile spectrum access markets," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 613–621, Apr. 2007.

[143] L. Duan, J. Huang, and B. Shou, "Competition with dynamic spectrum leasing," in *IEEE DySPAN*, Apr. 2010, pp. 1–11.

[144] ——, "Duopoly competition in dynamic spectrum leasing and pricing," *IEEE Trans. Mobile Comput.*, vol. 11, no. 11, pp. 1706–1719, Nov. 2012.

[145] ——, "Investment and pricing with spectrum uncertainty: A cognitive operator's perspective," *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1590–1604, Nov. 2011.

[146] A. Min, X. Zhang, J. Choi, and K. Shin, "Exploiting spectrum heterogeneity in dynamic spectrum market," *IEEE Trans. Mobile Comput.*, vol. 11, no. 12, pp. 2020–2032, Dec. 2012.

[147] H. Kim, J. Choi, and K. Shin, "Wi-fi 2.0: Price and quality competitions of duopoly cognitive radio wireless service providers with time-varying spectrum availability," in *IEEE INFOCOM*, Apr. 2011, pp. 2453–2461.

[148] K. Zhu, D. Niyato, P. Wang, and Z. Han, "Dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1136–1145, Mar. 2012.

[149] G. Zhang, K. Yang, J. Song, and Y. Li, "Fair and efficient spectrum splitting for unlicensed secondary users in cooperative cognitive radio networks," *Wireless Personal Communications*, vol. 71, no. 1, pp. 299–316, 2013.

[150] J. R. Vidal, V. Pla, L. Guijarro, and J. Martinez-Bauset, "Dynamic spectrum sharing in cognitive radio networks using truthful mechanisms and virtual currency," *Ad Hoc Networks*, vol. 11, no. 6, pp. 1858 – 1873, Aug. 2013.

[151] D. Niyato and E. Hossain, "A game-theoretic approach to competitive spectrum sharing in cognitive radio networks," in *IEEE WCNC*, Mar. 2007, pp. 16–20.

[152] D. Niyato, E. Hossain, and Z. Han, "Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: A game-theoretic modeling approach," *IEEE Trans. Mobile Comput.*, vol. 8, no. 8, pp. 1009–1022, Aug. 2009.

[153] X. Zhou, S. Gandhi, S. Suri, and H. Zheng, "eBay in the sky: Strategy-proof wireless spectrum auctions," in *ACM MobiCom*, 2008, pp. 2–13.

[154] L. Duan, L. Gao, and J. Huang, "Cooperative spectrum sharing: A contract-based approach," *IEEE Trans. Mobile Comput.*, vol. 13, no. 1, pp. 174–187, Jan. 2014.

[155] Y. Yan, J. Huang, and J. Wang, "Dynamic bargaining for relay-based cooperative spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1480–1493, Aug. 2013.

[156] X. Feng, G. Sun, X. Gan, F. Yang, and X. Tian, "Cooperative spectrum sharing in cognitive radio networks: A distributed matching approach," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2651–2644, Aug. 2014.

[157] L. Canzian, L. Badia, and M. Zorzi, "Promoting cooperation in wireless relay networks through stackelberg dynamic scheduling," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 700–711, Feb. 2013.

[158] T. T. Tran and H. Y. Kong, "Exploitation of diversity in cooperative spectrum sharing with the four-way relaying af transmission," *Wireless Personal Communications*, vol. 77, no. 4, pp. 2959–2980, Aug. 2014.

[159] C. Shao, H. Roh, and W. Lee, "Aspiration level-based strategy dynamics on the coexistence of spectrum cooperation and leasing," *IEEE Communications Letters*, vol. 18, no. 1, pp. 70–73, Jan. 2014.

[160] A. Goldsmith, S. Jafar Ali, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894 – 914, May 2009.

[161] Y. Han, S. H. Ting, and A. Pandharipande, "Secondary user selection," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2914–2923, Sep. 2010.

[162] C. Courcoubetis and R. Weber, *Pricing communication networks: economics, technology and modelling.* Wiley, 2003.

[163] K. Liu and B. Wang, *Cognitive radio networking and security: A game-theoretic view.* Cambridge Univ. Press, 2010.

[164] Y. Yan, J. Huang, and J. Wang, "Dynamic bargaining for relay-based cooperative spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1480–1493, Aug. 2013.

[165] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "Towards real-time dynamic spectrum auctions," *Comput. Netw.*, vol. 52, no. 4, pp. 879–897, Mar. 2008.

[166] F. R. Yu, *Cognitive radio mobile ad hoc networks*. Springer, 2011.

[167] S. Shen, X. Lin, and T. Lok, "Dynamic spectrum leasing under uncertainty: A stochastic variational inequality approach," in *IEEE WCNC*, Apr. 2013, pp. 727–732.

[168] Y. Tan, S. Sengupta, and K. Subbalakshmi, "Competitive spectrum trading in dynamic spectrum access markets: A price war," in *IEEE GLOBECOM*, Dec. 2010, pp. 1–5.

[169] L. Gao, J. Huang, Y.-J. Chen, and B. Shou, "An integrated contract and auction design for secondary spectrum trading," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 581–592, Mar. 2013.

[170] M. Pan, S. Liang, H. Xiong, J. Chen, and G. Li, "A novel bargaining based dynamic spectrum management scheme in reconfigurable systems," in *IC-SNC*, Oct. 2006, pp. 54–54.

[171] R. B. Myerson, *Game theory: analysis of conflict*. Hardvard Univ. Press, 1997.

[172] Z. Ji and K. Liu, "WSN03-3: Dynamic pricing approach for spectrum allocation in wireless networks with selfish users," in *IEEE GLOBECOM*, Nov. 2006, pp. 1–5.

[173] Y. Shoham, R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?" *Artificial Intelligence*, vol. 171, no. 7, pp. 365 – 377, 2007, foundations of Multi-Agent Learning.

[174] Y. Yan, J. Huang, X. Zhong, M. Zhao, and J. Wang, "Sequential bargaining in cooperative spectrum sharing: Incomplete information with reputation effect," in *IEEE GLOBECOM*, Dec. 2011, pp. 1–5.

[175] G. Hardin, "The tragedy of the commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, Dec. 1968.

[176] Z. Han, C. Pandana, and K. Liu, "Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning," in *IEEE WCNC*, Mar. 2007, pp. 11–15.

[177] M. Maskery, V. Krishnamurthy, and Q. Zhao, "Decentralized dynamic spectrum access for cognitive radios: cooperative design of a non-cooperative game," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 459–469, Feb. 2009.

[178] T. Levent-Levi, *Will MVNOs live long and prosper?* Amdocs blogs, Jun. 2012. [Online]. Available: http://blogs.amdocs.com/voices/2012/06/04/what-do-i-want-for-my-22nd-birthday-my-own-mvno-please/

[179] S. Kim, "A repeated bayesian auction game for cognitive radio spectrum sharing scheme," *Computer Communications*, vol. 36, no. 8, pp. 939 – 946, May 2013.

[180] B. Wang and K. Liu, "Advances in cognitive radio networks: A survey," *IEEE Sel. Topics in Signal Process.*, vol. 5, no. 1, pp. 5–23, Feb. 2011.

[181] D. Niyato, E. Hossain, and Z. Han, "Dynamic spectrum access in IEEE 802.22- based cognitive wireless networks: a game theoretic model for competitive spectrum bidding and pricing," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 16–23, Apr. 2009.

[182] K. Lee, O. Simeone, C.-B. Chae, and J. Kang, "Spectrum leasing via cooperation for enhanced physical-layer secrecy," in *IEEE ICC*, Jun. 2011, pp. 1–5.

[183] F. Pantisano, M. Bennis, W. Saad, and M. Debbah, "Spectrum leasing as an incentive towards uplink macrocell and femtocell cooperation," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 617–630, Apr. 2012.

[184] Y. Yi, J. Zhang, Q. Zhang, T. Jiang, and J. Zhang, "Cooperative communication-aware spectrum leasing in cognitive radio networks," in *IEEE DySPAN*, Apr. 2010, pp. 1–11.

[185] T. Galla and J. D. Farmer, "Complex dynamics in learning complicated games," *Proceedings of the National Academy of Sciences*, vol. 110, no. 4, pp. 1232–1236, Jan. 2013.

[186] J. Elias, F. Martignon, L. Chen, and E. Altman, "Joint operator pricing and network selection game in cognitive radio networks: Equilibrium, system dynamics and price of anarchy," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4576–4589, Nov. 2013.

[187] President's Council of Advisors on Science and Technology (PCAST), "Report to the president: Realizing the full potential of government-held spectrum to spur economic growth," Tech. Rep., 2012.

[188] O. Yu, E. Saric, and A. Li, "Dynamic control of open spectrum management," in *IEEE WCNC*, Mar. 2007, pp. 127–132.

[189] L. Kleinrock, *Queuing Systems*. Wiley, vol. 1: Theory.

[190] T. Nadkar, V. Thumar, G. Shenoy, A. Mehta, U. B. Desai, and S. N. Merchant, "A cross-layer framework for symbiotic relaying in cognitive radio networks," in *IEEE DySPAN*, May 2011, pp. 498–509.

[191] X. Yuan, Y. Shi, Y. T. Hou, W. Lou, and S. Kompella, "UPS: A united cooperative paradigm for primary and secondary networks," in *IEEE 10th International Conference on Mobile Ad-Hoc and Sensor Systems*, 2013, pp. 78–85.

[192] S. Jayaweera, M. Bkassiny, and K. A. Avery, "Asymmetric cooperative communications based spectrum leasing via auctions in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2716–2724, 2011.

[193] J. Alcaraz and M. van der Schaar, "Coalitional games with intervention: Application to spectrum leasing in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6166–6179, Nov. 2014.

[194] A. Calvo-Armengol, "On bargaining partner selection when communication is restricted," *International Journal of Game Theory*, vol. 30, no. 4, pp. 503–515, Jan. 2002.

[195] D. B. Brown and J. E. Smith, "Optimal Sequential Exploration: Bandits, Clairvoyants, and Wildcats," *Operations Research*, vol. 61, no. 3, pp. 644–665, Jun. 2013.

[196] S. Pandey, D. Chakrabarti, and D. Agarwal, "Multi-armed bandit problems with dependent arms," in *ICML*, 2007, pp. 721–728.

[197] P. Si, H. Ji, F. R. Yu, and V. C. M. Leung, "Optimal cooperative internetwork spectrum sharing for cognitive radio systems with spectrum pooling," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1760–1768, May 2010.

[198] L. Gavrilovska, V. Atanasovski, I. Macaluso, and L. A. Dasilva, "Learning and reasoning in cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1761–1777, Mar. 2013.

[199] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2-3, pp. 235–256, May 2002.

[200] J. N. Laneman and G. W. Wornell, "An efficient protocol for realizing distributed spatial diversity in wireless ad-hoc networks," in *ATIRP*, Washington, DC, 2001.

[201] S. Bubeck, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[202] J. Gittins and D. M. Jones, "A dynamic allocation index for new-product chemical research," Cambdrige University Engineering Department Report CUED/A, Cambdrige University, Tech. Rep. Mgt Stud/TR13, 1974.

[203] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. A Bradford Book, 1998.

[204] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed Bandit Allocation Indices*. Wiley, 2011.

[205] P. Whittle, "Multi-armed bandits and the Gittins index," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 42, no. 2, pp. 143–149, Dec. 1980.

[206] S. Gass and T. Saaty, "The computational algorithm for the parametric objective function," *Naval Research Logistics*, vol. 2, no. 1-2, pp. 39–45, Mar. 1955.

[207] J. Vermorel and M. Mohri, "Multi-armed bandit algorithms and empirical evaluation," in *Machine Learning: ECML*, ser. Lecture Notes in Computer Science. Springer, 2005, vol. 3720, pp. 437–448.

[208] B. X. Chen, *Carriers Warn of Crisis in Mobile Spectrum*. The New York Times, Apr. 2012. [Online]. Available: http://www.nytimes.com/2012/04/18/technology/mobile-carriers-warn-of-spectrum-crisis-others-see-hyperbole.html?pagewanted=all

[209] J. E. Nuechterlein, *Digital Crossroads*. The MIT Press, 2013.

[210] B. X. Chen, *Q.&A.: Wireless Industry Likes Spectrum Sharing, but Still Wants More*. The New York Times, May 2012. [Online]. Available: http://bits.blogs.nytimes.com/2012/05/30/qa-ctia-spectrum-sharing/?_r=0

[211] R. Roche, *Mobile Usage Continues to Increase + Projections Say Skyrocketing Demand = More Spectrum Required*. CTIA Blog, Feb. 2015. [Online]. Available: http://blog.ctia.org/2015/02/06/mobile-usage-more-spectrum-required/

[212] B. X. Chen, *Q.&A.: Martin Cooper, Father of the Cellphone, on Spectrum Sharing*. The New York Times, May 2012. [Online]. Available: http://bits.blogs.nytimes.com/2012/05/31/qa-marty-cooper-spectrum-sharing/

[213] M. A. Baker, *CTIA Statement After the Conclusion of the FCC's AWS-3 Spectrum Auction*. CTIA Blog, Jan. 2015. [Online]. Available: http://blog.ctia.org/2015/01/29/ctia-statement-after-the-conclusion-of-the-fccs-aws-3-spectrum-auction/

[214] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[215] ERG-RSPG, "Report on radio spectrum competition issues," Tech. Rep., 2009, ERG (09) 22. RSPG09-279 Rev.2.

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