ON THE ROLE OF COGNITION IN WIRELESS NETWORKS: AN INFORMATION THEORETIC PERSPECTIVE

BY

STEFANO RINI
B.A. in Computer Engineering, Universitá degli studi di Milano, 2004
M.A. in Statistics, University of Illinois at Chicago, 2009
M.A. in Electrical Engineering, University of Illinois at Chicago, 2010

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Chicago, 2010

Chicago, Illinois

TABLE OF CONTENTS

CHAPTER			PAGE
1	**INTRODUCTION**		**1**
	1.1	Notation	9
2	**THE DISCRETE MEMORYLESS COGNITIVE INTERFERENCE CHANNEL**		10
	2.1	Main contributions	10
	2.2	Organization	11
	2.3	Channel model, notation and definitions	11
	2.4	Available results for the DM-CIFC	13
	2.5	A new outer bound	18
	2.6	A new inner bound	24
	2.6.1	Rate splitting	28
	2.6.2	Codebook generation	29
	2.6.3	Encoding	30
	2.6.4	Decoding	32
	2.6.5	Error analysis	33
	2.6.6	Encoding Error	34
	2.6.7	Decoding Errors at decoder 2	41
	2.6.8	Decoding Errors at Decoder 1	46
	2.6.9	Two step binning	48
	2.7	Comparison with existing achievable regions	53
	2.7.1	Containment of region of Devroye et al. in $\mathcal{R}_{RTD}$	53
	2.7.2	Containment of the region of Biao et al. in $\mathcal{R}_{RTD}$	64
	2.7.3	Containment of the region of Jiang et al. in $\mathcal{R}_{RTD}$:	69
	2.8	New capacity results for the DM-CIFC	75
	2.9	Capacity for the semi-deterministic CIFC	82
	2.10	Capacity for the deterministic CIFC	88
	2.11	Examples	89
	2.11.1	Example I: the “Asymmetric Clipper”	91
	2.11.2	Example II: the “Symmetric Clipper”	97
3	**THE GAUSSIAN COGNITIVE INTERFERENCE CHANNEL**		101
	3.1	Main Contributions	101
	3.2	Organization	103
	3.3	Gaussian channel model and known results	104
	3.3.1	Gaussian CIFC	104
	3.4	The G-CIFC in standard form	105

TABLE OF CONTENTS (Continued)

CHAPTER		PAGE
3.4.1	Known results for the G-CIFC	107
3.5	Outer bounds	110
3.5.1	A unifying framework for Th. 3.4.1 and Th. 3.4.2	111
3.5.2	BC- based outer bounds	114
3.5.3	Outer bounds by transformation	130
3.6	Inner bounds	137
3.6.1	Achievable scheme with U_{2pb} and U_{1pb}: capacity achieving for the degraded broadcast channel.	143
3.6.2	Achievable scheme with X_2 and U_{1pb}: capacity achieving in the "weak interference" regime.	144
3.6.3	Achievable scheme with X_2, U_{1pb} and U_{2pb}: capacity achieving in the semi-deterministic DM-CIFC.	144
3.6.4	Achievable scheme with U_{1c} and U_{2c}: capacity achieving in "very strong interference" regime.	146
3.6.5	Achievable scheme with X_2, U_{1c}: capacity achieving in the primary decodes cognitive regime.	147
3.6.6	Achievable scheme with U_{2c}, X_2 and U_{1c}.	148
3.7	New capacity results	150
3.7.1	New capacity results for the C-CIFC.	153
3.7.2	New capacity results for the S-G-CIFC.	158
3.8	Capacity to within a constant gap	160
3.8.1	Additional constant gap results	166
3.8.1.1	Perfect interference cancellation	169
3.8.1.2	Non perfect interference cancellation	170
3.8.2	Further considerations on the utility of partial interference "pre-cancellation"	175
3.8.2.1	Cognitive broadcasting	176
3.8.2.2	Interference stripping	178
3.9	Numerical results	180
3.9.1	Numerical Results for Section 3.5: Outer bounds	181
3.9.2	Numerical Results for Section 3.6: Inner bounds	183
3.9.3	Numerical Results for Section 3.7: New capacity results	188
4	**CONCLUSIONS AND FUTURE WORK**	192
	CITED LITERATURE	196

LIST OF TABLES

TABLE		PAGE
I	ERROR EVENTS AT DECODER 2.	40
II	ERROR EVENTS AT DECODER 1.	46
III	ACHIEVABILITY FOR $(R_1, R_2) = (1, 3)$ IN EXAMPLE I.	94
IV	ACHIEVABILITY TABLE FOR THE RATE POINT $(R_1, R_2) = (2, 2)$ IN EXAMPLE I.	95
V	THE INPUT DISTRIBUTION FOR EXAMPLE II	98
VI	ACHIEVABILITY TABLE FOR THE RATE POINT $(R_1, R_2) = (1, 2)$ IN EXAMPLE II.	99

LIST OF FIGURES

FIGURE		PAGE
1	The CIFC model.	13
2	The achievable encoding scheme of Thm. 2.6.1. The ordering from left to right and the distributions demonstrate the codebook generation process. The dotted lines indicate binning. We see rate splits are used at both users, private messages $W_{1pb}, W_{2pa}, W_{2pb}$ are superimposed on common messages W_{1c}, W_{2c} and U_{1c}, is binned against (U_{2pa}, X_2) conditioned on U_{2c}, while U_{1pb} and U_{2pb} are binned against each and X_2 other in a Marton-like fashion (conditioned on other subsets of RV's). U_{1pb} is binned against U_{2pa} as well.	29
3	The region of the admissible binning rates R_{1pb} and R_{2pb} in Theorem 2.6.1.	53
4	The "asymmetric clipper" of Section 2.11.1.	90
5	The "symmetric clipper" of Section 2.11.2	97
6	The Gaussian cognitive interference channel (G-CIFC).	106
7	A representation of the capacity results Th. 3.4.1 and Th. 3.4.3 for $P_1 = P_2$ and $(a, \|b\|) \in [-5, 5] \times [0, 5]$. The regions for which capacity is known are shaded, while those for which capacity is unknown are white.	110
8	The "strong interference" outer bound of Th. 3.4.2 and the capacity region of the G-CIFC with $P_2 = 0$ and $\|b\| > 1$ (when the channel reduces to degraded BC).	118
9	The achievable schemes of Section 3.6.	142
10	The role of the different achievable schemes in the following Sections . .	143
11	The bound for R_1 in Equation 3.43a (bottom) and the bound for R_2 in Equation 3.43b (top) as a function of $\lambda \in \mathbb{R}$, for $P_1 = P_2 = 6$, $b = \sqrt{2}$, $a = \sqrt{0.3}$, $\alpha = 0.5$.	152

LIST OF FIGURES (Continued)

FIGURE		PAGE
12	A representation of the capacity result of Th. 3.7.1 for a G-CIFC with $a \in \mathbb{R}$, $P_1 = P_2 = 10$ and $(a, \|b\|) \in [-5, 5] \times [0, 5]$.	154
13	Condition Equation 3.50 for different values of $P_1 = P_2 = P$ for a G-CIFC with $a \in \mathbb{R}$ and $(a, \|b\|) \in [-0.5, 1, 5] \times [1, 5]$.	157
14	A schematic representation of the capacity results for the S-G-CIFC in Th. 3.7.3.	160
15	A graphical representation of Th. 3.8.1 and Th. 3.8.3.	167
16	G-CIFC with "strong interference": the "strong interference" outer bound, the piecewise linear "strong interference" outer bound and the achievable region by time sharing among points A and C;	179
17	The "strong interference" outer bound and the BC-based outer bound for the degraded G-CIFC.	182
18	The "strong interference" outer bound and the BC-based outer bound for the S-G-CIFC (right).	183
19	The "strong interference" outer bound and the MIMO BC outer bound for a general G-CIFC in proximity to point C.	184
20	The outer bounds of Lemma 3.5.9 alongside the "strong interference" outer bound and the BC-based outer bound.	185
21	The different schemes of Section 3.6 for the degraded G-CIFC.	186
22	The achievable schemes of Section 3.6 for the S-G-CIFC.	187
23	The achievable schemes of Section 3.6 for a general G-CIFC.	188
24	The achievable region of schemes (D), (E) and (F) for a general G-CIFC.	189
25	The "primary decodes cognitive" region for different powers $P_1 = P_2 = P$ for a G-CIFC with $a \in \mathbb{R}$ and $(a, \|b\|) \in [-5, 5] \times [0, 5]$.	190
26	The capacity results for the S-G-CIFC for the case $P_1 = P_2 = P$ for $(P, \|b\|) \in [0, 5] \times [0, 5]$.	191

LIST OF ABBREVIATIONS

RV	random variable
BC	broadcast channel
MAC	multiple access channel
IFC	interference channel
D-IFC	deterministic interference channel
DM-IFC	discrete memoryless interference channel
CIFC	cognitive interference channel
D-CIFC	deterministic cognitive interference channel
DM-CIFC	discrete cognitive interference channel
G-CIFC	Gaussian cognitive interference channel
SISO	single input-single output channel
SIMO	single input-multiple output channel
MISO	multiple input-single output channel
MIMO	multiple input-multiple output channel

SUMMARY

This doctoral thesis studies the effect of cognition in wireless networks from an information theoretic perspective. In a communication network, *cognition* is the ability of a device to sense the transmissions of neighboring nodes and adapt its own transmissions to the sensed environment: the concept is intrinsically linked to the wireless environment, where devices can overhear each others' transmissions. Although of crucial interest in the development of future technologies, cognition is still only starting to be studied and understood. In this thesis, we seek to understand cognition from a theoretical perspective, and to do so utilize statistical models of the wireless medium and consider a limited, and tractable, number of devices. We study the effect of cognition by characterizing the transmission rates that may be achieved with the use of cognitive devices. In some scenarios, we determine the optimal transmission strategies, which point at the relevance and improvements of cooperation among devices. This is of particular contemporary relevance, as spectrum availability and interference among devices are the two crucial factors in limiting the transmission rates in modern wireless networks. We show that cognition is able to effectively remove such hindrances by effectively managing and controlling the interference at the cost of an increase in the device's processing complexity.

We focus one a specific models of transmission channel: the *cognitive interference channel*. In the cognitive interference channel two transmitter/receiver pairs exchange messages over a common channel, and thus interfere with one another. Each transmitter wishes to send a message to its intended receiver. One of the transmitters, referred to as the *cognitive user*, in addition to its own message, has knowledge of the message of the other transmitter, the *primary user*. This assumption idealizes the case where one of the transmitters has the ability to sense the communication environment and smartly adapt to it. This channel could also model scenarios where a lower priority user and a higher priority user coexist on the same channel. For example, the low priority user would help the high priority user to transmit its message in exchange for the opportunity to use the channel when the higher priority user vacates it. This scenario arises in practical communication channels, especially wireless networks, when considering devices with the capability of understanding the surrounding environment.

We investigate the capacity regions of such networks – the largest transmission rates that may be achieved with arbitrarily low probability of error. To de so we derive outer bounds to the maximum attainable transmission rates and propose transmission schemes to approach these bounds. When it is possible prove that an outer bound may be achieved with a particular transmission scheme, it is said that *capacity* has been determined. Capacity therefore quantifies the maximum amount of information that can be conveyed reliably through the channel, and acts as a benchmark for the performance evaluation of practical codes and and transmission protocols.

We focus on two classes of channels: a) *deterministic* channels, and b) *additive Gaussian* channels. In a deterministic channels the channel outputs are a deterministic function of the input and no randomness is introduced by the channel. In additive Gaussian channels the channel outputs are obtained as linear combinations of the channel inputs plus an additive noise term. The channel inputs are subject to an average power constraint and the noise term is a random variable with a Gaussian distribution. While additive Gaussian channels constitute a relatively accurate model of a transmission channel, determining capacity of such channels is at times a complicated task. Deterministic channels, on the other hand, are more simplistic, but it is often easier to prove capacity under the deterministic channel assumption. This thesis' contributions are as follows:

- We propose a new outer bound and a new achievable scheme for the cognitive interference channel, improving on the previously known results.
- We derive a set of new capacity results, both in the deterministic and in the additive Gaussian noise channel model.
- We provide important insights on the new capacity results and stress their implications in the design of practical systems.
- When it not possible to prove exact capacity, we prove that our achievable rate regions lie at a bounded distance from our outer bounds regardless of channel conditions, thereby providing performance guarantees. This result can be thought of as an approximate solution to the problem of determining the capacity of a channel and for this reason is termed *approximate capacity.*

The rest of the thesis is structured as follows: Chapter 1 provides an introduction to the study of wireless networks. Chapter 2 presents a cognitive interference channel with discrete inputs and the deterministic cognitive interference channel. In Chapter 3 we study the additive Gaussian cognitive interference channel. Chapter 4 concludes the thesis.

CHAPTER 1

INTRODUCTION

Recent advancements in communication technology, particularly in wireless communications, has increased the availability of smart devices with enhanced communication and computational capabilities. A large number of nowadays communication systems comprise devices that are able to transmit and receive signals over a large spectrum and perform complex computational operations. The demand for high rate data services has also witnessed an exponential increase: it has been reported that the mobile data traffic is doubling every nine months (see (1)).

The frequency spectrum is a limited natural resource and currently the access to this transmission media is regulated by governmental agencies. In the United States the Federal Communications Commission (FCC) has the jurisdiction over radio licensing. The current legislation of the FCC assigns a considerable part of the frequency spectrum to different services in an exclusive fashion. Similar regulations are adopted in most countries. The inefficiency of such a strategy has become apparent when it was realized that the spectrum is under-utilized most of the time (see (2))

A more efficient use of the spectrum has been advocating by many. It has been suggested that *cooperation* among devices would be able to increase spectrum utilization (see (3)). Cooperation can be obtained by having the devices in the network overhear the neighboring transmission. The term *cognitive radio* is used to define this new generation of smart devices.

The term cognition was initially coined by Mitola (4) who described it as:

> The point in which wireless personal digital assistants (PDAs) and the related networks are sufficiently computationally intelligent about radio resources and related computer-to-computer communications to detect user communications needs as a function of use context, and to provide radio resources and wireless services most appropriate to those needs.

Mitola's original definition of cognitive radio is fairly broad and it envisions a fully reconfigurable wireless black-box that automatically changes its communication strategies in response to network topology and user demands.

Different research communities have dealt with specifics aspects of this broad concept.

In the networking community, the research on cognitive radio considers devices that are able to sense utilization of the spectrum at a network level and dynamically schedule their packet transmissions based on the available resources.

The theoretical framework used is often game theoretical where cognition is modeled as the knowledge at a node of the transmission strategies of the surrounding nodes and where the game objective is the successfully transmission of packets (see (5; 6; 7)).

In the field of wireless communication, research focuses mainly on the opportunistic use of the spectrum. Here cognitive devices continuously monitor the frequency spectrum and initiate a transmission in the available frequency-time slots. The starting moment and the duration of a transmission are random and thus a device has to be able to quickly detect both the beginning and the end of the transmission from neighboring user (see (8; 9; 10)).

For computer scientists the issue in cognitive devices is the design of an adaptable physical layer that is able to implement different transmission modes depending an the channel state and interference level at the different points of the spectrum (see (11; 12)).

This concept is an important paradigm also in military communications where cognition allows reliable communication between entities in case of natural disaster (see (13)), when communication links in the network fail randomly.

From an information theoretical prospective, the problem is to determine the capacity region of cognitive channels, i.e. the set of rates that can be simultaneously sustained by all users in the network with arbitrarily low probability of error .

The primary motivations of such study are two folded. Firstly, information theory provides theoretical tools to measure the advantages in spectral efficiency when allowing cooperation between users on a licensed frequency band. Secondly , it provides practical guidelines on which strategies are able to approach the theoretical limits.

The first information theoretic model of a cognitive radio channel is due to Devroye, Mitran and Tarok (14). In they seminal work in 2006, Devroye *et al.* proposed to model cognition as follows. Two pairs of transmitter/receiver share a common communication channel, as in the classical *interference channel* (15). In the classical interference channel, the two pairs are uncoordinated and do not have knowledge of the messages the interfering transmitter/receiver pair exchanges. In a cognitive channel, one of the transmitter (the *cognitive* one) has noncausal knowledge of the message sent by the transmitter (the *primary* one). The primary pair is supposed to be licensed users, while the cognitive pair is assumed to be smart devices who "profits" of the licensed spectrum without harming the legitimate/licensed users. This fairly simple model captures the basic trade-offs involved in cognitive communications: the cognitive user can use the knowledge of the primary user's message to help its own receiver and/or to cooperate with it. In exchange for this the cognitive user is allowed access the channel as long as it does not interfere with the transmission of the primary user.

The the capacity of the classical (non cognitive) interference channel has been an open problem for almost thirty years (see (16)), and much progress has been made in the last couple of years (see (17)). Upper ((18)) and lower (19; 20; 21)

bounds for the (non cognitive) interference channel are available, but they do not coincide in general. Similarly the capacity of the cognitive interference channel has been an open problem since its introduction. Partial results for the capacity region were due to Marić, Goldsmith , Kramer , Shamai and Yates in (22; 23; 24; 25) , to Wu, Vishwanath, and Arapostathis (26), and Jovich and Viswanath (27) . These results cover the special cases where either one or both transmitter do not incur any rate penalty in decoding both messages (see (**?**)), or when it is optimal for the primary receiver to treat the message from the cognitive transmitter as noise (see (26)).

The largest known achievable region is probably the one proposed in (25). However, the region derived in (25) cannot be directly compared with original scheme proposed in (14) because of large number of variable involved in the computation of those achievable regions.

The tightest outer bound was proposed in (25) and included the bounds of (26; 22) as special cases.

Variations of the cognitive radio channel have been proposed. A setting where the cognitive decoder is required to decode both messages but the message of the cognitive user has to be kept secrete at the primary decoder was solved in (28). using an achievable scheme proposed in (29). A setting where both decoders need to decode both messages, known as *compound multiple access channel*, was solved in (24). An extension to the case where receivers and transmitters have multiple antennas was studied in (30). For this MIMO cognitive channel the sum-rate capacity and part of the capacity region was determined.

Unfortunately the combination of multiple transmitting strategy such as binning and superposition coding in one transmission scheme introduces multiple parameters to be optimized. This makes it difficult to compare inner and outer bounds. Probably the most general achievable scheme available up to date is (25, Th. 1) the rate region is expressed by six equations to be optimized over the distribution of six random variables.

In general the task of determining the assignments that could show capacity is a difficult and time consuming. In the last couple of years a new, powerful alternative to this task has arose for Gaussian channels. When the energy of the transmitter is far greater then amplitude of the noise, the channel can be approximated by a deterministic channel, neglecting the effect of the noise.

This consideration suggested an alternative to the task of determining the capacity region of a multi user communication networks. Rather than proving an equality between inner and outer bounds, the authors (31) advocate a powerful new method for obtaining achievable rate regions that lie within a constant number of bits from capacity region outer bounds, thereby determining the capacity region to “within a constant number of bits”. This constant gap is independent in the channel parameters. In a series of papers, and inspired by (32), Avestimehr and Tse introduced the linear deterministic approximation of wireless networks (33; 34; 35).

A linear deterministic network approximates a linear Gaussian network at high SNR by capturing the dynamic range of the desired and interference signals, but neglecting additive noise. At high SNR, this is a valid approximation, and is able to effectively separate the role of interference and noise in a wireless network, allowing one to focus on the *s*ignal interaction. The linear model is often easier to analyze (capacity can often be determined exactly), and insights gained from it may be used to guide coding schemes for inner bounds and receiver side-information in outer bounds for the practically motivated Gaussian noise channel, which are then ideally shown to lie within a constant gap from each other. Proving a constant gap result is in general easier then proving capacity, but yields similar insight on the strategies that perform well in practical applications. The approach has allowed, for example, to solve within a constant gap the capacity regions of channels that have been long standing open problems, such as Gaussian interference channels (36; 17) and Gaussian relay channels (37).

Determining the capacity for a channel is crucial to understand what is the nature of the communication strategy that can efficiently guarantee reliable transmissions.

When capacity is known it is possible measure of the distance between capacity and the rate achieved by a given code. In proving capacity a theoretical code with infinite block-length is constructed and shown reliable. This passage gives an insight on the nature of a practical code that performs well over the channel. When considering channel with only one message set, this constrictive part of the prove guides in designing how the message is handled. When there are more then two messages and interaction between encoders, this actually indicate which forms of collaboration between encoders perform the best. Such a result, therefore, not only is of interest in the code design, but also in the networking part. Unfortunately determining capacity is a arduous task may require several years to be solved. An constant gap result is a very promising alternative to this task that provide just the same insight on policies at physical and network layer alike.

This is particularly important when considering networks with a multiple numbers of receivers and transmitters. Given that the capacity of that two transmitter/receiver pair cognitive channel is still open, it comes to no surprise that very few results are available for more than two pairs.

Some results are available for an deterministic interference channel with three transmitters and receivers in (38) and for particular topologies of the general case (see (39)). also, in (40) the maximum sum rate of a multiple access channel with cognition is determined. For the general multiuser setting some results are available for the relay channel with one transmitter and one receiver and k relays in (35) where a constant gap result is proven where the gap is a function of the number of relays in the network. The sum capacity is known for the Gaussian vector broadcasts channel: this result was proved in (41). Scaling laws for cognitive networks where determined in (42) and (43). In general cognitive network are still an open field of research given the general complexity of the transmission scheme to be employed.

The main research objective for general networks is to understand the scaling of the achievable rates in a cognition network in the number of user. This line of research is studied in (44) and (45) where some preliminary results are derived.

1.1 Notation

We use the following convention:

- The symbol X^N indicates a vector of size N. The symbol X_i^j indicates the subset of the original vector between the index i and the index j.
- The symbol $X \sim \mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$ indicates that the random variable (RV) X is a complex-valued proper Gaussian RV with mean μ and covariance σ^2.
- We define $\mathcal{C}(x) := \log(1 + x)$ for $x \in \mathbb{R}^+$.
- We define $\overline{x} := 1 - x$ for $x \in [0, 1]$.
- For any two RVs X and Y, the symbol $X|_Y$ denotes the conditional distribution of X given Y.
- We use $[1 : n]$ to denote the set of natural numbers from 1 to n.
- The notation $A \overset{(n)}{=} B$ to indicate that the expression B is obtained from A with the assignment given in equation number n.
- For an integer N, the symbol X^N indicates the length-N vector $(X_1, ..., X_N)$.
- For the plots, the logarithms are in base 2, i.e., rates are expressed in bits/s/Hz.
- $\boldsymbol{C}(a, |b|, P_1, P_2)$ indicates the capacity of a Gaussian cognitive interference channel with channel parameters a and $|b|$ and powers P_1 and P_2.
- X^* denotes the complex conjugate of the complex number (or vector) X.

CHAPTER 2

THE DISCRETE MEMORYLESS COGNITIVE INTERFERENCE CHANNEL

The content of this chapter appears in the Proceedings of ITW2009 in Taormina, IZS2010, ICC2011 and is submitted to the IEEE Transaction of Information Theory

2.1 Main contributions

In this chapter we establish a series of new results for the discrete memoryless cognitive interference channel.

1. **W**e derive a new outer bound to the capacity region of the discrete memoryless cognitive interference channel. This outer bound is looser than previously derived outer bounds but it does not include auxiliary random variables and thus it can be evaluated for given channels. The outer bound is derived using an idea originally introduced by Sato for the broadcast channel.

2. **W**e present a new inner bound that encompasses all known achievable regions. This inner bound is shown to include all the previously presented regions and provides new and interesting features.

3. **W**e show capacity in the "better cognitive decoding" regime This regime includes the "very weak interference" and the "very strong interference" regimes and is thus the largest set of channels for which capacity is known.

4. **W**e determine capacity for the semi-deterministic cognitive interference channel. In this particular channel model the output at the cognitive receiver is a deterministic function of the channel inputs. We determine capacity for this channel model by showing the achievability of a the outer bound in (26).

2.2 Organization

The rest of the chapter is organized as follows. Section 2.3 introduces the basic definitions and notation. Section 2.4 summarizes the all the known results for the discrete memoryless cognitive interference channel. The new outer bound is presented in Section 3.5. In Section 3.6 we present the new inner bound. We show the inclusion in the previously presented regions in Section 2.7. We show capacity in the "better cognitive decoding" regime in Section 2.8. Section 2.9 focuses on the semi-deterministic cognitive interference channel. In Section 2.10 we consider the deterministic cognitive interference channel. The chapter concludes with some examples in Section 2.11 which provide insight on the role of cognition.

2.3 Channel model, notation and definitions

A two user InterFerence Channel (IFC) is a multi-terminal network with two senders and two receivers. Each transmitter i wishes to communicate a message W_i to receiver i , $i = \{1, 2\}$. In the classical IFC the two transmitters operate independently having no knowledge of each others messages. Here we consider a variation of this set up assuming that transmitter 1 (also called cognitive transmitter) in addition to its own message, W_1, also knows the message W_2 of transmitter 2 (also called primary transmitter). We refer to transmitter/receiver 1 as the cognitive pair and to transmitter/receiver 2 as the primary pair This model is commonly known as Cognitive InterFerence Channel (CIFC) and it is an idealized model for the unilateral cooperation at the encoder side. A graphical representation of this model can be seen in Figure 6.

A Discrete Memoryless CIFC (DM-CIFC) is a CIFC with finite cardinality input and output alphabets. The channel is memoryless with transition probability $p_{Y_1,Y_2|X_1,X_2}(x_1, x_2)$.

Transmitter $i = \{1, 2\}$ wishes to communicate a message W_i uniformly distributed on $[1, \ldots, 2^{NR_i}]$ to receiver i in N channel uses at rate R_i. The two messages are independent. Transmitter 1 knows both messages and transmitter 2 knows only W_2. A rate pair (R_1, R_2) is achievable if there exists a sequence of encoding functions

$$X_1^N = X_1^N(W_1, W_2)$$
$$X_2^N = X_2^N(W_2),$$

and a sequence of decoding functions

$$\hat{W}_i = \widehat{W}_1(Y_i^N), \quad i = \{1, 2\}$$

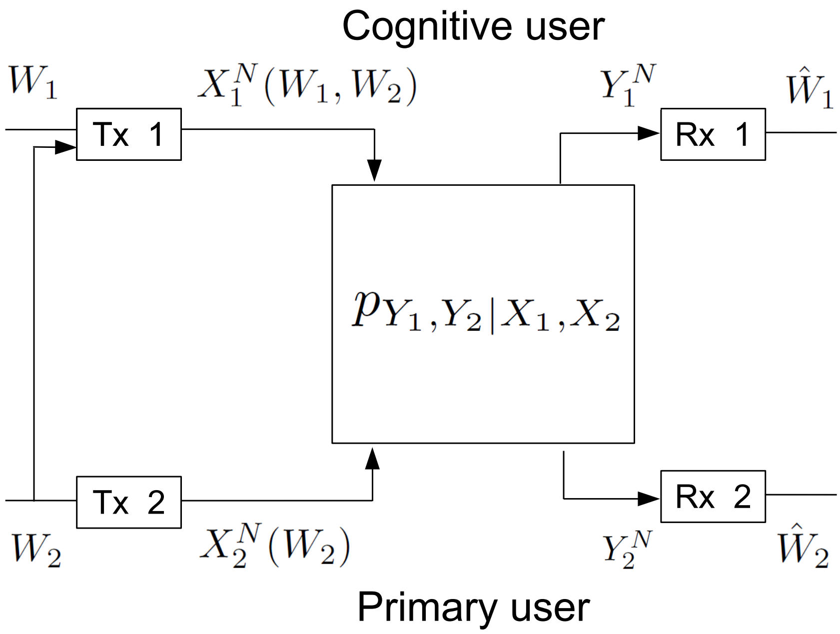

Figure 1. The CIFC model.

such that

$$\lim_{N\to\infty} \max_{i=\{1,2\}} \mathbb{P}\left[\hat{W}_i \neq W_i\right] \to 0.$$

The capacity region is defined as the closure of the region of all achievable (R_1, R_2) pairs (46).

2.4 Available results for the DM-CIFC

We now present the outer bounds and the capacity results available for the DM-CIFC. The first outer bound for the CIFC was obtained in (26, Thm 3.2) by the introduction of an auxiliary Random Variable (RV).

Theorem 2.4.1. ***One auxiliary RV outer bound*** *of (26, Thm 3.2):*

$$R_1 \leq I(X_1;Y_1|X_2) \tag{2.1a}$$

$$R_2 \leq I(X_2,U;Y_2) \tag{2.1b}$$

$$R_1+R_2 \leq I(X_2,U;Y_2)+I(X_1;Y_1|X_2,U), \tag{2.1c}$$

union over the distributions that factors as

$$p_{U,X_1,X_2}p_{Y_1,Y_2|X_1,X_2}.$$

Another general outer bound for the capacity region of the CIFC is provided in (25, Thm 4). This outer bound is derived using an argument originally devised in (47) for the Broadcast Channel (BC). The expression of the outer bound is identical to the outer bound in (47) but for the factorization of the auxiliary RV's differs.

Theorem 2.4.2. ***BC inspired outer bound*** *of (25, Thm. 4):*

$$R_1 \leq I(V,U_1;Y_1) \tag{2.2a}$$

$$R_2 \leq I(V,U_2;Y_2) \tag{2.2b}$$

$$R_1+R_2 \leq I(V,U_1;Y_1)+I(U_2;Y_2|U_1,V) \tag{2.2c}$$

$$R_1+R_2 \leq I(V,U_2;Y_2)+I(U_1;Y_1|U_2,V), \tag{2.2d}$$

union over the distribution that factors as

$$p_{U_1} p_{U_2} p_{V|U_1,U_2} p_{X_2|U_2,V} p_{X_1|U_1,U_2,V} p_{Y_1,Y_2|X_1,X_2}.$$

It is not possible to show in general the containment of the outer bound of Theorem 2.4.1,"one auxiliary RV outer bound", into the region of Theorem 2.4.2, "BC inspired outer bound".

The expression of the outer bound of Theorem 2.4.1,"one auxiliary RV outer bound", can be simplified in two instances called weak and strong interference.

Corollary 2.4.3. ***Weak interference outer bound*** *of (26, Thm 3.4):*

When the condition

$$I(U; Y_2|X_2) \leq I(U; Y_1|X_2) \quad \forall p_{U,X_1,X_2}, \tag{2.3}$$

is verified, the outer bound of Theorem 2.4.1 ,"one auxiliary RV outer bound", can be equivalently expressed as

$$R_1 \leq I(Y_1; X_1|U, X_2) \tag{2.4a}$$

$$R_2 \leq I(U, X_2; Y_2), \tag{2.4b}$$

union over all distributions p_{U,X_1,X_2}.

We refer to the condition in Equation 3.2 as "weak interference condition".

Corollary 2.4.4. ***Strong interference outer bound*** *of (20, Thm 5):*

When the condition

$$I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2) \quad \forall p_{X_1,X_2x}, \tag{2.5}$$

is verified, the outer bound of Theorem 2.4.1 ,"one auxiliary RV outer bound", can be equivalently expressed as

$$R_1 \leq I(Y_1;X_1|X_2) \tag{2.6a}$$

$$R_1 + R_2 \leq I(Y_2;X_1,X_2) \tag{2.6b}$$

union over all distributions p_{X_1,X_2}.

We refer to the condition in Equation 3.4 as "strong interference condition".

The outer bound of Theorem 2.4.1 ,"one auxiliary RV outer bound", can be shown to be achievable in a subset of the "weak interference" Equation 3.2 and of the "strong interference" Equation 3.4 conditions. We refer to such subsets as "very strong interference" and "very weak interference" regimes.

Theorem 2.4.5. ***Very weak interference capacity*** *of (26, Thm. 3.4) and (27, Thm. 4.1).*

The outer bound of Corollary 2.4.3, "weak interference outer bound", is the capacity region if the following holds

$$\begin{aligned} I(U;Y_2|X_2) \leq &\quad I(U;Y_1|X_2) \\ I(X_2;Y_2) \leq &\quad I(X_2;Y_1), \qquad \forall p_{U,X_1,X_2}. \end{aligned} \tag{2.7}$$

We refer to the condition in Equation 2.7 as "very weak interference". In this regime capacity is achieved by having encoder 2 transmitting as in a point-to-point channel and encoder 1 performing Gelf'and-Pinsker binning against the interference created by transmitter 2.

Theorem 2.4.6. ***Very strong interference capacity*** *of (20, Thm. 5). The outer bound of Corollary 3.4.2, "strong interference outer bound", is the capacity region if the following holds*

$$\begin{aligned} I(X_1;Y_1|X_2) &\leq I(X_1;Y_2|X_2) \\ I(Y_2;X_1,X_2) &\leq I(Y_1;X_1,X_2), \quad \forall p_{X_1,X_2}. \end{aligned} \tag{2.8}$$

We refer to the condition in Equation 3.6 as "very strong interference". In this regime, capacity is achieved by having both receivers decoding both messages.

The outer bounds presented in Theorem 2.4.1, "one auxiliary RV outer bound" and 2.4.2 , "BC inspired outer bound", cannot be evaluated in general since they include auxiliary RV's whose cardinality has yet not being bounded. In the following we propose a now outer bound, looser in general that the outer bound of Theorem 2.4.1 without auxiliary RV's. This bound is looser than the outer bound of Theorem 2.4.1,"one auxiliary RV outer bound", in the general case, but it is tight is the "very strong interference" regime.

2.5 A new outer bound

Theorem 2.5.1. *The capacity region of the CIFC is contained into the set*

$$R_1 \leq I(Y_1; X_1|X_2), \tag{2.9a}$$

$$R_2 \leq I(X_1, X_2; Y_2), \tag{2.9b}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y_2) + I(Y_1; X_1|Y_2', X_2), \tag{2.9c}$$

union over all distributions p_{X_1,X_2} *where* Y_2' *has the same marginal distribution of* Y_2*, i.e.,* $p_{Y_2'|X_1,X_2} = p_{Y_2|X_1,X_2}$*, but otherwise any joint distribution* $p_{Y_1,Y_2'|X_1,X_2}$*.*

The idea behind this outer bound is to exploit the fact that the capacity region only depends on the marginal distributions $P_{Y_1|X_1,X_2}$ and $P_{Y_2|X_1,X_2}$ because the receivers do not cooperate.

Proof. By Fano's inequality we have that $H(W_i|Y_i^N) \leq N\epsilon_N$, for some ϵ_N such that $\epsilon_N \to 0$ as $N \to 0$ for $i \in \{1,2\}$. The rate of user 1 can be bounded as

$$\begin{aligned}
N(R_1 - \epsilon_N) &\leq I(W_1; Y_1^N) \\
&\leq I(W_1; Y_1^N|W_2) \\
&= I(W_1, X_1^N(W_1, W_2); Y_1^N|W_2, X_2^N(W_2)) \\
&\leq H(Y_1^N|W_2, X_2^N) - H(Y_1^N|W_2, W_1, X_1^N, X_2^N) \\
&\leq H(Y_1^N|X_2^N) - H(Y_1^N|W_2, W_1, X_1^N, X_2^N) \\
&= H(Y_1^N|X_2^N) - H(Y_1^N|X_1^N, X_2^N) \\
&= \sum_{i=1}^{N} H(Y_{1i}|X_2^N, (Y_1)_1^{i-1}) - H(Y_{1i}|X_2^N, X_2^N, (Y_1)_1^{i-1}) \\
&\leq \sum_{i=1}^{N} H(Y_{1i}|X_{2i}) - H(Y_1^N|X_{1i}, X_{2i}) \\
&= NI(Y_{1T}; X_{1T}|X_{2T}, T) \\
&= N(H(Y_{1T}|X_{2T}, T) - H(Y_{1T}|X_{1T}, X_{2T}, T)) \\
&= N(H(Y_{1T}|X_{2T}, T) - H(Y_{1T}|X_{1T}, X_{2T})) \\
&\leq N(H(Y_{1T}|X_{2T}) - H(Y_{1T}|X_{1T}, X_{2T})) \\
&\leq I(Y_{1T}; X_{1T}|X_{2T}),
\end{aligned} \tag{2.10a}$$

where T is the time sharing RV, informally distributed over the set $\{1...N\}$.

The rate of user 2 can be bounded as

$$
\begin{aligned}
N(R_2-\epsilon_N) &\le I(Y_2^N;W_2)\\
&\le I(Y_2^N;W_2,W_1)\\
&= H(Y_2^N)-H(Y_2^N|W_1,W_2,X_2^N(W_2),X_1^N(W_1,W_2))\\
&= H(Y_2^N)-H(Y_2^N|X_2^N,X_1^N)\\
&= \sum_{i=1}^{N} H(Y_{2i}|(Y_2)_1^{i-1})-H(Y_{2i}|X_1^N,X_2^N,(Y_2)_1^{i-1})\\
&\le \sum_{i=1}^{N} H(Y_{2i})-H(Y_{2i}|X_{1i},X_{2i})\\
&\le NI(Y_{2T};X_{1T},X_{2T}|T)\\
&= N(H(Y_{2T}|T)-H(Y_{2T}|X_{1T},X_{2T},T))\\
&\le N(H(Y_{2T})-H(Y_{2T}|X_{1T},X_{2T}))\\
&\le I(Y_{2T};X_{1T},X_{2T}). \qquad (2.10b)
\end{aligned}
$$

Next let Y_2' be any RV such that $P_{Y_2'|X_1,X_2} = P_{Y_2|X_1,X_2}$ but with any joint distribution $P_{Y_1,Y_2'|X_1,X_2}$. The sum-rate can then be bounded as

$$
\begin{aligned}
N(R_1+R_2-2N\epsilon_N) &\le I(W_1;Y_1) + I(W_2;Y_2) \\
&\le I(W_1;Y_1^N|W_2) + I(W_2;Y_2^N) \\
&\le I(W_1;Y_1^N,Y_2'^N|W_2) + I(W_2;Y_2^N) \\
&= I(W_2;Y_2^N) + I(W_1;Y_2'^N|W_2) + I(W_1;Y_1^N|Y_2'^N,W_2) \\
&= H(Y_2^N) + \Big(- H(Y_2^N|W_2) + H(Y_2'^N|W_2)\Big) \\
&\quad - H(Y_2'^N|W_1,W_2) + H(Y_1^N|Y_2'^N,W_2) - H(Y_1^N|Y_2'^N,W_1,W_2) \\
&= H(Y_2^N) + H(Y_1^N|W_2,X_2^N,Y_2'^N) \\
&\quad - H(Y_2'^N|W_1,W_2,X_1^N,X_2^N) - H(Y_1^N|Y_2'^N,W_1,W_2,X_1^N,X_2^N) \\
&= H(Y_2^N) + H(Y_1^N|W_2,X_2^N,Y_2'^N) \\
&\quad - H(Y_2^N|X_1^N,X_2^N) - H(Y_1^N|Y_2'^N,X_1^N,X_2^N) \\
&\le H(Y_2^N) + H(Y_1^N|X_2^N,Y_2'^N) \\
&\quad - H(Y_2^N|X_1^N,X_2^N) - H(Y_1^N|Y_2'^N,X_1^N,X_2^N) \\
&\le I(Y_2^N;X_1^N,X_2^N) + \sum_{i=1}^{N} H(Y_{1i}|X_2^N,Y_2'^N,(Y_1)_1^{i-1}) - H(Y_{1i}|X_1^N,X_2^N,Y_2'^N,(Y_1)_1^{i-1}) \\
&\le I(Y_2^N;X_1^N,X_2^N) + \sum_{i=1}^{N} H(Y_{1i}|X_{2i},Y_{2i}') - H(Y_{1i}|X_{1i},X_{2i},Y_{2i}') \\
&\le I(Y_2^N;X_1^N,X_2^N) + \sum_{i=1}^{N} H(Y_{1i}|X_{2i},Y_{2i}') - H(Y_{1i}|X_{1i},X_{2i},Y_{2i}') \\
&= N\left(I(Y_{2T};X_{1T},X_{2T}) + H(Y_{1T}|X_{2T},Y_{2T}',T) - H(Y_{1T}|X_{1T},X_{2T},Y_{2T}')\right) \\
&\le N\left(I(Y_{2T};X_{1T},X_{2T}) + I(Y_{1T};X_{1T}|X_{2T},Y_{2T}')\right). \qquad (2.10c)
\end{aligned}
$$

$\square$

Remark 2.5.2. The outer bound of Theorem 2.5.1 contains the outer bound of Theorem 2.4.1,"one auxiliary RV outer bound",. Indeed, for a fixed distribution p_{X_1,X_2} $Equation\ 2.1a = Equation\ 2.9a$ and $Equation\ 2.1b \leq Equation\ 2.9b$ since

$$\begin{aligned} Equation\ 2.1b &= I(Y_2; X_2, U) \\ &\overset{(a)}{\leq} I(Y_2; X_2, U) + I(Y_2; X_1|U, X_2) \\ &= I(Y_2; X_1, X_2, U) \\ &= I(Y_2; X_1, X_2) = Equation\ 2.9b, \end{aligned}$$

where the last equality follows from the Markov chain $U - X_1, X_2 - Y_1, Y_2$.

Consider Y_2' such that $p_{Y_2'|U,X_1,X_2} = p_{Y_2|U,X_1,X_2}$, which also implies $p_{Y_2'|U,X_2} = p_{Y_2|U,X_2}$ since

$$\begin{aligned} p_{Y_2'|U,X_2} &= \tfrac{1}{p_{X_1}} \int_{|Y_2'|} p_{Y_2'|U,X_1,X_2} p_{U,X_1,X_2} dX_1 \\ &= \tfrac{1}{p_{X_1}} \int_{|Y_2|} p_{Y_2|U,X_1,X_2} p_{U,X_1,X_2} dX_1 \\ &= p_{Y_2|U,X_2}, \end{aligned}$$

then:

$$
\begin{aligned}
\textit{Equation } 2.1c \quad &= I(Y_2; X_2, U) + I(X_1; Y_1|U, X_2) \\
&= H(Y_2) + H(Y_2|X_1, X_2, U) - H(Y_2|U, X_1, X_2) - H(Y_2|U, X_2) + I(X_1; Y_1|U, X_2) \\
&= I(Y_2; X_1, X_2, U) + H(Y_2'|U, X_1, X_2) - H(Y_2'|U, X_2) + I(X_1; Y_1|U, X_2) \\
&\leq I(Y_2; X_1, X_2) - I(Y_2'; X_1|U, X_2) + I(X_1; Y_1|U, X_2) + I(Y_2'; Y_1|U, X_1, X_2) \\
&= I(Y_2; X_1, X_2) - I(Y_2'; X_1|U, X_2) + I(Y_2', X_1; Y_1|U, X_2) \\
&= I(Y_2; X_1, X_2) + I(Y_1; X_1|Y_2', U, X_2) \\
&= I(Y_2; X_1, X_2) + H(Y_1|Y_2', U, X_2) - H(Y_1|Y_2', U, X_1, X_2) \\
&\overset{(b)}{\leq} I(Y_2; X_1, X_2) + H(Y_1|Y_2', X_2) - H(Y_1|Y_2', X_1, X_2) \\
&= I(Y_2; X_1, X_2) + I(Y_1; X_1|Y_2', X_2) = \textit{Equation } 2.9c.
\end{aligned}
$$

Now the RV U does not appear in the outer bound expression Equation 2.9c and thus we can consider simply the RV's with $p_{\widetilde{Y}_2|X_1,X_2} = p_{Y_2|X_1,X_2}$ which corresponds to the definition of Y_2' in Theorem Equation 2.9.

Equality of the outer bounds is verified when conditions (a) and (b) hold: with equality, that is when

$$
\begin{aligned}
&I(Y_2; X_1|U, X_2) = 0 \\
&I(Y_1; X_1|\widetilde{Y}_2, U, X_2) = I(Y_1; X_1|\widetilde{Y}_2, X_2) \quad \forall p_U,
\end{aligned}
$$

and a given $\widetilde{Y}_2$. The first conditions implies the MC

$$Y_2 - U, X_2 - X_1$$

and the second condition the MC

$$Y_1, X_1 - \widetilde{Y}_2 X_2 - U$$

We currently cannot relate these conditions to any specific class of DM-CIFC or provide an intuition on their role in the outer bound expression.

Remark 2.5.3. The outer bound of Theorem 2.5.1 reduces to the strong interference outer bound in Equation 3.5, in fact

$$I(Y_1; X_1|X_2) \leq I(Y_2; X_1|X_2) \qquad \forall p_{X_1,X_2}$$

implies

$$I(Y_1; X_1|Y_2', X_2) \leq I(Y_2; X_1|Y_2', X_2) \qquad \forall p_{X_1,X_2,Y_2'}$$

Now let $Y_2' = Y_2$ to obtain that $I(Y_1; X_1|Y_2, X_2) = 0$ yielding *Equation* 2.9*c* = *Equation* 2.9*b* so that the two outer bounds coincide.

2.6 A new inner bound

As the DM-CIFC encompasses classical interference, multiple-access and broadcast channels, we expect to see a combination of their achievability proving techniques surface in any unified scheme for the CIFC:

• **R**ate-splitting. As in Han and Kobayashi (18) for the interference-channel and in the DM-CIFC regions of (25; 14; 29), rate-splitting is not necessary in the very weak (26) and very strong (22) interference regimes of Equation 2.7 and Equation 3.6.

• **S**uperposition-coding. Useful in multiple-access and broadcast channels (46), in the CIFC the superposition of private messages on top of common ones proposed in (25; 29) and is known to be capacity achieving in very strong interference (22).

• **B**inning. Gel'fand-Pinsker coding (48), often simply referred to as binning, allows a transmitter to "cancel" (portions of) the interference known to be experienced at a receiver. Binning is also used by Marton in deriving the largest known channel achievable rate region (49) for the broadcast channel.

We now present a new achievable region for the DM-CIFC which generalizes all the known achievable rate regions presented in (25; 26; 29; 14; 50) and (51).

Theorem 2.6.1. ***The*** $\mathcal{R}_{RTD}$ ***region*** *Region* $\mathcal{R}_{RTD}$*. A rate pair* (R_1, R_2) *such that*

$$\begin{aligned} R_1 &= R_{1c} + R_{1pb}, \\ R_2 &= R_{2c} + R_{2pa} + R_{2pb}. \end{aligned} \tag{2.11}$$

is achievable for a general DM-CIFC if $(R'_{1c}, R'_{1pb}, R'_{2pb}, R_{1c}, R_{1pb}, R_{2c}, R_{2pa}, R_{2pb}) \in \mathbb{R}^8_+$ *satisfies:*

$$
\begin{aligned}
R'_{1c} &= I(U_{1c}; X_2|U_{2c}) && \text{(2.12a)}\\
R'_{1c} + R'_{1pb} &\geq I(U_{1pb}; X_2|U_{1c}, U_{2c}) \\
&\quad +I(U_{1c}; X_2|U_{2c}) && \text{(2.12b)}\\
R'_{1c} + R'_{1pb} + R'_{2pb} &\geq I(U_{1pb}; X_2, U_{2pb}|U_{1c}, U_{2c}) \\
&\quad +I(U_{1c}; X_2|U_{2c}) && \text{(2.12c)}\\
R_{2c} + R_{2pa} + (R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}, X_2, U_{2c}) \\
&\quad +I(U_{1c}; X_2|U_{2c}) && \text{(2.12d)}\\
R_{2pa} + (R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}, X_2|U_{2c}) \\
&\quad +I(U_{1c}; X_2|U_{2c}) && \text{(2.12e)}\\
R_{2pa} + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, X_2|U_{1c}, U_{2c}) \\
&\quad +I(U_{1c}; X_2|U_{2c}) && \text{(2.12f)}\\
(R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}|X_2, U_{2c}) \\
&\quad +I(U_{1c}; X_2|U_{2c}) && \text{(2.12g)}\\
(R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}|U_{1c}, X_2, U_{2c}) && \text{(2.12h)}\\
R_{2c} + (R_{1c} + R'_{1c}) + (R_{1pb} + R'_{1pb}) &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}), && \text{(2.12i)}\\
(R_{1c} + R'_{1c}) + (R_{1pb} + R'_{1pb}) &\leq I(Y_1; U_{1pb}, U_{1c}|U_{2c}), && \text{(2.12j)}\\
(R_{1pb} + R'_{1pb}) &\leq I(Y_1; U_{1pb}|U_{1c}, U_{2c}), && \text{(2.12k)}
\end{aligned}
$$

for some input distribution

$$p_{Y_1,Y_2,X_1,X_2,U_{1c},U_{2c},U_{2pa},U_{1pb},U_{2pb}} = p_{U_{1c},U_{2c},U_{2pa},U_{1pb},U_{2pb},X_1,X_2} p_{Y_1,Y_2|X_1,X_2}.$$

Remark 2.6.2. Moreover:

- Equation 3.36d can be dropped when $R_{2c} = R_{2pa} = R_{2pb} = R'_{2pb} = 0$;
- Equation 3.36e can be dropped when $R_{2pa} = R_{2pb} = R'_{2pb} = 0$;
- Equation 3.36g can be dropped when $R_{2pb} = R'_{2pb} = 0$;
- Equation 3.36i can be dropped when $R_{1c} = R'_{1c} = R_{1pb} = R'_{1pb} = 0$,

since they correspond to the event that a common message from the non intended user is incorrectly decoded. This event is not an error event if no other intended message is incorrectly decoded.

Proof. The meaning of the RV's in Theorem 2.6.1 is as follows. Both transmitters perform superposition of two codewords: a common one (to be decoded at both decoders) and a private one (to be decoded at the intended decoder only). In particular:

- Rate R_1 is split into R_{1c} and R_{1pb} and conveyed through the RV's U_{1c} and U_{1pb}, respectively.
- Rate R_2 is split into R_{2c}, R_{2pa} and R_{2pb} and conveyed through the RV's U_{2c}, X_2 and U_{2pb}, respectively.
- U_{2c} is the common message of transmitter 2. The subscript "c" stands for "common".

- X_2 is the private message of transmitter 2 to be sent by transmitter 2 only. It superimposed to U_{2c}. The subscript "p" stands for "private" and the subscript "a" stands for "alone".
- U_{1c} is the common message of transmitter 1. It is superimposed to U_{2c} and–conditioned on U_{2c}–is binned against X_2.
- U_{1pb} and U_{2pb} are the private messages of transmitter 1 and transmitter 2, respectively, and are sent by transmitter 1 only. They are binned against one another conditioned on U_{2c}, as in Marton's achievable region for broadcast channels (49). The subscript "b" stands for "broadcast".
- X_1 is finally superimposed to all the previous RV's and transmitted over the channel.

A graphical representation of the encoding scheme of Theorem 2.6.1 can be found in Figure 2.

The formal description of the proposed encoding scheme is as follows:

2.6.1 Rate splitting

Let W_1 and W_2 be two independent RV's uniformly distributed on $[1...2^{NR_1}]$ and $[1...2^{NR_2}]$ respectively. Consider splitting the messages as follows:

$$W_1 = (W_{1c}, W_{1pb}),$$

$$W_2 = (W_{2c}, W_{2pb}, W_{2pa}),$$

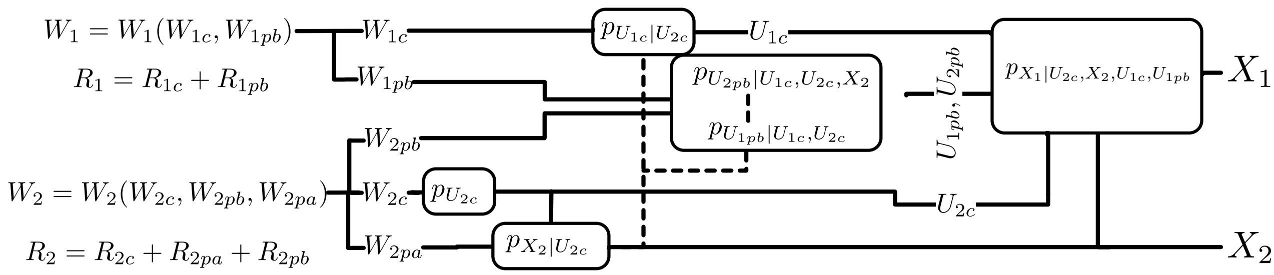

Figure 2. The achievable encoding scheme of Thm. 2.6.1. The ordering from left to right and the distributions demonstrate the codebook generation process. The dotted lines indicate binning. We see rate splits are used at both users, private messages $W_{1pb}, W_{2pa}, W_{2pb}$ are superimposed on common messages W_{1c}, W_{2c} and U_{1c}, is binned against (U_{2pa}, X_2) conditioned on U_{2c}, while U_{1pb} and U_{2pb} are binned against each and X_2 other in a Marton-like fashion (conditioned on other subsets of RV's). U_{1pb} is binned against U_{2pa} as well.

where the messages W_i, $i \in \{1c, 2c, 1pb, 2pb, 2pa\}$, are all independent and uniformly distributed on $[1...2^{NR_i}]$, so that the rate are

$$R_1 = R_{1c} + R_{1pb},$$

$$R_2 = R_{2c} + R_{2pa} + R_{2pb}.$$

as specified in Equation 3.35b.

2.6.2 Codebook generation

Consider a distribution $p_{U_{1c},U_{2c},X_2,U_{1pb},U_{1pb},X_1,X_2}$. The codebooks are generated as follows:

- Select uniformly at random $2^{NR_{2c}}$ length-N sequences $U_{2c}^N(w_{2c})$, $w_{2c} \in [1...2^{NR_{2c}}]$, from the typical set $T_\epsilon^N(p_{U_{2c}})$.

- for every $w_{2c} \in [1...2^{NR_{2c}}]$, select uniformly at random $2^{NR_{2pa}}$ length-N sequences $X_2^N(w_{2c}, w_{2pa})$, $w_{2pa} \in [1...2^{NR_{2pa}}]$, from the typical set $T_\epsilon^N(p_{X_2,U_{2c}}|U_{2c}^N(w_{2c}))$.

- for every $w_{2c} \in [1...2^{NR_{2c}}]$, select uniformly at random $2^{N(R_{1c}+R'_{1c})}$ length-N sequences $U_{1c}^N(w_{2c}, w_{1c}, b_0)$, $w_{1c} \in [1...2^{NR_{1c}}]$ and $b_0 \in [1...2^{NR'_{1c}}]$, from the typical set $T_\epsilon^N(p_{U_{1c}U_{2c}}|U_{2c}^N(w_{2c}))$

- for every $w_{2c} \in [1...2^{NR_{2c}}]$, $w_{2pa} \in [1...2^{NR_{2pa}}]$, $w_{1c} \in [1...2^{NR_{1c}}]$ and $b_0 \in [1...2^{NR'_{1c}}]$, select uniformly at random $2^{N(R_{2pb}+R'_{2pb})}$ length-N sequences $U_{2pb}^N(w_{2c}, w_{2pa}, w_{1c}, b_0, w_{2pb}, b_2)$, $w_{2pb} \in [1...2^{NR_{2pb}}]$ and $b_2 \in [1...2^{NR'_{2pb}}]$, from the typical set

  $T_\epsilon^N(p_{U_{2pb},U_{2c},U_{1c},X_2}|U_{2c}^N(w_{2c}), X_2^N(w_{2c}, w_{2pa}), U_{1c}^N(w_{2c}, w_{1c}, b_0))$.

- for every $w_{2c} \in [1...2^{NR_{2c}}]$, $w_{1c} \in [1...2^{NR_{1c}}]$ and $b_0 \in [1...2^{NR'_{1c}}]$, select uniformly at random $2^{N(R_{1pb}+R'_{1pb})}$ length-N sequences $U_{1pb}^N(w_{2c}, w_{1c}, b_0, w_{1pb}, b_1)$, $w_{1pb} \in [1...2^{NR_{1pb}}]$ and $b_1 \in [1...2^{NR'_{1pb}}]$, from the typical set $T_\epsilon^N(p_{U_{1pb},U_{2c},U_{1c}}|U_{2c}^N(w_{2c}), U_{1c}^N(w_{2c}, w_{1c}, b_0))$.

- for every $w_{2c} \in [1...2^{NR_{2c}}]$, $w_{2pa} \in [1...2^{NR_{2pa}}]$, $w_{1c} \in [1...2^{NR_{1c}}]$, $b_0 \in [1 : 2^{NR'_{1c}}]$, $w_{1pb} \in [1...2^{NR_{1pb}}]$, $b_1 \in [1 : 2^{NR'_{1pb}}]$, $w_{2pb} \in [1...2^{NR_{2pb}}]$, $b_2 \in [1 : 2^{NR'_{2pb}}]$, let the channel input $X_1^N(w_{2pa}, w_{2c}, w_{1c}, b_0, w_{1pb}, b_1, w_{2pb}, b_2)$ be any length-N sequence from the typical set

$$T_\epsilon^N(p_{X_1,U_{2c},U_{1c},X_2,U_{2pb},U_{1pb}}|U_{2c}^N(w_{2c}), X_2^N(w_{2c}, w_{2pa}), U_{1c}^N(w_{2c}, w_{1c}, b_0),$$
$$U_{2pb}^N(w_{2c}, w_{2pa}, w_{1c}, b_0, w_{2pb}, b_2), U_{1pb}^N(w_{2c}, w_{1c}, b_0, w_{1pb}, b_1)).$$

2.6.3 Encoding

Given the message $w_2 = (w_{2c}, w_{2pb}, w_{2pa})$, encoder 2 sends the codeword $X_2^N(w_{2c}, w_{2pa})$.

Given the message $w_2 = (w_{2c}, w_{2pb}, w_{2pa})$ and the message $w_1 = (w_{1c}, w_{1pb})$, encoder 1 looks for a triplet (b_0, b_1, b_2) such that:

$$(U_{2c}^N(w_{2c}), X_2^N(w_{2c}, w_{2pa}), U_{1c}^N(w_{2c}, w_{1c}, b_0), U_{1pb}^N(w_{2c}, w_{1c}, b_0, w_{1pb}, b_1), U_{2pb}^N(w_{2c}, w_{1c}, b_0, w_{2pb}, b_2))$$
$$\in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}}).$$

If not such a triplet exists, it sets $(b_0, b_1, b_2) = (1, 1, 1)$. If more than one such a triplet exists, it picks one uniformly at random from the found ones. For the selected (b_0, b_1, b_2), encoder 1 sends $X_1^N(w_{2pa}, w_{2c}, w_{1c}, b_0, w_{1pb}, b_1, w_{2pb}, b_2)$.

Since the codebooks are generated iid according to

$$p^{(codebook)} = p_{U_{2c}}\, p_{X_2|U_{2c}}\, p_{U_{1c}|U_{2c}}\, p_{U_{2pb}|U_{2c},U_{1c},X_2}\, p_{U_{1pb}|U_{2c},U_{1c}} \tag{2.13}$$

but the encoding forces the actual transmitted codewords to look as if they were generated iid according to

$$p^{(encoding)} = p_{U_{2c}}\, p_{X_2|U_{2c}}\, p_{U_{1c}|U_{2c},X_2}\, p_{U_{2pb}|U_{2c},U_{1c},X_2}\, p_{U_{1pb}|U_{2c},U_{1c},X_2,U_{2pb}}, \tag{2.14}$$

We expect the probability of encoding error to depends on

$$E\left[\frac{p^{(encoding)}}{p^{(codebook)}}\right] = E\left[\frac{p_{U_{1c}|U_{2c},X_2}\, p_{U_{1pb}|U_{2c},U_{1c},X_2,U_{2pb}}}{p_{U_{1c}|U_{2c}}\, p_{U_{1pb}|U_{2c},U_{1c}}}\right] = I(U_{1c}; X_2|U_{2c}) + I(U_{1pb}; X_2, U_{2pb}|U_{2c}, U_{1c}).$$

2.6.4 Decoding

Decoder 2 looks for a unique tuple $(w_{2c}, w_{2pa}, w_{2pb})$ and some (w_{1c}, b_0, b_2) such that

$$(U_{2c}^n(w_{2c}), X_2^n(w_{2c}, w_{2pa}), U_{1c}^n(w_{2c}, w_{1c}, b_0), U_{2pb}^n(w_{2c}, w_{1c}, b_0, w_{2pb}, b_2), Y_2^n) \in T_\epsilon^n(p_{U_{2c},X_2,U_{1c},U_{2pb},Y_2}).$$

Depending on which messages are wrongly decoded at decoder 2, the transmitted sequences and the received Y_2^n are generated iid according to

$$p_{2|\star} = p_{U_{2c}}\, p_{X_2|U_{2c}}\, p_{U_{1c}|U_{2c}}\, p_{U_{2pb}|U_{2c},U_{1c},X_2}\, p_{Y_2|\star}, \tag{2.15}$$

where "$\star$" indicates the messages decoded correctly. However, the actual transmitted sequences and the received Y_2^n considered at decoder 2 look as if they were generated iid according to

$$p_2 = p_{U_{2c}}\, p_{X_2|U_{2c}}\, p_{U_{1c}|U_{2c},X_2}\, p_{U_{2pb}|U_{2c},U_{1c},X_2}\, p_{Y_2|U_{2c},U_{1c},X_2,U_{2pb}}. \tag{2.16}$$

Hence we expect the probability of error at decoder 2 to depend on terms of the type

$$I_{2|\star} = E\left[\log \frac{p_2}{p_{2|\star}}\right] = E\left[\log \frac{p_{U_{1c}|U_{2c},X_2}\, p_{Y_2|U_{2c},U_{1c},X_2,U_{2pb}}}{p_{U_{1c}|U_{2c}}\, p_{Y_2|\star}}\right] = I(U_{1c}; X_2|U_{2c}) + I(Y_2; U_{2c}, U_{1c}, X_2, U_{2pb}|\star). \tag{2.17}$$

Decoder 1 looks for a unique pair (w_{1c}, w_{1pb}) and some (w_{2c}, b_0, b_1) such that

$$(U_{2c}^n(w_{2c}), U_{1c}^n(w_{2c}, w_{1c}, b_0), U_{1pb}^n(w_{2c}, w_{1c}, b_0, w_{1pb}, b_1), Y_1^n) \in T_\epsilon^n(p_{U_{2c},U_{1c},U_{1pb},Y_1}).$$

Depending on which messages are wrongly decoded at decoder 1, the transmitted sequences and the received Y_1^n are generated iid according to

$$p_{1|\star} = p_{U_{2c}}\, p_{U_{1c}|U_{2c}}\, p_{U_{1pb}|U_{2c},U_{1c}}\, p_{Y_1|\star}, \tag{2.18}$$

where "$\star$" indicates the messages decoded correctly. However, the actual transmitted sequences and the received Y_1^n considered at decoder 1 look as if they were generated iid according to

$$p_1 = p_{U_{2c}}\, p_{U_{1c}|U_{2c}}\, p_{U_{1pb}|U_{2c},U_{1c}}\, p_{Y_1|U_{2c},U_{1c},U_{1pb}}. \tag{2.19}$$

Hence we expect the probability of error at decoder 1 to depend on terms of the type

$$I_{1|\star} = E\left[\log \frac{p_1}{p_{1|\star}}\right] = E\left[\log \frac{p_{Y_1|U_{2c},U_{1c},U_{1pb}}}{p_{Y_1|\star}}\right] = I(Y_1; U_{2c}, U_{1c}, U_{1pb}|\star). \tag{2.20}$$

2.6.5 Error analysis

Without loss of generality assume that the message $(w_{1c}, w_{2c}, w_{2pa}, w_{1pb}, w_{2pb}) = (1,1,1,1,1)$ was sent and let $(\bar{b}_0, \bar{b}_1, \bar{b}_2)$ be the tuple (b_0, b_1, b_2) chosen at encoder 1. Let $(\widehat{w}_{1c}, \widehat{w}_{2c}, \widehat{w}_{2pa}, \widehat{w}_{2pb}, \hat{b}_0, \hat{b}_2)$ be the estimate at the decoder 2 and $(\widehat{\widehat{w}}_{1c}, \widehat{\widehat{w}}_{2c}, \widehat{\widehat{w}}_{1pb}, \hat{\hat{b}}_0, \hat{\hat{b}}_1)$ be the estimate at the decoder 1.

The probability of error at decoder u, $u \in \{1,2\}$, is bounded by

$$P[error\ u] \le P[error\ u|encoding\, successful] + P[encoding\, NOT\, successful].$$

An encoding error occurs if encoder 1 is not able to find a tuple $(\bar{b}_0, \bar{b}_1, \bar{b}_2)$ that guarantees typicality. A decoding error is committed at decoder 1 when $(\widehat{\widehat{w}}_{1c}, \widehat{\widehat{w}}_{1pb}) \neq (1,1)$. A decoding error is committed at decoder 2 when $(\widehat{w}_{2c}, \widehat{w}_{2pa}, \widehat{w}_{2pb}) \neq (1,1,1)$.

2.6.6 Encoding Error

The probability that the encoding fails can be bounded as:

$$P[encoding\, NOT\, successful] = P\Big[\bigcap_{b_0=1}^{2^{NR'_{1c}}} \bigcap_{b_1=1}^{2^{NR'_{1pb}}} \bigcap_{b_2=1}^{2^{NR'_{2pb}}}$$
$$\Big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2)\Big) \notin T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})\Big]$$
$$= P[K=0] \leq \frac{Var[K]}{E^2[K]}$$

where

$$K = \sum_{b_0=1}^{2^{NR'_{1c}}} \sum_{b_1=1}^{2^{NR'_{1pb}}} \sum_{b_2=1}^{2^{NR'_{2pb}}} K_{b_0,b_1,b_2}$$

and

$$K_{b_0,b_1,b_2} = 1_{\left\{\left(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2)\right) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})\right\}},$$

where $1_{\{x \in A\}} = 1$ if $x \in A$ and zero otherwise.

The mean value of K (neglecting all terms that depend on ϵ and that eventually go to zero) is:

$$E[K] = \sum_{b_0=1}^{2^{NR'_{1c}}} \sum_{b_1=1}^{2^{NR'_{1pb}}} \sum_{b_2=1}^{2^{NR'_{2pb}}} P[K_{b_0,b_1,b_2} = 1] = 2^{N(R'_{1c}+R'_{1pb}+R'_{2pb}-A)}$$

with

$$
\begin{aligned}
2^{-NA} &= P[K_{b_0,b_1,b_2} = 1] = E[K_{b_0,b_1,b_2}] \\
&= P[\big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2)\big) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})] \\
&= \sum_{(u_{1c}^N, u_{1pb}^N, u_{2pb}^N) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}} | u_{2c}^N, x_2^N)} p_{U_{1c}|U_{2c}}\, p_{U_{2pb}|U_{2c},U_{1c},X_2}\, p_{U_{1pb}|U_{2c},U_{1c}} \\
&\geq 2^{-N[I(U_{1c};X_2|U_{2c}) + I(U_{1pb};X_2,U_{2pb}|U_{1c},U_{2c})]}.
\end{aligned}
$$

The variance of K (neglecting all terms that depend on ϵ and that eventually go to zero) is:

$$
\begin{aligned}
&\mathrm{Var}[K] \\
&= \sum_{b_0=1}^{2^{NR'_{1c}}} \sum_{b_1=1}^{2^{NR'_{1pb}}} \sum_{b_2=1}^{2^{NR'_{2pb}}} \sum_{b'_0=1}^{2^{NR'_{1c}}} \sum_{b'_1=1}^{2^{NR'_{1pb}}} \sum_{b'_2=1}^{2^{NR'_{2pb}}} \Big(P[K_{b_0,b_1,b_2} = 1, K_{b'_0,b'_1,b'_2} = 1] - P[K_{b_0,b_1,b_2} = 1] P[K_{b'_0,b'_1,b`_2} = 1]\Big) \\
&= \sum_{b_0` = b_0, (b_1,b_2,b_1`,b_2`)} (P[K_{b_0,b_1,b_2} = 1, K_{b_0,b_1`,b_2`} = 1] - P[K_{b_0,b_1,b_2} = 1] P[K_{b_0,b_1`,b_2`} = 1]) \\
&\leq \sum_{b_0,(b_1,b_2,b_1`,b_2`)} P[K_{b_0,b_1,b_2} = 1, K_{b_0,b_1`,b_2`} = 1]
\end{aligned}
$$

because when $b_0 \neq b_0$‘, that is, $U_{1c}^N(..., b_0)$ and $U_{1c}^N(..., b_0')$ are independent, the RV's K_{b_0,b_1,b_2} and $K_{b‘_0,b_1‘,b_2‘}$ are independent and they do not contribute to the summation. We thus can focus only on the case $b_0 = b_0$‘. We can write:

$$
\begin{aligned}
\mathrm{Var}[K] \leq & \underbrace{\sum_{b_0,\, b_1 = b_1‘,\, b_2 = b_2‘} P[K_{b_0,b_1,b_2} = 1]}_{=E[K]} \\
& + \underbrace{\sum_{b_0,\, b_1 = b_1‘,\, b_2 \neq b_2‘} P[K_{b_0,b_1,b_2} = 1] P[K_{b_0,b_1,b_2‘} = 1 | K_{b_0,b_1,b_2} = 1]}_{=E[K]\, 2^{N(R'_{2pb} - B)}} \\
& + \underbrace{\sum_{b_0,\, b_1 \neq b_1‘,\, b_2 = b_2‘} P[K_{b_0,b_1,b_2} = 1] P[K_{b_0,b_1‘,b_2} = 1 | K_{b_0,b_1,b_2} = 1]}_{=E[K]\, 2^{N(R'_{1pb} - C)}} \\
& + \underbrace{\sum_{b_0,\, b_1 \neq b_1‘,\, b_2 \neq b_2‘} P[K_{b_0,b_1,b_2} = 1] P[K_{b_0,b_1‘,b_2‘} = 1 | K_{b_0,b_1,b_2} = 1]}_{=E[K]\, 2^{N(R'_{1pb} + NR'_{2pb} - D)}}
\end{aligned}
$$

and

$$
\begin{aligned}
&2^{-NB} \\
&= P[K_{b_0,b_1,b_2`} = 1 | K_{b_0,b_1,b_2} = 1] \\
&= P[\big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2`)\big) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})| \\
&\qquad \big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2)\big) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})] \\
&= \sum_{u_{2pb}^N \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}} | u_{2c}^N, x_2^N, u_{1c}^N, u_{1pb}^N)} p_{U_{2pb}|U_{2c},U_{1c},X_2} \\
&= 2^{-NI(U_{2pb};U_{1pb}|U_{2c},U_{1c},X_2)},
\end{aligned}
$$

and

$$
\begin{aligned}
2^{-NC} &= P[K_{b_0,b_1`,b_2} = 1 | K_{b_0,b_1,b_2} = 1] \\
&= P[\big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1`), U_{2pb}^N(1,1,b_0,1,b_2)\big) \\
&\quad \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}}) | \\
&\quad \big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2)\big) \\
&\quad \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})] \\
&= \sum_{u_{1pb}^N \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}} | u_{2c}^N, x_2^N, u_{1c}^N, u_{2pb}^N)} p_{U_{1pb}|U_{2c},U_{1c}} \\
&= 2^{-NI(U_{1pb};X_2,U_{2pb}|U_{1c},U_{2c})},
\end{aligned}
$$

and

$$
\begin{aligned}
2^{-ND} &= P[K_{b_0,b_1`,b_2`} = 1 | K_{b_0,b_1,b_2} = 1] \\
&= P[\big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1`), U_{2pb}^N(1,1,b_0,1,b_2`)\big) \\
&\quad\ \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})| \\
&\quad\ \big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{2pb}^N(1,1,b_0,1,b_2)\big) \\
&\quad\ \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})] \\
&= \sum_{(u_{1pb}^N, u_{2pb}^N) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}} | u_{2c}^N, x_2^N, u_{1c}^N)} p_{U_{2pb}|U_{2c},U_{1c},X_2} \, p_{U_{1pb}|U_{2c},U_{1c}} \\
&= 2^{-NI(U_{1pb};X_2,U_{2pb}|U_{1c},U_{2c})} = 2^{-NC}.
\end{aligned}
$$

Hence, we can bound $P[K=0]$ as:

$$
0 \le P[K=0] \;\; \le \frac{1 + 2^{N(R'_{1pb}-C)} + 2^{N(R'_{2pb}-B)} + 2^{N(R'_{1pb}+R'_{2pb}-C)}}{2^{N(R'_{1c}+R'_{1pb}+R'_{2pb}-I(U_{1c};X_2|U_{2c})-C)}}
$$

and $P[K=0] \to 0$ if

$$
\begin{aligned}
& R'_{1c} + R'_{1pb} + R'_{2pb} - I(U_{1c}; X_2|U_{2c}) - C > 0 \\
& R'_{1c} + R'_{1pb} + R'_{2pb} - I(U_{1c}; X_2|U_{2c}) - C - (R'_{2pb} - B) > 0 \\
& R'_{1c} + R'_{1pb} + R'_{2pb} - I(U_{1c}; X_2|U_{2c}) - C - (R'_{1pb} - C) > 0 \\
& R'_{1c} + R'_{1pb} + R'_{2pb} - I(U_{1c}; X_2|U_{2c}) - C - (R'_{1pb} + R'_{2pb} - C) > 0
\end{aligned}
$$

TABLE I

ERROR EVENTS AT DECODER 2.

Event	w_{2c}	(w_{1c}, b_1)	w_{2pa}	w_{2pb}	$p_{Y_2\|\star}$
$E_{2,1}$	X	$\cdots$	$\cdots$	$\cdots$	p_{Y_2}
$E_{2,2a}$	1	X	X	$\cdots$	$p_{Y_2\|U_{2c}}$
$E_{2,2b}$	1	1	X	$\cdots$	$p_{Y_2\|U_{2c},U_{1c}}$
$E_{2,3a}$	1	X	1	X	$p_{Y_2\|U_{2c},X_2}$
$E_{2,3b}$	1	1	1	X	$P_{Y_2\|U_{2c},U_{1c},X_2}$

that is, if

$$\begin{aligned} R'_{1c} + R'_{1pb} + R'_{2pb} &> I(U_{1c}; X_2|U_{2c}) + I(U_{1pb}; X_2, U_{2pb}|U_{1c}, U_{2c}) \\ &= I(U_{1c}, U_{1pb}; X_2|U_{2c}) + I(U_{1pb}; U_{2pb}|U_{1c}, U_{2c}, X_2) \\ R'_{1c} + R'_{1pb} &> I(U_{1c}; X_2|U_{2c}) + I(U_{1pb}; X_2|U_{1c}, U_{2c}) \\ &= I(U_{1c}, U_{1pb}; X_2|U_{2c}) \\ R'_{1c} + R'_{2pb} &> I(U_{1c}; X_2|U_{2c}), \\ R'_{1c} &> I(U_{1c}; X_2|U_{2c}) \end{aligned}$$

as in Equation 2.12a-Equation 2.12c, because the second to last equation is redundant.

2.6.7 Decoding Errors at decoder 2

If decoder 2 decodes $(\widehat{w}_{2c}, \widehat{w}_{2pa}, \widehat{w}_{2pb}) \neq (1,1,1)$, then an error is committed. The probability of error at decoder 2 is bounded as:

$$P[error\ 2|encoding\, successful] \leq \sum_{i\in\{1,2a,2b,3a,3b\}} P[E_{2,i}],$$

where $E_{2,i},\ i \in \{1,2a,2b,3a,3b\}$, are the error event defined in Table I. In Table I, an "X" means that the corresponding message is in error (when the header of the column contains two indices, an "X" indicates that at least one of the two indexes is wrong), a "1" means that the corresponding message is correct, while the dots "$\cdots$" indicates that "*it does not matter whether the corresponding message is correct or not; in this case the most restrictive case is when the message is actually wrong.*" The last column of Table I specifies the $p_{Y_2|\star}$ to be used in Equation 2.15.

We have that $P[error\ 2|encoding\, successful] \to 0$ when $N \to \infty$ if:

- When the event $E_{2,1}$ occurs we have $\widehat{w}_{2c} \neq 1$. In this case the received Y_2^N is independent of the transmitted sequences. This follows from the fact that the codewords U_{2c}^N are generated in an iid fashion and all the other codewords are generated independently conditioned on U_{2c}^N. Hence, when decoder 2 finds a wrong U_{2c}^N, all the decoded codewords are independent of the transmitted ones. We can bound the error probability of $E_{2,1}$ as:

$$
\begin{aligned}
&P[E_{2,1}] = P\Bigg[\bigcup_{\widetilde{w}_{2c}\neq 1,\widetilde{w}_{2pa},\widetilde{w}_{1c},\widetilde{w}_{2pb},b_0,b_2} \\
&(Y_2^N, U_{2c}^N(\widetilde{w}_{2c}), U_{1c}^N(\widetilde{w}_{1c},\widetilde{w}_{2c},b_0), X_2^N(\widetilde{w}_{2c},\widetilde{w}_{2pa}), U_{2pb}^N(\widetilde{w}_{2c},\widetilde{w}_{2pa},\widetilde{w}_{1c},b_0,\widetilde{w}_{2pb},b_2)) \\
&\in T_\epsilon^N\left(p_{Y_2,U_{2c},U_{1c},X_2,U_{2pb}}\right)\Bigg] \\
&\leq 2^{N(R_{2c}+R_{2pa}+R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb})} \sum_{(y_2^N,u_{2c}^N,u_{1c}^N,x_2^N,u_{2pb}^N)\in T_\epsilon^N\left(p_{Y_2,U_{2c},U_{1c},X_2,U_{2pb}}\right)} p_{2|\star}|_{\star=\emptyset} \\
&\leq 2^{N(R_{2c}+R_{2pa}+R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb}-I_{2|\star}|_{\star=\emptyset})}
\end{aligned}
$$

for $p_{2|\star}$ given in Equation 2.16 and $I_{2|\star}$ given in Equation 2.17. Hence $P[E_{2,1}] \to 0$ as $N \to \infty$ if Equation 3.36d is satisfied.

- When the event $E_{2,2}$ occurs, i.e., either $E_{2,2a}$ or $E_{2,2b}$, we have $\widehat{w}_{2c} = 1$ but $\widehat{w}_{2pa} \neq 1$. Whether $\widehat{w}_{1c}$ is correct or not, it does not matter since decoder 2 is not interested in $\widehat{w}_{1c}$. However we need to consider whether the pair $(\widehat{w}_{1c}, \widehat{b}_0)$ is equal to the transmitted one or not because this affect the way the joint probability among all involved RV's factorizes. We have:

  - Case $E_{2,2a}$: either $\widehat{w}_{1c} \neq 1$ or $\widehat{b}_0 \neq \overline{b}_0$. In this case, conditioned on the (correct) decoded sequence U_{2c}^N, the output Y_2^N is independent of the (wrong) decoded sequences U_{1c}^N, U_{2pa}^N and also of U_{2pb}^N (because U_{2pb}^N is superimposed to the wrong pair

$(U_{1c}^N, U_{2pa}^N))$. It is easy to see that the most stringent error event is when both $\widehat{w}_{1c} \neq 1$ and $\widehat{b}_0 \neq \bar{b}_0$. Thus we have

$$P[E_{2,2a}] = P\Bigg[\bigcup_{\widetilde{w}_{2pa}\neq 1, \widetilde{w}_{1c}\neq 1, b_0\neq \bar{b}_0, \widetilde{w}_{2pb}, b_2}$$

$$(Y_2^N, U_{2c}^N(1), U_{1c}^N(1, \widetilde{w}_{1c}, b_0), X_2^N(1, \widetilde{w}_{2pa}), U_{2pb}^N(1, \widetilde{w}_{2pa}, \widetilde{w}_{1c}, b_0, \widetilde{w}_{2pb}, b_2))$$

$$\in T_\epsilon^N \left(p_{Y_2, U_{2c}, U_{1c}, X_2, U_{2pb}}\right)\Bigg]$$

$$\leq 2^{N(R_{2pa}+R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb})} \sum_{(y_2^N, u_{2c}^N, u_{1c}^N, x_2^N, u_{2pb}^N)\in T_\epsilon^N\left(p_{Y_2, U_{2c}, U_{1c}, X_2, U_{2pb}}\right)} p_{2|\star}|_{\star=U_{2c}}$$

$$\leq 2^{N(R_{2pa}+R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb}-I_{2|\star}|_{\star=U_{2c}})}$$

for $p_{2|\star}$ given in Equation 2.16 and $I_{2|\star}$ given in Equation 2.17. Hence $P[E_{2,2a}] \to 0$ as $N \to \infty$ if Equation 3.36e is satisfied.

– Case $E_{2,2b}$: both $\widehat{w}_{1c} = 1$ and $\widehat{b}_0 = \bar{b}_0$. In this case, conditioned on the (correct) decoded (U_{2c}^N, U_{1c}^N), the output Y_2^N is independent of the (wrong) decoded sequences (X_2^N, U_{2pb}^N). Thus we have

$$P[E_{2,2b}] = P\Bigg[\bigcup_{\widetilde{w}_{2pa}\neq 1, \widetilde{w}_{2pb}, b_2}$$

$$(Y_2^N, U_{2c}^N(1), U_{1c}^N(1, 1, \bar{b}_0), X_2^N(1, \widetilde{w}_{2pa}), U_{2pb}^N(1, \widetilde{w}_{2pa}, 1, \bar{b}_0, \widetilde{w}_{2pb}, b_2)) \in T_\epsilon^N \left(p_{Y_2, U_{2c}, U_{1c}, X_2, U_{2pb}}\right)\Bigg]$$

$$\leq 2^{N(R_{2pa}+R_{2pb}+R'_{2pb})} \sum_{(y_2^N, u_{2c}^N, u_{1c}^N, x_2^N, u_{2pb}^N)\in T_\epsilon^N\left(p_{Y_2, U_{2c}, U_{1c}, X_2, U_{2pb}}\right)} p_{2|\star}|_{\star=(U_{2c}, U_{1c})}$$

$$\leq 2^{N(R_{2pa}+R_{2pb}+R'_{2pb}-I_{2|\star}|_{\star=(U_{2c}, U_{1c})})}$$

for $p_{2|\star}$ given in Equation 2.16 and $I_{2|\star}$ given in Equation 2.17. Hence $P[E_{2,2b}] \to 0$ as $N \to \infty$ if Equation 3.36f is satisfied.

- When the event $E_{2,3}$ occurs, i.e., either $E_{2,3a}$ or $E_{2,3b}$, we have $\widehat{w}_{2c} = 1$,$\widehat{w}_{2pa} = 1$ but $\widehat{w}_{2pb} \neq 1$. Again, whether $\widehat{w}_{1c}$ is correct or not, it does not matter since decoder 2 is not interested in $\widehat{w}_{1c}$. However we need to consider whether the pair $(\widehat{w}_{1c}, \widehat{b}_0)$ is equal to the transmitted one or not because this affect the way the joint probability among all involved RV's factorizes. The analysis proceeds similarly as for the even $E_{2,2}$.

  We have:

  - Case $E_{2,3a}$: either $\widehat{w}_{1c} \neq 1$ or $\widehat{b}_0 \neq \overline{b}_0$. In this case, conditioned on the (correct) decoded sequences (U_{2c}^N, X_2^N), the output Y_2^N is independent of the (wrong) decoded sequences $(U_{1c}^N, U_{2c}^n U_{2pb}^N)$. It is easy to see that the most stringent error event is when both $\widehat{w}_{1c} \neq 1$ and $\widehat{b}_0 \neq \overline{b}_0$. Thus we have

$$P[E_{2,3a}] = P\Bigg[\bigcup_{\widetilde{w}_{1c} \neq 1, b_0 \neq \overline{b}_0, \widetilde{w}_{2pb}, b_2}$$
$$(Y_2^N, U_{2c}^N(1), U_{1c}^N(1, \widetilde{w}_{1c}, b_0), X_2^N(1,1), U_{2pb}^N(1, 1, \widetilde{w}_{1c}, b_0, \widetilde{w}_{2pb}, b_2)) \in T_\epsilon^N\left(p_{Y_2, U_{2c}, U_{1c}, X_2, U_{2pb}}\right)\Bigg]$$
$$\leq 2^{N(R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb})} \sum_{(y_2^N, u_{2c}^N, u_{1c}^N, x_2^N, u_{2pb}^N) \in T_\epsilon^N\left(p_{Y_2, U_{2c}, U_{1c}, X_2, U_{2pb}}\right)} p_{2|\star}|_{\star=(U_{2c}, X_2)}$$
$$\leq 2^{N(R_{2pa}+R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb}-I_{2|\star}|_{\star=(U_{2c},X_2)})}$$

for $p_{2|\star}$ given in Equation 2.16 and $I_{2|\star}$ given in Equation 2.17. Hence $P[E_{2,3a}] \to 0$ as $N \to \infty$ if Equation 3.36g is satisfied.

– Case $E_{2,3b}$: both $\widehat{w}_{1c} = 1$ and $\widehat{b}_0 = \bar{b}_0$. In this case, conditioned on the (correct) decoded sequences $(U_{2c}^N, X_2^N, U_{1c}^N)$, the output Y_2^N is independent of the (wrong) decoded sequence U_{2pb}^N. However, since $(U_{2c}^N, X_2^N, U_{1c}^N)$ is the triplet that passed the encoding binning step, they are jointly typical. Hence, in this case we cannot use the factorization in $p_{2|\star}$ given in Equation 2.16, but we need to replace $p_{U_{1c}|U_{2c}}$ in Equation 2.16 with $p_{U_{1c}|U_{2c},X_2}$. Thus we have

$$
\begin{aligned}
P[E_{2,3b}] = P \Bigg[& \bigcup_{\widetilde{w}_{2pb}, b_2} \\
& (Y_2^N, U_{2c}^N(1), U_{1c}^N(1,1,\bar{b}_0), X_2^N(1,1), U_{2pb}^N(1,1,1,\bar{b}_0,\widetilde{w}_{2pb},b_2)) \in T_\epsilon^N \left(p_{Y_2,U_{2c},U_{1c},X_2,U_{2pb}}\right)\Bigg] \\
\le {} & 2^{N(R_{2pb}+R'_{2pb})} \sum_{(y_2^N,u_{2c}^N,u_{1c}^N,x_2^N,u_{2pb}^N)\in T_\epsilon^N\left(p_{Y_2,U_{2c},U_{1c},X_2,U_{2pb}}\right)} \\
& p_{U_{2c}}\, p_{X_2|U_{2c}}\, p_{U_{1c}|U_{2c},X_2}\, p_{U_{2pb}|U_{2c},U_{1c},X_2}\, p_{Y_2|U_{1c},U_{2c},X_2} \\
\le {} & 2^{N(R_{2pb}+R'_{2pb}-I(Y_2;U_{2pb}|U_{1c},U_{2c},X_2))}
\end{aligned}
$$

Hence $P[E_{2,3b}] \to 0$ as $N \to \infty$ if Equation 3.36h is satisfied.

TABLE II

ERROR EVENTS AT DECODER 1.

Event	w_{2c}	(w_{1c}, b_1)	w_{1pb}	$p_{Y_1\|\star}$
$E_{1,1}$	X	$\cdots$	$\cdots$	p_{Y_1}
$E_{1,2}$	1	X	$\cdots$	$p_{Y_1\|U_{2c}}$
$E_{1,3}$	1	1	X	$P_{Y_1\|U_{2c},U_{1c}}$

2.6.8 Decoding Errors at Decoder 1

The probability of error at decoder 1 is bounded as:

$$P[error\ 1|encoding\, successful] \leq \sum_{i=1}^{3} P[E_{1,i}],$$

where $P[E_{1,i}]$ is the error event defined in Table II. The meaning of the symbols in Table II is as for Table I. We have that $P[error\ 1|encoding\, successful] \to 0$ when $N \to \infty$ if:

- When the event $E_{1,1}$ occurs we have $\widehat{w}_{2c} \neq 1$. In this case the received Y_1^N is independent of the transmitted sequences. We can bound the error probability of $E_{1,1}$ as:

$$P[E_{1,1}] = P\Bigg[\bigcup_{\widetilde{w}_{2c}\neq 1, \widetilde{w}_{1c}, \widetilde{w}_{1pb}, b_0, b_1}$$

$$(Y_1^N, U_{2c}^N(\widetilde{w}_{2c}), U_{1c}^N(\widetilde{w}_{1c}, \widetilde{w}_{2c}, b_0), U_{1pb}^N(\widetilde{w}_{2c}, \widetilde{w}_{2pa}, \widetilde{w}_{1c}, b_0, \widetilde{w}_{2pb}, b_1)) \in T_\epsilon^N\left(p_{Y_1,U_{2c},U_{1c},U_{1pb}}\right)\Bigg]$$

$$\leq 2^{N(R_{2c}+R_{1c}+R'_{1c}+R_{1pb}+R'_{1pb})} \sum_{(y_1^N, u_{2c}^N, u_{1c}^N, u_{1pb}^N)\in T_\epsilon^N\left(p_{Y_1,U_{2c},U_{1c},U_{1pb}}\right)} p_{1|\star}|_{\star=\emptyset}$$

$$\leq 2^{N(R_{2c}+R_{2pa}+R_{1c}+R'_{1c}+R_{2pb}+R'_{2pb}-I_{1|\star}|_{\star=\emptyset})}$$

for $p_{1|\star}$ given in Equation 2.16 and $I_{1|\star}$ given in Equation 2.20. Hence $P[E_{1,1}] \to 0$ as $N \to \infty$ if Equation 3.36i is satisfied.

- When the event $E_{1,2}$ occurs, either $\widehat{w}_{1c} \neq 1$, $\widehat{b}_0 \neq \bar{b}_0$ or both. In this case, conditioned on the (correct) decoded sequence U_{2c}^N, the output Y_1^N is independent of the (wrong) decoded sequences U_{1c}^N and U_{1pb}^N . It is easy to see that the most stringent error event is when both $\widehat{w}_{1c} \neq 1$ and $\widehat{b}_0 \neq \bar{b}_0$. Thus we have

$$
\begin{aligned}
P[E_{1,2}] = P &\Bigg[\bigcup_{\widetilde{w}_{1c} \neq 1, b_0 \neq \bar{b}_0, \widetilde{w}_{1pb}, b_1} \\
&(Y_1^N, U_{2c}^N(1), U_{1c}^N(1, \widetilde{w}_{1c}, b_0), U_{1pb}^N(1, \widetilde{w}_{1c}, b_0, \widetilde{w}_{1pb}, b_1)) \in T_\epsilon^N \left(p_{Y_1,U_{2c},U_{1c},U_{1pb}}\right)\Bigg] \\
&\leq 2^{N(R_{1c}+R'_{1c}+R_{1pb}+R'_{1pb})} \sum_{(y_1^N, u_{2c}^N, u_{1c}^N, u_{1pb}^N) \in T_\epsilon^N \left(p_{Y_1,U_{2c},U_{1c},U_{1pb}}\right)} p_{1|\star}|_{\star=U_{2c}} \\
&\leq 2^{N(R_{1c}+R'_{1c}+R_{1pb}+R'_{1pb}-I_{1|\star}|_{\star=U_{2c}})}
\end{aligned}
$$

for $p_{1|\star}$ given in Equation 2.19 and $I_{1|\star}$ given in Equation 2.20. Hence $P[E_{1,2}] \to 0$ as $N \to \infty$ if Equation 3.36j is satisfied.

- When the event $E_{1,3}$ occurs, either $\widehat{w}_{1pb} \neq 1$, $\widehat{b}_1 \neq \bar{b}_1$ or both. In this case, conditioned on the (correct) decoded sequence U_{2c}^N and U_{1c}^N), the output Y_1^N is independent of the (wrong) decoded sequences U_{1pb}^N. It is easy to see that the most stringent error event is when both $\widehat{w}_{1pb} \neq 1$ and $\widehat{b}_1 \neq \bar{b}_1$. Thus we have

$$P[E_{1,3}] = P\left[\bigcup_{\widetilde{w}_{1pb}\neq 1, b_1 \neq \bar{b}_1}\right.$$

$$\left.(Y_1^N, U_{2c}^N(1), U_{1c}^N(1,1,\bar{b}_0), U_{1pb}^N(1,1,\bar{b}_0,\widetilde{w}_{1pb}, b_1)) \in T_\epsilon^N\left(p_{Y_1,U_{2c},U_{1c},U_{1pb}}\right)\right]$$

$$\leq 2^{N(R_{1pb}+R'_{1pb})} \sum_{(y_1^N, u_{2c}^N, u_{1c}^N, u_{1pb}^N)\in T_\epsilon^N\left(p_{Y_1,U_{2c},U_{1c},U_{1pb}}\right)} p_{1|\star}|_{\star=U_{2c},U_{1c}}$$

$$\leq 2^{N(R_{1c}+R'_{1c}+R_{1pb}+R'_{1pb}-I_{1|\star}|_{\star=U_{2c},U_{1c}})}$$

for $p_{1|\star}$ given in Equation 2.19 and $I_{1|\star}$ given in Equation 2.20. Hence $P[E_{1,3}] \to 0$ as $N \to \infty$ if Equation 3.36k is satisfied.

$\square$

2.6.9 Two step binning

It is also possible to perform binning in a sequential manner. First, U_{1c} is binned against X_1, and then U_{1pb} and U_{2pb} are binned against each other conditioned on (U_{2c}, U_{1c}) and (U_{2c}, X_2, U_{1c}) respectively.

With respect to the encoding operation of the previous section, this affects Section 2.6.3 as follows.

Given the message $w_2 = (w_{2c}, w_{2pb}, w_{2pa})$ and the message $w_1 = (w_{1c}, w_{1pb})$, encoder 1 looks for b_0 such that

$$\begin{aligned}&(U_{2c}^N(w_{2c}), X_2^N(w_{2c}, w_{2pa}), U_{1c}^N(w_{2c}, w_{1c}, b_0),\\ &\in T_\epsilon^N(p_{U_{2c},X_2,U_{1c}}).\end{aligned}$$

If not such a b_0 exists, it sets $b_0 = 1$. If more than one such b_0 exists, it picks one uniformly at random.

For the selected b_0, encoder 1 looks for (b_1, b_2) such that:

$$\begin{aligned}&(U_{2c}^N(w_{2c}), X_2^N(w_{2c}, w_{2pa}), U_{1c}^N(w_{2c}, w_{1c}, b_0), U_{1pb}^N(w_{2c}, w_{1c}, b_0, w_{1pb}, b_1), U_{2pb}^N(w_{2c}, w_{1c}, b_0, w_{2pb}, b_2))\\ &\in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}}).\end{aligned}$$

If not such a (b_1, b_2) exists, it sets $(b_1, b_2) = (1, 1)$. If more than one such a (b_1, b_2) exists, it picks one uniformly at random from the found ones.

For the selected (b_0, b_1, b_2), encoder 1 sends $X_1^N(w_{2pa}, w_{2c}, w_{1c}, b_0, w_{1pb}, b_1, w_{2pb}, b_2)$.

The next theorem states the condition under which this two step encoding procedure is successful with high probability.

Theorem 2.6.3. *The encoding procedure of Section 2.6.9 is successful if*

$$
\begin{aligned}
R'_{1c} &\geq I(U_{1c}; X_2 | U_{2c}), && (2.21a)\\
R'_{1pb} &\geq I(U_{1pb}; X_2 | U_{2c}, U_{1c}), && (2.21b)\\
R'_{1pb} + R'_{2pb} &\geq I(U_{1pb}; U_{2pa}, X_2, U_{2pb} | U_{2c}, U_{1c}).. && (2.21c)
\end{aligned}
$$

Proof. An encoding error is committed if we cannot find a b_0 in the first step or if, upon finding the correct b_0 in the first encoding step, we cannot find the correct (b_1, b_2) in the second step. Let $E_{e,0}$ the probability of the first event and $E_{e,12}$ of the latter, than:

$$P[encoding\, NOT\, successful] \leq P[E_{e,0}] + P[E_{e,12} | E^c_{e,0}]$$

where

$$
\begin{aligned}
P[E_{e,0}] &= P[\textstyle\bigcap_{b_0=1}^{2^{NR'_{1c}}} \big(U^N_{2c}(1), X^N_2(1,1), U^N_{1c}(1,1,b_0)\big) \notin T^N_\epsilon(p_{U_{2c},X_2,U_{1c}})]\\
&= (1 - P[\big(U^N_{2c}(1), X^N_2(1,1), U^N_{1c}(1,1,b_0)\big) \notin T^N_\epsilon(p_{U_{2c},X_2,U_{1c}})])^{2^{NR'_{1c}}}.
\end{aligned}
$$

Using standard typicality arguments we have

$$
\begin{aligned}
&P[\big(U^N_{2c}(1), X^N_2(1,1), U^N_{1c}(1,1,b_0)\big) \notin T^N_\epsilon(p_{U_{2c},X_2,U_{1c}})]\\
&= \textstyle\sum_{u_{1c} \in T^N_\epsilon(p_{U_{2c},X_2,U_{1c}} | U_{2c}, X_2)} \geq (1-\epsilon) 2^{N(I(U_{1c};X_2|U_{2c})+\delta)}.
\end{aligned}
$$

Now we can write

$$
\begin{aligned}
P[E_{e,0}] &\leq (1-(1-\epsilon)2^{N(I(U_{1c};X_2|U_{2c})+\delta)})^{2^{NR'_{1c}}} \\
&\leq \exp\left(1-(1-\epsilon)2^{N(R'_{1c}-I(U_{1c};X_2|U_{2c})+\delta)}\right)
\end{aligned}
$$

so that $P[E_{e,0}] \to 1$ when $N \to 0$ if Equation 2.21a is satisfied.

Now the error event $E_{e,12}$ can be divided in three distinct error events:

- $E_{e,21\ a}$: it is not possible to find b_1 such that $(U_{2c}^N, X_2^N, U_{1c}^N, U_{1pb}^N) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb}})$,
- $E_{e,21\ b}$: it is not possible to find b_2 such that $(U_{2c}^N, X_2^N, U_{1c}^N, U_{2pb}^N) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{2pb}})$.
- $E_{e,21\ c}$ Given that we can find b_1 and b_2 satisfy the first two equations, we cannot find a couple (b_1, b_2) such that $(U_{2c}^N, X_2^N, U_{1c}^N, U_{1pb}^N, U_{2pb}^N) \in T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})$.

We now establish the rate bounds that guarantee that the probability of error of each of these events goes to zero.

For $E_{e,21\ a}$ we have:

$$
P[E_{e,21\ a}] \;= (1-P[\left(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1)\right) \notin T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb}})])^{2^{NR'_{1pb}}}.
$$

where

$$
\begin{aligned}
&P[\left(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1)\right) \notin T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb}}) \\
&] \geq (1-\epsilon)2^{-N(I(X_2;U_{1pb}|U_{2c},U_{1c})+\delta)}
\end{aligned}
$$

As for $E_{e,0}$, this implies that $P[E_{e,21\ a}] \to 1$ when $N \to 0$ if Equation 2.21b is satisfied.

For $E_{e,21\ b}$, we have that the probability of this event goes to one for large N given that (U_{2c}, X_2, U_{1c}) appear to be generated according to the distribution $p_{U_{2c},X_2,U_{1c}}$ and U_{2pb} is generated according to $p_{U_{2pb}|U_{2c},X_2,U_{1c}}$.

For $E_{e,21\ c}$ we have:

$$\begin{aligned} P[E_{e,21\ c}] &= (1 - P[\Big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{1pb}^N(1,1,b_0,1,b_2)\Big) \\ &\quad \notin T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})])^{2^{N(R'_{1pb}+R'_{2pb})}}. \end{aligned}$$

where

$$\begin{aligned} &P[\Big(U_{2c}^N(1), X_2^N(1,1), U_{1c}^N(1,1,b_0), U_{1pb}^N(1,1,b_0,1,b_1), U_{1pb}^N(1,1,b_0,1,b_2)\Big) \\ &\notin T_\epsilon^N(p_{U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}})]) \leq 2^{I()+\delta} \end{aligned}$$

this implies that $P[E_{e,21\ c}] \to 1$ when $N \to 0$ if Equation 2.21c is satisfied. □

Remark 2.6.4. Since the binning rate Equation 2.12a of Theorem 2.6.1 can be taken with equality, the two step binning has the same performance of the joint binning. In fact by setting *Equation* 2.21*a* to hold with equality, we obtain the equality between the binning rate expression of the joint binning and the two step binning.

A plot of the permissable binning rates R_{1pb} and R_{2pb} is depicted in Figure 3

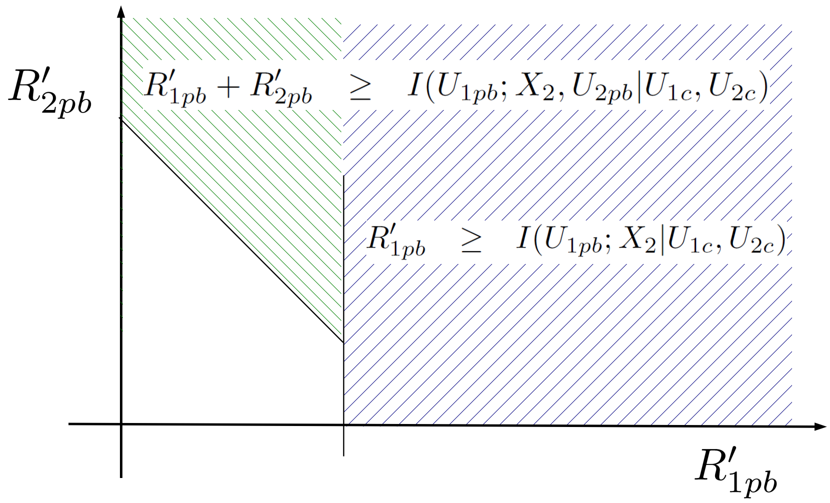

Figure 3. The region of the admissible binning rates R_{1pb} and R_{2pb} in Theorem 2.6.1.

2.7 Comparison with existing achievable regions

We now show that the region of Theorem 2.6.1 contains all other known achievable rate regions for the DM-CIFC. Showing inclusion of the rate regions (50, Thm. 2) , (52, Thm. 1) and (51, Thm. 4.1) is sufficient to demonstrate the largest known DM-CIFC region, since the region of (50) is shown to contain those of (25, Thm. 1) and (29).

2.7.1 Containment of region of Devroye et al. in $\mathcal{R}_{RTD}$

We refer to the region in (52, Thm. 1) as $\mathcal{R}_{DMT}$ for brevity. We show this inclusion of $\mathcal{R}_{DMT}$ in $\mathcal{R}_{RTD}$ with the following steps:

- We enlarge the region $\mathcal{R}_{DMT}$ by removing some rate constraints.
- We further enlarge the region by enlarging the set of possible input distributions. This allows

us to remove the V_{11} and Q from the inner bound. We refer to this region as $\mathcal{R}_{DMT}^{out}$ since is enlarges the original achievable region.

• We make a correspondence between the RV's and corresponding rates of $\mathcal{R}_{DMT}^{out}$ and $\mathcal{R}_{RTD}$.

• We choose a particular subset of $\mathcal{R}_{RTD}$, $\mathcal{R}_{RTD}^{in}$, for which we can more easily show $\mathcal{R}_{DMT} \subseteq \mathcal{R}_{DMT}^{out} \subset \mathcal{R}_{RTD}^{in} \subseteq \mathcal{R}_{RTD}$, since $\mathcal{R}_{DMT}^{out}$ and $\mathcal{R}_{RTD}^{in}$ have identical input distribution decompositions and similar rate bound equations.

Enlarge the region $\mathcal{R}_{DMT}$

We first enlarge the rate region of (52, Thm. 1), $\mathcal{R}_{DMT}$ by removing a number of constraints (specifically, we remove equations (2.6, 2.8, 2.10, 2.13, 2.14, 2.16 2.17) of (52, Thm. 1)) to obtain the region $\mathcal{R}_{DMT}^{out}$ defined as the set of all rate pairs satisfying:

$$R'_{21} = I(V_{21}; V_{11}, V_{12}|W) \tag{2.22a}$$

$$R'_{22} = I(V_{22}; V_{11}, V_{12}|W) \tag{2.22b}$$

$$R_{11} \leq I(Y_1, V_{12}, V_{21}; V_{11}|W) \tag{2.22c}$$

$$R_{21} + R'_{21} \leq I(Y_1, V_{11}, V_{12}; V_{21}|W) \tag{2.22d}$$

$$R_{11} + R_{21} + R'_{21} \leq I(Y_1, V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W) \tag{2.22e}$$

$$R_{11} + R_{21} + R'_{21} + R_{12} \leq I(Y_1; V_{11}, V_{21}, V_{12}|W) + I(V_{11}, V_{12}; V_{21}|W) \tag{2.22f}$$

$$R_{22} + R'_{22} \leq I(Y_2, V_{12}, V_{21}; V_{22}|W) \tag{2.22g}$$

$$R_{22} + R'_{22} + R_{21} + R'_{21} \leq I(Y_2, V_{12}; V_{22}, V_{21}|W) + I(V_{22}; V_{21}|W) \tag{2.22h}$$

$$R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} \leq I(Y_2; V_{22}, V_{21}, V_{12}|W) + I(V_{22}, V_{21}; V_{12}|W). \tag{2.22i}$$

taken over the union of distributions

$$p_W p_{V_{11}} p_{V_{12}} p_{X_1|V_{11},V_{12}} p_{V_{21}|V_{11}V_{12}} p_{V_{22}|V_{11},V_{12}} p_{X_2|V_{11},V_{12},V_{21},V_{22}}. \tag{2.23}$$

Following the line of thoughts in (53, Appendix D) it is possible to show that without loss of generality we can set X_1 to be a deterministic function of V_{11} and V_{12}, allowing us insert X_1 next to V_{11}, V_{12} as follows:

$$\begin{aligned}
R'_{21} &= I(V_{21}; X_1, V_{11}, V_{12}|W) && (2.24a)\\
R'_{22} &= I(V_{22}; X_1, V_{11}, V_{12}|W) && (2.24b)\\
R_{11} &\leq I(Y_1, V_{12}, V_{21}; V_{11}|W) && (2.24c)\\
R_{21} + R'_{21} &\leq I(Y_1, X_1, V_{11}, V_{12}; V_{21}|W) && (2.24d)\\
R_{11} + R_{21} + R'_{21} &\leq I(Y_1, V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W) && (2.24e)\\
R_{11} + R_{21} + R'_{21} + R_{12} &\leq I(Y_1; X_1, V_{11}, V_{12}, V_{21}|W) &&\\
&\quad + I(X_1, V_{11}, V_{12}; V_{21}|W) && (2.24f)\\
R_{22} + R'_{22} &\leq I(Y_2, V_{12}, V_{21}; V_{22}|W) && (2.24g)\\
R_{22} + R'_{22} + R_{21} + R'_{21} &\leq I(Y_2, V_{12}; V_{22}, V_{21}|W) + I(V_{22}; V_{21}|W) && (2.24h)\\
R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} &\leq I(Y_2; V_{22}, V_{21}, V_{12}|W) + I(V_{22}, V_{21}; V_{12}|W) && (2.24i)
\end{aligned}$$

Using the factorization of the auxiliary RV's in (52, Thm. 1), we may insert X_1 next to V_{11} in Equation 2.24f.

For Equation 2.24c:

$$\begin{aligned} R_{11} &\leq I(Y_1, V_{12}, V_{21}; V_{11}|W) \\ &= I(Y_1, V_{21}; V_{11}|V_{12}, W) + I(V_{12}; V_{11}|W) \\ &= I(Y_1, V_{21}; V_{11}|V_{12}, W) \\ &= I(Y_1, V_{21}; X_1, V_{11}|V_{12}, W) \\ &= I(Y_1; X_1, V_{11}|V_{12}, V_{21}, W) + I(V_{21}; X_1, V_{11}|V_{12}, W). \end{aligned}$$

For Equation 2.24e we have:

$$\begin{aligned} R_{11} + R_{21} + R'_{21} &\leq I(Y_1, V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W) \\ &= I(Y_1; V_{11}, V_{21}|V_{12}, W) + I(V_{12}; V_{11}, V_{21}|W) + I(V_{11}; V_{21}|W) \\ &= I(Y_1; V_{11}, V_{21}|V_{12}, W) + I(V_{12}; V_{21}|V_{11}, W) + I(V_{11}; V_{21}|W) \\ &= I(Y_1; V_{11}, V_{21}|V_{12}, W) + I(V_{11}, V_{12}; V_{21}|W) \\ &= I(Y_1; X_1, V_{11}, V_{21}|V_{12}, W) + I(X_1, V_{11}, V_{12}; V_{21}|W) \end{aligned}$$

The original region is thus equivalent to

$$
\begin{aligned}
R'_{21} &= I(V_{21}; X_1, V_{11}, V_{12}|W) && \text{(2.25a)} \\
R'_{22} &= I(V_{22}; X_1, V_{11}, V_{12}|W) && \text{(2.25b)} \\
R_{11} &\leq I(Y_1; X_1, V_{11}|V_{12}, V_{21}|W) + I(V_{21}; X_1|V_{12}, W) && \text{(2.25c)} \\
R_{21} + R'_{21} &\leq I(Y_1, X_1, V_{11}, V_{12}; V_{21}|W) && \text{(2.25d)} \\
R_{11} + R_{21} + R'_{21} &\leq I(Y_1; X_1, V_{11}, V_{21}|V_{12}, W) + I(X_1; V_{21}|W) && \text{(2.25e)} \\
R_{11} + R_{21} + R'_{21} + R_{12} &\leq I(Y_1; X_1, V_{11}, V_{21}, V_{12}|W) \\
&\qquad + I(X_1, V_{11}, V_{12}; V_{21}|W) && \text{(2.25f)} \\
R_{22} + R'_{22} &\leq I(Y_2, V_{12}, V_{21}; V_{22}|W) && \text{(2.25g)} \\
R_{22} + R'_{22} + R_{21} + R'_{21} &\leq I(Y_2, V_{12}; V_{22}, V_{21}|W) + I(V_{22}; V_{21}|W) && \text{(2.25h)} \\
R_{22} + R'_{22} + R_{21} + R'_{21} + R_{12} &\leq I(Y_2; V_{22}, V_{21}, V_{12}|W) + I(V_{22}, V_{21}; V_{12}|W) && \text{(2.25i)}
\end{aligned}
$$

union over all distributions that factor as in Equation 2.23.

Enlarge the class of input distribution and eliminate V_{11} and W

Now increase the set of possible input distribution of Equation 2.23 by letting V_{11} to have any joint distribution with V_{12}. This is done by substituting $p_{V_{11}}$ with $p_{V_{11}|V_{12}}$ in the expression of the input distribution. With this substitution we have:

$$
\begin{aligned}
& p_W p_{V_{11}|V_{12}} p_{V_{12}} p_{X_1|V_{11},V_{12}} p_{V_{21}|X_1,V_{11}V_{12}} p_{V_{22}|X_1,V_{11},V_{12}} p_{X_2|X_1,V_{11},V_{12},V_{21},V_{22}} \\
= \;& p_W p_{V_{12}} p_{V_{11},X_1|V_{12}} p_{V_{21}|X_1,V_{11}V_{12}} p_{V_{22}|X_1,V_{11},V_{12}} p_{X_2|X_1,V_{11},V_{12},V_{21},V_{22}} \\
= \;& p_W p_{V_{12}} p_{X_1'|V_{12}} p_{V_{21}|X_1',V_{12}} p_{V_{22}|X_1',V_{12}} p_{X_2|X_1',V_{12},V_{21},V_{22}}
\end{aligned}
$$

with $X_1' = (X_1, V_{11})$. Since V_{12} is decoded at both decoders, the time sharing random W may be incorporated with V_{12} without loss of generality and thus can be dropped. The region described in (Equation 2.25) is convex and thus time sharing is not needed. With these simplifications, the region $\mathcal{R}_{DMT}^{out}$ is now defined as

$$
\begin{aligned}
R_{21}' &= I(V_{21}; X_1', V_{12}) && \text{(2.26a)}\\
R_{22}' &= I(V_{22}; X_1', V_{12}) && \text{(2.26b)}\\
R_{11} &\leq I(Y_1; X_1'|V_{12}, V_{21}) + I(V_{21}; X_1|V_{12}) && \text{(2.26c)}\\
R_{21} + R_{21}' &\leq I(Y_1, X_1', V_{12}; V_{21}) && \text{(2.26d)}\\
R_{11} + R_{21} + R_{21}' &\leq I(Y_1; X_1', V_{21}|V_{12}) + I(X_1; V_{21}) && \text{(2.26e)}\\
R_{11} + R_{21} + R_{21}' + R_{12} &\leq I(Y_1; X_1', V_{21}, V_{12}) + I(X_1', V_{12}; V_{21}) && \text{(2.26f)}\\
R_{22} + R_{22}' &\leq I(Y_2, V_{12}, V_{21}; V_{22}) && \text{(2.26g)}\\
R_{22} + R_{22}' + R_{21} + R_{21}' &\leq I(Y_2, V_{12}; V_{22}, V_{21}) + I(V_{22}; V_{21}) && \text{(2.26h)}\\
R_{22} + R_{22}' + R_{21} + R_{21}' + R_{12} &\leq I(Y_2; V_{22}, V_{21}, V_{12}) + I(V_{22}, V_{21}; V_{12}) && \text{(2.26i)}
\end{aligned}
$$

RV, rate of Theorem 2.6.1	RV, rate of (52, Thm. 1)	Comments
U_{2c}, R_{2c}	V_{12}, R_{12}	TX 2 $\to$ RX 1, RX 2
U_{1c}, R_{1c}	V_{21}, R_{21}	TX 1 $\to$ RX 1, RX 2
U_{1pb}, R_{1pb}	V_{22}, R_{22}	TX 1 $\to$ RX 1
X_2, R_{2pa}	X_1', R_{11}	TX 2 $\to$ RX 2
$U_{2pb} = \emptyset, R_{2pb}' = 0$	–	TX 1 $\to$ RX 2
$R_{1c}' = I(U_{1c}; X_2 \vert U_{2c})$	$L_{21} - R_{21} = I(V_{21}; V_{11}, V_{12})$	Binning rate
$R_{1pb}' = I(U_{1pb}; X_2 \vert U_{1c}, U_{2c})$	$L_{22} - R_{22} = I(V_{22}; V_{11}, V_{12})$	Binning rates
X_1	X_2	

union over all the distributions

$$p_{V_{12}} p_{X_1'|V_{12}} p_{V_{21}|X_1',V_{12}} p_{V_{22}|X_1',V_{12}} p_{X_2|X_1',V_{12},V_{21},V_{22}}$$

Correspondence between the random variables and rates. When referring to (52) please note that the index of the primary and cognitive user are reversed with respect to our notation (i.e $1 \to 2$ and vice-versa). Consider the correspondences between the variables of (52, Thm. 1) and those of Theorem 2.6.1 in 2.7.1 to obtain the region $\mathcal{R}_{DMT}^{out}$ defined as the set of rate pairs satisfying

$$
\begin{aligned}
R'_{1c} &= I(U_{1c}; X_2, U_{2c}) && \text{(2.27a)}\\
R'_{1pb} &= I(U_{1pb}; X_2, U_{2c}) && \text{(2.27b)}\\
R_{2pa} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_2; U_{1c}, U_{2c}, X_2) + I(X_2, U_{2c}; U_{1c}) && \text{(2.27c)}\\
R_{2pa} + R_{1c} + R'_{1c} &\leq I(Y_2; X_2, U_{1c}|U_{2c}) + I(X_2; U_{1c}) && \text{(2.27d)}\\
R_{1c} + R'_{1c} &\leq I(Y_2, X_2, U_{2c}; U_{1c}) && \text{(2.27e)}\\
R_{2pa} &\leq I(Y_2; X_2|U_{2c}, U_{1c}) + I(U_{1c}; X_2|U_{2c}) && \text{(2.27f)}\\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}) + I(U_{1pb}, U_{1c}; U_{2c}) && \text{(2.27g)}\\
R_{1c} + R_{1pb} + R'_{1c} + R'_{1pb} &\leq I(Y_1, U_{2c}; U_{1pb}, U_{1c}) + I(U_{1pb}; U_{1c}) && \text{(2.27h)}\\
R_{1pb} + R'_{1pb} &\leq I(Y_1, U_{2c}, U_{1c}; U_{1pb}) && \text{(2.27i)}
\end{aligned}
$$

taken over the union of all distributions

$$p_{U_{2c}} p_{X_2|U_{2c}} p_{U_{1c}|X_2} p_{U_{1pb}|X_2} p_{X_1|X_2,U_{1c},U_{1pb}}. \tag{2.28}$$

Next, we using the correspondences of the table and restrict the fully general input distribution of Theorem 2.6.1 to match the more constrained factorization of (Equation 2.28), obtaining a region $\mathcal{R}_{RTD}^{in} \subseteq \mathcal{R}_{RTD}$ defined as the set of rate tuples satisfying

$$
\begin{aligned}
R'_{1c} &= I(U_{1c}; X_2|U_{2c}) && \text{(2.29a)}\\
R'_{1c} + R'_{1pb} &= I(X_2; U_{1c}, U_{1pb}|U_{2c}) && \text{(2.29b)}\\
R_{2c} + R_{1c} + R_{2pa} + R'_{1c} &\leq I(Y_2; U_{2c}, U_{1c}, X_2) + I(U_{1c}; X_2|U_{2c}) && \text{(2.29c)}\\
R_{2pa} + R_{1c} + R'_{1c} &\leq I(Y_2; U_{1c}, X_2|U_{2c}) + I(U_{1c}; X_2|U_{2c}) && \text{(2.29d)}\\
R_{1c} + R'_{1c} &\leq I(Y_2; U_{1c}|U_{2c}, X_2) + I(U_{1c}; X_2|U_{2c}) && \text{(2.29e)}\\
R_{2pa} &\leq I(Y_2; X_2|U_{2c}, U_{1c}) + I(U_{1c}; X_2|U_{2c}) && \text{(2.29f)}\\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_1; U_{2c}, U_{1c}, U_{1pb}) && \text{(2.29g)}\\
R_{1c} + R_{1pb} + R'_{1c} + R'_{1pb} &\leq I(Y_1; U_{1c}, U_{1pb}|U_{2c}) && \text{(2.29h)}\\
R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{1pb}|U_{2c}, U_{1c}) && \text{(2.29i)}
\end{aligned}
$$

union of all distributions that factor as

$$p_{U_{2c},X_2} p_{U_{1c}|X_2} p_{U_{1pb}|X_2} p_{X_1|X_2,U_{1c},U_{1pb}}.$$

Equation-by-equation comparison. We now show that $\mathcal{R}_{DMT}^{out} \subseteq \mathcal{R}_{RTD}^{in}$ by fixing an input distribution (which are the same for these two regions) and comparing the rate regions equation by equation. We refer to the equation numbers directly, and look at the difference between the corresponding equations in the two new regions.

- Equation 2.29c-Equation 2.29a vs Equation 2.27c-Equation 2.27a: Noting the cancelation / interplay between the binning rates, we see that

$$(Equation\ 2.29c - Equation\ 2.29a) - (Equation\ 2.27d - Equation\ 2.27a) = 0.$$

- Equation 2.29d-Equation 2.29a vs. Equation 2.27d-Equation 2.27a:

$$\begin{aligned}
&(Equation\ 2.29d - Equation\ 2.29a) - (Equation\ 2.27d - Equation\ 2.27a)\\
&\quad = -I(X_2; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = I(U_{2c}; U_{1c}|X_2)\\
&\quad = 0
\end{aligned}$$

- Equation 2.29e-Equation 2.29a vs. Equation 2.27e-Equation 2.27a: again noting the cancelations,

$$(Equation\ 2.29e - Equation\ 2.29a) - (Equation\ 2.27e - Equation\ 2.27a) = 0$$

- Equation 2.29f vs. Equation 2.27f:

$$Equation\ 2.29f - Equation\ 2.27f = 0$$

- Equation 2.29g-Equation 2.29b vs. Equation 2.27g-Equation 2.27b-Equation 2.27a

$$
\begin{aligned}
&(Equation\ 2.29g - Equation\ 2.29b) - (Equation\ 2.27g - Equation\ 2.27b - Equation\ 2.27a)\\
&\quad = -I(X_2; U_{1c}, U_{1pb}|U_{2c})\\
&\qquad -I(U_{1pb}, U_{1c}; U_{2c}) + I(U_{1c}; U_{2c}, X_2) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1pb}, U_{1c}; X_2, U_{2c}) + I(U_{1c}; U_{2c}, X_2) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1pb}; X_2, U_{2c}) - I(U_{1c}; X_2, U_{2c}|U_{1pb}) + I(U_{1c}; U_{2c}, X_2) + I(U_{1pb}; U_{2c}, X_2)\\
&\quad = -I(U_{1c}; X_2, U_{2c}|U_{1pb}) + I(U_{1c}; U_{2c}, X_2)\\
&\quad = -H(U_{1c}|U_{1pb}) + H(U_{1c}|X_2, U_{2c}, U_{1pb}) + H(U_{1c}) - H(U_{1c}|X_2, U_{2c})\\
&\quad = I(U_{1c}; U_{1pb}) > 0
\end{aligned}
$$

where we have used the fact that U_{1c} and U_{1pb} are conditionally independent given (U_{2c}, X_2).

- *Equation* 2.29*h* − *Equation* 2.29*b* vs. *Equation* 2.27*h* − *Equation* 2.27*b* − *Equation* 2.27*a*:

$$
\begin{aligned}
&(Equation\ 2.29h - Equation\ 2.29b) - (Equation\ 2.27h - Equation\ 2.27b - Equation\ 2.27a)\\
&\quad = -I(X_2; U_{1c}, U_{1pb}|U_{2c}) - I(U_{2c}; U_{1c}, U_{1pb}) + I(U_{1pb}; U_{2c}, X_2) - I(U_{1pb}; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(X_2, U_{2c}; U_{1c}, U_{1pb}) + I(U_{1pb}; U_{2c}, X_2) - I(U_{1pb}; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(X_2, U_{2c}; U_{1pb}) - I(U_{1c}; X_2, U_{2c}|U_{1pb}) + I(U_{1pb}; U_{2c}, X_2) - I(U_{1pb}; U_{1c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(U_{1c}; X_2, U_{2c}, U_{1pb}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = -I(U_{1c}; X_2, U_{2c}) - I(U_{1c}; U_{1pb}|X_2, U_{2c}) + I(U_{1c}; X_2, U_{2c})\\
&\quad = 0
\end{aligned}
$$

where we have used the fact that U_{1c} and U_{1pb} are conditionally independent given (U_{2c}, X_2).

- *Equation* 2.29*i* − *Equation* 2.29*b* + *Equation* 2.29*a* vs. *Equation* 2.27*i* − *Equation* 2.27*b*:

$$
\begin{aligned}
&(Equation\ 2.29i - Equation\ 2.29b + Equation\ 2.29a) - (Equation\ 2.27i - Equation\ 2.27b)\\
&= -I(U_{1pb}; X_2|U_{2c}, U_{1c}) - I(U_{1pb}; U_{2c}, U_{1c}) + I(U_{1pb}; X_2, U_{2c})\\
&= -I(U_{1pb}; X_2, U_{2c}, U_{1c}) + I(U_{1pb}; U_{2c}, X_2)\\
&= -I(U_{1pb}; U_{1c}|U_{2c}, X_2)\\
&= 0
\end{aligned}
$$

2.7.2 Containment of the region of Biao et al. in $\mathcal{R}_{RTD}$

The independently derived region in (50, Thm. 2) uses a similar encoding structure as that of $\mathcal{R}_{RTD}$ with two exceptions: a) the binning is done sequentially rather than jointly as in $\mathcal{R}_{RTD}$ leading to binning constraints (43)–(45) in (50, Thm. 2) as opposed to Equation 2.12a–Equation 2.12c in Thm.2.6.1. Notable is that both schemes have adopted a Marton-like binning scheme at the cognitive transmitter, as first introduced in the context of the CIFC in (50). b) While the cognitive messages are rate-split in identical fashions, the primary message is split into 2 parts in (50, Thm. 2) ($R_1 = R_{11} + R_{10}$, note the reversal of indices) while we explicitly split the primary message into three parts $R_2 = R_{2c} + R_{2pa} + R_{2pb}$. We show that the region of (50, Thm.2), denoted as $\mathcal{R}_{CC} \subseteq \mathcal{R}_{RTD}$ in two steps:

- We first show that we may WLOG set $U_{11} = \emptyset$ in (50, Thm.2), creating a new region R'_{CC}.

• We next make a correspondence between our RV's and those of (50, Thm.2) and obtain identical regions.

We note that the primary and cognitive indices are permuted in (50).

We first show that U_{11} in (50, Thm. 2) may be dropped WLOG. Consider the region $\mathcal{R}_{CC}$ of (50, Thm. 2), defined as the union over all distributions $p_{U_{10},U_{11},V_{11},V_{20},V_{22},X_1,X_2}p_{Y_1,Y_2|X_1,X_2}$ of all rate tuples satisfying:

$$
\begin{aligned}
R_1 &\leq I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) && \text{(2.30a)}\\
R_2 &\leq I(Y_2; V_{20}, V_{22}|U_{10}) - I(V_{22}, V_{20}; U_{11}|U_{10}) && \text{(2.30b)}\\
R_1 + R_2 &\leq I(Y_1; V_{11}, U_{11}|V_{20}, U_{10}) + I(Y_2; V_{22}, V_{20}, U_{10})\\
&\quad -I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) && \text{(2.30c)}\\
R_1 + R_2 &\leq I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10})\\
&\quad -I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) && \text{(2.30d)}\\
2R_2 + R_1 &\leq I(Y_1; V_{11}, U_{11}, V_{20}|U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) + I(Y_2; V_{20}, V_{22}, U_{10})\\
&\quad -I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) - I(V_{22}, V_{20}; U_{11}|U_{10}) && \text{(2.30e)}
\end{aligned}
$$

Now let $\mathcal{R}'_{CC}$ be the region obtained by setting $U'_{11} = \emptyset$ and $V'_{11} = (V_{11}, U_{11})$ while keeping all remaining RV's identical. Then $\mathcal{R}'_{CC}$ is the union over all distributions $p_{U_{10},V'_{11},V_{20},V_{22},X_1,X_2}p_{Y_1,Y_2|X_1,X_2}$, with $V'_{11} = (V_{11}, U_{11})$ in $\mathcal{R}_{CC}$, of all rate tuples satisfying:

$$
\begin{aligned}
R_1 &\leq I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) && \text{(2.31a)}\\
R_2 &\leq I(Y_2; V_{20}, V_{22}|U_{10}) && \text{(2.31b)}\\
R_1 + R_2 &\leq I(Y_1; V_{11}, U_{11}|V_{20}, U_{10}) + I(Y_2; V_{22}, V_{20}, U_{10}) \\
&\qquad -I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) && \text{(2.31c)}\\
R_1 + R_2 &\leq I(Y_1; V_{11}, U_{11}, V_{20}, U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) \\
&\qquad -I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) && \text{(2.31d)}\\
2R_2 + R_1 &\leq I(Y_1; V_{11}, U_{11}, V_{20}|U_{10}) + I(Y_2; V_{22}|V_{20}, U_{10}) + I(Y_2; V_{20}, V_{22}, U_{10}) \\
&\qquad -I(V_{22}; U_{11}, V_{11}|V_{20}, U_{10}) && \text{(2.31e)}
\end{aligned}
$$

Comparing the two regions equation by equation, we see that

- Equation 2.30a= Equation 2.31a
- Equation 2.30b < Equation 2.31b as this choice of RV's sets the generally positive mutual information to 0
- Equation 2.30c=Equation 2.31c
- Equation 2.30d=Equation 2.31d

RV, rate of Theorem 2.6.1	RV, rate of (52, Thm. 1)	Comments
U_{2c}, R_{2c}	U_{10}, R_{10}	TX 2 $\rightarrow$ RX 1, RX 2
$X_2 = U_{2c},\ R_{2pa} = 0$	$U_{11} = \emptyset,\ R_{11} = 0$	TX 2 $\rightarrow$ RX 2
U_{1c}, R_{1c}	V_{20}, R_{20}	TX 1 $\rightarrow$ RX 1, RX 2
U_{1pb}, R_{1pb}	V_{22}, R_{22}	TX 1 $\rightarrow$ RX 1
U_{2pb}, R_{2pb}	V_{11}	TX 1 $\rightarrow$ RX 2
R'_{1c}	$L_{20} - R_{20}$	
R'_{1pb}	$L_{22} - R_{22}$	
R'_{2pb}	$L_{11} - R_{11}$	
X_1	X_2	
X_2	X_1	

- Equation 2.30e < Equation 2.31e as this choice of RV's sets the generally positive mutual information to 0

From the previous, we may set $U_{11} = \emptyset$ in the region $\mathcal{R}_{CC}$ of (50, Thm. 2) without loss of generality, obtaining the region $\mathcal{R}'_{CC}$ defined in Equation 2.31a – Equation 2.31e. We show that $\mathcal{R}'_{CC}$ may be obtained from the region $\mathcal{R}_{RTD}$ with the assignment of RV's, rates and binning rates in 2.7.2 .

Evaluating $\mathcal{R}'_{CC}$ defined by Equation 2.31a – Equation 2.31e with the above assignment, translating all RV's into the notation used here, we obtain the region:

$$\begin{aligned}
R'_{1c} &\geq 0 \\
R'_{1pb} + R'_{2pb} &\geq I(U_{1pb}; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} &\leq I(Y_2; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} &\leq I(Y_2; U_{1c}, U_{2pb}|U_{2c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_2; U_{1c}, U_{2c}, U_{2pb}) \\
R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{1pb}|U_{2c}, U_{1c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} &\leq I(Y_1; U_{1pb}, U_{1c}|U_{2c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c})
\end{aligned}$$

Note that we may take binning rate equations $R'_{1c} \geq 0$ and $R'_{1pb} + R'_{2pb} \geq I(U_{1pb}; U_{2pb}|U_{2c}, U_{1c})$ to be equality without loss of generality - the largest region will take $R'_{1c}, R'_{1pb}, R'_{2pb}$ as small as possible. The region $\mathcal{R}_{RTD}$ with $R_{2pa} = 0$

$$\begin{aligned}
R'_{1c} &\geq 0 \\
R'_{1c} + R'_{1pb} &\geq 0 \\
R'_{1c} + R'_{1pb} + R'_{2pb} &\geq I(U_{1pb}; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} &\leq I(Y_2; U_{2pb}|U_{2c}, U_{1c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} &\leq I(Y_2; U_{1c}, U_{2pb}|U_{2c}) \\
R_{2pb} + R'_{2pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_2; U_{1c}, U_{2c}, U_{2pb}) \\
R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{1pb}|U_{2c}, U_{1c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} &\leq I(Y_1; U_{1pb}, U_{1c}|U_{2c}) \\
R_{1pb} + R'_{1pb} + R_{1c} + R'_{1c} + R_{2c} &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c})
\end{aligned}$$

For $R'_{1c} = 0$ these two regions are identical, showing that $\mathcal{R}_{RTD}$ is surely no smaller than $\mathcal{R}_{CC}$. For $R'_{1c} > 0$, $\mathcal{R}_{RTD}$, the binning rates of the region $\mathcal{R}_{RTD}$ are looser than the ones in $\mathcal{R}_{CC}$. This is probably due to the fact that the first one uses joint binning and latter one sequential binning. Therefore $\mathcal{R}_{RTD}$ may produce rates larger than $\mathcal{R}_{CC}$. However, in general, no strict inclusion of $\mathcal{R}_{CC}$ in $\mathcal{R}_{RTD}$ has been shown.

2.7.3 Containment of the region of Jiang et al. in $\mathcal{R}_{RTD}$:

In this scheme the common messages are created independently instead of having the common message from transmitter 1 being superposed to the common message from transmitter 2.

The former choice introduces more rate constraints than the latter and allows us to show inclusion in $\mathcal{R}_{RTD}$.

The region of (51) is expressed as the set of all rate tuples satisfying

$$\begin{aligned}
R'_{22} &\geq I(W_2; V_1|U_1, U_2) && \text{(2.32a)}\\
R'_{11} + R'_{22} &\geq I(W_2; W_1, V_1|U_1, U_2) && \text{(2.32b)}\\
R_{11} + R'_{11} &\leq I(V_1, W_1; Y_1|U_1, U_2) && \text{(2.32c)}\\
R_{12} + R_{11} + R'_{11} &\leq I(U_1, V_1, W_1; Y_1|U_2) && \text{(2.32d)}\\
R_{21} + R_{11} + R'_{11} &\leq I(U_2, V_1, W_1; Y_1|U_1) && \text{(2.32e)}\\
R_{12} + R_{21} + R_{11} + R'_{11} &\leq I(U_1, V_1, W_1, U_2; Y_1) && \text{(2.32f)}\\
R_{22} + R'_{22} &\leq I(W_2; Y_2|U_1, U_2) && \text{(2.32g)}\\
R_{21} + R_{22} + R'_{22} &\leq I(U_2, W_2; Y_2|U_1) && \text{(2.32h)}\\
R_{12} + R_{22} + R'_{22} &\leq I(U_1, W_2; Y_2|U_2) && \text{(2.32i)}\\
R_{12} + R_{21} + R_{22} + R'_{22} &\leq I(U_1, U_2, W_2; Y_2) && \text{(2.32j)}
\end{aligned}$$

union over all the distributions

$$p_{U_1} p_{V_1|U_1} p_{X_1|V_1,U_1} p_{U_2} p_{W_1,W_2|V_1,U_1,U_2} p_{X_0|W_1,W_2,V_1,U_1,U_2} p_{Y_1,Y_2|X_1,X_0}$$

for $(R'_{11}, R'_{22}, R_{11}, R_{12}, R_{21}, R_{22}) \in \mathbb{R}^6_+$.

Following the argument of (53, Appendix D) we can show that WLG we can take X_1 and X_2 to be deterministic functions, so that we can write

$$
\begin{aligned}
R'_{22} &\geq I(W_2;V_1,X_1|U_1,U_2) && \text{(2.33a)}\\
R'_{11}+R'_{22} &\geq I(W_2;W_1,V_1,X_1|U_1,U_2) && \text{(2.33b)}\\
R_{11}+R'_{11} &\leq I(V_1,X_1,W_1;Y_1|U_1,U_2) && \text{(2.33c)}\\
R_{12}+R_{11}+R'_{11} &\leq I(U_1,V_1,X_1,W_1;Y_1|U_2) && \text{(2.33d)}\\
R_{21}+R_{11}+R'_{11} &\leq I(U_2,V_1,X_1,W_1;Y_1|U_1) && \text{(2.33e)}\\
R_{12}+R_{21}+R_{11}+R'_{11} &\leq I(U_1,V_1,X_1W_1,U_2;Y_1) && \text{(2.33f)}\\
R_{22}+R'_{22} &\leq I(W_2;Y_2|U_1,U_2) && \text{(2.33g)}\\
R_{21}+R_{22}+R'_{22} &\leq I(U_2,W_2;Y_2|U_1) && \text{(2.33h)}\\
R_{12}+R_{22}+R'_{22} &\leq I(U_1,W_2;Y_2|U_2) && \text{(2.33i)}\\
R_{12}+R_{21}+R_{22}+R'_{22} &\leq I(U_1,U_2,W_2;Y_2). && \text{(2.33j)}
\end{aligned}
$$

We can now eliminate one RV by noticing that

$$
\begin{aligned}
&p_{U_1}p_{V_1|U_1}p_{X_1|V_1,U_1}p_{U_2}p_{W_1,W_2|V_1,U_1,U_2}p_{X_0|W_1,W_2,V_1,U_1,U_2}p_{Y_1,Y_2|X_1,X_0}\\
&\quad = p_{U_1}p_{V_1,X_1|U_1}p_{U_2}p_{W_1,W_2|V_1,U_1,X_1,U_2}p_{X_0|W_1,W_2,V_1,U_1,X_1,U_2}p_{Y_1,Y_2|X_1,X_0},
\end{aligned}
$$

and setting $V'_1 = [V_1, X_1]$, to obtain the region

$$\begin{aligned}
R'_{22} &\geq I(W_2; V_1'|U_1, U_2) && \text{(2.34a)}\\
R'_{11} + R'_{22} &\geq I(W_2; W_1, V_1'|U_1, U_2) && \text{(2.34b)}\\
R_{11} + R'_{11} &\leq I(V_1', W_1; Y_1|U_1, U_2) && \text{(2.34c)}\\
R_{12} + R_{11} + R'_{11} &\leq I(U_1, V_1', W_1; Y_1|U_2) && \text{(2.34d)}\\
R_{21} + R_{11} + R'_{11} &\leq I(U_2, V_1', W_1; Y_1|U_1) && \text{(2.34e)}\\
R_{12} + R_{21} + R_{11} + R'_{11} &\leq I(U_1, V_1' W_1, U_2; Y_1) && \text{(2.34f)}\\
R_{22} + R'_{22} &\leq I(W_2; Y_2|U_1, U_2) && \text{(2.34g)}\\
R_{21} + R_{22} + R'_{22} &\leq I(U_2, W_2; Y_2|U_1) && \text{(2.34h)}\\
R_{12} + R_{22} + R'_{22} &\leq I(U_1, W_2; Y_2|U_2) && \text{(2.34i)}\\
R_{12} + R_{21} + R_{22} + R'_{22} &\leq I(U_1, U_2, W_2; Y_2) && \text{(2.34j)}
\end{aligned}$$

taken over the union of all distributions of the form

$$p_{U_1} p_{V_1'|U_1} p_{U_2} p_{W_1,W_2|V_1',U_1,U_2} p_{X_0|W_1,W_2,V_1',U_1,U_2} p_{Y_1,Y_2|V_1',X_0}$$

We equate the RV's in the region of (51) with the RV's in Theorem 2.6.1 as in 2.7.3.

RV, rate of Theorem 2.6.1	RV, rate of (52, Thm. 1)	Comments
U_{2c}, R_{2c}	U_1, R_{12}	TX 2 → RX 1, RX 2
X_2, R_{2pa}	V_1', R_{11}'	TX 2 → RX 2
U_{1c}, R_{1c}	U_2, R_{21}	TX 1 → RX 1, RX 2
U_{1pb}, R_{1pb}	W_2, R_{22}	TX 1 → RX 1
$U_{2pb}, R_{2pb} = 0$	W_1	TX 1 → RX 2
R_{1c}'	$L_{20} - R_{20}$	
R_{1pb}'	$L_{11} - R_{11}$	
R_{2pb}'	$L_{22} - R_{22}$	
X_1	X_0	
X_2	X_1	

With the substitution in the achievable rate region of Equation 2.34, we obtain the region

$$R_{1pb}' \geq I(U_{1pb}; X_2 | U_{2c}, U_{1c}) \tag{2.35a}$$

$$R_{1pb}' + R_{2pb}' \geq I(U_{1pb}; U_{2pb}, X_2 | U_{2c}, U_{1c}) \tag{2.35b}$$

$$R_{2pa} + R_{2pb}' \leq I(X_2, U_{2pb}; Y_2 | U_{2c}, U_{1c}) \tag{2.35c}$$

$$R_{2c} + R_{2pa} + R_{2pb}' \leq I(U_{2c}, X_2, U_{2pb}; Y_2 | U_{1c}) \tag{2.35d}$$

$$R_{1c} + R_{2pa} + R_{2pb}' \leq I(U_{1c}, X_2, U_{2pb}; Y_2 | U_{2c}) \tag{2.35e}$$

$$R_{2c} + R_{1c} + R_{2pa} + R_{2pb}' \leq I(U_{2c}, X_2, U_{1c}, U_{1pb}; Y_2) \tag{2.35f}$$

$$R_{1pb} + R_{1pb}' \leq I(U_{1pb}; Y_1 | U_{2c}, U_{1c}) \tag{2.35g}$$

$$R_{1c} + R_{1pb} + R_{1pb}' \leq I(U_{1c}, U_{1pb}; Y_1 | U_{2c}) \tag{2.35h}$$

$$R_{2c} + R_{1pb} + R_{1pb}' \leq I(U_{2c}, U_{1pb}; Y_1 | U_{1c}) \tag{2.35i}$$

$$R_{2c} + R_{1c} + R_{1pb} + R_{1pb}' \leq I(U_{2c}, U_{1c}, U_{1pb}; Y_1) \tag{2.35j}$$

taken over the union of all distributions of the form

$$p_{U_{1c}}p_{U_{2c}}p_{X_2|U_{2c}}p_{U_{1pb},U_{2pb}|U_{1c},U_{2c},X_2}p_{X_1|U_{2c},U_{1c},U_{1pb},U_{2pb}}.$$

Set $R_{2pb} = 0$ and $R'_{1c} = I(U_{1c}; X_2|U_{2c})$ in the achievable scheme of Theorem 2.6.1 and consider the factorization of the remaining RV's as in the scheme of Equation 2.35, that is, according to

$$p_{U_{1c}}p_{U_{2c}}p_{X_2|U_{2c}}p_{U_{1pb},U_{2pb}|U_{1c},U_{2c},X_2}p_{X_1|U_{2c},X_2,U_{1c},U_{1pb},U_{2pb}}.$$

With this factorization of the distributions, we obtain the achievable region

$$\begin{aligned}
R'_{1c} &= I(U_{1c}; X_2|U_{2c}) && \text{(2.36a)}\\
R'_{1pb} &\geq I(U_{1pb}; X_2|U_{2c}, U_{1c}) && \text{(2.36b)}\\
R'_{1pb} + R'_{2pb} &\geq I(U_{1pb}; X_2, U_{2pb}|U_{2c}, U_{1c}) && \text{(2.36c)}\\
R_{2pa} + R'_{2pb} &\leq I(Y_2; X_2, U_{2pb}|U_{2c}, U_{1c}) + I(U_{1c}; X_2|U_{2c}) && \text{(2.36d)}\\
R_{1c} + R_{2pa} + R'_{2pb} &\leq I(Y_2; U_{1c}, X_2, U_{2pb}|U_{2c}) && \text{(2.36e)}\\
R_{2c} + R_{1c} + R_{2pa} + R'_{2pb} &\leq I(Y_2; U_{2pb}, U_{1c}, U_{2c}, X_2) && \text{(2.36f)}\\
R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{1pb}|U_{2c}, U_{1c}) && \text{(2.36g)}\\
R_{1c} + R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{1c}, U_{1pb}|U_{2c}) && \text{(2.36h)}\\
R_{2c} + R_{1c} + R_{1pb} + R'_{1pb} &\leq I(Y_1; U_{2c}, U_{1c}, U_{1pb}) && \text{(2.36i)}
\end{aligned}$$

Note that with this particular factorization we have that $I(U_{1c}; X_2|U_{2c}) = 0$, since X_2 is conditionally independent on U_{1c} given U_{2c}.

We now compare the region of Equation 2.35 and Equation 2.36 for a fixed input distribution, equation by equation:

$$Equation\ 2.36b = Equation\ 2.35a$$

$$Equation\ 2.36c = Equation\ 2.35b$$

$$Equation\ 2.36d = Equation\ 2.35c$$

$$Equation\ 2.36e = Equation\ 2.35e$$

$$Equation\ 2.36f = Equation\ 2.35f$$

$$Equation\ 2.36g = Equation\ 2.35g$$

$$Equation\ 2.36h = Equation\ 2.35h$$

$$Equation\ 2.36i = Equation\ 2.35j$$

clearly Equation 2.35d and Equation 2.35i are extra bounds that further restrict the region in (51) to be smaller than the region of Theorem 2.6.1.

2.8 New capacity results for the DM-CIFC

We now look at the expression of the outer bound (26, Thm. 3.1) to gain an insight on the achievable scheme that is potentially capacity achieving. In particular we look at the expression of the corner points of the outer bound region for a fixed p_{U,X_1,X_2} and try to interpret the RV's as private and common messages to be decoded at the transmitter side. We then consider an achievable scheme inspired by these observation and we show that such scheme achieves capacity for a particular class of channels. This class of channels contains the 'very strong' and the 'very weak' interference regime and thus it is the largest region where capacity is currently known.

The outer bound region of (26, Thm. 3.1) has at most two corner points where both R_1 and R_2 are non zero:

$$(R_1^{out\ (a)}, R_2^{out\ (a)}) = (I(Y_1; X_1|U, X_2), I(Y_2; U, X_2)) \tag{2.37}$$

$$(R_1^{out\ (b)}, R_2^{out\ (b)}) = (I(Y_1; X_1|U, X_2) + I(Y_2; U, X_2) - \Delta, \Delta) \tag{2.38}$$

$$\Delta = [I(Y_2; U, X_2) - I(Y_1; U|X_2)]^+,$$

since

$$\begin{aligned} R_2^{out\ (a)} &= \min\{I(Y_2; U, X_2), I(Y_2; U, X_2) + I(Y_1; X_1|U, X_2)\} = I(Y_2; U, X_2), \\ R_1^{out\ (a)} &= \min\{I(Y_1; X_1|U, X_2), I(Y_1; X_1|X_2)\} = I(Y_1; X_1|U, X_2), \end{aligned}$$

and

$$\begin{aligned} R_2^{out\ (b)} &= \min\{I(Y_2; U, X_2), I(Y_2; U, X_2) + I(Y_1; X_1|U, X_2) - I(Y_1; X_1|X_2)\} \\ &= [I(Y_2; U, X_2) + \min\{0, I(Y_1; X_1|U, X_2) - I(Y_1; X_1, U|X_2)\}]^+ \\ &= [I(Y_2; U, X_2) - I(Y_1; U|X_2)]^+ \triangleq \Delta, \\ R_1^{out\ (b)} &\leq \min\{I(Y_1; X_1|X_2), I(Y_2; U, X_2) + I(Y_1; X_1|U, X_2)\} \\ &= I(Y_1; X_1|U, X_2) + I(Y_2; U, X_2) - \max\{I(Y_2; U, X_2) - I(Y_1; U|X_2), 0\} \\ &= I(Y_1; X_1|U, X_2) + I(Y_2; U, X_2) - \Delta. \end{aligned}$$

Proving the achievability of both these corner points for any p_{U,X_1,X_2} shows capacity by a simple time sharing argument.

We can now look at the corner point expression and try to draw some intuition on the achievable schemes that can possibly achieve these rates.

For the corner point $(R_1^{(a)}, R_2^{(a)})$ we can interpret (U, X_2) as a common message from transmitter 2 to receiver 2 that is also decoded at receiver 1. X_1 is superposed to (U, X_2) since the decoding of X_1 follows the one of (U, X_2) at decoder 2.

The corner point $(R_1^{out\ (b)}, R_2^{out\ (b)})$ has two possible expressions.

If $I(Y_1; U|X_2) \leq I(Y_2; U, X_2)$ we have that

$$(R_1^{out\ (b)'}, R_2^{out\ (b)'}) = (I(Y_1; X_1, U|X_2), I(Y_2; U, X_2) - I(Y_1; U|X_2)) \tag{2.39}$$

which suggests the that X_2 is again the common primary message and the cognitive message is divided a in public and private part, U and X_1 respectively.

If $I(Y_1; U|X_2) > I(Y_2; U, X_2)$ we have

$$(R_1^{out\ (b)"}, R_2^{out\ (b)"}) = (I(Y_2; U, X_2) + I(Y_1; X_1, U|X_2), 0)\,. \tag{2.40}$$

In this case outer bound has only one corner point where both rates are non zero. Note that we can always achieve the point

$$(R_1^{in\ (b)"}, R_2^{in\ (b)"}) = (I(Y_1; X_1, U|X_2), 0)$$

by having transmitter 2 sending a known signal. In this case we have $R_2^{out\ (b)"} = R_2^{in\ (b)"}$ and $R_1^{out\ (b)"} \leq R_1^{in\ (b)"}$ since

$$I(Y_1; X_1, U|X_2) \geq I(Y_2; U, X_2) + I(Y_1; X_1, U|X_2)$$
$$I(Y_1; U|X_2) > I(Y_2; U, X_2).$$

So in this case showing the achievability of the point in Equation 2.38 is sufficient to show capacity.

Guided by these considerations, we consider a scheme that has only the components U_{2c}, U_{1c} and U_{1pb}, that is the primary message ω_2 is common and the cognitive message ω_1 is split in private and public part. With this scheme we are able to extend the capacity results in the "very weak interference" of Theorem 2.4.5 and the "very strong interference" of Theorem 3.4.3. This scheme coincides with the one in (54) which achieves capacity if cognitive receiver has to decode both messages (even with secrecy constraint).

Theorem 2.8.1. ***Capacity in the "better cognitive decoding" regime.***

When the following condition holds

$$I(Y_1; X_2, U) \geq I(Y_2; X_2, U) \qquad \forall p_{X_1,X_2,U} \tag{2.41}$$

the capacity of the DM-CIFC is given by region in Equation 2.1

Proof. Consider the achievable region of Theorem 2.6.1 when setting

$$X_1 = U_{1pb}$$
$$X_2 = U_{2c} = U_{2pb}$$

so that

$$R_2 = R_{2c}$$
$$R_{2pa} = R_{2pb} = 0$$
$$R'_{1c} = R'_{1pb} = R'_{2pb} = 0.$$

In the resulting scheme, the message from transmitter 2 to receiver 2 is all common while the message from transmitter 1 to receiver 1 is split between common and private part. The achievable region of this sub-scheme is:

$$R_2 + R_{1c} \leq I(Y_2; U_{1c}, X_2) \tag{2.42a}$$
$$R_2 + R_{1c} + R_{1pb} \leq I(Y_1; U_{1c}, X_2) \tag{2.42b}$$
$$R_{1c} + R_{1pb} \leq I(Y_1; U_{1c}, X_1|U_{2c}) \tag{2.42c}$$
$$R_{1pb} \leq I(Y_1; X_1|X_2, U_{1c}), \tag{2.42d}$$

by applying the Fourier-Motzkin elimination (55) we obtain the following region

$$R_1 \leq I(Y_1; U_{1c}, X_1|X_2) \tag{2.43a}$$

$$R_2 \leq I(Y_2; U_{1c}, X_2) \tag{2.43b}$$

$$R_1 + R_2 \leq I(Y_2; U_{1c}, X_2) + I(Y_1; X_1|X_2, U_{1c}) \tag{2.43c}$$

$$R_1 + R_2 \leq I(Y_1; X_2, U_{1c}, X_1) \tag{2.43d}$$

When letting $U_{1c} = U$ we have that, Equation 2.1a matches Equation 2.43a , Equation 2.1b matches Equation 2.43b and Equation 2.1c matches Equation 2.43c and Equation 2.43d is redundant when

$$I(Y_1; X_2, X_1, U) \geq I(Y_2; U, X_2) + I(Y_1; X_1|X_2, U)$$

or equivalently

$$I(Y_1; U, X_2) \geq I(Y_2; U, X_2). \tag{2.44}$$

□

We term the condition Equation 2.44 “better cognitive decoding” since decoder 1 has a higher mutual information among the receiver channel output and the RV's U and X_2 than the primary receiver.

Remark 2.8.2. The "better cognitive decoding" in Equation 2.44 is looser than both the "very weak interference" condition of Equation 2.7 and the "very strong interference" condition of Equation 3.6 In fact summing the two equations of the condition in Equation 2.7 we have

$$I(U;Y_1|X_2) + I(X_2;Y_1) \geq I(U;Y_2|X_2) + I(X_2;Y_2) I(Y_1;U,X_2) \geq I(Y_2;U,X_2)$$

which corresponds to condition Equation 2.44. Similarly by summing the two equation of the condition in Equation 3.6 we obtain

$$\begin{aligned}
I(Y_1;X_1,X_2) + I(X_1;Y_2|X_2) &\geq I(Y_2;X_1,X_2) + I(X_1;Y_1|X_2) \Longleftrightarrow \\
I(Y_1;X_1,X_2) - I(X_1;Y_1|X_2) &\geq I(Y_2;X_1,X_2) - I(X_1;Y_2|X_2) \Longleftrightarrow \\
I(Y_1;X_1,X_2,U) - I(X_1;Y_1|X_2) &\geq I(Y_2;X_1,X_2,U) - I(X_1;Y_2|X_2) \Longleftrightarrow \\
I(Y_1;X_2,U) &\geq I(Y_2;X_2,U)
\end{aligned}$$

which again corresponds to condition Equation 2.44.

Since both Equation 2.7 and Equation 3.6 imply the Equation 2.44, we conclude that Equation 2.44 is more general than the previous two.

The scheme that achieves capacity in very weak interference is obtained by setting $U_{1c} = X_2$ so that all the cognitive message is private and the primary message is common. The scheme that achieves capacity in very strong interference is obtained by setting $U_{1c} = X_1$ so that both transmitters send only public messages. The scheme that we use to show the achievability in the "strong cognitive decoding" regime mixes these two schemes by splitting the cognitive message in a public and a private part. This relaxes the strong interference achievability conditions as now the cognitive encoder needs to decode only part of the cognitive message. The scheme also relaxes the very weak achievability condition since it allows the cognitive encoder to decode part of the cognitive message and remove its unwanted effects. For this reason, the resulting achievability conditions are looser than both cases.

2.9 Capacity for the semi-deterministic CIFC

Consider the specific class of DM-CIFC for which the signal received at receiver 1 is a deterministic function of the channel inputs, that is

$$Y_1 = f_1(X_1, X_2). \tag{2.45}$$

This class of channels is termed semi-deterministic CIFC and it was first introduced in (56). In (56) capacity is derived for the case $I(Y_1; X_2) \geq I(Y_2; X_2)$, we extend this result by showing the capacity in the general case. Note that the authors of (56) consider the case where f_1 is invertible. We consider a more general case where such condition is not required.

Theorem 2.9.1. *The capacity of the semi-deterministic cognitive interference channel in Equation 2.45 is*

$$R_1 \leq H(Y_1|X_2) \tag{2.46a}$$

$$R_2 \leq I(Y_2; U, X_2) \tag{2.46b}$$

$$R_1 + R_2 \leq I(Y_2; U, X_2) + H(Y_1|U, X_2) \tag{2.46c}$$

union over all the distributions p_{U,X_1,X_2}.

Proof. Outer bound

The outer bound is obtained from Theorem 2.4.1 , “one auxiliary RV outer bound” , by using the deterministic condition in Equation 2.45.

Achievability

Consider the scheme with only the RV’s X_2, U_{1pb} and U_{2pb}, obtained by setting $U_{2c} = U_{1c} = \emptyset$. The achievable region of Theorem 2.6.1 becomes:

$$R'_{1pb} \geq I(U_{1pb}; X_2) \tag{2.47a}$$

$$R'_{1pb} + R'_{2pb} \geq I(U_{1pb}; U_{2pb}, X_2) \tag{2.47b}$$

$$R_{2pa} + R_{2pb} + R'_{2pb} \leq I(Y_2; U_{2pb}, X_2) \tag{2.47c}$$

$$R_{2pb} + R'_{2pb} \leq I(Y_2; U_{2pb}|X_2) \tag{2.47d}$$

$$R_{1pb} + R'_{1pb} \leq I(Y_1; U_{1pb}), \tag{2.47e}$$

union over all the input distributions $p_{U_{1pb},U_{2pb},X_1,X_2} p_{Y_1,Y_2|X_1,X_2}$.

From the Fourier Motzkin elimination of this sub-scheme, we have that we can set $R_{2pb} = 0$ without loss of generality and that the region can be rewritten as

$$
\begin{aligned}
\mathcal{R}_0(U_{1pb}, U_{2pb}, X_2) \triangleq \{ \quad R_1 &\leq I(Y_1; U_{1pb}) - I(U_{1pb}; X_2) && \text{(2.48a)}\\
R_1 &\leq I(Y_2; U_{2pb}|X_2) - I(U_{1pb}; U_{2pb}|X_2) + I(Y_1; U_{1pb}) \\
&\quad - I(U_{1pb}; X_2) && \text{(2.48b)}\\
R_2 &\leq I(Y_2; U_{2pb}, X_2) && \text{(2.48c)}\\
R_1 + R_2 &\leq I(Y_2; U_{2pb}, X_2) + I(Y_1; U_{1pb}) \\
&\quad - I(U_{1pb}; U_{2pb}, X_2)\} && \text{(2.48d)}
\end{aligned}
$$

union over all the distributions that factor as

$$p_{U_{1pb}, U_{2pb}, X_1, X_2} p_{Y_1, Y_2|X_1, X_2} \tag{2.49}$$

Let now

$$
\begin{aligned}
\mathcal{R}_1(U_{1pb}, U_{2pb}, X_2) \triangleq \{ \quad R_1 &\leq I(Y_1; U_{1pb}) - I(U_{1pb}; X_2) && \text{(2.50a)}\\
R_2 &\leq I(Y_2; U_{2pb}, X_2) && \text{(2.50b)}\\
R_1 + R_2 &\leq I(Y_2; U_{2pb}, X_2) + I(Y_1; U_{1pb}) \\
&\quad - I(U_{1pb}; U_{2pb}, X_2)\} && \text{(2.50c)}
\end{aligned}
$$

and

$$\mathcal{R}_2(U_{1pb}, X_2) \triangleq \{ \quad R_1 \quad \leq I(Y_1; U_{1pb}) - I(U_{1pb}; X_2) \tag{2.51a}$$

$$R_2 \quad \leq I(Y_2; X_2)\}. \tag{2.51b}$$

Notice that

$$\mathcal{R}_2(U_{1pb}, X_2) \subseteq \mathcal{R}_1(U_{1pb}, U_{2pb}, X_2) \subseteq \mathcal{R}_0(U_{1pb}, U_{2pb}, X_2),$$

since

$$\mathcal{R}_2(U_{1pb}, X_2) = \mathcal{R}_1(U_{1pb}, U_{2pb} = X_2, X_2) = \mathcal{R}_0(U_{1pb}, U_{2pb} = X_2, X_2),$$

and $\mathcal{R}_0(U_{1pb}, U_{2pb}, X_2)$ has one less constraint than $\mathcal{R}_1(U_{1pb}, U_{2pb}, X_2)$.

We now want to show that

$$\bigcup_{p_{X_2, U_{1pb}, U_{2pb}}} \mathcal{R}_0 = \bigcup_{p_{X_2, U_{1pb}, U_{2pb}}} \mathcal{R}_1,$$

that is, Equation 2.48b can be removed from the Fourier Motzkin eliminated region of Equation 2.47.

The proof of this equivalence follows the same line as (57, Lemma 2).

For those $P(U_{1pb}, U_{2pb}, X_2)$ such that

$$I(Y_2; U_{2pb}|X_2) - I(U_{1pb}; U_{2pb}|X_2) \geq 0$$

we have

$$\mathcal{R}_1(U_{1pb}, U_{2pb}, X_2) = \mathcal{R}_0(U_{1pb}, U_{2pb}, X_2).$$

For those $P(U_{1pb}, U_{2pb}, X_2)$ such that

$$I(Y_2; U_{2pb}|X_2) - I(U_{1pb}; U_{2pb}|X_2) < 0$$

we have that the point

$$(R_1, R_1) = (I(Y_1; U_{1pb} - I(U_{1pb}; X_2)), I(Y_2; X_2))$$

is achievable in $\mathcal{R}_2$. Such point lies inside $\mathcal{R}_1$ and $\mathcal{R}_0$ and satisfies all the rate constraints in Equation 2.48 but Equation 2.48b. In particular the sum rate Equation 2.48d given by

$$R_1 + R_2 \leq I(Y_2; U_{2pb}, X_2) + I(Y_1; U_{1pb}) - I(U_{1pb}; U_{2pb}, X_2),$$

which implies

$$R_2 \leq I(Y_2; X_2)$$

since

$$\begin{aligned} R_2 &\leq I(Y_2; U_{2pb}, X_2) + I(Y_1; U_{1pb}) - I(U_{1pb}; U_{2pb}, X_2) - R_1 \\ &= I(Y_2; X_2) + I(Y_2; U_{2pb}|X_2) I(U_{1pb}; U_{2pb}|X_2) \\ &\leq I(Y_2; X_2). \end{aligned}$$

Using time sharing we can show the achievability of all the region $\mathcal{R}_1 \cap \mathcal{R}_0$. This means that these rate points are are not in $\mathcal{R}_0(U_{1pb}, U_{2pb}, X_2)$ are in $\mathcal{R}_2(U_{1pb}, X_2)$. But since $\mathcal{R}_2(U_{1pb}, X_2)$ is special case of $\mathcal{R}_0(U_{1pb}, U_{2pb}, X_2)$, we conclude that

$$\mathcal{R}_1(U_{1pb}, U_{2pb}, X_2) = \mathcal{R}_0(U_{1pb}, U_{2pb}, X_2),$$

This means is that decoder 2 must not decode U_{2pb} if that imposes a more stringent rate constraint than the decoding of U_{1pb} at the intended decoder 1. For this reason U_{2pb} can be chosen so that $U_{2pb} = X_2$ without loss of generality in such case.

This shows that $\mathcal{R}_1$ is achievable and thus concludes the achievability proof

$\square$

Remark 2.9.2. The achievable scheme of Equation 2.47 cannot be obtained as a special case of any previously known achievable scheme but (51). The RV U_{2pb}, as a broadcasted private primary message from transmitter 1, appears in (50) as well. In this case in is possible to reobtain the scheme of Equation 2.47 with a specific choice of the RV's. Here the same message is embedded in U_{2pb} and the private primary message, this perform strictly worse than using only U_{2pb}.

2.10 Capacity for the deterministic CIFC

In the deterministic CIFC both outputs are deterministic functions of the channel inputs, that is

$$Y_1 = Y_1(X_1, X_2)$$
$$Y_2 = Y_2(X_1, X_2) \tag{2.52}$$

This class of channels is a subclass of the semi-deterministic CIFC of Section 2.9, so we already have capacity for this case. Here we now show the achievability of the outer bound of Theorem 2.5.1 when letting $Y_2' = Y_2$, instead of the outer bound of Theorem 2.4.1 , "one auxiliary RV outer bound". The achievability also differs since only one scheme is needed to achieve the outer bound.

Theorem 2.10.1. *The capacity of the deterministic cognitive interference channel is*

$$R_1 \leq H(Y_1|X_2) \tag{2.53a}$$
$$R_2 \leq H(Y_2) \tag{2.53b}$$
$$R_1 + R_2 \leq H(Y_2) + H(Y_1|Y_2, X_2) \tag{2.53c}$$

union over all the distributions p_{X_1,X_2}.

Proof. Outer bound

The outer bound is obtained from Theorem 2.5.1 by the deterministic conditions in Equation 2.52.

Achievability

Consider the scheme in Equation 2.50 and let $U_{1pb} = Y_1$, $U_{2pb} = Y_2$ to achieve the region

$$
\begin{aligned}
R_1 &\leq H(Y_1|X_2) && \text{(2.54a)}\\
R_2 &\leq H(Y_2) && \text{(2.54b)}\\
R_1 + R_2 &\leq H(Y_2; U, X_2) + H(Y_1|Y_2, X_2) && \text{(2.54c)}
\end{aligned}
$$

which corresponds to the outer bound in Equation 2.53. □

2.11 Examples

The scheme that achieve capacity in the deterministic and semi-deterministic CIFC uses the RV U_{2pb} to perform Gel'an Pinsker binning to achieve the most general distribution among (X_2, U_{1pb}, U_{2pb}), but it carries no message. This feature of the capacity achieving scheme does not provide a clear intuition on the role of this RV. For this reason we present two examples of deterministic channels where the encoders can choose their respective codebooks in a way that allows binning of the interference without rate splitting. To make these examples more

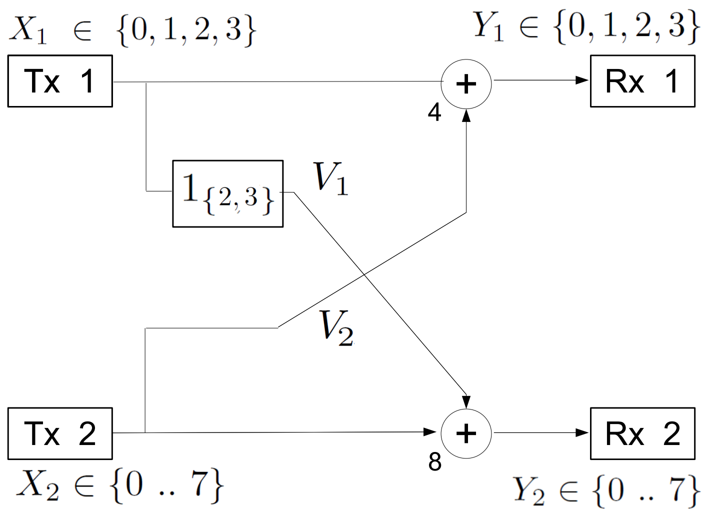

Figure 4. The "asymmetric clipper" of Section 2.11.1.

interesting we choose them so that they do not fall into the category of the "very strong interference regime" of Theorem 3.4.3 that in deterministic case reduces to

$$\begin{aligned} H(Y_1|X_2) &\leq H(Y_2|X_2) \\ H(Y_2) &\leq H(Y_1) \qquad \forall p_{X_1,X_2} \end{aligned} \tag{2.55}$$

Unfortunately, checking for the "very weak interference condition" of Theorem 2.4.5 is not possible as no cardinality bound on U are available.

2.11.1 Example I: the "Asymmetric Clipper"

Consider the channel in Figure 4. The input and output alphabets are $\mathcal{X}_1 = \mathcal{Y}_1 = \{0,1,2,3\}$ and $\mathcal{X}_2 = \mathcal{Y}_2 = \{0,1,2,3,4,5,6,7\}$ and the input/output relationships are

$$Y_1 = X_1 \oplus_4 X_2, \tag{2.56}$$

$$Y_2 = 1_{\{2,3\}}(X_1) \oplus_8 + X_2, \tag{2.57}$$

where $1_A(x) = 1$ if $x \in A$ and zero otherwise and $\oplus_N$ is the addition operation over the ring $\{1...N\}$. Also let $\mathcal{U}(\mathcal{S})$ be the uniform distribution over the set $\mathcal{S}$.

First we show that the channel in Equation 2.57 does not fall in the "very strong interference" class.

Consider the input distribution:

$$\begin{aligned} &X_2 \sim \mathcal{U}(1) \quad \Longrightarrow \quad P[X_1 = 0] = 1, \\ &X_2 \sim \mathcal{U}(\mathcal{X}_2). \end{aligned}$$

For this input distribution, we have $Y_1 \sim \mathcal{U}(\mathcal{Y}_1)$ and $Y_2 \sim \mathcal{U}(\mathcal{Y}_2)$, so that

$$H(Y_2) = \log(|\mathcal{Y}_2|) = 3 > 2 = \log(|\mathcal{Y}_1|) = H(Y_1).$$

which contradicts the condition in Equation 2.55 for the "very strong interference" condition to hold.

For this channel we have:

$$H(Y_1|X_2) \leq H(Y_1) \leq \log(|\mathcal{Y}_1|) = 2$$

$$H(Y_2) \leq \log(|\mathcal{Y}_2|) = 3$$

$$H(Y_1|X_2, Y_2) \leq H(X_1|1_{\{2,3\}}(X_1)) \leq 1.$$

where the last bound follows from the multiplicity of the solutions of an addition in a Galois field. This shows that the outer bound in Theorem 2.10.1 is included in

$$R_1 \leq 2 \tag{2.58a}$$

$$R_2 \leq 3 \tag{2.58b}$$

$$R_1 + R_2 \leq 4. \tag{2.58c}$$

We now show that the region in Equation 2.58 indeed corresponds to the Theorem 2.10.1 when considering the union over all input distributions. The corner point $(R_1, R_2) = (1, 3)$ in Equation 2.58 is obtained in Theorem 2.10.1 with the input distribution:

$$X_1 \sim \mathcal{U}(\{0, 1\})$$

$$X_2 \sim \mathcal{U}(\mathcal{X}_2).$$

The corner point $(R_1, R_2) = (2, 2)$ in Equation 2.58 is obtained in Theorem 2.10.1 considering the input distribution:

$$X_1 \sim \mathcal{U}(\mathcal{X}_1)$$

$$X_2 \sim \mathcal{U}(\mathcal{X}_2).$$

Time sharing shows that the region of Equation 2.58 and the region Theorem 2.10.1 indeed coincide.

We next show the achievability of the corner point $(R_1, R_2) = (1, 3)$: consider the following strategy:

- transmitter 2 sends symbols from $\mathcal{X}_2 = \{0...7\}$ with uniform probability,
- transmitter 1 transmits $[x_1 - x_2]_2$ (where the inverse of the difference operation is taken over the ring $\mathcal{G}_2$);
- receiver 1 decodes $\widehat{w}_1 = \lfloor \frac{y_2}{2} \rfloor$;
- receiver 2 decodes $\widehat{w}_2 = y_2$.

It can be verified by the inspection of Table III that the rate pair $(R_1, R_2) = (1, 3)$ is indeed achievable.

Now we show the achievability of the corner point $(R_1, R_2) = (2, 2)$: consider the following strategy:

- transmitter 2 sends symbols from $x_2 \in \{0, 2, 4, 6\}$ with uniform probability;
- transmitter 1 transmits $[x_1 - x_2]_4$ (where the inverse of the difference operation is taken over the ring $\mathcal{G}_4$);

TABLE III

ACHIEVABILITY FOR $(R_1, R_2) = (1, 3)$ IN EXAMPLE I.

	ω_1	ω_2	x_1	x_2	y_1	y_2	$\hat{\omega}_1$	$\hat{\omega}_2$
0	0	0	0	0	0	0	0	0
1	0	1	1	0	2	2	1	0
2	0	2	0	2	2	2	2	0
3	0	3	1	3	2	0	3	0
4	0	4	0	4	0	0	4	0
5	0	5	1	5	0	2	5	0
6	0	6	0	6	0	2	6	0
7	0	7	1	7	0	0	7	0
8	1	0	1	0	0	1	0	1
9	1	1	0	0	2	1	1	1
10	1	2	1	2	2	3	2	1
11	1	3	0	3	2	3	3	1
12	1	4	1	4	0	1	4	1
13	1	5	0	5	0	1	5	1
14	1	6	1	6	0	3	6	1
15	1	7	0	7	0	3	7	1

- receiver 1 decodes $\widehat{w}_1 = y_1$;
- receiver 2 decodes $\widehat{w}_2 = \lfloor \frac{y_2}{2} \rfloor$.

It can be verified by the inspection of Table IV that the rate pair $(R_1, R_2) = (2, 2)$ is indeed achievable.

TABLE IV

ACHIEVABILITY TABLE FOR THE RATE POINT $(R_1, R_2) = (2, 2)$ IN EXAMPLE I.

	ω_1	ω_2	x_1	x_2	y_1	y_2	$\hat{\omega}_1$	$\hat{\omega}_2$
0	0	0	0	0	0	0	0	0
1	0	1	2	2	0	3	0	1
2	0	2	0	4	0	4	0	2
3	0	3	2	6	0	7	0	3
4	1	0	1	0	1	0	1	0
5	1	1	3	2	1	3	1	1
6	1	2	1	4	1	4	1	2
7	1	3	3	6	1	7	1	3
8	2	0	2	0	2	0	2	0
9	2	1	0	2	2	2	2	1
10	2	2	2	4	2	5	2	2
11	2	3	0	6	2	6	2	3
12	3	0	3	0	3	1	3	0
13	3	1	1	2	3	2	3	1
14	3	2	3	4	3	5	3	2
15	3	3	1	6	3	6	3	3

In this example we see how the two senders jointly design the codebook to achieve the outer bound and in particular how the cognitive transmitter 1 adapts its strategy to the transmissions from the primary pair so avoid interfering with it .

In achieving the point $(R_1, R_2) = (1, 3)$, transmitter 2 sends as in a point to point channel to achieve its maximum rate over the primary link. Transmitter 1 chooses its codewords so not to interfere with the primary transmission. Only two codewords do not interfere: it alternatively pick one of the two codewords to produce the desired channel output. For example when the primary message is sending $\omega_2 = 0$ (line 0 and 8 in Table III) transmitter 1 can send either 1 or 2 without creating interference at receiver 2. On the other hand, these two values produce a different output at receiver 1, allowing the transmission of 1 bit.

In achieving the point $(R_1, R_2) = (2, 2)$, the primary receiver picks its codewords so as to tolerate 1 unit of interference. Transmitter 1 again chooses its input codewords in order to create at most 1 unit of interference at the primary decoder. By adapting its transmission to the primary symbol, the cognitive transmitter is able to always find four such codewords. It interesting to notice the tension at transmitter 1 between the interference it creates at the primary decoder and its own rate. There is an optimal trade off between these two quantities that is achieved by carefully picking the codewords at the primary transmitter. For example when the primary receiver is sending $\omega_2 = 0$, lines $0, 4, 8$ and 12, transmitter 1 can send $x_1 \in \{0, 1, 2, 3\}$ and create at most 1 bit of interference at receiver 2. Each of these for values produces a different output at receiver 1], thus allowing the transmission of 2 bits.

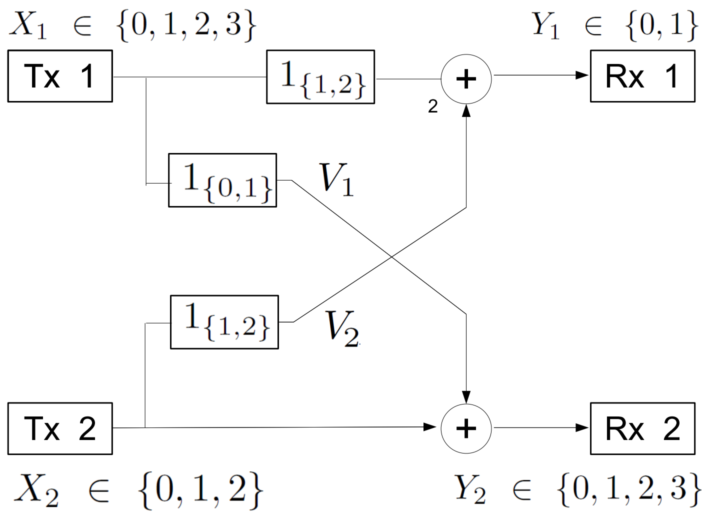

Figure 5. The "symmetric clipper" of Section 2.11.2

2.11.2 Example II: the "Symmetric Clipper"

Consider the now channel in Figure 5. The channel input and output alphabets are $\mathcal{X}_1 = \{0,1,2,3\} = \mathcal{Y}_2$, $\mathcal{X}_2 \in \{0,1,2\}$, and $\mathcal{Y}_1 = \{0,1\}$. The input/output relationships are:

$$Y_1 = 1_{\{1,2\}}(X_1) \oplus_2 1_{\{1,2\}}(X_2)$$

$$Y_2 = 1_{\{0,1\}}(X_1) \oplus X_2$$

TABLE V

THE INPUT DISTRIBUTION FOR EXAMPLE II

X_2 \ X_1	1	2	3	4	
0	1/8	1/8	1/8	1/8	1/2
1	1/8	1/8	0	0	1/4
2	1/8	1/8	0	0	1/4
	3/8	3/8	1/8	1/8	

Consider the input distribution: Consider the input distribution:

$$P[X_1 = 3] = 1,$$

$$X_2 \sim \mathcal{U}(\{1, 2\}),$$

in this case $H(Y_1) = 0$ and $H(Y_2) = 1$.This shows that there exists at least one input distribution for which $H(Y_2) > H(Y_1)$ and thus this channel is not is the "very strong interference" regime. The outer bound of Theorem 2.10.1 is achieved here by a single input distribution p_{X_1,X_2}: consider the distribution in Table V. This distribution produce $H(Y_1) = 1 = \log_2(|\mathcal{Y}_1|)$ and $H(Y_2) = 2 = \log(|\mathcal{Y}_2|)$ and clearly no possible larger outer bound can exists given the output cardinality. We therefore conclude that the region of Theorem 2.10.1 can be rewritten as:

$$R_1 \leq 1$$

$$R_2 \leq 2.$$

TABLE VI

ACHIEVABILITY TABLE FOR THE RATE POINT $(R_1, R_2) = (1, 2)$ IN EXAMPLE II.

	ω_1	ω_2	x_1	x_2	v_1	v_2	y_1	y_2
0	0	0	3	0	0	0	0	0
1	0	1	0	0	1	0	0	1
2	0	2	1	1	1	1	0	2
3	0	3	1	2	1	1	0	3
4	1	0	2	0	0	0	1	0
5	1	1	1	0	1	0	1	1
6	1	2	0	1	1	1	1	2
7	1	3	0	2	1	1	1	3

This region can be shown achievable using the transmission scheme described in Table VI. The decoding is simply $\hat{\omega}_i = Y_i, \quad i \in \{1, 2\}$. This transmission scheme achieves the proposed outer bound, thus showing capacity. The transmission scheme can be described as follows

- encoder 2 transmits $[x_2 - 1]^+$,
- encoder 1 transmits the value X_1 that simultaneously makes $Y_1 = \omega_1$ and $Y_2 = \omega_2$. For each ω_1 and ω_2 such value always exists because X_2 has only three possible values,
- receivers 1 and 2 decode $\hat{\omega}_1 = Y_1$ and $\hat{\omega}_2 = Y_2$.

This example is particularly interesting since both decoders obtain the transmitted symbol without suffering any interference from the other user. Here cognition allows the simultaneous cancelation of the interference at both decoders. Encoder 2 has only three codewords and relies on transmitter 1 to achieve its full rate of $R_2 = 2$. In fact encoder 1 is able to design is codebooks to transmit two codewords for its decoder and still contribute to the rate of primary user by making the codewords corresponding to $\omega_2 = \{2, 3\}$ distinguishable at the cognitive decoder.

This feature of the capacity achieving scheme is really intriguing. The primary transmitter needs the support of the cognitive transmitter to achieve $R_2 = 2$ since its input alphabet has cardinality three. The transmitters optimally design their codebooks so to make the effect X_1 on both outputs the desired one.

For example consider the transmission of $\omega_2 = 2$ or 3, lines $2, 3, 6$ and 7. In this case transmitter 1 sends $x_1 = 0$ or $x_1 = 1$ to simultaneously influence both channel output so that both decoders receive the desired symbol. This simultaneous cancelation can be implemental given the channel deterministic nature and the extra knowledge at the cognitive transmitter.

CHAPTER 3

THE GAUSSIAN COGNITIVE INTERFERENCE CHANNEL

The content of this chapter appears in the Proceedings of ITW2010 in Cairo, Allerton conference 2010, ICC2011 and is submitted to the IEEE Transaction of Information Theory

3.1 Main Contributions

In this chapter we focus on the Gaussian cognitive interference channel in a comprehensive and comparative manner. In particular, our main contributions are:

1. **W**e evaluate the outer bound of Th. 2.5.1 for the Gaussian cognitive interference channel. We show that it unifies the previously proposed outer bounds for the "weak interference" and the "strong interference" regimes of (26) and (25), respectively.

2. **W**e derive a new outer bound based on the broadcast channel and inspired by (14). The capacity region of the Gaussian MIMO (multi input multi output antenna) broadcast channel with degraded message sets is an outer bound for a channel in "strong interference". Interestingly, we show that the new bound may be strictly tighter than the "strong interference" outer bound of (25).

3. **D**erive new outer bounds by transformation / inclusion into channels with known capacity. We determine the conditions under which the capacity region of a Gaussian channel is

contained in that of a channel with known capacity. The capacity of the latter channel thus provides an outer bound for the former.

4. **W**e specialize the largest known inner bound of Th. 2.6.1 to the Gaussian channel. We utilize it as a unified framework to derive and compare various achievable schemes in this and prior work.

5. **W**e prove a new capacity result for the "primary decodes cognitive" regime. This regime is a subset of the "strong interference" regime that is not included in the "very strong interference" regime for which capacity was known (24). In this regime capacity is achieved by having the primary receiver decode the message of the cognitive user in addition to its own message.

6. **W**e prove a new capacity result for the S-channel, a channel in which the primary transmission does not interfere with the cognitive receiver. For this channel we show the achievability of the outer bound based on the capacity of the broadcast channel with degraded message sets.

7. **W**e show capacity to within half a bit/s/Hz per real dimension and to within a factor two regardless of channel parameters. These two results characterize the capacity region of the Gaussian channel both at high and low SNR, respectively. To this end, we use a transmission scheme inspired by the capacity achieving scheme for the semi-deterministic cognitive interference channel of Th. 2.9.1 where capacity is achieved by having the cognitive transmitter perform partial interference pre-coding for both decoders. The

multiplicative gap is shown by using a simple time sharing argument between achievable points.

8. **W**e provide insights on the capacity region of the Gaussian channel for the regimes in which capacity is still unknown. We do so by showing that very simple transmission strategies can achieve capacity to within a constant gap for large sets of parameters. We conclude by showing that a constant gap result may alternatively be proved by trading off interference pre-coding at the cognitive encoder and interference decoding at the primary receiver.

3.2 Organization

The rest of the chapter is organized as follows. Section 3.3 formally defines the cognitive interference channel model and summarizes known results for the Gaussian channel. Section 3.5 presents new outer bounds for the Gaussian channel. Section 3.6 lists the achievable schemes used in the rest of the chapter and shows how they may be obtained from the largest known inner bound of Chapter 2. Section 3.7 proves the two new capacity results. Section 3.8 characterizes the capacity of the Gaussian channel to within half a bit/s/Hz per real dimension and to within a factor two. Section 3.9 shows some relevant numerical results.

3.3 Gaussian channel model and known results

3.3.1 Gaussian CIFC

A Gaussian CIFC (G-CIFC) in *standard form* is described by the input/output relationship

$$Y_1 = X_1 + aX_2 + Z_1,$$

$$Y_2 = |b|X_1 + X_2 + Z_2,$$

where the channel gains a and b are complex-valued, constant, and known to all terminals, the channel inputs are subject to the power constraint

$$\mathbb{E}[|X_i|^2] \leq P_i, \qquad P_i \in \mathbb{R}^+, \qquad i \in \{1, 2\},$$

and the channel noise $Z_i \sim \mathcal{N}_{\mathbb{C}}(0, 1)$, $i \in \{1, 2\}$. Since the capacity only depends on the output conditional marginals, the correlation coefficient of Z_1 and Z_2 is irrelevant. A graphical representation of a G-CIFC is found in Figure 6.

A G-CIFC is said to be a:

- **Z**-channel if $|b| = 0$; we refer to it as a Z-G-CIFC. In this case the primary decoder does not experience interference from the cognitive transmitter. Capacity is trivially given by

$$R_1 \leq \mathcal{C}(P_1), \quad R_2 \leq \mathcal{C}(P_2).$$

- **S**-channel if $a = 0$; we refer to it as a S-G-CIFC. In this channel the cognitive decoder does not experience interference from the primary transmitter. For this channel capacity is only known for $|b| \leq 1$ (26).

- **D**egraded channel if $a|b| = 1$. In this case one channel output is a degraded version of the other. In particular, for $|b| > 1$, Y_1 is a degraded version of Y_2 since

$$Y_1 = X_1 + \frac{1}{|b|}X_2 + Z_1 \sim \frac{1}{|b|}Y_2 + Z_0,$$

for $Z_0 \sim \mathcal{N}_{\mathbb{C}}(0, |b|^2 - 1)$ independent of everything else. Similarly, when $|b| \leq 1$, Y_2 is a degraded version of Y_1. Capacity is known in the case $|b| \leq 1$ (26).

3.4 The G-CIFC in standard form

A general G-CIFC has outputs

$$\widetilde{Y}_1 = h_{11}\widetilde{X}_1 + h_{12}\widetilde{X}_2 + \widetilde{Z}_1$$

$$\widetilde{Y}_2 = h_{21}\widetilde{X}_1 + h_{22}\widetilde{X}_2 + \widetilde{Z}_2$$

where

$$\widetilde{Z}_i \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_i^2), \qquad \sigma_i^2 > 0, \qquad i \in \{1, 2\},$$

and the inputs are subject to the power constraint

$$\mathbb{E}[|\widetilde{X}_i|^2] \leq \widetilde{P}_i, \qquad \widetilde{P}_i \geq 0, \qquad i \in \{1, 2\}.$$

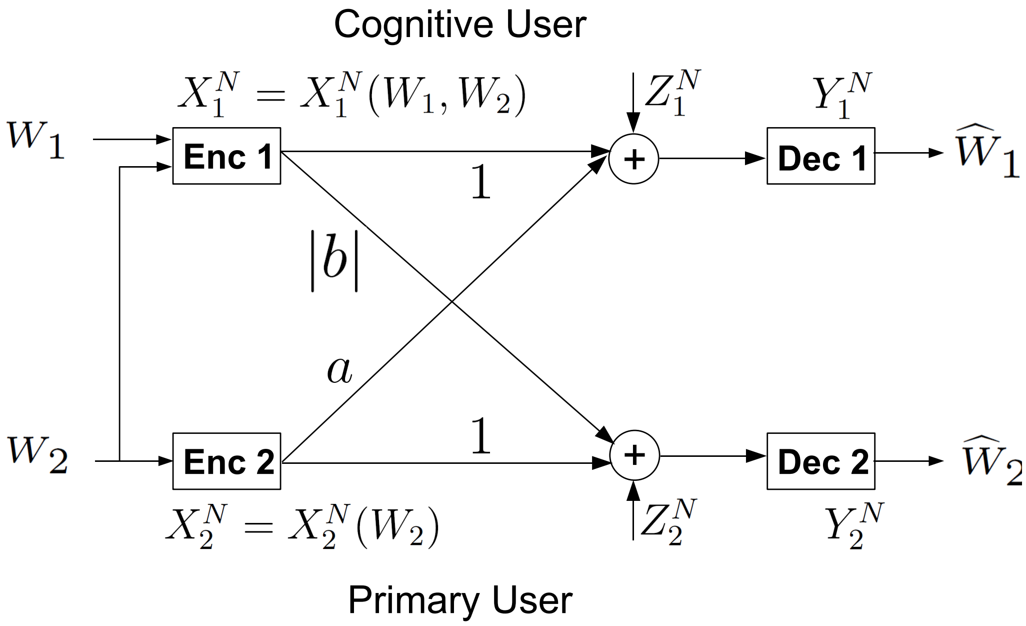

Figure 6. The Gaussian cognitive interference channel (G-CIFC).

When $h_{11} \neq 0$ and $h_{22} \neq 0$, we may scale each channel output by the standard deviation (assumed strictly positive) of the corresponding additive Gaussian noise and change the phase as

$$\begin{aligned}
Y_1 &\triangleq \frac{\widetilde{Y}_1}{\sigma_1} \\
Y_2 &\triangleq \frac{\widetilde{Y}_2}{\sigma_2} \mathrm{e}^{\mathrm{j}(\angle h_{11} - \angle h_{12})} \\
X_1 &\triangleq \frac{h_{11}}{\sigma_1} \widetilde{X}_1 \quad \text{such that} \quad \mathbb{E}[|X_1|^2] \leq P_1 \triangleq \frac{|h_{11}|^2}{\sigma_1^2} \widetilde{P}_1 \\
X_2 &\triangleq \frac{h_{22}}{\sigma_2} \mathrm{e}^{\mathrm{j}(\angle h_{11} - \angle h_{12})} \widetilde{X}_2 \quad \text{such that} \quad \mathbb{E}[|X_2|^2] \leq P_2 \triangleq \frac{|h_{22}|^2}{\sigma_2^2} \widetilde{P}_2 \\
a &\triangleq \frac{h_{12}}{\sigma_1} \frac{\sigma_2}{h_{22}} \mathrm{e}^{\mathrm{j}(-\angle h_{11} + \angle h_{12})} \in \mathbb{C} \\
b &\triangleq \frac{|h_{21}|}{\sigma_2} \frac{\sigma_1}{|h_{11}|} \in \mathbb{R}^+,
\end{aligned} \tag{3.1}$$

to obtain the equivalent channel outputs have additive noise of unit variance, unit gain on the direct link, as claimed in Section 3.3.1. To remind the reader that b is always real-valued and non-negative we use the notation $|b|$.

When $h_{22} = 0$, transmitter 2 can only create interference at receiver 1 and thus the channel reduces to a BC where the cognitive transmitter is sending both messages to both receivers. When $h_{22} = 0$ in Equation 3.1, we have $a = \infty$ and $P_2 = 0$ corresponds to the scenario above; the same in not true when $h_{11} = 0$.

If $h_{11} = 0$, the channel reduces to a MISO, point-to-point channel since decoder 1 can only receive interference from transmitter 2. For $h_{11} = 0$ the transformation in Equation 3.1 does not yield a MISO channel, since in this case $P_1 = 0$ and $b = \infty$. In (58, Sec. II.B), this fact is overlooked and the transformation in Equation 3.1 is considered to be without loss of generality.

Note that the equivalent transformation in standard form for a classical interference channel does not require $h_{11} > 0$, since the transmitters cannot cooperate.

3.4.1 Known results for the G-CIFC

The capacity of the G-IFC is not known in general. However several capacity results exist, as summarized next.

Theorem 3.4.1. *"Weak interference" capacity of (26, Lemma 3.6) and (27, Th. 4.1).* *If*

$$|b| \leq 1, \quad \text{(the "weak interference" regime/condition)} \tag{3.2}$$

the capacity of the G-CIFC is:

$$R_1 \le \mathcal{C}(\alpha P_1), \tag{3.3a}$$

$$R_2 \le \mathcal{C}\left(|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right) - \mathcal{C}(|b|^2 \alpha P_1), \tag{3.3b}$$

taken over the union of all $\alpha \in [0, 1]$.

Theorem 3.4.2. ***"Strong interference" outer bound of (25, Th. 4).*** *When*

$$|b| > 1, \quad \textit{(the "strong interference" regime/condition)} \tag{3.4}$$

the capacity region of the G-CIFC is included in the region $\mathcal{R}^{(SI)}$ *defined as:*

$$R_1 \le \mathcal{C}(\alpha P_1), \tag{3.5a}$$

$$R_1 + R_2 \le \mathcal{C}\left(|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right), \tag{3.5b}$$

taken over the union of all $\alpha \in [0, 1]$.

Theorem 3.4.3. ***"Very strong interference" capacity of (24, Th. 6)*** *extended to complex-valued channels. When*

$$(|a|^2 - 1)P_2 - (|b|^2 - 1)P_1 - 2\big|a - |b|\big|\sqrt{P_1 P_2} \ge 0,$$

$$\textit{and } |b| > 1 \textit{ ("very strong interference" regime/condition)} \tag{3.6}$$

the outer bound $\mathcal{R}^{(SI)}$ of Th. 3.4.2 is tight.

Proof. For a complex-valued G-CIFC with $|b| > 1$, the outer bound of Th. 3.4.2 is achievable by the superposition-only (scheme (D) of Section 3.6.4) if $I(Y_1; X_1, X_2) \geq I(Y_2; X_1, X_2)$ for all input distributions (24), that is, if

$$\begin{aligned} \mathbb{E}[|Y_1|^2] - \mathbb{E}[|Y_2|^2] &= (|a|^2 - 1)P_2 - (|b|^2 - 1)P_1 + \\ &+ 2\sqrt{P_1 P_2}(\mathrm{Re}\{a^* \rho\} - |b|\mathrm{Re}\{\rho\}) \geq 0, \quad \forall |\rho| \leq 1. \end{aligned} \tag{3.7}$$

Let $\rho = |\rho| e^{j\phi_\rho}$ and $a = |a| e^{j\phi_a}$. We have

$$\begin{aligned} \mathrm{Re}\{a^* \rho\} - |b|\mathrm{Re}\{\rho\} &= |\rho||a| \cos(\phi_\rho - \phi_a) - |\rho||b| \cos(\phi_\rho) \\ &= |\rho| \Big[|a| \cos(\phi_a) - |b| \Big] \cos(\phi_\rho) + |\rho| \Big[|a| \sin(\phi_a) \Big] \sin(\phi_\rho) \\ &= |\rho| \sqrt{(\Big[|a| \cos(\phi_a) - |b| \Big])^2 + \Big[|a| \sin(\phi_a) \Big]^2} \cos(\phi) \\ &= \Big| a - |b| \Big| \cdot |\rho| \cos(\phi), \end{aligned}$$

for some angle ϕ. The condition in Equation 3.7 is thus verified for all $|\rho| \cos(\phi) \in [-1, +1]$ if it is verified for $|\rho| \cos(\phi) = -1$. $\square$

A plot of the capacity results of Th. 3.4.1 and Th. 3.4.3 for $a \in \mathbb{R}$ and $P_1 = P_2$ is depicted in Figure 7. The channel gains a and $|b|$ for which capacity is known are shaded, while those for which capacity is unknown are white.

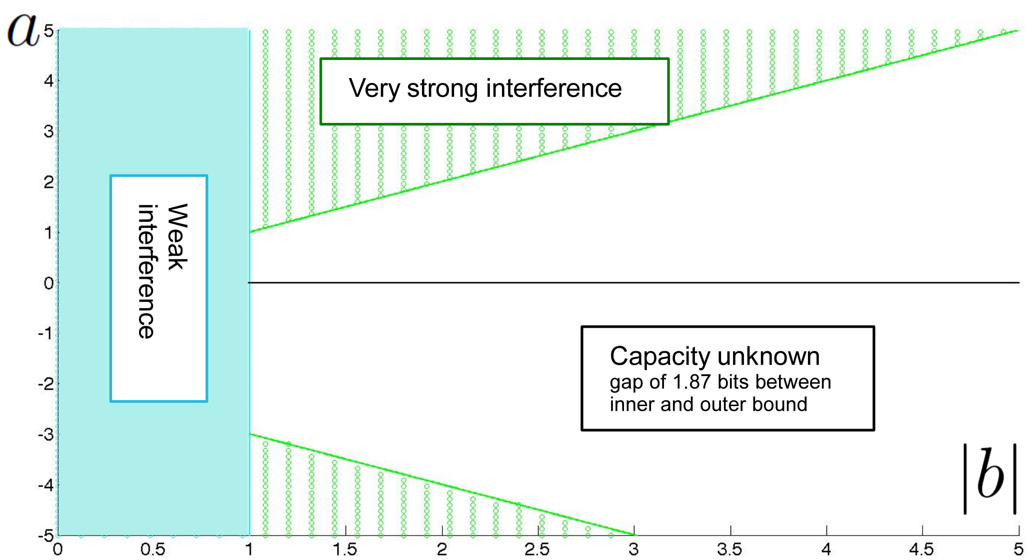

Figure 7. A representation of the capacity results Th. 3.4.1 and Th. 3.4.3 for $P_1 = P_2$ and $(a, |b|) \in [-5, 5] \times [0, 5]$. The regions for which capacity is known are shaded, while those for which capacity is unknown are white.

3.5 Outer bounds

In this section we prove several outer bounds:

1. First we evaluate the outer bound of Th. 2.5.1 for the Gaussian channel and show that it coincides with the outer bounds of Th. 3.4.1 and Th. 3.4.2 in "weak" and "strong interference" respectively.

2. Then we tighten it by using the observation of (14) that the capacity region of a G-CIFC is included into the capacity region of the Gaussian MIMO BC obtained by allowing full cooperation among the transmitters. We further tighten the outer bound in "strong

interference", where we show that the capacity region of a Gaussian broadcast channel with degraded message sets forms an outer bound to the capacity of the G-CIFC.

3. Finally, we propose outer bounds based on enhancing the original channel so as to transform it into a channel for which capacity is known.

3.5.1 A unifying framework for Th. 3.4.1 and Th. 3.4.2

Our objective is to obtain an outer bound for the G-CIFC with $|b| > 1$ that improves on the "strong interference" outer bound of Th. 3.4.2. Although the following theorem does not result in such a bound, it is of interest because it provides a simple unifying framework for Th. 3.4.1 and Th. 3.4.2. The proof of Th. 3.4.1 and the proof of Th. 3.4.2 use very different techniques. On the one hand, the bound in Th. 3.4.1 is valid for a general channel under the "weak interference" condition in (26, Th. 3.7) and is inspired by the converse for "less noisy BC". On the other hand, the bound in Th. 3.4.2 is valid for Gaussian channels with "strong interference" only and is inspired by the converse of "strong interference IFC". We will show next that both results may be derived within the framework proposed in Chapter 2. The proof of Th. 2.5.1 uses the argument originally devised by Sato for the BC (59) that, for channels without receiver cooperation, the capacity only depends on the output conditional marginals.

Theorem 3.5.1. ***Unifying outer bound.*** *The capacity region of the G-CIFC is contained in the region*

$$R_1 \le \mathcal{C}\left(\alpha P_1\right), \tag{3.8a}$$

$$R_2 \le \mathcal{C}\left(|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right), \tag{3.8b}$$

$$R_1 + R_2 \le \mathcal{C}\left(|b|^2 \bar{\alpha} P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right) + \left[\mathcal{C}\left(\alpha P_1\right) - \mathcal{C}\left(|b|^2 \alpha P_1\right)\right]^+ \tag{3.8c}$$

taken over the union of all $\alpha \in [0, 1]$. In "strong interference" ($|b| > 1$) the region in Equation 3.8 reduces to Th. 3.4.2, and in "weak interference" ($|b| \le 1$) to Th. 3.4.1.

Proof. In Th. 2.5.1, we showed that the capacity of a general CIFC is contained in the region

$$R_1 \le I(Y_1; X_1|X_2), \tag{3.9a}$$

$$R_2 \le I(X_1, X_2; Y_2), \tag{3.9b}$$

$$R_1 + R_2 \le I(X_1, X_2; Y_2) + I(Y_1; X_1|Y'_2, X_2), \tag{3.9c}$$

taken over the union of all joint distributions P_{X_1,X_2} and where Y'_2 has the same conditional marginal distribution as Y_2, i.e., $P_{Y'_2|X_1,X_2} = P_{Y_2|X_1,X_2}$. The result in Equation 3.9 specialized to the G-CIFC amounts to optimizing the correlation coefficient over the Gaussian additive noises, that is, optimizing with respect to $\gamma : \ |\gamma| \leq 1$ in

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}} \left(\mathbf{0}, \begin{bmatrix} 1 & \gamma \\ \gamma^* & 1 \end{bmatrix} \right).$$

First we show that a proper-complex Gaussian input exhausts the region in Equation 3.9. For any $\alpha \in [0, 1]$, let $\mathbf{S}$ be a covariance matrix defined as

$$\mathbf{S} \triangleq \begin{bmatrix} P_1 & \rho\sqrt{P_1 P_2} \\ \rho^* \sqrt{P_1 P_2} & P_2 \end{bmatrix} : \quad \rho = \sqrt{1-\alpha}\, \mathrm{e}^{\mathrm{j}\theta}, \ \theta \in \mathbb{R}, \tag{3.10}$$

and let $(X_{1G}, X_{2G}) \sim \mathcal{N}_{\mathbb{C}} (\mathbf{0}, \mathbf{S})$. By using the "Gaussian maximizes entropy" principle (see also (60, Eq.(3.29))), we conclude that for a given input covariance constraint $\mathbf{S}$ in Equation 3.10 for P_{X_1,X_2}, the regime in Equation 3.9c is upper bounded by

$$\textit{Equation } 3.9a \leq I(Y_1; X_{1G}|X_{2G}) = \textit{Equation } 3.8a, \tag{3.11}$$

$$\begin{aligned} \textit{Equation } 3.9b &\leq I(Y_2; X_{1G}, X_{2G}) \\ &= \log(1 + P_2 + |b|^2 P_1 + 2|b|\mathrm{Re}\{\rho\}\sqrt{P_1 P_2}) \leq \textit{Equation } 3.8b, \end{aligned} \tag{3.12}$$

$$\begin{aligned} \textit{Equation } 3.9c &\leq I(Y_2; X_{1G}, X_{2G}) + I(Y_1; X_{1G}|Y_2, X_{2G}) \\ &\leq \textit{Equation } 3.8b + \log \left(\frac{1 + (1 - |\rho|^2) P_1 \dfrac{|b|^2 + 1 - 2|b|\mathrm{Re}\{\gamma\}}{1 - |\gamma|^2}}{1 + |b|^2 (1 - |\rho|^2) P_1} \right). \end{aligned} \tag{3.13}$$

Since the bound in Equation 3.13 is valid for any $|\gamma| \leq 1$, the minimizing γ is

$$\underset{\gamma:\ |\gamma|\leq 1}{argmin} \frac{|b|^2 + 1 - 2|b|\mathrm{R}e\{\gamma\}}{1 - |\gamma|^2} = \min\left\{|b|, \frac{1}{|b|}\right\}. \tag{3.14}$$

After substituting the optimal value of γ given by Equation 3.14 in Equation 3.13 we obtain that the sum-rate in Equation 3.9c is bounded by Equation 3.8c. This shows that a Gaussian input is optimal in Equation 3.9 and that the worst conditional marginal is such that one of $Y_1|_{X_2}$ and $Y_2|_{X_2}$ is the degraded version of the other.

Finally, in "strong interference" the region in Equation 3.8 reduces to Th. 3.4.2 because the bound in Equation 3.8b is redundant due to Equation 3.8c, while in "weak interference" it reduces to Th. 3.4.1 because the closure of the region is determined by the rates pairs for which Equation 3.8a and Equation 3.8c are met with equality as argued in (61, Ex. 4.3). □

3.5.2 BC- based outer bounds

In this subsection we propose an outer bound that is tighter than the "strong interference" outer bound of Th. 3.4.2 in the "strong interference" regime. The following observation is key: if we provide the primary transmitter with the cognitive message, the G-CIFC becomes a Gaussian MIMO BC (with two antennas at the transmitter and one antenna at each receiver) where the input is subject to a per-antenna power constraint, as originally used in (14, page 1819). Thus, our proposed outer bound, valid for a fully general C-IFC is:

Theorem 3.5.2. ***BC-based outer bound.*** *The capacity of a general CIFC is contained in the following region*

$$\mathcal{R}^{(BC-PR)} \cap \mathcal{R}^{(2.5.1)}, \tag{3.15}$$

where $\mathcal{R}^{(BC-PR)}$ *is the capacity region (or an outer bound) for the* BC with private rates only *obtained by allowing the transmitters to fully cooperate is the outer bound in Th. 2.5.1 given in Equation 3.9.*

Proof. The theorem follows from the fact that allowing transmitter cooperation enlarges the capacity region of the CIFC and results in a BC.

The closed form expression of $\mathcal{R}^{(BC-PR)}$ was obtained in (62) and is presented here for completeness.

Consider an input covariance matrix defined as follows

$$\boldsymbol{S} \triangleq \begin{bmatrix} P_1 & \rho\sqrt{P_1 P_2} \\ \rho^* \sqrt{P_1 P_2} & P_2 \end{bmatrix} : \quad \rho = \sqrt{1-\alpha}\, \mathrm{e}^{\mathrm{j}\theta},\ \theta \in \mathbb{R},\ \alpha \in [0,1]. \tag{3.16}$$

The capacity region of a Gaussian *MIMO BC with private rates only* with a per-antenna power constraint is given by (62)

$$\mathcal{R}^{(BC-PR)} = \mathcal{CH} \bigcup_{\mathbf{S}} \mathcal{R}^{(BC-PR)}(\mathbf{S})$$

where $\mathcal{CH}$ denotes the convex-hull operation, $\bigcup_{\mathbf{S}}$ denotes the union over all input covariance matrices $\mathbf{S}$ that satisfy the per-antenna power constraint, and where

$$\mathcal{R}^{(BC-PR)}(\mathbf{S}) = \bigcup_{u\in\{1,2\}} \mathcal{R}^{(DPC\ u)}(\mathbf{S})$$

where $\mathcal{R}^{(DPC\ u)}(\mathbf{S})$ is the DPC region for the encoding order where user u is pre-coded against the interference created by the other user at its intended receiver, which is given by

$$\mathcal{R}^{(DPC\ u)}(\mathbf{S}) = \bigcup_{0\preceq\mathbf{B}_1,\ 0\preceq\mathbf{B}_2,\ \mathbf{B}_1+\mathbf{B}_2=\mathbf{S}} \mathcal{R}^{(DPC\ u)}(\mathbf{B}_1, \mathbf{B}_2), \quad u \in \{1, 2\},$$

and where, for

$$\mathbf{B}_1 = \begin{bmatrix} \alpha_1 P_1 & \rho_1\sqrt{\alpha_1 P_1\, \alpha_2 P_2} \\ \rho_1^*\sqrt{\alpha_1 P_1\, \alpha_2 P_2} & \alpha_2 P_2 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} \bar{\alpha}_1 P_1 & \rho_2\sqrt{\bar{\alpha}_1 P_1\, \bar{\alpha}_2 P_2} \\ \rho_2^*\sqrt{\bar{\alpha}_1 P_1\, \bar{\alpha}_2 P_2} & \bar{\alpha}_2 P_2 \end{bmatrix},$$

with

$$(\alpha_1, \alpha_2, |\rho_1|, |\rho_2|) \in [0, 1]^4 : \quad \rho_1\sqrt{\alpha_1\, \alpha_2} + \rho_2\sqrt{\bar{\alpha}_1\, \bar{\alpha}_2} = \rho,$$

the region $\mathcal{R}^{(DPC\ 1)}(\mathbf{B}_1, \mathbf{B}_2)$ is given by

$$R_1 \leq \mathcal{C}(\alpha_1 P_1 + |a|^2\alpha_2 P_2 + 2\mathrm{R}e\{a^*\rho_1\}\sqrt{\alpha_1\alpha_2 P_1 P_2}), \tag{3.17a}$$

$$R_2 \leq \mathcal{C}\left(\frac{\bar{\alpha}_1|b|^2 P_1 + \bar{\alpha}_2 P_2 + 2\mathrm{R}e\{\rho_2\}\sqrt{\bar{\alpha}_1\bar{\alpha}_2|b|^2 P_1 P_2}}{1 + |b|^2\alpha_1 P_1 + \alpha_2 P_2 + 2\mathrm{R}e\{\rho_1\}\sqrt{\alpha_1\alpha_2|b|^2 P_1 P_2}}\right), \tag{3.17b}$$

and $\mathcal{R}^{(DPC\ 2)}(\mathbf{B}_1, \mathbf{B}_2)$ is given by

$$R_1 \leq \mathcal{C}\left(\frac{\alpha_1 P_1 + |a|^2 \alpha_2 P_2 + 2\mathrm{Re}\{a^* \rho_1\}\sqrt{\alpha_1 \alpha_2 P_1 P_2}}{1 + \bar{\alpha}_1 P_1 + |a|^2 \bar{\alpha}_2 P_2 + 2\mathrm{Re}\{a^* \rho_2\}\sqrt{\bar{\alpha}_1 \bar{\alpha}_2 P_1 P_2}}\right), \tag{3.18a}$$

$$R_2 \leq \mathcal{C}(\bar{\alpha}_1 |b|^2 P_1 + \bar{\alpha}_2 P_2 + 2\mathrm{Re}\{\rho_2\}\sqrt{\bar{\alpha}_1 \bar{\alpha}_2 |b|^2 P_1 P_2}). \tag{3.18b}$$

The quantity α_u, $u \in \{1, 2\}$, represents the fraction of power P_u used to send the cognitive message W_1 on antenna u. The requirement $(\alpha_1, \alpha_2) \in [0, 1]^2$ guarantees that the per-antenna power constraints are verified. $\square$

Consider the G-CIFC with "strong interference" $|b| > 1$ and where the primary user is silent, i.e., $P_2 = 0$. This channel is equivalent to a (degraded) BC with input X_1 whose capacity $\boldsymbol{C}(a, |b|, P_1, 0)$ is given by (63)

$$R_1 \leq \mathcal{C}\left(\frac{\alpha P_1}{\bar{\alpha} P_1 + 1}\right),$$

$$R_2 \leq \mathcal{C}(\alpha |b|^2 P_1),$$

taken over the union of all $\alpha \in [0, 1]$. For $P_2 = 0$, the "strong interference" outer bound of Th. 3.4.2 reduces to

$$R_1 \leq \mathcal{C}(P_1),$$

$$R_1 + R_2 \leq \mathcal{C}(|b|^2 P_1).$$

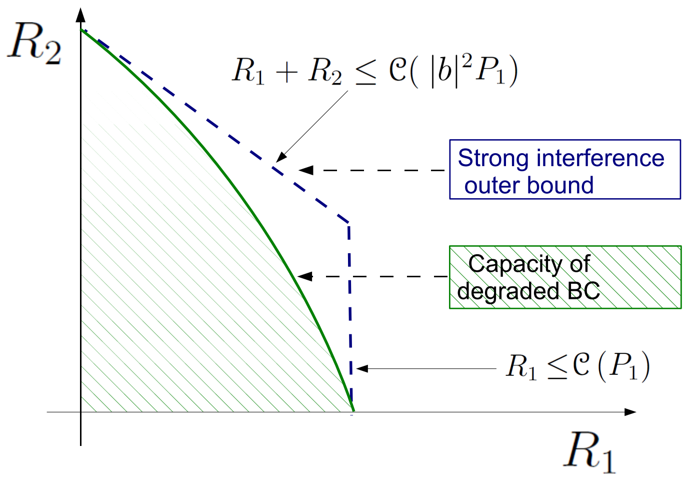

Figure 8. The "strong interference" outer bound of Th. 3.4.2 and the capacity region of the G-CIFC with $P_2 = 0$ and $|b| > 1$ (when the channel reduces to degraded BC).

These two regions are shown in Figure 8 where it is clear that the "strong interference" outer bound of Th. 3.4.2 fully contains the outer bound of the BC of Lemma 3.8.2. The two regions only coincide at the two Pareto optimal points A and B in Equation 3.56.

Th. 3.5.2 is valid for a general channel. It may be further tightened for the Gaussian channel in the "strong interference" regime. As previously noted in (25, Sec. 6.1), in the "strong interference" regime there is no loss of optimality in having the primary receiver decode the cognitive message in addition to its own message. Indeed, after decoding W_2, receiver 2 can reconstruct $X_2^N(W_2)$ and compute the following estimate of the receiver 1 output

$$\widetilde{Y}_1^N \triangleq \frac{Y_2^N - X_2^N}{|b|} + aX_2^N + \sqrt{1 - \frac{1}{|b|^2}} Z_0^N \sim Y_1^N, \tag{3.19}$$

where $Z_0^N \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I})$ and independent of everything else. Hence, if receiver 1 can decode W_1 from Y_1^N, so can receiver 2 from $\widetilde{Y}_1^N$. For this reason the capacity region of the G-CIFC for $|b| > 1$ is unchanged if receiver 2 is required to decoded both messages. If we further allow the two transmitters to fully cooperate, the resulting channel is a *Gaussian MIMO BC with degraded message sets*, with per-antenna power constraint, where message W_2 is to be decoded at receiver 2 only and message W_1 at both receivers. This implies that the bound in Th. 3.5.2 may be tightened for G-CIFC with $|b| > 1$ by using the capacity of the *Gaussian MIMO BC with degraded message sets* (BC-DMS) instead of the capacity of the *Gaussian MIMO BC with private rates only* (BC-PR):

Theorem 3.5.3. ***BC-DMS-based outer bound.*** *The capacity of a G-CIFC in "strong interference" ($|b| > 1$) satisfies*

$$\boldsymbol{C}(a, |b|, P_1, P_2) \subseteq \mathcal{R}^{(BC-DMS)} \cap \mathcal{R}^{(SI)}, \tag{3.20}$$

where $\mathcal{R}^{(BC-DMS)}$ is the capacity of the MIMO BC with degraded message sets *determined in (64; 62) and $\mathcal{R}^{(SI)}$ is the "strong interference" outer bound of Th. 3.4.2.*

Remark 3.5.4. The capacity of the of the general BC-DMS is derived in (64) and it is an outer bound for a general CIFC in "strong interference". This observation was also pointed out in the independent work of (65). It is possible to obtain the same outer bound by loosening the outer bound in (25, Th. 4), in particular by dropping (25, eq. (33)) and letting $U = [V, U_1]$. Our contribution is to determine a simpler expression for the capacity region of the Gaussian

MIMO BC-DMS, in particular by proving the optimality of Gaussian inputs in the region of (64).

The analytical evaluation of the outer bound region in Equation 3.15 of Th. 3.5.2 (or in Equation 3.20 of Th. 3.5.3) is quite involved in general. For the special cases of degraded G-CIFC and of S-G-CIFC a closed form expression may be obtained as follows.

Corollary 3.5.5. ***BC-based outer bound for the degraded G-CIFC****. For a degraded G-CIFC with $1/a = |b| \geq 1$, Th. 3.5.2 and Th. 3.5.3 coincide and reduce to*

$$R_1 \leq \mathcal{C}\left(\alpha P_1\right), \tag{3.21a}$$

$$R_2 \leq \mathcal{C}\left(\frac{P_2 + \bar{\alpha}|b|^2 P_1 + 2\sqrt{|b|^2 P_1 P_2}}{1 + \alpha P_1}\right), \tag{3.21b}$$

$$R_1 + R_2 \leq \mathcal{C}\left(P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right). \tag{3.21c}$$

Moreover, the R_2-bound from the MIMO BC capacity region (in Equation 3.21b) is more stringent than the R_2-bound from the "strong interference" outer bound (from the difference of Equation 3.21c and Equation 3.21a) if

$$|b| \geq \sqrt{\frac{P_2}{P_1}} + \sqrt{1 + \frac{P_2}{P_1}}.$$

Proof. When allowing full transmitter cooperation for a channel with $a|b| = 1$ and $|b| > 1$, we obtain an equivalent degraded BC with input $X_{\text{eq}} = |b|X_1 + X_2$ and outputs

$$Y_2 = (|b|X_1 + X_2) + Z_2 = X_{\text{eq}} + Z_2,$$
$$|b|Y_1 = (|b|X_1 + X_2) + |b|Z_1 \sim Y_2 + \sqrt{|b|^2 - 1}\, Z_0,$$

with $Z_0 \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ and independent of everything else. The input of the equivalent BC is subject to the power constraint

$$\mathbb{E}[|X_{\text{eq}}|^2] \leq (\sqrt{|b|^2 P_1} + \sqrt{P_2})^2 \triangleq P_{\text{eq}}.$$

For this order of degradedness among the users, the capacity region of the degraded BC with private rates equals the capacity with degraded message sets. In general $\mathcal{R}^{(BC-DMS)} \subseteq \mathcal{R}^{(BC-PR)}$, but since here Y_1 is a degraded version of Y_2, decoder 2 can decode the message of decoder 1 without imposing any rate penalty to user 1, thus $\mathcal{R}^{(BC-PR)}$ is achievable. This implies $\mathcal{R}^{(BC-DMS)} = \mathcal{R}^{(BC-PR)}$.

The capacity region of the equivalent BC is (63)

$$R_1 \leq R_1^{(BC-deg)}(\alpha') = \mathcal{C}\left(\frac{\alpha' P_{\text{eq}}}{(1-\alpha')P_{\text{eq}} + |b|^2}\right), \tag{3.22a}$$
$$R_2 \leq R_2^{(BC-deg)}(\alpha') = \mathcal{C}\left((1-\alpha')P_{\text{eq}}\right), \tag{3.22b}$$

taken over the union of all $\alpha' \in [0,1]$, i.e., that is $\rho_1 = \rho_2 = 1,\ \alpha_1 = \alpha_2 = \alpha'$ and $\mathcal{R}^{(BC-PR)} = \mathcal{R}^{(DPC\ 2)}$ in Equation 3.18.

To intersect the region in Equation 3.22 with the "strong interference" outer bound of Th. 3.4.2 we equate the R_1-bounds in Equation 3.22a and Equation 3.5a to obtain

$$\alpha' = \frac{\alpha P_1}{1+\alpha P_1}\left(1 + \frac{|b|^2}{(\sqrt{|b|^2 P_1} + \sqrt{P_2})^2}\right). \tag{3.23}$$

Notice that α' in Equation 3.23 satisfies $\alpha' \leq 1$ (the maximum value of 1 is obtained for $P_2 = 0$ and $\alpha = 1$). By substituting α' from Equation 3.23 in Equation 3.22b, we obtain the bound in Equation 3.21b.

The BC-based outer bound is more stringent than the "strong interference" outer bound if

$$\begin{aligned}
&R_1^{(BC-deg)}(\alpha) + R_2^{(BC-deg)}(\alpha) \leq R_{sum}^{(SI)}(\alpha) \quad \forall\alpha \in [0,1] \\
&\iff \alpha P_1 + P_2 + (1-\alpha)|b|^2 P_1 + 2\sqrt{|b|^2 P_1 P_2} \leq P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2} \quad \forall\alpha \in [0,1] \\
&\iff 2\sqrt{|b|^2 P_1 P_2}(1-\sqrt{\bar{\alpha}}) \leq \alpha P_1(|b|^2 - 1) \quad \forall\alpha \in [0,1] \\
&\iff 2\frac{\sqrt{|b|^2 P_1 P_2}}{P_1(|b|^2-1)} \leq 1 + \sqrt{\bar{\alpha}} \quad \forall\alpha \in [0,1] \quad \big(\text{since } \alpha = (1-\sqrt{\bar{\alpha}})(1+\sqrt{\bar{\alpha}})\big) \\
&\iff 2\frac{\sqrt{|b|^2 P_1 P_2}}{P_1(|b|^2-1)} \leq \min_{\bar{\alpha}\in[0,1]}\{1+\sqrt{\bar{\alpha}}\} = 1 \\
&\iff 1 + \frac{P_2}{P_1} \leq \left(|b| - \sqrt{\frac{P_2}{P_1}}\right)^2 \\
&\iff |b| \geq \sqrt{1+\frac{P_2}{P_1}} + \sqrt{\frac{P_2}{P_1}},
\end{aligned}$$

as claimed.

Remark 3.5.6. The capacity of the equivalent degraded BC may be achieved both by using superposition coding and binning. An achievable scheme inspired by the degraded BC and employing superposition coding is scheme (E) with $\lambda = 0$. An achievable scheme inspired by the degraded BC and employing binnig coding is scheme (B). Both schemes achieve the outer bound only in point A in Equation 3.56a. The capacity region of the degraded CIFC in therefore unknown in general it remains an interesting open problem.

□

Corollary 3.5.7. ***BC-DMS-based outer bound for the S-G-CIFC****. For a S-G-CIFC with $a = 0$ and $|b| \geq 1$ the outer bound of Th. 3.5.3 is contained in the region*

$$R_1 \leq \mathcal{C}\left(\alpha P_1\right), \tag{3.24a}$$

$$R_2 \leq \mathcal{C}\left(P_2 + \frac{|b|^2 P_1 \bar{\alpha}}{1+\alpha P_1} + 2\sqrt{\frac{\bar{\alpha}|b|^2 P_1 P_2}{1+\alpha P_1}}\right), \tag{3.24b}$$

$$R_1 + R_2 \leq \mathcal{C}\left(P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right). \tag{3.24c}$$

Moreover, the R_2-bound from the MIMO BC capacity region (from Equation 3.24b) is more stringent than the R_2-bound from the "strong interference" outer bound (from the difference of Equation 3.24c and Equation 3.24a) if

$$|b| \geq \sqrt{P_2 + 1}.$$

Proof. To establish the result in Corollary 3.5.7 we proceed as follows: first we prove that the capacity region of the Gaussian BC-DMS may be obtained form the region in (64) by considering Gaussian inputs and auxiliary RV. Successively we perform a partial optimization of the region in (64) in the Gaussian case and obtain a looser outer bound that may be expressed as a function of a single parameter. Finally we intersect this outer bound with the "strong interference" outer bound of Equation 3.5 to obtain the expression in Equation 3.24.

The capacity region of the general BC-DMS is found in (64) and is expressed as the union over all possible distributions of the input and one auxiliary RV. A closed form expression of the capacity region of the Gaussian BC-DMS is derived in (66) and is expressed as the intersection of the capacity region of a general BC and an additional sum rate constraint. We derive another simpler expression of the capacity region of the Gaussian BC-DMS and we do so by showing that we may restrict the union in (64) over all Gaussian inputs and auxiliary RV.

Consider the BC-DMS defined as:

$$Y_i = \boldsymbol{H}_i X + Z_i \qquad \forall\ i \in [1, 2] \tag{3.25}$$

where:

- X is a real valued input vector of size $n \times 1$ subject to the second moment constraint $\mathrm{C}ov[X] = \boldsymbol{K}_X \preceq \boldsymbol{S}$ for some $\boldsymbol{S} \succeq 0$,
- Y_i is a real valued output vector of size $m_i \times 1$ received by user $i \in [1, 2]$,

- $\boldsymbol{H}_i$ is a fixed real valued gain matrix imposed on user $i \in [1, 2]$. This is a matrix of size $m_i \times n$,
- Z_i is a real valued Gaussian random vector with zero mean and covariance matrix $\mathrm{C}ov[Z_i] = \boldsymbol{K}_Z \succ 0$.

As for the BC of (66), we consider real valued channels; the extension to complex valued channels is easily obtained by doubling the real dimensions. We first derive the capacity of a Gaussian BC-DMS for the case where $\boldsymbol{H}_i$ is square and invertible, we than argue that the case for a general $\boldsymbol{H}_i$ may be obtained by series of channel transformations originally devised for the BC in (62).

Theorem 3.5.8. *The capacity region of the Gaussian BC-DMS in Equation 3.25 is*

$$R_1 \leq I(U; Y_1), \tag{3.26a}$$

$$R_2 \leq I(X; Y_2|U), \tag{3.26b}$$

$$R_1 + R_2 \leq I(X; Y_2). \tag{3.26c}$$

taken over the union of all Gaussian U and X vectors of size n such that $\boldsymbol{K}_X \preceq \boldsymbol{S}$.

Proof. The region in Equation 3.26 was originally obtained in (64) for a general BC-DMS but considering the union over any distribution P_{UX}. To prove the theorem we need to show that only Gaussian U and X need to be considered.

First, we notice that Equation 3.26c is always maximized by having X Gaussian by the "Gaussian maximizes entropy" of (67). Since Equation 3.26c is maximized by Gaussian inputs, we have to show that the region obtained by considering Equation 3.26a and Equation 3.26b only is optimized by Gaussian inputs as well. To this end we write the region with Equation 3.26a and Equation 3.26b as

$$Equation\ 3.26a + (1-\mu) Equation\ 3.26b = \max_{P_{X|U}: Cov[X] \preceq \mathbf{S}} \quad \mu I(U;Y_1) + (1-\mu) I(X;Y_2|U)$$

$$\le \mu h(\boldsymbol{H}_1 X_G + Z_1) - (1-\mu) h(Z_2) +$$
$$(1-\mu) \max_{P_{X|U}:\ Cov[X|U] \preceq \mathbf{S}} \quad \left(h(\boldsymbol{H}_1 X + Z_1|U) - \frac{\mu}{(1-\mu)} h(\boldsymbol{H}_2 X + Z_2|U) \right), \quad (3.27)$$

for any $\mu \in [1/2, 1]$ and where X_G is a Gaussian vector with $\boldsymbol{K}_X \preceq \boldsymbol{S}$.

We need not consider $\mu \in [0, 1/2]$ because the region in Equation 3.26 is convex and contained in the triangular region

$$R_1, R_2 \quad \ge \quad 0, \tag{3.28a}$$

$$R_1 + R_2 \quad \le \quad I(\boldsymbol{H}_2 X_G + Z_2; X_G), \tag{3.28b}$$

see (64).

For these reasons, the region in Equation 3.26 cannot contain any rate point with tangent greater than -1 and thus there is no loss of generality in restricting μ in Equation 3.31 to the interval $[1/2, 1]$.

We now show that solution of the optimization problem

$$\max_{P_{X|U}:\ Cov[X|U]\preceq \mathbf{S},\ \mu\in[1/2,1]} \quad h(\boldsymbol{H}_1 X + Z_1|U) - \tfrac{\mu}{1-\mu} h(\boldsymbol{H}_2 X + Z_2|U).$$

must be Gaussian by using the extremal inequality of (68). We first focus on channels where $\boldsymbol{H}_i,\ i\in[1,2]$ is square and invertible, then show how this result may be used to establish a general channel using the perturbation techniques of (62).

If $\boldsymbol{H}_i,\ i\in[1,2]$, are square we may write

$$\begin{aligned}
&\max_{P_{X|U}:\ Cov[X|U]\preceq \mathbf{S}} \quad h(\boldsymbol{H}_1 X + Z_1|U) - \frac{\mu}{(1-\mu)} h(\boldsymbol{H}_2 X + Z_2|U) \\
&= \frac{\mu}{(1-\mu)} \left(\log|\boldsymbol{H}_2|\right)^{-1} - \left(\log|\boldsymbol{H}_1|\right)^{-1} \\
&\qquad + \max_{P_{X|U}:\ Cov[X|U]\preceq \mathbf{S}} \quad h(X + \boldsymbol{H}_1^{-1} Z_1|U) - \frac{\mu}{(1-\mu)} h(X + \boldsymbol{H}_2^{-1} Z_2|U). \qquad (3.29)
\end{aligned}$$

Th. 8 in (68) grants that the solution of the optimization problem in Equation 3.32 is Gaussian since $\mu/(1-\mu) > 1$ for $\mu \in [0, 1/2]$. Since we have established that both Equation 3.31 and Equation 3.26c are maximized by Gaussian X and U, we conclude that Equation 3.26 is also maximized by Gaussian X and U as well.

Finally the perturbation technique in (62, Section V.B) allows us to extend this result to a general channel where $\boldsymbol{H}_i$ in not necessarily square and invertible. The derivation in (62, Section V.B) was originally devised for the general BC but it extends in a straight-forward manner to the BC-DMS, since it solely relies on the channel matrix and the covariance of the noise and not on the message set.

□

The region in Equation 3.26 was originally obtained in (64) for a general BC-DMS but considering the union over any distribution P_{UX}. To prove the theorem we need to show that only Gaussian U and X need to be considered.

First, we notice that Equation 3.26c is always maximized by having X Gaussian by the "Gaussian maximizes entropy" of (67). Since Equation 3.26c is maximized by Gaussian inputs, we have to show that the region obtained by considering Equation 3.26a and Equation 3.26b only is optimized by Gaussian inputs as well. To this end we write the region with Equation 3.26a and Equation 3.26b as

$$\begin{aligned} &Equation\ 3.26a + (1-\mu)Equation\ 3.26b = \max_{P_{X|U}:Cov[X]\preceq \mathbf{S}} \quad \mu I(U;Y_1) + (1-\mu) I(X;Y_2|U) \\ &\quad \leq \mu h(\boldsymbol{H}_1 X_G + Z_1) - (1-\mu) h(Z_2) + \\ &\qquad (1-\mu) \max_{P_{X|U}:\ Cov[X|U]\preceq \mathbf{S}} \quad \left(h(\boldsymbol{H}_1 X + Z_1|U) - \frac{\mu}{(1-\mu)} h(\boldsymbol{H}_2 X + Z_2|U) \right), \end{aligned}$$

for any $\mu \in [1/2, 1]$ and where X_G is a Gaussian vector with $\boldsymbol{K}_X \preceq \boldsymbol{S}$.

We need not consider $\mu \in [0, 1/2]$ because the region in Equation 3.26 is convex and contained in the triangular region

$$R_1, R_2 \geq 0, \tag{3.30a}$$

$$R_1 + R_2 \leq I(\boldsymbol{H}_2 X_G + Z_2; X_G), \tag{3.30b}$$

see (64).

For these reasons, the region in Equation 3.26 cannot contain any rate point with tangent greater than -1 and thus there is no loss of generality in restricting μ in Equation 3.31 to the interval $[1/2, 1]$.

We now show that solution of the optimization problem

$$\max_{P_{X|U}:\ Cov[X|U] \preceq \mathbf{S},\ \mu \in [1/2,1]} \quad h(\boldsymbol{H}_1 X + Z_1|U) - \tfrac{\mu}{1-\mu} h(\boldsymbol{H}_2 X + Z_2|U).$$

must be Gaussian by using the extremal inequality of (68). We first focus on channels where $\boldsymbol{H}_i,\ i \in [1, 2]$ is square and invertible, then show how this result may be used to establish a general channel using the perturbation techniques of (62).

If $\boldsymbol{H}_i,\ i \in [1,2]$, are square we may write

$$\max_{P_{X|U}:\ Cov[X|U] \preceq \mathbf{S}} \quad h(\boldsymbol{H}_1 X + Z_1|U) - \frac{\mu}{(1-\mu)} h(\boldsymbol{H}_2 X + Z_2|U) \tag{3.31}$$

$$= \frac{\mu}{(1-\mu)} \left(\log|\boldsymbol{H}_2|\right)^{-1} - \left(\log|\boldsymbol{H}_1|\right)^{-1}$$

$$+ \max_{P_{X|U}:\ Cov[X|U] \preceq \mathbf{S}} \quad h(X + \boldsymbol{H}_1^{-1} Z_1|U) - \frac{\mu}{(1-\mu)} h(X + \boldsymbol{H}_2^{-1} Z_2|U). \tag{3.32}$$

Th. 8 in (68) grants that the solution of the optimization problem in Equation 3.32 is Gaussian since $\mu/(1-\mu) > 1$ for $\mu \in [0, 1/2]$. Since we have established that both Equation 3.31 and Equation 3.26c are maximized by Gaussian X and U, we conclude that Equation 3.26 is also maximized by Gaussian X and U as well.

Finally the perturbation technique in (62, Section V.B) allows us to extend this result to a general channel where $\boldsymbol{H}_i$ in not necessarily square and invertible. The derivation in (62, Section V.B) was originally devised for the general BC but it extends in a straight-forward manner to the BC-DMS, since it solely relies on the channel matrix and the covariance of the noise and not on the message set.

□

3.5.3 Outer bounds by transformation

Further outer bounds for the G-CIFC may be obtained by transforming the original G-CIFC into a different channel for which capacity is known. In the transformed channel the transmitters can reproduce the channel outputs of the original channel: this ensures that the transformation enlarges the capacity region thus providing an outer bound for the original channel.

Theorem 3.5.9. ***O****uter bound by channel transformations. For the capacity region* $\boldsymbol{C}(a,|b|,P_1,P_2)$ *we have*

$$\boldsymbol{C}(a,b,P_1,P_2) \subseteq \bigcap_{A,B,C\ :|A|\geq 1,|\frac{C}{1-B|b|}|\geq 1} \boldsymbol{C}\left(\frac{aA-B}{C},\frac{C|b|}{1-B|b|},(\sqrt{|A|^2P_1}+\sqrt{|B|^2P_2})^2,|C|^2P_2\right).$$

Proof. Let $X_1^N(W_1,W_2), X_2^N(W_2)$ be a good code for the channel $(a,|b|,P_1,P_2)$. Consider now the inputs

$$X_1' = AX_1 + BX_2,$$

$$X_2' = CX_2,$$

on a channel with parameters $(a',|b'|,P_1',P_2')$ resulting in the outputs

$$Y_1' = X_1' + a'X_2' + Z_1 \propto X_1 + \frac{B+a'C}{AX_2} + \frac{Z_1}{A},$$

and

$$Y_2' = |b'|X_1' + X_2' + Z_2 \propto \frac{|b'|}{|b'|B+C}X_1 + X_2 + \frac{Z_1}{|b'|B+C}.$$

If

$$
\begin{aligned}
a &= \frac{B + a'C}{A}, \\
|b| &= \frac{|b'|}{b'B + C}, \\
|A|^2 &\geq 1, \\
||b'|B + C|^2 &\geq 1, \\
P_1' &\geq (\sqrt{|A|^2 P_1} + \sqrt{|B|^2 P_2})^2, \\
P_2' &\geq |C|^2 P_2, \qquad (3.33)
\end{aligned}
$$

the output of the channel (a, b, P_1, P_2) may be reconstructed in the channel (a', b', P_1', P_2'). This implies

$$
\boldsymbol{C}(a, b, P_1, P_2) \subseteq \bigcap_{A,B,C:|A|\geq 1, |C/(1-B|b|)|\geq 1} \boldsymbol{C}\left(\frac{aA - B}{C}, \frac{C|b|}{1 - B|b|}, (\sqrt{|A|^2 P_1} + \sqrt{|B|^2 P_2})^2, |C|^2 P_2\right).
$$

S-G-CIFC.

By considering the transformation in (Equation 3.34) with

$$
\begin{aligned}
A &= 1 \\
B &= a \\
C &= 1 - a|b|
\end{aligned}
$$

we see that the capacity of a general G-CIFC $\boldsymbol{C}(a, |b|, P_1, P_2)$ is contained in the capacity region of S-G-CIFC $\boldsymbol{C}(0, |b|, |\sqrt{P_1} + a\sqrt{P_2}|^2, |1 - a|b||^2 P_2)$.

G-CIFC in "weak interference".

By considering the transformation in (Equation 3.34) with

$$A = |b|$$
$$B = \tfrac{a(1-|b|)}{a-1}$$
$$C = \tfrac{a|b|-1)}{a-1}$$

we have that the capacity of a general G-CIFC $\boldsymbol{C}(a, |b|, P_1, P_2)$ is contained in the capacity region of G-CIFC in "weak interference" $\boldsymbol{C}\left(a, 1, \left|\sqrt{|b|^2 P_1} + \tfrac{a(1-|b|)}{a-1}\sqrt{P_2}\right|^2, \left|\tfrac{a|b|-1}{a-1}\right|^2 P_2\right)$.

G-CIFC in "very strong interference".

By considering the transformation in (Equation 3.34) with

$$A = |b|$$
$$B = |b|\tfrac{1-a|b|}{|b|^2-1} - a$$
$$C = \tfrac{1-a|b|}{|b|^2-1}$$

we have that the capacity of a G-CIFC $\boldsymbol{C}(a, |b|, P, P)$ is contained in the capacity region of G-CIFC $\boldsymbol{C}\left(|b|, |b|, P', P'\right),\ P' = \tfrac{P}{(|b|^2-1)^2} \max\{||b|^2 - 1 + |b| - a|^2, |1 - a|b||^2\}$.

$\square$

Corollary 3.5.10. ***S****pecial cases of Th. 3.5.9. The capacity of the G-CIFC,* $\boldsymbol{C}(a,|b|,P_1,P_2)$ *is contained in the capacity region of the following channels:*

- ***S-G-CIFC:***

$$\boldsymbol{C}(a,|b|,P_1,P_2) \subseteq \boldsymbol{C}\left(0,|b|,|\sqrt{P_1}+a\sqrt{P_2}|^2,|1-a|b||^2P_2\right),$$

- ***G-CIFC in "weak interference":***

$$\boldsymbol{C}(a,|b|,P_1,P_2) \subseteq \boldsymbol{C}\left(a,1,\left|\sqrt{|b|^2P_1}+\frac{a(1-|b|)}{a-1}\sqrt{P_2}\right|^2,\left|\frac{a|b|-1}{a-1}\right|^2P_2\right),$$

- ***G-CIFC in "very strong interference":***

$$\boldsymbol{C}(a,|b|,P,P) \subseteq \boldsymbol{C}\left(|b|,|b|,P',P'\right),\ P'=\frac{P}{(|b|^2-1)^2}\max\{||b|^2-1+|b|-a|^2,|1-a|b||^2\}$$

Proof. Let $X_1^N(W_1,W_2),X_2^N(W_2)$ be a good code for the channel $(a,|b|,P_1,P_2)$. Consider now the inputs

$$X_1' = AX_1+BX_2,$$

$$X_2' = CX_2,$$

on a channel with parameters $(a',|b'|,P_1',P_2')$ resulting in the outputs

$$Y_1' = X_1'+a'X_2'+Z_1 \propto X_1+\frac{B+a'C}{AX_2}+\frac{Z_1}{A},$$

and

$$Y_2' = |b'|X_1' + X_2' + Z_2 \propto \frac{|b'|}{|b'|B+C}X_1 + X_2 + \frac{Z_1}{|b'|B+C}.$$

If

$$\begin{aligned}
a &= \frac{B + a'C}{A}, \\
|b| &= \frac{|b'|}{b'B+C}, \\
|A|^2 &\geq 1, \\
||b'|B+C|^2 &\geq 1, \\
P_1' &\geq (\sqrt{|A|^2P_1} + \sqrt{|B|^2P_2})^2, \\
P_2' &\geq |C|^2P_2, \qquad (3.34)
\end{aligned}$$

the output of the channel (a, b, P_1, P_2) may be reconstructed in the channel (a', b', P_1', P_2'). This implies

$$\boldsymbol{C}(a, b, P_1, P_2) \subseteq \bigcap_{A,B,C:|A|\geq 1, |C/(1-B|b|)|\geq 1} \boldsymbol{C}\left(\frac{aA-B}{C}, \frac{C|b|}{1-B|b|}, (\sqrt{|A|^2P_1} + \sqrt{|B|^2P_2})^2, |C|^2P_2\right).$$

S-G-CIFC.

By considering the transformation in Equation 3.34 with

$$A = 1$$
$$B = a$$
$$C = 1 - a|b|$$

we see that the capacity of a general G-CIFC $\boldsymbol{C}(a, |b|, P_1, P_2)$ is contained in the capacity region of S-G-CIFC $\boldsymbol{C}(0, |b|, |\sqrt{P_1} + a\sqrt{P_2}|^2, |1 - a|b||^2 P_2)$.

G-CIFC in "weak interference".

By considering the transformation in Equation 3.34 with

$$A = |b|$$
$$B = \tfrac{a(1-|b|)}{a-1}$$
$$C = \tfrac{a|b|-1)}{a-1}$$

we have that the capacity of a general G-CIFC $\boldsymbol{C}(a, |b|, P_1, P_2)$ is contained in the capacity region of G-CIFC in "weak interference" $\boldsymbol{C}\left(a, 1, \left|\sqrt{|b|^2 P_1} + \frac{a(1-|b|)}{a-1}\sqrt{P_2}\right|^2, \left|\frac{a|b|-1}{a-1}\right|^2 P_2\right)$.

G-CIFC in "very strong interference".

By considering the transformation in Equation 3.34 with

$$A = |b|$$
$$B = |b|\tfrac{1-a|b|}{|b|^2-1} - a$$
$$C = \tfrac{1-a|b|}{|b|^2-1}$$

we have that the capacity of a G-CIFC $\boldsymbol{C}(a, |b|, P, P)$ is contained in the capacity region of G-CIFC $\boldsymbol{C}\left(|b|, |b|, P', P'\right)$, $P' = \frac{P}{(|b|^2-1)^2} \max\{||b|^2 - 1 + |b| - a|^2, |1 - a|b||^2\}$.

□

Remark 3.5.11. The G-IFC with conferencing encoders of (69) encompasses the G-CIFC as a special case when $C_{12} = 0$ and $C_{21} = \infty$. The outer bound in (69, Lemma 4.1) with $C_{12} = 0$ and $C_{21} = \infty$ is an outer bound for the G-CIFC. This outer bound reduces to the "strong interference" outer bound of Th. 3.4.2 when the channel is a G-CIFC. In particular we notice that for a CIFC, unlike for a classical IFC and the IFC with conferencing encoders, no bounds of the form $2R_1 + R_2$ are known. In (69) the authors provide an interesting interpretation of this type of bound for a channel with and without conferencing transmitters. With regard to this interpretation we point out that, with full a priori knowledge of the primary message, the cognitive transmitter can always pre-code its message against the interference from the primary user and thus the strategy of the primary encoder never limits the rate of the cognitive receiver.

3.6 Inner bounds

In Chapter 2 we introduce a new inner bound for the Discrete Memoryless CIFC (DM-CIFC) and show that this scheme encompasses all previously proposed achievable schemes as special cases; it is thus the largest known achievable rate region to date. This achievable scheme also introduces new transmission features that were crucial in proving capacity for the semi-deterministic DM-CIFC of Th. 2.9.1. Here we use the inner bound of Th. 2.6.1 as a unified framework to present the achievable schemes used in the remainder of the chapter. In this section we introduce the general achievable scheme in Th. 2.6.1 and use it to obtain six simple sub-schemes that will be used in the following sections to prove capacity and constant gap results.

As the Gaussian CIFC encompasses classical interference, multiple-access and broadcast channels, the achievable rate region of Th. 2.6.1 incorporates a combination of the transmission techniques devised for these channels.

- **R**ate-splitting. Both the primary and the cognitive message are split into private and common parts, as in the Han and Kobayashi scheme (18) for the IFC. Although rate-splitting was shown to be unnecessary in the "weak interference" (26) and "very strong interference" (22) regimes of Equation 3.2 and Equation 3.6, respectively, it allows significant rate improvement in the "strong interference" regime.
- **S**uperposition-coding. The cognitive common message is superposed to the primary common message and parts of the cognitive message are superposed to parts of the primary message. Useful in multiple-access and broadcast channels (46), a simple superposition of the primary and cognitive messages (all common) is capacity achieving in the "very strong interference" regime (22).
- **P**re-coding. Gel'fand-Pinsker coding (48), often referred to as binning or Dirty Paper Coding (DPC), allows a transmitter to pre-code (portions of) the interference known to be experienced at the receiver. Binning is also used by Marton in (49) to derive the largest known achievable rate region for the BC. In the scheme of Th. 2.6.1, binning is performed at the cognitive encoder for both the common and the private message and it allows for the cancellation of interference from the primary transmitter.

- **B**roadcasting. In Chapter 2 we introduced the idea of having the cognitive encoder transmit part of the primary message. This is made possible by the perfect knowledge of the primary message at the cognitive transmitter, which is specific to this channel model. The additional primary message is superposed to the cognitive common message and also precoded against the cognitive private message. The incorporation of the broadcast feature at the cognitive transmitter was initially motivated by the fact that in certain regimes, this strategy was shown to be capacity achieving for the high-SNR linear deterministic approximation of the CIFC (70).

The achievable scheme may be described as follows:

- **R**ate-splitting. The independent messages W_1 and W_2, uniformly distributed on $\mathcal{M}_1 = [1 : 2^{nR_1}]$ and $\mathcal{M}_2 = [1 : 2^{nR_2}]$ respectively, are rate split into the messages W_i, $i \in \{1c, 2c, 1pb, 2pb, 2pa\}$, all independent and uniformly distributed on $[1 : 2^{nR_i}]$, each encoded using the RV U_i.

- **P**rimary encoder. Transmitter 2 sends X_2 that carries the private message W_{2pa} ("p" for private, "a" for alone) **superposed** to the common message W_{2c} carried by U_{2c} ("c" for common).

- **C**ognitive encoder. The common message of transmitter 1, encoded by U_{1c}, is **b**inned against X_2 conditioned on U_{2c}. The private message of transmitter 2, W_{2pb}, encoded by U_{2pb} ("b" for broadcast) and a portion of the private message of transmitter 1, W_{1pb}, encoded as U_{1pb}, are **b**inned against each other as in Marton's region (49) conditioned on U_{1c}, U_{2c}, X_2. Transmitter 1 sends X_1, which is a function of all the RVs.

- **P**rimary decoder. Receiver 2 jointly decodes U_{2c} (carrying W_{2c}), U_{1c} (carrying W_{1c}), U_{2pb} (carrying W_{2pb}), and X_2 (carrying W_{2pa}).

- **C**ognitive decoder. Receiver 1 jointly decodes U_{1c} (carrying W_{1c}), U_{2pb} (carrying W_{2pb}), and U_{1pb} (carrying W_{1pb}).

- **A**nalysis. The codebook generation, encoding, decoding and the error analysis are provided Chapter 2.

Corollary 3.6.1. ***Achievable region*** $\mathcal{R}^{(RTD)}$ ***in Th. 2.9.1.***

A rate pair (R_1, R_2) *such that*

$$R_1 = R_{1c} + R_{1pb}, \tag{3.35a}$$

$$R_2 = R_{2c} + R_{2pa} + R_{2pb} \tag{3.35b}$$

is achievable for a general DM-CIFC if $(R'_{1c}, R'_{1pb}, R'_{2pb}, R_{1c}, R_{1pb}, R_{2c}, R_{2pa}, R_{2pb}) \in \mathbb{R}_+^8$ *satisfies:*

$$\begin{aligned}
R'_{1c} &\geq I(U_{1c}; X_2|U_{2c}) && (3.36a)\\
R'_{1c} + R'_{1pb} &\geq I(U_{1pb}; X_2|U_{1c}, U_{2c}) + I(U_{1c}; X_2|U_{2c}) && (3.36b)\\
R'_{1c} + R'_{1pb} + R'_{2pb} &\geq I(U_{1pb}; X_2, U_{2pb}|U_{1c}, U_{2c})\\
&\quad + I(U_{1c}; X_2|U_{2c}) && (3.36c)\\
R_{2c} + R_{2pa} + (R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}, X_2, U_{2c})\\
&\quad + I(U_{1c}; X_2|U_{2c}) && (3.36d)\\
R_{2pa} + (R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}, X_2|U_{2c}) + I(U_{1c}; X_2|U_{2c}) && (3.36e)\\
R_{2pa} + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, X_2|U_{1c}, U_{2c}) + I(U_{1c}; X_2|U_{2c}) && (3.36f)\\
(R_{1c} + R'_{1c}) + (R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}, U_{1c}|X_2, U_{2c}) + I(U_{1c}; X_2|U_{2c}) && (3.36g)\\
(R_{2pb} + R'_{2pb}) &\leq I(Y_2; U_{2pb}|U_{1c}, X_2, U_{2c}) && (3.36h)\\
R_{2c} + (R_{1c} + R'_{1c}) + (R_{1pb} + R'_{1pb}) &\leq I(Y_1; U_{1pb}, U_{1c}, U_{2c}), && (3.36i)\\
(R_{1c} + R'_{1c}) + (R_{1pb} + R'_{1pb}) &\leq I(Y_1; U_{1pb}, U_{1c}|U_{2c}), && (3.36j)\\
(R_{1pb} + R'_{1pb}) &\leq I(Y_1; U_{1pb}|U_{1c}, U_{2c}), && (3.36k)
\end{aligned}$$

for some input distribution

$$P_{Y_1,Y_2,X_1,X_2,U_{1c},U_{2c},U_{2pa},U_{1pb},U_{2pb}} = P_{U_{1c},U_{2c},U_{2pa},U_{1pb},U_{2pb},X_1,X_2} P_{Y_1,Y_2|X_1,X_2}.$$

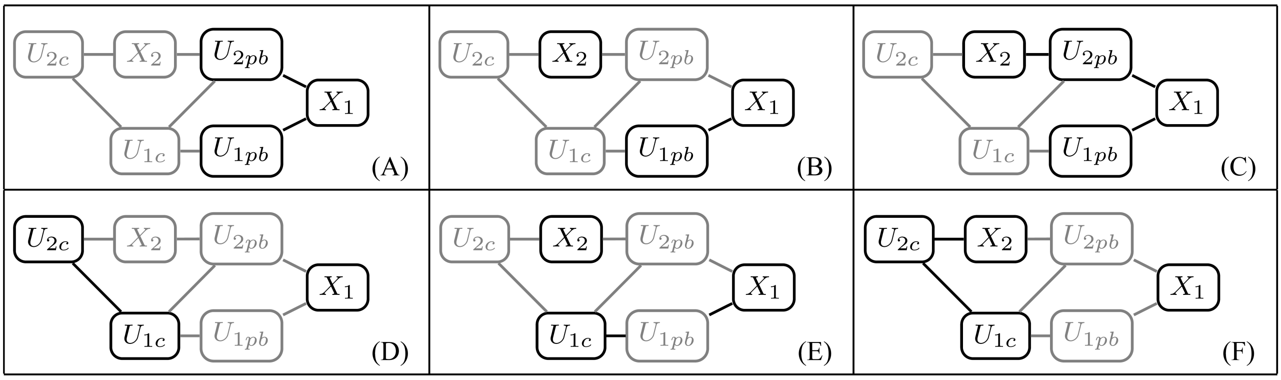

Figure 9. The achievable schemes of Section 3.6.

We now present six different sub-schemes obtained from the achievable scheme of Corollary 3.6.1 by reducing the number of rate splits to at most three rather than five. By setting some rates to zero we may drop the corresponding RVs and simplify the region in Equation 3.36. The resulting transmission schemes are used in the rest of the chapter for achievability proofs (for capacity and constant gap results) and numerical evaluations. Figure 9 and 3.6 help illustrate the different schemes: Figure 9 shows, for each scheme, which rate splits in the $\mathcal{R}^{(RTD)}$ are set to zero (the corresponding RV is in gray) and which ones are not (the corresponding RV is in black), while 3.6 indicates which result will be proved with the corresponding scheme.

	U_{2c}	U_{1c}	X_2	U_{1pb}	U_{2pb}	Role	Where
(A)				•	•	constant gap in a subspace	Th. 3.8.4
(B)			•	•		capacity in “weak interference”	Th. 3.4.1, 3.8.1, 3.8.3
(C)			•	•	•	constant gap in the whole parameter region	Th. 3.8.1
(D)	•	•				capacity in “very strong strong interference”	Th. 3.4.3
						constant gap in a subspace	Th. 3.8.4
(E)		•	•			capacity for the “primary decodes cognitive”	Th. 3.7.1, 3.7.3
						constant gap in a subspace	Th. 3.8.3, 3.8.4
(F)	•	•		•		numerical results	Sec. 3.9

Figure 10. The role of the different achievable schemes in the following Sections

3.6.1 Achievable scheme with U_{2pb} and U_{1pb}: capacity achieving for the degraded broadcast channel.

Motivation: *Achieve the capacity to within a finite gap in some parameter regime by having transmitter 2 silent.*

Consider the case where transmitter 2 is silent and transmitter 1 transmits to both decoders. In this case, the G-CIFC with $P_2 = 0$ reduces to a degraded BC with input X_1 (71). When $|b| < 1$, Y_2 is a degraded version of Y_1 and the maximum achievable rate region when transmitter 2 is silent is

$$R_1 \leq I(Y_1; U_{1pb}) - I(U_{1pb}; U_{2pb}), \tag{3.37a}$$

$$R_2 \leq I(Y_2; U_{2pb}) \tag{3.37b}$$

taken over the union over of all $\alpha \in [0, 1]$. When $|b| \geq 1$, Y_1 is a degraded version of Y_2 and the maximum achievable rate region when transmitter 2 is silent is

$$R_1 \leq I(Y_1; U_{1pb}), \tag{3.38a}$$

$$R_2 \leq I(Y_2; U_{2pb}|U_{1pb}) \tag{3.38b}$$

taken over the union of all $\alpha \in [0, 1]$.

3.6.2 Achievable scheme with X_2 and U_{1pb}: capacity achieving in the "weak interference" regime.

Motivation: Completeness.

In this scheme both messages are private and receiver 2 treats the interference from transmitter 1 as noise while transmitter 1 performs perfect DPC against the interference from transmitter 2. This scheme achieves capacity in the "weak interference regime" of Th. 3.4.1 (26).

3.6.3 Achievable scheme with X_2, U_{1pb} and U_{2pb}: capacity achieving in the semi-deterministi DM-CIFC.

Motivation: Achieve the "strong interference" outer bound to within a constant gap in the whole parameter regime.

This achievable strategy is obtained by combining the previous two transmission schemes, scheme (A) and (B), and it corresponds to the capacity achieving scheme for the semi-deterministic G-CIFC in Th. 2.9.1. The broadcasting RV U_{2pb} appears only in the $\mathcal{R}^{(RTD)}$ region and in (51; 50). The achievable rate region is

$$\begin{aligned} R_1 &\leq I(Y_1;U_{1pb}) - I(U_{1pb};X_2) \\ &\overset{Equation\ 3.40}{=} \log(\sigma^2_{1pb} + \alpha P_1) - \log\left(\sigma^2_{1pb} + \frac{\mathrm{V}ar[X_1 + aX_2]}{1 + \mathrm{V}ar[X_1 + aX_2]}\right), \quad (3.39a) \\ R_2 &\leq I(Y_2;U_{2pb},X_2) \\ &\overset{Equation\ 3.40}{=} \log(1 + \mathrm{V}ar[|b|X_1 + X_2]) - \log\left(1 + \frac{\sigma^2_{2pb}\mathrm{V}ar[|b|X_1|_{X_2}]}{\sigma^2_{2pb} + \mathrm{V}ar[|b|X_1|_{X_2}]}\right), \quad (3.39b) \\ R_1 + R_2 &\leq I(Y_2;U_{2pb},X_2) + I(Y_1;U_{1pb}) - I(U_{1pb};U_{2pb},X_2) \\ &\overset{Equation\ 3.40}{\leq} Equation\ 3.39a + Equation\ 3.39b + \log\left(1 - \frac{\left|[|b|P_1\alpha - \sqrt{\sigma^2_{1pb}\sigma^2_{2pb}}]^+\right|^2}{(|b|^2 P_1\alpha + \sigma^2_{2pb})(P_1\alpha + 1)}\right) \quad (3.39c) \end{aligned}$$

where

$$\begin{aligned} X_{1pb} &\sim \mathcal{N}_{\mathbb{C}}(0, \alpha P_1) \\ X_2 &\sim \mathcal{N}_{\mathbb{C}}(0, P_2), \text{ independent of } X_{1pb}, \\ X_1 &= X_{1pb} + \sqrt{\frac{\bar{\alpha}P_1}{P_2}}X_2 \\ U_{1pb} &= X_1 + aX_2 + Z_{1pb} \\ U_{2pb} &= |b|X_1 + X_2 + Z_{2pb}, \end{aligned} \quad (3.40)$$

and

$$\begin{bmatrix} Z_{1pb} \\ Z_{2pb} \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \begin{bmatrix} \sigma^2_{1pb} & \rho_{pb}\sqrt{\sigma^2_{1pb}\sigma^2_{2pb}} \\ \rho^*_{pb}\sqrt{\sigma^2_{1pb}\sigma^2_{2pb}} & \sigma^2_{2pb} \end{bmatrix}\right),$$

for $|\rho_{pb}| \leq 1$. The assignment in Equation 3.40 is inspired by the capacity achieving scheme for the semi-deterministic CIFC of Th. 2.9.1where U_{1pb} and U_{2pb} are set to be equal to Y_1 and Y_2 respectively. The inequality in Equation 3.39c is obtained by optimizing ρ_{pb} as detailed in Th. 3.8.1.

3.6.4 Achievable scheme with U_{1c} and U_{2c}: capacity achieving in "very strong interference" regime.

Motivation: Completeness.

This scheme achieves the "strong interference" outer bound of Th. 3.4.2 under the "very strong interference" conditions of Th. 3.4.3 (24). The achievable rate region is

$$\begin{aligned}
R_1 &\leq I(Y_1; X_1|X_2) \overset{\text{Equation 3.42}}{=} \mathcal{C}((1-|\rho|^2)P_1), && \text{(3.41a)} \\
R_1 &\leq I(Y_2; X_1|X_2) \overset{\text{Equation 3.42}}{=} \mathcal{C}((1-|\rho|^2)|b|^2P_1), && \text{(3.41b)} \\
R_1 + R_2 &\leq I(Y_1; X_1, X_2) \overset{\text{Equation 3.42}}{=} \mathcal{C}(P_1 + |a|^2P_2 + 2\text{Re}\{a^*\rho\}\sqrt{P_1P_2}), && \text{(3.41c)} \\
R_1 + R_2 &\leq I(Y_1; X_1, X_2) \overset{\text{Equation 3.42}}{=} \mathcal{C}(|b|^2P_1 + P_2 + 2|b|\text{Re}\{\rho\}\sqrt{P_1P_2}), && \text{(3.41d)}
\end{aligned}$$

where the RHS of Equation 3.41 is achieved with the assignment

$$\begin{aligned}
X_{1c} &\sim \mathcal{N}_{\mathbb{C}}(0, (1-|\rho|^2)P_1) \\
X_2 &\sim \mathcal{N}_{\mathbb{C}}(0, P_2), \text{ independent of } X_{1c}, \\
X_1 &= X_{1c} + \rho\sqrt{\frac{P_1}{P_2}}X_2, && \text{(3.42)}
\end{aligned}$$

for some $|\rho| \leq 1$. This scheme was originally proposed for real-valued channels in (24). Here we consider its extension to complex-valued valued channels.

3.6.5 Achievable scheme with X_2, U_{1c}: capacity achieving in the primary decodes cognitive regime.

Motivation: Achieve capacity in the "primary decodes cognitive" regime.

In this scheme the primary message is private while the cognitive message is public and binned against the interference created by the primary user at the cognitive decoder. This scheme can also be obtained as a special case of the scheme in (25) and (51). The achievable rate region is

$$\begin{aligned} R_1 &\leq I(Y_1; U_{1c}) - I(U_{1c}; X_2) \\ &\overset{\text{Equation 3.45}}{=} f\left(a + \sqrt{\frac{\bar{\alpha} P_1}{P_2}}, 1; \lambda\right), \end{aligned} \tag{3.43a}$$

$$\begin{aligned} R_2 &\leq I(Y_2; U_{1c}, X_2) - (I(Y_2; U_{1c}) - I(U_{1c}; X_2)) \\ &\overset{\text{Equation 3.45}}{=} \mathcal{C}(P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}) - f\left(\frac{1}{|b|} + \sqrt{\frac{\bar{\alpha} P_1}{P_2}}, \frac{1}{|b|^2}; \lambda\right), \end{aligned} \tag{3.43b}$$

$$\begin{aligned} R_1 + R_2 &\leq I(Y_2; U_{1c}, X_2) \\ &\overset{\text{Equation 3.45}}{=} \mathcal{C}(P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}), \end{aligned} \tag{3.43c}$$

for

$$f(h,\sigma^2;\lambda) \triangleq I(X_{1c}+hX_2+\sigma Z_1;U_{1c}) - I(U_{1c};X_2)$$
$$\stackrel{Equation\ 3.45}{=} \log\left(\frac{\sigma^2+\alpha P_1}{\sigma^2+\frac{\alpha P_1|h|^2P_2}{\alpha P_1+|h|^2P_2+\sigma^2}\left|\frac{\lambda}{\lambda_{Costa}(h,\sigma^2)}-1\right|^2}\right),$$

with

$$\lambda_{Costa}(h,\sigma^2) \triangleq \frac{\alpha P_1}{\alpha P_1+\sigma^2}\ h, \tag{3.44}$$

and where the RHS of Equation 3.43 is achieved with the assignment

$$\begin{aligned} X_{1c} &\sim \mathcal{N}_{\mathbb{C}}(0,\alpha P_1) \\ X_2 &\sim \mathcal{N}_{\mathbb{C}}(0,P_2) \\ X_1 &\sim X_{1c}+\sqrt{\frac{\bar{\alpha}P_1}{P_2}}X_2 \\ U_{1c} &= X_{1c}+\lambda X_2, \end{aligned} \tag{3.45a}$$

for some $\alpha \in [0,1]$ and $\lambda \in \mathbb{C}$. Note that $f(h,\sigma^2;\lambda) \geq 0$ if $\left|\frac{\lambda}{\lambda_{Costa}(h,\sigma^2)}-1\right|^2 \leq 1+\frac{\alpha P_1+\sigma^2}{|h|^2P_2}$.

3.6.6 Achievable scheme with U_{2c}, X_2 and U_{1c}.

Motivation: Achieve capacity in the largest subset of the "strong interference" regime.

As for scheme (C), this scheme is obtained by combining the previous two schemes, scheme (D) and (E). The achievable rate region is

$$
\begin{aligned}
R_1 &\leq I(Y_1; U_{1c}|U_{2c}) - I(X_2; U_{1c}|U_{2c}), && (3.46a)\\
R_1 &\leq I(Y_2; U_{1c}, X_2|U_{2c}), && (3.46b)\\
R_1 + R_2 &\leq I(Y_2; U_{2c}, X_2, X_{1c}), && (3.46c)\\
R_1 + R_2 &\leq I(Y_2; X_2|U_{1c}, U_{2c}) + I(Y_1; U_{1c}, U_{2c}), && (3.46d)\\
2R_1 + R_2 &\leq I(Y_2; U_{1c}, X_2|U_{2c}) + I(Y_1; U_{1c}, U_{2c}) - I(U_{1c}; X_2|U_{2c}). && (3.46e)
\end{aligned}
$$

In particular, we consider the choice of RVs

$$
\begin{aligned}
X_{2c}, X_{2pa}, X_{1c} &\sim \text{iid } \mathcal{N}_{\mathbb{C}}(0, 1) && (3.47a)\\
X_2 &= \sqrt{\beta P_2} X_{2c} + \sqrt{\bar{\beta} P_2} X_{2pa} && (3.47b)\\
X_1 &= \sqrt{\alpha P_1} X_{1c} + \sqrt{\bar{\alpha} P_1} \left(\sqrt{\gamma} X_{2c} + \sqrt{\bar{\gamma}} X_{2pa} \right) && (3.47c)\\
U_{1c} &= X_{1c} + \lambda X_{2pa} && (3.47d)\\
U_{2c} &= X_{2c}. && (3.47e)
\end{aligned}
$$

This scheme unifies the two schemes that achieve capacity in two different parameter regimes of of $|b| > 1$ and hence is the scheme that achieves capacity in the largest subset of the “strong interference” regime.

3.7 New capacity results

We now present two new capacity results for the G-CIFC. The first capacity result uses scheme (E) to achieve the "strong interference" outer bound in what we term the "primary decodes cognitive" regime, a subset of the "strong interference" regime that is not included in the "very strong interference" regime of Th. 3.4.3, for which capacity is already known. The second capacity result focuses on the S-G-CIFC where we show that the BC-DMS-based outer bound of Th. 3.5.2 is achieved by scheme (E) for a large set of parameters where capacity was previously unknown. Although the two results involve the same achievable scheme (E), in the first result the cognitive receiver performs Costa's "interference pre-cancellation" (or pre-coding) of the interference from the primary receiver while, in the second result, no pre-coding in necessary. In scheme (E) the pre-coding operation has an interesting effect on the rate region that we investigate in detail in Remark 3.7.2.

Before presenting the new results, we describe scheme (E) in more detail. The achievable rate region is expressed in two parameters: α and λ. The parameter α denotes the fraction of power that encoder 1 employs to transmit its own message versus the power to broadcast X_2. For $\alpha = 0$, transmitter 1 uses all its power to broadcast X_2 as in a virtual Multiple Input Single Output (MISO) channel. When $\alpha = 1$, transmitter 1 utilizes all its power to transmit X_{1c}. The parameter λ controls the amount of interference (created by X_2 at receiver 1) "pre-cancellation" achievable using DPC at transmitter 1. With $\lambda = 0$, no DPC is performed at transmitter 1 and the interference due to X_2 is treated as noise. On the other hand, with $\lambda = \lambda_{Costa}$ for

$$\lambda_{Costa}\left(a + \sqrt{\frac{\bar{\alpha}P_1}{P_2}},\ 1\right) \triangleq \lambda_{Costa\ 1},$$

with $\lambda_{Costa}(\cdot,\cdot)$ defined in Equation 3.44, the interference due to X_2 at receiver 1 is completely "pre-canceled", thus achieving the maximum possible rate R_1. Different values of λ are not usually investigated because, as long as the interference is a nuisance (i.e., no node in the network has information to extract from the interference), the best is to completely "pre-cancel" it by using $\lambda = \lambda_{Costa}(h, \sigma^2)$.

However, λ influences not only the rate R_1 in Equation 3.43a, but also the rate R_2 in Equation 3.43b. An interesting question is whether $\lambda \neq \lambda_{Costa\ 1}$, although it does not achieve the largest possible R_1, would improve the achievable rate region by sufficiently boosting the rate R_2. We comment on this question later on in Section **??**. At this point we make the following observation: R_1 is a concave function in λ, symmetric around $\lambda = \lambda_{Costa\ 1}$ and with a global maximum at $\lambda = \lambda_{Costa\ 1}$, while R_2 is a convex function in λ, symmetric around $\lambda = \lambda_{Costa\ 2}$ and with a global minimum at $\lambda = \lambda_{Costa\ 2}$, where

$$\lambda_{Costa}\left(\frac{1}{|b|} + \sqrt{\frac{\bar{\alpha}P_1}{P_2}},\ \frac{1}{|b|^2}\right) \triangleq \lambda_{Costa\ 2}.$$

Figure 11 shows R_1 in Equation 3.43a and R_2 in Equation 3.43b as a function of $\lambda \in \mathbb{R}$, for $P_1 = P_2 = 6$, $b = \sqrt{2}$, $a = \sqrt{0.3}$, and $\alpha = 0.5$. For the chosen parameters, we observe a

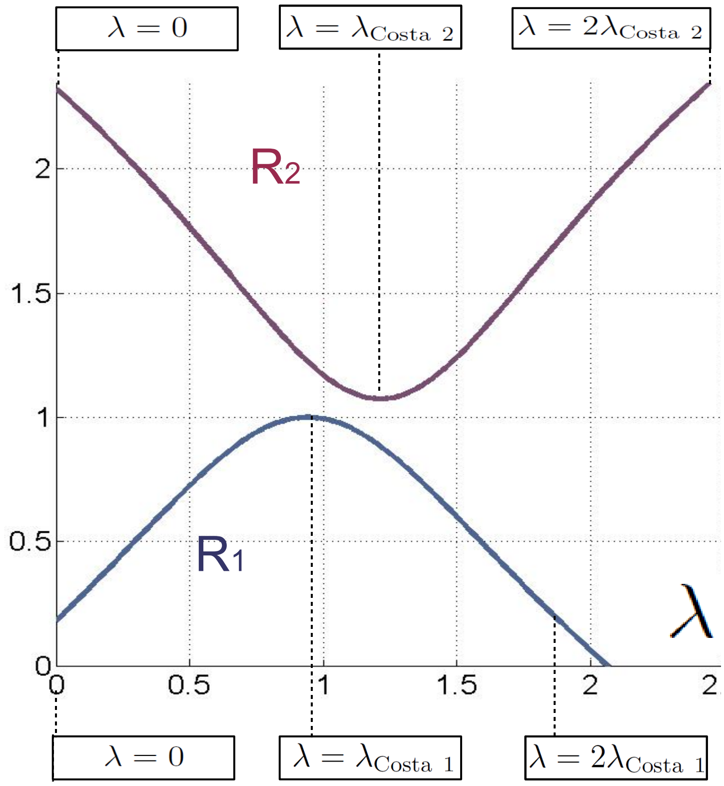

Figure 11. The bound for R_1 in Equation 3.43a (bottom) and the bound for R_2 in Equation 3.43b (top) as a function of $\lambda \in \mathbb{R}$, for $P_1 = P_2 = 6$, $b = \sqrt{2}$, $a = \sqrt{0.3}$, $\alpha = 0.5$.

trade-off among the rates: $\lambda = \lambda_{Costa\ 1}$ achieves the maximum for R_1, but it achieves close to the minimum for R_2. This observation will help in understanding why scheme (E) does not perform well in certain parameter regimes as will be pointed out in Remark 3.7.2.

3.7.1 New capacity results for the C-CIFC.

Theorem 3.7.1. ***Capacity in the "primary decodes cognitive" regime.*** *When $|b| > 1$ and*

$$P_2|1 - a|b||^2 \geq (|b|^2 - 1)(1 + P_1 + |a|^2 P_2) - P_1 P_2 \big|1 - a|b|\big|^2, \tag{3.48a}$$

$$P_2|1 - a|b||^2 \geq (|b|^2 - 1)(1 + P_1 + |a|^2 P_2 + 2\mathrm{Re}\{a\}\sqrt{P_1 P_2}), \tag{3.48b}$$

the "strong interference" outer bound of Th. 3.4.2 is tight and achieved by scheme (E).

The "primary decodes cognitive" regime, illustrated in Figure 12 in the $(a, |b|)$-plane for $a \in \mathbb{R}$ and $P_1 = P_2 = 10$, covers parts of the "strong interference" regime $|b| > 1$ where capacity was not known. It also shows that the scheme in Equation 3.43 (i.e., scheme (E)) is capacity achieving for part of the "very strong interference" region in Equation 3.6, thus providing an alternative capacity achieving scheme to superposition coding (24) (i.e., scheme (D)).

Proof. We compare the achievable scheme (E) in Section 3.6.5 with the "strong interference" outer bound of Th. 3.4.2. Scheme (E) for $|b| > 1$, $\lambda = \lambda_{Costa\ 1}$ and the assignment in Equation 3.43. This achieves Equation 3.8a=Equation 3.43a and Equation 3.8c=Equation 3.43c

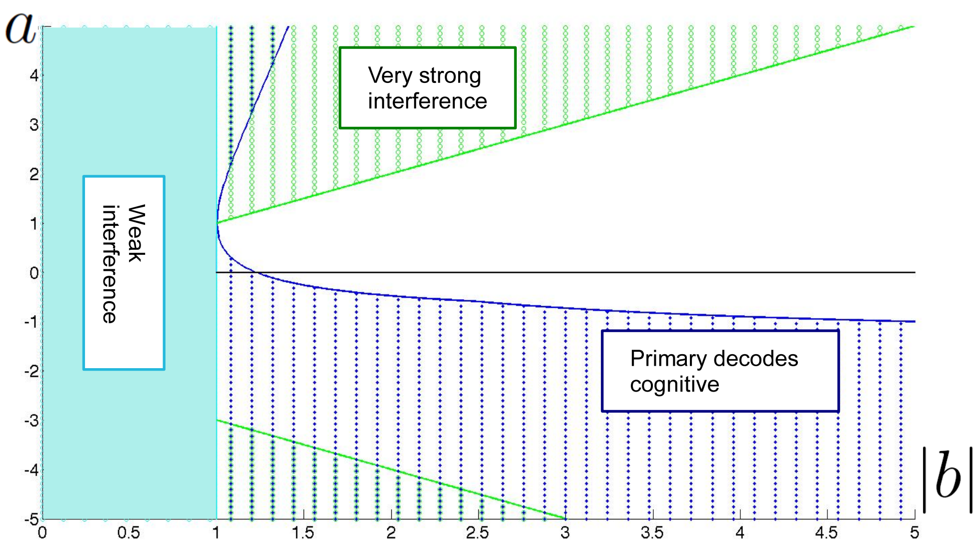

Figure 12. A representation of the capacity result of Th. 3.7.1 for a G-CIFC with $a \in \mathbb{R}$, $P_1 = P_2 = 10$ and $(a, |b|) \in [-5, 5] \times [0, 5]$.

(and Equation 3.8b is redundant). Therefore the "strong interference" outer bound is achievable when (Equation 3.43a+Equation 3.43b)$\geq$Equation 3.8b, i.e. when

$$\mathcal{C}(\alpha P_1) = f\left(a + \sqrt{\tfrac{\bar{\alpha} P_1}{P_2}}, 1; \lambda_{\text{Costa }1}\right) \geq f\left(\tfrac{1}{|b|} + \sqrt{\tfrac{\bar{\alpha} P_1}{P_2}}, \tfrac{1}{|b|^2}; \lambda_{\text{Costa }1}\right), \ \forall \alpha \in [0,1],$$

$$\iff \quad \alpha P_1 + |l_{\text{Costa}} a|^2 P_2 - \frac{\left| |b| \alpha P_1 + \lambda \left(P_2 + \sqrt{\bar{\alpha} |b|^2 P_1 P_2}\right)\right|^2}{|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha} |b|^2 P_1 P_2} + 1} \geq \frac{\alpha P_1}{\alpha P_1 + 1}, \ \forall \alpha \in [0,1],$$

$$\iff \quad \left(\frac{\alpha P_1}{\alpha P_1 + 1}\right)^2 \frac{Q(\alpha)}{1 + |b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha} |b|^2 P_1 P_2}} \geq 0, \ \forall \alpha \in [0,1], \qquad (3.49a)$$

where

$$Q(\alpha) \triangleq P_2\big|1-a|b|\big|^2(\alpha P_1+1)-(|b|^2-1)\Big(P_1+|a|^2P_2+2\mathrm{Re}\{a\}\sqrt{\bar{\alpha}P_1P_2}+1\Big).$$

Clearly the condition in Equation 3.49a is verified if for all $\alpha \in [0,1]$ we have $Q(\alpha) \geq 0$. $Q(\alpha)$ is a quadratic function in $x=\sqrt{1-\alpha}$ of the form $c_1x^2+c_2x+c_3$ with $c_1=-P_1P_2|1-a|b||^2 \leq 0$, which implies that $Q(\alpha)$ is concave in α. Hence, the inequality in Equation 3.49a is verified for every $\alpha \in [0,1]$ if it is verified for $\alpha=1$ and $\alpha=0$. The condition $Q(0) \geq 0$ corresponds to Equation 3.48b while the condition $Q(1) \geq 0$ corresponds to Equation 3.48a. □

Remark 3.7.2. Previous capacity results for the G-CIFC imposed conditions on the channel parameters that lent themselves well to "natural" interpretations. For example, the "weak interference" condition $I(Y_1;X_1|X_2) \geq I(Y_2;X_1|X_2)$ of (26) in Equation 3.2 suggests that decoding X_1 at receiver 2, even after having decoded the intended message in X_2, would constrain the rate R_1 too much, thus preventing it from achieving the interference-free rate in Equation 3.8a. The "very strong interference" condition $I(Y_1;X_1,X_2) \geq I(Y_2;X_1,X_2)$ of (24) in Equation 3.4 suggests that requiring receiver 1 to decode both messages should not prevent achieving the maximum sum-rate at receiver 2 given by Equation 3.8c. A similar intuition about the new "primary decodes cognitive" capacity condition in Equation 3.48 unfortunately does not emerge from the proof of Th. 3.7.1.

To provide some insight on the achievability conditions of Th. 3.7.1, we focus on the condition in Equation 3.48a. When Equation 3.48a is verified, scheme (E) in Section 3.6.5

achieves the "strong interference outer" bound point for $\alpha = 0$ in Equation 3.5: to achieve more points on the "strong interference" outer bound Th. 3.4.2 stricter conditions are necessary; to achieve all the points on the outer bound, both conditions Equation 3.48a and Equation 3.48b must be verified.

A representation of the region where the condition in Equation 3.48a holds is depicted in Figure 13 for the case $a \in \mathbb{R}$ and $P_1 = P_2 = P$ with increasing P, in which case Equation 3.48a becomes

$$P(P+1)|1 - a|b||^2 \geq (|b|^2 - 1)(P + 1 + |a|^2 P). \tag{3.50}$$

We observe that, as P increases, the region where the condition in Equation 3.50 is not verified shrinks. Indeed, as $P \to \infty$, the condition in Equation 3.50 is always verified unless the channel is degraded (i.e., $a|b| = 1$). For a degraded channel with "strong interference", the primary receiver is able to reconstruct Y_1 from Y_2 once W_2 has been decoded, as seen in Equation 3.19. This means that U_{1c} may be decoded at the primary receiver with no rate penalty for the cognitive user. Under this condition, the scheme with a common cognitive message and a private primary one seems a natural choice, reminiscent of the capacity achieving scheme in the degraded BC. Despite this intuition, in a degraded channel with large power P, $\lambda_{Costa\ 1}$ approaches $\lambda_{Costa\ 2}$ (similarly to the case depicted in Figure 11) and thus the maximum of the rate R_1 in Equation 3.43a approaches the minimum of the rate R_2 in Equation 3.43b. This

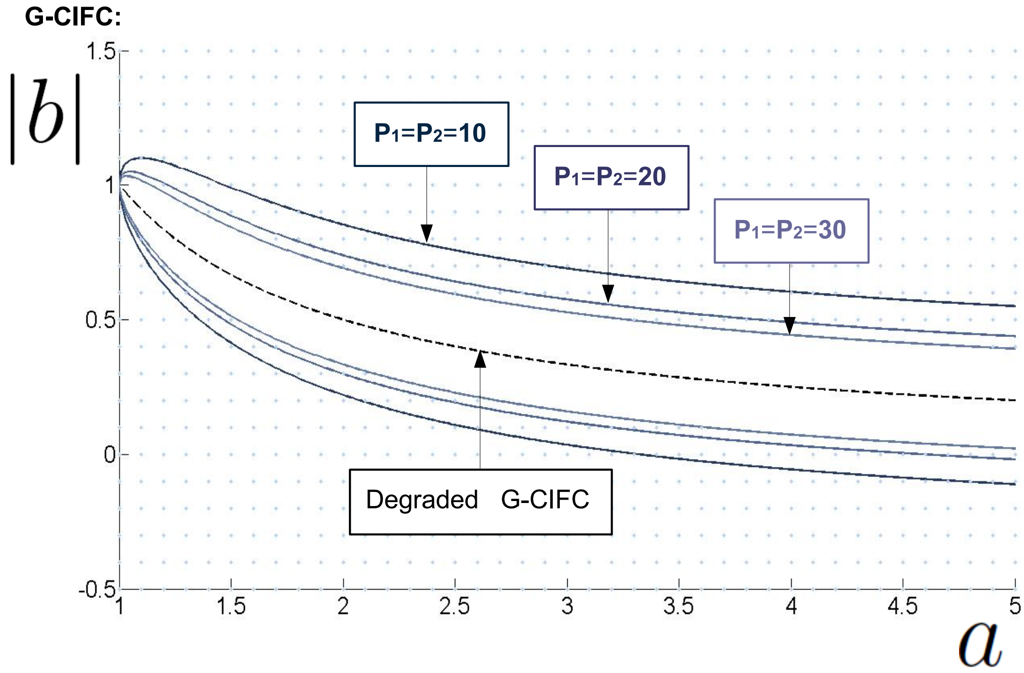

Figure 13. Condition Equation 3.50 for different values of $P_1 = P_2 = P$ for a G-CIFC with $a \in \mathbb{R}$ and $(a, |b|) \in [-0.5, 1, 5] \times [1, 5]$.

rate penalty for the R_2-bound prevents us from achieving the "strong interference" outer bound point for $\alpha = 0$ in Equation 3.5 when $a|b| = 1$.

Another consideration provides further insight on the condition in Equation 3.48a: take a channel where

$$\frac{1}{|b|} = a\frac{P_1}{P_1 + 1}. \tag{3.51}$$

Then, as $P_1 \to \infty$ in Equation 3.51 and for $\alpha = 0$, this condition approaches the degraded condition $a|b| = 1$. For this choice of a, Y_2 may be rewritten as $Y_2 = |b|U_{1c} + Z_2$, so that the R_2-bound of Equation 3.43b for $\alpha = 0$ becomes

$$R_2 \leq I(Y_2, U_{1c}; X_2) = I(U_{1c}; X_2) = \mathcal{C}\left(\frac{P_2}{|b|^2 P_1}\right).$$

This observation reveals an interesting aspect of the RV U_{1c}. U_{1c} is DPC coded against X_2 with the objective to remove (some of) the interference created by X_2 at Y_1. However, decoder 2 is not interested in removing X_2 from Y_2 (it must decode X_2!). Hence, for decoder 2, U_{1c} acts as "side information" when decoding X_2. Now, both U_{1c} and Y_2 contain X_2, but for this specific choice of parameters Y_2 is a noisy version of U_{1c}. This shows why the scheme performs poorly close to the degraded line: there is no gain for receiver 2 from having two observations (i.e., Y_2 and U_{1c}) of the intended message X_2 as they are noisy versions of each other.

3.7.2 New capacity results for the S-G-CIFC.

Theorem 3.7.3. ***C****apacity for S-G-CIFC. For an S-G-CIFC (i.e., $a = 0$) with*

$$|b| \;\leq\; \sqrt{1 + P_2\left(1 - \frac{P_1}{P_1 + 1}\right)} \tag{3.52}$$

or with

$$|b| \geq \sqrt{P_1 P_2 + P_2 + 1} + \sqrt{P_1 P_2} \tag{3.53}$$

Th. 3.5.7 is tight.

Proof. When $|b| \leq 1$, capacity is known so we focus only on the case $|b| > 1$. By setting $a = 0$ in Th. 3.7.1 we obtain that scheme (E) with $\lambda = \lambda_{Costa\ 1}$ achieves the "strong interference" outer bound for

$$(|b|^2 - 1)(1 + P_1) \leq \min\{P_2, P_2(1 + P_1)\} = P_2,$$

which is equivalent to Equation 3.52.

Scheme (E) with $\lambda = 0$ achieves

$$\begin{aligned}
R_1 &\leq I(Y_1; U_{1c}) - I(U_{1c}; X_2) = I(Y_1; U_{1c}) = \mathcal{C}\left(\tfrac{\alpha P_1}{1+\bar{\alpha} P_1}\right), \\
R_2 &\leq I(Y_2, U_{1c}; X_2) = I(Y_2; X_2 | U_{1c}) = \mathcal{C}\left((\sqrt{P_2} + \sqrt{\bar{\alpha}|b|^2 P_1})^2\right), \\
R_1 + R_2 &\leq I(Y_2; X_2, U_{1c}) = \mathcal{C}\left(\alpha|b|^2 P_1 + (\sqrt{P_2} + \sqrt{\bar{\alpha}|b|^2 P_1})^2\right).
\end{aligned}$$

In this case the MIMO-BC outer bound may be achieved when the sum rate outer bound Equation 3.24c is redundant, that is, if

$$\begin{aligned}
& 1 + P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2} \geq \frac{1 + P_1}{1 + \bar{\alpha} P_1}(1 + P_2 + |b|^2 P_1 - \alpha|b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}) \qquad \forall \alpha \in [0,1] \\
\iff \quad & |b|^2 \geq 1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2} \qquad \forall \alpha \in [0,1] \\
\iff \quad & |b|^2 \geq 1 + P_2 + 2\sqrt{|b|^2 P_1 P_2}
\end{aligned}$$

which corresponds to Equation 3.53.

□

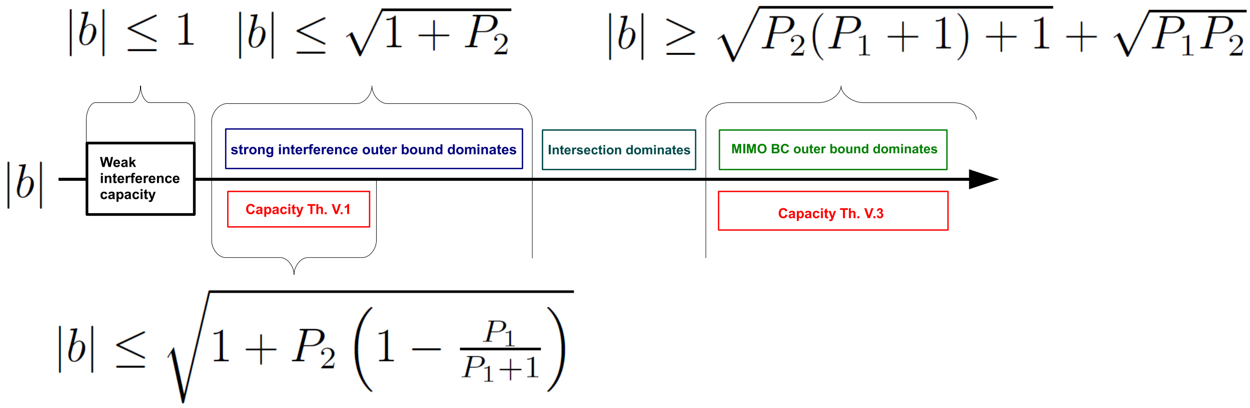

Figure 14. A schematic representation of the capacity results for the S-G-CIFC in Th. 3.7.3.

A representation of the region where capacity is known for the S-G CIFC is depicted in Figure 14. Capacity remains unknown for

$$\sqrt{1+P_2\left(1-\frac{P_1}{P_1+1}\right)} \le |b| \le \sqrt{1+P_2+P_1P_2}+\sqrt{P_1P_2}$$

3.8 Capacity to within a constant gap

In the last couple of years a novel approach to the difficult task of determining the capacity region of a multi-user Gaussian network has been suggested. Rather than proving an equality between inner and outer bounds, the authors of (35) (and references therein) advocate a powerful new method for obtaining achievable rate regions that lie within a bounded distance from capacity region outer bounds, thereby determining the capacity region to *within a constant gap* for any channel configuration. Two measures are used to determine the distance between inner and outer bounds: the additive gap and the multiplicative gap. An additive gap corresponds to a finite difference between inner and outer bound, while a multiplicative gap corresponds to a finite ratio. The additive gap is useful at high SNR, where the difference between inner and outer bound is small in comparison to the magnitude of the capacity region, while the multiplicative gap is useful at low SNR, where the ratio between inner and outer bounds is a more indicative measure of their distance. In this section we show the capacity to within an additive gap of half a bit/s/Hz per real dimension and to within a multiplicative gap of a factor two. We also determine additional constant gap results that suggest which strategies approach the "strong interference" outer bound in different parameter regimes. Since the expressions of the BC-based outer bound of Th. 3.5.2 and of the BC-DMS-based outer bound of Th. 3.5.3 involve many parameters over which to optimize, it is not analytically straightforward to determine conditions for achievability; for this reason we restrict our attention to the "strong interference" outer bound of Th. 3.4.2. These results are derived for the complex-valued channel and rather than for the real-valued channel as done in (72).

Theorem 3.8.1. ***Additive gap.*** *Capacity is known to within half a bit/s/Hz per real dimension.*

Proof. The capacity for weak interference ($|b| \leq 1$) was determined in (26), so we only need to concentrate on the strong interference regime ($|b| > 1$). We show the achievability of the "strong interference" outer bound in Equation 3.5 to within a constant additive gap using the scheme (E) of Section 3.6.3 with the assignment in Equation 3.40. The assignment proposed in Equation 3.40 is inspired by the capacity achieving scheme for deterministic channels in Th.

2.10.1, where we showed that setting $U_c = Y_c$, $c \in \{1,2\}$, is optimal. In a noisy channel, it is not possible to choose $U_c = Y_c$; we mimic this by setting $U_c \sim Y_c$, $c \in \{1,2\}$.

Consider the achievable rate region in Equation 3.39 and note that

$$\begin{aligned}
\mathrm{V}ar[X_1 + aX_2] &= P_1 + |a|^2 P_2 + 2\mathrm{R}e\{a\}\sqrt{\bar{\alpha}P_1P_2},\\
\mathrm{V}ar[|b|X_1 + X_2] &= |b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}.
\end{aligned}$$

The inequality in Equation 3.39c follows by choosing

$$\begin{aligned}
\rho_{pb} &= \mathrm{a}rg \min_{\rho} I(U_1;U_2|X_2)\\
&= \mathrm{a}rg \min_{\rho} |\mathbb{E}[U_1U_2^*|X_2]|^2\\
&= \mathrm{a}rg \min_{\rho} \left||b|P_1\alpha + \rho\sqrt{\sigma_{1pb}^2\sigma_{2pb}^2}\right|^2\\
&= -\min\left\{1, \frac{|b|P_1\alpha}{\sqrt{\sigma_{1pb}^2\sigma_{2pb}^2}}\right\}.
\end{aligned}$$

With $\sigma_{2pb}^2 = 0$ and $\sigma_{1pb}^2 = 1$ in Equation 3.39 we have

$$R_1 \leq \log(1 + \alpha P_1) - \mathrm{GAP}(\alpha), \tag{3.54a}$$

$$R_1 + R_2 \leq \log(1 + \mathrm{V}ar[|b|X_1 + X_2]) - \mathrm{GAP}(\alpha), \tag{3.54b}$$

with GAP(α) bounded as

$$\text{GAP}(\alpha) = \log\left(1 + \frac{\mathrm{V}ar[X_1 + aX_2]}{1 + \mathrm{V}ar[X_1 + aX_2]}\right) \le \log(2) = 1,$$

as claimed. Notice that with $\sigma_{2pb}^2 = 0$, the R_2-bound in Equation 3.39b is equivalent to the sum-rate outer bound in Equation 3.5b and it is thus redundant. □

To prove the multiplicative gap result, we utilize a looser version of Th. 3.5.1 that we present in the next lemma.

Lemma 3.8.2. ***"Piecewise linear strong interference" outer bound.*** *The outer bound of Th. 3.5.1 for $|b| > 1$ is contained in the region $\mathcal{R}^{(PL-SI)}$ defined as:*

$$\begin{aligned} R_1 &\le \mathcal{C}\left(P_1\right), & (3.55a)\\ R_1 + R_2 &\le \mathcal{C}\left((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2\right). & (3.55b)\end{aligned}$$

Proof. The bound in Equation 3.55a (respectively Equation 3.55b) is obtained by considering the maximum value of Equation 3.8a (respectively Equation 3.8c) over $\alpha \in [0, 1]$. □

The region $\mathcal{R}^{(PL-SI)}$ in Equation 3.55 has two Pareto optimal points:

$$\begin{aligned} A &= \left(0, \mathcal{C}\left((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2\right)\right), & (3.56a)\\ B &= \left(\mathcal{C}(P_1), \mathcal{C}((\sqrt{|b|^2 P_1} + \sqrt{P_2})^2) - \mathcal{C}(P_1)\right). & (3.56b)\end{aligned}$$

The point A is on the boundary of the "strong interference" outer bound region $\mathcal{R}^{(SI)}$ of Th. 3.4.2 while Point B has the same R_1-coordinate as the point for $\alpha = 0$ in $\mathcal{R}^{(SI)}$, given by

$$C = \left(\mathcal{C}(P_1), \mathcal{C}(|b|^2 P_1 + P_2) - \mathcal{C}(P_1)\right), \tag{3.57}$$

but lies outside $\mathcal{R}^{(SI)}$. However the two points are no more than one bit away, i.e., $R_2^{(B)} \leq \log(2) + R_2^{(C)}$, as we will show later.

Theorem 3.8.3. ***Multiplicative gap.*** *For a Gaussian C-IFC, the capacity is known to within a factor two.*

Proof. The capacity for weak interference ($|b| \leq 1$) was determined in (26), thus we only need to concentrate on the strong interference regime ($|b| > 1$).

Outer bound:

We use the "piecewise linear strong interference" outer bound of Lemma 3.8.2, in particular we rewrite the outer bound as

$$\begin{aligned} R_2 &\leq \log\left(1 + |b|^2 P_1 + P_2 + 2\sqrt{|b|^2 P_1 P_2}\right) - R_1 \\ &\triangleq R_2^{(PL-SI)}(R_1), \end{aligned} \tag{3.58}$$

for $R_1 \in [0, \log(1 + P_1)]$.

Achievability to within a factor two: Consider the following TDMA strategy. The rate-point

$$(R_1, R_2) = (\log(1 + P_1), 0),$$

is achievable by silencing the primary transmitter, while the rate-point A in Equation 3.56a is achievable by beamforming. Hence, the following region is achievable by time sharing

$$\begin{aligned} R_2 &\leq \left(1 - \frac{R_1}{\log(1+P_1)}\right)\log(1 + (\sqrt{|b|^2 P_1} + \sqrt{P_2})^2) \\ &\triangleq R_2^{(tdma)}(R_1). \end{aligned} \tag{3.59}$$

The multiplicative gap is given by the smallest $M \geq 1$ for which

$$M R_2^{(tdma)}(R_1/M) \geq R_2^{(PL-SI)}(R_1), \tag{3.60}$$

that is

$$\left(1 - \frac{R_1}{M\log(1+P_1)}\right) M \log(1 + (\sqrt{|b|^2 P_1} + \sqrt{P_2})^2) - \log\left(1 + (\sqrt{|b|^2 P_1} + \sqrt{P_2})^2\right) + R_1 \geq 0 \tag{3.61}$$

The LHS of Equation 3.61 is a linear function of R_1 and thus has at most one zero. From this, it follows that the inequality in Equation 3.61 is verified for every $R_1 \in [0, \log(1+P_1)]$ if it is verified at the boundary points of the interval. For $R_1 = 0$, the inequality is verified for $M \geq 1$ while for $R_1 = \log(1+P_1)$ it is verified if $M \geq 2$; thus the smallest M for which Equation 3.61 is verified for all channels is $M = 2$. □

We remark here that we consider the multiplicative gap as the ratio of the outer bound over the inner bound; as originally introduced in (36) the multiplicative gap is defined as the ratio between the inner bound over the outer bound.

What	Scheme	Regime	gap
perfect interference cancelation	(E)	$P_2(1+\lvert a\rvert^2-2\mathrm{Re}\{a\}\lvert b\rvert) \geq$ $(b^2-1)(P_1+1)-P_1P_2\lvert 1-\lvert a\rvert b\rvert^2$	.5
non perfect interference cancelation	(E)	$\lvert b\rvert>1$ and $\lvert b\rvert^2P_1 \leq P_2$	1.87
cognitive	(A)	$\lvert b\rvert \leq 1$ and $\lvert b\rvert^2P_1 > P_2$	1
broadcasting	(A)	$\lvert b\rvert > 1$ and $\lvert b\rvert^2P_1 > P_2$	1.5
interference stripping	(D)	$\lvert a\rvert \geq 1$, $\lvert b\rvert > 1$ and $\lvert b\rvert^2P_1 \leq P_2$	1.5

A schematic plot of the proofs of Th. 3.8.1, Th. 3.8.3 and Lemma 3.8.2 is provided in **??**. The green hatched area represents the achievable rate region with scheme (E) in Equation 3.39 which lies to within half a bit/s/Hz per real dimension from the "strong interference" outer bound of Equation 3.5, illustrated by a solid blue line. The green cross-hatched area represents the achievable rate region with time sharing in Equation 3.59 while the green dashed line is the region in Equation 3.59 multiplied by a factor two, which contains the "piecewise linear strong interference outer bound" in Equation 3.55, illustrated by a dotted blue line.

3.8.1 Additional constant gap results

In this section we provide additional additive gap results for specific subsets of the parameter region. Our aim is to provide insights on the relationship between inner and outer bounds for the region where capacity is still unknown.

Corollary 3.8.4. *The additive gaps between inner and outer bound in 3.8.1 are achievable under the prescribed conditions.*

Proof. The proof is detailed in the following sections. □

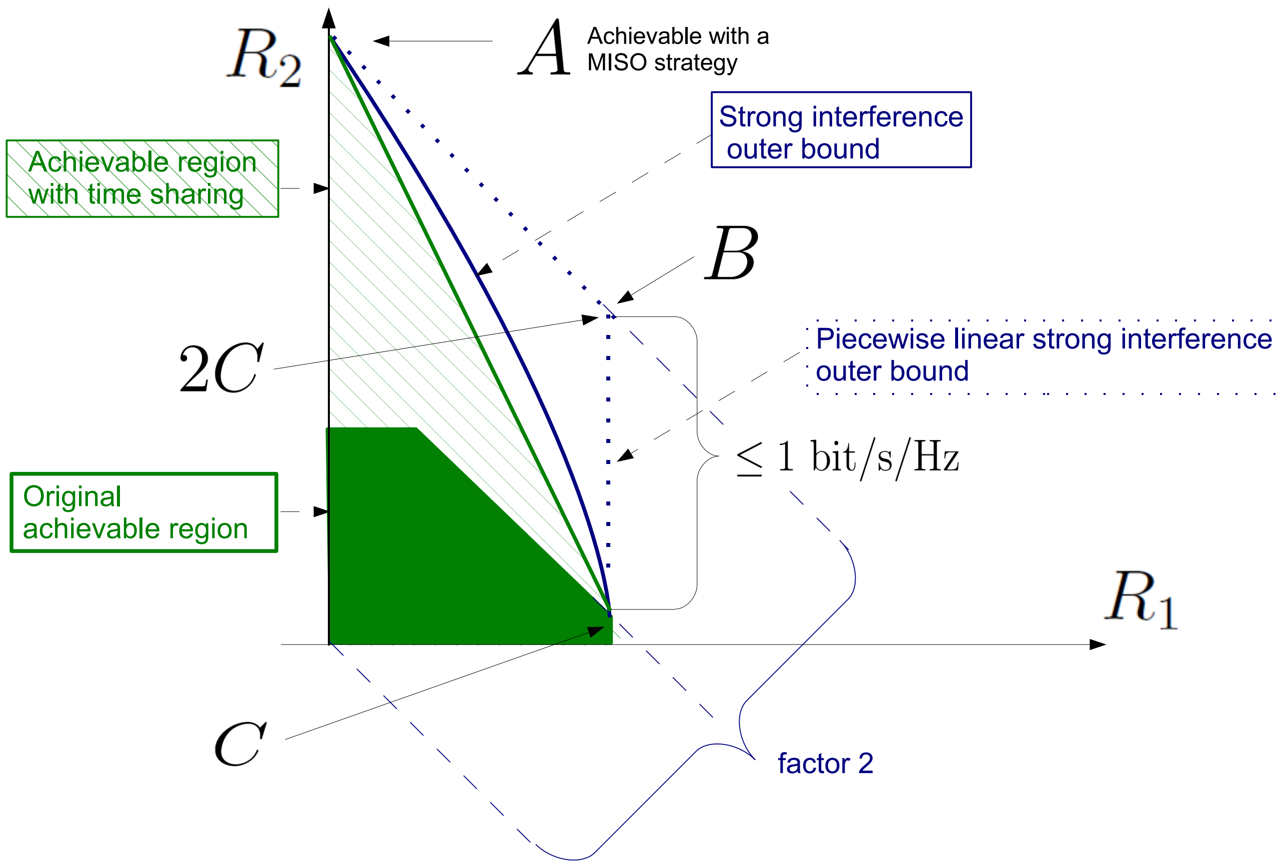

Figure 15. A graphical representation of Th. 3.8.1 and Th. 3.8.3.

In particular we consider four transmission strategies and show where they achieve capacity to within a constant gap:

- **P**erfect interference cancelation. By using the scheme (E) with Costa's DPC we can achieve the "strong interference" outer bound to within a constant gap in a larger parameter region than the "primary decodes cognitive" regime, where it achieves capacity. Scheme (E) with Costa's DPC achieves the "strong interference" outer bound to within a constant gap in a larger parameter region than the "primary decodes cognitive" regime, where it achieves capacity. In this region, both the additive and the multiplicative gaps are smaller than that of Th. 3.8.1 respectively.
- **N**on perfect interference cancelation. The scheme (E) with a specific DPC strategy achieves the "strong interference" outer bound to within a constant gap when the SNR is larger the INR at the primary receiver. The choice of DPC differs from Costa's and it favors the decoding of the common cognitive message at the primary decoder and enhances the performance for channel parameters close to the degraded G-CIFC.
- **C**ognitive broadcasting. When the INR is larger that the SNR the primary receiver, scheme (A) achieves a constant gap from the outer bound in both the "weak" and the "strong interference" regime. In this scheme the primary transmitter is silent and the cognitive transmitter acts as a broadcast trasmitter.
- **I**nterference stripping. Scheme (D) achieves the "strong interference" outer bound to within a constant gap in a larger parameter region than the "very strong interference"

regime, where it achieves capacity. In this scheme both decoders decode both messages as in a compound MAC.

In the following we use the fact that point B in Equation 3.56b is to within one bits/s/Hz and a factor two from point C in Equation 3.57. This is the case as, for the additive gap, $R_1^{(B)} = R_1^{(C)}$ and

$$R_2^{(B)} - R_2^{(C)} = \mathcal{C}\left(\frac{2\sqrt{|b|^2 P_1 P_2}}{1+|b|^2 P_1 + P_2}\right) \leq \mathcal{C}\left(\sqrt{\frac{P_2}{P_2+1}}\right) \leq \log(2) = 1, \tag{3.62}$$

where we use the fact that $R_2^{(B)} - R_2^{(C)}$ has a maximum in $|b|^2 P_1 = P_2 + 1$.

A representation of "strong interference" outer bound and the "piecewise linear strong interference" outer bound is shown in Figure 15. The "strong interference" outer bound coincides with the "piecewise linear strong interference" outer bound at point A and the largest distance between the two outer bounds is attained between points B and C. This figure also introduces a new corner point of the inner bound: point D, the inner bound point with the largest R_2 rate when $R_1 = \mathcal{C}(P_1) - \Delta_1$.

3.8.1.1 Perfect interference cancellation

In the proof of Th. 3.7.1 we have seen that under condition Equation 3.48a it is possible achieve point C in Equation 3.57 with scheme (E) with Costa's DPC. This result may be used to show achievability of the "strong interference" outer bound to within half a bit/s/Hz per real dimension.

Theorem 3.8.5. *If condition in Equation 3.48a holds, the "strong interference" outer bound of Th. 3.4.2 is achievable to within half a bit/s/Hz per real dimension.*

Proof. Under the condition in Equation 3.48a, point C is achievable. This point lies to within half a bit/s/Hz per real dimension from the outer bound. □

3.8.1.2 Non perfect interference cancellation

Although it is not possible to achieve point C using scheme (E) and perfect interference cancellation, it is possible to achieve this point to within a bounded distance using non perfect interference cancellation in the strong interference ($|b| > 1$) and strong signal ($P_2 \geq |b|^2 P_1$) regimes.

Theorem 3.8.6. *When* $|b| > 1$ *and* $P_2 \geq |b|^2 P_1$*, the outer bound of Th. 3.4.2 may be achieved to within* 1.87 *bits/s/Hz per real dimension.*

Proof. To prove this theorem we show the achievability of a point D in Figure 15 which lies at a bounded distance from point C using scheme (E) in Equation 3.43 for $\alpha = 0$. Figure 15 shows the different additive gaps between inner and outer bound points in the following proof. If equation Equation 3.43a is tight there are two possible scenarios: the corner point D is determined by 1) the intersection between Equation 3.43c and Equation 3.43a or by 2) the intersection of Equation 3.43b and Equation 3.43a. We choose λ so that both Equation 3.43a and Equation 3.43b lie within a finite distance from $R_1^{(B)}$ and $R_2^{(B)}$ respectively. The sum rate bound Equation 3.43c does not depend on the choice of λ and is always equal to $R_1^{(C)} + R_2^{(C)}$. We divide the proof in two subcases $\mathrm{Re}\{a\} \gtrless |b|^{-1}$.

*S*ub-case $\mathrm{Re}\{a\} \leq |b|^{-1}$: When $P_1 \leq 1$ a gap of 1 bit per dimension is achievable by having both transmitters transmit to receiver 2 at rate $R_2^{(C)}$. In this case the distance along the rate R_2 is zero and on the rate R_1 is $R_1^{(C)} - 0 \leq \log(1+1) < 2$. For $P_1 \geq 1$ let $\lambda = \frac{P_1 - \sqrt{P_1}}{P_1+1} a$, in Equation 3.43. The distance between inner and outer bound for R_1 is

$$\Delta_1 \triangleq R_1^{(C)} - R_1^{(D)} \quad = \log\left(\frac{1+P_1+2|a|^2P_2}{1+P_1+|a|^2P_2}\right) \leq 1,$$

where we have used the inequality $P_2 \geq |b|^2 P_1$. Similarly letting Equation 3.43b hold with equality, we obtain

$$\begin{aligned}
\Delta_2 \quad & \triangleq R_2^{(C)} - R_2^{(D)} \\
& \leq \max\nolimits_{a:\mathrm{Re}\{a\} \leq |b|^{-1}} \log\left(\frac{\frac{1+2P_2}{1+P_1}}{1+P_2\left|1-\frac{(P_1-\sqrt{P_1})a\,|b|}{1+P1}\right|^2}\right) \\
& \leq \log\left(\frac{(1+P_1)(1+2P_2)}{(1+P_1)(1+P_2+P_1)+2P_2\sqrt{P_1}}\right) \\
& \leq \log\left(\frac{1+2P_2}{1+P_2+P_1}\right) \leq 1,
\end{aligned}$$

where we have used that the expression has a global maximum in $a^* > \frac{1}{|b|}$. The largest gap between the inner bound and B is thus bounded by $\max\{1+\Delta_1, \Delta_1+\Delta_2\} = 2$, and so the overall gap between the specified achievable scheme of Equation 3.43 and the outer bound is within $1+2 = 3$ bits/s/Hz for a complex valued channel.

Sub-case $\mathrm{Re}\{a\} > |b|^{-1}$: When $P_1 \le 3$ a gap of 1 bit per dimension is achievable by having transmitter 1 remain silent (rate $R_1 = 0$) since in this case $R_1^{(B)} - 0 \le \log(3+1)$. When $P_1 > 3$ let $\lambda = \frac{P_1+2\sqrt{P_1}}{P_1+1}$ in Equation 3.43. The gap for R_1 may be bounded as

$$\Delta_1 \triangleq R_1^{(C)} - R_1^{(D)} = \log\left(\frac{1+P_1+5|a|^2P_2}{1+P_1+|a|^2P_2}\right) \le \log(5),$$

while that for the rate R_2 of transmitter 2 may be bounded as

$$\begin{aligned}
\Delta_2 &\triangleq R_2^{(C)} - R_2^{(D)} && \text{(3.63a)}\\
&\le \max_{a:\mathrm{Re}\{a\}\le|b|^{-1}} \log\left(\frac{1+2P_2}{(1+P_1)\left(P_2\left|1-\frac{P_1+2a|b|\sqrt{P_1}}{1+P1}\right|^2+1\right)}\right) && \text{(3.63b)}\\
&\le \log\left(\frac{(1+P_1)(1+2P_2)}{P_2-4P_2\sqrt{P_1}+4P_2P_1+(1+P_1)^2}\right) && \text{(3.63c)}\\
&\le \log\left(\frac{(1+P_1)(1+2P_2)}{2P_1P_2+(1+P_1)^2}\right) && \text{(3.63d)}\\
&\le \log\left(\frac{P_1+1}{P_1}\right) && \text{(3.63e)}\\
&\le \log\left(\frac{4}{3}\right), && \text{(3.63f)}
\end{aligned}$$

where Equation 3.63c follows since the expression has a global maximum for $a^{(opt)} < \frac{1}{|b|}$ and Equation 3.63d follows since $4P_1 - 4\sqrt{P_1} + 1 > 2P_1$ for $P_1 > 3$. Finally Equation 3.63e and Equation 3.63f follow since the expression is monotonically increasing in P_2 and decreasing in P_1. As in the sub-case $\mathrm{Re}\{a\} \le |b|^{-1}$, the maximum distance between points C and D is bounded by $\max\{1+\Delta_1, \Delta_1+\Delta_2\} \le \log\left(\frac{20}{3}\right)$ so that the overall gap is bounded by $\log\left(\frac{40}{3}\right) \approx 3.74$ bits/s/Hz for a complex valued channel.

$\square$

We note that in the proof of Th. 3.8.6, non perfect interference cancellation in crucial. The choice $\lambda = \lambda_{Costa\ 1} + \epsilon$, with ϵ an appropriate decreasing function of the transmit powers, yields a finite gap to capacity although in general this choice does not allow the achievability of point C. This choice of λ can be interpreted as follows. From Equation 3.43b, we see that U_{1c} plays the role of side information at receiver 2 when decoding X_2. The DPC coefficient $\lambda = \lambda_{Costa\ 1}+\epsilon$ favors the decoding of U_{1c} at receiver 2 while slightly degrading the rate of user 1 in Equation 3.43a. In particular, the achievable scheme in Section 3.6.5 for $\lambda = \lambda_{Costa\ 1} + \epsilon$ achieves:

$$R_1 \leq Equation\ 3.43a = \log\left(1+\alpha P_1\right) - \log\left(1 + \frac{(\alpha P_1+1)^2 P_2}{\alpha P_1(1+P_1+|a|^2P_2+2\mathrm{Re}\{a\}\sqrt{\bar{\alpha}P_1P_2})}|\epsilon|^2\right),$$

$$R_2 \leq Equation\ 3.43b = \log\left(1+|b|^2P_1+P_2+2\sqrt{\bar{\alpha}|b|^2P_1P_2}\right) +$$

$$+\log\left(\frac{1}{1+\alpha|b|^2P_1} + \frac{(\alpha|b|^2P_1+1)P_2}{\alpha P_1(1+|b|^2P_1+P_2+2\sqrt{\bar{\alpha}|b|^2P_1P_2}))}|\epsilon+\Delta_\lambda|^2\right),$$

where $\Delta_\lambda = \lambda_{Costa\ 1} - \lambda_{Costa\ 2}$. Let ϵ be small and with the same phase as Δ_λ. From the above expressions, we see that the decrease of R_1 is of the order $\mathcal{O}(|\epsilon|^2)$, while the increase of R_2 is of the order $\mathcal{O}(|\epsilon|)$. *This demonstrates how one can trade residual interference at the cognitive receiver for rate improvement at the primary receiver,* which makes use of U_{1c} to decode its own message.

A natural question at this point is whether it is possible to achieve capacity by using the strategy in Section 3.6.5 with some choice of $\lambda = \lambda_{Costa\ 1} + \epsilon$. Unfortunately, the answer is negative:

Lemma 3.8.7. *The "strong interference" outer bound of Th. 3.4.2 can be achieved by scheme (E) in Equation 3.43 only in the "primary decodes cognitive" regime of Th. 3.7.1.*

Proof. This result is shown by observing that only one choice of α and λ in Equation 3.43 achieves both the sum rate bound in Equation 3.5b and the R_1 bound in Equation 3.5a–the choice which corresponds to the "primary decodes cognitive" regime of Th. 3.7.1. To distinguish the parameters α in the inner and outer bound, let $\alpha^{(out)}$ be the α parameter in the outer bound in Equation 3.5 and $\alpha^{(in)}$ be the α parameter in Equation 3.43.

Consider the region Equation 3.5 for a fixed $\alpha^{(out)}$. We first notice that to achieve the sum rate outer bound in Equation 3.5b we have to pick the $\alpha^{(in)}$ in Equation 3.43 such that $\alpha^{(in)} \leq \alpha^{(out)}$ (because the expressions are monotonically decreasing in α). On the other hand, the maximum rate R_1 in Equation 3.43a for $\alpha^{(in)} \leq \alpha^{(out)}$ is always smaller than the outer bound for R_1 in Equation 3.5, with equality only if $\alpha^{(in)} = \alpha^{(out)}$ and $\lambda = \lambda_{Costa\ 1}$. So to achieve both the sum rate outer bound and the R_1 rate outer bound we must have $\alpha^{(in)} = \alpha^{(out)}$ and $\lambda = \lambda_{Costa\ 1}$, which are the specific assignments considered in Th. 3.7.1. $\square$

3.8.2 Further considerations on the utility of partial interference "pre-cancellation"

In spite of the result of Lemma 3.8.7, in the following lemma we show that the largest inner bound region with the scheme of Section 3.6.5 is achieved by $\lambda \neq \lambda_{Costa\ 1}$. Finding the λ that minimizes the distance between the inner and outer bounds is analytically involved. Instead, we consider the simpler problem of determining the value of λ that maximizes the sum rate.

Lemma 3.8.8. *When $|b| > 1$ and the "very strong interference" condition is not satisfied, setting λ to a solution of*

$$\begin{aligned}
&- P_2(|b|^2 H_2^{-1} - H_1^{-1} + \frac{H_2^{-1} - H_1^{-1}}{\alpha P_1})|\lambda|^2 + \\
&2\left(\sqrt{\bar{\alpha} P_1 P_2}(|b|^2 H_2^{-1} - H_1^{-1}) + P_2(|b| H_2^{-1} - \mathrm{Re}\{a\} H_1^{-1})\right)\mathrm{Re}\{\lambda\} + \\
&\alpha P_1(|b|^2 H_2^{-1} - H_1^{-1}) = 0. \qquad (3.64)
\end{aligned}$$

for

$$H_1 = \mathbb{E}[|Y_1|^2] = 1 + |a|^2 P_2 + P_1 + 2\mathrm{Re}\{a\}\sqrt{\bar{\alpha} P_1 P_2},$$

$$H_2 = \mathbb{E}[|Y_2|^2] = 1 + P_2 + |b|^2 P_1 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2},$$

yields the largest achievable sum rate in the scheme of Section 3.6.5 for a given $\alpha \in [0, 1]$.

Proof. This result is established by observing that Equation 3.43c is the maximum achievable sum rate for a fixed α. Then, the λ that maximizes the sum rate must satisfy $Equation\ 3.43a + Equation\ 3.43b = Equation\ 3.43c$, that is,

$$f\Big(a + \sqrt{\frac{\bar{\alpha}P_1}{P_2}}, 1;\ \lambda\Big) = f\Big(\frac{1}{|b|} + \sqrt{\frac{\bar{\alpha}P_1}{P_2}}, \frac{1}{|b|^2};\ \lambda\Big). \tag{3.65}$$

After some algebra, we rewrite the condition in Equation 3.65 as in Equation 3.64.

Notice that there always exist real-valued solutions (i.e., $\lambda \in \mathbb{R}$) for in Equation 3.64. Indeed, when the "very strong interference" condition is not verified, and $|b| \geq 1$, we have $H_1 \leq H_2$ for all $\alpha \in [0, 1]$, and thus it follows that: the coefficient of $|\lambda|^2$ is negative, the coefficient of $\mathrm{R}e\{\lambda\}$ is positive, and the constant term is positive. By choosing $\mathrm{I}m\{\lambda\} = 0$, the equation Equation 3.64 reduces to the quadratic function in $\mathrm{R}e\{\lambda\}$, which has positive definite determinant and thus has at least one real-valued solution. □

3.8.2.1 Cognitive broadcasting

The outer bound Thm 3.4.1 is achievable in "weak interference": the capacity achieving scheme in this regime is scheme (B) in Section 3.6.2 and it employs Costa's DPC at the cognitive transmitter to "pre-cancel" the known interference generated by the primary user. While capacity is known in this regime, we show that the very simple broadcast strategy of scheme (A) in Section 3.6.1 achieves capacity to within a constant gap from the outer bound when the INR is larger than the SNR at the primary receiver (i.e $|b|^2 P_1 > P_2$). When the INR is larger than the SNR at the primary receiver, scheme (A) achieves a constant gap from the outer bound in "strong interference" as well. Although the resulting gap does not improve on the result of Th. 3.8.1, this result suggests that, in a general scheme, rate improvement may be obtained by having the cognitive transmitter send part of the primary message.

Theorem 3.8.9. *When* $|b| < 1$ *and* $|b|^2 P_1 \geq P_2$*, the outer bound of Th. 3.4.1 may be achieved within 1 bit/s/Hz per real dimension.*

Consider the strategy (A) in Section 3.6.1 for $|b| \leq 1$. Then since Equation 3.3a and Equation 3.37a are the same for every α there is zero gap for the rate R_1. By considering the difference between Equation 3.3b and Equation 3.37b, the gap for the rate R_2 is bounded as

$$\begin{aligned}
\textit{Equation } 3.3b - \textit{Equation } 3.37b \ &\leq \mathcal{C}\left(|b|^2 P_1 + P_2 + 2\sqrt{\bar{\alpha}|b|^2 P_1 P_2}\right) - \mathcal{C}\left(|b|^2 P_1\right) \\
&\leq \mathcal{C}\left(\frac{P_2 + 2\sqrt{|b|^2 P_1 P_2}}{1+|b|^2 P_1}\right) \\
&\leq \mathcal{C}\left(\frac{3|b|^2 P_1}{1+|b|^2 P_1}\right) \\
&\leq \log(4) = 2.
\end{aligned}$$

Theorem 3.8.10. *When* $|b| > 1$ *and* $|b|^2 P_1 \geq P_2$*, the outer bound of Th. 3.4.2 may be achieved within 1.5 bits/s/Hz per real dimension.*

Proof. Consider scheme (A) in Section 3.6.1 for $|b| > 1$ and $\alpha = \min\{1, 1/P_1\}$ in Equation 3.38: the gap for user 1 is

$$\Delta_1 \triangleq R_1^{(B)} - R_1^{(C)} = \mathcal{C}(\min\{1, P_1\}) \leq \log(2) = 1,$$

while the gap for user 2 (using $P_2 \leq |b|^2 P_1$ and $|b|^2 \geq 1$) is

$$\begin{aligned}\Delta_2 \triangleq R_2^{(B)} - R_2^{(C)} &\leq \mathcal{C}\left(\frac{1+2|b|^2P_1}{(1+P_1)(1+|b|^2\min\{1,P_1\})}\right) \\ &\leq \max\left\{\log\left(\frac{2|b|^2}{1+|b|^2}\right), \log\left(\frac{2}{1+P_1}\right)\right\} \\ &\leq \log(2) = 1.\end{aligned}$$

As shown in Figure 15, the achievable point C in Equation 3.57 is at most at $1+\Delta_1+\Delta_2 \leq 3$ bits from the outer bound. By time sharing between points A and C, we have an achievable rate region that is at most at $\max\{1,3\} = 3$ bits/Hz/s from the outer bound for complex valued channel.

□

3.8.2.2 Interference stripping

Theorem 3.8.11. *When $|a| \geq 1$, $|b| \geq 1$ and $|b|^2 P_1 \leq P_2$, the outer bound of Th. 3.4.2 may be achieved within 1.5 bits/s/Hz per real dimension.*

Proof. We consider scheme (D)'s performance in the "strong interference" regime when $|b^2| > 1, |a|^2 \geq 1$. When we set $\alpha = 1$, it achieves the rate

$$\begin{aligned} R_1 &\leq \mathcal{C}(P_1) \\ R_1 + R_2 &\leq \mathcal{C}\left(\min\{|a|^2P_2 + P_1, P_2 + |b^2|P_1\}\right). \end{aligned}$$

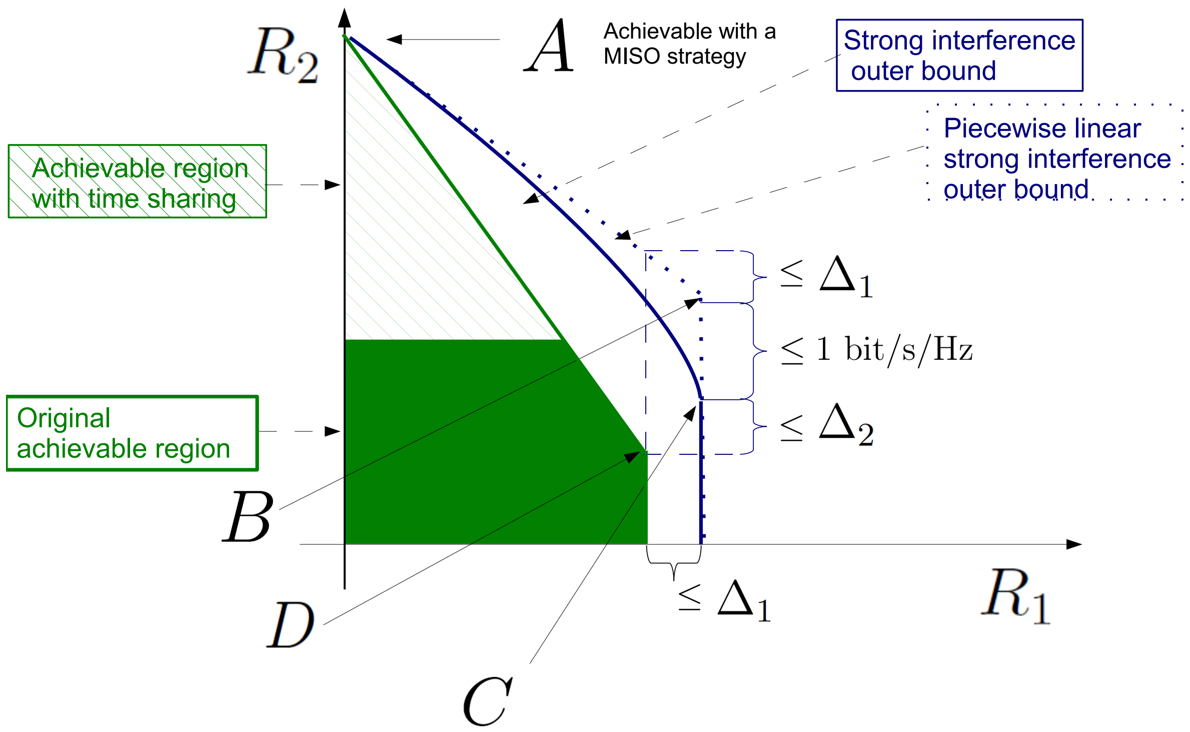

Figure 16. G-CIFC with “strong interference”: the “strong interference” outer bound, the piecewise linear “strong interference” outer bound and the achievable region by time sharing among points A and C;

Referring again to Figure 15, the gap between points B and C may be bounded as

$$\Delta_1 \triangleq R_1^{(B)} - R_1^{(C)} \leq \log(2) = 1,$$

and

$$\begin{aligned}
\Delta_2 \triangleq R_2^{(B)} - R_2^{(C)} &\leq \mathcal{C}\left(\frac{1+|b^2|P_1+P_2}{1+\min\{|a|^2P_2+P_1, P_2+|b|^2P_1\}}\right) \\
&\leq \mathcal{C}\left(\max\left\{1, \frac{1+|b|^2P_1+P_2}{1+|a|^2P_2+P_1}\right\}\right) \\
&\leq \mathcal{C}\left(\max\left\{1, \frac{1+2P_2}{1+|a|^2P_2+P_1}\right\}\right) \\
&\leq \mathcal{C}\left(\max\left\{1, \frac{1+2P_2}{1+|a|^2P_2+P_1}\right\}\right) \\
&\leq \log(2) = 1.
\end{aligned}$$

We thus achieve a rate pair that lies within $1 + \Delta_1 + \Delta_1 = 3$ bits/s/Hz of the outer bound for complex valued channel .

□

3.9 Numerical results

We now revisit each of the previous sections and provide numerical examples of the results therein. In the following we restrict ourselves to real-valued input/output G-CIFC so as to reduce the dimensionality of the search space for the optimal parameter values.

3.9.1 Numerical Results for Section 3.5: Outer bounds

In Section 3.5 we introduced the tightest outer bound for a GCIFC in "strong intereference", obtained as the intersection of the "strong interference" outer bound of Th. 3.4.2 and the BC based outer bound of Th. 3.5.2. This outer bound has a simple closed form expression for the degraded G-CIFC and the S-G-CIFC: Figure 17 and Figure 18 present the result of Corollaries 3.5.5 and 3.5.7 respectively, where the intersection of the "strong interference" outer bound and the BC-based outer bound for the degraded G-CIFC and the S-G-CIFC is derived. Note that we chose two channels where the two bounds intersect for some $R_1 \in (0, \mathcal{C}(P_1)]$ and neither bound strictly includes the other. The two outer bounds coincide at the point A in Equation 3.56a. The maximum rate R_1 in the "strong interference" outer bound and the BC-based outer bound for the S-G-CIFC are the same: in this channel transmitter 2 does not influence the output at receiver 1 and hence full receiver cooperation does not increase the maximum attainable rate R_1.

For a general G-CIFC the intersection between the "strong interference" and the BC-based outer bound has no simple closed form expression. Consequently, it is difficult to determine where one dominates and find their intersection analytically. In Figure 19 we show that the two bounds can intersect up to two times.

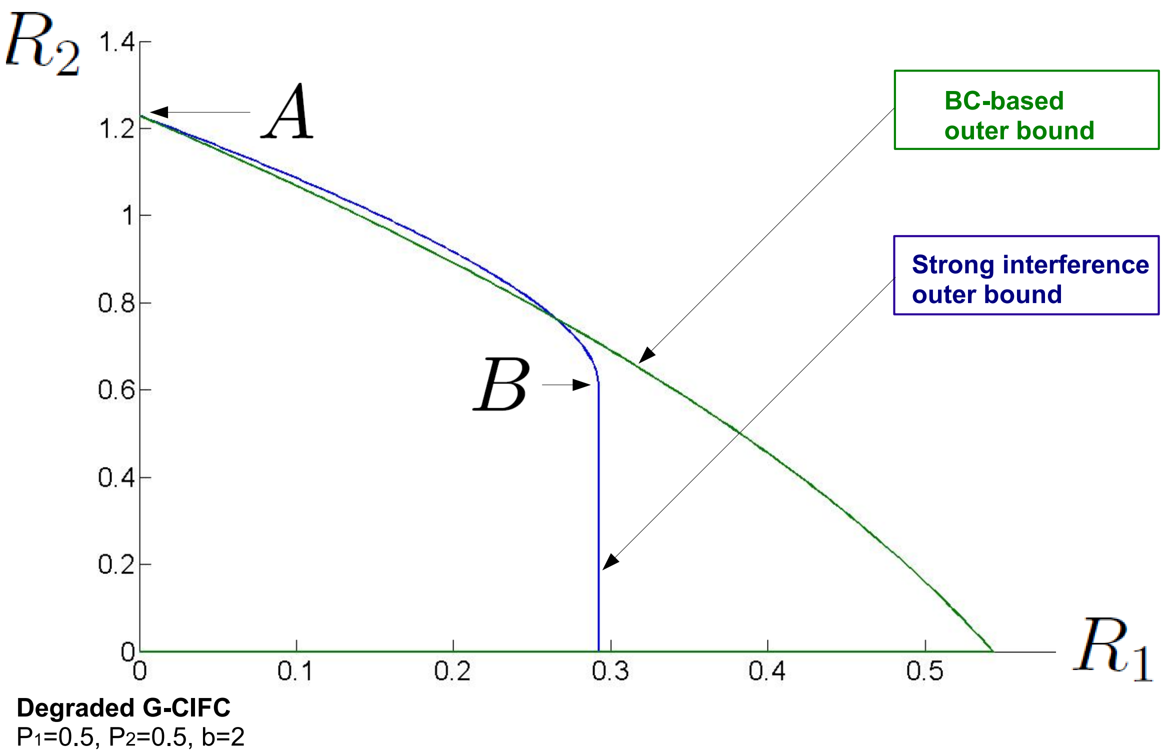

Figure 17. The "strong interference" outer bound and the BC-based outer bound for the degraded G-CIFC.

The outer bounds of Th. 3.5.9 are presented in Figure 20 which shows that these outer bounds may be tighter than either the "strong interference" or the BC-based outer bounds. Unfortunately, in the examples we considered, we did not find an instance where the outer bounds of Th. 3.5.9 were tighter than the intersection of the "strong interference" and the BC-based outer bound. Despite this, we believe that our approach in transforming the channel provides a general, useful tool to derive outer bounds for channels with cognition.

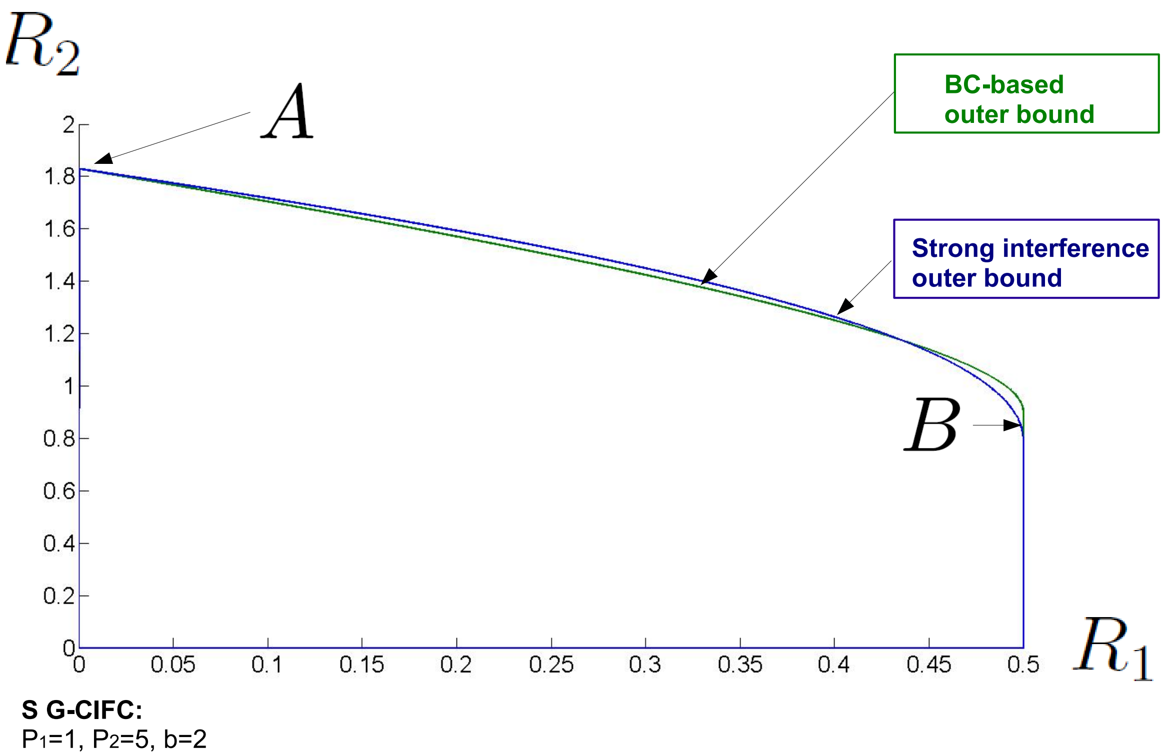

Figure 18. The "strong interference" outer bound and the BC-based outer bound for the S-G-CIFC (right).

3.9.2 Numerical Results for Section 3.6: Inner bounds

In Section 3.6 we introduced the $\mathcal{R}^{(RTD)}$ achievable rate region and derived six sub-schemes from this general inner bound: in the following we plot these sub-schemes for the degraded channel, the S channel and a general G-CIFC. The "strong interference" and the "weak interference" outer bounds are provided for reference. Note that both the achievable rate regions and the outer bounds are expressed as a function of one parameter only, $\alpha \in [0, 1]$, that controls the amount of cooperation between the cognitive and the primary transmitters.

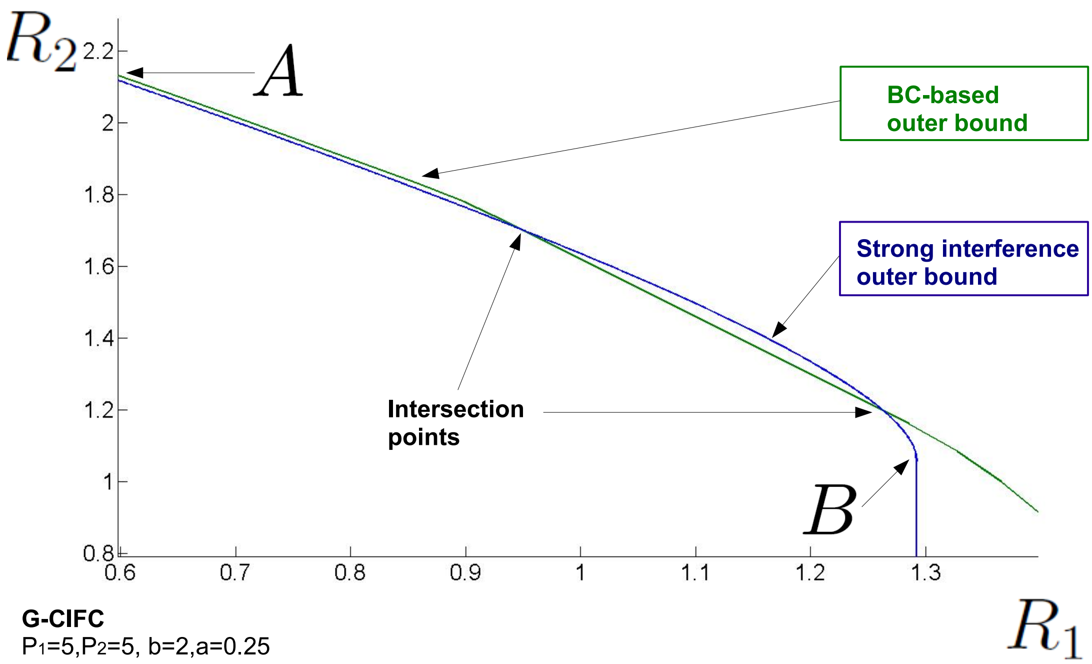

Figure 19. The "strong interference" outer bound and the MIMO BC outer bound for a general G-CIFC in proximity to point C.

We begin by considering the degraded G-CIFC in Figure 21. The scheme that yields the largest achievable rate region in scheme (E) with the choice $\lambda = \lambda_{Costa\ 1}$. Despite its superior performance (to other presented schemes) we may analytically show that this scheme cannot achieve either the "strong interference" or the BC-based outer bound.

Both schemes (A) and (B) treat the interference at noise at receiver 1 and thus the maximum R_1 may be achieved only by silencing transmitter 2. For this reason $R_2 \to 0$ as $R_1 \to \mathcal{C}(P_1)$ for these two schemes.

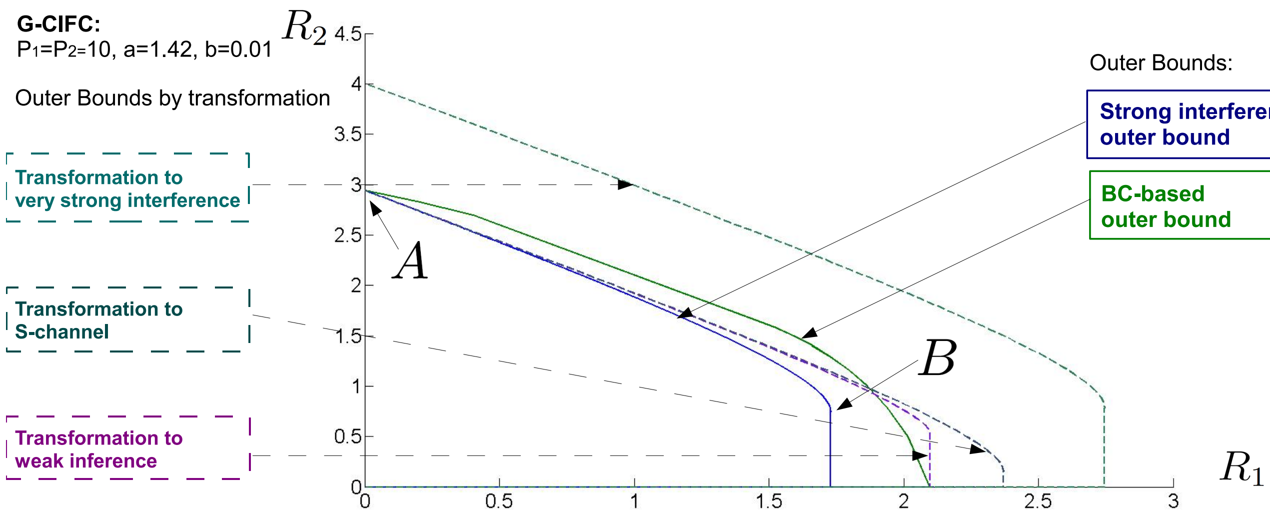

Figure 20. The outer bounds of Lemma 3.5.9 alongside the "strong interference" outer bound and the BC-based outer bound.

The channels parameters are chosen to show how scheme (E) with the choice $\lambda = \lambda_{Costa\ 1}$ achieves the "strong interference" outer bound for a subset of $R_1 \in (0, \mathcal{C}(P_1)]$ where the inequality in Equation 3.49a holds. The figure also shows how, in the S channel, it is possible to achieve the outer bound for $R_1 = \mathcal{C}(P_1)$ with scheme (E) without DPC. This is possible only in this channel, since X_2 does not influence Y_1 and no rate loss occurs at the cognitive receiver by treating the interference as noise. Note that scheme (D) performs the worst among all the achievable schemes: in this scheme the cognitive receiver is required to decode both messages – a very stringent constraint since Y_1 does not contain X_2. In particular, $R_2 \to 0$

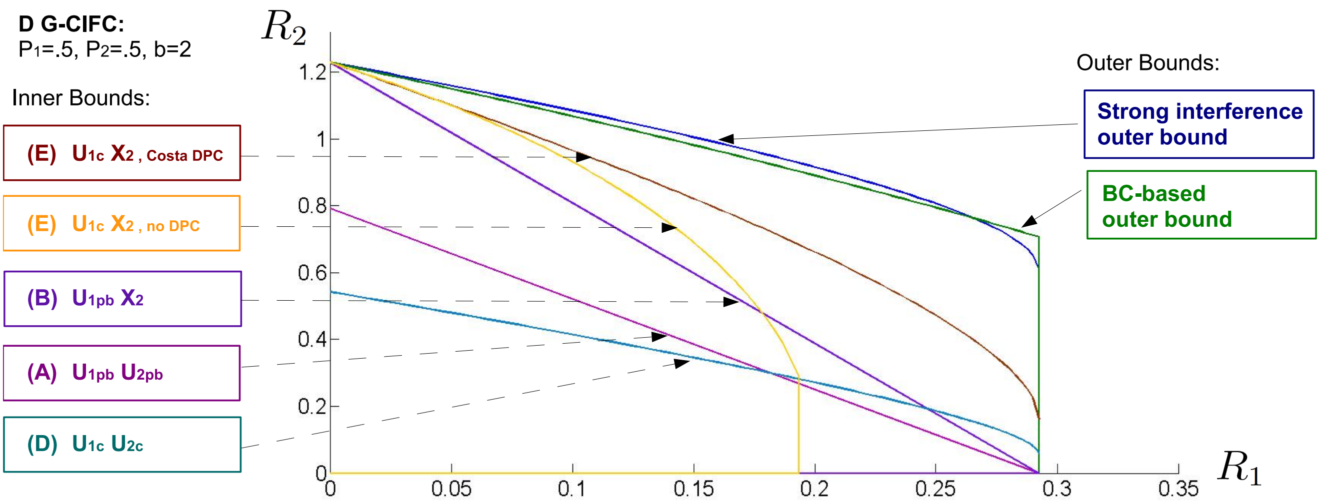

Figure 21. The different schemes of Section 3.6 for the degraded G-CIFC.

when $R_1 \to \mathcal{C}(P_1)$ as in schemes (A) and (B): this is so because $R_1 = \mathcal{C}(P_1)$ may be achieved with scheme (D) only for Y_1 independent of X_2.

A general G-CIFC in Figure 23. In this example, scheme (E) with $\lambda = 0$ performs better than the scheme with $\lambda = \lambda_{Costa\ 1}$ for small R_1 while the opposite is true for large R_1. This is the first instance in which we see that a single choice of λ does not yield the largest inner bound: for small INR, is better for the cognitive user to treat the interference as noise, while for large INR it is more advantageous to perform Costa's DPC.

From Section 3.5.2 we know that, for $|b| > 1$, the cognitive receiver can decode the primary message with no additional rate penalties; this may be observed by comparing scheme (E) with Costa's DPC and scheme (B). The primary message is private in both schemes while

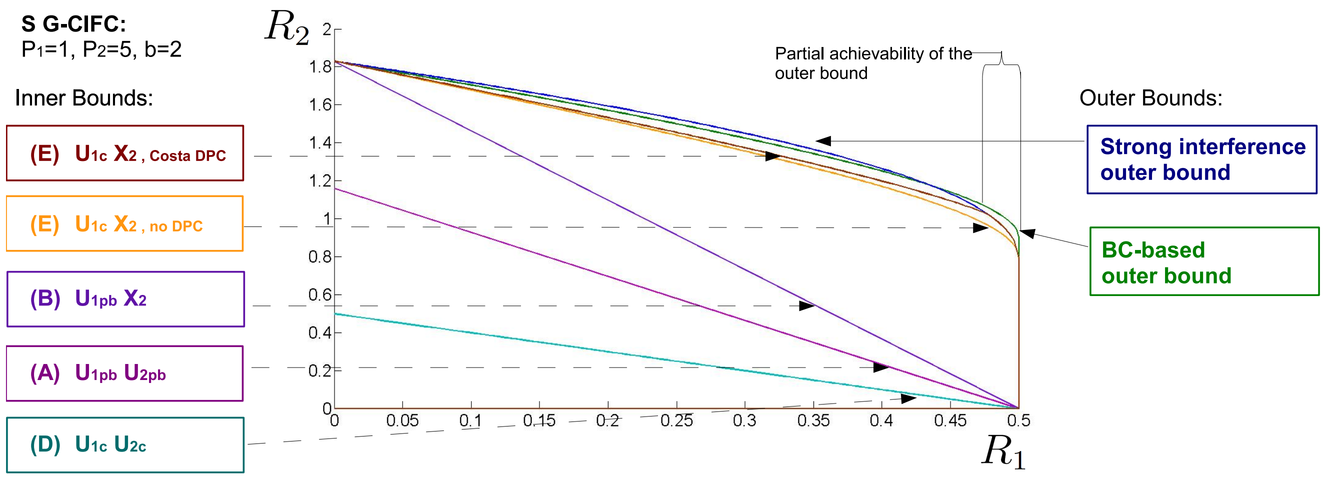

Figure 22. The achievable schemes of Section 3.6 for the S-G-CIFC.

the cognitive message is common in scheme (E) and private in scheme (B). Since the primary receiver can decode the cognitive message at no cost, scheme (E) with Costa's DPC achieves larger rates than scheme (B). When no DPC is used ($\lambda = 0$) in scheme (E), the cognitive receiver observes an equivalent additive Gaussian noise noise of variance $1 + |a|^2 P_2$: for this region rate R_1 is always bounded by $R_1 \leq \mathcal{C}(P_1/(1+|a|^2 P_2))$ and thus scheme (B) outperforms scheme (E) with no DPC in the interval $R_1 \in [\mathcal{C}(P_1/(1+|a|^2 P_2)), \mathcal{C}(P_1)]$.

The scheme (F) in Section 3.6.6 unifies capacity achieving schemes in the "very strong interference" and the "primary decodes cognitive" regimes. It is possible that by unifying the two schemes, we may show achievability in a larger region than the union of the two regimes. Unfortunately determining the achievability conditions in closed form is not straightforward as it requires the optimization of the four parameters in Equation 3.47. In Figure 24 we show

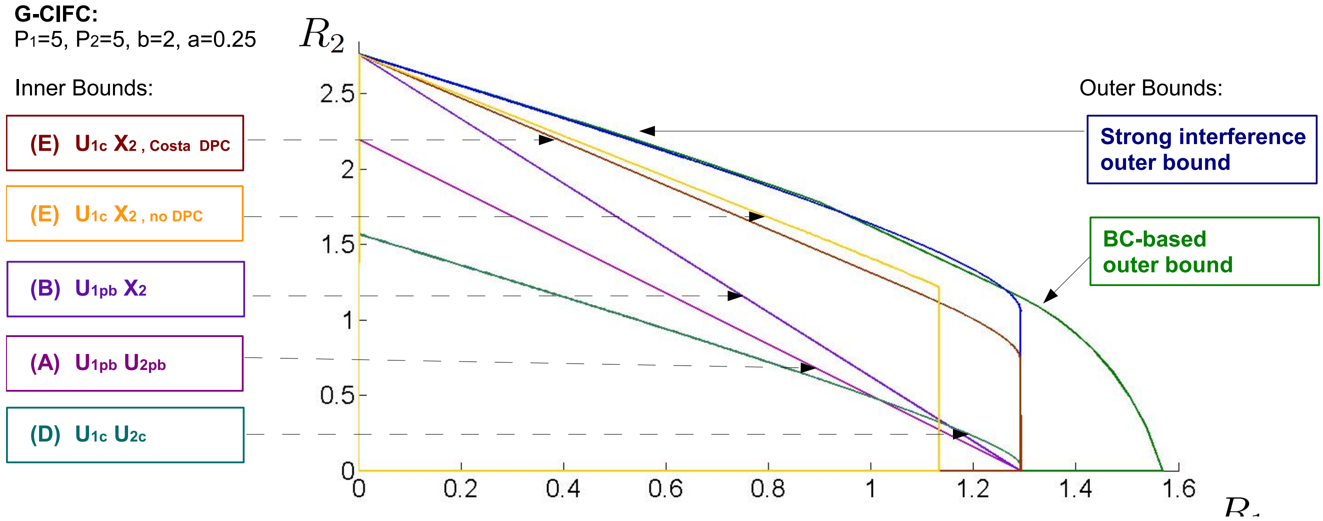

Figure 23. The achievable schemes of Section 3.6 for a general G-CIFC.

through numerical evaluation that scheme (F) indeed achieves a larger region of than the union of the schemes (E) and (D). Whether this scheme achieves capacity for a larger parameter region remains an open question.

3.9.3 Numerical Results for Section 3.7: New capacity results

In Section 3.7 we determine new capacity results for the "primary decodes cognitive" regime both for a general G-CIFC and the S-G-CIFC. In Figure 25 we plot the "primary decodes cognitive" regime in Equation 3.48 for different transmitter powers $P_1 = P_2 = P$. Note that the "weak interference" and the "very strong interference" regimes do not depend on P so their plot does not vary. As the power P increases, the "primary decodes cognitive" region expands from the line $|b| = 1$ to cover a larger region around the degraded line. Interestingly

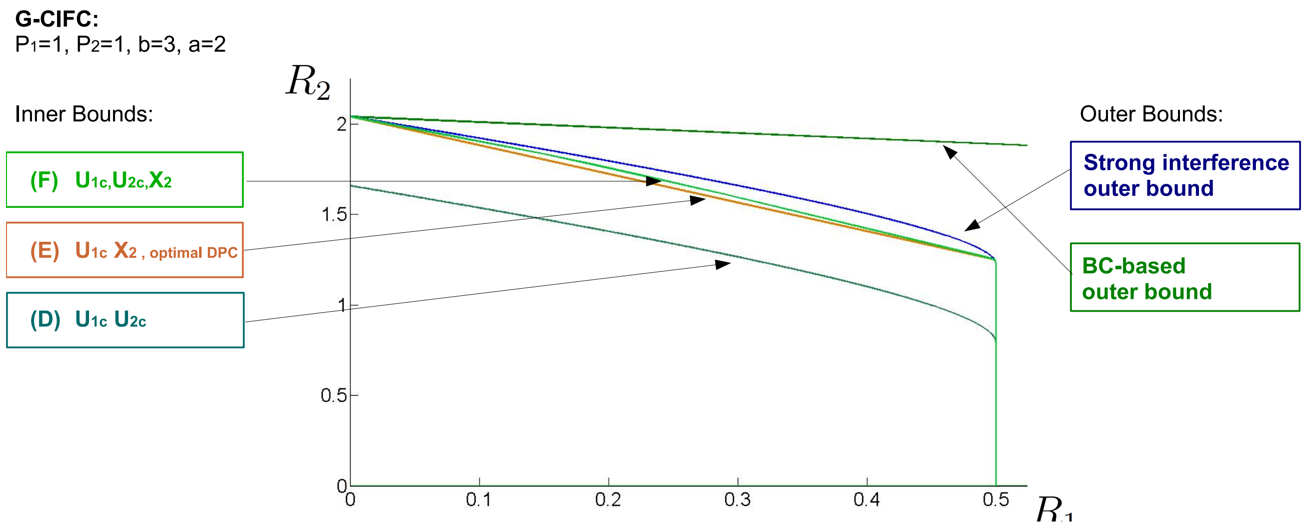

Figure 24. The achievable region of schemes (D), (E) and (F) for a general G-CIFC.

the "primary decodes cognitive" regime intersects with the "very strong interference" regime, thus showing that the "strong interference" outer bound may be achieved with two different transmission schemes for some channels.

In a similar fashion, Figure 26 shows the capacity results of Th. 3.7.3 for the case $P_1 = P_2 = P$ on the plane $P \times |b|$. For equal transmitter powers, the conditions in Equation 3.52 and in Equation 3.53 become

$$|b|^2 \leq \frac{2P+1}{P+1} \approx 2 \tag{3.66a}$$

$$|b|^2 \geq P + \sqrt{P^2+P+1} \approx 2P \tag{3.66b}$$

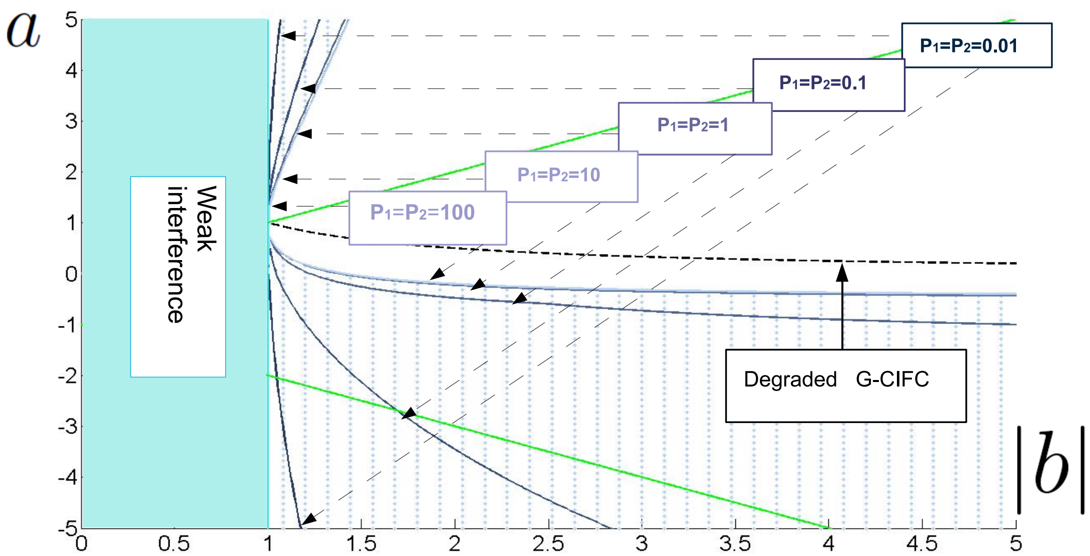

Figure 25. The “primary decodes cognitive” region for different powers $P_1 = P_2 = P$ for a G-CIFC with $a \in \mathbb{R}$ and $(a, |b|) \in [-5, 5] \times [0, 5]$.

and these two asymptotic behaviors are clearly visible in Figure 26.

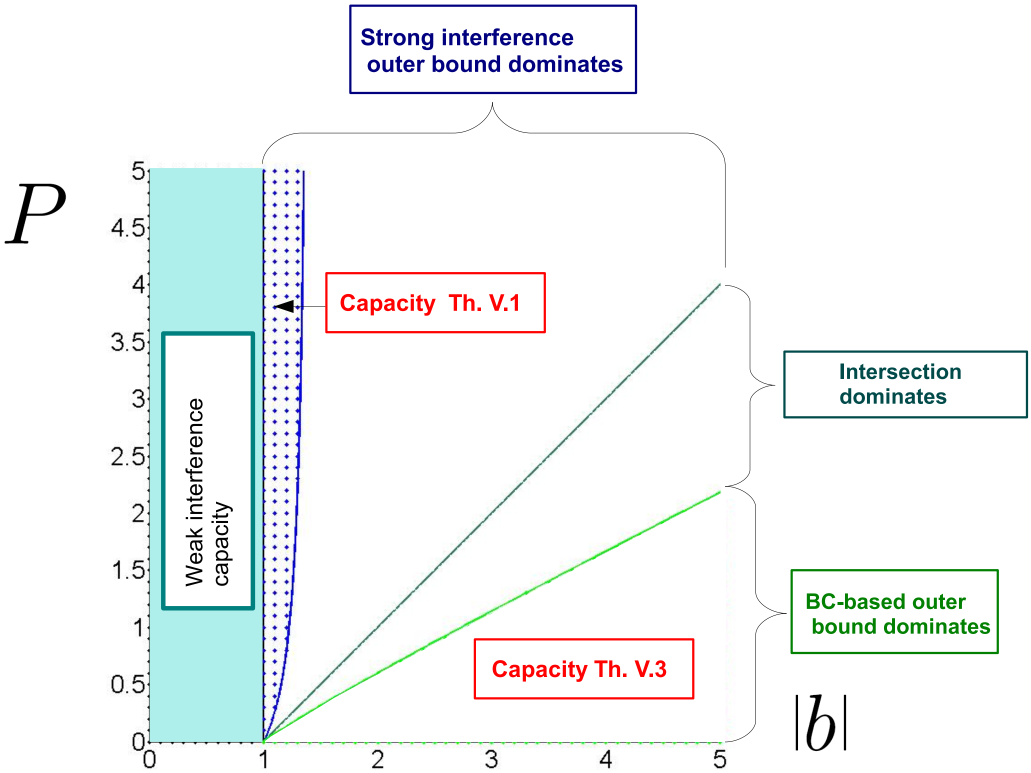

Figure 26. The capacity results for the S-G-CIFC for the case $P_1 = P_2 = P$ for $(P, |b|) \in [0, 5] \times [0, 5]$.

CHAPTER 4

CONCLUSIONS AND FUTURE WORK

In this thesis we studied the cognitive interference channel: a transmission channel model that captures the ability of wireless devices to overhear the transmission taking place over the medium.

In chapter one we focused on the discrete memoryless cognitive interference channel and derived new inner and outer bounds, derived the capacity region for a class of "better cognitive decoding" channels, and obtained the capacity region for the semi-deterministic cognitive interference channel. quantify the rate improvements that can be attained when one transmitter knows the message of another transmitter non-causally. Capacity for this channel was known in the "very weak interference regime" of (26) and in the "very strong interference regime" of (20). General outer bounds were presented in (26) and (25). These outer bounds are expressed as the union over the distribution of different auxiliary RVs for which no cardinality bounds are available. General inner bounds were presented in (25) and (50) to include and generalize the different achievable region proposed in literature. We proposed a new outer bound using an idea originally devised for the broadcast channel in (59). This outer bound does not involve auxiliary RVs and is thus more easily computable. evaluate for a specific channel. This (26) and they coincide We also proposed a new inner bound that generalizes all other known achievable rate regions. for this channel model and it is not clear what relationship exists between them. We propose a scheme that generalizes all the previously known scheme and, in

Section 2.7, (43; 52); it was previously unclear how the performance of the scheme in (43; 52) compared with that of other achievable rate regions. the structure of the outer bound of (26) We determined capacity for a class of channels that we term the "better cognitive decoding" regime. The conditions defining this regime are looser than the "very weak interference condition" of (26) and the "very strong interference condition" of (24) and is the largest region where capacity is known. We also determined the capacity region for the class of semi-deterministic cognitive interference channels. at the cognitive receiver result of Section 2.8 to the all class of channels. in (26) in two different conditions using two different achievable schemes. where capacity was determined in a subclass of the semi-deterministic cognitive interference channel. determine capacity in a specific class of the semi-deterministic cognitive interference channel: the deterministic cognitive interference channel. In this channel model Furthermore, for channels where both outputs are deterministic functions of the inputs, we showed the achievability of our new outer bound. outer bound in (26), is tight for certain channels. The scheme that achieves capacity in the deterministic cognitive interference channel uses Gelf'and-Pinsker binning against the interference created at the primary receiver. This binning is performed by the cognitive encoder for the primary decoder. This feature of the transmission scheme was never known before to be capacity achieving. interesting features of the capacity achieving scheme in the deterministic cognitive interference channel. For this reason in Section 2.11 we present two examples of deterministic cognitive interference channels to provide some insight on the capacity achieving scheme. allowing the coordination of the transmission of primary and cognitive encoders, so to control the interference and take advantage of the extra knowledge at the cognitive transmitter. modern day communication systems is potentially crucial. Only with the understanding of which forms of collaboration are most efficient and beneficial for the users, we can design systems and protocols that perform close to optimal. our understanding of the role of cognition in dealing with interference and cooperation in communication networks. Extensions of the results presented here to Gaussian channels will be presented in (73).

In Chapter 2 we presented outer bounds, inner bounds, and new capacity results for the Gaussian cognitive interference channel. We derived the tightest known outer bound for the cognitive interference channel in “strong interference”, which is based on the capacity of the MIMO BC with degraded message sets. We showed the achievability of this outer bound in the subset of the channel parameter space which we term the “primary decodes cognitive” regime. We also proved capacity to within both an additive and a multiplicative gap, thus providing a characterization of the capacity region in both high and low SNR.

Despite the new results presented, the capacity of th cognitive interference channel remains unknown in general. The achievable we presented provides a comprehensive inner bound that may yield new capacity results: only some specific choices of parameters for this region have been considered so far and we expect that new results may be derived from this region. We have shown that the tightest outer bound for the Gaussian cognitive interference channel in “strong interference” is obtained as the intersection of different bounds. The expression of this outer bound does not have a simple closed form expression except in some special cases like the S and the degraded channels. Even in these two subcases, capacity is not known in general. Another interesting open question is how much rate improvement is attainable with binning at the cognitive encoder: we have shown how dirty paper coding may be used to boost the rate of both the primary and the cognitive user; whether non perfect interference cancellation achieves capacity is still unknown.

Although we have established many significative results for the cognitive interference channel, many interesting problems remain open. We hope that further research in this area will mature into a widely deployed technology that will make a difference in the future development of wireless communication networks.

Acknowledgment The authors would like to thank Prof. Shlomo Shamai and Prof. Gerhard Kramer for insightful discussions on the problem during IZS 2010 and ITA 2010 respectively.

CITED LITERATURE

1. Cisco visual networking index: Global mobile data traffic forecast update.

2. Force, F.: Report of the spectrum efficiency working group. Federal Communications Commision, Technical Report, pages 02–155, 2002.

3. Haykin, S.: Cognitive radio: brain-empowered wireless communications. IEEE Journal on Selected Areas in Communications, 23(2):201–220, 2005.

4. Mitola, J. and G. Q. Maguire, J.: Cognitive radio: making software radios more personal. IEEE Personal Communications Magazine, 6(4):13–18, 1999.

5. Neel, J., Reed, J., and Gilles, R.: Convergence of cognitive radio networks. In Wireless Communications and Networking Conference, volume 4, pages 2250–2255, 2004.

6. Nie, N. and Comaniciu, C.: Adaptive channel allocation spectrum etiquette for cognitive radio networks. Mobile Networks and Applications, 11(6):779–797, 2006.

7. Akyildiz, I., Lee, W., Vuran, M., and Mohanty, S.: NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey. Computer Networks, 50(13):2127–2159, 2006.

8. Kim, H. and Shin, K.: Efficient discovery of spectrum opportunities with mac-layer sensing in cognitive radio networks. Mobile Computing, IEEE Transactions on, 7(5):533–545, May 2008.

9. Kim, H. and Shin, K.: Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks.

10. Ganesan, G. and Ye, L.: Cooperative spectrum sensing in cognitive radio, part II: Multiuser networks. IEEE transactions on wireless communications, 6(6):2214–2222, 2007.

11. Fette, B.: Cognitive radio technology. Newnes, 2006.

12. Cordeiro, C., Challapali, K., and Ghosh, M.: Cognitive PHY and MAC layers for dynamic spectrum access and sharing of TV bands. In Proceedings of the first international

workshop on Technology and policy for accessing spectrum. ACM New York, NY, USA, 2006.

13. Krishnamurthy, S., Thoppian, M., Venkatesan, S., and Prakash, R.: Control channel based MAC-layer configuration, routing and situation awareness for cognitive radio networks. In IEEE Military Communications Conference, 2005. MILCOM 2005, pages 455–460, 2005.

14. Devroye, N., Mitran, P., and Tarokh, V.: Achievable rates in cognitive radio channels. IEEE Trans. Inf. Theory, 52(5):1813–1827, May 2006.

15. Carleial, A.: On the capacity of multiple-terminal communication networks. Ph. D. Thesis Stanford Univ., CA., 1975.

16. Carleial, A.: Interference channels. IEEE Transactions on Information Theory, 24(1):60–70, 1978.

17. Bresler, G. and Tse, D.: The two-user Gaussian interference channel: a deterministic view. Draft of version in Euro. Trans. Telecomm., Volume 19, Issue 4, pp. 333-354, June 2008, 2008.

18. Han, T. and Kobayashi, K.: A new achievable rate region for the interference channel. Information Theory, IEEE Transactions on, 27(1):49–60, Jan 1981.

19. Suh, C. and Tse, D.: Feedback capacity of the Gaussian interference channel to within 1.7075 bits: the symmetric case. arXiv [http://arXiv.org/oai2] (United States), 2009.

20. Maric, I., Yates, R., and Kramer, G.: The capacity region of the strong interference channel with common information. In Proc. Asilomar Conferenece on Signal, Systems and Computers, pages 1737–1741, Nov. 2005.

21. Kramer, G.: Review of rate regions for interference channels. In Communications, 2006 International Zurich Seminar on, pages 162–165, March 2006.

22. Maric, I., Yates, R., and Kramer, G.: The strong interference channel with unidirectional cooperation. The Information Theory and Applications (ITA) Inaugural Workshop, Feb 2006. UCSD La Jolla, CA.

23. Maric, I., Goldsmith, A., Kramer, G., and Shamai, S.: On the capacity of interference channels with a cognitive transmitter. Information Theory and Applications Workshop, 2007, pages 268–273, 29 2007-Feb. 2 2007.

24. Maric, I., Yates, R., and Kramer, G.: Capacity of interference channels with partial transmitter cooperation. Information Theory, IEEE Transactions on, 53(10):3536–3548, Oct. 2007.

25. Maric, I., Goldsmith, A., Kramer, G., and Shamai, S.: On the capacity of interference channels with one cooperating transmitter. Information Theory, IEEE Transactions on, 19(4):405–420, 2008.

26. Wu, W., Vishwanath, S., and Arapostathis, A.: Capacity of a class of cognitive radio channels: Interference channels with degraded message sets. Information Theory, IEEE Transactions on, 53(11):4391–4399, Nov. 2007.

27. Jovicic, A. and Viswanath, P.: Cognitive radio: An information-theoretic perspective. Proc. IEEE International Symposium on Information Theory (ISIT), Washington, U.S.A., pages 2413–2417, July 2006.

28. Liang, Y., Somekh-Baruch, A., Poor, H. V., Shamai, S., and Verdú, S.: Cognitive interference channels with confidential messages. Proc. of the 45th Annual Allerton Conference, abs/0710.2018, 2007.

29. Jiang, J. and Xin, Y.: On the achievable rate regions for interference channels with degraded message sets. Information Theory, IEEE Transactions on, 54(10):4707–4712, 2008.

30. Sridharan, S. and Vishwanath, S.: On the capacity of a class of mimo cognitive radios, Sept. 2007.

31. Avestimehr, A., Diggavi, S., and Tse, D.: A deterministic approach to wireless relay networks. Allerton, Sept., 2007.

32. Gamal, A. E. and Costa, M.: The capacity region of a class of deterministic interference channels. IEEE Trans. Inf. Theory, 28(2):343–346, March 1982.

33. Avestimehr, A., Diggavi, S., and Tse, D.: A deterministic model for wreless relay networks an its capacity, July 2007.

34. Avestimehr, A., Diggavi, S., and Tse, D.: A deterministic approach to wireless relay networks. Proc. Allerton Conf. Commun., Control and Comp., Sep. 2007. Monticello.

35. Avestimehr, A., Diggavi, S., and Tse, D.: Wireless network information flow. Proc. Allerton Conf. Commun., Control and Comp., Sep. 2007. Monticello, IL, U.S.A.

36. Etkin, R., Tse, D., and Wang, H.: Gaussian interference channel capacity to within one bit. IEEE Trans. Inf. Theory, 54(12):5534–5562, December 2008.

37. Avestimehr, A., Diggavi, S., and Tse, D.: Wireless network information flow. Allerton, Sept., 2007.

38. Gou, T., Jafar, S., Aly, S., Kamal, A., Cuff, P., Su, H., El Gamal, A., Haddad, R., Smith, B., Vishwanath, S., et al.: Capacity of a Class of Symmetric SIMO Gaussian Interference Channels within O (1). Arxiv preprint arXiv:0905.1745, 2009.

39. Bandemer, B. and Vazquez-Vilar, G.: On the Sum Capacity of A Class of Cyclically Symmetric Deterministic Interference Channels.

40. Cheng, P., Yu, G., Zhang, Z., Chen, H., and Qiu, P.: On the Maximum Sum-rate Capacity of Cognitive Multiple Access Channel. Arxiv preprint cs/0612024, 2006.

41. Viswanath, P. and Tse, D.: Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. Information Theory, IEEE Transactions on, 49(8):1912–1921, 2003.

42. Devroye, N., Mitran, P., and Tarokh, V.: Cognitive decompiosition of wirless networks, June 2006.

43. Devroye, N., Mitran, P., and Tarokh, V.: Cognitive multiple access networks. In Proc. IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia, pages 57–61, 2005.

44. Vu, M., Devroye, N., Sharif, M., and Tarokh, V.: Scaling laws of cognitive networks. Submitted to IEEE Journal on Selected Topics in Signal Processing, 2007.

45. Jeon, S., Devroye, N., Vu, M., Chung, S., and Tarokh, V.: Cognitive networks achieve throughput scaling of a homogeneous network. In Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on, pages 1–5. IEEE, 2009.

46. Cover, T. and Thomas, J.: Elements of Information Theory. Wiley-Interscience, New York, 1991.

47. Nair, C. and Gamal, A. E.: An outer bound to the capacity region of the broadcast channel. Information Theory, IEEE Transactions on, 53(1):350–355, Jan. 2007.

48. Gel'fand, S. and Pinsker, M.: Coding for channel with random parameters. Problems of control and information theory, 9(1):19–31, 1980.

49. Marton, K.: A coding theorem for the discrete memoryless broadcast channel. Information Theory, IEEE Transactions on, 25(3):306–311, May 1979.

50. Cao, Y. and Chen, B.: Interference channel with one cognitive transmitter. In Signals, Systems and Computers, 2008 42nd Asilomar Conference on, pages 1593–1597. IEEE, 2009.

51. Jiang, J., Maric, I., Goldsmith, A., and Cui, S.: Achievable rate regions for broadcast channels with cognitive relays. Proc. IEEE Information Theory Workshop (ITW), Taormina, Italy, 2009.

52. Devroye, N.: Information Theoretic Limits of Cognition and Cooperation in Wireless Networks. Doctoral dissertation, Harvard University, 2007.

53. Willems, F. and Van der Meulen, E.: The discrete memoryless multiple-access channel with cribbing encoders. Information Theory, IEEE Transactions on, 31(3):313–327, 1985.

54. Jiang, J., Xin, Y., and Garg, H.: The capacity region of a class of deterministic interference channels with common information. IEEE ICASSP 2007, 3:III–681–III–684, April 2007.

55. Lall, S.: Advanced topics in computation for control. Lecture Notes for Engr210b at Stanford University, Stanford, CA.

56. Cao, Y. and Chen, B.: Interference channels with one cognitive transmitter. In Proc. Asilomar Conferenece on Signal, Systems and Computers, 2009.

57. Chong, H., Motani, M., Garg, H., and El Gamal, H.: On The Han–Kobayashi Region for theInterference Channel. Information Theory, IEEE Transactions on, 54(7):3188–3195, 2008.

58. Jovicic, A. and Viswanath, P.: Cognitive radio: An information-theoretic perspective. 55(9):3945–3958, September 2009.

59. Sato, H.: An outer bound to the capacity region of broadcast channels (Corresp.). Information Theory, IEEE Transactions on, 24(3):374–377, 1978.

60. Kramer, G.: Topics in Multi-User Information Theory. Foundations and Trends in Communications and Information Theory. Vol. 4: No 45, pp 265-444, 2008.

61. I Csiszar, J. K.: Information theory: Coding theorems for discrete memoryless channels. Budapest: Akademiai Kiado, 1981.

62. Weingarten, H., Steinberg, Y., and Shamai, S.: On the capacity region of the multi-antenna broadcast channel with common messages. In Proc. IEEE International Symposium on Information Theory (ISIT) Seattle, USA, pages 2195–2199. IEEE, 2006.

63. Bergmans, P.: Random coding theorem for broadcast channels with degraded components. Information Theory, IEEE Transactions on, 19(2):197–207, 1973.

64. Korner, J. and Marton, K.: General broadcast channels with degraded message sets. Information Theory, IEEE Transactions on, 23(1):60–64, Jan 1977.

65. Jiang, J., Maric, I., Goldsmith, A. J., Shamai, S., and Cui, S.: On the capacity of a class of cognitive z-interference channels. Arxiv preprint arXiv:1007.1811, 2010.

66. Weingarten, H., Steinberg, Y., and Shamai, S.: The capacity region of the Gaussian multiple-input multiple-output broadcast channel. Information Theory, IEEE Transactions on, 52(9):3936–3964, 2006.

67. Cover, T., Thomas, J., and Wiley, J.: Elements of information theory. Wiley Online Library, 1991.

68. Liu, T. and Viswanath, P.: An extremal inequality motivated by multiterminal information-theoretic problems. Information Theory, IEEE Transactions on, 53(5):1839–1851, 2007.

69. Wang, I. and Tse, D.: Interference mitigation through limited transmitter cooperation. In Proc. IEEE International Symposium on Information Theory (ISIT), 2010 Austin, USA, pages 380–384. IEEE, 2010.

70. Rini, S., Tuninetti, D., and Devroye, N.: The capacity region of deterministic cognitive radio channels. Proc. IEEE Information Theory Workshop (ITW), Taormina, Italy, October 2009.

71. Gallager, R.: Capacity and coding for degraded broadcast channels. Problems in the Transmission of Information, 10(3):3–14, 1974.

72. Rini, S., Tuninetti, D., and Devroye, N.: The capacity region of Gaussian cognitive radio channels to within 1.87 bits. Proc. IEEE Information Theory Workshop (ITW), Cairo, Egypt, January 2010.

73. Rini, S., Tuninetti, D., and Devroye, N.: Inner and outer bounds for the gaussian cognitive interference channel and new capacity results. Information Theory, IEEE Transactions on, 2010. to be submitted.