Universitá degli Studi di Padova

FACOLTÀ DI INGEGNERIA
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
Corso di Laurea Magistrale in Ingegneria delle Telecomunicazioni

TESI DI LAUREA MAGISTRALE

Access Policies for Two Cognitive Secondary Users under a Primary ARQ Process

LAUREANDA:
Giulia Agostini

RELATORE:
Dr. Michele Zorzi

Anno Accademico 2013-2014

A mamma, papà,
Andrea e Sara:
le colonne portanti del mio mondo;
e a chi, da lassù,
ha sempre vegliato su di me.

Contents

List of Figures

Chapter 1

Introduction

The fact that almost all the available frequencies are assigned to licensed users and that these frequencies are often under-utilized caused a lack of spectrum resources in wireless communication which gives rise to *cognitive radio* (CR) as a way to improve spectral efficiency in wireless networks. Cognitive radio enables the licensed primary user (PU) and unlicensed secondary users (SUs) to coexist and transmit in the same frequency band; in the underlay cognitive radio approach, the smart SUs are allowed to simultaneously transmit in the licensed frequency band allotted to PU, while PU is oblivious to the presence of SU, so SU needs to control the limited interference it causes at the PU receiver. The idea of exploiting the ARQ retransmissions implemented by the PU is employed in [1], [2] and [3]. Levorato et al. in [1] consider a cognitive radio network composed of one PU and one SU and does not utilize interference cancellation (IC) at the SU receiver. Tannious and Nosratinia in [2] apply Hybrid ARQ with incremental redundancy with at most one retransmission, where the SU receiver tries to decode the PU message in the first time slot and if successful, it removes this PU message in the second time slot to improve the SU throughput. Michelusi et al. in [3] propose to exploit the intrinsic redundancy, introduced by the Type-I HARQ implemented by the PU by enabling IC at the SU receiver; in particular, they consider an arbitrary number of retransmissions and apply backward and forward IC after decoding the PU message at the SU receiver. Forward IC (FIC) provides IC on SU transmissions performed in future time slots, since the SUrx, after decoding the PU message, performs IC in the next PU retransmission attempts, if these occur. Backward IC (BIC) provides IC on SU transmissions performed in previous time slots within the same primary ARQ transmission window, whose decoding failed due severe interference from the PU. However the number of SUs is limited to one in all these papers which leverage the PU ARQ retransmission. Joda and Zorzi in [4] consider an underlay cognitive radio network that consists of two SUs and one PU in which the PU employs Type-I HARQ; exploiting the redundancy in PU retransmissions, each SU receiver applies IC to remove a successfully decoded PU message in the subsequent PU retransmissions. Using a Constrained Markov Decision Process (CMDP) model, they propose centralized optimum access policies for the two SUs in order to maximize the average SUs sum throughput under a PU throughput constraint.

The scenario we consider in our work is a cognitive radio that consists of two secondary users and one primary user in which there are no centralized mechanisms, so the SUs make their decisions independently; every SU has only a partial view of the state of system that is based on what it can observe from its own perspective. The aim of our work is to analyze the described system and find decentralized access policies for the two SUs. In order to design optimum decentralized policies for the considered scenario we should model our system by a Decentralized Constrained Partially Observable Markov Decision Process (DEC-CPOMDP) model; however this represents a very hard challenge, so in our work we decide to concentrate our efforts in the development of some valid approximation (heuristic policy) and only in the last section we briefly try to introduce the DEC-POMDP solution to our problem and to suggest some future research topics.

Chapter 2

Cognitive Radio

2.1 The Birth of Cognitive Radio

Cognitive radio has opened up a new way of sensing and utilizing wireless spectrum resources; essentially, CR is a dynamically reconfigurable radio that can adapt its operating parameters to the surrounding environment, which has made been feasible by recent advances, such as software-defined radio (SDR) and smart antennas, that enable flexible and agile access to the wireless spectrum, and thus improve efficiency in spectrum utilization significantly. So far wireless networks are characterized by a static spectrum allocation policy, where governmental agencies assign wireless spectrum to license holders on a long-term basis for large geographical region. Recently, because of the increase in spectrum demand, this policy faces spectrum scarcity in particular spectrum bands: recent studies [5] have shown that the licensed spectrum bands are severely under-utilized mainly due to the traditional command-and-control type spectrum regulation that has prevailed for decades. Under such a spectrum policy, each spectrum band is assigned to a designated party, which is given an exclusive spectrum usage right for a specific type of service and radio device. Hence, dynamic spectrum access (DSA) techniques have been recently proposed to solve these spectrum inefficiency problems. In particular, the key enabling technology of DSA techniques is CR technology which allows unlicensed users/devices to identify the un/under-utilized portions of licensed spectrum and utilize them opportunistically as long as they do not cause any harmful interference to the legacy spectrum users' communications. The temporarily unused portions of spectrum are called spectrum white spaces (WS) or spectrum holes that may exist in time, frequency and space domains. Typically, spectrum holes are considered as the total or partial lack of power in the time-frequency plane. Thus, a DSA, depicted in Fig. 2.1, which consists in 'jumping' from a spectrum hole to another allows to improve the spectrum usage. In particular there are three different DSA access model:

- *Dynamic Exclusive Access Model*: the spectrum bands are allotted to licensed users for exclusive usage. In order to introduce flexibility in the spectrum employment two approaches are proposed: spectrum property rights and dynamic spectrum allocation. The first allows the licensed users to sell or lease the portions of the spectrum assigned to them and choose the desired technology. The second allots, for a certain period and in a given place, a portion of the spectrum for some service at exclusive usage.

- *Open Sharing Model*: also called common spectrum, it employs an open spectrum sharing between equal users, i.e., it is an access model in which there are no licensed and unlicensed users and the radio-frequency spectrum is available for use by all. However, sharing spectrum between unlicensed equipment requires that mitigation techniques (e.g. power limitation, dynamic frequency selection) are imposed to ensure that these devices operate without interference.

- *Hierarchical Access Model*: this model adopts a hierarchical access in the presence of primary users (PUs), the legacy users, and secondary users (SUs), with the aim of allowing SUs to exploit the spectrum under a constraint on the interference they cause to PUs. There are two possible spectrum sharing approaches: underlay spectrum and overlay spectrum that we describe below. This is the DSA approach we consider in our CR model.

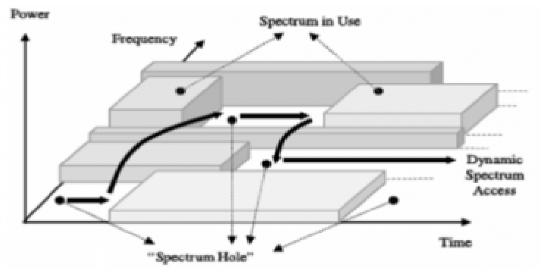


Figure 2.1: Spectrum holes and DSA

The concept of CR was first proposed in 1999 by Joseph Mitola III in his pioneering work; since then, there has been rapidly increasing interest in CR due to its potential for reshaping the way of utilizing spectrum resources: in the United States the regulations on exploiting spectrum WS have been developed by the Federal Communications Commission (FCC) that in 2000 released the first notice of proposed rulemaking, discussing the necessary actions to remove barriers to the development of the secondary spectrum market. After proposing to allow unlicensed operation in the TV white spaces (TVWS), in 2008 the FCC specified the rules in such unlicensed transmission in rural and urban areas for fixed and personal/portable devices, thus paving the way for the CR-based spectrum access. In the United Kingdom the Office of Communications (Ofcom) launched the Digital Dividend Review (DDR) project in 2005 to explore the options available after the digital TV switchover. The Ofcom proposed to allow license-exempt use of interleaved spectrum for cognitive devices and decided to allow cognitive access unless harmful interference is imposed on the licensed users. Recently, the Ofcom also proposed parameters for license-exempt CRs to provide PU protection, including those for spectrum sensing and geolocation databases ([5]).

In the industrial field, in 2008, to evaluate the potential of WS devices (WSDs), the FCC tested prototype WSDs in indoor and outdoor environments: each tested device was capable of performing a combination of functions including DTV sensing, wireless microphone sensing, transmission and geolocation. In 2009 the first public WS network was launched in Virginia using devices which led to the first large scale 'Smart City' network in North Carolina in 2010; such movement has shown that TVWS has real market value and thus draws significant attention from the industry ([5]).

There have also been efforts to create international standards to utilize TVWS using CR technology, in particular, IEEE 802.22 WRAN and Ecma 392. The former is designed for last-mile service in rural areas with fixed devices including the BS and the end-customer devices called customer premises equipment (CPE); the latter has been proposed more recently to create an international standard for the personal/portable use of TVWS in urban areas. IEEE 802.11af (also known as Wi-Fi 2.0 or White-Fi) has also been introduced as a potential application of CR that may enhance the capacity and services of current Wi-Fi systems by utilizing the TVWS, which provides better channel propagation characteristics ([5]).

2.2 Cognitive Radio and Spectrum Management in Cognitive Radio Networks

Although numerous definitions of cognitive radio exist, there are some common features which characterize a CR:

- observation capability: a CR is aware of its environment and is able to catch information from the latter;
- adaptability: a CR can dynamically and autonomously change its state and/or its operating mode according to changes in its environment;
- intelligence: a CR uses information collected by observation to make decisions in order to achieve an objective.

CR networks are envisioned to provide high bandwidth to mobile users via heterogeneous wireless architectures and DSA techniques; this goal can be realized only through dynamic and efficient spectrum management techniques. CR networks, however, impose unique challenges due to the high fluctuation in the available spectrum, as well as the diverse quality of service (QoS) requirements of various applications.

In order to address these challenges, each CR user in the CR network must:

- determine which portions of the spectrum are available;
- select the best available channel;
- coordinate access to this channel with other users;
- vacate the channel when a licensed user is detected.

These capabilities can be realized through spectrum management functions that address four main challenges: *spectrum sensing*, *spectrum decision*, *spectrum sharing* and *spectrum mobility*.

2.2.1 Cognitive Radio Technology and Network Architecture

Formally, a CR is defined as a radio that can change its tramitter parameters based on the interaction with its environment. Specifying what we have just pointed out, a CR has to have two main characteristics [6]:

- *cognitive capability*: through real-time interaction with the radio environment, the portions of the spectrum that are unused at a specific time or location (known as spectrum hole or white space) can be identified; conseguently, the best spectrum can be selected, shared with other users and exploited without interference with the licensed user.

- *Reconfigurability*: a CR can be programmed to transmit and receive on a variety of frequencies, and use different access technologies supported by its hardware design. In this way, the best spectrum and the most appropriate operating parameters can be selected and reconfigured.

As described in [6], in order to provide these capabilities, CR requires a novel radio frequency (RF) transceiver, whose main components are the radio front-end and the baseband processing unit. In the former, the received signal is amplified, mixed, and analo-to-digital converted; in the latter, the signal is modulated/demodulated. Each component can be reconfigured via a control bus to adapt to the time-varying RF environment. The novel characteristic of the CR transceiver is the wideband RF front-end that is capable of simultaneous sensing over a wide frequency range; this functionality is related mainly to the RF hardware technologies, such as wideband antenna, power amplifier and adaptive filter. RF hardware should be capable of being tuned to any part of a large range spectrum. However, because the CR transceiver receives signals from various transmitters operating at different power levels, bandwidths and locations, the RF front-end should have the capability to detect a weak signal in a large dynamic range, which is a major challenge in CR transceiver design.

As far as CR network architecture is concerned, its components are substantially two: the *primary network* and the *CR network*.

The primary (or licensed) network is referred to as an existing network, where the PUs have a license to operate in a certain spectrum band; if such a network has an infrastructure, PU activities are controlled through a primary base station. Due to their priority in spectrum access, the operations of PUs should not be affected by unlicensed users.

The CR network (or DSA network, or unlicensed network) does not have a license to operate in a desired band; hence, additional functionality is required for CR users to share the licensed spectrum band. CR networks also can be equipped with CR base stations that provide single-hop connection to CR users. Finally, they may include spectrum brokers that play a role in distributing the spectrum resources among different CR networks.

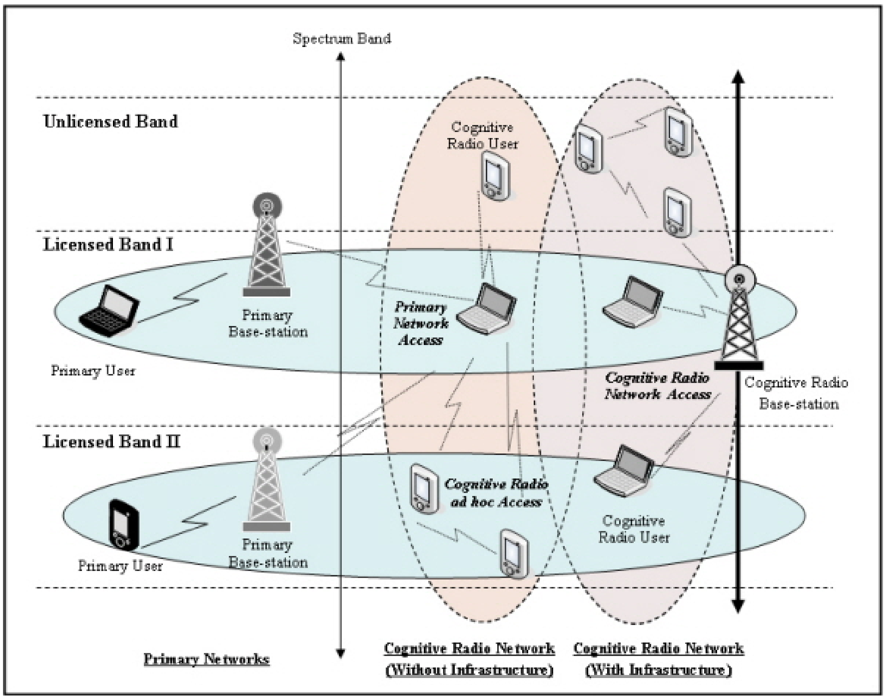


Figure 2.2: Cognitive radio network architecture

Due to spectrum heterogeneity, CR users are capable of both licensed band and unlicensed band operations. In the first case, they access the licensed portions of the spectrum that is primarily used by PUs; hence, they are focused mainly on the detection of PUs: if a PU appears in the spectrum band occupied by CR users, they should vacate that spectrum band and move to available spectrum immediately. In the second case, in the absence of PUs, the CR users have the same right to access the spectrum, so spectrum sharing methods are required to compete for the unlicensed band. Due to network heterogeneity, CR users can perform three different access types [6]:

- *CR network access*: a CR user can access its own CR base station, on both licensed and unlicensed spectrum bands. Since all interactions occur inside the CR network, its spectrum sharing policy can be independent of that of the primary network.
- *CR ad hoc access*: a CR user can communicate with other CR users through an ad hoc connection on both licensed and unlicensed spectrum bands.
- *Primary network access*: a CR user can also access the primary base station through the licensed band; unlike the other access type, CR user requires an adaptive medium access control (MAC) protocol, which enables roaming over multiple primary networks with different access technologies.

According to the CR architecture shown in Fig. 2.2 various functionalities are required to support spectrum management in CR networks.

2.2.2 Spectrum Management

CR networks impose unique challenges, so new spectrum management functions are required: in particular, CR networks should avoid interference with primary networks, should support QoS-aware communication considering the dynamic and heterogeneous environment in order to decide on an appropriate spectrum band, and should provide seamless communication regardless of the appearance of the primary users. Specifically, the spectrum management process consists of four major steps:

- *Spectrum sensing*: a CR user can allocate only an unused portion of the spectrum. Therefore, it has to monitor the available spectrum bands, capture their information and then detect spectrum holes without causing interference to the primary network. Generally, there are three type of spectrum sensing techniques [6]: primary transmitter detection, primary receiver detection and interference temperature management. Transmitter detection is based on the detection of a weak signal from a primary transmitter through the local observations of CR users; due to the lack of interactions between PUs and CR users, this type of techniques alone cannot avoid interference to primary receivers. Therefore, sensing information from other users, referred to as cooperative detection, grants more accuracy minimizing the uncertainty of a single user's detection and improving the detection probability in a heavily shadowed environment. The most efficient way to detect spectrum holes is to detect the primary users that are receiving data within the communication range of a CR user, i.e., using primary receiver detection. Interference temperature is a new model for measuring interference recently introduced by the FCC, it limits the interference at the receiver through an interference temperature threshold, which is the amount of new interference the receiver could tolerate: as long as the CR users do not exceed this limit, they can use the spectrum band.

- *Spectrum decision*: based on the spectrum availability, CR users can allocate a channel; thus, they have to be capable to decide which is the best spectrum band among the available bands according to the QoS requirements of the applications. Spectrum decision is closely related to the channel characteristics, such as interference at the PU receiver that influences the CR transmitting power, path loss, wireless link errors and link layer delay, and operations of primary users. Finally, also the activities of the other CR users in the network affect the spectrum decision. The decision procedure consists of two steps: first, each spectrum band is characterized, based on not only local observations of CR users but also statistical information of primary network. Then, based on this characterization, the most appropriate spectrum band can be chosen. In particular, to describe the dynamic nature of CR networks, a new metric is considered: the primary user activity [6], which represents the probability of a primary user appearance during a CR user transmission. Because of the PUs' operation, CR users cannot obtain a reliable communication channel for a long time period; therefore, multiple noncontiguous spectrum bands can be simultaneously used for CR users' transmissions, this method can create a signal that is immune to the interference of the PU's activity.

- *Spectrum sharing*: because there may be multiple CR users trying to access the spectrum, CR network access should be coordinated to prevent multiple users colliding in overlapping portions of the spectrum. There are three possible classification of the spectrum sharing based respectively on network architecture, allocation behaviour and access technology [6]. In a *centralized spectrum sharing* the spectrum allocation and access procedures are controlled by a central entity, while in *decentralized spectrum sharing* they are based on local (or possibly global) policies that are performed by each node distributively. From the allocation behaviour point of view, in a *cooperative spectrum sharing* collaborative solutions exploit the interference measurements of each node such that the effect of the communication of one node on other nodes is considered. Tipically, clusters are formed to share interference information locally. In a *non-cooperative spectrum sharing* only a single node is considered in non-collaborative (selfish) solutions. Because interference in other CR nodes is not considered, this type of solutions may result in reduced spectrum utilization; however, they do not require frequent message exchanges between neighbors as in cooperative solutions. Considering the access technology, in *overlay spectrum sharing* nodes access the network using a portion of the spectrum that has not been used by licensed users; this minimizes interference to the primary network. In *underlay spectrum sharing* the spread spectrum techniques are exploited such that the transmission of a CR node is regarded as noise by licensed users. Thus, underlay techniques can utilize higher bandwidth at the cost of a slight increase in complexity. Finally, spectrum sharing techniques are generally focused on two types of solutions: spectrum sharing inside a CR network (intranetwork spectrum sharing) and among multiple coexisting CR networks (internetwork spectrum sharing).

- *Spectrum mobility*: CR users are regarded as visitors to the spectrum. Hence, if the specific portion of the spectrum in use is required by a PU, the communication must be continued in another vacant portion of the spectrum, referred as spectrum handoff [6]. Since each time a CR user changes its frequency of operation, protocols for different layers of the network stack must adapt to the channel parameters for the operating frequency. The purpose of the spectrum mobility management in CR networks is to ensure smooth and fast transition leading to minimum performance degradation during a spectrum handoff.

It is evident from the significant number of interactions that the spectrum management functions require a cross-layer design approach.

2.2.3 Markov Decision Processes and Reinforcement Learning in Cognitive Radio

Markov decision processes (MDP) represent a mathematical model useful to describe and analyze decisional processes when the results are random and/or under the control of a decision maker. In CR the decision process is crucial to choose the best action in response to changes in the environment. Another useful technique in CR is reinforcement learning which concerns how an agent should make decisions in a certain environment with the aim of maximizing its reward in a long term horizon. The premise behind this approach is that an agent receives external responses to the actions it perform, i.e. a 'good' action will involve a reward, while a 'bad' action will cause a cost. Thus, the reinforcement learning algorithms search some rules which could map external environment states into actions that the agent should select to improve its reward. Typically, the environment is modelled by an MDP with a finite number of states.

Our work is structured as follows: in Chapter 3 we report the centralized access policies design developed in paper [4] which represents a reference bound for our work performance, i.e., a comparing instrument to evaluate the 'goodness' of the access policies we propose; in Chapter 4 we extend the work of [4] to the decentralized case and we find the optimal access policies for two independent SUs in a cognitive radio network under a primary ARQ process. Then, in Chapter 5, we focus on the search of decentralized heuristic access policies for the same CR scenario, in particular, we propose three offline and one online policies considering both the symmetric and the asymmetric case, i.e., since the SUs have the same transmission parameters (e.g. transmission rates, average SNRs on channels from SU_{tx} to possible network receivers), first we suppose they adopt the same access strategy with the same transmission probabilities and then we consider the possibility they adopt different transmitting behavior. Furthermore, for each heuristic proposed we present numerical results in order to analyze and compare the system performance in all considered cases. In Chapter 6 we suggest some interesting starting point to find an efficient solution to the DEC-POMDP problem; finally, we conclude with some considerations and we point out some open research problems which can be investigated in future works.

Chapter 3

Centralized Access Policies in a Cognitive Radio Network with two SUs

The design of centralized access policy for two secondary users under a primary ARQ process has jet been studied and discussed in [4] where the authors design an optimum access policy for two SUs, which exploits the redundancy introduced by the HARQ protocol in transmitting copies of the same PU message and interference cancellation at the SU receivers. The aim of the paper is the same we have, i.e., to maximize the average long term sum throughput of the SUs under a constraint on the average long term PU throughput degradation. The basic assumption they make is that the number of retransmissions is limited and both SUs have a new packet to transmit in each time slot. Noting the PU message knowledge state at each of the SU receivers and also the ARQ retransmission time, the network is modeled using a Markov Decision Process (MDP). Due to the constraint on the average long term PU throughput, they then have a Constrained Markov Decision Process (CMDP).

3.1 System Model and Policy Definition

In the system we consider, there exist one primary and two secondary transmitters denoted by PU_{tx}, SU_{tx1} and SU_{tx2}, respectively. These transmitters transmit their messages with constant power over block fading channels and, in each time slot, the channels are considered to be constant. The signal to noise ratios (SNRs) of the channels $PU_{tx} \rightarrow PU_{rx}$, $PU_{tx} \rightarrow SU_{rx1}$, $PU_{tx} \rightarrow SU_{rx2}$, $SU_{tx1} \rightarrow PU_{rx}$, $SU_{tx1} \rightarrow SU_{rx1}$, $SU_{tx1} \rightarrow SU_{rx2}$, $SU_{tx2} \rightarrow PU_{rx}$, $SU_{tx2} \rightarrow SU_{rx1}$, $SU_{tx2} \rightarrow SU_{rx2}$ are denoted by γ_{pp}, γ_{ps1}, γ_{ps2}, γ_{s1p}, γ_{s1s1}, γ_{s1s2}, γ_{s2p}, γ_{s2s1} and γ_{s2s2}, respectively. We assume that no channel State Information (CSI) is available at the transmitters. Thus, transmissions are under outage, when the selected rates are greater than the current channel capacity.

PU is unaware of the presence of the SUs and employs Type-I HARQ with at most T transmissions of the same PU message. We assume that the ARQ feedback is received by the PU transmitter at the end of the time-slot and a

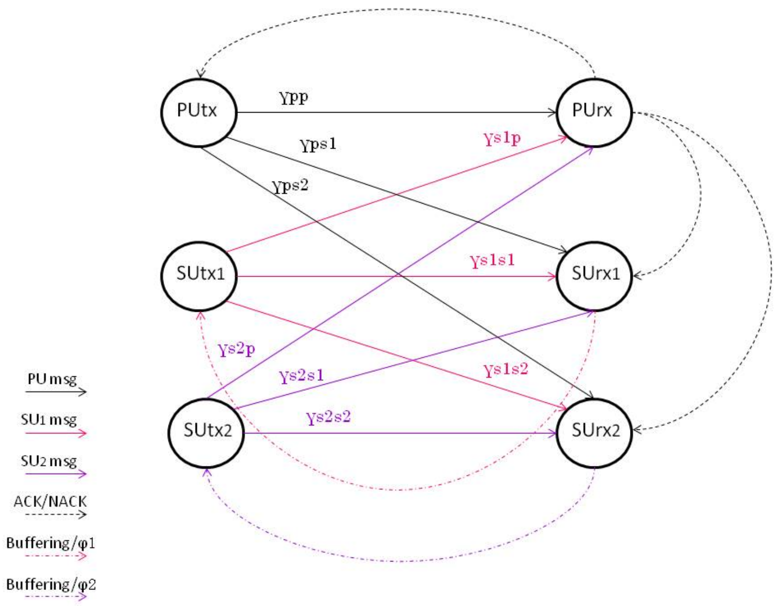


Figure 3.1: Cognitive radio system model

retransmission can be performed in the next time-slot. Retransmission of the PU message is performed if it is not successfully decoded at the PU receiver until the PU message is correctly decoded or the maximum number of transmissions allowed, T, is reached. In each time-slot, each SU, if it accesses the channel, transmits its own message, otherwise stays idle and does not transmit. This decision is based on its access policy. The activity of the SUs affects the outage performance of the PU, by creating interference to the PU receiver. The objective is to find access policies for the two SUs to maximize the average sum throughput of the SUs under a constraint on the PU average throughput degradation.

In paper [4] it is assumed that there is a central unit which controls the activities of the SUs. The central unit sends the ARQ transmission time, PU codebook, maximum transmission deadline T and feedback from PU_{rx} (ACK/NACK message). This unit also computes the secondary access probabilities and provides them to the two SUs.

In the system we want to model, we have four different combinations of the accessibility of the SUs to the channel, listed in the accessibility vector $\varphi = [\{0,0\}, \{0,1\}, \{1,0\}, \{1,1\}]$; the l^{th} element of the accessibility vector is referred to as accessibility action $l \in \mathcal{A}$, where $\mathcal{A} = \{0,1,2,3\}$. For example, $\varphi(1) = \{0,1\}$ shows that only SU_2 accesses the channel.

If SU_{rx1} or SU_{rx2} succeds to decode the PU message, it can cancel the PU message from the received signal in the future retransmissions. We refer to

this as Forward Interference Cancellation (FIC) [5]. We call the PU message kowledge state as $\phi = [\{U,U\}, \{U,K\}, \{K,U\}, \{K,K\}]$, which denotes the knowledge of the PU message at the two SU receivers. We suppose PU message knowledge state, ARQ transmission time, maximum transmission deadline, T, and feedback from PU are known to SU_{tx1} and SU_{tx2}.
Based on PU message knowledge state φ and accessibility actions of the two SUs, the rate of the secondary user i can be adapted and it is denoted by $R_{si,l,\phi}$, $i = 1, 2$ and $l \in \{1, 2, 3\}$. (If $l = 0$ the rate is zero.) Therefore, we have:

$$R_{s1,2,\{K,K\}} = R_{s1,2,\{K,U\}} = R_{s1,2,K} \tag{3.1}$$

$$R_{s1,2,\{U,K\}} = R_{s1,2,\{U,U\}} = R_{s1,2,U} \tag{3.2}$$

$$R_{s2,1,\{K,K\}} = R_{s2,1,\{K,U\}} = R_{s2,1,K} \tag{3.3}$$

$$R_{s2,1,\{U,K\}} = R_{s2,1,\{U,U\}} = R_{s1,1,U} \tag{3.4}$$

We also define $R_{s1,3,\{K,K\}} = R_{s1,3,K}$ and $R_{s2,3,\{K,K\}} = R_{s2,3,K}$. Note that we can use (1) to (4) for action 3 if the channels from SU_{tx1} to SU_{rx2} and viceversa are interference free.

The outage probabilities of the channel $PU_{tx} \to PU_{rx}$ in SU accessibility action 0, 1, 2 and 3 are denoted by $\rho_{p,0}$, $\rho_{p,1}$, $\rho_{p,2}$ and $\rho_{p,3}$, respectively. Noting that the SU_1 and SU_2 transmissions are considered as background noise at the PU_{rx}, we have:

$$\rho_{p,0} = 1 - Pr(R_p \leq C(\gamma_{pp})) \tag{3.5}$$

$$\rho_{p,i} = 1 - Pr\Big(R_p \leq C\Big(\frac{\gamma_{pp}}{1 + \gamma_{sip}}\Big)\Big) \quad i \in \{1, 2\} \tag{3.6}$$

$$\rho_{p,3} = 1 - Pr\Big(R_p \leq C\Big(\frac{\gamma_{pp}}{1 + \gamma_{s1p} + \gamma_{s2p}}\Big)\Big) \tag{3.7}$$

where R_p denotes the PU transmission rate in bits/s/Hz, $C(x) = \log_2(1 + x)$ is the (normalized) capacity of the Gaussian channel with SNR x at the receiver.

The outage probability of the channel $SU_{txi} \to SU_{rxi}$ at the PU message knowledge state ϕ and accessibility action l is denoted by $\rho_{si,l,\phi}$, $i \in \{1, 2\}$. At PU knowledge state $\{K, K\}$ or $\{K, U\}$, the PU message is known at SU_{rx1} and therefore the PU message may be canceled at this receiver. Thus, at accessibility action 2, i.e., when only SU_1 transmits its message, we have $\rho_{s1,2,\{K,K\}} = \rho_{s1,2,\{K,U\}} = \rho_{s1,2,K}$, where:

$$\rho_{s1,2,K} = Pr(R_{s1,2,K} > C(\gamma_{s1s1})) \tag{3.8}$$

In contrast, at PU knowledge $\{U, K\}$ or $\{U, U\}$, where the PU message is not decoded at SU_{rx1}, the outage probability of the channel from SU_{tx1} to SU_{rx1} is under the influence of the received PU message. Thus, at accessibility action 2, we have $\rho_{s1,2,\{U,K\}} = \rho_{s1,2,\{U,U\}} = \rho_{s1,2,U}$, where:

$$\rho_{s1,2,U} = Pr(R_{s1,2,U} \notin \Gamma_{s1}(R_{s1,2,U}, R_p)) \tag{3.9}$$

Similarly, at accessibility action 1, i.e., when only SU_2 transmits its message, we obtain $\rho_{s2,1,\{K,K\}} = \rho_{s2,1,\{K,U\}} = \rho_{s2,1,K}$ and $\rho_{s2,1,\{U,K\}} = \rho_{s2,1,\{U,U\}} = \rho_{s2,1,U}$, where:

$$\rho_{s2,1,K} = Pr(R_{s2,1,K} > C(\gamma_{s2s2})) \tag{3.10}$$

$$\rho_{s2,1,U} = Pr(R_{s2,1,U} \notin \Gamma_{s2}(R_{s2,1,U}, R_p)) \tag{3.11}$$

The SNR region $\Gamma_{si}(R_{si,j,U}, R_p)$, $i, j \in \{1, 2\}$ and $i \neq j$, is the union of two region: the first region guarantees that SU_i and PU messages, transmitted at rates $R_{si,j,U}$ and R_p, respectively, are correctly decoded at SU_{rxi} via joint decoding; on the other hand, in the second region, only SU_i message can be successfully decoded by assuming the interference from PU as background noise. Note that the other source is idle.

For accessibility action 3, i.e., when both the SUs transmit their own message, we have:

$$\rho_{s1,3,\{K,\phi_2\}} = Pr(R_{s1,3,\{K,\phi_2\}} \notin \dot{\Gamma}_{s1}(R_{s1,3,\{K,\phi_2\}}, R_{s2,3,\{K,\phi_2\}})) \tag{3.12}$$

$$\rho_{s2,3,\{\phi_1,K\}} = Pr(R_{s2,3,\{\phi_1,K\}} \notin \dot{\Gamma}_{s2}(R_{s1,3,\{\phi_1,K\}}, R_{s2,3,\{\phi_1,K\}})) \tag{3.13}$$

$$\rho_{s1,3,\{U,\phi_2\}} = Pr(R_{s1,3,\{U,\phi_2\}} \notin \ddot{\Gamma}_{s1}(R_{s1,3,\{U,\phi_2\}}, R_{s2,3,\{U,\phi_2\}}, R_p)) \tag{3.14}$$

$$\rho_{s2,3,\{\phi_1,U\}} = Pr(R_{s2,3,\{\phi_1,U\}} \notin \ddot{\Gamma}_{s2}(R_{s1,3,\{\phi_1,U\}}, R_{s2,3,\{\phi_1,U\}}, R_p)) \tag{3.15}$$

The SNR region $\dot{\Gamma}_{si}(R_{si,3,\phi}, R_{sj,3,\phi})$, $\{i, j\} \in \{\{1, 2\}, \{2, 1\}\}$, guarantees that the SU_i message transmitted at rate $R_{si,3,\phi}$ is successfully decoded at SU_{rxi} when another SU message is transmitted at rate $R_{sj,3,\phi}$. Note that the PU message received at SU_{rxi} is canceled using FIC. On the other hand, if PU message is not decoded at SU_{rxi}, the SNR region $\ddot{\Gamma}_{si}(R_{si,3,\phi}, R_{sj,3,\phi}, R_p)$, $\{i, j\} \in \{\{1, 2\}, \{2, 1\}\}$, guarantees that the SU_i message transmitted at rate $(R_{si,3,\phi}$ is successfully decoded at SU_{rxi} when other SU and PU messages are transmitted at rates $R_{sj,3,\phi}$ and R_p, respectively.
Since the value of $R_{si,j,K}$ does not affect the outage performance at PU_{rx} and SU_{rxj}, $\{i, j\} \in \{\{1, 2\}, \{2, 1\}\}$, this rate is chosen so as to maximize the SU_i throughput. Rate $R_{si,3,K}$ does not affect the outage performance at PU_{rx}; thus, the value of $R_{si,3,K}$ nad $R_{sj,3,K}$ are selected such that the SUs sum throughput is maximized, whereas the same argument can not be applied for the states in which the PU message is unknown, because in this case there is a tradeoff between the SUs sum throughput and helping the SU receivers to decode the PU message.

3.2 Optimal Access Policies for two SUs

The state of the system can be modeled by an MDP $s = (t, \phi(s))$, where $t \in \{1, 2, ..., T\}$ is the primary ARQ state and $\phi(s) \in \{\{U, U\}, \{U, K\}, \{K, U\}, \{K, K\}\}$ denotes the PU message knowledge state. The set of all states is indicated by $\mathcal{S}$.

The policy μ maps the state of the network, s, to the probability that the secondary users take accessibility action $l \in \{0, 1, 2, 3\}$. The probability that action l is selected in state s is denoted by $\mu_l(s)$. If accessibility action l is selected, the expected throughputs of SU_1 and SU_2 in state s are respectively computed as:

$$T_{s1,l,\phi(s)} = \begin{cases} R_{s1,l,\phi(s)}(1 - \rho_{s1,l,\phi(s)}) & \text{for } l \in \{2,3\} \\ 0 \quad \text{for } l \in \{0,1\} \end{cases} \tag{3.16}$$

$$T_{s2,l,\phi(s)} = \begin{cases} R_{s2,l,\phi(s)}(1 - \rho_{s2,l,\phi(s)}) & \text{for } l \in \{1,3\} \\ 0 \quad \text{for } l \in \{0,2\} \end{cases} \tag{3.17}$$

Since the model considered is a stationary Markov chain, the average long term SUs sum throughput can be obtained as:

$$\begin{aligned} \bar{T}_s(\mu) &= E_{l,s=(t,\phi(s))}[\bar{T}_{s1,l,\phi(s)} + \bar{T}_{s2,l,\phi(s)}] \\ &= E_{s=(t,\phi(s))}\Big[\sum_{l=1}^{2} \mu_l(s) R_{si,l,\phi(s)}(1 - \rho_{si,l,\phi(s)}) + \\ &\mu_3(s)(R_{s1,3,\phi(s)}(1 - \rho_{s1,3,\phi(s)}) + R_{s2,3,\phi(s)}(1 - \rho_{s2,3,\phi(s)}))\Big] \end{aligned} \tag{3.18}$$

where $\rho_{s1,2,\phi(s)}$, $\rho_{s2,1,\phi(s)}$, $\rho_{s1,3,\phi(s)}$ and $\rho_{s2,3,\phi(s)}$ are given in (3.8) to (3.15).

The average long term PU throughput is given by:

$$\bar{T}_p = R_p\Big(1 - \sum_{l=0}^{3} E_{s=(t,\phi(s))}[\mu_l(s)]\rho_{p,l}\Big) \tag{3.19}$$

Using $\mu_0 = 1 - \mu_1 - \mu_2 - \mu_3$, $\bar{T}_p$ can be rewritten as follows:

$$\begin{aligned} \bar{T}_p &= R_p\Big(1 - \sum_{l=1}^{3} E_{s=(t,\phi(s))}[\pi_l(s)]\rho_{p,l}\Big) \\ &- R_p\Big(\rho_{p,0} - \sum_{l=1}^{3} E_{s=(t,\phi(s))}[\pi_l(s)]\rho_{p,0}\Big) \\ &= T_p^I - R_p(E_{s=(t,\phi(s))}[\rho_{p,l} - \rho_{p,0}]) \end{aligned} \tag{3.20}$$

where

$$T_p^I = R_p(1 - \rho_{p,0}) \tag{3.21}$$

$\rho_{p,0}$, $\rho_{p,1}$, $\rho_{p,2}$ and $\rho_{p,3}$ are given in (3.5) to (3.7).

Thus, if we request that $\bar{T}_p \geq T_p^I(1 - \epsilon_{PU})$, the PU throughput degradation constraint can be computed as follows:

$$T_p^I - \bar{T}_p = R_p E_{l,s=(t,\phi(s))}[\rho_{p,l} - \rho_{p,0}] \leq R_p(1 - \rho_{p,0})\epsilon_{PU} \tag{3.22}$$

Now we can formalize the optimization problem as follows:

$$\text{maximize}_{\mu_l(s)} \bar{T}_s = E_{l,s=(t,\phi(s))}[\bar{T}_{s1,l,\phi(s)} + \bar{T}_{s2,l,\phi(s)}] \tag{3.23}$$

$$\text{s.t. } E_{l,s=(t,\phi(s))}[\rho_{p,l} - \rho_{p,0}] \leq (1 - \rho_{p,0})\epsilon_{PU} = \epsilon_\omega \tag{3.24}$$

$$\tag{3.25}$$

where $\pi_l(s)$ is the probability that accessibility action l is selected in state s.

In paper [4], the authors first compute an upper bound to the average long term SUs sum throughput, then they give a solution to (3.25) in low SU access rate regime and in high access rate regime; in order to do this, they provide the following definition, which identifies the boundary between low and high access rate regimes.

Definition 1: Let $\mu_{init} = \{\mu_{0,init}, \mu_{1,init}, \mu_{2,init}, \mu_{3,init}\}$ be the policy such that SU_1 or/and SU_2 in all states $s \in \mathcal{S}_\mathcal{K} = \{(t, \{K, K\}) : t \in \{1, 2, ..., T\}\}$ access the channel as follows:

$$\mu_{init} = \begin{cases} \{0,0,1,0\} & \text{if max(a,b,c)=a} \\ \{0,1,0,0\} & \text{if max(a,b,c)=b} \\ \{0,0,0,1\} & \text{if max(a,b,c)=c} \end{cases} \tag{3.26}$$

and for all other states, $s \notin \mathcal{S}_\mathcal{K}$, $\mu_{init} = \{1, 0, 0, 0\}$, where

$$a = \frac{R_{s1,2,K}(1 - \rho_{s1,2,K})}{\rho_{p,2} - \rho_{p,0}} \tag{3.27}$$

$$b = \frac{R_{s2,1,K}(1 - \rho_{s2,1,K})}{\rho_{p,1} - \rho_{p,0}} \tag{3.28}$$

$$c = \frac{R_{s1,3,K}(1 - \rho_{s1,3,K}) + R_{s2,3,K}(1 - \rho_{s2,3,K})}{\rho_{p,3} - \rho_{p,0}} \tag{3.29}$$

For access probability π_{init}, the constraint given in (3.24) can be computed and referred to as ω_{init}. Hence, replacing (3.26) in (3.24) and then computing the expectation with respect to l and s, ω_{init} can be obtained as follows:

$$\omega_{init} = \begin{cases} (\rho_{p,2} - \rho_{p,0}) \sum_{t=1}^{T} \mu(t, \{K, K\}) & \text{if max(a,b,c)=a} \\ (\rho_{p,1} - \rho_{p,0}) \sum_{t=1}^{T} \mu(t, \{K, K\}) & \text{if max(a,b,c)=b} \\ (\rho_{p,3} - \rho_{p,0}) \sum_{t=1}^{T} \mu(t, \{K, K\}) & \text{if max(a,b,c)=c} \end{cases} \tag{3.30}$$

where $\mu(t, \{K, K\})$ is the steady-state probability of being in state $s = (t, \{K, K\})$, and a, b, c are given in (3.27) to (3.29).

3.2.1 Upper Bound to the Average Long Term SUs Sum Throughput

An upper bound to the average long term SUs sum throughput is achieved when the receivers are assumed to be aware of the PU message, so that they can always cancel the PU interference. Since each SU always knows the PU message, there exist an optimal access policy which is independent of the ARQ state and therefore is the same in each slot. Thus, in this case problem (3.25) may be rewritten as follows:

$$\max_{\mu_1,\mu_2,\mu_3} \bar{T}_s = \sum_{l=1}^{2} \mu_l R_{si,l,K}(1-\rho_{si,l,K})$$

$$+ \mu_3(R_{s1,3,K}(1-\rho_{s1,3,K}) + R_{s2,3,K}(1-\rho_{s2,3,K})) \tag{3.31}$$

$$\text{s.t.} \sum_{l=1}^{3} \mu_l(\rho_{p,l} - \rho_{p,0}) \leq \epsilon_\omega \tag{3.32}$$

$$\tag{3.33}$$

In paper [4] the solution to problem (3.33) is given by the following proposition:

Proposition 1: An access policy to achieve the upper bound is given by

$$\mu^{up} = \begin{cases} \left\{1 - \frac{\epsilon_\omega}{\rho_{p,2}-\rho_{p,0}}, 0, \frac{\epsilon_\omega}{\rho_{p,2}-\rho_{p,0}}, 0\right\} & \text{if max(a,b,c)=a} \\ \left\{1 - \frac{\epsilon_\omega}{\rho_{p,1}-\rho_{p,0}}, \frac{\epsilon_\omega}{\rho_{p,1}-\rho_{p,0}}, 0, 0\right\} & \text{if max(a,b,c)=b} \\ \left\{1 - \frac{\epsilon_\omega}{\rho_{p,3}-\rho_{p,0}}, 0, 0, \frac{\epsilon_\omega}{\rho_{p,3}-\rho_{p,0}}\right\} & \text{if max(a,b,c)=c} \end{cases} \tag{3.34}$$

Furthermore, the upper bound to the average long term SUs sum throughput is obtained as:

$$\bar{T}_s^{up} = \begin{cases} \frac{\epsilon_\omega}{\rho_{p,2}-\rho_{p,0}} R_{s1,2,K}(1-\rho_{s1,2,K}) & \text{if max(a,b,c)=a} \\ \frac{\epsilon_\omega}{\rho_{p,1}-\rho_{p,0}} R_{s2,1,K}(1-\rho_{s2,1,K}) & \text{if max(a,b,c)=b} \\ \frac{\epsilon_\omega}{\rho_{p,3}-\rho_{p,0}} \sum_{i=1}^{2} R_{si,3,K}(1-\rho_{si,3,K}) & \text{if max(a,b,c)=c} \end{cases} \tag{3.35}$$

3.2.2 Low SU Access Rate Regime

In low SU access rate regime $\epsilon_\omega \leq \omega_{init}$; in paper [4] the optimum access policy is characterized by the following proposition:

Proposition 2: In the low SU rate regime $\epsilon_\omega \leq \omega_{init}$, the optimal access policy $\forall s \in \mathcal{S}_{\mathcal{K}}$ is given by

$$\mu^* = \begin{cases} \left\{1-\frac{\epsilon_\omega}{\omega_{init}}, 0, \frac{\epsilon_\omega}{\omega_{init}}, 0\right\} & \text{if max(a,b,c)=a} \\ \left\{1-\frac{\epsilon_\omega}{\omega_{init}}, \frac{\epsilon_\omega}{\omega_{init}}, 0, 0\right\} & \text{if max(a,b,c)=b} \\ \left\{1-\frac{\epsilon_\omega}{\omega_{init}}, 0, 0, \frac{\epsilon_\omega}{\omega_{init}}\right\} & \text{if max(a,b,c)=c} \end{cases} \tag{3.36}$$

and

$$\mu^* = \{1, 0, 0, 0\} \quad \forall s \notin \mathcal{S}_{\mathcal{K}} \tag{3.37}$$

Furthermore, the average long term SUs sum throughput is obtained as:

$$\bar{T}_s^* = \begin{cases} \frac{\epsilon_\omega}{\omega_{init}} R_{s1,2,K}(1-\rho_{s1,2,K}) & \text{if max(a,b,c)=a} \\ \frac{\epsilon_\omega}{\omega_{init}} R_{s2,1,K}(1-\rho_{s2,1,K}) & \text{if max(a,b,c)=b} \\ \frac{\epsilon_\omega}{\omega_{init}} \sum_{i=1}^{2} R_{si,3,K}(1-\rho_{si,3,K}) & \text{if max(a,b,c)=c} \end{cases} \tag{3.38}$$

3.2.3 High SU Access Rate Regime

In paper [4] to obtain a solution to the CMDP problem described in (3.25) in high access rate regime, i.e., $\epsilon_\omega > \omega_{init}$, the authors employ the equivalent LP formulation corresponding to CMDP. To provide the equivalent LP, it is necessary the transition probability matrix of the Markov process denoted by $\mathbf{P}$, where $P_{ss',l}$ is the probability of moving from state s to s' if the accessibility action l is chosen.
For any unichain Constrained Markov Decision Process, there exists an equivalent Linear Programming (LP) formulation, where an MDP is considered unichain if it contains a single recurrent class plus a (perhaps empty) set of transient states. Thus, the equivalent LP problem for problem (3.25) is:

$$\max_x \sum_{s\in\mathcal{S}} \sum_{l\in\mathcal{A}\times\mathcal{A}} (\bar{T}_{s1,l,\phi(s)} + \bar{T}_{s2,l,\phi(s)}) x(s,l) \tag{3.39}$$

$$\text{s.t.} \quad \sum_{s\in\mathcal{S}} \sum_{l\in\mathcal{A}\times\mathcal{A}} (\rho_{p,l} - \rho_{p,0}) x(s,l) \leq \epsilon_\omega \tag{3.40}$$

$$\sum_{l\in\mathcal{A}\times\mathcal{A}} x(s',l) - \sum_{s\in\mathcal{S}} \sum_{l\in\mathcal{A}\times\mathcal{A}} P(ss',l) x(s,l) = 0 \quad \forall s' \in \mathcal{S} \tag{3.41}$$

$$\sum_{s\in\mathcal{S}} \sum_{l\in\mathcal{A}\times\mathcal{A}} x(s,l) = 1 \tag{3.42}$$

$$x(s,l) \geq 0 \quad \forall s \in \mathcal{S}, \quad l \in \mathcal{A}\times\mathcal{A} \tag{3.43}$$

$$\tag{3.44}$$

The relationship between the optimal solution of LP problem (3.44) and the optimal solution to problem (3.25) is obtained as follows:

$$\mu_l(s) = \begin{cases} \frac{x(s,l)}{\sum_{l' \in \mathcal{A}\times\mathcal{A}} x(s,l')}, & \text{if } \sum_{l' \in \mathcal{A}\times\mathcal{A}} x(s,l') > 0 \\ \text{arbitrary}, & \text{otherwise} \end{cases} \tag{3.45}$$

3.3 Numerical results

Now we report some numerical results of paper [4]. The channels considered are Rayleigh fading, thus, the SNRs γ_x, $x \in \{pp, ps1, ps2, s1s1, s1p, s1s2, s2s2, s2p, s2s1\}$, are exponentially distributed random variables with mean $\bar{\gamma}_x$. For mathematical convenience the links from SU_{tx1} to SU_{rx2} and viceversa are assumed interference free. They consider the following parameters: the average SNRs are $\bar{\gamma}_{pp} = 10$, $\bar{\gamma}_{si} = 5$, $\bar{\gamma}_{psi} = 5$, $\bar{\gamma}_{sip} = 2$, $i \in \{1, 2\}$, and the ARQ deadline is $T = 5$. The PU rate R_p is selected such that PU throughput is maximized when both SUs are idle, i.e., $R_p = \text{argmax}_R T_p^I(R)$. The SU_i rate $R_{si,l,U}$ under PU message unknown to SU_{rxi} is computed as $R_{si,l,U} = \text{argmax}_{R_{si}} T_{si,l,U}(R_{si}, R_p)$, where $i = 1$, $l \in \{2, 3\}$ or $i = 2$, $l \in \{1, 3\}$, so as to maximize the SUs sum throughput. The SU_i rate $R_{si,l,K}$ under PU message known to the SU_{rxi} is computed as $R_{si,l,K} = \text{argmax}_R T_{si,l,K}(R)$, where $i = 1$, $l \in \{2, 3\}$ or $i = 2$, $l \in \{1, 3\}$. The PU throughput constraint is set to $(1 - \epsilon_{PU})T_p^I$, where $\epsilon_{PU} \in \{0.01, 0.05, 0.08, 0.1, 0.13, 0.15, 0.18, 0.2, 0.25, 0.3, 0.4, 0.4861, 0.6, 0.7, 0.8, 1\}$. FIC is exploited at the SU_{rxi}, $i \in \{1, 2\}$.

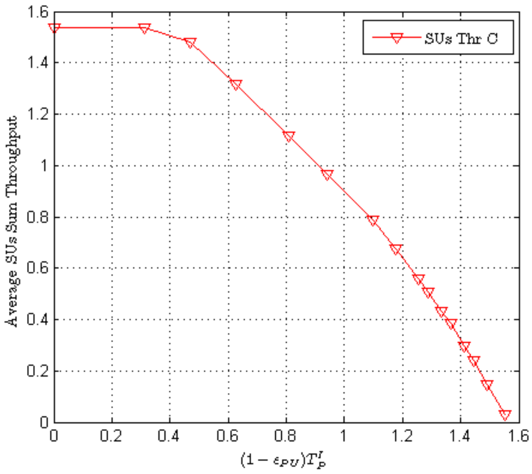


Figure 3.2: MMDP: Average SUs sum throughput with respect to PU throughput constraint

The SUs sum throughput with respect to the PU throughput constraint by varying the value of ϵ_{PU} is depicted in Fig. 3.2. Obviously, as the PU throughput increases, the average SUs sum throughput decreases.

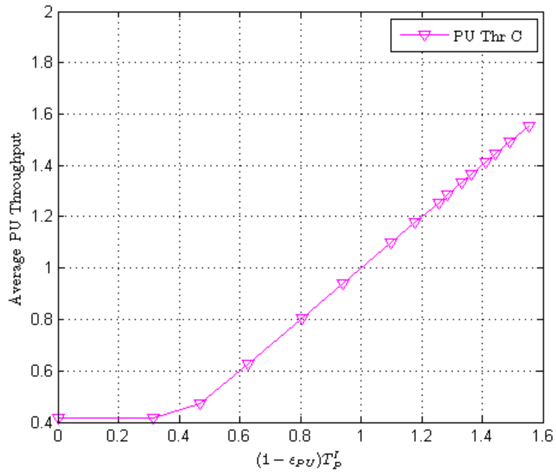


Figure 3.3: MMDP: Average PU throughput with respect to PU throughput constraint

Fig. 3.3 depicts the PU throughput with respect to the PU throughput constraint by varying the value of ϵ_{PU}. Obviously, as ϵ_{PU} decreases the constraint increases and the PU throughput degradation decreases.

Chapter 4

Decentralized Access Policies in a Cognitive Radio Network with two SUs and a Completely Observable System

The design of decentralized access policies for two secondary users under a primary ARQ process has not been studied jet. Thus, the purpose of our work is to study a decentralized CR network with two SUs and one PU exploiting the redundancy introduced by the HARQ protocol in transmitting copies of the same PU message and interference cancellation at the SU receivers. Our aim remains the same, i.e., we want to maximize the average long term sum throughput of the SUs under a constraint on the average long term PU throughput degradation. To start with a simple scenario, we suppose that the system is completely observable for the two SUs, which means that in each time-slot they know the system state, $s = (t, \phi(s)) = (t, \{\phi_1, \phi_2\})$, as in the centralized case, but in this case there is no central unit which controls the activities of the SUs, i.e., they select their own action in a certain state independently. Again we assume that the number of retransmissions is limited and both SUs have a new packet to transmit in each time slot. Noting the PU message knowledge state at each of the SU receivers and also the ARQ retransmission time, the network can again be modeled using a Markov Decision Process (MDP), and due to the constraint on the average long term PU throughput, using a Constrained Markov Decision Process (CMDP).

4.1 System Model and Policy Definition

The system we consider is the same as in the centralized case. Thus, there are the same transmitters, receivers and channels, whose signal to noise ratio is denoted in the same way, i.e., γ_{pp}, γ_{ps1}, γ_{ps2}, γ_{s1p}, γ_{s1s1}, γ_{s1s2}, γ_{s2p}, γ_{s2s1} and γ_{s2s2} are the SNR of the channels $PU_{tx} \rightarrow PU_{rx}$, $PU_{tx} \rightarrow SU_{rx1}$, $PU_{tx} \rightarrow SU_{rx2}$,

$SU_{tx1} \to PU_{rx}$, $SU_{tx1} \to SU_{rx1}$, $SU_{tx1} \to SU_{rx2}$, $SU_{tx2} \to PU_{rx}$, $SU_{tx2} \to SU_{rx1}$, $SU_{tx2} \to SU_{rx2}$, respectively.

We mantain the same assumptions as in the centralized case about the PU transmissions and retransmissions, and about the SUs activities and their influence on the outage performance of the PU. The substantial difference consists in the absence of a central unit that monitors the activities of the SUs and provides them with the optimal access policies.

In the system we want to model, we have the same accessibility vector, φ, and the same PU message knowledge state, ϕ, as in the centralized scenario; in particular, each element of φ can be considered as the couple of actions performed by the two SUs, $\{a_1, a_2\} \in \mathcal{A} \times \mathcal{A}$, in the considered time-slot, where $a_i = 1$ means that SU_i accesses the channel and transmits its message, $i = 1, 2$. The relationship between the accessibility action l in the MMDP and $\mathbf{a}= \{a_1, a_2\}$ in the DEC-MMDP is simple, since l corresponds to $\mathbf{a}$ in decimal notation. For example, $l = 1$ corresponds to $\{a_1, a_2\} = \{0, 1\}$, which means that only SU_2 accesses the channel. Similarly, each element of ϕ represents the couple of PU message knowledge at the two $SU_{rx}s$, i.e. $\{\phi_1, \phi_2\} \in \Phi \times \Phi$, where $\Phi = \{U, K\}$. The transmission rates for the possible accessible action and the outage probabilities are the same as in the centralized case, given in (3.1) to (3.4), and (3.5) to (3.15), respectively.

4.2 Optimal Access Policies for Two Independent SUs

As in the centralized case, the state of the system can be modeled by an MDP $s = (t, \phi(s))$, where $t \in \{1, 2, ..., T\}$ is the primary ARQ state and $\phi(s) \in \{\{U, U\}, \{U, K\}, \{K, U\}, \{K, K\}\}$ denotes the PU message knowledge state. Since we want to analize a decentralized scenario in which each SU selects its own action independently, there are two policies, one for each SU, denoted by π_1 and π_2, that map the state of the network s to the probability that the SU takes one of the two possible action, $a_i \in \mathcal{A}$, where $\mathcal{A} = \{0, 1\}$, i.e., each SU can accesses or not the channel, independently of the action of the other SU. The probability that action a_i is selected by SU_i in state s is denoted by $\pi_i(a_i|s)$, $i \in \{1, 2\}$. Since the SUs are independent, the relationship between the policy μ of the MMDP case and the policies π_1 and π_2 of the DEC-MMDP $\forall s \in \mathcal{S}$ is given by:

$$\begin{aligned} \mu_1(s) &= \pi_1(0|s)\pi_2(1|s) \\ \mu_2(s) &= \pi_1(1|s)\pi_2(0|s) \\ \mu_3(s) &= \pi_1(1|s)\pi_2(1|s) \end{aligned}$$

Since there is no central unit which controls the SUs activities and provides them the optimal access policy, we can adopt the following optimization strategy: first we assume that SU_2 has a fixed stochastic policy, $\pi_2(a_2|s)$ $\forall a_2 \in \mathcal{A}$, $\forall s \in \mathcal{S}$, which is known to SU_1, and we try to find SU_1's optimal stochastic policy, given that the system state, $s = (t, \phi_1, \phi_2)$, is known to both the SUs. Then we invert the perspective, we assume that SU_1 has the fixed stochastic policy just found,

$\pi_1(a_1|s)$ $\forall a_1 \in \mathcal{A}$, $\forall s \in \mathcal{S}$, which is known to SU_2, and we try to find SU_2's optimal stochastic policy, given that the system state, $s = (t, \phi_1, \phi_2)$, is known to both the SUs, and so on until the policies of the two SUs converge to the optimal ones.
Specifically:

- at optimization round k we fix SU_2's stochastic policy, $\pi_{2,k-1}(a_2|s)$ (for $k = 0$ it is randomly chosen), and find SU_1's optimal stochastic policy, $\pi_{1,k}(a_1|s)$, by solving a LP problem;
- then, we invert the perspective: we fix SU_1's stochastic policy just found, $\pi_{1,k}(a_1|s)$, and find SU_2's optimal stochastic policy, $\pi_{2,k}(a_2|s)$, by solving another LP problem.
- At this point we can check the stopping conditions: $\pi_{1,k} = \pi_{1,k-1}$, $\pi_{2,k} = \pi_{2,k-1}$; if they are satisfied we have convergence and the SU stochastic policies found are the optimal ones, i.e., $\pi_1^* = \pi_{1,k}$ and $\pi_2^* = \pi_{2,k}$, otherwise we have to start another optimization round, i.e., $k = k + 1$.

Once we have found the optimal stochastic policy for both the SUs to evaluate the performace we have to calculate the average long term SUs sum throughput, $\bar{T}_s$, and the average long term PU throughput, $\bar{T}_p$:

$$\bar{T}_s = E_{s,a}[\bar{T}_{s1,a,\phi(s)} + \bar{T}_{s2,a,\phi(s)}]$$

$$= \sum_{s\in\mathcal{S}} \Big[\sum_{a\in\mathcal{A}\times\mathcal{A}} \Big(R_{s1,a,\phi_1}(1-\rho_{s1,a,\phi_1}) + R_{s2,a,\phi_2}(1-\rho_{s2,a,\phi_2}) \Big) \pi_1(a_1|s)\pi_2(a_2|s) \Big] Pr(s) \tag{4.1}$$

$$\bar{T}_p = R_p\Big(1 - \sum_{a\in\mathcal{A}\times\mathcal{A}} E_s[\pi_1(a|s)\pi_2(a|s)]\rho_{p,a}\Big)$$

$$= R_p\Big(1 - \sum_{a\in\mathcal{A}\times\mathcal{A}} \Big(\sum_{s\in\mathcal{S}} \pi_1(a_1|s)\pi_2(a_2|s)Pr(s)\Big)\rho_{p,a}\Big) \tag{4.2}$$

where $Pr(s)$ is the stationary probability of being in state s. We know that for a regular Markov chain the stationary equations are valid: if we let $P_s(j)$ be the stationary probability of state j, we have

$$P_s(j) = \sum_{k\in\mathcal{S}} P_s(k)P_{kj} \quad \forall j \in \mathcal{S} \tag{4.3}$$

$$\sum_{k\in\mathcal{S}} P_s(k) = 1 \tag{4.4}$$

$$P_s(k) \geq 0 \quad k \in \mathcal{S} \tag{4.5}$$

where P_{kj} is the transition probability from state k to state j. In matrix form we have $\mathbf{P}_s$ =($\mathbf{I}$ - $\mathbf{P}$)$^{-1}$, where $\mathbf{P}$ is the transition probability matrix and $\mathbf{I}$ is the identity matrix. The transition probability $P_{ss',a} = Pr(s|s', a)$ if accessibility action **a**= $\{a_1, a_2\}$ is chosen in state s can be computed as:

$$Pr(s' = (t-1, U, U)|s = (t, U, U), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}\rho_{ps1,\{i,j\}}\rho_{ps2,\{i,j\}} \tag{4.6}$$

$$Pr(s' = (t-1, U, U)|s = (t, U, K), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}\rho_{ps1,\{i,j\}}(1 - \rho_{ps2,\{i,j\}}) \tag{4.7}$$

$$Pr(s' = (t-1, U, U)|s = (t, K, U), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}(1 - \rho_{ps1,\{i,j\}})\rho_{ps2,\{i,j\}} \tag{4.8}$$

$$Pr(s' = (t-1, U, U)|s = (t, K, K), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}(1 - \rho_{ps1,\{i,j\}})(1 - \rho_{ps2,\{i,j\}}) \tag{4.9}$$

$$Pr(s' = (t-1, U, K)|s = (t, U, U), a = \{i, j\}) = 0 \tag{4.10}$$

$$Pr(s' = (t-1, U, K)|s = (t, U, K), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}\rho_{ps1,\{i,j\}} \tag{4.11}$$

$$Pr(s' = (t-1, U, K)|s = (t, K, U), a = \{i, j\}) = 0 \tag{4.12}$$

$$Pr(s' = (t-1, U, K)|s = (t, K, K), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}(1 - \rho_{ps1,\{i,j\}}) \tag{4.13}$$

$$Pr(s' = (t-1, K, U)|s = (t, U, U), a = \{i, j\}) = 0 \tag{4.14}$$

$$Pr(s' = (t-1, K, U)|s = (t, U, K), a = \{i, j\}) = 0 \tag{4.15}$$

$$Pr(s' = (t-1, K, U)|s = (t, K, U), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}\rho_{ps2,\{i,j\}} \tag{4.16}$$

$$Pr(s' = (t-1, K, U)|s = (t, K, K), a = \{i, j\}) \\ = \rho_{pp,\{i,j\}}(1 - \rho_{ps2,\{i,j\}}) \tag{4.17}$$

$$Pr(s' = (t-1, K, K)|s = (t, U, U), a = \{i, j\}) = 0 \tag{4.18}$$

$$Pr(s' = (t-1, K, K)|s = (t, U, K), a = \{i, j\}) = 0 \tag{4.19}$$

$$Pr(s' = (t-1, K, K)|s = (t, K, U), a = \{i, j\}) = 0 \tag{4.20}$$

$$Pr(s' = (t-1, K, K)|s = (t, K, K), a = \{i, j\}) = \rho_{pp,\{i,j\}} \tag{4.21}$$

$$Pr(s' = (t-1, \phi_1, \phi_2)|s = (1, U, U), a = \{i, j\}) \\ = 1 - \rho_{pp,\{i,j\}}, \quad t-1 \in \{0, ..., T-2\} \tag{4.22}$$

$$Pr(s' = (T, \phi_1, \phi_2)|s = (1, U, U), a = \{i, j\}) = 1 \tag{4.23}$$

where $\rho_{pp,\{i,j\}}$ are the PU outage probabilities of the channel $PU_{tx} \rightarrow PU_{rx}$ given in (3.5) to (3.7), $\rho_{ps1,\{i,j\}}$ are the PU outage probabilities of the channel $PU_{tx} \rightarrow SU_{rx1}$ and $\rho_{ps2,\{i,j\}}$ are the PU outage probabilities of the channel $PU_{tx} \rightarrow SU_{rx2}$; the last two are given by:

$$\rho_{ps1,\{i,j\}} = Pr\Big(R_p > C\Big(\frac{\gamma_{ps1}}{1 + i \cdot \gamma_{s1s1} + j \cdot \gamma_{s2s1}}\Big)\Big) \tag{4.24}$$

$$\rho_{ps2,\{i,j\}} = Pr\Big(R_p > C\Big(\frac{\gamma_{ps2}}{1 + j \cdot \gamma_{s2s2} + i \cdot \gamma_{s1s2}}\Big)\Big) \tag{4.25}$$

The relationship between $P_{ss'}$ and $P_{ss',a}$ is given by:

$$P_{ss'} = \sum_{a \in \mathcal{A} \times \mathcal{A}} P_{ss',a} \pi_1(a_1|s) \pi_2(a_2|s) \tag{4.26}$$

Fig. 4.1 shows the Markov chain behind our MMDP model.

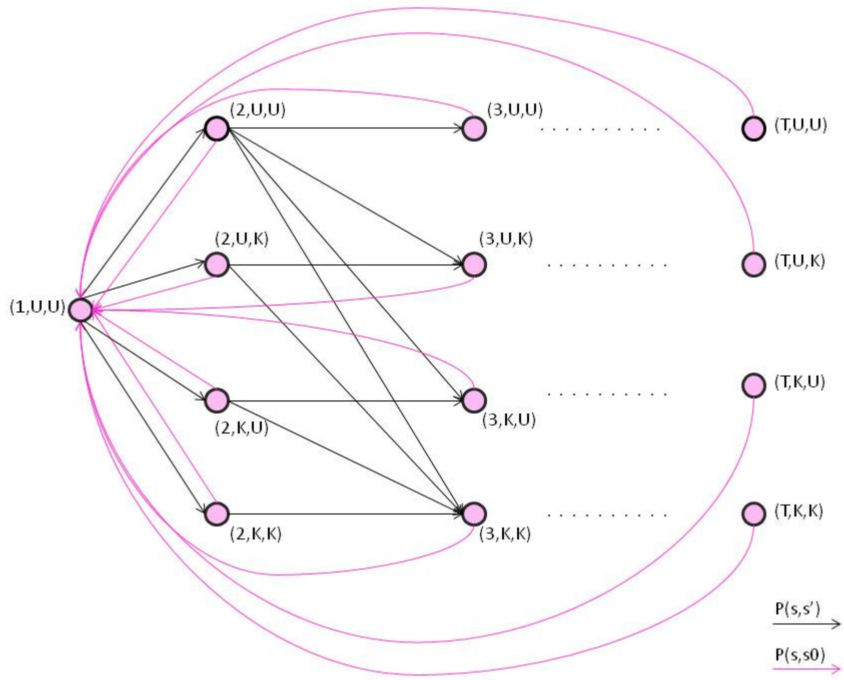


Figure 4.1: Markov Chain behind the model

4.2.1 SU_1 Optimal Access Policy

If we consider SU_1's perspective, under the assumptions that it knows the system state s and SU_2's policy, in each time slots it has two possible actions to select, $a_1 \in \mathcal{A}$. Our aim is to maximize the average long term SUs sum throughput under the average long term PU throughput degradation constraint.

If accessibility action a_1 is selected by SU_1, the expected SUs sum throughput in state s, denoted by $R_1(s, a_1)$, can be computed as:

$$R_1(s, a_1) = \sum_{a_2 \in \mathcal{A}} R_1(s, a_1, a_2)\pi_2(a_2|s) \tag{4.27}$$

where

$$R_1(s, a_1, a_2) = \bar{T}_{s1,\{a_1,a_2\},\phi(s)} + \bar{T}_{s2,\{a_1,a_2\},\phi(s)} \tag{4.28}$$

Specifically, we have:

$$R_1(s = (t, U, U), a_1 = 0, a_2 = 0) = 0$$

$$R_1(s = (t, U, U), a_1 = 0, a_2 = 1) = R_{s2,1,U}(1 - \rho_{s2,1,U})$$

$$R_1(s = (t, U, U), a_1 = 1, a_2 = 0) = R_{s1,2,U}(1 - \rho_{s1,2,U})$$

$$R_1(s = (t, U, U), a_1 = 1, a_2 = 1) = R_{s1,3,U}(1 - \rho_{s1,3,U}) + R_{s2,3,U}(1 - \rho_{s2,3,U})$$

$$R_1(s = (t, U, K), a_1 = 0, a_2 = 0) = 0$$

$$R_1(s = (t, U, K), a_1 = 0, a_2 = 1) = R_{s2,1,K}(1 - \rho_{s2,1,K})$$

$$R_1(s = (t, U, K), a_1 = 1, a_2 = 0) = R_{s1,2,U}(1 - \rho_{s1,2,U})$$

$$R_1(s = (t, U, K), a_1 = 1, a_2 = 1) = R_{s1,3,U}(1 - \rho_{s1,3,U}) + R_{s2,3,K}(1 - \rho_{s2,3,K})$$

$$R_1(s = (t, K, U), a_1 = 0, a_2 = 0) = 0$$

$$R_1(s = (t, K, U), a_1 = 0, a_2 = 1) = R_{s2,1,U}(1 - \rho_{s2,1,U})$$

$$R_1(s = (t, K, U), a_1 = 1, a_2 = 0) = R_{s1,2,K}(1 - \rho_{s1,2,K})$$

$$R_1(s = (t, K, U), a_1 = 1, a_2 = 1) = R_{s1,3,K}(1 - \rho_{s1,3,K}) + R_{s2,3,U}(1 - \rho_{s2,3,U})$$

$$R_1(s = (t, K, K), a_1 = 0, a_2 = 0) = 0$$

$$R_1(s = (t, K, K), a_1 = 0, a_2 = 1) = R_{s2,1,K}(1 - \rho_{s2,1,K})$$

$$R_1(s = (t, K, K), a_1 = 1, a_2 = 0) = R_{s1,2,K}(1 - \rho_{s1,2,K})$$

$$R_1(s = (t, K, K), a_1 = 1, a_2 = 1) = R_{s1,3,K}(1 - \rho_{s1,3,K}) + R_{s2,3,K}(1 - \rho_{s2,3,K})$$

where R_{s1,l,ϕ_1}, R_{s2,l,ϕ_2} are the SUs transmission rates given in (3.1) to (3.4), and ρ_{s1,l,ϕ_1}, ρ_{s2,l,ϕ_2} are the SUs outage probabilities given in (3.8) to (3.15).

Similarly, if accessibility action a_1 is selected by SU_1, the expected PU throughput degradation in state s, denoted by $C_1(s, a_1)$, can be computed as:

$$C_1(s, a_1) = \sum_{a_2 \in \mathcal{A}} C_1(s, a_1, a_2)\pi_2(a_2|s) \tag{4.29}$$

where

$$C_1(s, a_1, a_2) = \rho_{p,\{a_1,a_2\}} - \rho_{p,0} \tag{4.30}$$

and $\rho_{p,l}$ are the PU outage probabilities given in (3.5) to (3.7).

The optimization problem we want to solve can be formalized as follows:

$$\begin{gathered} \text{maximize}_{\pi_1(a_1|s)} E_{a_1,s}[R_1(s, a_1)] \\ \text{s.t. } E_{a_1,s}[C_1(s, a_1)] \leq (1 - \rho_{p,0})\epsilon_{PU} = \epsilon_\omega \end{gathered} \tag{4.31}$$

where $\pi_1(a_1|s)$ is the probability that accessibility action a_1 is selected by SU_1 in state s.

As in the centralized case, for high SU access rate regime, we can employ the equivalent Linear Programming (LP) formulation corresponding to CMDP. To provide the equivalent LP, we need the transition probability matrix of the Markov process denoted by P, where $P_{ss',a_1} = Pr(s'|s, a_1)$ is the probability of moving from state s to s' if accessibility action a_1 is chosen by SU_1. The transition probability from state $s = (t, \phi_1, \phi_2)$ to $s' = (t + 1, \phi_1', \phi_2')$ when SU_1 selects action a_1 is given by:

$$Pr(s'|s, a_1) = \sum_{a_2 \in \mathcal{A}} Pr(s'|s, a_1, a_2)\pi_2(a_2|s) \tag{4.32}$$

Since for every unichain Constrained Markov Decision Process there exists an equivalent Linear Programming (LP) formulation, where a MDP is considered unichain if it contains a single recurrent class plus a (perhaps empty) set of transient states, the equivalent LP problem of problem (4.6) is the following:

$$\max_x \sum_{s \in \mathcal{S}} \sum_{a_1 \in \mathcal{A}} R_1(s, a_1)x(s, a_1) \tag{4.33}$$

$$\text{s.t.} \quad \sum_{s \in \mathcal{S}} \sum_{a_1 \in \mathcal{A}} C_1(s, a_1)x(s, a_1) \leq \epsilon_\omega \tag{4.34}$$

$$\sum_{a_1 \in \mathcal{A}} x(s', a_1) - \sum_{s \in \mathcal{S}} \sum_{a_1 \in \mathcal{A}} P(s'|s, a_1)x(s, a_1) = 0 \quad \forall s' \in \mathcal{S} \tag{4.35}$$

$$\sum_{s \in \mathcal{S}} \sum_{a_1 \in \mathcal{A}} x(s, a_1) = 1 \tag{4.36}$$

$$x(s, a_1) \geq 0 \quad \forall s \in \mathcal{S}, \quad a_1 \in \mathcal{A} \tag{4.37}$$

(4.38)

The relationship between the optimal solution of LP problem (4.13) and the optimal solution to our problem is obtained as follows:

$$\pi_1(a_1|s) = \begin{cases} \frac{x(s,a_1)}{\sum_{a_1' \in \mathcal{A}} x(s,a_1')}, & \text{if } \sum_{a_1' \in \mathcal{A}} x(s,a_1') > 0 \\ \text{arbitrary}, & \text{otherwise} \end{cases} \tag{4.39}$$

The symmetric case, i.e., SU_2's optimal access policy search, can be dealt with in the same way simply by inverting the role of the SUs and exchanging the indexes.

4.2.2 Numerical results

Now we present some numerical results. The channels considered are Rayleigh fading, like in the centralized case, thus, the SNRs γ_x, $x \in \{pp, ps1, ps2, s1s1, s1p, s1s2, s2s2, s2p, s2s1\}$, are exponentially distributed random variables with mean $\bar{\gamma}_x$. For mathematical convenience the links from SU_{tx1} to SU_{rx2} and viceversa are assumed interference free. We consider the same parameters as in the centralized case, i.e., the average SNRs are $\bar{\gamma}_{pp} = 10$, $\bar{\gamma}_{si} = 5$, $\bar{\gamma}_{psi} = 5$, $\bar{\gamma}_{sip} = 2$, $i \in \{1,2\}$, and the ARQ deadline is $T = 5$. The PU rate R_p is selected such that PU throughput is maximized when both SUs are idle, i.e. $R_p = \text{argmax}_R T_p^I(R)$. The SU_i rate $R_{si,l,U}$ under PU message unknown for SU_{rxi} is computed as $R_{si,l,U} = \text{argmax}_{R_{si}} T_{si,l,U}(R_{si}, R_p)$, where $i = 1$, $l \in \{2,3\}$ or $i = 2$, $l \in \{1,3\}$ so as to maximize the SUs sum throughput. The SU_i rate $R_{si,l,K}$ under PU message known for the SU_{rxi} is computed as $R_{si,l,K} = \text{argmax}_R T_{si,l,K}(R)$, where $i = 1$, $l \in \{2,3\}$ or $i = 2$, $l \in \{1,3\}$. The PU throughput constraint is set to $(1-\epsilon_{PU})T_p^I$, where $\epsilon_{PU} \in \{0.01, 0.05, 0.08, 0.1, 0.13, 0.15, 0.18, 0.2, 0.25, 0.3, 0.4, 0.4861, 0.6, 0.7, 0.8, 1\}$. FIC is exploited at the SU_{rxi}, $i \in \{1,2\}$.

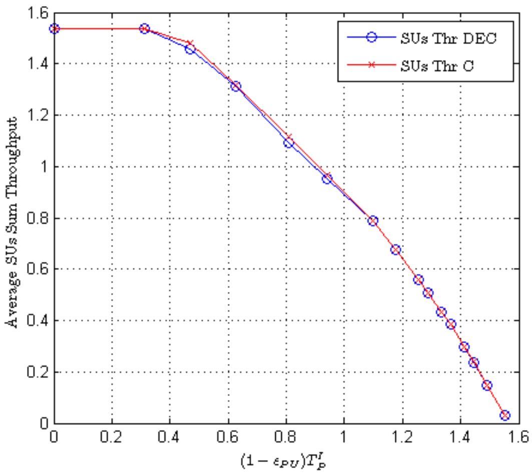


Figure 4.2: DEC-MMDP: Average SUs sum throughput with respect to PU throughput constraint

The SUs sum throughput with respect to the PU throughput constraint by varying the value of ϵ_{PU} is depicted in Fig. 4.2 in contrast with the SUs sum throughput in the centralized case (MMDP) that can be considered as an upper bound for the network performance in the decentralized case. Obviously, as the PU throughput increases, the average SUs sum throughput decreases. Furthermore, we can note that the performance in the decentralized case (DEC-MMDP) is almost the same as in the centralized ones.

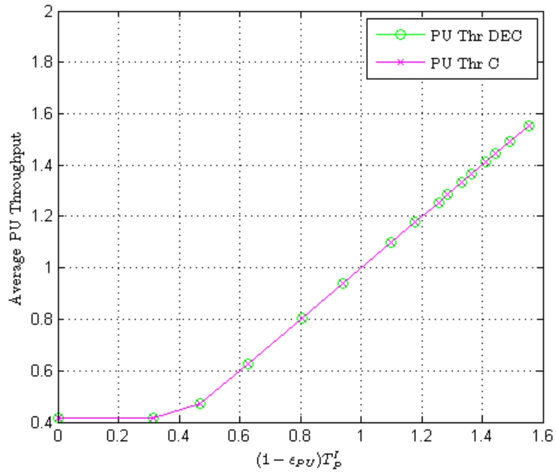
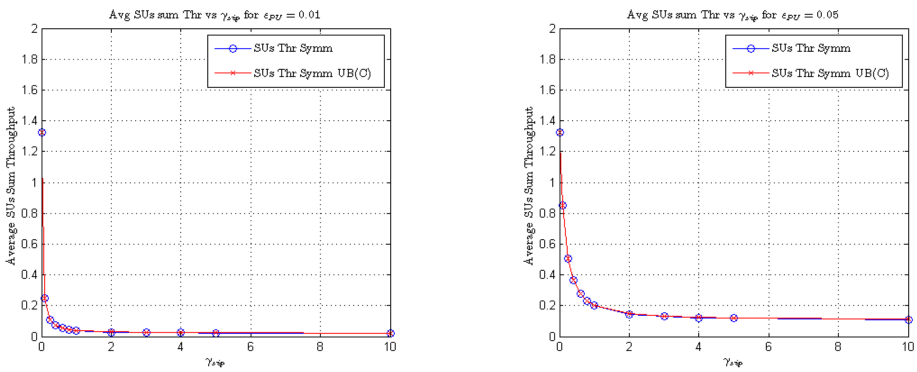
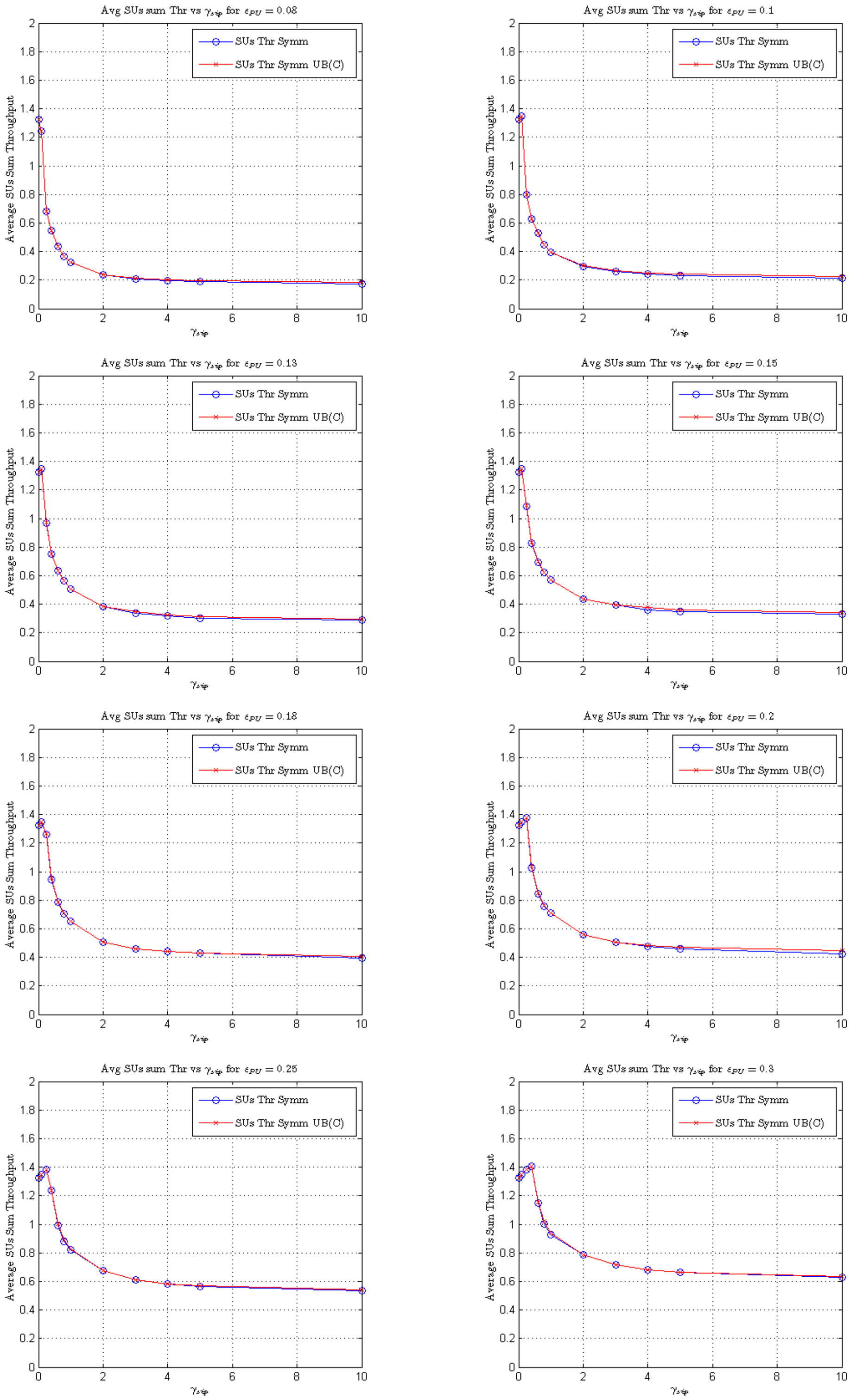


Figure 4.3: DEC-MMDP: Average PU throughput with respect to PU throughput constraint

Fig. 4.3 depicts the PU throughput with respect to the PU throughput constraint by varying the value of ϵ_{PU}. Obviously, as ϵ_{PU} decreases the constraint increases and the PU throughput degradation decreases.




Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.08$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.1$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.13$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.15$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.18$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.2$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.25$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.3$
SUs Thr Symm
SUs Thr Symm UB(C)
Average SUs Sum Throughput
γ_{sip}

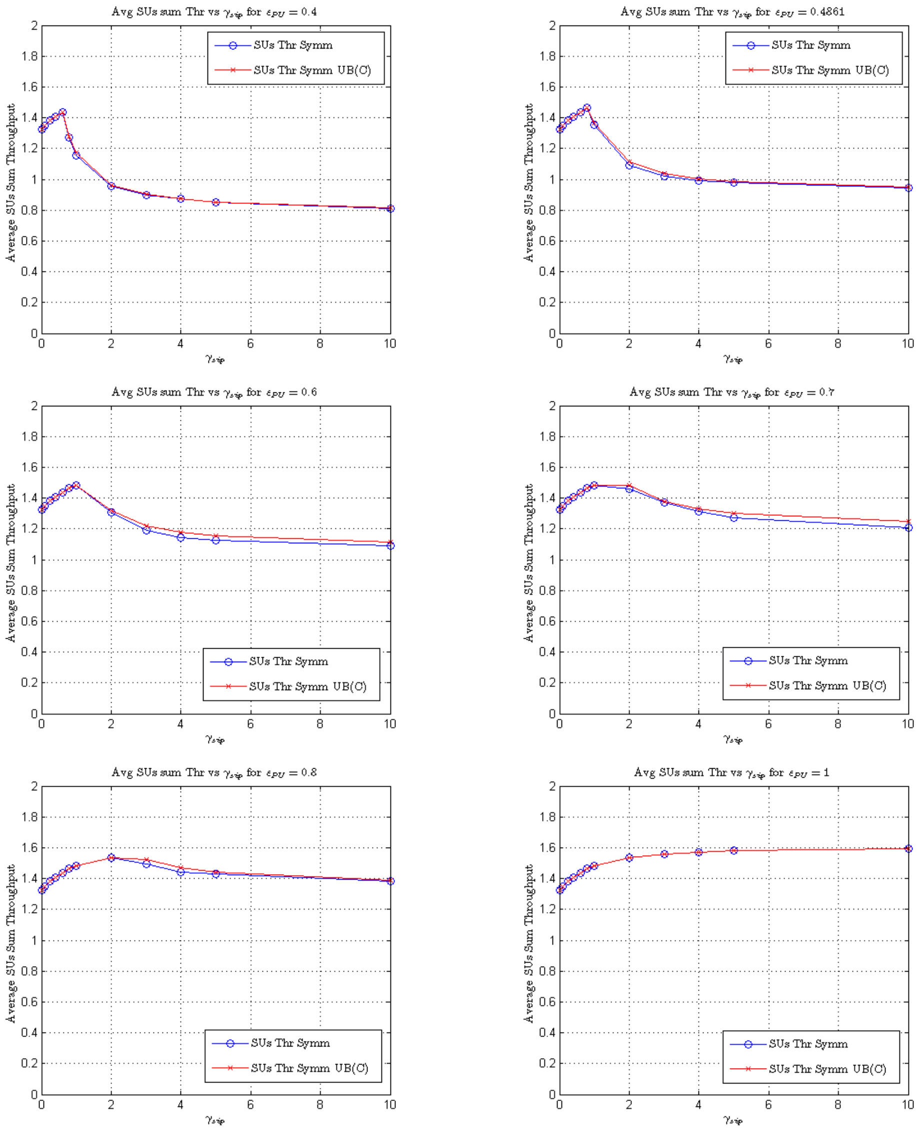


Figure 4.4: DEC-MMDP: Average SUs sum Throughput vs γ_{sip}

So far we have considered only one specific SNR value, $\gamma_{sip} = 2$; it is interesting to analyze the perfomance of the system for varying SNRs, γ_{sip}. Fig. 4.4 depicts the average SUs sum throughput with respect to γ_{sip} for varying ϵ_{PU}; it shows different evolutions based on the PU constraint: for $\epsilon_{PU} \leq 0.08$ the average SUs sum throughput decreases as γ_{sip} increases, this is because as the SNR grows the interference power of the SUs affects more significantly the PU performace, so they have to limit their channel accesses in order to respect the PU constraint with a consequent reduction of the maximum achievable throughput. For $0.1 < \epsilon_{PU} < 0.8$ the average SUs sum throughput increases for γ_{sip} below a certain value and decreases for γ_{sip} over it; this is due to the fact that the PU degradation constraint is not active for small values of the SNR, i.e., for a low interference level at PU_{tx}, thus, the SUs can exploit the transmitting chances much more and utilize their transmitting power to gain a higher reward. On the other hand, when the interference power of the SUs become too high they affect significantly the PU performace, so they have to limit their channel accesses in order to respect the PU constraint with a consequent reduction of the maximum achievable throughput. For $\epsilon_{PU} = 1$ instead the average SUs sum throughput tends to grow as the SNR increases, i.e., the PU degradation constraint is not active for the considered SNRs and the SUs can exploit the transmitting chances as much as possible.

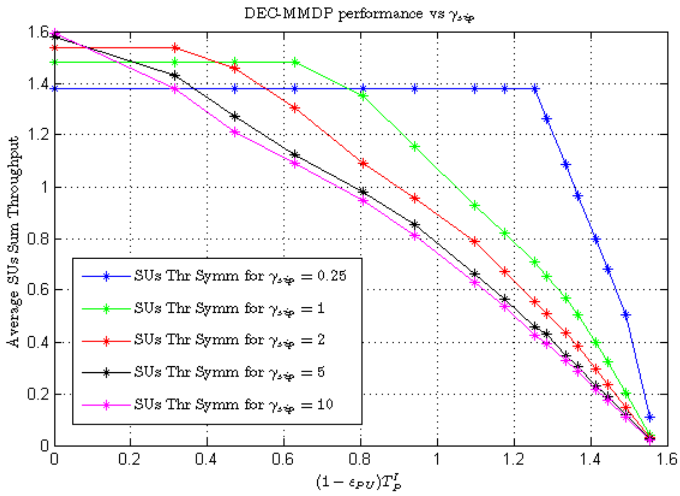


Figure 4.5: DEC-MMDP: Average SUs sum throughput with respect to PU throughput constraint

Fig. 4.5 depicts the average SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} for different values of γ_{sip}; we can note that for small values of ϵ_{PU} as γ_{sip} increases there is a performance degradation since the SUs create a higher level of interference at PU_{rx} (see Fig. 4.6), thus, since the PU constraint is active, they have to limit their own channel accesses and consequently the chances to increase their average reward. On the other hand, as γ_{sip} decreases there is an improvement since the SUs' interference power affects less the PU performance, thus, they can exploit more transmitting chances. Considerations developed so far do not hold for high values of ϵ_{PU}: performance results degrade as the SNR decreases and improve as the SNR increases; this is reasonable since for small values of SUs' SNR the interference level is very low, so the PU retransmits rarely and the SUs have less chances to transmit successfully and increase their own reward, whereas for high values of SUs' SNR the situation is inverted, i.e., the PU retransmits more frequently and so the SUs have more chances to exploit FIC and improve their own throughput. In effect, the curves corresponding to $\gamma_{sip} < 5$ have an initial flat part that corresponds to ϵ_{PU} values for which the PU constraint is not active, i.e., the

SUs behavior does not degrade further the PU performance and so the SUs can exploit the transmitting chances as much as possible with a consequent maximization of their own throughput. Then, in correspondence to the activation of the PU constraint the performace begins to degrade in order to satisfy the transmission limitations. For $\gamma_{sip} \geq 5$ the curves are decreasing, but $\epsilon_{PU} = 1$ is probably a borderline case in which the PU constraint has an unperceivable effect on SUs' reward that consequently is very close to the maximum they can achieve.

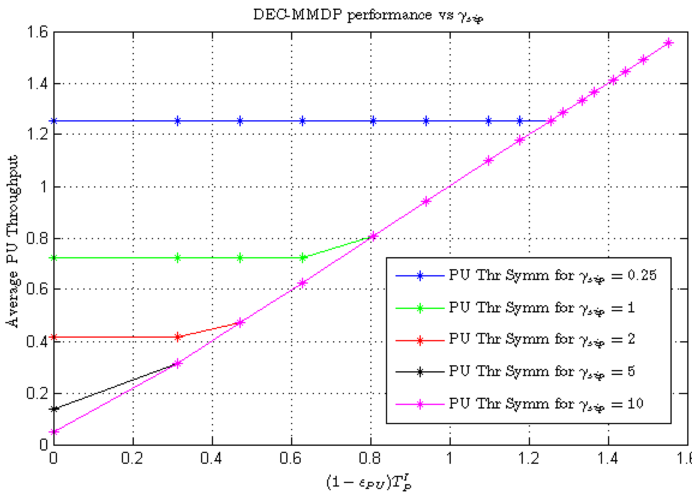
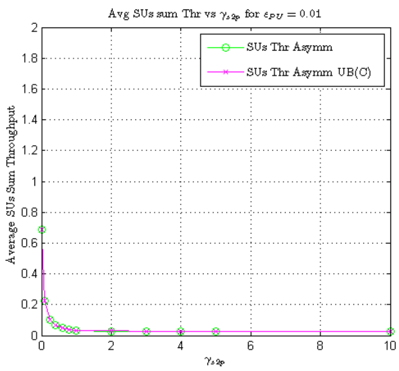
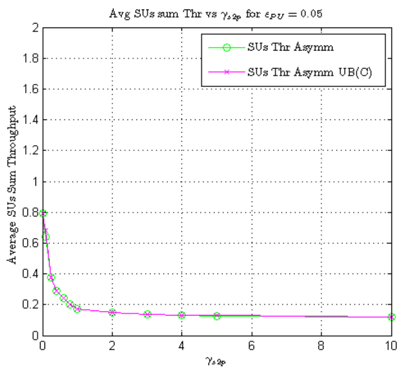
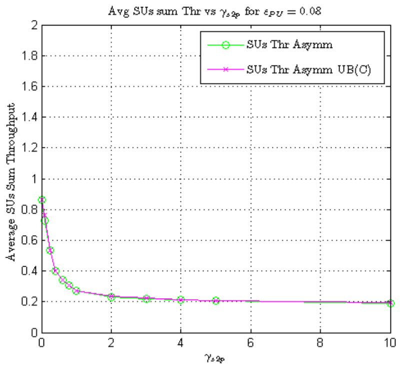
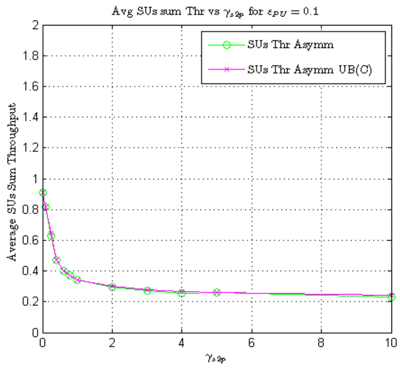


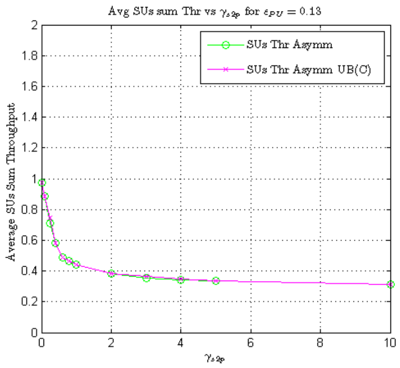
Figure 4.6: DEC-MMDP: Average PU throughput with respect to PU throughput constraint

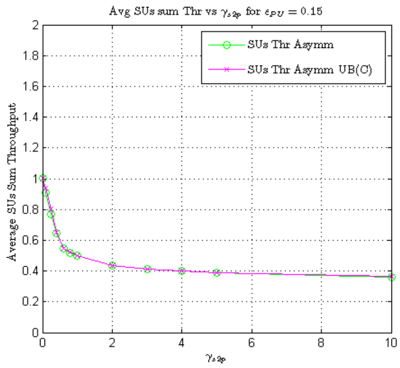


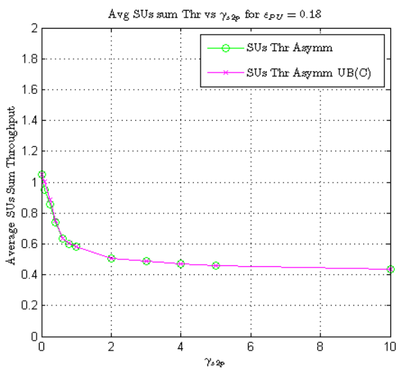


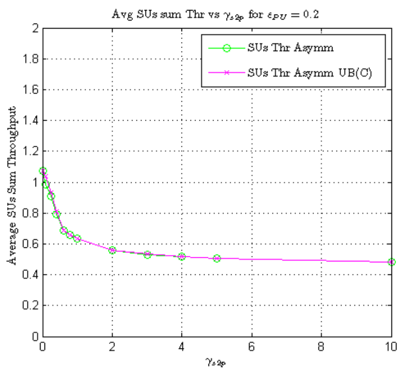

Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.08$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}

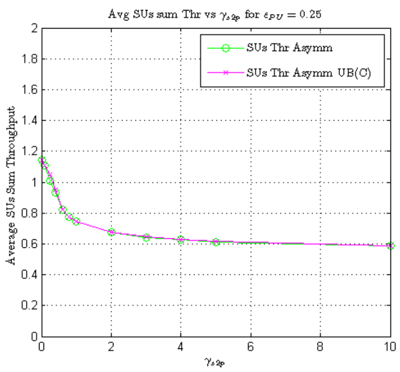

Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.1$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}

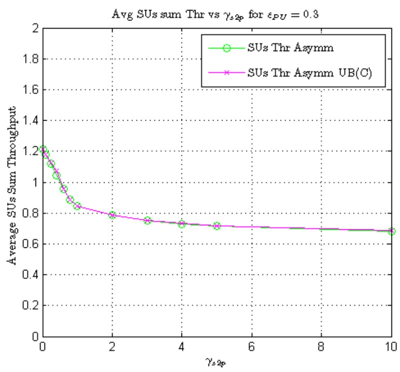

Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.13$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}


Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.15$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}


Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.18$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}


Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.2$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}


Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.25$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}


Avg SUs sum Thr vs γ_{s2p} for $\epsilon_{PU} = 0.3$
SUs Thr Asymm
SUs Thr Asymm UB(C)
Average SUs Sum Throughput
γ_{s2p}

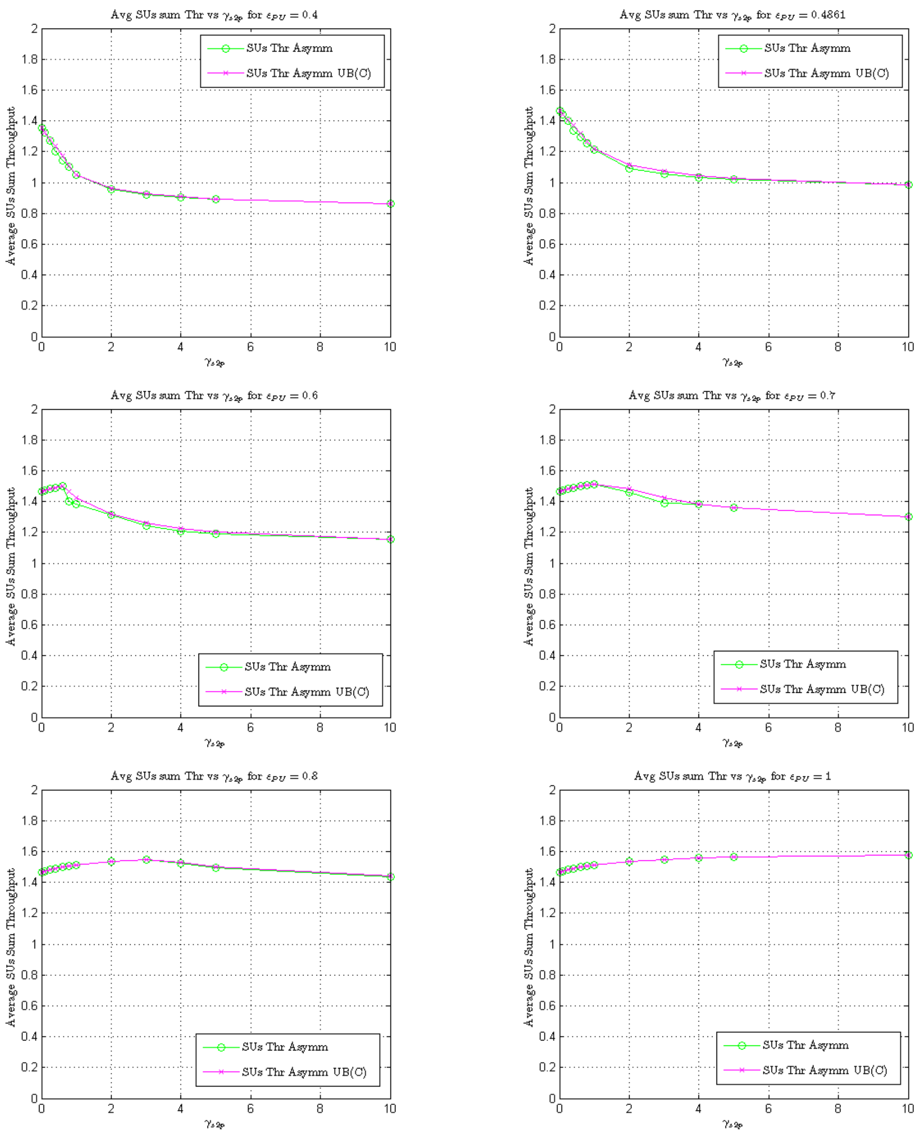


Figure 4.7: DEC-MMDP$_{ASY}$: Average SUs sum Throughput vs γ_{s2p}

Another interesting aspect to analyze is the influence of the asymmetries in the SNR on the system performance, i.e., we suppose that $\gamma_{s1p} = 2$ whereas $\gamma_{s2p} \neq \gamma_{s1p}$ and varying. Fig. 4.7 depicts the average SUs sum throughput with respect to γ_{s2p} for varying ϵ_{PU}; as in the symmetric case we can note three different evolutions, obviously the ϵ_{PU} bounds change, in particular the region in which the average SUs sum throughput decreases as γ_{s2p} increases is larger (until $\epsilon_{PU} = 0.4861$); this is reasonable since SU_1 creates a constant interference level at PU_{rx} with a significant effect on the PU performance especially when the constraint is very tight, i.e., for small ϵ_{PU}, whereas in the absence of asymmetries γ_{s1p} grows in parallel to γ_{s2p} and the effect of SUs interference power is weaker at

PU_{rx}. Consequently the region in which the SUs first exploit as much as possible the opportunities to increase their reward and then limit their accesses to the channel in order to satisfy the constraint is narrower, but for $\epsilon_{PU} = 1$ the PU degradation constraint is not active, thus the SUs can exploit the transmitting chances as much as possible and the average SUs sum throughput tends to grow as the SNR increases.

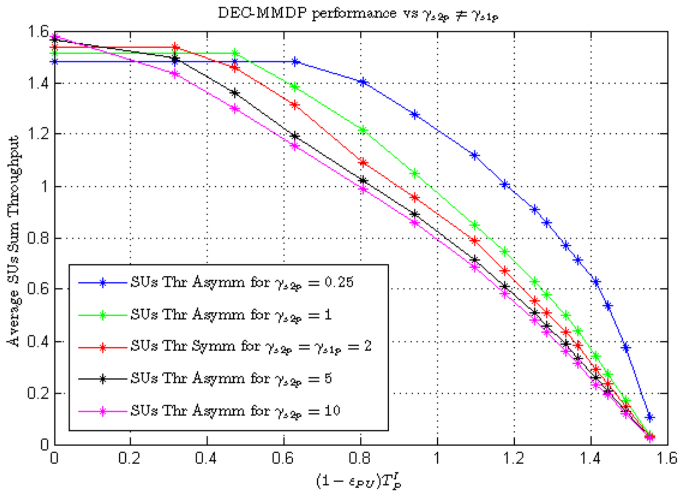


Figure 4.8: DEC-MMDP$_{ASY}$: Average SUs sum throughput with respect to PU throughput constraint

Fig. 4.8 depicts the average SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} in different asymmetric situations, i.e., for different values of γ_{s2p}. We can note that for $\gamma_{s2p} > \gamma_{s1p}$ there is a performance degradation since SU_2 creates a higher level of interference at PU_{rx} than in the symmetric case (see Fig. 4.9), thus it has to limit its own channel accesses and consequently the chances to increase the average reward. On the other hand, for $\gamma_{s2p} < \gamma_{s1p}$ there is an improvement since SU_2 interference power affects less the PU performance than in the symmetric case, so can mainly exploit the transmitting chances. Finally, it is important to note that the considerations done so far do not hold for $\epsilon_{PU} = 1$: performance results are worse than the symmetric case for $\gamma_{s2p} < \gamma_{s1p}$, whereas they are better than the symmetric case for $\gamma_{s2p} > \gamma_{s1p}$; this is probably due to the fact that for small values of SU_2's SNR the interference level is very low, so the PU retransmits rarely and the SU_2 has less chances to transmit successfully and increase its own reward, whereas for high values of SU_2's SNR the situation is inverted, i.e., the PU retransmits more frequently and so SU_2 has more chances to exploit FIC and improve its own throughput.

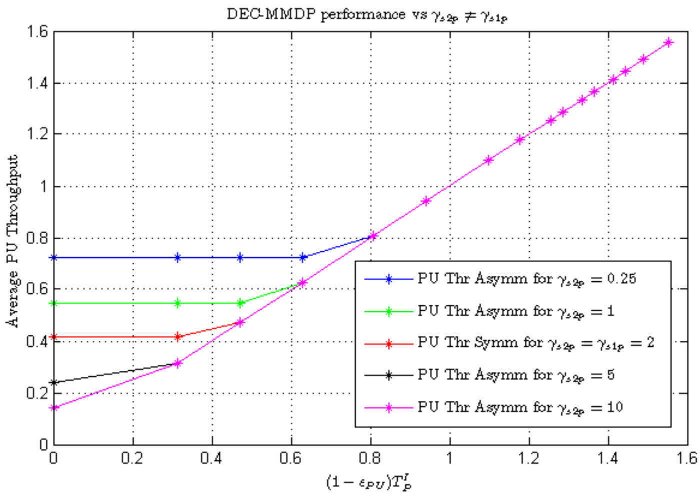
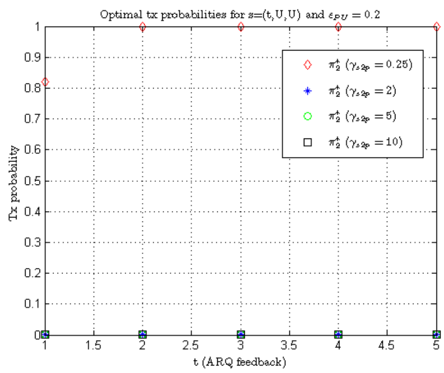
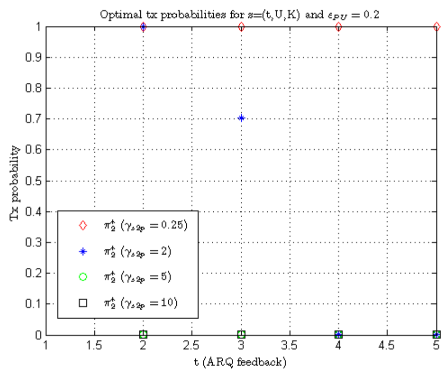


Figure 4.9: DEC-MMDP$_{ASY}$: Average PU throughput with respect to PU throughput constraint

Fig. 4.9 depicts the average PU throughput with respect to the PU throughput constraint by varying the value of ϵ_{PU} in the same asymmetric situations illustrated in Fig 4.5. We can note that for $\epsilon_{PU} \geq 0.4861$ the asymmetries in the SNR influence the PU throughput degradation, in particular, for $\gamma_{s2p} > \gamma_{s1p}$ SU_2's interference power at PU_{rx} causes a degradation of PU performance with respect to the symmetric case, whereas for $\gamma_{s2p} < \gamma_{s1p}$ there is an improvement in PU performance due to the lower interference level SU_2 creates at PU_{rx}.

Fig. 4.10 and 4.11 depict SU_2's optimal transmission probabilities, denoted by π_2^*, for $\epsilon_{PU} = 0.2$ and $\epsilon_{PU} = 1$, respectively, in different system states.





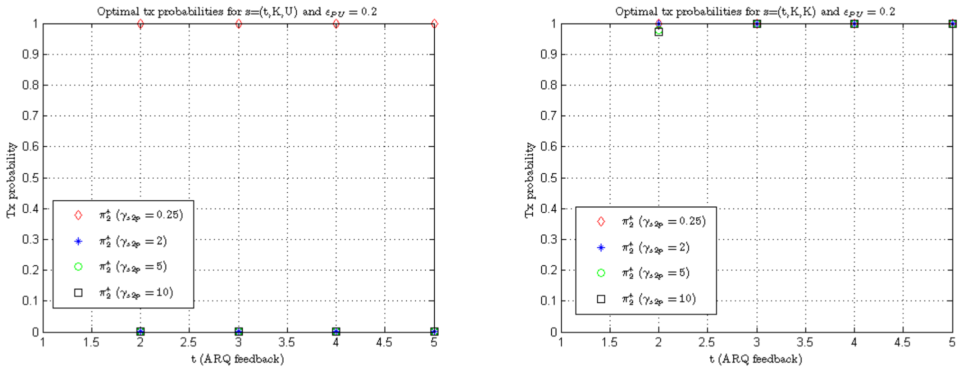


Figure 4.10: DEC-MMDP$_{ASY}$: Tx probabilities vs γ_{s2p} for $\epsilon_{PU} = 0.2$

In Fig. 4.10 we can note that SU_2 tends to transmit more than in the symmetric case for $\gamma_{s2p} = 0.25$, i.e., when its interference power is almost unperceivable at PU_{rx}, instead for $\gamma_{s2p} > 2$ it mantains the same transmitting behavior as in the symmetric case because, even if the interference it creates at PU_{rx} is higher, the PU performance is not affected significantly, as shown in Fig. 4.9.

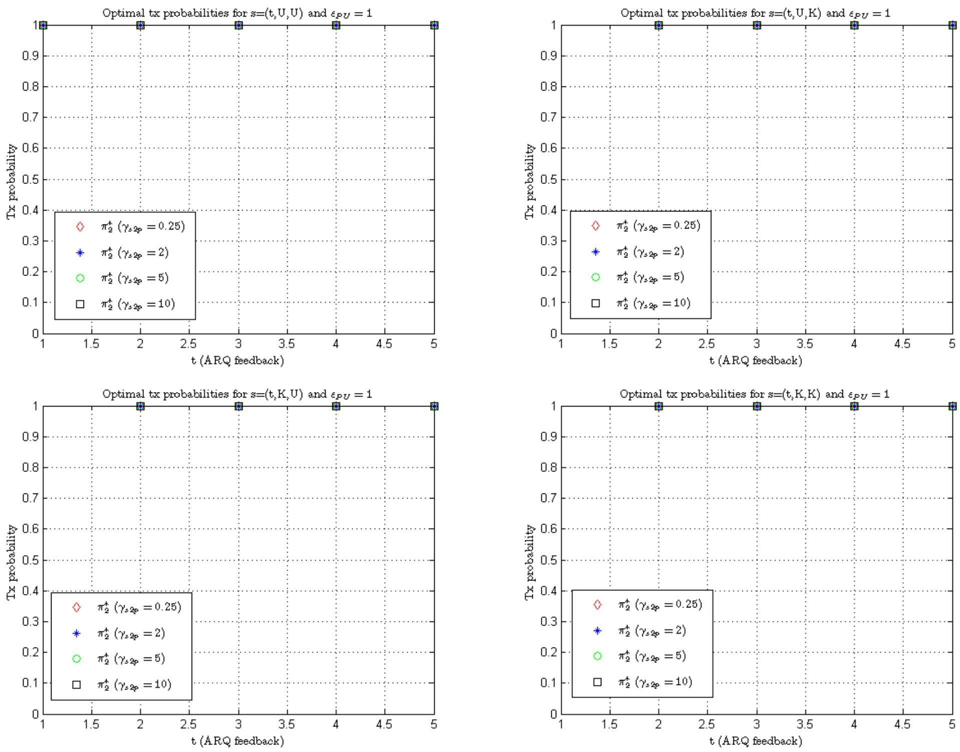


Figure 4.11: DEC-MMDP$_{ASY}$: Tx probabilities vs γ_{s2p} for $\epsilon_{PU} = 1$

In Fig. 4.11 it is evident that SU_2 mantains the same transmitting behavior as in the symmetric case irrespective of its SNR changes; we suppose this is related to the relaxation of the PU constraint which induces SU_2 to access the channel as much as possible.

Chapter 5

Heuristic Decentralized Access Policies in a Cognitive Radio Network with Two Independent SUs

In the previous chapter we found the optimal access policies for the decentralized case of two independent SUs in a CR network and we obtained that their performance was almost the same as that of the optimal centralized policy described in Chapter 3. Now we are interested in developing some heuristic policies for the same decentralized scenario and in analyzing their performance. In order to achieve our purpose, we developed a network simulator which reproduces the behavior of a decentralized cognitive radio scenario with one PU and two independent SUs.

In simulation our work consists in monitoring the evolution of the system we model in order to evaluate the average long term SUs sum throughput and the corresponding average long term PU throughput degradation due to the SUs interference. In the system we consider, there exist one primary and two secondary transmitters denoted by PU_{tx}, SU_{tx1} and SU_{tx2}, respectively. These transmitters transmit their messages with constant power over block fading channels and, in each time slot, the channels are considered to be constant. The signal to noise ratios (SNR) of the channels $PU_{tx} \to PU_{rx}$, $PU_{tx} \to SU_{rx1}$, $PU_{tx} \to SU_{rx2}$, $SU_{tx1} \to PU_{rx}$, $SU_{tx1} \to SU_{rx1}$, $SU_{tx1} \to SU_{rx2}$, $SU_{tx2} \to PU_{rx}$, $SU_{tx2} \to SU_{rx1}$, $SU_{tx2} \to SU_{rx2}$ are denoted by γ_{pp}, γ_{ps1}, γ_{ps2}, γ_{s1p}, γ_{s1s1}, γ_{s1s2}, γ_{s2p}, γ_{s2s1} and γ_{s2s2}, respectively. We assume that no channel State Information (CSI) is available at the transmitters. Thus, transmissions are under outage, when the selected rates are greater than the current channel capacity. The system model, trasmission rates and outage probabilities are described in Section 4.1.

The harder step in building a CR network simulator consists in evaluating the decoding event at PU_{rx}, SU_{rx1} and SU_{rx2}. Thus, it is necessary to identify some outage conditions which help to characterize whether or not decoding is successful We denote by α_{UiUj} the fading coefficient on the channel $Ui_{tx} \rightarrow Uj_{rx}$.

We first consider the PU message correct decoding at the PU_{rx}. The conditions to have a successful transmission of the PU message in SU accessibility actions 0, 1, 2 and 3 are respectively:

$$R_p \leq C(\gamma_{pp}\alpha_{pp}) = \log_2(1 + \gamma_{pp}\alpha_{pp}) \quad \text{if } l = 0 \tag{5.1}$$

$$R_p \leq C\Big(\frac{\gamma_{pp}\alpha_{pp}}{1 + \gamma_{sip}\alpha_{sip}}\Big) = \log_2\Big(1 + \frac{\gamma_{pp}\alpha_{pp}}{1 + \gamma_{sip}\alpha_{sip}}\Big) \quad \text{if } l = 1, 2, \quad i \in \{1, 2\} \tag{5.2}$$

$$R_p \leq C\Big(\frac{\gamma_{pp}\alpha_{pp}}{1 + \gamma_{s1p}\alpha_{s1p} + \gamma_{s2p}\alpha_{s2p}}\Big) = \log_2\Big(1 + \frac{\gamma_{pp}\alpha_{pp}}{1 + \gamma_{s1p}\alpha_{s1p} + \gamma_{s2p}\alpha_{s2p}}\Big) \quad \text{if } l = 3 \tag{5.3}$$

From the simulation point of view it is interesting to consider the SNRs which influence the evolution of the system in each time-slot and the messages of interest for the various receivers. PU_{rx} is interested in decoding only its own message and it is oblivious of the presence of the SUs in the network; thus, their messages are considered as background noise when they transmit. The conditions to check to establish the possible success of the PU message transmission are given in (4.52) to (4.54).

SU_{rxi} is interested in decoding its own message, but also the PU message, if it is unknown, in order to perform FIC. At PU knowledge state $\{K, K\}$ or $\{K, U\}$ ($\{U, K\}$), the PU message is known at SU_{rx1} (SU_{rx2}) and therefore the PU message may be canceled at this receiver. The condition to have a successful transmission of the SU_i message, $i \in \{1, 2\}$, in SU accessibility action $l \in \{1, 2\}$, $l \neq i$, i.e. when the other SU is idle, is:

$$R_{si,l,K} \leq C(\gamma_{sisi}\alpha_{sisi}) = \log_2(1 + \gamma_{sisi}\alpha_{sisi}), \quad i \in \{1, 2\} \tag{5.4}$$

In contrast, at PU knowledge state $\{U, U\}$ or $\{U, K\}$ ($\{K, U\}$), where the PU message is not decoded at SU_{rx1} (SU_{rx2}), the outage probability of the channel from SU_{tx1} (SU_{tx2}) to SU_{rx1} (SU_{rx2}) is under the influence of the received PU message. Thus, we have two different SNRs to consider corresponding to the two messages of interest:

$$SNR_P = \frac{\gamma_{psi}\alpha_{psi}}{1 + \gamma_{sisi}\alpha_{sisi}}$$

$$SNR_{Si} = \frac{\gamma_{sisi}\alpha_{sisi}}{1 + \gamma_{psi}\alpha_{psi}}$$

Hence, the conditions to have a successful decoding of the SU_i message when the other SU is idle are:

$$\begin{aligned} R_{si,l,U} \leq & C\Big(\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{psi}\alpha_{psi}}\Big) \\ = & \log_2\Big(1+\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{psi}\alpha_{psi}}\Big) \quad \text{if } R_p > C\Big(\gamma_{psi}\alpha_{psi}\Big) \end{aligned} \tag{5.5}$$

$$\begin{aligned} R_{si,l,U} \leq & C(\gamma_{sisi}\alpha_{sisi}) \\ = & \log_2(1+\gamma_{sisi}\alpha_{sisi}) \quad \text{if } R_p \leq C\Big(\gamma_{psi}\alpha_{psi}\Big) \end{aligned} \tag{5.6}$$

Similarly, the conditions to have a successful decoding of the PU message at the SU_{rxi} when the other SU is idle are:

$$\begin{aligned} R_p \leq & C\Big(\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{sisi}\alpha_{sisi}}\Big) \\ = & \log_2\Big(1+\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{sisi}\alpha_{sisi}}\Big) \quad \text{if } R_{si,l,U} > C\Big(\gamma_{sisi}\alpha_{sisi}\Big) \end{aligned} \tag{5.7}$$

$$\begin{aligned} R_p \leq & C(\gamma_{psi}\alpha_{psi}) \\ = & \log_2(1+\gamma_{psi}\alpha_{psi}) \quad \text{if } R_{si,l,U} \leq C\Big(\gamma_{sisi}\alpha_{sisi}\Big) \end{aligned} \tag{5.8}$$

For accessibility action 3 and PU message known by SU_i, SU_{rxi} is interested in decoding its own message like for $l = 1, 2$, the only difference is the presence of the other SU message. In this case, we have only one SNR corresponding to the message of interest:

$$SNR_{Si} = \frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}$$

Hence, the condition to have a successfully decoding of the SU_i message is:

$$R_{si,l,K} \leq C\Big(\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big) = \log_2\Big(1+\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big) \tag{5.9}$$

For accessibility action 3 and PU message unknown by SU_i, SU_{rxi} is interested in decoding its own message, but also the PU message in order to perform FIC, like for $l = 1, 2$, the only difference is the presence of the other SU message. Again we have two different SNRs corresponding to the two messages of interest:

$$SNR_P = \frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{sisi}\alpha_{sisi}+\gamma_{slsi}\alpha_{slsi}}$$

$$SNR_{Si} = \frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{psi}\alpha_{psi}+\gamma_{slsi}\alpha_{slsi}}$$

Hence, the conditions to have a successful decoding of the SU_i message when the other SU is not idle are:

$$\begin{aligned}R_{si,l,U} \leq & C\Big(\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{psi}\alpha_{psi}+\gamma_{slsi}\alpha_{slsi}}\Big)\\ =& \log_2\Big(1+\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{psi}\alpha_{psi}+\gamma_{slsi}\alpha_{slsi}}\Big) \quad \text{if } R_p > C\Big(\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big)\end{aligned} \tag{5.10}$$

$$\begin{aligned}R_{si,l,U} \leq & C\Big(\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big)\\ =& \log_2\Big(1+\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big) \quad \text{if } R_p \leq C\Big(\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big)\end{aligned} \tag{5.11}$$

Similarly, the conditions to have a successful decoding of the PU message at the SU_{rxi} when the other SU is not idle are:

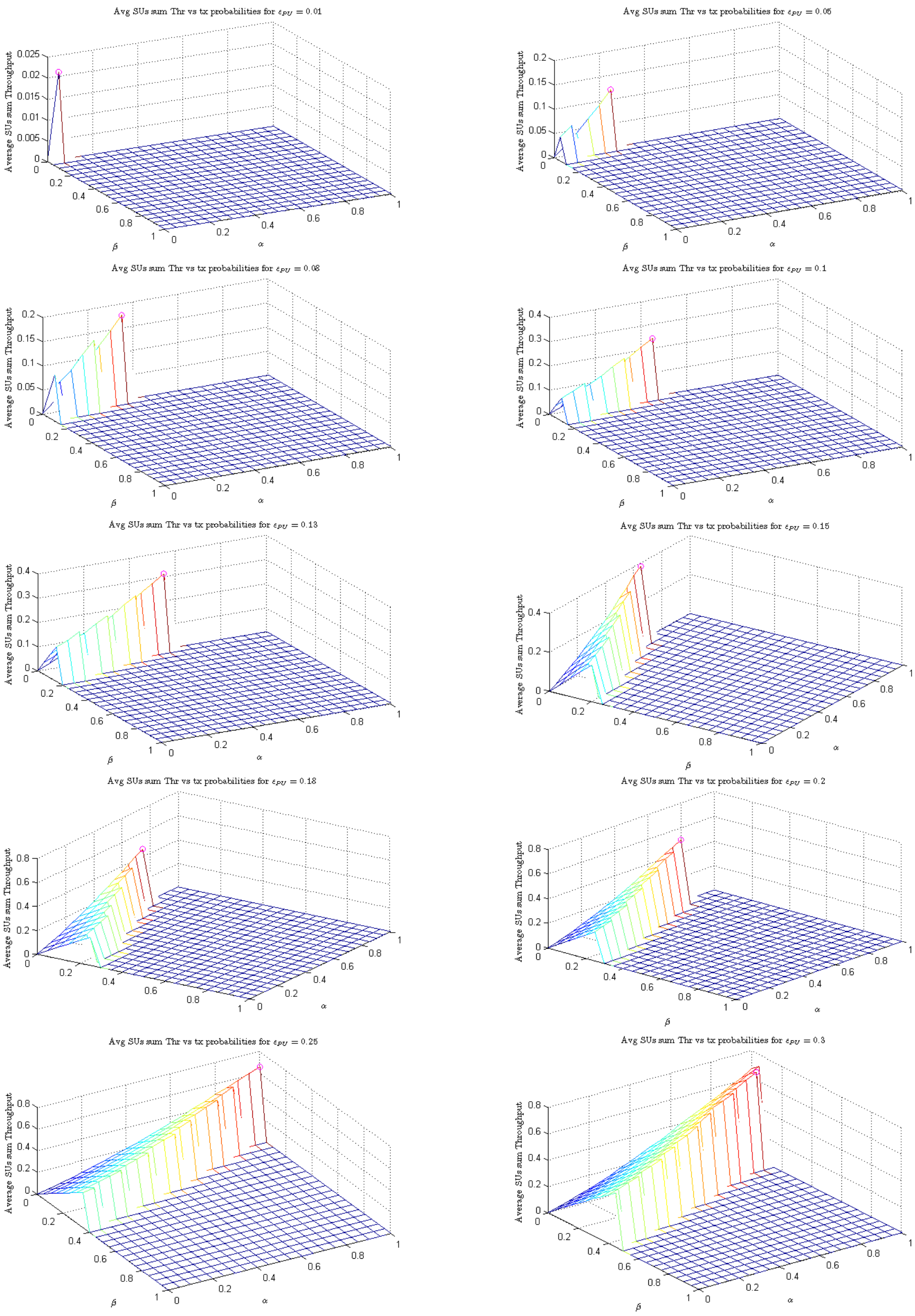
$$\begin{aligned}R_p \leq & C\Big(\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{sisi}\alpha_{sisi}+\gamma_{slsi}\alpha_{slsi}}\Big)\\ =& \log_2\Big(1+\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{sisi}\alpha_{sisi}+\gamma_{slsi}\alpha_{slsi}}\Big) \quad \text{if } R_{si,l,U} > C\Big(\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big)\end{aligned} \tag{5.12}$$

$$\begin{aligned}R_p \leq & C\Big(\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big)\\ =& \log_2\Big(1+\frac{\gamma_{psi}\alpha_{psi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big) \quad \text{if } R_{si,l,U} \leq C\Big(\frac{\gamma_{sisi}\alpha_{sisi}}{1+\gamma_{slsi}\alpha_{slsi}}\Big)\end{aligned} \tag{5.13}$$

5.1 Heuristic Access Policy H_1

The evolution of a CR network is deeply influenced by the state of the system. In the model we develop, the state of the system is represented by the PU message knowledge state, ϕ, and the primary ARQ state, t; since the PU message knowledge state of the single SU can assume only two possible values, $\phi_i \in \{U, K\}$, with $i \in \{1, 2\}$, the first idea we have in designing a heuristic access policy for the two SUs consists in identifying two different transmission probabilities for the two possible PU message knowledge state, denoted by α and β, respectively. We mean that α is the probability that the SU accesses the channel and transmits its own message when it knows the PU message in the current time-slot, whereas β is the probability that the SU accesses the channel and transmits its own message when it does not know the PU message in the current time-slot. Obviously, $\alpha \geq \beta$, since the probability of success for the SU is higher when it knows the PU message and so can perform FIC. We suppose that the two SUs are symmetric, i.e., they have the same transmission parameters, thus, it is reasonable to suppose they adopt the same access policy.
The aim of our policy design is to develop a heuristic access policy which has to maximize the SUs sum throughput as much as possible under the PU throughput degradation constraint. Thus, it is useful to identify which values of α and

β guarantee the higher possible SUs sum throughput. Considering the same average SNRs, transmission rates and PU throughput constraints used to obtain numerical results in the optimal centralized and decentralized case, we examine the performance of the CR network for different values of the probability couple (α, β).



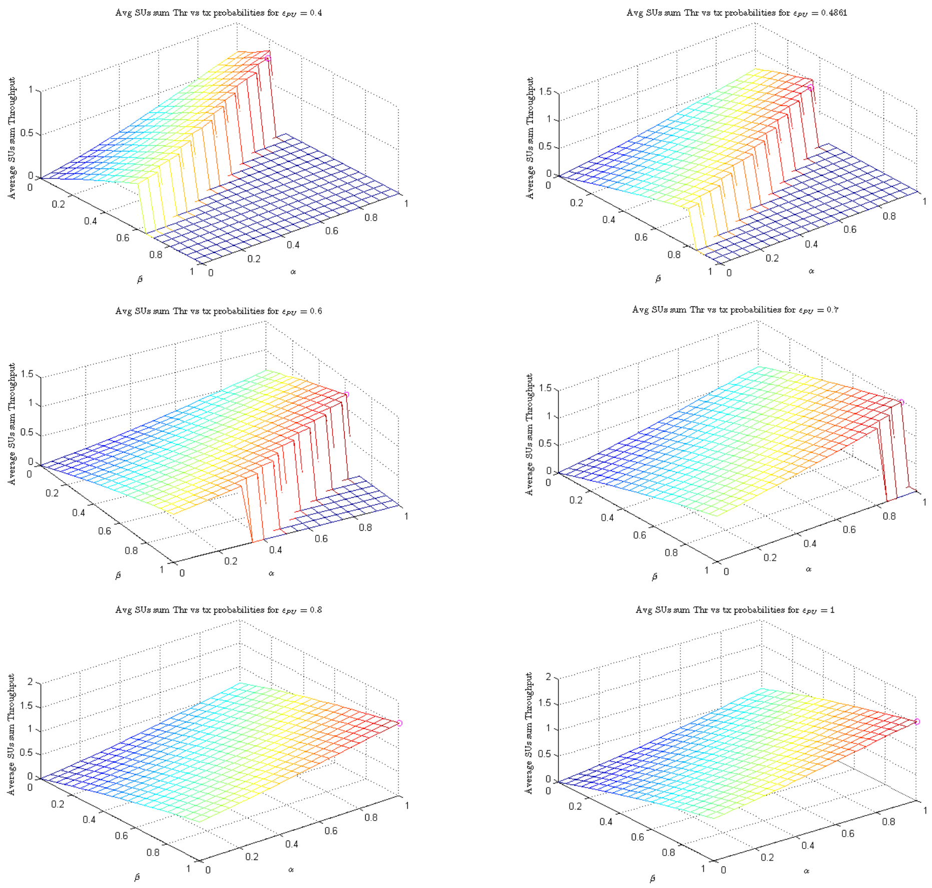


Figure 5.1: H_1: Average SUs sum Throughput vs (α, β)

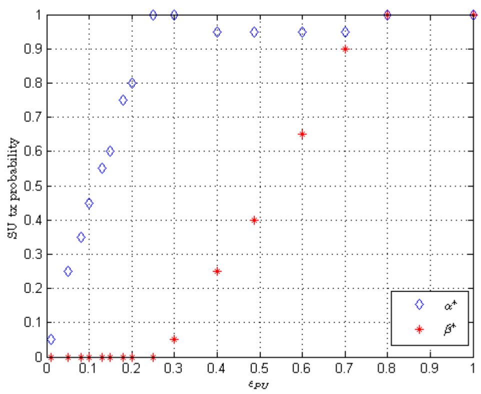


Figure 5.2: H_1: (α^*, β^*) vs ϵ_{PU}

Fig. 5.1 illustrates the $\alpha - \beta$ search for various values of ϵ_{PU}; it shows that the higher ϵ_{PU}, the bigger the tendency to grow of the α and β which grant the maximum SUs sum throughput, denoted by α^* and β^*, i.e., as ϵ_{PU} increases and the PU throughput constraint is relaxed, the two SUs adopt a more aggressive policy to exploit the transmitting chances as much as possible (see Fig. 5.2).

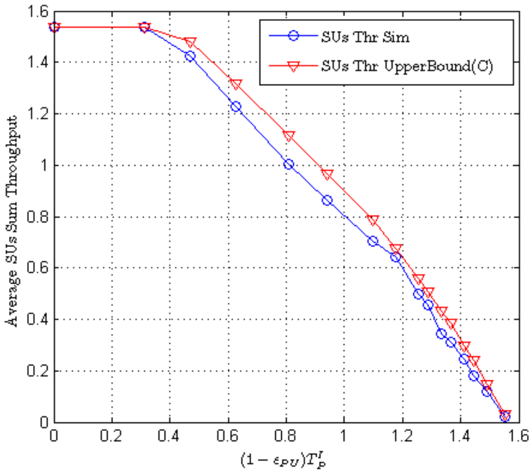


Figure 5.3: H_1: Average SUs sum throughput with respect to PU throughput constraint

The SUs sum throughput with respect to the PU throughput constraint for various values of ϵ_{PU} is depicted in Fig. 5.3 and compared with the upper bound represented by the centralized case (MMDP). Obviously, as the PU throughput increases, the average SUs sum throughput decreases. The numerical results are obtained by selecting for each value of ϵ_{PU} considered the couple (α, β) which guarantees the maximum SUs sum throughput achievable under this heuristic.

Fig. 5.4 depicts the average PU throughput with respect to the PU throughput constraint for various values of ϵ_{PU}. Obviously, as ϵ_{PU} decreases the constraint, $(1 - \epsilon_{PU})\bar{T}_p^I$, increases and the PU throughput degradation decreases. In other words, as ϵ_{PU} increases and the PU thoughput constraint is relaxed, the SUs can mainly exploit the transmitting chances and gain a higher throughput; although, in doing so they create more interference at the PU_{rx} and conseguently cause a degradation of the PU throughput which is more significant as the transmitting chances grow, i.e., as ϵ_{PU} increases. Furthermore, since the MMDP case represents the upper bound for the maximum achievable SUs sum throughput, on the other hand it represents the lower bound for the maximum allowable PU throughput degradation.

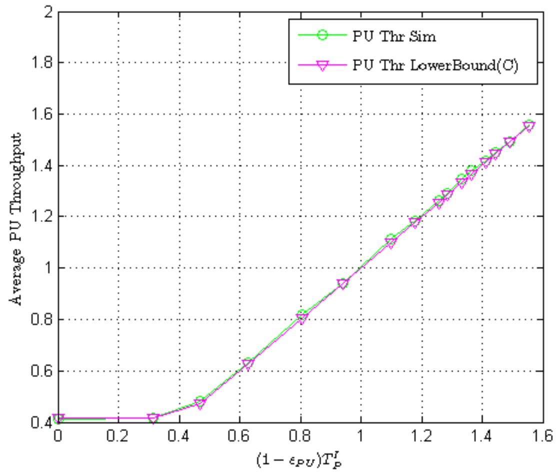
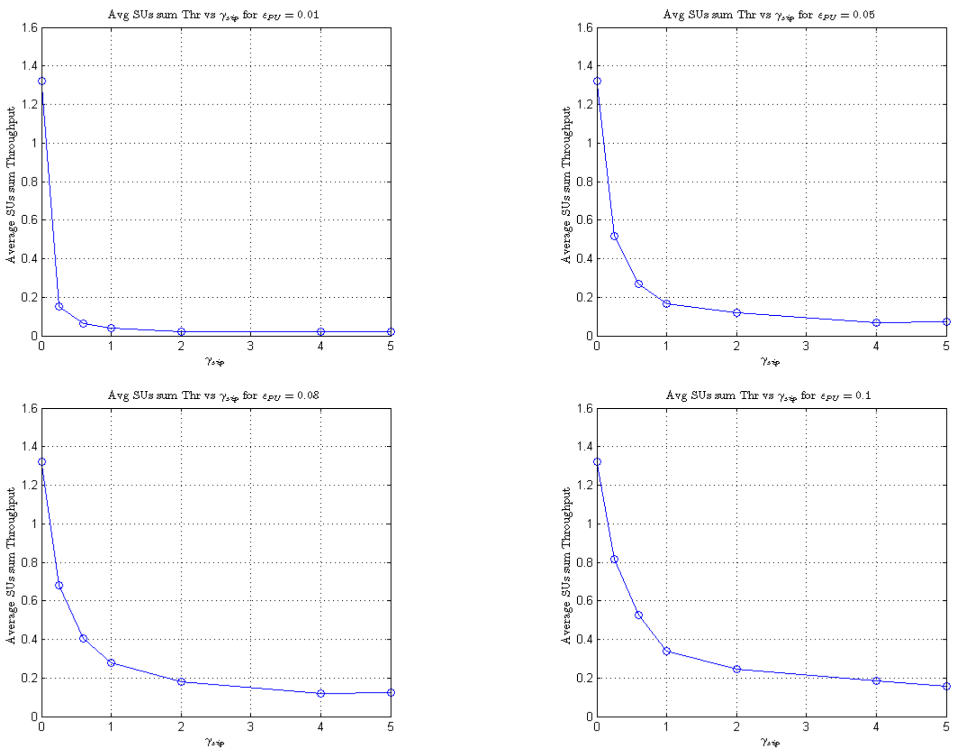
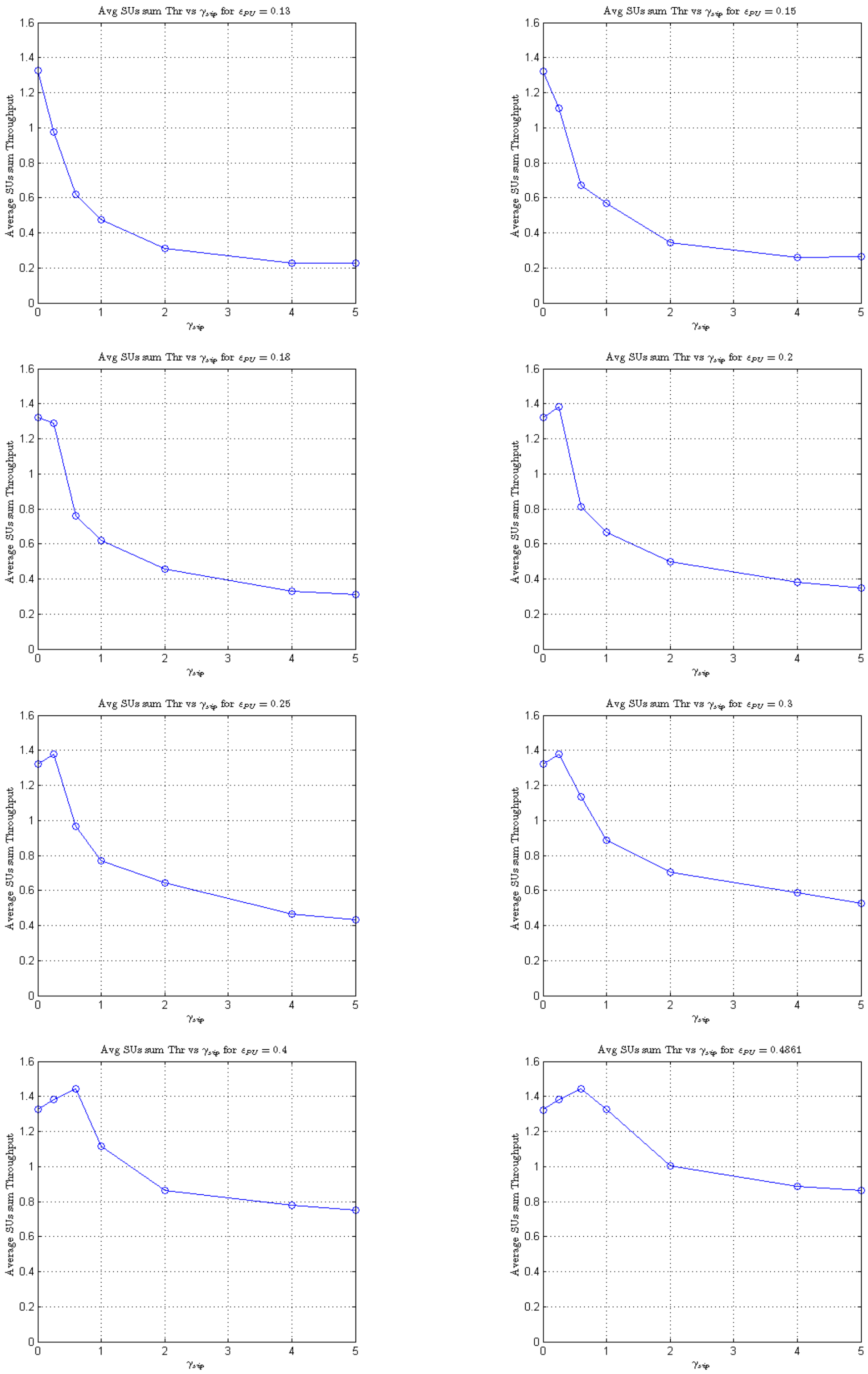


Figure 5.4: H_1: Average PU throughput with respect to PU throughput constraint





Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.13$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.15$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.18$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.2$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.25$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.3$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.4$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.4861$
Average SUs sum Throughput
γ_{sip}

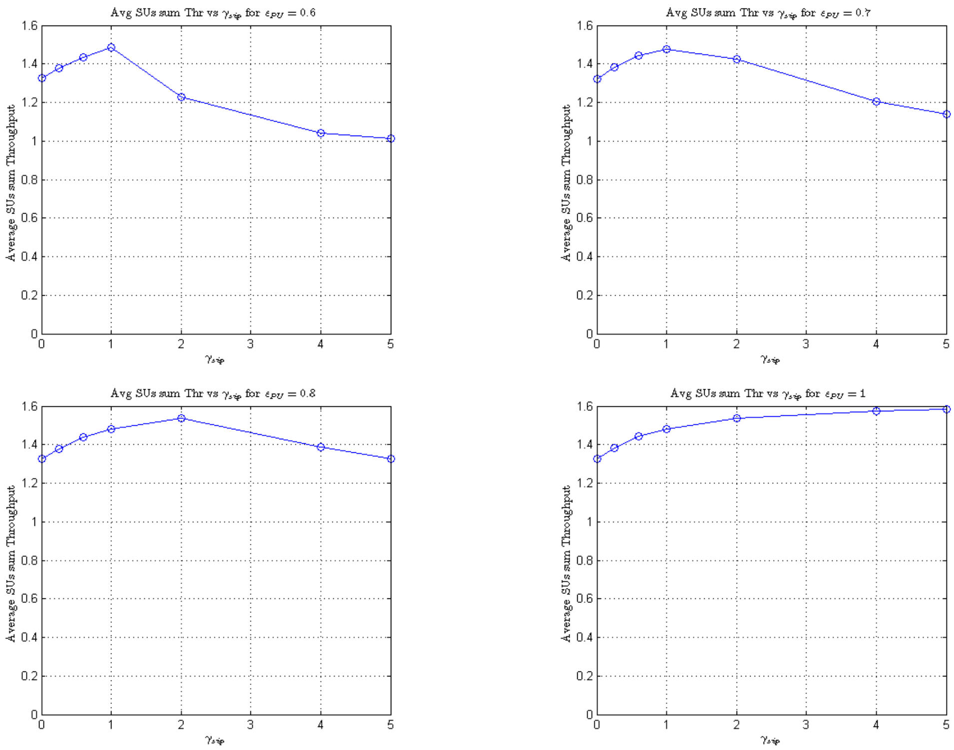
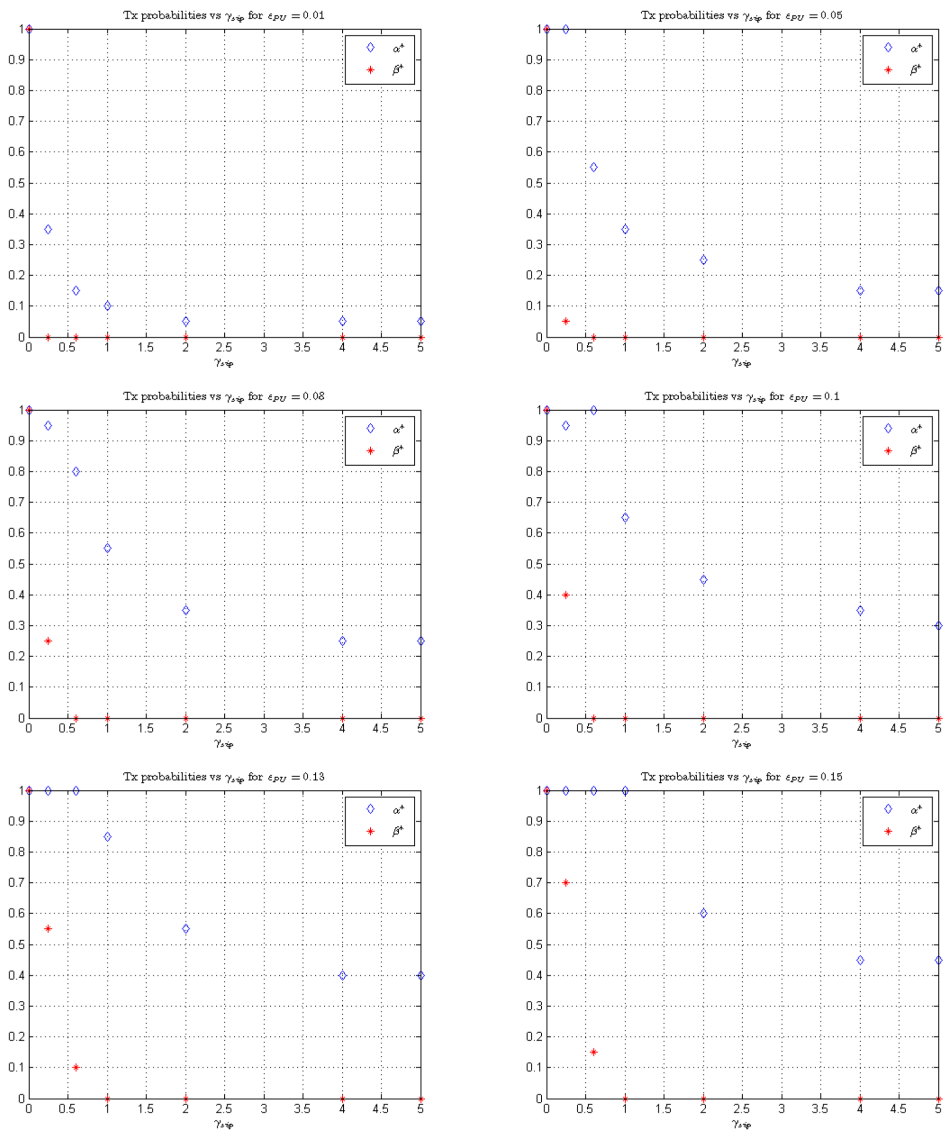


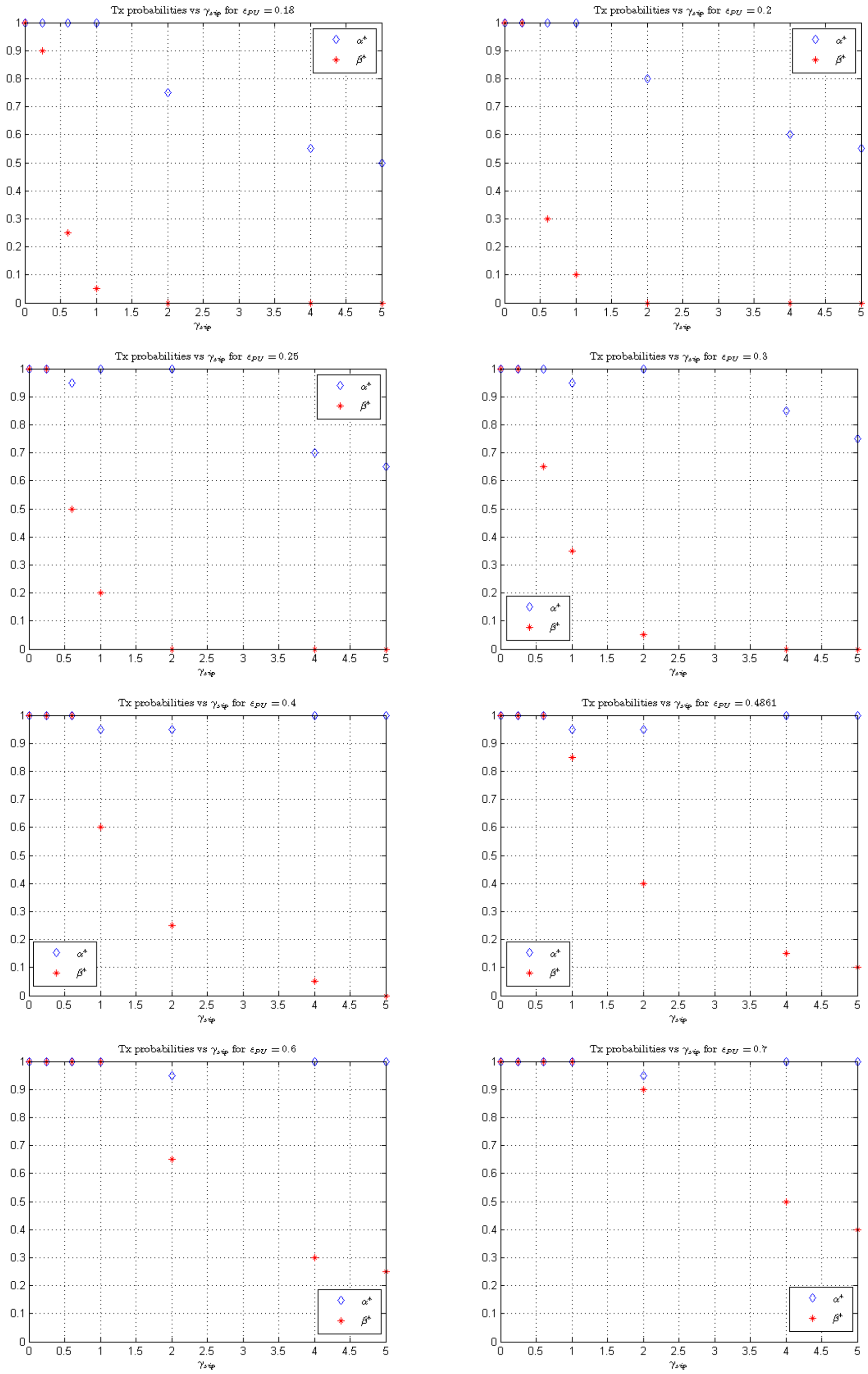
Figure 5.5: H_1: Average SUs sum Throughput vs γ_{sip}

The numerical results obtained so far consider only one specific value of SNR for the SUs, $\gamma_{sip} = 2$, $i \in \{1, 2\}$. Fig. 5.5 depicts the average SUs sum throughput with respect to γ_{sip} for varying ϵ_{PU}; it shows different evolutions based on the PU constraint: for $\epsilon_{PU} < 0.2$ the bigger the SNR, the lower the maximum SUs sum throughput; this is reasonable, since small values of ϵ_{PU} correspond to a tight constraint, i.e., harder to satisfy, which implies a decrease in the SUs sum throughput as they create a higher interference level at PU_{rx} since they have to limit their transmissions. For $0.2 \leq \epsilon_{PU} \leq 0.8$ the average SUs sum throughput increases for γ_{sip} below a certain value and decreases for γ_{sip} over it; this is due to the fact that the PU degradation constraint is not active for small values of the SNR, i.e., for a low interference level at PU_{tx}, thus, the SUs can exploit the transmitting chances much more and utilize their transmitting power to gain a higher reward. On the other hand, when the interference powers of the SUs become too high they affect significantly the PU performace, so they have to limit their channel accesses in order to respect the PU constraint with a consequent reduction of the maximum achievable throughput. For $\epsilon_{PU} = 1$ instead the average SUs sum throughput tends to grow as the SNR increases, i.e., the PU degradation constraint is not active for the considered SNRs and the SUs can exploit the transmitting chances as much as possible.

Fig. 5.6 depicts the α and β which grant the maximum SUs sum throughput with respect to γ_{sip} for varying ϵ_{PU}; as we already observed, for $\gamma_{sip} = 2$

they tend to increase as ϵ_{PU} increases since the SUs exploit the transmitting chances as much as possible. If we consider their evolution with respect to the SNR they tend to decrease as γ_{sip} grows and this is reasonable since the level of interference they cause at PU_{rx} increases, so they have to limit their access to the channel to satisfy the PU constraint; only for $\epsilon_{PU} = 1$ do the SUs always transmit since the PU constraint is not active for the considered SNRs.




Tx probabilities vs γ_sip for ε_PU = 0.18
Tx probabilities vs γ_sip for ε_PU = 0.2
Tx probabilities vs γ_sip for ε_PU = 0.25
Tx probabilities vs γ_sip for ε_PU = 0.3
Tx probabilities vs γ_sip for ε_PU = 0.4
Tx probabilities vs γ_sip for ε_PU = 0.4861
Tx probabilities vs γ_sip for ε_PU = 0.6
Tx probabilities vs γ_sip for ε_PU = 0.7
α*
β*
γ_sip

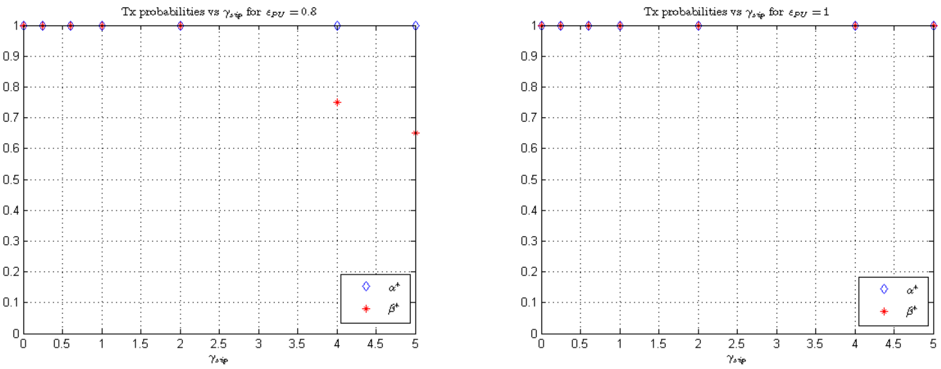


Figure 5.6: H_1: (α^*, β^*) vs γ_{sip}

In order to analyze the degradation of H_1 performance, we test it by using the α^* and β^* obtained for $\gamma_{s1p} = \gamma_{s2p} = 2$ and imposing $\gamma_{s1p} = \gamma_{s2p} \neq 2$.

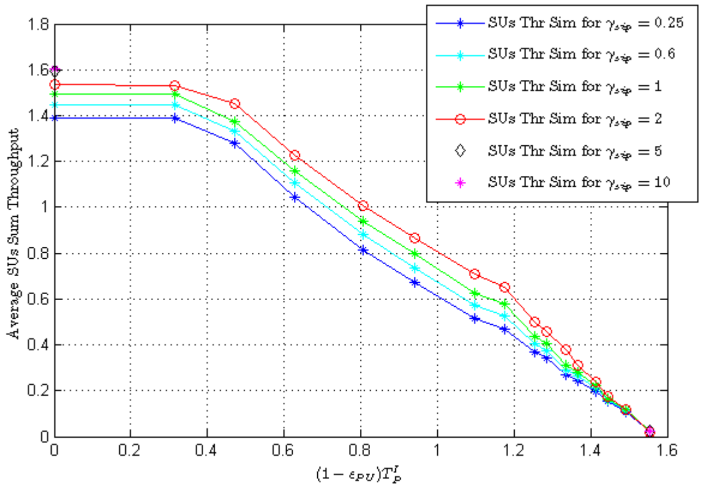


Figure 5.7: H_1 Robustness: Average SUs sum throughput vs γ_{sip}

Fig. 5.7 depicts the degradation of the average SUs sum throughput with respect to the PU throughput constraint for various value of γ_{sip}; it shows the gap between performance obtained for $\gamma_{sip} = 2$ and performance obtained for different SNRs values grows as the SNR decreases. This is reasonable since for $\gamma_{sip} < 2$ the SUs create a lower interference level at PU_{rx}, thus the PU retransmits less and they have less chances to exploit FIC and increase their reward, whereas for $\gamma_{sip} > 2$ the transmitting behavior adopted, which is the optimal one for $\gamma_{sip} = 2$, is too aggressive and we are not able to find a result because the SUs do not succeed in satisfying the PU throughput constraint except for $\epsilon_{PU} = 1$ for which the constraint is not active (see Fig. 5.8) and for $\epsilon_{PU} = 0.01$ case in which probably the transmission probabilities adopted are very low or the transmitting chances very sporadic. We can note that for high SNRs, i.e., for $\gamma_{sip} > 2$, and $\epsilon_{PU} = 1$ there is a performance improvement; we suppose it is due to the fact that the higher SUs interference power at PU_{rx} makes PU to retransmit frequently so gives them more chances to increase their throughput. This consideration clearly appear in Fig. 5.8.

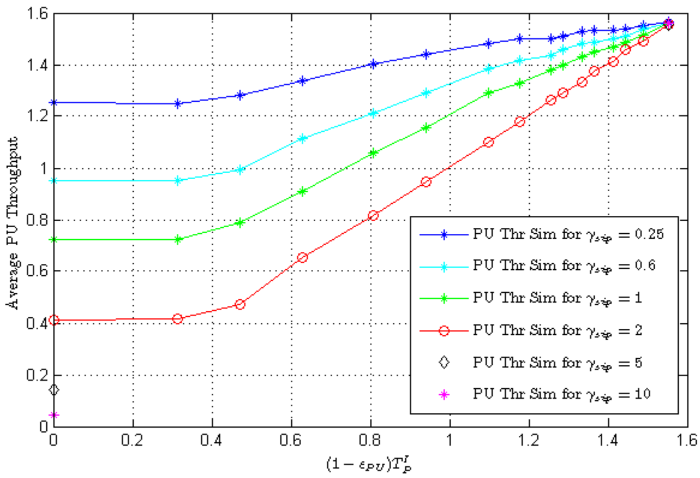


Figure 5.8: H_1 Robustness: Average PU throughput vs γ_{sip}

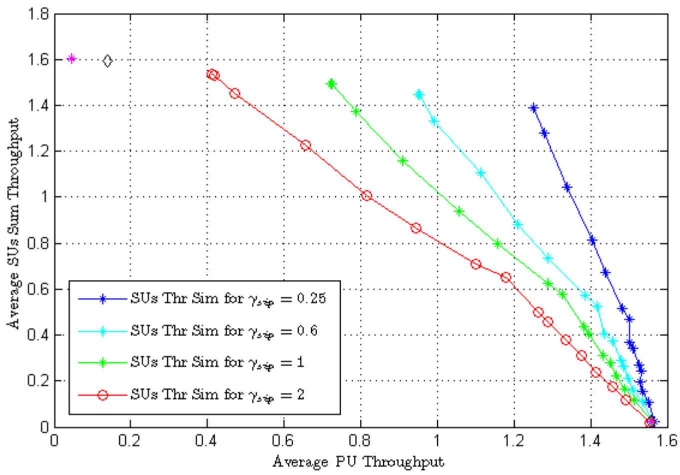


Figure 5.9: H_1 Robustness: Average SUs sum throughput-Average PU throughput tradeoff

Fig. 5.9 is very useful to better understand the tradeoff that affects the heuristic performance: it depicts the average long term SUs sum throughput with respect to the average PU throughput in the same situations examined in

Fig 5.7. We can note that as the SNR of the interference channel, γ_{sip}, decreases the performance generally improves and the ranking reverses with respect to the results shown in Fig. 5.7. It is evident that greater SNR values allow to reach a higher average SUs sum throughput but at price of a more significant degradation of PU performance, i.e., the tradeoff between what we can gain and the cost we have to pay clearly appears, whereas in Fig. 5.7 the PU advantage due to the interference level reduction as γ_{sip} decreases is not evident because it underlines only the effect of using a transmitting behavior too conservative in respect to the real interference level at PU_{rx}.

In the performance analysis it is interesting to consider also the asymmetric case, i.e., to suppose that the SUs have different transmission probabilities, denoted by α_i and β_i, which represents SU_i transmission probability when the PU message is known and unknown, respectively, $i \in \{1, 2\}$. As in the symmetric case, we explore the space of all possible transmission probabilities and find the α_1, β_1, α_2 and β_2 which maximize the SUs sum throughput.

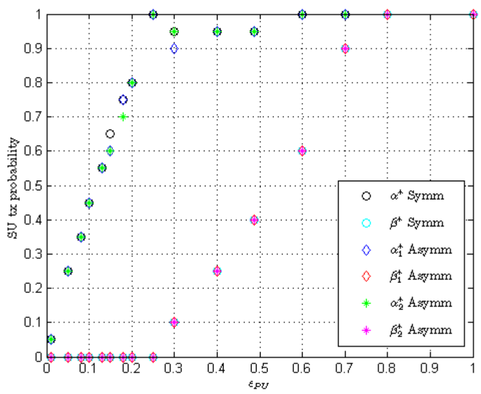


Figure 5.10: H_1: $(\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*)$ vs ϵ_{PU}

Fig. 5.10 illustrates the best transmissions probabilities for varying ϵ_{PU} compared with the optimal ones in the symmetric case; it is interesting to note that in almost all the considered cases the best choice is represented by the symmetric behavior. This is reasonable since the SUs have the same transmission parameters, i.e., the same transmission rates and the same average SNRs on the accessible channels, so they can exchange their role and access the channel with the same frequency in order to maximize their reward.

Fig. 5.11 instead illustrates the best transmission probabilities for varying ϵ_{PU}, again compared with the optimal ones in the symmetric case, but imposing that $\alpha_1 \neq \alpha_2$ and $\beta_1 \neq \beta_2$, i.e., considering the first suboptimal case; we can note that for intermediate values of ϵ_{PU} the optimal transmission probabilities in the asymmetric case are distributed around the symmetric ones, in particular one is bigger and the other is smaller than α^* or β^*, i.e., there is a sort of balance between the SU activities. This is reasonable since the SUs have the same

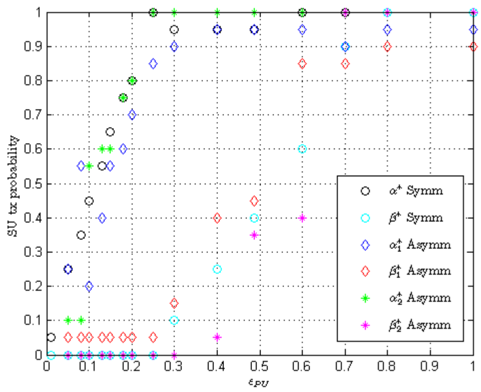


Figure 5.11: $H_{1,subOpt}$: $(\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*)$ vs ϵ_{PU}

transmission paremeters, thus, if one of them is a little more aggressive the other has to be a little more 'idle' in order to gain a reward close to the symmetric one which represent the best case. For low and high values of ϵ_{PU}, instead, one of the SU reaches the symmetric transmission probability while the other adopts the closest one to respect the asymmetric conditions; this is reasonable since for extreme values of ϵ_{PU} the SUs in order to gain the maximum reward have to transmit as much as possible (for high ϵ_{PU}, i.e., when the PU constraint is more relaxed) or the least as possible (for low ϵ_{PU}, i.e., when the PU constraint is very tight), thus they try to exploit or not the transmitting chances in the best way they are allowed to. Finally, in both figures we can note that the higher ϵ_{PU}, the bigger the tendency to grow of the α_i and β_i which maximize the SUs sum throughput, denoted by α_i^* and β_i^*, i.e., as usual as ϵ_{PU} increases the two SUs adopt a more aggressive policy to exploit the transmitting chances as much as possible, $i \in \{1, 2\}$.

The SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} is depicted in Fig. 5.12 and compared with the upper bound represented by the centralized case (MMDP). As in the symmetric case, as the PU throughput increases, the average SUs sum throughput decreases and the numerical results are obtained by selecting for each value of ϵ_{PU} considered the values of α_i and β_i which maximize the SUs sum throughput under this heuristic, $i \in \{1, 2\}$. In Fig. 5.12 is also depicts the symmetric case; we can note that the performace in the first suboptimal case almost coincides with the symmetric one for low and intermidiate ϵ_{PU}, whereas are below the latter for high ϵ_{PU}, this is because as we have just evidence examining the transmission probabilities, the SUs manifest a tendency to balance their actions in order to reach the best reward for intermediate value of ϵ_{PU}; when the PU constraint is relaxed instead the best performance can be reached only with a symmetric approach, i.e., to completely exploit the transmitting chances, the SUs have to transmit with the same probability, and thus there exist an unavoidable degradation due to the asymmetry imposed to the transmission probabilities.

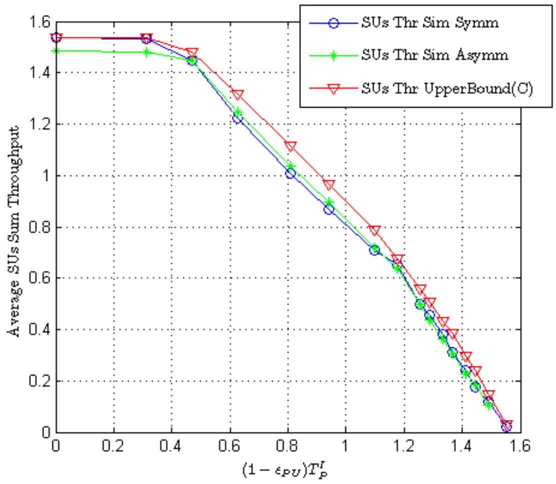


Figure 5.12: $H_{1,subOpt}$: Average SUs sum throughput with respect to PU throughput constraint

A final aspect interesting to analyze is the effect of the asymmetries in the SNR on the heuristic performance; Fig. 5.13 depicts the average SUs sum throughput with respect to the PU throughput constraint for three specific values of γ_{s2p} and compares them with the symmetric case.

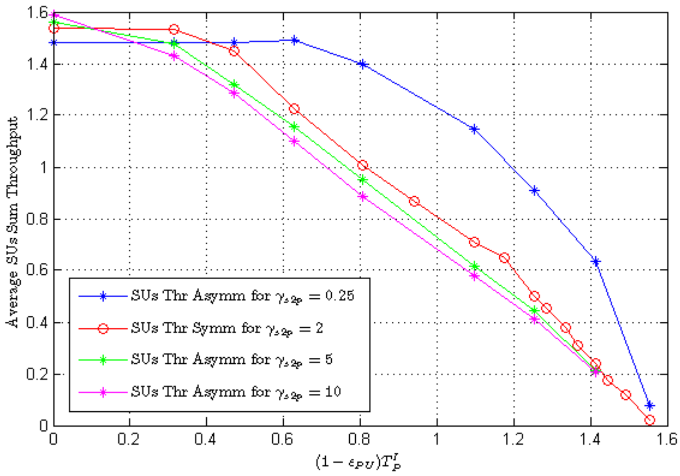


Figure 5.13: $H_{1,ASY}$: Average SUs sum throughput with respect to PU throughput constraint

In all the examined scenarios there is SU_1 which creates a constant average interference level at PU_{rx} and affects the PU performance more or less in dependence of the tightness of the constraint, then there is SU_2 whose SNR changes: for $\gamma_{s2p} = 0.25$ SU_2 creates an interference level almost unperceivable at PU_{rx}, thus it can exploit its transmitting chances more than in the symmetric case, i.e., for $\gamma_{s2p} = 2$, and increase the average SUs reward. Obviously, once it reaches its maximum transmission capability there is a saturation of the SUs throughput which correspond to the maximum degradation of PU throughput (see Fig. 5.14). On the other hand, for $\gamma_{s2p} > 2$ the total interference caused by the SUs at PU_{rx} is higher than the symmetric case with a consequent degradation of the performace, in effect as γ_{s2p} increases SU_2 has to limit its accesses to the channel in order to respect the PU constraint. As we already unerlined in the DEC-MMDP analysis, the considerations done so far do not hold for $\epsilon_{PU} = 1$: performance results are worse than the symmetric case for $\gamma_{s2p} = 0.25$, whereas they are better than the symmetric case for $\gamma_{s2p} > 2$; this is probably due to the fact that for small values of SU_2's SNR the interference level is very low, so the PU retransmits rarely and the SU_2 has less chances to transmit successfully and increase its own reward, whereas for high values of SU_2's SNR the situation is inverted, i.e., the PU retransmits more frequently and so SU_2 has more chances to exploit FIC and improve its own throughput.

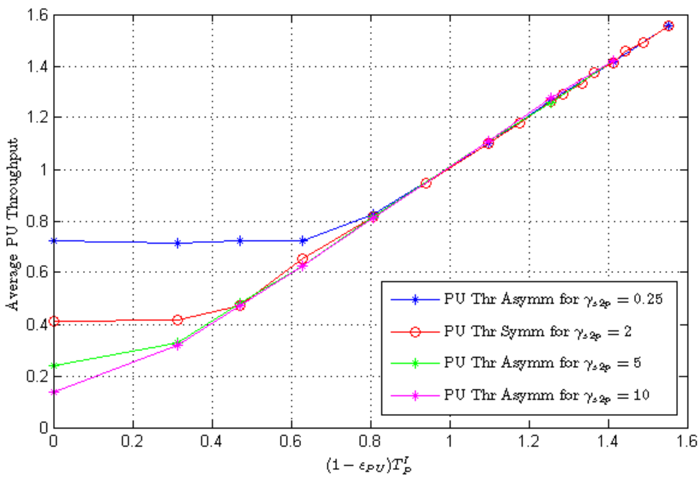


Figure 5.14: $H_{1,ASY}$: Average PU throughput with respect to PU throughput constraint

Fig. 5.14 depicts the average PU throughput with respect to the PU throughput constraint by varying the value of ϵ_{PU} in the same asymmetric situations illustrated in Fig 5.13. We can note that for $\epsilon_{PU} \geq 0.6$ the asymmetries in the SNR influence the PU throughput degradation, in particular, for $\gamma_{s2p} > 2$ SU_2's interference power at PU_{rx} causes a degradation of PU performance with respect to the symmetric case, i.e., for $\gamma_{s2p} = 2$, whereas for $\gamma_{s2p} = 0.25$ there is an improvement in PU performance due to the lower interference level SU_2 creates at PU_{rx}.

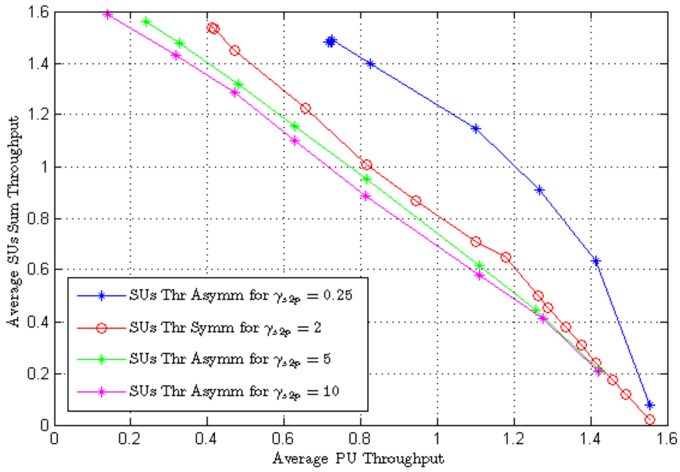


Figure 5.15: $H_{1,ASY}$: Average SUs sum throughput-Average PU throughput tradeoff

Fig. 5.15 is very useful to better understand the tradeoff that affects the heuristic performance: it depicts the average long term SUs sum throughput with respect to the average PU throughput in the same asymmetric situations examined in Fig. 5.13. We can note that as the SNR of the interference channel, γ_{s2p}, decreases the performance generally improves; furthermore, it is evident that greater SNR values allow to reach a higher average SUs sum throughput but at price of a more significant degradation of PU performance, i.e., the tradeoff between what we can gain and the cost we have to pay clearly appears, whereas in Fig. 5.13 the PU advantage due to the interference level reduction as γ_{s2p} decreases is not so evident.

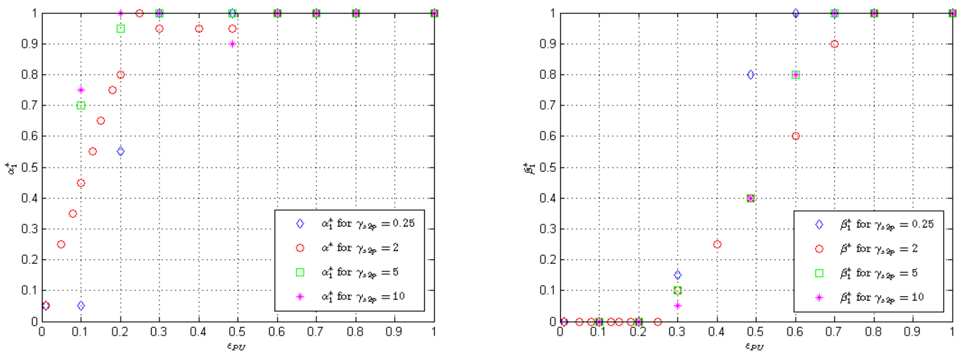


Figure 5.16: $H_{1,ASY}$: α_1^* and β_1^* vs γ_{s2p}

Fig. 5.16 and 5.17 depict SU_1's and SU_2's optimal transmission probabilities, respectively, for the same values of γ_{s2p} we just considered compared with the optimal ones in the symmetric case. We can note that for $\gamma_{s2p} = 0.25$ SU_2

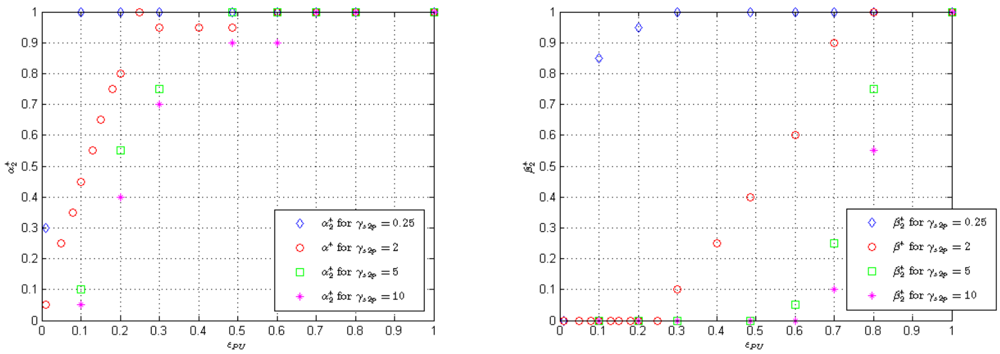
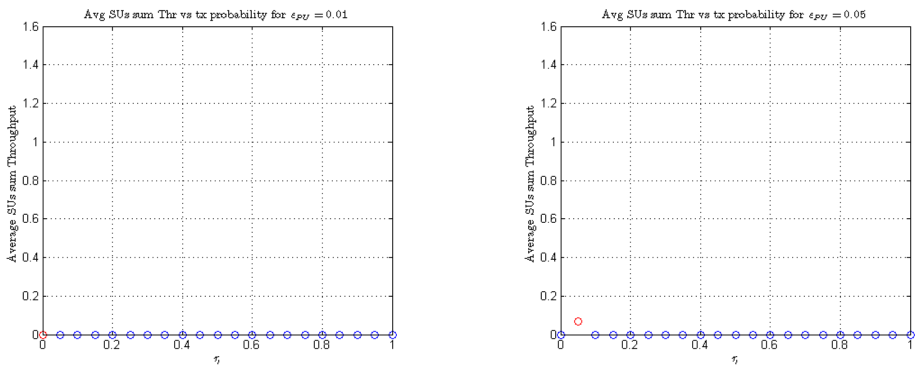


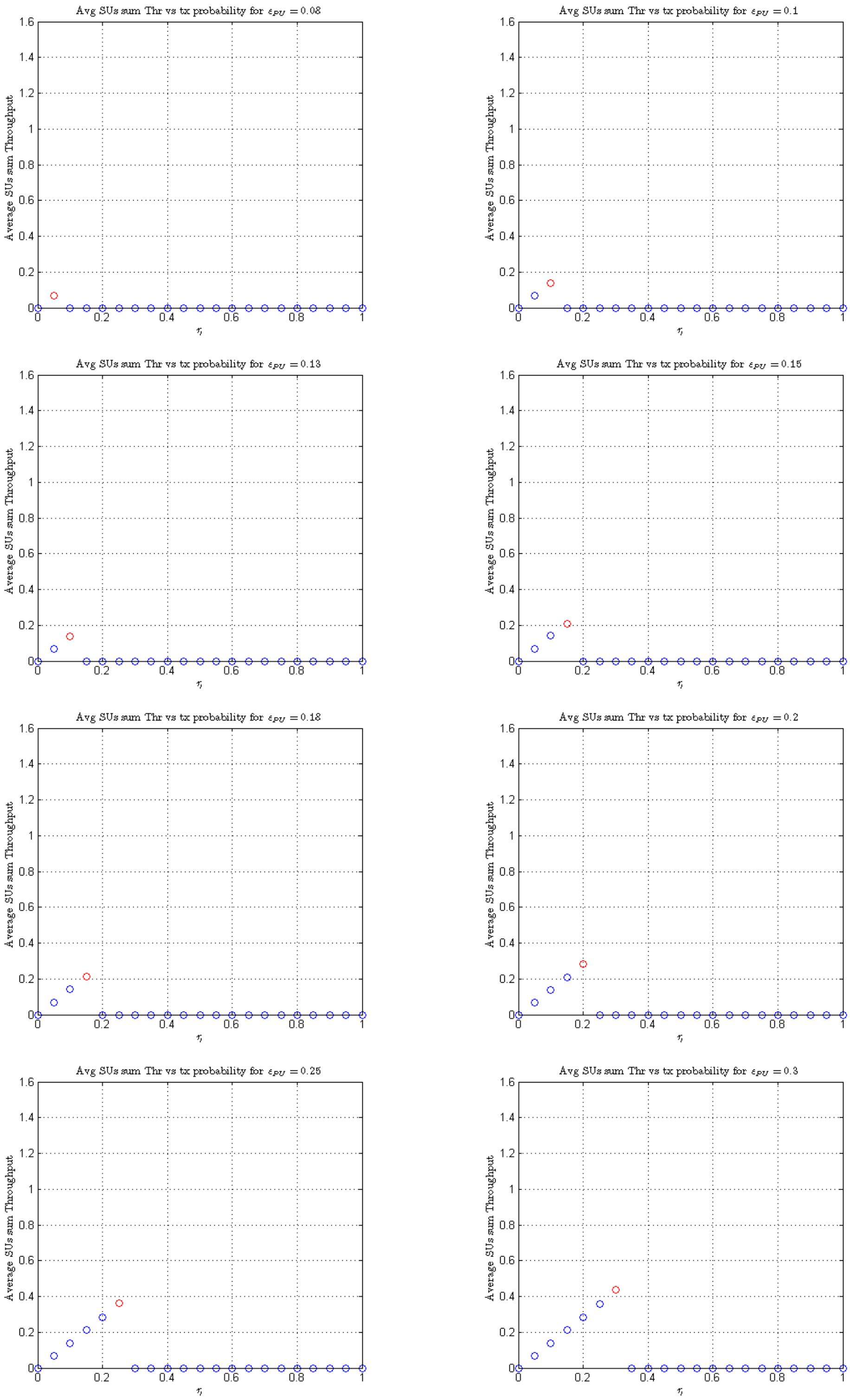
Figure 5.17: $H_{1,ASY}$: α_2^* and β_2^* vs γ_{s2p}

adopts a more aggressive behavior than in the symmetric case as we supposed before, in particular for small values of ϵ_{PU} and if it does not know the PU message, whereas SU_1 limits its accesses to the channel; on the other hand, as γ_{s2p} increases the SUs exchange their roles, i.e., SU_2 tends to limit its accesses to the channel as SU_1 exploits as much as possible the transmitting chances. In other word, there is a sort of balance of the SUs behavior: according to its SNR SU_2 adopts a certain trasmitting behavior and SU_1, whose SNR does not change, tries to balance the effects of SU_2's interference at PU_{rx}.

5.2 Heuristic Access Policy H_2

A simpler heuristic policy we analyze, denoted by H_2, does not consider the influence of the PU message knowledge in the activities of the SUs. We simply suppose that there is only one transmission probability, η, irrespective of the PU knowledge state of the system. As in heuristic H_1, we need to identify which value of η guarantees the highest possible SUs sum throughput. Considering the same average SNRs, transmission rates and PU throughput constraints used to obtain numerical results in the optimal centralized and decentralized case, we examine the performance of the CR network for different values of η.




Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.08$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.1$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.13$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.15$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.18$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.2$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.25$
Avg SUs sum Thr vs tx probability for $\epsilon_{PU} = 0.3$
Average SUs sum Throughput
τ

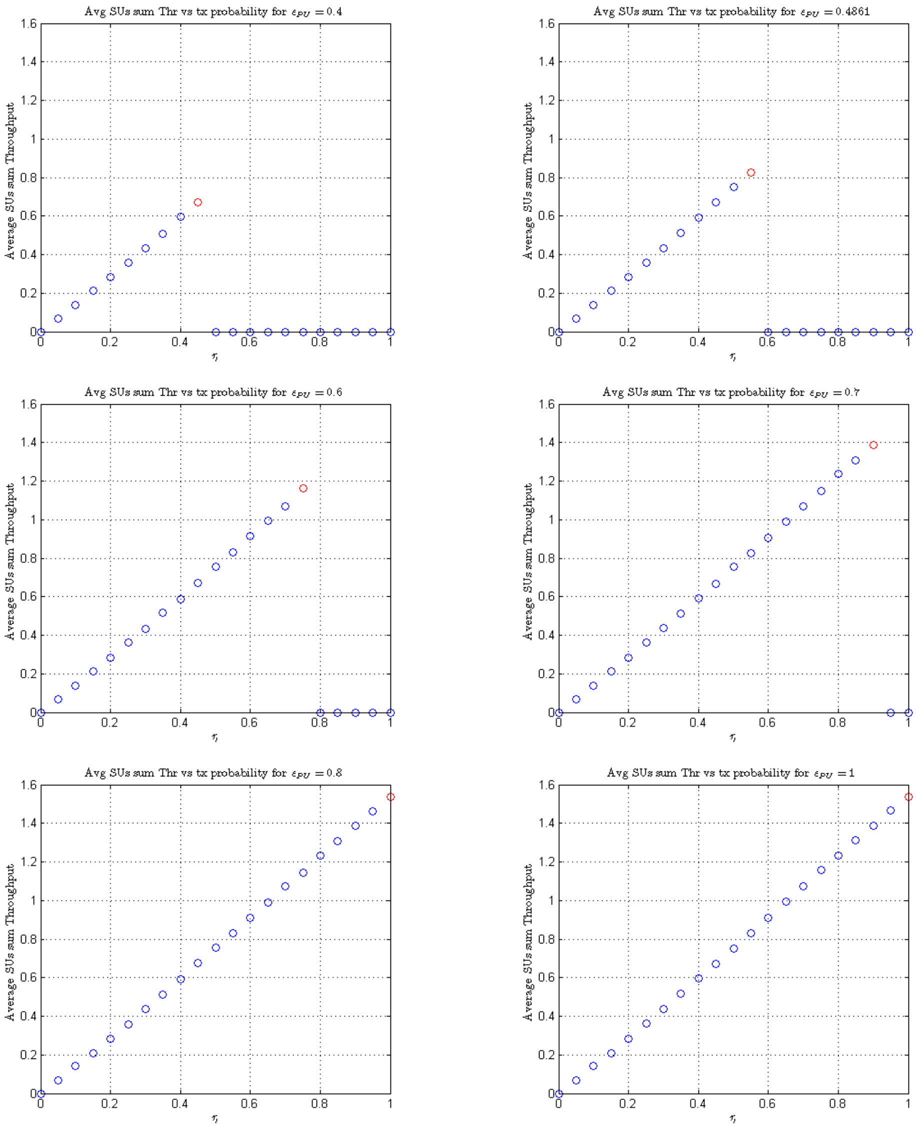


Figure 5.18: H_2: Average SUs sum Throughput vs η

Fig. 5.18 illustrates the η search for varying ϵ_{PU}; it shows again that the higher ϵ_{PU}, the bigger the tendency to grow of the η which grants the maximum SUs sum throughput, denoted by η^*, i.e., as ϵ_{PU} increases the two SUs adopt a more aggressive policy to exploit the transmitting chances as much as possible (see Fig. 5.19). Note that when $\eta = 0$ or the SUs sum throughput is zero means that it is impossible to respect the PU throughput constraint.

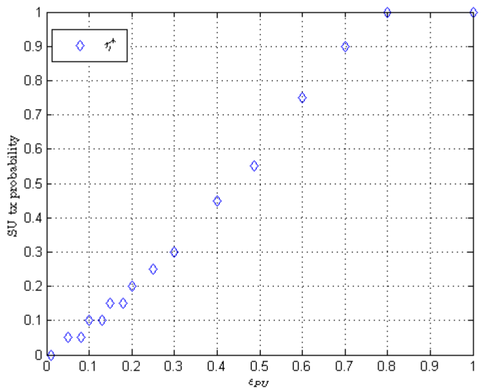


Figure 5.19: H_2: η^* vs ϵ_{PU}

The SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} is depicted in Fig. 5.20 and compared with the upper bound represented by the centralized case (MMDP). Obviously, as the PU throughput increases, the average SUs sum throughput decreases. The numerical results are obtained by selecting for each value of ϵ_{PU} considered the value of η which maximizes the SUs sum throughput under this heuristic.

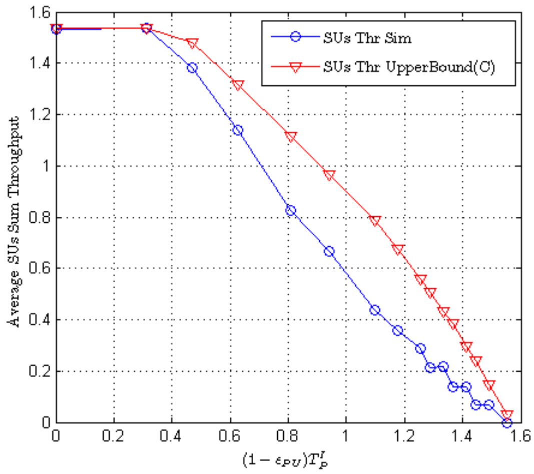
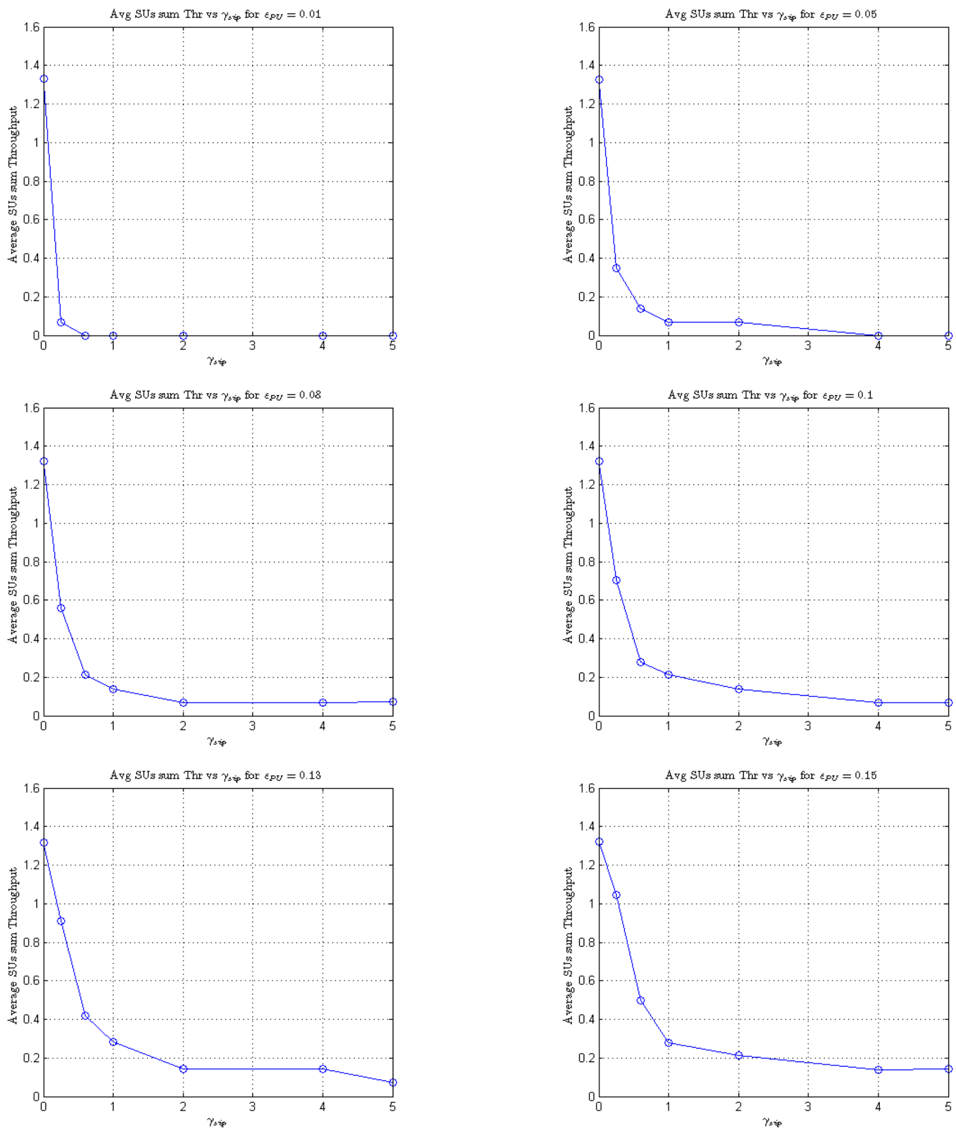
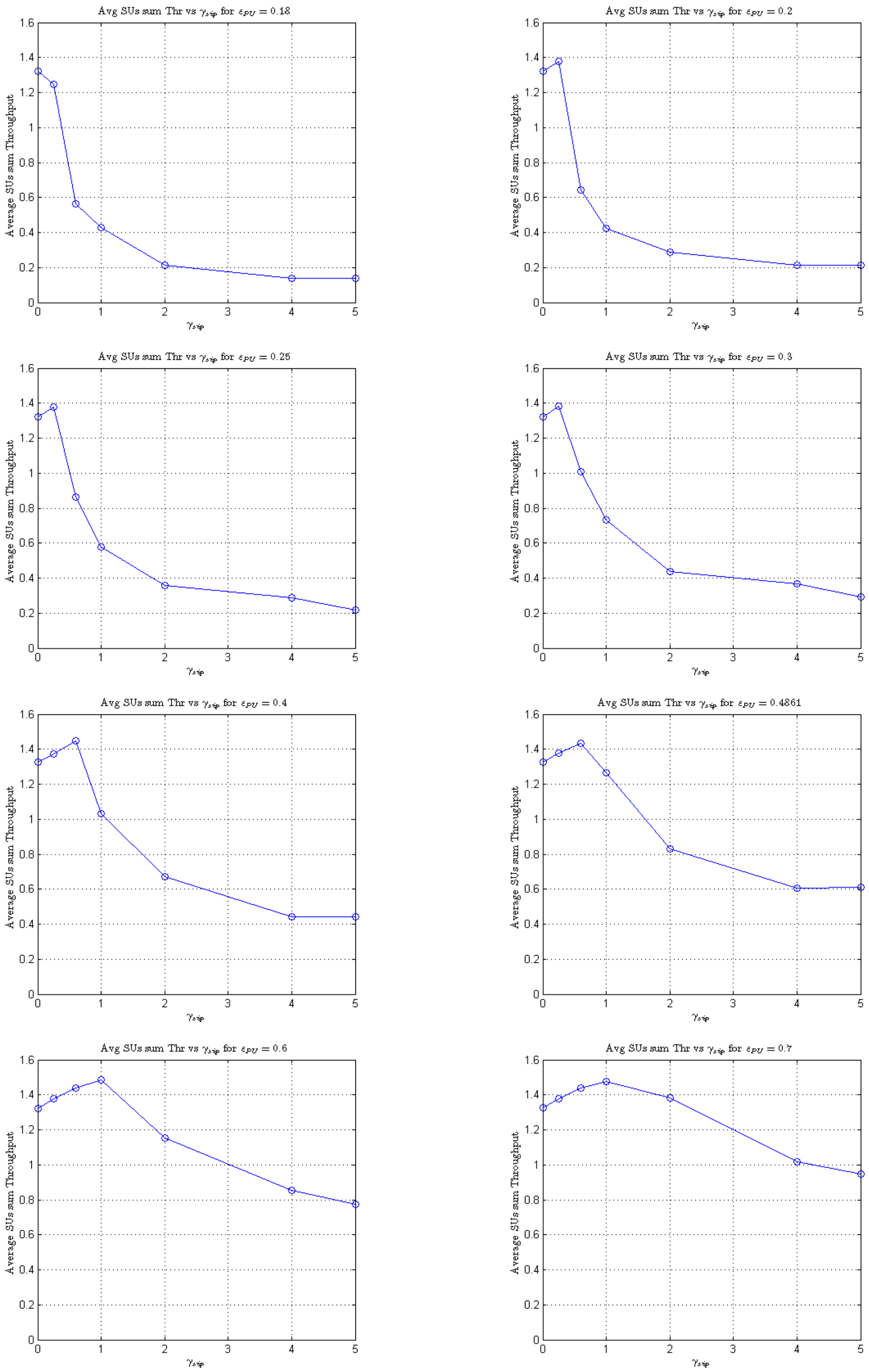


Figure 5.20: H_2: Average SUs sum throughput with respect to PU throughput constraint

As for H_1, in Fig. 5.21 we consider the average SUs sum throughput with respect to γ_{sip} for varying ϵ_{PU}; again, it shows different evolutions based on the PU constraint: for $\epsilon_{PU} < 0.2$ the bigger the SNR, the lower the maximum SUs sum throughput, for $0.2 \leq \epsilon_{PU} \leq 0.8$ the SUs reward increases for γ_{sip} below a certain value and decreases for γ_{sip} over it, and for $\epsilon_{PU} = 1$ the average SUs sum throughput tends to grow as the SNR increases. As in H_1 case, we can explain the different behaviors repeating the same considerations on the effect of the SUs interference power on the PU performance and the consequent limitation of the SUs accessibility actions to satisfy the PU degradation constraint.




Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.18$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.2$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.25$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.3$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.4$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.4861$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.6$
Avg SUs sum Thr vs γ_{sip} for $\epsilon_{PU} = 0.7$
Average SUs sum Throughput
γ_{sip}

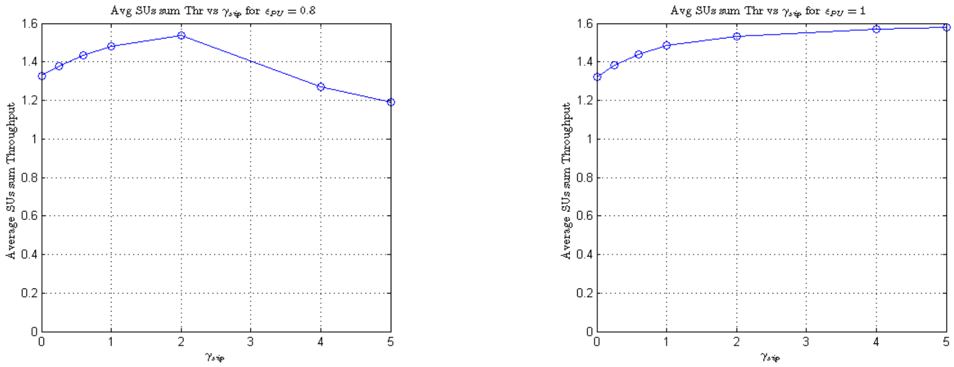
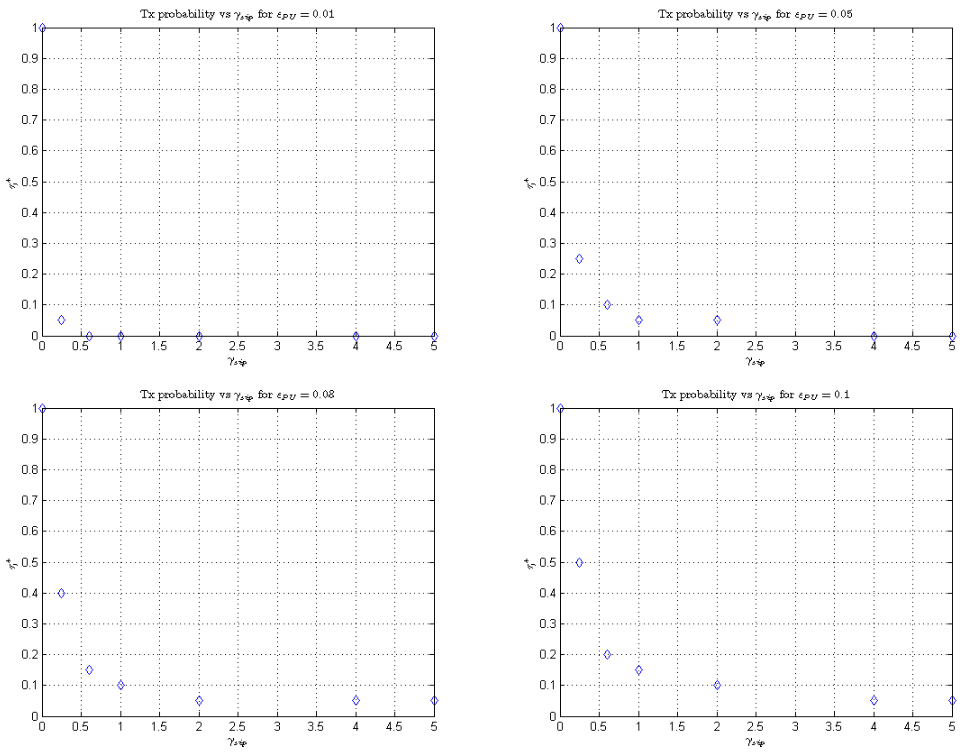
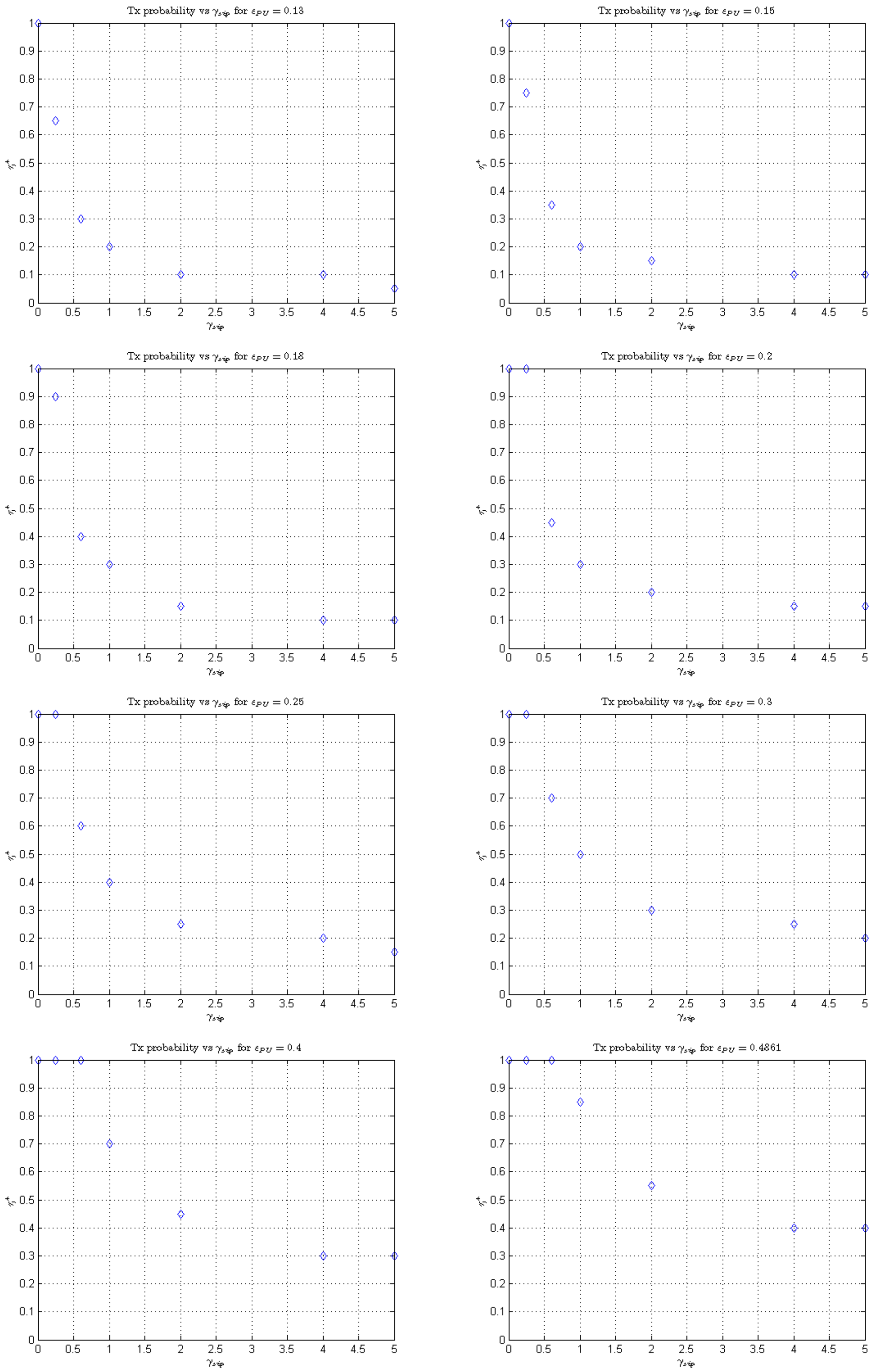


Figure 5.21: H_2: Average SUs sum Throughput vs γ_{sip}

Fig. 5.22 depicts the η which grants the maximum SUs sum throughput with respect to γ_{sip} for varying ϵ_{PU}; as we already observed, for $\gamma_{sip} = 2$ it tends to increase as ϵ_{PU} increases since the SUs exploit the transmitting chances as much as possible. As α and β in H_1, if we consider its evolution with respect to the SNR it tends to decrease as γ_{sip} grows because the interference SUs cause at PU_{rx} increases, so they have to limit their access to the channel to satisfy the PU constraint; only for $\epsilon_{PU} = 1$ do the SUs always transmit since the PU constraint is not active for the considered SNRs.





Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.13$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.15$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.18$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.2$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.25$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.3$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.4$
Tx probability vs γ_{sip} for $\epsilon_{PU} = 0.4861$
γ_{sip}

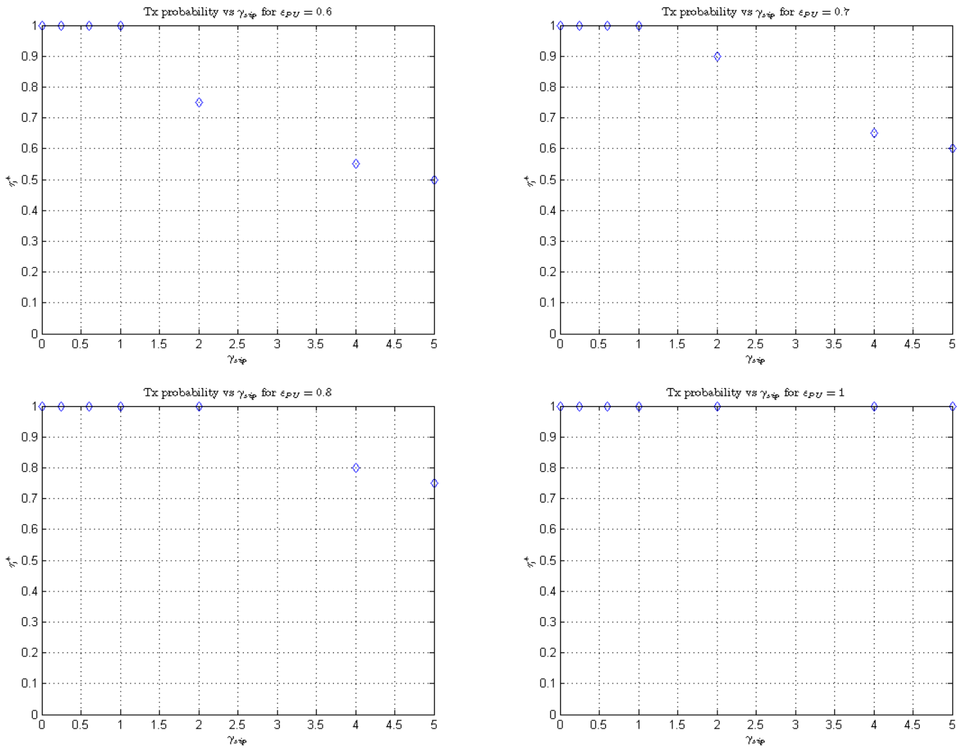


Figure 5.22: H_2: η^* vs γ_{sip}

As in H_1 case, in order to analyze the degradation of H_2 performance, we test it by using the η^* obtained for $\gamma_{s1p} = \gamma_{s2p} = 2$ and imposing $\gamma_{s1p} = \gamma_{s2p} \neq 2$. Fig. 5.23 depicts the degradation of the average SUs sum throughput with respect to the PU throughput constraint for various value of γ_{sip}; while in H_1 case the gap between curves corresponding to different γ_{sip} is almost costant as the PU constraint changes, in this case there is no gap for small values of ϵ_{PU}, while as the PU constraint is relaxed the gap grows. This is because for small value of ϵ_{PU} the PU constraint is more difficul to satisfy, thus SUs limit their transmission and even if the η we adopt is lower than the best one for $\gamma_{sip} < 2$, the SUs performance degradation is unperceivable due to the scarsity of transmitting chances created by the PU that retransmits rarely. On the other hand, as ϵ_{PU} increases SUs tend to mainly exploit the transmitting chances and, since they adopt a transmission probability that is ideal for a higher SNR, they do not exploit as much as possible their trasmitting power with a consequent degradation of the maximum achievable reward which becomes more evident for high ϵ_{PU} values, i.e., when they waste much more transmitting opportunities.

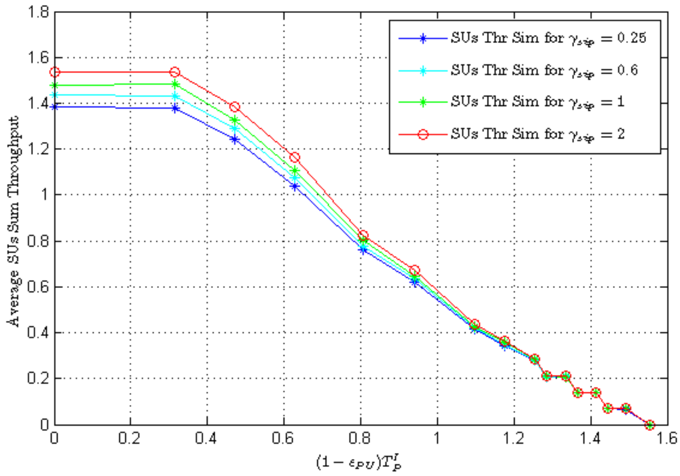


Figure 5.23: H_2 Robustness: Average SUs sum throughput vs γ_{sip}

Fig. 5.24 is very useful to better understand the tradeoff that affects the heuristic performance: it depicts the average long term SUs sum throughput with respect to the average PU throughput in the same situations examined in Fig 5.23. As in H_1 case, we can note that as the SNR of the interference channel, γ_{sip}, decreases the performance generally improves and the ranking reverses with respect to the results shown in Fig. 5.23.

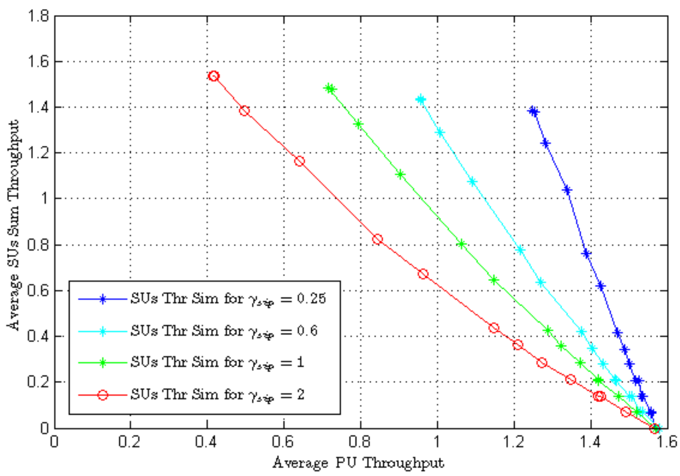
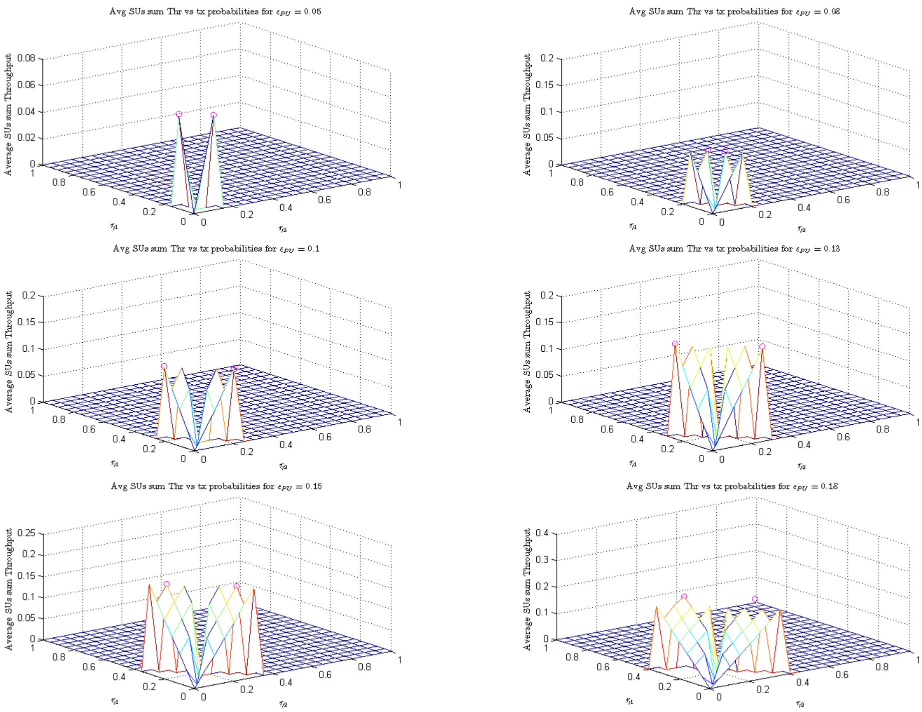


Figure 5.24: H_2 Robustness: Average SUs sum throughput-Average PU throughput tradeoff

It is evident that greater SNR values allow to reach a higher average SUs sum throughput but at price of a more significant degradation of PU performance, i.e., the tradeoff between what we can gain and the cost we have to pay clearly appears, whereas in Fig. 5.23 the PU advantage due to the interference level reduction as γ_{sip} decreases is not evident because it underlines only the effect of using a transmitting behavior too conservative in respect to the real interference level at PU_{rx}.

As for H_1, it is interesting to analyze also the asymmetric case, i.e., to suppose that the SUs have two different transmission probabilities, denoted by η_1 and η_2, respectively, irrespective of their own PU message knowledge. Fig. 5.25 illustrates the $\eta_1 - \eta_2$ search for varying ϵ_{PU}; while in the symmetric case for each considered value of ϵ_{PU} there exists only one η which grants the maximum SUs sum throughput, in the asymmetric case there exist two couples of transmission probabilities, (η_1', η_2') and (η_1'', η_2''), which maximize the SUs reward; in particular, the elements of each one of these couples are 'symmetric', i.e., $\eta_1' = \eta_2''$ and $\eta_2' = \eta_1''$. This result is reasonable because the two SUs have the same transmission parameters, i.e., the same transmission rates and the same average SNRs on the accessible channels, thus we can exchange their role without altering the performance of the system.



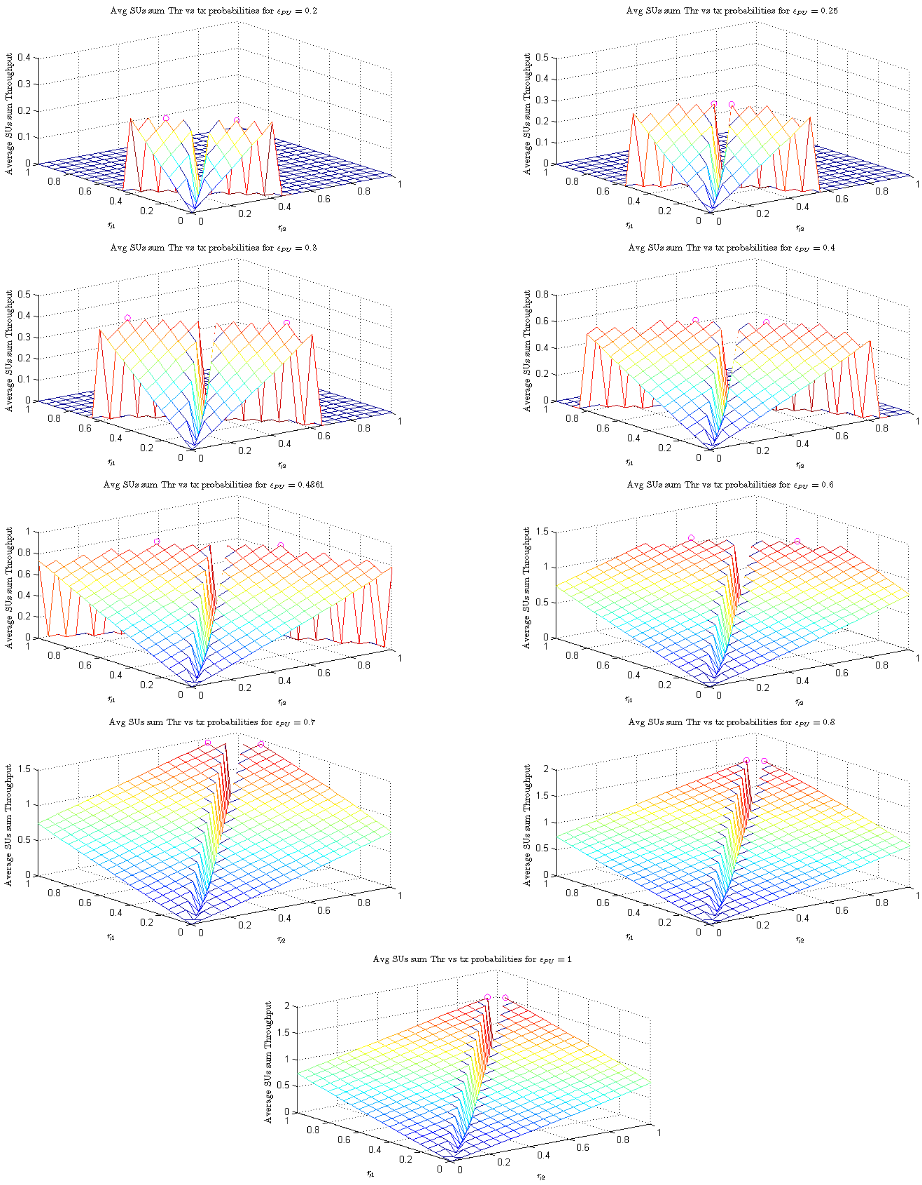


Figure 5.25: $H_{2,ASY}$: Average SUs sum Throughput vs (η_1, η_2)

As in the symmetric case, we can note that the higher ϵ_{PU}, the bigger the tendency to grow of the η_1 and η_2 which maximize the SUs sum throughput, denoted by η_1^* and η_2^*, i.e., as ϵ_{PU} increases the two SUs adopt a more aggressive policy to exploit the transmitting chances as much as possible (see Fig. 5.26). Note that when $\epsilon_{PU} = 0.01$ there exist no values of the transmission probabilities which satisfy the PU constraint, i.e., the heuristic access policy analyzed creates too much interference at PU_{rx}.

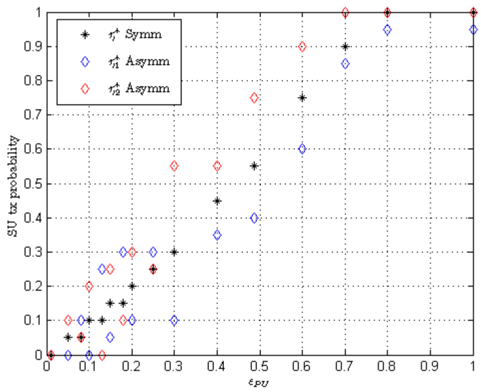


Figure 5.26: $H_{2,ASY}$: (η_1^*, η_2^*) vs ϵ_{PU}

Fig. 5.26 also represents the optimal transmission probability in the symmetric case, η^*; we can note that the optimal transmission probabilities in the asymmetric case are distributed around the symmetric one. The fact that one is a little bigger and the other a litte smaller than η^* reveal a balance between the SU activities, this is reasonable since the SUs have the same transmission parameters so if one of them is a little more aggressive the other has to be a little more 'idle' in order to gain a reward close to the symmetric one.

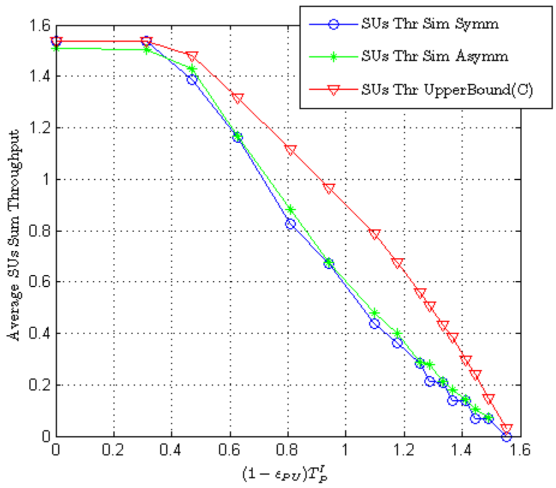
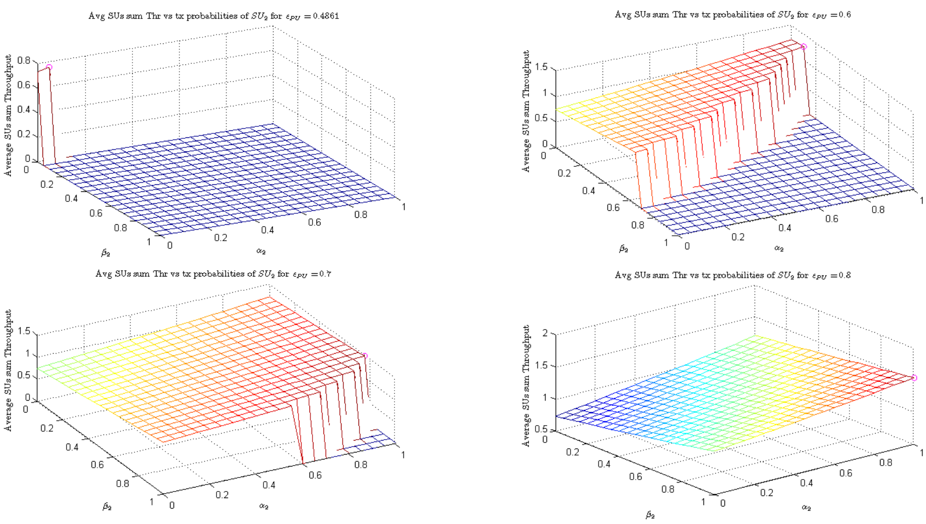


Figure 5.27: $H_{2,ASY}$: Average SUs sum throughput with respect to PU throughput constraint

The SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} is depicted in Fig. 5.27 and compared with the upper bound represented by the centralized case (MMDP). As in the symmetric case, as the PU throughput increases, the average SUs sum throughput decreases and the numerical results are obtained by selecting for each value of ϵ_{PU} considered the values of η_1 and η_2 which maximize the SUs sum throughput under this heuristic. Fig. 5.27 also depicts the symmetric case; we can note that the performace in the asymmetric case almost coincides with the symmetric one, indeed is slightly above the latter, since, as we have just evidenced examining the transmission probabilities, the SUs manifest a tendency to balance their actions in order to reach the best reward. Only for high ϵ_{PU} is the symmetric approach better than the asymmetric one since, to completely exploit the transmitting chances, the SUs have to transmit with the same probability, i.e., always.

5.3 Heuristic Access Policy H_3

The third heuristic policy we analyze, denoted by H_3, is a mix between H_1 and H_2, and represents an interesting case for the performance analysis: we suppose that one of the two SUs, for example SU_1, has a fixed transmission probability, η_1, irrespective of its own PU message knowledge state, while the other SU has two different transmission probabilities, one selected if it knows the PU message and the other chosen otherwise, denoted by α_2 and β_2, respectively. In particular, we consider three possible values of η_1 (1, 0.5 and 0.05), which correspond to an aggressive, a moderate or a weak SU_1, respectively, and for each possible case we identify which values of α_2 and β_2 guarantee the higher possible SUs sum throughput. In other words, considering the same average SNRs, transmission rates and PU throughput constraints used to obtain numerical results in the optimal centralized and decentralized case, we examine the performance of the CR network for different values of the probability couple (α_2, β_2), given a certain value of η_1, and we choose the one that gives us the maximum achievable SUs sum throughput.



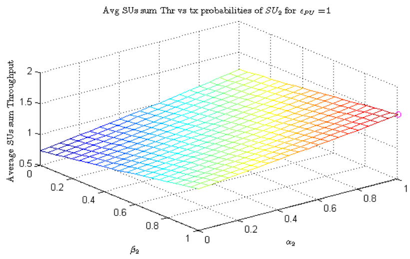


Figure 5.28: H_3: Average SUs sum Throughput vs (α_2, β_2) given $\eta_1 = 1$

Fig. 5.28 illustrates the $\alpha_2 - \beta_2$ search for varying ϵ_{PU} and given $\eta_1 = 1$; as for the previous heuristic, we can do the same considerations about the increase of ϵ_{PU} and the parallel growth of the α_2 and β_2 which grant the maximum SUs sum throughput, denoted by α_2^* and β_2^*, respectively, (see Fig. 5.29). It is important to note that there exists no couple (α_2, β_2) which satisfies the PU throughput constraint when $\epsilon_{PU} < 0.4861$. This is because the value of η_1 chosen determines a too aggressive channel access of SU_1, which creates a costant level of interference at PU_{rx} that makes it impossible to satisfy the PU constraint when the latter is very tight, i.e., for small values of ϵ_{PU}.

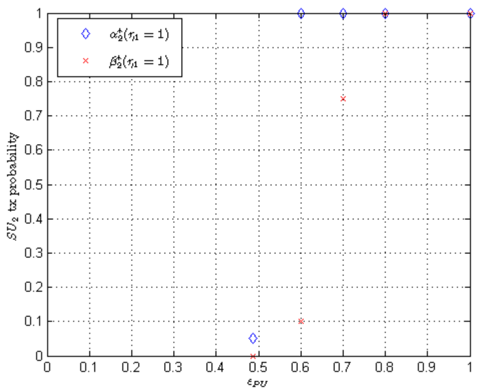


Figure 5.29: $H_3(\eta_1 = 1)$: (α_2^*, β_2^*) vs ϵ_{PU}

The SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} is depicted in Fig. 5.30 and compared with the upper bound represented by the centralized case (MMDP). Obviously, as the PU throughput increases, the average SUs sum throughput decreases. As before, the numerical results are obtained by selecting for each value of ϵ_{PU} considered the couple (α_2, β_2) which maximizes SUs sum throughput under this heuristic, given $\eta_1 = 1$.

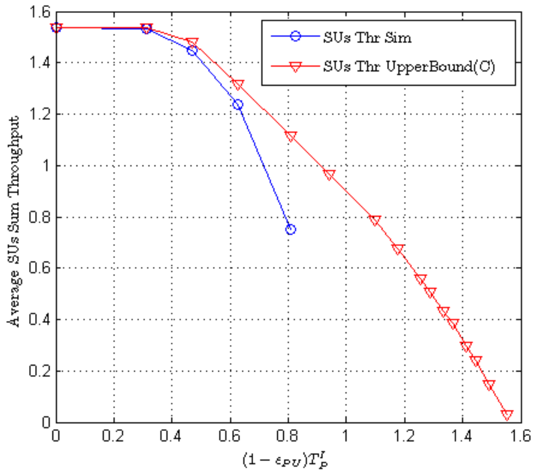


Figure 5.30: $H_3(\eta_1 = 1)$: Average SUs sum throughput with respect to PU throughput constraint

Fig. 5.31 depicts the average PU throughput with respect to the PU throughput constraint for varying ϵ_{PU}. Again, as ϵ_{PU} decreases the constraint increases and the PU throughput degradation decreases and the same considerations developed for previous heuristics are valid.

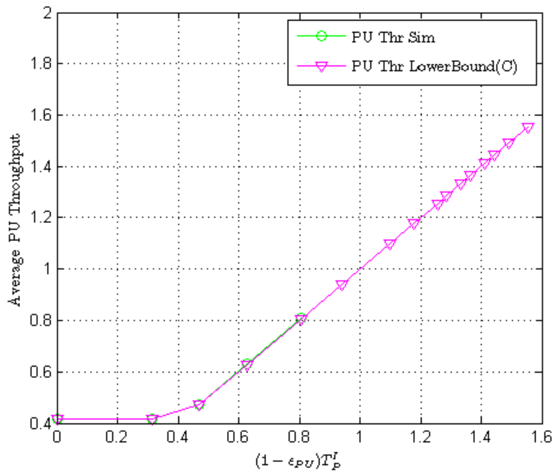


Figure 5.31: $H_3(\eta_1 = 1)$: Average PU throughput with respect to PU throughput constraint

Fig. 5.32 illustrates the $\alpha_2 - \beta_2$ search for varying ϵ_{PU} and given $\eta_1 = 0.5$; as before, it shows that as ϵ_{PU} increases SU_2 adopts a more aggressive policy to exploit the transmitting chances as much as possible, i.e., α_2^* and β_2^* grow (see

Fig. 5.33). It is important to note that there exists no couple (α_2, β_2) which satisfies the PU throughput constraint when $\epsilon_{PU} < 0.25$. This is because the value of η_1 chosen determines a too aggressive channel access of SU_1 for $\epsilon_{PU} < 0.25$, which creates a costant level of interference at PU_{rx} that makes it impossible to satisfy the PU constraint. On the other side, when $\epsilon_{PU} \geq 0.25$ SU_1 does not exploit as much as possible the opportunity of accessing the channel and thus the SUs sum throughput achieved is lower than the optimal case, in particular for high values of ϵ_{PU}.

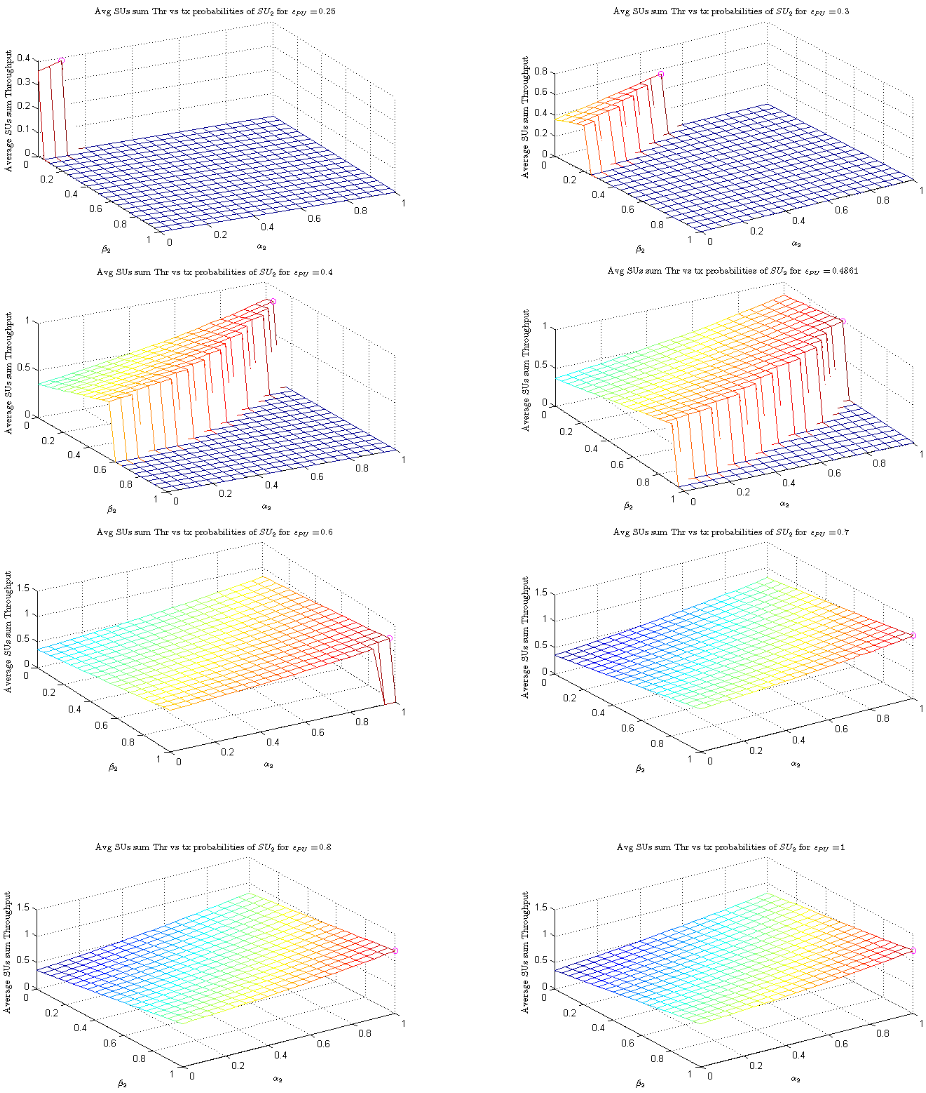


Figure 5.32: H_3: Average SUs sum Throughput vs (α_2, β_2) given $\eta_1 = 0.5$

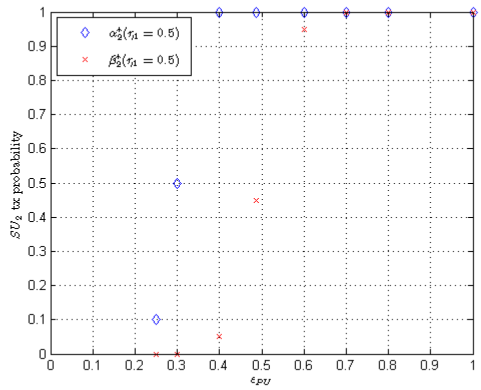


Figure 5.33: $H_3(\eta_1 = 0.5)$: (α_2^*, β_2^*) vs ϵ_{PU}

The SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} is depicted in Fig. 5.34 and compared with the upper bound represented by the centralized case (MMDP). Again, as the PU throughput increases, the average SUs sum throughput decreases, and the numerical results are obtained by selecting for each value of ϵ_{PU} considered the couple (α_2, β_2) which maximizes the SUs sum throughput under this heuristic, given $\eta_1 = 0.5$.

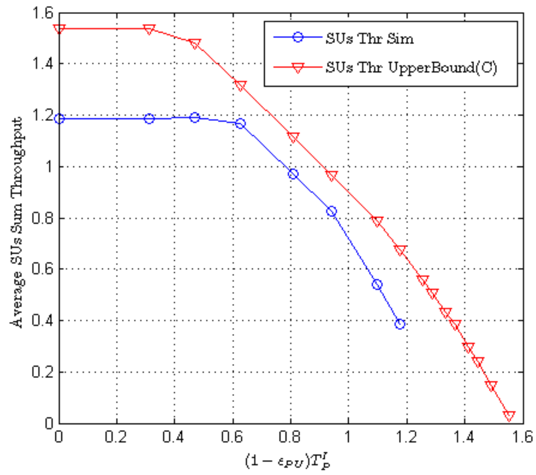


Figure 5.34: $H_3(\eta_1 = 0.5)$: Average SUs sum throughput with respect to PU throughput constraint

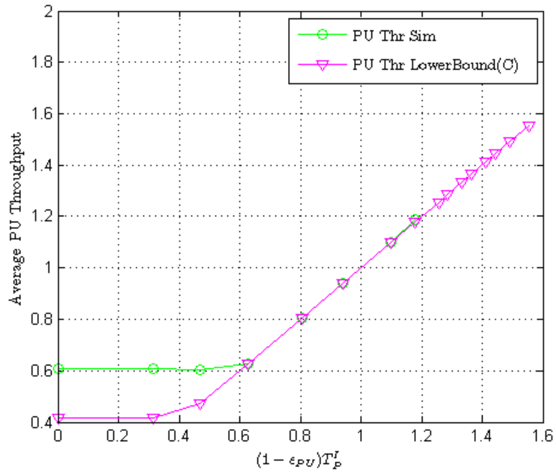
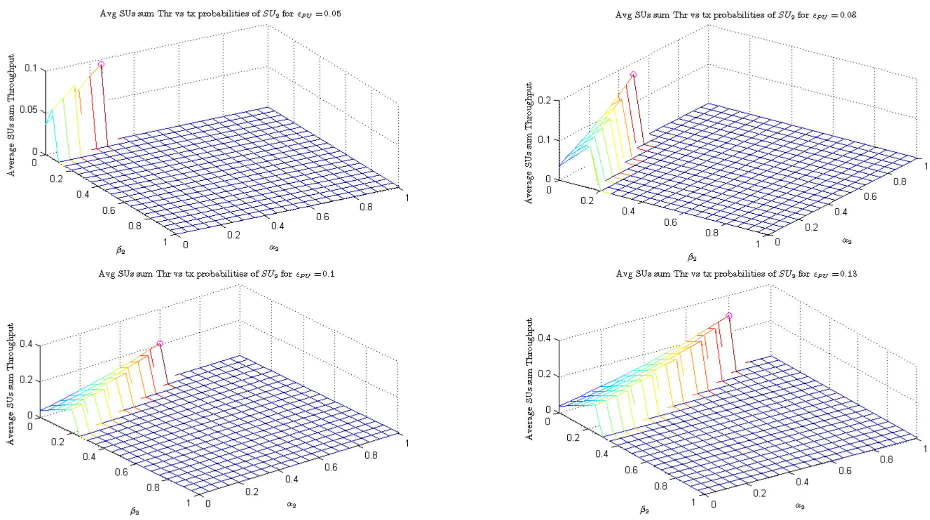


Figure 5.35: $H_3(\eta_1 = 0.5)$: Average PU throughput with respect to PU throughput constraint

Fig. 5.35 depicts the average PU throughput with respect to the PU throughput constraint for varying ϵ_{PU}. As in previous cases, we can do the same condiderations about the ϵ_{PU} increase, the resulting constraint relaxation and the PU throughput degradation increase.



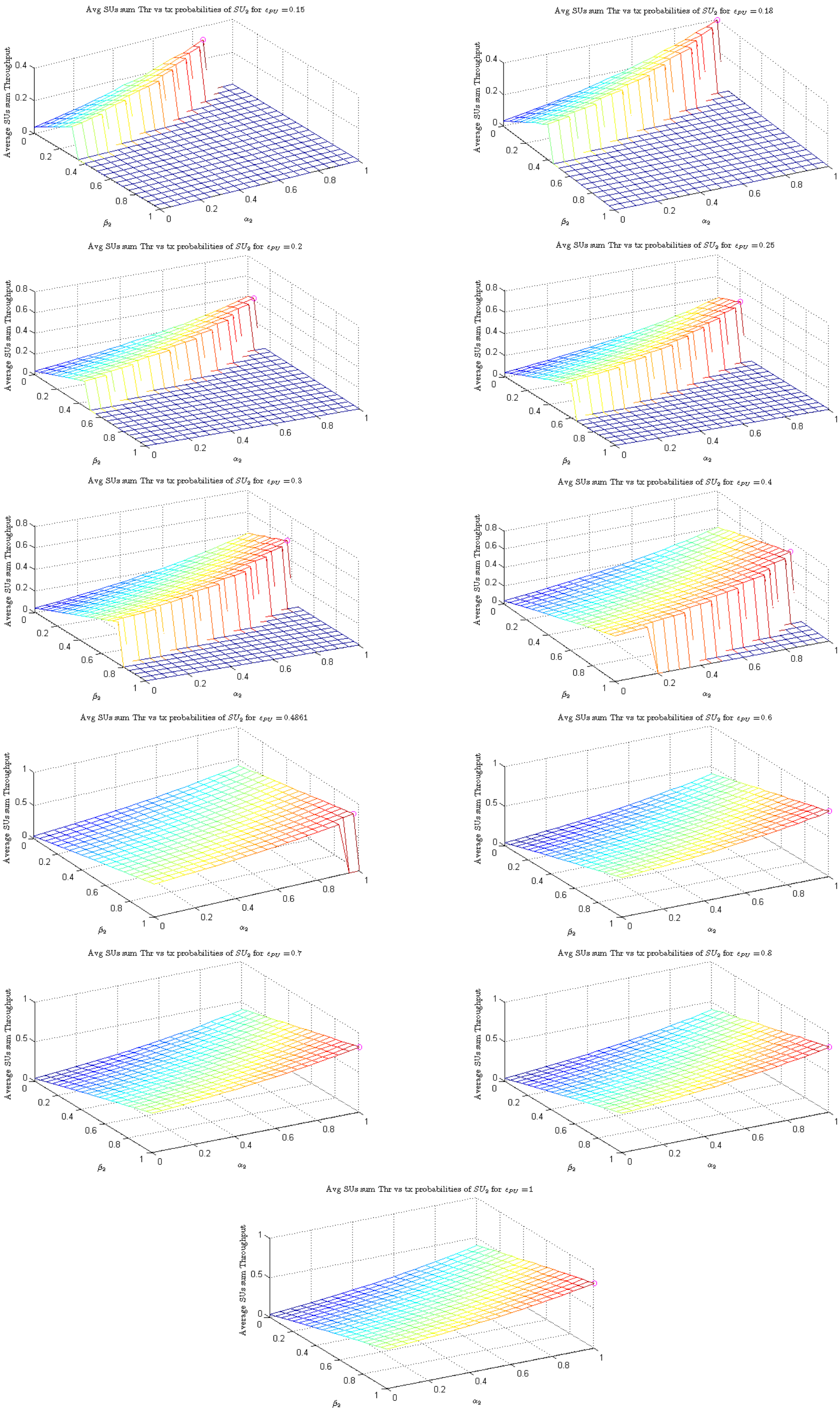


Figure 5.36: H_3: Average SUs sum Throughput vs (α_2, β_2) given $\eta_1 = 0.05$

Fig. 5.36 illustrates the $\alpha_2 - \beta_2$ search for varying ϵ_{PU} and given $\eta_1 = 0.05$; as before, it shows the growth of the α_2 and β_2 which grant the maximum SUs sum throughput and simultaneously the increase of ϵ_{PU}, i.e., the relaxation of the PU throughput constraint (see Fig. 5.37). It is important to note that there exists no couple (α_2, β_2) which satisfies the PU throughput constraint when $\epsilon_{PU} = 0.01$. This is because the value of η_1 chosen determines a too aggressive channel access of SU_1 for $\epsilon_{PU} = 0.01$, which creates a costant level of interference at PU_{rx} that makes it impossible to satisfy the PU constraint. On the other side, for high values of ϵ_{PU} SU_1 does not exploit as much as possible the opportunity of accessing the channel and thus the SUs sum throughput achieved is lower than the optimal case.

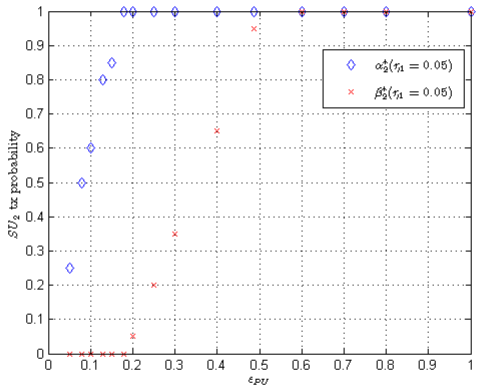


Figure 5.37: $H_3(\eta_1 = 0.05)$: (α_2^*, β_2^*) vs ϵ_{PU}

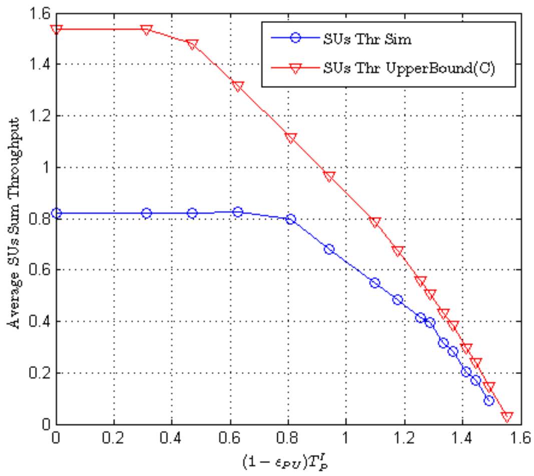


Figure 5.38: $H_3(\eta_1 = 0.05)$: Average SUs sum throughput with respect to PU throughput constraint

The SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} is depicted in Fig. 5.38 and compared with the upper bound represented by the centralized case (MMDP). Obviously, as the PU throughput increases, the average SUs sum throughput decreases. As for previous cases, the numerical results are obtained by selecting for each value of ϵ_{PU} considered the couple (α_2, β_2) which maximizes the SUs sum throughput under this heuristic, given $\eta_1 = 0.05$.

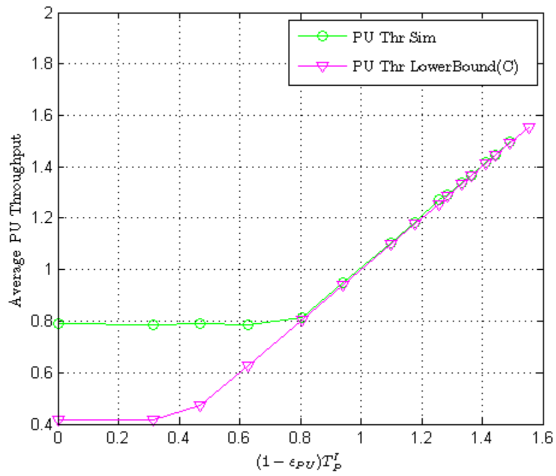


Figure 5.39: $H_3(\eta_1 = 0.05)$: Average PU throughput with respect to PU throughput constraint

Fig. 5.39 depicts the average PU throughput with respect to the PU throughput constraint for varying ϵ_{PU}. As in previous cases, the increase of ϵ_{PU} leads to a more significant PU throughput degradation.

Fig. 5.40 offers a comparison of the performance given by heuristic H_3 in the three different cases considered. Since SU_1 transmission probability, η_1, is fixed, it creates a constant level of interference at the PU receiver. Thus, if η_1 is too high, for example $\eta_1 = 1$, the PU throughput degradation constraint can not be satisfy since the latter is very tight; this is why it is not possible to find a valid (α_2, β_2) couple for ϵ_{PU} under a certain value. On the other side, the fact that SU_1 has a fixed transmission probability affects negatively the SUs sum throughput for high values of ϵ_{PU}; in these cases if $\eta_1 < 1$ SU_1 does not exploit the transmitting chances as much as possible at the expense of the maximum achievable SUs sum throughput.

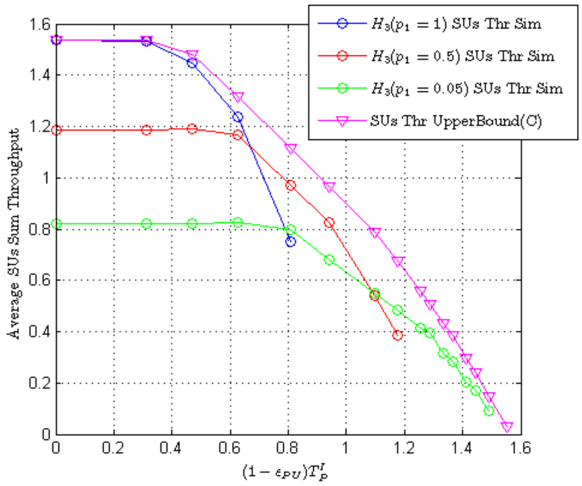


Figure 5.40: H_3 comparison: Average SUs sum throughput with respect to PU throughput constraint

5.4 Heuristic Access Policy H_4

The heuristic access policies proposed so far are offline strategies since, after a research of the optimal transmission probability values, we analyze the performance of a CR network where the SUs select their action according to an established policy strictly depending on the identified optimal transmission probabilities. The last heuristic policy we propose, denoted by H_4, is an online strategy, i.e., the SUs do not follow an established policy, but try to rearrange their transmission probability during the evolution of the system according to the PU feedback they receive and their own PU message knowledge with the aim of maximizing their own throughput under the PU throughput constraint. Since the state of the system deeply influences the network evolution, as in H_1, we consider two different transmission probabilities for each one of the SUs in each time-slot; thus, we have α_i and β_i, $i \in \{1, 2\}$, which are the SU_i probability of accessing the channel when it knows and does not know the PU message, respectively. At the beginning of the simulation, α_i and β_i are initialized randomly with $\alpha_i > \beta_i$ and we suppose they are the same for the two SUs; instead, during the evolution of the system they are not bound to assume the same value. The aim of each SU is to maximize its own throughput under the PU throughput constraint. The primary ARQ feedback gives the SUs the notion of how the system is evolving; in addition, they know the maximum PU throughput, T_p^I, and the constraint they have to satisfy. In each time-slot the SU can check if the PU throughput constraint is satisfied, i.e., if $T_p^I - T_p \leq R_p \epsilon_\omega$, and thus rearrange its own transmission probability to regulate the level of interference at the PU receiver. In other words, if the primary ARQ feedback, t, is equal to 1 and the PU message is successfully decoded at PU_{rx}, each SUs can compute the PU throughput collected until this moment, T_p; if it is too high, i.e., $T_p > T_p^I$, the SUs transmission probabilities are too small so they can increase them by an arbitrary small quantity (in simulation we use an increment of 0.05). On the other hand, if $t > 1$, i.e., the PU is retransmitting, and the PU message is not successfully decoded at PU_{rx}, each SU can check the PU throughput constraint; if it is not satisfied, i.e., $T_p^I - T_p > R_p \epsilon_\omega$, the SUs create too much interference at the PU receiver, so they have to decrease their transmission probability by the same arbitrary small quantity. This kind of approach allows to rearrange the SU access policy according to the evolution of the system.

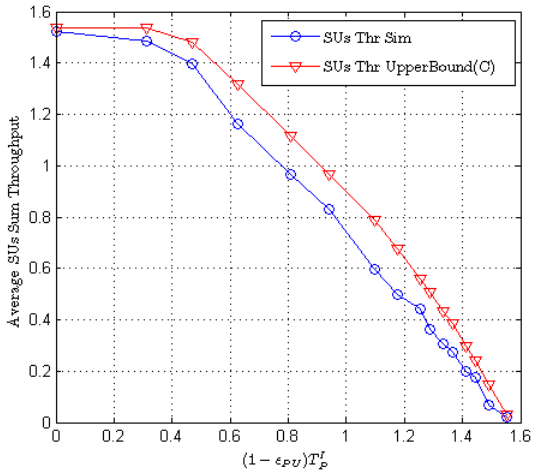
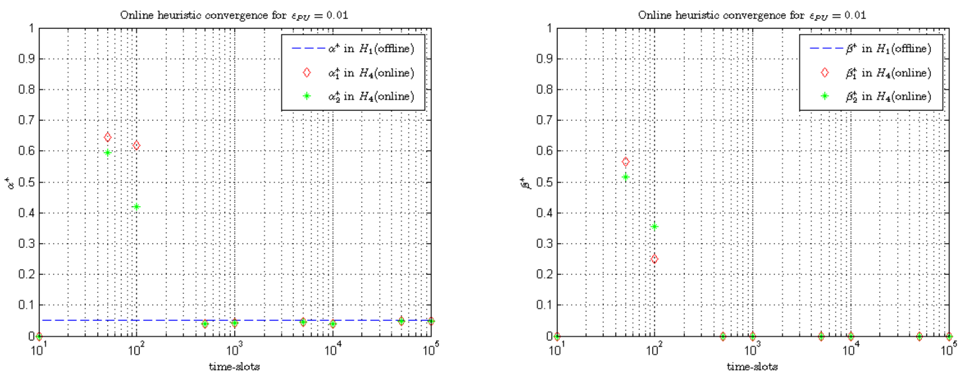
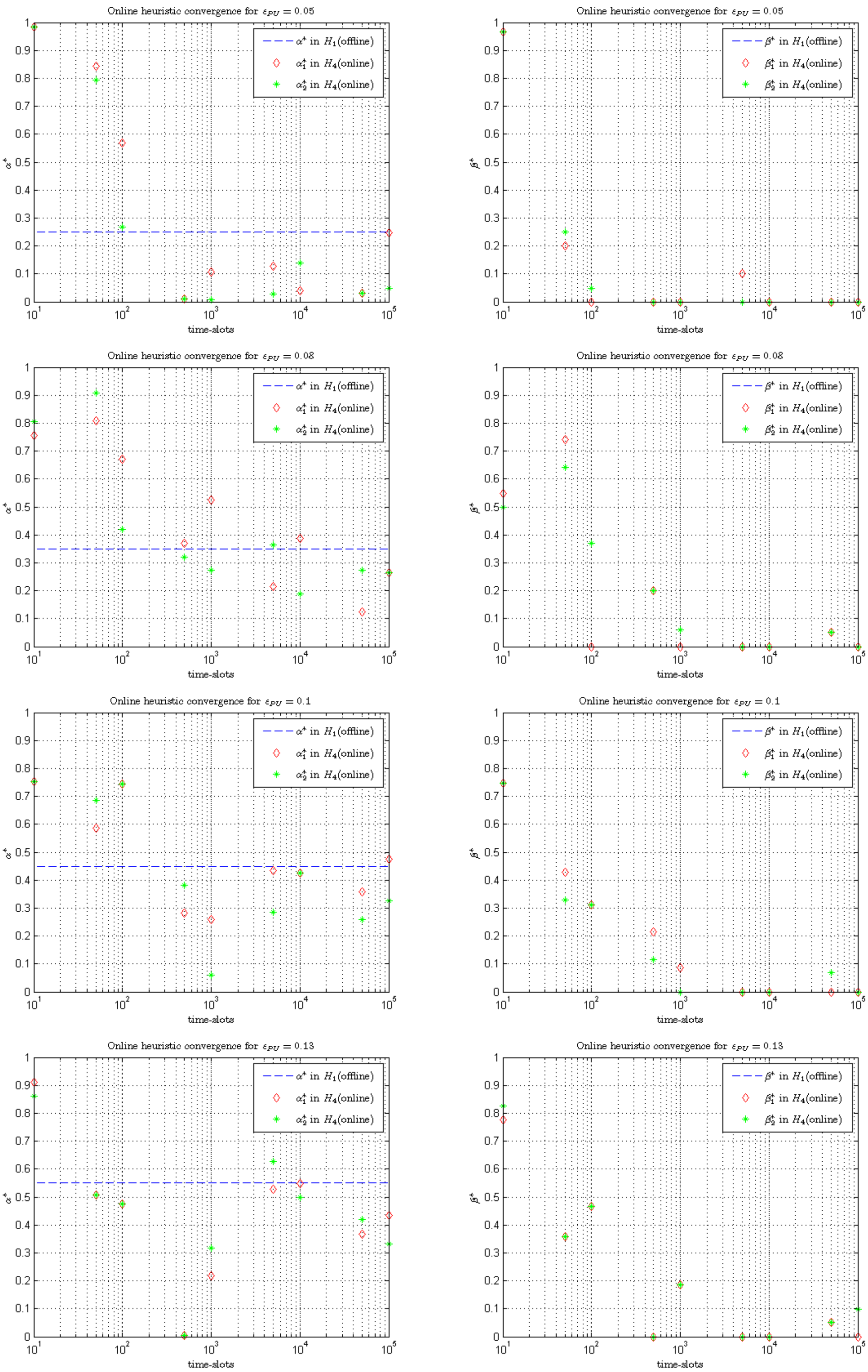
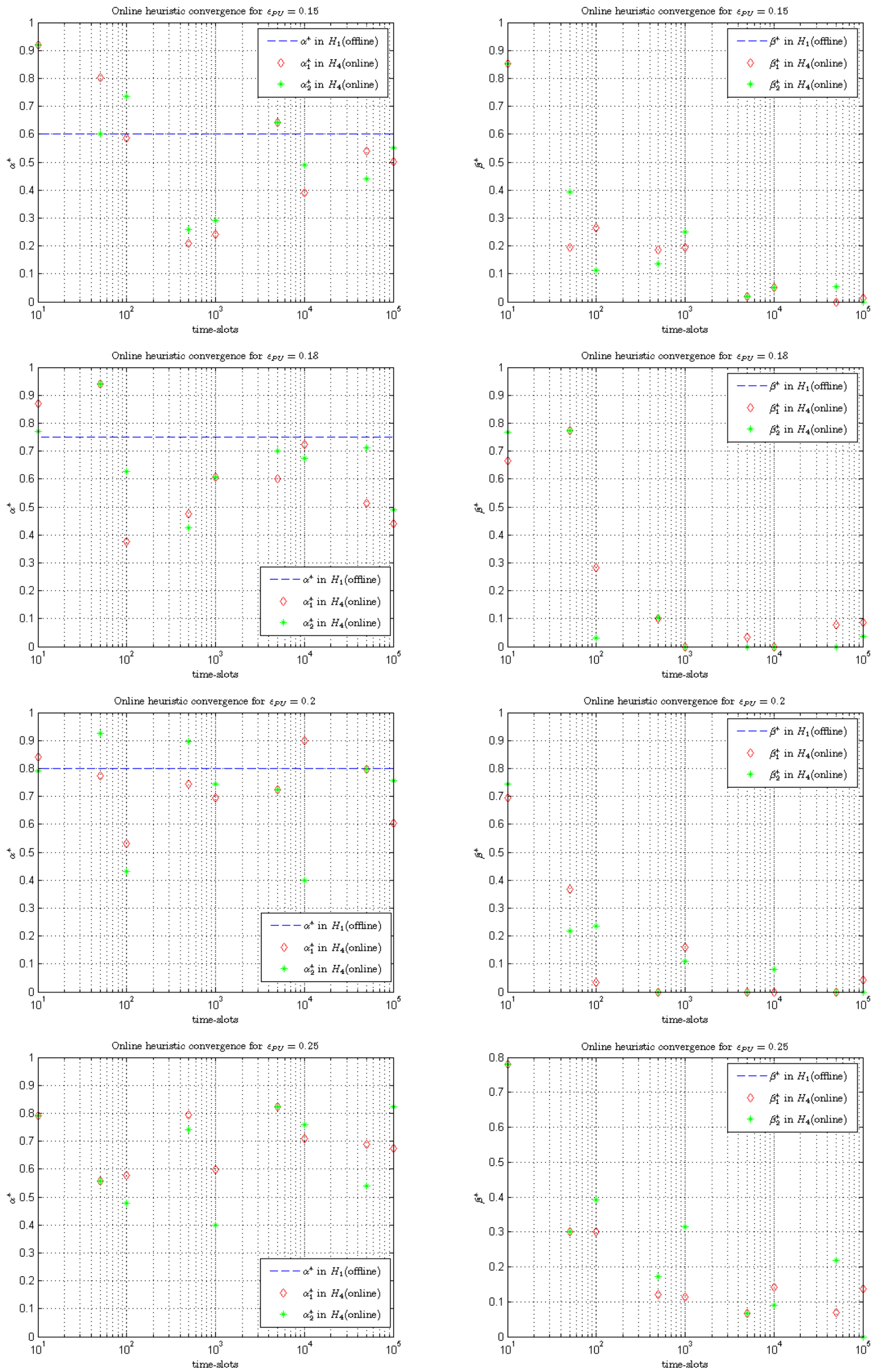


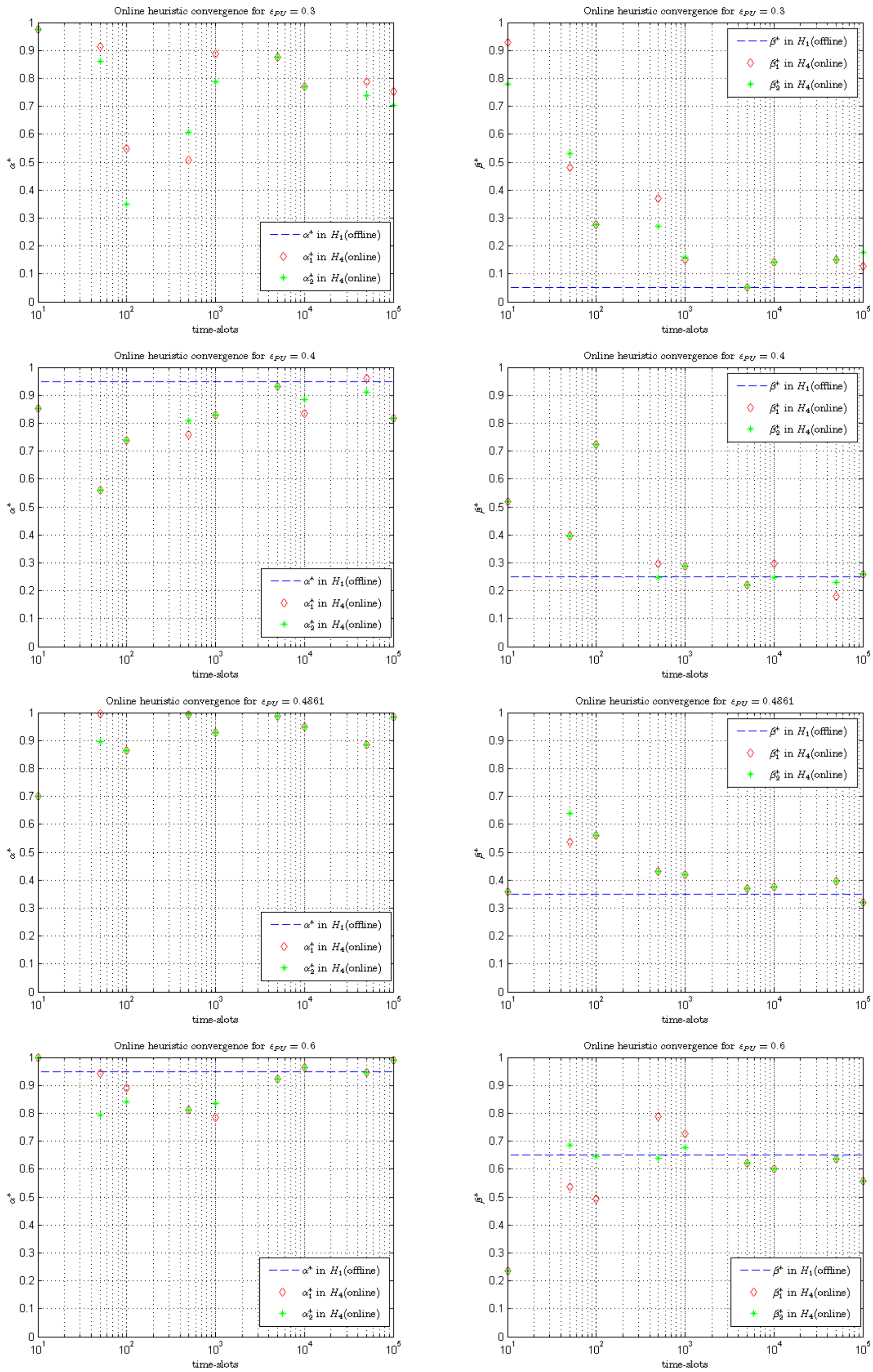
Figure 5.41: H_4: Average SUs sum throughput with respect to PU throughput constraint

Fig. 5.41 depicts the SUs sum throughput with respect to the PU throughput constraint for varying ϵ_{PU} and compared with the upper bound represented by the centralized case (MMDP). Obviously, as the PU throughput increases, the average SUs sum throughput decreases. The gap between the performance of H_4 and the upper bound is not very large; thus, we can consider this heuristic approach as a good approximation of the optimal decentralized case.




Online heuristic convergence for $\epsilon_{PU} = 0.05$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
Online heuristic convergence for $\epsilon_{PU} = 0.08$
Online heuristic convergence for $\epsilon_{PU} = 0.1$
Online heuristic convergence for $\epsilon_{PU} = 0.13$
α^*
β^*
time-slots


Online heuristic convergence for $\epsilon_{PU} = 0.15$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
α^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.15$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
β^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.18$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
α^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.18$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
β^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.2$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
α^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.2$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
β^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.25$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
α^*
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.25$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
β^*
time-slots


Online heuristic convergence for $\epsilon_{PU} = 0.3$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.3$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.4$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.4$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.4861$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.4861$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.6$
α^* in H_1(offline)
α_1^* in H_4(online)
α_2^* in H_4(online)
time-slots
Online heuristic convergence for $\epsilon_{PU} = 0.6$
β^* in H_1(offline)
β_1^* in H_4(online)
β_2^* in H_4(online)
time-slots

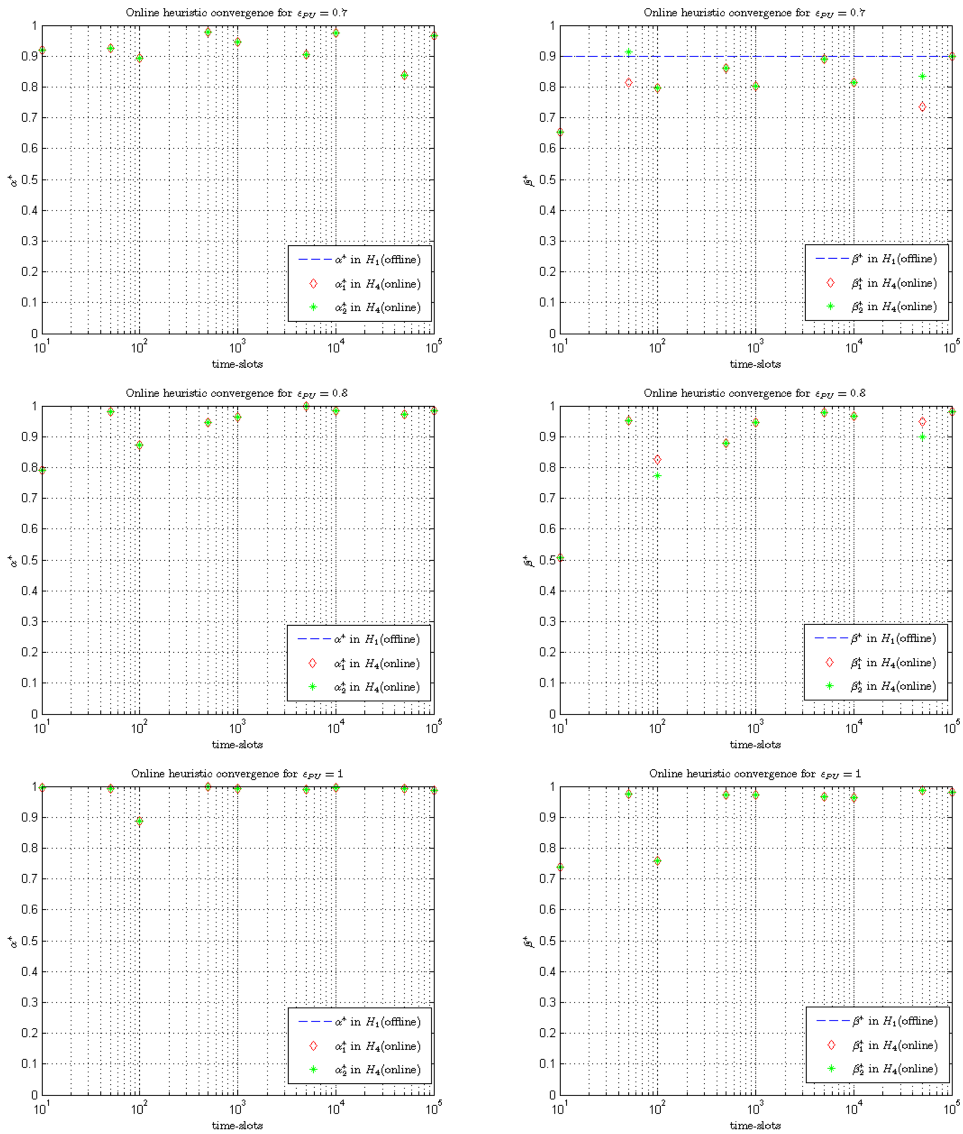


Figure 5.42: H_4: Transmission probabilities convergence

The offline heuristic, H_1, represents the best we can do since after a search among all possible cases we choose the optimal transmission probabilities which maximize the SUs sum throughput under the PU throughput constraint; the online heuristic, H_4, instead is an adaptive algorithm which rearranges the transmission probabilities as the system evolves with the aim to gain the higher reward and satisfy the PU constraint. Thus, H_4 transmission probabilities have to converge to the optimal ones or at least to get them as close as possible.
Fig. 5.42 depicts the evolution of α_1^*, α_2^*, β_1^* and β_2^* compared with the optimal ones for varying ϵ_{PU}. We can note that for small and high values of ϵ_{PU} the transmission probabilities tend to easily converge to the optimal values, while

for intermediate values the convergence is not completely reached. This clarifies why the gap between H_1 and H_4 is larger for intermediate values of ϵ_{PU}; if the SUs adopt transmission probabilities far from the optimal ones the best reward can not be achieved with a consequent degradation of the performance.

As in H_1 case, it is interesting to analyze the effect of the asymmetries in the SNR on the heuristic performance; Fig. 5.43 depicts the average SUs sum throughput with respect to the PU throughput constraint for three specific values of γ_{s2p} and compares them with the symmetric case. It is important to underline that the online heuristic is based on the rearrangement of the SUs transmission probabilities according to the fluctuations of the PU throughput with the aim of maximizing the channel accesses, but under the PU constraint. In all the examined scenarios there is SU_1 which creates a constant average interference level at PU_{rx} and affects the PU performance more or less in dependence of the tightness of the constraint, then there is SU_2 whose SNR changes: for $\gamma_{s2p} = 0.25$ SU_2 creates an interference level almost unperceivable at PU_{rx}, thus it can exploit its transmitting chances more than in the symmetric case, i.e., for $\gamma_{s2p} = 2$, and increase the average SUs reward. Obviously, once it reaches its maximum transmission capability there is a saturation of the SUs throughput. For $\gamma_{s2p} > 2$ the total interference caused by the SUs at PU_{rx} is higher than the symmetric case with a consequent degradation of the performace, in effect as γ_{s2p} increases SU_2 has to limit its accesses to the channel in order to respect the PU constraint. We can note that the performance is almost the same for $\gamma_{s2p} = 5$ and $\gamma_{s2p} = 10$ until $\epsilon_{PU} < 0.6$, then the performance degradation SU_2 causes is more significant as its SNR increases; this is reasonable if we observe Fig. 5.44: it clearly shows that for $\epsilon_{PU} \geq 0.6$ the degradation of PU performance grows as γ_{s2p} increases and thus SU_2 has to rearrange its transmitting behavior in a different way according to the interference it creates at PU_{rx}.

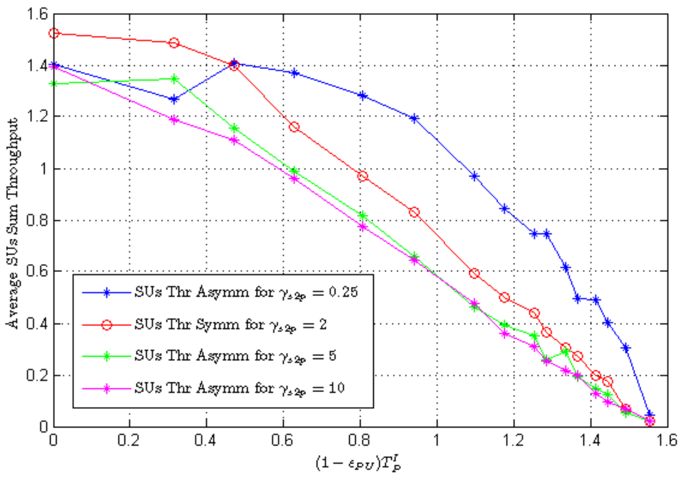


Figure 5.43: $H_{4,ASY}$: Average SUs sum throughput with respect to PU throughput constraint

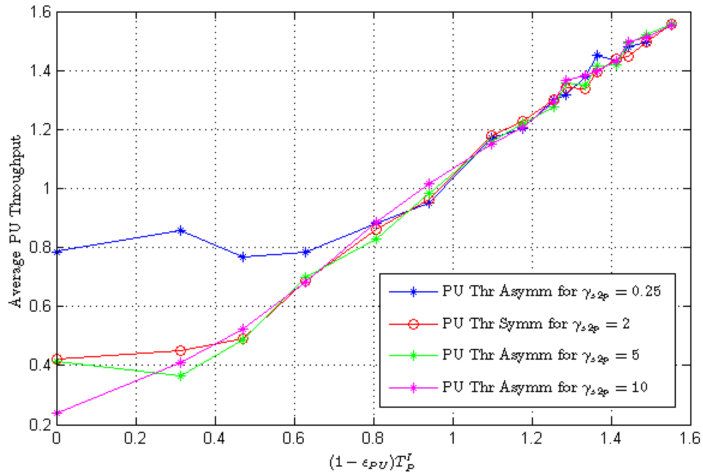


Figure 5.44: $H_{4,ASY}$: Average PU throughput with respect to PU throughput constraint

Fig. 5.45 is very useful to better understand the tradeoff that affects the heuristic performance: it depicts the average long term SUs sum throughput with respect to the average PU throughput in the same asymmetric situations examined in Fig. 5.43.

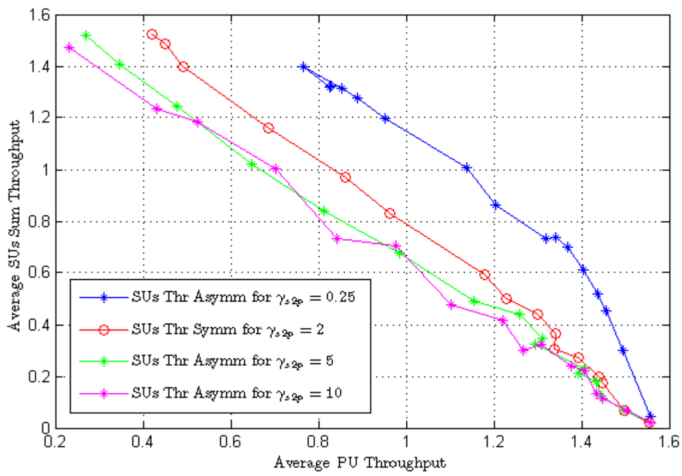


Figure 5.45: $H_{4,ASY}$: Average SUs sum throughput-Average PU throughput tradeoff

As in the offline case, we can note that as the SNR of the interference channel, γ_{s2p}, decreases the performance generally improves; furthermore, it is evident that greater SNR values allow to reach a higher average SUs sum throughput but at price of a more significant degradation of PU performance, i.e., the tradeoff between what we can gain and the cost we have to pay clearly appears, whereas in Fig. 5.43 the PU advantage due to the interference level reduction as γ_{s2p} decreases is not so evident.

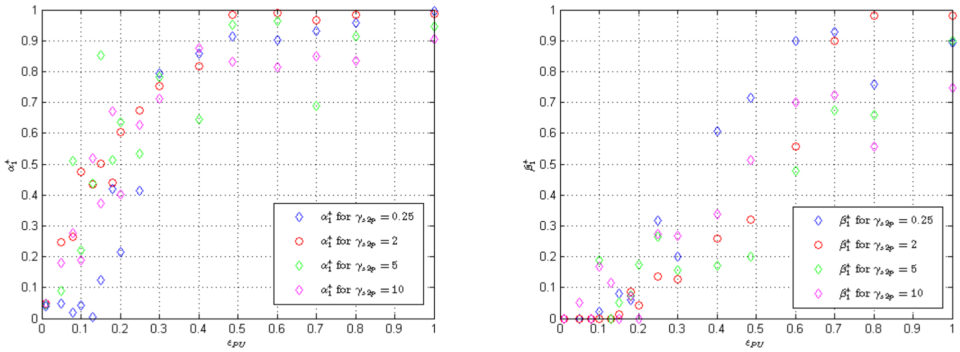


Figure 5.46: $H_{4,ASY}$: α_1^* and β_1^* vs γ_{s2p}

Fig. 5.46 and 5.47 depict SU_1's and SU_2's optimal transmission probabilities, respectively, for the same values of γ_{s2p} we just considered compared with the optimal ones in the symmetric case. As we already underlined in the offline case, for $\gamma_{s2p} = 0.25$ SU_2 adopts a more aggressive behavior than in the symmetric case, in particular for small values of ϵ_{PU} and if it does not know the PU message, whereas SU_1 limits its accesses to the channel; on the other hand, as γ_{s2p} increases the SUs exchange their roles, i.e., SU_2 tends to limit its accesses to the channel as SU_1 exploits as much as possible the transmitting chances. In other word, there is a sort of balance of the SUs behavior: according to its SNR SU_2 adopts a certain trasmitting behavior and SU_1, whose SNR does not change, tries to balance the effects of SU_2's interference at PU_{rx}.

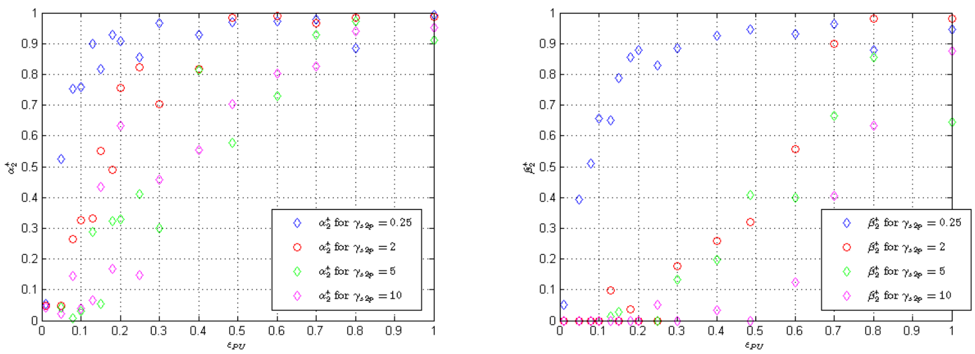


Figure 5.47: $H_{4,ASY}$: α_2^* and β_2^* vs γ_{s2p}

Fig. 5.48 offers a comparison of the performance given by the offline heuristic H_1 and H_2, and the online heuristic H_4.

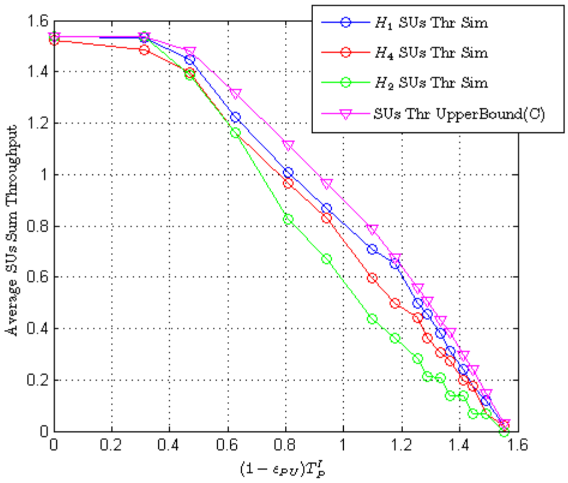


Figure 5.48: $H_1 - H_2 - H_4$ comparison: Average SUs sum throughput with respect to PU throughput constraint

If we consider the two offline heuristic, H_1 and H_2, the best is H_1. This result is reasonable since the access policy proposed in H_1 depends on the PU message knowledge of the SUs and thus exploits the information which the SUs know about the state of the system, while the access policy proposed in H_2 does not exploit the PU message knowledge, which leads to inferior performance in general.

The online heuristic curve, H_4, stays a little below the offline heuristic H_1; in particular, it shows slightly worse performance for intermediate values of ϵ_{PU}. This means that the auto-rearrangement of the SUs transmission probabilities does not allow them to completely exploit the transmitting chances in these cases. In our opinion this is due to the fact that for intermediate ϵ_{PU} the system evolution affects more significantly the SUs behavior, i.e., there is more diversification in SU transmission probabilities than when ϵ_{PU} is high and the SUs try to transmit as much as possible or when ϵ_{PU} is small and the SUs limit their channel accesses as much as possible. In other words for intermediate values of ϵ_{PU} the system possible changes and evolutions are more heterogeneous, thus simulative results are more approximate.

In terms of control informations used by the heuristics, the offline policies only exploit the knowledge of the state of the system, i.e., the primary ARQ feedback and the PU message knowledge, while the online one is more complex, since in each time- slot it has to evaluate the instantaneous PU throughput, T_p, and check if the PU constraint is satisfied.

Chapter 6

Decentralize Access Policies in a Cognitive Radio Network with two SUs and a Partially Observable System

The problem to efficiently model a partially observable decentralized system is a very hard challenge; it can not be reduced to the adaptation of a POMDP model to a decentralized scenario because a lot of factors affect the system evolution and the lack of information is an obstacle very difficult to overcome. As a final contribute to our work we suggest some starting strategy which combined or re-elaborated could offer some interesting idea for a solution to the DEC-POMDP problem.

A first interesting paper is represented by [25] which offers a new framework for learning without state estimation in a partially observable scenario: typically in an MDP the value of the state of the system, s, under a specific policy, π, can be written recursively as follows:

$$V^{\pi}(s) = \sum_{a \in \mathcal{A}} P(a|\pi, s)[R^{a}(s) + \sum_{s' \in \mathcal{S}} P^{a}(s, s') \cdot V^{\pi}(s')]$$

but in a POMDP the value of an observation, x, under a specific policy, π, cannot be defined in a similar form. However, the value of a state in the underlying MDP does not change just because it is inacessible; therefore the value of an observation, x, under a specific policy, π, can be define as follows:

$$V^{\pi}(x) = \sum_{s \in \mathcal{S}} P^{\pi}(s|x) \cdot V^{\pi}(s)$$

where $P^{\pi}(s|x)$ is the asymptotic probability that the state of the underlying

MDP is s when the observation is known to be x and can be defined as:

$$P^{\pi}(s|x) = \frac{P(x|s)P^{\pi}(s)}{P^{\pi}(x)} = \frac{P(x|s)P^{\pi}(s)}{\sum_{s' \in \mathcal{S}} P(x|s')P^{\pi}(s')}$$

Note that the last equation is only a definition of $V^{\pi}(x)$ because the states are not observable in POMDPs. The authors use a Q-learning algorithm applied with a fixed stationary persistent exitation learning policy, i.e. a policy that assigns a non-zero probability to every action in every state and for which the underlying Markov chain is ergodic; in POMDPs that satisfy the assumption that the underlying MDPs are ergodic for every stationary policy all stationary policies with the characteristics just illustrated are persistent exciting. In the paper is demonstrated that in a POMDP of the type just described, if a persistent excitation policy π is followed during learning, the Q-learnimg algorithm will converge to the solution of the following system of equations with probability one: $\forall x \in \mathcal{X}$

$$Q(x,a) = \sum_{s \in \mathcal{S}} P^{\pi_1}(s|x,a)\big[R^a(s) + \gamma \sum_{x' \in \mathcal{X}} P^a(s,x') \cdot \max_{a' \in \mathcal{A}} \{Q(x',a')\}\big] \quad (6.1)$$

where $P^{\pi}(s|x,a)$ is the asymptotic probability, under policy π, that the underlying state is s given that the observation-action pair is (x,a), and $P^a(s,x') = \sum_{s' \in \mathcal{S}} P^a(s,s')P(x'|s')$. Since the scenario we are considering is decentralized, i.e. the SUs are independent and take their action independently to each other on the base of their own observation of the state of the system, our approach will consists on fixing the policy of one SU and finding the optimum policy for the other, and viceversa. A possible way to exploit this approach can be the following: since we aim to find the optimal policy of agent SU_i assuming that policy of agent SU_j is fixed and known to agent SU_i, with $i,j \in \{1,2\}$ and $i \neq j$, we have to use equation (6.1) assuming that $P^{\pi_i}(s|x,a)$ is the asymptotic probability, under policy π_i, that the underlying state is s given that the observation-action pair is (x,a), and $P^a(s,x') = \sum_{s' \in \mathcal{S}} P^a(s,s')P(x'|s')$. To implement equation (6.1) we have to calculate previously the single components; we assume that $a = a_i \in \{0,1\}$, $s = (t, \phi_i, \phi_j)$ with $t \in \{1,...,T\}$ and $\phi_i, \phi_j \in \{U,K\}$, $x = (t, \phi_i)$ is the observation on PU message knowledge of agent SU_i.

The probability that SU_i observes x when the system is in state s, $P(x|s)$, can be calculated as follows:

$$\begin{aligned} P[x_i = (t,U)|s = ((t,U,\phi_j)] &= 1 \\ P[x_i = (t,U)|s = ((t,K,\phi_j)] &= 0 \\ P[x_i = (t,K)|s = ((t,K,\phi_j)] &= 1 \\ P[x_i = (t,K)|s = ((t,U,\phi_j)] &= 0 \end{aligned} \quad (6.2)$$

The instantaneous SUs sum throughput when the system is in state s and SU_i take action a, $R_i^a(s)$, can be computed as follows:

$$\begin{aligned} R_i^a(s) &= \textstyle\sum_{a_j \in \mathcal{A}} R_i(s, a_i, a_j)\pi_j(a_j|x_j) \\ &= \textstyle\sum_{a_j \in \mathcal{A}} R_i(x_i, a_i, a_j)\pi_j(a_j|x_j) \end{aligned} \quad (6.3)$$

where $R_i(s, a_i, a_j) = R_i(x_i, a_i, a_j)$ since in simulations we suppose that there is no interference between SU_i trasmitter and SU_j receiver.
The instantaneous PU throughput degradation when the system is in state s and SU_i take action a, $C_i^a(s)$, can be computed as follows:

$$\begin{aligned} C_i^a(s) &= \textstyle\sum_{a_j \in \mathcal{A}} C_i(s, a_i, a_j)\pi_j(a_j|x_j) \\ &= \textstyle\sum_{a_j \in \mathcal{A}} C_i(x_i, a_i, a_j)\pi_j(a_j|x_j) \end{aligned} \tag{6.4}$$

where $C_i(s, a_i, a_j) = \rho_{p,(a_1,a_2)} - \rho_{p,0} = C_i(x_i, a_i, a_j)$ for the same reason as before.
The transition probability from state s to s' when SU_i take action a, $P^a(s, s') = P(s'|s, a)$, can be calculated as follows:

$$P^{a_i}(s, s') = P(s'|s, a_i) = \sum_{a_j \in \mathcal{A}} P(s'|s, a_i, a_j)\pi_j(a_j|x_j) \tag{6.5}$$

The asymptotic probability under policy π_i that the underlying state is s given that the observation-action pair of SU_i is (x, a), $P^{\pi_i}(s|x, a)$, can be compute in the following way:

$$\begin{aligned} P^{\pi_i}(s|x, a) &= P^{\pi_i}(s|x_i, a_i) \\ &= \frac{P(x_i|s,a_i)P(s|a_i)}{\sum_{s' \in \mathcal{S}} P(x_i|s',a_i)P(s'|a_i)} \end{aligned} \tag{6.6}$$

where

$$P(x_i|s, a_i) = P(x_i|s) \tag{6.7}$$

and $P(s|a_i)$ is:

$$\begin{aligned} P(s|a_i) &= \frac{P(a_i|s)P(s)}{\sum_{s' \in \mathcal{S}} P(a_i|s')P(s')} \\ &= \frac{\pi_i(a_i|x_i)P(s)}{\sum_{s' \in \mathcal{S}} \pi_i(a_i|x_i')P(s')} \end{aligned} \tag{6.8}$$

Let be $s = (t, \phi_1, \phi_2)$ the actual state, $s' = (t-1, \phi_1', \phi_2')$ the previous state and $a = (a_1, a_2)$ the actions selected by the two agents in state s'; the probability of being in state s, $P(s)$, can be compute as follows:

$$P(s) = \sum_{s' \in \mathcal{S}} \sum_{a \in \mathcal{A}x\mathcal{A}} P^a(s', s)\pi_1(a_1|x_1')\pi_2(a_2|x_2') \tag{6.9}$$

where $P^a(s', s)$ is the transition probability from state s' to state s under the action pair a.
However, this procedure does not consider any constraints and thus the optimum policies we can obtain are deterministic, even if we fix stochastic starting policies at convergence the optimum policies are deterministic.

A second interesting perspective is represented by paper [28] in which the authors demonstrate that a constrained optimization problem can often be reduced to one with no constraints through the introduction of parameters, called *Lagrange multipliers*. They illustrates as a discrete-time Markovian system with a finite state space, a compact action space and an average reward subject to a

global constraint can be optimized by simple dynamic programming equations and the use of Lagrangian multipliers; moreover, they demonstrate that there is always a stationary optimal policy and that it can be simple (non-randomized) or a mixed policy, i.e. a mix of two non-randomized policies. Therefore, the solution of our problem can be divided in two steps:
1) solve the Lagrangian unconstrained POMDP problem using a suitable policy iteration algorithm;
2) compute the optimal randomized policy as a probabilistic mixture of two pure policies.

The problem we have to solve is the following:

$$\begin{cases} J = \max_\pi \left\{ E^\pi_{s,a} \left[\sum_{s,a} R(s,a) \right] \right\} \\ \text{s.t. } E^\pi_{s,a} \left[\sum_{s,a} C(s,a) \right] \le \epsilon_W \end{cases} \tag{6.10}$$

where $R(s, a)$ is the SUs sum throughput for the state $s = (s_1, s_2)$ and the SUs joint action $a = (a_1, a_2)$, and $C(s, a)$ is the PU throughput degradation in the state $s = (s_1, s_2)$ when the SUs take the joint action $a = (a_1, a_2)$. Since the analyzed scenario is distributed, i.e., agents SU_1 and SU_2 take their action indipendently from each other and have only a partial view of the state of the system, the observation x, we can use the same strategy adopted before, i.e., first we consider the perspective of agent SU_1, so we fix agent SU_2 stochastic policy and try to find the optimum stochastic policy for agent SU_1 with the Lagrangian procedure, then we exchange the role of the two agents and after fixing the policy of agent SU_1 just found we try to find the optimum stochastic policy for agent SU_2 in the same way; henceforward when we refer to action, a, and observation, x, we intend action and observation of the agent SU_i for which we want to find the optimum policy. In the Lagrangian perspective the problem above described can be reformulated as follows:

$$B^\lambda(s, a) = R(s, a) - \lambda C(s, a) \tag{6.11}$$

$$V^\lambda(x, a) = \max_{a\in A}\Big\{ \sum_{s\in S} P^\pi(s|x, a)\Big[B^\lambda(s, a) + \sum_{x'\in X} P^a(s, x') \cdot V^\lambda(x')\Big]\Big\} \tag{6.12}$$

$$\pi^\lambda(x) = \text{argmax}_{a\in A} \left\{ V^\lambda(x, a) \right\} \tag{6.13}$$

where, since each agent does not know the state of the system, s, but only its own perspective, the calculation of V^λ depends on the actual observation of agent SU_i, x. $R(s, a) = R^a(s)$ given in (6.3) and $C(s, a) = C^a(s)$ given in (6.4) are respectively the SUs sum throughput and the PU throughput degradation when the system is in state $s = (s_1, s_2)$ and agent SU_i take action $a = a_i$, for $i \in \{1, 2\}$, and $P^{\pi_i^\lambda}(s|x, a)$ is the asymptotic probability under policy π_i^λ that the underlying state is s given that the observation-action pair is $(x, a) = (x_i, a_i)$. Thus for every possible observation, $x \in X_i$, and SU_i action , $a \in A_i$, we have $V^\lambda(x, a)$ and we can find the deterministic corresponding policy by policy iteration.
Finding a policy iteration algortithm that solve problem (6.10) is anything but

simple, beacuse the update of $V^{\lambda}(x,a)$ has not to depend on the policies to grant convergence. Supposing to have an efficient policy iteration algorithm, at the end of this procedure we have π_i^{λ} optimum associated to a certain λ, thus we can evaluate the corresponding PU throughput degradation:

$$C^{\lambda} = \sum_{x \in X_i} C^{\lambda}(x) \cdot P(x) \tag{6.14}$$

where

$$C^{\lambda}(x) = \sum_{a \in A_i} \Big(\sum_{s \in S} P^{\pi_i^{\lambda}}(s|x,a) \cdot C^a(s) \Big) \pi_i^{\lambda}(a|x) \tag{6.15}$$

and

$$P(x) = \sum_{s \in S} P(x|s) \cdot P(s) \tag{6.16}$$

where $C^a(s)$, $P(x|s)$, $P(s)$ are given in (6.4), (6.2) and (6.9), respectively.

Since V^{λ} and C^{λ} are monotone non-increasing function in λ (as demonstrated in [28]), we can find two different values of λ (λ_1 and λ_2), such that the corrisponding PU throughput degradations (C^{λ_1} and C^{λ_2}) are respectively $C^{\lambda_1} > \epsilon_W$ and $C^{\lambda_2} < \epsilon_W$ in the most strictly way, i.e., λ_1 is the greater value of λ for which $C^{\lambda_1} > \epsilon_W$ and λ_2 is the smaller value of λ for which $C^{\lambda_2} < \epsilon_W$, where ϵ_W is the constraint we have to satisfy in the optimization procedure. The optimum policy we expected to find is an optimal constrained policy in the form:

$$\mu_i^*(x) = q\mu_i^{\lambda_1}(x) + (1-q)\mu_i^{\lambda_2}(x) \quad \forall x \in X_i \tag{6.17}$$

where $\mu_i^{\lambda_1}$ and $\mu_i^{\lambda_2}$ are respectively the trasmission probability for λ_1 and λ_2 associated to the policy obtained by policy iteration, $q \in [0,1]$ and it has to satisfy $C_{\mu_i^*}(x) = \epsilon_W$ $\forall x \in X_i$ that can be calculate as follows:

$$C_{\mu_i^*}(x) = \sum_{a \in A_i} \Big(\sum_{s \in S} P^{\pi_i^*}(s|x,a) \cdot C^a(s) \Big) \pi_i^*(a|x), \quad \forall x \in X_i \tag{6.18}$$

where π_i^* is the SU_i's optimum stochastic policy we aimed to find and is related to μ_i^* simply by

$$\pi_i^*(a|x) = \begin{cases} \pi_i^*(0|x) & = 1 - \mu_i^*(x), \quad \forall x \in X_i \\ \pi_i^*(1|x) & = \mu_i^*(x), \quad \forall x \in X_i \end{cases}$$

The last approach we suggest is the optimal one from the analitical point of view: an optimization procedure based on belief states. A POMDP model can be formally defined by the 6-tuple $\Xi = (\mathcal{S}, \mathcal{A}, \mathcal{Z}, \mathcal{R}, \mathcal{T}, \mathcal{O})$ where: $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action set, $\mathcal{T}$ is the transition probability from one state to another given the action, $\mathcal{Z}$ and $\mathcal{O}$ are the observation set and related probabilities and $\mathcal{R}$ generalizes the COMDP rewards to add a dependency on the observation. As in any MDP our goal is to find the optimal policy and the only way for a policy to specify the truly optimal behavior is to remember the entire history of the process. In a POMDP it is possible to derive a summary statistic for the entire history of the process, called *information state* or *belief state*, that unlike the entire history is of fixed dimension. A belief state is a sufficient statistic for the history, which means that optimal behavior can be achieved using the belief state in place of the history. A belief state, b, is simply a probability distribution over the set of the states, $\Pi(S)$, with $b(s)$ being the probability of occupying state s; thus, we can define $\mathcal{B} = \Pi(S)$ to be the space of all probability distributions over $\mathcal{S}$. A single belief state can capture the relevant aspects of the entire previous history of the process and more importantly can be easily updated after each transition to incorporate one additional step into the history.

A starting point we suggest for a complete formulation is the following: assuming that SU_j has a fixed stochastic policy, π_j, and we want to find SU_i's optimal stochastic policy, with $i, j \in \{1, 2\}$ and $i \neq j$, our state is $s_{MDP} = (t, \phi_i, \phi_j) = (x_i, \phi_j)$ in underlay MDP, where $x_i = (t, \phi_i)$ and $\phi_i, \phi_j \in \Phi = \{U, K\}$. The unknow part of the state is ϕ_j, while the action selected by SU_j given $x_j = (t, \phi_j)$ is known to SU_i, since it knows SU_j's fixed stochastic policy. Hence, in new model, we can use the belief on ϕ_j and our new state is given by $s = (t, \phi_i, Pr(\phi_j|h_i)) = (x_i, b_i(\phi_j))$, where $b_i(\phi_j) = Pr(\phi_j|h_i)$ and h_i is the history of previous actions and observations in SU_i. The new MDP model for SU_i is a tuple $(\dot{S}, \dot{P}, \dot{R}_i, \dot{C}_i)$, where $\dot{S}$ is the state space, $\dot{P}$ is the transition probability matrix, $\dot{R}_i$ is the instantaneous reward function and $\dot{C}_i$ is the instantaneous cost function.

In order to characterize our new model we need to find a formula which given a belief state vector $b_i(\phi_j)$ computes the resulting belief state, $b^{a_i}_{i,x'_i}(\phi'_j)$, after a transition in the process, then we have to compute $\dot{P}$, $\dot{R}_i$, $\dot{C}_i$. Once we have all this elements, since for every unichain Constrained Markov Decision Process there exists an equivalent Linear Programming (LP) formulation, where an MDP is considered unichain if it contains a single recurrent class plus a (perhaps empty) set of transient states. The equivalent LP problem of the problem we want to solve is the following:

$$\max_z \sum_{s\in\dot{\mathcal{S}}} \sum_{a_i\in\mathcal{A}} \dot{R}_i(s, a_i) z(a_i, s) \tag{6.19}$$

$$\text{s.t.} \quad \sum_{s\in\dot{\mathcal{S}}} \sum_{a_i\in\mathcal{A}} \dot{C}_i(s, a_i) z(a_i, s) \leq \epsilon_W \tag{6.20}$$

$$\sum_{a_i\in\mathcal{A}} z(a_i, s') - \sum_{s\in\dot{\mathcal{S}}} \sum_{a_i\in\mathcal{A}} \dot{P}(s'|s, a_i) z(a_i, s) = 0 \quad \forall s' \in \dot{\mathcal{S}} \tag{6.21}$$

$$\sum_{s\in\dot{\mathcal{S}}} \sum_{a_i\in\mathcal{A}} z(a_i, s) = 1 \tag{6.22}$$

$$z(a_i, s) \geq 0 \quad \forall s \in \dot{\mathcal{S}}, \quad a_i \in \mathcal{A} \tag{6.23}$$

(6.24)

The relationship between the optimal solution of LP problem (6.24) and the optimal solution to our problem is obtained as follows:

$$\pi_i(a_i|x_i, b_i) = \begin{cases} \frac{z(a_i,x_i,b_i)}{\sum_{a_i \in \mathcal{A}} z(a_i,x_i,b_i)}, & \text{if } \sum_{a_i \in \mathcal{A}} z(a_i, x_i, b_i) > 0 \\ \text{arbitrary}, & \text{otherwise} \end{cases}$$

The development of an efficient DEC-POMDP formulation based on belief states remain an open research problem that could be investigated in future works.

Chapter 7

Conclusions

Cognitive radio represents an innovative and intelligent way to exploit the frequency spectrum and offers a lot of idea for interesting future research work. In this thesis work we focus on the decentralized access policies design in a CR network with one PU and two independent SUs under a primary ARQ process; exploiting FIC we extend previous research works which considered only the centralized scenario and/or a network with only one secondary agent. The decentralized scenario is very actual and recurrent in many applications and real contests, thus our work can be considered as a little stone in the wall of future reasearch on cognitive radio network. Cognitive radio represents a very hard challenge for complexity, lack of information, problems in system modelling and expecially in analytical formulation; for these reason we concentrate our efforts on the decentralized heuristic policies design. We suggest some interesting offline and online access strategies and we point out that in such a scenario with independent secondary users characterized by the same transmission parameters a symmetric approach gives good results very closed to the optimal ones. Since the only partial knowledge of the state of the system represents a crucial element which affects significantly the SUs performance, we show how important could be exploiting this information in the development of a good channel access strategy. Finally, we point out the lack of an efficient analytical formulation of the DEC-POMDP problem and suggest some starting point for future research work, like Q-learning exploitation, Lagrangian multipliers utilization and formulation based on belief states. Cognitive radio represents a continuous challenge and thus a rich inspiration source for future research works: in addition to the development of a correct analytical formulation, another interesting field to investigate is the multiple CR scenario, i.e., to study how the growth in SUs number affects the performance of system and thus design suitable access policies which let to efficiently employ the spectrum bands.

Appendix A

MATLAB Code

In this appendix we report some fragment of the Matlab code developed to implement the access policies proposed.

A.1 DecMmdp.m

```
% ..Transmission paramiters initialization..
% T = maximum number of primary transmission
% Ns = cardinality of states space
% Na = number of possible actions
% gammaP = max power of PU
% gammaS1max = max power of SU1
% gammaS1min = min power of SU1
% gammaS2max = max power of SU2
% gammaS2min = min power of SU2
% alphaBarS1S2 = ch. coeff. from SU1 to SU2
% alphaBarS2S1 = ch. coeff. from SU2 to SU1
% alphaBarS1S1 = ch. coeff. from SU1 to SU1
% alphaBarS2S2 = ch. coeff. from SU2 to SU2
% alphaBarS1P = ch. coeff. from SU1 to PU
% alphaBarS2P = ch. coeff. from SU2 to PU
% alphaBarPS1 = ch. coeff. from PU to SU1
% alphaBarPS2 = ch. coeff. from PU to SU2
% alphaBarPP = ch. coeff. from PU to PU
% gammaPS1 = average SNR on ch. PUtx->SU1rx;
% gammaPS2 = average SNR on ch. PUtx->SU2rx;
% gammaPP = average SNR on ch. PUtx->PUrx;
% R_P = PU tx rate
% gammaS1 = power of SU1
% gammaS2 = power of SU2
% R_SK1 = SU1 tx rate if PU msg known
% R_SK2 = SU2 tx rate if PU msg known
% R_SU1 = SU1 tx rate if PU msg unknown
% R_SU2 = SU2 tx rate if PU msg unknown
% rhoP = PU outage on ch. PUtx->PUrx
```

```
% rhoPS1 = PU outage on ch. PUtx->SU1rx
% rhoPS2 = PU outage on ch. PUtx->SU2rx
% rhoS1K = SU1 outage when PU msg known
% rhoS2K = SU2 outage when PU msg known
% rhoS1U = SU1 outage when PU msg unknown
% rhoS2U = SU2 outage when PU msg unknown

...

EpsPU = [.01 .05 .08 .1 .13 .15 .18 .2 .25 .3 .4 .4861 .6 .7 .8 1];
Nsim = length(EpsPU);
% Optimization procedure..
for ns = 1:Nsim
fprintf('Optimization for epsPU=%f..\n',EpsPU(ns));
epsPU = EpsPU(ns);
epsW = (1-rhoP(1,1))*epsPU;  % PU thr degradation constraint
...
ThrSmax = 0;
ThrPmin = 0;
pi1opt = zeros(Na,Ns);
pi2opt = zeros(Na,Ns);
Nit = 1000;
for iter = 1:Nit
 if mod(iter,1000)==0
   fprintf('Simulation #%d..\n',iter);
 end
 ThrS = 0;
 ThrP = 0;
% initialization of SU1's policy
 pi1 = zeros(Na,Ns);
 mu1 = zeros(1,Ns);
% initialization of SU2's fixed policy
pi2 = zeros(Na,Ns);
mu2 = zeros(1,Ns);
for s = 1:Ns
    mu2(s) = random('unif',0,1);
end
for s=1:Ns
    for j=1:Na
        if j==1
          pi2(j,s) = 1-mu2(s);
        else
          pi2(j,s) = mu2(s);
        end
    end
end
stop = 0;
pi1_old = zeros(Na,Ns);
pi2_old = zeros(Na,Ns);
round = 1;
```

```
while(stop~=1)
if round==10
   break;
end
found1 = 0;
% LP problem formulation for SU1..
% transition probability from state s to state sprim
% when agent SU1 selects action a1=i (P^{a1}(s,s')=Pr(s'|s,a1))
Ptran1 = zeros(Ns,Ns,Na);
for i = 1:Na
    for s = 1:Ns
        for sprim = 1:Ns
           PR = 0;
           for j = 1:Na
              PR = PR+Ptran(s,sprim,i,j)*pi2(j,s);
           end
           Ptran1(s,sprim,i) = PR;
        end
    end
end
% reward when SU1 is in state s
% and selects action a1=i (R^{a1}(s)=R1(s,a1))
R1 = zeros(Ns,Na);
for i = 1:Na
    for s = 1:Ns
        R1(s,i) = rwd1(s,i,mu2);
    end
end
% cost in terms of PU thr degradation when SU1 is in state s
% and selects action a1=i (C^{a1}(s)=C1(s,a1))
C1 = zeros(Ns,Na);
for i = 1:Na
    for s = 1:Ns
        C1(s,i) = cost1(s,i,mu2);
    end
end
Ts = zeros(Ns*Na,1); % SUs reward
for s = 1:Ns
   for i = 1:Na
      Ts((s-1)*Na+i) = R1(s,i);
   end
end
Tc = zeros(1,Ns*Na); % PU thr degradation(cost)
for s = 1:Ns
   for i = 1:Na
      Tc((s-1)*Na+i) = C1(s,i);
   end
end
DeltaMinusP = zeros(Ns,Ns*Na);
for sprim = 1:Ns
```

```
   for s = 1:Ns
      delta = 1*(s==sprim)+0*(s~=sprim);
      for i = 1:Na
         DeltaMinusP(sprim,(s-1)*Na+i) = DeltaMinusP(sprim,(s-1)*Na+i)
         +(delta-Ptran1(s,sprim,i));
      end
   end
end
sumZ = zeros(1,Ns*Na);
for s = 1:Ns
   for i = 1:Na
      sumZ(1,(s-1)*Na+i) = 1;
   end
end
Aeq=[DeltaMinusP;sumZ];
beq=[zeros(Ns,1);1];
lb=zeros(Ns*Na,1);
options=optimset('LargeScale','on','Simplex','off','Display','off');
[z,fval,exitflag,output,lambda]=linprog(-1*Ts,Tc,epsW,Aeq,beq,lb,[],[],options);
if length(z)==0
   break;
end
pi1_old = pi1;
for s = 1:Ns
    sum = 0;
    for i = 1:Na
        sum = sum+z((s-1)*Na+i);
    end
    if sum>0
       for i = 1:Na
           pi1(i,s) = z((s-1)*Na+i)/sum;
           if i==2
              mu1(s) = pi1(i,s);
           end
       end
    else
       pi1(1,s) = random('unif',0,1);
       pi1(2,s) = 1-pi1(1,s);
       mu1(s) = pi1(2,s);
    end
end
if exitflag==1
   found1 = 1;
elseif exitflag==-2
     fprintf('No feasible solution found!\n');
end
found2 = 0;
% ..LP problem formulation for SU2 identical to SU1 case..
if exitflag==1
   found2 = 1;
```

```
elseif exitflag==-2
     fprintf('No feasible solution found!\n');
end
% check stopping condition..
eq1 = 0;
eq2 = 0;
for s = 1:Ns
    if abs(pi1_old(1,s)-pi1(1,s))<0.001
        eq1 = eq1+1;
    end
    if abs(pi2_old(1,s)-pi2(1,s))<0.001
        eq2 = eq2+1;
    end
end
if (eq1==Ns) && (eq2==Ns)
    if found1==1 && found2==1
      % Average SUs sum throughput calculation:
      [ThrS,ThrP] = thr(pi1,pi2,Ptran);
      stop = 1;
       if ThrS>ThrSmax
          ThrSmax = ThrS;
          ThrPmin = ThrP;
          pi1opt = pi1;
          pi2opt = pi2;
       end
    else
        break;
    end
else
    round = round+1;
end
end
end
...
end

end
```

A.2 System evolution in CR network simulator

```
% ..Transmission paramiters initialization (as in DecMmdp.m)..

...

% Rp = PU tx rate
% R1 = SU1 tx rate
% R2 = SU2 tx rate
% a1 = SU1 action in current time-slot
% a2 = SU2 action in current time-slot
```

```
% fpp = fading coeff. on ch. PUtx->PUrx
% fps1 = fading coeff. on ch. PUtx->SU1rx
% fps2 = fading coeff. on ch. PUtx->SU2rx
% fs1p = fading coeff. on ch. SU1tx->PUrx
% fs2p = fading coeff. on ch. SU2tx->PUrx
% fs1s1 = fading coeff. on ch. SU1tx->SU1rx
% fs1s2 = fading coeff. on ch. SU1tx->SU2rx
% fs2s2 = fading coeff. on ch. SU2tx->SU2rx
% fs2s1 = fading coeff. on ch. SU2tx->SU1rx
% t = primary ARQ feedback
% S1k = SU1's knowledge of PU msg
% S2k = SU2's knowledge of PU msg

if a1==0 && a2==0  % only PU tx
  snrPP = gammaPP*fpp;
  snrPS1 = gammaPS1*fps1;
  snrPS2 = gammaPS2*fps2;
  if Rp<=log2(1+snrPP) % PU success
    succPU = succPU+1;
    t = 1;
    S1k = 0;
    S2k = 0;
    else
      if t==T
        t = 1;
        S1k = 0;
        S2k = 0;
      else
        t = t+1;
        if S1k==0
          if Rp<=log2(1+snrPS1)
            S1k = 1;
          end
        end
        if S2k==0
          if Rp<=log2(1+snrPS2)
            S2k = 1;
          end
        end
      end
    end
elseif a1==1 && a2==0  % PU and only SU1 tx
      % SU1 tx success/insuccess
      if S1k==0
        R1 = R_SU1(a1+1,a2+1);
        snrPS1 = gammaPS1*fps1;
        cP = (Rp<=log2(1+snrPS1));
        if cP==1
          snrS1S1 = gammaS1S1(a1+1)*fs1s1;
          cS1 = (R1<=log2(1+snrS1S1));
```

```
    cPS1 = ((R1+Rp)<=log2(1+snrS1S1+snrPS1));
    if cS1==1 && cPS1==1 % SU1 success;
      succSU1 = succSU1+1;
      thrS1 = thrS1+R1;
      S1k = 1;
    end
  elseif cP==0
    snrS1S1 = (gammaS1S1(a1+1)*fs1s1)/(1+gammaPS1*fps1);
    cS1 = (R1<=log2(1+snrS1S1));
    if cS1==1 % SU1 success
      succSU1 = succSU1+1;
      thrS1 = thrS1+R1;
    end
  end
  if S1k==0
    snrS1S1 = gammaS1S1(a1+1)*fs1s1;
    cS1 = (R1<=log2(1+snrS1S1));
    if cS1==0
      snrPS1 = (gammaPS1*fps1)/(1+gammaS1S1(a1+1)*fs1s1);
      cP = (Rp<=log2(1+snrPS1));
      if cP==1
        S1k = 1;
      end
    end
   end
elseif S1k==1
   R1 = R_SK1(a1+1,a2+1);
   snrS1S1 = gammaS1S1(a1+1)*fs1s1;
   if R1<=log2(1+snrS1S1) % SU1 success
     succSU1 = succSU1+1;
     thrS1 = thrS1+R1;
   end
end
% PU tx success/insuccess
snrPP = (gammaPP*fpp)/(1+gammaS1P(a1+1)*fs1p);
snrPS2  = (gammaPS2*fps2)/(1+gammaS1S2(a1+1)*fs1s2);
if Rp<=log2(1+snrPP) % PU success
  succPU = succPU+1;
  t = 1;
  S1k = 0;
  S2k = 0;
else
  if t==T
    t = 1;
    S1k = 0;
    S2k = 0;
  else
    t = t+1;
    if S2k==0
      if Rp<=log2(1+snrPS2)
```

```
            S2k = 1;
          end
        end
      end
    end
elseif a1==0 && a2==1  % PU and only SU2 tx
    % SU2 tx success/insuccess
    if S2k==0
      R2 = R_SU2(a1+1,a2+1);
      snrPS2 = gammaPS2*fps2;
      cP = (Rp<=log2(1+snrPS2));
      if cP==1
        snrS2S2 = gammaS2S2(a2+1)*fs2s2;
        cS2 = (R2<=log2(1+snrS2S2));
        cPS2 = ((R2+Rp)<=log2(1+snrS2S2+snrPS2));
        if cS2==1 && cPS2==1 % SU2 success
          succSU2 = succSU2+1;
          thrS2 = thrS2+R2;
          S2k = 1;
        end
      elseif cP==0
         snrS2S2 = (gammaS2S2(a2+1)*fs2s2)/(1+gammaPS2*fps2);
         cS2 = (R2<=log2(1+snrS2S2));
         if cS2==1 % SU2 success
           succSU2 = succSU2+1;
           thrS2 = thrS2+R2;
         end
      end
      if S2k==0
        cS2 = (R2<=log2(1+gammaS2S2(a2+1)*fs2s2));
        if cS2==0
          snrPS2 = (gammaPS2*fps2)/(1+gammaS2S2(a2+1)*fs2s2);
          cP = (Rp<=log2(1+snrPS2));
          if cP==1
            S2k = 1;
          end
        end
      end
    elseif S2k==1
        R2 = R_SK2(a1+1,a2+1);
        snrS2S2 = gammaS2S2(a2+1)*fs2s2;
        if R2<=log2(1+snrS2S2) % SU2 success
          succSU2 = succSU2+1;
          thrS2 = thrS2+R2;
        end
    end
    % PU tx success/insuccess
    snrPP = (gammaPP*fpp)/(1+gammaS2P(a2+1)*fs2p);
    snrPS1 = (gammaPS1*fps1)/(1+gammaS2S1(a2+1)*fs2s1);
    if Rp<=log2(1+snrPP) % PU success
```

```
        succPU = succPU+1;
        t = 1;
        S1k = 0;
        S2k = 0;
        if t==T
          t = 1;
          S1k = 0;
          S2k = 0;
        else
          t = t+1;
          if S1k==0
            if Rp<=log2(1+snrPS1)
              S1k = 1;
            end
          end
        end
      end
elseif a1==1 && a2==1  % PU and both SUs tx
      % SU1 tx success/insuccess
      if S1k==0
        R1 = R_SU1(a1+1,a2+1);
        snrPS1 = (gammaPS1*fps1)/(1+gammaS2S1(a2+1)*fs2s1);
        cP = (Rp<=log2(1+snrPS1));
        if cP==1
          snrS1S1 = (gammaS1S1(a1+1)*fs1s1)/(1+gammaS2S1(a2+1)*fs2s1);
          cS1 = (R1<=log2(1+snrS1S1));
          cPS1 = ((R1+Rp)<=log2(1+snrPS1+snrS1S1));
          if cS1==1 && cPS1==1 % SU1 success
            succSU1 = succSU1+1;
            thrS1 = thrS1+R1;
            S1k = 1;
          end
        else
          snrS1S1 = (gammaS1S1(a1+1)*fs1s1)/(1+gammaPS1*fps1+gammaS2S1(a2+1)*fs2s1);
          cS1 = (R1<=log2(1+snrS1S1));
          if cS1==1 % SU1 success
            succSU1 = succSU1+1;
            thrS1 = thrS1+R1;
          end
        end
        if S1k==0
          snrS1S1 = (gammaS1S1(a1+1)*fs1s1)/(1+gammaS2S1(a2+1)*fs2s1);
          cS1 = (R1<=log2(1+snrS1S1));
          if cS1==0
            snrPS1 = (gammaPS1*fps1)/(1+gammaS1S1(a1+1)*fs1s1+gammaS2S1(a2+1)*fs2s1);
            cP = (Rp<=log2(1+snrPS1));
            if cP==1
              S1k = 1;
            end
          end
```

```
  end
elseif S1k==1
   R1 = R_SK1(a1+1,a2+1);
   snrS1S1 = (gammaS1S1(a1+1)*fs1s1)/(1+gammaS2S1(a2+1)*fs2s1);
   if R1<=log2(1+snrS1S1) % SU1 success
     succSU1 = succSU1+1;
     thrS1 = thrS1+R1;
   end
end
% SU2 tx success/insuccess
if S2k==0
  R2 = R_SU2(a1+1,a2+1);
  snrPS2 = (gammaPS2*fps2)/(1+gammaS1S2(a1+1)*fs1s2);
  cP = (Rp<=log2(1+snrPS2));
  if cP==1
    snrS2S2 = (gammaS2S2(a2+1)*fs2s2)/(1+gammaS1S2(a1+1)*fs1s2);
    cS2 = (R2<=log2(1+snrS2S2));
    cPS2 = ((R2+Rp)<=log2(1+snrPS2+snrS2S2));
    if cS2==1 && cPS2==1 % SU2 success
      succSU2 = succSU2+1;
      thrS2 = thrS2+R2;
      S2k = 1;
    end
  else
    snrS2S2 = (gammaS2S2(a2+1)*fs2s2)/(1+gammaPS2*fps2+gammaS1S2(a1+1)*fs1s2);
    cS2 = (R2<=log2(1+snrS2S2));
    if cS2==1 % SU2 success
      succSU2 = succSU2+1;
      thrS2 = thrS2+R2;
    end
  end
  if S2k==0
    snrS2S2 = (gammaS2S2(a2+1)*fs2s2)/(1+gammaS1S2(a1+1)*fs1s2);
    cS2 = (R2<=log2(1+snrS2S2));
    if cS2==0
      snrPS2 = (gammaPS2*fps2)/(1+gammaS2S2(a2+1)*fs2s2+gammaS1S2(a1+1)*fs1s2);
      cP = (Rp<=log2(1+snrPS2));
      if cP==1
        S2k = 1;
      end
    end
  end
elseif S2k==1
  R2 = R_SK2(a1+1,a2+1);
  snrS2S2 = (gammaS2S2(a2+1)*fs2s2)/(1+gammaS1S2(a1+1)*fs1s2);
  if R2<=log2(1+snrS2S2) % SU2 success
    succSU2 = succSU2+1;
    thrS2 = thrS2+R2;
  end
end
```

```
      % PU tx success/insuccess
      snrPP = (gammaPP*fpp)/(1+gammaS1P(a1+1)*fs1p+gammaS2P(a2+1)*fs2p);
      if Rp<=log2(1+snrPP) % PU success
        succPU = succPU+1;
        t = 1;
        S1k = 0;
        S2k = 0;
      else
        if t==T
          t = 1;
          S1k = 0;
          S2k = 0;
        else
          t = t+1;
        end
      end
end

...
```

A.3 H_4: transmission probabilities rearrangement

```
% ..Transmission paramiters initialization (as in DecMmdp.m)..

...

% alpha1 = SU1 tx prob when PU msg known
% alpha2 = SU2 tx prob when PU msg known
% beta1 = SU1 tx prob when PU msg unknown
% beta2 = SU2 tx prob when PU msg unknown
% Rp = PU tx rate
% t = primary ARQ feedback
% S1k = SU1's knowledge of PU msg
% S2k = SU2's knowledge of PU msg
% phi1 = S1k in current time-slot
% phi2 = S2k in current time-slot

if Rp<=log2(1+snrPP) % PU success
  succPU = succPU+1;
  % E4 tx prob. updating
  if t==1
    Tp = (Rp*succPU)/n;
    if Tp>TpI  % too much PU success
      if phi1==1
        if alpha1<0.96
          alpha1 = alpha1+0.05;
        else
          alpha1 = 1;
        end
```

```
      else
        if beta1<0.95
          if (beta1+0.05)<alpha1
             beta1 = beta1+0.05;
          end
        else
          if alpha1==1
            beta1 = 1;
          end
        end
      end
      if phi2==1
        if alpha2<0.96
          alpha2 = alpha2+0.05;
        else
          alpha2 = 1;
        end
      else
        if beta2<0.95
          if (beta2+0.05)<alpha2
             beta2 = beta2+0.05;
          end
        else
          if alpha2==1
            beta2 = 1;
          end
        end
      end
    end
  end
  t = 1;
  S1k = 0;
  S2k = 0;
else
  % E4 tx prob. updating
  if t>1
    Tp = (Rp*succPU)/n;
    if (TpI-Tp)>Rp*epsW  % too much SUs interference
       if phi1==1
         if (alpha1>0.05) && ((alpha1-0.05)>beta1)
            alpha1 = alpha1-0.05;
         end
       else
         if beta1>0.05
           beta1 = beta1-0.05;
         else
           beta1 = 0;
         end
       end
       if phi2==1
```

```
        if (alpha2>0.05) && ((alpha2-0.05)>beta2)
           alpha2 = alpha2-0.05;
        end
      else
        if beta2>0.05
          beta2 = beta2-0.05;
        else
          beta2 = 0;
        end
      end
    end
  end
  if t==T
    t = 1;
    S1k = 0;
    S2k = 0;
  else
    t = t+1;
    if S1k==0
      if Rp<=log2(1+snrPS1)
        S1k = 1;
      end
    end
    if S2k==0
      if Rp<=log2(1+snrPS2)
        S2k = 1;
      end
    end
  end
end

...
```

Bibliography

[1] M. Levorato, U. Mitra, and M. Zorzi, *"Cognitive interference management in retransmission-based wireless networks"*, IEEE Trans. Inform. Theory., vol. 58, pp. 3023 – 3046, May 2012.

[2] R. Tannious and A. Nosratinia, *"Cognitive radio protocols based on exploiting hybrid ARQ retransmissions"*, IEEE Trans. Wireless Commun., vol. 9, pp. 2833.2841, Sept. 2010.

[3] N. Michelusi, P. Popovski, O. Simeone, M. Levorato, and M. Zorzi, *"Cognitive access policies under a primary ARQ process via forward-backward interference cancellation"*, IEEE J. Sel. Areas Commun., vol. 31, pp. 2374 – 2386, Nov. 2013.

[4] M. Zorzi and R. Joda, *"Centralized access policy design for two cognitive secondary users under a primary ARQ process"*, IEEE Int. Conf. Commun. Worksh. ICCW , pp. 268 – 273 , June 2014 .

[5] K. G. Shin, H. Kim, A. W. Min, and A. Kumar, *"Cognitive radios for dynamic spectrum access: from concept to reality"*, IEEE Trans. Wireless Commun., vol. 17, pp. 64 – 74, Dec. 2010.

[6] I. Akyildiz, W. Y. Lee, M. Vuran, and S. Mohanty, *"A survey on spectrum management in cognitive radio networks"*, IEEE Commun. Mag., vol. 46, pp. 40 – 48, Apr. 2008.

[7] M. L. Puterman, *"Markov Decision Processes. Discrete Sthochastic Dynamic Programming"*, Wiley, 1994, chapter 8.

[8] R. Tajan, *"Mécamismes de retransmission Hybrid-ARQ en radio-cognitive"*, PhD Thesis, Université de Cergy-Pontoise, 2013.

[9] B. Makki, and T. Eriksson, *"On the AverageRate of HARQ-based Quasi-static Spectrum Sharing Networks"*, IEEE Trans. Wireless Commun., vol. 11, pp. 65 – 77, Jan. 2012.

[10] H.-S. Ahn, and R. Righter, *"Multi-Actor Markov Decision Processes"*, J. Appl. Prob. 42, pp. 15 – 26, 2005.

[11] C. Boutilier, *"Planning, Learning and Coordination in Multiagent Decision Processes"*, Proceedings of the 6th Conference on Theoretical Aspects of Rationality and Knowledge, pp. 195 – 210, 1996.

[12] P. Xuan, V. Lesser, and S. Zilberstein, *"Communication in Multi-agent Markov Decision Processes"*, Technical Report TR-2000-01, Department of Computer Science, University of Massachusetts, 2000.

[13] P. Xuan, and V. Lesser, *"Multi-Agent Policies: from Centralized Ones to Decentralized Ones"*, Proceedings of the 1st International Joint Conference on Autonomous Agents and Multi Agents Systems, pp. 1098 − 1105, 2002.

[14] M. Petrik, and S. Zilberstein, *"Average-Reward Decentralized Markov Decision Processes"*, Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-07), pp.1997 − 2002, 2007.

[15] R. Becker, S. Zilberstein, V. Lesser, and C. V. Goldman, *"Solving Transition Independent Decentralized Markov Decison Processes"*, Journal of Artificial Intelligence Research 22, pp. 423 − 455, 2004.

[16] D. S. Bernstein, S. Zilberstein, and N. Immerman, *"The Complexity of Decentralized Control of Markov Decision Processes"*, Proceedings of the Conference on Uncertainty in Artificial Intelligence, pp. 32 − 37, 2000.

[17] C. Amato, D. S. Bernstein, and s. Zilberstein, *"Optimizing fixed-size stochastic controllers for POMDPs and decentralized POMDPs"*, Journal of Autonomous Agents and Multi Agents Systems, vol. 21, pp. 293 − 310, 2010.

[18] D. S. Bernstein, E. A. Hansen, and S. Zilberstein, *"Bounded Policy Iteration for Decentralized POMDPs"*, Proceedings of the 15th International Conference on Automated Planning and Scheduling, pp. 52 − 57, 2005.

[19] T. Jaakkola, S. P. Singh, and M. I. Jordan, *"Reinforcement Learning Algorithm for Partially Observable Markov Decision Problems"*, Advances in Neural Information Processing Systems 7, MIT Press, Cambridge, MA, pp. 345 − 352, 1995.

[20] S. P. Singh, *"Reinforcement Learning Algorithms for Average-Payoff Markovian Decision Processes"*, Proceedings of the 12th National Conference on Artificial Intelligence, pp. 700 − 704, 1994.

[21] C. Amato, G. Chowdhary, A. Geramifard, N. K. *Ü*re, and M. J. Kochenderfer, *"Decentralized Control of Partially Observable Markov Decision Processes'*, Proceedings of the 52nd IEEE Conference on Decision and Control, pp. 2398 − 2405, Dec. 2013.

[22] J. S. Dibangoye, C. Amato, O. Buffet, and F. Charpillet, *"Optimally Solving Dec-POMDPs as Continuous-State MDPs"*, Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI-13), Aug. 2013.

[23] R. Nair, M. Tambe, M. Yokoo, D. Pynadath, and S. Marsella, *"Taming Decentralized POMDPs: Towards Efficient Policy Computation for Multiagent Settings"*, Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-03), pp. 705 − 711, Aug. 2003.

[24] I. Chades, B. Scherrer, and F. Charpillet, *"A Heuristic Approach for Solving Decentralized-POMDP: Assessment on the Pursuit Problem"*, Proceedings of the 2002 ACM Symposium on Applied Computing (SAC), March 2002.

[25] S. P. Singh, T. Jaakkola, and M. I. Jordan, *"Learning Without State-Estimation in Partially Observable Markovian Decision Processes"*, Proceedings of the 11th International Conference on Machine Learning, pp. 284 – 292, 1994.

[26] D. Kim, J. Lee, K.-E. Kim, and P. Poupart, *"Point-Based Value Iteration for Constrained POMDPs"*, Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-11), July 2011.

[27] R. C. Chen, and G. L. Blankenship, *"Dynamic Programming Equations for Discounted Constrained Stochastic Control"*, IEEE Trans. on Automatic Control, vol. 49, pp. 699 – 709, May 2004.

[28] F. J. Beutler, and K. W. Ross, *"Optimal Policies for Controlled Markov Chains with a Constraint"*, Journal of Mathematical Analysis and Applications 112, pp. 236 – 252, 1985.

[29] D. V. Djonin, and V. Krishnamurthy, *"MIMO Transmission Control in Fading Channels: A Constrained Markov Decision Process Formulation With Monotone Randomized Policies"*, IEEE Trans. on Signal Processing, vol. 55, pp. 5069 – 5083, Oct. 2007.

[30] K. P. Murphy, *"A Survey of POMDP Solution Techniques"*, Sept. 2000.