The Dissertation Committee for Goochul Chung
certifies that this is the approved version of the following dissertation:

Cognitive Radios: Fundamental Limits and Applications to Cellular and Wireless Local Networks

Committee:

Sriram Vishwanath, Supervisor

Sanjay Shakkottai

Jeffrey Andrews

Constantine Caramanis

Jihwan Choi

Cognitive Radios: Fundamental Limits and Applications to Cellular and Wireless Local Networks

by

Goochul Chung, B.S.E.E.; M.S.E.E.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2012

Cognitive Radios: Fundamental Limits and Applications to Cellular and Wireless Local Networks

Publication No. ____________________

Goochul Chung, Ph.D.
The University of Texas at Austin, 2012

Supervisor: Sriram Vishwanath

An ever increasing number of wirelessly-enabled applications places a very high demand on stringent spectral resources. Cognitive radios have the potential of enhancing spectral efficiency by improving the usage of channels that are already licensed for a specific purpose. Research on cognitive radios involves answering questions such as: how can a cognitive radio transmit at a high data rate while maintaining the same quality of service for the licensed user? There are multiple forms of cognition studied in literature, and each of these models must be studied in detail to understand its impact on the overall system performance. Specifically, the information-theoretic capacity of such systems is of great interest. Also, the design of cognitive radio is necessary to achieve those capacities in real applications.

In this dissertation, we formulate different problems that relate to the performance of such systems and methods to increase their efficiency. This

dissertation discusses, firstly, the means of "sensing" in cognitive systems, secondly, the optimal resource allocation algorithms for interweave cognitive radio, and finally, the fundamental limits of partially and overly cognitive overlay systems.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Cognitive Radios and Related work

Due to the spectrum scarcity in a wireless channel, it is hard to establish a new wireless application. Cognitive radio draws a lot of attention because of its ability to alleviate this spectrum scarcity problem by sharing spectrum with legitimate users. A cognitive radio is a transmitter that possesses information about its environment that allows it to adapt and tailor its transmission to maximize network throughput while meeting constraints imposed on it [1]. More specifically, a cognitive radio uses the wireless channel which is licensed to legitimate users without interfering with them by utilizing the cognitive information that it possesses. Spectral efficiency increases, because the performances of legitimate radios in the licensed channel are not affected, and there is an additional transmission of the cognitive radio.

There are multiple notions of cognition in literature [1], [2]. In a very broad sense, cognitive radio can be divided into three different groups which seeks to underlay, interweave, or overlay the cognitive user's signals with the legitimate users' signals in such a way that the legitimate users of the spectrum are as unaffected as possible [3]. We plan to study each of these three classes of cognitive radios, for example, the fundamental limits on the performance of different classes of cognitive radios, means of achieving those practically, and their applications.

In the underlay model, a cognitive radio is allowed to share the channel with the legitimate user, even when there is a legitimate transmission. In [3], this underlay technique is defined to be one where the cognitive radio spreads its signal over the wide range of the spectral band, so that the level of interference to the legitimate user is below the acceptable threshold. The cognitive information for the underlay model is the interference level that the cognitive radio causes to the legitimate receiver [4]. Discussion on how to obtain this information can be found in [5]. Also, resource allocation over the wide range of spectral band is established in [6]. The capacity in the underlay model can be characterized by translating constraint on average receive power into a transmit power constraint at the cognitive transmitter [7], [8], [9].

The interweave cognition technique enables a cognitive radio to exploit the channel only when it is not occupied by a legitimate user so that the transmission of the legitimate users are guaranteed not to experience any interference from the cognitive radios. Thus, the knowledge of the existence of the legitimate transmission represents the cognitive information desired by the interweave cognitive radio. Dynamic radio spectrum sensing, access, and sharing algorithms are needed to increase the data rate in the interweave cognitive radio system [10].

The vast body of literature on cognitive radios addresses multiple issues in studying such radios. One of the main research issues in studying the interweave cognitive radio is obtaining reliable information about the availability of the channel. Since an inaccurate detection leads a cognitive radio to attempt transmitting over the channel in use by the legitimate users believing that it is empty, an accurate information about the existence of the legitimate user is an essential element in the cognitive radio. This information is achieved either from the sensing or from the geographical information. Although spectrum sensing is not strictly required by the standard on the cognitive radio, 802.22 standard (rural area network standard), it is still an important building block for the cognitive radio.

[11] suggests three sensing techniques for a detection of the legitimate radios: matched filter detection, energy detection, and cyclostationarity feature matched detection. Matched filter being the optimal way for any signal detection, it requires a cognitive radio to have a priori knowledge of the legitimate user's signal at both PHY and MAC layers. The process of the matched filter detection is also cumbersome because timing and carrier synchronization and channel equalization are needed there. Energy detection is much considered for use with a cognitive radio, as in [12], [13], and [14]. [12] describes the simplest energy detection of an unknown deterministic signal in AWGN channel. In [13], performance of the energy detection in a multi-path channel is analyzed. A cyclostationarity feature matched detection utilizes the periodicity in signals' statistics in sensing the legitimate signal [15]. There are other detection algorithms which utilize signals' statistics in the sensing to in-

crease the reliability of the sensing [78], [79], [63]. Feature detection is known to have better performance than energy detection in terms of the probability of detection error and false alarm, but requires a longer time to be finished. Thus, it is efficient when the sensing requires more accuracy, for example in the severe fading condition where a false conclusion about the presence of a legitimate transmitter is more prevalent [76]. To increase the reliability of the sensing information even further, additional information can be introduced in the sensing or the sensing can be made collaboratively. In [14], the additional side information is considered in analyzing the performance of cognitive sensing. Side information that the cognitive radio can use includes spatial locations of the cognitive and legitimate receivers, received power of the legitimate signal at the cognitive user, and a priori transmission probability of the legitimate user. With the help of side information, performance improvement can be made. Also, cooperative spectrum sensing in cognitive radio networks has been studied intensely in recent literature [80]-[82]. [80] analyzes cooperative sensing with simple energy detection and establishes combining methods, [81] and [83] design optimal detector for sensing, and [82] studies distance-user tradeoff in correlated fading environment. Given many different types of sensing techniques, it is necessary to find or develop a sensing mechanism suitable for a specific cognitive radio and legitimate network.

Another research issue in interweave cognitive radio is to find its capacity, and to determine the optimal manner in which the resource allocation is performed. In an interweave cognitive radio setting, the limit of cognitive radio's data rate is directly related to the probability of detection error and false alarm. [20], [21], and [22] study the region of operation and the limits in this setting. In [23], parameters in the interweaved cognitive radio setting is optimized to maximize its throughput. However, due to the technological advancement of the sensing algorithm, we assume that sensing is one hundred percent accurate, and focus on finding the optimal resource allocation (power allocation and channel selection) to achieve the maximum throughput. Cognitive radio does not necessarily need to explore one channel at a time. It can sense more than one channel and make a transmission over multiple channels. Also, resource allocation among those channels can vary from one to the other. Thus, throughput of the interweave cognitive radio depends on the selection of the channel to sense and the resource allocation among those channels. The problem of channel selection for cognitive radios is studied in isolation in [43] and [44]. Also, by itself, the resource allocation problem for multi-band radios is studied in [45].

In an overlay cognitive radio setting, cognitive and legitimate radios transmit messages in the same frequency band simultaneously (as in the underlay case). However, the main difference is that, in the overlay case, the cognitive radio has access to information about the legitimate user so as to mitigate network interference and thus increase network throughput [48]. The information-theoretic capacity of cognitive radio in this overlay cognitive radio setting is explored in [48]-[55]. In [48], achievable rate for an overlay cognitive radio is shown. [51] characterizes the capacity region for the class of "strong" interference channel, and [52] and [53] study the capacity region of this channel for "weak" interference channel. There is also a work which studies capacity region of the overlay cognitive radio setting with the degree of freedom perspective [54]. It finds the degree of freedom of the overlay cognitive radio where multiple antennas are deployed. These papers assume perfect and complete information about the legitimate radio's message at the cognitive radio. In the case that the complete cognitive information is not obtainable, [49] and [55] study capacity region of this partially cognitive radio with "strong" and "weak" interference channel respectively. Meanwhile, [56] considers the cognitive radio in the opposite case. A cognitive radio has access to a message that a legitimate transmitter has, and to additional legitimate message that a legitimate transmitter does not have an access to. The capacity of this class of cognitive radio is analyzed.

In the next section, we demonstrate the need for each class of algorithms studied in this thesis.

1.2 Motivation

1.2.1 Interweave Cognitive Radio

As the number of wireless (multimedia) applications increases, so do the stringent requirements they impose on the wireless medium. Thus, it is essential that we determine efficient means of utilizing the limited spectral resources available to us. Currently, bandwidth resources are divided into frequency bands and allocated to different users exclusively in order to ensure the quality of service (QoS) of multiple wireless systems, and the FCC's frequency allocation chart [60] shows that almost all frequency bands are currently allocated to different groups for varying purposes. According to recent surveys [61], most of these allocated radio frequency spectrums are vastly under-utilized by the groups they are given to. The latest spectrum occupancy measurement from SSC (Spectrum Sharing Company) gives more detailed information about occupancy rate [62]. The occupancy rates of frequency bands from 30-2900MHz, which include TV, Air Traffic Control, Amateur, and Unlicensed bands, are measured in 4 different places: Maine, West Virginia, Chicago, and New York City. Spectrum in these bands are shown to be under-utilized. This lowly utilized spectrum and possibility of increasing the spectral efficiency by reusing white spaces in the frequency channel motivates the use of a cognitive radio.

An interweave cognitive radio exploits the channel when it is not utilized by the legitimate user. The idea of spectrum reuse received regulatory support in the form of the FCC white space ruling, authorizing cautious reuse of under-utilized spectrum in the licensed TV bands [63]. In its Notice of Proposed Rulemaking, released in May 2004 [64], and its latest R&O, released in November 2008 [65], the FCC indicates that TV channels 5-13 in the VHF band and 14-51 in the UHF band can be used for fixed broadband access systems. The IEEE 802.22 [66], [67] is a standard which is designed to operate in the TV broadcast bands while ensuring that no harmful interference is caused to the incumbent operation, by formalizing a solution that will meet FCC approval. It aims to use the TV broadcast bands to bring broadband access to a rural areas of typically 17.30 km or more in radius [67]. Also, the possibility of opening up the bands other than TV channels, such as ISM bands, is mentioned in literatures [3],[68], and the advantages of using 3-5GHz channel for cognitive radio is mentioned in [69]. Also, cognitive radio has more possibilities of expanding its application into many other legitimate channels, such as 2.4GHz ISM band and 5GHz channel, where there exists WLAN. FCC's white space ruling, standard activity of 802.22, and the prospects of cognitive radio mentioned in literature, indicate that the concept of cognitive radio in practical use is gaining popularity.

1.2.1.1 Sensing in Interweave Cognitive Radio

Interweave cognitive radio requires sensing, which gives information about availability of the channel. There are stringent requirements for a cognitive radio sensing. First, sensing must be accurate. Poor sensing leads to either detection error, where the cognitive radio acknowledges that the channel is available when it is not, or false alarm, where the sensing indicates that the channel is occupied when it is vacant. Detection error results in undesirable interference to the legitimate user, and false alarm reduces spectrum efficiency. Second, sensing must detect the returning legitimate user quickly. This property, delay of notice, is important in in-band spectrum sensing. In-band spectrum sensing monitors in-band channels, which are the channels that are reused by the cognitive radio. The late notice leads to the delayed evacuation of the cognitive radio, thus the legitimate radio faces longer interference from the cognitive radio.

Different types of sensing are required for different purposes and different applications. If the cognitive radio attempts to enter the licensed channel, it requires a very accurate sensing to avoid the interference to the legitimate users. Meanwhile, less accuracy is required when the cognitive radio is accessing the ISM band. As aforementioned, in-band sensing is required to have short delay of notice [16]. For the secondary cognitive radio system in the primary 802.11 WLAN, the delay of notice is even more important [19]. It is necessary to find or develop a sensing mechanism suitable for a specific cognitive radio and legitimate network. Here, we are interested in the sensing technique which has short delay of notice.

1.2.2 Fundamental Limits of Interweave Cognitive Radio

With reasonably accurate sensing information, it is possible to bring an interweave cognitive radio into practical usage. And, it is important to make the full use of the capacity that the cognitive radio provides. For this reason, it is important to find the fundamental limit of the interweave cognitive radio. By establishing the capacity of cognitive radio, we know the limit of the increased data rate (spectral efficiency) from using the cognitive radio. It gives intuition on how useful the cognitive radio is, and on how to utilize the interweave cognitive radio. To make best use of the interweave cognitive radio, it needs to select channels to use properly, and optimize its power allocation among those channel. For an interweave cognitive radio, it can sense multiple frequency channels, and exploit them. For example, even though 802.22 standard indicates that each channel is sensed independently, it does not preclude an implementation that senses multiple channels simultaneously [67]. When more than one channel are available for transmission, and each channel has dissimilar channel statistics, it is especially difficult to determine channel selection and power allocation over available channels. Our goal is to determine which channels should be sensed when. In addition, we desire to perform a resource-allocation problem across multiple channels which may or may not be available to the cognitive radio. Overall, we ask the question "Given that there are multiple dissimilar channels available for us to sense, which channels should we sense and, if they are available, what rate/power should you assign to them?"

The dissimilarity between different channels arises from various factors. The properties of the propagation environment depend on frequency and thus can be significantly different from channel to channel. Just as any other multi-band radio, the cognitive radio must allocate resources across different bands it uses while simultaneously determining which ones it is permitted to exploit. Note that, in isolation, the problem of channel selection for cognitive radios [43], [44] is well studied. Also, by itself, the resource allocation problem for multi-band radios is well-understood [45]. However, bringing the two together is both important and challenging as they are tightly coupled in the context of interweaved cognitive radios. Therefore, designing channel selection and allocation jointly is essential for cognitive radios. Note that the focus is on the fundamental limits of joint selection and resource allocation in cognitive networks to provide a benchmark on performance. Thus, aspects such as sensing error, delay, device and network non-linearities etc. are not incorporated into the analysis. In [46], joint optimization of the channel selection and power allocation is made to obtain the maximum average throughput.

1.2.3 Overlay Cognitive Radio

An overlay cognitive radio is the most sophisticated form among three cognitive radio classes. It received much attention due to its ability to increase the spectrum efficiency of the cognitive radio by enabling it to transmit simultaneously with legitimate users. In the overlay paradigm, the cognitive transmitter has knowledge of the legitimate users' codebooks and possibly their messages as well. The codebook information can be obtained, for example, if the legitimate users follow a uniform standard for communication based on a publicized codebook. Alternatively, they can broadcast their codebooks periodically [4]. Also, the cognitive transmitter has legitimate users' messages before its transmission. This assumption may hold when the legitimate user retransmits the messages and the cognitive transmitter has knowledge about them from overhearing. Alternatively, the legitimate user may send its message to the cognitive user prior to its transmission [4]. Furthermore, there may exist a backbone network that provides the legitimate user's messages to both legitimate and cognitive transmitters. Note that the amount of the legitimate user's message that the legitimate transmitter and the cognitive transmitter carry can differ for various reasons.

The main application of this overlay cognitive radio in our considerations is the cognitive radio in a cellular system. Consider a cognitive base-station placed within a cellular system (a proprietary band of operation) by the cellular provider. There are multiple legitimate transmitters, whose design can only be minimally changed. However, the design of the cognitive base-station and the receivers (the handhelds) can be changed. This assumption is based on the idea that people purchase new handsets frequently, while changing a legitimate station is a fairly expensive and time-consuming effort. Further, a backbone network exists that enables information transfer to and between base-stations.

As the cognitive base-station can be newly designed, it can enable one-way cooperation, which is the main idea behind overlay cognitive radio capacity analysis. The data to be transmitted by the legitimate radio is made available through the backbone to the cognitive radio. Now, overlay cognition reduces to the analysis of a two-user interference channel with degraded message sets.

1.2.3.1 Fundamental Limits of Partially Overlay Cognitive Radio

Here, we focus on the scenario that the cognitive radio is only partially knowledgable of legitimate user's messages, which can happen because there are not enough resources for the cognitive user to obtain the legitimate user's messages. For example, a resource from the legitimate user to the cognitive user may not be enough, or a backbone network connection to the cognitive transmitter may be inferior to a connection to the legitimate transmitter. As the portion of the messages that the cognitive radio has access to ranges from nothing to everything that the legitimate user has access to, the channel model includes the interference channel (IFC) [72], [74], [75], and IFC with fully-degraded message set [52] as special cases. This channel is referred to as an interference channel with a partially cognitive transmitter. Also, the weak interference channel to the legitimate receiver is assumed, where the interference from the cognitive transmitter to the legitimate receiver is weak compared to the signal strength from the cognitive transmitter to the cognitive receiver. This setting is of practical interest since it models the realistic scenario in which the cognitive transmitter is closer to the cognitive receiver than to the legitimate receiver [53]. This channel model is motivated by practical constraints, where the cognitive transmitter is only able to garner limited information about the legitimate transmitter's message. It is important to know the limit of the rate region of the legitimate and cognitive radios, as well as encoding and decoding strategies which can achieve the rates close to that limit.

1.2.3.2 Fundamental Limits of Overly Overlay Cognitive Radio

On the other hand, the cognitive transmitter may obtain more legitimate user's messages than the legitimate transmitter does. This is possible when the backbone network channel to the cognitive transmitter is superior, and the backbone network is willing to transmit more messages to the cognitive radio strategically. By including this class of overlay cognitive radio, our research on the capacity of overlay cognitive radio is complete. It is necessary to find the capacity of this class of overlay cognitive radio in order to gain knowledge on the fundamental limit of the overlay cognitive radio.

1.3 Contributions

Contributions of the research conducted during my Ph.D. program is summarized as follows.

1.3.1 Sensing for Interweave Cognitive Radio

1. We develop the new sensing mechanisms that are good for the in-band sensing, which requires a short delay in noticing the legitimate transmission. [71]

2. We apply the cognitive radio with new sensing technique in the WiFi network [19].

1.3.2 Fundamental Limits of Interweave Cognitive Radios

1. We verify the optimal power allocation and channel sensing strategy and corresponding capacity with the exact probability of the channel being available.

2. We establish a practical solution to obtain the capacity [55].

1.3.3 Capacity of Overlay Cognitive Radios

1. We characterize the capacity of overlay cognitive radio with partial knowledge in the interference channel by obtaining inner and outer bounds for it. [46]

2. We characterize the capacity of overlay cognitive radio with the additional legitimate messages in the interference channel [56].

1.4 Organization

Each of the following chapters describes the completed work of each research topic. Chapter 2 is on the sensing of the interweave cognitive radio; Chapter 3 deals with the resource allocation of the interweave cognitive radio; Chapter 4 works on the fundamental limits on the overlay cognitive radio with partial information. Chapter 5 solves the problem in the capacity of the overlay cognitive radio with additional information. Finally, Chapter 6 concludes with a summary.

Chapter 2

Sensing in Interweave Cognitive Radios

2.1 Introduction

In this chapter we introduce the new sensing technique which is suitable for in-band sensing. We focus on reducing the delay of notice in the sensing, and this is achieved by bringing the capability of the cognitive radio to cancel the self interference. This new technique is applied to WRAN cognitive radios and the cognitive radio system which shares the channel with WLAN and provide incumbent protection to the WLAN users.

2.1.1 Our Contributions

Our main contributions in this chapter are as follows:

1. We develop a new sensing technique, which is called full duplex cognitive radio.

2. We compare the performance of a full duplex cognitive radio numerically with that of a conventional energy detector.

3. We build a full duplex cognitive radio, and determine its performance.

4. We apply a full duplex cognitive radio to different groups of network.

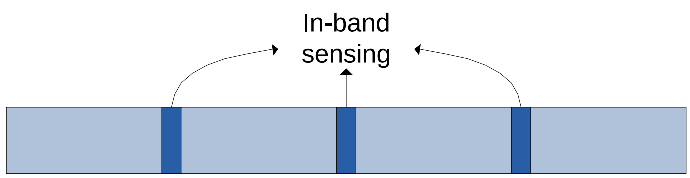


Figure 2.1: In-band sensing

2.2 Conventional Half duplex cognitive radio and its limitation

Figure 2.1 illustrates the operation of the conventional half duplex cognitive radio. It halts its transmission awhile and senses the in-band channel, and this in-band sensing is conducted periodically. Thus, the cognitive radio can empty the channel if the in-band sensing detects the legitimate signal. Figure 2.2 shows the timing for the transmission and in-band sensing of the cognitive radio and entrance of the legitimate radio. The time interval between two consecutive in-band sensings in a half duplex cognitive radio is called sensing interval, and we denote it by $T_{P,half}$. The sensing time for the half duplex radio denoted by $T_{S,half}$ is the time duration taken for in-band sensing. We use the energy detector here because it is simple and its short sensing time is desirable for an in-band sensing. Figure 2.3 shows the operation of the half duplex cognitive radio sensing. The detection ability of an energy detector, which is represented by the probability of detection error and false alarm, is well known from [12]. We obtain the probability of false alarm for the half

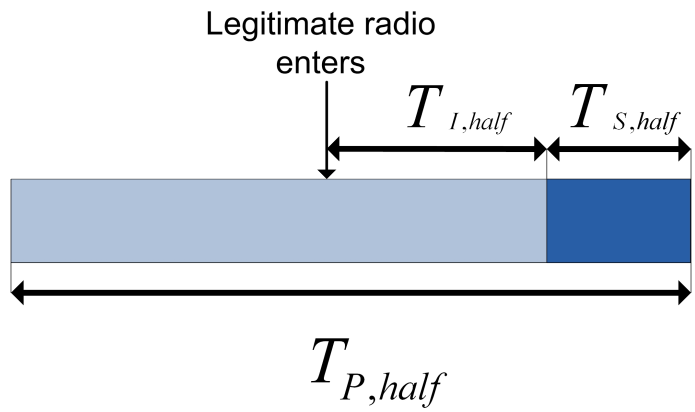


Figure 2.2: Sensing timing

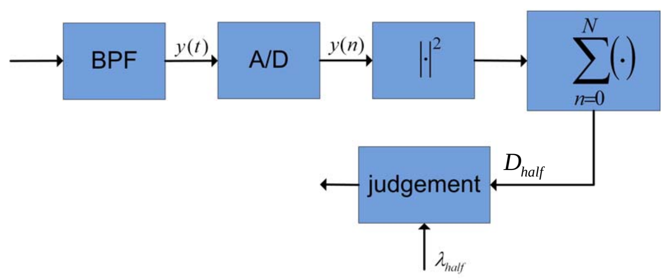


Figure 2.3: Half duplex cognitive radio with energy detection

duplex cognitive radio with an energy detector as follows:

$$P_{f,half} = P(D_{half} > \lambda_{half}|H_0) = \frac{\Gamma\left(T_{S,half}W, \frac{\lambda_{half}}{2}\right)}{\Gamma\left(T_{S,half}W\right)}. \tag{2.1}$$

Here, W is the bandwidth of the legitimate radio's signal, D_{half} is the detection value which is obtained as is shown in Figure 2.3, λ_{half} is the threshold value for a half duplex cognitive radio, and H_0 indicates that there is no legitimate transmission. $\Gamma(\cdot)$ and $\Gamma(\cdot,\cdot)$ are the complete and upper incomplete gamma functions, respectively. The probability of detection error for half duplex cognitive radio, $P_{d,half}$, is given by

$$\begin{aligned} P_{d,half} &= P(D_{half} < \lambda_{half}|H_1) \\ &= 1 - Q_{T_{S,half}W}\left(\sqrt{2\gamma_{half}}, \sqrt{\lambda_{half}}\right) \\ &= e^{-\frac{\gamma_{half}}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\gamma_{half}}{2}\right)^j}{j!} Q\left(\lambda_{half}; 2T_{S,half}W + 2j\right), \end{aligned} \tag{2.2}$$

where H_1 indicates that there exists legitimate radio's transmission, and γ_{half} is the non-central parameter, which is signal energy to noise power ratio. $Q_{T_{S,half}W}(\cdot, \cdot)$ is the generalized $T_{S,half}W$ order Marcum Q function, and $Q\left(x; k\right)$ is the cumulative distribution function of the central chi-squared distribution with k degrees of freedom, which is given by

$$Q\left(x; k\right) = \frac{\gamma\left(\frac{k}{2}, \frac{x}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}. \tag{2.3}$$

Here, $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function. It is known that there is a trade-off between the probability of detection error and the probability of false alarm, which depend on the threshold value, λ_{half}. Also, trade-off curve shrinks as $T_{S,half}W$ increases [24]. Thus, the sensing accuracy of the half duplex cognitive radio with the energy detector increases as the sensing time $T_{S,half}$ becomes larger.

We then calculate the delay of notice of the half duplex cognitive radio. As shown in Figure 2.2, a legitimate radio accesses the channel regardless of the existence of the legitimate radio expecting that the cognitive radio will leave the channel when the cognitive radio detects the signal. However, the cognitive radio does not stop its transmission until the start of the in-band sensing. The delay of notice is also called the interference time, and it is denoted as $T_{I,half}$. The value of it is the time from the entrance of the legitimate radio to the start of the in-band sensing, as is shown in Figure 2.2. We protect incumbent transmission by restricting the maximum interference time:

$$\max(T_{I,half}) = T_{P,half} - T_{S,half} \leq T_{\max}. \tag{2.4}$$

We can surely reduce the maximum interference time for the half duplex cognitive radio by reducing $T_{P,half}$ or increasing $T_{S,half}$. However, it reduces the efficiency of the half duplex cognitive radio, which is defined by

$$\eta_{half} = \frac{T_{P,half} - T_{S,half}}{T_{P,half}}. \tag{2.5}$$

Usually, the half duplex cognitive radio senses the in-band channel at the end of the frame. Thus, the sensing period is at least greater than the frame size plus the sensing time. Thus, the maximum delay of notice, interference time, is very large. Even though the sensing time is kept small by using the energy detector, the maximum delay of notice is still large.

2.3 Full Duplex Cognitive Radio Sensing

In this section, we propose the full duplex cognitive radio sensing, which can reduce the delay of notice significantly. A full duplex cognitive radio can cancel the self interference caused by its own transmission. Reduction of the self interference is done by using antenna, RF, and digital cancelation techniques, as in [70, 71]. Again, the cognitive radio senses the in-band channel

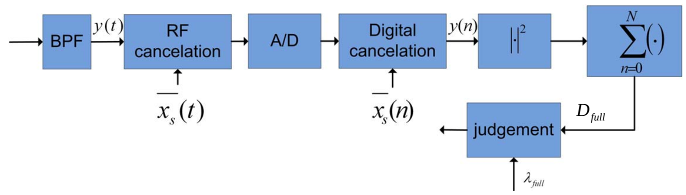


Figure 2.4: Full duplex cognitive radio's energy detection

in order to evacuate from that channel if there is a returning legitimate user. Figure 2.4 shows the block diagram of the in-band sensor for the full duplex cognitive radio which uses an energy detector. An antenna configuration which utilizes antenna placement is not included in this diagram. The cognitive radio receives the intended signal along with the self interference and white Gaussian noise. A band pass filter is applied to extract the signal of the band in interest, and the self interference is reduced using the RF canceler, and further canceled digitally after sampling. Then, energy is calculated, and detection is made. The benefit which comes from using the full duplex cognitive radio is that it can constantly transmit its signal while sensing the channel at the same time. There is no separation of sensing and transmission, since the self interference cancelation enables the cognitive radio to detect the legitimate transmission seamlessly while transmitting. Since the full duplex cognitive radio does not stop its transmission for in-band sensing, its efficiency, $\eta_{full} = 1$.

Furthermore, the interference time (delay of notice) for the full duplex radio, $T_{I,full}$, can be reduced as well. Figure 2.5 shows the timing of the legitimate signal and in-band sensing. A sliding window with the size of $T_{S,full}$ is used in calculating the energy such that the energy detector calculates the

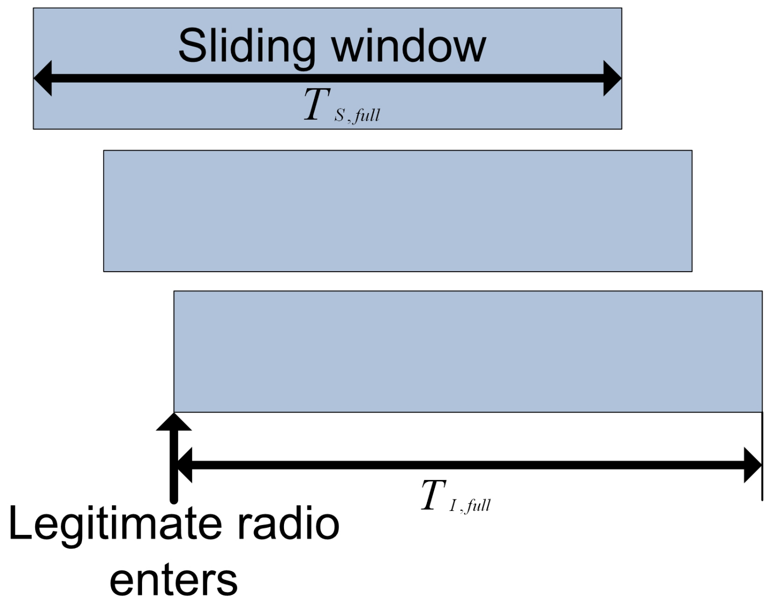


Figure 2.5: Full duplex cognitive radio's interference time

sum energy of the signal within the window, and slides the window after the detection is made. The full duplex radio can detect the legitimate transmission, if the window is completely filled with a legitimate signal, which makes the maximum interference time to be

$$\max T_{I,full} = T_{S,full}.$$

If the full duplex cognitive radio is able to remove self-interference completely, the sensing ability is not affected by its transmission. Then, we can set the sensing time of the full duplex cognitive radio same as that of the half duplex cognitive radio, i.e.

$$T_{S,full} = T_{S,half}.$$

Thus, by using the full duplex cognitive radio, we can ideally reduce the interference time to $T_{S,half}$ while maintaining the detection ability. However, the interference cancelation is not perfect in practice, and there is a residue of interference, as in [71]. The remaining interference affects the performance of the cognitive detection. Thus, it is necessary to analyze the probability of detection error and false alarm, and design $T_{S,full}$ accordingly.

[71] models the received signal after three steps of cancelations. Due to the frequency distortion in the antenna cancelation stage, mismatch in the parameters in the RF cancelation stage, and estimation error in the digital cancelation, interference cannot be canceled out entirely. The signal after three steps of cancelation can be described as

$$y(n) = \begin{cases} x_{s,e}\left(\frac{n}{W}\right) + z\left(\frac{n}{W}\right), & H_0 \\ x\left(\frac{n}{W}\right) + x_{s,e}\left(\frac{n}{W}\right) + z\left(\frac{n}{W}\right), & H_1 \end{cases}, \tag{2.6}$$

where $x_{s,e}\left(\frac{n}{W}\right)$ is the residue of self interference left due to limitation of interference cancelation. Also, self interference may be stronger than the maximum level of interference cancelation capability of a full duplex radio, which is about 55dB in [71]. After passing through the energy detector, the decision value for the full duplex cognitive radio, D_{full}, becomes

$$D_{full} = \begin{cases} \displaystyle\sum_{n=1}^{T_{full}W} \left| x_{s,e}\left(\frac{n}{W}\right) + z\left(\frac{n}{W}\right)\right|^2, & H_0 \\ \displaystyle\sum_{n=1}^{T_{full}W} \left| x\left(\frac{n}{W}\right) + x_{s,e}\left(\frac{n}{W}\right) + z\left(\frac{n}{W}\right)\right|^2, & H_1 \end{cases}. \tag{2.7}$$

Again, the received signal is normalized such that $z\left(\frac{n}{W}\right) \sim N(0,1)$. The signal energy to noise power ratio of the legitimate signal, γ_{full}, and the interference energy to noise power ratio, γ_I, are expressed as follows:

$$\gamma_{full} = \sum_{n=1}^{T_{S,full}W} \left| x\left(\frac{n}{W}\right)\right|^2, \quad \gamma_I = \sum_{n=1}^{T_{S,full}W} \left| x_{s,e}\left(\frac{n}{W}\right)\right|^2.$$

Also, we define the signal and interference to noise ratio, Γ, which takes the form

$$\begin{aligned}\Gamma = & \sum_{n=1}^{T_{S,full}W} \left| x\left(\frac{n}{W}\right) + x_{s,e}\left(\frac{n}{W}\right)\right|^2 \\ = & \gamma_{full} + \gamma_I \\ & + \sum_{n=1}^{T_{S,full}W} \left(2\Re\left(x\left(\frac{n}{W}\right)\right)\Re\left(x_{s,e}\left(\frac{n}{W}\right)\right) + 2\Im\left(x\left(\frac{n}{W}\right)\right)\Im\left(x_{s,e}\left(\frac{n}{W}\right)\right)\right).\end{aligned} \tag{2.8}$$

Even with constant SNR and INR, the signal and interference to noise ratio is randomly distributed, and follows the distribution $P_\Gamma(\gamma)$. The decision statistic follows the distribution of

$$D_{full} \sim \begin{cases} \chi^2_{2T_{S,full}W}(\gamma_I), & H_0 \\ \int_0^\infty \chi^2_{2T_{S,full}W}(\gamma) P_\Gamma(\gamma)\, d\gamma, & H_1 \end{cases}. \tag{2.9}$$

Finally, we can obtain the expression for the probability of false alarm and detection error for a full duplex cognitive radio. The probability of false alarm for the full duplex radio, $P_{f,full}$, can be expressed as follows:

$$P_{f,full} = P(D_{full} > \lambda_{full}|H_0) = Q_{T_{S,full}W}\left(\sqrt{2\gamma_I}, \sqrt{\lambda_{full}}\right). \tag{2.10}$$

We calculate the probability of detection error for the full duplex cognitive radio, $P_{d,full}$ as follows:

$$\begin{aligned} P_{d,full} &= P(D_{full} < \lambda_{full}|H_1) \\ &= 1 - \int_0^{\infty} Q_{T_{S,full}W}\left(\sqrt{2\gamma}, \sqrt{\lambda_{full}}\right) P_{\Gamma}(\gamma)\, d\gamma \\ &= \int_0^{\infty} e^{-\frac{\gamma}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\gamma}{2}\right)^j}{j!} Q\left(\lambda_{full}; 2T_{S,half}W + 2j\right) P_{\Gamma}(\gamma)\, d\gamma. \end{aligned} \tag{2.11}$$

Before we further investigate $P_{d,full}$, we present the following lemma.

Lemma 1. The function $f(\gamma)$, which is defined by

$$f(\gamma) = e^{-\gamma} \sum_{j=0}^{\infty} \frac{\gamma^j}{j!} Q\left(2\lambda; 2k + 2j\right), \tag{2.12}$$

is convex if $\gamma \geq \lambda$.

Proof: We take the second derivative of $f(\gamma)$; then it becomes

$$\begin{aligned} f''(\gamma) &= e^{-\gamma} \sum_{j=0}^{\infty} \frac{\gamma^j}{j!} \left(\begin{array}{l} Q\left(2\lambda; 2k+2j\right) + Q\left(2\lambda; 2k+2j+4\right) \\ -2Q\left(2\lambda; 2k+2j+2\right) \end{array} \right) \\ &\overset{(a)}{=} e^{-(\gamma+\lambda)} \sum_{j=0}^{\infty} \frac{\gamma^j}{j!} \left(\frac{\lambda^{k+j}}{(k+j)!} - \frac{\lambda^{k+j+1}}{(k+j+1)!} \right) \\ &\overset{(b)}{=} e^{-(\gamma+\lambda)} \left(\left(\frac{\lambda}{\gamma}\right)^{\frac{k}{2}} I_k\left(2\sqrt{\lambda\gamma}\right) - \left(\frac{\lambda}{\gamma}\right)^{\frac{k+1}{2}} I_{k+1}\left(2\sqrt{\lambda\gamma}\right) \right) \\ &\overset{(c)}{>} e^{-(\gamma+\lambda)} \left(\left(\frac{\lambda}{\gamma}\right)^{\frac{k}{2}} I_{k+1}\left(2\sqrt{\lambda\gamma}\right) - \left(\frac{\lambda}{\gamma}\right)^{\frac{k+1}{2}} I_{k+1}\left(2\sqrt{\lambda\gamma}\right) \right) \\ &\geq 0 \qquad\qquad \forall \gamma \geq \lambda, \end{aligned}$$

where $I_k(x)$ is a modified bessel function. (a) comes from reducing the function $Q(;)$ to gamma functions and by finding recurrence relations in the low incomplete gamma functions. (b) is obtained by finding the closed form equation of the regularized hypergeometric function. Inequality in (c) is proven in [25]. [25] provides useful bounds for the modified bessel function that can be applied to the lemma above. The bound in (c) is not tight enough, and multiple simulations with varying parameters show the result of $f''(\gamma)$ being greater than 0 if $\gamma \geq \lambda - k$. This is not a surprising result considering the expected value of the random variable with noncentral chi-squared distribution with the degree of freedom, k, and the noncentral parameter, γ, is $\gamma + k$. Since it is densely distributed around the mean, the tail probability decreases more with increasing γ around the mean.

From Lemma 1, we have that $P_{f,full}(\gamma)$, the probability of detection error function of noncentral parameter, is a convex function in the following region:

$$\gamma > \lambda_{full}. \tag{2.13}$$

If being more aggressive, we use the condition

$$\gamma > \lambda_{full} - 2T_{S,full}W \tag{2.14}$$

, where the probability of detection error function is convex. Next, we find the distribution of the random variable, Γ. The distribution of this signal and interference to noise radio is approximated in the following lemma.

Lemma 2. The probability distribution of the signal interference to noise ratio,

Γ, can be approximated as

$$P_\Gamma(\gamma) = \begin{cases} \frac{1}{\sigma_{full}\sqrt{2\pi}} e^{-\frac{\left(\gamma - m_{full}\right)^2}{2\sigma_{full}^2}} & \text{if } -\sqrt{\frac{\gamma_{full}\gamma_I}{\sqrt{T_{S,full}W}}} \le \gamma - m_{full} \le \sqrt{\frac{\gamma_{full}\gamma_I}{\sqrt{T_{S,full}W}}} \\ 0 & \text{o/w} \end{cases}$$

where $\sigma_{full}^2 = \frac{2\gamma_{full}\gamma_I}{T_{S,full}W}$, and $m_{full} = \gamma_{full} + \gamma_I$.

Proof: (2.8) shows that Γ is the sum of $2T_{S,full}W$ number of independent and identically distributed random samples with the offset $\gamma_{full} + \gamma_I$. Each sample has a mean of zero, and its variance is as follows:

$$\begin{aligned} \mathbb{E}\left[\left(2\Re\left(x\left(\frac{n}{W}\right)\right)\Re\left(x_{s,e}\left(\frac{n}{W}\right)\right)\right)^2\right] &= \mathbb{E}\left[\left(2\Im\left(x\left(\frac{n}{W}\right)\right)\Im\left(x_{s,e}\left(\frac{n}{W}\right)\right)\right)^2\right] \\ &= \frac{\gamma_{full}\gamma_I}{\left(T_{S,full}W\right)^2}. \end{aligned}$$

According to the central limit theorem, signal and interference to noise ratio, Γ, obeys the Gaussian distribution.

$$\Gamma \sim \; N\left(\gamma_{full} + \gamma_I, \frac{2\gamma_{full}\gamma_I}{T_{S,full}W}\right).$$

Also, from the Chernoff bounds, we have

$$Pr\left(\left|\Gamma - \left(\gamma_{full} + \gamma_I\right)\right| \ge \sqrt{\frac{\gamma_{full}\gamma_I}{\sqrt{T_{S,full}W}}}\right) \le 2e^{-\frac{\sqrt{T_{S,full}W}}{4}} \approx 0.$$

Thus, we can approximate the probability distribution of Γ as in Lemma 2. With Lemma 1 and Lemma 2, we obtain the bounds for $P_{d,full}$ as in the next two theorems.

Theorem 1. $P_{d,full}$ is lower bounded as

$$P_{d,full} \ge 1 - Q_{T_{S,full}W}\left(\sqrt{2\left(\gamma_{full} + \gamma_I\right)}, \sqrt{\lambda_{full}}\right),$$

when

$$\lambda_{full} \leq \gamma_{full} + \gamma_I - \sqrt{\frac{\gamma_{full}\gamma_I}{\sqrt{T_{S,full}W}}}. \tag{2.15}$$

Proof: From Lemma 2, Γ is larger than λ_{full} under the condition (2.15). Then, using the convexity from Lemma 1, we apply Jensen's inequality rule, and obtain

$$\begin{aligned} P_{d,full}\gamma_{full} &= \mathbb{E}\left[1 - Q_{T_{S,full}W}\left(\sqrt{2\Gamma}, \sqrt{\lambda_{full}}\right)\right] \\ &\geq 1 - Q_{T_{S,full}W}\left(\sqrt{2\mathbb{E}[\Gamma]}, \sqrt{\lambda_{full}}\right) \\ &= 1 - Q_{T_{S,full}W}\left(\sqrt{2\left(\gamma_{full} + \gamma_I\right)}, \sqrt{\lambda_{full}}\right). \end{aligned}$$

We also obtain the following theorem on the upper bound.

Theorem 2. $P_{d,full}$ is upper bounded as

$$\begin{aligned} P_{d,full} \leq& 1 - Q_{T_{S,full}W}\left(\sqrt{2\left(\gamma_{full} + \gamma_I\right)}, \sqrt{\lambda_{full}}\right) \\ &+ \begin{pmatrix} 2Q_{T_{S,full}W}\left(\sqrt{2\left(\gamma_{full} + \gamma_I\right)}, \sqrt{\lambda_{full}}\right) \\ -Q_{T_{S,full}W}\left(\sqrt{2\left(\gamma_{full} + \gamma_I - d\right)}, \sqrt{\lambda_{full}}\right) \\ -Q_{T_{S,full}W}\left(\sqrt{2\left(\gamma_{full} + \gamma_I + d\right)}, \sqrt{\lambda_{full}}\right) \end{pmatrix} \frac{1}{\sqrt{\pi}\left(T_{S,full}W\right)^{\frac{1}{4}}}, \end{aligned}$$

where

$$d = \sqrt{\frac{\gamma_{full}\gamma_I}{\sqrt{T_{S,full}W}}},$$

and when

$$\lambda_{full} \leq \gamma_{full} + \gamma_I - \sqrt{\frac{\gamma_{full}\gamma_I}{\sqrt{T_{S,full}W}}}.$$

Proof: Again, Γ is in the region which makes the function, $f(\gamma) = 1 - Q_{T_{S,full}W}\left(\sqrt{2(\gamma)}, \sqrt{\lambda_{full}}\right)$, convex. We define the function, $g(\gamma)$, which linearly connects two extreme points and center point in $f(\gamma)$ that are $(\gamma_{full}+\gamma_I-d, f(\gamma_{full}+\gamma_I-d))$, $(\gamma_{full}+\gamma_I, f(\gamma_{full}+\gamma_I))$, and $(\gamma_{full}+\gamma_I+d, f(\gamma_{full}+\gamma_I+d))$:

$$g(\gamma) = \begin{cases} r_1\left(\gamma - (\gamma_{full}+\gamma_I)\right) + f(\gamma_{full}+\gamma_I) & \text{if } \gamma \le \gamma_{full}+\gamma_I \\ r_2\left(\gamma - (\gamma_{full}+\gamma_I)\right) + f(\gamma_{full}+\gamma_I) & \text{o/w} \end{cases},$$

where

$$\begin{aligned} r_1 &= \frac{f(\gamma_{full}+\gamma_I) - f(\gamma_{full}+\gamma_I-d)}{d} \\ &= \frac{Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I-d)}, \sqrt{\lambda_{full}}\right) - Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I)}, \sqrt{\lambda_{full}}\right)}{d} \\ r_2 &= \frac{f(\gamma_{full}+\gamma_I+d) - f(\gamma_{full}+\gamma_I)}{d} \\ &= \frac{Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I)}, \sqrt{\lambda_{full}}\right) - Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I+d)}, \sqrt{\lambda_{full}}\right)}{d} \end{aligned}$$

$g(\gamma) \ge f(\gamma)$, because $f(\gamma)$ is convex. And, we have

$$\begin{aligned} P_{d,full} =& \mathbb{E}\left[1 - Q_{T_{S,full}W}\left(\sqrt{2\Gamma}, \sqrt{\lambda_{full}}\right)\right] \\ \le& 1 - Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I)}, \sqrt{\lambda_{full}}\right) + \int_0^d (r_2 - r_1)\gamma P_\Gamma(\gamma) d\gamma \\ \overset{(d)}{\approx}& 1 - Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I)}, \sqrt{\lambda_{full}}\right) \\ &+ \begin{pmatrix} 2Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I)}, \sqrt{\lambda_{full}}\right) \\ -Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I-d)}, \sqrt{\lambda_{full}}\right) \\ -Q_{T_{S,full}W}\left(\sqrt{2(\gamma_{full}+\gamma_I+d)}, \sqrt{\lambda_{full}}\right) \end{pmatrix} \frac{1}{\sqrt{\pi}\,(T_{S,full}W)^{\frac{1}{4}}}, \end{aligned}$$

where (d) is because $1 - e^{\sqrt{T_{S,full}W}/2}$ approximates to 1.

2.4 Numerical Analysis

In this section, we compare the performance of the full duplex cognitive radio with that of the half duplex cognitive radio. The comparison is made numerically. We bring some parameters from in-banding sensing in the 802.22 standard. For this standard, we have a bandwidth $W = 6MHz$, and the delay of notice, $T_{P,half} - T_{S,half}$, is a multiple of the frame size (i.e., $n \cdot 10$ms). We set the delay of notice to be the minimum allowed time, 10ms. For the half duplex radio, we test with the sensing time which makes time bandwidth product, $2T_{S,half}W$=100. We want to detect the legitimate signal of -116dBm with both the probability of detection error and false alarm being less than 0.1. Assuming noise floor to be formed at around -110dBm, we find the trade-off curve between the probability of detection error and false alarm by varying the detection threshold γ_{half}.

For full duplex radio, we assume that the interference power is 3dB higher than the noise floor, which is a very conservative approach to the self interference cancelation. We test with different time bandwidth products (i.e. $2T_{S,full}W$ =1000, 1200, 120000). Again, we obtain the trade-off curve between the probability of detection error and false alarm. Figure 2.6 shows the performance of the half and full duplex cognitive radio when SNR of legitimate signal is -6dB, and INR of the self interference of full duplex radio is 3dB. In the figure, ub and lb stands for upper bound and lower bound respectively. First, we match the interference time for the full duplex cognitive radio, $T_{I,full}$ to be the same as that of the half duplex cognitive radio by having $2T_{S,full}W = 120000$.

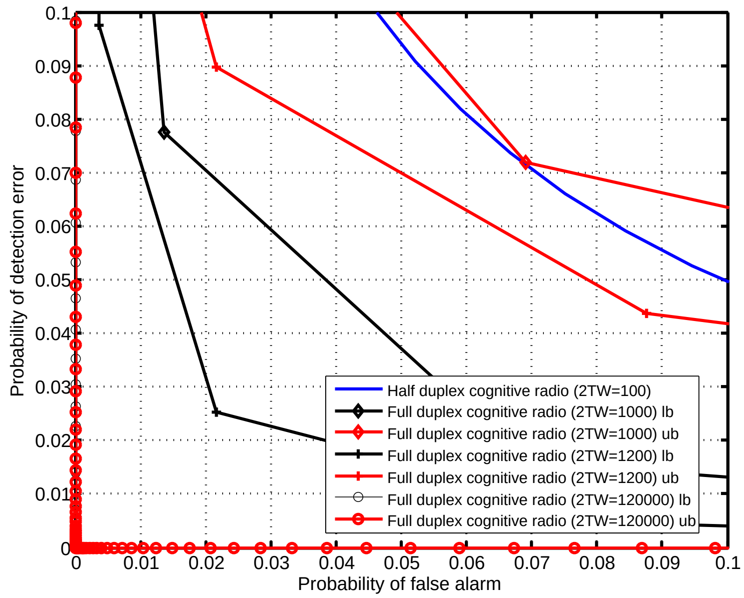


Figure 2.6: Sensing performance comparison with SNR=-6dB and INR=3dB

Thus, interference time for the full duplex radio is 10ms. Then, it shows that the probability of detection error and false alarm decays to 0, because the full duplex cognitive radio can collect large amounts of data. Thus, it is known that the full duplex cognitive radio can increase sensing ability while keeping the interference time the same as that of the half duplex cognitive radio. Next, we reduce the sensing time of the full duplex cognitive radio. By having $2T_{S,full}W = 1000$, the upper bound of probability of the detection error and the false alarm region includes the point where $P_{d,full} \approx 0.07 \approx P_{f,full}$. This can also be obtained by the half duplex cognitive radio. Thus, we can conclude that the interference time can be reduced by a factor of 120. Next, we increase the interference power. Some full duplex cognitive radios may not be able to reduce self interference enough, and some may strategically overpower

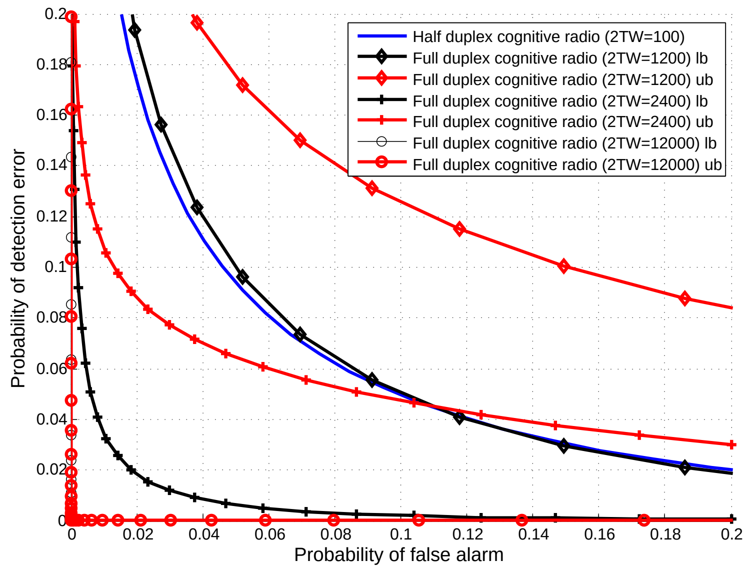


Figure 2.7: Sensing performance comparison with SNR=-6dB and INR=6dB

themselves in their transmission to the degree that there is much leftover self interference. Figure 2.7 shows the sensing ability when SNR of legitimate signal is -6dB, and INR of the self interference of the full duplex cognitive radio is 6dB. The performance of the full duplex cognitive radio reduces significantly. However, full duplex radio can still have a slightly lower probability of detection and false alarm, and reduce the interference time 100 times. Thus, we have $T_{I,full} = 0.1$ms. From this analysis, we can conclude that the full duplex cognitive radio can theoretically reduce the delay of notice significantly.

2.5 Applications and Experimental Results

A full duplex cognitive radio is implemented on our testbed. This testbed is similar to that in [71]. This radio is designed to transmit and

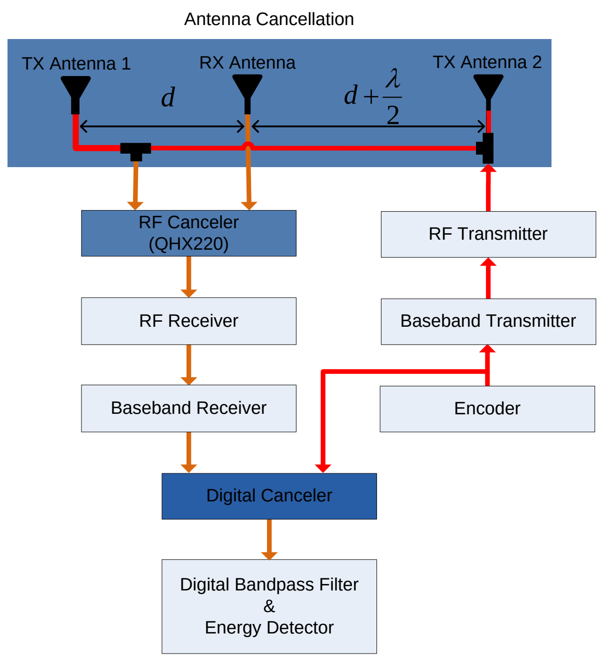


Figure 2.8: Sensing performance comparison with SNR=0dB and INR=3dB

sense at the same time in the same frequency. We use the 2.4GHz ISM band for this experiment. Figure 2.8 demonstrates the block diagram for our full duplex cognitive radio testbed. We split the transmit signal into two, and forward them to two transmit antennas, which are separated such that there is half a wavelength's difference between the distances from the receive antenna to each transmit antenna. In the receive chain, the signal first passes through the analog RF cancelation, which is implemented in our testbed using the QHX220 chipset. QHX220 takes in the received signal and reference interference (noise), and it outputs the received signal with an attenuated self-interference. The reference transmit signal (which is self-interference) is made available to this chipset. As illustrated in Figure 2.8, a power splitter takes the transmit signal, and relays it to QHX220 as a reference interference. Then, the output of QHX220 is down converted to baseband signal, and passes through an analogue to digital converter (ADC). Finally, a Digital Signal Processor (DSP) performs the digital cancelation in real-time (RT) and passes the signal through a bandpass filter. A bandpass filter is used instead of a lowpass filter to take away the severe DC offset which presents in our testbed. The energy of the samples is calculated, and the decision is made whether the legitimate user's transmission exists or not. We use a real-time National Instruments system (the NI PXIe-8133 RT) with a Xilinx SX-100 FPGA and baseband transceiver (NI 5781) combined with RF transceiver (XCVR2450) to transmit and sense a signal.

2.5.1 Application in 802.22 WRAN

A cognitive radio in 802.22 WRAN detects the signals in the TV bands. TV (DTV) signals typically have a bandwidth of 6MHz [16]. It is required to sense the signal of -116dBm with both probability of detection error and false alarm less than 0.1. We transmit a pseudo DTV signal of 6MHz bandwidth and constant power level. Transmit power of the pseudo DTV signal is carefully adjusted manually such that the signal to noise ratio at the sensor (cognitive radio) is -116dBm. First, a conventional sensing is developed with a simple energy detector. Second, a full duplex cognitive radio's sensing is implemented with antenna setup, an RF canceler, and a digital canceler. We use a digital Kaiser bandpass filter with 100KHz low cut off frequency, 3MHz high cut

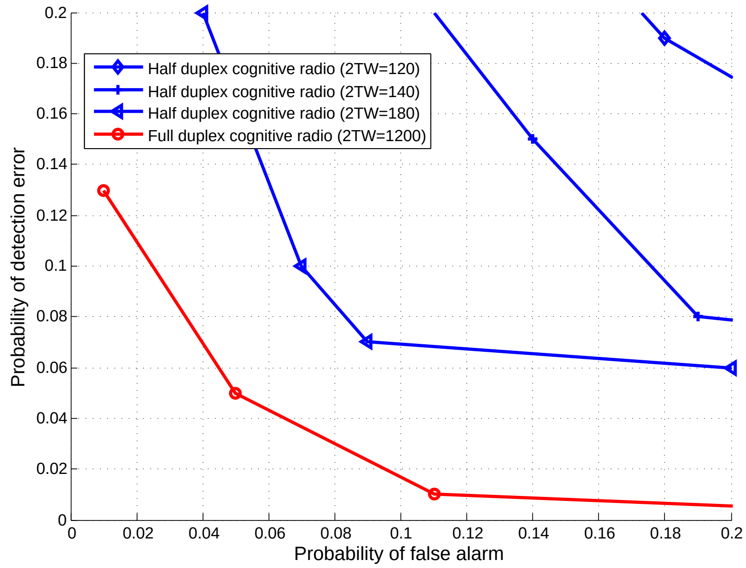


Figure 2.9: Sensing performance comparison with SNR=-6dB and INR=2dB

off frequency, and the filter order of 100. A transmit signal of the cognitive radio uses the same channel with pseudo DTV signal. Interference to noise ratio in our set up is about 2dB. Also, line of sight is provided in the set up to eliminate fading effect as much as possible. Figure 2.9 demonstrates the performance of the half duplex cognitive radio and full duplex cognitive radio in our experiment. Different time bandwidth products are tried in the experiment. The probability of detection error and false alarm is kept under 0.1 if we have $2T_{S,full}W = 1200$ for full duplex cognitive radio, and the delay of notice is 0.1msec. A similar but slightly higher probability of detection error and false alarm can be achieved by the half duplex cognitive radio sensing with $2T_{S,half}W = 180$. Here the delay of notice is close to the sensing period, which is 10msec. It is shown by this experiment that the full duplex cognitive radio sensing can reduce the delay of notice one hundred times as much as the half duplex cognitive radio sensing, while maintaining a slightly lower probability of detection error and false alarm. Therefore, we conclude that the full duplex cognitive radio is desirable for sensing in the 802.22 WRAN device. Note that the experimental performance of the half duplex cognitive radio is not as good as the result of numerical analysis. There can be several reasons for that; for example, the channel may suffer from the fading effect.

2.5.2 Application in 802.11 WLAN

We apply the cognitive radio in the 802.11 WLAN. Here, the cognitive radio tries not to enter the channel if the WLAN user is already using it and evacuates from the channel if the WLAN user re-enters the channel. However, under the current 802.11 protocol, WLAN users cannot re-enter the channel if the cognitive radio is using the channel. In 802.11, media access control is governed by a CSMA/CA DCF. The detailed DCF transmission process is as follows [29]:

1. The sending station determines if the channel is idle before transmitting.
2. If the receiver senses no activity on the channel, the station assumes the channel is idle.
    - The station selects a random back-off interval.
    - The station decreases the back-off interval counter while the channel is idle.

3. If the channel is determined to be busy, the sending station defers until the end of the current transmission.

If the cognitive radio is occupying the channel, the WLAN sending station defers though deferring is not necessary, because the WLAN sending station cannot differentiate if it is the other WLAN user or the cognitive radio that is accessing the channel currently. Thus, we modify the WLAN protocol to signal the cognitive radio to evacuate from the channel. The modified CSMA/CA distributed control function is as follows:

1. The sending station determines if the channel is idle before transmitting.
2. If the receiver senses no activity on the channel, the station assumes the channel is idle.
    - The station selects a random back-off interval.
    - The station decreases the back-off interval counter while the channel is idle.
3. If the receiver senses an activity on the channel, the station assumes the channel is busy.
    - The station transmits the banning control signal.
    - After sending the banning control signal, the station senses the channel again.

Thus, if the WLAN user try to access the channel, and finds the channel to be busy, it transmits the banning control signal. The banning control signal is detected by the cognitive radio when it is conducting in-band sensing. An important thing to note in the banning control signal design is that the length of the banning control signal needs to be larger than the delay of notice. Otherwise, the control signal may be left undetected. We build the cognitive radio system with the WLAN radio which transmits the control signal if the channel is busy. 802.11 standards on WLAN specifies the channel bandwidth to be 20 or 40MHz, depending on the different bodies of the standard [29]. We assume that the WLAN user has a bandwidth of 20MHz, and the control signal uses the same bandwidth. However, the cognitive radio uses the 10MHz bandwidth because the self interference cancelation technique of the signal more than 10MHz has not been studied well yet [71]. We break the wideband channel into pieces, and make the bandwidth of the cognitive radio at most 10MHz. Therefore, even though the signal of 20MHz is present, we obtain the signal within the bandwidth of interest by using the bandpass filter. The rest of the detection algorithm is the same as that of the cognitive radio in 802.22 WRAN. We design that the bandwidth of the cognitive radio's signal is 10MHz. In this case, interference cancelation is not as good as when 6MHz bandwidth is used, as is shown in [71]. We have an interference to noise ratio of 2.3dB this time. Figure 2.10 shows the performances of the full and half duplex cognitive radio sensing. The sensing ability of the full duplex cognitive radio is not as good as those in 6Mhz bandwidth radio system, because there

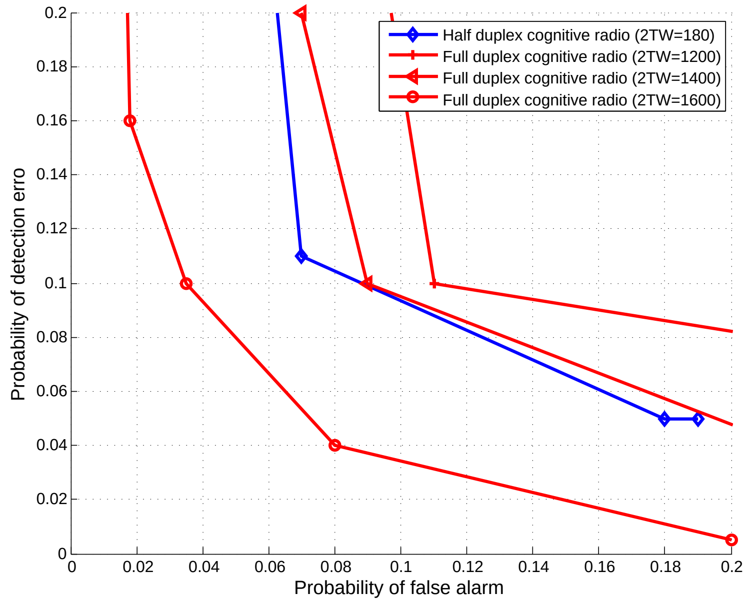


Figure 2.10: Sensing performance comparison with SNR=-6dB and INR=2.3dB

is an increase in the interference. However, a full duplex cognitive radio is capable of sensing the WLAN user's banning control signal with an error rate of 0.1. Even though sensing time for the half duplex cognitive radio is shorter than that of the full duplex radio, it is delay of notice that is important. The full duplex radio has a short delay of notice, 70μsec.

Thus, the WLAN station, upon finding that the channel is busy, transmits a control sequence which has a length of 70μsec. The cognitive radio can detect the channel with the probability of 0.1 when the received power of the control signal is -116dBm. The WLAN station then can attempt to access the channel without the presence of the cognitive radio. Assuming that there is no propagation and processing delay, the WLAN user has a cognitive radio free channel in 70μsec.

On the other hand, it may be the other WLAN user who is occupying the channel. If the new WLAN user attempts to use the channel, and transmits the control signal, it acts as an interference to the transmission of the WLAN user who is already using the channel. An 802.11 frame consists of frame header, frame body, and FCS. The frame header and FCS is about 30 and 4 octets in size, respectively. An 802.11b frame can carry 0 to 2312 octets, while 802.11a and 802.11g frames can carry 0 to 4095 octets in the frame body. We consider that the WLAN user has infinite buffer, thus it can bring as many data to the frame as it wants. However, most drivers set the maximum frame body size to 1500 octets to match the maximum Ethernet frame size. We modify this a little bit such that the frame body can have 1500 or 3000 octets, because it is easy for a driver to concatenate frames. As a result, frame size is 1534 octets for 802.11b, and 3034 for 802.11a and 802.11g. Maximum line rate is 72 Mbps for 802.11a and 802.11g, and 11 Mbps for 802.11b, which makes the length of 802.11b frame 1.12msec, and the 802.11a and 802.11g frames 0.32msec. Thus, 802.11 WLAN users may face interference up to 22 percent of the time. It is controversial if this interference from using the control signal is sustainable.

We compromise the sensing ability of the cognitive radio such that it can detect the signal of -110dBm power with the probability of false alarm and detection error of 0.1. Figure 2.11 illustrates the performance of the full duplex radio and half duplex radio when the detecting signal has -110dBm power. It

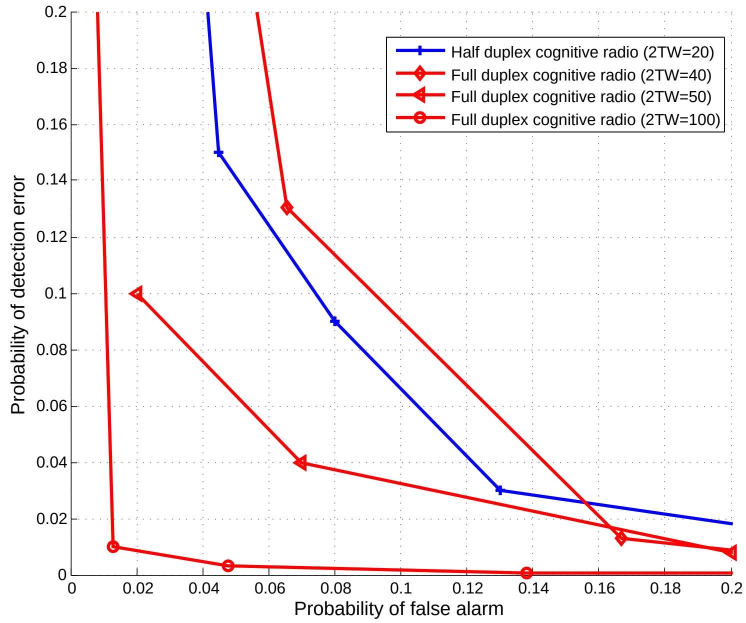


Figure 2.11: Sensing performance comparison with SNR=0dB and INR=2.3dB

shows that detection can be done in 2μsec, making the delay of notice and the length of the control signal to be 2μsec. 2μsec of control signal interferes with the other WLAN user's frame 6.25% of the time. In the extreme case (i.e., when two WLAN users are close enough), this interference from the control signal is so high that the WLAN user's data cannot be decoded. We throw a random bit for the bits that face the interference, and obtain a 0.003125 raw bit error rate. With the help of error correction code, this frame can be detected. From this experiment, we prove that the secondary full duplex cognitive radio system with primary 802.11 WLAN can increase the spectral efficiency without compromising the network throughput of the WLAN users.

Chapter 3

Resource Allocation in Interweave Cognitive Radio

3.1 Introduction

In this chapter, we focus on the interweave cognitive radio with multiple sub-channels. In this setting, there are two parameters that determine system performance: a. which subset of sub-channels should be sensed and b. what power must be allocated to each sub-channel. Selecting which channel to sense is typically an integer program, and thus hard to solve exactly. This is further compounded by the fact that the power allocation is tightly coupled with the selection process.

3.1.1 Our Contributions

Our main contributions in this chapter are as follows:

1. We establish the power control and thus, the fundamental capacity limit of cognitive radio with multiple channels.

2. We establish a practical but approximate algorithm for joint power allocation and sub-channel selection.

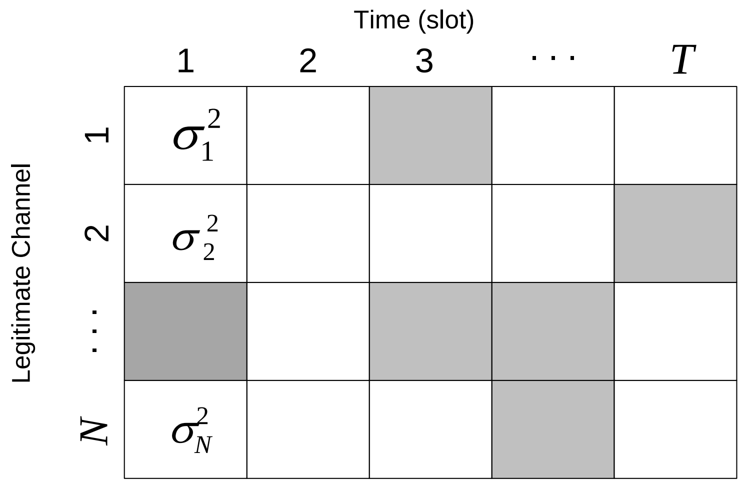


Figure 3.1: Channel model

3.2 System Model and Problem Statement

The channel model is shown in Figure 3.1. We consider N parallel legitimate channels with equal bandwidth. In each time slot, a channel $n, where 1 \leq n \leq N$, is occupied by a legitimate user with probability q_n. There is one cognitive transmitter and cognitive receiver pair. The cognitive transmitter is allowed to transmit over channel n, if it is not occupied by any legitimate user. In legitimate channel n, cognitive radio's channel is characterized mathematically as:

$$Y_n = X_n + Z_n$$

where Z_n is additive Gaussian noise of variance σ_n^2. This noise variance can be different from channel to channel, as it represents the fading state of a particular channel. At the start of every time slot, the cognitive transmitter senses a *subset* of channels, and is allowed to exploit those channels that are unoccupied; in this paper, we assume that the sensing is performed perfectly. As N is large, we require it to cleverly choose a subset of bands on which to focus its efforts. In order to guarantee the transmission of the legitimate users, the cognitive transmitter cannot use the channel which is not sensed. The capacity of the cognitive radio depends on its choice of sensing channels (from N parallel channels) and its power allocation among the available channels. Average total transmission power of cognitive transmitter is constrained to P.

First, define the $I_n(t)$ and $I_{E,n}(t)$ to be the indicator function for the channel selected for sensing and an indicator function for the occupied channel at time instance t, respectively, i.e.,

$$I_n(t) = \begin{cases} 0 & \text{if channel } n \text{ is not to be sensed} \\ 1 & \text{if channel } n \text{ is to be sensed} \end{cases} \tag{3.1}$$

$$I_{E,n}(t) = \begin{cases} 0 & \text{if channel } n \text{ is occupied} \\ 1 & \text{if channel } n \text{ is unoccupied} \end{cases} . \tag{3.2}$$

The time average capacity of the cognitive radio with the selection of the sensing channel $I^{N,T} = (I_1(1), ..., I_N(1), ..., I_1(T), ..., I_N(T))$ and power allocation $P^{N,T} = (P_1(1), ..., P_N(1), ..., P_1(T), ..., P_N(T))$ in one time block, $\mathcal{C}\left(I^{N,T}, P^{N,T}\right)$, can be numerically expressed as

$$\mathcal{C}\left(I^{N,T}, P^{N,T}\right) = \frac{1}{T}\sum_{n=1}^{N}\sum_{t=1}^{T}\frac{I_n(t)I_{E,n}(t)}{2}\log\left(1+\frac{P_n(t)}{\sigma_n^2}\right),$$

where T is the number of time slots in each time block.

We assume two constraints on the cognitive radio:

1. An average power constraint of P,
2. L, number of channels to be sensed at any given time, $\leq N$.

The resulting optimization problem can be stated as follows:

$$\max_{I^{N,T},P^{N,T}} \mathcal{C}\left(I^{N,T}, P^{N,T}\right) \tag{3.3a}$$

such that

$$\frac{1}{T}\sum_{n=1}^{N}\sum_{t=1}^{T} I_n(t) I_{E,n}(t) P_n(t) \leq P, \tag{3.3b}$$

$$\sum_{n=1}^{N} I_n(t) \leq L, \tag{3.3c}$$

$$I_n(t) \in \{0,1\}, \qquad P_n(t) \geq 0, \qquad \texttt{for all } (n = 1, ..., N), (t = 1, ..., T). \tag{3.3d}$$

The optimization problem given by (3.3) determines the maximum empirical average rate achieved by the cognitive radio given constraints on the system. It is an integer programming (IP) due to the constraints in (3.3d), and multi-dimensional due to its dependence on time t.

3.3 Optimal Power Allocation and Selection of Sensing Channel

As a first step, we assume that our policy is ergodic and "static", i.e., that our sensing and power allocation policies are only functions of the channel statistics and do not evolve with time. $\mathcal{C}\left(I^N, P^N\right)$, the average capacity of the cognitive radio with the selection of the sensing channel $I^N = (I_1, ..., I_N)$ and power allocation $P^N = (P_1, ..., P_N)$, is calculated as follows:

$$\mathcal{C}\left(I^N, P^N\right) = \sum_{n=1}^{N} \frac{I_n q_n}{2} \log\left(1 + \frac{P_n}{\sigma_n^2}\right). \tag{3.4}$$

This results in the following (simplified) optimization problem:The resulting optimization problem of interest is to maximize $\mathcal{C}\left(I^N, P^N\right)$ such that

$$\sum_{n=1}^{N} I_n q_n P_n \leq P, \quad \sum_{n=1}^{N} I_n \leq L, \tag{3.5}$$

$$I_n \in \{0,1\}, \quad P_n(t) \geq 0, \quad \texttt{for all } (n = 1, ..., N). \tag{3.6}$$

3.4 Optimal Power Allocation and Selection of Sensing Channel

First, we find the optimal power allocation strategy with a given channel selection, I^N . The optimal power allocation is given by the following theorem:

Theorem 3. $\mathcal{C}\left(I^N, P^N\right)$ is maximized when:

$$P_n = \left\lceil \lambda - \sigma_n^2 \right\rceil^+ I_n,$$

where

$$\sum_{n=1}^{N} \left\lceil \lambda - \sigma_n^2 \right\rceil^+ I_n q_n = P,$$

and $\lceil w \rceil^+$ is a maximum value of 0 and w.

Proof: Since $P_n = 0$ where $I_n = 0$, the first constraint in (3.5) can further be relaxed to

$$\sum_{n=1}^{N} q_n P_n \leq P. \tag{3.7}$$

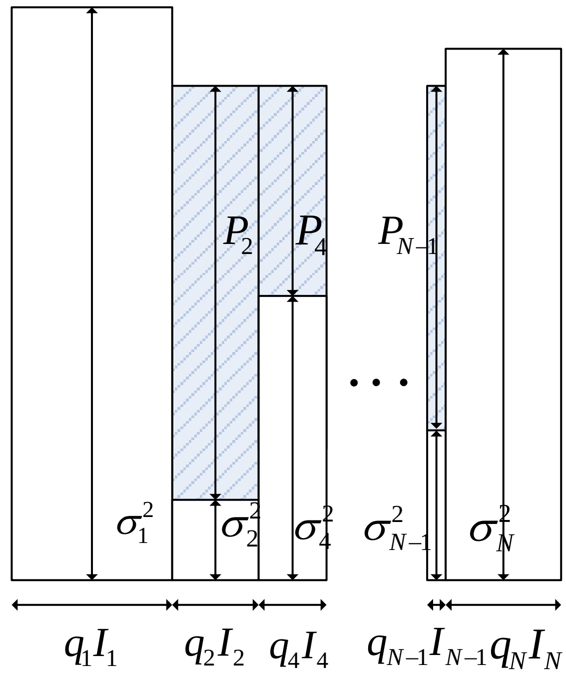


Figure 3.2: Modified water-filling

By applying the Karush-Kuhn-Tucker (KKT) conditions, we obtain the following equations on P_n:

$$\frac{I_n q_n \log e}{2\left(P_n + \sigma_n^2\right)} - \lambda' q_n + \lambda_n = 0,$$
$$\lambda' \left(\sum_{n=1}^{N} I_n^* q_n P_n - P \right) = 0, \quad \lambda_n P_n = 0.$$

As a result, we obtain

$$P_n = \left\lceil \frac{I_n \log e}{2\lambda'} - \sigma_n^2 \right\rceil^+ = \left\lceil \lambda - \sigma_n^2 \right\rceil^+ I_n, \tag{3.8}$$

where $\sum_{n=1}^{N} I_n q_n \left\lceil \lambda - \sigma_n^2 \right\rceil^+ = P$. This optimal power allocation is very similar to the water-filling as shown in Figure 3.2. A little difference is that we only allocate the power where the sensing is performed, and more power tends toward the channel with a higher q_n and lower σ_n^2 whereas the water-filling solution allocates more power to the channel where it has lower σ_n^2. We refer to this policy as *modified* water-filling throughout this paper. Given that we understand the structure of the power allocation policy, we now desire to determine I_n for $n = 1, ..., N$. Note again that the optimization problem with respect to I_n is an integer programming (IP). It can be found by an exhaustive search, but is computationally very hard to solve.

3.5 Joint Selection and Power Control

A typical IP is non-polynomial in complexity. Our focus is on developing a complexity-wise practical algorithm customized to this problem setting. We perform this in two steps, which we call "coarse" and "fine" optimization. The coarse optimization step determines a set of L channels which gives us the lowest possible water-level, λ_{min}. The fine optimization step uses λ_{min}, which we obtain from coarse optimization to further optimize the choice of the L channels. First, we describe the coarse optimization step:

Coarse Optimization: We iteratively find the channels to sense which incur the lowest water level in modified water-filling. Let λ_{min} denote the lowest water level, and $I_c^N = (I_{c,1}, ..., I_{c,N})$ and $P_c^N = (P_{c,1}, ..., P_{c,N})$ indicate the selection of the channel and power allocation which result in λ_{min}. Procedures to find λ_{min}, I_n^c, and P_n^c are as follows:

1. Start with L random initial channels. For example, $I_{c,n} = 1$ for $n = 1, ..., L$, and $I_{c,n} = 0$ otherwise.

2. Perform the modified water-filling with I_c^N, and obtain λ_{min} and P_c^N, which satisfy equations in Theorem 3.

3. Calculate $q_n(\lambda_{min} - \sigma_n^2)$, and select the largest L channels. Update I_c^N with those channels.

   - If I_c^N is the same from the previous iteration, terminate the iteration.
   - Otherwise, repeat from 2).

The optimality of the coarse optimization in one special case is given by the following Lemma.

Theorem 4. Define S_c to be the set of the channels which are selected from coarse optimization:

$$S_c = \{n \in [1, N] | I_{c,n} = 1\}.$$

If the noise variances of all the channels which are not selected in the coarse optimization are greater than the lowest water level λ_{min}, i.e.,

$$\sigma_n^2 \geq \lambda_{min}, \ \ \forall n \in [1, N], n \notin S_c$$

then the coarse optimization is optimal.

Proof: Consider any set of L channels $S \neq S_c$. Assume that average capacity with channels in S with optimal power allocation, $\mathcal{C}(S)$, is greater than the average capacity with channels in S_c with optimal power allocation, $\mathcal{C}(S_c)$. Take the union of S and S_c, and perform the modified water filling among those channel. Resulting average capacity $\mathcal{C}(S \cup S_c)$ is supposed to be greater than $\mathcal{C}(S)$. However, modified water filling with the channels in $S \cup S_c$ allocates power to the channels in S_c only. Thus, $\mathcal{C}(S \cup S_c) = \mathcal{C}(S)$, and it contradicts the earlier assumption. This concludes the proof. Theorem 4 indicates that the coarse optimization is optimal when the λ_{min} is less than the noise variances of the unselected channels, and it happens when we have low signal to noise ratio (SNR). Thus, the coarse optimization is optimal in the low SNR region, but further optimization is required when SNR is high, in which case we perform the fine optimization.

Fine Optimization: We assume that the channels with noise variance larger than λ_{min} do not contribute in increasing the average capacity. We rearrange the useful channels by indexing from 1 to M, where M is the number of channels that has noise variance smaller than λ_{min}, and solve the optimization problem. The optimization problem can be rewritten as follows:

$$
\begin{aligned}
\max_{\lambda, I^M} \mathcal{C}\left(\lambda, I^M\right) &= \max_{\lambda, I^M} \sum_{n=1}^{M} \frac{q_n}{2} \log\left(1 + \frac{\left\lceil \lambda - \sigma_n^2 \right\rceil^+ I_n}{\sigma_n^2}\right) \\
&= \max_{\lambda, I^M} \sum_{n=1}^{M} \frac{q_n I_n}{2} \left\lceil \log \frac{\lambda}{\sigma_n^2} \right\rceil^+, \\
&\overset{(a)}{=} \max_{\lambda, I^M} \sum_{n=1}^{M} \frac{q_n I_n}{2} \log \frac{\lambda}{\sigma_n^2},
\end{aligned}
$$

such that

$$\sum_{n=1}^{M} \left\lceil \lambda - \sigma_n^2 \right\rceil^+ I_n q_n \overset{(b)}{=} \left(\lambda - \sigma_n^2\right) I_n q_n \leq P,$$
$$\sum_{n=1}^{M} I_n \leq L, \qquad \lambda \geq \lambda_{min}$$

where (a) and (b) result from constraining $\lambda \geq \lambda_{min}$. Then, the sub-optimal channel selection and power allocation can be determined by using the following theorem:

Theorem 5.

$$\lambda = \frac{\sum_{n=1}^{M} q_n I_n \sigma_n^2 + P}{\sum_{n=1}^{M} q_n I_n}$$
$$I_n = 0 \quad \texttt{if} \ \lambda > \sigma_n^2 e^{1-\frac{\sigma_n^2}{\lambda}}$$

Proof: Relax the constraint on I_n, such that the I_n can take the value in the region $[0, 1]$. We aim to find the optimal channel selection and power allocation by applying KKT conditions. However, direct application does not provide useful conditions. Thus, we modify the optimization as follows:

$$\max_{\lambda, I^M} \sum_{n=1}^{M} \frac{q_n I_n^k}{2} \log \frac{\lambda}{\sigma_n^2}, \tag{3.9}$$

such that

$$\left(\lambda - \sigma_n^2\right) I_n^k q_n \leq P, \qquad \sum_{n=1}^{M} I_n \leq L, \tag{3.10}$$
$$\lambda \geq \lambda_{min}, \qquad 0 \leq I_n \leq 1.$$

We set the k value to be smaller than 1, which narrows the region of I^M in conditions (3.10). We wish that this results in the channel selection solution, I^M, to become integer value at the end. Also, by having k to be smaller than

$$\min_{\sigma_n^2}\left(\frac{\log\frac{\lambda_{min}}{\sigma_n^2}}{\log\frac{\lambda_{min}}{\sigma_n^2}+\log e}\right),$$

(3.9) becomes concave. However, by setting k to be less than 1, conditions in (3.10) loses its convexity. Thus, a solution from KKT condition is not guaranteed to produce the global maximum. From the KKT condition, we have

$$\begin{matrix} kq_nI_n^{k-1}\log\frac{\lambda}{\sigma_n^2}-\mu_0q_n(\lambda-\sigma_n^2) \\ -\mu_1+\mu_{2,i}-\mu_{3,i} \end{matrix} = 0 \tag{3.11}$$

$$\sum_{n=1}^{M} q_nI_n^k\frac{\log e}{\lambda}-\mu_0\sum_{n=1}^{M} q_nI_n+\mu_4=0 \tag{3.12}$$

$$\mu_0\left(\sum_{n=1}^{M} q_nI_n(\lambda-\sigma_n^2)-P\right)=0, \tag{3.13}$$

$$\mu_1\left(\sum_{n=1}^{M} I_n-L\right)=0, \tag{3.14}$$

$$\mu_{2,i}I_n=0, \tag{3.15}$$

$$\mu_{3,i}(I_n-1)=0, \tag{3.16}$$

$$\mu_4(\lambda-\lambda_{min})=0, \tag{3.17}$$

where μ_0, μ_1, $\mu_{2,i}$, $\mu_{3,i}$, and μ_4 are non-negative values. From the condition (3.12), μ_0 is non-zero in order for λ to be positive. Then, from the condition (3.13), we have

$$\sum_{n=1}^{M} q_nI_n^k(\lambda-\sigma_n^2)=P. \tag{3.18}$$

From the condition (3.11), we find that

$$I_n = \left(\frac{\left(q_n \log \frac{\lambda}{\sigma_n^2} - \mu_0 q_n(\lambda - \sigma_n^2) - \mu_1 \right) k}{\mu_{3,i} - \mu_{2,i}} \right)^{\frac{1}{1-k}}$$

If I_n is not either 0 or 1, from conditions (3.15) and (3.16), $\mu_{2,i}$ and $\mu_{3,i}$ becomes 0, which will result in making I_n an infinite number. Thus, I_n takes either 0,1 value, which gives the desirable solution, such that the optimization of the relaxed condition coincides with the condition for the original problem, and

$$\lambda = \frac{\sum_{n=1}^{M} q_n I_n \sigma_n^2 + P}{\sum_{n=1}^{M} q_n I_n}. \tag{3.19}$$

We set μ_4 to be zero, and we obtain

$$I_n = \frac{\left(q_n \log \frac{\lambda}{\sigma_n^2 e^{1-\frac{\sigma_n^2}{\lambda}}} - \mu_1 \right) k}{\mu_{3,i} - \mu_{2,i}}^{\frac{1}{1-k}}. \tag{3.20}$$

As a result, I_n can be 1 only if

$$\lambda > \sigma_n^2 e^{1-\frac{\sigma_n^2}{\lambda}}. \tag{3.21}$$

This concludes the proof. With the Theorem 5, we can design an iterative algorithm to find the selection of channels to sense and water-level (water-level is directly related to the power allocation). We denote the channel selection from fine optimization as I_f^M, and the water-level result as λ_f.

1. Bring the channels from coarse optimization to be the initial channels.

$$I_{f,n} = \begin{cases} 1 & \texttt{if } n \in S_c \\ 0 & \texttt{otherwise} \end{cases}$$

2. Calculate the water-level λ_f from (3.19)

3. Calculate $\lambda_f - \sigma_n^2 e^{1-\frac{\sigma_n^2}{\lambda_f}}$, and select the largest L channels. Update I_f^M

   - If I_f^M is the same from the previous iteration, terminate the iteration, and set the power allocation accordingly from λ_f and channel selection values, I_f^M.

   - Otherwise, repeat from 2).

Following from Theorem 5, this algorithm gives the efficient joint channel selection and power allocation algorithm. It may result in the local maximum. However, we start the iteration from the coarse optimization value, and believe that water-level increase from λ_c is not significant, and converges into the global optimal solution.

3.6 Numerical Analysis

In this section, we present numerical result of capacities for coarse and fine optimization along with optimal solution. Frequency selective channel, $N = 16, L = 8$, is made by adapting multi-path fading, and occupation of the legitimate channel is modeled by having q_n be uniform i.i.d. in $[0, 1]$. Figure 3.3 shows that performance of the fine optimization meets with that of optimal one as Theorem 5 states. Coarse optimization also performs optimally in the low SNR region. Overall, computationally practical joint channel selection and power allocation are shown to perform as well as the optimal one.

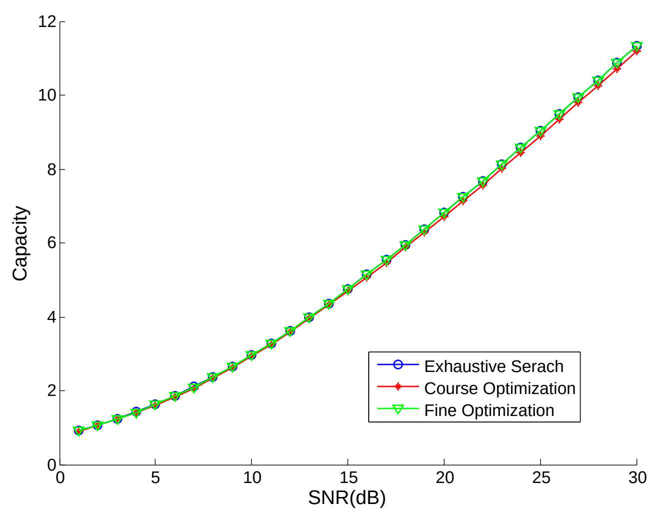


Figure 3.3: Performance Analysis

Chapter 4

Capacity on the Overlay Cognitive Radio with Partial Information

4.1 Introduction

Chapter 4 and 5 study the capacity region of classes of overlay cognitive radio. In this overlay cognitive radio setting, there is a unidirectional cooperation from the cognitive radio to the legitimate radios. It is different from the wireless systems with cooperation, where transmitters cooperate with each other [57]. Unidirectional cooperation of the overlay cognitive radio is assumed because we do not intend to have legitimate radios to change their structures or protocols due to the presence of the cognitive radio.

In this chapter, we focus on the overlay cognitive radio channel with imperfect cognitive information. Even though the capacity region of the overlay cognitive radio with perfect cognitive information is well established, it is hard to assume that perfect message sets of legitimate radio are provided to the cognitive radio in practice. We find the capacity region of this overlay cognitive radio with the partial information.

4.1.1 Our Contributions

Our main contribution in this chapter is as follows:

1. We establish outer bound for capacity of overlay cognitive radio with partial information

2. We establish achievable scheme, and inner bound for capacity of overlay cognitive radio with partial information

3. We analyze gap between outer and inner bound.

4.2 System Model and Preliminaries

First, we describe the notations that are used in this paper. Random variables are written in capital letters, and their realizations are denoted by the corresponding lower-case letters. X_m^n denotes the random vector $(X_m, ..., X_n)$, X^n denotes the random vector $(X_1, ..., X_n)$, and $X^{n\backslash m}$ denotes the random vector $(X_1, ..., X_{m-1}, X_{m+1}, ..., X_n)$. Also, for any set S, $\overline{S}$ denotes the convex hull of S, and $\widetilde{S}$ means the complementary set of S. Finally, the notation $X \Leftrightarrow Y \Leftrightarrow Z$ is used to denote that X and Z are conditionally independent given Y.

4.2.1 Discrete Memoryless Partially Cognitive Radio Channels

A two-user interference channel as in Figure 4.1 is a quintuple $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, p)$, where $\mathcal{X}_1, \mathcal{X}_2$ are two input alphabet sets; $\mathcal{Y}_1, \mathcal{Y}_2$ are two output alphabet sets; $p(y_1, y_2|x_1, x_2)$ is a transition probability. Since we confine

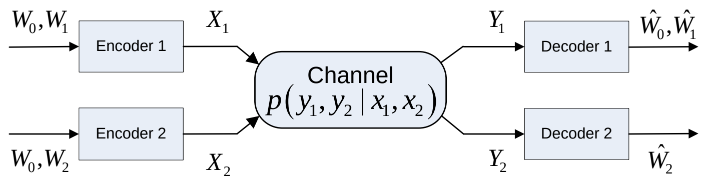


Figure 4.1: Discrete memoryless partially cognitive radio model

channel to be memoryless, the transition probability of y_1^n, y_2^n given x_1^n, x_2^n is

$$p(y_1^n, y_2^n | x_1^n, x_2^n) = \prod_{i=1}^{n} p(y_{1,i}, y_{2,i} | x_{1,i}, x_{2,i}).$$

This channel model is similar to that of an interference channel with difference being the message sets at each transmitter. Transmitter 1 is the legitimate user, which communicates messages from the sets $W_0 \in \{1, ..., M_0\}$ and $W_1 \in \{1, ..., M_1\}$ to Receiver 1, the legitimate receiver. Transmitter 2, the cognitive transmitter communicates a message $W_2 \in \{1, ..., M_2\}$ to Receiver 2, the cognitive receiver. The unique feature of this channel is that the realization of W_0 is known to *both* transmitters 1 and 2, which allows partial and unidirectional cooperation between the transmitters. Difference between this channel model and interference channel with common message is that W_0 does not need to be decoded in Receiver 2, the cognitive receiver. Also, it is different from interference channel with transmitter cooperation, where all the message sets are shared between transmitters. An $(R_0, R_1, R_2, n, P_{e,0}, P_{e,1}, P_{e,2})$ code is any code with the rate vector (R_0, R_1, R_2) and block size n, where $R_t \triangleq \log(M_t)/n$ bits per usage for $t = 0, 1, 2$. As discussed earlier, W_0 and W_1 are the messages that Receiver 1 must decode with (average) probabilities of error of at most $P_{e,0}, P_{e,1}$ respectively, and W_2 is the message that Receiver 2 must decode with an error probability of at most $P_{e,2}$. Rate triplet (R_0, R_1, R_2) is said to be achievable if the error probabilities $P_{e,t}$ for $t = 0, 1, 2$ can be made arbitrarily small as the block size n grows. The capacity region of the interference channel with a partially cognitive transmitter is the closure of the set of all achievable rate triplets (R_0, R_1, R_2). The main goal of the users, legitimate and cognitive, is to maximize in general $\mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2$ for some nonnegative numbers μ_0, μ_1, and μ_2.

Note that if the optimization problem above was unconstrained, the optimal value will also set R_1 to zero. In essence, the optimization problem would transform the system into a fully cognitive model. To obtain a viable partially cognitive solution, we place a restriction on the pair (R_0, R_1), requiring that $R_1 \geq \mu R_0$ for some positive number μ. In some sense, μ represents the degree to which the cognitive radio is cognizant of the legitimate radio's message. When $\mu \to 0$, it represents full cognition as it removes any restriction on R_1, and when $\mu \to \infty$, it represents no cognition as it sets R_0 to zero.

4.2.2 Gaussian Partially Cognitive Radio Channel

In the Gaussian IFC, input and output alphabets are the real $\mathbb{R}$, and outputs are the linear combination of the inputs and additive white Gaussian noise. A Gaussian IFC model in Figure 4.2 is characterized mathematically as

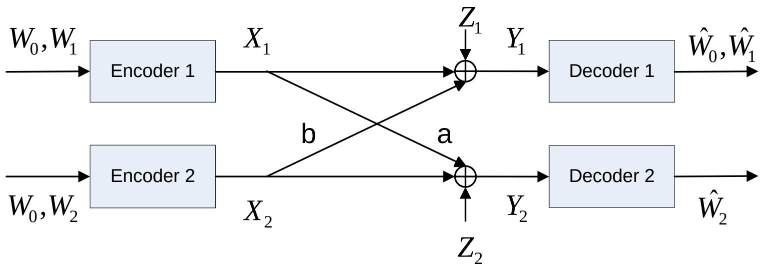


Figure 4.2: Gaussian partially cognitive radio channel

follows:

$$\begin{aligned} Y_1 &= X_1 + bX_2 + Z_1, \\ Y_2 &= aX_1 + X_2 + Z_2, \end{aligned} \tag{4.1}$$

where a and b are real numbers and Z_1 and Z_2 are independent, zero-mean, unit-variance Gaussian random variables. Further, each transmitter has a power constraint

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_{t,i}^2] \leq P_t, t = 1, 2.$$

This concludes our description of the models considered in this paper. The next section describes the outer bound on the capacity region for these channels under "weak" interference condition.

4.3 The Outer Bound Region

4.3.1 Discrete Memoryless Partially Cognitive Radio Channels

For a discrete memoryless channel, under the condition

$$X_2|X_1 \Leftrightarrow Y_2|X_1 \Leftrightarrow Y_1|X_1, \tag{4.2}$$

we say that the legitimate receiver is observing weak interference [53], [52]. In this setting, we present an outer bound on the rate region using the following theorem:

Theorem 6. The convex closure of the following inequalities defines an outer bound on the capacity region of "weak" partially cognitive radio channels:

$$R_0 \leq I(U, X_1; Y_1|V), \tag{4.3}$$

$$R_1 \leq I(X_1; Y_1|X_2), \tag{4.4}$$

$$R_0 + R_1 \leq I(U, X_1; Y_1), \tag{4.5}$$

$$R_2 \leq I(X_2; Y_2|U, X_1), \tag{4.6}$$

$$R_1 \geq \mu R_0, \tag{4.7}$$

for any $p(u,v)p(x_1|u,v)p(x_2|u)$ such that:

1. V and X_2 are independent,
2. X_1 is a function of U and V,
3. $(U,V) \Leftrightarrow (X_1,X_2) \Leftrightarrow (Y_1,Y_2)$.

Proof: First, we restate a lemma from [50] which is used in constituting the outer bound.

Lemma 3 ([50]). The following forms a Markov chain for the partially cognitive radio channel:

$$(W_0, W_t) \Leftrightarrow (W_0, X_t) \Leftrightarrow Y_t, \tag{4.8}$$

where $t = 1, 2$.

We start the main proof by verifying the outer bound for R_0, R_1, and R_2. We have

$$
\begin{aligned}
nR_0 =& H(W_0|W_1) \\
\leq& I(W_0; Y_1^n|W_1) + n\epsilon_0 \\
=& \sum_{i=1}^{n} [H(Y_{1,i}|Y_1^{i-1}, W_1) - H(Y_{1,i}|Y_1^{i-1}, W_0, W_1)] + n\epsilon_0 \\
\leq& \sum_{i=1}^{n} [H(Y_{1,i}|W_1) - H(Y_{1,i}|Y_1^{i-1}, X_1^{n\setminus i}, W_0, W_1, X_{1,i})] + n\epsilon_0 \\
\overset{(a)}{\leq}& \sum_{i=1}^{n} [H(Y_{1,i}|W_1) - H(Y_{1,i}|Y_2^{i-1}, X_1^{n\setminus i}, W_0, W_1, X_{1,i})] + n\epsilon_0 \\
\overset{(b)}{=}& \sum_{i=1}^{n} [H(Y_{1,i}|V_i) - H(Y_{1,i}|U_i, V_i, X_{1,i})] + n\epsilon_0 \\
=& \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{1,i}|V_i) + n\epsilon_0,
\end{aligned}
$$

where (a) results from the conditional Markov chain for the weak interference channel, $X_2|X_1 \Leftrightarrow Y_2|X_1 \Leftrightarrow Y_1|X_1$ in (4.2). (b) results from identifying auxiliaries $U_i = (Y_2^{i-1}, X_1^{n\backslash i}, W_0)$ and $V_i = W_1$.

$$
\begin{aligned}
nR_1 =& H(W_1) \\
\leq& I(W_1; Y_1^n) + n\epsilon_0 \\
=& I(W_1; Y_1^n | X_2^n) + n\epsilon_0 \\
=& \sum_{i=1}^{n} [H(Y_{1,i} | Y_1^{i-1}, X_2^n) - H(Y_{1,i} | Y_1^{i-1}, X_2^n, W_1)] + n\epsilon_0 \\
\leq& \sum_{i=1}^{n} [H(Y_{1,i} | X_{2,i}) - H(Y_{1,i} | X_{1,i}, X_{2,i})] + n\epsilon_0 \\
=& \sum_{i=1}^{n} I(Y_{1,i}; X_{1,i} | X_{2,i}) + n\epsilon_0,
\end{aligned}
$$

and

$$
\begin{aligned}
nR_2 &= H(W_2 | W_0) \\
&\leq I(W_2; Y_2^n | W_0) + n\epsilon_2 \\
&\leq I(W_2; Y_2^n, X_1^n | W_0) + n\epsilon_2 \\
&\overset{(a)}{=} I(W_2; Y_2^n | X_1^n, W_0) + n\epsilon_2 \\
&= H(Y_2^n | X_1^n, W_0) - H(Y_2^n | X_1^n, W_0, W_2) + n\epsilon_2 \\
&\overset{(b)}{\leq} H(Y_2^n | X_1^n, W_0) - H(Y_2^n | X_1^n, W_0, X_2^n) + n\epsilon_2 \\
&\overset{(c)}{\leq} \sum_{i=1}^{n} [H(Y_{2,i} | U_i, X_{1,i}) - H(Y_{2,i} | U_i, X_{1,i}, X_{2,i})] + n\epsilon_2 \\
&= \sum_{i=1}^{n} I(X_{2,i}; Y_{2,i} | U_i, X_{1,i}) + n\epsilon_2,
\end{aligned}
$$

where (a) is due to the independence of W_2 and X_1^n, (b) is from $(W_0, W_2) \Leftrightarrow (W_0, X_2^n) \Leftrightarrow (Y_2^n)$ in Lemma 3, and (c) comes from the same definition aforesaid of $U_i = Y_2^{i-1}, X_1^{n\backslash i}, W_0$.

Next, we prove the outer bound for the sum rate $R_0 + R_1$. We have

$$\begin{aligned}
nR_0 + nR_1 =& H(W_0, W_1) \\
\leq& I(W_0, W_1; Y_1^n) + n\epsilon_1 \\
=& H(Y_1^n) - H(Y_1^n|W_0, W_1) + n\epsilon_1 \\
\overset{(a)}{=}& H(Y_1^n) - H(Y_1^n|W_0, X_1^n) + n\epsilon_1 \\
=& \sum_{i=1}^{n} \left[\begin{array}{l} H(Y_{1,i}|Y_1^{i-1}) \\ -H(Y_{1,i}|Y_1^{i-1}, X_1^{n\backslash i}, W_0, X_{1,i}) \end{array} \right] + n\epsilon_1 \\
\overset{(b)}{=}& \sum_{i=1}^{n} \left[\begin{array}{l} H(Y_{1,i}|Y_1^{i-1}) \\ -H(Y_{1,i}|Y_2^{i-1}, X_1^{n\backslash i}, W_0, X_{1,i}) \end{array} \right] + n\epsilon_1 \\
\overset{(c)}{\leq}& \sum_{i=1}^{n} [H(Y_{1,i}) - H(Y_{1,i}|U_i, X_{1,i})] + n\epsilon_1 \\
=& \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{1,i}) + n\epsilon_1.
\end{aligned}$$

(a) results from $(W_0, W_1) \Leftrightarrow (W_0, X_1^n) \Leftrightarrow (Y_1^n)$ (Lemma 3), (b) results from $X_2 \Leftrightarrow Y_2 \Leftrightarrow Y_1$, given X_1 in (4.2), and (c) results from the aforementioned definition of $U_i = Y_2^{i-1}, X_1^{n\backslash i}, W_0$. Note that the choice of auxiliary random variables automatically satisfies the constraints imposed on them in Theorem 6.

Finally, (4.7) comes from the restriction on the (R_0, R_1), which is described in Section 4.2. As discussed earlier, this constraint captures the extent of (partial) cognitive information available at the cognitive transmitter in the system.

An intuitive understanding of the variables in this theorem is as follows: We have T, which is an auxiliary time-sharing variable. And, U represents

(W_0, T), where W_0 is the shared message between cognitive and legitimate radios, and V corresponds to (W_1, T), where message W_1 is available only to the legitimate user. Then, the outer bound presented here is a generalized version of the one in [52] with two auxiliary random variables (U and V) instead of one.

4.3.2 Gaussian Partially Cognitive Radio Channel

For the Gaussian case, the weak interference constraint can be translated into the requirement that $b < 1$ in (4.1). With the condition $b < 1$, the conditional Markov chain for the weak interference channel, $X_2 \Leftrightarrow Y_2 \Leftrightarrow Y_1$, given X_1 in (4.2) is satisfied. Thus, a proof methodology analogous to the one adopted for the discrete memoryless case will result in a rate region similar to that in Theorem 6 for the Gaussian case.

Next, we present three lemmas that will prove essential in obtaining a closed-form evaluation of the outer bound.

Lemma 4 (Lemma 1 in [84]). Let $X_1, X_2, ..., X_k$ be arbitrarily distributed zero-mean random variables with covariance matrix K, and $X_1^*, X_2^*, ..., X_k^*$ be zero-mean Gaussian distributed random variables with the same covariance matrix K. Let S be any subset of $\{1, 2, ..., k\}$ and $\widetilde{S}$ be its complement. Then,

$$h(X_S|X_{\widetilde{S}}) \leq h(X_S^*|X_{\widetilde{S}}^*). \tag{4.9}$$

Lemma 5. Let X_1, X_2, V be an arbitrarily distributed zero-mean random variables with covariance matrix K, where X_2 and V are independent of each other.

Let X_1^*, X_2^*, V^* be the zero-mean Gaussian distributed random variables with the same covariance matrix as X_1, X_2, V. Then,

$$\mathbb{E}[X_1X_2] = \mathbb{E}[X_1^*X_2^*|V^*]. \tag{4.10}$$

Proof: Without loss of generality, X_1^* can be written as $X_1^* = W^* + cV^*$, where W^* is the zero-mean Gaussian random variable independent of V^*. Then

$$\begin{aligned}
\mathbb{E}[X_1X_2] &= \mathbb{E}[X_1^*X_2^*] \\
&= \mathbb{E}[\mathbb{E}[X_1^*X_2^*|V^*]] \\
&= \mathbb{E}[\mathbb{E}[(W^* + cV^*)X_2^*|V^*]] \\
&= \mathbb{E}[\mathbb{E}[W^*X_2^*|V^*]] + c\mathbb{E}[\mathbb{E}[V^*X_2^*|V^*]] \\
&\overset{(a)}{=} \mathbb{E}[X_1^*X_2^*|V^*] + c\mathbb{E}[V^*\mathbb{E}[X_2^*]] \\
&\overset{(b)}{=} \mathbb{E}[X_1^*X_2^*|V^*],
\end{aligned}$$

where (a) results from the independence of X_2^* and V^*. And, (b) results from the fact that X_2^* is zero-mean.

Lemma 6. Random variables in Lemma 5, X_1^*, X_2^*, and V^* satisfy the following equation:

$$\mathbb{E}[X_1^*X_2^*|V^*] \leq (\mathbb{E}[(X_1^*)^2|V^*])^{\frac{1}{2}}(\mathbb{E}[(\mathbb{E}[X_2^*|X_1^*])^2])^{\frac{1}{2}}.$$

Proof: Note that

$$\begin{aligned}\mathbb{E}[X_1^*X_2^*|V^*] &\overset{(a)}{=} \mathbb{E}[\mathbb{E}[X_1^*X_2^*|V^*, X_1^*]] \\ &\overset{(b)}{=} \mathbb{E}[X_1^*\mathbb{E}[X_2^*|V^*, X_1^*]|V^*] \\ &\overset{(c)}{\leq} (\mathbb{E}[(X_1^*)^2|V^*])^{\frac{1}{2}}(\mathbb{E}[(\mathbb{E}[X_2^*|V^*, X_1^*])^2])^{\frac{1}{2}} \\ &\overset{(d)}{\leq} (\mathbb{E}[(X_1^*)^2|V^*])^{\frac{1}{2}}(\mathbb{E}[(\mathbb{E}[X_2^*|X_1^*])^2])^{\frac{1}{2}},\end{aligned}$$

where (a) comes from the law of iterated expectations, (b) comes from the independence of X_2^* and V^*, (c) from the Cauchy-Schwartz inequality, and (d) comes from the fact that entropy can only be reduced by conditioning.

Definition 1. *Define the rate region* $\mathcal{R}_{out}^{\alpha,\beta_1,\beta_2}$ *to be the convex hull of all rate triplets* (R_0, R_1, R_2) *satisfying*

$$\begin{aligned}&R_0 \leq \frac{1}{2}\log\left(1 + \frac{\beta_1 P_1 + b^2(1-\alpha)P_2 + 2b\sqrt{(\beta_2(1-\alpha)P_1P_2)}}{(1+b^2\alpha P_2)}\right), \\ &R_1 \leq \frac{1}{2}\log\left(1 + (1-\beta_2(1-\alpha))\,P_1\right), \\ &R_0 + R_1 \leq \frac{1}{2}\log\left(1 + \frac{P_1 + b^2(1-\alpha)P_2 + 2b\sqrt{(\beta_2(1-\alpha)P_1P_2)}}{(1+b^2\alpha P_2)}\right), \\ &R_2 \leq \frac{1}{2}\log(1+\alpha P_2), \\ &R_1 \geq \mu R_0,\end{aligned} \tag{4.11}$$

for some $\alpha \in [0,1]$*,* $\beta_1 \in [0,1]$*, and* $\beta_2 \in [0, \beta_1]$*.*

Definition 2. *Define the rate region* $\mathcal{R}_{out}$ *to be convex hull of the union of rate region* $\mathcal{R}_{out}^{\alpha,\beta}$*:*

$$\mathcal{R}_{out} \triangleq \overline{\bigcup_{0\leq\alpha,\beta_1\leq 1, 0\leq\beta_2\leq\beta_1} \mathcal{R}_{out}^{\alpha,\beta_1,\beta_2}}. \tag{4.12}$$

We denote $\mathcal{C}$ to be the capacity region of the Gaussian weak partially cognitive radio channel. An outer bound for $\mathcal{C}$ is obtained as follows.

Theorem 7. $\mathcal{R}_{out}$ is an outer bound of the capacity region for the Gaussian weak partially cognitive radio channel

$$\mathcal{C} \subset \mathcal{R}_{out}.$$

Proof: We start from the rate region in Theorem 6:

$$\begin{aligned} R_0 &\le I(U, X_1; Y_1|V) = h(Y_1|V) - h(Y_1|V, U, X_1) \\ &= h(Y_1|V) - h(Y_1|U, X_1), && (4.13) \\ R_1 &\le I(X_1; Y_1|X_2) = h(Y_1|X_2) - h(N_1), && (4.14) \\ R_0 + R_1 &\le I(U, X_1; Y_1) = h(Y_1) - h(Y_1|U, X_1), && (4.15) \\ R_2 &\le I(X_2; Y_2|U, X_1) = h(Y_2|U, X_1) - h(N_2). && (4.16) \end{aligned}$$

(4.13) follows from the Markov chain, $V \Leftrightarrow (U, X_1) \Leftrightarrow Y_1$. First, we set

$$h(Y_2|U, X_1) = \frac{1}{2}\log(2\pi e(1 + \alpha P_2)), \tag{4.17}$$

without loss of generality for some $\alpha \in [0, 1]$. Note that

$$\begin{aligned} Y_1 &= b(X_2 + Z_1) + X_1 + Z', \\ h(Y_1|U, X_1) &= h(b(X_2 + Z_1) + Z'|U, X_1), && (4.18) \end{aligned}$$

where $b < 1$ because legitimate receiver faces a weak interference, and Z' is a Gaussian distributed random variable with variance $1 - b^2$. By entropy power inequality (EPS)[87], we have

$$
\begin{aligned}
2^{2h(Y_1|U,X_1)} &\geq 2^{2h(bY_2|U,X_1)} + 2^{2h(Z')}. \\
&= b^2 2^{2h(Y_2|U,X_1)} + 2\pi e(1-b^2) \\
&= 2\pi e(1 + b^2 \alpha P_2),
\end{aligned}
$$

which yields

$$h(Y_1|U,X_1) \geq \frac{1}{2}\log(2\pi e(1 + b^2\alpha P_2)). \tag{4.19}$$

Next, we need to bound $h(Y_1)$, $h(Y_1|V)$, and $h(Y_1|X_2)$. Note that, by setting $h(Y_2|U,X_1) = \frac{1}{2}\log(2\pi e(1+\alpha P_2))$, we have the following result:

$$
\begin{aligned}
h(Y_2|U,X_1) &\leq h(X_2 + Z_2|X_1) \\
&\leq h(X_2^* + Z_2|X_1^*) \\
&= \frac{1}{2}\log(2\pi e(1 + \mathrm{Var}(X_2^*|X_1^*))),
\end{aligned} \tag{4.20}
$$

where $\mathrm{Var}(\cdot|\cdot)$ denotes the conditional covariance. Combining (4.17) with (4.20), we obtain the bound

$$\mathrm{Var}(X_2^*|X_1^*) \geq \alpha P_2. \tag{4.21}$$

Also,

$$\mathrm{Var}(X_2^*|X_1^*) = \mathbb{E}[(X_2^*)^2] - \mathbb{E}[(\mathbb{E}[X_2^*|X_1^*])^2]. \tag{4.22}$$

From (4.21) and (4.22), we obtain

$$\mathbb{E}[(\mathbb{E}[X_2^*|X_1^*])^2] \leq (1-\alpha)P_2. \tag{4.23}$$

Note that

$$\mathbb{E}[(X_1^*)^2|V^*] \leq P_1, \tag{4.24}$$

since conditioning only reduces the entropy. Again, we set $\mathbb{E}[(X_1^*)^2|V^*] = \beta_1 P_1$ for some $\beta_1 \in [0,1]$ without loss of generality. Now combining Lemma 5, Lemma 6, and the aforementioned result, (4.23), we have

$$\mathbb{E}[X_1 X_2] \leq \sqrt{\beta_1 P_1}\sqrt{(1-\alpha)P_2}. \tag{4.25}$$

We can set

$$\mathbb{E}[X_1 X_2] = \sqrt{\beta_2 P_1}\sqrt{(1-\alpha)P_2}, \tag{4.26}$$

where $\beta_2 \in [0, \beta_1]$. Therefore, we obtain the bound for $h(Y_1)$ as

$$\begin{aligned} h(Y_1) &\leq \frac{1}{2}\log\left(2\pi e\left(\begin{array}{l} 1+\mathrm{Var}(X_1)+b^2\,\mathrm{Var}(X_2) \\ +2b\mathbb{E}[X_1X_2] \end{array}\right)\right) \\ &= \frac{1}{2}\log\left(2\pi e\left(\begin{array}{l} 1+P_1+b^2P_2 \\ +2b\sqrt{\beta_2(1-\alpha)P_1P_2} \end{array}\right)\right). \end{aligned} \tag{4.27}$$

For $h(Y_1|V)$, note that (Y_1^*, V^*) has the same covariance matrix as (Y_1, V) if $Y_1 = X_1^* + bX_2^*$. Also, Y_1 is a mean zero Gaussian distributed random variable. Thus,

$$\begin{aligned} h(Y_1|V) \leq& h(Y_1^*|V^*) \\ =& h(X_1^* + bX_2^* + Z_1|V^*) \\ =& \frac{1}{2}\log\left(2\pi e\left(\begin{array}{l} 1+\mathrm{Var}(X_1^*|V^*) \\ +b^2\,\mathrm{Var}(X_2^*|V^*) \\ +2b\mathbb{E}[X_1^*X_2^*|V^*] \end{array}\right)\right) \\ \leq& \frac{1}{2}\log\left(2\pi e\left(\begin{array}{l} 1+\beta_1P_1+b^2P_2 \\ +2b\sqrt{(\beta_2(1-\alpha)P_1P_2)} \end{array}\right)\right). \end{aligned} \tag{4.28}$$

For $h(Y_1|X_2)$,

$$\begin{aligned}
h(Y_1|X_2) =& h(X_1 + bX_2 + Z_1|X_2) \\
=& h(X_1 + Z_1|X_2) \\
\leq& h(X_1^* + Z_1|X_2^*) \\
=& \frac{1}{2}\log\left(2\pi e\left(1 + \mathrm{Var}(X_1^*|X_2^*)\right)\right) && (4.29) \\
=& \frac{1}{2}\log\left(2\pi e\left(1 + P_1 - \frac{\mathbb{E}[X_1^* X_2^*]^2}{P_2}\right)\right) && (4.30) \\
=& \frac{1}{2}\log\left(2\pi e\left(1 + \left(1 - \beta_2(1-\alpha)\right)P_1\right)\right) && (4.31)
\end{aligned}$$

The intuition behind this region is as follows: β represents the fraction of power assigned to the message W_0 at the legitimate transmitter, and α_2 represents the fraction of power assigned to the message W_1 at the cognitive transmitter. The outer bound structure dictates that W_2 be decoded without interference, and that W_0 and W_1 be decoded treating W_2 as "interference" when being decoded.

It is necessary that we examine whether the outer bound on the capacity region of the Gaussian channel model is tight. A Gaussian partially cognitive radio channel includes the Gaussian interference channel. As the capacity region of the Gaussian interference channel remains as an open problem, we expect that obtaining the capacity region of the partially cognitive radio with the Gaussian interference channel is difficult. Given this, we analyze our outer bound result, which is represented in Theorem 7, in two extremes: when $\mu = 0$ and $\mu \to \infty$. If $\mu = 0$, by setting both β_1 and β_2 to be 1, we obtain the rate region, which encompass all other rate regions

$$R_{out}^{\alpha,1,1} \supseteqq R_{out}^{\alpha,\beta_1,\beta_2} \quad \forall \beta_1 \in [0,1], \beta_2 \in [0,\beta_1].$$

It implies that the legitimate transmitter only transmits the shared message, W_0, and no power is allocated for transmitting W_1. Thus, the channel becomes the Gaussian fully cognitive radio channel when μ value becomes 0. Let R_L be the sum rate of the legitimate transmission, which is $R_0 + R_1$ in 4.11, and R_C be the rate for the cognitive transmission, which corresponds to R_2. Then, we develop an outer bound on the capacity region for rate pair (R_L, R_C) of the Gaussian partially cognitive radio. We denote $\mathcal{C}_{\mu=0}$ to be the capacity region of the Gaussian weak partially cognitive radio channel, when $\mu = 0$.

Definition 3. *Define the rate region $\mathcal{R}_{\mu=0}^{\alpha}$ to be the convex hull of all rate pair (R_L, R_C) satisfying*

$$\begin{aligned} R_L &\le \frac{1}{2}\log\left(1 + \frac{P_1 + b^2(1-\alpha)P_2 + 2b\sqrt{(1-\alpha)P_1P_2}}{(1+b^2\alpha P_2)}\right), \\ R_C &\le \frac{1}{2}\log(1+\alpha P_2), \end{aligned} \tag{4.32}$$

for some $\alpha \in [0,1]$.

Also, we have the following.

Definition 4. *Define the rate region $\mathcal{R}_{\mu=0}$ to be convex hull of the union of rate region $\mathcal{R}_{\mu=0}^{\alpha}$:*

$$\mathcal{R}_{\mu=0} \triangleq \overline{\bigcup_{0\le\alpha\le1} \mathcal{R}_{\mu=0}^{\alpha}}. \tag{4.33}$$

Then, we present an outer bound on the rate region of partially cognitive radio in Gaussian interference channel when $\mu = 0$ in the following corollary:

Corollary 1. $\mathcal{R}_{\mu=0}$ is an outer bound of the capacity region for the Gaussian weak partially cognitive radio channel when $\mu = 0$:

$$\mathcal{C}_{\mu=0} \subset \mathcal{R}_{\mu=0}.$$

Proof: Proof of this can be obtained by substituting $R_0 + R_1$ with R_L and R_C with R_2, and removing the $R_1 > \mu R_0$ requirement in (4.11).

The outer bound on the capacity region, $\mathcal{R}_{\mu=0}$, is equivalent to the outer bound on the capacity region for the fully cognitive radio in [52]. The other extreme of partially cognitive radio is when $\mu \to \infty$. In this case, constraint $R_1 \geq \mu R_0$, becomes $R_0 = 0$. This means that there is no shared message W_0. Since there is no cognitive information, the model reduces to a Gaussian interference channel. Since R_0 is zero, the rate region is defined from the rates corresponding to W_1 and W_2. We develop an outer bound on the capacity region in this extreme as follows:

Definition 5. *Define the rate region* $\mathcal{R}^{\alpha,\beta}_{\mu\to\infty}$ *to be the convex hull of all rate pair* (R_1, R_2) *satisfying*

$$\begin{aligned} R_1 &\leq \min\left(\begin{array}{l} \frac{1}{2}\log\left(1+(1-\beta)P_1\right), \\ \frac{1}{2}\log\left(1+\frac{P_1+b^2(1-\alpha)P_2+2b\sqrt{\beta P_1 P_2}}{(1+b^2\alpha P_2)}\right) \end{array} \right), \\ R_2 &\leq \frac{1}{2}\log(1+\alpha P_2), \end{aligned} \tag{4.34}$$

for some $\alpha \in [0,1]$ *and* $\beta \in [0, 1-\alpha]$.

In addition, we have the following.

Definition 6. *Define the rate region $\mathcal{R}_{\mu\to\infty}$ to be convex hull of the union of rate region $\mathcal{R}^{\alpha,\beta}_{\mu\to\infty}$:*

$$\mathcal{R}_{\mu\to\infty} \triangleq \overline{\bigcup_{0\le\alpha\le 1, 0\le\beta\le 1-\alpha} \mathcal{R}^{\alpha,\beta}_{\mu\to\infty}}. \tag{4.35}$$

Then, we develop an outer bound on the capacity region of the Gaussian interference channel when $\mu \to \infty$ using the following corollary:

Corollary 2. $\mathcal{R}_{\mu\to\infty}$ is an outer bound of the capacity region for the Gaussian weak partially cognitive radio channel when $\mu \to \infty$:

$$\mathcal{C}_{\mu\to\infty} \subset \mathcal{R}_{\mu\to\infty}.$$

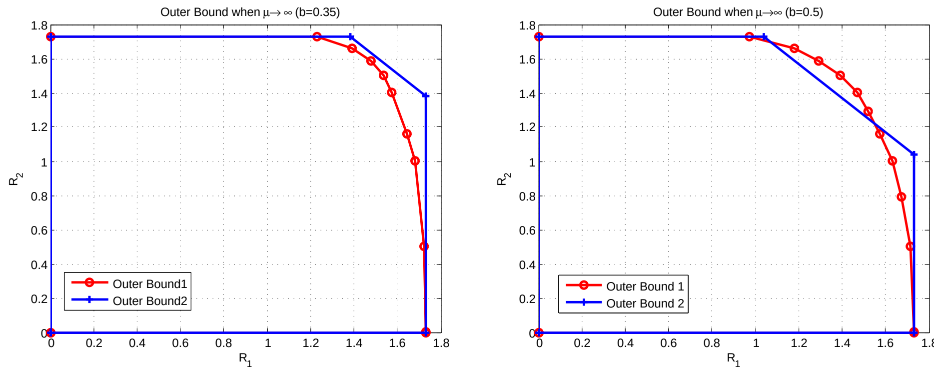
Proof: Proof can be obtained by making R_0 to be 0 in Equation (4.11).

We compare this outer bound with the known outer bound in [75]. Since $b < 1$, we have the outer bound for Gaussian interference channel as follows:

$$\mathcal{C}_{\mu\to\infty} \subset \mathcal{R}_{Carleial}, \tag{4.36}$$

where $\mathcal{R}_{carleial}$ is defined as the rate region of the rate pair (R_1, R_2) which satisfies

$$\begin{aligned} R_1 &\le \frac{1}{2}\log(1+P_1), \\ R_2 &\le \frac{1}{2}\log(1+P_2), \\ R_1 + R_2 &\le \frac{1}{2}\log\left(1 + \frac{1}{b^2}P_1 + P_2\right). \end{aligned} \tag{4.37}$$



(a) $b = 0.35$

(b) $b = 0.5$

Figure 4.3: Outer bound comparison when $\mu \to \infty$

Note that a comparison between the two depends on the channel parameters. $\mathcal{R}_{Carleial}$ can include our outer bound in some cases, and our outer bound can include $\mathcal{R}_{Carleial}$ in some other cases. Next, we present examples of these two different cases.

We compare our outer bound when $\mu \to \infty$ with the outer bound in [75] when $P_1 = P_2 = 10$. In Figure 4.3, outer bound 1 represents our outer bound, $\mathcal{R}_{\mu\to\infty}$, and outer bound 2 is Carleial's outer bound, $\mathcal{R}_{Carleial}$. It is not always the case that our outer bound is tighter than the Carleial outer bound as seen in Figure 4.3(b), which is when $b = 0.5$. However, we show that the outer bound $\mathcal{R}_{\mu\to\infty}$ is tighter than Carleial's outer bound in some cases. Figure 4.3(a) makes analysis when $b = 0.35$. In this case, our outer bound encompasses Carleial's outer bound.

From analyzing these two extreme cases, we conclude that we have a

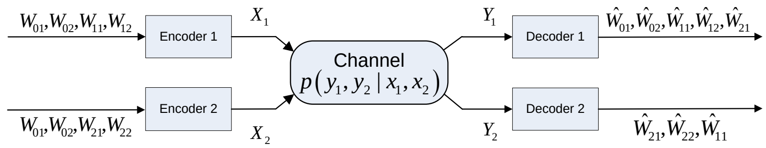


Figure 4.4: Discrete memoryless partial cognitive radio channel

nontrivial outer bound on the capacity region of a Gaussian partially cognitive radio channel.

4.4 The Achievable Region

In this section, we study an achievable region for the partially cognitive radio in the interference channel.

4.4.1 Discrete Memoryless Partially Cognitive Radio Channels

Here, we first develop an achievable region for a general discrete memoryless partially cognitive radio channel. We subdivide W_0, W_1 and W_2 into two components each. We split W_0 into W_{01} and W_{02}, W_1 into W_{11} and W_{12}, and W_2 into W_{21} and W_{22}, where W_{01}, W_{11} and W_{21} are decoded at both receivers, and W_{02}, W_{12} and W_{22} are decoded only at the intended receiver. Figure 4.4 shows the modified channel model. Here, M_0, N_0, M_1, N_1, M_2, and N_2 are auxiliary random variables, which bear information on W_{01}, W_{02}, W_{11}, W_{12}, W_{21}, and W_{22} respectively, and we define $R_{01}, R_{02}, R_{11}, R_{12}, R_{21}$, and R_{22} to be the rate for W_{01}, W_{02}, W_{11}, W_{12}, W_{21}, and W_{22} respectively. Then, we have following lemma that characterizes our achievable region:

Lemma 7. Define Q as

$$Q \triangleq (Y_1, Y_2, X_1, X_2, M_0, N_0, M_1, N_1, M_2, N_2)$$

, and let $\mathcal{P}$ be the set of distributions on Q that can be decomposed into the form as follows:

$$\begin{aligned}\mathcal{P} =& p(m_0)p(n_0)p(m_1)p(n_1)p(m_2|m_0,n_0)p(n_2|m_0,n_0)\\ & p(x_1|m_0,n_0,m_1,n_1)p(x_2|m_2,n_2)\\ & p(y_1|x_1,x_2)p(y_2|x_1,x_2).\end{aligned}$$

For any $Q \in \mathcal{P}$, let $S(Q)$ be the set of all rate 6-tuple $(R_{01}, R_{02}, R_{11}, R_{12}, R_{21}, R_{22})$ of nonnegative real numbers such that there exists a nonnegative real-valued pair (L_{21}, L_{22}) satisfying

$$\begin{aligned}
&R_{21} \leq L_{21} - I(M_2; M_0, N_0, M_1),\\
&R_{22} \leq L_{22} - I(N_2; M_0, N_0, M_1),\\
&R_{01} \leq I(M_0; Y_1, N_0, M_1, N_1, M_2),\\
&R_{02} \leq I(N_0; Y_1, M_0, M_1, N_1, M_2),\\
&R_{11} \leq I(M_1; Y_1, M_0, N_0, N_1, M_2),\\
&R_{12} \leq I(N_1; Y_1, M_0, N_0, M_1, M_2),\\
&L_{21} \leq I(M_2; Y_1, M_0, N_0, M_1, N_1),\\
&R_{11} \leq I(M_1; Y_2, M_2, N_2),\\
&L_{21} \leq I(M_2; Y_2, M_1, N_2),\\
&L_{22} \leq I(N_2; Y_2, M_1, M_2),\\
&R_{01} + R_{02} \leq I(M_0, N_0; Y_1, M_1, N_1, M_2),\\
&R_{01} + R_{11} \leq I(M_0, M_1; Y_1, N_0, N_1, M_2),\\
&R_{01} + R_{12} \leq I(M_0, N_1; Y_1, N_0, M_1, M_2),\\
&R_{01} + L_{21} \leq I(M_0, M_2; Y_1, N_0, M_1, N_1),\\
&R_{02} + R_{11} \leq I(N_0, M_1; Y_1, M_0, N_1, M_2),\\
&R_{02} + R_{12} \leq I(N_0, N_1; Y_1, M_0, M_1, M_2),\\
&R_{02} + L_{21} \leq I(N_0, M_2; Y_1, M_0, M_1, N_1),\\
&R_{11} + R_{12} \leq I(M_1, N_1; Y_1, M_0, N_0, M_2),
\end{aligned}$$

$$
\begin{aligned}
&R_{11} + L_{21} \leq I(M_1, M_2; Y_1, M_0, N_0, N_1),\\
&R_{12} + L_{21} \leq I(N_1, M_2; Y_1, M_0, N_0, M_1),\\
&R_{11} + L_{21} \leq I(M_1, M_2; Y_2, N_2),\\
&R_{11} + L_{22} \leq I(M_1, N_2; Y_2, M_2),\\
&L_{21} + L_{22} \leq I(M_2, N_2; Y_2, M_1),\\
&R_{01} + R_{02} + R_{11} \leq I(M_0, N_0, M_1; Y_1, N_1, M_2),\\
&R_{01} + R_{02} + R_{12} \leq I(M_0, N_0, N_1; Y_1, M_1, M_2),\\
&R_{01} + R_{02} + L_{21} \leq I(M_0, N_0, M_2; Y_1, M_1, N_1),\\
&R_{01} + R_{11} + R_{12} \leq I(M_0, M_1, N_1; Y_1, N_0, M_2),\\
&R_{01} + R_{11} + L_{21} \leq I(M_0, M_1, M_2; Y_1, N_0, N_1),\\
&R_{01} + R_{12} + L_{21} \leq I(M_0, N_1, M_2; Y_1, N_0, M_1),\\
&R_{02} + R_{11} + R_{12} \leq I(N_0, M_1, N_1; Y_1, M_0, M_2),\\
&R_{02} + R_{11} + L_{21} \leq I(N_0, M_1, M_2; Y_1, M_0, N_1),\\
&R_{02} + R_{12} + L_{21} \leq I(N_0, N_1, M_2; Y_1, M_0, M_1),\\
&R_{11} + R_{12} + L_{21} \leq I(M_1, N_1, M_2; Y_1, M_0, N_0),\\
&R_{11} + L_{21} + L_{22} \leq I(M_1, M_2, N_2; Y_2),\\
&R_{01} + R_{02} + R_{11} + R_{12} \leq I(M_0, N_0, M_1, N_1; Y_1, M_2),\\
&R_{01} + R_{02} + R_{11} + L_{21} \leq I(M_0, N_0, M_1, M_2; Y_1, N_1),\\
&R_{01} + R_{02} + R_{12} + L_{21} \leq I(M_0, N_0, N_1, M_2; Y_1, M_1),\\
&R_{01} + R_{11} + R_{12} + L_{21} \leq I(M_0, M_1, N_1, M_2; Y_1, N_0),\\
&R_{02} + R_{11} + R_{12} + L_{21} \leq I(N_0, M_1, N_1, M_2; Y_1, M_0),\\
&R_{01} + R_{02} + R_{11} + R_{12} + L_{21}\\
&\leq I(M_0, N_0, M_1, N_1, M_2; Y_1),\\
&R_{11} + R_{12} \geq \mu(R_{01} + R_{02}).
\end{aligned}
\tag{4.38}
$$

Let S be the closure of $\bigcup_{Q \in \mathcal{P}} S(Q)$. Then, any element of S is achievable.

Proof: We prove the lemma by showing the achievability of the interior elements of $S(Z)$ for each $Q \in \mathcal{P}$ by fixing $Q = (Y_1, Y_2, X_1, X_2, M_0, N_0, M_1, N_1, M_2, N_2)$ and taking any $(R_{01}, R_{02}, R_{11}, R_{12}, R_{21}, R_{22})$ and (L_{21}, L_{22}) satisfying the constraints of the lemma. And, let some distribution on Q which satisfy the form in the theorem is given.

Encoding Strategy at the Legitimate Transmitter: The legitimate transmitter randomly generates the codebook, following the given distribution. Specifically, we generate the following codebooks:

1. For every message $W_{01} \in \{1, \ldots, 2^{n(R_{01}-6\epsilon)}\}$, generate n-sequences m_0^n i.i.d. according to $\prod_{t=1}^{n} p\left(m_0^{(t)}\right)$.
2. For every message $W_{02} \in \{1, \ldots, 2^{n(R_{02}-6\epsilon)}\}$, generate n-sequences n_0 i.i.d. according to $\prod_{t=1}^{n} p\left(n_0^{(t)}\right)$.
3. For every message $W_{11} \in \{1, \ldots, 2^{n(R_{11}-6\epsilon)}\}$, generate n-sequences m_1^n i.i.d. according to $\prod_{t=1}^{n} p\left(m_1^{(t)}\right)$.
4. For every message $W_{12} \in \{1, \ldots, 2^{n(R_{12}-6\epsilon)}\}$, generate n-sequences n_1^n i.i.d. according to $\prod_{t=1}^{n} p\left(n_1^{(t)}\right)$.

From these codebooks, we look up the message set $(w_{01}, w_{02}, w_{11}, w_{12})$, and choose a codeword for each message. The Legitimate transmitter forms the net transmit vector $x_1^n = f(m_0^n, m_1^n, m_1^n, n_1^n)$, and communicates it.

Encoding Strategy at the Cognitive Transmitter: The cognitive transmitter generates a random codebook, following the given distribution. Specifically, we generate the following codebooks:

1. Generate $2^{n(L_{21}-6\epsilon)}$ n-sequences m_1^n i.i.d. according to $\prod_{t=1}^{n} p\left(m_2^{(t)}\right)$, and place in $2^{n(R_{21}-6\epsilon)}$ bins uniformly.

2. Generate $2^{n(L_{22}-6\epsilon)}$ n-sequences m_1^n i.i.d. according to $\prod_{t=1}^{n} p\left(n_2^{(t)}\right)$, and place in $2^{n(R_{22}-6\epsilon)}$ bins uniformly.

Note that

$$p\left(m_2^{(t)}\right) = \sum_{m_0^{(t)},n_0^{(t)}} p\left(m_2^{(t)}|m_0^{(t)}, n_0^{(t)}\right),$$
$$p\left(n_2^{(t)}\right) = \sum_{m_0^{(t)},n_0^{(t)}} p\left(n_2^{(t)}|m_0^{(t)}, n_0^{(t)}\right).$$

Here, the codebooks corresponding to the common messages are assumed be known to the cognitive transmitter, and message sets at the cognitive transmitter are $W_{21} \in \{1, \ldots, 2^{n(R_{21}-6\epsilon)}\}$ and $W_{22} \in \{1, \ldots, 2^{n(R_{22}-6\epsilon)}\}$. The cognitive transmitter, upon obtaining the message set (w_{21}, w_{22}), looks in bins w_{21} and w_{22} for sequences m_2^n and n_2^n such that (m_0^n, m_1^n, m_2^n) and (m_0^n, m_1^n, n_2^n) are jointly typical. Then, it generates x_2^n i.i.d. according to $\prod_{t=1}^{n} p\left(x_2^{(t)}|m_2^{(t)}, n_2^{(t)}\right)$, and transmits it. Next, we describe the decoding strategy and the rate constraints associated with the two receivers.

Decoding Strategy at the Legitimate Receiver: The legitimate receiver decodes W_{01}, W_{02}, W_{11}, W_{12}, and W_{21} based on strong joint typicality. Upon receiving y_1^n, the legitimate receiver performs jointly typical decoding as:

$$\left\{ \ (m_0^n, n_0^n, m_1^n, n_1^n, m_2^n) : (y_1^n, m_0^n, n_0^n, m_1^n, n_1^n, m_2^n) \in A_\epsilon^n(Y_1, M_0, N_1, M_1, N_1, M_2) \ \right\}$$

if all $(m_0^n, \cdot, \cdot, \cdot, \cdot)$ in this set have the same message index, it decodes $\hat{w}_{01}$ to be $B(m_0^n)$, where $B(m_0^n)$ is a message index for m_0^n. Similarly, it decode $\hat{w}_{02}$, $\hat{w}_{11}$, $\hat{w}_{12}$, and $\hat{w}_{21}$ to be $B(n_0^n)$, $B(m_1^n)$, $B(n_1^n)$, and $B(m_2^n)$ respectively. If this typicality test fails, it declare an errors.

Decoding strategy at cognitive receiver: The cognitive receiver decodes W_{21}, W_{22}, and W_{11} based on strong joint typicality. Upon receiving y_2^n, the cognitive receiver examines joint typicality by finding the set of codeword

$$\{(m_2^n, n_2^n, m_1^n) : (y_2^n, m_2^n, n_2^n, m_1^n) \in A_\epsilon^n(Y_2, M_2, N_2, M_1)\}$$

if all $(m_2^n, \cdot, \cdot, \cdot)$ in this set have the same message index, it decodes $\hat{w}_{21}$ to be $B(m_2^n)$, where $B(m_2^n)$ bin index for m_2^n. Likewise, it decodes $\hat{w}_{22}$ and $\hat{w}_{11}$ to be $B(n_2^n)$ and $B(m_1^n)$ respectively. If this typicality test fails, it declares an error. We defer the probability of error analysis to the Appendix. Note that $R_{11} + R_{12} \geq \mu(R_{01} + R_{02})$ comes from the partially cognitive radio condition.

In addition, we have the following lemma:

Lemma 8. Any rate 6-tuple $(R_{01}, R_{02}, R_{11}, R_{12}, R_{21}, R_{22})$ that satisfies

$$\begin{aligned}
&R_{11} + R_{12} \leq I(X_1; Y_1) \\
&R_{01} + R_{02} + R_{11} + R_{12} \leq I(X_1, X_2; Y_1) \\
&R_{21} = 0 \\
&R_{22} = 0 \\
&R_{11} + R_{12} \geq \mu(R_{01} + R_{02}),
\end{aligned}$$

is achievable for the discrete memoryless partially cognitive radio.

Proof of this lemma is straightforward and thus skipped. From this, we have the following theorem:

Theorem 8. The convex hull of the points of Lemma 7 and Lemma 8 is achievable.

Proof: Proof follows using standard time-sharing techniques and the fact that the achievable region is the closure of achievable rates.

This achievable rate region can be compared with achievable rate region in [48] and that in [72] in two extremes. If we allow $H(M_1, N_1) = 0$, our achievable rate region becomes the achievable rate region in [48], i.e., the one for fully cognitive radio. Thus, we can match the achievable rate region of fully cognitive radio when $\mu = 0$. Also, if we set M_0 and M_1 such that $H(M_0, N_0) = 0$, our achievable rate region becomes the same as that in [72]. Thus, we can match nontrivial achievable rate regions in two extreme cases.

4.4.2 Gaussian Partially Cognitive Radio Channel

In this section, we describe an achievable region for the Gaussian channel model described in (4.1). In deriving the achievable region, we combine dirty paper coding [86], and Han-Kobayashi coding [72]. The reason for using this combination is to bring the regular interference channel results together with those for interference channels with degraded message sets. Thus, as $\mu \to \infty$, the channel becomes an interference channel, and we desire that our coding scheme reduces to Han-Kobayashi coding [73]. Also, as $\mu \to 0$, the channel resembles a cognitive radio

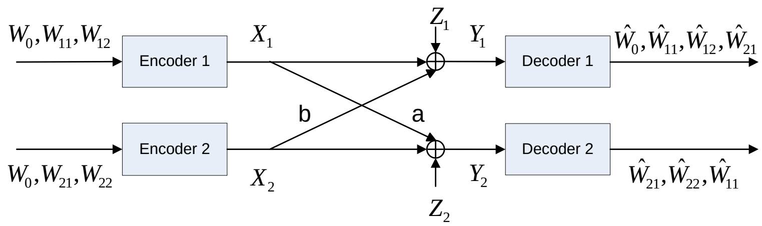


Figure 4.5: The Gaussian partially cognitive radio channel

with full knowledge of legitimate transmitter's message sets. In such a case, we desire that our strategy reduces to dirty paper coding which is known to be optimal [48], [52]. Figure 4.5 presents the messages sets to be encoded and decoded at each transmitter and receiver in this system.

In our achievable strategy, we require that the legitimate transmitter encode messages W_0, W_{11}, and W_{12} using Gaussian codebooks and then superpose them to obtain its transmit sequence. Here, W_0 is the common message shared between legitimate and cognitive transmitters. Further, W_{11} and W_{12} correspond to a split of W_1 (as shown in Figure 4.1 and 4.2). W_{12} is a public message intended to be decoded by both the legitimate and cognitive receivers. W_{11} is a private message intended for the legitimate receiver alone. The cognitive transmitter allocates a portion of its power to aid the communication of W_0 to the legitimate receiver. The remaining power is used to communicate its own message W_2. Again, W_2 is subdivided into a public message W_{21}, and a private message W_{22}. The cognitive transmitter encodes message W_{22} using dirty paper coding treating the codewords corresponding to W_0 as noncausally known interference.

Let $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3 > 0$ such that

$$\alpha_1 + \alpha_2 + \alpha_3 = 1, \qquad \beta_1 + \beta_2 + \beta_3 = 1.$$

We define the function $L : R^+ \to R^+$ as $L(x) = \frac{1}{2}\log(1+x)$. Let $Q = \left(1 + \sqrt{\frac{\beta_1 P_2}{\alpha_1 P_1}}\right)^2 \alpha_1 P_1$ and $S = \left(a + \sqrt{\frac{\beta_1 P_2}{\alpha_1 P_1}}\right)^2 \alpha_1 P_1$.

We define the constants $r_0, r_1, r_2, \ldots r_{17}$ as follows:

$$
\begin{aligned}
r_0 &= L\left(\tfrac{Q}{1+b^2\beta_3P_2}\right), & r_1 &= L\left(\tfrac{\alpha_2P_1}{1+b^2\beta_3P_2}\right),\\
r_2 &= L\left(\tfrac{\alpha_3P_1}{1+b^2\beta_3P_2}\right), & r_3 &= L\left(\tfrac{b^2\beta_2P_2}{1+b^2\beta_3P_2}\right),\\
r_4 &= L\left(\tfrac{Q+\alpha_2P_1}{1+b^2\beta_3P_2}\right), & r_5 &= L\left(\tfrac{Q+\alpha_3P_1}{1+b^2\beta_3P_2}\right),\\
r_6 &= L\left(\tfrac{Q+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right), & r_7 &= L\left(\tfrac{(\alpha_2+\alpha_3)P_1}{1+b^2\beta_3P_2}\right),\\
r_8 &= L\left(\tfrac{\alpha_2P_1+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right), & r_9 &= L\left(\tfrac{\alpha_3P_1+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right),
\end{aligned}
$$

$$
\begin{aligned}
r_{10} &= L\left(\tfrac{Q+(\alpha_2+\alpha_3)P_1}{1+b^2\beta_3P_2}\right),\\
r_{11} &= L\left(\tfrac{Q+\alpha_2P_1+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right),\\
r_{12} &= L\left(\tfrac{Q+\alpha_3P_1+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right),\\
r_{13} &= L\left(\tfrac{(\alpha_2+\alpha_3)P_1+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right),\\
r_{14} &= L\left(\tfrac{Q+(\alpha_2+\alpha_3)P_1+b^2\beta_2P_2}{1+b^2\beta_3P_2}\right),\\
r_{15} &= L\left(\tfrac{a^2\alpha_3P_1}{1+S+a^2\alpha_2P_1+beta_3P_2}\right),\\
r_{16} &= L\left(\tfrac{\beta_2P_2}{1+S+a^2\alpha_2P_1+\beta_3P_2}\right),\\
r_{17} &= L\left(\tfrac{a^2\alpha_3P_1+\beta_2P_2}{1+S+a^2\alpha_2P_1+\beta_3P_2}\right),\\
r_{18} &= L\left(\tfrac{\beta_3P_2}{1+a^2\alpha_2P_1}\right).
\end{aligned}
\tag{4.39}
$$

Define the rate region $\mathcal{R}_i^{\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3}$ to be the convex hull of all rate triplets

(R_0, R_1, R_2) satisfying

$$\begin{aligned}
&R_0 \leq r_0, \\
&R_1 \leq \min(r_7, r_1 + r_{15}), \\
&R_2 \leq \min(r_3 + r_{18}, r_{16} + r_{18}), \\
&R_0 + R_1 \leq \min(r_{10}, r_4 + r_{15}), \\
&R_0 + R_2 \leq r_6 + r_{18}, \\
&R_1 + R_2 \leq \min(r_{13} + r_{18}, r_8 + r_{15} + r_{18}, r_1 + r_{17} + r_{18}), \\
&R_0 + R_1 + R_2 \\
&\leq \min(r_{14} + r_{18}, r_{11} + r_{15} + r_{18}, r_4 + r_{17} + r_{18}), \\
&2R_0 + R_1 \leq r_4 + r_5, \\
&R_1 + 2R_2 \leq \min(r_8 + r_9 + 2r_{18}, r_8 + r_{17} + 2r_{18}), \\
&2R_0 + R_1 + R_2 \leq \min(r_5 + r_{11} + r_{18}, r_4 + r_{12} + r_{18}), \\
&R_0 + R_1 + 2R_2 \\
&\leq \min(r_9 + r_{11} + 2r_{18}, r_8 + r_{!2} + 2r_{18}, r_{11} + r_{17} + 2r_{18}), \\
&2R_0 + R_1 + 2R_2 \leq r_{11} + r_{12} + 2r_{18}
\end{aligned} \tag{4.40}$$

Define the rate region $\mathcal{R}_i$ to be convex hull of the union of rate region $\mathcal{R}_i^{\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3}$:

$$\mathcal{R}_i \triangleq \overline{\bigcup_{\substack{\alpha_1+\alpha_2+\alpha_3=1 \\ \beta_1+\beta_2+\beta_3=1}} \mathcal{R}_i^{\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3}}. \tag{4.41}$$

Theorem 9. For the Gaussian channel with partially cognitive radio as described in (4.1), the region described by

$$\mathcal{R}_{in} = \{(R_0, R_1, R_2) \in \mathcal{R}_i : R_1 \geq \mu R_0\} \tag{4.42}$$

is achievable.

Proof: In establishing the result, we use a combination of dirty paper coding with Han-Kobayashi coding. We first describe the encoding strategy at the two transmitters. We fix $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ such that $\alpha_1+\alpha_2+\alpha_3 = 1$ and $\beta_1+\beta_2+\beta_3 = 1$.

Encoding Strategy at the Legitimate Transmitter: For every message $W_0 \in \{1, \ldots, M_0\}$, the legitimate transmitter generates a codeword $X_{10}^n(W_0)$ from the distribution

$p(X_{10}^n) = \Pi_{i=1}^n p(X_{10}(i))$, where $X_{10}(i) \sim \mathcal{N}(0, \alpha_1 P_1)$. For every message $W_{11} \in \{1, \ldots, M_{11}\}$, the legitimate transmitter generates a codeword $X_{11}^n(W_1)$ from the distribution $p(X_{11}^n) = \Pi_{i=1}^n p(X_{11}(i))$, where $X_{11}(i) \sim \mathcal{N}(0, \alpha_2 P_1)$. For every message $W_{12} \in \{1, \ldots, M_{12}\}$, the legitimate transmitter generates a codeword $X_{12}^n(W_2)$ from the distribution $p(X_{12}^n) = \Pi_{i=1}^n p(X_{12}(i))$, where $X_{12}(i) \sim \mathcal{N}(0, \alpha_3 P_1)$. The legitimate transmitter then superimposes these codewords to form the net transmit vector X_1^n as

$$X_1^n = X_{10}^n + X_{11}^n + X_{12}^n.$$

Encoding Strategy at the Cognitive Transmitter: The cognitive transmitter allocates a portion of its power in communicating the message W_0 to the legitimate receiver. For message W_0, the cognitive transmitter generates a codeword $X_{20}^n(W_0)$ as follows:

$$X_{20}^n(W_0) = \sqrt{\frac{\beta_1 P_2}{\alpha_1 P_1}} X_{10}^n(W_0).$$

That is, the cognitive transmitter uses the same codeword for encoding message W_0 as used by the legitimate transmitter except that it is scaled to power $\beta_1 P_2$. Next, the cognitive transmitter encodes message W_{21} to codeword X_{21}^n. The cognitive transmitter generates a codeword $X_{21}^n(W_1)$ from the distribution $p(X_{21}^n) = \Pi_{i=1}^n p(X_{21}(i))$, where $X_{21}(i) \sim \mathcal{N}(0, \beta_2 P_2)$. Then, the cognitive transmitter encodes message W_{22} to codeword X_{22}^n using dirty paper coding treating $aX_{10}^n + X_{20}^n$ as non-causally known interference. X_{22}^n is independent of the interference, $aX_{10}^n + X_{20}^n$, and is distributed as $p(X_{22}^n) = \Pi_{i=1}^n p(X_{22}(i))$ and $X_{22}(i) \sim \mathcal{N}(0, \beta_3 P_2)$. The cognitive transmitter superimposes the three codewords X_{20}^n, X_{21}^n, and X_{22}^n to form its net codeword X_2^n. That is

$$X_2^n = X_{20}^n + X_{21}^n + X_{22}^n.$$

Next, we describe the decoding strategy and the rate constraints associated at the two receivers.

Decoding Strategy at the Legitimate Receiver: The legitimate receiver obtains the signal

$$Y_1^n = X_{10}^n + X_{11}^n + X_{12}^n + bX_{20}^n + bX_{21}^n + bX_{22}^n + Z_1^n.$$

The licensed receiver decodes the messages $W_0, W_{11}, W_{12}, W_{21}$ jointly treating X_{22}^n as noise. The decoding is successful if the rates satisfy the constraints given by

$$
\begin{aligned}
&R_0 \leq r_0, && R_{11} \leq r_1, \\
&R_{12} \leq r_2, && R_{21} \leq r_3, \\
&R_0 + R_{11} \leq r_4, && R_0 + R_{12} \leq r_5, \\
&R_0 + R_{21} \leq r_6, && R_1 \leq r_7, \\
&R_{11} + R_{21} \leq r_8, && R_{12} + R_{21} \leq r_9, \\
&R_0 + R_1 \leq r_{10}, && R_0 + R_{11} + R_{21} \leq r_{11}, \\
&R_0 + R_{12} + R_{21} \leq r_{12}, && R_1 + R_{21} \leq r_{13}, \\
&R_0 + R_1 + R_{21} \leq r_{14}.
\end{aligned} \tag{4.43}
$$

Decoding Strategy at the Cognitive Receiver: The cognitive receiver obtains the signal

$$Y_2^n = aX_{10}^n + aX_{11}^n + aX_{12}^n + X_{20}^n + X_{21}^n + X_{22}^n + Z_2^n.$$

The cognitive receiver decodes message W_{12} and W_{21} jointly treating X_{10}^n, X_{20}^n, X_{11}^n and X_{22}^n as Gaussian noise. The receiver can decode message W_{12} and W_{21} successfully if

$$
\begin{aligned}
R_{12} &\leq r_{15}, \\
R_{21} &\leq r_{16}, \\
R_{12} + R_{21} &\leq r_{17}.
\end{aligned} \tag{4.44}
$$

Finally, the cognitive receiver decodes W_{22} using Costa's dirty paper decoding. In decoding W_{22}, X_{10}^n and X_{20}^n do not impact rate due to the dirty paper coding employed at the encoder. The decoding is successful if

$$R_{22} \leq r_{18}. \tag{4.45}$$

Using Fourier-Motzkin elimination, we find that the region given by $\mathcal{R}_i^{\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3}$ is achievable. By taking the closure of the convex hull over the set of α's and β's, we find that the region given by $\mathcal{R}_i$ is achievable,. This completes the achievability part of this paper.

As μ grows to infinity, the channel resembles the independent-message interference channel with no cognitive message sets. Our achievable scheme then enforces β_1 and $1 - \alpha_1 - \alpha_2$ to be fixed at 0, and the rate region reduces to the one corresponding to the Han-Kobayashi coding strategy. At the other extreme, the channel becomes a cognitive radio channel - an interference channel with degraded message sets. In this case, β_2 and α_2 are reduced to zero, and the cognitive user now utilizes

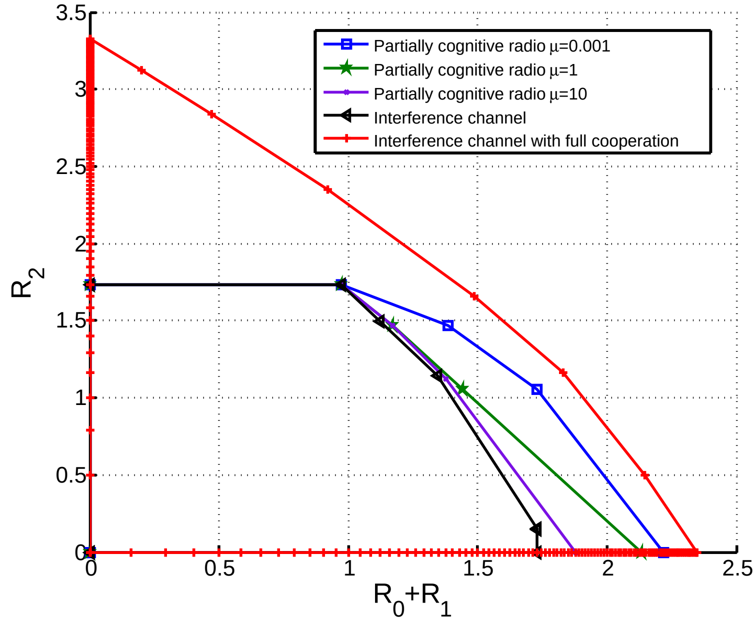


Figure 4.6: Achievable region

dirty paper coding which is already known to be optimal for this class of channels. We compare the achievable rate region of the partially cognitive radio, where the cognitive radio is unidirectionally cooperating, with that of interference channel, where there is no cooperation, and with interference channel with full transmitter cooperation. For an interference channel, Han-Kobayashi coding strategy is used. And, the interference channel with full cooperation can be converted to a broadcast channel with multiple transmit antennas. Thus, we use the 'Dirty Paper Region' in [58]. This comparison is presented in Figure 4.6. As μ becomes smaller, the rate region of the partially cognitive cognitive radio asymptotically approaches that of interference channel. And, as μ grows, the rate region of the partially cognitive radio expands as well. However, the full cooperation can still further increase the rate region for obvious reasons.

4.5 Numerical Analysis

We compare the achievable region and outer bound derived in this paper in this section. This comparison is presented in Figure 4.7. For our numerical calculation, we set both transmit powers P_1 and P_2 to 10dB, and interference gains

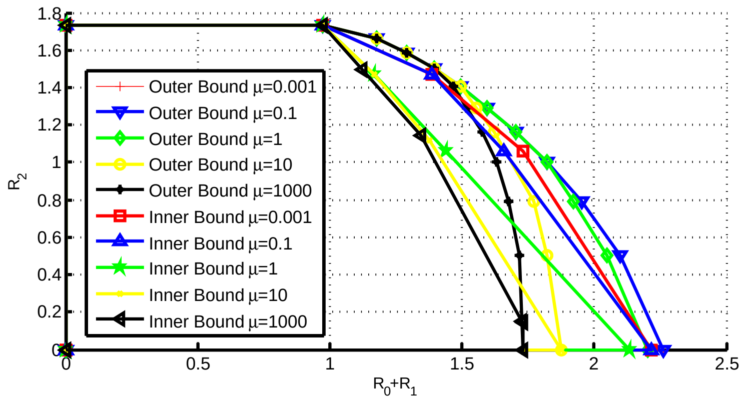


Figure 4.7: Achievable region and outer bound

a and b to 2 and 0.5 respectively. We vary μ from the traditional interference channel extreme to the degraded message set extreme and plot both the achievable region and the outer bound in each case. As intuition suggests, the achievable region and outer bound match in the case when μ is small (full cognition) and differ the most when μ is large (the traditional interference channel).

Notice that as the value μ grows, achievable region asymptotically approaches the outer bound. It is desirable to show constant gap between inner and outer bounds as is done for fully cognitive radio and interference channel. However, our approach is to find efficient achievable region that is optimal in two extremes. Thus, the bounds are not analytical to give a constant gap. Such an approach to find a constant gap between two bounds is also valuable, and we leave it to the future work.

Chapter 5

Capacity on the Overlay Cognitive Radio with Additional Information

5.1 Introduction

In this chapter, an overlay cognitive radio is assumed to possess extra information about the legitimate radio's message sets. When the cognitive radio has more message sets about the legitimate radio than the legitimate radio itself has, the capacity region is analyzed in the interference channel.

5.1.1 Our Contributions

Our main contribution in this chapter is as follows:

1. We obtain the capacity region of overlay cognitive radio with additional information in the "weak" interference case.

2. We establish the outer bound for the capacity region of overlay cognitive radio with additional information in the "strong" interference case.

3. We obtain the achievability scheme and capacity region of overlay cognitive radio with additional information in the "strong" interference case.

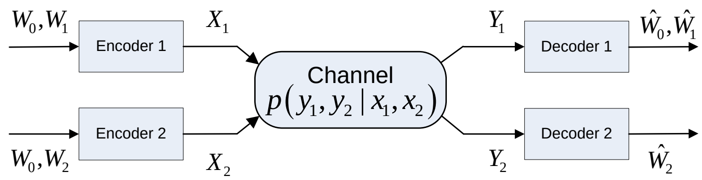


Figure 5.1: The discrete memoryless interference channel model with cognitive radio with additional information

5.2 System Model and Preliminaries

Random variables (RVs) are denoted using capital letters, and their realizations using the corresponding lower case letters. X_m^n denotes the vector $(X_m, ..., X_n)$, and $X^{n\backslash m}$ denotes the vector $(X_1, ..., X_{m-1}, X_{m+1}, ..., X_n)$. For any set S, $\overline{S}$ denotes its convex hull and $\widetilde{S}$ the complementary set of S. Finally, the notation $X \leftrightarrow Y \leftrightarrow Z$ is used to denote that X and Z are conditionally independent given Y.

5.2.1 Discrete Memoryless Interference Channel with Overlay Cognitive Radio with Additional Information

A two-user interference channel with cognitive radio with additional information is a quintuple $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, p)$, where $\mathcal{X}_1, \mathcal{X}_2$ are two input alphabet sets, $\mathcal{Y}_1, \mathcal{Y}_2$ are two output alphabet sets, and $p(y_1, y_2|x_1, x_2)$ is the transition probability. Since we restrict our study to memoryless channels, the transition probability of y_1^n, y_2^n given x_1^n, x_2^n satisfies $p(y_1^n, y_2^n|x_1^n, x_2^n) = \prod_{i=1}^{n} p(y_{1,i}, y_{2,i}|x_{1,i}, x_{2,i})$.

Transmitter 1 henceforth is referred to as the *legitimate transmitter* that communicates W_0 to Receiver 1, the *legitimate receiver*. Transmitter 2, henceforth called the *cognitive transmitter* desires to communicate two messages W_1 and W_2, one to the legitimate receiver and the other to Receiver 2, the *cognitive receiver*, respectively. Transmitter 2 gets the name cognitive as it has access to W_0, the legitimate transmitter's message. Thus, overall, W_0 is known to *both* transmitters, making this an overlay cognitive radio setting. An $(R_0, R_1, R_2, n, P_{e,0}, P_{e,1}, P_{e,2})$ code is one with the rate vector (R_0, R_1, R_2) and block size n, where $R_t = \log(M_t)/n$ bits per usage for $t = 0, 1, 2$. As aforementioned, W_0, and W_1 are messages intended for the legitimate receiver with (average) probabilities of error of at most $P_{e,0}, P_{e,1}$ respectively, and W_2 must be retrieved at the cognitive receiver while suffering an error probability that is no more than $P_{e,2}$. The rate vector (R_0, R_1, R_2) is said to be achievable if the error probabilities $P_{e,t}$ for $t = 0, 1, 2$ can be made arbitrarily small for a large enough block size n.

5.2.2 Gaussian Interference Channel with Overlay Cognitive Radio with Additional Information

A Gaussian interference channel with cognitive radio is characterized mathematically in a manner similar to the two-user interference channel as:

$$\begin{aligned} Y_1 &= X_1 + bX_2 + Z_1 \\ Y_2 &= aX_1 + X_2 + Z_2, \end{aligned} \tag{5.1}$$

where a and b are real numbers and Z_1 and Z_2 are independent, zero-mean, unit-variance Gaussian random variables. Further, each transmitter has a power constraint

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_{t,i}^2] \leq P_t, t = 1, 2.$$

It differs from the conventional interference channel in the way messages are allocated and their intended destinations. In this respect, they are exactly

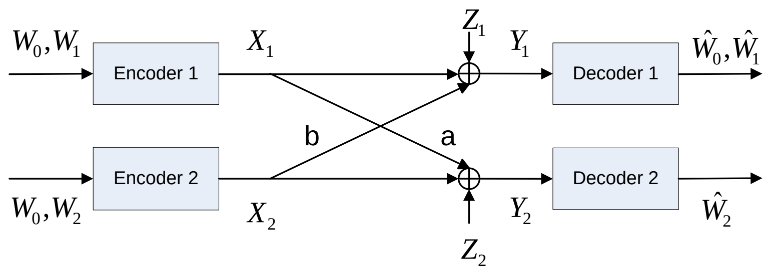


Figure 5.2: The Gaussian over-cognitive radio channel

like the discrete memoryless channel with overlay cognitive radio with additional information. To avoid repetition, we do not reproduce the encoding and decoding definitions here.

The achievable strategy for this channel comes from the idea behind the achievable strategies of the conventional cognitive radio model fairly straight-forwardly. Thus we do not make it a focal point. The outer bound, however, is considerably more challenging and thus, the next section focuses on the outer bound.

5.3 The Outer Bound

5.3.1 Discrete Memoryless Interference Channel with Cognitive Radio with Additional Information

In this section, we find outer bounds for two classes of overlay cognitive radio channels - under "weak" and "strong" interference conditions. Such a distinction into classes is necessary to find meaningful (nontrivial) bounds for cognitive models in general, and our problem in particular.

1) "Weak" Interference Radios: The following Markov chain requirement characterizes those channels that satisfy the "weak" interference requirement: Given

X_1, we require that

$$X_2 \leftrightarrow Y_2 \leftrightarrow Y_1. \tag{5.2}$$

Note that although the condition above specifies a physical degradation in signal quality between the two receivers, a stochastic version of it is sufficient. The outer bound for this class is stated in the following theorem:

Theorem 10. The convex closure of the following inequalities defines an outer bound on the capacity region of "weak" interference channel with cognitive radio with additional information:

$$\begin{aligned} R_1 &\leq I(V; Y_1 | U, X_1) \\ R_0 + R_1 &\leq I(U, V, X_1; Y_1) \\ R_2 &\leq I(X_2; Y_2 | U, V, X_1) \end{aligned}$$

for all $p(u,v))p(x_1|u)p(x_2|u,v)$ such that:

1. V and X_1 are independent,
2. $(U, V) \rightarrow (X_1, X_2) \rightarrow (Y_1, Y_2)$.

Proof: We have

$$
\begin{aligned}
nR_1 &\le I(W_1; Y_1^n | W_0) + n\epsilon_1 \\
&= \sum_{i=1}^{n} \Big[\; H(Y_{1,i}|Y_1^{i-1}, W_0) - H(Y_{1,i}|Y_1^{i-1}, W_0, W_1) \; \Big] + n\epsilon_1 \\
&\overset{(a)}{\le} \sum_{i=1}^{n} \Big[\; H(Y_{1,i}|Y_1^{i-1}, W_0, X_1^i) - H(Y_{1,i}|Y_1^{i-1}, W_0, W_1, X_1^i) \; \Big] + n\epsilon_1 \\
&\overset{(b)}{\le} \sum_{i=1}^{n} [H(Y_{1,i}|U_i, X_{1,i}) - H(Y_{1,i}|U_i, V_i, X_{1,i})] + n\epsilon_1 \\
&\le \sum_{i} I(V_i; Y_{1i}|U_i, X_{1i}) + n\epsilon_1
\end{aligned}
$$

where (a) results from the fact that X_1^i is the function of W_0, and (b) results from identifying auxiliaries $U_i = (Y_1^{i-1}, X_1^{i-1}, W_0)$ and $V_i = W_1$. For $R_0 + R_1$,

$$
\begin{aligned}
n\,(R_0 + R_1) &= H(W_0, W_1) \\
&\le I(W_0, W_1; Y_1^n) + n\epsilon_2 \\
&\overset{(a)}{\le} \sum_{i=1}^{n} \Big[\; H(Y_{1,i}) - H(Y_{1,i}|Y_1^{i-1}, W_0, W_1, X_1^i) \; \Big] + n\epsilon_2 \\
&\overset{(b)}{\le} \sum_{i=1}^{n} [H(Y_{1,i}) - H(Y_{1,i}|U_i, V_i, X_{1,i})] + n\epsilon_2 \\
&\le \sum_{i} I(U_i, V_i, X_{1i}; Y_{1i}) + n\epsilon_2
\end{aligned}
$$

where (a) is due to conditioning and (b) results from the same aforementioned definitions of $U_i = (Y_1^{i-1}, W_0)$ and $V_i = W_1$. Next, we establish the outer bound on

R_2. We have:

$$
\begin{aligned}
nR_2 &\le I(W_2; Y_2^n | W_0, W_1) + n\epsilon_3 \\
&= \sum_{i=1}^{n} \left[\ H(Y_{2,i}|Y_2^{i-1}, W_0, W_1) - H(Y_{2,i}|Y_2^{i-1}, W_0, W_1)\ \right] + n\epsilon_3 \\
&\overset{(a)}{\le} \sum_{i=1}^{n} \left[\ H(Y_{2,i}|Y_2^{i-1}, W_0, W_1, X_1^i) - H(Y_{2,i}|Y_2^{i-1}, W_0, W_1, X_{1,i}, X_{2,i})\ \right] + n\epsilon_3 \\
&\overset{(b)}{\le} \sum_{i=1}^{n} \left[\begin{array}{l} H(Y_{2,i}|Y_1^{i-1}, W_0, W_1, X_1^{i-1}, X_{1,i}) \\ -H(Y_{2,i}|Y_1^{i-1}, W_0, W_1, X_{1,i}, X_{2,i}) \end{array} \right] + n\epsilon_3 \\
&\le nI(X_2; Y_2 | U, V, X_1) + n\epsilon_3
\end{aligned}
$$

where (a) results from the fact that X_1^i is the function of W_0, and (b) is due to the weak interference condition as specified by (5.2) and the memoryless nature of the channel.

2) "Strong" Interference Radios: Intuitively, the channel-gain between the cognitive transmitter and receiver pair is stronger than the channel-gain from the cognitive transmitter to the cognitive receiver. This translates in the discrete memoryless case to the following Markov chain: Given X_1, we have

$$X_2 \leftrightarrow Y_1 \leftrightarrow Y_2. \tag{5.3}$$

Again, this represents a physical degradation that can be relaxed to a stochastic one. Under this strong interference condition, we have the following Lemma:

Lemma 9. For the cognitive radio channel with addition information in the "strong" interference condition, we have: $I(W_i; Y_1^n | W_0, X_1^n) \ge I(W_i; Y_2^n | W_0, X_1^n)$, where $i = 1, 2$.

Proof: First, W_i, $i = 0, 1$, is independent of W_0 and X_1^n. Also, X_2 is a function of W_0, W_1, and W_2. From these and strong interference condition we have the Markov chain $W_i \leftrightarrow X_2^n \leftrightarrow Y_1^n \leftrightarrow Y_2^n$, given W_0, X_1^n. Thus,

$$\begin{aligned} I(W_i; Y_1^n | W_0, X_1^n) &= I(W_i; Y_1^n, Y_2^n | W_0, X_1^n) \\ &\geq I(W_i; Y_2^n | W_0, X_1^n). \end{aligned}$$

We present the outer bound under the strong interference condition in the following:

Theorem 11. The convex closure of the following inequalities defines an outer bound on the capacity region of a "strong" interference channel with cognitive radio with additional information:

$$\begin{aligned} R_2 &\leq I(V; Y_2 | U, X_1) \\ R_0 + R_2 &\leq I(U, V, X_1; Y_1) \\ R_1 &\leq I(X_2; Y_1 | U, V, X_1) \end{aligned}$$

for any $p(u, v))p(x_1|u)p(x_2|u, v)$ such that:

1. V and X_1 are independent,
2. $(U, V) \rightarrow (X_1, X_2) \rightarrow (Y_1, Y_2)$.

Proof: First we prove the outer bound for R_2 and the sum rate $R_0 + R_2$. We have

$$
\begin{aligned}
nR_2 &\overset{(a)}{\leq} I(W_2; Y_2^n | W_0, X_1^n) + n\epsilon_1 \\
&= \sum_{i=1}^{n} \quad H(Y_{2,i}|Y_2^{i-1}, W_0, X_1^n) - H(Y_{2,i}|Y_2^{i-1}, W_0, W_2, X_1^n) \quad + n\epsilon_1 \\
&\leq \sum_{i=1}^{n} \Big[\ H(Y_{2,i}|Y_2^{i-1}, W_0, X_1^n) - H(Y_{2,i}|Y_2^{i-1}, Y_1^{i-1}, W_0, W_2, X_1^n) \ \Big] + n\epsilon_1 \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} [H(Y_{2,i}|U_i, X_{1,i}) - H(Y_{2,i}|U_i, V_i, X_{1,i})] + n\epsilon_1 \\
&\leq \sum_{i} I(V_i; Y_{2i}|U_i, X_{1i}) + n\epsilon_1,
\end{aligned}
$$

where (a) results from the fact that X_1^n is the function of W_0, and (b) results from identifying auxiliaries $U_i = (Y_2^{i-1}, W_0, X_1^{n\setminus i})$ and $V_i = (Y_1^{i-1}, W_2)$. For $R_0 + R_2$,

$$
\begin{aligned}
n\,(R_0 + R_2) &\leq I(W_0; Y_1^n) + I(W_2; Y_2^n | W_0) + n\epsilon_2 \\
&= I(W_0, X_1^n; Y_1^n) + I(W_2; Y_2^n | W_0, X_1^n) + n\epsilon_2 \\
&\overset{(a)}{\leq} I(W_0, X_1^n; Y_1^n) + I(W_2; Y_1^n | W_0, X_1^n) + n\epsilon_2 \\
&= I(W_0, W_2, X_1^n; Y_1^n) + n\epsilon_2 \\
&= \sum_{i=1}^{n} \Big[\ H(Y_{1,i}|Y_1^{i-1}) - H(Y_{1,i}|Y_1^{i-1}, Y_2^{i-1}, W_0, W_2, X_1^n) \ \Big] \\
&\quad + n\epsilon_2 \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} [H(Y_{1,i}) - H(Y_{1,i}|U_i, V_i, X_{1,i})] + n\epsilon_2 \\
&\leq \sum_{i} I(U_i, V_i, X_{1i}; Y_{1i}) + n\epsilon_2
\end{aligned}
$$

where (a) is from the Lemma 9, and (b) comes from the same aforementioned definitions of $U_i = (Y_2^{i-1}, W_0, X_1^{n\setminus i})$ and $V_i = (Y_1^{i-1}, W_2)$. Last, we establish the outer bound expression for the rate R_1. We have

$$
\begin{aligned}
nR_1 &= H(W_1|W_0, W_2) \\
&\leq I(W_1; Y_1^n|W_0, W_2) + n\epsilon_3 \\
&= \sum_{i=1}^{n} \left[\ H(Y_{1,i}|Y_1^{i-1}, W_0, W_2) - H(Y_{1,i}|Y_1^{i-1}, W_0, W_1, W_2)\ \right] + n\epsilon_3 \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} \left[\ H(Y_{1,i}|Y_1^{i-1}, W_0, W_2, X_1^n) - H(Y_{1,i}|Y_1^{i-1}, W_0, W_1, W_2, X_1^n, X_{2,i})\ \right] + n\epsilon_3 \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} \left[\begin{array}{l} H(Y_{1,i}|Y_1^{i-1}, Y_2^{i-1}, W_0, W_2, X_1^n) \\ -H(Y_{1,i}|Y_1^{i-1},, Y_2^{i-1}, W_0, W_1, W_2, X_1^n, X_{2,i}) \end{array} \right] + n\epsilon_3 \\
&\leq \sum_{i} I(X_{2i}; Y_{2i}|U_i, V_i, X_{1i}) + n\epsilon_3
\end{aligned}
$$

(a) results from the fact that X_1^n is the function of W_0, and conditioning reduces the entropy, and (b) is due to the strong interference condition that we have in (5.3). Given this framework, we proceed to analyzing the outer bound for the Gaussian case.

5.3.2 Gaussian Interference Channel with Cognitive Radio with Additional Information

1) "Weak" Interference Radios: First, we consider the case where the channel is defined as "weak". Let $\mathcal{R}_{out}^{weak}(\rho_1, \rho_2)$ denote the set of rate triplet $(R_0, R_1, R_2) \in$

$\mathbb{R}_+^3$ given by

$$\begin{aligned}
R_0 &\leq \frac{1}{2}\log\left(\frac{1+P_1+b^2(1-\rho_2^2)P_2+2b\rho_1\sqrt{P_1P_2}}{1+b^2(1-\rho_1^2-\rho_2^2)P_2}\right)\\
R_1 &\leq \frac{1}{2}\log\left(\frac{1+b^2(1-\rho_1^2)P_2}{1+b^2(1-\rho_1^2-\rho_2^2)P_2}\right)\\
R_0+R_1 &\leq \frac{1}{2}\log\left(\frac{1+P_1+b^2P_2+2b\rho\sqrt{P_1P_2}}{1+b^2(1-\rho_1^2-\rho_2^2)P_2}\right)\\
R_2 &\leq \frac{1}{2}\log\left(1+(1-\rho_1^2-\rho_2^2)P_2\right).
\end{aligned}$$

Then, we have the following theorem:

Theorem 12. An outer bound on the capacity region for the Gaussian "weak" interference channel with cognitive radio with additional information is given by a convex hull of the union of the rate region $\mathcal{R}_{out}^{weak}(\rho_1,\rho_2)$ with constraints $\rho_1 \in [0,1]$, $\rho_2 \in [0,1]$, and $\rho_1^2+\rho_2^2 \leq 1$:

$$\mathcal{C}^{weak} \subset \overline{\bigcup_{\rho_1,\rho_2\in[0,1],\rho_1^2+\rho_2^2\leq 1} \mathcal{R}_{out}^{weak}(\rho_1,\rho_2)},$$

Proof: The proof of this closely resembles that in [89] and is skipped to avoid repetition.

2) "Strong" Interference Radios: Next, we find the outer bound of the capacity region for the "strong" interference case; $b > 1$. Define $\mathcal{R}_{out}^{strong}(\alpha,\rho)$ as the set of rate triplet $(R_0,R_1,R_2) \in \mathbb{R}_+^3$ given by

$$\begin{aligned}
R_2 &\leq \frac{1}{2}\log\left(\frac{1+(1-\rho^2)P_2}{1+\alpha P_2}\right)\\
R_0+R_2 &\leq \frac{1}{2}\log\left(\frac{1+P_1+b^2P_2+2b\rho\sqrt{P_1P_2}}{1+b^2\alpha P_2}\right)\\
R_1 &\leq \frac{1}{2}\log\left(1+b^2\alpha P_2\right)
\end{aligned}$$

Then, we have the outer bound of the capacity region is given by:

Theorem 13. An outer bound of the capacity region for the Gaussian "strong" interference channel with cognitive radio with additional information is given by the convex hull of the rate region $\mathcal{R}_{out}^{strong}(\alpha, \rho)$ with constraints $\alpha \in [0, 1]$, $\rho \in [0, 1]$, and $\alpha \in [0, 1]$, $\alpha + \rho^2 \leq 1$:

$$\mathcal{C}^{strong} \subset \overline{\bigcup_{\alpha,\rho \in [0,1], \alpha+\rho^2 \leq 1} \mathcal{R}_{out}^{strong}(\alpha, \rho)}.$$

Proof: We have:

$$\begin{aligned}
R_2 &\leq I(V; Y_2|U, X_1) \\
&= h(Y_2|U, X_1) - h(Y_2|U, V, X_1) \\
R_0 + R_2 &\leq I(U, V, X_1; Y_1) \\
&= h(Y_1) - h(Y_1|U, V, X_1) \\
R_1 &\leq I(X_2; Y_1|U, V, X_1) \\
&= h(Y_1|U, V, X_1) - h(Y_1|U, V, X_1, X_2)
\end{aligned}$$

Note that $h(Y_1|U, V, X_1, X_2) = h(Z_1) = \frac{1}{2} \log 2\pi e$, and that

$$\begin{aligned}
h(Y_1|U, V, X_1) &= h(bX_2 + Z_1|U, V, X_1) \\
&\leq \frac{1}{2} \log 2\pi e \left(1 + b^2 P_2\right),
\end{aligned}$$

Thus, without loss of generality for some $\alpha \in [0, 1]$, we can set $h(Y_1|U, V, X_1) = \frac{1}{2} \log 2\pi e \left(1 + b^2 \alpha P_2\right)$. Then, we use entropy power inequality to obtain a lower bound on $h(Y_2|U,V,X_1)$.

$$
\begin{aligned}
2^{2h(Y_2|U,V,X_1)} &\geq 2^{2h\left(\frac{1}{b}Y_1|U,V,X_1\right)} + 2^{2h(Z')} \\
&= \frac{1}{b^2}2^{2h(Y_1|U,V,X_1)} + 2\pi e\left(1-\frac{1}{b^2}\right) \\
&= 2\pi e(1+\alpha P_2),
\end{aligned}
$$

where $b > 1$ given the "strong" interference, and Z' is a Gaussian distributed random variable with variance $1-\frac{1}{b^2}$. Thus, we have $h(Y_2|U,V,X_1) \geq \frac{1}{2}\log 2\pi e(1+\alpha P_2)$.

Next, we bound $h(Y_1)$ and $h(Y_2|U,X_1)$. Let X_1^*, X_2^* be the arbitrarily distributed zero-mean random variables with the same covariance matrix of X_1, X_2, where $\mathbb{E}[X_1X_2] = \rho\sqrt{P_1P_2}$. Note that by setting $h(Y_1|U,V,X_1) = \frac{1}{2}\log 2\pi e\left(1+b^2\alpha P_2\right)$ we have the following result:

$$
\begin{aligned}
h(Y_1|U,V,X_1) &\leq h(bX_2+Z_1|X_1) \\
&\leq h(bX_2^*+Z_2|X_1^*) \\
&= \frac{1}{2}\log(2\pi e(1+b^2(1-\rho^2)P_2)),
\end{aligned}
$$

which yields $\alpha \leq 1-\rho^2$. Hence, we can bound $h(Y_1)$ and $h(Y_2|U,X_1)$ to

$$
\begin{aligned}
h(Y_1) &= h(X_1+bX_2+Z_1) \\
&\leq h(X_1^*+bX_2^*+Z_1) \\
&= \frac{1}{2}\log(2\pi e(1+P_1+b^2P_2+2b\rho\sqrt{P_1P_2})),
\end{aligned}
$$

and

$$
\begin{aligned}
h(Y_2|U, X_1) &\leq h(X_2 + Z_2|X_1) \\
&\leq h(X_2^* + Z_2|X_1^*) \\
&= \frac{1}{2}\log(2\pi e(1 + (1-\rho^2)P_2)).
\end{aligned}
$$

In summary, the outer bound can be expressed in terms of $0 \leq \alpha \leq 1$,

$$
\begin{aligned}
R_2 &\leq \frac{1}{2}\log\left(\frac{1 + (1-\rho^2)P_2}{1 + \alpha P_2}\right) \\
R_0 + R_2 &\leq \frac{1}{2}\log\left(\frac{1 + P_1 + b^2 P_2 + 2b\rho\sqrt{P_1 P_2}}{1 + b^2 \alpha P_2}\right) \\
R_1 &\leq \frac{1}{2}\log\left(1 + b^2 \alpha P_2\right).
\end{aligned}
$$

5.4 Achievable Region for the Gaussian Channel

The achievable region for the "weak" interference case is a straightforward generalization of the one in [89], and so it is not repeated here. The achievable region for the "strong" interference case is presented next.

The legitimate transmitter uses an i.i.d Gaussian codebook to communicated W_0. The cognitive transmitter splits its power into three parts: one part is used to aid in the transmission of W_0, the second part to communicate W_1 and the final to communicate W_2. Finally, dirty paper coding is used to eliminate interference at the cognitive receiver.

Define $\mathcal{R}_{in}^{strong}(\alpha, \rho)$ as the set of rate triplet $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ given by

$$
\begin{aligned}
R_0 &\leq \frac{1}{2}\log\left(\frac{1 + P_1 + b^2 P_2 + 2b\rho\sqrt{P_1 P_2}}{1 + b^2(1-\rho^2)P_2}\right) \\
R_1 &\leq \frac{1}{2}\log(1 + b^2\alpha P_2) \\
R_2 &\leq \frac{1}{2}\log\left(\frac{1 + (1-\rho^2)P_2}{1 + \alpha P_2}\right).
\end{aligned}
$$

Then, we have

Theorem 14. An inner bound on the capacity of the Gaussian overlay cognitive MAC channel as given by (5.1) is the convex hull of $\mathcal{R}_{in}^{strong}(\alpha, \rho)$ with constraints $\alpha \in [0, 1]$, $\rho \in [0, 1]$, and $\alpha \in [0, 1]$, $\alpha + \rho^2 \leq 1$:

$$
\overline{\bigcup_{\alpha,\rho \in [0,1], \alpha+\rho^2 \leq 1} \mathcal{R}_{in}^{strong}(\alpha, \rho)} \subset \mathcal{C}^{strong}
$$

Proof: This proof is skipped due to space restrictions, and it straightforwardly follows by using a combination of Gaussian codebooks with dirty paper coding.

5.5 Optimality of Achievable Region

For the "weak" interference case, a simple inspection shows that the achievable region and outer bound meet, and thus we have a capacity region characterization. For the "strong" interference case, we have a match if

$$
\frac{1}{2}\log\left(\frac{1 + P_1 + b^2 P_2 + 2b\rho\sqrt{P_1 P_2}}{1 + b^2\alpha P_2}\right)
= \left[\begin{array}{l} \frac{1}{2}\log\left(\frac{1+P_1+b^2P_2+2b\rho\sqrt{P_1P_2}}{1+b^2(1-\rho^2)P_2}\right) \\ +\frac{1}{2}\log\left(\frac{1+(1-\rho^2)P_2}{1+\alpha P_2}\right). \end{array}\right]
$$

A sufficient condition for this to hold is:

$$(1-b^2)(1-\alpha-\rho^2)=0.$$

Note that in general, this may not be true, and thus an exact capacity characterization in the strong interference case still eludes us.

Chapter 6

Summary

6.1 Summary

A cognitive radio is known to provide a vast potential in increasing the spectral efficiency by sharing the spectrum with the legitimate radios. It is starting to be put into practical use, and expected to expand its applications in wireless communications. It is important to study its limit, and find a method to maximize its efficiency.

Firstly, the new sensing technique, which can sense and transmit simultaneously using self interference cancellation, is proposed. Spectral efficiency increases with this proposed technique, and the interweave cognitive radio can expand its application to the WiFi network.

Secondly, the fundamental limit of this interweave cognitive radio is also studied, and the joint channel selection and power allocation algorithm is proposed under the condition of multiple legitimate channel. A modified waterfilling algorithm together with an approximate selection of channel to sense is derived that performs close to the limits on performance of these radios. With this algorithm, the throughput of the cognitive radio can increase.

Finally, an overlay cognitive radio is studied. This sophisticated cognitive radio, which can be applied to the cognitive radio in the cellular network, can bring about additional spectral efficiency in the network. The capacity region of this cognitive radio is studied under the assumption of partial and additional message sets. In the partially cognitive setup, the transmitter of the cognitive radio has only a portion of the legitimate user's message. As the extent of cognition reduces, the channel becomes a conventional interference channel. As the extent of cognition increases, the channel resembles an interference channel with degraded message sets. Thus, the partially cognitive radio model we consider in this paper lies in between these two extremes and encompasses both as special cases. For the general discrete memoryless IFC setting, we obtain an outer bound for the capacity region and achievable rate region under assumptions of "weak" interference. We also determine an outer bound on the capacity region of a Gaussian partially cognitive radio channel. We determine an achievable region by combining Han-Kobayashi coding strategy and dirty paper coding for the Gaussian channel. When the cognitive radio has more message sets than the legitimate radio, it resembles the interference channel with three messages and overlay cognition. In the "weak" interference case, we find an exact characterization of the capacity region, while in the "strong" interference case, we determine inner and outer bounds for the channel.

Overall, possible advantages that different classes of cognitive radios can bring about are studied, and the practical solutions to achieve them are proposed.

Appendix

Appendix 1

Partially Cognitive Radio

1.1 Proof of the Achievable Rate Region (Probability of Error Analysis)

Here, we analyze the probability of error for the discrete memoryless partially cognitive radio, and complete the proof of the achievable rate region. Let P_e be the total average error probability. We assume the equiprobable message set. Without loss of generality, it is assumed that $(w_{01}, w_{02}, w_{11}, w_{12}, w_{21}, w_{22}) = (1, 1, 1, 1, 1, 1)$ is transmitted. Then, we bound the probability of error as follows:

$$P_e \leq \Pr\left\{ \begin{array}{l} (\hat{w}_{01}, \hat{w}_{02}, \hat{w}_{11}, \hat{w}_{12}, \hat{w}_{21}) \neq (1, 1, 1, 1, 1) \\ |(w_{01}, w_{02}, w_{11}, w_{12}, w_{21}, w_{22}) = (1, 1, 1, 1, 1, 1) \end{array} \right\} + \Pr\left\{ \begin{array}{l} (\hat{w}_{01}, \hat{w}_{11}, \hat{w}_{21}, \hat{w}_{22}) \neq (1, 1, 1, 1) \\ |(w_{01}, w_{02}, w_{11}, w_{12}, w_{21}, w_{22}) = (1, 1, 1, 1, 1, 1) \end{array} \right\}.$$

We consider that the codewords $m_0^n(w_{01}) = m_0^n(1)$, $n_0^n(w_{02}) = n_0^n(1)$, $m_1^n(w_{11}) = m_1^n(1)$ and $n_1^n(w_{12}) = n_1^n(1)$ are used at the legitimate transmitter, and $m_2^n(w_0, w_{21}, k) = m_2^n(1, 1, \hat{k})$ and $n_2^n(w_0, w_{22}, l) = n_2^n(1, 1, \hat{l})$ are used at the cognitive transmitter for sending $(w_{01}, w_{02}, w_{11}, w_{12}, w_{21}, w_{22}) = (1, 1, 1, 1, 1, 1)$, where l and k are the sequence numbers in bins w_{21} and w_{22}, respectively. x_1^n is computed from $m_0^n(1)$, $n_0^n(1)$, $m_1^n(1)$, and $n_1^n(1)$. And, x_2^n is derived from $m_2^n(1, 1, \hat{k})$ and $n_2^n(1, 1, \hat{l})$. Then, the probabilities of error can be upper bounded as

$$\Pr\left\{\begin{array}{l}(\hat{w}_{01},\hat{w}_{02},\hat{w}_{11},\hat{w}_{12},\hat{w}_{21})\neq(1,1,1,1,1)\\ |(w_{01},w_{02},w_{11},w_{12},w_{21},w_{22})=(1,1,1,1,1,1)\end{array}\right\}$$
$$\leq \Pr\left\{\begin{array}{l}(u^n(1),m_1^n(1),n_1^n(1))\\ \text{ is not the only element}\\ \text{in jointly typical set in legitimate receiver}\\ |(w_0,w_{11},w_{12},w_{21},w_{22})=(1,1,1,1,1)\end{array}\right\},$$
$$\Pr\left\{\begin{array}{l}(\hat{w}_{21},\hat{w}_{22},\hat{w}_{11})\neq(1,1,1)\\ |(w_0,w_{11},w_{12},w_{21},w_{22})=(1,1,1,1,1)\end{array}\right\}$$
$$\leq \Pr\left\{\begin{array}{l}\left(m_2^n(1,1,\hat{k}),n_2^n(1,1,\hat{l}),m_1^n(1)\right)\\ \text{ is not the only element}\\ \text{in jointly typical set in cognitive receiver}\\ |(w_0,w_{11},w_{12},w_{21},w_{22})=(1,1,1,1,1)\end{array}\right\}.$$

Next, we define several error events, which indicate error in decoding.

$$E_0^1=\left\{\begin{array}{ll}\nexists\hat{k} & \\ & \left(m_0^n(1),n_0^n(1),m_1^n(1),n_1^n(1),m_2^n(1,1,\hat{k})\right)\\ \text{s.t.} & \notin A_\epsilon^n(M_0,N_0,M_1,N_1,M_2),\\ & 1\leq\hat{k}\leq 2^{n(L_{21}-R_{21})}\end{array}\right\},$$
$$E_1^1=\left\{\begin{array}{l}\left(y_1^n,m_0^n(1),n_0^n(1),m_1^n(1),n_1^n(1),m_2^n(1,1,\hat{k})\right)\\ \notin A_\epsilon^n(Y_1,M_0,N_0,M_1,N_1,M_2)\end{array}\right\},$$
$$E_{ghijs_{21}k}^1=\left\{\begin{array}{l}(y_1^n,m_0^n(g),n_0^n(h),m_1^n(i),n_1^n(j),m_2^n(1,s_{21},k))\\ \in A_\epsilon^n(Y_1,M_0,N_0,M_1,N_1,M_2)\end{array}\right\},$$
$$E_0^2=\left\{\begin{array}{ll}\nexists\hat{l} & \\ & \left(m_0^n(1),n_0^n(1),m_1^n(1),m_2^n(1,1,\hat{k}),n_2^n(1,1,\hat{l})\right)\\ \text{s.t.} & \notin A_\epsilon^n(M_0,N_0,M_1,M_2,N_2),\\ & 1\leq\hat{l}\leq 2^{n(L_{22}-R_{22})}\end{array}\right\},$$
$$E_1^2=\left\{\begin{array}{l}\left(y_2^n,m_1^n(1),m_2^n(1,1,\hat{k}),n_2^n(1,1,\hat{l})\right)\\ \notin A_\epsilon^n(Y_2,M_1,M_2,N_2)\end{array}\right\},$$
$$E_{is_{21}ks_{22}l}^2=\left\{\begin{array}{l}\left(y_2^n,m_1^n(i),m_2^n(1,s_{21},\hat{k}),n_2^n(1,s_{22},\hat{l})\right)\\ \in A_\epsilon^n(Y_1,M_1,M_2,N_2)\end{array}\right\}.$$

Then, we have

$$
\begin{aligned}
P_e \le & \Pr\left\{E_0^1\right\} + \Pr\left\{E_1^1 | \widetilde{E_0^1}\right\} \Pr\left\{\widetilde{E_0^1}\right\} \\
& + \sum_{hijs_{21}k \neq 111\hat{l}} \Pr\left\{E_{hijs_{21}k}^1 | \widetilde{E_1^1}\right\} \Pr\left\{\widetilde{E_1^1}\right\} \\
& + \Pr\left\{E_0^2\right\} + \Pr\left\{E_1^2 | \widetilde{E_0^2}\right\} \Pr\left\{\widetilde{E_0^2}\right\} \\
& + \sum_{is_{21}ks_{22}l \neq 11\hat{k}1\hat{l}} \Pr\left\{E_{is_{21}ks_{22}l}^2 | \widetilde{E_1^2}\right\} \Pr\left\{\widetilde{E_1^2}\right\}.
\end{aligned} \tag{1.1}
$$

Next, we examine the probability of each error event. First, we find the probability of E_0^1 as follows:

$$
\begin{aligned}
& \Pr\left\{E_0^1\right\} \\
& \le \Pr\left\{(m_0^n(1), n_0^n(1), m_1^n(1), n_1^n(1)) \notin A_\epsilon^n(M_0, N_0, M_1, N_1)\right\} \\
& \quad + \prod_{1 \le k \le 2^{n(L_{21}-R_{21})}} \Pr\left\{\begin{array}{l} \begin{pmatrix} m_0^n(1), n_0^n(1), m_1^n(1), \\ n_1^n(1), m_2^n(1,1,1,k) \end{pmatrix} \\ \notin A_\epsilon^n(M_0, N_0, M_1, N_1, M_2) \\ | \, (m_0^n(1), n_0^n(1), m_1^n(1), n_1^n(1)) \\ \in A_\epsilon^n(M_0, N_0, M_1, N_1) \end{array}\right\} \\
& = \Pr\left\{(m_0^n(1), n_0^n(1), m_1^n(1), n_1^n(1)) \notin A_\epsilon^n(M_0, N_0, M_1, N_1)\right\} \\
& \quad + \prod_{1 \le k \le 2^{n(L_{21}-R_{21})}} \Pr\left\{\begin{array}{l} \begin{pmatrix} m_0^n(1), n_0^n(1), m_1^n(1), \\ n_1^n(1), x_1^n m_2^n(1,1,1,k) \end{pmatrix} \\ \notin A_\epsilon^n(M_0, N_0, M_1, N_1, X_1, M_2) \\ | \, (m_0^n(1), n_0^n(1), m_1^n(1), n_1^n(1), x_1^n) \\ \in A_\epsilon^n(M_0, N_0, M_1, N_1, X_1) \end{array}\right\} \\
& \le \epsilon + \left(1 - 2^{n(I(M_2; M_0, N_0, M_1, N_1, X_1) + 3\epsilon)}\right)^{2^{n(L_{21}-R_{21})}} \\
& \le \epsilon + e^{-2^{n(I(M_2; M_0, N_0, M_1, N_1, X_1) + 3\epsilon - L_{21} + R_{21})}} \\
& = \epsilon + e^{-2^{n(I(M_2; X_1) + 3\epsilon - L_{21} + R_{21})}}
\end{aligned}
$$

Thus, provided that

$$
L_{21} - R_{21} > I(M_2; X_1) + 3\epsilon,
$$

$\Pr\left\{E_0^1\right\}$ reduces to 0 as $n \to \infty$. For an error probability $\Pr\left\{E_0^2\right\}$, we have the following result:

$$\begin{aligned}
&\prod_{1 \le k \le 2^{n(L_{21}-R_{21})}} \Pr\left\{\begin{array}{l} (m_0^n(1), n_0^n(1), m_1^n(1), m_2^n(1,1,1,k)) \\ \notin A_\epsilon^n(M_0, N_0, M_1, M_2) \\ \mid (m_0^n(1), n_0^n(1), m_1^n(1)) \\ \notin A_\epsilon^n(M_0, N_0, M_1) \end{array}\right\} \\
&\le \left(1 - 2^{n(I(M_2;M_0,N_0,M_1)+3\epsilon)}\right)^{2^{n(L_{21}-R_{21})}} \\
&\le e^{-2^{n(I(M_2;M_0,N_0,M_1)+3\epsilon-L_{21}+R_{21})}},
\end{aligned}$$

and

$$\begin{aligned}
&\prod_{1 \le l \le 2^{n(L_{22}-R_{22})}} \Pr\left\{\begin{array}{l} (m_0^n(1), n_0^n(1), m_1^n(1), n_2^n(1,1,1,l)) \\ \notin A_\epsilon^n(M_0, N_0, M_1, N_2) \\ \mid (m_0^n(1), n_0^n(1), m_1^n(1)) \\ \notin A_\epsilon^n(M_0, N_0, M_1) \end{array}\right\} \\
&\le \left(1 - 2^{n(I(N_2;M_0,N_0,M_1)+3\epsilon)}\right)^{2^{n(L_{22}-R_{22})}} \\
&\le e^{-2^{n(I(N_2;M_0,N_0,M_1)+3\epsilon-L_{22}+R_{22})}}.
\end{aligned}$$

These two probabilities decay to 0 as $n \to \infty$ if the following two conditions satisfy.

$$L_{21} - R_{21} > I(M_2; M_0, N_0, M_1) + 3\epsilon, \tag{1.2}$$

$$L_{22} - R_{22} > I(N_2; M_0, N_0, M_1) + 3\epsilon.. \tag{1.3}$$

From Lemma 5 in [48], we find that $\Pr\left\{E_0^2\right\}$ reduces to 0 with (1.2) and (1.3) as n tends to infinity. Also for $\Pr\left\{E_1^1\right\}$ and $\Pr\left\{E_1^2\right\}$, we observe that the probability goes to 0 as n approaches infinity, which comes from the Markov lemma. Next, we make an error analysis for $\Pr\left\{E_{ghijs_{21}k}^1\right\}$ and $\Pr\left\{E_{js_{21}ks_{22}l}^2\right\}$. We suppose that indices $\hat{k}$ and $\hat{l}$ are chosen and that

$$T_1 = \left\{ \begin{array}{l} \left(y_1^n, m_0^n(1), n_0^n(1), m_1^n(1), n_1^n(1), m_2^n(1,1,\hat{k})\right) \\ \in A_\epsilon^n(Y_1, M_0, N_0, M_1, N_1, M_2) \end{array} \right\}$$

$$T_2 = \left\{ \begin{array}{l} \left(y_2^n, m_1^n(1), m_2^n(1,1,\hat{k}), n_2^n(1,1,\hat{l})\right) \\ \in A_\epsilon^n(Y_2, M_1, M_2, N_2) \end{array} \right\}$$

.

Then, for any $\bar{k} \neq \hat{k}$ and $\bar{l} \neq \hat{l}$, we have

$$\begin{aligned} \Pr\left\{E^1_{11111\bar{k}}|T_1\right\} &= \Pr\left\{E^1_{11112\hat{k}}|T_1\right\} = \Pr\left\{E^1_{11112\bar{k}}|T_1\right\}, \\ \Pr\left\{E^2_{11\bar{k}1\hat{l}}|T_2\right\} &= \Pr\left\{E^2_{12\hat{k}1\hat{l}}|T_2\right\} = \Pr\left\{E^2_{12\bar{k}1\hat{l}}|T_2\right\}, \\ \Pr\left\{E^2_{11\hat{k}1\bar{l}}|T_2\right\} &= \Pr\left\{E^2_{11\hat{k}2\hat{l}}|T_2\right\} = \Pr\left\{E^2_{11\hat{k}2\bar{l}}|T_2\right\}. \end{aligned}$$

Then,

$$\begin{aligned} &\sum_{ghijs_{21}k \neq 11111\hat{k}} \Pr\left\{E^1_{ghijs_{21}k}|T_1\right\} \\ &= \sum_{ghijs_{21}k \neq 11111\hat{k}} \Pr\left\{ \begin{array}{l} \left(\begin{array}{l} y_1^n, m_0^n(g), n_0^n(h), \\ m_1^n(i), n_1^n(j), m_2^n(1, s_{21}, k) \end{array} \right) \\ \in A_\epsilon^n(Y_1, M_0, N_0, M_1, N_1, M_2)|T_1 \end{array} \right\} \\ &\leq \left(2^{n(R_{01}-6\epsilon)} - 1\right) \Pr\left\{E^1_{21111\hat{k}}|T_1\right\} \\ &\quad + \left(2^{n(R_{02}-6\epsilon)} - 1\right) \Pr\left\{E^1_{12111\hat{k}}|T_1\right\} \\ &\quad + \left(2^{n(R_{11}-6\epsilon)} - 1\right) \Pr\left\{E^1_{1211\hat{k}}|T_1\right\} \\ &\quad + \left(2^{n(R_{12}-6\epsilon)} - 1\right) \Pr\left\{E^1_{1121\hat{k}}|T_1\right\} \\ &\quad + \left(2^{n(L_{21}-6\epsilon)} - 1\right) \Pr\left\{E^1_{1112\bar{k}}|T_1\right\} \\ &\quad + \left(2^{n(R_{01}-6\epsilon)} - 1\right)\left(2^{n(R_{02}-6\epsilon)} - 1\right) \Pr\left\{E^1_{22111\hat{k}}|T_1\right\} \\ &\quad + \left(2^{n(R_{01}-6\epsilon)} - 1\right)\left(2^{n(R_{11}-6\epsilon)} - 1\right) \Pr\left\{E^1_{21211\hat{k}}|T_1\right\} \end{aligned}$$

$$
\begin{aligned}
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{21121\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{21112\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\Pr\left\{E^1_{12211\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{12121\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{12112\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{11221\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{11212\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{11122\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\Pr\left\{E^1_{22211\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{22121\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{22112\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{21221\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{21212\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{21122\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{12221\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{12212\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{12122\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{11222\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\Pr\left\{E^1_{22221\hat{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{22212\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{22122\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{21222\bar{k}}|T_1\right\}\\
&+\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{12222\bar{k}}|T_1\right\}\\
&+\left(\begin{array}{c}\left(2^{n(R_{01}-6\epsilon)}-1\right)\left(2^{n(R_{02}-6\epsilon)}-1\right)\left(2^{n(R_{11}-6\epsilon)}-1\right)\\ \left(2^{n(R_{12}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right)\Pr\left\{E^1_{22222\bar{k}}|T_1\right\}.\end{array}\right)
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&\sum_{ghijs_{21}k \neq 11111\hat{k}} \Pr\left\{E^1_{ghijs_{21}k} | T_1\right\} \\
\leq\ & 2^{n(R_{01}-I(M_0;Y_1,N_0,M_1,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{02}-I(N_0;Y_1,M_0,M_1,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{11}-I(M_1;Y_1,M_0,N_0,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{12}-I(N_1;Y_1,M_0,N_0,M_1,M_2)-4\epsilon)} \\
&+ 2^{n(L_{21}-I(M_2;Y_1,M_0,N_0,M_1,N_1)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{02}-I(M_0,N_0;Y_1,M_1,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{11}-I(M_0,M_1;Y_1,N_0,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{12}-I(M_0,N_1;Y_1,N_0,M_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{01}+L_{21}-I(M_0,M_2;Y_1,N_0,M_1,N_1)-4\epsilon)} \\
&+ 2^{n(R_{02}+R_{11}-I(N_0,M_1;Y_1,M_0,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{02}+R_{12}-I(N_0,N_1;Y_1,M_0,M_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{02}+L_{21}-I(N_0,M_2;Y_1,M_0,M_1,N_1)-4\epsilon)} \\
&+ 2^{n(R_{11}+R_{12}-I(M_1,N_1;Y_1,M_0,N_0,M_2)-4\epsilon)} \\
&+ 2^{n(R_{11}+L_{21}-I(M_1,M_2;Y_1,M_0,N_0,N_1)-4\epsilon)} \\
&+ 2^{n(R_{12}+L_{21}-I(N_1,M_2;Y_1,M_0,N_0,M_1)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{02}+R_{11}-I(M_0,N_0,M_1;Y_1,N_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{02}+R_{12}-I(M_0,N_0,N_1;Y_1,M_1,M_2)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{02}+L_{21}-I(M_0,N_0,M_2;Y_1,M_1,N_1)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{11}+R_{12}-I(M_0,M_1,N_1;Y_1,N_0,M_2)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{11}+L_{21}-I(M_0,M_1,M_2;Y_1,N_0,N_1)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{12}+L_{21}-I(M_0,N_1,M_2;Y_1,N_0,M_1)-4\epsilon)} \\
&+ 2^{n(R_{02}+R_{11}+R_{12}-I(N_0,M_1,N_1;Y_1,M_0,M_2)-4\epsilon)} \\
&+ 2^{n(R_{02}+R_{11}+L_{21}-I(N_0,M_1,M_2;Y_1,M_0,N_1)-4\epsilon)} \\
&+ 2^{n(R_{02}+R_{12}+L_{21}-I(N_0,N_1,M_2;Y_1,M_0,M_1)-4\epsilon)} \\
&+ 2^{n(R_{11}+R_{12}+L_{21}-I(M_1,N_1,M_2;Y_1,M_0,N_0)-4\epsilon)} \\
&+ 2^{n(R_{01}+R_{02}+R_{11}+R_{12}-I(M_0,N_0,M_1,N_1;Y_1,M_2)-4\epsilon)}
\end{aligned}
$$

$$+ 2^{n(R_{01}+R_{02}+R_{11}+L_{21}-I(M_0,N_0,M_1,M_2;Y_1,N_1)-4\epsilon)}$$
$$+ 2^{n(R_{01}+R_{02}+R_{12}+L_{21}-I(M_0,N_0,N_1,M_2;Y_1,M_1)-4\epsilon)}$$
$$+ 2^{n(R_{01}+R_{11}+R_{12}+L_{21}-I(M_0,M_1,N_1,M_2;Y_1,N_0)-4\epsilon)}$$
$$+ 2^{n(R_{02}+R_{11}+R_{12}+L_{21}-I(N_0,M_1,N_1,M_2;Y_1,M_0)-4\epsilon)}$$
$$+ 2^{n(R_{01}+R_{02}+R_{11}+R_{12}+L_{21}-I(M_0,N_0,M_1,N_1,M_2;Y_1)-4\epsilon)},$$

and

$$\sum_{js_{21}ks_{22}l\neq 11\hat{k}1\hat{l}} \Pr\left\{E^1_{js_{21}ks_{22}l}|T_2\right\}$$
$$= \sum_{js_{21}ks_{22}l\neq 11\hat{k}1\hat{l}} \Pr\left\{ \begin{array}{l} (y_2^n, m_1^n(j), m_2^n(1,s_{21},k), n_2^n(1,s_{22},l)) \\ \in A_\epsilon^n(Y_2,M_1,M_2,N_2)|T_2 \end{array} \right\}$$
$$\leq \left(\begin{array}{l} \left(2^{n(R_{11}-6\epsilon)}-1\right) \Pr\left\{E^1_{21\hat{k}1\hat{l}}|T_2\right\} \\ + \left(2^{n(L_{21}-6\epsilon)}-1\right) \Pr\left\{E^1_{12\bar{k}1\hat{l}}|T_2\right\} \\ + \left(2^{n(L_{22}-6\epsilon)}-1\right) \Pr\left\{E^1_{11\hat{k}2\bar{l}}|T_2\right\} \\ + \left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right) \Pr\left\{E^1_{22\bar{k}1\hat{l}}|T_2\right\} \\ + \left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{22}-6\epsilon)}-1\right) \Pr\left\{E^1_{21\hat{k}2\bar{l}}|T_2\right\} \\ + \left(2^{n(L_{21}-6\epsilon)}-1\right)\left(2^{n(L_{22}-6\epsilon)}-1\right) \Pr\left\{E^1_{12\bar{k}2\bar{l}}|T_2\right\} \\ + \begin{array}{l} \left(2^{n(R_{11}-6\epsilon)}-1\right)\left(2^{n(L_{21}-6\epsilon)}-1\right) \\ \left(2^{n(L_{22}-6\epsilon)}-1\right) \Pr\left\{E^1_{22\bar{k}2\bar{l}}|T_2\right\} \end{array} \end{array} \right)$$
$$\leq \left(\begin{array}{l} 2^{n(R_{11}-I(M_1;Y_2,M_2,N_2)-4\epsilon)} \\ +2^{n(L_{21}-I(M_2;Y_2,M_1,N_2)-4\epsilon)} \\ +2^{n(L_{22}-I(N_2;Y_2,M_1,M_2)-4\epsilon)} \\ +2^{n(R_{11}+L_{21}-I(M_1,M_2;Y_2,N_2)-4\epsilon)} \\ +2^{n(R_{11}+L_{22}-I(M_1,N_2;Y_2,M_2)-4\epsilon)} \\ +2^{n(L_{21}+L_{22}-I(M_2,N_2;Y_2,M_1)-4\epsilon)} \\ +2^{n(R_{11}+L_{21}+L_{22}-I(M_1,M_2,N_2;Y_2)-4\epsilon)} \end{array} \right).$$

If all the equation in (4.38) satisfies, probability or error decays to 0 as n becomes large.

1.2 Proof of the Achievable Rate Region by Fourier Motzkin Elimination

From (4.43), (4.44), and (4.45), we have the following rate region with rate quintuple $(R_0, R_{11}, R_{12}, R_{21}, R_{22})$ satisfying

$$\begin{aligned}
&R_0 \leq r_0, && R_{11} \leq r_1, \\
&R_{12} \leq r_2, && R_{21} \leq r_3, \\
&R_0 + R_{11} \leq r_4, && R_0 + R_{12} \leq r_5, \\
&R_0 + R_{21} \leq r_6, && R_1 \leq r_7, \\
&R_{11} + R_{21} \leq r_8, && R_{12} + R_{21} \leq r_9, \\
&R_0 + R_1 \leq r_{10}, && R_0 + R_{11} + R_{21} \leq r_{11}, \\
&R_0 + R_{12} + R_{21} \leq r_{12}, && R_1 + R_{21} \leq r_{13}, \\
&R_0 + R_1 + R_{21} \leq r_{14}, && R_{12} \leq r_{15}, \\
&R_{21} \leq r_{16}, && R_{12} + R_{21} \leq r_{17}, \\
&R_{22} \leq r_{18}.
\end{aligned} \tag{1.4}$$

Since $R_1 = R_{11} + R_{12}$, replace R_{12} with $R_1 - R_{11}$ in inequalities. Also, since $R_2 = R_{21} + R_{22}$, replace R_{22} with $R_2 - R_{21}$ in inequalities. By collecting all inequalities with R_{21} in it, we have

$$\begin{aligned}
&0 \leq R_{21}, \\
&R_2 - r_{18} \leq R_{21},
\end{aligned} \tag{1.5}$$

and

$$\begin{aligned}
&R_{21} \leq r_3,\\
&R_{21} \leq r_6 - R_0,\\
&R_{21} \leq r_8 - R_{11},\\
&R_{21} \leq r_9 - R_1 + R_{11},\\
&R_{21} \leq r_1 1 - R_1 - R_{11},\\
&R_{21} \leq r_1 2 - R_0 - R_1 + R_{11},\\
&R_{21} \leq r_1 3 - R_1,\\
&R_{21} \leq r_1 4 - R_0 - R_1,\\
&R_{21} \leq r_1 6,\\
&R_{21} \leq r_1 2 - R_0 + R_{11},\\
&R_{21} \leq R_2.
\end{aligned} \tag{1.6}$$

All the left-hand sides of (1.5) are less than equal to the right-hand sides of (1.6). Thus, we have

$$\begin{aligned}
&R_2 \leq r_3 + r_{18},\\
&R_2 + R_0 \leq r_6 + r_{18},\\
&R_2 + R_{11} \leq r_8 + r_{18},\\
&R_2 + R_1 - R_{11} \leq r_9 + r_{18},\\
&R_2 + R_1 + R_{11} \leq r_1 1 + r_{18},\\
&R_2 + R_0 + R_1 - R_{11} \leq r_1 2 + r_{18},\\
&R_2 + R_1 \leq r_1 3 + r_{18},\\
&R_2 + R_0 + R_1 \leq r_1 4 + r_{18},\\
&R_2 \leq r_1 6 + r_{18},\\
&R_2 + R_0 - R_{11} \leq r_1 2 + r_{18}.
\end{aligned} \tag{1.7}$$

Then, we accumulate the inequalities with R_{11} and compare the lower bounds and upper bounds on R_{11}. Then, we have

$$\begin{aligned}
&0 \leq R_{11},\\
&R_1 - r_2 \leq R_{11},\\
&R_1 - r_{15} \leq R_{11},\\
&R_0 + R_1 - r_5 \leq R_{11},\\
&R_1 + R_2 - r_9 - r_{18} \leq R_{11},\\
&R_1 + R_2 - r_{17} - r_{18} \leq R_{11},\\
&R_0 + R_1 + R_2 - r_{12} - r_{18} \leq R_{11},
\end{aligned} \tag{1.8}$$

and

$$\begin{aligned}
&R_{11} \le r_1, \\
&R_{11} \le r_4 - R_0, \\
&R_{11} \le r_8 + r_{18} - R_2, \\
&R_{11} \le r_1 1 + r_{18} - R_0 - R_2, \\
&R_{11} \le R_1.
\end{aligned} \tag{1.9}$$

We eliminate R_{11}, and have inequalities that all the left-hands side of (1.8) are less than equal to the right-hand sides of (1.9). Then, by collecting all the inequalities and removing the redundant ones, we have

$$\begin{aligned}
&R_0 \le r_0, \\
&R_1 \le \min(r_7, r_1 + r_{15}), \\
&R_2 \le \min(r_3 + r_{18}, r_{16} + r_{18}), \\
&R_0 + R_1 \le \min(r_{10}, r_4 + r_{15}), \\
&R_0 + R_2 \le r_6 + r_{18}, \\
&R_1 + R_2 \le \min(r_{13} + r_{18}, r_8 + r_{15} + r_{18}, r_1 + r_{17} + r_{18}), \\
&R_0 + R_1 + R_2 \\
&\le \min(r_{14} + r_{18}, r_{11} + r_{15} + r_{18}, r_4 + r_{17} + r_{18}), \\
&2R_0 + R_1 \le r_4 + r_5, \\
&R_1 + 2R_2 \le \min(r_8 + r_9 + 2r_{18}, r_8 + r_{17} + 2r_{18}), \\
&2R_0 + R_1 + R_2 \le \min(r_5 + r_{11} + r_{18}, r_4 + r_{12} + r_{18}), \\
&R_0 + R_1 + 2R_2 \\
&\le \min(r_9 + r_{11} + 2r_{18}, r_8 + r_{!2} + 2r_{18}, r_{11} + r_{17} + 2r_{18}), \\
&2R_0 + R_1 + 2R_2 \le r_{11} + r_{12} + 2r_{18}.
\end{aligned} \tag{1.10}$$

This ends the proof.

Bibliography

[1] J. Mitola, "Cognitive Radio," Ph.D. dissertation, Royal Institute of Technology (KTH), Stockholm, Sweden, 2000.

[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas in Commun.*, vol. 23, pp. 201-220, Feb. 2005.

[3] S. Srinivasa and S. Jafar, "The Throughput Potential of Cognitive Radio: A Theoretical Perspective," *Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2006.

[4] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective," *IEEE Proceedings*, vol. 97, pp. 894-914, May. 2009.

[5] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, pp. 79-89, May. 2007.

[6] K. Son, B. Jung, S. Chong, and D. Sung, "Opportunistic Underlay Transmission in Multi-Carrier Cognitive Radio Systems," *Wireless Communications and Networking Conf. (WCNC)*, Budapest, Hungary, Apr. 2009.

[7] M. Gastpar, "On Capacity under Receive and Spatial Spectrum-Sharing Constraints," *IEEE Trans. Info. Theory*, vol. 53, pp. 471-487, Feb. 2007.

[8] A. Ghasemi and E. Sousa, "Capacity of Fading Channels Under Spectrum-Sharing Constraints," *IEEE International Conf. on Communications*, Istanbul, Turkey, Jun. 2006.

[9] H. Cho and J. G. Andrews, "Upper Bound on the Capacity of Cognitive Radio Without Cooperation" *IEEE Trans. Wireless Communications*, vol. 8, pp. 5314-5327, Oct. 2009.

[10] J. Mitola, "Cognitive Radio Architecture Evolution," *IEEE Proceedings*, vol. 97, pp. 626-641, Apr. 2009.

[11] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," *Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Apr. 2004.

[12] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, pp. 523.531, Apr. 1967

[13] F. Digham, M. Alouini, and M. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Communications*, vol. 55, pp. 21-24, Jan. 2007.

[14] S. Hong, M. Vu, and V. Tarokh, "Cognitive Sensing Based on Side Information," *IEEE Sarnoff Symp.*, Princeton, NJ, Apr. 2008.

[15] W. Gardner, "Signal Interception: A Unifying Theoretical Framework for Feature Detection", *IEEE Trans. on Communications*, vol. 36, pp. 897-906, Aug. 1988

[16] H. Kim and K. Shin, "In-band spectrum sensing in cognitive radio networks: energy detection or feature detection?" in *Proceedings of the 14th ACM international conference on Mobile computing and networking* pp. 14-25, 2008.

[17] J. Bazerque and G. Giannakis, "Distributed spectrum sensing for cognitive radios by exploiting sparsity," *Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Apr. 2008.

[18] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio networks," *IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD, Nov. 2005.

[19] G. Chung and S. Vishwanath, "Sensing while Transmission for Cognitive Radios: Limits and Strategies," Submitted to *IEEE Trans. Communications.*

[20] A. Sahai, N. Hoven, S. Mishra, and R. Tandra, "Fundamental Tradeoffs in Robust Spectrum Sensing for Opportunistic Frequency Reuse," *Technical Report* Mar. 2006. Available at
http://www.eecs.berkeley.edu/ sahai/Papers/CognitiveTechReport06.pdf.

[21] R. Tandra and A. Sahai, "SNR Walls for signal detection," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, pp. 4-17, Feb. 2008.

[22] R. Tandra and A. Sahai, "SNR Walls for Feature Detectors," *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Dublin, Island, Apr. 2007.

[23] Y. Pei, A. Hoang, and Y. Liang, "Sensing-Throughput Tradeoff in Cognitive Radio Networks: How Frequently Should Spectrum Sensing be Carried Out?", *IEEE Symposium on Personal, Indoor and Mobile Radio Communications*, Athens, Greece, Dec. 2007.

[24] G. Li, J. Fang, H. Tan, and J. Li, "The impact of time-bandwidth product on the energy detection in the cognitive radio," *IEEE Intl. Conf. on Broadband Network and Multimedia Technology* (IC-BNMT), Beijing, China, Oct, 2010.

[25] A. Baricz, "Bounds for modified Bessel functions of the first and second kind," *Proceedings of the Edinburgh Mathematical Society*, In press.

[26] M. Duarte and A. Sabharwal, "full duplex Wireless Communications Using Off-The-Shelf Radios: Feasibility and First Results," *Asilomar Conference on Signals, Systems, and Computers*, Asilomar, CA, Nov. 2010.

[27] D. Bliss, P. Parker, and A. Margetts, "Simultaneous Transmission and Reception for Improved Wireless Network Performance," *IEEE Workshop on Statistical Signal Processing*, Madison, WI, Aug. 2007.

[28] T. Schmid, "Gnu radio 802.15.4 en-and decoding," available at http://nesl.ee.ucla.edu/fw/thomas/thomas_project_report.pdf.

[29] IEEE Standards Association, "*IEEE 802.11: Wireless Local Area Networks Specification*", available at http://standards.ieee.org/about/get/802/802.11.html.

[30] 3GPP, "*3GPP LTE Specification*", available at http://www.3gpp.org/-specifications-

[31] Y. Li, J.Zhou, and Y. Onozato, "Performance Study for STC-OFDM Systems Based on IEEE802.11a Standard," *IEEE Intl. Conf. on Wireless Communications, Networking and Mobile Computing*, Wuhan, China, Sep. 2009.

[32] P. Zhou and R. Guo"Design of Video NLOS Communication System Using MIMO-OFDM Based on IEEE802.16a Standard," *IEEE Intl. Conf. on Wireless Communications, Networking and Mobile Computing*, Wuhan, China, Sep. 2006.

[33] C. Shin, J. G. Andrews, and E. Powers, "An Efficient Design of Doubly Selective Channel Estimation for OFDM Systems," *IEEE Trans. on Communications*, vol. 6 , pp. 3790-3802, Oct. 2007.

[34] Quellan Inc. "Qhx220 narrowband noise canceller ic.," available at http://www.quellan.com/products/qhx220_ic.php.

[35] K. Jung, C. Park, and K. Lee, "Bit and Power Allocation for MIMO-OFDM Systems with Spatial Mode Selection over Frequency-space-time Selective Channels," *IEEE Conf. on Vehicular Technolgy*, Los Angeles, CA, Sep. 2004.

[36] S. Yun and C. Caramanis, "Reinforcement Learning for Link Adaptation in MIMO-OFDM Wireless Systems," *IEEE Global Telecommunications Conf.*, Miami, FL, Dec. 2010.

[37] S. Vishwanath, S. Jafar, and A. Goldsmith, "Adaptive resource allocation in composite fading environments," *IEEE Global Telecommunications Conf.*, San Antonio, TX, Nov. 2001.

[38] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2008.

[39] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "Overview of Limited Feedback in Wireless Communication Systems," *IEEE Journal on Sel. Areas in Comm., Special Issue on Exploiting Limited Feedback in Tomorrow Wireless Communication Networks,* vol. 26, no. 8, pp. 1341-1365, Oct. 2008.

[40] D. Landon and C. Furse, "The MIMO transmission equation," *IEEE Symp. on Antennas and Propagation Society*, San Diego, CA, Jul. 2008.

[41] D. Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media," U.S. Patent 4,833,706, May, 1989.

[42] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view", *IEEE Transactions on Communications*, pp. 1561-1571, Sept. 2001.

[43] Y. Song, Y. Fang, and Y. Zhang, "Stochastic Channel Selection in Cognitive Radio Networks," *IEEE Global Communications Conf.*, Washington, DC, Nov. 2007.

[44] X. Yang, Z. Yang, and D. Liao, "Adaptive Spectrum Selection for Cognitive Radio Networks," *International Conf. on Computer Science and Software Engineering*, Wuhan, China, Dec. 2008.

[45] D. Huang, C. Miao, C. Leung, and Z. Shen, "Resource Allocation of MU-OFDM Based Cognitive Radio Systems Under Partial Channel State Information," http://arxiv.org/abs/0808.0549

[46] G. Chung, S. Vishwanath, and C. S. Hwang, "On the Fundamental Limits of Interweaved Cognitive Radios," *IEEE Radio and Wireless Symposium*, New Orleans, LA, Jan. 2010.

[47] L. Lai, E. Gamal, H. Jiang, and V. Poor, " Optimal medium access protocols for cognitive radio networks," *Internaltional Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (Wiopt)*, Berlin, Germany, Mar. 2008.

[48] N. Devroye, P. Mitran, and V. Tarokh, "Achievable Rates in Cognitive Rado Channels," *IEEE Trans. Info. Theory*, vol. 52, pp. 1813-1827, May 2006.

[49] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai (Shitz), "On the Capacity of Interference Channel with a Partially-Cognitive Transmitter," *IEEE International Symp. on Information Theory*, Nice, France, Jun. 2007.

[50] I. Maric and R. Yates, "The Strong Interference Channel with Common Information," *Allerton Conf. Communications*, Monticello, Illinois, Sep. 2005.

[51] I. Maric, R. Yates, and G. Kramer, "The strong interference channel with unidirectional cooperation," presented at the Information Theory and Applications (ITA) Inaugural Workshop, Feb. 2006.

[52] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a Class of Cognitive Radio Channels: Interference Channels With Degraded Message Sets," *IEEE Trans. Info. Theory*, vol. 53, pp. 4391-4399, Nov. 2007.

[53] A. Jovicic and P. Viswanath, "Cognitive radio: an information-theoretic perspective," *IEEE Trans. Info. Theory*, vol. 55, pp. 3945-3958, Sep. 2009.

[54] C. Huang and S. Jafar, "Degrees of freedom of the MIMO interference channel with cooperation and cognition," *IEEE International Symp. on Information Theory*, Toronto, ON, Jul. 2008.

[55] G. Chung, S. Sridharan, S. Vishwanath, and C. S. Hwang, "On the Capacity of Partially Cognitive Radios," *IEEE International Symp. on Information Theory*, Seoul, Korea, Jun. 2009.

[56] G. Chung and S. Vishwanath, "Capacity of a class of overlay cognitive MAC radios," *IEEE International Symp. on Information Theory*, St. Petersburg, Russia, Jul. 2011.

[57] A. Lozano, R. W. Heath Jr., and J. G. Andrews, "Fundamental Limits of Cooperation," Submitted to *IEEE Trans. Info. Theory*.

[58] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, Achievable Rates and Sum-Rate Capacityof Gaussian MIMO Broadcast Channels," *IEEE Trans. Communications*, vol. ,pp. , .

[59] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative Sensing among Cognitive Radios," *International Conf. on Communication*, Istanbul, Turkey, Jun. 2006

[60] National Telecommunications and Information Administration (NTIA), "FCC Frequency Allocation Chart, 2003" available at www.ntia.doc.gov/osmhome/allochrt.pdf.

[61] Federal Communications Commision, “Spectrum Policy Task Force,” Rep. ET Docket No. 02-135, Nov. 2002.

[62] Spectrum Sharing Company “Spectrum Occupancy Measurements Loring Commerce Centre Limestone, Maine, September 18-20, 2007” available at www.sharedspectrum.com/measurements/download/Loring_Spectrum_Occupancy_Measurements_v2_3.pdf

[63] J. Kim, J. G. Andrews, “Spectrum sensing with spectral covariance for cognitive radio”, *IEEE Trans. on Signal Processing, in preparation.* Available: http://arxiv.org/abs/0910.1787

[64] U.S. FCC, ET Docket 04-186, “Notice of Proposed Rule Making, in the matter of Unlicensed Operation in the TV Broadcast Bands,” May 25, 2004.

[65] U.S. FCC, ET Docket 08-260, “Second Report and Order and Memorandum Opinion and Order, in the Matter of Unlicensed Operation in the TV Broadcast Bands Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band,” Oct. 18, 2006.

[66] IEEE 802.22/D1.0, “Draft Standard for Wireless Regional Area Networks Part 22: Cognitive Wireless RAN Medium Access Control (MAC) specifications: Policies and procedures for operation in the TV bands,” Apr. 2008.

[67] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell, “IEEE 802.22: The first cognitive radio wireless regional area network standard,” IEEE Communications Mag., vol. 47, pp. 130-138, Jan. 2009.

[68] B. Chen, H. Huang, Z. Chao, X. Yang, L. Lei, "Future Architecture with Dynamic Spectrum Access and Cross-layer Design: An Overview", *Wireless World Research Forum*, Ottawa, ON, Apr. 2008.

[69] D. Cabric, S.M. Mishra, D. Willkomm, R. Brodersen, A. Wolisz, "A Cognitive radio approach for usage of virtual unlicensed spectrum," *14th IST Mobile and Wireless Communications Summit*, Dresden, Germany, Jun. 2005.

[70] J. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving Single Channel, Full Duplex Wireless Communication," *Intl. Conf. on Mobile Computing and Networking*, Chicago, IL, Sep. 2010.

[71] G. Chung, J. Ryu, K. Srinivasan, S. Vishwanath, "Link Adaptation in Wideband Full Duplex Radios," Submitted to *IEEE Trans. Wireless Communications.*

[72] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Info. Theory*, vol. 27, pp. 49-60, Jan. 1981.

[73] H. Chong, M. Motani, H. Garg, and H. El Gamal, "On the Han-Kobayashi region for the interference channel" submitted to *IEEE Trans. Info. Theory*, Aug. 2006.

[74] H. Sato, "Two-user communication channels," *IEEE Trans. Info. Theory*, vol. 23, pp. 295-304, May 1977.

[75] A. B. Carleial, "Outer bounds on the capacity of interference channels," *IEEE Trans. Info. Theory*, vol. 29, pp. 602-606, Jul. 1983.

[76] M. Shahid and J. Kamruzzaman, "Weighted Soft Decision for Cooperative Sensing in Cognitive Radio Networks," *IEEE International conf. on Network,* New Delhi, India, Dec. 2008.

[77] S. Shellhammer, "Spectrum Sensing in IEEE 802.22," *IAPR Wksp. Cognitive Info. Processing*, June 2008.

[78] Y. Zeng and Y. Liang, "Spectrum-sensing algorithms for cognitive radio based on statistical covariances," *IEEE Trans. Vehicular Technology*, vol. 58, pp. 1804-1815, May 2009.

[79] N. Han, S. Shon, J. Chung, and J. Kim, "Spectral correlation based signal detection method for spectrum sensing in ieee 802.22 wran systems," *International Conf. Advanced Communication Technology,* Korea, Feb. 2006.

[80] F. Visser, G. Janssen, and P. Pawelczak, "Multinode Spectrum Sensing Based on Energy Detection for Dynamic Spectrum Access," *IEEE Vehicular Technology Conf.*, Marina Bay, Singapore, May 2008.

[81] H. Uchiyama, K. Umebayashi, T. Fujii, F. Ono, K. Sakaguchi, Y. Kamiya, and Y. Suzuki, "Study on Soft Decision Based Cooperative Sensing for Cognitive Radio Networks," *IEICE Trans. Communications*, vol. E91-B, pp. 95-101, Jan. 2008.

[82] E. Visotsky, S. Kuffner and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," *IEEE International*

Symp. on New Frontiers in Dynamic Spectrum Access Networks, Baltimoren, MD, Nov. 2005.

[83] J. Tsitsiklis and M. Athans, "On the Complexity of Decentralized Decision Making and Detection Problems", *IEEE Trans. Automatic Control,* Vol. 30, pp. 440-446, 1998.

[84] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Info. Theory*, vol. 33, pp. 711-716, Sep. 1987.

[85] H. Weingarten, Y. Steinberg and S. Shamai (Shitz), "The Capacity region of the Gaussian MIMO broadcast channel," *IEEE Trans. Info. Theory*, vol. 52, pp. 3936-3964, Sep. 2006.

[86] M. Costa, "Writing on dirty paper," *IEEE Trans. Info. Theory*, vol. 29, pp. 439-441, May 1983.

[87] T. M. Cover and J. A. Thomas, *Elements of information theory*, John Wiley & Sons Inc., 1991

[88] R. Etkin, D. Tse, and H. Wang. "Gaussian interference channel capacity to within one bit: the general case," Proc. IEEE International Symp. on Information Theory, 2007.

[89] A. Somekh-Baruch, S. Sridharan, S. Vishwanath, S. Verdu, S. Shamai, "On the Capacity of Cognitive Radios in Multiple Access Networks," *Asilomar Conference on Signal, Systems and Computers*, Pacific Grove, California, Nov. 2008.

IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications

Vita

Goochul Chung is currently pursuing a Ph.D. in the department of Electrical and Computer Engineering at The University of Texas at Austin (UT Austin). He is advised by Prof. Sriram Vishwanath. He received his B.S. and M.E. degrees in Electrical Engineering from Seoul National University in 2002 and 2004, respectively. His industry experience includes the work as Engineering Staff at Pantech Co., Seoul, Korea, as Researching Staff at Samsung Advanced Institute of Technology, Yong-in, Korea.

Permanent address: 2501 Lake Austin Blvd. G102
Austin, Texas 78703

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.