2015-ENST-0020

EDITE - ED 130

Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Electronique et Communications »

présentée et soutenue publiquement par

Martina CARDONE

24 Avril 2015

Techniques de coopération appliquées aux futurs réseaux cellulaires

Directeur de thèse : **Raymond KNOPP**
Co-encadrement de la thèse : **Daniela TUNINETTI**

Jury

M. Giuseppe CAIRE, Professeur, Technische Universität Berlin	Président du jury
M. Gerhard KRAMER, Professeur, Technische Universität München	Rapporteur
M. Aydin SEZGIN, Professeur, Ruhr-Universität Bochum	Rapporteur
M. Pablo PIANTANIDA, Maître de Conférences, Supélec	Examinateur
M. Petros ELIA, Maître de Conférences, EURECOM	Examinateur
M. Umer SALIM, Docteur, Intel Mobile Communications	Examinateur

TELECOM ParisTech
école de l'Institut Télécom - membre de ParisTech

THÈSE

DISSERTATION

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from TELECOM ParisTech

Specialization: Communication and Electronics

Martina Cardone

Cooperation Strategies for Next Generation Cellular Systems

Successfully defended on the 24th of April 2015 before a committee composed of:

President of the Jury		
Professor	Giuseppe Caire	Technische Universität Berlin
Reviewers		
Professor	Gerhard Kramer	Technische Universität München
Professor	Aydin Sezgin	Ruhr-Universität Bochum
Examiners		
Assistant Professor	Pablo Piantanida	Supélec
Assistant Professor	Petros Elia	EURECOM
Doctor	Salim Umer	Intel Mobile Communications
Thesis Supervisors		
Professor	Raymond Knopp	EURECOM
Associate Professor	Daniela Tuninetti	University of Illinois at Chicago

THÈSE

présentée pour l'obtention du grade de

Docteur de TELECOM ParisTech

Spécialité: Communication et Electronique

Martina Cardone

Techniques de Coopération Appliquées aux Futurs Réseaux Cellulaires

Soutenue publiquement avec succès le 24 Avril 2015 devant le jury composé de :

Président du jury		
Professeur	Giuseppe Caire	Technische Universität Berlin
Rapporteurs		
Professeur	Gerhard Kramer	Technische Universität München
Professeur	Aydin Sezgin	Ruhr-Universität Bochum
Examinateurs		
Maître de Conférences	Pablo Piantanida	Supélec
Maître de Conférences	Petros Elia	EURECOM
Docteur	Salim Umer	Intel Mobile Communications
Directeurs de thèse		
Professeur	Raymond Knopp	EURECOM
Professeur Associée	Daniela Tuninetti	University of Illinois at Chicago

Acknowledgements

If you're not failing every now and again, it's a sign you're not doing anything very innovative.

– Woody Allen

First and foremost, I would like to express my deepest gratitude to my two Ph.D. advisors, Prof. Raymond Knopp and Prof. Daniela Tuninetti for their constant support and guidance. Their joint supervision, complementary research approaches and positive thinking gave me the confidence to master challenging problems and allowed me to benefit from an amazing research experience. Also, their precious advices and encouraging words in any moment of my Ph.D. greatly helped me to overcome the 'down' periods every Ph.D. candidate faces during the thesis. I am extremely thankful to Daniela for having welcomed me for six months within her research group at UIC in Fall 2012 and Fall 2013. This experience considerably enriched me in many ways, both from personal and technical points of view.

I would like to thank Prof. Giuseppe Caire, Prof. Gerhard Kramer, Prof. Aydin Sezgin, Prof. Pablo Piantanida, Prof. Petros Elia and Dr. Umer Salim for having agreed to be part of the jury committee. I feel honored to have had the opportunity to present my Ph.D. work in front of such prominent scientists. Also, their precious and detailed comments certainly helped to improve the quality of this dissertation.

I would like to gratefully acknowledge Intel Mobile Communications, which financially supported part of my Ph.D.; my special thanks go to Dr. Umer Salim for the insightful discussions especially at the beginning of my Ph.D. and to Jean-Xavier Canonici for his kindness and for having spent a lot of time to deal with bureaucratic stuffs. I would like to profusely thank Qualcomm as well which awarded a fellowship to me in April 2014.

My experience at EURECOM was extremely simplified by the helpfulness of all the staff; in particular, I am deeply grateful to Gwenaelle for having taken care of all the bureaucracy related to my Ph.D., to Audrey for having organized my trips to the conferences and to the IT team for having helped me everytime I had an issue with my computer.

I started my Ph.D. fifteen days before my two 'brothers' Orso and Johnny, with who I had the pleasure to share the real value of friendship. There was not a single moment in which they were not present: ready to make a toast to celebrate a success and ready to cheer me up with a word and a smile in the dark moments. Their frankness, generosity and liberal thoughts immensely enriched me and left memories that I will keep forever with me.

At EURECOM and, more generally in Côte d'Azur, I was extremely lucky to be surrounded by extraordinary friends: Andrea E., Paul, Claudia, Michela, Laurent, Francesco, Pavlos, Miriam, Pierre, Quentin, Arash, Katrin, Adriano, Rim, Laura, Francesca, Paolo, Andrea M., Mariano, Alberto. I will sorely miss all of you! Special thanks go also to my Ph.D. fellow Robin with who I had the pleasure to collaborate and who helped me to see a practical aspect of my work with my own eyes.

A note of sincere gratitude goes to my parents Bruno and Livia and to my brother Davide and his family for always supporting me, believing in me and encouraging me. Even if for them it was quite hard to fully understand many aspects of my Ph.D., such as the meaning of a deadline, their generosity of heart and constant love were fundamental pillars for me. I owe most of my achievements to them: grazie!

My last heartfelt thanks go to Matteo for his 'unbounded' patience, love and kindness. Looking back on these three years I am astounded on how many sacrificies he made and how many energies he spent for our relationship; always whispering me that 'even if we are distant, we are together', always at my side whenever I had to make a choice, always finding the right word to brighten my days. I feel blessed to have such an extraordinary person at my side!

Abstract

A uniform mobile user quality of service and a distributed use of the spectrum represent the key-ingredients for next generation cellular networks. Toward this end, physical layer cooperation among the network infrastructure and the wireless nodes has emerged as a potential technique. Cooperation leverages the broadcast nature of the wireless medium, that is, the same transmission can be heard by multiple nodes, thus opening up the possibility that nodes help one another to convey the messages to their intended destination. Cooperation also promises to offer a smart way to manage interference, instead of just simply disregarding it and treating it as noise. Understanding how to properly design such cooperative wireless systems so that the available resources are fully utilized is of fundamental importance.

The objective of this thesis is to conduct an information theoretic study on practically relevant wireless systems where the network infrastructure nodes cooperate among themselves in an attempt to enhance the network performance in many critical aspects, such as throughput, robustness and coverage. Wireless systems with half-duplex relay stations as well as scenarios where a base station overhears another base station and consequently helps serving this other base station's associated mobile users, represent the wireless cooperative networks under investigation in this thesis.

The first part of the thesis is dedicated to the study of half-duplex relay networks, where the downlink communication from a base station to a mobile user is assisted by a series of relay stations, operating in time-division duplexing (at each point in time each relay either receives or transmits). First, the single relay case is analyzed and its channel capacity is studied. In particular, the exact capacity of the linear deterministic channel is determined and several transmission strategies are designed. These techniques, when evaluated for the practically relevant Gaussian noise channel, are proven to achieve the cut-set outer bound to within a constant gap, uniformly over all channel gains. This analysis presents interesting insights and might be used as a guideline to deploy a half-duplex relay station. Then, a network with a general number N of half-duplex relays is considered for which fundamental intrinsic structural properties are indentified that allow for a drastic (from exponential in N to linear in N) simplification of the analysis. In such a network, since each relay can either transmit or receive, there are 2^N possible listen / transmit configuration states. It is proven that for any memoryless half-duplex N-relay network for which the cut-set bound is approximately optimal to within a constant gap under some conditions (satisfied for example by Gaussian noise networks), at most $N + 1$ states have a strictly positive probability.

The second part of the thesis focuses on the study of the two-user causal cognitive interference channel, where two transmitters aim to communicate independent messages to two different receivers via a common channel. One source, referred to as the *cognitive* source, is capable of overhearing the other source, referred to as the *primary* source, through a noisy in-band link and can hence assist in sending the primary's data. Two different modes of operation at the cognitive source are considered, namely full-duplex, that is, when it can simultaneously transmit and receive over the same time-frequency-space resources, and half-duplex. Different network topologies are considered, corresponding to different interference scenarios: the interference-symmetric scenario, where both destinations are in the coverage area of the two sources and hence experience interference, and the interference-asymmetric scenario, where one destination does not suffer from interference. Novel outer bounds on the capacity region are derived and several transmission strategies are designed. For each topology and mode of operation at the cognitive source, the outer and inner bounds are evaluated for the Gaussian noise channel and shown to be a constant number of bits apart from one another.

Contents

List of Figures

Acronyms

Here are the main acronyms used in this document. The meaning of an acronym is also indicated when it is first used.

AF	Amplify-and-Forward.
AWGN	Additive White Gaussian Noise.
BC	Broadcast Channel.
CIC	Cognitive Interference Channel.
CCIC	Causal Cognitive Interference Channel.
cdf	cumulative density function.
CF	Compress-and-Forward.
CTx	Cognitive Transmitter.
CRx	Cognitive Receiver.
DF	Decode-and-Forward.
DPC	Dirty Paper Coding.
FD	Full-Duplex.
FDD	Frequency-Division Duplexing.
FDG	Functional Dependence Graph.
FME	Fourier-Motzkin Elimination.
gDoF	generalized Degrees-of-Freedom.
GF	Galois Field.
HD	Half-Duplex.
i.i.d.	independent and identically distributed.
IC	Interference Channel.
ISD	Injective Semi-Deterministic.
LDA	Linear Deterministic Approximation.
LP	Linear Program.
MAC	Multiple Access Channel.
MGN	Multicast Gaussian Network.
MIMO	Multiple Input Multiple Output.

MISO	Multiple Input Single Output.
MWBM	Maximum Weighted Bipartite Matching.
NNC	Noisy Network Coding.
pdf	probability density function.
PDF	Partial Decode-and-Forward.
PTx	Primary Transmitter.
PRx	Primary Receiver.
QMF	Quantize-reMap-and-Forward.
RHS	Right Hand Side.
SIMO	Single Input Multiple Output.
SNR	Signal-to-Noise Ratio.
SISO	Single Input Single Output.
TDD	Time-Division Duplexing.
WF	Water Filling.
ZFBF	Zero Forcing BeamForming.

Notations

$[n_1 : n_2]$	Set of integers from n_1 to $n_2 \geq n_1$.
$[n_1, n_2]$	Set of all real numbers greater than or equal to n_1 and less than or equal to $n_2 \geq n_1$.
$\lfloor a \rfloor$	Floor operator, which gives the largest integer less than or equal to a.
$\emptyset$	Empty set.
$\mathbf{1}_j$	Column vector of length j of all ones.
$\mathbf{I}_j$	Identity matrix of dimension j.
$\mathbf{0}_j$	All-zero column vector of length j.
$\mathbf{0}_{i \times j}$	All-zero matrix of dimension $i \times j$.
$\arg\max$	Argument of the maximum.
$\arg\min$	Argument of the minimum.
$\mathcal{A}^c$	Complement of the set $\mathcal{A}$.
$\mathcal{A}_1 \backslash \mathcal{A}_2$	Set of elements in $\mathcal{A}_1$ but not in $\mathcal{A}_2$.
$\mathcal{A}_1 \subseteq \mathcal{A}_2$	The set $\mathcal{A}_1$ is a subset of the set $\mathcal{A}_2$.
$\mathcal{A}_1 \cup \mathcal{A}_2$	Union of the sets $\mathcal{A}_1$ and $\mathcal{A}_2$.
$\mathcal{A}_1 \cap \mathcal{A}_2$	Intersection of the sets $\mathcal{A}_1$ and $\mathcal{A}_2$.
$\lvert\mathcal{A}\rvert$	Cardinality of the set $\mathcal{A}$.
$\lvert X\rvert$	Cardinality of the random variable X.
$\lvert a\rvert$	Absolute value of the complex number a.
a^*	Complex conjugate of the complex number a.
$\angle a$	Phase angle of the complex number a.
$\lVert\mathbf{a}\rVert$	Euclidean norm of the vector $\mathbf{a}$.
$\mathbf{A}^T$	Transpose of the matrix $\mathbf{A}$.
$\mathbf{A}^H$	Hermitian transpose of the matrix $\mathbf{A}$.
$[\mathbf{A}]_{ij}$	Element in the i-th row and j-th column of the matrix $\mathbf{A}$.

$\mathbf{A}_{\mathcal{R},\mathcal{C}}$	Submatrix of the matrix $\mathbf{A}$ where only the blocks in the rows indexed by the set $\mathcal{R}$ and the blocks in the columns indexed by the set $\mathcal{C}$ are retained.
$\lvert\mathbf{A}\rvert$	Determinant of the square matrix $\mathbf{A}$.
$\mathrm{Tr}[\mathbf{A}]$	Trace of the matrix $\mathbf{A}$.
$[\mathbf{A}_{1,1}, \mathbf{A}_{1,2}; \mathbf{A}_{2,1}, \mathbf{A}_{2,2}]$	Block matrix $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} \end{bmatrix}$.
$\mathrm{diag}[\mathbf{A}_{[1:N]}]$	Block diagonal matrix with the matrix $\mathbf{A}_i$ in position (i, i) for $i \in [1:N]$.
$\mathrm{rank}\,[\mathbf{A}]$	Rank of the matrix $\mathbf{A}$.
$\delta[n]$	Kronecker delta function.
$\mathbb{E}[\cdot]$	Expected value.
$\log$	Unless otherwise specified, logarithms are in base 2.
$\log^+(x)$	Maximum between zero and $\log(x)$.
$\max, \min$	Maximum, minimum.
$\mathrm{Var}[X]$	Variance of the random variable X.
$\mathrm{Cov}[X]$	Covariance of the random variable X.
$[x]^+$	Maximum between the real number x and 0.
$f(x) = o\left(g(x)\right)$	Represents the fact that $\lim\limits_{x \to +\infty} f(x)/g(x) = 0$.
$f(x) = O\left(g(x)\right)$	Represents the fact that $\lim\limits_{x \to +\infty} \lvert f(x)\rvert/\lvert g(x)\rvert \leq M$, with $M > 0$.
$f(x) \doteq g(x)$	Represents the fact that $\lim_{x\to\infty} f(x)/g(x) = 1$.
$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$	X is a proper-complex Gaussian random variable with mean μ and variance σ^2.
Y^j	Vector of length j with component $(Y_1, \ldots, Y_j)$.
$Y_{\mathcal{A}}$	Set of all Y_j such that $j \in \mathcal{A}$, with $\mathcal{A}$ being an index set.

Chapter 1

Introduction

In this chapter, we first briefly intoduce the two network models analyzed in this dissertation, namely the half-duplex relay network and the causal cognitive interference channel, or the interference channel with unilateral source cooperation. We then summarize relevant past information theoretic results on these two scenarios and finally conclude the chapter with the thesis outline and the list of contributions.

1.1 Motivation

The next major upgrade of fourth generation cellular networks will consist of a massive deployment of radio infrastructure nodes, i.e., base stations and relay stations. Radio infrastructure nodes will come in several flavors, characterized primarily by their available bandwidth and number of concurrent frequency channels on which they can simultaneously operate (spectrum aggregation), the capacity of their backhaul links to the operator's core network (e.g., wireless, high throughput / low-latency wired interconnect, non carrier-grade wired backhaul), their ability to collaborate with other similar nodes, and their coverage area and tolerance to interference. Enabling physical layer cooperation among the infrastructure and the wireless nodes is envisaged to be the key-ingredient of future cellular networks. The broadcast nature of the wireless medium, in fact, allows the same transmission to be heard by multiple nodes, hence opening up the possibility the nodes assist one another to relay their message to the destination. Cooperation promises to offer smart solutions to cope with and manage the interference, to guarantee a fair and uniform mobile user quality of service within the cell and to allow for a distributed and aggressive use of the spectrum. All these factors are of extreme importance and it becomes therefore critical to understand how to properly design such cooperative networks.

Since Shannon's landmark work "A mathematical theory of communications", information theory has played a central role in the evolution of wireless communication systems. The core of information theory for wireless networks is to provide fundamental insights for several key problems (such as interference), by determining the ultimate performance limits of these systems. This then, for many years, has motivated wireless researchers to design techniques and transmission strategies through which these limits can be as closely as possible approached.

In this thesis, we conduct an information theoretic study on two practically relevant classes of cooperative wireless systems, where the various radio infrastructure nodes (base stations and relay stations), by leveraging the broadcast nature of the wireless medium, cooperate between themselves in an attempt to increase the network performance (e.g., throughput, coverage, robustness). In particular, we focus on the half-duplex multi-relay network and on the Causal Cognitive Interference Channel (CCIC), or the Interference Channel (IC) with unilateral source cooperation.

The multi-relay network represents a fundamental example of a cooperative wireless system [2], where several relay stations assist the over-the-air communication from a source base station (connected to a network infrastructure) to a mobile user. Adding relaying stations to today's cellular infrastructure promises, in fact, to boost network performance in terms of coverage, network throughput and robustness. Actually, relay nodes provide extended coverages in targeted areas, offering a way through which the base station can communicate with cell-edge users. Moreover, the use of relay nodes may offer a cheaper and lower energy consumption alternative to installing new base stations, especially for regions where deployment of fiber fronthaul solutions are impossible. Depending on the mode of operation, relays are classified into two categories: Full-Duplex (FD) and Half-Duplex (HD). A relay is said to operate in FD mode if it can receive and transmit simultaneously over the same time-frequency-space resource, and in HD mode otherwise. Although higher performances can be attained with FD relays, in commercial wireless networks the HD modeling assumption is at present more practical than the FD one. This is so because practical restrictions arise when a node can simultaneously transmit and receive, such as for example how well the self-interference can be canceled, making the implementation of FD relays challenging [3–5]. It is therefore more realistic to assume that the relay stations operate in HD mode, either in Frequency-Division Duplexing (FDD) or Time-Division Duplexing (TDD). In FDD, the relays use one frequency band to transmit and another one to receive, while in TDD, the relays listen for a fraction of time and then transmit in the remaining time. We first analyze the single relay case, i.e., the classical relay channel for which we seek to derive the ultimate capacity performance in the spirit of [6]. Many interesting insights are provided on how the design of a HD relay station should be properly carried out, which is an important practical task for future cellular networks. We then consider a general number N of HD relay stations. For such a network there are 2^N possible listen-transmit configuration states whose probability must be optimized. Due to the prohibitively large complexity of this optimization problem (i.e., exponential in the number of relays N) it is critical to identify, if any, structural properties of such networks that can be leveraged in order to find optimal solutions with limited complexity. By using properties of submodular functions and Linear Programs (LPs), we seek to show that a practically relevant class of HD multi-relay networks has indeed structural intrinsic properties which allow for a remarkable (from exponential in N to linear in N) simplification of the analysis.

The CCIC, or the IC with unilateral source cooperation, represents a particular aspect of future wireless networks, namely, a practical application of the cognitive overlay paradigm [7]. It consists of one *primary* source PTx (Primary Transmitter) and one *cognitive / capable* source CTx (Cognitive Transmitter) which aim to reliably communicate with two different receivers, namely the PRx (Primary Receiver) and the CRx (Cognitive Receiver), via a common channel. Differently from the classical non-cooperative IC, in the CCIC the CTx (thanks to advanced radio capabilities) is able to overhear the PTx through a noisy in-band link; the CTx can therefore exploit this side information to enhance the rate performance of the two (primary and cognitive) systems. The major and novel feature of the CCIC is the concept of *causal cognition / source cooperation*, which represents both an interference management tool and a practical model for the cognitive radio technology. Actually, unilateral source cooperation offers a way to 'smartly' manage and cope with the interference. In today's wireless systems, the general approach to deal with interference is either to avoid it, by trying to 'orthogonalize' (in time / frequency / space) the users' transmission, or to simply treat it as noise. However, these approaches may severely limit the system capacity since a perfect user orthogonalization is not possible in practice [1]. In contrast, in the CCIC the CTx, which can causally learn the primary's data through a noisy link, may protect both its own (by precoding against some known interference) and the primary's (by allocating some of its transmission resources to assist the PTx to convey data to the PRx) information from interference. Thus, the transmission techniques designed for the CCIC aim to leverage the structure of the interference, instead of just simply disregarding it and treating it as noise. The CCIC also represents a more practically relevant model for the cognitive overlay paradigm, compared to the case where the CTx is assumed to a priori (before the transmission begins) know the message of the PTx [11], which may be granted only in limited scenarios. In contrast, in the CCIC the CTx causally learns the PTx's data through a noisy link. Thus, the transmission techniques designed for the CCIC account for the time the CTx needs for decoding and for the (possible) further rate losses that may incur in decoding the PTx's message though a limited capacity and noisy link. We study different deployment configurations, which correspond to different interference scenarios. In the interference-symmetric scenario both destinations are in the coverage area of the two sources; this implies that both destinations are interfered. In the interference-asymmetric scenario, one destination does not suffer from interference; in this case one of the interfering links is absent. Due to the asymmetry in the cooperation, two interference-asymmetric scenarios must be considered: the Z-channel, where the link from the PTx to the CRx is non-existent (i.e., the CRx is out of the range of the PTx) and the S-channel, where the link from the CTx to the PRx is non-existent (i.e., the PRx is out of the range of the CTx). We further assume two different modes of operation at the CTx, namely FD (i.e., the CTx can simultaneously receive and transmit over the same time-frequency-space resource) and HD TDD (i.e., in each time slot, the CTx listens for a fraction of time and then transmits

[1] A well-known example on how 'treating interference as noise' severely limits the system capacity is given by an ad-hoc network where n randomly located pairs of devices aim to communicate. In [8] the authors showed that, if each node decodes only the signal of the closest neighbor (by treating all the other signals as noise) the rate per source-destination pair decreases to zero as $O\left(1/\sqrt{n}\right)$ for *dense* networks (i.e., when the area is fixed and $n \to \infty$). In [9], this scaling law was proven to be information theoretically optimal for *extended* networks (when n is fixed and the area increases linearly with n) in high attenuation (i.e., for a path loss exponent $\alpha > 4$). In [10], Özgür et al. showed a novel scaling law for dense networks and extended networks in low attenuation: if nodes can cooperate, then the total capacity of the network scales with n, i.e., the rate of each source-destination pair is not impaired as n increases. This was accomplished through a novel hierarchical cooperation architecture, where nodes within the same cluster cooperate in delivering the messages to their destinations.

in the remaining time). For each topology we study the ultimate capacity performance in the spirit of [6, 12], by deriving novel outer bounds on the capacity region and by designing transmission strategies which are provably approximately optimal for the Gaussian noise channel.

1.2 Background

1.2.1 Half-Duplex Relay Networks

The relay channel model, where a source communicates with a destination with the help of one relay station, was first introduced by van der Meulen [13] in 1971. Despite the significant research efforts, the capacity of the general memoryless relay channel is still unknown. In their seminal work [14], Cover and El Gamal proposed a general outer bound, now known as the max-flow min-cut outer bound or cut-set for short, and two achievable schemes: Decode-and-Forward (DF) and Compress-and-Forward (CF). In DF, the relay fully decodes the message sent by the source and then coherently cooperates with the source to communicate this information to the destination. In CF, the relay does not attempt to recover the source message, but it just compresses the received signal and then sends it to the destination. The combination of DF and CF is still the largest known achievable rate for a general memoryless relay channel. The cut-set outer bound was shown to be tight for the degraded relay channel, the reversely degraded relay channel and the semi-deterministic relay channel [14], but it is not tight in general [15]. The pioneering work of [14] has been extended to networks with multiple relays. In [16], the authors proposed several inner and outer bounds for FD relay networks as a generalization of DF, CF and the cut-set bound; it was shown that DF achieves the ergodic capacity of a wireless Gaussian network with uniform phase fading if the phase information is locally available and the relays are close to the source node.

Although more study has been conducted for FD relays, there are some important references treating HD ones. In [17], the author studied the TDD relay channel. Both an outer bound, based on the cut-set argument, and an inner bound, based on Partial DF (PDF), a generalization of DF where the relay only decodes part of the message sent by the source, were derived. In [17], the time instants at which the relay switches from listen to transmit and vice versa were assumed to be fixed, i.e., a priori known by all the nodes; we refer to this mode of operation as *deterministic* switch. In [18], Kramer showed that higher rates can be achieved by considering a *random* [2] switch at the relay. In this way the randomness that lies into the switch may be used to transmit (at most 1 bit per channel use of) further information to the destination. In [18], it was also shown how the memoryless FD framework incorporates the HD one as a special case, and as such there is no need to develop a separate theory for networks with HD nodes.

The exact characterization of the capacity region of a general memoryless network is challenging. Recently it has been advocated that progress can be made towards understanding the capacity by showing that achievable strategies are provably close to (easily computable) outer bounds [6]. As an example, in [19], the authors studied FD Gaussian relay networks with $N+2$ nodes (i.e., N relays, a source and a destination) and showed that the capacity can be achieved to within $\sum_{k=1}^{N+2} 5\min\{M_k, N_k\}$ bits with a network generalization of CF named Quantize-reMap-and-Forward (QMF), where M_k and N_k are the number of transmit and receive antennas, respectively, of node k. Recently, for single antenna networks with N FD relays, the $5(N+2)$ bits gap of [19] was reduced to $2 \times 0.63(N+2)$ bits (where the factor 2 accounts for complex-valued inputs) thanks to a novel ingenious generalization of CF named Noisy Network Coding (NNC) [20]. The gap characterization of [20] is valid for a general Multicast Gaussian Network (MGN) with FD nodes; the gap grows linearly with the number of nodes in the network, which could be a too coarse capacity characterization for networks with a large number of nodes. Smaller gaps can be obtained for more structured networks. For example, a diamond network [21] consists of a source, a destination and N relays where the source and the destination can not communicate directly and the relays can not communicate among themselves. In other words, a general Gaussian relay network with N relays is characterized by $(N+2)(N+1)$ generic channel link gains, while a diamond network has only $2N$ non-zero channel link gains. In [21] the case of $N=2$ relays was studied and an achievable region based on time sharing between DF and Amplify-and-Forward (AF) was proposed. In [22], the authors considered two specific configurations of a diamond network with a general number of relays (agents), where the relay-destination links were assumed to be lossless; in the first scenario the relays do not have decoding capabilities, while in the second scenario they do. Upper and lower bounds on the capacity were derived and evaluated for the Gaussian noise channel. Moreover the capacity of the deterministic channel when the relays can decode

[2] Since the relay's state (either listen or transmit) is part of the codebook, *random* switch can equivalently be referred to as *coded* switch.

was characterized. The scenarios of [22] were further studied in [23] under the assumption of lossy relay-destination links and where each source-relay link and relay-destination link is a binary-symmetric channel. In [24], the authors analyzed the Gaussian diamond network with a direct link between the source and the destination and showed that 'uncoded forwarding' at the relays asymptotically achieves the cut-set upper bound when the number of relays goes to infinity. This strategy simply requires that each relay delays the input of one time unit and scales it to satisfy the power constraint. In general, the capacity of the Gaussian FD diamond network is known to within $2\log(N+1)$ bits [25], [26], i.e., the simplified (and sparse) diamond topology allows for a gap reduction from linear [20] to logarithmic [25], [26]. If, in addition, the network is symmetric, that is, all source-relay links are of the same strength and all relay-destination links are of the same strength, the gap is less than 3.6 bits for any N [27].

Interestingly, the gap result of [19] remains valid for static and ergodic fading networks where the nodes operate either in FD mode or in HD mode; however [19] did not account for random switch in the outer bound. In [28], the authors demonstrated that the QMF scheme can be realized with nested lattice codes. Moreover, they showed that for HD networks with N relays, by following the approach of [18], i.e., by also accounting for random switch in the outer bound, the cut-set outer bound is achievable to within $N + 4\sum_{k=1}^{N} M_k + (2+\log(2))\sum_{k=1}^{N} N_k$ bits, with M_k and N_k being the number of antennas used to receive and transmit at the k-th relay; in the special case of single-antenna nodes this gap reduces to $5N$. In [29], the authors established capacity expressions of the error-free half-duplex line network, i.e., a relay network where a source, a certain number of relays and a destination are arranged on a line and communication takes place only between adjacent nodes. In particular, in [29, Theorem 1] they characterized the capacity of the line network with a single source-destination pair, in [29, Theorem 2] they found an explicit capacity expression when the number of relays goes to infinity and in [29, Theorem 3] they characterized the capacity of the line network where each relay can act as a source if the rates of the relay sources fall below certain thresholds. All these capacity results were proved by using a random switch at each relay. In [30, Theorems 3.1, 3.2], the capacity of the deterministic line network with two sources, i.e., when either the second relay (in [30, Theorem 3.1]) or the last relay in the line (in [30, Theorem 3.2]) is the second source, was characterized; also in these scenarios the cut-set upper bound is achieved if the relays randomly switch from listen to transmit. In general, finding the capacity of a single-antenna HD multi-relay network is a combinatorial problem since the cut-set upper bound is the minimum between 2^N bounds (one for each possible cut in the network), each of which is a linear combination of 2^N relay states (since each relay can either transmit or receive). For a diamond network with $N = 2$ relays, [31] showed that, out of the $2^N = 4$ possible states, at most $N + 1 = 3$ states suffice to achieve the cut-set bound to within less than 4 bits. We refer to the states with a strictly positive probability as active states. The achievable scheme of [31] is a clever extension of the two-hop DF strategy of [32]. In [31] a closed-form expression for the aforementioned active states, by assuming no power control and deterministic switch, was derived by solving the dual LP associated with the LP derived from the cut-set bound. The work in [33] studied a Gaussian diamond network with $N = 2$ relays and an 'antisymmetric' Gaussian diamond network with $N = 3$ relays and showed that a significant fraction of the capacity can be achieved by: (i) selecting a single relay, or (ii) selecting two relays and allowing them to work in a complementary fashion as in [31]. Inspired by [31], the authors of [33] also showed that, for a specific HD diamond network with $N = 3$ relays, at most $N + 1 = 4$ states, out of the $2^N = 8$ possible ones, are active. The authors also numerically verified that for a general Gaussian HD diamond network with $N \leq 7$ relays, at most $N + 1$ states are active and conjectured that the same holds for any number N of relays. In [34], this conjecture was proved for single-antenna Gaussian HD diamond networks with $N \leq 6$ relays; the proof is by contradiction and uses properties of submodular functions and LP duality but requires numerical evaluations; for this reason the authors could only prove the conjecture for $N \leq 6$, since for larger values of N "the computational burden becomes prohibitive" [34].

HD relay networks were also studied in [35], where an iterative algorithm was proposed to determine the optimal fraction of time each HD relay transmits/receives by using DF with deterministic switch. In [36] the authors proposed a 'grouping' technique to find the relay schedule that maximizes the approximate capacity of certain Gaussian HD relay networks, including for example layered networks; since finding a good node grouping is computationally complex, the authors proposed an heuristic approach based on tree decomposition that results in polynomial-time algorithms; as for diamond networks in [33], the low-complexity algorithm of [36] relies on the 'simplified' topology of certain networks.

1.2.2 The Interference Channel with Source Cooperation

The presence of a lossy communication link between the PTx and the CTx enables the CTx to cooperate with the PTx. The CTx, in fact, through this noisy channel overhears the signal sent by the PTx and gathers information about the PTx's message, which serves as the basis for unilateral cooperation between the two sources. Unilateral source cooperation is a special case of the IC with generalized feedback, or bilateral source cooperation. The CCIC also represents a practical scenario for cognitive radios, where one source has superior capabilities with respect to the other source. Moreover, closely related to the IC with unilateral source cooperation is the classical output feedback model, where the received signal is sent back through a perfect or noisy channel from one receiver to the corresponding transmitter. Lately, these scenarios have received significant attention, as summarized next.

Non-cooperative IC. The capacity region of the classical non-cooperative IC is not known in general. The only case for which the capacity is known is the strong interference regime [37,38], where the interfering / cross links are of a better quality with respect to the direct links. The largest achievable rate region is due to Han and Kobayashi [39]. In the transmission strategy proposed in [39], each source splits its message into two parts, i.e., a common message, decoded also at the non-intended receiver and a private message, treated as noise at the non-intended receiver. In [40], the Han-Kobayashi scheme was shown to be optimal for a class of deterministic discrete memoryless ICs for which the receiver outputs and the interferences are a deterministic function of the channel inputs. In [12], the authors evaluated the rate region of [39] for the practically relevant Gaussian noise channel. They showed that, by setting the power of the private message in such a way it is received at most at the level of the noise at the non-intended receiver, the corresponding achievable rate region is to within 1 bit of the capacity.

FD IC with bilateral source cooperation. Bilateral source cooperation has been actively investigated recently. Host-Madsen [41] first studied outer and inner bounds on the capacity for the Gaussian IC with either source or destination bilateral cooperation. Regarding the outer bound, the author in [41] evaluated the different cut-set upper bounds and then tightened the sum-rate upper bound by extending the sum-rate outer bounds originally developed by Kramer [42] for the Gaussian non-cooperative IC in weak and strong interference to the cooperative case. The lower bound region of [41] was derived by designing a scheme based on Gelfand-Pinsker's binning [43] (i.e., Dirty Paper Coding (DPC) in Gaussian noise [44]) and superposition encoding, DF relaying and joint decoding. Tuninetti [45] derived a general outer bound for the IC with bilateral source cooperation by extending Kramer's Gaussian noise sum-rate upper bounds in [42, Theorem 1] to any memoryless IC with source cooperation, and more recently to any form of source and destination cooperation [46]. Prabhakaran and Viswanath [47] extended the idea of [12, Theorem 1] to derive a sum-rate outer bound for a class of Injective Semi-Deterministic (ISD) IC with bilateral source cooperation in the spirit of the work by Telatar and Tse [48], and evaluated it for the Gaussian channel with independent noises (this assumption is not without loss of generality when cooperation and feedback are involved). Tandon and Ulukus [49] derived an outer bound for the IC with bilateral source cooperation based on the dependence-balance idea of Hekstra and Willems [50] and proposed a novel method to evaluate it for the Gaussian channel with independent noises.

The largest known achievable rate region for general bilateral source cooperation, to the best of our knowledge, is the one presented in [51, Section V]. In [51, Section V] each source splits its message into two parts, i.e., a common and a private message, as in the Han-Kobayashi's scheme for the non-cooperative IC [39]; these two messages are further sub-divided into a non-cooperative and a cooperative part. The non-cooperative messages are transmitted as in the non-cooperative IC [39], while the cooperative messages are delivered to the destinations by exploiting the cooperation among the two sources. In [51, Section V] each source, e.g. source 1, after learning the cooperative messages of source 2, sends the common cooperative message of source 2 and uses Gelfand-Pinsker's binning [43] against the private cooperative message of source 2 in an attempt to rid its own receiver of this interference. The achievable scheme in [51, Section V] uses PDF for cooperation. A possibly larger achievable region could be obtained by including CF as cooperation mechanism as in [14] for the relay channel.

For the two-user Gaussian noise IC with bilateral source cooperation, under the assumption that the cooperation links have the same strength, the scheme of [51, Section V] was sufficient to match the sum-capacity upper bounds of [45, 47] to within a constant gap [47, 52]. In particular, [47] characterized the sum-capacity to within 19 bits of the IC with bilateral source cooperation under the condition that the cooperation links have the same strength, but otherwise arbitrary direct and interfering links. The gap was reduced to 2 bits in the 'strong cooperation regime' in [52] with symmetric direct links, symmetric interfering links and symmetric cooperation links.

FD IC with unilateral source cooperation. Unilateral source cooperation is clearly a special case of the general bilateral cooperation case where the cooperation capabilities of the two sources are not restricted to be the same. This case has been specifically considered in [53] where the cooperating transmitter works either in FD or in HD mode. The authors of [53] evaluated the performance of two achievable schemes: one that exploits PDF and binning and a second one that extends the first by adding rate splitting. It was observed, through numerical evaluations, that the proposed inner bounds are not too far from the outer bound of [49] for certain Gaussian noise channels. An extension of the IC with unilateral source cooperation was studied in [54], where it was assumed that at any given time instant the cognitive source has a non-causal access to $L \geq 0$ future channel outputs. The case $L = 0$ corresponds to the strictly causal case, while the case $L \to \infty$ to the ideal non-causal Cognitive Interference Channel (CIC) [11]. The authors of [54] derived potentially tighter outer bounds on the capacity of the CCIC channel (i.e., case $L = 0$) than those of [45, 47] specialized to unilateral source cooperation; unfortunately it is not clear how to evaluate these bounds in Gaussian noise because they are expressed as a function of auxiliary random variables jointly distributed with the inputs and for which no cardinality bounds on the corresponding alphabets are known. The achievable region in [54, Corollary 1] is also no smaller than the region in [51, Section V] specialized to the case of unilateral source cooperation (see [54, Remark 2, point 6]). Although [54, Corollary 1] is, to the best of our knowledge, the largest known achievable region for the general memoryless IC with unilateral cooperation, its evaluation in general is quite involved as the rate region is specified by 9 jointly distributed auxiliary random variables and by 30 rate constraints. In [54] inner bounds were compared numerically to the 2×2 Multiple Input Multiple Output (MIMO) outer bound for the Gaussian CCIC; the 2×2 MIMO outer bound is loose in general compared to the bounds in [41,45,47]. Although it was noted in [54] that, for the simulated set of channel gains, the proposed bounds are not far away from one another, a performance guarantee in terms of (sum-)capacity to within a constant gap was not given.

HD IC with source cooperation. HD cooperation can be studied as a special case of FD cooperation by using the formalism of [18]. This approach is usually not followed in the literature, often making imprecise claims about capacity and Gaussian capacity to within a constant gap. In [55], the sum-capacity of the Gaussian IC with HD source cooperation and deterministic switch was characterized to within 20 bits and 31 bits for the case of symmetric (direct, interference and cooperation links) bilateral and general unilateral cooperation, respectively. These approximately optimal schemes are inspired by the Linear Deterministic Approximation (LDA) of the Gaussian noise channel at high Signal-to-Noise Ratio (SNR). The LDA, first proposed in [19] in the context of relay networks, captures in a simple deterministic way the interaction between interfering signals of different strengths. In the LDA the effect of the noise is neglected and the signal interaction is modeled as bit-wise additions. Thereby, this simplification allows for a complete characterization of the capacity region in many instances where the capacity of the noisy channel counterpart is a long standing open problem.

IC with output feedback. In [56], Suh and Tse studied the Gaussian IC where each source has a perfect output feedback from the intended destination. The authors characterized the capacity of this system to within 2 bits and showed that feedback provides a generalized Degrees-of-Freedom (gDoF) gain, and thus an unbounded rate gain, with respect to the classical (i.e., with no feedback) IC. It was proved, see [56, Theorems 2-3], that the capacity region has constraints on the single rates and on the sum-rate, but not bounds of the type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ and $R_{\mathsf{p}}+2R_{\mathsf{c}}$ (where R_{p}, respectively R_{c}, is the transmission rate for the PTx, respectively CTx), which appear in the capacity region of the classical Gaussian IC [12]. The authors interpreted the bounds on $2R_{\mathsf{p}}+R_{\mathsf{c}}$ and $R_{\mathsf{p}}+2R_{\mathsf{c}}$ in the capacity region of the classical IC as a measure of the amount of 'resource holes', or system underutilizations, due to the distributed nature of the non-cooperative IC. In other words, output feedback eliminates these 'resource holes' and the system resources are fully utilized. In [57], the symmetric Gaussian IC with all 9 possible output feedback configurations was analyzed. The authors proved that the bounds derived in [56] suffice to approximately (i.e., to within a constant gap) characterize the capacity of all the 9 configurations except for the case where only one source receives feedback from the corresponding destination, i.e., the 'single direct-link feedback model / model (1000)'. For this model, in [57] it was shown that an outer bound of the type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ is needed to capture the fact that the second source (whose transmission rate is R_{c}) does not receive feedback. In the language of [56] we thus have that the 'single direct-link feedback' does not suffice to cover all the 'resource holes' whose presence is captured by the bound on $2R_{\mathsf{p}}+R_{\mathsf{c}}$. The authors of [57] derived a novel outer bound on $2R_{\mathsf{p}}+R_{\mathsf{c}}$ for the ISD model (1000) with independent noises and showed it is active for the Gaussian noise case. In [58], the authors characterized the capacity of the two-user 'symmetric linear deterministic IC with partial feedback', where only some bits are received at the transmitter as feedback from the corresponding receiver. In [59], the same authors evaluated the bounds for the symmetric Gaussian noise channel and proved that they are at most 11.7 bits far from one another, universally over all channel parameters. The capacity characterization was accomplished by deriving novel outer bounds on $2R_{\mathsf{p}}+R_{\mathsf{c}}$ and $R_{\mathsf{p}}+2R_{\mathsf{c}}$ that rely on carefully chosen side information random variables tailored to the symmetric Gaussian setting and whose generalization to non-symmetric or non-Gaussian scenarios does not apper straightforward.

Non-causal cognitive radio channel. The cognitive radio channel is commonly modeled following the pioneering work of Devroye et al. [11] in which the superior capabilities of the cognitive source are modeled as perfect non-causal, i.e., before transmission begins, knowledge of the PTx's message at the CTx. For this non-causal model, the capacity is exactly known when the PRx experiences weak interference [60, 61] and in the strong interference regime [62]. For the other operating regimes, to the best of our knowledge, the largest known achievable rate region is the one presented in [63, Theorem 7], which in [64] was evaluated for the Gaussian noise case and shown to be at most 1 bit apart from an outer bound region characterized by constraints on the single rates and on the sum-rate. In other words, the capacity region of the non-causal model does not have bounds on $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$, i.e., the assumption of full a priori knowledge of the PTx's message at the CTx allows to fully exploit the available system resources.

1.3 Contributions of this dissertation

In this thesis we analyze two practically relevant wireless channel models with nodes cooperation, namely the HD relay network and the CCIC, or the IC with unilateral source cooperation. These two scenarios are studied into two different parts, namely Part I and Part II, respectively. In particular, our analysis makes use of information theoretic and graph theoretic tools. Properties on submodular functions and linear programming are also used.

This thesis resulted in 13 conference papers and 6 journal papers, all currently under submission or already published by IEEE. Parts of these works are reprinted next with permission from IEEE.

1.3.1 Part I

In Part I, we study the HD relay network where the communication between a source and a destination is assisted by N relay stations operating in HD.

In particular,

Chapter 2. In Chapter 2, we analyze the practically relevant Gaussian noise case for $N = 1$, i.e., the Gaussian relay channel, whose exact capacity is unknown. We make progress toward determining its capacity by characterizing its gDoF in closed-form and proving a constant gap result. We also propose a scheme inspired by the LDA, which is provably asymptotically optimal. Our main contributions can be summarized as follows:

1. We determine the exact capacity of the LDA channel: we show that random switch and correlated non-uniform input bits at the relay are optimal. We also show that deterministic switch is at most 1 bit from optimal.

2. We derive the gDoF for the Gaussian relay channel in closed-form: we show that both PDF and CF are gDoF optimal, both with deterministic and with random switch at the relay. We also show that a scheme inspired by the LDA with deterministic switch is gDoF optimal.

3. For the Gaussian noise case, we prove that the above transmission strategies are optimal to within a constant gap, uniformly over all channel parameters. In particular, PDF is optimal to within 1 bit, CF to within 1.61 bits, and the scheme inspired by the LDA to within 3 bits. In all cases, the gap is smaller than the one of 5 bits available in the literature for the case of one relay [28].

4. For the three coding schemes, we obtain a closed-form expression for the approximately optimal schedule (i.e., duration of the transmit- and receive-phases at the relay) with deterministic switch. This result sheds light on the design of a HD relay node in future wireless networks.

5. We prove that PDF with random switch is exactly optimal for the general memoryless line network, i.e., when the direct link between the source and the destination is absent. A closed-form expression for the optimal input distribution with random switch policy is however not available.

Publications related to this chapter are:

- [65] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "Gaussian half-duplex relay channels: generalized degrees of freedom and constant gap result", in 2013 IEEE International Conference on Communications (ICC 2013), Budapest (Hungary), June 2013.

- [66] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "The capacity to within a constant gap of the Gaussian half-duplex relay channel", in 2013 IEEE International Symposium on Information Theory (ISIT 2013), Istanbul (Turkey), July 2013.

- [67] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the Gaussian half-duplex relay channel", in IEEE Transactions on Information Theory, Volume 60, Issue n.5, May 2014, Pages 2542-2562.

A practical implementation of the transmission strategy inspired by the LDA can be found in

- [68] R. Thomas, M. Cardone, R. Knopp, D. Tuninetti, B. T. Maharaja, "An LTE implementation of a novel strategy for the Gaussian half-duplex relay channel", to appear in 2015 IEEE International Conference on Communications (ICC 2015), London (United Kingdom), June 2015.

Chapter 3. In Chapter 3, we study the HD relay network with a general number N of relays, by following the approach proposed in [18]. Our main contributions can be summarized as follows:

1. For the practically relevant Gaussian noise case, we prove that NNC with deterministic switch achieves the cut-set bound (properly evaluated to account for random switch) to within $1.96(N + 2)$ bits. This gap is smaller than the $5N$ bits gap available in the literature [28]. Our gap result for a HD relay network is obtained as a special case of a more general result for a HD MGN, which extends the 1.26 bits/node gap for the FD case [20] to a 1.96 bits/node gap for the HD case. We also show that this gap result extends to the case of multi-antenna nodes and is of 1.96 bits per channel use per antenna.

2. In order to determine the gDoF of the Gaussian channel, one needs to find a tight high-SNR approximation for the different mutual information terms involved in the cut-set upper bound. As a result of independent interest, beyond its application to the Gaussian relay network studied in this chapter, we show that such tight approximations can be found as the solution of Maximum Weighted Bipartite Matching (MWBM) problems, or assignment problems [69], for which efficient polynomial-time algorithms, such as the Hungarian algorithm [70], exist. As an example, we show that this technique is useful to derive the

gDoF of Gaussian broadcast networks with and without relays and to solve user scheduling problems.

3. We prove Brahma et al.'s conjecture [33] beyond Gaussian networks with a diamond topology. In particular, we show that for any HD network with N relays, with independent noises and for which the cut-set bound is approximately optimal to within a constant under certain assumptions, the (approximately) optimal relay policy is simple, i.e., at most $N + 1$ states (out of the 2^N possible ones) have a strictly positive probability. The key idea is to use the Lovász extension and the greedy algorithm for submodular polyhedra to highlight structural properties of the minimum of a submodular function. Then, by using the saddle-point property of min-max problems and the existence of optimal basic feasible solutions for LPs, an (approximately) optimal relay policy with the claimed number of active states can be shown. Gaussian noise relay networks satisfy all the assumptions and thus admit a simple schedule. More importantly, when the nodes are equipped with multiple antennas and the antennas at the relays may be switched between transmit and receive modes independently of one another, the schedule has at most $N+1$ active states (as in the single-antenna case), regardless of the total number of antennas in the system.

4. We finally consider two network examples: for the first scenario, consisting of $N = 2$ single-antenna relays, we highlight under which channel conditions a best-relay selection scheme is strictly suboptimal in terms of gDoF and we gain insights into the nature of the rate gain attainable in networks with multiple relays; for the second scenario, consisting of $N = 1$ relay equipped with 2 antennas, we show that independently switching the 2 antennas at the relays not only achieves in general strictly higher rates compared to using the antennas for the same purpose, but can actually provide a strictly larger pre-log factor.

Publications related to this chapter are:

- [71] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "Gaussian half-duplex relay networks: improved gap and a connection with the assignment problem", in 2013 IEEE Information Theory Workshop (ITW 2013), Seville (Spain), September 2013.

- [72] M. Cardone, D. Tuninetti, R. Knopp, "On user scheduling for maximum throughput in K-user MISO broadcast channels", to appear in 2015 IEEE International Conference on Communications (ICC 2015), London (United Kingdom), June 2015.

- [73] M. Cardone, D. Tuninetti, R. Knopp, "The approximate optimality of simple schedules for half-duplex multi-relay networks", to appear in 2015 IEEE Information Theory Workshop (ITW 2015), Jerusalem (Israel), May 2015.
- [74] M. Cardone, D. Tuninetti, R. Knopp, "Gaussian MIMO half-duplex relay networks: approximate optimality of simple schedules", to appear in 2015 IEEE International Symposium on Information Theory (ISIT 2015), Hong Kong, June 2015.
- [75] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "Gaussian half-duplex relay networks: improved constant gap and connections with the assignment problem", in IEEE Transactions on Information Theory, Volume 60, Issue n.6, June 2014, Pages 3559-3575.
- [76] M. Cardone, D. Tuninetti, R. Knopp, "On the optimality of simple schedules for networks with multiple half-duplex relays", submitted to IEEE Transactions on Information Theory, December 2014.

1.3.2 Part II

In Part II, we study the CCIC, or the IC with unilateral source cooperation, which consists of two source-destination pairs sharing the same channel and where the CTx overhears the PTx through a lossy communication link and can hence allocate some of its transmission resources to assist the communication of the primary pair. In particular,

Chapter 4. In Chapter 4, we consider FD mode of operation at the cognitive source, i.e., the CTx can receive and transmit simultaneously over the same time-frequency-space resources. Our main contributions can be summarized as follows:

1. We develop a general framework to derive outer bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ on the capacity of the general ISD CCIC when the noises at the different source-destination pairs are independent; this framework includes for example feedback from the intended destination. As a special case, we recover and strengthen the bounds derived in [47, 57]. The key technical ingredient is the proof of two Markov chains.

2. We design a transmission strategy for the general memoryless CCIC and we derive its achievable rate region. The proposed scheme uses superposition and binning encoding, PDF relaying and simultaneous decoding at the receivers. Since the CCIC shares common features with the classical non-cooperative IC [39], both common and private messages are used. Moreover, we use both cooperative and non-cooperative messages for the PTx, while the messages of the CTx are only non-cooperative.

3. We evaluate the outer bound and the achievable rate regions for the practically relevant Gaussian noise channel. We prove that for the symmetric case, i.e., when the two direct links and the two cross / interfering links are of the same strength, for the Z-channel, i.e., when the link from the PTx to the CRx is absent, and for the S-channel, i.e., when the link from the CTx to the PRx is absent, the achievable region is a constant (uniformly over all channel gains) number of bits apart from the outer bound region. Interestingly, we show that the capacity regions of the two asymmetric scenarios (i.e., the Z-channel and the S-channel) do not have bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$, i.e., unilateral cooperation allows for a full utilization of the channel resources. On the other hand, we prove that the two novel outer bounds of the type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ and $R_{\mathsf{p}}+2R_{\mathsf{c}}$ are active for the symmetric channel in weak interference and when the cooperation link is weaker than the direct link, i.e., for this regime unilateral cooperation is too weak and leaves some system resources underutilized.

4. The constant gap results imply the exact knowledge of the gDoF for the Z-, S- and symmetric channels. We identify the parameter regimes where the Gaussian CCIC (both with symmetric and asymmetric configurations) is equivalent in terms of gDoF to the non-cooperative Gaussian IC [12] (i.e., unilateral cooperation might not be worth implementing in practical systems) and to the Gaussian non-causal CIC [64] (i.e., unilateral causal cooperation attains the ultimate limit of cognitive radio technology). These comparisons shed lights into the parameter regimes and network topologies that in practice might provide an unbounded throughput gain compared to currently available (non-cognitive) technologies.

Publications related to this chapter are:

- [77] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "Approximate sum-capacity of full- and half-duplex asymmetric interference channels with unilateral source cooperation", in 2013 Information Theory and Applications Workshop (ITA 2013), San Diego (USA), February 2013.

- [78] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the interference channel with causal cognition", in 2013 IEEE International Conference on Communications (ICC 2013), Budapest (Hungary), June 2013.

- [79] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the Gaussian interference channel with unilateral generalized feedback", in 6th International Symposium on Communications, Control and Signal Processing (ISCCSP 2014), Athens (Greece), May 2014.

- [80] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the capacity of full-duplex causal cognitive interference channels to within a constant gap", in 2014 IEEE International Conference on Communications (ICC 2014), Sydney (Australia), June 2014.

- [81] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "New outer bounds for the interference channel with unilateral source cooperation", in 2014 IEEE International Symposium on Information Theory (ISIT 2014), Honolulu (Hawaii), July 2014.

- [82] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the capacity of the two-user Gaussian causal cognitive interference channel", in IEEE Transactions on Information Theory, Volume 60, Issue n.5, May 2014, Pages 2512-2541.

- [83] M. Cardone, D. Tuninetti, R. Knopp, "The two-user causal cognitive interference channel: novel outer bounds and constant gap result for the symmetric Gaussian noise channel in weak Interference", submitted to IEEE Transactions on Information Theory, March 2014.

Chapter 5. In Chapter 5, we consider HD mode of operation at the cognitive source, i.e., in each time slot the CTx listens for a fraction of time and then transmits in the remaining time. Our main contributions can be summarized as follows:

1. We characterize the sum-capacity to within a constant gap for the Gaussian symmetric Z-channel, the Gaussian symmetric S-channel and the symmetric fully-connected Gaussian HD-CCIC; this is accomplished by adapting the sum-capacity outer bounds for FD unilateral

cooperation in Chapter 4 to the case of HD unilateral cooperation by using the framework of [18], i.e., by properly accounting for random switch at the CTx, and by designing novel transmission strategies inspired by the LDA of the Gaussian noise channel at high SNR. In particular, the gap is of 5 bits/user for the symmetric case and of 3 bits/user for the symmetric Z-channel and the symmetric S-channel. We remark that these gap results not only, differently from [55], are derived by properly accounting for random switch at the CTx, but they are also smaller than those derived in [55].

2. Using the LDA model, we obtain a closed-form expression for the gDoF and for the different optimization variables (e.g., schedule, power splits, coding schemes and corresponding decoding orders, etc.). This result sheds light on how the design of the HD CTx should be properly carried out, which is an important practical task for future wireless networks.

3. As done for the FD case in Chapter 4, we compare the gDoF of the Gaussian HD-CCIC with that of: (i) the classical non-cooperative IC, i.e., where there is no cooperation among the nodes [12], and (ii) the non-causal CIC, i.e., where the CTx has a non-causal knowledge of the PTx's message [64]. In particular, we find the parameter regimes where HD unilateral cooperation does not yield benefits compared to the non-cooperative IC [12], and those where it attains the ultimate performance limits of the non-causal CIC [64]. Interestingly, we show that in the regimes where the Gaussian HD-CCIC outperforms the non-cooperative IC the cooperation link must be able to reliably convey a rate larger than the sum-capacity of the corresponding non-cooperative IC.

4. We finally identify the regimes where a loss, in terms of gDoF, incurs by using HD mode of operation at the CTx with respect to the FD case analyzed in Chapter 4. These losses might motivate the use of a more expensive CTx with FD capabilities in future wireless networks in these regimes.

Publications related to this chapter are:

- [84] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "The symmetric sum-capacity of the Gaussian half-duplex causal cognitive interference channel to within a constant gap", in 2013 IEEE International Symposium on Information Theory (ISIT 2013), Istanbul (Turkey), July 2013.

- [85] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the Gaussian interference channel with half-duplex causal cognition", in IEEE Journal on Selected Areas in Communications, Volume 32, Issue n.11, November 2014, Pages 2177-2189.

Part I

Half-Duplex Relay Networks

Chapter 2

Half-Duplex Relay Channel

In this chapter, we study the HD relay channel. Our main contributions can be summarized as follows: (i) we determine the exact capacity of the LDA channel; (ii) we show that, for the Gaussian noise case, the cut-set outer bound is achievable to within a constant gap by PDF and CF, evaluated both with deterministic and random switch; (iii) we design an 'optimal to within a constant gap' scheme inspired by the LDA of the Gaussian noise channel at high SNR; (iv) we prove that PDF with random switch is exactly optimal when the direct link is absent.

2.1 System model

2.1.1 General memoryless channel

A general memoryless relay network has one source (node 0), one destination (node $N+1$), and N [1] relays indexed from 1 to N. It consists of $N+1$ input alphabets $(\mathcal{X}_1, \cdots, \mathcal{X}_N, \mathcal{X}_{N+1})$ (here $\mathcal{X}_i$ is the input alphabet of node i except for the source / node 0 where, for notation convenience, we use $\mathcal{X}_{N+1}$ rather than $\mathcal{X}_0$), $N+1$ output alphabets $(\mathcal{Y}_1, \cdots, \mathcal{Y}_N, \mathcal{Y}_{N+1})$ (here $\mathcal{Y}_i$ is the output alphabet of node i), and a transition probability $\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]}}$. The source has a message W uniformly distributed on $[1 : 2^{nR}]$ for the

[1] Even if this chapter focuses on the single relay case, we here define the channel model for the general case of N relays, since we will adopt the same model in the next chapter.

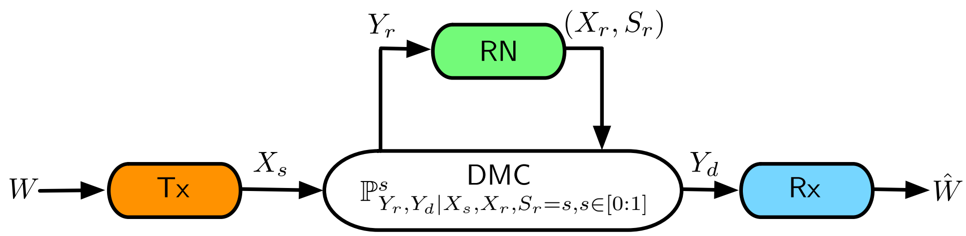


Figure 2.1: The general memoryless HD relay channel.

destination, where n denotes the codeword length and R the transmission rate in bits per channel use. At time i, $i \in [1:n]$, the source maps its message W into a channel input symbol $X_{N+1,i}(W)$, and the k-th relay, $k \in [1:N]$, maps its past channel observations into a channel input symbol $X_{k,i}\left(Y_k^{i-1}\right)$. The channel is assumed to be memoryless, that is, the following Markov chain holds for all $i \in [1:n]$

$$\left(W, Y_{[1:N+1]}^{i-1}, X_{[1:N+1]}^{i-1}\right) - X_{[1:N+1],i} - Y_{[1:N+1],i}.$$

At time n, the destination makes an estimate of the message W based on all its channel observations Y_d^N as $\widehat{W}(Y_{N+1}^n)$. A rate R is said to be ϵ-achievable if, for some block length n, there exists a code such that $\mathbb{P}[\widehat{W} \neq W] \leq \epsilon$ for any $\epsilon > 0$. The capacity is the largest non-negative rate that is ϵ-achievable.

In this general memoryless framework, each relay can listen and transmit at the same time, i.e., it is a FD node. HD channels are a special case of the memoryless FD framework in the following sense [18]. With a slight abuse of notation compared to the previous paragraph, we let the channel input of the k-th relay, $k \in [1:N]$, be the pair (X_k, S_k), where $X_k \in \mathcal{X}_k$ as before and $S_k \in [0:1]$ is the *state* random variable that indicates whether the k-th relay is in receive-mode ($S_k = 0$) or in transmit-mode ($S_k = 1$). In the HD case the transition probability is specified as $\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]},S_{[1:N]}}$. In particular, when the k-th relay, $k \in [1:N]$, is listening ($S_k = 0$) the outputs are independent of X_k, while when the k-th relay is transmitting ($S_k = 1$) its output Y_k is independent of all other random variables.

For the particular case of $N = 1$ studied in this chapter, the general memoryless channel is shown in Figure 2.1, where for notation convenience, we use the subscripts s for the source, r for the relay, and d for the destination; the memoryless HD channel transition probability for $N = 1$ is hence defined by $\mathbb{P}_{Y_r,Y_d|X_s,X_r,S_r=0} := \mathbb{P}^{(0)}_{Y_r,Y_d|X_s,S_r=0}$ and $\mathbb{P}_{Y_r,Y_d|X_s,X_r,S_r=1} := \mathbb{P}^{(1)}_{Y_d|X_s,X_r,S_r=1}\mathbb{P}^{(1)}_{Y_r|S_r=1}$.

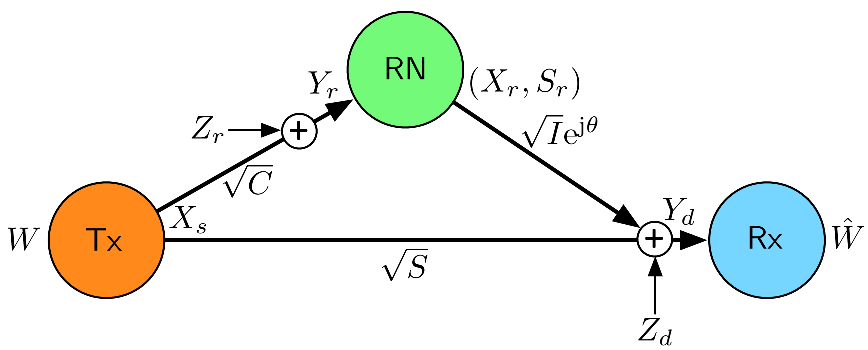


Figure 2.2: The Gaussian HD relay channel.

2.1.2 The Gaussian noise channel

The single-antenna complex-valued power-constrained Gaussian HD relay channel, shown in Figure 2.2, is described by the input/output relationship

$$Y_r = \sqrt{C} X_s \ (1 - S_r) + Z_r \in \mathbb{C}, \tag{2.1a}$$

$$Y_d = \sqrt{S} X_s + e^{j\theta} \sqrt{I} X_r \ S_r + Z_d \in \mathbb{C}, \tag{2.1b}$$

where the real-valued and non-negative channel power gains C, S, I and the phase θ are constant and therefore known to all terminals. Since a node can compensate for the phase of one of its channel gains, we can assume without loss of generality that the channel gains from the source to the other two terminals are real-valued and nonnegative. The channel inputs are subject to unitary average power constraints without loss of generality, i.e., $\mathbb{E}[|X_u|^2] \leq 1, \ u \in \{s, r\}$. The *switch* random variable S_r is binary. In our model, both X_r and S_r at any given time, are functions of the past received channel outputs. The noise (Z_d, Z_r) is a zero-mean proper-complex Gaussian random vector with, without loss of generality, unit entries on the main diagonal of the covariance matrix. In particular, but not without loss of generality [86], we assume that Z_d and Z_r are independent. In the following we consider the Gaussian HD relay channel for which $C > 0$ and $I > 0$, since for either $C = 0$ or $I = 0$ the relay is disconnected from either the source or the destination, respectively, so the channel reduces to a point-to-point channel with capacity equal to the direct-link capacity $\log(1 + S)$.

2.1.3 The deterministic / noiseless channel

The LDA approximates the Gaussian noise HD relay channel in (2.1) at high SNR. It is a deterministic channel with input-output relationship

$$Y_r = \mathbf{S}^{n-\beta_{\rm sr}} X_s \ (1 - S_r), \tag{2.2a}$$

$$Y_d = \mathbf{S}^{n-\beta_{\rm sd}} X_s + \mathbf{S}^{n-\beta_{\rm rd}} X_r \ S_r, \tag{2.2b}$$

for some non-negative integers $\beta_{\rm sr}, \beta_{\rm sd}, \beta_{\rm rd}$, where the vectors Y_r, Y_d, X_r, X_s are of length $n := \max\{\beta_{\rm sr}, \beta_{\rm sd}, \beta_{\rm rd}\}$ and take value in GF(2) (with GF we indicate the Galois Field), the sum is understood bit-wise on GF(2), $\mathbf{S}$ is the $n \times n$ shift matrix [19], and S_r is the relay binary-valued state random variable. The model has the following interpretation. The source sends a length-n vector X_s, whose top $\beta_{\rm sd}$ bits are received at the destination (through the source-destination link) and the top $\beta_{\rm sr}$ bits are received at the relay (through the source-relay link); similarly the relay sends a length-n vector X_r, whose top $\beta_{\rm rd}$ bits are received at the destination (through the relay-destination link). The fact that only a certain number of bits are observed at a given node is a consequence of the 'down shift' operation through the matrix $\mathbf{S}$. The bits not observed at a node are said to be 'below the noise floor'.

2.2 Overview of the main results

The capacity $C^{(\rm HD-RC)}$ of the Gaussian HD relay channel in (2.1) is unknown. Here we make progress toward determining its capacity by first establishing its gDoF, i.e., an exact "pre-log" capacity characterization in the limit for high SNR, and then by characterizing its capacity to within a constant gap at any finite SNR. Consider $\mathsf{SNR} > 0$ and the parameterization

$$S := \mathsf{SNR}^{\beta_{\rm sd}}, \text{source-destination link}, \tag{2.3a}$$

$$I := \mathsf{SNR}^{\beta_{\rm rd}}, \text{relay-destination link}, \tag{2.3b}$$

$$C := \mathsf{SNR}^{\beta_{\rm sr}}, \text{source-relay link}, \tag{2.3c}$$

for some non-negative real-valued triplet $(\beta_{\rm sd}, \beta_{\rm rd}, \beta_{\rm sr})$[2]. We define:

[2] We use the symbols $(\beta_{\rm sd}, \beta_{\rm rd}, \beta_{\rm sr})$ for both the LDA in (2.2) and the SNR parameterization in (2.3) for the channel power gains of the Gaussian HD relay channel in (2.1). In the former case $(\beta_{\rm sd}, \beta_{\rm rd}, \beta_{\rm sr}) \in \mathbb{N}^3$, while in the latter case $(\beta_{\rm sd}, \beta_{\rm rd}, \beta_{\rm sr}) \in \mathbb{R}_+^3$. This choice is motivated by the fact that the capacity of the LDA is related to the gDoF of the Gaussian HD relay channel.

Definition 1. *The gDoF of the Gaussian HD relay channel is defined as*

$$\mathsf{d}^{(\mathrm{HD-RC})} := \lim_{\mathsf{SNR}\to+\infty} \frac{C^{(\mathrm{HD-RC})}}{\log(1+\mathsf{SNR})}.$$

Definition 2. *The capacity $C^{(\mathrm{HD-RC})}$ is said to be known to within* GAP *bits if one can show achievable rates $R^{(\mathrm{in})}$ and outer bound $R^{(\mathrm{out})}$ such that*

$$R^{(\mathrm{in})} \le C^{(\mathrm{HD-RC})} \le R^{(\mathrm{out})} \le R^{(\mathrm{in})} + \mathsf{GAP}.$$

Our main results of this chapter are summarized as follows:

Theorem 1. *The gDoF of the Gaussian HD relay channel in* (2.1) *is*

$$\mathsf{d}^{(\mathrm{HD-RC})} = \begin{cases} \beta_{\mathrm{sd}} + \frac{(\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}})(\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}})}{(\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}})+(\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}})} & \text{for } \beta_{\mathrm{sr}} > \beta_{\mathrm{sd}},\ \beta_{\mathrm{rd}} > \beta_{\mathrm{sd}} \\ \beta_{\mathrm{sd}} & \text{otherwise} \end{cases}, \quad (2.4)$$

and the cut-set upper bound is achieved to within

Achievable scheme	*LDAi*	*CF*	*PDF*
analytical gap	*3 bits*	*1.61 bits*	*1 bit*
numerical gap	*1.32 bits*	*1.16 bits*	*1 bit*

where LDAi is an achievable scheme inspired by the LDA.

The result of Theorem 1 should be compared to a similar result for the FD case. The gDoF of the Gaussian FD relay channel is

$$\mathsf{d}^{(\mathrm{FD-RC})} = \beta_{\mathrm{sd}} + \min\{[\beta_{\mathrm{sr}} - \beta_{\mathrm{sd}}]^+, [\beta_{\mathrm{rd}} - \beta_{\mathrm{sd}}]^+\}, \quad (2.5)$$

and its capacity $C^{(\mathrm{FD-RC})}$ is achievable to within 1 bit by either DF or CF [19]. We notice that HD achieves the same gDoF of FD if $\min\{\beta_{\mathrm{rd}}, \beta_{\mathrm{sr}}\} \le \beta_{\mathrm{sd}}$, in which case the relay channel behaves gDoF-wise like a point-to-point channel from the source to the destination with gDoF given by β_{sd}. In both FD and HD the gDoF has a *routing* interpretation [19]: if the weakest link from the source to the destination through the relay is smaller than the direct link from the source to the destination, then direct transmission is optimal and the relay can be kept silent, otherwise it is optimal to communicate with the help of the relay, i.e., route part of the information through the relay.

Regarding gaps, we note that Theorem 1 improves on the 5 bits gap of [28]. Moreover, we note a tradeoff between the coding scheme complexity and the gap, with lower gaps for more complex schemes (for example, compare the gap of PDF with that of LDAi).

In an attempt to design simple and asymptotically optimal achievable schemes for the Gaussian HD relay channel, by following the footsteps of [19], we study the capacity of the LDA. We show:

Theorem 2. *The capacity of the LDA in* (2.2) *is given by*

$$C^{(\mathrm{HD})} = \begin{cases} \beta_{\mathrm{sd}} & \text{if } \beta_{\mathrm{sd}} \le \max\{\beta_{\mathrm{sr}}, \beta_{\mathrm{rd}}\} \\ \beta_{\mathrm{sd}} + \max_{\gamma\in[0,1]} \min\left\{\mathsf{A}(\gamma), \gamma(\beta_{\mathrm{sr}} - \beta_{\mathrm{sd}})\right\} & \text{otherwise} \end{cases} \tag{2.6}$$

where

$$\mathsf{A}(\gamma) := (1-\theta^*(\gamma)) \log \frac{1}{1-\theta^*(\gamma)} + \theta^*(\gamma) \log \frac{L-1}{\theta^*(\gamma)},$$
$$\theta^*(\gamma) := 1 - \max\left\{\frac{1}{L}, \gamma\right\}, \ L := 2^{(\beta_{\mathrm{rd}} - \beta_{\mathrm{sd}})},$$

and is achieved with random switch and correlated non-uniform input bits at the relay. Moreover, a scheme with deterministic switch and independent and identically distributed (i.i.d.) Bernoulli(1/2) *bits at the relay is at most 1 bit from the capacity in* (2.6).

2.3 The gDoF for the Gaussian HD relay channel

In this section, by adapting known bounds for the general memoryless FD relay channel [87] to the HD case with the methodology introduced by [18], we derive the gDoF of the Gaussian HD relay channel in (2.1).

2.3.1 Cut-set upper bounds

We now prove a number of upper bounds that we shall use for the converse part of Theorem 1. From the cut-set bound we have:

Proposition 1. *The capacity $C^{(\mathrm{HD-RC})}$ of the Gaussian HD relay channel is upper bounded as*

$$C^{(\mathrm{HD-RC})}$$
$$\le \min\left\{I(X_s, X_r, S_r; Y_d), I(X_s; Y_r, Y_d | X_r, S_r)\right\}\Big|_{(X_s,X_r,S_r)\sim\mathbb{P}^*_{X_s,X_r,S_r}} \tag{2.7a}$$
$$\le \max\min\left\{\mathcal{H}(\gamma) + \gamma I_1 + (1-\gamma) I_2, \gamma I_3 + (1-\gamma) I_4\right\} =: r^{(\mathrm{CS-HD})} \tag{2.7b}$$
$$\le 2 + \log(1+S)\left(1 + \frac{(b_1-1)(b_2-1)}{(b_1-1)+(b_2-1)}\right), \tag{2.7c}$$

where:

- *In* (2.7a)*: the distribution* $\mathbb{P}^*_{X_s,X_r,S_r}$ *is the one that maximizes the cut-set upper bound, i.e.,*

$$\mathbb{P}^*_{X_s,X_r,S_r} := \arg \max_{\mathbb{P}_{X_s,X_r,S_r}} \min \Big\{ I(X_s, X_r, S_r; Y_d), I(X_s; Y_r, Y_d | X_r, S_r) \Big\}.$$

- *In* (2.7b)*: the parameter* $\gamma := \mathbb{P}[S_r = 0] \in [0, 1]$ *represents the fraction of time the relay node listens,* $\mathcal{H}(\gamma)$ *is the binary entropy function*

$$\mathcal{H}(\gamma) := -\gamma \log(\gamma) - (1-\gamma) \log(1-\gamma), \tag{2.8}$$

the maximization is over the set

$$\gamma \in [0, 1], \tag{2.9a}$$
$$|\alpha_1| \le 1, \tag{2.9b}$$
$$(P_{u,0}, P_{u,1}) \in \mathbb{R}^4_+ : \gamma P_{u,0} + (1-\gamma) P_{u,1} \le 1, \ u \in \{s, r\}, \tag{2.9c}$$

and the mutual information terms $I_1, \ldots, I_4$ *are defined as*

$$I_1 := \log\left(1 + S \ P_{s,0}\right), \tag{2.10}$$
$$I_2 := \log\left(1 + SP_{s,1} + IP_{r,1} + 2|\alpha_1|\sqrt{SP_{s,1} \ IP_{r,1}}\right), \tag{2.11}$$
$$I_3 := \log\left(1 + (C + S) P_{s,0}\right), \tag{2.12}$$
$$I_4 := \log\left(1 + (1 - |\alpha_1|^2) S \ P_{s,1}\right). \tag{2.13}$$

- *In* (2.7c)*: the terms* b_1 *and* b_2 *are defined as*

$$b_1 := \frac{\log\left(1 + (\sqrt{I} + \sqrt{S})^2\right)}{\log(1+S)} > 1 \text{ since } I > 0, \tag{2.14}$$
$$b_2 := \frac{\log(1 + C + S)}{\log(1+S)} > 1 \text{ since } C > 0. \tag{2.15}$$

Proof. The proof can be found in Appendix 2.A. □

The upper bound in (2.7a) will be used to prove that PDF with random switch achieves the capacity to within 1 bit, the one in (2.7b) to prove that PDF with deterministic switch also achieves the capacity to within 1 bit and for numerical evaluations (since we do not know the distribution $\mathbb{P}^*_{X_s,X_r,S_r}$ that maximizes the cut-set upper bound in (2.7a)), and the one in (2.7c) for analytical computations such as the derivation of the gDoF. With the upper bound in Proposition 1 we can show:

Proposition 2. *The gDoF of the Gaussian HD relay channel is upper bounded by the Right Hand Side (RHS) of* (2.4)*.*

Proof. The proof can be found in Appendix 2.B. □

2.3.2 PDF lower bounds

In this section we prove a number of lower bounds that we shall use for the direct part of Theorem 1. From the achievable rate with PDF we have:

Proposition 3. *The capacity of the Gaussian HD relay channel is lower bounded as*

$$
\begin{aligned}
&C^{(\mathrm{HD-RC})}\\
&\geq \min\left\{I(U;Y_r|X_r,S_r)+I(X_s;Y_d|X_r,S_r,U), I(X_s,X_r,S_r;Y_d)\right\} && (2.16)\\
&\geq \max\min\left\{I_0^{(\mathrm{PDF})}+\gamma I_5+(1-\gamma)I_6, \gamma I_7+(1-\gamma)I_8\right\} =: r^{(\mathrm{PDF-HD})} && (2.17)\\
&\geq \log(1+S)\left(1+\frac{(c_1-1)(c_2-1)}{(c_1-1)+(c_2-1)}\right), && (2.18)
\end{aligned}
$$

where:

- *In* (2.16)*: we fix the input $\mathbb{P}_{U,X_s,X_r,S_r}$ to evaluate the PDF lower bound; in particular we set $\mathbb{P}_{X_s,X_r,S_r}$ to be the same distribution that maximizes the cut-set upper bound in* (2.7a) *and we choose either $U = X_r$ or $U = X_rS_r + X_s(1 - S_r)$.*

- *In* (2.17)*: the parameter $\gamma := \mathbb{P}[S_r = 0] \in [0, 1]$ represents the fraction of time the relay node listens, the maximization is over the set* (2.9a)-(2.9c) *as for the cut-set upper bound in* (2.7b)*, the mutual information terms $I_5, \dots, I_8$ are*

$$
\begin{aligned}
&I_5 := I_1 \textit{ given in } (2.10), && (2.19)\\
&I_6 := I_2 \textit{ given in } (2.11), && (2.20)\\
&I_7 := \log(1+\max\{C,S\}P_{s,0}) \leq I_3 \textit{ given in } (2.12), && (2.21)\\
&I_8 := I_4 \text{ given in } (2.13), && (2.22)
\end{aligned}
$$

and $I_0^{(\mathrm{PDF})} := I(S_r; Y_d)$ is computed from the density

$$
f_{Y_d}(t) = \frac{\gamma}{\pi v_0}\mathrm{e}^{-|t|^2/v_0} + \frac{1-\gamma}{\pi v_1}\mathrm{e}^{-|t|^2/v_1},\ t \in \mathbb{C}, \tag{2.23}
$$

with $v_0 = 2^{I_5}$ where I_5 is given in (2.19)*, and $v_1 = 2^{I_6}$ where I_6 is given in* (2.20)*.*

- *In* (2.18)*: the terms c_1 and c_2 are*

$$c_1 := \frac{\log(1+I+S)}{\log(1+S)} > 1 \text{ since } I > 0, \tag{2.24}$$

$$c_2 := \frac{\log(1+\max\{C,S\})}{\log(1+S)} > 1 \text{ since } C > 0. \tag{2.25}$$

Proof. The proof can be found in Appendix 2.C. □

The lower bound in (2.16) will be compared to the upper bound in (2.7a) to prove that PDF with random switch achieves capacity to within 1 bit, the one in (2.17) with the one in (2.7b) to prove that PDF with deterministic switch also achieves capacity to within 1 bit and for numerical evaluations, and the one in (2.18) will be used for analytical computations such as the evaluation of the achievable gDoF.

Remark 1. In Appendix 2.C, we found that the approximately optimal schedule for PDF with deterministic switch is

$$\gamma^*_{\rm PDF} := \frac{(c_1-1)}{(c_1-1)+(c_2-1)} \ \in \ [0,1],$$

where c_1 is given in (2.24) and c_2 is given in (2.25). The expression for $\gamma^*_{\rm PDF}$ can be understood as follows. Suppose that $\min\{C,I\} \geq S$, otherwise the relay is not used in the transmission and setting either $\gamma^*_{\rm PDF} = 0$ or $\gamma^*_{\rm PDF} = 1$ is approximately optimal. Notice that $\gamma^*_{\rm PDF}$ is a decreasing function in C and increasing in I. This implies that the stronger C compared to I the lesser the time the relay needs to listen to the channel to (partially) decode the source message. On the other hand, if $C < I$, more time is needed to learn the message and less time to convey the message to the destination.

With the lower bound in Proposition 3 we can show:

Proposition 4. *The gDoF of the Gaussian HD relay channel is lower bounded by the RHS of* (2.4)*.*

Proof. The proof can be found in Appendix 2.D. □

Propositions 2 and 4 prove that the gDoF of the Gaussian HD relay channel is given by (2.4) and that PDF achieves the gDoF.

Figure 2.3 shows the difference between the gDoF of the Gaussian FD relay channel in (2.5) and that of the Gaussian HD relay channel in (2.4) as a function of $\beta_{\rm sr}$ and $\beta_{\rm rd}$, where without loss of generality we fixed $\beta_{\rm sd} = 1$. This difference is zero when $\min\{\beta_{\rm rd}, \beta_{\rm sr}\} \leq \beta_{\rm sd} = 1$, in which case both

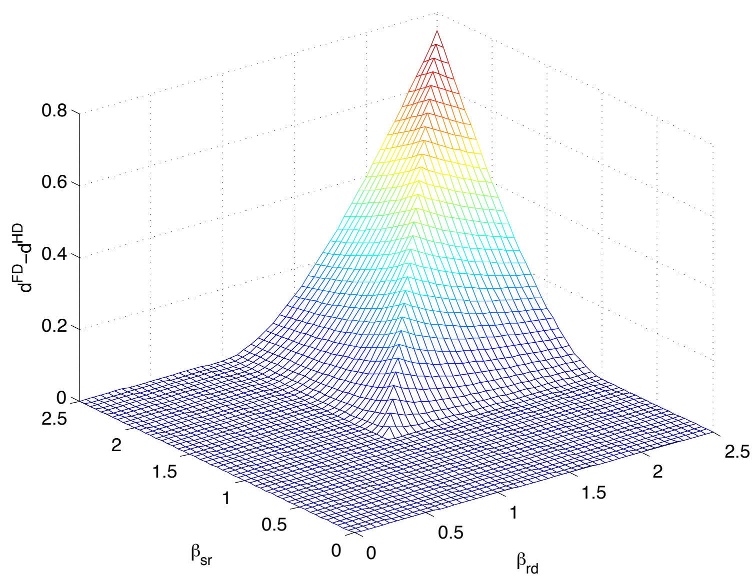


Figure 2.3: Difference between the gDoF of the Gaussian FD and of the Gaussian HD relay channels, for $\beta_{\rm sd} = 1$, as a function of $\beta_{\rm sr}$ and $\beta_{\rm rd}$.

the FD and the HD channels are gDoF-wise equivalent to a point-to-point channel without relay. When $\min\{\beta_{\rm rd}, \beta_{\rm sr}\} > \beta_{\rm sd} = 1$, the point-to-point communication channel is outperformed by the relay channel since now using the relay to convey the information is optimal. Moreover, as expected, the difference is always greater than or equal to zero because in the Gaussian FD relay channel the relay can simultaneously listen and transmit; therefore, the Gaussian FD relay channel represents an outer bound for the Gaussian HD relay channel. The largest difference occurs when $\beta_{\rm rd} = \beta_{\rm sr} := \beta_{\rm sd}\alpha$ in which case $\frac{\mathsf{d}^{(\rm FD-RC)}}{\beta_{\rm sd}} = \max\{1, \alpha\}$, while $\frac{\mathsf{d}^{(\rm HD-RC)}}{\beta_{\rm sd}} = \max\left\{1, \frac{1+\alpha}{2}\right\}$, in other words, for $\alpha > 1$ the rate difference between FD and HD grows unboundedly as SNR increases. This might motivate the use of more expensive FD relays in future wireless networks in this regime.

2.4 Capacity of the LDA and a simple achievable strategy for the Gaussian noise channel

In the previous section we showed that PDF achieves the gDoF of the Gaussian HD relay channel. PDF is based on block Markov encoding and joint decoding [87], which can be too complex to realize in practical systems. For this reason we seek now to design schemes that are simpler than PDF and that are still gDoF optimal. In order to do so, we consider the LDA in (2.2). Based on the many recent success stories, such as [19], we first determine the capacity achieving scheme for the LDA and we then try to 'translate' it into a gDoF-optimal scheme for the Gaussian HD relay channel. The rational is the "folk's theorem" that the capacity of the LDA gives the gDoF of the corresponding Gaussian noise channel.

2.4.1 Capacity of the LDA

The capacity of the general memoryless deterministic relay channel is given by the cut-set bound [14]. For the LDA the cut-set bound evaluates to (2.6) in Theorem 2, which is proved next.

Proof. The capacity of a HD channel is upper bounded by the capacity of the corresponding FD channel. Therefore for the capacity of the LDA we have $C^{(\mathrm{HD})} \le C^{(\mathrm{FD})}$ where $C^{(\mathrm{HD})}$ and $C^{(\mathrm{FD})}$ are defined as

$$\begin{aligned}
C^{(\mathrm{HD})} &:= \max_{\mathbb{P}_{X_s,X_r,S_r}} \min\Big\{I(X_s,X_r,S_r;Y_d), I(X_s;Y_r,Y_d|X_r,S_r)\Big\} \\
&= \max_{\mathbb{P}_{X_s,X_r,S_r}} \min\Big\{H(Y_d), H(Y_r,Y_d|X_r,S_r)\Big\} && (2.26) \\
C^{(\mathrm{FD})} &:= \max_{\mathbb{P}_{X_s,X_r}} \min\Big\{I(X_s,X_r;Y_d), I(X_s;Y_r,Y_d|X_r)\Big\} \\
&= \beta_{\mathrm{sd}} + \min\{[\beta_{\mathrm{rd}} - \beta_{\mathrm{sd}}]^+, [\beta_{\mathrm{sr}} - \beta_{\mathrm{sd}}]^+\}, && (2.27)
\end{aligned}$$

where $C^{(\mathrm{FD})}$ in (2.27) is achieved by i.i.d. Bernoulli(1/2) input bits for the source and the relay [19]. In order to evaluate $C^{(\mathrm{HD})}$ we distinguish two cases:

Regime 1: $\beta_{\mathrm{rd}} \le \beta_{\mathrm{sd}}$ or $\beta_{\mathrm{sr}} \le \beta_{\mathrm{sd}}$ in which case $C^{(\mathrm{HD})} \le C^{(\mathrm{FD})} = \beta_{\mathrm{sd}}$. Since the rate $C^{(\mathrm{HD})} = \beta_{\mathrm{sd}}$ can be achieved by silencing the relay and using i.i.d. Bernoulli(1/2) input bits for the source, we conclude that $C^{(\mathrm{HD})} = C^{(\mathrm{FD})} = \beta_{\mathrm{sd}}$ in this regime.

Regime 2: $\beta_{\mathrm{rd}} > \beta_{\mathrm{sd}}$ and $\beta_{\mathrm{sr}} > \beta_{\mathrm{sd}}$. Here we need to evaluate the expression in (2.26), for which we need to determine the optimal $H(Y_d)$ and

$$\begin{aligned}
H(Y_r,Y_d|X_r,S_r) &= \mathbb{P}[S_r=0]H(Y_r,Y_d|X_r,S_r=0) \\
&\quad + \mathbb{P}[S_r=1]H(Y_r,Y_d|X_r,S_r=1) \\
&\le \gamma\max\{\beta_{\mathrm{sr}},\beta_{\mathrm{sd}}\} + (1-\gamma)\beta_{\mathrm{sd}}.
\end{aligned}$$

To upper bound $H(Y_d)$, we write $Y_d = [Y_{d,u}, Y_{d,l}]$, where

- $Y_{d,l}$ contains the lower $\beta_{\rm sd}$ bits of Y_d. These bits are a combination of the bits of X_s and the lower bits of X_r. The lower bits of X_r are indicated as $X_{r,l}$. With reference to Figure 2.4(b), $Y_{d,l}$ corresponds to the portion of Y_d containing the "orange bits" labeled $b_1[2]$.

- $Y_{d,u}$ contains the upper $\beta_{\rm rd} - \beta_{\rm sd}$ bits of Y_d. These bits only depend on the upper bits of X_r. The upper bits of X_r are indicated as $X_{r,u}$. With reference to Figure 2.4(b), $Y_{d,u}$ corresponds to the portion of Y_d containing the "green bits" labeled a.

Hence we have

$$H(Y_d) = H(Y_{d,u}, Y_{d,l}) \leq H(Y_{d,u}) + H(Y_{d,l}) \leq H(Y_{d,u}) + \beta_{\rm sd},$$

since $Y_{d,l}$ contains $\beta_{\rm sd}$ bits and where $H(Y_{d,u})$ is computed from

$$\begin{aligned}\mathbb{P}[Y_{d,u} = y] &= \mathbb{P}[S_r = 0]\mathbb{P}[Y_{d,u} = y|S_r = 0] + \mathbb{P}[S_r = 1]\mathbb{P}[Y_{d,u} = y|S_r = 1] \\ &= \gamma\delta[y] + (1-\gamma)\mathbb{P}[X_{r,u} = y|S_r = 1],\end{aligned}$$

for $y \in [0 : L-1]$, $L := 2^{\beta_{\rm rd}-\beta_{\rm sd}} > 1$, where $\delta[y] = 1$ if $y = 0$ and zero otherwise, and where $\gamma := \mathbb{P}[S_r = 0]$. Let $\mathbb{P}[X_{r,u} = y|S_r = 1] = p_y \in [0, 1] : \sum_y p_y = 1$. Then, we have that

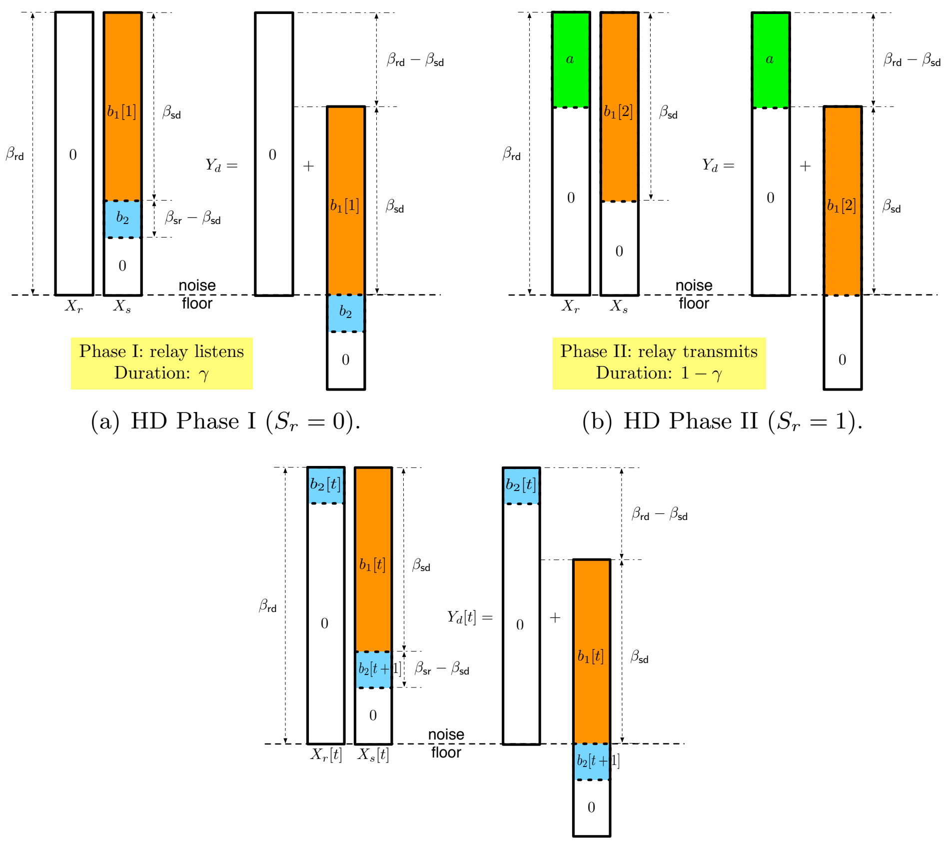
$$\begin{aligned}H(Y_{d,u}) &= H\Big(\big[\gamma + (1-\gamma)p_0, (1-\gamma)p_1, \ldots, (1-\gamma)p_{L-1}\big]\Big) \\ &\leq H\left(\left[\gamma + (1-\gamma)p_0,\ \underbrace{(1-\gamma)\frac{1-p_0}{L-1}, \ldots, (1-\gamma)\frac{1-p_0}{L-1}}_{L-1 \text{ times}}\right]\right) \\ &= (1-\theta)\log\frac{1}{1-\theta} + \theta\log\frac{L-1}{\theta}\Big|_{\theta:=(1-\gamma)(1-p_0)\in[0,1-\gamma]},\end{aligned} \tag{2.28}$$

which is maximized by

$$\theta^* = 1 - \max\{1/L, \gamma\} \Longleftrightarrow p_0^* = \frac{[1/L-\gamma]^+}{1-\gamma}. \tag{2.29}$$

Thus, collecting all the bounds, we have that $C^{(\rm HD)}$ in (2.26) is upper bounded as

$$C^{(\rm HD)} \leq \beta_{\rm sd} + \max_{\gamma\in[0,1]} \min\left\{(1-\theta^*)\log\frac{1}{1-\theta^*} + \theta^*\log\frac{L-1}{\theta^*}, \gamma(\beta_{\rm sr} - \beta_{\rm sd})\right\}. \tag{2.30}$$



(a) HD Phase I ($S_r = 0$).

(b) HD Phase II ($S_r = 1$).

(c) FD.

Figure 2.4: Achievable strategy for the LDA with $\beta_{\text{sd}} \leq \beta_{\text{sr}} \leq \beta_{\text{rd}}$.

In order to show the achievability of (2.30) consider the following inputs: the state S_r is Bernoulli$(1-\gamma)$ independent of any other random variable, and X_s and X_r are independent. The source uses i.i.d. Bernoulli$(1/2)$ bits. The relay uses i.i.d. Bernoulli(0) bits for $X_{r,l}$ and $\mathbb{P}[X_{r,u} = y] = p_0^*$ if $y = 0$ and $\mathbb{P}[X_{r,u} = y] = (1-p_0^*)/(L-1)$ otherwise, for p_0^* in (2.29), i.e., the components of $X_{r,u}$ are neither independent nor uniformly distributed. Notice that the distribution of $X_{r,u}$ in state $S_r = 0$ is irrelevant because its contribution at the destination is zero anyway, so we can assume that the input distribution for X_r is independent of the state S_r. It is straightforward to verify that this choice of input distribution achieves the upper bound in (2.30) thereby showing capacity in this regime. $\square$

Our motivation to determine the capacity of the LDA was to get 'inspiration' to design a simple achievable scheme for the Gaussian HD relay channel. While proving Theorem 2 we found that the capacity achieving distribution of the LDA has two fundamental features that can not be straightforwardly translated into a strategy for the Gaussian HD relay channel, namely: (i) the relay employs random switch, and (ii) correlated non-uniform inputs at the relay are optimal. Therefore next we further upper bound the capacity in (2.30) in the hope to get finally 'inspired'. Consider

$$\begin{aligned}
C^{(\mathrm{HD})} &= \max_{\mathbb{P}_{X_s,X_r,S_r}} \min\left\{H(Y_d), H(Y_r,Y_d|X_r,S_r)\right\} \\
&\leq \max_{\mathbb{P}_{X_s,X_r,S_r}} \min\left\{H(Y_d|S_r), H(Y_r,Y_d|X_r,S_r)\right\} + H(S_r) \\
&\leq \max_{\gamma\in[0,1]} \min\{\gamma\beta_{\mathrm{sd}}+(1-\gamma)\max\{\beta_{\mathrm{sd}},\beta_{\mathrm{rd}}\}, \gamma\max\{\beta_{\mathrm{sd}},\beta_{\mathrm{sr}}\}+(1-\gamma)\beta_{\mathrm{sd}}\}+1 \\
&= \beta_{\mathrm{sd}} + \gamma^*_{\mathrm{LDA}}[\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}}]^+ + 1, \qquad (2.31)
\end{aligned}$$

where γ^*_{LDA} is the optimal $\gamma := \mathbb{P}[S_r = 0] \in [0,1]$ obtained by equating the two arguments within the min and is given by

$$\gamma^*_{\mathrm{LDA}} := \begin{cases} \frac{(\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}})}{(\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}})+(\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}})} & \text{if } \beta_{\mathrm{rd}} > \beta_{\mathrm{sd}}, \beta_{\mathrm{sr}} > \beta_{\mathrm{sd}} \\ 0 & \text{otherwise.} \end{cases} \qquad (2.32)$$

We now show that the upper bound in (2.31) is achievable to within 1 bit. This 1 bit represents the maximum amount of information $I(S_r; Y_d)$ that could be conveyed to the destination through a random switch at the relay. If we neglect this 1 bit we can achieve the upper bound in (2.31) with the scheme shown in Figure 2.4(a) and Figure 2.4(b) for the case $\min\{\beta_{\mathrm{sr}}, \beta_{\mathrm{rd}}\} > \beta_{\mathrm{sd}}$, which is the case where the upper bound differs from direct transmission and for which $X_r \neq 0$. In Phase I / Figure 2.4(a) the relay listens and the source sends b_1 (of length β_{sd} bits) directly to the destination and b_2 (of length $\beta_{\mathrm{sr}} - \beta_{\mathrm{sd}}$ bits) to the relay; note that b_2 is below the noise floor at the destination; the duration of Phase I is γ, hence the relay has accumulated $\gamma(\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}})$ bits to forward to the destination. In Phase II / Figure 2.4(b) the relay forwards the bits learnt in Phase I to the destination by 'repackaging' them into a (of length $\beta_{\mathrm{rd}} - \beta_{\mathrm{sd}}$ bits); the source keeps sending a new b_1 (of length β_{sd} bits) directly to the destination; note that a does not interfere with b_2 at the destination; the duration of Phase II is such that all the bits accumulated by the relay in Phase I can be delivered to the destination, i.e.,

$$\gamma(\beta_{\mathrm{sr}} - \beta_{\mathrm{sd}}) = (1-\gamma)(\beta_{\mathrm{rd}} - \beta_{\mathrm{sd}}),$$

giving precisely the optimal γ^*_{LDA} in (2.32). The total number of bits decoded at the destination is $1 \cdot \beta_{\text{sd}} + \gamma^*_{\text{LDA}} \cdot (\beta_{\text{sr}} - \beta_{\text{sd}})$, which shows that the rate in (2.31) is achievable to within 1 bit. Notice that the LDA-rate in (2.31), besides the 1 bit term, looks formally the same as the gDoF in (2.4) after straightforward manipulations.

The scheme that is optimal within 1 bit for the LDA uses deterministic switch and i.i.d. Bernoulli(1/2) input bits, similarly to the FD optimal scheme in [19]; therefore, similarly to the FD case, we are now in the position to obtain a scheme for the original Gaussian HD relay channel. Before we describe the scheme for the Gaussian noise channel, let us compare the results obtained for the LDA. The HD optimal strategy in Figure 2.4(a) and Figure 2.4(b) should be compared with the FD optimal strategy in Figure 2.4(c). In Figure 2.4(c), in a given time slot t, the source sends $b_1[t]$ (of length β_{sd} bits) directly to the destination and $b_2[t+1]$ (of length at most $\beta_{\text{sr}} - \beta_{\text{sd}}$ bits) to the relay; the relay decodes both $b_1[t]$ and $b_2[t+1]$ and forwards $b_2[t+1]$ in the next slot; in slot t the relay sends $b_2[t]$ (of length at most $\beta_{\text{rd}} - \beta_{\text{sd}}$ bits) to the destination; the number of bits the relay forwards must be the minimum among the number of bits the relay can decode (given by $\beta_{\text{sr}} - \beta_{\text{sd}}$) and the number of bits that can be decoded at the destination without harming the direct transmission from the source (given by $\beta_{\text{rd}} - \beta_{\text{sd}}$). Therefore, the total number of bits decoded at the destination is $\beta_{\text{sd}} + \min\{\beta_{\text{rd}} - \beta_{\text{sd}}, \beta_{\text{sr}} - \beta_{\text{sd}}\}$, which formally looks exactly as the optimal gDoF for the Gaussian FD relay channel in (2.5) in the case the relay is actually used.

Figure 2.5 compares the capacities of the FD and HD LDA channels; it also shows some achievable rates for the HD LDA channel. In particular, the capacity of the FD channel is given by (2.5) (dotted black curve labeled “FD”), the capacity of the HD channel is given by (2.6) (solid black curve labeled “HD” obtained with the optimal p_0^* in (2.29)) and its upper bound by (2.31) (red curve labeled “HDlda upper”). For comparison we also show the performance when the source uses i.i.d. Bernoulli(1/2) bits and the relay uses one of the following strategies: i.i.d. Bernoulli(q) bits and random switch (blue curve labeled “HDiid q+rand” obtained by numerically optimizing $q \in [0,1]$), i.i.d. Bernoulli(1/2) bits and random switch (green curve labeled “HDiid 1/2+rand” obtained with $p_0 = 1/L$ in (2.28)), and i.i.d. Bernoulli(1/2) bits and deterministic switch (magenta curve labeled “HDiid 1/2+det” and given by $\beta_{\text{sd}} + \min\{\gamma[\beta_{\text{sr}} - \beta_{\text{sd}}]^+, (1-\gamma)[\beta_{\text{rd}} - \beta_{\text{sd}}]^+\}$). We can draw some interesting conclusions from Figure 2.5:

- With deterministic switch: i.i.d. Bernoulli(1/2) bits for the relay are

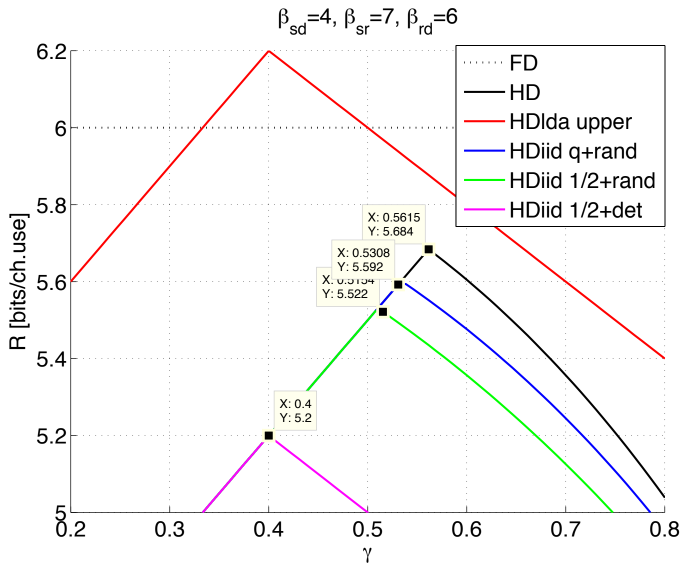


Figure 2.5: Comparison of the capacities of the LDA for both HD and FD modes of operation at the relay.

optimal but this choice is quite far from capacity (magenta curve vs. solid black curve); this choice however is at most 1 bit from optimal (magenta curve vs. red curve).

- With random switch: the optimal input distribution for the relay is not i.i.d. bits; i.i.d. inputs incur a rate loss (blue curve vs. solid black curve); if in addition we insist on i.i.d. Bernoulli(1/2) bits for the relay we incur a further loss (green curve vs. blue curve).

This shows that for optimal performance the relay inputs are correlated and that random switch must be used.

2.4.2 LDAi: an achievable strategy for the Gaussian HD relay channel inspired by the LDA

We mimic the LDA strategy with deterministic switch from Section 2.4.1 so as to get an achievable rate for the Gaussian HD relay channel. We assume $S < C$, otherwise we use direct transmission to achieve $R = \log(1 + S)$. The transmission is divided into two phases (it might help to refer to Figure 2.4(a) and Figure 2.4(b)):

• Phase I of duration γ: the transmit signals are

$$X_s[1] = \sqrt{1-\delta}X_{b_1[1]} + \sqrt{\delta}X_{b_2},\ \delta := \frac{1}{1+S},$$
$$X_r[1] = 0.$$

The relay applies successive decoding of $X_{b_1[1]}$ followed by X_{b_2} from

$$Y_r[1] = \sqrt{C}\ \sqrt{1-\delta}X_{b_1[1]} + \sqrt{C}\ \sqrt{\delta}X_{b_2} + Z_r[1],$$

which is possible if (rates are normalized by the total duration of the two phases)

$$R_{b_1[1]} \le \gamma \log(1+C) - \gamma \log\left(1 + \frac{C}{1+S}\right),$$
$$R_{b_2} \le \gamma \log\left(1 + \frac{C}{1+S}\right). \tag{2.33}$$

The destination decodes $X_{b_1[1]}$ treating X_{b_2} as noise from

$$Y_d[1] = \sqrt{S}\ \sqrt{1-\delta}X_{b_1[1]} + \sqrt{S}\ \sqrt{\delta}X_{b_2} + Z_d[1],$$

which is possible if

$$R_{b_1[1]} \le \gamma \log(1+S) - \gamma \log\left(1 + \frac{S}{1+S}\right). \tag{2.34}$$

Finally, since we assume $S < C$, Phase I is successful if (2.33) and (2.34) are satisfied.

• Phase II of duration $1-\gamma$: the transmit signals are

$$X_s[2] = X_{b_1[2]},$$
$$X_r[2] = X_{b_2},$$

recall that the bits in a in Figure 2.4(b) are the exact same bits in b_2 in Figure 2.4(a) just 'repacked' to form a vector with different length, which we mimic here by setting $X_r[2] = X_{b_2}$.

The destination applies successive decoding of X_{b_2} (by exploiting also the information about b_2 that it gathered in the first phase) followed by $X_{b_1[2]}$ from

$$Y_d[2] = \sqrt{S}X_{b_1[2]} + \mathrm{e}^{+\mathrm{j}\theta}\sqrt{I}X_{b_2} + Z_d[2],$$

which is possible if

$$R_{b_2} \le (1-\gamma)\log\left(1+\frac{I}{1+S}\right)+\gamma\log\left(1+\frac{S}{1+S}\right), \tag{2.35}$$

$$R_{b_1[2]} \le (1-\gamma)\log(1+S). \tag{2.36}$$

• By imposing that the rate R_{b_2} is the same in both phases, that is, that (2.33) and (2.35) are equal, we get that γ should be chosen equal to γ^*

$$\gamma^* = \frac{\log\left(1+\frac{I}{1+S}\right)}{\log\left(1+\frac{I}{1+S}\right)+\log\left(1+\frac{C}{1+S}\right)-\log\left(1+\frac{S}{1+S}\right)}. \tag{2.37}$$

Note that γ^* in (2.37) tends to γ^*_{LDA} in (2.32) as SNR increases by using the parameterization in (2.3). Moreover we give here an explicit closed form expression for the optimal duration of the time the relay listens to the channel.

The rate sent directly from the source to the destination, that is, the sum of (2.34) and (2.36), is

$$R_{b_1[1]} + R_{b_1[2]} = \log(1+S) - \underbrace{\gamma^*\log\left(1+\frac{S}{1+S}\right)}_{\in[0,1]}. \tag{2.38}$$

Therefore the total rate decoded at the destination through the two phases is $r^{(\mathrm{LDAi-HD})} := R_{b_1[1]} + R_{b_1[2]} + R_{b_2}$ as in Proposition 5 below:

Proposition 5. *The capacity of the Gaussian HD relay channel is lower bounded as* $C^{(\mathrm{HD-RC})} \ge r^{(\mathrm{LDAi-HD})}$*, with*

$$r^{(\mathrm{LDAi-HD})} = \log(1+S) + \frac{\log\left(1+\frac{I}{1+S}\right)\left[\log\left(1+\frac{C}{1+S}\right)-\log\left(1+\frac{S}{1+S}\right)\right]^+}{\log\left(1+\frac{I}{1+S}\right)+\left[\log\left(1+\frac{C}{1+S}\right)-\log\left(1+\frac{S}{1+S}\right)\right]^+}. \tag{2.39}$$

We notice that the rate expression for $r^{(\mathrm{LDAi-HD})}$ in (2.39) (please notice the operator $[\cdot]^+$), which was derived under the assumption $C > S$, is valid for all C since for $C < S$ it reduces to direct transmission from the source to the destination. Moreover we can show that:

Proposition 6. *The LDAi strategy achieves the gDoF in* (2.4).

Proof. The proof can be found in Appendix 2.E. □

Remark 2. The LDAi scheme can be seen as a specialization of PDF with deterministic switch at the relay combined with the scheduling and power splits inspired by the analysis of the LDA channel. The specialization consists of the classical PDF with sliding window decoding and without coherent codebooks [87]. Thus, the same observations drawn for γ^*_{PDF} in Remark 1 also hold for the LDAi schedule γ^* in (2.37).

Before concluding this section, we point out some important practical aspects of the LDAi that are worth noticing:

1. The proposed scheme is not the classical block Markov encoding scheme with backward decoding; in particular, the destination uses sliding window decoding, which simplifies the decoding procedure and incurs no delay; a further simplification would be to consider a slot-by-slot decoding scheme.

2. The destination uses successive decoding, which is simpler than joint decoding.

3. No power allocation is applied at the source or at the relay across the two phases; this simplifies the encoding procedure and can be used for time-varying channels as well. The source uses superposition coding, i.e., power split, only to 'route' part of its data through the relay.

2.5 Analytical gaps

In Sections 2.3 and 2.4 we described upper and lower bounds to determine the gDoF of the Gaussian HD relay channel. In Section 2.4 we proposed a scheme inspired by the analysis of the LDA channel that also achieves the optimal gDoF. We now show that the same upper and lower bounds are to within a constant gap of one another thereby concluding the proof of Theorem 1. We consider both the case of random switch and of deterministic switch for the relay. For completeness we also consider the CF lower bound.

Proposition 7. *PDF with random switch is optimal to within 1 bit.*

Proof. The proof can be found in Appendix 2.F. □

Proposition 8. *PDF with deterministic switch is optimal to within 1 bit.*

Proof. The proof can be found in Appendix 2.G. □

The intuition of why the gap does not improve with random switch is that there exist channel parameters for which direct transmission is approximately optimal (when $\min\{C, I\} \leq S$); in the case of direct transmission there are no benefits to use the relay at all and silencing the relay is a case of deterministic switch.

Proposition 9. *LDAi is optimal to within 3 bits.*

Proof. The proof can be found in Appendix 2.H. □

For completeness, we conclude this section with a discussion on the gap that can be obtained with CF. For the Gaussian FD relay channel, it is known that CF represents a good alternative to PDF in the case when the link between the source and the relay is weaker than the direct link [87]. The CF achievable rate is presented in Appendix 2.I. By using Remark 5 in Appendix 2.I we have:

Proposition 10. *CF with deterministic switch is optimal to within 1.61 bits.*

Proof. The proof can be found in Appendix 2.J. □

Remark 3. In Appendix 2.I, we found that the approximately optimal schedule with CF and deterministic switch is given by

$$\begin{aligned}
\gamma^*_{\mathrm{CF}} &:= \frac{(c_5 - 1)}{(c_5 - 1) + (c_6 - 1)} \;\in\; [0, 1], \\
c_5 &:= \frac{\log(1 + I + S)}{\log(1 + S)} > 1 \text{ since } I > 0, \\
c_6 &:= \frac{\log\left(1 + \frac{C}{1+\sigma_0^2} + S\right)}{\log(1 + S)} > 1 \text{ since } C > 0.
\end{aligned}$$

Suppose, as in Remark 1 for PDF, that $\min\{C, I\} \geq S$, otherwise setting either $\gamma^*_{\mathrm{CF}} = 0$ or $\gamma^*_{\mathrm{CF}} = 1$ is approximately optimal. Notice that, although the same observations drawn from the analysis of γ^*_{PDF} in Remark 1 hold, γ^*_{CF} here also depends on the variance of the quantization noise at the relay, i.e., σ_0^2. The schedule γ^*_{CF} is an increasing function of σ_0^2, meaning that the higher σ_0^2 the longer the time the relay should listen to the channel. Therefore, differently from PDF, the approximately optimal schedule does not only depend on the channel gains, but also on the level at which the signal at the relay is quantized.

Proposition 11. *CF with random switch is optimal to within 1.61 bits.*

Proof. Random switch improves on deterministic switch, since at most 1 bit of further information may be conveyed to the destination by randomly switching between the transmit- and receive-phases. Thus, it follows that any rate achievable with deterministic switch is also achievable with random switch, i.e., random switch can not increase the gap. □

2.6 Numerical gaps

In this section we show that the gap results obtained in Section 2.5 are pessimistic and are due to crude bounding of the upper and lower bounds, which was necessary in order to obtain rate expressions that can be handled analytically. In order to illustrate our point, we first consider a relay network without the source-destination link, that is, with $S = 0$, and then we show that the same observations are valid for any network.

2.6.1 Gaussian HD relay channel without a source-destination link (single-relay line network)

Upper Bound: We start by showing that the (upper bound on the) cut-set upper bound in (2.7b) can be improved upon. Note that we were not able to evaluate the actual cut-set upper bound in (2.7a) so we further bounded it as in (2.7b), which for $S = 0$ reduces to

$$r^{(\mathrm{CS-HD})}|_{S=0} = \max_{\gamma\in[0,1]} \min\left\{\mathcal{H}(\gamma)+(1-\gamma)\log\left(1+\frac{I}{1-\gamma}\right),\ \gamma\log\left(1+\frac{C}{\gamma}\right)\right\}.$$

The capacity of the Gaussian FD relay channel for $S = 0$ is known exactly and is given by the cut-set upper bound, i.e., $C^{(\mathrm{FD})}|_{S=0} = \log(1+\min\{C, I\})$. $C^{(\mathrm{FD})}$ is a trivial upper bound for the capacity of the Gaussian HD relay channel. Now we show that our upper bound $r^{(\mathrm{CS-HD})}|_{S=0}$ can be larger than $C^{(\mathrm{FD})}|_{S=0}$. For the case $C = 15/2 > I = 3/2$ we have

$$\begin{aligned} r^{(\mathrm{CS-HD})}|_{S=0} &\geq \min\left\{\mathcal{H}\left(\frac{1}{2}\right) + \frac{1}{2}\log(1+2I), \frac{1}{2}\log(1+2C)\right\} \\ &= \log(4) > C^{(\mathrm{FD})}|_{S=0} = \log(2.5). \end{aligned}$$

The reason why the capacity of the FD channel can be smaller than our upper bound $r^{(\mathrm{CS-HD})}|_{S=0}$ is the crude bound $I(S_r; Y_d) \leq H(S_r) = \mathcal{H}(\gamma)$. As mentioned earlier, we needed this bound in order to have an analytical expression for the upper bound. Actually for $S = 0$ the cut-set upper bound in (2.7a) is tight, as we show next.

Exact capacity with PDF:

Corollary 1. *In absence of direct link between the source and the destination PDF with random switch achieves the cut-set upper bound.*

Proof. The single-relay line network represents an example of *degraded* relay channel since $X_s - (X_r, S_r, Y_r) - Y_d$ forms a Markov chain. The capacity of the general memoryless degraded relay channel is exactly known [14, Theorem 1], i.e., for this network the cut-set upper bound is tight. Therefore, our result is a special case of [14, Theorem 1]. □

Improved gap for the LDAi lower bound: Despite knowing the capacity expression for $S = 0$ from Corollary 1, its actual evaluation is elusive as it is not clear what the optimal input distribution $\mathbb{P}^*_{X_s,X_r,S_r}$ in (2.7a) is. For this reason we next specialize the LDAi strategy to the case $S = 0$ and evaluate its gap from the (upper bound on the) cut-set bound in (2.7b).

The LDAi achievable rate in (2.39) with $S = 0$ is

$$r^{(\mathrm{LDAi-HD})}|_{S=0} = \max_{\gamma\in[0,1]} \min\{\gamma \log(1 + C), (1 - \gamma)\log(1 + I)\},$$

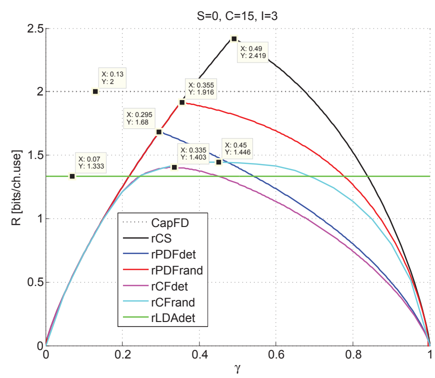
where we left intentionally explicit the optimization with respect to γ, and where we note that $r^{(\mathrm{LDAi-HD})}|_{S=0}$ coincides with the PDF lower bound with deterministic switch at the relay and without optimizing the powers between the relay transmit- and receive-phases. The gap between the outer bound and $r^{(\mathrm{LDAi-HD})}|_{S=0}$ is less than 3 bits since

$$\begin{aligned}
\mathsf{GAP} &\le r^{(\mathrm{CS-HD})}|_{S=0} - r^{(\mathrm{LDAi-HD})}|_{S=0}\\
&\le \max_{\gamma\in[0,1]} \Bigg\{\gamma \log\left(1 + \frac{C}{\gamma}\right) - \gamma\log(1 + C),\\
&\qquad \mathcal{H}(\gamma) + (1 - \gamma)\log\left(1 + \frac{I}{1-\gamma}\right) - (1-\gamma)\log(1 + I)\Bigg\}\\
&\le \max_{\gamma\in[0,1]} \left\{\gamma \log\left(\frac{1}{\gamma}\right), \mathcal{H}(\gamma) + (1 - \gamma)\log\left(\frac{1}{1-\gamma}\right)\right\}\\
&= \max_{\gamma\in[0,1]} \left\{\mathcal{H}(\gamma) + (1 - \gamma)\log\left(\frac{1}{1-\gamma}\right)\right\} = 1.5112 \text{ bits.}
\end{aligned}$$

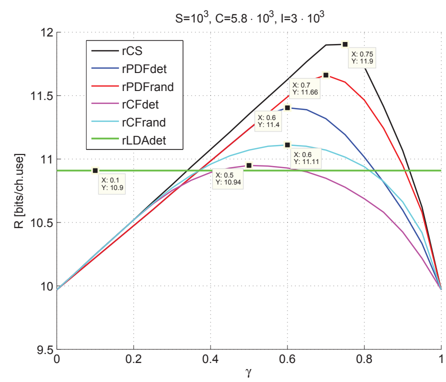
Note that the actual gap is even less than 1.5 bits. In fact, by numerically evaluating

$$\mathsf{GAP} = \min\{C^{(\mathrm{FD})}, r^{(\mathrm{CS-HD})}\}|_{S=0} - r^{(\mathrm{LDAi-HD})}|_{S=0}$$

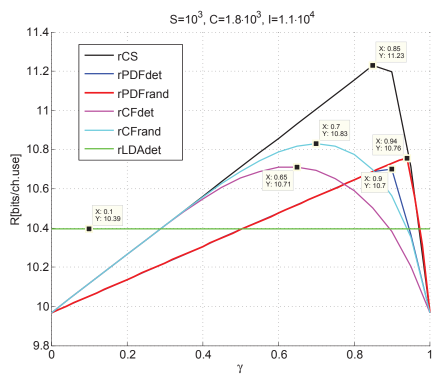
one can found that the gap is at most 1.11 bits.



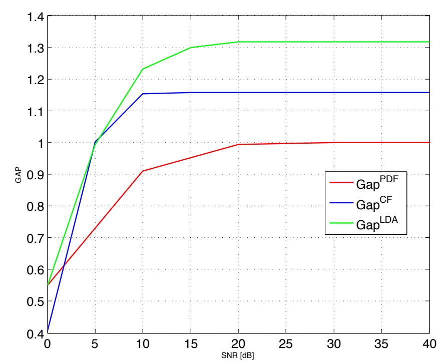
(a) $S\!=\!0,\ C\!=\!15,\ I\!=\!3.$



(b) $S\!=\!30$ dB, C=37.63 dB, I=34.77 dB.



(c) $S\!=\!30$ dB, C=32.55 dB, I=40.41 dB.



(d) Gap for $\beta_{\rm sd}\!=\!1$ and $(\beta_{\rm rd},\beta_{\rm sr})\!\in\![0,2.4]$.

Figure 2.6: Numerical evaluation of the various achievable schemes.

Numerical gaps with deterministic switch: Similarly to what done for the LDAi, by numerical evaluations one can find that the PDF strategy with deterministic switch in Remark 4-Appendix 2.C and the CF strategy with deterministic switch in Remark 5-Appendix 2.I are to within 0.80 bits and 1.01 bits, respectively, of the improved bound $\min\{C^{(\rm FD)}, r^{(\rm CS-HD)}\}|_{S=0}$. Note that in these cases there is no information conveyed by the relay to the destination through the switch.

Figure 2.6(a) shows different upper an lower bounds for the Gaussian HD relay channel for $S = 0,\ C = 15,\ I = 3$ versus $\gamma = \mathbb{P}[S_r = 0]$. We see that the cut-set upper bound (solid black curve) exceeds the capacity of the Gaussian FD relay channel (dashed black curve). Different achievable strategies are also shown, whose order from the most performing to the least performing is: PDF with random switch (red curve with maximum rate 1.916 bits/ch.use), PDF with deterministic switch (blue curve with maximum rate 1.68 bits/ch.use), CF with random switch (cyan curve with maximum rate 1.446 bits/ch.use), CF with deterministic switch (magenta curve with maximum rate 1.403 bits/ch.use), and LDAi (green curve with maximum rate 1.333 bits/ch.use). In this particular setting, the maximum rate using the CF strategy with random switch (cyan curve with maximum rate 1.446 bits/ch.use) is achieved for $\mathbb{P}[Q=0,S_r=0]=0, \mathbb{P}[Q=0,S_r=1]\approx 0.33, \mathbb{P}[Q=1,S_r=0]\approx 0.45, \mathbb{P}[Q=1,S_r=1]\approx 0.22$. This is due to the absence of the direct link ($S=0$) between the source and the destination. Actually, since the source can communicate with the destination only through the relay, it is necessary a coordination between the transmissions of the source and those of the relay. This coordination is possible thanks to the time-sharing random variable Q, i.e., when $Q=0$ the source stays silent while when $Q=1$ the source transmits.

2.6.2 Gaussian HD relay channel with direct link

Figure 2.6(b) and Figure 2.6(c) show the rates achieved by using the different achievable schemes presented in the previous sections for a channel with $S>0$. In Figure 2.6(b) the channel conditions are such that PDF outperforms CF, while in Figure 2.6(c) the opposite holds. In Figure 2.6(b) the PDF strategy with random switch (red curve with maximum rate 11.66 bits/ch.use) outperforms both the CF with random switch (cyan curve with maximum rate 11.11 bits/ch.use) and the PDF with deterministic switch (blue curve with maximum rate 11.4 bits/ch.use); then the PDF with deterministic switch outperforms the CF with deterministic switch (magenta curve with maximum rate 10.94 bits/ch.use), which is also encompassed by the CF with random switch. Differently from the case without direct link, we observe that the maximum CF rates both in Figure 2.6(b) and in Figure 2.6(c) are achieved with the choice $Q=\emptyset$, i.e., the time-sharing random variable Q is a constant. This is due to the fact that the source is always heard by the destination even when the relay transmits so there is no need for the source to remain silent when the relay sends.

Figure 2.6(d) shows, as a function of SNR and for $\beta_{\rm sd}=1, (\beta_{\rm rd},\beta_{\rm sr})\in[0,2.4]$, the maximum gap between the cut-set upper bound $r^{(\rm CS-HD)}$ in (2.7b) and the following lower bounds with deterministic switch: the PDF lower bound obtained from $r^{(\rm PDF-HD)}$ in (2.17) with $I_0^{(\rm PDF)}=0$, the CF lower bound in Remark 5 in Appendix 2.I, and the LDAi lower bound in (2.39). From Figure 2.6(d) we observe that the maximum gap with PDF is 1 bit as

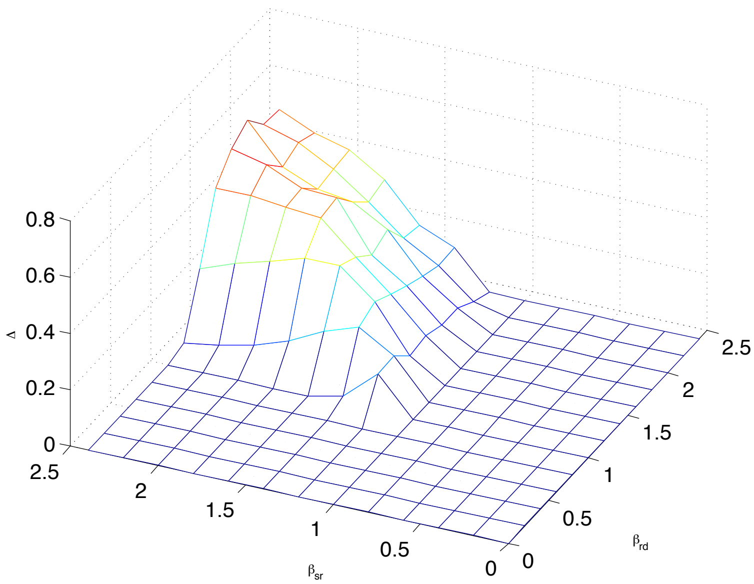


Figure 2.7: $\Delta = r^{(\mathrm{PDF-HD})}|_{I_0^{(\mathrm{PDF})}=I_0^{\mathrm{opt}}} - r^{(\mathrm{PDF-HD})}|_{I_0^{(\mathrm{PDF})}=0}$ at $\mathrm{SNR} = 20\mathrm{dB}$ for $\beta_{\mathrm{sd}} = 1$ as a function of $(\beta_{\mathrm{rd}}, \beta_{\mathrm{sr}}) \in [0, 2.4]$.

in Proposition 8, but with CF the gap is around 1.16 bits and with LDAi around 1.32 bits, which are lower than the analytical gaps found in Propositions 10 and 9, respectively.

The lower bounds can be improved upon by considering that information can be transmitted through a random switch. However, this improvement depends on the channel gains. If the information can not be routed through the relay because $\min\{C, I\} \leq S$, then the system can not exploit the randomness of the switch, and so $I_0^{(\mathrm{PDF})} = 0$ and $I_0^{(\mathrm{CF})} = 0$ are approximately optimal (in this case the relay can remain silent). This behavior for the PDF strategy is represented in Figure 2.7. In this figure we numerically evaluate the difference between the analytical gap, i.e., the one computed with $I_0^{(\mathrm{PDF})} = 0$, and the numerical one, i.e., computed with the optimal $I_0^{(\mathrm{PDF})}$ indicated as I_0^{opt} (i.e., I_0^{opt} is the actual value of $I_0^{(\mathrm{PDF})}$), at a fix $\mathrm{SNR} = 20$ dB and by varying $(\beta_{\mathrm{rd}}, \beta_{\mathrm{sr}})$. We observe that when the information can not be conveyed through the relay, i.e., $\min\{\beta_{\mathrm{rd}}, \beta_{\mathrm{sr}}\} \leq 1$, then $I_0^{(\mathrm{PDF})} = 0$ is optimal, since the information only flows through the direct link. On the other hand, when $\min\{\beta_{\mathrm{rd}}, \beta_{\mathrm{sr}}\} > 1$, random switch outperforms deterministic switch. Moreover, from Figure 2.7 we observe that, the

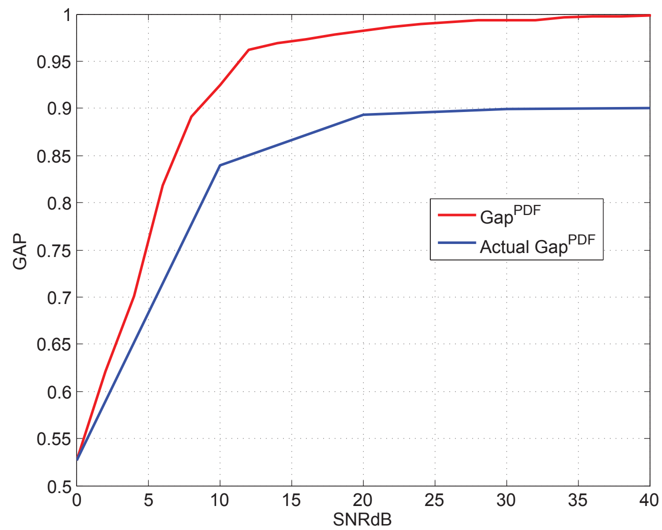


Figure 2.8: Numerical evaluation of the maximum gap varying the SNR for $\beta_{\rm sd} = 1$ and $(\beta_{\rm rd}, \beta_{\rm sr}) \in [1.2, 2.4]$ with deterministic (red curve) and random switch (blue curve).

stronger the channel gains along the path through the relay the larger the amount of information conveyed by random switch.

In Figure 2.8 the channel gains are set such that the use of the relay increases the gDoF of the channel ($\beta_{\rm sd} = 1$ and $(\beta_{\rm rd}, \beta_{\rm sr}) \in [1.2, 2.4]$). Here the relay uses PDF. We observe that we have a further improvement in terms of gap by using a random switch (blue curve) instead of using a deterministic switch (red curve). We notice that at high SNR, where the gap is maximum, this improvement is around 0.1 bits. As mentioned earlier, the rate advantage of random switch over deterministic switch depends on the channel gains.

2.7 Conclusions and future directions

In this chapter we considered a system where a source communicates with a destination across a Gaussian channel with the help of a HD relay node. We determined the capacity of the LDA of the Gaussian noise channel at high SNR, by showing that random switch and correlated non-uniform input bits at the relay are optimal. We then analyzed the Gaussian noise channel at finite SNR; we derived its gDoF and showed several schemes that achieve the cut-set upper bound on the capacity to within a constant finite gap, uniformly for all channel parameters. We considered both the case of deterministic switch and of random switch at the relay. We showed that random switch is optimal and for the case without a direct link it achieves the exact capacity. In general random switch increases the achievable rate at the expense of more complex coding and decoding schemes. For each scheme, we determined in closed form the approximately optimal schedule, i.e., duration of the transmit- and receive-phases at the relay, to shed light into practical HD relays for future wireless networks.

With respect to the results presented in this chapter, interesting future research directions may include: (i) designing the switch so that further (at most 1 bit per channel use) information can be conveyed to the destination when CF is used; (ii) implementing the LDA-inspired scheme on an LTE simulation test bench and study the impact of using codes of finite length and discrete input constellations in contrast to asymptotically large block-length Gaussian codes in the spirit of [68].

Appendix

2.A Proof of Proposition 1

An outer bound on the capacity of the memoryless relay channel is given by the cut-set outer bound [87, Theorem 16.1] that specialized to our Gaussian HD relay channel gives

$$C^{(\mathrm{HD-RC})} \leq \max_{\mathbb{P}_{X_s,[X_r,S_r]}} \min \Big\{ I(X_s,[X_r,S_r];Y_d), I(X_s;Y_r,Y_d|[X_r,S_r]) \Big\} \tag{2.40a}$$

$$= \max_{\mathbb{P}_{X_s,X_r,S_r}} \min \Big\{ I(S_r;Y_d) + I(X_s,X_r;Y_d|S_r), I(X_s;Y_r,Y_d|X_r,S_r) \Big\} \tag{2.40b}$$

$$\leq \max_{\mathbb{P}_{X_s,X_r,S_r}} \min \Big\{ H(S_r) + I(X_s,X_r;Y_d|S_r), I(X_s;Y_r,Y_d|X_r,S_r) \Big\} \tag{2.40c}$$

$$\leq \max \min \Big\{ \mathcal{H}(\gamma)+\gamma I_1+(1-\gamma)I_2, \gamma I_3+(1-\gamma)I_4 \Big\} =: r^{(\mathrm{CS-HD})}, \tag{2.40d}$$

where the different steps follow since:

- We indicate the (unknown) distribution that maximizes (2.40a) as $\mathbb{P}^*_{X_s,X_r,S_r}$ in order to get the bound in (2.7a).
- In order to obtain the bound in (2.40c) we used the fact that, for a

discrete binary-valued random variable S_r, we have

$$I(S_r; Y_d) = H(S_r) - H(S_r|Y_d) \leq H(S_r) = \mathcal{H}(\gamma),$$

for some $\gamma := \mathbb{P}[S_r = 0] \in [0, 1]$ that represents the fraction of time the relay listens and where $\mathcal{H}(\gamma)$ is the binary entropy function in (2.8). In (2.40d) the maximization is over the set defined by (2.9a)-(2.9c) and is obtained as an application of the 'Gaussian maximizes entropy' principle as follows. Given any input distribution $\mathbb{P}_{X_s,X_r,S_r}$, the covariance matrix of (X_s, X_r) conditioned on S_r can be written as

$$\mathrm{Cov}\left.\begin{bmatrix} X_s \\ X_r \end{bmatrix}\right|_{S_r=\ell} = \begin{bmatrix} P_{s,\ell} & \alpha_\ell \sqrt{P_{s,\ell}P_{r,\ell}} \\ \alpha_\ell^* \sqrt{P_{s,\ell}P_{r,\ell}} & P_{r,\ell} \end{bmatrix},$$

with $|\alpha_\ell| \leq 1$ for some $(P_{s,0}, P_{s,1}, P_{r,0}, P_{r,1}) \in \mathbb{R}_+^4$ satisfying the average power constraint in (2.9c). Then, a zero-mean jointly Gaussian input with the above covariance matrix maximizes the different mutual information terms in (2.40c). In particular, we obtain

$$\begin{aligned}
&I(X_s, X_r; Y_d|S_r = 0) \leq \log(1 + SP_{s,0}) =: I_1, \\
&I(X_s, X_r; Y_d|S_r = 1) \leq I_2 \\
&\quad := \log\left(1 + SP_{s,1} + IP_{r,1} + 2|\alpha_1|\sqrt{SP_{s,1}\ IP_{r,1}}\right), \\
&I(X_s; Y_r, Y_d|X_r, S_r = 0) \leq \log\big(1 + (C + S)(1 - |\alpha_0|^2)P_{s,0}\big) \\
&\quad \leq \log(1 + (C + S)P_{s,0}) =: I_3, \\
&I(X_s; Y_r, Y_d|X_r, S_r = 1) \leq \log\big(1 + S(1 - |\alpha_1|^2)P_{s,1}\big) =: I_4,
\end{aligned}$$

as defined in (2.10)-(2.13) thereby proving the upper bound in (2.7b), which is the same as $r^{(\mathrm{CS-HD})}$ in (2.40d).

- Regarding (2.7c), the average power constraints at the source and at the relay given in (2.9c) can be expressed as follows. Since the source transmits in both phases we define, for some $\beta \in [0, 1]$, the power split $P_{s,0} = \frac{\beta}{\gamma},\ P_{s,1} = \frac{1-\beta}{1-\gamma}$. Since the relay transmission only affects the destination output for a fraction $(1-\gamma)$ of the time, i.e., when $S_r = 1$, the relay must exploit all its available power when $S_r = 1$; we thus split the relay power as $P_{r,0} = 0,\ P_{r,1} = \frac{1}{1-\gamma}$. The cut-set upper bound

$r^{(\mathrm{CS-HD})}$ in (2.40d) can be rewritten as

$$
\begin{aligned}
r^{(\mathrm{CS-HD})} = & \max_{(\gamma,|\alpha_1|,\beta)\in[0,1]^3} \min\Bigg\{\mathcal{H}(\gamma)+\gamma\log\left(1+\frac{S\beta}{\gamma}\right) \\
& +(1-\gamma)\log\left(1+\frac{I}{1-\gamma}+\frac{S(1-\beta)}{1-\gamma}+2|\alpha_1|\sqrt{\frac{I}{1-\gamma}\frac{S(1-\beta)}{1-\gamma}}\right), \\
& \gamma\log\left(1+\frac{C\beta}{\gamma}+\frac{S\beta}{\gamma}\right)+(1-\gamma)\log\left(1+(1-|\alpha_1|^2)\frac{S(1-\beta)}{1-\gamma}\right)\Bigg\} \\
\le & \max_{\gamma\in[0,1]} \min\Bigg\{\mathcal{H}(\gamma)+\gamma\log\left(1+\frac{S}{\gamma}\right) \\
& +(1-\gamma)\log\left(1+\left(\sqrt{\frac{I}{1-\gamma}}+\sqrt{\frac{S}{1-\gamma}}\right)^2\right), \\
& \gamma\log\left(1+\frac{C}{\gamma}+\frac{S}{\gamma}\right)+(1-\gamma)\log\left(1+\frac{S}{1-\gamma}\right)\Bigg\} \\
= & \max_{\gamma\in[0,1]} \min\Bigg\{2\mathcal{H}(\gamma)+\gamma\log(\gamma+S)+(1-\gamma)\log\left(1-\gamma+\left(\sqrt{I}+\sqrt{S}\right)^2\right), \\
& \mathcal{H}(\gamma)+\gamma\log(\gamma+C+S)+(1-\gamma)\log(1-\gamma+S)\Bigg\} \\
\le & \, 2+\max_{\gamma\in[0,1]} \min\Bigg\{\gamma\log(1+S)+(1-\gamma)\log\left(1+\left(\sqrt{I}+\sqrt{S}\right)^2\right), \\
& \gamma\log(1+C+S)+(1-\gamma)\log(1+S)\Bigg\} \\
= & \, 2+\log(1+S)\max_{\gamma\in[0,1]} \min\{\gamma+(1-\gamma)b_1, \gamma b_2+(1-\gamma)\} \\
= & \, 2+\log(1+S)\left(1+\max_{\gamma\in[0,1]} \min\{(1-\gamma)(b_1-1), \gamma(b_2-1)\}\right) \\
= & \, 2+\log(1+S)\left(1+\frac{(b_1-1)(b_2-1)}{(b_1-1)+(b_2-1)}\right),
\end{aligned} \tag{2.41}
$$

where we defined b_1 and b_2 as in (2.14)-(2.15), namely

$$
\begin{aligned}
b_1 &:= \frac{\log\left(1+(\sqrt{I}+\sqrt{S})^2\right)}{\log(1+S)} > 1 \text{ since } I>0, \\
b_2 &:= \frac{\log(1+C+S)}{\log(1+S)} > 1 \text{ since } C>0.
\end{aligned}
$$

Note that the optimal γ is found by equating the two arguments of the min and is given by $\gamma^*_{\mathrm{CS}} := \frac{(b_1-1)}{(b_1-1)+(b_2-1)}$.

2.B Proof of Proposition 2

The upper bound in (2.7c) implies

$$
\begin{aligned}
\mathsf{d}^{(\mathrm{HD-RC})} \leq & \lim_{\mathsf{SNR}\to+\infty} \frac{\log(1+S)}{\log(1+\mathsf{SNR})}\left(1+\frac{(b_1-1)(b_2-1)}{(b_1-1)+(b_2-1)}\right) \\
=& \beta_{\mathrm{sd}}\left(1+\frac{[\beta_{\mathrm{rd}}/\beta_{\mathrm{sd}}-1]^+\ [\beta_{\mathrm{sr}}/\beta_{\mathrm{sd}}-1]^+}{[\beta_{\mathrm{rd}}/\beta_{\mathrm{sd}}-1]^+ + [\beta_{\mathrm{sr}}/\beta_{\mathrm{sd}}-1]^+}\right) \\
=& \beta_{\mathrm{sd}} + \frac{[\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}}]^+\ [\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}}]^+}{[\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}}]^+ + [\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}}]^+},
\end{aligned}
$$

since $b_1 \to \max\{\beta_{\mathrm{sd}}, \beta_{\mathrm{rd}}\}/\beta_{\mathrm{sd}}$ and $b_2 \to \max\{\beta_{\mathrm{sd}}, \beta_{\mathrm{sr}}\}/\beta_{\mathrm{sd}}$ at high SNR, which is equivalent to the RHS of (2.4).

2.C Proof of Proposition 3

The PDF scheme in [87, Theorem 16.3] adapted to the HD model gives the following rate lower bound

$$
\begin{aligned}
C^{(\mathrm{HD-RC})} \geq & \max_{\mathbb{P}_{U,X_s,X_r,S_r}} \min\Big\{I(S_r;Y_d) + I(X_s,X_r;Y_d|S_r), \\
& \qquad I(U;Y_r|X_r,S_r) + I(X_s;Y_d|U,X_r,S_r)\Big\} \\
\geq & \max\min\Big\{I_0^{(\mathrm{PDF})} + \gamma I_5 + (1-\gamma)I_6, \gamma I_7 + (1-\gamma) I_8\Big\} \\
=& r^{(\mathrm{PDF-HD})} \text{ in (2.17)},
\end{aligned}
$$

where for the last inequality we let $\gamma := \mathbb{P}[S_r = 0] \in [0,1]$ be the fraction of time the relay listens and, conditioned on $S_r = \ell$, $\ell \in \{0,1\}$, we consider the following jointly Gaussian inputs

$$
\left.\begin{pmatrix} U \\ \frac{X_s}{\sqrt{P_{s,\ell}}} \\ \frac{X_r}{\sqrt{P_{r,\ell}}} \end{pmatrix}\right|_{S_r=\ell} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho_{s|\ell} & \rho_{r|\ell} \\ \rho_{s|\ell}^* & 1 & \alpha_\ell \\ \rho_{r|\ell}^* & \alpha_\ell^* & 1 \end{bmatrix}\right) : \begin{bmatrix} 1 & \rho_{s|\ell} & \rho_{r|\ell} \\ \rho_{s|\ell}^* & 1 & \alpha_\ell \\ \rho_{r|\ell}^* & \alpha_\ell^* & 1 \end{bmatrix} \succeq \mathbf{0}.
$$

In particular, we use specific values for the parameters $\{\rho_{s|\ell}, \rho_{r|\ell}, \alpha_\ell\}_{\ell\in\{0,1\}}$, namely

$$\angle\alpha_1 + \theta = 0, \tag{2.42a}$$

$$
\begin{aligned}
&\alpha_0 = 0 \text{ and either } |\rho_{s|0}|^2 = 1 - |\rho_{r|0}|^2 = 0 \\
&\qquad \text{or } |\rho_{r|0}|^2 = 1 - |\rho_{s|0}|^2 = 0,
\end{aligned} \tag{2.42b}
$$

$$\rho_{s|1} = \alpha_1^*,\ \rho_{r|1} = 1. \tag{2.42c}$$

With these definitions, the mutual information terms $I_0^{(\mathrm{PDF})}, I_5, \ldots, I_8$ in (2.17) are

$$I(X_s, X_r; Y_d | S_r = 0) = \log\left(1 + SP_{s,0}\right) =: I_5,$$
$$I(X_s, X_r; Y_d | S_r = 1) = \log\left(1 + SP_{s,1} + IP_{r,1} + 2|\alpha_1|\sqrt{SP_{s,1}\ IP_{r,1}}\right) =: I_6,$$

(note $I_5 = I_1$ and $I_6 = I_2$ because of the assumption in (2.42a)); next, by using the assumption in (2.42b), that is, in state $S_r = 0$ the inputs X_s and X_r are independent, and that either $U = X_s$ or $U = X_r$, we have: if $U = X_s$ independent of X_r

$$\begin{aligned}
&I(U; Y_r | X_r, S_r = 0) + I(X_s; Y_d | U, X_r, S_r = 0) \\
&= I(X_s; \sqrt{C}X_s + Z_r | X_r, S_r = 0) + I(X_s; \sqrt{S}X_s + Z_d | X_s, X_r, S_r = 0) \\
&= \log\left(1 + CP_{s,0}\right),
\end{aligned}$$

and if $U = X_r$ independent of X_s

$$\begin{aligned}
&I(U; Y_r | X_r, S_r = 0) + I(X_s; Y_d | U, X_r, S_r = 0) \\
&= I(X_r; \sqrt{C}X_s + Z_r | X_r, S_r = 0) + I(X_s; \sqrt{S}X_s + Z_d | X_r, S_r = 0) \\
&= \log\left(1 + SP_{s,0}\right);
\end{aligned}$$

therefore under the assumption in (2.42b) we have

$$I(U; Y_r | X_r, S_r = 0) + I(X_s; Y_d | U, X_r, S_r = 0) = \log\left(1 + \max\{C, S\} P_{s,0}\right) =: I_7;$$

next, by using the assumption in (2.42c), that is, in state $S_r = 1$ we let $U = X_r$, we have

$$\begin{aligned}
&I(U; Y_r | X_r, S_r = 1) + I(X_s; Y_d | U, X_r, S_r = 1) \\
&= I(X_r; Z_r | X_r, S_r = 1) + I(X_s; \sqrt{S}X_s + Z_d | X_r, S_r = 1) \\
&= I(X_s; \sqrt{S}X_s + Z_d | X_r, S_r = 1) \\
&= \log\left(1 + S(1 - |\alpha_1|^2) P_{s,1}\right) =: I_8,
\end{aligned}$$

(note $I_7 \leq I_3$ and $I_8 = I_4$); finally

$$I(S_r; Y_d) = \mathbb{E}\left[\log \frac{1}{f_{Y_d}(Y_d)}\right] - \left[\gamma \log(v_0) + (1-\gamma)\log(v_1) + \log(\pi \mathrm{e})\right] =: I_0^{(\mathrm{PDF})},$$

where $f_{Y_d}(\cdot)$ is the density of the destination output Y_d, which is a mixture of (proper complex) Gaussian random variables, i.e.,

$$f_{Y_d}(t) = \frac{\gamma}{\pi v_0}\mathrm{e}^{-|t|^2/v_0} + \frac{1-\gamma}{\pi v_1}\mathrm{e}^{-|t|^2/v_1},\ t \in \mathbb{C},$$
$$v_0 := \mathrm{Var}[Y_d|S_r = 0] = 2^{I_5},\ v_1 := \mathrm{Var}[Y_d|S_r = 1] = 2^{I_6}.$$

Note that $I_0^{(\mathrm{PDF})} = I(S_r; Y_d) \leq H(S_r) = \mathcal{H}(\gamma)$. This proves the lower bound in (2.17).

Next we show how to further lower bound the rate in (2.17) to obtain the rate expression in (2.18). With the same parameterization of the powers as in Appendix 2.A, namely $P_{s,0} = \frac{\beta}{\gamma},\ P_{s,1} = \frac{1-\beta}{1-\gamma},\ P_{r,0} = 0,\ P_{r,1} = \frac{1}{1-\gamma}$, we have that

$$\begin{aligned}
r^{(\mathrm{PDF-HD})} &= \max_{\gamma\in[0,1],|\alpha|\leq 1,\beta\in[0,1]} \min\Big\{ I_0^{(\mathrm{PDF})} + \gamma\log\left(1+\frac{\beta S}{\gamma}\right) + \\
&\quad +(1-\gamma)\log\left(1+\frac{S(1-\beta)}{1-\gamma}+\frac{I}{1-\gamma}+2|\alpha|\sqrt{\frac{S(1-\beta)}{1-\gamma}\frac{I}{1-\gamma}}\right), \\
&\quad \gamma\log\left(1+\frac{1}{\gamma}\max\{C\beta, S\beta\}\right) + (1-\gamma)\log\left(1+(1-|\alpha|^2)\frac{S(1-\beta)}{1-\gamma}\right)\Big\} \\
&\geq \max_{\gamma\in[0,1],\beta\in[0,1]} \min\Big\{\gamma\log\left(1+\frac{\beta S}{\gamma}\right) + (1-\gamma)\log\left(1+\frac{S(1-\beta)}{(1-\gamma)}+\frac{I}{1-\gamma}\right), \\
&\quad \gamma\log\left(1+\frac{1}{\gamma}\max\{\beta C, \beta S\}\right) + (1-\gamma)\log\left(1+\frac{S(1-\beta)}{(1-\gamma)}\right)\Big\} \\
&\geq \max_{\gamma\in[0,1]} \min\{\gamma\log(1+S) + (1-\gamma)\log(1+S+I), \\
&\qquad \gamma\log(1+\max\{C,S\}) + (1-\gamma)\log(1+S)\} \\
&= \log(1+S)\max_{\gamma\in[0,1]}\min\{\gamma+(1-\gamma)c_1, \gamma c_2 + (1-\gamma)\} \\
&= \log(1+S)\left(1+\max_{\gamma\in[0,1]}\min\{(1-\gamma)(c_1-1), \gamma(c_2-1)\}\right) \\
&= \log(1+S)\left(1+\frac{(c_1-1)(c_2-1)}{(c_1-1)+(c_2-1)}\right),
\end{aligned} \tag{2.43}$$

where we defined c_1 and c_2 as in (2.24)-(2.25), namely

$$c_1 := \frac{\log(1+I+S)}{\log(1+S)} > 1 \text{ since } I > 0,$$
$$c_2 := \frac{\log(1+\max\{C,S\})}{\log(1+S)} > 1 \text{ since } C > 0.$$

Notice that $c_i \le b_i, i = 1, 2$, where $b_i, i = 1, 2$, are defined in (2.14)-(2.15). The optimal γ, indicated by $\gamma^*_{\rm PDF}$ is given by

$$\gamma^*_{\rm PDF} := \frac{(c_1 - 1)}{(c_1 - 1) + (c_2 - 1)} \ \in \ [0, 1].$$

Remark 4. A further lower bound on the PDF rate $r^{(\rm PDF-HD)}$ in (2.17) can be obtained by trivially lower bounding $I_0^{(\rm PDF)} \ge 0$, which corresponds to a fixed transmit/receive schedule for the relay.

2.D Proof of Proposition 4

The lower bound in (2.18) implies

$$\begin{aligned}
\mathsf{d}^{(\rm HD-RC)} \ge & \lim_{\mathsf{SNR}\to+\infty} \frac{\log(1+S)}{\log(1+\mathsf{SNR})} \left(1 + \frac{(c_1-1)(c_2-1)}{(c_1-1)+(c_2-1)}\right) \\
=& \beta_{\rm sd} \left(1 + \frac{[\beta_{\rm rd}/\beta_{\rm sd} - 1]^+ \ [\beta_{\rm sr}/\beta_{\rm sd} - 1]^+}{[\beta_{\rm rd}/\beta_{\rm sd} - 1]^+ + [\beta_{\rm sr}/\beta_{\rm sd} - 1]^+}\right) \\
=& \beta_{\rm sd} + \frac{[\beta_{\rm rd} - \beta_{\rm sd}]^+ \ [\beta_{\rm sr} - \beta_{\rm sd}]^+}{[\beta_{\rm rd} - \beta_{\rm sd}]^+ + [\beta_{\rm sr} - \beta_{\rm sd}]^+},
\end{aligned}$$

since $c_1 \to \max\{\beta_{\rm sd}, \beta_{\rm rd}\}/\beta_{\rm sd}$ and $c_2 \to \max\{\beta_{\rm sd}, \beta_{\rm sr}\}/\beta_{\rm sd}$ at high SNR, which is equivalent to the RHS of (2.4).

2.E Proof of Proposition 6

The rate in (2.39) can be further lower bounded as

$$r^{(\rm LDAi-HD)} \ge -1 + \log(1+S) \left(1 + \frac{(c_3-1)(c_4-1)}{(c_3-1)+(c_4-1)}\right),$$

where $c_3 := c_1 = \frac{\log(1+I+S)}{\log(1+S)}$ and $c_4 := b_2 = \frac{\log(1+C+S)}{\log(1+S)}$. The rate above implies

$$\begin{aligned}
\mathsf{d} \ge & \lim_{\mathsf{SNR}\to+\infty} \frac{\log(1+S)}{\log(1+\mathsf{SNR})} \left(1 + \frac{(c_3-1)(c_4-1)}{(c_3-1)+(c_4-1)}\right) \\
=& \ \beta_{\rm sd} \left(1 + \frac{[\beta_{\rm rd}/\beta_{\rm sd} - 1]^+ \ [\beta_{\rm sr}/\beta_{\rm sd} - 1]^+}{[\beta_{\rm rd}/\beta_{\rm sd} - 1]^+ + [\beta_{\rm sr}/\beta_{\rm sd} - 1]^+}\right) \\
=& \ \beta_{\rm sd} + \frac{[\beta_{\rm rd} - \beta_{\rm sd}]^+ \ [\beta_{\rm sr} - \beta_{\rm sd}]^+}{[\beta_{\rm rd} - \beta_{\rm sd}]^+ + [\beta_{\rm sr} - \beta_{\rm sd}]^+},
\end{aligned}$$

since $c_3 \to \max\{\beta_{\rm sd}, \beta_{\rm rd}\}/\beta_{\rm sd}$ and $c_4 \to \max\{\beta_{\rm sd}, \beta_{\rm sr}\}/\beta_{\rm sd}$ at high SNR, which is equivalent to the RHS of (2.4).

2.F Proof of Proposition 7

Consider the upper bound in (2.7a) and the lower bound in (2.16). Since the term $I(X_s, X_r, S_r; Y_d)$ is the same in the upper and lower bounds, the gap is given by

$$\mathsf{GAP} \le I(X_s; Y_r, Y_d|X_r, S_r) - I(U; Y_r|X_r, S_r) - I(X_s; Y_d|X_r, S_r, U).$$

Next we consider two different choices for U:

- For $C \le S$ we choose $U = X_r$ and

$$\begin{aligned}
\mathsf{GAP} &\le I(X_s; Y_r, Y_d|X_r, S_r) - I(X_s; Y_d|X_r, S_r) \\
&= I(X_s; Y_r|X_r, S_r, Y_d) \\
&= \mathbb{P}[S_r = 0] I(X_s; \sqrt{C}X_s + Z_r|X_r, S_r = 0, \sqrt{S}X_s + Z_d) \\
&\quad + \mathbb{P}[S_r = 1] I(X_s; Z_r|X_r, S_r = 1, \ \sqrt{S}X_s + Z_d) \\
&= \mathbb{P}[S_r = 0] \log\left(1 + \frac{CP_{s,0}}{1 + SP_{s,0}}\right) + \mathbb{P}[S_r = 1] \cdot 0 \\
&\le 1 \cdot \log\left(1 + \frac{SP_{s,0}}{1 + SP_{s,0}}\right) \\
&\le 1 \text{ bit.}
\end{aligned}$$

- For $C > S$ we choose $U = X_r S_r + X_s(1 - S_r)$ and

$$\begin{aligned}
\mathsf{GAP} &\le I(X_s; Y_r, Y_d|X_r, S_r) - I(X_r S_r + X_s(1 - S_r); Y_r|X_r, S_r) \\
&\quad - I(X_s; Y_d|X_r, S_r, \ X_r S_r + X_s(1 - S_r)) \\
&= \mathbb{P}[S_r = 0]\Big(I(X_s; Y_r, Y_d|X_r, S_r = 0) - I(X_s; Y_r|X_r, S_r = 0)\Big) \\
&\quad + \mathbb{P}[S_r = 1]\Big(I(X_s; Y_r, Y_d|X_r, S_r = 1) - I(X_s; Y_d|X_r, S_r = 1)\Big) \\
&= \mathbb{P}[S_r = 0] \ I(X_s; Y_d|X_r, S_r = 0, Y_r) + \mathbb{P}[S_r = 1] \ I(X_s; Y_r|X_r, S_r = 1, Y_d) \\
&= \mathbb{P}[S_r = 0] I(X_s; \sqrt{S}X_s + Z_d|X_r, S_r = 0, \sqrt{C}X_s + Z_r) \\
&\quad + \mathbb{P}[S_r = 1] \ I(X_s; Z_r|X_r, S_r = 1, \ \sqrt{S}X_s + Z_d) \\
&= \mathbb{P}[S_r = 0] \ \log\left(1 + \frac{SP_{s,0}}{1 + CP_{s,0}}\right) + \mathbb{P}[S_r = 1] \cdot 0 \\
&\le 1 \cdot \log\left(1 + \frac{CP_{s,0}}{1 + CP_{s,0}}\right) \\
&\le 1 \text{ bit.}
\end{aligned}$$

2.G Proof of Proposition 8

Consider the upper bound in (2.7b) and the lower bound in (2.17). Recall that $I_1 = I_5$ $I_2 = I_6$ $I_3 \geq I_7$ $I_4 = I_8$ and therefore

$$\begin{aligned}
\mathsf{GAP} &\leq \max\Big\{\mathcal{H}(\gamma) + \gamma I_1 + (1-\gamma)I_2 - \gamma I_5 - (1-\gamma)I_6, \\
&\qquad \gamma I_3 + (1-\gamma)I_4 - \gamma I_7 - (1-\gamma)I_8\Big\} \\
&\leq \max\left\{1, \log\left(\frac{1 + CP_{s,0} + SP_{s,0}}{1 + \max\{C,S\}P_{s,0}}\right)\right\} \\
&\leq \max\left\{1, \log\left(\frac{1 + 2\max\{C,S\}P_{s,0}}{1 + \max\{C,S\}P_{s,0}}\right)\right\} \leq 1 \text{ bit.}
\end{aligned}$$

2.H Proof of Proposition 9

Consider the upper bound in (2.7c) and the lower bound in (2.39). We distinguish two cases:

- Case 1: $S > C$. In this case $r^{(\mathrm{LDAi-HD})} = \log(1+S)$. The gap is

$$\begin{aligned}
\mathsf{GAP} &\leq r^{(\mathrm{CS-HD})} - r^{(\mathrm{LDAi-HD})} \\
&\leq 2 + \log(1+S)\,\frac{(b_1-1)(b_2-1)}{(b_1-1)+(b_2-1)} \\
&\leq 2 + \log(1+S)\,(b_2-1) \\
&= 2 + \log\left(1 + \frac{C}{1+S}\right) \leq 3 \text{ bits.}
\end{aligned}$$

- Case 2: $S \leq C$. First, by noticing that $\log\left(1 + \left(\sqrt{I} + \sqrt{S}\right)^2\right) \leq \log(1+I+S) + 1$, we further upper bound the expression in (2.7c) as

$$r^{(\mathrm{CS-HD})} \leq 2 + \log(1+S) + \frac{\left(\log\left(1+\frac{I}{1+S}\right)+1\right)\log\left(1+\frac{C}{1+S}\right)}{\log\left(1+\frac{I}{1+S}\right) + 1 + \log\left(1+\frac{C}{1+S}\right)}.$$

  Next we further lower bound $r^{(\mathrm{LDAi-HD})}$ in (2.39) as

$$r^{(\mathrm{LDAi-HD})} \geq \log(1+S) + \frac{\log\left(1+\frac{I}{1+S}\right)\left(\log\left(1+\frac{C}{1+S}\right)-1\right)}{\log\left(1+\frac{I}{1+S}\right) + \log\left(1+\frac{C}{1+S}\right)}.$$

Hence, with $x = \log\left(1 + \frac{I}{1+S}\right), y = \log\left(1 + \frac{C}{1+S}\right)$, we have

$$
\begin{aligned}
\mathsf{GAP} &\leq r^{(\mathrm{CS-HD})} - r^{(\mathrm{LDAi-HD})} \\
&\leq 2 + \frac{(x+1)y}{x+1+y} - \frac{x(y-1)}{x+y} \leq 3 \text{ bits.}
\end{aligned}
$$

2.I Achievable rate with CF

Proposition 12. *The capacity of the Gaussian HD relay channel is lower bounded as*

$$
C^{(\mathrm{HD-RC})} \geq r^{(\mathrm{CF-HD})} := \max\min\left\{I_0^{(\mathrm{CF})} + \sum_{(i,j)\in[0:1]^2} \gamma_{ij} I_{9,ij}, \sum_{(i,j)\in[0:1]^2} \gamma_{ij} I_{10,ij}\right\}, \tag{2.44a}
$$

where the maximization is over

$$
\gamma_{ij} \in [0,1] : \sum_{(i,j)\in[0:1]^2} \gamma_{ij} = 1, \tag{2.44b}
$$

$$
P_{s,i} \geq 0 : \sum_{(i,j)\in[0:1]^2} \gamma_{ij}\ P_{s,i} \leq 1, \tag{2.44c}
$$

$$
P_{r,ij} \geq 0 : \sum_{(i,j)\in[0:1]^2} \gamma_{ij}\ P_{r,ij} \leq 1, \tag{2.44d}
$$

and where the different mutual information terms in (2.44) *are defined next.*

Proof. The CF scheme in [87, Theorem 16.4] adapted to the HD model gives

$$
\begin{aligned}
C^{(\mathrm{HD-RC})} &\geq \max_{\mathbb{P}_Q \mathbb{P}_{X_s|Q} \mathbb{P}_{[X_r,S_r]|Q} \mathbb{P}_{\widehat{Y}_r|[X_r,S_r],Y_r,Q} : |Q|\leq 2} \min\Big\{ I(X_s; \widehat{Y}_r, Y_d | [X_r, S_r], Q), \\
&\qquad I(X_s, [X_r, S_r]; Y_d | Q) - I(Y_r; \widehat{Y}_r | X_s, [X_r, S_r], Y_d, Q) \Big\} \\
&= \max_{\mathbb{P}_Q \mathbb{P}_{S_r|Q} \mathbb{P}_{X_s|Q} \mathbb{P}_{X_r|S_r,Q} \mathbb{P}_{\widehat{Y}_r|X_r,Y_r,S_r,Q} : |Q|\leq 2} \min\Big\{ I(X_s; \widehat{Y}_r, Y_d | Q, S_r, X_r), \\
&\qquad I(S_r; Y_d | Q) + I(X_s, X_r; Y_d | S_r, Q) - I(Y_r; \widehat{Y}_r | X_s, X_r, Y_d, S_r, Q) \Big\} \\
&\geq r^{(\mathrm{CF-HD})} \text{ in (2.44a)},
\end{aligned} \tag{2.45}
$$

where the mutual information terms $\{I_{9,ij}, I_{10,ij}\}$, $(i,j) \in [0:1]^2$ and $I_0^{(\mathrm{CF})}$ in (2.44a) are obtained as follows. We consider the following assignment on the inputs and on the auxiliary random variables for each $(i,j) \in [0:1]^2$

$$\mathbb{P}[Q=i, S_r=j] = \gamma_{ij} \text{ such that (2.44b) is satisfied,}$$
$$\begin{pmatrix} X_s \\ X_r \end{pmatrix}\Bigg|_{Q=i,S_r=j} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} P_{s,i} & 0 \\ 0 & P_{r,ij} \end{bmatrix}\right)$$
$$\text{such that (2.44c) and (2.44d) are satisfied,}$$
$$\widehat{Y}_r|_{X_r,Y_r,Q=i,S_r=j} = Y_r + \widehat{Z}_{r,ij},$$
$$\widehat{Z}_{r,ij} \sim \mathcal{N}(0, \sigma_{ij}^2) \text{ and independent of everything else,}$$

and in order to meet the constraint that X_s can not depend on S_r conditioned on Q we must impose the constraint that in state $Q=i, S_r=j$ the power of the source only depends on the index i. Then for each $(i,j) \in [0:1]^2$

$$I(X_s; \widehat{Y}_r, Y_d | X_r, Q=i, S_r=j) =$$
$$= \log\left(1 + \left(S + \frac{C(1-j)}{1+\sigma_{ij}^2}\right) P_{s,i}\right) =: I_{10,ij}, \tag{2.46}$$
$$I(X_s, X_r; Y_d | Q=i, S_r=j) - I(Y_r; \widehat{Y}_r | X_s, X_r, Y_d, Q=i, S_r=j)$$
$$= \log\left(1 + SP_{s,i} + IjP_{r,ij}\right) - \log\left(1 + \frac{1}{\sigma_{ij}^2}\right) =: I_{9,ij}, \tag{2.47}$$
$$I(S_r; Y_d | Q) = I_0^{(\mathrm{CF})} =: -\sum_{(i,j)} \gamma_{ij} \log(v_{ij}) - \log(\pi \mathrm{e})$$
$$+ (\gamma_{00} + \gamma_{01})\, \mathbb{E}\left[\log \frac{1}{f_0(Y)} \middle| Q=0\right] + (\gamma_{10} + \gamma_{11})\, \mathbb{E}\left[\log \frac{1}{f_1(Y)} \middle| Q=1\right],$$

where

$$Y_d|_{Q=0} \sim f_0(t) := \frac{\gamma_{00}}{\gamma_{00}+\gamma_{01}} \frac{1}{\pi v_{00}} \mathrm{e}^{-|t|^2/v_{00}} + \frac{\gamma_{01}}{\gamma_{00}+\gamma_{01}} \frac{1}{\pi v_{01}} \mathrm{e}^{-|t|^2/v_{01}},\ t \in \mathbb{C},$$
$$Y_d|_{Q=1} \sim f_1(t) := \frac{\gamma_{10}}{\gamma_{10}+\gamma_{11}} \frac{1}{\pi v_{10}} \mathrm{e}^{-|t|^2/v_{10}} + \frac{\gamma_{11}}{\gamma_{10}+\gamma_{11}} \frac{1}{\pi v_{11}} \mathrm{e}^{-|t|^2/v_{11}},\ t \in \mathbb{C},$$
$$v_{ij} := \mathrm{Var}[Y_d | Q=i, S_r=j] = 1 + S\ P_{s,i} + I\ j\ P_{r,ij}.$$

This proves the lower bound in (2.44) as a function of $\sigma_{ij}^2, (i,j) \in \{0,1\}^2$.

In order to find the optimal $\sigma_{ij}^2, (i,j) \in \{0,1\}^2$ we reason as follows. $I_{10,ij}$ in (2.46) is decreasing in σ_{ij}^2 while $I_{9,ij}$ in (2.47) is increasing. At the optimal point these two rates are the same. Let

$$C_i := 1 + \frac{CP_{s,i}}{1+SP_{s,i}},\ x_i := \frac{1}{\sigma_{i0}^2},\ I' := I(S_r, X_r; Y_d | Q),$$

and rewrite the lower bound in (2.44) as

$$\begin{aligned} r^{(\mathrm{CF-HD})} &= (\gamma_{00}+\gamma_{01})\log(1+SP_{s,0}) + (\gamma_{10}+\gamma_{11})\log(1+SP_{s,1}) \\ &- \gamma_{00}\log\left(1+x_0\right) - \gamma_{10}\log\left(1+x_1\right) \\ &+ \min\Big\{\gamma_{00}\log\left(1+x_0C_0\right) + \gamma_{10}\log\left(1+x_1C_1\right), I'\Big\}. \end{aligned}$$

The solution of

$$\begin{aligned} &\min_{(x_0,x_1)\in\mathbb{R}_+^2} \Big\{\gamma_{00}\log\left(1+x_0\right) + \gamma_{10}\log\left(1+x_1\right)\Big\} \\ &\text{subject to} \quad \gamma_{00}\log\left(1+x_0C_0\right) + \gamma_{10}\log\left(1+x_1C_1\right) = I' \end{aligned}$$

can be found to be $x_i = \frac{[\eta C_i - 1]^+}{(1-\eta)C_i}$, $i \in [1:2]$, with $\eta \le 1$ such that $\gamma_{00}\log\left(1+x_0C_0\right) + \gamma_{10}\log\left(1+x_1C_1\right) = I'$. □

Remark 5. For the special case of $Q = S_r$, that is, $I_0^{(\mathrm{CF})} = I(S_r;Y_d|Q) = I(Q;Y_d|Q) = 0$, the achievable rate in Proposition 12 reduces to

$$r^{(\mathrm{CF-HD})} \ge \max_{(\gamma,\beta)\in[0,1]^2} \min\Big\{\gamma I_9 + (1-\gamma)I_{10}, \gamma I_{11} + (1-\gamma)I_{12}\Big\}, \tag{2.48a}$$

$$I_9 := \log\left(1+SP_{s,0}\right) - \log\left(1+\frac{1}{\sigma_0^2}\right), \tag{2.48b}$$

$$I_{10} := \log\left(1+SP_{s,1}+IP_{r,1}\right) - \log\left(1+\frac{1}{\sigma_1^2}\right), \tag{2.48c}$$

$$I_{11} := \log\left(1+SP_{s,0}+\frac{C}{1+\sigma_0^2}P_{s,0}\right), \tag{2.48d}$$

$$I_{12} := \log\left(1+SP_{s,1}\right), \tag{2.48e}$$

$$\sigma_0^2 := \frac{B+1}{(1+A)^{\frac{1}{\gamma}-1}-1}, \quad \sigma_1^2 = +\infty, \tag{2.48f}$$

$$A := \frac{IP_{r,1}}{1+SP_{s,1}}, \quad B := \frac{CP_{s,0}}{1+SP_{s,0}}, \tag{2.48g}$$

$$P_{s,0} = \frac{\beta}{\gamma}, \quad P_{s,1} = \frac{1-\beta}{1-\gamma}, \quad P_{r,1} = \frac{1}{1-\gamma}, \tag{2.48h}$$

where the optimal value for σ_0^2 in (2.48f) is obtained by equating the two expressions within the min in (2.48a).

Proposition 13. *CF with deterministic switch achieves the gDoF upper bound in* (2.4).

Proof. With the achievable rate in Remark 5 (where here we explicitly write the optimization with respect to σ_0^2) we have that

$$\begin{aligned}
r^{(\mathrm{CF-HD})} &\geq \max_{\gamma\in[0,1],\sigma_0^2\geq 0,\beta\in[0,1]} \min\Big\{\gamma\log\left(1+\frac{\beta S}{\gamma}\right)-\gamma\log\left(1+\frac{1}{\sigma_0^2}\right)+ \\
&\quad +(1-\gamma)\log\left(1+\frac{(1-\beta)S}{1-\gamma}+\frac{I}{1-\gamma}\right), \\
&\quad \gamma\log\left(1+\frac{C\beta}{(1+\sigma_0^2)\gamma}+\frac{S\beta}{\gamma}\right)+(1-\gamma)\log\left(1+\frac{(1-\beta)S}{1-\gamma}\right)\Big\} \\
&\overset{\beta=\gamma}{\geq} \max_{\gamma\in[0,1],\sigma_0^2\geq 0} \min\left\{\gamma\log(1+S)+(1-\gamma)\log(1+S+I),\right. \\
&\quad \left.\gamma\log\left(1+\frac{C}{1+\sigma_0^2}+S\right)+(1-\gamma)\log(1+S)\right\}-\gamma\log\left(1+\frac{1}{\sigma_0^2}\right) \\
&= \max_{\gamma\in[0,1],\sigma_0^2\geq 0}\left[\log(1+S)\min\{\gamma+(1-\gamma)c_5,\gamma c_6+(1-\gamma)\}-\gamma\log\left(1+\frac{1}{\sigma_0^2}\right)\right] \\
&\overset{\gamma=\gamma^*_{\mathrm{CF}}}{\geq} \max_{\sigma_0^2\geq 0}\log(1+S)\left(1+\frac{(c_5-1)(c_6-1)}{(c_5-1)+(c_6-1)}\left(1-\frac{\log\left(1+\frac{1}{\sigma_0^2}\right)}{\log\left(1+\frac{C}{(1+\sigma_0^2)(1+S)}\right)}\right)\right) \\
&\overset{\sigma_0^2=1}{\geq} -1+\log(1+S)\left(1+\frac{(c_5-1)(c_6-1)}{(c_5-1)+(c_6-1)}\right), \qquad (2.49)
\end{aligned}$$

where we defined c_5 and c_6 as

$$\begin{aligned}
c_5 = c_1 &:= \frac{\log(1+I+S)}{\log(1+S)} > 1 \text{ since } I>0 \text{ and as in (2.24)}, \\
c_6 &:= \frac{\log\left(1+\frac{C}{1+\sigma_0^2}+S\right)}{\log(1+S)} > 1 \text{ since } C>0,
\end{aligned}$$

and where

$$\gamma^*_{\mathrm{CF}} := \frac{(c_5-1)}{(c_5-1)+(c_6-1)} \in [0,1].$$

By reasoning as for the PDF in Appendix 2.D, it follows from the last rate bound that CF also achieves the gDoF in (2.4). □

Remark 6. For the special case of $Q=\emptyset$, i.e., the time-sharing variable Q

is a constant, the achievable rate in Proposition 12 reduces to

$$
\begin{aligned}
r^{(\mathrm{CF-HD})} \geq & \max_{\mathbb{P}_{X_s}\mathbb{P}_{X_r,S_r}\mathbb{P}_{\widehat{Y}_r|[X_r,S_r],Y_r}} \min\left\{I\left(X_s;\widehat{Y}_r,Y_d|S_r,X_r\right),\right.\\
& \left. I\left(X_s,X_r,S_r;Y_d\right)-I\left(Y_r;\widehat{Y}_r|S_r,X_r,X_s,Y_d\right)\right\}\\
\geq & \max_{\gamma\in[0,1],\sigma^2} \min\left\{\gamma\log\left(1+S+\frac{C}{1+\sigma^2}\right)+(1-\gamma)\log\left(1+S\right),\right.\\
& I\left(S_r;Y_d\right)+\gamma\log\left(1+S\right)-\gamma\log\left(1+\frac{1}{\sigma^2}\right)\\
& \left.+(1-\gamma)\log\left(1+S+\frac{I}{1-\gamma}\right)\right\}.
\end{aligned}
$$

With $Q=\emptyset$ the source always transmits with constant power, regardless of the state of the relay, while the relay sends only when in transmitting mode. Thus with $Q=\emptyset$ there is no coordination between the source and the relay.

2.J Proof of Proposition 10

With CF we have that

$$
\begin{aligned}
\mathsf{GAP} \leq & \max\Big\{\mathcal{H}(\gamma)+\gamma I_1+(1-\gamma)I_2-\gamma I_9-(1-\gamma)I_{10},\\
& \gamma I_3+(1-\gamma)I_4-\gamma I_{11}-(1-\gamma)I_{12}\Big\}\\
\leq & \max\Big\{\mathcal{H}(\gamma)+\gamma\log\left(1+SP_{s,0}\right)+\gamma\log\left(1+\frac{1}{\sigma_0^2}\right)-\gamma\log\left(1+SP_{s,0}\right)\\
& +(1-\gamma)\log\left(1+(\sqrt{SP_{s,1}}+\sqrt{IP_{r,1}})^2\right)-(1-\gamma)\log\left(1+SP_{s,1}+IP_{r,1}\right),\\
& \gamma\log\left(1+(C+S)P_{s,0}\right)+(1-\gamma)\log\left(1+SP_{s,1}\right)+\\
& -\gamma\log\left(1+SP_{s,0}+\frac{CP_{s,0}}{1+\sigma_0^2}\right)-(1-\gamma)\log\left(1+SP_{s,1}\right)\Big\}\\
\leq & \max\Big\{\mathcal{H}(\gamma)+(1-\gamma)+\gamma\log\left(1+\frac{1}{\sigma_0^2}\right),\\
& \gamma\log\left(1+\frac{\frac{\sigma_0^2}{1+\sigma_0^2}CP_{s,0}}{1+SP_{s,0}+\frac{1}{1+\sigma_0^2}CP_{s,0}}\right)\Big\}\\
\leq & \max\left\{\mathcal{H}(\gamma)+(1-\gamma)+\gamma\log\left(1+\frac{1}{\sigma_0^2}\right),\gamma\log\left(1+\sigma_0^2\right)\right\}\\
\leq & \ 1.6081 \text{ bits},
\end{aligned}
$$

where for σ_0^2 we chose the value $\sigma_0^2 = 2^{\frac{\mathcal{H}(\gamma)+(1-\gamma)}{\gamma}}$ by equating the two arguments of the max (this is so because $\mathcal{H}(\gamma) + (1-\gamma) + \gamma \log\left(1 + \frac{1}{\sigma_0^2}\right)$ is decreasing in σ_0^2, while $\log\left(1+\sigma_0^2\right)$ is increasing in σ_0^2). Numerically one can find that with the chosen σ_0^2 the maximum over $\gamma \in [0,1]$ is 1.6081 for $\gamma = 0.3855$. Note that by choosing $\sigma_0^2 = 1$ the gap would be upper bounded by 2 bits.

Chapter 3

The Half-Duplex Multi-Relay Network

In this chapter, we study HD relay networks where the communication between a source and a destination is assisted by N HD relays. Our main contributions can be summarized as follows: (i) we show that, for the Gaussian noise case, the cut-set outer bound is achievable to within a constant gap by NNC; (ii) we prove that, for any memoryless HD N-relay network with independent noises and for which the cut-set outer bound is achievable to within a constant gap under certain assumptions, the (approximately) optimal schedule has at most $N+1$ states, out of the 2^N possible ones, with a strictly positive probability; (iii) we show that the gDoF of the Gaussian network is the solution of a LP, where the coefficients of the linear inequality constraints are the solution of several LPs referred to as the MWBM problem; this result also allows to characterize the gDoF of broadcast networks with relays and to solve user scheduling problems; (iv) we apply the results to networks with multi-antenna nodes, where the antennas at the relays can be switched between listen and transmit state independently of one another.

3.1 System model

The general multi-relay network, defined in Section 2.1.1, consists of N HD relay nodes (numbered 1 through N) assisting the communication between a source (node 0) and a destination (node $N+1$), through a shared memoryless channel. The input-output relationship of a multi-antenna complex-valued power-constrained Gaussian HD relay network generalizes (2.1) as follows [1]

$$\mathbf{y} = \mathbf{H}_{\mathrm{eq}}\mathbf{x} + \mathbf{z} \in \mathbb{C}^{(m_{\mathrm{tot}}+m_{N+1})\times 1}, \tag{3.1a}$$

$$\mathbf{H}_{\mathrm{eq}} := \begin{bmatrix} \mathbf{I}_{m_{\mathrm{tot}}} - \mathbf{S} & \mathbf{0}_{m_{\mathrm{tot}}\times m_{N+1}} \\ \mathbf{0}_{m_{N+1}\times m_{\mathrm{tot}}} & \mathbf{I}_{m_{N+1}} \end{bmatrix} \mathbf{H} \begin{bmatrix} \mathbf{S} & \mathbf{0}_{m_{\mathrm{tot}}\times m_0} \\ \mathbf{0}_{m_0\times m_{\mathrm{tot}}} & \mathbf{I}_{m_0} \end{bmatrix}, \tag{3.1b}$$

where

- m_0 is the number of antennas at the source, m_k is the number of antennas at relay $k \in [1 : N]$ with $m_{\mathrm{tot}} := \sum_{k=1}^{N} m_k$ (i.e., m_{tot} is the total number of antennas at the relays), and m_{N+1} is the number of antennas at the destination.

- $\mathbf{y} := [\mathbf{y}_1; \ldots; \mathbf{y}_N; \mathbf{y}_{N+1}] \in \mathbb{C}^{(m_{\mathrm{tot}}+m_{N+1})\times 1}$ is the vector of the received signals with $\mathbf{y}_i \in \mathbb{C}^{m_i\times 1}, i \in [1 : N+1]$ being the received signal at node i.

- $\mathbf{x} := [\mathbf{x}_1; \ldots; \mathbf{x}_N; \mathbf{x}_{N+1}] \in \mathbb{C}^{(m_{\mathrm{tot}}+m_0)\times 1}$ is the vector of the transmitted signals where $\mathbf{x}_i \in \mathbb{C}^{m_i\times 1}, i \in [0 : N]$ is the signal transmitted by node i ($\mathbf{x}_{N+1}$ is the channel input of the source).

- $\mathbf{z} := [\mathbf{z}_1; \ldots; \mathbf{z}_N; \mathbf{z}_{N+1}] \in \mathbb{C}^{(m_{\mathrm{tot}}+m_{N+1})\times 1}$ is the jointly Gaussian noise vector which is assumed to have i.i.d. $\mathcal{N}(0, 1)$ components.

- $\mathbf{S}$ is the block diagonal matrix of dimension $m_{\mathrm{tot}} \times m_{\mathrm{tot}}$ to account for the state (either transmit or receive) of the relay antennas; in particular

$$\mathbf{S} := \begin{bmatrix} \mathbf{S}_1 & \mathbf{0}_{m_1\times m_2} & \ldots & \mathbf{0}_{m_1\times m_N} \\ \mathbf{0}_{m_2\times m_1} & \mathbf{S}_2 & \ldots & \mathbf{0}_{m_2\times m_N} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{m_N\times m_1} & \mathbf{0}_{m_N\times m_2} & \ldots & \mathbf{S}_N \end{bmatrix},$$
$$\mathbf{S}_i := \mathrm{diag}[S_{i,1}, \ldots, S_{i,m_i}] \in [0 : 1]^{m_i},$$

  where $S_{i,j} = 1$ if the j-th antenna of the i-th relay is transmitting and $S_{i,j} = 0$ if it is receiving, with $j \in [1 : m_i]$, $i \in [1 : N]$. In this model the antennas of each relay can be switched independently of one another to transmit or receive mode for a total of $2^{m_{\mathrm{tot}}}$ possible states.

[1] Recall that, for notation convenience, the input at the source / node 0 is denoted as X_{N+1} rather than X_0.

- $\mathbf{H} \in \mathbb{C}^{(m_{N+1}+m_{\rm tot})\times(m_0+m_{\rm tot})}$ is the constant, hence known to all nodes, channel matrix defined as

$$\mathbf{H} := \begin{bmatrix} \mathbf{H}_{\rm r\to r} & \mathbf{H}_{\rm s\to r} \\ \mathbf{H}_{\rm r\to d} & \mathbf{H}_{\rm s\to d} \end{bmatrix}, \tag{3.2}$$

where:

  - $\mathbf{H}_{\rm r\to r} \in \mathbb{C}^{m_{\rm tot}\times m_{\rm tot}}$ is the block matrix which defines the network connections among the relays. In particular,

$$\mathbf{H}_{\rm r\to r} := \begin{bmatrix} \star & \mathbf{H}_{1,2} & \dots & \mathbf{H}_{1,N} \\ \mathbf{H}_{2,1} & \star & \dots & \mathbf{H}_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_{N,1} & \mathbf{H}_{N,2} & \dots & \star \end{bmatrix},$$

    with $\mathbf{H}_{i,j} \in \mathbb{C}^{m_i\times m_j}$, $(i,j) \in [1:N]^2$, being the channel matrix from the j-th relay to the i-th relay. Notice that the matrices on the main diagonal of $\mathbf{H}_{\rm r\to r}$ do not matter for the channel capacity since the relays operate in HD mode.

  - $\mathbf{H}_{\rm s\to r} := [\mathbf{H}_{1,N+1};\mathbf{H}_{2,N+1};\dots;\mathbf{H}_{N,N+1}] \in \mathbb{C}^{m_{\rm tot}\times m_0}$ is the matrix which contains the channel gains from the source to the relays. In particular, $\mathbf{H}_{i,N+1} \in \mathbb{C}^{m_i\times m_0}$, $i \in [1:N]$, is the channel matrix from the source to the i-th relay.

  - $\mathbf{H}_{\rm r\to d} := [\mathbf{H}_{N+1,1},\mathbf{H}_{N+1,2},\dots,\mathbf{H}_{N+1,N}] \in \mathbb{C}^{m_{N+1}\times m_{\rm tot}}$ is the matrix which contains the channel gains from the relays to the destination. In particular, $\mathbf{H}_{N+1,i} \in \mathbb{C}^{m_{N+1}\times m_i}$, $i \in [1:N]$, is the channel matrix from the i-th relay to the destination.

  - $\mathbf{H}_{\rm s\to d} \in \mathbb{C}^{m_{N+1}\times m_0}$ is the channel matrix between the source and the destination.

3.2 Background and overview of the main results

In this section we first briefly overview some general definitions and properties on submodular functions [88], LPs [89] and graph theory [69, 70, 90] that are crucial for the proof of our main results, which are outlined at the end of this section.

Definition 3 (Submodular function, Lovász extension and greedy solution for submodular polyhedra)**.** *A set-function $f : 2^N \to \mathbb{R}$ is submodular if* *and only if, for all subsets* $\mathcal{A}_1, \mathcal{A}_2 \subseteq [1:N]$*, we have* $f(\mathcal{A}_1) + f(\mathcal{A}_2) \geq f(\mathcal{A}_1 \cup \mathcal{A}_2) + f(\mathcal{A}_1 \cap \mathcal{A}_2)$ [2]*.*

Submodular functions are closed under non-negative linear combinations.

For a submodular function f *such that* $f(\emptyset) = 0$*, the Lovász extension is the function* $\widehat{f} : \mathbb{R}^N \to \mathbb{R}$ *defined as*

$$\widehat{f}(\mathbf{w}) := \max_{\mathbf{x} \in P(f)} \mathbf{w}^T \mathbf{x}, \quad \forall \mathbf{w} \in \mathbb{R}_+^N, \tag{3.3}$$

where $P(f)$ *is the submodular polyhedron defined as*

$$P(f) := \left\{ \mathbf{x} \in \mathbb{R}^N : \sum_{i \in \mathcal{A}} x_i \leq f(\mathcal{A}),\ \forall \mathcal{A} \subseteq [1:N] \right\}. \tag{3.4}$$

The optimal $\mathbf{x}$ *in* (3.3) *can be found by the greedy algorithm for submodular polyhedra and has components*

$$x_{\pi_i} = f(\{\pi_1, \ldots, \pi_i\}) - f(\{\pi_1, \ldots, \pi_{i-1}\}), \forall i \in [1:N], \tag{3.5}$$

where π *is a permutation of* $[1:N]$ *such that the weights* $\mathbf{w}$ *are ordered as* $w_{\pi_1} \geq w_{\pi_2} \geq \ldots \geq w_{\pi_N}$*, and where by definition* $\{\pi_0\} = \emptyset$*.*

The Lovász extension is a piecewise linear convex function.

Proposition 14 (Minimum of submodular functions)**.** *Let* f *be a submodular function and* $\widehat{f}$ *its Lovász extension. The minimum of the submodular function satisfies*

$$\min_{\mathcal{A} \subseteq [1:N]} f(\mathcal{A}) = \min_{\mathbf{w} \in [0:1]^N} \widehat{f}(\mathbf{w}) = \min_{\mathbf{w} \in [0,1]^N} \widehat{f}(\mathbf{w}),$$

i.e., $\widehat{f}(\mathbf{w})$ *attains its minimum at a vertex of the cube* $[0,1]^N$*.*

Definition 4 (Basic feasible solution)**.** *Consider the LP*

$$\begin{array}{ll} \text{maximize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \leq \mathbf{b} \quad \mathbf{x} \geq 0, \end{array}$$

where $\mathbf{x} \in \mathbb{R}^n$ *is the vector of unknowns,* $\mathbf{b} \in \mathbb{R}^m$ *and* $\mathbf{c} \in \mathbb{R}^n$ *are vectors of known coefficients, and* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is a known matrix of coefficients. If* $m < n$*, a solution for the LP with at most* m *non-zero values is called a basic feasible solution.*

[2] A set-function f is supermodular if and only if $-f$ is submodular, and it is modular if it is both submodular and supermodular.

Proposition 15 (Optimality of basic feasible solutions)**.** *If a LP is feasible, then an optimal solution is at a vertex of the (non-empty and convex) feasible set* $S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge 0\}$*. Moreover, if there is an optimal solution, then an optimal basic feasible solution exists as well.*

Proposition 16 (Saddle-point property)**.** *Let* $\phi(x, y)$ *be a function of two vector variables* $x \in \mathcal{X}$ *and* $y \in \mathcal{Y}$*. By the minimax inequality we have*

$$\max_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} \phi\left(x, y\right) \le \min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \phi\left(x, y\right)$$

and equality holds if the following three conditions hold: (i) $\mathcal{X}$ *and* $\mathcal{Y}$ *are both convex and one of them is compact, (ii)* $\phi\left(x, y\right)$ *is convex in* x *and concave in* y*, and (iii)* $\phi\left(x, y\right)$ *is continuous.*

Definition 5 (Bigraph, matching, assignment problem, Hungarian algorithm)**.** *A weighted bipartite graph, or bigraph, is a graph whose vertices can be separated into two sets such that each edge in the graph has exactly one endpoint in each set. A non-negative weight is associated with each edge in the bigraph. The weight matrix* $\mathbf{B}$ *is defined as follows: there is one set of* n_1 *nodes, where* n_1 *is the number of rows in* $\mathbf{B}$*, and another set of* n_2 *nodes, where* n_2 *is the number of columns in* $\mathbf{B}$*; the element* $[\mathbf{B}]_{ij}$ *is the weight of the edge between nodes* i *and* j*. A matching, or independent edge set, is a set of edges without common vertices. The MWBM problem, or assignment problem, is defined as a matching where the sum of the edge weights in the matching has the maximal value. The Hungarian algorithm is a polynomial time algorithm that efficiently solves the assignment problem.*

In the following we overview our main results on the HD multi-relay network. In particular, for simplicity of presentation, we state the results for the particular case of single antenna nodes, i.e., $m_i = 1, \ \forall i \in [0 : N+1]$; however, in the rest of the chapter we show how each of these results generalizes to the case of multi-antenna nodes. Our main results of this chapter are summarized as follows [3]:

Theorem 3. *The cut-set upper bound on the capacity of the Gaussian HD relay network with* N *relays is achievable by NNC with deterministic switch to within*

$$\mathsf{GAP} \le 1.96(N + 2) \text{ bits.} \tag{3.6}$$

[3] We refer to Definition 1 and Definition 2 in Section 2.2 for the concepts of gDoF and capacity to within a constant gap, respectively.

Theorem 4. *For any general memoryless HD relay network for which:*

1. *independent inputs are approximately (i.e., to within a constant gap) optimal in the cut-set outer bound, that is there exists a product input distribution*

$$\mathbb{P}_{X_{[1:N+1]}|S_{[1:N]}} = \prod_{i\in[1:N+1]} \mathbb{P}_{X_i|S_{[1:N]}} \tag{3.7}$$

for which we can bound the capacity $C^{(\mathrm{HD-RN})}$ *as*

$$\mathsf{C}' - \mathsf{G}_1 \leq C^{(\mathrm{HD-RN})} \leq \mathsf{C}' + \mathsf{G}_2, : \ \mathsf{C}' := \max_{\mathbb{P}_{S_{[1:N]}}} \min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(\mathit{fix})}, \tag{3.8}$$

where G_1 *and* G_2 *are non-negative constants that may depend on* N *but not on the channel transition probability and where*

$$\begin{aligned} I_{\mathcal{A}}^{(\mathit{fix})} &:= I\left(X_{N+1}, X_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{[1:N]}\right) & (3.9)\\ &= \sum_{s\in[0:1]^N} \lambda_s \ f_s(\mathcal{A}), & (3.10) \end{aligned}$$

with

$$\begin{aligned} \lambda_s &:= \mathbb{P}[S_{[1:N]} = s] \in [0,1] : \sum_{s\in[0:1]^N} \lambda_s = 1, & (3.11)\\ f_s(\mathcal{A}) &:= I\left(X_{N+1}, X_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{[1:N]} = s\right), & (3.12) \end{aligned}$$

2. *the "noises are independent", that is*

$$\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]},S_{[1:N]}} = \prod_{i\in[1:N+1]} \mathbb{P}_{Y_i|X_{[1:N+1]},S_{[1:N]}}, \tag{3.13}$$

3. *the functions in* (3.12) *are not a function of* $\{\lambda_s, s \in [0:1]^N\}$*, i.e., they can depend on the state* s *but not on the* $\{\lambda_s, s \in [0:1]^N\}$,

then simple relay policies are (approximately) optimal in (3.8)*, i.e., the (approximately) optimal probability mass function* $\mathbb{P}_{S_{[1:N]}}$ *has at most* $N+1$ *non-zero entries / active states.*

Theorem 5. *The gDoF* $\mathsf{d}^{(\mathrm{HD-RN})}$ *of the Gaussian HD multi-relay network is the solution of the following LP*

$$\text{maximize}\{\mathbf{f}^T\ \mathbf{x}\} \tag{3.14}$$

$$\text{subject to} \begin{bmatrix} -\mathbf{A} & \mathbf{1}_{2^N} \\ \mathbf{1}_{2^N}^T & 0 \end{bmatrix} \mathbf{x} \leq \mathbf{f}, \quad \mathbf{x} \geq 0, \tag{3.15}$$

where $\mathbf{x}^T := [\lambda_{\mathrm{vect}}, \mathsf{d}^{(\mathrm{HD-RN})}]$ *with* $\lambda_{\mathrm{vect}} := [\lambda_s] \in \mathbb{R}_+^{1\times 2^N}$, $\mathbf{f}^T := [\mathbf{0}_{2^N}^T, 1]$ *and where the entries of the non-negative matrix* $\mathbf{A} \in \mathbb{R}^{2^N \times 2^N}$ *can be found by solving* $2^{N-1}\left(2^N + 1\right)$ *independent assignment problems.*

Note that, in the theorems above and in the rest of this chapter as well, we use interchangeably the notation $s \in [0:1]^N$ to index all possible binary vectors of length N, as well as, $s \in [0 : 2^N - 1]$ to indicate the decimal representation of a binary vector of length N.

3.3 Capacity to within a constant gap

This section is devoted to the proof of Theorem 3. We first adapt the cut-set upper bound [16] and the NNC lower bound [20] to the HD case by following the approach proposed in [18]. We then show that these bounds are at most a constant number of bits apart. In particular, since the unicast Gaussian network with HD relays is a special MGN with $K = N + 2$ nodes (one source, N relays, and one destination), we first prove that for a single-antenna complex-valued MGN with HD power-constrained nodes the cut-set upper bound can be achieved to within 1.96 bits/node (while for the FD case the gap is 1.26 bits/node [20, Theorem 4]).

3.3.1 Channel Model

A MGN with K nodes [4] is defined similarly to the multi-relay network introduced in Section 3.1 except that now each node $k \in [1 : K]$, with channel input (X_k, S_k) and channel output Y_k, has an independent message of rate R_k to be decoded by the nodes indexed by $\mathcal{D} \subseteq [1 : K]$. The channel input/output relationship of this HD GMN reads $\mathbf{Y} = (\mathbf{I}_K - \mathrm{diag}[\mathbf{S}])\ \mathbf{H}\ \mathrm{diag}[\mathbf{S}]\ \mathbf{X} + \mathbf{Z}$. We let $\mathcal{C}_{\mathrm{multicast}}$ be the capacity region.

[4] Here, for notation convenience, we number the nodes from 1 to K, rather than from 0 to $N + 1$.

3.3.2 Inner Bound

The capacity of a HD GMN can be lower bounded by adapting the NNC scheme for the general memoryless network from [20] to the HD case by following the approach of [18]. In particular, NNC achieves the rate region

$$\mathcal{C}_{\text{multicast}} \supseteq \bigcup \left\{ \sum_{i \in \mathcal{A}^c} R_i \leq I(X_{\mathcal{A}^c}; \widehat{Y}_{\mathcal{A}} | X_{\mathcal{A}}, S_{[1:K]}, Q) \right. \\ - I(Y_{\mathcal{A}^c}; \widehat{Y}_{\mathcal{A}^c} | \widehat{Y}_{\mathcal{A}}, X_{[1:K]}, S_{[1:K]}, Q) \\ \left. \text{such that } \mathcal{A} \subseteq [1:K],\ \mathcal{A}^c \neq \emptyset,\ \mathcal{A} \cap \mathcal{D} \neq \emptyset \right\},$$

where $\widehat{Y}_k$ represents a compressed version of Y_k, for $k \in [1:K]$, and where the union is over all input distributions that factorize as

$$\mathbb{P}_Q \prod_{k=1}^{K} \mathbb{P}_{X_k, S_k | Q} \mathbb{P}_{\widehat{Y}_k | Y_k, X_k, S_k, Q}$$

and satisfy the power constraints. We consider jointly Gaussian inputs so as to get a rate region similar to [20, eq.(20)]. In all states $s \in [0:1]^K$, we consider i.i.d. $\mathcal{N}(0,1)$ inputs, time sharing random variable Q set to $Q = S_{[1:K]}$ (with this choice the nodes can coordinate), and compressed channel output $\widehat{Y}_k := Y_k + \widehat{Z}_k$, $k \in [1:K]$, for $\widehat{Z}_k \sim \mathcal{N}(0, \sigma^2)$ independent of all other random variables and where the variance of $\widehat{Z}_k$ does not depend on the user index k. With this the NNC achievable region evaluates to

$$\mathcal{C}_{\text{multicast}} \supseteq \bigcup \left\{ \sum_{i \in \mathcal{A}^c} R_i \ \leq \sum_{s \in [0:1]^K} \lambda_s \ \log \left| \mathbf{I}_{|\mathcal{A}|} + \frac{1}{1+\sigma^2}\ \mathbf{H}_{\mathcal{A},s} \mathbf{H}_{\mathcal{A},s}^H \right| \right. \\ - |\mathcal{A}^c| \log \left(1 + \frac{1}{\sigma^2} \right) \\ \left. \text{such that } \mathcal{A} \subseteq [1:K],\ \mathcal{A}^c \neq \emptyset,\ \mathcal{A} \cap \mathcal{D} \neq \emptyset \right\}, \quad (3.16)$$

where the union is over all $\lambda_s := \mathbb{P}[S_{[1:K]} = s] \in [0,1]$, $\forall s \in [0:1]^K$: $\sum_{s \in [0:1]^K} \lambda_s = 1$ and over all $\sigma^2 \in \mathbb{R}^+$, and where the matrix $\mathbf{H}_{\mathcal{A},s} \in \mathbb{C}^{|\mathcal{A}| \times |\mathcal{A}^c|}$ is defined as $\mathbf{H}_{\mathcal{A},s} := \left[(\mathbf{I}_K - \text{diag}[s])\ \mathbf{H}\ \text{diag}[s] \right]_{\mathcal{A},\mathcal{A}^c}$.

3.3.3 Outer Bound

The cut-set upper bound, adapted to the HD case by following [18], gives

$$\mathcal{C}_{\text{multicast}} \subseteq \bigcup \Big\{ \sum_{i\in\mathcal{A}^c} R_i \leq I(X_{\mathcal{A}^c}, S_{\mathcal{A}^c}; Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{\mathcal{A}}) \\ \text{such that } \mathcal{A} \subseteq [1:K], \mathcal{A}^c \neq \emptyset, \mathcal{A}\cap\mathcal{D}\neq\emptyset \Big\},$$

where the union is over all joint input distributions $\mathbb{P}_{X^K,S^K}$ and satisfy the power constraints. Similarly to [20, eq.(19)], we upper bound each mutual information term as

$$\begin{aligned}
&I(X_{\mathcal{A}^c}, S_{\mathcal{A}^c}; Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{\mathcal{A}}) \\
&= I(S_{\mathcal{A}^c}; Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{\mathcal{A}}) + I(X_{\mathcal{A}^c}; Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{[1:K]}) \\
&\leq H(S_{\mathcal{A}^c}) + \sum_{s\in[0:1]^K} \lambda_s \log\left|\mathbf{I}_{|\mathcal{A}|} + \mathbf{H}_{\mathcal{A},s}\mathbf{K}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right| &(3.17\text{a}) \\
&\leq |\mathcal{A}^c|\log(2) + \sum_{s\in[0:1]^K} \lambda_s \log\left|\mathbf{I}_{|\mathcal{A}|} + \frac{1}{\gamma}\mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right| \\
&\quad + \sum_{s\in[0:1]^K} \lambda_s |\mathcal{A}^c| \frac{\log\left(\mathrm{e}\max\left\{1, \frac{\gamma}{\mathrm{e}}\frac{|\mathcal{A}^c|}{\mathrm{rank}[\mathbf{H}_{\mathcal{A},s}]}\right\}\right)}{\max\left\{\frac{\mathrm{e}}{\gamma}, \frac{|\mathcal{A}^c|}{\mathrm{rank}[\mathbf{H}_{\mathcal{A},s}]}\right\}}, &(3.17\text{b})
\end{aligned}$$

where: (i) $\mathbf{K}_{\mathcal{A},s}$ represents the covariance matrix of $X_{\mathcal{A}^c}$ conditioned on $S_{[1:K]} = s$; (ii) the inequality in (3.17a) follows since conditioning reduces the entropy, since the entropy of a discrete random variable is non-negative, and by using the 'Gaussian maximizes entropy' principle; (iii) the inequality in (3.17b) follows since the entropy of a discrete random variable can be upper bounded as a function of the size of its support and from [20, Lemma 1] for all $\gamma \geq \mathrm{e} - 1$. Finally, since the function $\frac{\log(\mathrm{e}\ \max\{1,x\})}{\max\{1,x\}}$ in (3.17b) is decreasing in x, the function in (3.17b) attains its maximum value when $\mathrm{rank}[\mathbf{H}_{\mathcal{A},s}]$ is maximum, i.e., when $x = \frac{\gamma}{\mathrm{e}}\frac{|\mathcal{A}^c|}{\mathrm{rank}[\mathbf{H}_{\mathcal{A},s}]} = \frac{\gamma}{\mathrm{e}}\frac{|\mathcal{A}^c|}{\min\{|\mathcal{A}|,|\mathcal{A}^c|\}}$, from which it thus follows that

$$\begin{aligned}\mathcal{C}_{\text{multicast}} \subseteq \bigcup \Bigg\{ & \sum_{i\in\mathcal{A}^c} R_i \leq |\mathcal{A}^c|\log(2) + \sum_{s\in[0:1]^K} \lambda_s \ \log\left|\mathbf{I}_{|\mathcal{A}|} + \frac{1}{\gamma}\ \mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right| \\ & + |\mathcal{A}^c| \ \frac{\log\left(\mathrm{e}\max\left\{1, \frac{\gamma}{\mathrm{e}}\ \frac{|\mathcal{A}^c|}{\min\{|\mathcal{A}|,|\mathcal{A}^c|\}}\right\}\right)}{\max\left\{\frac{\mathrm{e}}{\gamma}, \frac{|\mathcal{A}^c|}{\min\{|\mathcal{A}|,|\mathcal{A}^c|\}}\right\}} \\ & \text{such that } \mathcal{A} \subseteq [1:K],\ \mathcal{A}^c \neq \emptyset,\ \mathcal{A}\cap\mathcal{D} \neq \emptyset \Bigg\} \qquad (3.18)\end{aligned}$$

holds, where the union is over all $\lambda_s := \mathbb{P}[S_{[1:K]} = s] \in [0,1],\ \forall s \in [0:1]^K :$ $\sum_{s\in[0:1]^K} \lambda_s = 1$ and where the parameter $\gamma \geq \mathrm{e}-1$ can be chosen so as to tighten the RHS of (3.18).

3.3.4 Gap

We now proceed to bound the worst case gap (over $\mathcal{A}$) between the cut-set upper bound in (3.18) and the NNC lower bound in (3.16) (recall that the parameters γ and σ^2 can be chosen so as to tighten the bound). By choosing $\sigma^2 = \gamma - 1$ in (3.16) and by defining $\mu = \frac{|\mathcal{A}^c|}{K} \in [0,1]$, the gap is given by

$$\begin{aligned}\frac{\mathsf{GAP}}{K} \leq \min_{\gamma\geq \mathrm{e}-1} \max_{\mu\in[0,1]} & \left\{\mu \ \log\left(\frac{2\gamma}{\gamma-1}\right)\right. \\ & \left. + \mu \ \min\left\{\frac{\gamma}{\mathrm{e}}, \frac{\min\{\mu, 1-\mu\}}{\mu}\right\} \log\left(\max\left\{\mathrm{e}, \frac{\gamma\ \mu}{\min\{\mu,1-\mu\}}\right\}\right)\right\} \\ \leq & \ 1.96 \text{ bits/node},\end{aligned}$$

where the last inequality follows by numerical evaluations. The gap result in Theorem 3 follows by substituting $K = N + 2$.

The difference between the HD and the FD case is the factor 2 (inside the logarithm) for the HD case. Also notice that the HD gap of 1.96 bits/node is smaller than $(1 + 1.26)$ bits/node where 1.26 bits/node is the FD gap [20] and the extra 1 bit/node is due to random switch.

Remark 7. The gap in Theorem 3 improves on the previously known gap result of $5N$ bits [28]. □

Remark 8 (Single relay case). The gap result in (3.6) for $N = 1$ gives $\mathsf{GAP} \leq$ 5.88 bits, which is greater than the 1.61 bits gap we found in Chapter 2. This is due to the fact that the bounding steps in the special case of $N = 1$ are tighter than those we used here for a general MGN with K nodes. Notice also

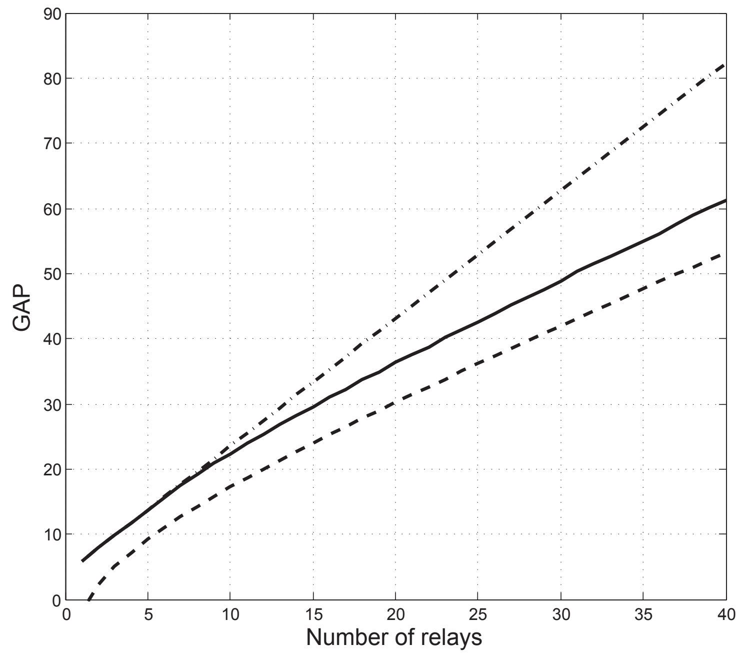


Figure 3.1: Gap in (3.6) (dash-dotted curve), gap in (3.6) specialized to the HD diamond network (solid curve) and gap in [1] (dashed curve) for the HD diamond network.

that for a single relay, PDF is optimal to within 1 bit (see Chapter 2). PDF has been extended to a general HD multi-relay network in [91]. However, to analytically evaluate this achievable rate and show that it achieves the cut-set upper bound to within a constant gap seems to be a challenging task, which is the main motivation for considering NNC here. □

Remark 9 (Diamond networks). A smaller gap than the one in (3.6) may be obtained for specific network topologies. For example, in [25] and [26] it was found that for a Gaussian FD diamond network with N relays the gap is of the order $\log(N)$, rather than linear in N [20]. Moreover, for a symmetric FD diamond network with N relays the gap does not depend on the number of relays and it is upper bounded by 3.6 bits [27]. The key difference between a general relay network and a diamond network is that for each subset $\mathcal{A}$ we have that $\text{rank}[\mathbf{H}_{\mathcal{A}}] \leq 2$; hence in (3.17b) we can use $\text{rank}[\mathbf{H}_{\mathcal{A}}] \leq \min\{|\mathcal{A}|, |\mathcal{A}^c|, 2\}$. With this and by numerically evaluating the resulting gap we obtain the result plotted in Figure 3.1. From Figure 3.1, we observe that the gap for the HD diamond network is in general smaller than the one computed for the general HD relay network; this is in line with what happens in FD. However, in FD for the diamond network the gap is logarithmic in N [25], [26], while the gap in Figure 3.1 (solid curve) still grows linearly with N. This is due to the fact that the HD cut-set outer bound, as opposed to the FD one, contains the entropy of the state vector, which is upper bounded by the uniform distribution over all the possible states; this term contributes linearly in the number of nodes to the overall gap. Moreover, from Figure 3.1 we observe that our gap (solid curve) is larger than the gap of order $N + 3\log(N)$ from [1] (dashed curve). We believe that the reason is because our gap has been computed as a special case of a general HD MGN while the one in [1] has been specifically derived for HD diamond relay networks. □

Remark 10. We argue here that Theorem 3, valid for Gaussian HD relay networks with single-antenna nodes, gives a constant gap result also for the case of multiple-antenna nodes. Actually, Theorem 3 holds for the more general MGN in which one has $K = N+2$ HD nodes (N relays, 1 source and 1 destination); thus, we shall argue that the gap result for the general single-antenna MGN extends to the multiple-antenna case. The key observation is to consider a MGN with multiple-antenna nodes as a new MGN with single-antenna nodes, where: (i) each node in the new MGN corresponds to a different antenna in the original MGN model and (ii) in the new MGN, the links connecting the nodes corresponding to different antennas at the same node in the original MGN are of infinite capacity. Now, since our original gap result applies to the new MGN (as the gap result in Theorem 3 holds uniformly over all channel gains), then for the original MGN we have that $\mathsf{GAP} \leq 1.96 M_{\rm tot}$ bits per channel use, with $M_{\rm tot}$ being the total number of nodes in the new MGN, that is the total number of antennas in the original MGN, i.e., $M_{\rm tot} := m_0 + m_{\rm tot} + m_{N+1}$.

3.4 Simple schedules for a class of HD multi-relay networks

The goal of this section is to prove Theorem 4, i.e., to show that simple relay policies are (approximately) optimal in (3.8).

We start by noting that the capacity $C^{(\mathrm{HD-RN})}$ of the HD multi-relay network is not known in general, but can be upper bounded by the cut-set bound

$$C^{(\mathrm{HD-RN})} \leq \max_{\mathbb{P}_{X_{[1:N+1]},S_{[1:N]}}} \min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(\mathrm{rand})}, \tag{3.19}$$

where

$$I_{\mathcal{A}}^{(\mathrm{rand})} := I\left(X_{N+1}, X_{\mathcal{A}^c}, S_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}} | X_{\mathcal{A}}, S_{\mathcal{A}}\right) \tag{3.20}$$

$$\leq H(S_{\mathcal{A}^c}) + I_{\mathcal{A}}^{(\mathrm{fix})}, \tag{3.21}$$

where $I_{\mathcal{A}}^{(\mathrm{fix})}$ is defined in (3.9). In particular, $I_{\mathcal{A}}^{(\mathrm{rand})}$ in (3.20) is the mutual information across the network cut $\mathcal{A} \subseteq [1:N]$ when a *random schedule* is employed, i.e., information is conveyed from the relays to the destination by switching between listen and transmit modes of operation at random times [18] (see the term $H(S_{\mathcal{A}^c}) \leq |\mathcal{A}^c| \leq N$ in (3.21)); $I_{\mathcal{A}}^{(\mathrm{fix})}$ in (3.9) is the mutual information with a *fixed schedule*, i.e., the time instants at which a relay transitions between listen and transmit modes of operation are fixed and known to all nodes in the network [18] (see the term $S_{[1:N]}$ in the conditioning in (3.9)).

The proof of Theorem 4 consists of the following steps:

1. We first show that the function $I_{\mathcal{A}}^{(\mathrm{fix})}$ defined in (3.9) is submodular under the three assumptions in Theorem 4.

2. By using Proposition 14, we show that the problem in (3.8) can be recast into an equivalent max-min problem.

3. With Proposition 16 we show that the max-min problem is equivalent to solve a min-max problem. The min-max problem is then shown to be equivalent to solve $N!$ max-min problems, for each of which we obtain an optimal basic feasible solution by Proposition 15 with the claimed maximum number of non-zero entries.

We now give the details for each step in a separate subsection.

3.4.1 Proof Step 1

We show that $I_{\mathcal{A}}^{(\mathrm{fix})}$ in (3.9) is submodular. The result in [92, Theorem 1] showed that $f_s(\mathcal{A})$ in (3.12) is submodular for each relay state $s \in [0:1]^N$ under the assumption of independent inputs and independent noises (the same work provides an example of a diamond network with correlated inputs for which the cut-set bound is neither submodular nor supermodular). Since submodular functions are closed under non-negative linear combinations (see Definition 3), this implies that $I_{\mathcal{A}}^{(\text{fix})} = \sum_{s\in[0:1]^N} \lambda_s \ f_s(\mathcal{A})$ is submodular under the assumptions of Theorem 4. For completeness, we provide the proof of this result in Appendix 3.A, where we use Definition 3 as opposed to the "diminishing marginal returns" property of a submodular function used in [92].

Example for $N = 2$: In this setting we have $2^2 = 4$ possible cuts, each of which is a linear combination of $2^2 = 4$ possible listen/transmission configuration states. In particular, from (3.10) we have

$$\begin{aligned}
&\mathcal{A} = \emptyset, && I_{\emptyset}^{(\text{fix})} := \lambda_0 f_0(\emptyset) + \lambda_1 f_1(\emptyset) + \lambda_2 f_2(\emptyset) + \lambda_3 f_3(\emptyset), \\
&\mathcal{A} = \{1\}, && I_{\{1\}}^{(\text{fix})} := \lambda_0 f_0(\{1\}) + \lambda_1 f_1(\{1\}) + \lambda_2 f_2(\{1\}) + \lambda_3 f_3(\{1\}), \\
&\mathcal{A} = \{2\}, && I_{\{2\}}^{(\text{fix})} := \lambda_0 f_0(\{2\}) + \lambda_1 f_1(\{2\}) + \lambda_2 f_2(\{2\}) + \lambda_3 f_3(\{2\}), \\
&\mathcal{A} = \{1,2\}, && I_{\{1,2\}}^{(\text{fix})} := \lambda_0 f_0(\{1,2\}) + \lambda_1 f_1(\{1,2\}) \\
& && \qquad\qquad + \lambda_2 f_2(\{1,2\}) + \lambda_3 f_3(\{1,2\}),
\end{aligned}$$

where, $\forall s \in [0:3]$, we have that the functions in (3.12) are given by

$$\begin{aligned}
f_s(\emptyset) &:= I\left(X_3, X_1, X_2; Y_3 | S_{[1:2]} = s\right), \\
f_s(\{1\}) &:= I\left(X_3, X_2; Y_3, Y_1 | X_1, S_{[1:2]} = s\right), \\
f_s(\{2\}) &:= I\left(X_3, X_1; Y_3, Y_2 | X_2, S_{[1:2]} = s\right), \\
f_s(\{1,2\}) &:= I\left(X_3; Y_3, Y_2, Y_1 | X_2, X_1, S_{[1:2]} = s\right),
\end{aligned}$$

and are submodular under the three assumptions in Theorem 4.

3.4.2 Proof Step 2

Given that $I_{\mathcal{A}}^{(\text{fix})}$ in (3.9) is submodular, we would like to use Proposition 14 to replace the minimization over the subsets of $[1:N]$ in (3.8) with a minimization over the cube $[0:1]^N$. Since $I_{\emptyset}^{(\text{fix})} = I\left(X_{[1:N+1]}; Y_{N+1} | S_{[1:N]}\right) \geq 0$ in general, we define a new submodular function

$$g(\mathcal{A}) := I_{\mathcal{A}}^{(\text{fix})} - I_{\emptyset}^{(\text{fix})} \tag{3.22}$$

and proceed as follows

$$
\begin{aligned}
\min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(\mathrm{fix})} &= I_{\emptyset}^{(\mathrm{fix})} + \min_{\mathcal{A}\subseteq[1:N]} g\left(\mathcal{A}\right) \\
&= I_{\emptyset}^{(\mathrm{fix})} + \min_{\mathbf{w}\in[0,1]^N} \begin{bmatrix} w_{\pi_1} & w_{\pi_2} & \dots & w_{\pi_N} \end{bmatrix} \begin{bmatrix} g\left(\{\pi_1\}\right) - g\left(\emptyset\right) \\ \vdots \\ g\left(\{\pi_1,\dots,\pi_N\}\right) - g\left(\{\pi_1,\dots,\pi_{N-1}\}\right) \end{bmatrix} \\
&= I_{\emptyset}^{(\mathrm{fix})} + \min_{\mathbf{w}\in[0,1]^N} \begin{bmatrix} w_{\pi_1} & w_{\pi_2} & \dots & w_{\pi_N} \end{bmatrix} \begin{bmatrix} I_{\{\pi_1\}}^{(\mathrm{fix})} - I_{\emptyset}^{(\mathrm{fix})} \\ \vdots \\ I_{\{\pi_1,\dots,\pi_N\}}^{(\mathrm{fix})} - I_{\{\pi_1,\dots,\pi_{N-1}\}}^{(\mathrm{fix})} \end{bmatrix} \\
&= \min_{\mathbf{w}\in[0,1]^N} \begin{bmatrix} 1 & w_{\pi_1} & w_{\pi_2} & \dots & w_{\pi_N} \end{bmatrix} \begin{bmatrix} I_{\emptyset}^{(\mathrm{fix})} \\ I_{\{\pi_1\}}^{(\mathrm{fix})} - I_{\emptyset}^{(\mathrm{fix})} \\ \vdots \\ I_{\{\pi_1,\dots,\pi_N\}}^{(\mathrm{fix})} - I_{\{\pi_1,\dots,\pi_{N-1}\}}^{(\mathrm{fix})} \end{bmatrix} \\
&=: \min_{\mathbf{w}\in[0,1]^N} \left\{ [1, \mathbf{w}^T] \ \mathbf{H}_{\pi,f} \right\}, \qquad (3.23)
\end{aligned}
$$

which implies that the problem in (3.8) is equivalent to

$$
\mathsf{C}' = \max_{\lambda_{\mathrm{vect}}} \min_{\mathbf{w}\in[0,1]^N} \left\{ [1, \mathbf{w}^T] \ \mathbf{H}_{\pi,f} \lambda_{\mathrm{vect}} \right\}, \qquad (3.24)
$$

where λ_{vect} is the probability mass function of $S_{[1:N]}$ in (3.11), $\mathbf{H}_{\pi,f}$ is defined as

$$
\mathbf{H}_{\pi,f} := \mathbf{P}_{\pi} \underbrace{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & -1 & 1 \end{bmatrix}}_{(N+1)\times(N+1)} \mathbf{F}_{\pi} \in \mathbb{R}^{(N+1)\times 2^N}, \qquad (3.25)
$$

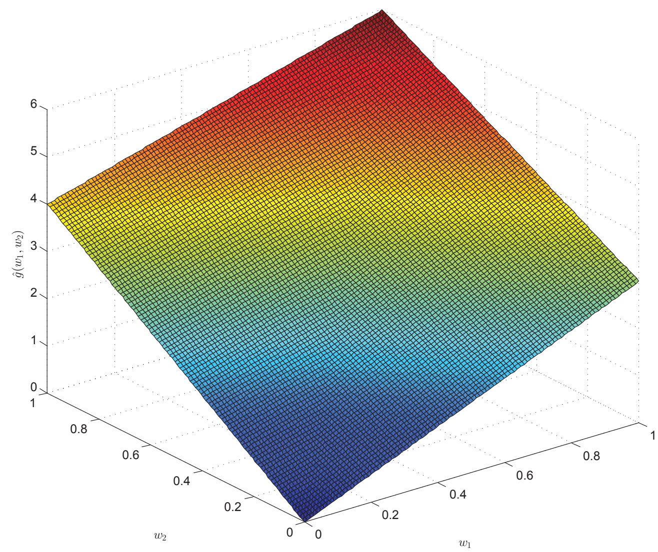


Figure 3.2: Lovász extension $\widehat{g}(w_1, w_2)$ in (3.27), with $g(\{1\}) = 3$, $g(\{2\}) = 4$ and $g(\{1, 2\}) = 6$.

where $\mathbf{P}_\pi \in \mathbb{R}^{(N+1)\times(N+1)}$ is the permutation matrix that maps $[1, w_1, \ldots, w_N]$ into $[1, w_{\pi_1}, \ldots, w_{\pi_N}]$, and $\mathbf{F}_\pi$ is defined as

$$\mathbf{F}_\pi := \begin{bmatrix} f_0(\emptyset) & \ldots & f_{2^N-1}(\emptyset) \\ f_0(\{\pi_1\}) & \ldots & f_{2^N-1}(\{\pi_1\}) \\ f_0(\{\pi_1, \pi_2\}) & \ldots & f_{2^N-1}(\{\pi_1, \pi_2\}) \\ \ldots & & \\ f_0(\{\pi_1, \ldots, \pi_N\}) & \ldots & f_{2^N-1}(\{\pi_1, \ldots, \pi_N\}) \end{bmatrix} \in \mathbb{R}^{(N+1)\times 2^N}, \tag{3.26}$$

with $f_s(\mathcal{A})$ being defined in (3.12). We thus expressed our original optimization problem in (3.8) as the max-min problem in (3.24).

Example for $N = 2$**:** With $N = 2$, we have $g(\mathcal{A}) = I_{\mathcal{A}}^{(\text{fix})} - I_{\emptyset}^{(\text{fix})}, \mathcal{A} \subseteq [1 : 2]$ and the Lovász extension (see Definition 3) is

$$\widehat{g}(w_1, w_2) = \begin{cases} w_1 g(\{1\}) + w_2 [g(\{1, 2\}) - g(\{1\})] & \text{if } w_1 \geq w_2 \\ w_2 g(\{2\}) + w_1 [g(\{1, 2\}) - g(\{2\})] & \text{if } w_2 \geq w_1 \end{cases}. \tag{3.27}$$

A visual representation of the Lovász extension $\widehat{g}(w_1, w_2)$ in (3.27) on $[0, 1]^2$ is given in Figure 3.2, where we considered $g(\{1\}) = 3$, $g(\{2\}) = 4$ and $g(\{1, 2\}) = 6$ (recall $g(\emptyset) = 0$).

Let

$$i_{\mathrm{M}} := \arg\max\{w_1, w_2\} \quad \text{and} \quad i_{\mathrm{m}} := \arg\min\{w_1, w_2\}. \tag{3.28}$$

The optimization problem in (3.23) for $N = 2$ can be written as

$$\begin{aligned}
&\min_{0 \le w_{i_{\mathrm{m}}} \le w_{i_{\mathrm{M}}} \le 1} \left\{ \begin{bmatrix} 1 & w_{i_{\mathrm{M}}} & w_{i_{\mathrm{m}}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \mathbf{F}_\pi \right\} \\
&= \min_{0 \le w_{i_{\mathrm{m}}} \le w_{i_{\mathrm{M}}} \le 1} \left\{ \begin{bmatrix} 1 - w_{i_{\mathrm{M}}} & w_{i_{\mathrm{M}}} - w_{i_{\mathrm{m}}} & w_{i_{\mathrm{m}}} \end{bmatrix} \mathbf{F}_\pi \right\},
\end{aligned} \tag{3.29}$$

with

$$\mathbf{F}_\pi = \begin{bmatrix} f_0(\emptyset) & f_1(\emptyset) & f_2(\emptyset) & f_3(\emptyset) \\ f_0(\{i_{\mathrm{M}}\}) & f_1(\{i_{\mathrm{M}}\}) & f_2(\{i_{\mathrm{M}}\}) & f_3(\{i_{\mathrm{M}}\}) \\ f_0(\{1,2\}) & f_1(\{1,2\}) & f_2(\{1,2\}) & f_3(\{1,2\}) \end{bmatrix}, \tag{3.30}$$

and finally the optimization problem in (3.24) is

$$\mathsf{C}' = \max_{\lambda_{\mathrm{vect}}} \min_{0 \le w_{i_{\mathrm{m}}} \le w_{i_{\mathrm{M}}} \le 1} \left\{ \begin{bmatrix} 1 - w_{i_{\mathrm{M}}} & w_{i_{\mathrm{M}}} - w_{i_{\mathrm{m}}} & w_{i_{\mathrm{m}}} \end{bmatrix} \mathbf{F}_\pi \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} \right\}. \tag{3.31}$$

3.4.3 Proof Step 3

In order to solve (3.24) we would like to reverse the order of min and max. We note that the function $\phi(\lambda_{\mathrm{vect}}, \mathbf{w}) := [1, \mathbf{w}^T]\, \mathbf{H}_{\pi,f} \lambda_{\mathrm{vect}}$ satisfies the properties in Proposition 16 (it is continuous; it is convex in $\mathbf{w}$ by the convexity of the Lovász extension and linear (under the assumption in item 3 in Theorem 4), thus concave, in λ_{vect}; the optimization domain in both variables is compact). Thus, we now focus on the problem

$$\mathsf{C}' = \min_{\mathbf{w} \in [0,1]^N} \max_{\lambda_{\mathrm{vect}}} \left\{ [1, \mathbf{w}^T]\, \mathbf{H}_{\pi,f} \lambda_{\mathrm{vect}} \right\}, \tag{3.32}$$

which can be equivalently rewritten as

$$\begin{aligned}
\mathsf{C}' &= \min_{\pi \in \mathcal{P}_N} \min_{\mathbf{w}_\pi \in [0:1]^N} \max_{\lambda_{\mathrm{vect}}} \left\{ [1, \mathbf{w}_\pi^T]\, \mathbf{H}_{\pi,f} \lambda_{\mathrm{vect}} \right\} && (3.33) \\
&= \min_{\pi \in \mathcal{P}_N} \max_{\lambda_{\mathrm{vect}}} \min_{\mathbf{w}_\pi \in [0:1]^N} \left\{ [1, \mathbf{w}_\pi^T]\, \mathbf{H}_{\pi,f} \lambda_{\mathrm{vect}} \right\}, && (3.34)
\end{aligned}$$

where $\mathcal{P}_N$ is the set of all the $N!$ permutations of $[1:N]$. In (3.33), for each permutation $\pi \in \mathcal{P}_N$, we first find the optimal λ_{vect}, and then find the optimal $\mathbf{w}_\pi : w_{\pi_1} \geq w_{\pi_2} \geq \ldots w_{\pi_N}$. This is equivalent to (3.34), where again by Proposition 16, for each permutation $\pi \in \mathcal{P}_N$, we first find the optimal $\mathbf{w}_\pi : w_{\pi_1} \geq w_{\pi_2} \geq \ldots w_{\pi_N}$, and then find the optimal λ_{vect}.

Let now consider the inner optimization in (3.34), that is, the problem

$$P_1 : \max_{\lambda_{\text{vect}}} \min_{\mathbf{w}_\pi \in [0:1]^N} \left\{ [1, \mathbf{w}_\pi^T] \, \mathbf{H}_{\pi,f} \lambda_{\text{vect}} \right\}. \tag{3.35}$$

From Proposition 14 we know that, for a given $\pi \in \mathcal{P}_N$, the optimal $\mathbf{w}_\pi$ is a vertex of the cube $[0:1]^N$. For a given $\pi \in \mathcal{P}_N$, there are $N+1$ vertices whose coordinates are ordered according to π. In (3.35), for each of the $N+1$ feasible vertices of $\mathbf{w}_\pi$, it is easy to see that the product $[1, \mathbf{w}_\pi^T] \, \mathbf{H}_{\pi,f}$ is equal to a row of the matrix $\mathbf{F}_\pi$. By considering all possible $N+1$ feasible vertices compatible with π we obtain all the $N+1$ rows of the matrix $\mathbf{F}_\pi$. Hence, P_1 is equivalent to

$$\begin{array}{lll} P_2 : & \text{maximize} & \tau \\ & \text{subject to} & \mathbf{1}_{(N+1)} \tau \leq \mathbf{F}_\pi \lambda_{\text{vect}} \\ & \text{and} & \mathbf{1}_{2^N}^T \lambda_{\text{vect}} = 1, \ \lambda_{\text{vect}} \geq \mathbf{0}_{2^N}, \ \tau \geq 0. \end{array} \tag{3.36}$$

The LP P_2 in (3.36) has $n = 2^N + 1$ optimization variables (2^N values for λ_{vect} and one value for τ), $m = N+2$ constraints, and is feasible (consider for example the uniform distribution of λ_{vect} and $\tau = 0$). Therefore, by Proposition 15, P_2 has an optimal basic feasible solution with at most $m = N+2$ non-zero values. Since $\tau > 0$ (otherwise the channel capacity would be zero), it means that λ_{vect} has at most $N+1$ non-zero entries.

Since for each $\pi \in \mathcal{P}_N$ the optimal λ_{vect} in (3.34) has at most $N+1$ non-zero values, then also for the optimal permutation the corresponding optimal λ_{vect} has at most $N+1$ non-zero values. This shows that the (approximately) optimal schedule in the original problem in (3.8) is simple.

This concludes the proof of Theorem 4.

Example for $N = 2$: For $N = 2$, we have $|\mathcal{P}_2| = 2! = 2$ possible permutations. From Proposition 14, the optimal $\mathbf{w}$ is one of the vertices $(0,0), (0,1), (1,0), (1,1)$. Let now focus on the case $i_{\text{M}} = 1$ and $i_{\text{m}} = 2$ (a similar reasoning holds for $i_{\text{M}} = 2$ and $i_{\text{m}} = 1$ as well). Under this condition P_1 in (3.35) is the problem in (3.31) with $i_{\text{M}} = 1$ and $i_{\text{m}} = 2$. The vertices compatible with this permutation are $(w_1, w_2) \in \{(0,0), (1,0), (1,1)\}$, which result in $(1-w_1, w_1-w_2, w_2) \in \{(1,0,0),(0,1,0),(0,0,1)\}$. This implies that P_2 in (3.36) is

$$\begin{array}{ll} \text{maximize} & \tau \\ \text{subject to} & \tau \le f_0(\emptyset)\lambda_0 + f_1(\emptyset)\lambda_1 + f_2(\emptyset)\lambda_2 + f_3(\emptyset)\lambda_3, \\ & \tau \le f_0(\{1\})\lambda_0 + f_1(\{1\})\lambda_1 + f_2(\{1\})\lambda_2 + f_3(\{1\})\lambda_3, \\ & \tau \le f_0(\{1,2\})\lambda_0 + f_1(\{1,2\})\lambda_1 + f_2(\{1,2\})\lambda_2 + f_3(\{1,2\})\lambda_3, \\ & \lambda_0+\lambda_1+\lambda_2+\lambda_3 = 1,\ \lambda_i \ge 0\ i \in [0:3],\ \tau \ge 0, \end{array} \tag{3.37}$$

where each of the three inequality constraints correspond to a different row of $\mathbf{F}_\pi$ multiplied by $\lambda_{\text{vect}} = [\lambda_0, \lambda_1, \lambda_2, \lambda_3]^T$. Therefore, P_2 in (3.37) has four constraints (three from the rows of $\mathbf{F}_\pi$ and one from λ_{vect}) and five unknowns (one value for τ and four entries of λ_{vect}). Thus, by Proposition 15, P_2 has an optimal basic feasible solution with at most four non-zero values, of which one is τ and thus the other (at most) three belong to λ_{vect}. In particular, in Appendix 3.B, we show that either λ_0 or λ_3 is zero, thus giving the desired (approximately) optimal simple schedule.

Remark 11. In order to apply the saddle-point property (see Proposition 16) and hence cast our optimization problem as a LP, the proof of Step 3 requires that the matrix $\mathbf{F}_\pi$ does not depend on λ_{vect}; this is the reason of our assumption in item 3 in Theorem 4. In Gaussian noise this assumption excludes the possibility of power allocation across the relay states because power allocation makes the optimization problem non-linear in λ_{vect}.

Remark 12. As stated in Theorem 4, our three assumptions provide a set of sufficient conditions for the existence of an (approximately) optimal simple schedule. As those conditions are not necessary, there might exist networks for which these assumptions are not satisfied, but for which the (approximately) optimal schedule is still simple. Determining necessary conditions for optimality of simple schedules is an interesting challenging open question.

Remark 13. For FD relays, it was shown in [92] that wireless erasure networks, Gaussian networks with single-antenna nodes and their linear deterministic high-SNR approximations are examples for which the cut-set bound (or an approximation to it) is submodular. Since submodular functions are closed under non-negative linear combinations (see Definition 3), this implies that the cut-set bound (or an approximation to it) is submodular when evaluated for these networks with HD relays. As a consequence, Theorem 4 holds for wireless erasure networks, Gaussian networks with single-antenna nodes and their linear deterministic high-SNR approximations with HD relays.

Remark 14. Gaussian relay networks with multi-antenna nodes, where each antenna at the relays can be switched independently of one another, satisfy all the conditions in Theorem 4. Actually, as highlighted in Remark 10, the NNC strategy, which uses independent inputs, achieves the cut-set upper bound to within a constant gap; moreover, as we shall see in the example in Section 3.6.2, a constant power allocation across the relay states is optimal to within a constant gap. As we showed for the single-antenna nodes case, what dictates the number of active states of the relay scheduling policy is related to the minimization over $\mathcal{A} \subseteq [1:N]$ and not to the maximization over the $2^{m_{\mathrm{tot}}}$ possible relay configurations. This extends the result in Theorem 4 to Gaussian HD multi-relay networks with multi-antenna nodes, i.e., the (approximately) optimal schedule has at most $N+1$ active states (out of the $2^{2^{m_{\mathrm{tot}}}}$ possible ones), independently of the total number of antennas. This result implies that for Gaussian relay networks, the cut-set upper bound can be achieved to within a constant gap by employing the NNC strategy with a time-sharing among the $N+1$ active states.

3.5 The gDoF and its relation to the MWBM problem.

In order to determine the gDoF of the Gaussian HD multi-relay network we must find a tight high-SNR approximation for the different MIMO-type mutual information terms involved in the cut-set upper bound (see eq.(3.17a)). As a result of independent interest beyond the application to the Gaussian HD relay network studied in this chapter, we first show that such an approximation can be found as the solution of a MWBM problem.

In particular, equipped with the definitions in Definition 5, we now show the following high-SNR approximation of the MIMO capacity:

Theorem 6. *Let the channel matrix $\mathbf{H} \in \mathbb{R}^{k \times n}$ be a full-rank matrix, where without loss of generality $k \leq n$. Let $\mathcal{S}_{n,k}$ be the set of all k-combinations of the integers in $[1:n]$ and $\mathcal{P}_{n,k}$ be the set of all k-permutations of the integers in $[1:n]$*[5].

[5] The k-combinations and the k-permutations of the integers in $[1:n]$ are defined as sequences of a fixed length k of elements taken from a given set of size n such that no elements occurs more than once. Then, over this k-length sequence all the possible combinations $\mathcal{S}_{n,k}$ and all the possible permutations $\mathcal{P}_{n,k}$ are computed. With $\pi(i)$ we indicate the element in the i-th position of the permutation $\pi \in \mathcal{P}_{n,k}$.

Then,

$$|\mathbf{I}_k + \mathbf{H}\mathbf{H}^H| = \sum_{\varsigma \in \mathcal{S}_{n,k}} \sum_{\pi \in \mathcal{P}_{n,k}} \prod_{i=1}^{k} \left|[\mathbf{H}_\varsigma]_{i,\pi(i)}\right|^2 + \quad T \doteq \mathsf{SNR}^{\mathsf{MWBM}(\mathbf{B})},$$

where

$$\mathsf{MWBM}(\mathbf{B}) := \max_{\varsigma \in \mathcal{S}_{n,k}} \max_{\pi \in \mathcal{P}_{n,k}} \sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi(i)}, \tag{3.38}$$

where $\mathbf{B}$ *is the SNR-exponent matrix defined as* $[\mathbf{B}]_{ij} = \beta_{ij} \geq 0 : |h_{ij}|^2 = \mathsf{SNR}^{\beta_{ij}}$ *(with* h_{ij} *being the channel gain from the* j*-th antenna at the transmitter to the* i*-th antenna at the receiver),* $\mathbf{H}_\varsigma$ *and* $\mathbf{B}_\varsigma$ *are the square matrices obtained from* $\mathbf{H}$ *and* $\mathbf{B}$*, respectively, by retaining all rows and the columns indexed by* ς*, and* T *is the sum of terms that overall behave as* $o\left(\mathsf{SNR}^{\mathsf{MWBM}(\mathbf{B})}\right)$.

Proof. The proof can be found in Appendix 3.C. The expression in (3.38) is a possible way of writing the MWBM problem. □

Theorem 6 establishes an interesting connection between the gDoF of a MIMO channel (with independent inputs) and graph theory. Notice that the high-SNR expression found in Theorem 6 holds for correlated inputs as well, as long as the average power constraint is a finite constant. More importantly, Theorem 6 allows to move from DoF, where all exponents β_{ij} have the same value, to gDoF, where different channel gains have different exponential behavior. DoF is essentially a characterization of the rank of the channel matrix; gDoF captures the potential advantage due to 'asymmetric' channel gains. In Section 3.7 we will show, through some network examples, that Theorem 6 is an efficient tool to characterize the gDoF region for any Gaussian network whose capacity can be approximated to within a constant gap by linear combinations of $\log|\ldots|$ terms and it also represents an useful tool to solve user scheduling problems.

With Theorem 6 we can now express the gDoF $\mathsf{d}^{(\mathrm{HD-RN})}$ of a HD relay network as in Theorem 5, where the non-negative matrix $\mathbf{A} \in \mathbb{R}^{2^N \times 2^N}$ in (3.15) has entries

$$[\mathbf{A}]_{ij} := \lim_{\mathsf{SNR} \to +\infty} \frac{I(X_{\mathcal{A}_i^c \cup \{N+1\}}; Y_{\mathcal{A}_i \cup \{N+1\}} | X_{\mathcal{A}_i}, S_{[1:N]} = s_j)}{\log(1 + \mathsf{SNR})}, \tag{3.39}$$

where $\mathcal{A}_i$ and s_j are defined right after Theorem 7. In other words, each row of the matrix $\mathbf{A}$ refers to a possible cut in the network, while each column of $\mathbf{A}$ refers to a possible listening/transmitting configuration state.

By a simple application of Theorem 6 we have that each entry of the matrix $\mathbf{A}$ can be evaluated by solving the corresponding MWBM problem. More formally

Theorem 7. *For the LP in Theorem 5*

$$[\mathbf{A}]_{ij} = \mathsf{MWBM}\left(\mathbf{B}_{\{N+1\}\cup(\mathcal{A}_i\cap\mathcal{A}_j),\ \{N+1\}\cup(\mathcal{A}_i^c\cap\mathcal{A}_j^c\}}\right).$$

The notation in eq.(3.39) and in Theorem 7 is as follows. $\mathbf{B}$ indicates the SNR-exponent matrix defined as $[\mathbf{B}]_{ij} = \beta_{ij} \geq 0 : |h_{ij}|^2 = \mathsf{SNR}^{\beta_{ij}}$, and the indices (i, j) have the following meaning. Index i refers to a "cut" in the network and index j to a "state of the relays". Both indices range in $[1 : 2^N]$ and must be seen as the decimal representation of a binary number with N bits. $\mathcal{A}_i^c$, $i \in [1 : 2^N]$, is the set of those relays who have a one in the corresponding binary representation of $i - 1$ and s_j, $j \in [1 : 2^N]$, sets the state of a relay to the corresponding bit in the binary representation of $j - 1$. Finally, we evaluate the MWBM of the bigraph with weight matrix

$$\begin{aligned}&\left[\begin{bmatrix}\mathbf{I}_N - \mathrm{diag}[s_j] & \mathbf{0}_N \\ \mathbf{0}_N^T & 1\end{bmatrix}\mathbf{B}\begin{bmatrix}\mathrm{diag}[s_j] & \mathbf{0}_N \\ \mathbf{0}_N^T & 1\end{bmatrix}\right]_{\{N+1\}\cup\mathcal{A}_i, \{N+1\}\cup\mathcal{A}_i^c} \\ &= \mathbf{B}_{\{N+1\}\cup(\mathcal{A}_i\cap\mathcal{A}_j),\ \{N+1\}\cup(\mathcal{A}_i^c\cap\mathcal{A}_j^c\}},\end{aligned}$$

where the equality follows from the following observation. Among the relays 'on the side of the destination' (indexed by $\mathcal{A}_i$) only those in receive mode matter (indexed by $\mathcal{A}_j$), therefore we can reduce the set of 'receiving nodes' from $\mathcal{A}_i$ to $\mathcal{A}_i \cap \mathcal{A}_j$. Similarly, among the relays 'on the side of the source' (indexed by $\mathcal{A}_i^c$) only those in transmit mode matter (indexed by $\mathcal{A}_j^c$), therefore we can reduce the set of 'transmitting nodes' from $\mathcal{A}_i^c$ to $\mathcal{A}_i^c \cap \mathcal{A}_j^c$. Notice that $\mathbf{B}_{\{N+1\}\cup(\mathcal{A}_i\cap\mathcal{A}_j),\ \{N+1\}\cup(\mathcal{A}_i^c\cap\mathcal{A}_j^c\}}$ does not change if the roles of i and j are swapped, which implies that $[\mathbf{A}]_{ij} = [\mathbf{A}]_{ji}$, i.e., the matrix $\mathbf{A}$ is symmetric. To better understand the notation, consider the following example.

Example: $N = 3$, $i = 7$, and $j = 5$. From $i - 1 = 6 = 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0$ we have $\mathcal{A}_7 = \{3\} = \{1, 2\}^c$, meaning that relay 1 and relay 2 lie in the cut of the source and relay 3 lies in the cut of the destination. From $j - 1 = 4 = 1 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0$ we have $s_5 = [1, 0, 0]$, meaning that relay 1 is transmitting, and relays 2 and 3 are receiving (also $\mathcal{A}_5 = \{2,3\} = \{1\}^c$). With this we have

$$\begin{aligned}
&\{N+1\} \cup \mathcal{A}_i = \{4\} \cup \{3\} = \{3,4\}, \\
&\{N+1\} \cup \mathcal{A}_i^c = \{4\} \cup \{1,2\} = \{1,2,4\}, \\
&\{N+1\} \cup (\mathcal{A}_i^c \cap \mathcal{A}_j^c) = \{4\} \cup (\{1,2\} \cap \{1\}) = \{1,4\}, \\
&\{N+1\} \cup (\mathcal{A}_i \cap \mathcal{A}_j\} = \{4\} \cup (\{3\} \cap \{2,3\}) = \{3,4\},
\end{aligned}$$

and

$$\begin{aligned}
[\mathbf{A}]_{7,5} &= \mathsf{MWBM}\left(\left[\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \\ \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right]_{\{3,4\},\{1,2,4\}}\right) \\
&= \mathsf{MWBM}\left(\begin{bmatrix} \beta_{31} & 0 & \beta_{34} \\ \beta_{41} & 0 & \beta_{44} \end{bmatrix}\right) = \max\left\{\beta_{31}+\beta_{44}, \beta_{34}+\beta_{41}\right\}.
\end{aligned}$$

Also

$$\begin{aligned}
[\mathbf{A}]_{7,5} &= \mathsf{MWBM}\left(\begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \\ \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{bmatrix}_{\{3,4\},\{1,4\}}\right) \\
&= \mathsf{MWBM}\left(\begin{bmatrix} \beta_{31} & \beta_{34} \\ \beta_{41} & \beta_{44} \end{bmatrix}\right) = \max\left\{\beta_{31}+\beta_{44}, \beta_{34}+\beta_{41}\right\}. \square
\end{aligned}$$

3.6 Network examples

In order to gain insights into how relays are best utilized, in this section we analyze two network examples. In particular, the first example, shown in Figure 3.3, consists of $N = 2$ single-antenna HD relays (RN1 and RN2) assisting the communication between a source (Tx) and a destination (Rx), while in the second example, shown in Figure 3.5, there is $N = 1$ relay (RN) equipped with $m_r = 2$ antennas. For the scenario in Figure 3.3 we seek to find under which channel conditions a best-relay selection scheme is strictly suboptimal in terms of gDoF with respect of using both relays, while for the scenario in Figure 3.5 we aim to show that independently switching the $m_r = 2$ antennas at the relay not only achieves in general strictly higher rates compared to using the $m_r = 2$ antennas for the same purpose, but can actually provide a strictly larger pre-log factor. We now analyze these two scenarios separately.

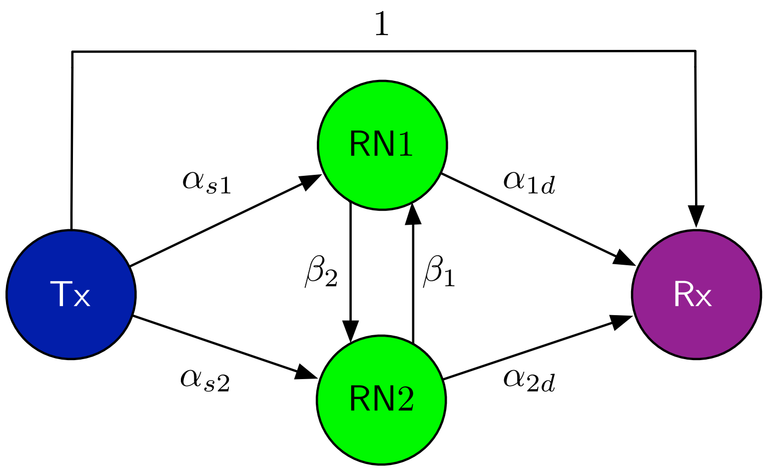


Figure 3.3: Example of a network with $N = 2$ relays with single-antenna nodes.

3.6.1 Example 1: HD relay network with $N = 2$ relays

We consider the network in Figure 3.3 where, in order to increase the readability, the SNR-exponents are indicated as

$$\left[\frac{\log(|h_{ij}|^2)}{\log(\mathsf{SNR})}\right]_{(i,j)\in[1:3]^2} = \begin{pmatrix} \star & \beta_1 & \alpha_{s1} \\ \beta_2 & \star & \alpha_{s2} \\ \alpha_{1d} & \alpha_{2d} & 1 \end{pmatrix}, \tag{3.40}$$

where $\star$ denotes an entry that does not matter for channel capacity, α_{si} is the SNR-exponent on the link from the source to relay i, $i \in [1:2]$, α_{id} is the SNR-exponent on the link from relay i, $i \in [1:2]$, to the destination, β_i is the SNR-exponent on the link from relay j to relay i, $(i,j) \in [1:2]^2$ with $j \neq i$, and the direct link from the source to the destination (entry in position (3,3) in (3.40)) has SNR-exponent normalized to 1 without loss of generality. Notice that in order to consider a network without a direct link it suffices to consider all the other SNR-exponents to be larger than 1, or simply replace '1' with '0' in the discussion in the rest of the section.

We next derive the gDoF in both the FD and HD cases.

The full-duplex case: For the FD case, the cut-set bound is achievable to within $2 \times 0.63 \times 4 = 5.04$ bits with NNC [20]. As a consequence, it can be verified that the gDoF for the FD case is

$$\begin{aligned}\mathsf{d}^{(\mathrm{FD})}_{N=2} = \min\Big\{ &\max\{1,\alpha_{s1},\alpha_{s2}\}, \max\{\alpha_{s2}+\alpha_{1d},\beta_2+1\}, \\ &\max\{\alpha_{s1}+\alpha_{2d},\beta_1+1\}, \max\{1,\alpha_{1d},\alpha_{2d}\}\Big\}. \end{aligned} \tag{3.41}$$

Note that $\mathsf{d}^{(\mathrm{FD})}_{N=2} \geq 1$, i.e., the gDoF in (3.41) is no smaller than the gDoF that could be achieved without using the relays, that is, by communicating directly through the direct link to achieve gDoF $= 1$. Notice also that the gDoF in (3.41) does not change if we exchange α_{s1} with α_{2d}, and α_{s2} with α_{1d}, i.e., if we swap the role of the source and destination. We aim to identify the channel conditions under which using both relays strictly improves the gDoF compared to the best-relay selection policy (which includes direct transmission from the source to the destination as a special case) that achieves

$$\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} = \max\left\{1, \min\{\alpha_{s1},\alpha_{1d}\}, \min\{\alpha_{s2},\alpha_{2d}\}\right\} \in [1, \mathsf{d}^{(\mathrm{FD})}_{N=2}]. \tag{3.42}$$

We distinguish the following cases:

1. *Case 1*: if
$$\text{either} \begin{cases} \alpha_{s1} \geq \alpha_{s2} \\ \alpha_{1d} \geq \alpha_{2d} \end{cases} \text{or} \begin{cases} \alpha_{s1} < \alpha_{s2} \\ \alpha_{1d} < \alpha_{2d} \end{cases}$$
then, since one of the relays is 'uniformly better' than the other, we immediately see that $\mathsf{d}^{(\mathrm{FD})}_{N=2} = \mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}}$, so in this regime selecting the best relay for transmission is gDoF optimal.

2. *Case 2*: if not in Case 1, then we are in
$$\text{either} \begin{cases} \alpha_{s1} \geq \alpha_{s2} \\ \alpha_{1d} < \alpha_{2d} \end{cases} \text{or} \begin{cases} \alpha_{s1} < \alpha_{s2} \\ \alpha_{1d} \geq \alpha_{2d} \end{cases}.$$
Consider the case $\alpha_{s2} \leq \alpha_{s1},\ \alpha_{1d} < \alpha_{2d}$ (the other one is obtained essentially by swapping the role of the relays). This corresponds to an 'asymmetric' situation where relay 1 has the best link from the source but relay 2 has the best link to the destination. In this case we would like to exploit the inter relay communication links (which are not present in a diamond network) to create a route source $\to$ relay1 $\to$ relay2 $\to$ destination in addition to the direct link source$\to$destination. Indeed, in this case $\mathsf{d}^{(\mathrm{FD})}_{N=2}$ in (3.41) can be rewritten as
$$\mathsf{d}^{(\mathrm{FD})}_{N=2} = \min\left\{\max\{\alpha_{s2}+\alpha_{1d},\beta_2+1\}, \max\{1,\min\{\alpha_{s1},\alpha_{2d}\}\}\right\}, \tag{3.43}$$

where the term $\max\{1, \min\{\alpha_{s1}, \alpha_{2d}\}\}$ in (3.43) corresponds to the gDoF of a virtual single-relay channel such that the link from the source to the "virtual relay" has SNR-exponent α_{s1} and the link from the "virtual relay" to the destination has SNR-exponent α_{2d}. We aim to determine the subset of the channel parameters $\alpha_{s2} \le \alpha_{s1},\ \alpha_{1d} < \alpha_{2d}$ for which the gDoF in (3.43) is strictly larger than the 'best relay' gDoF in (3.42). The case $\alpha_{s2} \le \alpha_{s1},\ \alpha_{1d} < \alpha_{2d}$ subsumes the following possible orders of the channel gains:

case i	α_{1d}	α_{2d}	α_{s2}			α_{s1}		
case ii	α_{1d}		α_{s2}	α_{2d}		α_{s1}		
case iii	α_{1d}		α_{s2}			α_{s1}	α_{2d}	
case iv			α_{s2}	α_{1d}	α_{2d}	α_{s1}		
case v			α_{s2}	α_{1d}		α_{s1}	α_{2d}	
case vi			α_{s2}			α_{s1}	α_{1d}	α_{2d}

We partition the set of channel parameters $\alpha_{s2} \le \alpha_{s1},\ \alpha_{1d} < \alpha_{2d}$ as follows:

- *Sub-case 2a)* (all but cases i and vi in the table): if

$$\max\{\alpha_{s2}, \alpha_{1d}\} < \min\{\alpha_{s1}, \alpha_{2d}\}, \tag{3.44}$$

then

$$\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} = \max\{1, \alpha_{s2}, \alpha_{1d}\}, \tag{3.45}$$

which is strictly less than $\mathsf{d}^{(\mathrm{FD})}_{N=2}$ in (3.43) if either

$$\max\{1, \alpha_{s2}, \alpha_{1d}\} < \min\{\alpha_{s1}, \alpha_{2d}\} \le \max\{\alpha_{s2} + \alpha_{1d}, \beta_2 + 1\}$$

or

$$\Big\{ \max\{\alpha_{s2} + \alpha_{1d}, \beta_2 + 1\} < \min\{\alpha_{s1}, \alpha_{2d}\} \Big\} \cap \mathcal{O}^c$$

where

$$\begin{aligned} \mathcal{O} :=& \{\beta_2 = 0, \alpha_{s2} + \alpha_{1d} \le 1\} \cup \{\alpha_{1d} = 0, \beta_2 + 1 \le \alpha_{s2}\} \\ & \cup \{\alpha_{s2} = 0, \beta_2 + 1 \le \alpha_{1d}\}, \end{aligned}$$

that is for

$$\max\{1, \alpha_{s2}, \alpha_{1d}\} < \min\{\alpha_{s1}, \alpha_{2d}\} \text{ except in region } \mathcal{O}.$$

- *Sub-case 2b)* (case i in the table above): if $\alpha_{1d} < \alpha_{2d} \le \alpha_{s2} \le \alpha_{s1}$, then the condition

$$\begin{aligned}\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} &= \max\{1, \alpha_{2d}\} < \mathsf{d}^{(\mathrm{FD})}_{N=2} \\ &= \min\Big\{ \max\left\{\alpha_{s2} + \alpha_{1d}, \beta_2 + 1\right\}, \max\{1, \alpha_{2d}\}\Big\}\end{aligned}$$

  is never verified, i.e., in this case $\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} = \mathsf{d}^{(\mathrm{FD})}_{N=2}$.

- *Sub-case 2c)* (case vi in the table above): if $\alpha_{s2} \le \alpha_{s1} \le \alpha_{1d} < \alpha_{2d}$, then

$$\begin{aligned}\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} &= \max\{1, \alpha_{s1}\} < \mathsf{d}^{(\mathrm{FD})}_{N=2} \\ &= \min\Big\{ \max\left\{\alpha_{s2} + \alpha_{1d}, \beta_2 + 1\right\}, \max\{1, \alpha_{s1}\}\Big\}\end{aligned}$$

  is never verified, i.e., in this case $\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} = \mathsf{d}^{(\mathrm{FD})}_{N=2}$.

To summarize, for a 2-relay network where the single-antenna relays operate in FD, using both relays gives a strictly larger gDoF compared to only exploiting the best one if

$$\max\{1, \alpha_{s2}, \alpha_{1d}\} < \min\{\alpha_{s1}, \alpha_{2d}\} \text{ except for} \tag{3.46a}$$

$$\begin{aligned}\mathcal{O} := &\{\beta_2 = 0, \alpha_{s2} + \alpha_{1d} \le 1\} \cup \{\alpha_{1d} = 0, \beta_2 + 1 \le \alpha_{s2}\} \\ &\cup \{\alpha_{s2} = 0, \beta_2 + 1 \le \alpha_{1d}\}.\end{aligned} \tag{3.46b}$$

Recall that there is also a regime similar to (3.46) where the role of the relays is swapped.

In Figure 3.3 consider the case of $\alpha_{s1} = \alpha_{2d} = x$, $\alpha_{s2} = \alpha_{1d} = y$, $\beta_1 = \beta_2 = z$ with $0 < y < x$. With these parameters, the network in Figure 3.3 satisfies the conditions in (3.44). This is an 'asymmetric' network, i.e., one relay has the best link from the source and the other relay has the best link to the destination. By exploiting both relays, the system attains

$$\begin{aligned}\mathsf{d}^{(\mathrm{FD})}_{N=2} &= \min\Big\{ \max\{1, x, y\}, \max\{2x, z+1\}, \max\{2y, z+1\}\Big\} \\ &= \min\Big\{ \max\{1, x\}, \max\{2y, z+1\}\Big\},\end{aligned}$$

while, by using only the best relay, it achieves

$$\mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}} = \max\big\{1, \min\{x, y\}\big\} = \max\big\{1, y\big\}.$$

By (3.46), we have $\mathsf{d}^{(\mathrm{FD})}_{N=2} > \mathsf{d}^{(\mathrm{FD})}_{N=2,\text{best relay}}$ if

$$x > \max\{1, y\} \text{ except for } \left\{z = 0, y \le \frac{1}{2}\right\}. \tag{3.47}$$

Note that a non-zero link between relay1 and relay2 allows to route the information through the path source→relay1→relay2→destination, which leads to an increase in terms of gDoF with respect to the best relay selection strategy.

The half-duplex case: With HD, the gDoF is given by

$$\begin{aligned}\mathsf{d}^{(\mathrm{HD})}_{N=2} = \max\min\Big\{&\lambda_0 D_1^{(0)} + \lambda_1 D_1^{(1)} + \lambda_2 D_1^{(2)} + \lambda_3 D_1^{(3)},\\ &\lambda_0 D_2^{(0)} + \lambda_1 D_2^{(1)} + \lambda_2 D_2^{(2)} + \lambda_3 D_2^{(3)},\\ &\lambda_0 D_3^{(0)} + \lambda_1 D_3^{(1)} + \lambda_2 D_3^{(2)} + \lambda_3 D_3^{(3)},\\ &\lambda_0 D_4^{(0)} + \lambda_1 D_4^{(1)} + \lambda_2 D_4^{(2)} + \lambda_3 D_4^{(3)}\Big\},\end{aligned} \tag{3.48}$$

where the maximization is over $\lambda_s, \ \forall s \in [0:1]^2$, with $\lambda_s = \mathbb{P}[S_{[1:2]} = s] \ge 0$, such that $\sum_{s\in[0:3]}\lambda_s = \lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 = 1$ [6], and

$$\begin{aligned}&D_1^{(0)} := \max\{1, \alpha_{s1}, \alpha_{s2}\}, && D_1^{(1)} = D_2^{(0)} := \max\{1, \alpha_{s1}\},\\ &D_4^{(3)} := \max\{1, \alpha_{1d}, \alpha_{2d}\}, && D_1^{(2)} = D_3^{(0)} := \max\{1, \alpha_{s2}\},\\ &D_2^{(1)} := \max\{\alpha_{s1} + \alpha_{2d}, \beta_1 + 1\}, && D_2^{(3)} = D_4^{(1)} := \max\{1, \alpha_{2d}\},\\ &D_3^{(2)} := \max\{\alpha_{s2} + \alpha_{1d}, \beta_2 + 1\}, && D_3^{(3)} = D_4^{(2)} := \max\{1, \alpha_{1d}\},\\ &D_1^{(3)} = D_2^{(2)} = D_3^{(1)} = D_4^{(0)} := 1.\end{aligned}$$

For future reference, if only one relay helps the communication between the source and the destination then the achievable gDoF is given by (2.4) in Chapter 2, which with the notation in (3.40)

$$\mathsf{d}^{(\mathrm{HD})}_{N=2,\text{best relay}} = 1 + \max_{i\in[1:2]} \frac{[\alpha_{si} - 1]^+ \ [\alpha_{id} - 1]^+}{[\alpha_{si} - 1]^+ + [\alpha_{id} - 1]^+} \in [1, \mathsf{d}^{(\mathrm{HD})}_{N=2}]. \tag{3.49}$$

An analytical closed form solution for the optimal $\{\lambda_s\}$ in (3.48) is complex to find for general channel gain assignments. However, numerically it

[6] Recall that we use interchangeably the notation $s \in [0 : 1]^N$ to index all possible binary vectors of length N, as well as, $s \in [0 : 2^N - 1]$ to indicate the decimal representation of a binary vector of length N.

is a question of solving a LP, for which efficient numerical routines exist. Moreover, see Appendix 3.B, we can set, without loss of optimality, either λ_0 or λ_3 to zero.

For the example in Figure 3.3 with $\alpha_{s1} = \alpha_{2d} = x$, $\alpha_{s2} = \alpha_{1d} = y$, $\beta_1 = \beta_2 = z$ and $0 < y < x$, the (approximately) optimal schedule has $\lambda_0 = \lambda_3 = 0$ without loss of optimality (see Appendix 3.B). By letting $\lambda_1 = \gamma \in [0,1]$ and $\lambda_2 = 1-\gamma$ (recall $0 < y < x$ without loss of generality), the gDoF in (3.48) can be written as

$$\begin{aligned}
\mathsf{d}_{N=2}^{(\mathrm{HD})} &= \max_{\gamma\in[0,1]} \min\Big\{\gamma \max\{1,x\} + (1-\gamma)\max\{1,y\}, \\
&\qquad\qquad \gamma\max\{2x, z+1\} + (1-\gamma), \\
&\qquad\qquad \gamma + (1-\gamma)\max\{2y, z+1\}\Big\} \\
&= 1 + \min\left\{ \frac{[x-1]^+ \max\{2y-1, z\}}{[x-1]^+ + \max\{2y-1,z\} - [y-1]^+}, \right. \\
&\qquad\qquad \left. \frac{\max\{2x-1,z\}\max\{2y-1,z\}}{\max\{2x-1,z\}+\max\{2y-1,z\}} \right\}. \qquad (3.50)
\end{aligned}$$

By using only the best relay as in (3.49), we would achieve

$$\mathsf{d}_{N=2,\text{best relay}}^{(\mathrm{HD})} = 1 + \frac{[x-1]^+[y-1]^+}{[x-1]^+ + [y-1]^+}. \qquad (3.51)$$

It can be easily seen that the best relay selection policy is strictly suboptimal if (3.47) is verified, as for the FD case. Considerations similar to those made for the FD case, can be made for the HD case as well. Figure 3.4 shows, for different values of z, i.e., strength of the links between the two relays, the behaviors of $\mathsf{d}_{N=2,\text{best relay}}^{(\mathrm{HD})}$ in (3.51) and $\mathsf{d}_{N=2}^{(\mathrm{HD})}$ in (3.50). Regarding the curves with $y = 0.4$, since we have $y < \frac{1}{2}$ and hence $\max\{2y-1, z\} = z, \forall z \geq 0$, $\mathsf{d}_{N=2}^{(\mathrm{HD})}$ in (3.50) is an increasing function of z. On the other hand, since $y < 1$, $\mathsf{d}_{N=2,\text{best relay}}^{(\mathrm{HD})}$ in (3.51) is always equal to 1, i.e., direct transmission is gDoF optimal. We also notice that for $z = 0$, the two curves overlap since the condition in (3.47) holds. Regarding the curves with $y = 1.2$, we notice that $\mathsf{d}_{N=2}^{(\mathrm{HD})}$ in (3.50) is always strictly greater than $\mathsf{d}_{N=2,\text{best relay}}^{(\mathrm{HD})}$ in (3.51), i.e., the channel conditions are such that the synergies between the two relays bring to an unbounded rate gain with respect to best relay selection. Moreover, $\mathsf{d}_{N=2}^{(\mathrm{HD})}$ in (3.50) starts to increase with z, when $\min\{\max\{2y-1,z\}, \max\{2x-1,z\}\} = \max\{2y-1,z\} = z$, i.e., $z = 1.4$ and best relay selection is always gDoF-wise greater than direct transmission, i.e., $\mathsf{d}_{N=2,\text{best relay}}^{(\mathrm{HD})} > 1$, since $\min\{x,y\} > 1$.

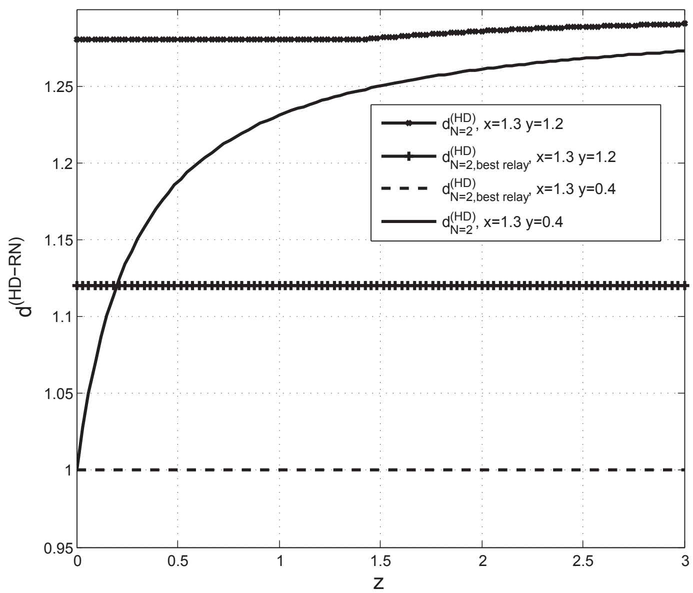


Figure 3.4: $\mathsf{d}^{(\mathrm{HD})}_{N=2}$ in (3.50) and $\mathsf{d}^{(\mathrm{HD})}_{N=2,\text{best relay}}$ in (3.51) for different values of $z \in [0, 3]$ and for $x = 1.3$, $y = 0.4, 1.2$ in Figure 3.3

3.6.2 Example 2: HD relay network with $N = 1$ relay equipped with $m_r = 2$ antennas

We consider the network in Figure 3.5, which consists of a single-antenna source (Tx), a single-antenna destination (Rx) and $N = 1$ relay (RN) equipped with $m_r = 2$ antennas. For readability, we use here a different convention for the subscripts compared to the rest of the chapter and indicate the input-output relationship as

$$\mathbf{y}_r = \begin{bmatrix}(1 - S_1)h_{rs,1} \\ (1 - S_2)h_{rs,2}\end{bmatrix} x_0 + \mathbf{z}_r, \tag{3.52a}$$

$$y_d = \begin{bmatrix}h_{ds} & h_{dr,1} & h_{dr,2}\end{bmatrix} \begin{bmatrix}x_0 \\ S_1 x_1 \\ S_2 x_2\end{bmatrix} + z_d, \tag{3.52b}$$

where: (i) x_0 and $\mathbf{x}_r = [x_1;\ x_2]$ are the signals transmitted by the source and the relay, respectively; (ii) $\mathbf{y}_r = [y_1;\ y_2]$ and y_d are the signals received at the relay and destination, respectively; (iii) $\mathbf{z}_r = [z_1;\ z_2]$ and z_d are the noises at the relay and destination, respectively; (iv) $\mathbf{s}_r = [S_1;\ S_2]$ is the state of

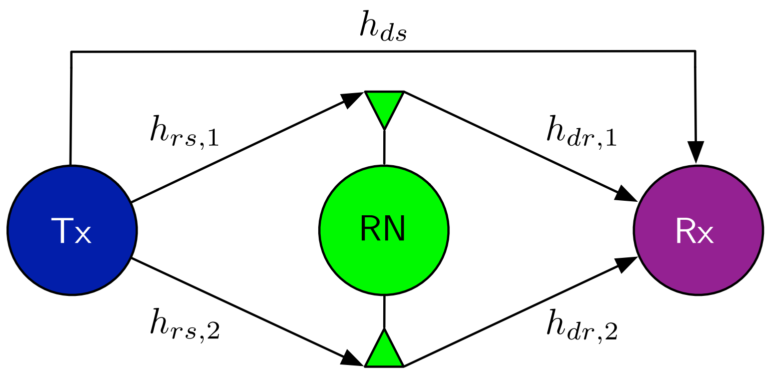


Figure 3.5: Example of network with $N = 1$ relay with $m_r = 2$ antennas, and single-antenna source and destination.

the relay antennas; (v) the inputs are subject to the power constraints

$$\mathbb{E}[|x_0|^2] = \sum_{s\in[0:1]^2} \lambda_s \mathbb{E}[|x_0|^2|\mathbf{s}_r = s] = \sum_{s\in[0:1]^2} \lambda_s P_{0|s} \leq 1, \tag{3.53a}$$

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{x}_r\|^2\right] &= \mathrm{Tr}\left[\sum_{s\in[0:1]^2} \lambda_s \mathbb{E}\left[\mathbf{x}_r\mathbf{x}_r^H|\mathbf{s}_r = s\right]\right] \\ &= \mathrm{Tr}\left[\sum_{s\in[0:1]^2} \lambda_s \begin{bmatrix} P_{1|s} & \rho_s\sqrt{P_{1|s}P_{2|s}} \\ \rho_s^*\sqrt{P_{1|s}P_{2|s}} & P_{2|s}\end{bmatrix}\right] \leq 1, \end{aligned} \tag{3.53b}$$

where $\rho_s : |\rho_s| \in [0, 1]$ is the correlation coefficient among the relay antennas in state $s \in [0 : 1]^2$ and $P_{k|s}$ is the power allocated on x_k, $k \in [0 : 2]$, in state $s \in [0 : 1]^2$.

In what follows we consider two different possible switching strategies at the relay: (i) $\mathbf{s}_r \in [0 : 1]^2$: the $m_r = 2$ antennas at the relay are switched independently of one another, and (ii) $\mathbf{s}_r = S\mathbf{1}_2 : S \in [0 : 1]$: the $m_r = 2$ antennas at the relay are used for the same purpose, either transmit or receive. We now analyze these two cases separately.

1. *Case (i): independent use of the relay antennas.* For the cut-set upper bound, two cuts must be considered, namely, $\mathcal{A} = \emptyset$ (the relay is in the cut of the source) and $\mathcal{A} = \{1\}$ (the relay is in the cut of the

destination). In this case the capacity $\mathsf{C}_{\text{case (i)}}$ is upper bounded as

$$\begin{aligned}\mathsf{C}_{\text{case (i)}} &\leq \max_{\mathbb{P}_{x_0,\mathbf{x}_r,\mathbf{s}_r}} \min\left\{I\left(x_0,\mathbf{x}_r,\mathbf{s}_r;y_d\right), I\left(x_0;y_d,\mathbf{y}_r|\mathbf{x}_r,\mathbf{s}_r\right)\right\}\\ &\leq H\left(\mathbf{s}_r\right) + \max_{\mathbb{P}_{x_0,\mathbf{x}_r}} \min\left\{I\left(x_0,\mathbf{x}_r;y_d|\mathbf{s}_r\right), I\left(x_0;y_d,\mathbf{y}_r|\mathbf{x}_r,\mathbf{s}_r\right)\right\},\end{aligned}$$

where the last inequality follows since $I\left(\mathbf{s}_r;y_d\right) \leq H\left(\mathbf{s}_r\right) \leq 2$ bits. Note that, in general, Gaussian inputs are not optimal for Gaussian networks with HD relays since useful information can be conveyed to the destination through random switch [18]. However, as seen in Remark 10, to within a constant gap a fixed switching policy between receive and transmit states is optimal, in which case a Gaussian input for each state is optimal. Moreover, the optimal choice of the correlation coefficients is $\rho_{00} = \rho_{01} = \rho_{10} = 0$ and $\rho_{11} = \mathrm{e}^{\mathrm{j}\angle(h_{dr,1}^H h_{dr,2})}$. With this we have

$$\begin{aligned}&I\left(x_0,\mathbf{x}_r;y_d|\mathbf{s}_r\right) \leq I_{\emptyset}^{(\text{fix})}\\ &:= m_d\log(2) + \lambda_0\log\left(1+|h_{ds}|^2P_{0|00}\right)\\ &+\lambda_1\log\left(1+|h_{ds}|^2P_{0|01}+|h_{dr,2}|^2P_{2|01}\right)\\ &+\lambda_2\log\left(1+|h_{ds}|^2P_{0|10}+|h_{dr,1}|^2P_{1|10}\right)\\ &+\lambda_3\log\left(1+|h_{ds}|^2P_{0|11}+\left(\sqrt{|h_{dr,1}|^2P_{1|11}}+\sqrt{|h_{dr,2}|^2P_{2|11}}\right)^2\right),\end{aligned} \tag{3.54}$$

where the term $m_d\log(2)$ (with m_d being the number of antennas at the destination) accounts for the loss of considering independent inputs at Tx and at RN. Similarly, we have

$$\begin{aligned}&I\left(x_0;y_d,\mathbf{y}_r|\mathbf{x}_r,\mathbf{s}_r\right) \leq I_{\{1\}}^{(\text{fix})}\\ &:= \lambda_0\log\left(1+(|h_{ds}|^2+|h_{rs,1}|^2+|h_{rs,2}|^2)P_{0|00}\right)\\ &+\lambda_1\log\left(1+(|h_{ds}|^2+|h_{rs,1}|^2)P_{0|01}\right)\\ &+\lambda_2\log\left(1+(|h_{ds}|^2+|h_{rs,2}|^2)P_{0|10}\right)\\ &+\lambda_3\log\left(1+|h_{ds}|^2P_{0|11}\right).\end{aligned} \tag{3.55}$$

Note that to determine the NNC achievable rate it suffices to remove the term $I\left(\mathbf{y}_r;\hat{\mathbf{y}}_r|x_0,\mathbf{x}_r,\mathbf{s}_r,y_d\right) = m_r\log(1+1/\sigma^2)$ from $I_{\emptyset}$ and the term $I\left(x_0;\mathbf{y}_r|\hat{\mathbf{y}}_r,y_d,\mathbf{x}_r,\mathbf{s}_r\right) \leq \log(1+\sigma^2)$ from $I_{\{1\}}$, with σ^2 being the variance of the quantization noise. We let $\sigma^2 = 1$ for simplicity. Note also that the expressions for $I_{\emptyset}^{(\text{fix})}$ and $I_{\{1\}}^{(\text{fix})}$ should be optimized with respect to the power allocation across the relay states, which makes the optimization problem non-linear in $\lambda_s, s \in [0:1]^{m_r}$. As pointed out in Remark 11 (see also the assumption in item 3 in Theorem 4), in order to apply Theorem 4 (see also Remark 14) we must further bound the two expressions so that to obtain a new optimization problem with constant powers across the relay states, i.e., we need to obtain a LP in $\{\lambda_s\}$. In Appendix 3.D we show $\mathsf{C}_{\text{case (i)}} \leq \mathsf{GAP} + \mathsf{C}'_{\text{case (i)}}$ where

$$
\begin{aligned}
\mathsf{C}'_{\text{case (i)}} = \max_{\lambda_s} \min \Big\{ & \lambda_0 \log\left(1+|h_{ds}|^2\right) + \lambda_1 \log\left(1+|h_{ds}|^2+|h_{dr,2}|^2\right) \\
& +\lambda_2 \log\left(1+|h_{ds}|^2+|h_{dr,1}|^2\right) \\
& +\lambda_3 \log\left(1+|h_{ds}|^2+\left(\sqrt{|h_{dr,1}|^2}+\sqrt{|h_{dr,2}|^2}\right)^2\right), \\
& \lambda_0 \log\left(1+|h_{ds}|^2+|h_{rs,1}|^2+|h_{rs,2}|^2\right) + \lambda_1 \log\left(1+|h_{ds}|^2+|h_{rs,1}|^2\right) \\
& +\lambda_2 \log\left(1+|h_{ds}|^2+|h_{rs,2}|^2\right) + \lambda_3 \log\left(1+|h_{ds}|^2\right) \Big\},
\end{aligned}
$$

and where $\mathsf{GAP} \leq 8$ bits to account for deterministic switch, independent inputs at the source and at the relay, constant power allocation across the states and NNC transmission strategy. Now, by applying Theorem 4 (see also Remark 14) $\mathsf{C}'_{\text{case (i)}}$, which can be straightforwardly cast into a LP as in (3.36), has at most $N+1=2$ active states.

2. *Case (ii): same use of the relay antennas.* In this case the $m_r = 2$ antennas at the relay are used for the same purpose so it suffices to set $\lambda_1=\lambda_2=0$ in $\mathsf{C}'_{\text{case (i)}}$ and optimize over $\lambda_0 = 1-\lambda_3 = \lambda \in [0,1]$. With this we get that $\mathsf{C}_{\text{case (ii)}} \leq \mathsf{GAP} + \mathsf{C}'_{\text{case (ii)}}$ where

$$
\mathsf{C}'_{\text{case (ii)}} = \log\left(1+|h_{ds}|^2\right) + \lambda^\star \log\left(1+\frac{|h_{rs,1}|^2+|h_{rs,2}|^2}{1+|h_{ds}|^2}\right)
$$

$$
\lambda^\star = \frac{\log\left(1+\frac{\left(\sqrt{|h_{dr,1}|^2}+\sqrt{|h_{dr,2}|^2}\right)^2}{1+|h_{ds}|^2}\right)}{\log\left(1+\frac{\left(\sqrt{|h_{dr,1}|^2}+\sqrt{|h_{dr,2}|^2}\right)^2}{1+|h_{ds}|^2}\right)+\log\left(1+\frac{|h_{rs,1}|^2+|h_{rs,2}|^2}{1+|h_{ds}|^2}\right)}
$$

and where again $\mathsf{GAP} \leq 8$ bits. The optimal $\lambda^\star$ for $\mathsf{C}'_{\text{case (ii)}}$ was found by equating the two expressions within the max min.

We now show through some simple examples that not only $\mathsf{C}'_{\text{case (i)}} \geq \mathsf{C}'_{\text{case (ii)}}$, i.e., independently switching the antennas at the relay brings achievable rate gains compared to using the antennas for the same purpose, but that the difference between the two can be *unbounded*. In other words, at high SNR $\mathsf{C}'_{\text{case (i)}}$ and $\mathsf{C}'_{\text{case (ii)}}$ have different pre-logs / multiplexing gains / degrees of freedom.

Example 1: let $|h_{ds}| = |h_{rs,2}| = |h_{dr,1}| = 0$ and $|h_{rs,1}|^2 = |h_{dr,2}|^2 = \gamma > 0$ in Figure 3.5. With this choice of the channel parameters we get

$$\begin{aligned}\mathsf{C}'_{\text{case (i)}} &= \max_{\lambda_s} \min \{\lambda_1 \log(1+\gamma) + \lambda_3 \log(1+\gamma), \\ &\quad \lambda_0 \log(1+\gamma) + \lambda_1 \log(1+\gamma)\} = \log(1+\gamma),\end{aligned}$$

where the last equality follows since the optimal choice of λ_s is given by $\lambda_0 = \lambda_2 = \lambda_3 = 0$ and $\lambda_1 = 1$, i.e., there is $1 < N+1 = 2$ active state. For $\mathsf{C}'_{\text{case (ii)}}$ the optimal λ is $1/2$ and

$$\mathsf{C}'_{\text{case (ii)}} = \frac{\log(1+\gamma)}{2}.$$

From the two expressions above not only we have $\mathsf{C}'_{\text{case (i)}} > \mathsf{C}'_{\text{case (ii)}}, \forall \gamma > 0$, but independently switching the $m_r = 2$ antennas also provides a pre-log factor that is twice of the one provided by using the antennas for the same purpose. This can be interpreted as follows. By independently switching the $m_r = 2$ antennas at the relay, the achievable rate $\mathsf{C}'_{\text{case (i)}}$ equals (to within a constant gap) the capacity of a single-antenna relay channel with a FD relay with the source-relay and relay-destination channel gains of strength equal to γ. On the other hand, by using the $m_r = 2$ antennas for the same purpose, the achievable rate $\mathsf{C}'_{\text{case (ii)}}$ reduces to the capacity of a single-antenna HD relay channel.

Example 2: let $|h_{ds}| = 0$ and $|h_{rs,1}|^2 = |h_{rs,2}|^2 = |h_{dr,1}|^2 = |h_{dr,2}|^2 = \gamma > 0$ in Figure 3.5. With this choice of the channel parameters we get

$$\begin{aligned}\mathsf{C}'_{\text{case (i)}} &= \max_{\lambda_s} \min \{\lambda_1 \log(1+\gamma) + \lambda_2 \log(1+\gamma) + \lambda_3 \log(1+4\gamma), \\ &\quad \lambda_0 \log(1+2\gamma) + \lambda_1 \log(1+\gamma) + \lambda_2 \log(1+\gamma)\} \\ &\overset{\text{(a)}}{=} \max \left\{ \log(1+\gamma), \frac{\log(1+2\gamma)\log(1+4\gamma)}{\log(1+2\gamma)+\log(1+4\gamma)} \right\} \\ &\overset{\text{(b)}}{=} \begin{cases} \log(1+\gamma) & \text{if } \gamma \geq 0.752 \\ \frac{\log(1+2\gamma)\log(1+4\gamma)}{\log(1+2\gamma)+\log(1+4\gamma)} & \text{otherwise} \end{cases}, \end{aligned} \tag{3.56}$$

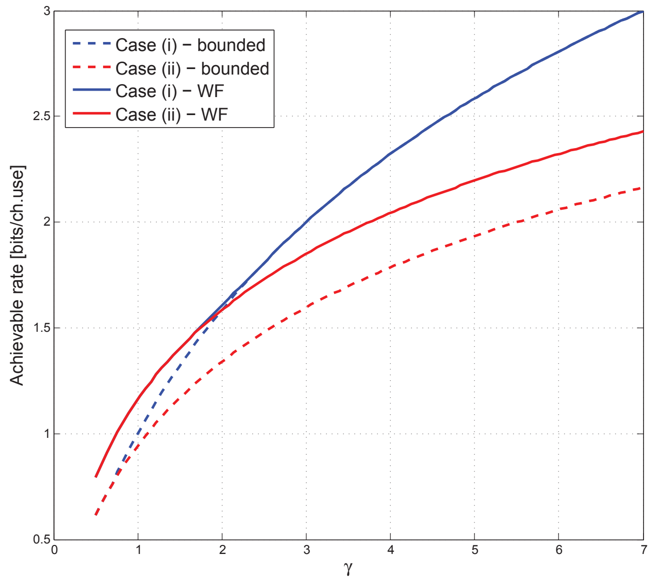


Figure 3.6: $\mathsf{C}'_{\text{case (i)}}$,$\mathsf{C}'_{\text{case (ii)}}$,$\mathsf{C}''_{\text{case (i)}}$, $\mathsf{C}''_{\text{case (ii)}}$ versus different values of γ.

where the equality in (a) follows since among the ten possible (approximately) optimal simple schedules λ_s (six possible λ_s with two active states plus four possible λ_s with one active state), it is easy to see that only the two cases $\lambda_s = [0,0,1,0]$ and $\lambda_s = [\lambda,0,0,1-\lambda]$, with $\lambda = \frac{\log(1+4\gamma)}{\log(1+2\gamma)+\log(1+4\gamma)}$, have to be considered and the equality in (b) follows from numerical evaluations. Thus, if $\gamma \geq 0.752$ the (approximately) optimal schedule has $1 < N+1 = 2$ active state (i.e., λ_2 only), otherwise it has $N+1 = 2$ active states (i.e., λ_0 and λ_3).

For $\mathsf{C}'_{\text{case (ii)}}$ we obtain that the optimal $\lambda = \frac{\log(1+4\gamma)}{\log(1+2\gamma)+\log(1+4\gamma)}$ and

$$\mathsf{C}'_{\text{case (ii)}} = \frac{\log(1+2\gamma)\log(1+4\gamma)}{\log(1+2\gamma)+\log(1+4\gamma)}. \tag{3.57}$$

It hence follows that $\mathsf{C}'_{\text{case (i)}} > \mathsf{C}'_{\text{case (ii)}}, \forall\gamma \geq 0.752$, as can also be observed from Figure 3.6 (blue dashed line for $\mathsf{C}'_{\text{case (i)}}$ versus red dashed line for $\mathsf{C}'_{\text{case (ii)}}$). Moreover, in the high-SNR regime, the pre-log factor for $\mathsf{C}'_{\text{case (i)}} = \log(1+\gamma)$ is again twice of the one of $\mathsf{C}'_{\text{case (ii)}} \approx \frac{1}{2}\log(1+\gamma)$. This example (as also Example 1) highlights the importance of smartly switching the relay antennas in order to fully exploit the available system resources. Figure 3.6 also shows the achievable rates $\mathsf{C}''_{\text{case (i)}} = \max_{\lambda_s}\min\{I^{(\text{fix})}_{\emptyset}, I^{(\text{fix})}_{\{1\}}\}$ (solid blue line) and $\mathsf{C}''_{\text{case (ii)}}$ (solid red line) obtained by optimizing the powers in $I_{\emptyset}^{(\text{fix})}$ in (3.54) and $I_{\{1\}}^{(\text{fix})}$ in (3.55) across the different states by Water Filling (WF), as described in Appendix 3.E. In particular, under the channel conditions considered in this example, from Appendix 3.E we get that the optimal power allocation can be found by solving

$$\begin{aligned} \mathsf{C}''_{\text{case (i)}} &= \max_{\lambda\in[0,1],\nu\geq 0} \left\{ \lambda \log^{+}(\gamma\nu) + \frac{1-\lambda}{2}\log^{+}(2\gamma\nu) \right\} \\ \nu :\ & \lambda\left(\nu - \frac{1}{\gamma}\right)^{+} + \frac{1-\lambda}{2}\left(\nu - \frac{1}{2\gamma}\right)^{+} = 1, \end{aligned}$$

where $\lambda_1 + \lambda_2 = \lambda \in [0,1],\ \lambda_0 = \lambda_3 = \frac{1-\lambda}{2}$, which is equal to

$$\begin{aligned} \mathsf{C}''_{\text{case (i)}} = \max_{\lambda\in[0,1]} &\left\{ \lambda \log\left(\frac{3\lambda+1}{2(\lambda+1)} + \frac{2}{\lambda+1}\gamma \right) \right. \\ &\left. + \frac{1-\lambda}{2} \log\left(\frac{3\lambda+1}{\lambda+1} + \frac{4}{\lambda+1}\gamma \right) \right\}, \end{aligned} \tag{3.58}$$

which is represented by the blue solid line in Figure 3.6. For case (ii) it suffices to set $\lambda = 0$ in $\mathsf{C}''_{\text{case (i)}}$; with this we obtain

$$\mathsf{C}''_{\text{case (ii)}} = \frac{1}{2}\log(1+4\gamma), \tag{3.59}$$

which is represented by the red solid line in Figure 3.6.

From Figure 3.6 we observe that the highest rates are achieved by optimizing the powers across the different states (solid lines versus dashed lines). However, as also highlighted in Remark 11 (see also the assumption in item 3 in Theorem 4), with optimal power allocation there are no guarantees that the (approximately) optimal schedule is simple. This is exactly what we observe in this example for which the optimal $\lambda \in [0,1]$ that maximizes $\mathsf{C}''_{\text{case (i)}}$ in (3.58) is neither zero nor one, i.e., the schedule has $3 > N+1 = 2$ active states. From Figure 3.6 we also notice that the difference between the solid lines (obtained by optimizing the powers across the states) and the dashed lines (obtained with a constant / fixed power allocation) is at most 0.1977 bits for case (i) (blue lines) and 0.2636 bits for case (ii) (red lines). These differences are far smaller than the 3 bits computed analytically in Appendix 3.D, showing that the theoretical gap of 3 bits is very conservative, at least for this choice of the channel parameters.

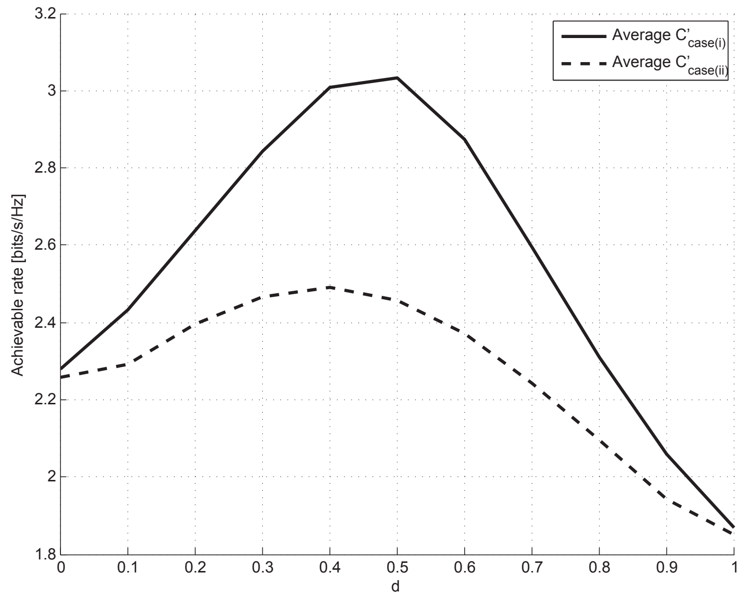


Figure 3.7: $\mathbb{E}\left[\mathsf{C}'_{\text{case (i)}}\right]$ (solid curve) and $\mathbb{E}\left[\mathsf{C}'_{\text{case (ii)}}\right]$ (dashed curve) versus different values of $d \in [0,1]$.

Example 3: we consider the case of Rayleigh fading, where $h_{ds} \sim \mathcal{N}\left(0, \sigma^2_{ds}\right)$, $h_{rs,i} \sim \mathcal{N}\left(0, \sigma^2_{rs}\right)$ and $h_{dr,i} \sim \mathcal{N}\left(0, \sigma^2_{dr}\right)$ with $i \in [1:2]$ in Figure 3.5 are assumed to be constant over the whole slot (block-fading model) and we let $\sigma^2_{ds} = \mathbb{E}\left[|h_{ds}|^2\right] = \frac{\mathsf{c}}{1^\alpha}$, $\sigma^2_{rs} = \mathbb{E}\left[|h_{rs,i}|^2\right] = \frac{\mathsf{c}}{d^\alpha}$ and $\sigma^2_{dr} = \mathbb{E}\left[|h_{dr,i}|^2\right] = \frac{\mathsf{c}}{(1-d)^\alpha}$, where c is a constant, $d \in [0,1]$ is the distance between the source and the relay and $(1-d)$ is the distance between the relay and the destination, and $\alpha \geq 2$ is the path loss exponent.

Figure 3.7 shows the average $\mathsf{C}'_{\text{case (i)}}$ (solid curve) and the average $\mathsf{C}'_{\text{case (ii)}}$ (dashed curve) versus $d \in [0,1]$, with fixed $\alpha = 3$ and $\mathsf{c} = 1$. The average was taken over $5 \cdot 10^4$ different realizations of the channel gains for each value of $d \in [0,1]$. From Figure 3.7 we observe again that in general $\mathbb{E}\left[\mathsf{C}'_{\text{case (i)}}\right] > \mathbb{E}\left[\mathsf{C}'_{\text{case (ii)}}\right]$, with a maximum difference of around 0.6 bits at $d = 0.5$. Note, in fact, that for $d = 0.5$ we have $\sigma^2_{ds} = 1$ and $\sigma^2_{rs} = \sigma^2_{dr} = 8$. Under these channel conditions, by independently switching the $m_r = 2$ antennas at the relay we (approximately) achieve the FD performance, i.e., $\mathbb{E}\left[\mathsf{C}'_{\text{case (i)}}\right] \approx \log\left(\sigma^2_{rs}\right) = 3$ bits/s/Hz, while by using the $m_r = 2$ antennas for the same purpose the rate performance reduces to the capacity of a single-antenna HD relay channel, i.e., $\mathbb{E}\left[\mathsf{C}'_{\text{case (ii)}}\right] \approx \frac{\log(32)\log(16)}{\log(32)+\log(16)} \approx 2.2$ bits/s/Hz.

3.7 Applications of Theorem 6

In this section we show that the result in Theorem 6 is an efficient tool to characterize the gDoF region for any Gaussian network whose capacity can be approximated to within a constant gap by linear combinations of $\log|\ldots|$ terms and it also represents an useful tool to solve user scheduling problems. In particular, in what follows, we analyze separately the MIMO point-to-point channel, the relay-aided Broadcast Channel (BC) and the Multiple Input Single Output (MISO) BC.

3.7.1 The MIMO point-to-point channel

The gDoF, to the best of our knowledge, has been investigated so far only for Single Input Single Output (SISO) networks with very few number of nodes; we believe that the reason is that in these cases one has only to consider equivalent MISO and Single Input Multiple Output (SIMO) channels, or to explicitly deal with determinants of matrices with small dimensions. Our result extends the gDoF analysis to any MIMO channel as we explain through some examples.

MISO and SIMO channels, i.e., the case $k = 1 \leq n$: In a MISO or SIMO channel, with channel vector $\mathbf{h} := [h_1, \ldots, h_n]$ such that $|h_i|^2 = \mathsf{SNR}^{\beta_i}, i \in [1:n]$, one trivially has

$$\log(1+\|\mathbf{h}\|^2) = \log\left(1+\sum_{i=1}^{n}\mathsf{SNR}^{\beta_i}\right) \overset{\mathsf{SNR}\gg 1}{\doteq} \log\left(\mathsf{SNR}^{\max_{i\in[1:n]}\{\beta_i\}}\right).$$

The corresponding MWBM problem has one set of vertices $\mathcal{A}_1$ consisting of $k = |\mathcal{A}_1| = 1$ node and the other set of vertices $\mathcal{A}_2$ consisting of $n = |\mathcal{A}_2| \geq 1$ nodes. The weights of the edges connecting the single vertex in $\mathcal{A}_1$ to the n vertices in $\mathcal{A}_2$ can be represented as the non-negative vector $\mathbf{B} = [\beta_1, \ldots, \beta_n]$. Clearly, the optimal $\mathsf{MWBM}(\mathbf{B}) = \max_{i\in[1:n]}\{\beta_i\}$ assigns the single vertex in $\mathcal{A}_1$ to the vertex in $\mathcal{A}_2$ that is connected to it through the edge with the maximum weight.

2×2 MIMO channels, i.e., the case $k = n = 2$: As another example from the 2-user interference channel literature, consider the cut-set sum-rate upper bound

$$\begin{aligned}
&\mathbf{H} := \begin{bmatrix} h_{13} & h_{23} \\ h_{14} & h_{24} \end{bmatrix} = \begin{bmatrix} \sqrt{\mathsf{SNR}^{\beta_{13}}}\ \mathrm{e}^{\mathrm{j}\theta_{13}} & \sqrt{\mathsf{SNR}^{\beta_{23}}}\ \mathrm{e}^{\mathrm{j}\theta_{23}} \\ \sqrt{\mathsf{SNR}^{\beta_{14}}}\ \mathrm{e}^{\mathrm{j}\theta_{14}} & \sqrt{\mathsf{SNR}^{\beta_{24}}}\ \mathrm{e}^{\mathrm{j}\theta_{24}} \end{bmatrix} \\
&\Longrightarrow \mathbf{B} := \begin{bmatrix} \beta_{13} & \beta_{23} \\ \beta_{14} & \beta_{24} \end{bmatrix}, \\
&\log\left|\mathbf{I}_2 + \mathbf{H}\mathbf{H}^H\right| \overset{\mathsf{SNR}\gg 1}{\doteq} \log\left(\mathsf{SNR}^{\max\{\beta_{13}+\beta_{24},\beta_{23}+\beta_{14}\}}\right).
\end{aligned}$$

The corresponding MWBM problem has one set of vertices $\mathcal{A}_1$ consisting of $k = |\mathcal{A}_1| = 2$ nodes (for future references let us refer to these vertices as nodes 1 and 2 – see first subscript in the channel gains) and the other set of vertices $\mathcal{A}_2$ consisting also of $n = |\mathcal{A}_2| = 2$ nodes (for future references let us refer to these vertices as nodes 3 and 4 – see second subscript in the channel gains). The weights of the edges connecting the vertices in $\mathcal{A}_1$ to the vertices in $\mathcal{A}_2$ can be represented as the non-negative weights $\beta_{ji},\ i \in [3:4],\ j \in [1:2]$. In this example, one possible matching assigns node 1 to node 3 and node 2 to node 4 (giving total weight $\beta_{13} + \beta_{24}$), while the other possible matching assigns node 2 to node 3 and node 1 to node 4 (giving total weight $\beta_{23} + \beta_{14}$); the best assignment is the one that gives the largest total weight.

Notice that the MWBM is a tight approximation of the 2×2 MIMO capacity only if the channel matrix is full rank, see [47, eq.(5) 1st line], but it is loose when the channel matrix is rank deficient, see [47, eq.(5) 2nd line, and compare with eq.(11)]. The reason is that the MWBM can not capture the impact of phases in MIMO situations. To exclude the case of a rank deficient channel matrix from our general setting for any value of k and n, we may proceed as in [93, page 2925]. Namely, we pose a reasonable distribution, such as for example the i.i.d. uniform distribution, on the phases $\theta_{ji},\ i \in [3:4],\ j \in [1:2]$, so that almost surely the channel matrix is full rank.

3.7.2 The relay-aided BC

The relay-aided BC consists of one source communicating with K destinations with the help of L FD relays. The cut-set outer bound on the capacity region of such a network was shown to be achievable to within $O\left(N\log(N)\right)$ bits, where $N = K + L + 1$ is the total number of nodes [94]. This constant gap result implies the exact knowledge of the gDoF region. As an example of application of Theorem 6, we next show how to derive the sum-gDoF of the relay-aided BC.

Consider a relay-aided BC with one source, K destinations, and $L = 1$ relay (the result can be straightforwardly extended to the case of multiple relays, of cooperation among destinations and with generalized feedback at the source). The source has input X_0, the relay has input X_1, the k-th destination has output

$$Y_k = \sqrt{\mathsf{SNR}^{\beta_{k,0}}} X_0 + \sqrt{\mathsf{SNR}^{\beta_{k,1}}}\ \mathrm{e}^{\mathrm{j}\theta_{k,1}} X_1 + Z_k,\ k \in [1:K],$$

and the relay has output

$$Y_{\mathrm{R}} = \sqrt{\mathsf{SNR}^{\beta_{\mathrm{R}}}} X_0 + Z_{\mathrm{R}},$$

where, since the channel is known to all nodes, each receiving node compensates for the phase of the link from the source. We assume that the phases $\{\theta_{k,1},\ k \in [1:K]\}$ are such that all the involved channel (sub)matrices are full rank almost surely. Without loss of generality, we let

$$\beta_{1,0} = \max_{k \in [1:K]} \{\beta_{k,0}\},$$

i.e., destination 1 has the strongest link from the source. We define the gDoF of destination k as $\mathsf{d}_k = \lim_{\mathsf{SNR}\to\infty} \frac{R_k}{\log(1+\mathsf{SNR})}$, $k \in [1:K]$. The capacity region of this relay-aided BC is to within a constant gap from the cut-set upper bound [94]. The cut-set outer bound yields for all $\mathcal{A} \subseteq [1:K]$, $\mathcal{A} \neq \emptyset$,

$$\sum_{k \in \mathcal{A}} R_k \leq I(X_0, X_1; Y_{\mathcal{A}}), \tag{3.60a}$$

$$\sum_{k \in \mathcal{A}} R_k \leq I(X_0; Y_{\mathcal{A}}, Y_{\mathrm{R}} | X_1). \tag{3.60b}$$

The sum gDoF (and similarly for any other bounds) is the minimum of two terms: the first term from (3.60a) with $\mathcal{A} = [1:K]$ is

$$\sum_{k=1}^{K} \mathsf{d}_k \leq \mathsf{MWBM}\left(\begin{bmatrix} \beta_{1,0} & \beta_{1,1} \\ \vdots & \vdots \\ \beta_{K,0} & \beta_{K,1} \end{bmatrix}\right) = \max_{j \in [2:K]} \{\beta_{1,0} + \beta_{j,1}, \beta_{j,0} + \beta_{1,1}\}, \tag{3.61a}$$

and the second term from (3.60b) with $\mathcal{A} = [1:K]$ is

$$\sum_{k=1}^{K} \mathsf{d}_k \leq \mathsf{MWBM}\left(\begin{bmatrix} \beta_{1,0} & \dots & \beta_{K,0} & \beta_{\mathrm{R}} \end{bmatrix}\right) = \max\{\beta_{1,0}, \beta_{\mathrm{R}}\}, \tag{3.61b}$$

from the assumption $\beta_{1,0} \geq \beta_{k,0},\ k \in [2:K]$. The closed-form expression for the gDoF in (3.61) sheds light into approximately optimal achievable schemes: if $\beta_{\mathrm{R}} \leq \beta_{1,0} = \max_{k\in[1:K]}\{\beta_{k,0}\}$ the sum-gDoF is as for the BC without a relay (i.e., in practical wireless broadcast networks it might not be worth using a relay if the source-relay link is weaker than the strongest source-destination link), while if $\beta_{1,0} < \beta_{\mathrm{R}}$ it is sum-gDoF optimal to serve at most one extra destination in addition to destination 1 (see eq.(3.61a)). With L relays, it is sum-gDoF optimal to serve at most $L+1$ destinations; which subset of destinations to serve can be found by examining the 2^L MWBM-based bounds, in the spirit of (3.61). This simple example shows that the result in Theorem 6 represents an useful tool to solve user scheduling problems, i.e., to understand which is the best subset of $L+1$ destinations which has to be served.

3.7.3 The MISO K-user BC

The static K-user MISO BC consists of one source equipped with N antennas and K single-antenna destinations. The source has an independent message for each of the K destinations. The input-output relationship reads

$$y_k = \mathbf{h}_k\mathbf{x} + z_k, \quad k \in [1:K], \tag{3.62}$$

where the input $\mathbf{x} \in \mathbb{C}^{N\times 1}$ is constrained to satisfy the average power constraint $\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right]\right) \leq 1$ (a non-unitary power constraint can be incorporated into the channel gains), the vector $\mathbf{h}_k \in \mathbb{C}^{1\times N}$ contains the channel gain coefficients from each transmit antenna at the base station to the k-th user and z_k is the zero-mean unit-variance proper-complex white Gaussian noise. We assume $N < K$, i.e., the number of transmit antennas at the base station is strictly smaller than the number of users.

The sum-capacity or throughput for the K-user MISO BC is given by the "Sato's cooperative upper bound with least favorable noise correlation" [95]

$$C(\mathbf{H}) = \min_{\substack{0\preceq\mathbf{S}_z:\\ [\mathbf{S}_z]_{k,k}=1,\\ \forall k\in[1:K]}} \max_{\substack{0\preceq\mathbf{S}_x:\\ \mathrm{Tr}(\mathbf{S}_x)\leq 1}} I\left(\mathbf{x};\mathbf{H}\mathbf{x}+\mathbf{z}\right), \tag{3.63}$$

where $\mathbf{H} = [\mathbf{h}_1;\ldots;\mathbf{h}_K] \in \mathbb{C}^{K\times N}$ is the overall channel matrix and $\mathbf{z} = [z_1;\ldots;z_K]$ is the overall noise vector with covariance matrix $\mathbf{S}_z = \mathbb{E}\left[\mathbf{z}\mathbf{z}^H\right] \in \mathbb{C}^{K\times K}$. By exploiting the Multiple Access Channel (MAC)-BC duality [96–98], the sum-capacity in (3.63) can be equally obtained by solving $C(\mathbf{H}) = \max_{\mathbf{D}\in\mathcal{D}} \log\left|\mathbf{I}_N + \mathbf{H}^H\mathbf{D}\mathbf{H}\right|$, where $\mathcal{D}$ is the set of $K\times K$ non-negative diagonal matrices $\mathbf{D}$ with $\mathrm{Tr}(\mathbf{D})\leq 1$.

We now show that the result in Theorem 6 allows us to find the sum-gDoF for the K-user MISO BC and inspires a user selection algorithm which outputs a set of N users, out of the K possible ones, which has to be served (recall that we are assuming $N < K$).

For some $\mathsf{SNR} > 0$ we parameterize the channel gains as $|h_{k,n}|^2 = \mathsf{SNR}^{\beta_{k,n}}, \beta_{k,n} \geq 0$, for all $k \in [1:K]$ and $n \in [1:N]$ and assume that the phases of the fading channel gains are such that all involved channel (sub)matrices are full rank almost surely. The sum-gDoF, as a function of $\{\beta_{k,n}\}$, is defined as $\mathsf{d} := \lim_{\mathsf{SNR}\to\infty} \frac{C(\mathbf{H})}{\log(1+\mathsf{SNR})}$.

Since the constraint $\mathrm{Tr}(\mathbf{D}) \leq 1$ implies $\mathbf{D} \preceq \mathbf{I}_K$, we have that the sum-capacity is upper bounded by

$$C(\mathbf{H}) \leq \log\left|\mathbf{I}_K + \mathbf{H}\mathbf{H}^H\right| = \log\left|\mathbf{I}_N + \mathbf{H}^H\mathbf{H}\right|. \tag{3.64}$$

By applying Theorem 6 to the RHS of (3.64) we immediately find that the sum-gDoF is upper bounded by

$$\mathsf{d} \leq \mathsf{MWBM}(\mathbf{B}), \ \mathbf{B} = \begin{bmatrix} \beta_{1,1} & \dots & \beta_{1,N} \\ \vdots & \ddots & \vdots \\ \beta_{K,1} & \dots & \beta_{K,N} \end{bmatrix}, \tag{3.65}$$

that amounts to solve a MWBM problem with weight matrix $\mathbf{B}$ given by the SNR-exponents $\{\beta_{k,n},\ k \in [1:K],\ n \in [1:N]\}$.

To gain insights into the result in (3.65), we next consider the case $N = 2$ (the result can be straightforwardly extended to a general N). Without loss of generality, let the antennas and the users be numbered in such a way that

$$|h_{1,1}| \geq \max_{k\in[1:K],\ n\in[1:N]} |h_{k,n}| \iff \beta_{1,1} \geq \max_{k\in[1:K],\ n\in[1:N]} \beta_{k,n}, \tag{3.66}$$

i.e., the link from antenna 1 to user 1 is the strongest among all links to any user from any of the antennas; then, by using (3.66) in (3.65), it is easy to see that

$$\begin{aligned} \mathsf{d}^{(N=2)} &\leq \mathsf{MWBM}\begin{bmatrix} \beta_{1,1} & \beta_{1,2} \\ \vdots & \vdots \\ \beta_{K,1} & \beta_{K,2} \end{bmatrix} & (3.67) \\ &= \mathsf{MWBM}\begin{bmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{k^*,1} & \beta_{k^*,2} \end{bmatrix}, & (3.68) \\ k^* &:= \arg\max_{k\in[2:K]} \{\beta_{1,1} + \beta_{k,2}, \beta_{1,2} + \beta_{k,1}\}, & (3.69) \end{aligned}$$

or in other words, destinations 1 and k^* form the best set of $N = 2$ users to be served in order to attain the gDoF upper bound in (3.67). Let $\mathbf{H}_\pi \in \mathbb{C}^{2\times 2}$ be the channel matrix that contains the channel gains of user 1 and user k^*. Since the constraint $\mathrm{Tr}\,(\mathbf{D}) \leq 1$ allows $\mathbf{D} = \frac{1}{2}\mathbf{I}_2$ (by allocating equal power among users 1 and k^*), we have that the sum-capacity is lower bounded by

$$C(\mathbf{H}) \geq \log\left|\mathbf{I}_2 + \frac{1}{2}\mathbf{H}_\pi \mathbf{H}_\pi^H\right| \geq \log\left|\mathbf{I}_2 + \mathbf{H}_\pi^H \mathbf{H}_\pi\right| - 2\log(2).$$

By applying Theorem 6 to the RHS of the above equation we immediately find that the sum-gDoF $\mathsf{d}^{(N=2)}$ is lower bounded by (3.68). This implies that the sum-gDoF $\mathsf{d}^{(N=2)}$ is given by (3.68) (the upper and lower bounds coincide). Thus, from (3.69) it is easy to see that sum-gDoF-wise just serving $N = 2$ users, among the possible K, is optimal. Moreover, it is simple to understand which $N = 2$ users have to be served: user 1, i.e., the user who has the strongest link from the source, has to be always served, and the "second best" user is the one defined in (3.69).

By extending the above reasoning to any N and K, it is straightforward to prove that the solution of the MWBM problem in (3.65), which outputs the $N = \min\{N, K\}$ users to be scheduled, represents the sum-gDoF of the K-user MISO BC as long as the channel matrix is full rank. One appealing feature of the proposed algorithm is that it runs in polynomial time.

We now numerically assess the performance of the MWBM-based algorithm for different values of N and K. We consider the case of Rayleigh fading, where $h_{k,n} \sim \mathcal{N}\left(0, \sigma_k^2\right)$, $k \in [1:K]$ and $n \in [1:N]$, is assumed to be constant over the whole slot (block-fading model), i.e.,

$$h_{k,n} = \sigma_k \; g_{k,n}, \tag{3.70}$$

where $g_{k,n} \sim \mathcal{N}(0, 1)$. We define

$$\sigma_k^2 = \mathbb{E}[|h_{k,n}|^2] = \frac{\mathsf{c}}{d_k^\alpha}, \tag{3.71}$$

where c is a constant that depends on the model parameters (e.g., base station's transmit power), d_k is the distance of the k-th user from the base station and $\alpha \geq 2$ is the path loss exponent. We assume a short-term average power constraint on the inputs. With this model, we start by considering a *dynamic scheduling* that depends on $|h_{k,n}|^2$ (note that our proposed algorithm does not make use of phase information), which later on will be compared to a *static scheduling* based on $\mathbb{E}[|h_{k,n}|^2]$ only. We set

$$\beta_{k,n} = 10\log_{10}\left(|h_{k,n}|^2\right). \tag{3.72}$$

Note that multiplying the weight matrix $\mathbf{B}$ in (3.65) by a constant and/or adding a constant to each matrix entry does not change the nature of the matching in the MWBM problem. We assume that the K users are independently uniformly distributed on an annulus with minimum radius equal to $r_{\min}$ and maximum radius equal to $r_{\max}$. Moreover, we consider that the model parameters are such that the average SNR at the cell edges is $\mathsf{SNR}\,(r_{\max})$, that is, the average SNR at distance d_k is given by

$$\mathsf{SNR}\,(d_k) := \mathbb{E}\left[\sum_{n=1}^{N} |h_{k,n}|^2 \middle| d_k\right] = \frac{N\mathsf{c}}{d_k^\alpha} = \mathsf{SNR}\,(r_{\max}) \left(\frac{d_k}{r_{\max}}\right)^{-\alpha}.$$

Let (X, Y) be the coordinate of the random position of a user; then the cumulative density function (cdf) of its position is

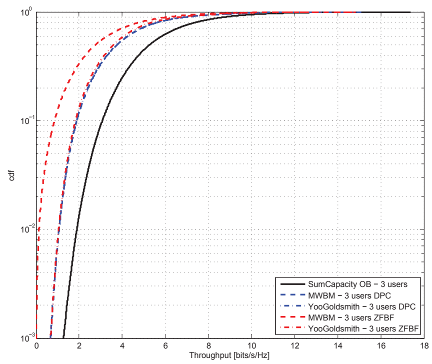
$$F_{\frac{d}{r_{\max}}}(x) = \mathbb{P}\left[\sqrt{X^2+Y^2} \leq x\ r_{\max}\right] = \frac{\left[\min(1,x^2) - \frac{r_{\min}^2}{r_{\max}^2}\right]^+}{1-\frac{r_{\min}^2}{r_{\max}^2}}, \tag{3.73}$$

for $x \geq 0$ and hence the probability density function (pdf) is

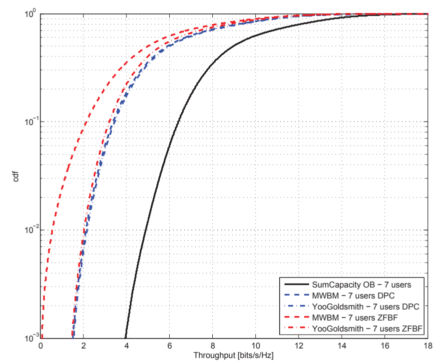
$$f_{\frac{d}{r_{\max}}}(x) = \frac{2x}{1-\frac{r_{\min}^2}{r_{\max}^2}} \text{ for } x \in \left[\frac{r_{\min}}{r_{\max}}, 1\right]. \tag{3.74}$$

Figure 3.8 shows the cdf of the throughput for different values of N and K, with fixed $\alpha = 3$, $\mathsf{SNR}\,(r_{\min}) = 40$ dB, and $\mathsf{SNR}\,(r_{\max}) = 0$ dB. The cfd was estimated with MATLAB command ecdf with a confidence level of 0.05 (default value) whose input was generated by considering $N_{\text{iter},1} = 100$ different user positions (i.e., for each $k \in [1:K]$ we consider $N_{\text{iter},1} = 100$ different values of d_k in (3.71)), for each of which we considered $N_{\text{iter},2} = K \cdot 10^3$ different realizations of $g_{k,n}$ in (3.70), $k \in [1:K], n \in [1:N]$. In Figure 3.8, the average throughput $\mathbb{E}\left[C(\mathbf{H}_\pi)\right]$ achieved by our MWBM-based algorithm is also reported for all values of K and N.

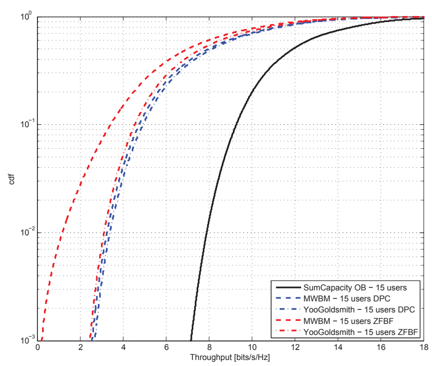
From Figure 3.8, we observe that the throughput performance of our MWBM-based algorithm is very close to the one of [99] when DPC is used in both cases (blue dashed lines versus dash-dotted lines). Differently from [99], our scheduling algorithm does not use the knowledge of the channel phases. This means that, in a practical scenario, less information has to be fed back to the base station for the purpose of scheduling users. Once our MWBM-based algorithm has selected the N users to serve, only the channel phases of the N selected users need to be fed back to the base station



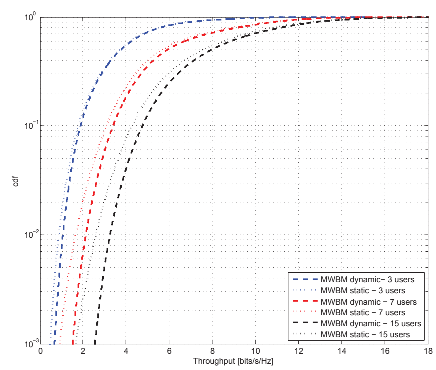
(a) $K = 3, N = 2, \mathbb{E}\left[C(\mathbf{H}_\pi)\right] = 4.1840$ bits/s/Hz.



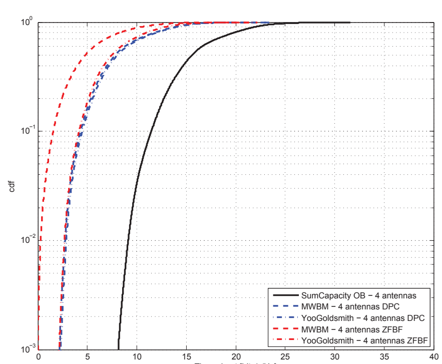
(b) $K = 7, N = 2, \mathbb{E}\left[C(\mathbf{H}_\pi)\right] = 6.6346$ bits/s/Hz.



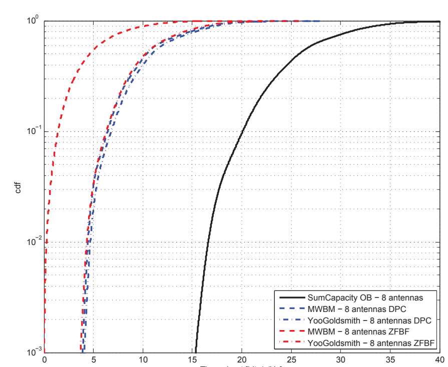
(c) $K = 15, N = 2, \mathbb{E}\left[C(\mathbf{H}_\pi)\right] = 8.4949$ bits/s/Hz.



(d) $K \in \{3, 7, 15\}, N = 2$.



(e) $K = 10, N = 4, \mathbb{E}\left[C(\mathbf{H}_\pi)\right] = 8.5260$ bits/s/Hz.



(f) $K = 10, N = 8, \mathbb{E}\left[C(\mathbf{H}_\pi)\right] = 11.1568$ bits/s/Hz.

Figure 3.8: cdf of the throughput with $N \in \{2, 4, 8\}$ and $K \in \{3, 7, 15\}$.

in order to implement the DPC strategy. In other words, given a fixed amount of bits on the feedback link, the base station can get a more accurate representation of the phases of the N selected users, as opposed to [99] that requires phases from all the K users. From Figure 3.8 we also observe that, if Zero Forcing BeamForming (ZFBF) is used instead of DPC, our algorithm does not perform as well as the one in [99] (red dashed lines versus dash-dotted lines). This is because ZFBF is most effective when the selected users have nearly orthogonal channel vectors. Hence, when ZFBF is used, it becomes essential to schedule those users whose channel gains are as orthogonal as possible. Thus, the knowledge of the channel phases becomes critical. Our MWBM-based algorithm, which is based on a "coarse" approximation of the channel gains (since only the magnitude of the channel gains is considered while the phases are neglected), does not capture this aspect. This appears to cost in performance at low-SNR if ZFBF is employed. Indeed we expect our MWBM-based algorithm to be nearly optimal at high-SNR, where the phases become negligible; in the simulated scenario the average SNR, averaged over the random positions of the users, is

$$\mathbb{E}\left[\mathsf{SNR}\left(d\right)\right] = \mathsf{SNR}\left(r_{\max}\right) \int_{\frac{r_{\min}}{r_{\max}}}^{1} \frac{2x^{1-\alpha}}{1-\frac{r_{\min}^2}{r_{\max}^2}} \mathrm{d}x = 16.1481 \text{ dB},$$

which is far from being in the high-SNR regime, thus explaining the better performance of [99] if ZFBF is used.

Figure 3.8(d) shows that the throughput increases when the number of users increases for a fixed value of N. This is due to *multiuser diversity*: as K increases for a fixed N, the base station has a larger pool of users to choose from and it is therefore more likely to find a subset of users with 'good' channels thereby attaining a larger throughput. Figure 3.8(d) also shows the throughput performance of our MWBM-based algorithm when a *static scheduling* is performed, i.e., a schedule which is based only on the fading expected value. We observe that the *dynamic scheduling* (dashed lines) outperforms the *static scheduling* (dotted lines), since the former is adapted to each instantaneous channel realization. Figure 3.8(e) and Figure 3.8(f) show that the throughput increases when the number of antennas increases for a fixed K. This is due to the *multiplexing gain*: for a fixed K, as the number of transmit antennas increases (always considering $N < K$), more users can be served leading to a throughput's boost.

Finally, we remark that in Figure 3.8 the black curves represent the sum-capacity outer bound in (3.64) and not the exact sum-capacity in (3.63).

Numerically, we notice that the gap between the black curves (outer bounds to the sum-capacity) and the achievable throughputs grows with K and N.

3.8 Conclusions and future directions

In this chapter we analyzed a network where a source communicates with a destination and is assisted by N relays operating in HD mode. For such networks, the capacity achieving scheme must be optimized over the 2^N possible listen-transmit relay configurations. We first characterized the capacity of the Gaussian noise network to within a constant gap by using NNC as achievable scheme and we proved that the gDoF is the solution of a LP, where the coefficients of the linear inequality constraints are the solution of several LPs referred to as the MWBM problem in graph theory. More generally, we showed that the high-SNR approximation of several practically relevant Gaussian networks, such as the MIMO point-to-point channel, the MISO BC and the relay-aided BC, can be found by solving several MWBM problems. We then proved that, if the noises are independent and independent inputs are approximately optimal in the cut-set bound, then the approximately optimal schedule is simple in the sense that at most $N+1$ relay configurations have a non-zero probability. Finally we showed how these results generalize to the case of multi-antenna nodes, where the antennas at the relays can be switched between listen and transmit state independently of one another. We also analyzed two network examples; for the first scenario with $N = 2$ single-antenna relays, we showed under which channel conditions by exploiting both relays a strictly greater gDoF can be attained compared to a network where best-relay selection is used; for the second scenario with $N = 1$ relay equipped with 2 antennas, we showed that independently switching the antennas at the relay can provide a strictly larger multiplexing gain compared to using the antennas for the same purpose.

With respect to the results presented in this chapter, interesting future research directions may include: (i) understanding which are the $N+1$ states with a strictly positive probability and (ii) determining necessary conditions for optimality of simple schedules.

Appendix

3.A Proof that $I_{\mathcal{A}}^{(\mathrm{fix})}$ in (3.9) is submodular

Consider two possible cuts of the network represented by $\mathcal{A}_1, \mathcal{A}_2 \subseteq [1:N]$ and let

$$\begin{aligned} &B_0 := \mathcal{A}_1 \cap \mathcal{A}_2, && B_1 := \mathcal{A}_1 \backslash \mathcal{A}_2, \\ &B_2 := \mathcal{A}_2 \backslash \mathcal{A}_1, && B_3 := [1:N] \backslash (\mathcal{A}_1 \cup \mathcal{A}_2), \end{aligned}$$

so that, $B_j, j \in [0:3]$ is a partition of $[1:N]$ and thus

$$\begin{aligned} &\mathcal{A}_1 = B_0 \cup B_1, && \mathcal{A}_2 = B_0 \cup B_2, \\ &\mathcal{A}_1 \cap \mathcal{A}_2 = B_0, && [1:N] \backslash (\mathcal{A}_1 \cup \mathcal{A}_2) = B_3. \end{aligned}$$

Let $X_{\mathcal{A}} := \{X_i : i \in \mathcal{A}\}$ and $X_{(n)} := \{X_i : i \in B_n\}$, $n \in [0:3]$. We write $I_{\mathcal{A}}^{(\mathrm{fix})} = H\left(Y_{N+1}, Y_{\mathcal{A}} | X_{\mathcal{A}}, S_{[1:N]}\right) - H\left(Y_{N+1}, Y_{\mathcal{A}} | X_{[1:N+1]}, S_{[1:N]}\right)$. We next show that, under the assumption of "independent noises" in (3.13), the function $h_1(\mathcal{A}) := H\left(Y_{N+1}, Y_{\mathcal{A}} | X_{[1:N+1]}, S_{[1:N]}\right)$ is modular and that, under the assumption of independent inputs in (3.7), the function $h_2(\mathcal{A}) := H\left(Y_{N+1}, Y_{\mathcal{A}} | X_{\mathcal{A}}, S_{[1:N]}\right)$ is submodular; these two facts imply that $I_{\mathcal{A}}^{(\mathrm{fix})}$ in (3.9) is submodular.

For $h_1(\mathcal{A})$ we have

$$\begin{aligned} &h_1(\mathcal{A}_1) + h_1(\mathcal{A}_2) - h_1(\mathcal{A}_1 \cup \mathcal{A}_2) - h_1(\mathcal{A}_1 \cap \mathcal{A}_2) \\ &= H\left(Y_{N+1}, Y_{(0)}, Y_{(1)} | X_{[1:N+1]}, S_{[1:N]}\right) + H\left(Y_{N+1}, Y_{(0)}, Y_{(2)} | X_{[1:N+1]}, S_{[1:N]}\right) \\ &\quad - H\left(Y_{N+1}, Y_{(0)}, Y_{(1)}, Y_{(2)} | X_{[1:N+1]}, S_{[1:N]}\right) - H\left(Y_{N+1}, Y_{(0)} | X_{[1:N+1]}, S_{[1:N]}\right) \\ &= H\left(Y_{(1)} | Y_{N+1}, Y_{(0)}, X_{[1:N+1]}, S_{[1:N]}\right) + H\left(Y_{(2)} | Y_{N+1}, Y_{(0)}, X_{[1:N+1]}, S_{[1:N]}\right) \\ &\quad - H\left(Y_{(1)}, Y_{(2)} | Y_{N+1}, Y_{(0)}, X_{[1:N+1]}, S_{[1:N]}\right) \\ &= I\left(Y_{(1)}; Y_{(2)} | Y_{N+1}, Y_{(0)}, X_{[1:N+1]}, S_{[1:N]}\right) = 0, \end{aligned}$$

where the last equality follows because of the assumption of "independent noises" in (3.13). Therefore $h_1(\mathcal{A})$ is modular.

For $h_2(\mathcal{A})$ we have

$$
\begin{aligned}
& h_2(\mathcal{A}_1) + h_2(\mathcal{A}_2) - h_2(\mathcal{A}_1 \cup \mathcal{A}_2) - h_2(\mathcal{A}_1 \cap \mathcal{A}_2) \\
&= H\left(Y_{N+1}, Y_{(0)}, Y_{(1)} | X_{(0)}, X_{(1)}, S_{[1:N]}\right) \\
&\quad + H(Y_{N+1}, Y_{(0)}, Y_{(2)} | X_{(0)}, X_{(2)}, S_{[1:N]}) \\
&\quad - H(Y_{N+1}, Y_{(0)}, Y_{(1)}, Y_{(2)} | X_{(0)}, X_{(1)}, X_{(2)}, S_{[1:N]}) \\
&\quad - H(Y_{N+1}, Y_{(0)} | X_{(0)}, S_{[1:N]}) \\
&= H(Y_{N+1}, Y_{(0)} | X_{(1)}, S_{[1:N]}, X_{(0)}) + H(Y_{N+1}, Y_{(0)} | X_{(2)}, S_{[1:N]}, X_{(0)}) \\
&\quad - H(Y_{N+1}, Y_{(0)} | X_{(1)}, X_{(2)}, S_{[1:N]}, X_{(0)}) - H(Y_{N+1}, Y_{(0)} | S_{[1:N]}, X_{(0)}) \\
&\quad + H(Y_{(1)} | X_{(1)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&\quad + H(Y_{(2)} | X_{(2)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&\quad - H(Y_{(1)}, Y_{(2)} | X_{(1)}, X_{(2)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&= I(Y_{N+1}, Y_{(0)}; X_{(2)} | X_{(1)}, S_{[1:N]}, X_{(0)}) - I(Y_{N+1}, Y_{(0)}; X_{(2)} | S_{[1:N]}, X_{(0)}) \\
&\quad + I(Y_{(1)}; X_{(2)} | X_{(1)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&\quad + I(Y_{(2)}; Y_{(1)}, X_{(1)} | X_{(2)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&= I(X_{(1)}; X_{(2)} | S_{[1:N]}, X_{(0)}, Y_{N+1}, Y_{(0)}) \\
&\quad + I(Y_{(1)}; X_{(2)} | X_{(1)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&\quad - I(X_{(1)}; X_{(2)} | S_{[1:N]}, X_{(0)}) \\
&\quad + I(Y_{(2)}; Y_{(1)}, X_{(1)} | X_{(2)}, S_{[1:N]}, Y_{N+1}, X_{(0)}, Y_{(0)}) \\
&\quad \geq 0,
\end{aligned}
$$

where the last inequality follows because the "independent inputs" assumption in (3.7) implies $I(X_{(1)}; X_{(2)} | S_{[1:N]}, X_{(0)}) = 0$. This shows that $h_2(\mathcal{A})$ is submodular.

3.B (Approximately) Optimal simple schedule for $N = 2$.

In a HD relay network with $N = 2$, we have $2^N = 4$ possible states that may arise with probabilities λ_s, $\forall s \in [0:3]$, with $\lambda_s = \mathbb{P}[S_{[1:2]} = s] \geq 0$, such that $\sum_{s\in[0:3]} \lambda_s = \lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 = 1$. Here we aim to demonstrate that a schedule with $\lambda_0\lambda_3 = 0$ is optimal.

Consider the following LP

$$\max_{\lambda' \mathrm{s}} \min \left\{ \begin{bmatrix} \max\{a_1, a_2\} + D_1 & a_2 & a_1 & 0 \\ a_2 & a_2 + b_1 + D_2 & 0 & b_1 \\ a_1 & 0 & a_1 + b_2 + D_3 & b_2 \\ 0 & b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_2 \\ \lambda_1 \\ \lambda_3 \end{bmatrix} \right\}, \tag{3.75}$$

where the different quantities $(a_u, b_u),\ u \in [1:2]$, and $D_v,\ v \in [1:4]$, are non-negative and will be defined later.

The proof is by contradiction. Assume that $\begin{bmatrix} \hat{\lambda}_0 & \hat{\lambda}_1 & \hat{\lambda}_2 & \hat{\lambda}_3 \end{bmatrix}$ is the optimal solution with $\hat{\lambda}_0 > 0$. This implies that for any $(\alpha, \beta, \gamma) \in [0,1]^3$ such that $\alpha + \beta + \gamma = 1$ we must have that

$$\min \left\{ \begin{bmatrix} \max\{a_1, a_2\} + D_1 & a_2 & a_1 & 0 \\ a_2 & a_2 + b_1 + D_2 & 0 & b_1 \\ a_1 & 0 & a_1 + b_2 + D_3 & b_2 \\ 0 & b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \hat{\lambda}_0 \\ \hat{\lambda}_2 \\ \hat{\lambda}_1 \\ \hat{\lambda}_3 \end{bmatrix} \right\}$$
$$\geq \min \left\{ \begin{bmatrix} \max\{a_1, a_2\} + D_1 & a_2 & a_1 & 0 \\ a_2 & a_2 + b_1 + D_2 & 0 & b_1 \\ a_1 & 0 & a_1 + b_2 + D_3 & b_2 \\ 0 & b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} 0 \\ \hat{\lambda}_2 + \hat{\lambda}_0 \alpha \\ \hat{\lambda}_1 + \hat{\lambda}_0 \beta \\ \hat{\lambda}_3 + \hat{\lambda}_0 \gamma \end{bmatrix} \right\}$$

holds. Since $\hat{\lambda}_0 > 0$ by assumption, we can rewrite the above problem as

$$0 = \min \left\{ \begin{bmatrix} \max\{a_1, a_2\} + D_1 & a_2 & a_1 & 0 \\ a_2 & a_2 + b_1 + D_2 & 0 & b_1 \\ a_1 & 0 & a_1 + b_2 + D_3 & b_2 \\ 0 & b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$
$$\geq \min \left\{ \begin{bmatrix} \max\{a_1, a_2\} + D_1 & a_2 & a_1 & 0 \\ a_2 & a_2 + b_1 + D_2 & 0 & b_1 \\ a_1 & 0 & a_1 + b_2 + D_3 & b_2 \\ 0 & b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} 0 \\ \alpha \\ \beta \\ \gamma \end{bmatrix} \right\}$$
$$= \min \left\{ \begin{bmatrix} a_2 & a_1 & 0 \\ a_2 + b_1 + D_2 & 0 & b_1 \\ 0 & a_1 + b_2 + D_3 & b_2 \\ b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \right\}$$

for all $(\alpha, \beta, \gamma) \in [0,1]^3$ such that $\alpha + \beta + \gamma = 1$. If we can find a triplet $(\alpha, \beta, \gamma) \in [0,1]^3 : \alpha + \beta + \gamma = 1$ for which

$$\min \left\{ \begin{bmatrix} a_2 & a_1 & 0 \\ a_2 + b_1 + D_2 & 0 & b_1 \\ 0 & a_1 + b_2 + D_3 & b_2 \\ b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \right\} > 0$$

holds, we reach a contradiction; hence, for this set of values we must have $\hat{\lambda}_0 = 0$. Assume $b_1 b_2 \le a_1 a_2$ and define

$$\alpha = \frac{b_2}{a_2 + b_2}, \ \beta = \frac{b_1}{a_1 + b_1},$$
$$\gamma = 1 - \alpha - \beta = \frac{a_1 a_2 - b_1 b_2}{(a_1 + b_1)(a_2 + b_2)},$$

which is a valid assignment since all coefficients are non-negative and sum to one. With this we have that

$$\max_{(\alpha,\beta,\gamma)\in[0,1]^3:\alpha+\beta+\gamma=1} \min\left\{\begin{bmatrix} a_2 & a_1 & 0 \\ a_2 + b_1 + D_2 & 0 & b_1 \\ 0 & a_1 + b_2 + D_3 & b_2 \\ b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}\right\}$$
$$\ge \min\left\{\begin{bmatrix} a_2 & a_1 & 0 \\ a_2 + b_1 + D_2 & 0 & b_1 \\ 0 & a_1 + b_2 + D_3 & b_2 \\ b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \frac{b_2}{a_2+b_2} \\ \frac{b_1}{a_1+b_1} \\ \frac{a_1 a_2 - b_1 b_2}{(a_1+b_1)(a_2+b_2)} \end{bmatrix}\right\}$$
$$= \min\left\{\begin{bmatrix} a_2 & a_1 & 0 \\ b_1 & b_2 & \max\{b_1, b_2\} + D_4 \end{bmatrix} \begin{bmatrix} \frac{b_2}{a_2+b_2} \\ \frac{b_1}{a_1+b_1} \\ \frac{a_1 a_2 - b_1 b_2}{(a_1+b_1)(a_2+b_2)} \end{bmatrix}\right\}$$
$$> 0 \text{ if } (a_1, a_2, b_1, b_2) \neq (0, 0, 0, 0).$$

Hence, for $b_1 b_2 \le a_1 a_2$ and $(a_1, a_2, b_1, b_2) \neq (0, 0, 0, 0)$ we must have $\hat{\lambda}_0 = 0$. A similar reasoning shows that if $b_1 b_2 \ge a_1 a_2$ and $(a_1, a_2, b_1, b_2) \neq (0, 0, 0, 0)$ we must have $\hat{\lambda}_3 = 0$. It is easy to show that if $\min\{a_1, a_2\} = 0$ then $\hat{\lambda}_0 = 0$, without loss of optimality. Similarly if $\min\{b_1, b_2\} = 0$ then $\hat{\lambda}_3 = 0$, without loss of optimality. This is because, under these conditions, one of the constraints in (3.75) becomes redundant and therefore, by contradiction, it is easy to show that either $\hat{\lambda}_0 = 0$ or $\hat{\lambda}_3 = 0$ is optimal.

We now define the different non-negative quantities (a_u, b_u), $u \in [1:2]$, and D_v, $v \in [1:4]$. From our previous discussion, we restrict attention to the case $\min\{a_1, a_2, b_1, b_2\} \neq 0$. After straightforward manipulations the cut-set bound, for $N = 2$, can be further upper bounded as

$$R^{(\text{cut-set},N=2)} \le 2\log(2) + \max_{S_1,S_2}\{I(X_3; Y_3|X_1, X_2, S_1, S_2)\} + \text{eq.}(3.75),$$

where the term $\max_{S_1,S_2}\{I(X_3; Y_3|X_1, X_2, S_1, S_2)\} \le \log(1 + |h_{33}|^2)$ (with h_{33} being the channel gain from the source to the destination) and where

$$
\begin{aligned}
a_2' &:= I(X_3; Y_2|Y_3, X_1, X_2, S_1 = 0, S_2 = 0),\\
c_2' &:= I(X_3; Y_2|Y_3, X_1, X_2, S_1 = 1, S_2 = 0),\\
a_1' &:= I(X_3; Y_1|Y_3, X_1, X_2, S_1 = 0, S_2 = 0),\\
d_1' &:= I(X_3; Y_1|Y_3, X_1, X_2, S_1 = 0, S_2 = 1),\\
a_2 &:= \max\{a_2', c_2'\},\\
a_1 &:= \max\{a_1', d_1'\},\\
b_1' &:= I(X_1; Y_3|X_2, S_1 = 1, S_2 = 1),\\
c_1' &:= I(X_1; Y_3|X_2, S_1 = 1, S_2 = 0),\\
b_2' &:= I(X_2; Y_3|X_1, S_1 = 1, S_2 = 1),\\
d_2' &:= I(X_2; Y_3|X_1, S_1 = 0, S_2 = 1),\\
b_1 &:= \max\{b_1', c_1'\},\\
b_2 &:= \max\{b_2', d_2'\},\\
D_1 &:= I(X_3; Y_1, Y_2|Y_3, X_1, X_2, S_1{=}0, S_2{=}0)\\
&\quad - \max\{I(X_3; Y_1|Y_3, X_1, X_2, S_1 = 0, S_2 = 0),\\
&\qquad\qquad I(X_3; Y_2|Y_3, X_1, X_2, S_1 = 0, S_2 = 0)\},\\
D_2 &:= I(X_1; Y_2|Y_3, X_2, S_1 = 1, S_2 = 0),\\
D_3 &:= I(X_2; Y_1|Y_3, X_1, S_1 = 0, S_2 = 1),\\
D_4 &:= I(X_1, X_2; Y_3|S_1 = 1, S_2 = 1)\\
&\qquad - \max\{I(X_1; Y_3|X_2, S_1 = 1, S_2 = 1), I(X_2; Y_3|X_1, S_1 = 1, S_2 = 1)\}.
\end{aligned}
$$

If one is interested in the gDoF for the Gaussian noise case, it suffices to consider

$$\mathsf{d} \leq 1 + \text{eq.}(3.75),$$

which is the high-SNR approximation of $R^{(\text{cut}-\text{set},N=2)}$ where the direct link from the source to the destination has SNR-exponent normalized to 1, i.e., $|h_{33}|^2 = \mathsf{SNR}^1$, without loss of generality. In this case the different quantities in (3.75) can be simply found by evaluating the different mutual information terms above and by using the definition of gDoF in Definition 1. We obtain

$$
\begin{aligned}
&a_2' = c_2' = a_2 := [\alpha_{s2} - 1]^+, && a_1' = d_1' = a_1 := [\alpha_{s1} - 1]^+,\\
&b_1' = c_1' = b_1 := [\alpha_{1d} - 1]^+, && b_2' = d_2' = b_2 := [\alpha_{2d} - 1]^+,\\
&D_1 := 0, && D_4 := 0,
\end{aligned}
$$

$$
\begin{aligned}
D_2 &:= \max\{\alpha_{1d} + \alpha_{s2} - 1, \beta_2\} - [\alpha_{1d} - 1]^+ - [\alpha_{s2} - 1]^+,\\
D_3 &:= \max\{\alpha_{2d} + \alpha_{s1} - 1, \beta_1\} - [\alpha_{2d} - 1]^+ - [\alpha_{s1} - 1]^+,
\end{aligned}
$$

where α_{si} is the SNR-exponent on the link from the source to relay i, $i \in [1:2]$, α_{id} is the SNR-exponent on the link from relay i, $i \in [1:2]$, to the destination and β_i is the SNR-exponent on the link from relay j to relay i, $(i,j) \in [1:2]^2$ with $j \neq i$.

3.C Proof of Theorem 6

Let $\mathcal{S}_{n,k}$ be the set of all k-combinations of the integers in $[1:n]$ and $\mathcal{P}_{n,k}$ be the set of all k-permutations of the integers in $[1:n]$. Let $\sigma(\pi)$ be the sign / signature of the permutation π.

We start by demonstrating that the asymptotic behavior of $|\mathbf{I}_k + \mathbf{H}\mathbf{H}^H|$ is as that of $|\mathbf{H}\mathbf{H}^H|$, i.e., the identity matrix can be neglected. By using the determinant Leibniz formula [100], in fact we have,

$$
\begin{aligned}
|\mathbf{I}_k + \mathbf{H}\mathbf{H}^H| &= \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \prod_{i=1}^{k} \left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} \\
&= \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \left\{ \left(\prod_{i=2}^{k} \left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} \right) \left(\left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{1,\pi(1)} \right) \right\} \\
&= \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \left(\prod_{i=2}^{k} \left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} \right) \delta[1-\pi(1)] \\
&+ \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \prod_{i=2}^{k} \left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} \left[\mathbf{H}\mathbf{H}^H\right]_{1,\pi(1)}.
\end{aligned}
$$

Let

$$
\begin{aligned}
\mathsf{A}(\mathsf{SNR}) &:= \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \left(\prod_{i=2}^{k} \left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} \right) \delta[1-\pi(1)] \\
\mathsf{B}(\mathsf{SNR}) &:= \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \prod_{i=2}^{k} \left[\mathbf{I}_k + \mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} \left[\mathbf{H}\mathbf{H}^H\right]_{1,\pi(1)},
\end{aligned}
$$

we have that $\mathsf{A}(\mathsf{SNR}) = o(\mathsf{B}(\mathsf{SNR}))$, because $\lim_{\mathsf{SNR}\to+\infty} \frac{\mathsf{A}(\mathsf{SNR})}{\mathsf{B}(\mathsf{SNR})} = 0$ where the SNR parameterizes the channel gains as $|h_{ij}|^2 = \mathsf{SNR}^{\beta_{ij}}$, for some non-negative β_{ij}. This is so because, as a function of SNR, $\mathsf{B}(\mathsf{SNR})$ grows faster than $\mathsf{A}(\mathsf{SNR})$ due to the term $\left[\mathbf{H}\mathbf{H}^H\right]_{1,\pi(1)}$. By induction it is possible to show that this reasoning holds $\forall i \in [1:k]$ and hence

$$|\mathbf{I}_k + \mathbf{H}\mathbf{H}^H| \doteq \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \prod_{i=1}^{k} \left[\mathbf{H}\mathbf{H}^H\right]_{i,\pi(i)} = |\mathbf{H}\mathbf{H}^H|.$$

Therefore, we now focus on the study of $|\mathbf{H}\mathbf{H}^H|$. We have that

$$\begin{aligned}
|\mathbf{H}\mathbf{H}^H| &\overset{(a)}{=} \sum_{\varsigma \in \mathcal{S}_{n,k}} |\mathbf{H}_\varsigma||\mathbf{H}_\varsigma^H| = \sum_{\varsigma \in \mathcal{S}_{n,k}} |\mathbf{H}_\varsigma|^2 \overset{(b)}{=} \sum_{\varsigma \in \mathcal{S}_{n,k}} \left| \sum_{\pi \in \mathcal{P}_{n,k}} \sigma(\pi) \prod_{i=1}^{k} [\mathbf{H}_\varsigma]_{i,\pi(i)} \right|^2 \\
&= \sum_{\varsigma \in \mathcal{S}_{n,k}} \left\{ \left(\sum_{\pi_1 \in \mathcal{P}_{n,k}} \sigma(\pi_1) \prod_{i=1}^{k} [\mathbf{H}_\varsigma]_{i,\pi_1(i)} \right) \left(\sum_{\pi_2 \in \mathcal{P}_{n,k}} \sigma(\pi_2) \prod_{j=1}^{k} [\mathbf{H}_\varsigma]_{j,\pi_2(j)} \right)^* \right\} \\
&= \sum_{\varsigma \in \mathcal{S}_{n,k}} \left\{ \left(\sum_{\pi \in \mathcal{P}_{n,k}} \prod_{i=1}^{k} \left| [\mathbf{H}_\varsigma]_{i,\pi(i)} \right|^2 \right) \right. \\
&\quad \left. + \left(\sum_{\pi_1,\pi_2 \in \mathcal{P}_{n,k}, \pi_1 \neq \pi_2} \sigma(\pi_1)\sigma(\pi_2) \prod_{i=1}^{k} \prod_{j=1}^{k} [\mathbf{H}_\varsigma]_{i,\pi_1(i)} \left([\mathbf{H}_\varsigma]_{j,\pi_2(j)} \right)^* \right) \right\} \\
&\overset{(c)}{\leq} \sum_{\varsigma \in \mathcal{S}_{n,k}} \left\{ \left(\sum_{\pi \in \mathcal{P}_{n,k}} \prod_{i=1}^{k} \left| [\mathbf{H}_\varsigma]_{i,\pi(i)} \right|^2 \right) \right. \\
&\quad \left. + \left(\sum_{\pi_1,\pi_2 \in \mathcal{P}_{n,k}, \pi_1 \neq \pi_2} \prod_{i=1}^{k} \prod_{j=1}^{k} \sqrt{ \left| [\mathbf{H}_\varsigma]_{i,\pi_1(i)} \right|^2 \left| [\mathbf{H}_\varsigma]_{j,\pi_2(j)} \right|^2 } \right) \right\} \\
&= \sum_{\varsigma \in \mathcal{S}_{n,k}} \left\{ \left(\sum_{\pi \in \mathcal{P}_{n,k}} \mathsf{SNR}^{\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi(i)}} \right) \right. \\
&\quad \left. + \left(\sum_{\pi_1,\pi_2 \in \mathcal{P}_{n,k}, \pi_1 \neq \pi_2} \mathsf{SNR}^{\frac{1}{2}\left(\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi_1(i)} + \sum_{j=1}^{k} [\mathbf{B}_\varsigma]_{j,\pi_2(j)} \right)} \right) \right\} \\
&\overset{(d)}{\doteq} \sum_{\varsigma \in \mathcal{S}_{n,k}} \left(\sum_{\pi \in \mathcal{P}_{n,k}} \mathsf{SNR}^{\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi(i)}} \right) \doteq \mathsf{SNR}^{\max_{\varsigma \in \mathcal{S}_{n,k}} \max_{\pi \in \mathcal{P}_{n,k}} \sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi(i)}},
\end{aligned}$$

where the equalities / inequalities above are due to the following facts:

- equality (a): by applying the Cauchy-Binet formula [100] where $\mathbf{H}_\varsigma$ is the square matrix obtained from $\mathbf{H}$ by retaining all rows and those columns indexed by ς;

- equality (b): by applying the determinant Leibniz formula [100];
- inequality (c): by applying the Cauchy-Schwarz inequality [101];
- equality (d): when $\mathsf{SNR} \to \infty$, we have

$$\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi(i)} \geq \frac{1}{2}\left(\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi_1(i)} + \sum_{j=1}^{k} [\mathbf{B}_\varsigma]_{j,\pi_2(j)}\right).$$

Consider the following example. Let $|a|^2 = \mathsf{SNR}^{\beta_a}$, $|b|^2 = \mathsf{SNR}^{\beta_b}$, $|c|^2 = \mathsf{SNR}^{\beta_c}$, $|d|^2 = \mathsf{SNR}^{\beta_d}$

$$|ab - cd|^2 \leq |a|^2|b|^2 + |c|^2|d|^2 + 2|a||b||c||d|.$$

Now apply the gDoF formula, i.e.,

$$\begin{aligned} \mathsf{d} &:= \lim_{\mathsf{SNR}\to+\infty} \frac{\log\left(|a|^2|b|^2 + |c|^2|d|^2 + 2|a||b||c||d|\right)}{\log(1+\mathsf{SNR})} \\ &= \max\left\{\beta_a + \beta_b,\ \beta_c + \beta_d,\ \frac{\beta_a+\beta_b+\beta_c+\beta_d}{2}\right\}, \end{aligned}$$

but

$$\frac{\beta_a+\beta_b+\beta_c+\beta_d}{2} \leq \frac{2\max\left\{\beta_a+\beta_b, \beta_c+\beta_d\right\}}{2} = \max\left\{\beta_a+\beta_b, \beta_c+\beta_d\right\}.$$

Therefore, the term $\frac{\beta_a+\beta_b+\beta_c+\beta_d}{2}$ does not contribute in characterizing the gDoF. By direct induction, the above reasoning may be extended to a general number of terms leading to $\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi(i)} \geq \frac{1}{2}\left(\sum_{i=1}^{k} [\mathbf{B}_\varsigma]_{i,\pi_1(i)} + \sum_{j=1}^{k} [\mathbf{B}_\varsigma]_{j,\pi_2(j)}\right)$.

3.D Upper and lower bounds for $I_{\emptyset}^{(\mathbf{fix})}$ in (3.54) and $I_{\{1\}}^{(\mathbf{fix})}$ in (3.55)

In (3.53) we assume, without loss of optimality that $P_{1|00} = P_{1|01} = 0$ (respectively $P_{2|00} = P_{2|10} = 0$), since for the HD constraint when the first (respectively second) antenna at the relay is receiving the relay's transmit power on that antenna is zero. With this, we let

$$\begin{aligned} &P_{0|00} = \frac{\alpha_0}{\lambda_0},\ P_{0|01} = \frac{\beta_0}{\lambda_1},\ P_{0|10} = \frac{\gamma_0}{\lambda_2},\ P_{0|11} = \frac{\delta_0}{\lambda_3}, \\ &P_{2|01} = \frac{\alpha_1}{\lambda_1},\ P_{1|10} = \frac{\beta_1}{\lambda_2},\ P_{1|11} = \frac{\gamma_1}{\lambda_3},\ P_{2|11} = \frac{\delta_1}{\lambda_3}, \end{aligned}$$

where $\alpha_i + \beta_i + \gamma_i + \delta_i \le 1, i \in [0:1]$ in order to meet the power constraints in (3.53). We now upper bound $I_{\emptyset}^{(\text{fix})}$ in (3.54) and $I_{\{1\}}^{(\text{fix})}$ in (3.55) separately. We have

$$\begin{aligned}
I_{\emptyset}^{(\text{fix})} &= m_d \log(2) + \lambda_0 \log\left(1 + |h_{ds}|^2 \frac{\alpha_0}{\lambda_0}\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 \frac{\beta_0}{\lambda_1} + |h_{dr,2}|^2 \frac{\alpha_1}{\lambda_1}\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 \frac{\gamma_0}{\lambda_2} + |h_{dr,1}|^2 \frac{\beta_1}{\lambda_2}\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2 \frac{\delta_0}{\lambda_3} + \left(\sqrt{|h_{dr,1}|^2 \frac{\gamma_1}{\lambda_3}} + \sqrt{|h_{dr,2}|^2 \frac{\delta_1}{\lambda_3}}\right)^2\right) \\
&\le m_d \log(2) + H(\lambda_s) + \lambda_0 \log\left(1 + |h_{ds}|^2 \alpha_0\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 \beta_0 + |h_{dr,2}|^2 \alpha_1\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 \gamma_0 + |h_{dr,1}|^2 \beta_1\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2 \delta_0 + \left(\sqrt{|h_{dr,1}|^2 \gamma_1} + \sqrt{|h_{dr,2}|^2 \delta_1}\right)^2\right) \\
&\le m_d \log(2) + 2\log(2) + \lambda_0 \log\left(1 + |h_{ds}|^2\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 + |h_{dr,2}|^2\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 + |h_{dr,1}|^2\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2 + \left(\sqrt{|h_{dr,1}|^2} + \sqrt{|h_{dr,2}|^2}\right)^2\right),
\end{aligned}$$

where the two inequalities follow because: (i) the entropy of a discrete random variable can be upper bounded by the logarithm of the size of its support (i.e., $H(\lambda_s) \le \log(4)$); (ii) by further upper bounding the power splits by setting $\alpha_i = \beta_i = \gamma_i = \delta_i = 1$, $i \in [0:1]$; (iii) by further upper bounding all the $\lambda_s, s \in [0:3]$ inside the logarithms by one. With similar steps we obtain

$$\begin{aligned}
I_{\{1\}}^{(\text{fix})} &\le 2\log(2) + \lambda_0 \log\left(1 + |h_{ds}|^2 + |h_{rs,1}|^2 + |h_{rs,2}|^2\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 + |h_{rs,1}|^2\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 + |h_{rs,2}|^2\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2\right).
\end{aligned}$$

We now lower bound $I_{\emptyset}^{(\text{fix})}$ in (3.54) and $I_{\{1\}}^{(\text{fix})}$ in (3.55) separately. We have

$$\begin{aligned}
I_{\emptyset}^{(\text{fix})} &= m_d \log(2) + \lambda_0 \log\left(1 + |h_{ds}|^2 \frac{\alpha_0}{\lambda_0}\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 \frac{\beta_0}{\lambda_1} + |h_{dr,2}|^2 \frac{\alpha_1}{\lambda_1}\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 \frac{\gamma_0}{\lambda_2} + |h_{dr,1}|^2 \frac{\beta_1}{\lambda_2}\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2 \frac{\delta_0}{\lambda_3} + \left(\sqrt{|h_{dr,1}|^2 \frac{\gamma_1}{\lambda_3}} + \sqrt{|h_{dr,2}|^2 \frac{\delta_1}{\lambda_3}}\right)^2\right) \\
&\geq m_d \log(2) - \log(2) + \lambda_0 \log\left(1 + |h_{ds}|^2\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 + |h_{dr,2}|^2\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 + |h_{dr,1}|^2\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2 + \left(\sqrt{|h_{dr,1}|^2} + \sqrt{|h_{dr,2}|^2}\right)^2\right),
\end{aligned}$$

by setting $\alpha_1 = \lambda_1$, $\beta_1 = \lambda_2$, $\gamma_1 = \delta_1 = \frac{\lambda_3}{2}$, $\alpha_0 = \lambda_0$, $\beta_0 = \lambda_1$, $\gamma_0 = \lambda_2$ and $\delta_0 = \lambda_3$ (note that with these power splits the power constraints in (3.53) are satisfied) and by using the further bound $\log\left(1 + \left(\sqrt{\frac{a}{2}} + \sqrt{\frac{c}{2}}\right)^2\right) = \log\left(1 + \frac{1}{2}\left(\sqrt{a} + \sqrt{c}\right)^2\right) \geq \log\left(\frac{1}{2} + \frac{1}{2}\left(\sqrt{a} + \sqrt{c}\right)^2\right) = \log\left(1 + \left(\sqrt{a} + \sqrt{c}\right)^2\right) - \log(2)$. With similar steps we obtain

$$\begin{aligned}
I_{\{1\}}^{(\text{fix})} &\geq \lambda_0 \log\left(1 + |h_{ds}|^2 + |h_{rs,1}|^2 + |h_{rs,2}|^2\right) \\
&+ \lambda_1 \log\left(1 + |h_{ds}|^2 + |h_{rs,1}|^2\right) \\
&+ \lambda_2 \log\left(1 + |h_{ds}|^2 + |h_{rs,2}|^2\right) \\
&+ \lambda_3 \log\left(1 + |h_{ds}|^2\right).
\end{aligned}$$

3.E Water filling power allocation for $I_{\emptyset}^{(\text{fix})}$ in (3.54) and $I_{\{1\}}^{(\text{fix})}$ in (3.55)

By optimizing the powers in the different relay states subject to the power constraints in (3.53) we have

$$\mathsf{C}''_{\text{case (i)}} = \max_{\lambda_s} \min\left\{I_{\emptyset}^{(\text{fix})}, I_{\{1\}}^{(\text{fix})}\right\},$$

where $I_{\emptyset}^{(\mathrm{fix})}$ and $I_{\{1\}}^{(\mathrm{fix})}$ are defined in (3.54) and in (3.55), respectively. By writing the Lagrangian of the optimization problem above (subject to the power constraints in (3.53) and by considering $|h_{ds}| = 0$, i.e., the direct link is absent) we obtain

$$\begin{aligned}
I_{\emptyset}^{(\mathrm{fix})} &= \lambda_1 \log^+\left(\nu_0 |h_{dr,2}|^2\right) + \lambda_2 \log^+\left(\nu_0 |h_{dr,1}|^2\right) \\
&\quad + \lambda_3 \log^+\left(\nu_0 (|h_{dr,1}|^2 + |h_{dr,2}|^2)\right) \\
\nu_0 &: \lambda_1 \left[\nu_0 - \frac{1}{|h_{dr,2}|^2}\right]^+ + \lambda_2 \left[\nu_0 - \frac{1}{|h_{dr,1}|^2}\right]^+ \\
&\quad + \lambda_3 \left[\nu_0 - \frac{1}{|h_{dr,1}|^2 + |h_{dr,2}|^2}\right]^+ = 1, \\
I_{\{1\}}^{(\mathrm{fix})} &= \lambda_0 \log^+\left(\nu_1 (|h_{rs,1}|^2 + |h_{rs,2}|^2)\right) + \lambda_1 \log^+\left(\nu_1 |h_{rs,1}|^2\right) \\
&\quad + \lambda_2 \log^+\left(\nu_1 |h_{rs,2}|^2\right), \\
\nu_1 &: \lambda_0 \left[\nu_1 - \frac{1}{|h_{rs,1}|^2 + |h_{rs,2}|^2}\right]^+ + \lambda_1 \left[\nu_1 - \frac{1}{|h_{rs,1}|^2}\right]^+ \\
&\quad + \lambda_2 \left[\nu_1 - \frac{1}{|h_{rs,2}|^2}\right]^+ = 1.
\end{aligned}$$

For case (ii), it suffices to set $\lambda_1 = \lambda_2 = 0$ in case (i). Let $\lambda_3 = 1 - \lambda_0 = \lambda \in [0,1]$, and $\|\mathbf{h}_{dr}\|^2 = |h_{dr,1}|^2 + |h_{dr,2}|^2$, $\|\mathbf{h}_{rs}\|^2 = |h_{rs,1}|^2 + |h_{rs,2}|^2$. With this we get

$$\begin{aligned}
\mathsf{C}''_{\text{case (ii)}} &= \max_{\lambda \in [0,1]} \min \left\{ \lambda \log\left(1 + \frac{\|\mathbf{h}_{dr}\|^2}{\lambda}\right), (1-\lambda) \log\left(1 + \frac{\|\mathbf{h}_{rs}\|^2}{1-\lambda}\right) \right\} \\
&\in \left[\frac{\log(1+\|\mathbf{h}_{rs}\|^2)\log(1+\|\mathbf{h}_{dr}\|^2)}{\log(1+\|\mathbf{h}_{rs}\|^2)+\log(1+\|\mathbf{h}_{dr}\|^2)}, \frac{\log(1+\|\mathbf{h}_{rs}\|^2)\log(1+\|\mathbf{h}_{dr}\|^2)}{\log(1+\|\mathbf{h}_{rs}\|^2)+\log(1+\|\mathbf{h}_{dr}\|^2)} + 1 \right],
\end{aligned}$$

where the optimal λ is obtained by equating the two expressions within the min.

Part II

The Causal Cognitive Interference Channel, or the Interference Channel with Unilateral Source Cooperation

Chapter 4

Case I: Full-Duplex CTx

In this chapter, we study the CCIC, or the IC with unilateral source cooperation, when the CTx operates in FD. Our main contributions can be summarized as follows: (i) we develop a general framework to derive outer bounds of the type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ *and* $R_{\mathsf{p}}+2R_{\mathsf{c}}$ *on the capacity of the general ISD channel when the noises at the different source-destination pairs are independent; (ii) we design a transmission strategy for the general memoryless channel and we derive its achievable rate region; (iii) we evaluate the outer bound and the achievable rate regions for the Gaussian noise channel and we prove a constant gap result for the Z-, the S- and the symmetric fully-connected channels; (iv) we identify the regimes where the Gaussian channel attains the same gDoF of the non-cooperative Gaussian IC and those where the gDoF performance equals that of the non-causal CIC.*

4.1 System Model

Throughout this chapter we adopt the following notation convention. The subscript c (in sans serif font) is used for quantities related to the cognitive pair, while the subscript p (in sans serif font) for those related to the primary pair. The subscript f or F (in sans serif font) is used to refer to generalized feedback information received at the CTx. The subscript c (in roman font) is used to denote both common and cooperative messages, the subscript p (in roman font) to denote private messages and the subscript n (in roman

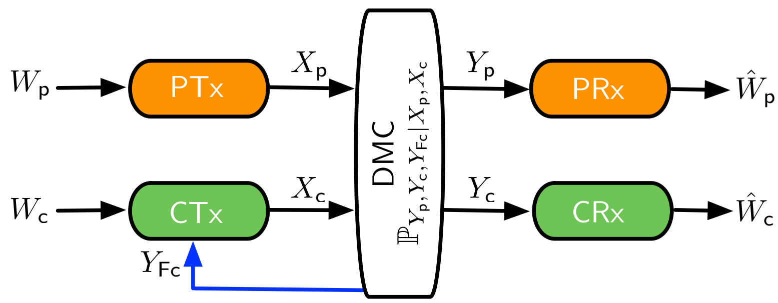


Figure 4.1: The general memoryless CCIC.

font) to denote non-cooperative messages.

4.1.1 General memoryless channel

A general memoryless CCIC, shown in Figure 4.1, consists of two input alphabets $(\mathcal{X}_{\mathsf{p}}, \mathcal{X}_{\mathsf{c}})$, three output alphabets $(\mathcal{Y}_{\mathsf{Fc}}, \mathcal{Y}_{\mathsf{p}}, \mathcal{Y}_{\mathsf{c}})$ and a memoryless transition probability $P_{Y_{\mathsf{Fc}},Y_{\mathsf{p}},Y_{\mathsf{c}}|X_{\mathsf{p}},X_{\mathsf{c}}}$. PTx has a message $W_{\mathsf{p}} \in [1 : 2^{NR_{\mathsf{p}}}]$ for PRx and CTx has a message $W_{\mathsf{c}} \in [1 : 2^{NR_{\mathsf{c}}}]$ for CRx, where $N \in \mathbb{N}$ denotes the codeword length and $R_{\mathsf{p}} \in \mathbb{R}_+$ and $R_{\mathsf{c}} \in \mathbb{R}_+$ the transmission rates for PTx and CTx, respectively, in bits per channel use. The messages W_{p} and W_{c} are independent and uniformly distributed on their respective domains. At time i, $i \in [1 : N]$, PTx maps its message W_{p} into a channel input symbol $X_{\mathsf{p}i}(W_{\mathsf{p}})$ and CTx maps its message W_{c} and its past channel observations into a channel input symbol $X_{\mathsf{c}i}(W_{\mathsf{c}}, Y_{\mathsf{Fc}}^{i-1})$. At time N, PRx outputs an estimate of its intended message based on all its channel observations as $\widehat{W}_{\mathsf{p}}({Y_{\mathsf{p}}}^N)$, and similarly CRx outputs $\widehat{W}_{\mathsf{c}}({Y_{\mathsf{c}}}^N)$. The capacity region is the convex closure of all non-negative rate pairs $(R_{\mathsf{p}}, R_{\mathsf{c}})$ such that $\max_{u\in\{\mathsf{c},\mathsf{p}\}} \mathbb{P}[\widehat{W}_u \neq W_u] \to 0$ as $N \to +\infty$.

4.1.2 ISD channel

The ISD model, shown in Figure 4.2 and first introduced in [48] for the classical IC, assumes that the input X_{p}, respectively X_{c}, before reaching the destinations, is passed through a memoryless channel to obtain T_{p}, respec-

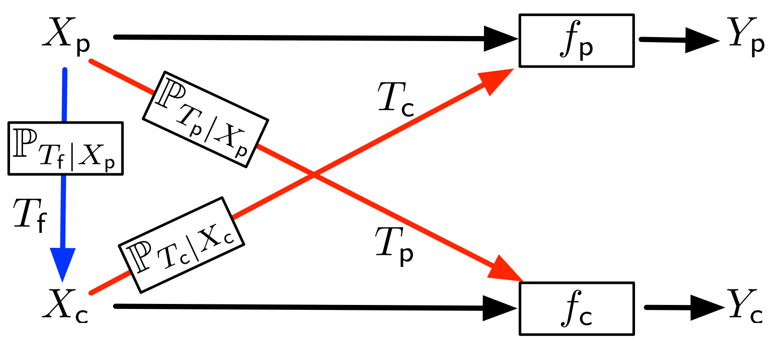


Figure 4.2: The ISD CCIC.

tively T_{c}. The channel outputs are therefore given by

$$Y_{\mathsf{p}} = f_{\mathsf{p}}\left(X_{\mathsf{p}}, T_{\mathsf{c}}\right), \tag{4.1a}$$
$$Y_{\mathsf{c}} = f_{\mathsf{c}}\left(X_{\mathsf{c}}, T_{\mathsf{p}}\right), \tag{4.1b}$$

where f_u, $u \in \{\mathsf{p}, \mathsf{c}\}$, is a deterministic function that is invertible given X_u, or in other words, T_{p}, respectively T_{c}, is a deterministic function of $(Y_{\mathsf{c}}, X_{\mathsf{c}})$, respectively $(Y_{\mathsf{p}}, X_{\mathsf{p}})$.

In the CCIC, the "generalized feedback signal" at the CTx satisfies

$$Y_{\mathsf{Fc}} = f_{\mathsf{f}}\left(X_{\mathsf{c}}, T_{\mathsf{f}}\right), \tag{4.1c}$$

for some deterministic function f_{f} that is invertible given X_{c}, i.e., T_{f} is a deterministic function of $(Y_{\mathsf{Fc}}, X_{\mathsf{c}})$, where T_{f} is obtained by passing X_{p} through a noisy channel [47].

We further assume that the noises seen by the different source-destination pairs are independent, i.e.,

$$\mathbb{P}_{Y_{\mathsf{Fc}}, Y_{\mathsf{p}}, Y_{\mathsf{c}}|X_{\mathsf{p}}, X_{\mathsf{c}}} = \mathbb{P}_{Y_{\mathsf{p}}|X_{\mathsf{p}}, X_{\mathsf{c}}} \mathbb{P}_{Y_{\mathsf{Fc}}, Y_{\mathsf{c}}|X_{\mathsf{p}}, X_{\mathsf{c}}}. \tag{4.2}$$

In other words, we assume that the noises at the PRx and at the CRx are independent, but we do not impose any constraint on the noises at the CTx and CRx.

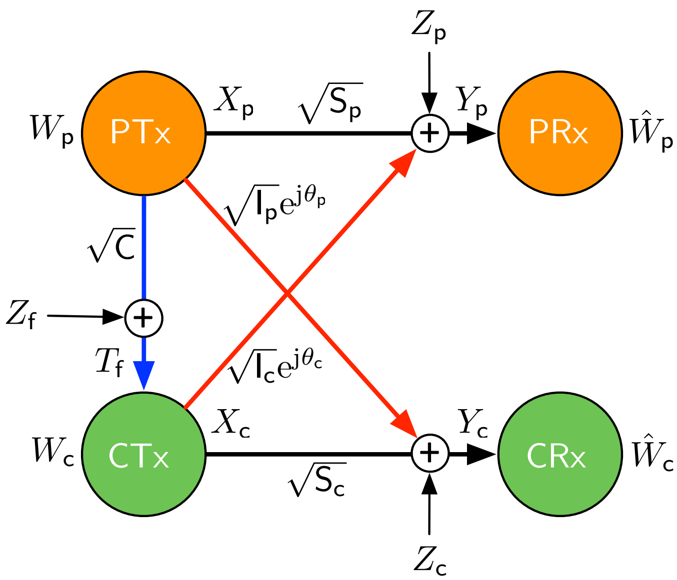


Figure 4.3: The Gaussian CCIC.

4.1.3 The Gaussian noise channel

A single-antenna Gaussian CCIC, shown in Figure 4.3, is a special case of the ISD model and it is defined by the input / output relationship

$$T_{\mathsf{p}} := \sqrt{\mathsf{I}_{\mathsf{p}}}e^{j\theta_{\mathsf{p}}}X_{\mathsf{p}} + Z_{\mathsf{c}}, \tag{4.3a}$$
$$T_{\mathsf{c}} := \sqrt{\mathsf{I}_{\mathsf{c}}}e^{j\theta_{\mathsf{c}}}X_{\mathsf{c}} + Z_{\mathsf{p}}, \tag{4.3b}$$
$$T_{\mathsf{f}} = \sqrt{\mathsf{C}}X_{\mathsf{p}} + Z_{\mathsf{f}}, \tag{4.3c}$$
$$Y_{\mathsf{p}} = \sqrt{\mathsf{S}_{\mathsf{p}}}X_{\mathsf{p}} + T_{\mathsf{c}}, \tag{4.3d}$$
$$Y_{\mathsf{c}} = T_{\mathsf{p}} + \sqrt{\mathsf{S}_{\mathsf{c}}}X_{\mathsf{c}}, \tag{4.3e}$$
$$Y_{\mathsf{Fc}} = T_{\mathsf{f}}, \tag{4.3f}$$

where $T_{\mathsf{f}} = Y_{\mathsf{Fc}}$ in (4.3f) is without loss of generality since the CTx can remove the contribution of its transmit signal X_{c} from its received signal Y_{Fc}. The channel gains are assumed to be constant for the whole transmission duration, and hence known to all nodes. Without loss of generality, certain channel gains can be taken to be real-valued and non-negative since a node can compensate for the phase of one of its channel gains. The channel inputs are subject to a unitary average power constraint, i.e., $\mathbb{E}\left[|X_i|^2\right] \leq 1, i \in \{\mathsf{p}, \mathsf{c}\}$. This assumption is without loss of generality, since non-unitary power constraints can be incorporated into the channel gains. The noises are circularly symmetric Gaussian random variables with, without loss of generality, zero mean and unitary variance. We assume that the noise Z_{p} is independent of $(Z_{\mathsf{c}}, Z_{\mathsf{f}})$, while $(Z_{\mathsf{c}}, Z_{\mathsf{f}})$ can be arbitrarily correlated.

The non-cooperative Gaussian IC is obtained as a special case of the Gaussian CCIC by setting $\mathsf{C} = 0$ and the Gaussian non-causal CIC in the limit for $\mathsf{C} \to +\infty$. A Gaussian CCIC is said to be a Z-channel if $\mathsf{I}_{\mathsf{p}} = 0$, i.e., the CRx does not experience interference from PTx, and an S-channel if $\mathsf{I}_{\mathsf{c}} = 0$, i.e., the PRx does not experience interference from CTx.

As already remarked for the HD relay channel in Chapter 2, for the Gaussian noise case, it is customary to approximate the channel capacity as follows.

Definition 6. *The capacity region of the Gaussian CCIC is said to be known to within* GAP *bits if one can show an inner bound region* $\mathcal{I}$ *and an outer bound region* $\mathcal{O}$ *such that*

$$(R_{\mathsf{p}}, R_{\mathsf{c}}) \in \mathcal{O} \Longrightarrow ([R_{\mathsf{p}} - \mathsf{GAP}]^+, [R_{\mathsf{c}} - \mathsf{GAP}]^+) \in \mathcal{I}.$$

For the two particular cases of $\mathsf{C} = 0$ (i.e., non-cooperative IC) and of $\mathsf{C} \to +\infty$ (non-causal CIC), the capacity is known to within 1 bit [12, 64]. The approximate (i.e., to within a constant gap) characterization of the capacity region implies the exact knowledge of its gDoF region. The gDoF metric, first introduced in [12] for the non-cooperative IC, captures the high-SNR behavior of the capacity as a function of the relative strengths of the direct, cooperation and interfering links. The gDoF represents a more refined characterization of the capacity in the high-SNR regime compared to the classical DoF since it captures the fact that, in wireless networks, the channel gains can differ by several orders of magnitude. Let $\mathsf{S} > 1$ and parameterize

$$\begin{aligned}
\mathsf{S}_{\mathsf{p}} &:= \mathsf{S}^1, \text{ primary direct link}, && (4.4\text{a})\\
\mathsf{S}_{\mathsf{c}} &:= \mathsf{S}^1, \text{ cognitive direct link}, && (4.4\text{b})\\
\mathsf{I}_{\mathsf{p}} &:= \mathsf{S}^{\alpha_{\mathsf{p}}}, \ \alpha_{\mathsf{p}} \geq 0, \text{ interference at CRx from PTx}, && (4.4\text{c})\\
\mathsf{I}_{\mathsf{c}} &:= \mathsf{S}^{\alpha_{\mathsf{c}}}, \ \alpha_{\mathsf{c}} \geq 0, \text{ interference at PRx from CTx}, && (4.4\text{d})\\
\mathsf{C} &:= \mathsf{S}^{\beta}, \ \beta \geq 0, \text{ cooperation link}, && (4.4\text{e})
\end{aligned}$$

where α_{p} and α_{c} measure the strength of the interference links compared to the direct link, while β the strength of the cooperation link compared to the direct link. We remark that the parameterization in (4.4), with direct links of the same strength, will be used only for evaluation of the gDoF. Moreover, in order to capture different network topologies, we focus on

1. interference-symmetric channel: $\alpha_{\mathsf{p}} = \alpha_{\mathsf{c}} = \alpha$;
2. Z-channel: $\alpha_{\mathsf{p}} = 0,\ \alpha_{\mathsf{c}} = \alpha$;
3. S-channel: $\alpha_{\mathsf{p}} = \alpha,\ \alpha_{\mathsf{c}} = 0$.

The case $\alpha_{\mathsf{p}} = \alpha_{\mathsf{c}} = 0$ is not interesting since in this case the Gaussian CCIC reduces to two parallel point-to-point links for which cooperation is useless. For the above three cases, the system is parameterized by the triplet $(\mathsf{S}, \alpha, \beta)$, where S is referred to as the (direct link) SNR, α as the interference exponent and β as the cooperation exponent. [1] Following the naming convention of the non-cooperative IC [12], we say that the Gaussian CCIC for the above three cases has strong interference if $\mathsf{S} \le \mathsf{I}$, that is $1 \le \alpha$, and weak interference otherwise. Similarly, we say that the Gaussian CCIC has strong cooperation if $\mathsf{S} \le \mathsf{C}$, that is $1 \le \beta$, and weak cooperation otherwise.

Definition 7. *Given the parameterization in* (4.4)*, the gDoF is defined as*

$$\mathsf{d}(\alpha, \beta) := \lim_{\mathsf{S} \to +\infty} \frac{\max\{R_{\mathsf{p}} + R_{\mathsf{c}}\}}{2\log(1+\mathsf{S})}, \tag{4.5}$$

where the maximization is intended over all possible achievable rate pairs $(R_{\mathsf{p}}, R_{\mathsf{c}})$.

The gDoF of the classical IC ($\mathsf{C} = 0$) is the "W-curve" first characterized in [12] and given by $\mathsf{d}(\alpha, 0) = \min\{\max\{1-\alpha, \alpha\},\ \max\{1-\alpha/2, \alpha/2\},\ 1\}$. The gDoF of the non-causal CIC ($\mathsf{C} \to \infty$) is the "V-curve", which can be evaluated from the capacity characterization to within 1 bit of [64], and is given by $\mathsf{d}(\alpha, \infty) = \max\{1-\alpha/2, \alpha/2\}$. An interesting question we seek to answer in this chapter is whether there are values of $\beta > 0$ such that $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, 0)$ — in which case unilateral causal cooperation is not helpful in terms of gDoF — or values of $\beta < \infty$ such that $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, \infty)$ — in which case unilateral causal cooperation is equivalent to non-causal message knowledge in terms of gDoF.

[1] In principle the system performance also depends on the phases of the interfering links $(\theta_{\mathsf{c}}, \theta_{\mathsf{p}})$. However, as far as gDoF and capacity to within a constant gap are concerned, the phases $(\theta_{\mathsf{c}}, \theta_{\mathsf{p}})$ only matter if the IC channel matrix $\begin{bmatrix} \sqrt{\mathsf{S}_{\mathsf{p}}} & \sqrt{\mathsf{I}_{\mathsf{c}}} e^{\mathrm{j}\theta_{\mathsf{c}}} \\ \sqrt{\mathsf{I}_{\mathsf{p}}} e^{\mathrm{j}\theta_{\mathsf{p}}} & \sqrt{\mathsf{S}_{\mathsf{c}}} \end{bmatrix}$ is rank deficient [56], in which case one received signal is a noisier version of the other. In this work, we assume that the phases are such that the IC channel matrix is full rank.

4.2 Overview of the main results

The exact capacity of the Gaussian CCIC described in (4.3) is unknown. In this chapter, we characterize the capacity to within a constant gap (see Definition 6) and, hence, the gDoF (see Definition 7), for the symmetric case (i.e., $\mathsf{I_p} = \mathsf{I_c}$ and $\mathsf{S_c} = \mathsf{S_p}$ in (4.3)), for the Z-channel (i.e., $\mathsf{I_p} = 0$ in (4.3)) and for the S-channel (i.e., $\mathsf{I_c} = 0$ in (4.3)) for the case of independent noises.

In order to show the constant gap results an outer and an inner bound regions on the capacity of the Gaussian CCIC are needed. Concerning the outer bound region, we use some outer bounds on the single rates R_p and R_c and on the sum-rate $R_\mathsf{p} + R_\mathsf{c}$ known in the literature [45, 47, 87]. Moreover, in order to show capacity to within a constant gap for the symmetric Gaussian CCIC in weak interference, we develop a general framework to derive outer bounds of the type $2R_\mathsf{p} + R_\mathsf{c}$ and $R_\mathsf{p} + 2R_\mathsf{c}$ on the capacity of the general ISD CCIC when the noises at the different source-destination pairs are independent; this framework includes for example feedback from the intended destination. In particular, our first main result is

Theorem 8. *For the ISD CCIC satisfying* (4.2) *the capacity region is outer bounded by*

$$2R_\mathsf{p} + R_\mathsf{c} \leq I\left(Y_\mathsf{p}; X_\mathsf{p}, X_\mathsf{c}\right) + I\left(Y_\mathsf{p}; X_\mathsf{p}|Y_\mathsf{c}, T_\mathsf{f}, X_\mathsf{c}\right) + I\left(Y_\mathsf{c}, T_\mathsf{f}; X_\mathsf{p}, X_\mathsf{c}|T_\mathsf{c}\right), \tag{4.6}$$

$$R_\mathsf{p} + 2R_\mathsf{c} \leq I\left(Y_\mathsf{c}; X_\mathsf{p}, X_\mathsf{c}\right) + I\left(Y_\mathsf{c}; X_\mathsf{c}|Y_\mathsf{p}, T_\mathsf{f}, X_\mathsf{p}\right) + I\left(Y_\mathsf{p}, T_\mathsf{f}; X_\mathsf{p}, X_\mathsf{c}|T_\mathsf{p}\right), \tag{4.7}$$

for some input distribution $\mathbb{P}_{X_\mathsf{p}, X_\mathsf{c}}$.

The key technical ingredient is the proof of the two following Markov chains.

Lemma 1. *For the ISD CCIC with the noise structure in* (4.2)*, the following Markov chains hold for all* $i \in [1:N]$*:*

$$(W_\mathsf{p}, {T_\mathsf{p}}^{i-1}, {X_\mathsf{p}}^{i}) - ({T_\mathsf{c}}^{i-1}, {T_\mathsf{f}}^{i-1}) - (T_{\mathsf{c}i}), \tag{4.8a}$$

$$(W_\mathsf{c}, {T_\mathsf{c}}^{i-1}, {X_\mathsf{c}}^{i}) - ({T_\mathsf{p}}^{i-1}, {T_\mathsf{f}}^{i-1}) - (T_{\mathsf{p}_i}, T_{\mathsf{f}i}). \tag{4.8b}$$

Concerning the inner bound region, we use the *superposition+binning* transmission strategy from [51, Section V]. This scheme was originally designed for the general memoryless IC with generalized feedback, or bilateral source cooperation. In this chapter we adapt this strategy to the case of unilateral source cooperation. In particular, the PTx's message is split into four parts: the non-cooperative common message and the non-cooperative private message are sent as in the Han-Kobayashi's scheme for the non-cooperative IC [39]; the cooperative common message and the cooperative private message are decoded at CTx in a given slot and retransmitted in the next slot by using a PDF based block-Markov scheme. The CTx's message is split into two parts: the non-cooperative common message and the non-cooperative private message that are sent as in the Han-Kobayashi's scheme for the non-cooperative IC [39]. The common messages are decoded at both destinations while non-intended private messages are treated as noise. For cooperation, the two sources 'beam form' the PTx's cooperative common message to the destinations as in a distributed MIMO system, and the CTx precodes its private messages against the interference created by the PTx's cooperative private message as in a MIMO BC. The achievable region in [51, Section V] is quite complex to evaluate because it is a function of 11 auxiliary random variables and is described by about 30 rate constraints per source-destination pair. In this chapter we use a small subset of these 11 auxiliary random variables in each parameter regime and show that the corresponding schemes are to within a constant gap from the outer bound region described above. In particular, our constant gap results are stated in the three following theorems.

Theorem 9. *The capacity region outer bound of the symmetric Gaussian CCIC (i.e., when $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_p} = \mathsf{I_c} = \mathsf{I}$) is achievable to within* 5 *bits. In particular,*

1. *When $\mathsf{I} \geq \mathsf{S}$ (i.e., $\alpha \geq 1$), then $\mathsf{GAP} \leq 1$ bit,*
2. *When $\mathsf{I} < \mathsf{S}$ (i.e., $\alpha < 1$) and $\mathsf{C} \leq \mathsf{S}$ (i.e., $\beta \leq 1$), then $\mathsf{GAP} \leq 5$ bits,*
3. *When $\mathsf{I} < \mathsf{S}$ (i.e., $\alpha < 1$) and $\mathsf{S} < \mathsf{C}$ (i.e., $1 < \beta$), then $\mathsf{GAP} \leq 2$ bits.*

Theorem 10. *The capacity region outer bound of the Z-channel (i.e., $\mathsf{I_p} = 0$, the link PTx$\rightarrow$CRx is non-existent) is characterized to within* 2 *bits. In particular,*

1. *When $\mathsf{C} \leq \mathsf{S_p}$, then $\mathsf{GAP} \leq 2$ bits,*
2. *When $\mathsf{C} > \mathsf{S_p}$ and $\mathsf{S_c} \leq \mathsf{I_c}$, then $\mathsf{GAP} \leq 1.5$ bits,*
3. *When $\mathsf{C} > \mathsf{S_p}$ and $\mathsf{S_c} > \mathsf{I_c}$, then $\mathsf{GAP} \leq 1$ bit.*

Theorem 11. *The capacity region outer bound of the S-channel (i.e.,* $\mathsf{I_c} = 0$*, the link CTx→PRx is non-existent) is achievable to within* 3 *bits. In particular,*

1. *When* $\mathsf{C} \leq \max\{\mathsf{S_p}, \mathsf{I_p}\}$*, then* $\mathsf{GAP} \leq 2.5$ *bits,*
2. *When* $\max\{\mathsf{S_p}, \mathsf{I_p}\} < \mathsf{C} \leq \mathsf{I_p}\mathsf{S_p}$*, then* $\mathsf{GAP} \leq 3$ *bits,*
3. *When* $\mathsf{C} > \mathsf{I_p}\mathsf{S_p}$*, then* $\mathsf{GAP} \leq 1$ *bit.*

The rest of this chapter is dedicated to the proof of Theorems 8-11.

4.3 Outer bounds on the capacity region for the CCIC

This section is dedicated to the study of outer bounds on the capacity region for the CCIC. First, in Section 4.3.1, some known outer bounds are summarized. Moreover, the outer bound originally derived in [47] for the ISD CCIC with independent noises at all terminals, is generalized to the case where the noises at the different source-destination pairs are independent as in (4.2). Then, in Section 4.3.2, the two novel outer bounds of the type $R_\mathsf{p} + 2R_\mathsf{c}$ and $2R_\mathsf{p} + R_\mathsf{c}$ in Theorem 8 are derived for the ISD CCIC when (4.2) holds. Finally, in Section 4.3.3, all the outer bounds are evaluated for the practically relevant Gaussian noise case.

4.3.1 Known outer bounds and some generalizations

In the literature, several outer bounds are known for the IC with bilateral source cooperation [45, 47], which we specialize here to the CCIC. In particular, for an input distribution $\mathbb{P}_{X_\mathsf{p},X_\mathsf{c}}$, we have:

1. For the general memoryless CCIC, described in Section 4.1.1, the cut-set upper bound [87] gives

$$R_\mathsf{p} \leq I\left(X_\mathsf{p}; Y_\mathsf{p}, Y_{\mathsf{Fc}}|X_\mathsf{c}\right), \tag{4.9a}$$
$$R_\mathsf{p} \leq I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p}\right), \tag{4.9b}$$
$$R_\mathsf{c} \leq I\left(X_\mathsf{c}; Y_\mathsf{c}|X_\mathsf{p}\right), \tag{4.9c}$$

   and from [45] we have

$$R_\mathsf{p} + R_\mathsf{c} \leq I\left(X_\mathsf{p}; Y_\mathsf{p}, Y_{\mathsf{Fc}}|Y_\mathsf{c}, X_\mathsf{c}\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{c}\right), \tag{4.9d}$$
$$R_\mathsf{p} + R_\mathsf{c} \leq I\left(X_\mathsf{c}; Y_\mathsf{c}|Y_\mathsf{p}, X_\mathsf{p}\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p}\right). \tag{4.9e}$$

Notice that in the bounds in (4.9a)-(4.9e), Y_{Fc} always appears conditioned on X_{c}. This implies that, for the ISD channel described in Section 4.1.2, Y_{Fc} can be replaced with T_{f} without loss of generality.

2. For the memoryless ISD CCIC, described in Section 4.1.2, with independent noises at the different source-destination pairs as in (4.2), we have

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{p}}, T_{\mathsf{f}}\right) + I\left(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{c}}\right). \tag{4.9f}$$

The details of the proof of the bound in (4.9f) can be found in Appendix 4.B. We note that a bound as the one in (4.9f) was originally derived in [47, Appendix IV pages 177-179] for the ISD IC with bilateral source cooperation when all noises are independent; for the case of unilateral source cooperation, this follows from the two following Markov chains

$$(W_{\mathsf{p}}, X_{\mathsf{p}}{}^{i}) - (T_{\mathsf{f}}{}^{i-1}) - (W_{\mathsf{c}}, X_{\mathsf{c}}{}^{i}, T_{\mathsf{c}}{}^{i}), \quad \forall i \in [1:N], \tag{4.10}$$

$$(W_{\mathsf{c}}, X_{\mathsf{c}}{}^{i}) - (T_{\mathsf{f}}{}^{i-1}) - (W_{\mathsf{p}}, X_{\mathsf{p}}{}^{i}, T_{\mathsf{p}}{}^{i}), \quad \forall i \in [1:N]. \tag{4.11}$$

A careful analysis of the bounding steps in [47, Appendix IV pages 177-179] shows that the derivation of the bound in (4.9f) is valid even when $\mathbb{P}_{Y_{\mathsf{Fc}},Y_{\mathsf{c}},Y_{\mathsf{p}}|X_{\mathsf{p}},X_{\mathsf{c}}}$ factors as in (4.2), i.e., the independent noises assumption at all terminals captured by the product distribution $\mathbb{P}_{Y_{\mathsf{Fc}},Y_{\mathsf{c}},Y_{\mathsf{p}}|X_{\mathsf{p}},X_{\mathsf{c}}} = \mathbb{P}_{Y_{\mathsf{Fc}}|X_{\mathsf{p}},X_{\mathsf{c}}}\mathbb{P}_{Y_{\mathsf{c}}|X_{\mathsf{p}},X_{\mathsf{c}}}\mathbb{P}_{Y_{\mathsf{p}}|X_{\mathsf{p}},X_{\mathsf{c}}}$ is not necessary for the bound to hold by suitably modifying the Markov chains in (4.10) and (4.11) – see Lemma 1. An advantage of the bound in (4.9f) is that the case of output feedback from the intended destination is a special case of the more general framework and can be obtained by $Y_{\mathsf{Fc}} = Y_{\mathsf{c}}$. We note that our bound in (4.9f) not only is more general but it is also tighter than the one in [47, Appendix IV pages 177-179] since $I\left(T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{c}}\right) \leq I\left(T_{\mathsf{f}}; X_{\mathsf{p}}\right)$; moreover, thanks to more careful bounding steps compared to [47], in (4.9f) we obtained the term $H\left(Y_{\mathsf{p}} | T_{\mathsf{p}}, T_{\mathsf{f}}\right) \leq H\left(Y_{\mathsf{p}} | T_{\mathsf{p}}\right)$. The key step of the proof for the bound in (4.9f) is Lemma 1, which is proved in Appendix 4.A by using the Functional Dependence Graph (FDG) [102].

3. For the memoryless ISD-IC with output feedback $Y_{\mathsf{Fc}} = Y_{\mathsf{c}}$ in (4.2), from [57, model-(1000)] we have

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq I\left(Y_{\mathsf{c}}; X_{\mathsf{p}}, X_{\mathsf{c}}\right) + I\left(Y_{\mathsf{c}}; X_{\mathsf{c}} | Y_{\mathsf{p}}, X_{\mathsf{p}}\right) + I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{p}}\right). \tag{4.12}$$

To the best of our knowledge, (4.12) is the only upper bound of the type $R_{\mathsf{p}}+2R_{\mathsf{c}}$ available in the literature for the cooperative IC (which includes feedback models as a special case), but it is only valid for the case of output feedback. Our goal in the next section is to derive bounds of the type of (4.12) for the class of ISD CCICs described in Section 4.1.2 with independent noises at the different source-destination pairs as in (4.2).

4.3.2 Novel outer bounds

In this section we prove Theorem 8, i.e., we derive two novel outer bounds of the type $R_{\mathsf{p}}+2R_{\mathsf{c}}$ and $2R_{\mathsf{p}}+R_{\mathsf{c}}$ on the capacity region for the ISD CCIC described in Section 4.1.2 with independent noises at the different source-destination pairs. These two outer bounds generalize those of [48, Theorem 1], derived for the classical non-cooperative IC, to the CCIC. Note also that, when evaluated for the case of output feedback with independent noises, i.e., $T_{\mathsf{f}}=T_{\mathsf{p}}$, the outer bound in (4.7) reduces to the one in (4.12).

By Fano's inequality, by considering that the messages W_{p} and W_{c} are independent and by giving side information similarly to [47], we have

$$
\begin{aligned}
&N(2R_{\mathsf{p}}+R_{\mathsf{c}}-3\epsilon_N)\\
\leq&2I\left(W_{\mathsf{p}};{Y_{\mathsf{p}}}^N\right)+I\left(W_{\mathsf{c}};{Y_{\mathsf{c}}}^N\right)\\
\leq&I\left(W_{\mathsf{p}};{Y_{\mathsf{p}}}^N\right)+I\left(W_{\mathsf{p}};{Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{c}}\right)+I\left(W_{\mathsf{c}};{Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N\right)\\
=&H\left({Y_{\mathsf{p}}}^N\right)-H\left({Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{p}},W_{\mathsf{c}}\right) && (4.13a)\\
&+H\left({Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{c}}\right)-H\left({Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{c}}\right) && (4.13b)\\
&+H\left({Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N\right)-H\left({Y_{\mathsf{p}}}^N|W_{\mathsf{p}}\right). && (4.13c)
\end{aligned}
$$

We now analyze and bound each pair of terms.

Pair in (4.13a)**:** We have

$$
\begin{aligned}
&H\left({Y_{\mathsf{p}}}^N\right)-H\left({Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{p}},W_{\mathsf{c}}\right)\\
\overset{(a)}{=}&\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i}|{Y_{\mathsf{p}}}^{i-1}\right)-H\left(Y_{\mathsf{p}_i},T_{\mathsf{p}_i},T_{\mathsf{f}i}|W_{\mathsf{p}},W_{\mathsf{c}},{Y_{\mathsf{p}}}^{i-1},{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{f}}}^{i-1},{X_{\mathsf{p}}}^{i},{X_{\mathsf{c}}}^{i}\right)\\
\overset{(b)}{\leq}&\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i}\right)-H\left(Y_{\mathsf{p}_i},T_{\mathsf{p}_i},T_{\mathsf{f}i}|W_{\mathsf{p}},W_{\mathsf{c}},{Y_{\mathsf{p}}}^{i-1},{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{f}}}^{i-1},{X_{\mathsf{p}}}^{i},{X_{\mathsf{c}}}^{i}\right)\\
\overset{(c)}{=}&\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i}\right)-H\left(Y_{\mathsf{p}_i},T_{\mathsf{p}_i},T_{\mathsf{f}i}|X_{\mathsf{p}_i},X_{\mathsf{c}i}\right),
\end{aligned}
$$

where: the equality in (a) follows by applying the chain rule of the entropy and since, for the ISD CCIC, the encoding function $X_{\mathsf{c}i}(W_\mathsf{c}, Y_{\mathsf{Fc}}^{i-1})$ is equivalent to $X_{\mathsf{c}i}(W_\mathsf{c}, {T_\mathsf{f}}^{i-1})$ and since, given W_p, X_p is uniquely determined; the inequality in (b) is due to the conditioning reduces entropy principle; the equality in (c) follows because of the ISD property of the channel and since the channel is memoryless.

Pair in (4.13b)**:** We have

$$
\begin{aligned}
& H\left({Y_\mathsf{p}}^N, {T_\mathsf{p}}^N, {T_\mathsf{f}}^N | W_\mathsf{c}\right) - H\left({Y_\mathsf{c}}^N, {T_\mathsf{c}}^N, {T_\mathsf{f}}^N | W_\mathsf{c}\right) \\
\overset{\text{(d)}}{=} & \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {Y_\mathsf{p}}^{i-1}, {T_\mathsf{p}}^{i-1}, {T_\mathsf{f}}^{i-1}, W_\mathsf{c}, {X_\mathsf{c}}^i\right) \\
& - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | {Y_\mathsf{c}}^{i-1}, {T_\mathsf{c}}^{i-1}, {T_\mathsf{f}}^{i-1}, W_\mathsf{c}, {X_\mathsf{c}}^i\right) \\
\overset{\text{(e)}}{=} & \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {Y_\mathsf{p}}^{i-1}, {T_\mathsf{p}}^{i-1}, {T_\mathsf{f}}^{i-1}, W_\mathsf{c}, {X_\mathsf{c}}^i\right) \\
& - H\left(T_{\mathsf{p}_i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | {T_\mathsf{p}}^{i-1}, {T_\mathsf{c}}^{i-1}, {T_\mathsf{f}}^{i-1}, W_\mathsf{c}, {X_\mathsf{c}}^i\right) \\
\overset{\text{(f)}}{\leq} & \sum_{i\in[1:N]} \underbrace{H\left(T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {T_\mathsf{p}}^{i-1}, {T_\mathsf{f}}^{i-1}, W_\mathsf{c}, {X_\mathsf{c}}^i\right) - H\left(T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {T_\mathsf{p}}^{i-1}, {T_\mathsf{c}}^{i-1}, {T_\mathsf{f}}^{i-1}, W_\mathsf{c}, {X_\mathsf{c}}^i\right)}_{=\,0 \text{ because of (4.8b)}} \\
& + \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i} | T_{\mathsf{p}_i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}\right) - H\left(T_{\mathsf{c}i} | {T_\mathsf{p}}^i, {T_\mathsf{c}}^{i-1}, {T_\mathsf{f}}^i, W_\mathsf{c}, {X_\mathsf{c}}^i, {X_\mathsf{p}}^i\right) \\
\overset{\text{(g)}}{=} & \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i} | T_{\mathsf{p}_i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}\right) - H\left(Y_{\mathsf{p}_i} | T_{\mathsf{p}_i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}, X_{\mathsf{p}_i}\right),
\end{aligned}
$$

where: the equality in (d) follows by applying the chain rule of the entropy and since, for the ISD CCIC, the encoding function $X_{\mathsf{c}i}(W_\mathsf{c}, Y_{\mathsf{Fc}}^{i-1})$ is equivalent to $X_{\mathsf{c}i}(W_\mathsf{c}, {T_\mathsf{f}}^{i-1})$; the equality in (e) is due to the fact that Y_c is a deterministic function of $(X_\mathsf{c}, T_\mathsf{p})$, which is invertible given X_c; the inequality in (f) is due to the conditioning reduces entropy principle; the equality in (g) follows because of the ISD property of the channel and since the channel is memoryless.

Pair in (4.13c)**:** Since

$$
\begin{aligned}
& H\left({Y_\mathsf{p}}^N | W_\mathsf{p}\right) \\
\overset{\text{(h)}}{=} & \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i} | {Y_\mathsf{p}}^{i-1}, W_\mathsf{p}, {X_\mathsf{p}}^i\right) \overset{\text{(i)}}{=} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i} | {T_\mathsf{c}}^{i-1}, W_\mathsf{p}, {X_\mathsf{p}}^i\right)
\end{aligned}
$$

$$\stackrel{\text{(j)}}{\geq} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}{}^{i}, T_{\mathsf{f}}{}^{i-1}\right)$$
$$\stackrel{\text{(k)}}{=} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1}, T_{\mathsf{f}}{}^{i-1}\right) - \underbrace{I\left(T_{\mathsf{c}i}; W_{\mathsf{p}}, X_{\mathsf{p}}{}^{i}|T_{\mathsf{c}}{}^{i-1}, T_{\mathsf{f}}{}^{i-1}\right)}_{=0 \text{ because of (4.8a)}},$$

where: the equality in (h) follows by applying the chain rule of the entropy and since, given W_{p}, X_{p} is uniquely determined; the equality in (i) is due to the fact that Y_{p} is a deterministic function of $(X_{\mathsf{p}}, T_{\mathsf{c}})$, which is invertible given X_{p}; the inequality in (j) follows since conditioning reduces the entropy; the equality in (k) follows from the definition of mutual information; we have

$$H\left(Y_{\mathsf{c}}{}^{N}, T_{\mathsf{c}}{}^{N}, T_{\mathsf{f}}{}^{N}\right) - H\left(Y_{\mathsf{p}}{}^{N}|W_{\mathsf{p}}\right)$$
$$\stackrel{\text{(l)}}{\leq} \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i}|Y_{\mathsf{c}}{}^{i-1}, T_{\mathsf{c}}{}^{i-1}, T_{\mathsf{f}}{}^{i-1}\right) - H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1}, T_{\mathsf{f}}{}^{i-1}\right)$$
$$\stackrel{\text{(m)}}{\leq} \sum_{i\in[1:N]} H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1}, T_{\mathsf{f}}{}^{i-1}\right) - H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}{}^{i-1}, T_{\mathsf{f}}{}^{i-1}\right)$$
$$+ H\left(Y_{\mathsf{c}i}, T_{\mathsf{f}i}|Y_{\mathsf{c}}{}^{i-1}, T_{\mathsf{c}}{}^{i}, T_{\mathsf{f}}{}^{i-1}\right)$$
$$\stackrel{\text{(n)}}{\leq} \sum_{i\in[1:N]} 0 + H\left(Y_{\mathsf{c}i}, T_{\mathsf{f}i}, |T_{\mathsf{c}i}\right),$$

where the inequality in (l) is a consequence of the inequality in (k) above and the inequalities in (m) and (n) are due to the conditioning reduces entropy principle.

Final step: By combining everything together, by introducing the time sharing random variable uniformly distributed over $[1:N]$ and independent of everything else, by dividing both sides by N and taking the limit for $N \to \infty$ we get the bound in (4.6). We finally notice that by dropping the time sharing we do not decrease the bound. Note also that, since for the ISD model defined in (4.1) T_{p}, respectively T_{c}, is a deterministic function of $(Y_{\mathsf{c}}, X_{\mathsf{c}})$, respectively $(Y_{\mathsf{p}}, X_{\mathsf{p}})$, we have $H\left(T_{\mathsf{p}}, T_{\mathsf{f}}|Y_{\mathsf{p}}, X_{\mathsf{p}}, X_{\mathsf{c}}\right) = H\left(Y_{\mathsf{c}}, T_{\mathsf{f}}|T_{\mathsf{c}}, X_{\mathsf{p}}, X_{\mathsf{c}}\right)$.

By following similar steps, see Appendix 4.C, as in the derivation of (4.6) and by using the Markov chains in (4.8a) and (4.8b), one can derive the upper bound in (4.7).

4.3.3 Outer bounds evaluated for the Gaussian CCIC

We evaluate the bounds in (4.9), (4.6) and (4.7) for the Gaussian noise channel in (4.3). We define $\mathbb{E}\left[X_{\mathsf{p}}X_{\mathsf{c}}^{*}\right] := \rho : |\rho| \in [0,1]$. We also assume that all the noises are independent, which represents a particular case for which our outer bounds hold. By the 'Gaussian maximizes entropy' principle, jointly Gaussian inputs exhaust the outer bounds in (4.9), (4.6) and (4.7). Thus, we start by evaluating each mutual information term in (4.9), (4.6) and (4.7) by using jointly Gaussian inputs. Then, we further upper bound each mutual information term over the input correlation coefficient $\rho : |\rho| \in [0,1]$. By doing so, see Appendix 4.D, we obtain:

Lemma 2. *The capacity region of the Gaussian CCIC is contained into*

$$R_{\mathsf{p}} \leq \log\left(1+\mathsf{C}+\mathsf{S_p}\right), \tag{4.14a}$$

$$R_{\mathsf{p}} \leq \log\left(1+\left(\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}}\right)^2\right), \tag{4.14b}$$

$$R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}\right), \tag{4.14c}$$

$$\begin{aligned} R_{\mathsf{p}}+R_{\mathsf{c}} \leq{}& \log\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right)+\log\left(1+\left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right) \\ &+\log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}+\mathsf{S_p}}\right), \end{aligned} \tag{4.14d}$$

$$R_{\mathsf{p}}+R_{\mathsf{c}} \leq \log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right)+\log\left(1+\left(\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}}\right)^2\right), \tag{4.14e}$$

$$R_{\mathsf{p}}+R_{\mathsf{c}} \leq \log\left(1+\mathsf{C}+\mathsf{I_p}+\frac{\mathsf{S_p}}{1+\mathsf{I_c}}\right)+\log\left(1+\mathsf{I_c}+\mathsf{S_c}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I_p}}\right)+2\log(2), \tag{4.14f}$$

$$\begin{aligned} 2R_{\mathsf{p}}+R_{\mathsf{c}} \leq{}& \log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right)+\log\left(1+\left(\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}}\right)^2\right) \\ &+\log\left(\frac{1+\mathsf{I_p}+\mathsf{S_p}}{1+\mathsf{I_c}+\mathsf{S_c}}\right)+\log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}+\mathsf{S_p}}\right) \\ &+\log\left(1+\mathsf{I_c}+\mathsf{S_c}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I_p}}\right)+\log(2), \end{aligned} \tag{4.14g}$$

$$\begin{aligned} R_{\mathsf{p}}+2R_{\mathsf{c}} \leq{}& \log\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right)+\log\left(1+\left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right) \\ &+\log\left(\frac{1+\mathsf{I_c}+\mathsf{S_c}}{1+\mathsf{I_p}+\mathsf{S_p}}\right)+\log\left(1+\mathsf{C}+\mathsf{I_p}+\frac{\mathsf{S_p}}{1+\mathsf{I_c}}\right)+\log(2). \end{aligned} \tag{4.14h}$$

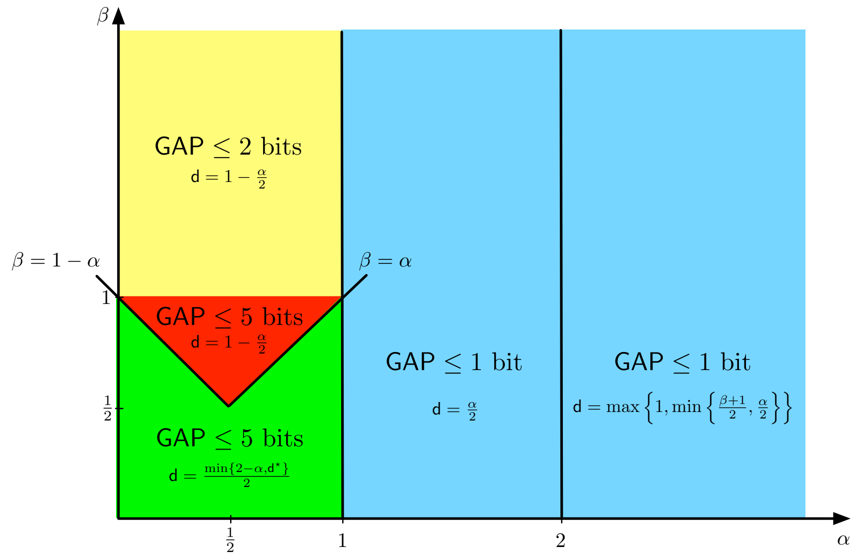


Figure 4.4: Different regimes depending on the values of α and β, with $\mathsf{d}^{\star} := \max\{\alpha, 1-\alpha\} + \max\{\alpha, 1+\beta-\max\{\alpha,\beta\}\}$.

4.4 The capacity region to within a constant gap for the symmetric Gaussian CCIC

In this section, we prove Theorem 9, i.e., we characterize the capacity to within a constant gap for the symmetric Gaussian CCIC defined by $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_p} = \mathsf{I_c} = \mathsf{I} = \mathsf{S}^{\alpha}$. Figure 4.4 shows the gDoF $\mathsf{d}(\alpha,\beta)$ (abbreviated with d) and the gap (per user) for the symmetric Gaussian CCIC for the different regions in the (α,β) plane, where the whole set of parameters has been partitioned into multiple sub-regions depending upon different levels of cooperation (β) and interference (α) strengths.

At a high level, the approximately optimal coding schemes are as follows. In the strong interference and weak cooperation regime both users employ a non-cooperative common message. In the strong interference and strong cooperation regime, PTx's common message becomes cooperative and is forwarded to PRx with the help of CTx. In the weak interference regime, each user splits its message into a common and a private part; for CTx the two message parts are non-cooperative while for PTx are both cooperative and non-cooperative depending on the strength of the cooperation link; in particular, PTx's cooperative private message is the 'known interference' against which CTx's message is precoded in a DPC-based scheme. Binning/DPC is used in the weak interference and 'sufficiently' strong cooperation regime where CTx can easily decode the signal from PTx because of the good quality of the cooperation link, but CRx cannot because of weak interference; therefore in this regime it makes sense that the best use of CTx's knowledge of PTx's message is to treat it as a 'known state' to precode its message against it.

We shall now discuss different operating regimes separately.

4.4.1 Regime 1 (strong interference I)

This regime corresponds to very strong interference (i.e., $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ or $\alpha \geq 2$) and weak cooperation (i.e., $\mathsf{C} \leq \mathsf{S}$ or $\beta \leq 1$), i.e., part of the blue region in Figure 4.4. In the non-cooperative IC with very strong interference it is exactly optimal to use only (non-cooperative) common messages in order to achieve the whole capacity region; since the interference is very strong, it can be decoded by treating the intended signal as noise, after which each receiver is left with an interference-free point-to-point channel from its transmitter; this non-cooperative strategy achieves

$$\mathcal{I}^{4.4.1}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{S}), \tag{4.15a}$$
$$R_{\mathsf{c}} \leq \log(1+\mathsf{S}), \tag{4.15b}$$

or $\mathsf{d}(\alpha,\beta) \leq (1+1)/2 = 1$. Since in this regime the cooperation link is weak, the amount of data PTx could communicate to CTx for cooperation is very limited. As a result in this regime unilateral cooperation does not improve the gDoF performance compared to the non-cooperative case. In other words cooperation provides a 'beamforming gain' but not a gDoF gain. To see this, the cut-set upper bounds on the individual rates in (4.14a) and (4.14c), in the symmetric case for $\beta \leq 1 \iff \mathsf{C} \leq \mathsf{S}$, give the following upper bounds on the individual rates

$$\mathcal{O}^{4.4.1}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{S}+\mathsf{C}) \leq \log(1+\mathsf{S}) + \log(2), \tag{4.16a}$$
$$R_{\mathsf{c}} \leq \log(1+\mathsf{S}). \tag{4.16b}$$

From the upper bound on R_{p} in (4.16a), we see that unilateral cooperation can at most double the SNR on the primary direct link, which can at most increase the rate by 1 bit compared to the non-cooperative case. As a result, the gDoF with unilateral cooperation is $\mathsf{d}(\alpha,\beta) = 1$ and the rate pair in (4.15) is optimal to within 1 bit, i.e., $\max\{\text{eq}(4.16\text{a}) - \text{eq}(4.15\text{a}), \text{eq}(4.16\text{b}) - \text{eq}(4.15\text{b})\} \leq \max\{\log(2), 0\} = 1$ bit.

4.4.2 Regime 2 (strong interference II)

In this regime the interference is very strong (i.e., $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ or $\alpha \geq 2$) and the cooperation is strong (i.e., $\mathsf{C} > \mathsf{S}$ or $\beta > 1$), i.e., part of the blue region in Figure 4.4. Similarly to the non-cooperative IC in very strong interference regime, the transmitters send a common message only. As opposed to regime 1, where both messages were sent non-cooperatively, here the PTx takes advantage of the strong cooperation link and sends its message to PRx with the help of the CTx. In order to enable cooperation, a block Markov coding scheme is used as follows. Transmission is over a frame of $B \gg 1$ slots. In slot $t \in [1 : B]$, the PTx sends its old (cooperative common) message $W_{\mathsf{p},t-1}$ and superposes to it the new (cooperative common) message $W_{\mathsf{p},t}$, while the CTx forwards the primary old (cooperative common) message $W_{\mathsf{p},t-1}$ and superposes to it its (non-cooperative common) message $W_{\mathsf{c},t}$. At the end of slot t, CTx decodes the new message $W_{\mathsf{p},t}$ after subtracting the contribution of the old message $W_{\mathsf{p},t-1}$. The destinations wait until the whole frame has been received and then proceed to jointly backward decode all messages. The details can be found in Appendix 4.E (more in particular in Appendix 4.E.1) and the achievable region is given in (4.51), which evaluated for $|a_1| = |c_1| = |b_2| = 0$ (since $U_1 = T_1 = T_2 = \emptyset$), and hence $|b_1| = |a_2| = 1$, for the symmetric Gaussian CCIC in very strong interference (note that all the constraints in (4.51), except (4.51b), (4.51c) and (4.51d), are redundant) reduces to

$$\mathcal{I}^{4.4.2}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{C}), \tag{4.17a}$$

$$R_{\mathsf{c}} \leq \log(1+\mathsf{S}), \tag{4.17b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log(1+\mathsf{S}+\mathsf{I}). \tag{4.17c}$$

The region in (4.17) is strictly larger than the non-cooperative capacity region in very strong interference given by (4.15) for $\mathsf{S}(1+\mathsf{S}) \leq \mathsf{I}$, or $\alpha \geq 2$, and $\mathsf{C} > \mathsf{S}$, or $\beta > 1$, which is precisely the definition of regime 2. The sum-capacity from (4.17) can take two possible values, depending on which one among the MAC sum-rate bound in (4.17c) and the sum of the bounds on the individual rates in (4.17a)-(4.17b) is the most stringent. In particular, the following sum-rate is achievable

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \begin{cases} \log(1+\mathsf{C}) + \log(1+\mathsf{S}) & \text{if } \mathsf{C}(1+\mathsf{S}) \leq \mathsf{I} \\ \log(1+\mathsf{S}+\mathsf{I}) & \text{if } \mathsf{C}(1+\mathsf{S}) > \mathsf{I} \end{cases},$$

that is, $\mathsf{d}(\alpha,\beta) \leq (\beta+1)/2$ if $\beta+1 \leq \alpha$ and $\mathsf{d}(\alpha,\beta) \leq \alpha/2$ otherwise; in both cases the gDoF is larger than the one of the non-cooperative IC

$\mathsf{d}(\alpha, 0) = 1$; moreover, when $\beta + 1 > \alpha$ the gDoF equals the one of the ideal non-causal CIC, i.e., $\mathsf{d}(\alpha, \beta) = \alpha/2$, i.e., unilateral source cooperation attains the ultimate performance limits of non-causal cognitive radio.

From the outer bound region obtained from the cut-set upper bounds on the individual rates in (4.14a) and (4.14c) and the sum-rate upper bound in (4.14e), under the condition $\beta > 1 \iff \mathsf{C} > \mathsf{S}$, we have that any achievable rate pair must satisfy

$$\mathcal{O}^{4.4.2}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{S}+\mathsf{C}) \leq \log(1+\mathsf{C}) + \log(2), \tag{4.18a}$$

$$R_{\mathsf{c}} \leq \log(1+\mathsf{S}), \tag{4.18b}$$

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) \\ &\overset{\mathsf{S}\leq\mathsf{I}}{\leq} \log\left(1+\mathsf{S}+\mathsf{I}\right) + 2\log(2), \end{aligned} \tag{4.18c}$$

since $(\sqrt{x}+\sqrt{y})^2 \leq 2(x+y), \ \forall (x,y) \in \mathbb{R}_+^2$. The upper bound in (4.18) and the achievable region in (4.17) are to within 1 bit of one another since

$$\begin{aligned} \mathsf{GAP} \leq \max \{&\text{eq}(4.18\text{a}) - \text{eq}(4.17\text{a}), \text{eq}(4.18\text{b}) - \text{eq}(4.17\text{b}), \\ &\left.\frac{\text{eq}(4.18\text{c}) - \text{eq}(4.17\text{c})}{2}\right\} \leq \log(2). \end{aligned}$$

This shows that the whole capacity region, and therefore the gDoF $\mathsf{d}(\alpha, \beta) = \min\{\beta+1, \ \alpha\}/2$ too, is achievable to within 1 bit in regime 2.

4.4.3 Regime 3 (strong interference III)

This regime corresponds to strong but not very strong interference (i.e., $\mathsf{S} \leq \mathsf{I} < \mathsf{S}(1+\mathsf{S})$ or $\alpha \in [1, 2)$), i.e., part of the blue region in Figure 4.4. Note that there are no restrictions on the cooperation exponent β in this regime. Similarly to regimes 1 and 2, here we use only common messages – a strategy that is capacity achieving in the corresponding non-cooperative IC. The difference between regime 1 and regime 3 is that stripping decoding is no longer optimal and the receivers must instead jointly decode the intended and non-intended messages as in a MAC. By taking the largest between the achievable region developed for regime 2 in (4.17) and the non-cooperative achievable region for this regime (i.e., common messages only), which has $R_{\mathsf{p}} \leq \log(1+\mathsf{S})$ as a bound on the primary rate rather than $R_{\mathsf{p}} \leq \log(1+\mathsf{C})$, we obtain the following achievable region

$$\begin{aligned}\mathcal{I}^{4.4.3}: \quad R_{\mathsf{p}} &\le \log(1+\max\{\mathsf{C},\mathsf{S}\}), && (4.19a)\\ R_{\mathsf{c}} &\le \log(1+\mathsf{S}), && (4.19b)\\ R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log(1+\mathsf{S}+\mathsf{I}), && (4.19c)\end{aligned}$$

which implies $\mathsf{d}(\alpha,\beta) \le \min\{1+\max\{1,\beta\},\ \max\{1,\alpha\}\}/2 = \alpha/2$, i.e., the sum-rate bound in (4.19c) is the tightest. Note also that in regime 3 $\mathsf{d}(\alpha,\beta) = \mathsf{d}(\alpha,0) = \mathsf{d}(\alpha,\infty)$, i.e., no matter how strong the cooperation link is, cooperation does not increase the gDoF of the non-cooperative IC.

From the outer bound region obtained from the cut-set upper bounds on the individual rates in (4.14a) and (4.14c) and the sum-rate upper bound in (4.14e), we have that any achievable rate pair must satisfy

$$\begin{aligned}\mathcal{O}^{4.4.3}: \quad R_{\mathsf{p}} &\le \log(1+\mathsf{S}+\mathsf{C})\\ &\le \log(1+\max\{\mathsf{C},\mathsf{S}\})+\log(2), && (4.20a)\\ R_{\mathsf{c}} &\le \log(1+\mathsf{S}), && (4.20b)\\ R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right)\\ &\le \log(1+\mathsf{S}+\mathsf{I})+2\log(2). && (4.20c)\end{aligned}$$

The upper bound in (4.20) and the achievable region in (4.19) are to within 1 bit of one another since

$$\begin{aligned}\mathsf{GAP} \le \max\Big\{&\mathrm{eq}(4.20\mathrm{a})-\mathrm{eq}(4.19\mathrm{a}), \mathrm{eq}(4.20\mathrm{b})-\mathrm{eq}(4.19\mathrm{b}),\\ &\frac{\mathrm{eq}(4.20\mathrm{c})-\mathrm{eq}(4.19\mathrm{c})}{2}\Big\} \le \log(2).\end{aligned}$$

This shows that the whole capacity region, and therefore the gDoF $\mathsf{d}(\alpha,\beta) = \alpha/2$ too, is achievable to within 1 bit in regime 3.

4.4.4 Regime 4 (weak interference I)

In this regime the interference is weak ($\mathsf{I} < \mathsf{S}$ or $\alpha < 1$) and the cooperation is strong (i.e., $\mathsf{C} > \mathsf{S}$ or $\beta > 1$), i.e., the cooperation link is stronger than both the interfering and the direct links (yellow region in Figure 4.4). The PTx takes advantage of the strong cooperation link and sends its message to the PRx with the help of the CTx, i.e., the messages of the PTx are only cooperative. Moreover, since the interference is weak, the messages of the CTx and of the PTx are both common and private. Also for this regime we use binning at the CTx. In other words, for this regime, the PTx does not make use of the non-cooperative messages, i.e., with reference to the transmission strategy in Appendix 4.E, we set $U_1 = T_1 = \emptyset$, i.e., we set $|d_1| = 0$ in the achievable rate region in (4.53) in Appendix 4.E.2. Moreover, motivated by the observation in [12] that all the terms that appear as noise should be at most at the level of the noise, we set $|a_1|^2 = |b_1|^2 = \frac{\mathsf{I}}{2(1+\mathsf{I})}$, $|b_2|^2 = \frac{\mathsf{I}}{1+\mathsf{I}}$ and $|c_1|^2 = |c_2|^2 = \frac{1}{1+\mathsf{I}}$ in the achievable rate region in (4.53). With these choices the achievable rate region in (4.53), evaluated for the symmetric channel, can be further lower bounded (by considering $\min\{k_1, k_2\} \geq 0$ and that the sum-rate constraint in (4.53e) is redundant in (4.53)) as

$$\mathcal{I}^{4.4.4}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{C}) - \log(2), \tag{4.21a}$$
$$R_{\mathsf{p}} \leq \log(1+\mathsf{S}+\mathsf{I}) - \log(2), \tag{4.21b}$$
$$R_{\mathsf{c}} \leq \log(1+\mathsf{S}) - \log(2), \tag{4.21c}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log(1+\mathsf{C}) + \log(1+\mathsf{S}+\mathsf{I}) - 3\log(2), \tag{4.21d}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\frac{\mathsf{C}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}) - \log(2), \tag{4.21e}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}) - 2\log(2), \tag{4.21f}$$
$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log(1+\mathsf{S}) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}+\mathsf{I}) - 4\log(2), \tag{4.21g}$$
$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(1+\frac{\mathsf{C}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}+\mathsf{I}) + \log(1+\mathsf{S}) - 3\log(2), \tag{4.21h}$$
$$R_{\mathsf{p}} + 3R_{\mathsf{c}} \leq 2\log(1+\mathsf{S}+\mathsf{I}) + \log(1+\mathsf{S}) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 6\log(2), \tag{4.21i}$$

which implies $\mathsf{d}(\alpha,\beta) \leq \min\{\beta+1, 2, \beta-\alpha+1, 2-\alpha\}/2 = 1-\alpha/2$, i.e., the sum-rate bound in (4.21f) is the tightest. In this regime the gDoF is larger than the one of the classical IC $\mathsf{d}(\alpha,0) = \min\{\max\{1-\alpha,\alpha\}, 1-\alpha/2\}$ everywhere except for $2/3 \leq \alpha < 1$; moreover, the gDoF equals everywhere the one of the ideal non-causal CIC, i.e., $\mathsf{d}(\alpha,\beta) = 1-\alpha/2$, i.e., unilateral source cooperation attains the ultimate performance limits of non-causal cognitive radio.

From the outer bound region obtained from the cut-set upper bounds on the individual rates in (4.14a), (4.14b) and (4.14c) and the sum-rate upper bound in (4.14e), we have that any achievable rate pair must satisfy

$$\mathcal{O}^{4.4.4}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{S}+\mathsf{C}) \leq \log(1+\mathsf{C}) + \log(2), \tag{4.22a}$$

$$R_{\mathsf{p}} \leq \log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) \leq \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log(2), \tag{4.22b}$$

$$R_{\mathsf{c}} \leq \log(1+\mathsf{S}), \tag{4.22c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}+\mathsf{I}) + \log(2). \tag{4.22d}$$

The upper bound in (4.22) and the achievable region in (4.21) are to within 2 bits of one another since

$$\begin{aligned} \mathsf{GAP} \leq \max \{ & \mathrm{eq}(4.22\mathrm{a}) - \mathrm{eq}(4.21\mathrm{a}), \mathrm{eq}(4.22\mathrm{b}) - \mathrm{eq}(4.21\mathrm{b}), \\ & \mathrm{eq}(4.22\mathrm{c}) - \mathrm{eq}(4.21\mathrm{c}), \frac{\mathrm{eq}(4.22\mathrm{a}) + \mathrm{eq}(4.22\mathrm{c}) - \mathrm{eq}(4.21\mathrm{d})}{2}, \\ & \frac{\mathrm{eq}(4.22\mathrm{d}) - \mathrm{eq}(4.21\mathrm{e})}{2}, \frac{\mathrm{eq}(4.22\mathrm{d}) - \mathrm{eq}(4.21\mathrm{f})}{2}, \\ & \frac{\mathrm{eq}(4.22\mathrm{d}) + \mathrm{eq}(4.22\mathrm{c}) - \mathrm{eq}(4.21\mathrm{g})}{3}, \\ & \frac{\mathrm{eq}(4.22\mathrm{d}) + \mathrm{eq}(4.22\mathrm{c}) - \mathrm{eq}(4.21\mathrm{h})}{3}, \\ & \left.\frac{\mathrm{eq}(4.22\mathrm{d}) + 2\mathrm{eq}(4.22\mathrm{c}) - \mathrm{eq}(4.21\mathrm{i})}{4} \right\} \leq 2\log(2). \end{aligned}$$

This shows that the whole capacity region, and therefore the gDoF $\mathsf{d}(\alpha, \beta) = 1 - \alpha/2$ too, is achievable to within 2 bits in regime 4.

4.4.5 Regime 5 (weak interference II)

In this regime the interference is weak (i.e., $\mathsf{I} < \mathsf{S}$ or $\alpha < 1$) and the cooperation link is stronger than the interfering link, but weaker than the direct link, i.e., $\max\left\{\mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}}\right\} \leq \mathsf{C} \leq \mathsf{S}$ or $\max\{\alpha, 1-\alpha\} \leq \beta \leq 1$ (red region in Figure 4.4). Thus, on the one hand the PTx takes advantage of these channel conditions by using cooperative messages; on the other hand, the cooperation link is not strong enough to allow the CTx to fully decode the PTx's message, and hence the PTx also uses a non-cooperative message. In particular, the non-cooperative message of the PTx is private; this is because the interference is too weak and forcing the CRx to fully decode the PTx's message would constrain the rate too much. The CTx can also benefit from the strength of the cooperation link and boost its rate performance by 'smartly' precoding its message against the private cooperative message of the PTx, i.e., the scheme is based both on superposition and binning. Thus, for this regime, the PTx does not make use of the common non-cooperative message, i.e., with reference to the transmission strategy in Appendix 4.E, we set $U_1 = \emptyset$ (see Appendix 4.E.2). After Fourier-Motzkin Elimination (FME) we obtain the rate region in (4.52), which evaluated for the practically relevant Gaussian noise case gives (4.53) in Appendix 4.E.2. In (4.53) we further set $|a_2| = 0$ and $|c_2|^2 = 1 - |b_2|^2 = \frac{1}{1+\mathsf{I}}$ so that the private message of the CTx (conveyed by T_2) is received below the noise level at the PRx in the spirit of [12]. Note also that the CTx does not cooperate with the PTx in conveying information to the PRx, but it just exploits the information it learns through the cooperation link to smartly pre-encode its messages. For the PTx we let $|a_1|^2 = |b_1|^2 = \frac{\mathsf{I}+\mathsf{C}+2\mathsf{IC}}{4(1+\mathsf{I})(1+\mathsf{C})}$, $|c_1|^2 = \frac{1}{2(1+\mathsf{I})}$ and $|d_1|^2 = \frac{1}{2(1+\mathsf{C})}$ in (4.53); with this choice of the power splits and since we are in the regime $\mathsf{C} > \mathsf{I}$, the two private messages of the PTx (i.e., the cooperative one carried by Z_1 and the non-cooperative one carried by T_1) are received at most at the level of the noise at the CRx. Moreover, the non-cooperative private message (carried by T_1) is received at the level of the noise at the CTx. With these choices we get that the achievable rate region in (4.53) is contained into (by considering $\min\{k_1, k_2\} \geq 0$ in (4.53))

$$
\begin{aligned}
\mathcal{I}^{4.4.5}: \quad & R_{\mathsf{p}} \leq \log(1+\mathsf{C}+\mathsf{S}) - 5\log(2), && \text{(4.23a)}\\
& R_{\mathsf{p}} \leq \log(1+\mathsf{S}+\mathsf{I}) - \log(2), && \text{(4.23b)}\\
& R_{\mathsf{c}} \leq \log(1+\mathsf{S}) - \log(3), && \text{(4.23c)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} \leq{} & \log(1+\mathsf{C}) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{C}}+\mathsf{I}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) \\
& - 5\log(2) - \log(3), && \text{(4.23d)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} \leq{} & \log(1+\mathsf{S}+\mathsf{I}) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - \log(2) - \log(3), && \text{(4.23e)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} \leq{} & \log\left(1+\frac{\mathsf{C}}{1+\mathsf{I}}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{C}}\right) + \log(1+\mathsf{S}) \\
& - 4\log(2) - \log(3), && \text{(4.23f)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} \leq{} & \log(1+\mathsf{S}) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}+\frac{\mathsf{S}}{1+\mathsf{C}}\right) \\
& - 2\log(2) - \log(3), && \text{(4.23g)}
\end{aligned}
$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \le \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{S}\right) - 3\log(2) - 2\log(3), \tag{4.23h}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \le \log\left(1+\mathsf{S}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{C}}\right) + \log\left(1+\frac{\mathsf{C}}{1+\mathsf{I}}\right) - 4\log(2) - 2\log(3), \tag{4.23i}$$

$$R_{\mathsf{p}} + 3R_{\mathsf{c}} \le \log\left(1+\mathsf{S}+\mathsf{I}\right) + 2\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{C}}\right) + \log\left(1+\mathsf{S}\right) - 5\log(2) - 3\log(3), \tag{4.23j}$$

which implies $\mathsf{d}\left(\alpha,\beta\right) \le \min\{\beta+1, 2, 2-\alpha\}/2 = 1-\alpha/2$, i.e., the sum-rate bound in (4.23e) is the tightest. In this regime the gDoF is larger than the one of the classical IC $\mathsf{d}\left(\alpha,0\right) = \min\left\{\max\left\{1-\alpha,\alpha\right\}, 1-\alpha/2\right\}$ everywhere except for $2/3 \le \alpha < 1$; moreover, the gDoF equals everywhere the one of the ideal non-causal CIC, i.e., $\mathsf{d}\left(\alpha,\beta\right) = 1-\alpha/2$, i.e., unilateral source cooperation attains the ultimate performance limits of non-causal cognitive radio.

From the outer bound region obtained from the cut-set upper bounds on the individual rates in (4.14a), (4.14b) and (4.14c), the sum-rate upper bound in (4.14e) and the upper bound on $R_{\mathsf{p}}+2R_{\mathsf{c}}$ in (4.14h), we have that any achievable rate pair must satisfy

$$\mathcal{O}^{4.4.5}: \quad R_{\mathsf{p}} \le \log(1+\mathsf{S}+\mathsf{C}), \tag{4.24a}$$

$$R_{\mathsf{p}} \le \log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) \le \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log(2), \tag{4.24b}$$

$$R_{\mathsf{c}} \le \log(1+\mathsf{S}), \tag{4.24c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}+\mathsf{I}) + \log(2), \tag{4.24d}$$

$$\begin{aligned} R_{\mathsf{p}} + 2R_{\mathsf{c}} &\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) \\ &\quad + \log\left(1+\mathsf{C}+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(2) \\ &\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{I}+\mathsf{S}\right) + \log\left(1+\mathsf{C}\right) + 4\log(2), \end{aligned} \tag{4.24e}$$

where the last inequality follows since $\mathsf{I} + \frac{\mathsf{S}}{1+\mathsf{I}} \le 2\max\left\{\mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}}\right\} \le 2\mathsf{C}$.

The upper bound in (4.24) and the achievable region in (4.23) are to within 5 bits of one another since

$$\begin{aligned}\mathsf{GAP} \leq \max\Big\{&\text{eq}(4.24\text{a}) - \text{eq}(4.23\text{a}), \text{eq}(4.24\text{b}) - \text{eq}(4.23\text{b}),\\ &\text{eq}(4.24\text{c}) - \text{eq}(4.23\text{c}), \frac{\text{eq}(4.24\text{d}) - \text{eq}(4.23\text{d})}{2},\\ &\frac{\text{eq}(4.24\text{d}) - \text{eq}(4.23\text{e})}{2}, \frac{\text{eq}(4.24\text{d}) - \text{eq}(4.23\text{f})}{2},\\ &\frac{\text{eq}(4.24\text{d}) - \text{eq}(4.23\text{g})}{2}, \frac{\text{eq}(4.24\text{d}) + \text{eq}(4.24\text{c}) - \text{eq}(4.23\text{h})}{3},\\ &\frac{\text{eq}(4.24\text{e}) - \text{eq}(4.23\text{i})}{3}, \frac{\text{eq}(4.24\text{e}) + \text{eq}(4.24\text{c}) - \text{eq}(4.23\text{j})}{4}\Big\}\\ \leq\ & 5\log(2).\end{aligned}$$

This shows that the whole capacity region, and therefore the gDoF $\mathsf{d}(\alpha, \beta) = 1 - \alpha/2$ too, is achievable to within 5 bits in regime 5.

4.4.6 Regime 6 (weak interference III)

In this regime the interference is weak (i.e., $\mathsf{I} < \mathsf{S}$ or $\alpha < 1$) and the cooperation link is also weak, i.e., $\mathsf{C} \leq \max\left\{\mathsf{I}, \frac{\mathsf{S}}{1+\mathsf{I}}\right\}$ or $\beta \leq \max\{\alpha, 1-\alpha\}$ (green region in Figure 4.4); we hence expect the CCIC to 'behave' as the classical non-cooperative IC [12] for which both private and common non-cooperative messages are approximately optimal. Differently from the classical IC, the PTx also conveys part of its message through the CTx. This cooperative message is common, and thus also decoded at the CRx. Actually, since the cooperation link is weak, the amount of information that can be decoded, and hence delivered, by the CTx is limited. Thus, there is no need to employ binning, i.e., the scheme is based on superposition coding only. In other words, with reference to the transmission strategy in Appendix 4.E, we set $S_1 = Z_1 = \emptyset$ (see Appendix 4.E.1). After FME we obtain the rate region in (4.50), which evaluated for the practically relevant Gaussian noise case gives (4.51) in Appendix 4.E.1. In (4.51) we further set $|b_2|^2 = 1 - |a_2|^2 = \frac{1}{1+\mathsf{I}}$ so that the private message of CTx (conveyed by T_2) is received below the noise level at the PRx in the spirit of [12]. Regarding the choice of the power splits for the PTx, we further split the green region into two subregions: subregion (i) for which $\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$ (i.e., $\beta \leq [2\alpha - 1]^+$) and subregion (ii) for which $\mathsf{C} > \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$ (i.e., $\beta > [2\alpha - 1]^+$). We now analyze these two subregions separately.

Subregion (i): when $\beta \leq [2\alpha-1]^+$, the cooperation link is very weak and thus we expect the Gaussian CCIC to behave as the non-cooperative Gaussian IC [12]. Therefore, we set the power of the cooperative common message (carried by V_1) to $|b_1|^2 = 0$ in (4.51) and $|c_1|^2 = 1-|a_1|^2 = \frac{1}{1+\mathsf{I}}$ so that the private message of PTx (conveyed by T_1) is received below the noise level at the CRx in the spirit of [12]. With these choices and by removing the redundant constraints in (4.51) (i.e., eq(4.51a), eq(4.51e), eq(4.51f), eq(4.51g), eq(4.51j) and eq(4.51l)), we get that the achievable rate region in (4.51) is contained into

$$
\begin{aligned}
\mathcal{I}^{4.4.6(i)}: \quad & R_{\mathsf{p}} \leq \log\left(1+\mathsf{S}\right) - \log(2), && (4.25a)\\
& R_{\mathsf{c}} \leq \log\left(1+\mathsf{S}\right) - \log(2), && (4.25b)\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2), && (4.25c)\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2), && (4.25d)\\
2R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{S}+\mathsf{I}\right)\\
&\quad + \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 3\log(2), && (4.25e)\\
R_{\mathsf{p}} + 2R_{\mathsf{c}} &\leq \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)\\
&\quad + \log\left(1+\mathsf{S}+\mathsf{I}\right) - 3\log(2), && (4.25f)
\end{aligned}
$$

which implies $\mathsf{d}(\alpha,\beta) \leq \min\left\{2, 2-\alpha, \max\left\{2\alpha, 2-2\alpha\right\}\right\}/2 = \alpha$, i.e., the sum-rate bound in (4.25d) is the tightest. Note that the rate region in (4.25) is (up to a constant gap) the achievable rate region for the classical symmetric non-cooperative IC in weak interference, which is optimal up to a gap of 1 bit/user [12]. This implies that for this regime $\mathsf{d}(\alpha,\beta) = \mathsf{d}(\alpha,0)$.

For this regime, we have that any achievable rate pair must satisfy (by considering all the constraints except (4.14b) and (4.14d))

$$
\begin{aligned}
\mathcal{O}^{4.4.6(i)}: \quad & R_{\mathsf{p}} \leq \log(1+\mathsf{S}+\mathsf{C}) \leq \log(1+\mathsf{S}) + \log(2), && (4.26a)\\
& R_{\mathsf{c}} \leq \log(1+\mathsf{S}), && (4.26b)\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right)\\
&\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}+\mathsf{I}) + \log(2), && (4.26c)
\end{aligned}
$$

$$
\begin{aligned}
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\mathsf{C}+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{I}+\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I}}\right)+2\log(2)\\
&\stackrel{0\leq \mathsf{C}\leq \mathsf{I}}{\leq} \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{I}+\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{I}}\right)+3\log(2)\\
&\stackrel{\mathsf{C}\leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}}{\leq} 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+4\log(2), && (4.26\text{d})\\
2R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{S}+\mathsf{I}\right)\\
&\quad +\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+4\log(2), && (4.26\text{e})\\
R_{\mathsf{p}} + 2R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right)\\
&\quad +\log\left(1+\mathsf{C}+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log(2)\\
&\leq \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{S}+\mathsf{I}\right)\\
&\quad +\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+3\log(2). && (4.26\text{f})
\end{aligned}
$$

The upper bound in (4.26) and the achievable region in (4.25) are to within 3 bits of one another since

$$
\begin{aligned}
\mathsf{GAP} \leq \max\Big\{&\mathrm{eq}(4.26\mathrm{a})-\mathrm{eq}(4.25\mathrm{a}), \mathrm{eq}(4.26\mathrm{b})-\mathrm{eq}(4.25\mathrm{b}),\\
&\frac{\mathrm{eq}(4.26\mathrm{c})-\mathrm{eq}(4.25\mathrm{c})}{2}, \frac{\mathrm{eq}(4.26\mathrm{d})-\mathrm{eq}(4.25\mathrm{d})}{2},\\
&\frac{\mathrm{eq}(4.26\mathrm{e})-\mathrm{eq}(4.25\mathrm{e})}{3}, \frac{\mathrm{eq}(4.26\mathrm{f})-\mathrm{eq}(4.25\mathrm{f})}{3}\Big\}\\
\leq 3\log(2).&
\end{aligned}
$$

This shows that the whole capacity region, and therefore the gDoF $\mathsf{d}(\alpha,\beta) = \alpha$ too, is achievable to within 3 bits in regime 6 subregion (i).

Subregion (ii): when $\beta > [2\alpha-1]^+$, the Gaussian CCIC starts to benefit from cooperation and indeed the outer bound region depends on C. Therefore the cooperative common message carried by V_1 can boost the rate performance of the system. In (4.51), we set the power of the common non-cooperative message (carried by U_1) to $|a_1|^2 = \frac{1}{2(1+\min\{\mathsf{C},\mathsf{I}\})}$. This choice is motivated by the fact that, in order to approximately match the outer bound, the single rate constraint on R_{p} in (4.51b) must approximately behave as an interference-free point-to-point channel. Therefore, the fact that CTx can now decode part of the message of PTx (carried by V_1) must not limit (up to a constant gap) the performance of the PTx. In other words, since C is 'quite large' but not 'huge', the rate of V_1 cannot be too large. Moreover, we set $|c_1|^2 = \frac{1}{2(1+\mathsf{I})}$ so that the private message of PTx (conveyed by T_1) is received below the noise level at the CRx in the spirit of [12]. Thus, if $\mathsf{I} \leq \mathsf{C}$ we have $|b_1|^2 = \frac{\mathsf{I}}{1+\mathsf{I}}$, while if $\mathsf{I} > \mathsf{C}$ we have $|b_1|^2 = \frac{\mathsf{C}+\mathsf{I}+2\mathsf{CI}}{2(1+\mathsf{C})(1+\mathsf{I})}$. With these choices and by removing the redundant constraints in (4.51) (i.e., eq(4.51a), eq(4.51d), eq(4.51f), eq(4.51g), and eq(4.51m)), we get that the achievable rate region in (4.51) is contained into

$$
\begin{aligned}
\mathcal{I}^{4.4.6(ii)}: \quad R_{\mathsf{p}} &\leq \log\left(1+\mathsf{S}\right) - 4\log(2), && \text{(4.27a)}\\
R_{\mathsf{c}} &\leq \log\left(1+\mathsf{S}\right) - \log(2), && \text{(4.27b)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 3\log(2), && \text{(4.27c)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)\\
&\quad + \log\left(1+\min\left\{\mathsf{I},\mathsf{C}\right\}\right) - 5\log(2), && \text{(4.27d)}\\
2R_{\mathsf{p}} + R_{\mathsf{c}} &\leq 2\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{S}+\mathsf{I}\right)\\
&\quad + \log\left(1+\min\{\mathsf{I},\mathsf{C}\}\right) - 6\log(2), && \text{(4.27e)}\\
2R_{\mathsf{p}} + R_{\mathsf{c}} &\leq 2\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\min\{\mathsf{I},\mathsf{C}\}}\right)\\
&\quad + 2\log\left(1+\min\{\mathsf{I},\mathsf{C}\}\right) - 9\log(2), && \text{(4.27f)}\\
R_{\mathsf{p}} + 2R_{\mathsf{c}} &\leq \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)\\
&\quad + \log\left(1+\mathsf{S}+\mathsf{I}\right) - 4\log(2), && \text{(4.27g)}
\end{aligned}
$$

which implies $\mathsf{d}\left(\alpha,\beta\right) \leq \min\left\{2-\alpha, \max\left\{1, 2-2\alpha\right\} + \min\{\alpha,\beta\}\right\}/2$, i.e., the minimum between the sum-rate bound in (4.27c) and the one in (4.27d) is the tightest.

For this regime, we have that any achievable rate pair must satisfy (by considering all the constraints except (4.14b) and (4.14d))

$$\mathcal{O}^{4.4.6(ii)}: \quad R_{\mathsf{p}} \le \log(1+\mathsf{S}+\mathsf{C}) \le \log(1+\mathsf{S})+\log(2), \tag{4.28a}$$

$$R_{\mathsf{c}} \le \log(1+\mathsf{S}), \tag{4.28b}$$

$$\begin{aligned}
R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right)\\
&\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log(1+\mathsf{S}+\mathsf{I})+\log(2),
\end{aligned} \tag{4.28c}$$

$$\begin{aligned}
R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\mathsf{C}+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{I}+\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I}}\right)+2\log(2)\\
&\overset{\mathsf{C}\le\max\left\{\mathsf{I},\frac{\mathsf{S}}{1+\mathsf{I}}\right\}}{\le} \log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{I}+\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{I}+\mathsf{C}}\right)\\
&\quad+3\log(2),
\end{aligned} \tag{4.28d}$$

$$\begin{aligned}
2R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right)\\
&\quad+\log\left(1+\frac{\mathsf{C}}{1+\mathsf{I}+\mathsf{S}}\right)+\log\left(1+\mathsf{I}+\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I}}\right)+\log(2)\\
&\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{S}+\mathsf{I}\right)\\
&\quad+\log\left(1+\mathsf{I}+\mathsf{S}\frac{1+\mathsf{C}}{1+\mathsf{C}+\mathsf{I}}\right)+3\log(2),
\end{aligned} \tag{4.28e}$$

$$\begin{aligned}
R_{\mathsf{p}}+2R_{\mathsf{c}} &\le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right)\\
&\quad+\log\left(1+\mathsf{C}+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log(2)\\
&\overset{\mathsf{C}\le\max\left\{\mathsf{I},\frac{\mathsf{S}}{1+\mathsf{I}}\right\}}{\le} \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+\log\left(1+\mathsf{S}+\mathsf{I}\right)\\
&\quad+\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+3\log(2).
\end{aligned} \tag{4.28f}$$

The upper bound in (4.28) and the achievable region in (4.27) are to within

5 bits of one another since

$$\begin{aligned}\mathsf{GAP} \le \max\Big\{&\text{eq}(4.28\text{a}) - \text{eq}(4.27\text{a}), \text{eq}(4.28\text{b}) - \text{eq}(4.27\text{b}),\\ &\frac{\text{eq}(4.28\text{c}) - \text{eq}(4.27\text{c})}{2}, \frac{\text{eq}(4.28\text{d}) - \text{eq}(4.27\text{d})}{2},\\ &\frac{\text{eq}(4.28\text{e}) - \text{eq}(4.27\text{e})}{3}, \frac{\text{eq}(4.28\text{e}) - \text{eq}(4.27\text{f})}{3},\\ &\frac{\text{eq}(4.28\text{f}) - \text{eq}(4.27\text{g})}{3}\Big\} \le 5\log(2).\end{aligned}$$

This shows that the whole capacity region, and therefore the gDoF too, is achievable to within 5 bits in regime 6 subregion (ii).

4.4.7 Implication of the gap result

In the symmetric case, the two novel outer bounds $2R_{\mathsf{p}} + R_{\mathsf{c}}$ in (4.14g) and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ in (4.14h) are active when $\mathsf{S} \ge \max\{\mathsf{C}, \mathsf{I}\}$, which corresponds to the red and green regions in Figure 4.4. In [56], the authors interpreted the need of this type of bounds as a measure of the amount of the 'resource holes', or inefficiency, due to the distributed nature of the non-cooperative classical IC [12]. Thus, in line with the work in [56], we conclude that when $\mathsf{S} \ge \max\{\mathsf{C}, \mathsf{I}\}$ unilateral cooperation is too weak to allow for a full utilization of the channel resources, i.e., it leaves some system resources underutilized.

We conclude this section with some observations on the regimes where the bounds $2R_{\mathsf{p}} + R_{\mathsf{c}}$ in (4.14g) and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ in (4.14h) are active for the symmetric Gaussian CCIC.

- **Strong interference** (i.e., $\mathsf{I} > \mathsf{S}$, blue region in Figure 4.4): in this regime both the capacity region of the non-cooperative Gaussian IC [12] and the capacity region of the Gaussian non-causal CIC [64], do not have bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$; since the capacity region of the Gaussian CCIC is 'sandwiched' between these two regions, it hence follows that these bounds are not necessary. In other words, in this regime, even for the non-cooperative Gaussian IC the system resources are fully utilized [56].

- **Weak interference and strong cooperation** (i.e., $\mathsf{I} \le \mathsf{S} < \mathsf{C}$, yellow region in Figure 4.4): for this regime the outer bound in (4.22) equals (to within a constant gap) the outer bound on the capacity region for the non-causal Gaussian CIC [64, Theorem III.1], which does not have bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$. In other words, in this regime the ideal non-causal cognition assumption at the CTx just provides a 'beamforming gain' compared to the more practical case of causal learning for the CTx through a noisy link. It hence follows that for this regime unilateral cooperation allows to fully utilize the channel resources [56], i.e., the bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ are not active.

- **Weak interference and weak cooperation** (i.e., $\mathsf{S} \geq \max\{\mathsf{C}, \mathsf{I}\}$, green and red regions in Figure 4.4): for this regime the capacity region of the non-cooperative Gaussian IC has bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ [12], while the one of the non-causal Gaussian CIC does not [64]. From our constant gap result in these two regions, it follows that: (i) $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ in (4.14h) is active in the red region, (ii) $2R_{\mathsf{p}} + R_{\mathsf{c}}$ in (4.14g) and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ in (4.14h) are both active in the green region. In other words, in this regime unilateral cooperation does not allow enough coordination among the sources which results in some 'resource holes' as in the non-cooperative Gaussian IC.

Remark 15. From our discussion, it follows that causal unilateral source cooperation does not improve on the gDoF of the non-cooperative Gaussian IC, i.e., $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, 0)$, when

$$\alpha \in \left[\frac{2}{3}, 2\right] \text{ or } \beta \leq \min\left\{1, [2\alpha - 1]^+\right\}.$$

Thus, for this set of parameters, unilateral cooperation might not be worth implementing in practical systems since the same gDoF is achieved without explicit cooperation, i.e., unilateral cooperation only provides a power / beamforming gain. Moreover, the gDoF of the Gaussian CCIC is equal to that of the non-causal Gaussian CIC, i.e., $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, \infty)$, everywhere except for

$$\alpha \leq \frac{2}{3} \text{ with } \beta \leq \min\{\alpha, 1 - \alpha\} \text{ and } \alpha \geq \max\{2, \beta + 1\}.$$

For this set of parameters unilateral cooperation attains the ultimate performance limits of non-causal cognitive radio and hence represents the ideal channel condition for cognitive radio.

4.5 The capacity region to within a constant gap for the Gaussian Z-channel

In this section, we prove Theorem 10, i.e., we characterize the capacity to within a constant gap for the Z Gaussian CCIC defined by $\mathsf{I_p} = 0$. In particular, we show that the upper bound

$$R_{\mathsf{p}} \le \log\left(1 + \mathsf{C} + \mathsf{S_p}\right), \tag{4.29a}$$

$$R_{\mathsf{p}} \le \log\left(1 + \left(\sqrt{\mathsf{S_p}} + \sqrt{\mathsf{I_c}}\right)^2\right), \tag{4.29b}$$

$$R_{\mathsf{c}} \le \log\left(1 + \mathsf{S_c}\right), \tag{4.29c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1 + \frac{\mathsf{S_c}}{1 + \mathsf{I_c}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S_p}} + \sqrt{\mathsf{I_c}}\right)^2\right), \tag{4.29d}$$

obtained from (4.14) by setting $\mathsf{I_p} = 0$, is achievable to within a constant gap. The region in (4.29) without the bound in (4.29a) (i.e., the only one that depends on C) is the capacity upper bound for the non-causal cognitive IC in [64, Theorem III.1], which unifies previously known outer bounds for the weak ($\mathsf{S_c} > \mathsf{I_c}$) and strong ($\mathsf{S_c} \le \mathsf{I_c}$) interference regimes and is achievable to within 1 bit. Hence, we interpret the bound in (4.29a) as the 'cost' of causal cooperation on the Z-channel.

Moreover, as we shall see later in more details, the capacity region of the Z-channel, differently from that of the symmetric Gaussian CCIC, does not have bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$.

4.5.1 Case $\mathsf{C} \le \mathsf{S_p}$: when unilateral cooperation might not be useful

For the case $\mathsf{C} \le \mathsf{S_p}$ we further upper bound the capacity outer bound in (4.29) as

$$\mathcal{O}^{4.5.1}: \quad R_{\mathsf{p}} \le \log\left(1 + \mathsf{S_p}\right) + \log(2), \tag{4.30a}$$

$$R_{\mathsf{c}} \le \log\left(1 + \mathsf{S_c}\right), \tag{4.30b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1 + \frac{\mathsf{S_c}}{1 + \mathsf{I_c}}\right) + \log\left(1 + \mathsf{S_p} + \mathsf{I_c}\right) + \log(2). \tag{4.30c}$$

The region in (4.30) is at most 1 bit away from

$$\begin{aligned}\mathcal{I}^{4.5.1}: \quad R_{\mathsf{p}} &\le \log\left(1+\mathsf{S}_{\mathsf{p}}\right), && (4.31a)\\ R_{\mathsf{c}} &\le \log\left(1+\mathsf{S}_{\mathsf{c}}\right), && (4.31b)\\ R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log^{+}\left(\frac{1+\mathsf{S}_{\mathsf{c}}}{1+\mathsf{I}_{\mathsf{c}}}\right)+\log\left(1+\mathsf{S}_{\mathsf{p}}+\mathsf{I}_{\mathsf{c}}\right), && (4.31c)\end{aligned}$$

which is achievable to within 1 bit by a non-cooperative scheme [12]. Therefore, for this set of parameters we have that the outer bound in (4.30) is achievable to within 2 bits.

4.5.2 Case $\mathsf{C} > \mathsf{S}_{\mathsf{p}},\ \mathsf{S}_{\mathsf{c}} \le \mathsf{I}_{\mathsf{c}}$ (i.e., strong interference at PRx): when unilateral cooperation is useful

In this case, we further outer bound the region in (4.29) as

$$\begin{aligned}\mathcal{O}^{4.5.2}: \quad R_{\mathsf{p}} &\le \log\left(1+\mathsf{C}\right)+\log(2), && (4.32a)\\ R_{\mathsf{c}} &\le \log\left(1+\mathsf{S}_{\mathsf{c}}\right), && (4.32b)\\ R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\mathsf{S}_{\mathsf{p}}+\mathsf{I}_{\mathsf{c}}\right)+2\log(2). && (4.32c)\end{aligned}$$

In this regime, the PTx takes advantage of the strong cooperation link and sends its message with the help of the CTx. Moreover, since the PTx does not create interference at the CRx ($\mathsf{I}_{\mathsf{p}} = 0$), it sends a (cooperative) private message only. On the other hand, since the interference at the PRx is strong, the CTx sends a (non-cooperative) common message only. This is exactly the strategy described in Appendix 4.E, by setting $V_1 = U_1 = T_1 = T_2 = \emptyset$ and the resulting achievable region is given by (4.53). In (4.53), we further set $|a_1| = |b_1| = |d_1| = 0$ (private message only for the PTx carried by Z_1), $|c_2| = 0$ (common non-cooperative message only for CTx carried by U_2) and $\mathsf{I}_{\mathsf{p}} = 0$. With the possible suboptimal choice $|a_2|^2 = \frac{1}{1+\mathsf{S}_{\mathsf{c}}}$, we have $k_1 = 0$, $k_2 \le \log\left(1+\frac{\mathsf{I}_{\mathsf{c}}|a_2|^2}{1+\mathsf{S}_{\mathsf{p}}}\right)$ in (4.53) and the achievable region in (4.53) is contained into

$$\begin{aligned}\mathcal{I}^{4.5.2}: \quad R_{\mathsf{p}} &\le \log\left(1+\mathsf{C}\right), && (4.33a)\\ R_{\mathsf{c}} &\le \log\left(1+\mathsf{S}_{\mathsf{c}}\right)-\log(2), && (4.33b)\\ R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\mathsf{S}_{\mathsf{p}}+\mathsf{I}_{\mathsf{c}}\right), && (4.33c)\\ R_{\mathsf{p}}+R_{\mathsf{c}} &\le \log\left(1+\mathsf{S}_{\mathsf{p}}+\frac{\mathsf{I}_{\mathsf{c}}}{1+\mathsf{S}_{\mathsf{c}}}\right)+\log\left(1+\mathsf{S}_{\mathsf{c}}\right)-\log(2). && (4.33d)\end{aligned}$$

It is not difficult to see that the outer bound in (4.32) and the inner bound in (4.33) are at most 1.5 bits away.

4.5.3 Case $\mathsf{C} > \mathsf{S_p}$, $\mathsf{S_c} > \mathsf{I_c}$ (i.e., weak interference at PRx): when unilateral cooperation is useful

For this regime, an outer bound for the Z-channel is given by the capacity of the non-causal CIC in weak interference [61, Theorem 4.1], [60, Lemma 3.6] together with the cut-set bound in (4.14a), i.e.,

$$\mathcal{O}^{4.5.3}: \quad R_{\mathsf{p}} \leq \log\left(1+\mathsf{C}\right) + \log(2), \tag{4.34a}$$

$$R_{\mathsf{p}} \leq \log\left(1 + \frac{\mathsf{S_p} + |\gamma_{\mathsf{c}}|^2 \mathsf{I_c} + 2|\gamma_{\mathsf{c}}|\sqrt{\mathsf{S_p I_c}}}{1+(1-|\gamma_{\mathsf{c}}|^2)\mathsf{I_c}}\right), \tag{4.34b}$$

$$R_{\mathsf{c}} \leq \log\left(1+(1-|\gamma_{\mathsf{c}}|^2)\mathsf{S_c}\right), \tag{4.34c}$$

union over all $|\gamma_{\mathsf{c}}| \leq 1$. Since $\mathsf{C} > \mathsf{S_p}$, the PTx takes advantage of the strong cooperation link and sends its message with the help of the CTx. Moreover, since the PTx does not create interference at the CRx ($\mathsf{I_p} = 0$), it sends a (cooperative) private message only. The outer bound in (4.34b) suggests that the PRx should treat as noise the message of the CTx, while the bound in (4.34c) tells us that the CRx should decode its own message without experiencing interference. In order to model this last observation, we use a DPC-based scheme. In this strategy the CTx precodes its message against the 'known interference' so that the CRx decodes its own message as if the interference was not present [44].

This is exactly the strategy described in Appendix 4.E, by setting $V_1 = U_1 = T_1 = T_2 = \emptyset$ and the resulting achievable region is given by (4.53). In (4.53), we further set $|a_1| = |b_1| = |d_1| = 0$ (private message only for the PTx carried by Z_1), $|b_2| = 0$ (private non-cooperative message only for CTx carried by T_2) and $\mathsf{I_p} = 0$. We further let $a_2 = \gamma_c$ and $|c_2| = \sqrt{1-|\gamma_c|^2}$, with $|\gamma_c| \in [0,1]$ in (4.53); with these choices we have $k_1 = 0$, $k_2 = \log\left(\frac{1+\mathsf{I_c}+\mathsf{S_p}}{1+\mathsf{S_p}+\mathsf{I_c}(1-|\gamma_c|^2)}\right)$ in (4.53) and the achievable region in (4.53) is contained into

$$\mathcal{I}^{4.5.3}: \quad R_{\mathsf{p}} \leq \log\left(1+\mathsf{C}\right), \tag{4.35a}$$

$$R_{\mathsf{p}} \leq \log\left(1 + \frac{\mathsf{S_p} + |\gamma_{\mathsf{c}}|^2 \mathsf{I_c}}{1+(1-|\gamma_{\mathsf{c}}|^2)\mathsf{I_c}}\right), \tag{4.35b}$$

$$R_{\mathsf{c}} \leq \log\left(1+(1-|\gamma_{\mathsf{c}}|^2)\mathsf{S_c}\right), \tag{4.35c}$$

for all $|\gamma_{\mathsf{c}}| \leq 1$. By simple computations, the achievable region in (4.35) can be shown to be at most 1 bit away from the upper bound in (4.34).

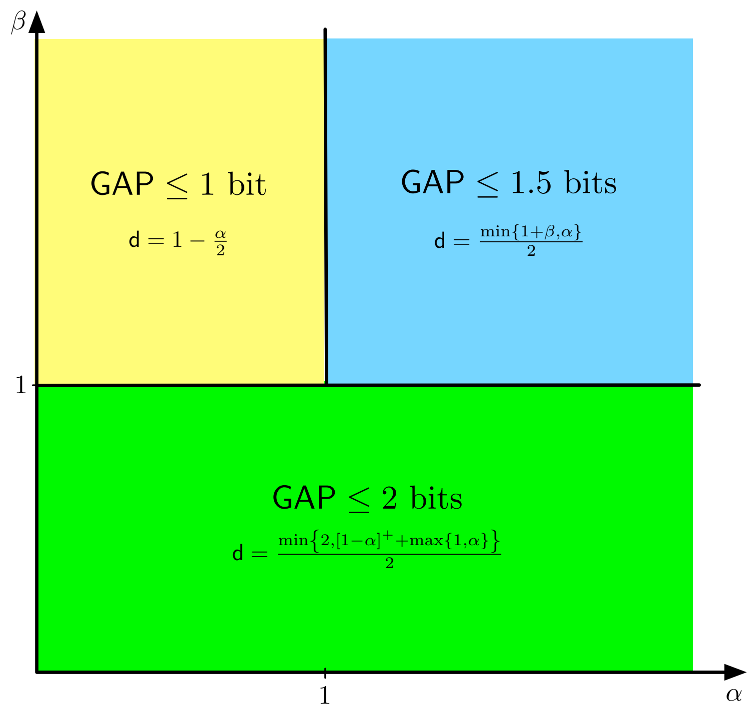


Figure 4.5: Optimal gDoF and constant gap for the Z-channel in the different regimes of (α, β).

4.5.4 Comparisons

We conclude this section by comparing the performance of unilateral cooperation on the Z-channel with other forms of cooperation. Moreover, we also consider whether the absence of an interfering link is beneficial in the Gaussian CCIC. We shall use as performance metric the gDoF, or high SNR throughput. In order to reduce the number of parameters, we restrict our attention to the case where the direct links have the same strength. For future reference, the gDoF of the non-cooperative Z-channel is given by [103]

$$\mathsf{d}^{\mathrm{Z}}(\alpha, 0) = \min\{\max\{1 - \alpha/2, \alpha/2\},\ 1\},$$

and that of the non-causal cognitive Z-channel, which can be evaluated from [64], is

$$\mathsf{d}^{\mathrm{Z}}(\alpha, \infty) = \max\{1 - \alpha/2, \alpha/2\}.$$

Figure 4.5 shows the gDoF and the gap for the Z-channel for different regions in the (α, β) plane. The whole set of parameters has been partitioned into multiple sub-regions depending upon different level of cooperation (β) and interference (α) strengths.

When comparing unilateral cooperation with other channel models in terms of gDoF we observe:

- For the non-cooperative IC, it is well known that removing an interference link cannot degrade the performance and the sum-capacity is known exactly for all channel parameters [103]. The same cannot be said in full generality for the cooperative channel because "useful cooperative information" can flow through the interference link. In particular, for the Z-channel unilateral cooperation improves the gDoF with respect to the non-cooperative case, i.e., $\mathsf{d}^{\mathrm{Z}}(\alpha,\beta) > \mathsf{d}^{\mathrm{Z}}(\alpha,0)$, only in the regime $\alpha \geq 2$ and $\beta \geq 1$, i.e., in very strong interference and strong cooperation.

- The Z-channel achieves the same gDoF of the ideal non-causal cognitive channel, i.e., $\mathsf{d}^{\mathrm{Z}}(\alpha,\beta) = \mathsf{d}^{\mathrm{Z}}(\alpha,\infty)$, everywhere except in $\alpha > \max\{2,\beta+1\}$.

- By comparing Figure 4.4 and Figure 4.5 we observe that the gDoF of the Z-channel is always greater than or equal to that of the interference-symmetric Gaussian CCIC. This is due to the fact that the PTx does not cooperate in sending the cognitive signal. Therefore, by removing the link between PTx and CRx we rid CRx of only an interfering signal and this leads to an improvement in terms of gDoF.

  The regimes where the Z-channel strictly outperforms the interference-symmetric Gaussian CCIC, i.e., $\mathsf{d}^{\mathrm{Z}}(\alpha,\beta) > \mathsf{d}(\alpha,\beta)$, are when $0 \leq \alpha \leq \frac{2}{3}$ and $\beta \leq \min\{\alpha, 1-\alpha\}$, i.e., weak interference and fairly weak cooperation. This regime can be thought of as the one where interference is the most harmful for the interference-symmetric Gaussian CCIC.

- From our constant gap result, it follows that the capacity region of the Z-channel, differently from the one of the symmetric Gaussian CCIC, does not have bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$. In other words, when the link between the PTx and the CRx is absent, unilateral cooperation allows for a full utilization of the channel resources, i.e., there are no 'resource holes' in the system [56].

4.6 The capacity region to within a constant gap for the Gaussian S-channel

In this section, we prove Theorem 11, i.e., we characterize the capacity to within a constant gap for the S Gaussian CCIC defined by $\mathsf{I_c} = 0$. In particular, we distinguish two cases, depending on whether the following upper bound

$$R_{\mathsf{p}} \leq \log\left(1+\mathsf{S_p}\right), \tag{4.36a}$$

$$R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}\right), \tag{4.36b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}})^2\right) + \log\left(1+\frac{\mathsf{C}+\mathsf{S_p}}{1+\mathsf{I_p}}\right), \tag{4.36c}$$

from (4.14) with $\mathsf{I_c} = 0$, can be achieved with a non-cooperative scheme or not. Note that the bounds on R_{p} and R_{c} in (4.36) are the capacity region of the corresponding non-causal CIC; we hence interpret the sum-rate bound in (4.36c) as the 'cost' for causally learning the primary message at the CTx through a noisy channel.

4.6.1 Case $\mathsf{C} \leq \max\{\mathsf{I_p}, \mathsf{S_p}\}$: when unilateral cooperation might not be useful

For the case $\mathsf{C} \leq \max\{\mathsf{I_p}, \mathsf{S_p}\}$ we can further upper bound (4.36) as

$$\mathcal{O}^{4.6.1}: \quad R_{\mathsf{p}} \leq \log\left(1+\mathsf{S_p}\right), \tag{4.37a}$$

$$R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}\right), \tag{4.37b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}+\mathsf{I_p}\right) + \log\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right) + 2\log(2). \tag{4.37c}$$

The region in (4.37) is at most 1.5 bits (per user) away from

$$\mathcal{I}^{4.6.1}: \quad R_{\mathsf{p}} \leq \log\left(1+\mathsf{S_p}\right), \tag{4.38a}$$

$$R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}\right), \tag{4.38b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}+\mathsf{I_p}\right) + \log^{+}\left(\frac{1+\mathsf{S_p}}{1+\mathsf{I_p}}\right), \tag{4.38c}$$

which is achievable to within 1 bit by a non-cooperative scheme [12]. Therefore, we conclude that for $\mathsf{C} \leq \max\{\mathsf{I_p}, \mathsf{S_p}\}$ a non-cooperative scheme is optimal to within 2.5 bits.

4.6.2 Case $\mathsf{C} > \max\{\mathsf{I_p}, \mathsf{S_p}\}$: when unilateral cooperation is useful

When $\mathsf{C} > \max\{\mathsf{I_p}, \mathsf{S_p}\}$, a sufficient condition for the sum-rate upper bound in (4.36c) to be redundant is that

$$1 + \mathsf{S_p} \leq 1 + \frac{\mathsf{C} + \mathsf{S_p}}{1 + \mathsf{I_p}} \Longleftrightarrow \mathsf{C} \geq \mathsf{I_p}\mathsf{S_p}. \tag{4.39}$$

For the set of parameters in (4.39), we use the achievable region in Appendix 4.E, by further setting $V_1 = U_1 = T_1 = T_2 = \emptyset$ and the resulting achievable region is given by (4.53). In (4.53), we further set $|b_1| = |d_1| = 0$ (private message only for the PTx carried by the pair (S_1, Z_1)), $|a_2| = |b_2| = 0$ (private non-cooperative message only for CTx carried by T_2) and $\mathsf{I_c} = 0$. We further let $|c_1|^2 = \frac{\mathsf{S_p}}{\mathsf{C}}$ in (4.53); with these choices we have $k_1 = 0$, $k_2 = \log\left(\frac{1+\mathsf{S_p}}{1+\mathsf{S_p}|c_1|^2}\right)$ in (4.53) and the achievable region in (4.53) is contained into

$$\mathcal{I}^{4.6.2}: \quad R_{\mathsf{p}} \leq \log\left(1 + \mathsf{S_p}\right), \tag{4.40a}$$

$$R_{\mathsf{c}} \leq \log\left(1 + \mathsf{S_c}\right) - \log(2). \tag{4.40b}$$

By comparing the rate bounds in (4.40) with those in (4.36), it is easy to see that when (4.39) holds the gap is at most of 1 bit.

This shows that, when the condition in (4.39) holds, not only the upper bound is achievable to within 1 bit but we can also achieve to within 1 bit the ultimate capacity of the corresponding non-causal cognitive channel. This result agrees with the intuition that, as the strength of the cooperation link increases, the performance of the causal cognitive channel should approach that of the corresponding non-causal model. The condition in (4.39) can thus be interpreted as a sufficient condition on the strength of the cooperation link to achieve the capacity region of the corresponding non-causal model to within a constant gap.

In the regime $\mathsf{C} < \mathsf{S_p}\mathsf{I_p}$ we use the DPC-based achievable scheme in Appendix 4.E with $U_1 = \emptyset$ (see Appendix 4.E.2). In this scheme the CTx sends a private message only since X_{c} is not received at the PRx; since the cooperation link is quite strong, the PTx sends a private and a common message (carried by the pairs (S_1, Z_1) and (Q, V_1), respectively), both with the help of the CTx. In other words, with reference to the scheme in Appendix 4.E, we set $U_1 = T_1 = \emptyset$, i.e., no non-cooperative messages for PTx and the resulting achievable rate region is given by (4.53). The PTx's common message is forwarded by the CTx to facilitate decoding at both receivers. The

PTx's private message is decoded at the CTx and its effect is 'pre-canceled' at the CRx thanks to DPC. In (4.53), we further set $|d_1| = 0$ (no private non-cooperative message for PTx carried by T_1), $|a_2| = |b_2| = 0$ (private non-cooperative message only for CTx carried by T_2) and $\mathsf{I_c} = 0$. We further let $|a_1|^2 = \frac{\mathsf{C}}{2(1+\mathsf{I_p})(1+\mathsf{S_p})}$ and $|c_1|^2 = \frac{1}{2(1+\mathsf{I_p})}$ in (4.53); with these choices we have $k_1 = \log\left(\frac{1+\mathsf{S_c}+\mathsf{I_p}}{1+\mathsf{S_c}+\mathsf{I_p}(1-|b_1|^2)}\right)$, $k_2 = \log\left(\frac{1+\mathsf{S_p}\left(1-|b_1|^2\right)}{1+\mathsf{S_p}|c_1|^2}\right)$ in (4.53) and the achievable region in (4.53) is contained into

$$\begin{aligned}
\mathcal{I}^{4.6.2}: \quad & R_\mathsf{p} \le \log(1+\mathsf{C}) - \log(2), && \text{(4.41a)}\\
& R_\mathsf{p} \le \log(1+\mathsf{S_p}), && \text{(4.41b)}\\
& R_\mathsf{c} \le \log\left(1+\mathsf{S_c}\right) - \log(2), && \text{(4.41c)}
\end{aligned}$$

$$\begin{aligned}
R_\mathsf{p} + R_\mathsf{c} \le & \log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}}\right) + \log\left(1+\mathsf{S_c}\right)\\
& + \log\left(\frac{1+\mathsf{S_c}+\mathsf{I_p}}{1+\mathsf{S_c}+\frac{1+\mathsf{C}}{1+\mathsf{S_p}}}\right) - 2\log(2), \qquad \text{(4.41d)}\\
R_\mathsf{p} + R_\mathsf{c} \le & \log\left(1+\mathsf{S_p}\frac{1+\mathsf{C}}{\left(1+\mathsf{I_p}\right)\left(1+\mathsf{S_p}\right)}\right) + \log\left(1+\mathsf{S_c}\right)\\
& + \log\left(\frac{1+\mathsf{S_c}+\mathsf{I_p}}{1+\mathsf{S_c}+\frac{1+\mathsf{C}}{1+\mathsf{S_p}}}\right) - 2\log(2). \qquad \text{(4.41e)}
\end{aligned}$$

By comparing the rate bounds in (4.41) with those in (4.36), it is easy to see that the gap is at most of 3 bits. Notice that, in order to find the gap, we compared (4.41d) and (4.41e) with (4.36a) + (4.36b) if $\mathsf{S_c} \le \frac{1+\mathsf{C}}{1+\mathsf{S_p}}$ and with (4.36c) otherwise.

4.6.3 Comparisons

We conclude this section by comparing the performance of unilateral cooperation on the S-channel with other forms of cooperation and by considering whether the absence of an interfering link is beneficial in the Gaussian CCIC. In order to reduce the number of parameters, we restrict our attention to the case where the direct links have the same strength. For future reference, the gDoF of the non-cooperative S-channel is given by [103]

$$\mathsf{d}^{\mathrm{S}}\left(\alpha, 0\right) = \min\{\max\{1-\alpha/2, \alpha/2\},\ 1\},$$

and that of the non-causal cognitive S-channel is given by [64]

$$\mathsf{d}^{\mathrm{S}}\left(\alpha, \infty\right) = 1.$$

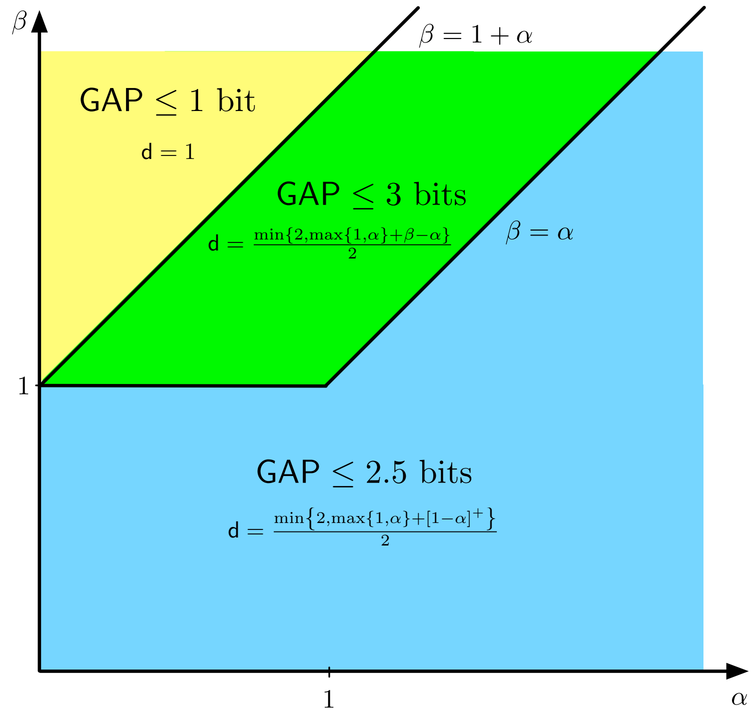


Figure 4.6: Optimal gDoF and constant gap for the S-channel in the different regimes of (α, β).

Figure 4.6 shows the gDoF and the gap for the S-channel in the (α, β) plane. The whole set of parameters has been partitioned into multiple subregions depending upon different levels of cooperation (β) and interference (α) strengths. We observe:

- Unilateral cooperation achieves the same gDoF of the non-cooperative IC, i.e., $\mathsf{d}^{\mathrm{S}}(\alpha, \beta) = \mathsf{d}^{\mathrm{S}}(\alpha, 0)$, when $\alpha \geq 2$ or $\beta \leq \max\{1, \alpha\}$. In other words, unilateral cooperation is worth implementing in practice when the interference is not very strong and the cooperation link is the strongest among all links.

- The S-channel achieves the same gDoF of the non-causal CIC, i.e., $\mathsf{d}^{\mathrm{S}}(\alpha, \beta) = \mathsf{d}^{\mathrm{S}}(\alpha, \infty)$, everywhere except in $\alpha \leq 2$ and $\beta \leq \min\{2, \alpha+1\}$.

- The S-channel outperforms the interference-symmetric Gaussian CCIC, i.e., $\mathsf{d}^{\mathrm{S}}(\alpha, \beta) > \mathsf{d}(\alpha, \beta)$, when either $0 \leq \alpha \leq \frac{2}{3}$ and $\beta \leq \min\{\alpha, 1 - \alpha\}$ or when $\alpha \leq 2$ and $\beta \geq \max\{1, \alpha\}$. On the other hand, the interference-symmetric Gaussian CCIC outperforms the S-channel, i.e., $\mathsf{d}^{\mathrm{S}}(\alpha,\beta) < \mathsf{d}(\alpha,\beta)$, in very strong interference and strong cooperation, i.e., $\alpha \geq 2$ and $\beta \geq 1$. This is due to the fact that the information for the PRx can no longer be routed through the CTx since $\sqrt{\mathsf{I_c}}e^{\mathrm{j}\theta_c} = 0$.

- As for the Z-channel, also the capacity region of the S-channel, differently from that of the symmetric Gaussian CCIC, does not have bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$. In other words, when the link between the CTx and the PRx is absent, unilateral cooperation allows for a full utilization of the channel resources, i.e., there are no 'resource holes' in the system [56].

4.7 Extension to the general Gaussian CCIC

In this section we seek to extend our gap results to the general Gaussian CCIC, which is more complex to analyze due to the fact that one has to deal with 5 different channel parameters. Following the naming convention of the non-cooperative IC, we say that the general Gaussian CCIC has strong interference if $\{\mathsf{S_p} \leq \mathsf{I_p},\ \mathsf{S_c} \leq \mathsf{I_c}\}$, weak interference if $\{\mathsf{S_p} > \mathsf{I_p},\ \mathsf{S_c} > \mathsf{I_c}\}$, and mixed interference otherwise. Moreover, we say that the general Gaussian CCIC has strong cooperation if $\mathsf{C} > \mathsf{S_p}$ and weak cooperation otherwise. As we shall see in what follows, this section provides a capacity characterization to within a constant gap for the general Gaussian CCIC when, roughly speaking, the two receivers do not experience weak interference simultaneously. In particular,

- **Case A**: When $\mathsf{C} \leq \mathsf{S_p}$, i.e., weak cooperation regime, we can further upper bound (4.14) as

$$\begin{aligned}
\mathcal{O}^{\text{Case A}}: \ R_{\mathsf{p}} &\leq \log(1+\mathsf{S_p}) + \log(2), && \text{(4.42a)}\\
R_{\mathsf{c}} &\leq \log(1+\mathsf{S_c}), && \text{(4.42b)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right) + \log(1+\mathsf{S_c}+\mathsf{I_p}) + 2\log(2), && \text{(4.42c)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right) + \log(1+\mathsf{S_p}+\mathsf{I_c}) + \log(2). && \text{(4.42d)}
\end{aligned}$$

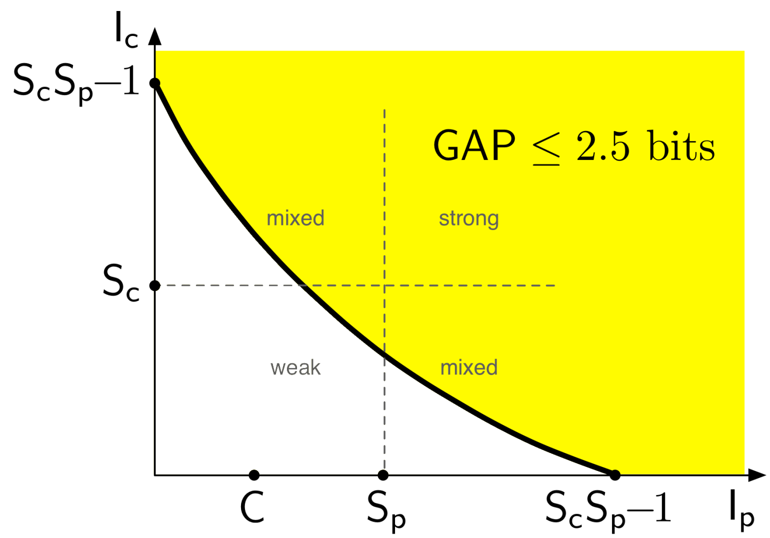


Figure 4.7: Regime identified as **Case A**, with $\mathsf{GAP} \leq 2$ bits.

The bounds in (4.42) are to within 1.5 bits of

$$\mathcal{I}^{\text{Case A}}: \quad R_{\mathsf{p}} \leq \log(1+\mathsf{S_p}), \tag{4.43a}$$
$$R_{\mathsf{c}} \leq \log(1+\mathsf{S_c}), \tag{4.43b}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log^+\left(\frac{1+\mathsf{S_p}}{1+\mathsf{I_p}}\right) + \log(1+\mathsf{S_c}+\mathsf{I_p}), \tag{4.43c}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log^+\left(\frac{1+\mathsf{S_c}}{1+\mathsf{I_c}}\right) + \log(1+\mathsf{S_p}+\mathsf{I_c}), \tag{4.43d}$$

which is achievable to within 1 bit for the non-cooperative IC [12] when $\mathsf{S_cS_p} \leq (1+\mathsf{I_p})(1+\mathsf{I_c})$. Thus, for this regime, depicted in yellow in Figure 4.7, we have $\mathsf{GAP} \leq 2.5$ bits.

- **Case B**: When $\mathsf{S_p} < \mathsf{C} \leq \mathsf{I_p}$, we further upper bound (4.14) as

$$\mathcal{O}^{\text{Case B}}: \tag{4.44a}$$
$$R_{\mathsf{p}} \leq \log(1+\mathsf{C}) + \log(2), \tag{4.44b}$$
$$R_{\mathsf{p}} \leq \log\left(1+\mathsf{S_p}+\mathsf{I_c}\right) + \log(2), \tag{4.44c}$$
$$R_{\mathsf{c}} \leq \log(1+\mathsf{S_c}), \tag{4.44d}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\mathsf{S_c}+\mathsf{I_p}\right) + 2\log(2), \tag{4.44e}$$
$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right) + \log\left(1+\mathsf{S_p}+\mathsf{I_c}\right) + \log(2). \tag{4.44f}$$

In this regime, unilateral cooperation helps increasing the rate of the primary user. In the symmetric case, this regime corresponds to part of the blue region in Figure 4.4 with $1 < \beta \leq \alpha$; we therefore consider the generalization of the achievable scheme we used for the symmetric case in this regime to the case of general channel gains. Here the PTx takes advantage of the strong cooperation link and sends its message with the help of the CTx. The sum-rate upper bound in (4.44e) suggests that the CRx should decode the PTx's message in addition to its intended message, that is, the PTx should use a (cooperative) common message only. The sum-rate upper bound in (4.44f), suggests that the PRx should decode the CTx's message only when $\mathsf{I_c} > \mathsf{S_c}$, that is, the CTx should use both a (non-cooperative) common and a (non-cooperative) private message. This is exactly the strategy described in Appendix 4.E.1 with $U_1 = T_1 = \emptyset$ (i.e., no non-cooperative messages for PTx) and the resulting achievable region is given in (4.51). In (4.51), we further set $|b_1| = 1$ (common cooperative message only for PTx carried by V_1) and $|b_2|^2 = \frac{1}{1+\mathsf{I_c}}$ so that the private message of the CTx (carried by T_2) is received below the noise level at the PRx in the spirit of [12]. With these choices the achievable rate region in (4.51) is contained into

$$
\begin{aligned}
\mathcal{I}^{\text{Case B}}: \quad & R_\mathsf{p} \leq \log\left(1+\mathsf{C}\right), && \text{(4.45a)}\\
& R_\mathsf{p} \leq \log\left(1+\mathsf{S_p}+\mathsf{I_c}\right) - \log(2), && \text{(4.45b)}\\
& R_\mathsf{c} \leq \log\left(1+\mathsf{S_c}\right), && \text{(4.45c)}\\
R_\mathsf{p}+R_\mathsf{c} & \leq \log(1+\mathsf{S_c}+\mathsf{I_p}), && \text{(4.45d)}\\
R_\mathsf{p}+R_\mathsf{c} & \leq \log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right)+\log(1+\mathsf{S_p}+\mathsf{I_c})-\log(2). && \text{(4.45e)}
\end{aligned}
$$

By comparing the upper bound in (4.44) with the achievable region in (4.45) we conclude that the capacity region is known to within 2 bits for a general Gaussian CCIC where the channel gains satisfy $\mathsf{S_p} < \mathsf{C} \leq \mathsf{I_p}$ (see blue regime Figure 4.8). Notice that we did not impose any condition on the strength of $\mathsf{I_c}$ compared to $\mathsf{S_c}$, i.e., in other words this gap result holds regardless of whether the interference at PRx is strong ($\mathsf{I_c} \geq \mathsf{S_c}$) or weak ($\mathsf{I_c} < \mathsf{S_c}$).

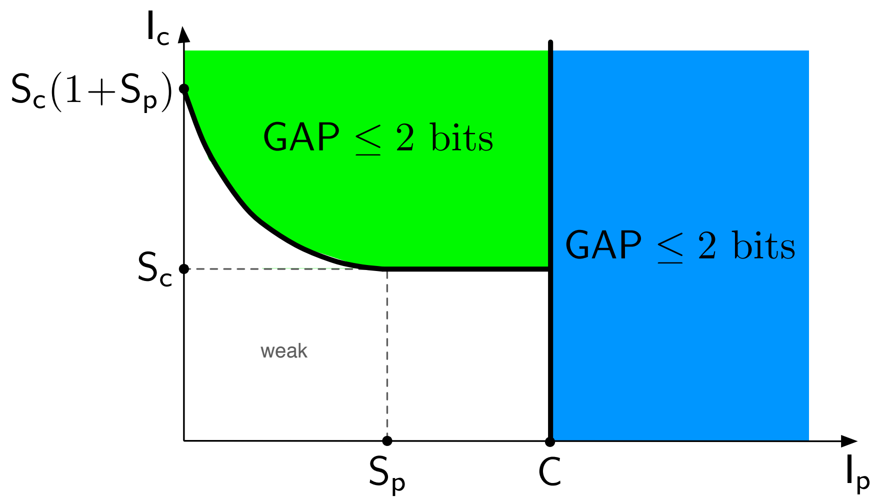


Figure 4.8: Blue and green regimes identified as **Case B** and **Case C**, respectively with $\mathsf{GAP} \leq 2$ bits.

- **Case C**: When $\max\{\mathsf{S_p}, \mathsf{I_p}\} < \mathsf{C}$ and $\mathsf{S_c} \leq \mathsf{I_c}$, we further upper bound (4.14) as

$$\mathcal{O}^{\text{Case C}}: \ R_\mathsf{p} \leq \log(1+\mathsf{C}) + \log(2), \tag{4.46a}$$

$$R_\mathsf{c} \leq \log(1+\mathsf{S_c}), \tag{4.46b}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1 + \frac{\mathsf{C}}{1+\mathsf{I_p}}\right) + \log\left(1+\mathsf{S_c}+\mathsf{I_p}\right) + 2\log(2), \tag{4.46c}$$

$$R_\mathsf{p} + R_\mathsf{c} \leq \log\left(1+\mathsf{S_p}+\mathsf{I_c}\right) + 2\log(2). \tag{4.46d}$$

In this regime, unilateral cooperation helps increasing both the rate of the primary user and the sum-capacity. In the symmetric case, this regime corresponds to part of the blue region in Figure 4.4 with $1 < \alpha < \beta$. Here the PTx takes advantage of the strong cooperation link and sends its message with the help of the CTx. The sum-rate upper bound in (4.46d) suggests that PRx should decode the CTx's message in addition to its intended message, that is, CTx should use a (non-cooperative) common message only; this is so because the condition $\mathsf{S_c} \leq \mathsf{I_c}$ corresponds to strong interference at the PRx. The sum-rate upper bound in (4.46c), suggests that PTx should use both a (cooperative) common and a (cooperative) private message; this is so because here we do not specify which one among $\mathsf{S_p}$ and $\mathsf{I_p}$ is the largest, and therefore the interference at CRx could be either strong or weak. This is exactly the strategy described in Appendix 4.E.2 with $T_1 = T_2 = \emptyset$, i.e., no private non-cooperative messages. The achievable region is given in (4.53). In (4.53), we further set $|a_1| = |d_1| = |c_2| = 0$ (i.e., no private non-cooperative messages) and $|c_1|^2 = \frac{1}{1+\mathsf{I_p}}$ and $|a_2|^2 = \frac{1}{1+\mathsf{S_c}}$ inspired by [12]. Moreover, we do not pre-encode U_2 against the private (cooperative) message of PTx, i.e., $\lambda_U = 0$. With these choices we have $k_1 = \log\left(\frac{1+\mathsf{S_c}+\mathsf{I_p}}{1+\mathsf{S_c}+\mathsf{I_p}(1-|b_1|^2)}\right)$, $k_2 = \log\left(1+\frac{\mathsf{I_c}|a_2|^2}{1+\mathsf{S_p}|c_1|^2}\right)$ and the achievable rate region in (4.53) is contained into

$$\mathcal{I}^{\text{Case C}}: \ R_{\mathsf{p}} \le \log\left(1+\mathsf{C}\right), \tag{4.47a}$$

$$R_{\mathsf{c}} \le \log\left(1+\mathsf{S_c}\right) - 2\log(2), \tag{4.47b}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}}\right) + \log\left(1+\mathsf{S_c}+\mathsf{I_p}\right) - 2\log(2), \tag{4.47c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1+\mathsf{S_p}+\mathsf{I_c}\right), \tag{4.47d}$$

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} \le \log&\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}+\frac{\mathsf{I_c}}{1+\mathsf{S_c}}\right) \\ &+ \log\left(1+\mathsf{S_c}+\mathsf{I_p}\right) - 2\log(2), \end{aligned} \tag{4.47e}$$

if $\mathsf{S_c}\frac{1+\mathsf{I_p}+\mathsf{S_p}}{1+2\mathsf{I_p}} \le \mathsf{I_c}$. We imposed this condition in order to get rid of the bound on $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ in (4.53h). Actually, from (4.52) we have that

$$\begin{aligned} &\text{eq}(4.52\text{c}) + \text{eq}(4.52\text{g}) \le \text{eq}(4.52\text{h}) \iff \\ &\text{eq}(4.49\text{r}) + \text{eq}(4.49\text{g}) \le \text{eq}(4.49\text{f}) + \text{eq}(4.49\text{t}) \iff \\ &I\left(Y_{\mathsf{c}}; U_2|Q, V_1\right) \le I\left(Y_{\mathsf{p}}; U_2|Q, V_1\right) \iff \\ &\frac{\mathsf{S_c}}{1+\mathsf{S_c}|a_2|^2+\mathsf{I_p}|c_1|^2} \le \frac{\mathsf{I_c}}{1+\mathsf{I_c}|a_2|^2+\mathsf{S_p}|c_1|^2} \iff \\ &\mathsf{S_c} + \mathsf{S_c}\mathsf{S_p}|c_1|^2 \le \mathsf{I_c} + \mathsf{I_c}\mathsf{I_p}|c_1|^2 \iff \mathsf{S_c}\frac{1+\mathsf{I_p}+\mathsf{S_p}}{1+2\mathsf{I_p}} \le \mathsf{I_c}. \end{aligned}$$

By comparing the upper bound in (4.46) with the inner bound in (4.47) it easy to see that they are at most 2 bits away from one another. Thus, when $\max\{\mathsf{S_p}, \mathsf{I_p}\} < \mathsf{C}$, $\mathsf{S_c} \le \mathsf{I_c}$ and if $\mathsf{S_c}\frac{1+\mathsf{I_p}+\mathsf{S_p}}{1+2\mathsf{I_p}} \le \mathsf{I_c}$, then $\mathsf{GAP} \le 2$ bits (see green regime Figure 4.8).

Remark 16. The analysis above provides the characterization of the capacity to within a constant gap for the general Gaussian CCIC when, roughly speaking, the two receivers do not suffer weak interference simultaneously.

Although not a trivial task (since there are 5 different channel gains), characterizing the capacity (to within a constant gap) in the regimes left open in this chapter is an interesting future research direction.

4.8 Conclusions and future directions

In this chapter we considered the two-user CCIC where the CTx is constrained to operate in FD mode. We first derived two novel outer bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ on the capacity region of the ISD channel with independent noises at the two destinations. We then designed a transmission strategy based on binning and superposition encoding, PDF relaying and jointly decoding and we derived its achievable rate region. We evaluated the outer and lower bounds on the capacity for the Gaussian noise case and we proved that these bounds are a constant number of bits (universally over all channel gains) apart from one another for the symmetric case (i.e., the two direct links and the two interfering links are of the same strength) and for the case where one interfering link is absent (i.e., Z-channel and S-channel). In particular, we showed that, for the symmetric case, the two novel outer bounds of the type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ and $R_{\mathsf{p}}+2R_{\mathsf{c}}$ are active in weak interference when the cooperation link is weaker than the direct link. We also considered the general Gaussian CCIC and we proved a constant gap result when, roughly speaking, the two receivers do not experience weak interference simultaneously. Finally, we identified the set of parameters where causal cooperation achieves the same gDoF of the non-cooperative IC and of the ideal non-causal CIC.

With respect to the results presented in this chapter, interesting future research directions may include: (i) characterization of the capacity to within a constant gap for the general Gaussian CCIC in the left open regimes and (ii) derivation of tighter outer bounds and design of novel transmission strategies in order to reduce the gap.

Appendix

4.A Proof of the Markov chains in (4.8a) and (4.8b)

We prove the two Markov chains in (4.8a)-(4.8b) by using the FDG [102]. Figure 4.9 proves the Markov chain in (4.8a), while Figure 4.10 the one in (4.8b). The two proofs, without loss of generality, consider the time instant $i = 3$. According to [102], we proceed through the following steps.

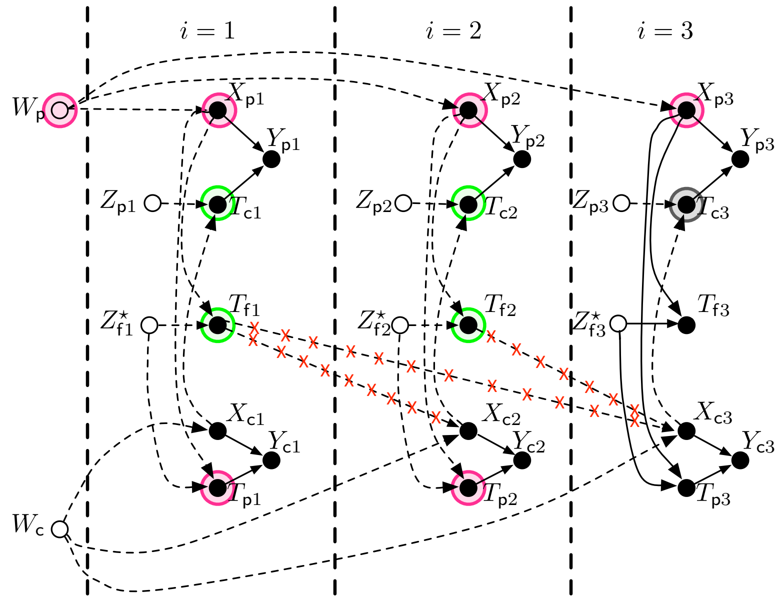


Figure 4.9: Proof of the Markov chain in (4.8a) using the FDG.

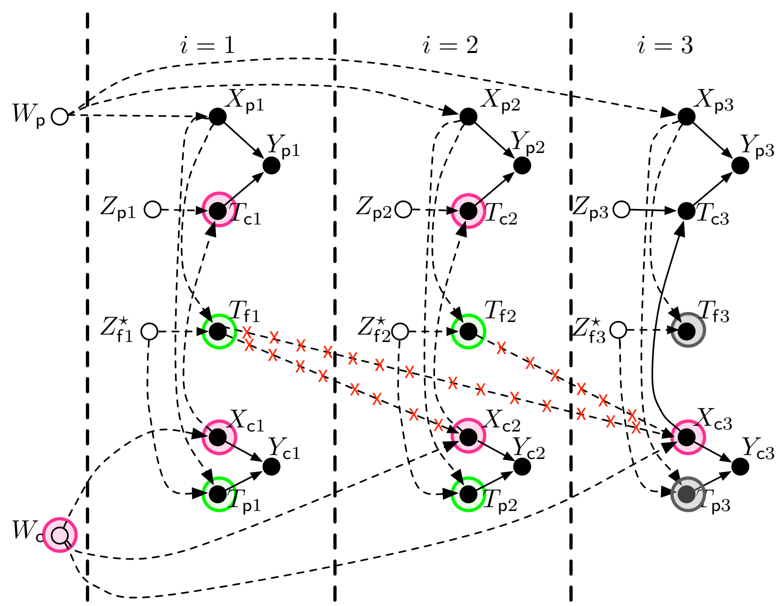


Figure 4.10: Proof of the Markov chain in (4.8b) using the FDG.

1. Draw the directed graph $\mathcal{G}_1$, which takes into consideration the dependence between the different random variables involved in the ISD CCIC considered. In particular, we define

$$Z^{\star}_{\mathsf{f}i} = \begin{bmatrix} Z_{\mathsf{f}i} \\ Z_{\mathsf{c}i} \end{bmatrix}$$

to consider the fact that the noises at the CTx and at the CRx can be arbitrally correlated and we have

$$X_{\mathsf{p}_i} = f(W_{\mathsf{p}}),\ Y_{\mathsf{p}_i} = f(X_{\mathsf{p}_i}, T_{\mathsf{c}i}),\ T_{\mathsf{c}i} = f(X_{\mathsf{c}i}, Z_{\mathsf{p}_i}),\ T_{\mathsf{f}i} = f(X_{\mathsf{p}_i}, Z^{\star}_{\mathsf{f}i}),$$
$$X_{\mathsf{c}i} = f(W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1}),\ Y_{\mathsf{c}i} = f(X_{\mathsf{c}i}, T_{\mathsf{p}_i}),\ T_{\mathsf{p}_i} = f(X_{\mathsf{p}_i}, Z^{\star}_{\mathsf{f}i}),$$

where with f we indicate that the left-hand side of the equality is a function of the random variables into the bracket.

2. In $\mathcal{G}_1$, highlight all the different nodes / random variables involved in the two Markov chains in (4.8a)-(4.8b) we aim to prove. In particular, the random variables circled in magenta, given those circled in green, should be proved to be independent of those circled in grey.

3. From the graph $\mathcal{G}_1$, consider the subgraph $\mathcal{G}_2$ which contains those edges and vertices encountered when moving backwards one or more edges starting from the colored (magenta, green and grey) random variables. The edges of the subgraph $\mathcal{G}_2$ are depicted with dashed black lines in Figure 4.9 and Figure 4.10 and the vertices in $\mathcal{G}_2$ are all those touched by a dashed black line.

4. From the graph $\mathcal{G}_2$, remove all the edges coming out from the random variables in green (those which are supposed to d-separate the random variables colored in magenta and grey). In Figure 4.9 and Figure 4.10, this step is highlighted with red crosses on the edges which are removed. We let $\mathcal{G}_3$ be the subgraph obtained from $\mathcal{G}_2$ by removing all the edges with red crosses.

5. From $\mathcal{G}_3$, remove all the arrows on the edges, and obtain the undirected subgraph $\mathcal{G}_4$. In $\mathcal{G}_4$ it is easy to see that, by starting from any grey node, it is not possible to reach any magenta node. This concludes the proof of the two Markov chains in (4.8a)-(4.8b).

4.B Proof of the sum-rate outer bound in (4.9f)

By using the two Markov chains in (4.8a)-(4.8b) we can now derive the sum-rate outer bound in (4.9f). This bound was originally derived in [47] for the case of independent noises; here we extend it to the case when only the noises at the different source-destination pairs are independent, i.e., $\mathbb{P}_{Y_{\mathsf{Fc}},Y_{\mathsf{c}}|X_{\mathsf{p}},X_{\mathsf{c}}}$ in (4.2) is not a product distribution. By using Fano's inequality and by providing the same genie side information as in [47], we have

$$\begin{aligned}
&N(R_{\mathsf{p}}+R_{\mathsf{c}}-2\epsilon_N)\\
\leq& I\left(W_{\mathsf{p}};{Y_{\mathsf{p}}}^N\right)+I\left(W_{\mathsf{c}};{Y_{\mathsf{c}}}^N\right)\\
\leq& I\left(W_{\mathsf{p}};{Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N\right)+I\left(W_{\mathsf{c}};{Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N\right)\\
=& H\left({Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N\right)-H\left({Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{c}}\right)\\
&+H\left({Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N\right)-H\left({Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{p}}\right).
\end{aligned}$$

We now analyze and bound the two pairs of terms. First pair:

$$\begin{aligned}
&H\left({Y_{\mathsf{p}}}^N,{T_{\mathsf{p}}}^N,{T_{\mathsf{f}}}^N\right)-H\left({Y_{\mathsf{c}}}^N,{T_{\mathsf{c}}}^N,{T_{\mathsf{f}}}^N|W_{\mathsf{c}}\right)\\
\overset{(a)}{=}&\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i},T_{\mathsf{p}_i},T_{\mathsf{f}i}|{Y_{\mathsf{p}}}^{i-1},{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{f}}}^{i-1}\right)\\
&-H\left(Y_{\mathsf{c}i},T_{\mathsf{c}i},T_{\mathsf{f}i}|{Y_{\mathsf{c}}}^{i-1},{T_{\mathsf{c}}}^{i-1},{T_{\mathsf{f}}}^{i-1},W_{\mathsf{c}},{X_{\mathsf{c}}}^{i}\right)\\
\overset{(b)}{=}&\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i},T_{\mathsf{p}_i},T_{\mathsf{f}i}|{Y_{\mathsf{p}}}^{i-1},{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{f}}}^{i-1}\right)\\
&-H\left(T_{\mathsf{p}_i},T_{\mathsf{c}i},T_{\mathsf{f}i}|{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{c}}}^{i-1},{T_{\mathsf{f}}}^{i-1},W_{\mathsf{c}},{X_{\mathsf{c}}}^{i}\right)\\
\overset{(c)}{\leq}&\sum_{i\in[1:N]}\underbrace{H\left(T_{\mathsf{p}_i},T_{\mathsf{f}i}|{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{f}}}^{i-1}\right)-H\left(T_{\mathsf{p}_i},T_{\mathsf{f}i}|{T_{\mathsf{p}}}^{i-1},{T_{\mathsf{f}}}^{i-1},W_{\mathsf{c}},{T_{\mathsf{c}}}^{i-1},{X_{\mathsf{c}}}^{i}\right)}_{=0\text{ because of (4.8b)}}\\
&+\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i}|T_{\mathsf{p}_i},T_{\mathsf{f}i}\right)-H\left(T_{\mathsf{c}i}|{T_{\mathsf{p}}}^{i},{T_{\mathsf{f}}}^{i},W_{\mathsf{c}},{T_{\mathsf{c}}}^{i-1},{X_{\mathsf{c}}}^{i},{X_{\mathsf{p}}}^{i}\right)\\
\overset{(d)}{=}&\sum_{i\in[1:N]}H\left(Y_{\mathsf{p}_i}|T_{\mathsf{p}_i},T_{\mathsf{f}i}\right)-H\left(Y_{\mathsf{p}_i}|T_{\mathsf{p}_i},T_{\mathsf{f}i},X_{\mathsf{p}_i},X_{\mathsf{c}i}\right),
\end{aligned}$$

where: the equality in (a) follows by applying the chain rule of the entropy and since, for the ISD CCIC, the encoding function $X_{\mathsf{c}i}(W_{\mathsf{c}},Y_{\mathsf{Fc}}^{i-1})$ is equivalent to $X_{\mathsf{c}i}(W_{\mathsf{c}},{T_{\mathsf{f}}}^{i-1})$; the equality in (b) is due to the fact that Y_{c} is a deterministic function of $(X_{\mathsf{c}},T_{\mathsf{p}})$, which is invertible given X_{c}; the inequality in (c) is due to the conditioning reduces entropy principle; the equality in

(d) follows because of the ISD property of the channel and since the channel is memoryless. Second pair:

$$
\begin{aligned}
& H\left({Y_{\mathsf{c}}}^N, {T_{\mathsf{c}}}^N, {T_{\mathsf{f}}}^N\right) - H\left({Y_{\mathsf{p}}}^N, {T_{\mathsf{p}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}\right) \\
\overset{\text{(e)}}{=} & \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | {Y_{\mathsf{c}}}^{i-1}, {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}\right) \\
& - H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {Y_{\mathsf{p}}}^{i-1}, {T_{\mathsf{p}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {X_{\mathsf{p}}}^i\right) \\
\overset{\text{(f)}}{=} & \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | {Y_{\mathsf{c}}}^{i-1}, {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}\right) \\
& - H\left(T_{\mathsf{c}i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{p}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {X_{\mathsf{p}}}^i\right) \\
\overset{\text{(g)}}{\leq} & \sum_{i\in[1:N]} \underbrace{H\left(T_{\mathsf{c}i} | {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}\right) - H\left(T_{\mathsf{c}i} | {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {T_{\mathsf{p}}}^{i-1}, {X_{\mathsf{p}}}^i\right)}_{=0 \text{ because of (4.8a)}} \\
& + \sum_{i\in[1:N]} H\left(T_{\mathsf{f}i} | T_{\mathsf{c}i}\right) - H\left(T_{\mathsf{f}i} | {T_{\mathsf{c}}}^i, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {T_{\mathsf{p}}}^{i-1}, {X_{\mathsf{p}}}^i, {X_{\mathsf{c}}}^i\right) \\
& + \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}\right) - H\left(T_{\mathsf{p}_i} | {T_{\mathsf{c}}}^i, {T_{\mathsf{f}}}^i, W_{\mathsf{p}}, {T_{\mathsf{p}}}^{i-1}, {X_{\mathsf{p}}}^i, {X_{\mathsf{c}}}^i\right) \\
\overset{\text{(h)}}{=} & \sum_{i\in[1:N]} H\left(T_{\mathsf{f}i} | T_{\mathsf{c}i}\right) - H\left(T_{\mathsf{f}i} | T_{\mathsf{c}i}, X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right) \\
& + \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}\right) - H\left(Y_{\mathsf{c}i} | T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right),
\end{aligned}
$$

where: the equality in (e) follows by applying the chain rule of the entropy and since, given W_{p}, X_{p} is uniquely determined; the equality in (f) is due to the fact that Y_{p} is a deterministic function of $(X_{\mathsf{p}}, T_{\mathsf{c}})$, which is invertible given X_{p}; the inequality in (g) is due to the conditioning reduces entropy principle; the equality in (h) follows because of the ISD property of the channel and since the channel is memoryless.

By combining all the terms together, by introducing the time sharing random variable uniformly distributed over $[1 : N]$ and independent of everything else, by dividing both sides by N and taking the limit for $N \to \infty$ we get the bound in (4.9f). We finally notice that by dropping the time sharing we do not decrease the bound.

4.C Proof of the outer bound in (4.7)

By Fano's inequality, by considering that the messages W_{p} and W_{c} are independent and by giving side information similarly to [47], we have

$$\begin{aligned}
& N(R_{\mathsf{p}} + 2R_{\mathsf{c}} - 3\epsilon_N) \\
\leq & I\left(W_{\mathsf{p}}; {Y_{\mathsf{p}}}^N\right) + 2I\left(W_{\mathsf{c}}; {Y_{\mathsf{c}}}^N\right) \\
\leq & I\left(W_{\mathsf{p}}; {Y_{\mathsf{p}}}^N, {T_{\mathsf{p}}}^N, {T_{\mathsf{f}}}^N\right) + I\left(W_{\mathsf{c}}; {Y_{\mathsf{c}}}^N\right) + I\left(W_{\mathsf{c}}; {Y_{\mathsf{c}}}^N, {T_{\mathsf{c}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}\right) \\
\leq & H\left({Y_{\mathsf{c}}}^N\right) - H\left({Y_{\mathsf{c}}}^N, {T_{\mathsf{c}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}, W_{\mathsf{c}}\right) \\
& + H\left({Y_{\mathsf{c}}}^N, {T_{\mathsf{c}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}\right) - H\left({Y_{\mathsf{p}}}^N, {T_{\mathsf{p}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}\right) \\
& + H\left({Y_{\mathsf{p}}}^N, {T_{\mathsf{p}}}^N, {T_{\mathsf{f}}}^N\right) - H\left({Y_{\mathsf{c}}}^N | W_{\mathsf{c}}\right).
\end{aligned}$$

We now analyze each pair of terms. In particular, we proceed similarly as we did to prove the outer bound $2R_{\mathsf{p}} + R_{\mathsf{c}}$ in (4.6).
First pair:

$$\begin{aligned}
& H\left({Y_{\mathsf{c}}}^N\right) - H\left({Y_{\mathsf{c}}}^N, {T_{\mathsf{c}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}, W_{\mathsf{c}}\right) \\
= & \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i} | {Y_{\mathsf{c}}}^{i-1}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | W_{\mathsf{p}}, W_{\mathsf{c}}, {Y_{\mathsf{c}}}^{i-1}, {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, {X_{\mathsf{p}}}^i, {X_{\mathsf{c}}}^i\right) \\
\leq & \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | W_{\mathsf{p}}, W_{\mathsf{c}}, {Y_{\mathsf{c}}}^{i-1}, {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, {X_{\mathsf{p}}}^i, {X_{\mathsf{c}}}^i\right) \\
= & \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}\right) - H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | X_{\mathsf{p}_i}, X_{\mathsf{c}i}\right).
\end{aligned}$$

Second pair:

$$\begin{aligned}
& H\left({Y_{\mathsf{c}}}^N, {T_{\mathsf{c}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}\right) - H\left({Y_{\mathsf{p}}}^N, {T_{\mathsf{p}}}^N, {T_{\mathsf{f}}}^N | W_{\mathsf{p}}\right) \\
= & \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | {Y_{\mathsf{c}}}^{i-1}, {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {X_{\mathsf{p}}}^i\right) \\
& - H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {Y_{\mathsf{p}}}^{i-1}, {T_{\mathsf{p}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {X_{\mathsf{p}}}^i\right) \\
= & \sum_{i \in [1:N]} H\left(Y_{\mathsf{c}i}, T_{\mathsf{c}i}, T_{\mathsf{f}i} | {Y_{\mathsf{c}}}^{i-1}, {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {X_{\mathsf{p}}}^i\right) \\
& - H\left(T_{\mathsf{c}i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i} | {T_{\mathsf{c}}}^{i-1}, {T_{\mathsf{p}}}^{i-1}, {T_{\mathsf{f}}}^{i-1}, W_{\mathsf{p}}, {X_{\mathsf{p}}}^i\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{i\in[1:N]} \underbrace{H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - H\left(T_{\mathsf{c}i}|T_{\mathsf{c}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{i}\right)}_{=0 \text{ because of (4.8a)}} \\
&\quad + \sum_{i\in[1:N]} H\left(T_{\mathsf{f}i}|T_{\mathsf{c}i}, X_{\mathsf{p}_i}\right) - H\left(T_{\mathsf{f}i}|T_{\mathsf{c}}^{i}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}, W_{\mathsf{p}}, X_{\mathsf{p}}^{i}\right) \\
&\quad + \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}\right) - H\left(T_{\mathsf{p}_i}|T_{\mathsf{c}}^{i}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i}, W_{\mathsf{p}}, X_{\mathsf{p}}^{i}, X_{\mathsf{c}}^{i}\right) \\
&= \sum_{i\in[1:N]} H\left(T_{\mathsf{f}i}|T_{\mathsf{c}i}, X_{\mathsf{p}_i}\right) - H\left(T_{\mathsf{f}i}|T_{\mathsf{c}i}, X_{\mathsf{p}_i}\right) \\
&\quad + \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}\right) - H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}, X_{\mathsf{p}_i}\right) \\
&= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{p}_i}\right) - H\left(Y_{\mathsf{c}i}|T_{\mathsf{c}i}, T_{\mathsf{f}i}, X_{\mathsf{c}i}, X_{\mathsf{p}_i}\right).
\end{aligned}
$$

Third pair: since

$$
\begin{aligned}
&H\left(Y_{\mathsf{c}}^{N}|W_{\mathsf{c}}\right) \\
&= \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}|Y_{\mathsf{c}}^{i-1}, W_{\mathsf{c}}\right) \\
&\geq \sum_{i\in[1:N]} H\left(Y_{\mathsf{c}i}|Y_{\mathsf{c}}^{i-1}, W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{c}}^{i}\right) \\
&= \sum_{i\in[1:N]} H\left(T_{\mathsf{p}_i}|T_{\mathsf{p}}^{i-1}, W_{\mathsf{c}}, T_{\mathsf{f}}^{i-1}, X_{\mathsf{c}}^{i}\right) \\
&= \sum_{i\in[1:N]} H\left(T_{\mathsf{p}_i}|T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}\right),
\end{aligned}
$$

where the last equality follows because of (4.8b), then

$$
\begin{aligned}
&H\left(Y_{\mathsf{p}}^{N}, T_{\mathsf{p}}^{N}, T_{\mathsf{f}}^{N}\right) - H\left(Y_{\mathsf{c}}^{N}|W_{\mathsf{c}}\right) \\
&\leq \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i}, T_{\mathsf{p}_i}, T_{\mathsf{f}i}|Y_{\mathsf{p}}^{i-1}, T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) - H\left(T_{\mathsf{p}_i}|T_{\mathsf{p}}^{i-1}, T_{\mathsf{f}}^{i-1}\right) \\
&\leq \sum_{i\in[1:N]} H\left(Y_{\mathsf{p}_i}, T_{\mathsf{f}i}|T_{\mathsf{p}_i}\right).
\end{aligned}
$$

By combining everything together, by introducing the time sharing random variable uniformly distributed over $[1:N]$ and independent of everything else, by dividing both sides by N and taking the limit for $N \to \infty$ we get the bound in (4.7). We finally notice that by dropping the time sharing we do not decrease the bound.

4.D Evaluation of the outer bounds in (4.9), (4.6) and (4.7) for the Gaussian CCIC

By defining $\mathbb{E}\left[X_{\mathsf{p}} X_{\mathsf{c}}^*\right] := \rho : |\rho| \in [0,1]$ we obtain: from the cut-set bounds in (4.9a)-(4.9c)

$$
\begin{aligned}
&R_{\mathsf{p}} \leq \log\left(1 + (\mathsf{C}+\mathsf{S_p})\left(1-|\rho|^2\right)\right) \overset{|\rho|=0}{\leq} \log\left(1+\mathsf{C}+\mathsf{S_p}\right),\\
&R_{\mathsf{p}} \leq \log\left(1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\,\Re\left\{\rho \mathrm{e}^{-\mathrm{j}\theta_{\mathsf{c}}}\right\}\right)\\
&\overset{\rho=\mathrm{e}^{\mathrm{j}\theta_{\mathsf{c}}}}{\leq} \log\left(1 + \left(\sqrt{\mathsf{S_p}} + \sqrt{\mathsf{I_c}}\right)^2\right),\\
&R_{\mathsf{c}} \leq \log\left(1 + \left(1-|\rho|^2\right)\mathsf{S_c}\right) \overset{|\rho|=0}{\leq} \log\left(1+\mathsf{S_c}\right).
\end{aligned}
$$

From the bounds in (4.9d)-(4.9e) we get

$$
\begin{aligned}
&R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{(\mathsf{S_p}+\mathsf{C})\left(1-|\rho|^2\right)}{1+\mathsf{I_p}\left(1-|\rho|^2\right)}\right) + \log\left(1 + \mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\,\Re\left\{\rho \mathrm{e}^{\mathrm{j}\theta_{\mathsf{p}}}\right\}\right)\\
&\overset{(a)}{\leq} \log\left(1 + \frac{\mathsf{S_p}+\mathsf{C}}{1+\mathsf{I_p}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right)\\
&= \underbrace{\log\left(1 + \frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right) + \log\left(1 + \left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right)}_{\text{as no cooperation / } \mathsf{C}=0}\\
&+ \underbrace{\log\left(1 + \frac{\mathsf{C}}{1+\mathsf{I_p}+\mathsf{S_p}}\right)}_{\text{increasing in } \mathsf{C}},
\end{aligned}
$$

and

$$
\begin{aligned}
&R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S_c}\left(1-|\rho|^2\right)}{1+\mathsf{I_c}\left(1-|\rho|^2\right)}\right) + \log\left(1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\,\Re\left\{\rho \mathrm{e}^{-\mathrm{j}\theta_{\mathsf{c}}}\right\}\right)\\
&\overset{(b)}{\leq} \underbrace{\log\left(1 + \frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right) + \log\left(1 + (\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}})^2\right)}_{\text{as no cooperation / } \mathsf{C}=0},
\end{aligned}
$$

where the inequality in (a) follows by evaluating the first logarithm in $|\rho| = 0$ and the second logarithm in $\rho = \mathrm{e}^{-\mathrm{j}\theta_{\mathsf{p}}}$ and the inequality in (b) follows by evaluating the first logarithm in $|\rho| = 0$ and the second logarithm in $\rho = \mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}$. Finally, from the bound in (4.9f) we obtain

$$
\begin{aligned}
R_\mathsf{p} + R_\mathsf{c} &\leq \log\left(1 + \frac{\mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\Re\left\{\rho \mathrm{e}^{-\mathrm{j}\theta_\mathsf{c}}\right\} + \left(1 - |\rho|^2\right)\left(\mathsf{I_p I_c} + \mathsf{C I_c}\right)}{1 + \mathsf{C} + \mathsf{I_p}}\right) \\
&+ \log\left(1 + \frac{\mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\Re\left\{\rho \mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\right\} + \left(1 - |\rho|^2\right)\left(\mathsf{I_p I_c} + \mathsf{C S_c}\right)}{1 + \mathsf{C} + \mathsf{I_c} + \mathsf{I_c C}\left(1 - |\rho|^2\right)}\right) \\
&+ \log\left(1 + \frac{\mathsf{C} + \mathsf{C I_c}\left(1 - |\rho|^2\right)}{1 + \mathsf{I_c}}\right) \\
&= \log\left(1 + \frac{\mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\Re\left\{\rho \mathrm{e}^{-\mathrm{j}\theta_\mathsf{c}}\right\} + \left(1 - |\rho|^2\right)\left(\mathsf{I_p I_c} + \mathsf{C I_c}\right)}{1 + \mathsf{C} + \mathsf{I_p}}\right) \\
&+ \log\left(1 + \frac{\mathsf{C} + \mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\Re\left\{\rho \mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\right\} + \left(1 - |\rho|^2\right)\left(\mathsf{I_p I_c} + \mathsf{C S_c} + \mathsf{I_c C}\right)}{1 + \mathsf{I_c}}\right) \\
&\overset{\text{(c)}}{\leq} \log\left(1 + \mathsf{I_c} + \frac{\mathsf{S_p}}{1 + \mathsf{C} + \mathsf{I_p}}\right) + \log\left(1 + \mathsf{C} + \mathsf{I_p} + \frac{\mathsf{S_c}\left(1 + \mathsf{C}\right)}{1 + \mathsf{I_c}}\right) + 2\log(2) \\
&= \underbrace{\log\left(1 + \mathsf{C} + \mathsf{I_p} + \frac{\mathsf{S_p}}{1 + \mathsf{I_c}}\right) + \log\left(1 + \mathsf{I_c} + \mathsf{S_c}\frac{1 + \mathsf{C}}{1 + \mathsf{C} + \mathsf{I_p}}\right)}_{\text{increasing in } \mathsf{C}} + 2\log(2),
\end{aligned}
$$

where the inequality in (c) follows by: (i) evaluating the first term of the first logarithm in $\rho = \mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}$ and the second term of the first logarithm in $|\rho| = 0$; (ii) evaluating the first term of the second logarithm in $\rho = \mathrm{e}^{-\mathrm{j}\theta_\mathsf{p}}$ and the second term of the second logarithm in $|\rho| = 0$; (iii) since $\log\left(1 + \left(\sqrt{|a|^2} + \sqrt{|b|^2}\right)^2\right) \leq \log\left(1 + |a|^2 + |b|^2\right) + \log(2)$.

We now evaluate the new outer bounds in Theorem 8 and we get

$$
\begin{aligned}
2R_\mathsf{p} + R_\mathsf{c} &\leq \log\left(1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p I_c}}\Re\left\{\rho \mathrm{e}^{-\mathrm{j}\theta_\mathsf{c}}\right\}\right) \\
&+ \log\left(1 + \frac{\mathsf{S_p}\left(1 - |\rho|^2\right)}{1 + \left(\mathsf{C} + \mathsf{I_p}\right)\left(1 - |\rho|^2\right)}\right) \\
&+ \log\left(1 + \frac{\mathsf{C} + \mathsf{S_c} + \mathsf{I_p} + 2\sqrt{\mathsf{S_c I_p}}\Re\left\{\rho \mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\right\} + \left(1 - |\rho|^2\right)\left(\mathsf{I_p I_c} + \mathsf{C S_c} + \mathsf{I_c C}\right)}{1 + \mathsf{I_c}}\right)
\end{aligned}
$$

$$\begin{aligned}
&\overset{\text{(d)}}{\leq} \log\left(1+\left(\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}}\right)^2\right)+\log\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}+\mathsf{C}}\right)\\
&+\log\left(1+\mathsf{C}+\mathsf{I_p}+\frac{\mathsf{S_c}\left(1+\mathsf{C}\right)}{1+\mathsf{I_c}}\right)+\log(2)\\
&=\underbrace{\log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right)+\log\left(1+\left(\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}}\right)^2\right)}_{\text{as no cooperation / }\mathsf{C}=0}+\log(2)\\
&+\underbrace{\log\left(\frac{1+\mathsf{I_p}+\mathsf{S_p}}{1+\mathsf{I_c}+\mathsf{S_c}}\right)+\log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}+\mathsf{S_p}}\right)+\log\left(1+\mathsf{I_c}+\mathsf{S_c}\frac{1+\mathsf{C}}{1+\mathsf{I_p}+\mathsf{C}}\right)}_{\text{increasing in }\mathsf{C}},
\end{aligned}$$

where the inequality in (d) follows by (i) evaluating the first logarithm in $\rho=\mathrm{e}^{\mathrm{j}\theta_{\mathsf{c}}}$, (ii) evaluating the second logarithm in $|\rho|=0$, (iii) evaluating the first term of the third logarithm in $\rho=\mathrm{e}^{-\mathrm{j}\theta_{\mathsf{p}}}$ and the second term of the third logarithm in $|\rho|=0$ and (iv) since $\log\left(1+\left(\sqrt{|a|^2}+\sqrt{|b|^2}\right)^2\right)\leq \log\left(1+|a|^2+|b|^2\right)+\log(2)$. Similarly,

$$\begin{aligned}
R_{\mathsf{p}}+2R_{\mathsf{c}} &\leq \log\left(1+\mathsf{S_c}+\mathsf{I_p}+2\sqrt{\mathsf{S_c}\mathsf{I_p}}\Re\left\{\rho\mathrm{e}^{\mathrm{j}\theta_{\mathsf{p}}}\right\}\right)\\
&+\log\left(1+\frac{\mathsf{S_c}\left(1-|\rho|^2\right)}{1+\mathsf{I_c}\left(1-|\rho|^2\right)}\right)\\
&+\log\left(1+\frac{\mathsf{S_p}+\mathsf{I_c}+2\sqrt{\mathsf{S_p}\mathsf{I_c}}\Re\left\{\rho\mathrm{e}^{-\mathrm{j}\theta_{\mathsf{c}}}\right\}+\left(1-|\rho|^2\right)\left(\mathsf{I_p}\mathsf{I_c}+\mathsf{C}\mathsf{I_c}\right)}{1+\mathsf{C}+\mathsf{I_p}}\right)\\
&+\log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}}\right)\\
&\overset{\text{(e)}}{\leq} \log\left(1+\left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right)+\log\left(1+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right)\\
&+\log\left(1+\mathsf{I_c}+\frac{\mathsf{S_p}}{1+\mathsf{C}+\mathsf{I_p}}\right)+\log\left(1+\frac{\mathsf{C}}{1+\mathsf{I_p}}\right)+\log(2)\\
&=\underbrace{\log\left(1+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right)+\log\left(1+\left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right)}_{\text{as no cooperation / }\mathsf{C}=0}\\
&+\underbrace{\log\left(\frac{1+\mathsf{I_c}+\mathsf{S_c}}{1+\mathsf{I_p}+\mathsf{S_p}}\right)+\log\left(1+\mathsf{C}+\mathsf{I_p}+\frac{\mathsf{S_p}}{1+\mathsf{I_c}}\right)}_{\text{increasing in }\mathsf{C}}+\log(2),
\end{aligned}$$

where the inequality in (e) follows by: (i) evaluating the first logarithm in $\rho = \mathrm{e}^{-\mathrm{j}\theta_\mathsf{p}}$, (ii) evaluating the second logarithm in $|\rho| = 0$ and (iii) evaluating the first term of the third logarithm in $\rho = \mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}$ and the second term of the third logarithm in $|\rho| = 0$ and again (iv) since $\log\left(1 + \left(\sqrt{|a|^2} + \sqrt{|b|^2}\right)^2\right) \leq \log\left(1 + |a|^2 + |b|^2\right) + \log(2)$.

4.E Achievable Scheme Based on Superposition Coding and Binning

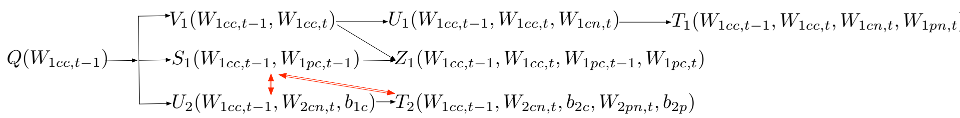


Figure 4.11: Achievable scheme based on binning and superposition coding.

We specialize the 'binning+superposition' achievable scheme in [51, Section V]. In [51, Thereom V.1] the network consists of four nodes numbered from 1 to 4; nodes 1 and 2 are sources and nodes 3 and 4 destinations; source node $j \in [1:2]$, with input to the channel X_j and output from the channel Y_j, has a message W_j for node $j+2$; destination node $j \in [3:4]$ has channel output Y_j from which it decodes message W_{j-2}. Both users use rate splitting, where the messages of user 1 / primary are both non-cooperative and cooperative, while the messages of user 2 / cognitive are non-cooperative. In [51, Section V], we set $Y_1 = S_2 = V_2 = Z_2 = \emptyset$, i.e., then $R_1 = R_{11c} + R_{10c} + R_{10n} + R_{11n}$, $R_2 = R_{22n} + R_{20n}$, to obtain a scheme that comprises: a cooperative common message (carried by the pair (Q, V_1) at rate R_{10c}) for user 1, a cooperative private message (carried by the pair (S_1, Z_1) at rate R_{11c}) for user 1, a non-cooperative common message (carried by U_1 at rate R_{10n}) for user 1, a non-cooperative private message (carried by T_1 at rate R_{11n}) for user 1, a non-cooperative common message (carried by U_2 at rate R_{20n}) for user 2 and a non-cooperative private message (carried by T_2 at rate R_{22n}) for user 2. Here the pair (Q, S_1) carries the 'past cooperative messages', and the pair (V_1, Z_1) the 'new cooperative messages' in a block Markov encoding scheme. The channel inputs are functions of the auxiliary random variables, where X_1 is a function of $(Q, S_1, Z_1, V_1, U_1, T_1)$ and X_2 is a function of (Q, S_1, U_2, T_2).

Input distributions: The set of possible input distributions is

$$\begin{aligned}
&\mathbb{P}_{Q,S_1,V_1,Z_1,U_1,T_1,X_1,U_2,T_2,X_2}\\
=&\mathbb{P}_Q P_{V_1|Q} P_{U_1,T_1|Q,V_1} \mathbb{P}_{S_1|Q} P_{Z_1|Q,S_1,V_1} \mathbb{P}_{U_2,T_2|S_1,Q}\\
&\mathbb{P}_{X_1|Q,S_1,Z_1,V_1,U_1,T_1} \mathbb{P}_{X_2|Q,S_1,U_2,T_2}.
\end{aligned} \tag{4.48}$$

A schematic representation of the achievable scheme is given in Figure 4.11, where a black arrow indicates superposition coding and a red arrow indicates binning.

Encoding: The codebooks are generated as follows: first the codebook Q is generated; then the codebook V_1 is superposed to Q, after the codebook U_1 is superposed to (Q, V_1) and finally the codebook T_1 is superposed to (Q, V_1, U_1); independently of (V_1, U_1, T_1), the codebook S_1 is superposed to Q and then the codebook Z_1 is superposed to (Q, S_1, V_1); independently of $(S_1, Z_1, V_1, U_1, T_1)$, the codebook U_2 is superposed to Q and then the codebook T_2 is superposed to (Q, U_2). With this random coding codebook generation, the pair (U_2, T_2) is independent of S_1 conditioned on Q. [51, Theorem V.1] involves several binning steps to allow for a large set of input distributions. Here the only binning steps are for (U_2, T_2) against S_1.

We use a block Markov coding scheme to convey the message of user 1 to user 2. In particular, at the end of any given time slot in a block Markov coding scheme, encoder 2 knows (Q, S_1, U_2, T_2) and decodes (V_1, Z_1) from its channel output; the decoded pair (V_1, Z_1) becomes the pair (Q, S_1) of the next time slot; then, at the beginning of each time slot, encoder 2, by binning, finds the new pair (U_2, T_2) that is jointly typical with (Q, S_1); for this to be possible, we must generate several (U_2, T_2) sequences for each message of user 2 so as to be able to find one pair to send with the correct joint distribution with (Q, S_1); this entails the rate penalties in [51, eq(20)] for user 1 and then again [51, eq(20)] for user 2 by swapping the role of the subscripts 1 and 2, with $S_2 = Z_2 = V_2 = \emptyset$, i.e.,

$$R'_{20n} + R'_{22n} \geq I(U_2, T_2; S_1|Q), \tag{4.49a}$$

$$R'_{20n} \geq I(U_2; S_1|Q). \tag{4.49b}$$

Decoding: The cooperative source uses the PDF strategy and the destinations backward decoding. There are three decoding nodes in the network and therefore three groups of rate constraints. These are:

- Node2 / CTx jointly decodes (V_1, Z_1) from its channel output with knowledge of the indices in (Q, S_1, U_2, T_2, X_2). Successful decoding is possible if (use [51, eq(21)] by swapping the role of the subscripts 1 and 2, with $S_2 = Z_2 = V_2 = \emptyset$ and with V_1 independent of S_1)

$$R_{10c} + R_{11c} \leq I(Y_2; Z_1, V_1 | U_2, T_2, X_2, S_1, Q), \tag{4.49c}$$

$$R_{11c} \leq I(Y_2; Z_1 | U_2, T_2, X_2, S_1, Q, V_1). \tag{4.49d}$$

- Node3 / PRx jointly decodes (Q, S_1, U_2, U_1, T_1) from its channel output, with knowledge of some message indices in (V_1, Z_1), by treating T_2 as noise. Successful decoding is possible if (see [51, eq(22)] with $S_2 = Z_2 = V_2 = \emptyset$)

$$R_{10c} + R_{10n} + R_{11n} + R_{20n} + R_{11c} \leq I(Y_3; Q, V_1, U_1, T_1, S_1, Z_1, U_2) - (R'_{20n} - I(U_2; S_1|Q)), \tag{4.49e}$$

$$R_{10n} + R_{11n} + R_{20n} + R_{11c} \leq I(Y_3; U_1, T_1, S_1, Z_1, U_2 | Q, V_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{4.49f}$$

$$R_{10n} + R_{11n} + R_{11c} \leq I(Y_3; U_1, T_1, S_1, Z_1 | Q, V_1, U_2), \tag{4.49g}$$

$$R_{11n} + R_{20n} + R_{11c} \leq I(Y_3; T_1, S_1, Z_1, U_2 | Q, V_1, U_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{4.49h}$$

$$R_{10n} + R_{11n} + R_{20n} \leq I(Y_3; U_1, T_1, U_2 | Q, S_1, Z_1, V_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{4.49i}$$

$$R_{11n} + R_{11c} \leq I(Y_3; T_1, S_1, Z_1 | Q, V_1, U_1, U_2), \tag{4.49j}$$

$$R_{20n} + R_{11c} \leq I(Y_3; S_1, Z_1, U_2 | Q, V_1, U_1, T_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{4.49k}$$

$$R_{10n} + R_{11n} \leq I(Y_3; U_1, T_1 | Q, S_1, Z_1, V_1, U_2), \tag{4.49l}$$

$$R_{11n} + R_{20n} \leq I(Y_3; T_1, U_2 | Q, S_1, Z_1, V_1, U_1) - (R'_{20n} - I(U_2; S_1|Q)), \tag{4.49m}$$

$$R_{11c} \leq I(Y_3; S_1, Z_1 | Q, V_1, U_1, T_1, U_2), \tag{4.49n}$$

$$R_{11n} \leq I(Y_3; T_1 | Q, S_1, Z_1, V_1, U_1, U_2). \tag{4.49o}$$

- Node4 / CRx jointly decodes (Q, U_1, U_2, T_2) from its channel output, with knowledge of some message index in V_1, by treating Z_1 and T_1 as noise (recall that the pair (U_2, T_2) has been precoded/binned against S_1). Successful decoding is possible if (see [51, eq(22)], with the role of the users swapped, where only the bounds in [51, eq(22a)], [51, eq(22h)], [51, eq(22i)], [51, eq(22j)], and [51, eq(22k)] remain after removing the redundant constraints)

$$
\begin{aligned}
&R_{10c} + R_{20n} + R_{22n} + R_{10n} \le I(Y_4; Q, U_2, T_2, V_1, U_1) \\
&\qquad - (R'_{20n} + R'_{22n}), && (4.49p) \\
&R_{20n} + R_{22n} + R_{10n} \le I(Y_4; U_2, T_2, U_1 | Q, V_1) - (R'_{20n} + R'_{22n}), && (4.49q) \\
&R_{20n} + R_{22n} \le I(Y_4; U_2, T_2 | Q, V_1, U_1) - (R'_{20n} + R'_{22n}), && (4.49r) \\
&R_{22n} + R_{10n} \le I(Y_4; T_2, U_1 | Q, U_2, V_1) - R'_{22n}, && (4.49s) \\
&R_{22n} \le I(Y_4; T_2 | Q, U_2, V_1, U_1) - R'_{22n}. && (4.49t)
\end{aligned}
$$

Compact region: Instead of applying the FME directly on the general achievable rate region, in the following we apply the FME on two particular cases, namely the case when $S_1 = Z_1 = \emptyset$ and the case $U_1 = \emptyset$. For both cases we take the constraints in (4.49a) and (4.49b) to hold with equality.

4.E.1 FME on the achievable rate region when $S_1 = Z_1 = \emptyset$

We set $S_1 = Z_1 = \emptyset$ in the achievable rate region in (4.49). After FME of the achievable region in (4.49) with $S_1 = Z_1 = \emptyset$ (see also [51, eq(8)]), we get

$$
\begin{aligned}
&R_1 \le \text{eq}(4.49\text{e}), && (4.50a) \\
&R_1 \le \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{g}), && (4.50b) \\
&R_2 \le \text{eq}(4.49\text{r}), && (4.50c) \\
&R_1 + R_2 \le \text{eq}(4.49\text{e}) + \text{eq}(4.49\text{t}), && (4.50d) \\
&R_1 + R_2 \le \text{eq}(4.49\text{j}) + \text{eq}(4.49\text{p}), && (4.50e) \\
&R_1 + R_2 \le \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{i}) + \text{eq}(4.49\text{t}), && (4.50f) \\
&R_1 + R_2 \le \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{j}) + \text{eq}(4.49\text{q}), && (4.50g) \\
&R_1 + R_2 \le \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{m}) + \text{eq}(4.49\text{s}), && (4.50h) \\
&2R_1 + R_2 \le \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{j}) + \text{eq}(4.49\text{e}) + \text{eq}(4.49\text{s}), && (4.50i) \\
&2R_1 + R_2 \le 2 \cdot \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{j}) + \text{eq}(4.49\text{i}) + \text{eq}(4.49\text{s}), && (4.50j) \\
&R_1 + 2R_2 \le \text{eq}(4.49\text{m}) + \text{eq}(4.49\text{s}) + \text{eq}(4.49\text{p}), && (4.50k) \\
&R_1 + 2R_2 \le \text{eq}(4.49\text{c}) + \text{eq}(4.49\text{m}) + \text{eq}(4.49\text{t}) + \text{eq}(4.49\text{q}), && (4.50l)
\end{aligned}
$$

for all distributions that factor as (4.48) and by setting $S_1 = Z_1 = \emptyset$ in all the mutual information terms.

We identify Node1 with the PTx (i.e., $X_{\mathsf{p}} = X_1$), Node2 with the CTx (i.e., $X_{\mathsf{c}} = X_2, T_{\mathsf{f}} = Y_2$), Node3 with the PRx (i.e., $Y_{\mathsf{p}} = Y_3$) and Node4 with the CRx (i.e., $Y_{\mathsf{c}} = Y_4$). For the Gaussian noise channel, in the region in (4.50), we choose $Q = \emptyset$, we let V_1, U_1, T_1, U_2, T_2 be i.i.d. $\mathcal{N}(0,1)$, and

$$\begin{aligned} X_{\mathsf{p}} &= a_1 U_1 + b_1 V_1 + c_1 T_1 : && |a_1|^2 + |b_1|^2 + |c_1|^2 = 1, \\ X_{\mathsf{c}} &= a_2 U_2 + b_2 T_2 : && |a_2|^2 + |b_2|^2 = 1. \end{aligned}$$

With these choices, the channel outputs are

$$\begin{aligned} T_{\mathsf{f}} &= \sqrt{\mathsf{C}}\,(a_1 U_1 + b_1 V_1 + c_1 T_1) + Z_{\mathsf{f}}, \\ Y_{\mathsf{p}} &= \sqrt{\mathsf{S_p}}\,(a_1 U_1 + b_1 V_1 + c_1 T_1) + \sqrt{\mathsf{I_c}}\mathrm{e}^{\mathrm{j}\theta_{\mathsf{c}}}\,(a_2 U_2 + b_2 T_2) + Z_{\mathsf{p}}, \\ Y_{\mathsf{c}} &= \sqrt{\mathsf{I_p}}\mathrm{e}^{\mathrm{j}\theta_{\mathsf{p}}}\,(a_1 U_1 + b_1 V_1 + c_1 T_1) + \sqrt{\mathsf{S_c}}\,(a_2 U_2 + b_2 T_2) + Z_{\mathsf{c}}, \end{aligned}$$

and the achievable region in (4.50) becomes

$$R_{\mathsf{p}} \le \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right), \tag{4.51a}$$

$$R_{\mathsf{p}} \le \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\left(|a_1|^2 + |c_1|^2\right)}\right) + \log\left(1 + \frac{\mathsf{S_p}\left(|a_1|^2 + |c_1|^2\right)}{1 + \mathsf{I_c}|b_2|^2}\right), \tag{4.51b}$$

$$R_{\mathsf{c}} \le \log\left(1 + \frac{\mathsf{S_c}}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{4.51c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{4.51d}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \le \log\left(1 + \frac{\mathsf{S_p}|c_1|^2}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(\frac{1 + \mathsf{S_c} + \mathsf{I_p}}{1 + \mathsf{I_p}|c_1|^2}\right), \tag{4.51e}$$

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} \le{}& \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\left(|a_1|^2 + |c_1|^2\right)}\right) \\ &+ \log\left(\frac{1 + \mathsf{S_p}\left(|a_1|^2 + |c_1|^2\right) + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \end{aligned} \tag{4.51f}$$

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} \le{}& \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\left(|a_1|^2 + |c_1|^2\right)}\right) + \log\left(1 + \frac{\mathsf{S_p}|c_1|^2}{1 + \mathsf{I_c}|b_2|^2}\right) \\ &+ \log\left(1 + \frac{\mathsf{S_c} + \mathsf{I_p}|a_1|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \end{aligned} \tag{4.51g}$$

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} \le{}& \log\left(\frac{1 + \mathsf{C}}{1 + \mathsf{C}\left(|a_1|^2 + |c_1|^2\right)}\right) + \log\left(\frac{1 + \mathsf{S_p}|c_1|^2 + \mathsf{I_c}}{1 + \mathsf{I_c}|b_2|^2}\right) \\ &+ \log\left(1 + \frac{\mathsf{I_p}|a_1|^2 + \mathsf{S_c}|b_2|^2}{1 + \mathsf{I_p}|c_1|^2}\right), \end{aligned} \tag{4.51h}$$

$$\tag{4.51i}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(\frac{1+\mathsf{C}}{1+\mathsf{C}\left(|a_1|^2+|c_1|^2\right)}\right) + \log\left(1+\frac{\mathsf{S_p}|c_1|^2}{1+\mathsf{I_c}|b_2|^2}\right)$$
$$+ \log\left(\frac{1+\mathsf{S_p}+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right) + \log\left(1+\frac{\mathsf{I_p}|a_1|^2+\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right), \tag{4.51j}$$

$$2R_{\mathsf{p}} + R_{\mathsf{c}} \leq 2\cdot\log\left(\frac{1+\mathsf{C}}{1+\mathsf{C}\left(|a_1|^2+|c_1|^2\right)}\right) + \log\left(1+\frac{\mathsf{S_p}|c_1|^2}{1+\mathsf{I_c}|b_2|^2}\right)$$
$$+\log\left(\frac{1+\mathsf{S_p}\left(|a_1|^2+|c_1|^2\right)+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right)+\log\left(1+\frac{\mathsf{I_p}|a_1|^2+\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right), \tag{4.51k}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(\frac{1+\mathsf{S_p}|c_1|^2+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right)+\log\left(1+\frac{\mathsf{I_p}|a_1|^2+\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right)$$
$$+ \log\left(\frac{1+\mathsf{S_c}+\mathsf{I_p}}{1+\mathsf{I_p}|c_1|^2}\right), \tag{4.51l}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(\frac{1+\mathsf{C}}{1+\mathsf{C}\left(|a_1|^2+|c_1|^2\right)}\right) + \log\left(\frac{1+\mathsf{S_p}|c_1|^2+\mathsf{I_c}}{1+\mathsf{I_c}|b_2|^2}\right)$$
$$+ \log\left(1+\frac{\mathsf{S_c}|b_2|^2}{1+\mathsf{I_p}|c_1|^2}\right) + \log\left(1+\frac{\mathsf{S_c}+\mathsf{I_p}|a_1|^2}{1+\mathsf{I_p}|c_1|^2}\right). \tag{4.51m}$$

4.E.2 FME on the achievable rate region when $U_1 = \emptyset$

After FME of the achievable region in (4.49) with $U_1 = \emptyset$, we get

$$R_1 \leq \text{eq}(4.49c) + \text{eq}(4.49o), \tag{4.52a}$$
$$R_1 \leq \text{eq}(4.49e), \tag{4.52b}$$
$$R_2 \leq \text{eq}(4.49r), \tag{4.52c}$$
$$R_1 + R_2 \leq \text{eq}(4.49c) + \text{eq}(4.49i) + \text{eq}(4.49t), \tag{4.52d}$$
$$R_1 + R_2 \leq \text{eq}(4.49e) + \text{eq}(4.49t), \tag{4.52e}$$
$$R_1 + R_2 \leq \text{eq}(4.49d) + \text{eq}(4.49o) + \text{eq}(4.49p), \tag{4.52f}$$
$$R_1 + R_2 \leq \text{eq}(4.49g) + \text{eq}(4.49p), \tag{4.52g}$$
$$R_1 + 2R_2 \leq \text{eq}(4.49f) + \text{eq}(4.49p) + \text{eq}(4.49t), \tag{4.52h}$$
$$R_1 + 2R_2 \leq \text{eq}(4.49d) + \text{eq}(4.49i) + \text{eq}(4.49p) + \text{eq}(4.49t), \tag{4.52i}$$
$$R_1 + 3R_2 \leq \text{eq}(4.49k) + \text{eq}(4.49i) + \text{eq}(4.49p) + 2\cdot\text{eq}(4.49t), \tag{4.52j}$$

for all distributions that factor as (4.48) and by setting $U_1 = \emptyset$ in all the mutual information terms.

We identify Node1 with the PTx (i.e., $X_\mathsf{p} = X_1$), Node2 with the CTx (i.e., $X_\mathsf{c} = X_2, T_\mathsf{f} = Y_2$), Node3 with the PRx (i.e., $Y_\mathsf{p} = Y_3$) and Node4 with the CRx (i.e., $Y_\mathsf{c} = Y_4$). For the Gaussian noise channel, in the achievable region in (4.52), we choose $Q = \emptyset$, we let $S_1, V_1, T_1, Z_1, U_2', T_2'$ be i.i.d. $\mathcal{N}(0,1)$, and

$$\begin{aligned}
&X_\mathsf{p} = |a_1|\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}} S_1 + b_1 V_1 + c_1 Z_1 + d_1 T_1 \ : |a_1|^2 + |b_1|^2 + |c_1|^2 + |d_1|^2 = 1,\\
&X_\mathsf{c} = |a_2| S_1 + b_2 U_2' + c_2 T_2' \ : |a_2|^2 + |b_2|^2 + |c_2|^2 = 1,\\
&U_2 = U_2' + \lambda_U S_1,\\
&T_2 = T_2' + \lambda_T S_1,
\end{aligned}$$

where

$$\begin{aligned}
\lambda_U &= \frac{\mathsf{S_c}|b_2|^2}{\mathsf{S_c}|b_2|^2 + \mathsf{S_c}|c_2|^2 + 1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)} \ \frac{\sqrt{\mathsf{I_p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}|a_1| + \sqrt{\mathsf{S_c}}|a_2|}{\sqrt{\mathsf{S_c}} b_2},\\
\lambda_T &= \frac{\mathsf{S_c}|c_2|^2}{\mathsf{S_c}|c_2|^2 + 1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)} \ \frac{\sqrt{\mathsf{I_p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}|a_1| + \sqrt{\mathsf{S_c}}|a_2| - \sqrt{\mathsf{S_c}} b_2 \lambda_U}{\sqrt{\mathsf{S_c}} c_2},
\end{aligned}$$

i.e., the choice of λ_U is so as to "pre-cancel" S_1 from Y_c in decoding U_2, i.e., so as to have $I(Y_\mathsf{c}; U_2|V_1, Q) - I(S_1; U_2|Q) = I(Y_\mathsf{c}; U_2|V_1, Q, S_1)$ and the choice of λ_T is so as to "pre-cancel" S_1 from Y_c in decoding T_2, i.e., so as to have $I(Y_\mathsf{c}; T_2|V_1, Q, U_2) - I(S_1; T_2|Q, U_2) = I(Y_\mathsf{c}; T_2|V_1, Q, U_2, S_1)$. With these choices, the channel outputs are

$$\begin{aligned}
T_\mathsf{f} &= \sqrt{\mathsf{C}}\left(|a_1|\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}} S_1 + b_1 V_1 + c_1 Z_1 + d_1 T_1\right) + Z_\mathsf{f},\\
Y_\mathsf{p} &= (\sqrt{\mathsf{S_p}}|a_1| + \sqrt{\mathsf{I_c}}|a_2|)\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}} S_1 + \sqrt{\mathsf{S_p}}\,(b_1 V_1 + c_1 Z_1 + d_1 T_1)\\
&\quad + \sqrt{\mathsf{I_c}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}\left(b_2 U_2' + c_2 T_2'\right) + Z_\mathsf{p},\\
Y_\mathsf{c} &= (\sqrt{\mathsf{I_p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{c}}|a_1| + \sqrt{\mathsf{S_c}}|a_2|) S_1 + \sqrt{\mathsf{I_p}}\mathrm{e}^{\mathrm{j}\theta_\mathsf{p}}\,(b_1 V_1 + c_1 Z_1 + d_1 T_1)\\
&\quad + \sqrt{\mathsf{S_c}}\left(b_2 U_2' + c_2 T_2'\right) + Z_\mathsf{c},
\end{aligned}$$

and the achievable region in (4.52) becomes

$$R_\mathsf{p} \le \log\left(1 + \frac{\mathsf{C}(|b_1|^2 + |c_1|^2)}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2}{1 + \mathsf{I_c}|c_2|^2}\right), \tag{4.53a}$$

$$R_\mathsf{p} \le \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p}\mathsf{I_c}|a_1|^2|a_2|^2}}{1 + \mathsf{I_c}|c_2|^2}\right), \tag{4.53b}$$

$$R_\mathsf{c} \le \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right), \tag{4.53c}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{C}(|b_1|^2 + |c_1|^2)}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right), \tag{4.53d}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(\frac{1 + \mathsf{S_p} + \mathsf{I_c} + 2\sqrt{\mathsf{S_p}\mathsf{I_c}|a_1|^2|a_2|^2}}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right), \tag{4.53e}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{C}|c_1|^2}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{4.53f}$$

$$R_{\mathsf{p}} + R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S_p}(|c_1|^2 + |d_1|^2)}{1 + \mathsf{I_c}|c_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right)$$
$$+ k_1 + k_2, \tag{4.53g}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S_p}(|c_1|^2 + |d_1|^2) + \left(\sqrt{\mathsf{S_p}}|a_1| + \sqrt{\mathsf{I_c}}|a_2|\right)^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+\log\left(1+\frac{\mathsf{S_c}(|b_2|^2+|c_2|^2)}{1+\mathsf{I_p}(|c_1|^2+|d_1|^2)}\right)+\log\left(1+\frac{\mathsf{S_c}|c_2|^2}{1+\mathsf{I_p}(|c_1|^2+|d_1|^2)}\right)+k_1, \tag{4.53h}$$

$$R_{\mathsf{p}} + 2R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{C}|c_1|^2}{1 + \mathsf{C}|d_1|^2}\right) + \log\left(1 + \frac{\mathsf{S_p}|d_1|^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + \log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{4.53i}$$

$$R_{\mathsf{p}} + 3R_{\mathsf{c}} \leq \log\left(1 + \frac{\mathsf{S_p}|c_1|^2 + \left(\sqrt{\mathsf{S_p}}|a_1| + \sqrt{\mathsf{I_c}}|a_2|\right)^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right)$$
$$+ \log\left(1 + \frac{\mathsf{S_p}|d_1|^2 + \mathsf{I_c}|b_2|^2}{1 + \mathsf{I_c}|c_2|^2}\right) + \log\left(1 + \frac{\mathsf{S_c}(|b_2|^2 + |c_2|^2)}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right)$$
$$+ 2\log\left(1 + \frac{\mathsf{S_c}|c_2|^2}{1 + \mathsf{I_p}(|c_1|^2 + |d_1|^2)}\right) + k_1, \tag{4.53j}$$

where we defined $k_1 := I\left(Y_{\mathsf{c}}; V_1\right)$ and $k_2 := I\left(Y_{\mathsf{p}}; S_1|V_1, U_2\right)$ without evaluating them for the Gaussian noise case.

Chapter 5

Case II: Half-Duplex CTx

In this chapter, we analyze the CCIC when the CTx operates in HD. Our main contributions can be summarized as follows: (i) we characterize the sum-capacity to within a constant gap for the Gaussian Z-, S- and symmetric fully-connected channels; this is accomplished by adapting the sum-capacity outer bounds of Chapter 4 to the HD case and by designing transmission strategies inspired by the LDA of the Gaussian noise channel at high SNR; (ii) we derive in closed form the (approximately) optimal schedule, i.e., the fraction of time the cognitive source listens to the channel; (iii) we highlight the regimes where the gDoF equals: (a) the one of the Gaussian non-cooperative IC, (b) the one of the ideal Gaussian non-causal CIC and (c) the one attained with a FD cognitive source.

5.1 System Model

5.1.1 General memoryless channel

The general memoryless CCIC with the CTx operating in HD mode is defined as in Section 4.1.1 (see Figure 4.1), with the difference that now the channel input at the CTx is the pair $(X_{\mathsf{c}}, M_{\mathsf{c}})$, where $M_{\mathsf{c}} \in [0:1]$ is the *state* random variable that indicates whether the CTx is in receive-mode ($M_{\mathsf{c}} = 0$) or in transmit-mode ($M_{\mathsf{c}} = 1$). As we already pointed out in Section 2.1.1 for the HD relay network, by following this approach, first proposed in [18], there is no need to develop a separate theory for memoryless networks with HD nodes as the HD constraints can be incorporated inside the FD framework.

5.1.2 Gaussian noise channel

The single-antenna Gaussian CCIC with the CTx operating in HD mode is defined similarly to the FD counterpart in Section 4.1.3 (see Figure 4.3). In particular, the input / output relationship is given by

$$\begin{bmatrix} T_{\mathsf{f}} \\ Y_{\mathsf{p}} \\ Y_{\mathsf{c}} \end{bmatrix} = \begin{bmatrix} 1-M_{\mathsf{c}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\mathsf{C}} & \star \\ \sqrt{\mathsf{S_p}} & \sqrt{\mathsf{I_c}}e^{j\theta_{\mathsf{c}}} \\ \sqrt{\mathsf{I_p}}e^{j\theta_{\mathsf{p}}} & \sqrt{\mathsf{S_c}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & M_{\mathsf{c}} \end{bmatrix} \begin{bmatrix} X_{\mathsf{p}} \\ X_{\mathsf{c}} \end{bmatrix} + \begin{bmatrix} Z_{\mathsf{f}} \\ Z_{\mathsf{p}} \\ Z_{\mathsf{c}} \end{bmatrix}. \quad (5.1)$$

The channel inputs are subject, without loss of generality, to the average power constraint $\mathbb{E}\left[|X_i|^2\right] \le 1,\ i \in \{\mathsf{p},\mathsf{c}\}$ (i.e., non-unitary power constraints can be incorporated into the channel gains) and M_{c} is the binary random variable that indicates the state of the CTx. A $\star$ in the channel transfer matrix indicates the channel gain that does not affect the capacity region because of the HD constraint. The channel parameters $(\mathsf{C},\mathsf{S_p},\mathsf{S_c},\mathsf{I_p},\mathsf{I_c},\theta_{\mathsf{p}},\theta_{\mathsf{c}}) \in \mathbb{R}_+^7$ are fixed and so known to all nodes. Some of the channel gains can be taken to be real-valued and non-negative since a node can compensate for the phase of one of its channel gains. In the following we assume that the channel sub-matrix $\begin{bmatrix} \sqrt{\mathsf{S_p}} & \sqrt{\mathsf{I_c}}e^{j\theta_{\mathsf{c}}} \\ \sqrt{\mathsf{I_p}}e^{j\theta_{\mathsf{p}}} & \sqrt{\mathsf{S_c}} \end{bmatrix}$ is full-rank (otherwise one channel output is a degraded version of the other and hence one receiver can, without loss of generality, decode all messages). The noises are independent proper-complex Gaussian random variables with, without loss of generality, zero mean and unit variance.

5.1.3 Deterministic / noiseless channel

As already introduced for the HD relay channel in Section 2.1.3, the LDA model approximates the Gaussian noise channel in (5.1) at high-SNR [19]. The LDA model is a deterministic channel which has input / output relationship

$$T_{\mathsf{f}} = (1 - M_{\mathsf{c}})\ \mathbf{S}^{n-n_{\mathsf{f}}} X_{\mathsf{p}}, \quad (5.2a)$$
$$Y_{\mathsf{p}} = \mathbf{S}^{n-n_{\mathsf{dp}}} X_{\mathsf{p}} + M_{\mathsf{c}}\ \mathbf{S}^{n-n_{\mathsf{ic}}} X_{\mathsf{c}}, \quad (5.2b)$$
$$Y_{\mathsf{c}} = \mathbf{S}^{n-n_{\mathsf{ip}}} X_{\mathsf{p}} + M_{\mathsf{c}}\ \mathbf{S}^{n-n_{\mathsf{dc}}} X_{\mathsf{c}}, \quad (5.2c)$$

where: $(n_{\mathsf{f}}, n_{\mathsf{d}k}, n_{\mathsf{i}k})$ are non-negative integers with $n := \max\{n_{\mathsf{f}}, n_{\mathsf{d}k}, n_{\mathsf{i}k}\}$ and $k \in \{\mathsf{p},\mathsf{c}\}$; M_{c} is the binary random variable that indicates the state of the CTx, all input and output vectors have length n and take value in GF(2), the sum is understood bit-wise on GF(2), and $\mathbf{S}$ is the down-shift matrix of dimension n. The model has the following interpretation. The PTx sends a length-n vector X_{p}, whose top n_{dp} bits are received at the PRx through the direct link, the top n_{ip} bits are received at the CRx through the interference link, and the top n_{f} bits are received at the CTx through the cooperation / feedback link; similarly for X_{c}. The fact that only a certain number of bits are observed at a given node is a consequence of the 'down shift' operation through the matrix $\mathbf{S}$. The bits not observed at a node are said to be 'below the noise floor'.

5.2 Overview of the main results

The exact capacity of the Gaussian HD-CCIC described in (5.1) is unknown. In this chapter, we characterize the sum-capacity to within a constant gap, which is defined as

Definition 8. *The sum-capacity of the Gaussian HD-CCIC in* (5.1) *is said to be known to within* GAP *bits per user if one can show an inner bound* $(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}$ *and an outer bound* $(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}}$ *such that*

$$\frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}}{2} \leq \mathsf{GAP},$$

where GAP *is a constant with respect to the channel parameters.*

As already noticed for the FD case, the knowledge of the sum-capacity to within a constant gap implies the exact knowledge of the gDoF, i.e., the sum-capacity pre-log factor at high SNR [12]. The gDoF is defined in Definition 7, where we use the parameterization in (4.4).

Before stating our main results of this chapter, it is worth noting that the parameters of the LDA model in (5.2) can be related to those of the Gaussian HD-CCIC in (5.1) with the parameterization in (4.4) as

$$\begin{aligned}
n_{\mathsf{dp}} &= n_{\mathsf{dc}} = n_{\mathsf{d}} = \lfloor \log(1 + \mathsf{S}_{\mathsf{p}}) \rfloor = \lfloor \log(1 + \mathsf{S}_{\mathsf{c}}) \rfloor, \\
n_{\mathsf{ip}} &= \lfloor \log(1 + \mathsf{I}_{\mathsf{p}}) \rfloor, \\
n_{\mathsf{ic}} &= \lfloor \log(1 + \mathsf{I}_{\mathsf{c}}) \rfloor, \\
n_{\mathsf{f}} &= \lfloor \log(1 + \mathsf{C}) \rfloor,
\end{aligned}$$

where we indicate $\alpha_{\mathsf{p}} := \frac{n_{\mathsf{ip}}}{n_{\mathsf{d}}}$, $\alpha_{\mathsf{c}} := \frac{n_{\mathsf{ic}}}{n_{\mathsf{d}}}$ and $\beta := \frac{n_{\mathsf{f}}}{n_{\mathsf{d}}}$ as they play the same role of the corresponding quantities in (4.4). The simplicity of the LDA model allows for the exact sum-capacity characterization in many instances where the capacity of the Gaussian counterpart is open. Moreover, the sum-capacity of the LDA normalized by $2n_{\mathsf{d}}$ equals the Gaussian gDoF defined in Definition 7.

The main contribution of this chapter is the sum-capacity characterization to within a constant gap (see Definition 8) for the symmetric case (i.e., $\mathsf{I_p} = \mathsf{I_c}$ and $\mathsf{S_c} = \mathsf{S_p}$ in (5.1)), for the symmetric Z-channel (i.e., $\mathsf{S_c} = \mathsf{S_p}$ and $\mathsf{I_p} = 0$ in (5.1)) and for the symmetric S-channel (i.e., $\mathsf{S_c} = \mathsf{S_p}$ and $\mathsf{I_c} = 0$ in (5.1)) for the case of independent noises. This constant gap result implies the closed-form characterization of the gDoF (see Definition 7) and the derivation in closed-form of the (approximately) optimal schedule, i.e., the fraction of time the CTx listens to the channel.

In order to show the constant gap results an outer and an inner bounds on the sum-capacity of the Gaussian HD-CCIC are needed. Concerning the sum-rate outer bound, we make use of some outer bounds known in the literature, namely those in [45, 47, 87]. These outer bounds were originally derived for the case of FD cooperation; in this chapter, we specialize them to the case of HD cooperation by following the approach of [18]. Concerning the inner bound our 'optimal to within a constant gap' schemes for the Gaussian HD-CCIC are inspired by the LDA model in (5.2). In particular, depending on the operating regime, we design different transmission strategies that, similarly to the FD case in Chapter 4, involve:

- Use of common (i.e., decoded also at the non-intended receiver) and private (i.e., treated as noise at the non-intended receiver) messages for both the PTx and the CTx.

- Use of cooperative (i.e., relayed to the PRx with the help of the CTx) and non-cooperative (i.e., sent without the help of the CTx) messages for the PTx and use of non-cooperative messages only for the CTx.

- Binning and superposition encoding, PDF relaying at the CTx and successive decoding both at the CTx and at the receivers, which is simpler than the joint decoding we used for the FD case.

By using these outer and inner bounds on the sum-capacity of the Gaussian HD-CCIC, in this chapter we will prove the three following theorems.

Theorem 12. *The sum-capacity of the symmetric Gaussian HD-CCIC (i.e., when* $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ *and* $\mathsf{I_p} = \mathsf{I_c} = \mathsf{I}$*) is achievable to within* 5 *bits/user.*

Theorem 13. *The sum-capacity of the symmetric Z-channel (i.e., $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_p} = 0$, the link PTx→CRx is non-existent) is characterized to within 3 bits/user.*

Theorem 14. *The sum-capacity of the symmetric S-channel (i.e., $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_c} = 0$, the link CTx→PRx is non-existent) is achievable to within 3 bits/user.*

5.3 Outer bounds on the sum-capacity for the Gaussian HD-CCIC

In this section we specialize the sum-capacity outer bounds for FD unilateral cooperation in (4.9) to the case of HD unilateral cooperation by following the approach of [18]. In particular (see Appendix 5.A for the details), we have

$$
\begin{aligned}
R_\mathsf{p} &\leq \gamma I\left(X_\mathsf{p}; Y_\mathsf{p}, T_\mathsf{f} | X_\mathsf{c}, M_\mathsf{c}=0\right) + (1-\gamma)\, I\left(X_\mathsf{p}; Y_\mathsf{p}, T_\mathsf{f} | X_\mathsf{c}, M_\mathsf{c}=1\right), && \text{(5.3a)}\\
R_\mathsf{p} &\leq H\left(M_\mathsf{c}\right) + \gamma I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p} | M_\mathsf{c}=0\right)\\
&\quad + (1-\gamma)\, I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p} | M_\mathsf{c}=1\right), && \text{(5.3b)}\\
R_\mathsf{c} &\leq H\left(M_\mathsf{c}\right) + \gamma I\left(X_\mathsf{c}; Y_\mathsf{c} | X_\mathsf{p}, M_\mathsf{c}=0\right)\\
&\quad + (1-\gamma)\, I\left(X_\mathsf{c}; Y_\mathsf{c} | X_\mathsf{p}, M_\mathsf{c}=1\right), && \text{(5.3c)}\\
R_\mathsf{p} + R_\mathsf{c} &\leq H\left(M_\mathsf{c}\right) + \gamma \left[I\left(X_\mathsf{p}; Y_\mathsf{p}, T_\mathsf{f} | Y_\mathsf{c}, X_\mathsf{c}, M_\mathsf{c}=0\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{c} | M_\mathsf{c}=0\right)\right]\\
&\quad + (1-\gamma)\left[I\left(X_\mathsf{p}; Y_\mathsf{p}, T_\mathsf{f} | Y_\mathsf{c}, X_\mathsf{c}, M_\mathsf{c}=1\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{c} | M_\mathsf{c}=1\right)\right], && \text{(5.3d)}\\
R_\mathsf{p} + R_\mathsf{c} &\leq 2H\left(M_\mathsf{c}\right) + \gamma \left[I\left(X_\mathsf{c}; Y_\mathsf{c} | Y_\mathsf{p}, X_\mathsf{p}, M_\mathsf{c}=0\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p} | M_\mathsf{c}=0\right)\right]\\
&\quad + (1-\gamma)\left[I\left(X_\mathsf{c}; Y_\mathsf{c} | Y_\mathsf{p}, X_\mathsf{p}, M_\mathsf{c}=1\right) + I\left(X_\mathsf{p}, X_\mathsf{c}; Y_\mathsf{p} | M_\mathsf{c}=1\right)\right], && \text{(5.3e)}\\
R_\mathsf{p} + R_\mathsf{c} &\leq 2H\left(M_\mathsf{c}\right)\\
&\quad + \gamma\left[I\left(Y_\mathsf{p}; X_\mathsf{p}, X_\mathsf{c} | T_\mathsf{p}, T_\mathsf{f}, M_\mathsf{c}=0\right) + I\left(Y_\mathsf{c}, T_\mathsf{f}; X_\mathsf{p}, X_\mathsf{c} | T_\mathsf{c}, M_\mathsf{c}=0\right)\right]\\
&\quad + (1-\gamma)\left[I\left(Y_\mathsf{p}; X_\mathsf{p}, X_\mathsf{c} | T_\mathsf{p}, T_\mathsf{f}, M_\mathsf{c}=1\right) + I\left(Y_\mathsf{c}, T_\mathsf{f}; X_\mathsf{p}, X_\mathsf{c} | T_\mathsf{c}, M_\mathsf{c}=1\right)\right], && \text{(5.3f)}
\end{aligned}
$$

where $\gamma := \mathbb{P}[M_\mathsf{c}=0] \in [0,1]$, i.e., γ indicates the fraction of time the CTx listens to the channel.

We now evaluate the sum-capacity outer bounds in (5.3) for the practically relevant Gaussian noise channel in (5.1), under the assumption of independent noises. We define $\begin{bmatrix} P_{\mathsf{p},\ell} & \rho_\ell \sqrt{P_{\mathsf{p},\ell}P_{\mathsf{c},\ell}} \\ \rho_\ell^* \sqrt{P_{\mathsf{p},\ell}P_{\mathsf{c},\ell}} & P_{\mathsf{c},\ell} \end{bmatrix}$ for $|\rho_\ell| \leq 1$, $\ell \in [0:1]$ and $(P_{\mathsf{p},0}, P_{\mathsf{p},1}, P_{\mathsf{c},0}, P_{\mathsf{c},1}) \in \mathbb{R}_+^4$ satisfying the power constraint

$\gamma P_{u,0} + (1-\gamma) P_{u,1} \leq 1,\ u \in \{\mathsf{p}, \mathsf{c}\}$. In particular, since all the mutual information terms in (5.3) are conditioned on $M_{\mathsf{c}} = \ell$, the 'Gaussian maximizes entropy' principle guarantees that in order to exhaust all possible input distributions it suffices to consider jointly Gaussian proper-complex inputs. By further upper bounding each mutual information term over $(\rho_0, \rho_1) \in [0,1]^2$, as well as over the phases of the channel gains and over the power allocation across the two phases, see Appendix 5.A for the details, we obtain

$$
\begin{aligned}
R_{\mathsf{p}} &\leq \gamma \log\left(1 + \mathsf{C} + \mathsf{S_p}\right) + (1-\gamma)\log\left(1+\mathsf{S_p}\right) + \log(2), && \text{(5.4a)}\\
R_{\mathsf{p}} &\leq \gamma \log\left(1+\mathsf{S_p}\right) + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S_p}} + \sqrt{\mathsf{I_c}}\right)^2\right) \\
&\quad + 2\log(2), && \text{(5.4b)}\\
R_{\mathsf{c}} &\leq (1-\gamma)\log\left(1+\mathsf{S_c}\right) + 1.52\log(2), && \text{(5.4c)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \gamma\log\left(1+\mathsf{S_p}+\mathsf{C}+\mathsf{I_p}\right) + (1-\gamma)\log\left(1+\frac{\mathsf{S_p}}{\mathsf{I_p}}\right) \\
&\quad + (1-\gamma)\log\left(1+\left(\sqrt{\mathsf{S_c}}+\sqrt{\mathsf{I_p}}\right)^2\right) + 2\log(2), && \text{(5.4d)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \gamma\log\left(1+\mathsf{S_p}\right) + (1-\gamma)\log\left(1+\frac{\mathsf{S_c}}{\mathsf{I_c}}\right) \\
&\quad + (1-\gamma)\log\left(1+\left(\sqrt{\mathsf{S_p}}+\sqrt{\mathsf{I_c}}\right)^2\right) + 3\log(2), && \text{(5.4e)}\\
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \gamma\log\left(1+\mathsf{I_p}+\mathsf{C}+\mathsf{S_p}\right) \\
&\quad + (1-\gamma)\log\left(1+\mathsf{I_c}+\frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right) \\
&\quad + (1-\gamma)\log\left(1+\mathsf{I_p}+\frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right) + 3.51\log(2). && \text{(5.4f)}
\end{aligned}
$$

In the rest of the chapter we will show that

$$
\begin{aligned}
(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} = \min\{&\mathrm{eq}(5.4\mathrm{a}) + \mathrm{eq}(5.4\mathrm{c}), \mathrm{eq}(5.4\mathrm{b}) + \mathrm{eq}(5.4\mathrm{c}), \\
&\mathrm{eq}(5.4\mathrm{d}), \mathrm{eq}(5.4\mathrm{e}), \mathrm{eq}(5.4\mathrm{f})\} && \text{(5.5)}
\end{aligned}
$$

is achievable to within a constant gap for the Z-, the S- and the symmetric Gaussian HD-CCIC. In all the three scenarios we will consider the case $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$, i.e., the two direct links are of the same strength.

5.4 Sum-capacity to within a constant gap for the symmetric Gaussian HD-CCIC

In this section, we prove Theorem 12, i.e., we characterize the sum-capacity to within a constant gap for the symmetric Gaussian HD-CCIC defined by $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_p} = \mathsf{I_c} = \mathsf{I} = \mathsf{S}^\alpha$. Figure 5.1 shows the gDoF $\mathsf{d}(\alpha, \beta)$ (abbreviated with d) and the gap (per user) for the symmetric Gaussian HD-CCIC for the different regions in the (α, β) plane, where the whole set of parameters has been partitioned into multiple sub-regions depending upon different levels of cooperation (β) and interference (α) strengths.

The proof of the constant gap result can be found in Appendix 5.B. This (approximate) sum-capacity characterization implies that the gDoF can be obtained from Definition 7 with (5.5) and equals

$$\mathsf{d}(\alpha, \beta) = \max_{\gamma \in [0,1]} \frac{1}{2} \min \Big\{ \gamma \max\{1, \beta\} + 2(1-\gamma), \tag{5.6a}$$

$$\gamma + (1-\gamma)\left(\max\{1, \alpha\} + [1-\alpha]^+\right), \tag{5.6b}$$

$$\gamma \max\{\alpha, \beta, 1\} + 2(1-\gamma)\max\{\alpha, 1-\alpha\} \Big\} \tag{5.6c}$$

$$= \begin{cases} 1 - \alpha + \frac{1}{2}\frac{[\beta-2+2\alpha]^+ \ \alpha}{\beta+\alpha-1} & \alpha \in [0, 1/2) \\ \alpha + \frac{1}{2}\frac{[\beta-2\alpha]^+ \ (2-3\alpha)}{\beta-3\alpha+1} & \alpha \in [1/2, 2/3) \\ \max\left\{1 - \frac{1}{2}\alpha, \frac{1}{2}\alpha\right\} & \alpha \in [2/3, 2) \\ 1 + \frac{1}{2}\frac{[\beta-2]^+ \ (\alpha-2)}{\beta+\alpha-3} & \alpha \in [2, \infty) \end{cases}. \tag{5.6d}$$

The gDoF expression in (5.6d), to be compared with

$$\mathsf{d}(\alpha, 0) = \min\{1, \max\{1-\alpha, \alpha\}, \max\{1-\alpha/2, \alpha/2\}\},$$

i.e., non-cooperative IC [12], and

$$\mathsf{d}(\alpha, \infty) = \max\{1-\alpha/2, \alpha/2\},$$

i.e., ideal non-causal CIC [64], has an interesting interpretation, which we shall discuss in details for the different interference regimes in the following. In particular, we will use the LDA model described in (5.2) with $\alpha_{\mathsf{p}} = \alpha_{\mathsf{c}} = \alpha$ to get insights into approximately optimal achievable schemes.

Very weak interference regime: $\alpha \in [0, 2/3)$. Without cooperation, i.e., $\beta = 0$, the tightest upper bound in this regime is (5.6c) [12]. Recall that no-cooperation in equivalent to $\gamma = 0$, i.e., the CTx never listens to

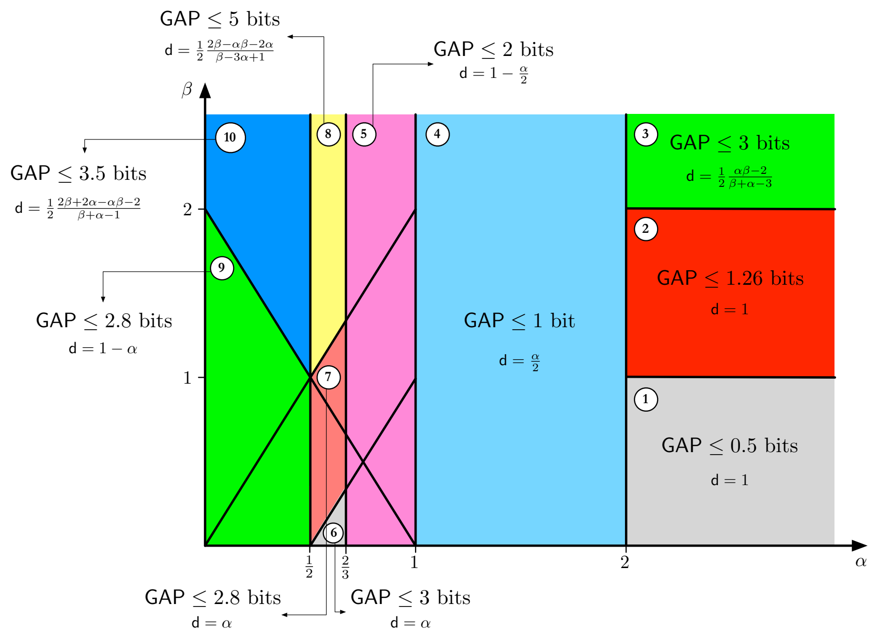


Figure 5.1: Different regimes depending on the values of α and β.

the channel. For a $\beta > 0$, the bound in (5.6c) is optimized by $\gamma = 0$, that is, by no-cooperation, whenever $\max\{1, \alpha, \beta\} \leq 2\max\{1-\alpha, \alpha\} = 2\mathsf{d}(\alpha, 0)$, which is equivalent to $\beta \leq 2\mathsf{d}(\alpha, 0)$. Intuitions suggest that the cooperation link gain should be "sufficiently strong" for HD unilateral cooperation to be beneficial. We can precisely quantify the statement "sufficiently strong" as follows: $\mathsf{d}(\alpha, \beta) > \mathsf{d}(\alpha, 0)$ if $\beta > 2\mathsf{d}(\alpha, 0)$. Recall that a strict inequality in the gDoF, or sum-capacity pre-log at high SNR, implies that the difference between the sum-capacities with HD unilateral cooperation and without cooperation becomes unbounded when SNR increases. In other words, when the cooperation link can reliably convey a rate larger than the sum-capacity of the non-cooperative IC ($\beta > 2\mathsf{d}(\alpha, 0)$), HD unilateral cooperation provides an unbounded sum-rate gain compared to the non-cooperative IC.

The optimal γ is obtained by equating the bounds in (5.6c) and (5.6b) and is given by

$$\gamma^\star = \frac{\min\{2-3\alpha, \alpha\}}{\min\{2-3\alpha, \alpha\} + \beta - 1}. \tag{5.7}$$

We now give an intuitive argument for the optimal γ in (5.7). For the case $2 - 3\alpha \leq \alpha$, i.e., $\alpha \in [1/2, 2/3)$, an achievable scheme for the LDA

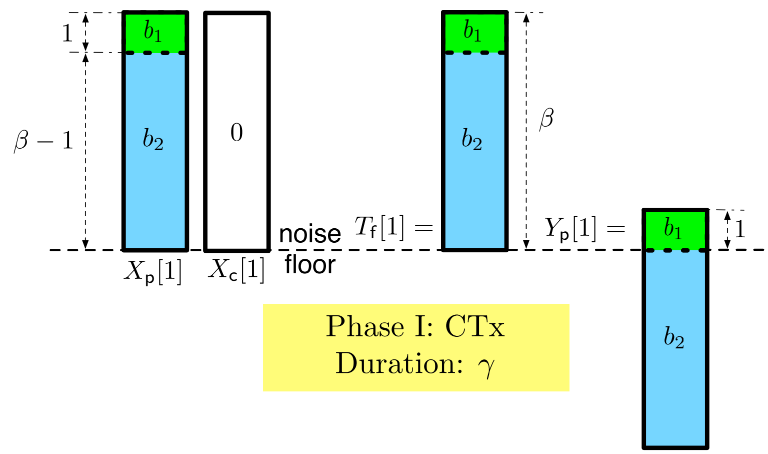


Figure 5.2: Phase I ($M_{\mathsf{c}} = 0$) common to the symmetric and asymmetric channels.

is represented in Figure 5.2 and Figure 5.3 for the case $\beta \geq 1$.[1] The bit vectors (b_1, b_2, b_3) are from the PTx to the PRx, and the bit vector b_4 from the CTx to the CRx. Since the CTx can only either receive or transmit at any point in time, we divide the transmission into two phases. *Phase I*: for a fraction $\gamma \in [0, 1]$ of the time the CTx listens to the channel and the PTx sends (b_1, b_2); b_1 is decoded by the PRx and b_2 is decoded only at the CTx. *Phase II*: for the remaining fraction $1-\gamma$ of the time the CTx transmits; the PTx sends (b_2, b_3) and the CTx sends (b_2, b_4) – notice that the PTx and the CTx cooperate in sending b_2, which hence is a cooperative message. The vectors $b_i, i \in \{3, 4\}$, are split into a common message (b_{ic}) and a private message (b_{ip}).

More specifically, in Phase I in Figure 5.2 the CTx listens to the channel and the PTx sends the vector $[b_1, b_2]$, where b_1 has normalized (by the direct link gain n_{d}) length 1 and b_2 has normalized length $\beta - 1$ (for a total normalized length of $1 + (\beta - 1) = \beta = \max\{1, \alpha, \beta\}$). Hence, over a fraction γ of the transmission time, the CTx receives $\gamma(\beta - 1)$ b_2-bits that the PRx has not received yet. In Phase II, the CTx assists the PTx to deliver these

[1]In each figure, on the LHS we represent the transmitted signals X_{p} and X_{c}, which are vectors of normalized length $n/n_{\mathsf{d}} = \max\{1, \alpha, \beta\}$, and on the RHS the received signals Y_{p} and Y_{c}, which are vectors of normalized length $\max\{1, \alpha\}$ and are the sum of a certain down shifted version of the transmitted vectors. After the down-shift operation, the top part of a vector would be populated by zero; we do not represent these zeros and instead leave an empty space in order not to clutter the figure. Note that the bits received at the same level at a node must be summed modulo-two.

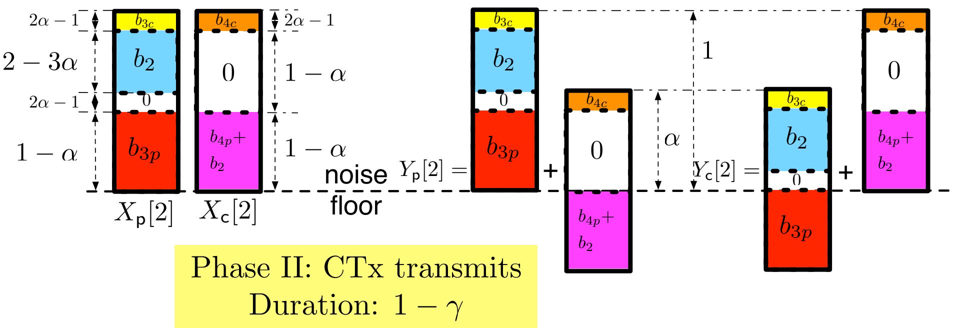


Figure 5.3: Phase II ($M_{\mathsf{c}} = 1$) for $\alpha \in [1/2, 2/3)$.

b_2-bits to the PRx in either of the two following cooperation modes: (i) the CTx relays the b_2-bits to the PRx on behalf of the PTx by spending some of its own resources, (ii) the CTx treats the b_2-bits as a 'state non-causally known at the transmitter but unknown at the receiver' and precodes its transmitted signal against it.

Phase II in Figure 5.3: the CTx sends the vector $[b_{4c}, 0, b_{4p} + b_2]$, whose components have normalized lengths $2\alpha - 1$, $1 - \alpha$ and $1 - \alpha$, respectively. In the LDA, the linear combination $b_{4p} + b_2$ can be thought of as pre-coding the signal b_{4p} against the interference caused by b_2. The PTx sends the vector $[b_{3c}, b_2, 0, b_{3p}]$, whose components have normalized lengths $2\alpha - 1$, $2 - 3\alpha$, $2\alpha - 1$ and $1 - \alpha$ (with an abuse of notation, here b_2 indicates the bits that have been received in Phase I at CTx), respectively. The CRx successively decodes b_{4c}, b_{3c}, b_{4p} in this order, while the PRx successively decodes $b_{3c}, b_2, b_{4c}, b_{3p}$ in this order. Notice that the CRx does not experience interference from b_2 when decoding b_{4p} (recall that on GF(2) $1+1 = 0+0 = 0$). The achievable rates are

$$\frac{R_{\mathsf{p}}}{n_{\mathsf{d}}} = \gamma \cdot 1 + (1-\gamma) \cdot (2 - 2\alpha),$$
$$\frac{R_{\mathsf{c}}}{n_{\mathsf{d}}} = \gamma \cdot 0 + (1-\gamma) \cdot \alpha,$$

thus giving the sum-rate

$$\frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{n_{\mathsf{d}}} = \gamma \cdot 1 + (1-\gamma) \cdot (2 - \alpha).$$

This sum-rate is larger than that without cooperation, given by $2\mathsf{d}(\alpha, 0) = 2\alpha$ [12], if $\gamma \leq \frac{2-3\alpha}{1-\alpha}$. Next, $\gamma^\star$ in (5.7) is smaller than $\frac{2-3\alpha}{1-\alpha}$ only if $\beta > 2\alpha$.

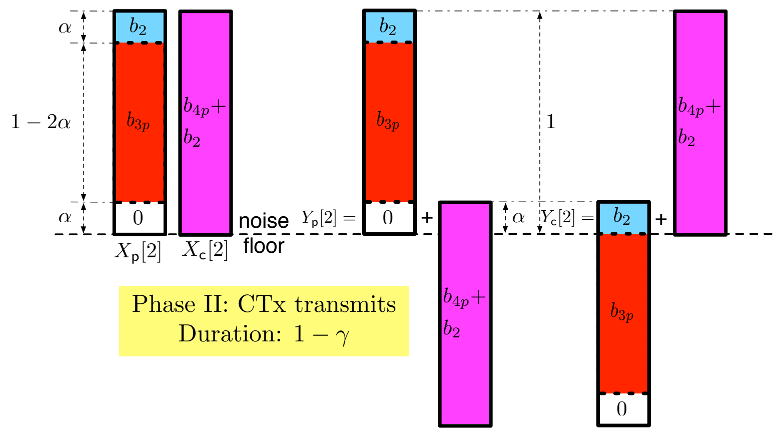


Figure 5.4: Phase II ($M_c = 1$) for $\alpha \in [0, 1/2)$.

Thus, when $\beta \leq 2\alpha$, it would take too much time for the CTx to learn the message of the PTx and it is therefore better to not cooperate at all. The last observation gives an intuitive interpretation of why the gDoF in (5.6d) contains the term $[\beta-2\alpha]^+$ for $\alpha \in [1/2, 2/3)$: the gDoF without cooperation is improved by HD unilateral cooperation only when $\beta > 2\mathsf{d}(\alpha, 0) = 2\alpha$.

A similar reasoning may be done for the case $2 - 3\alpha > \alpha$, which corresponds to $\alpha \in [0, 1/2)$. For this regime an achievable scheme is given in Figure. 5.2 and Figure 5.4, and the gDoF without cooperation is improved by HD unilateral cooperation only when $\beta > 2\mathsf{d}(\alpha, 0) = 2-2\alpha$. The scheme for $\alpha \in [0, 1/2)$ is simpler than the one for $\alpha \in [1/2, 2/3)$ in that it only involves private messages. In particular, the CTx sends the vector $[b_{4p} + b_2]$ (i.e., b_{4p} is DPC-ed against b_2), of normalized length 1; the PTx sends the vector $[b_2, b_{3p}, 0]$, whose components have normalized lengths α, $1-2\alpha$ and α, respectively; the CRx decodes b_{4p} interference free because of DPC; the PRx decodes b_2 and b_{3p} in this order; the optimal γ is such that the amount of b_2-bits received by the CTx in Phase I can be delivered to the PRx in Phase II, that is, $\gamma(\beta-1) = (1-\gamma)\alpha$ thus giving the $\gamma^\star$ in (5.7) for $\alpha < 1/2$; the achievable sum-rate is

$$\frac{(R_p + R_c)^{(\mathrm{IB})}}{n_d} = \gamma \cdot 1 + (1-\gamma) \cdot (2-\alpha).$$

Very strong interference regime: $\alpha \geq 2$. Without cooperation, i.e., $\beta = 0$, the tightest upper bound in this regime is (5.6a) [12]. For a general

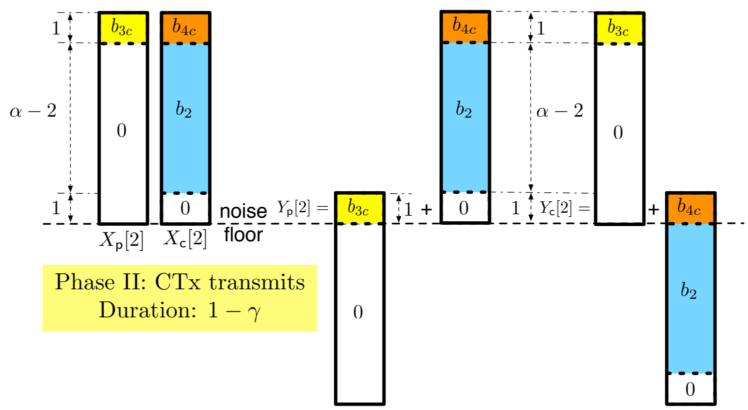


Figure 5.5: Phase II ($M_{\mathsf{c}} = 1$) for $\alpha \in [2, \infty)$.

$\beta > 0$, the bound in (5.6a) is optimized by $\gamma = 0$, which is equivalent to no-cooperation, whenever $\max\{1, \beta\} \le 2 = 2\mathsf{d}(\alpha, 0)$, which is equivalent to $\beta \le 2\mathsf{d}(\alpha, 0)$. Again we see that HD unilateral cooperation is beneficial in terms of gDoF only when β is larger than the sum-gDoF without cooperation. Here the optimal γ is obtained by equating the bounds in (5.6a) and (5.6b) and given by

$$\gamma^\star = \frac{\alpha - 2}{\beta + \alpha - 3}. \tag{5.8}$$

To see why the optimal γ is given by (5.8), we again first analyze the LDA. Phase I is the same as in Figure 5.2. In Phase II / Figure 5.5, the CTx sends $[b_{4c}, b_2, 0]$, whose components have normalized lengths 1, $\alpha - 2$, and 1 (here b_2 indicates again, with an abuse of notation, the bits that have been received in Phase I at the CTx), respectively. The PTx sends $[b_{3c}, 0]$, whose components have normalized lengths 1 and $\alpha - 1$, respectively. The CRx successively decodes b_{3c}, b_{4c} in this order. The PRx successively decodes b_{4c}, b_2, b_{3c} in this order. The achievable rates are

$$\frac{R_{\mathsf{p}}}{n_{\mathsf{d}}} = \gamma \cdot 1 + (1-\gamma)\cdot(\alpha - 1),$$
$$\frac{R_{\mathsf{c}}}{n_{\mathsf{d}}} = \gamma \cdot 0 + (1-\gamma)\cdot 1,$$

giving a sum-rate of

$$\frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{n_{\mathsf{d}}} = \gamma \cdot 1 + (1-\gamma)\cdot \alpha.$$

This sum-rate is larger than that without cooperation, given by $2\mathsf{d}(\alpha, 0) = 2$ [12], if $\gamma \leq \frac{\alpha-2}{\alpha-1}$. Next, $\gamma^\star$ in (5.8) is smaller than $\frac{\alpha-2}{\alpha-1}$ only if $\beta > 2$. Again, the interpretation is that, if $\beta \leq 2$, it takes too long to transfer bits from PTx to CTx and hence it is preferable not to cooperate. This last observation gives an intuitive interpretation of why the gDoF in (5.6d) contains the term $[\beta-2]^+$ for $\alpha \in [2, \infty)$: the gDoF without cooperation is improved only when $\beta > 2\mathsf{d}(\alpha, 0) = 2$.

Moderately weak and strong interference regimes: For $\alpha \in [2/3, 2)$ and without cooperation $\beta = 0$ the upper bound in (5.6b) is the tightest [42]. The bound in (5.6b) is always optimized by $\gamma = 0$, which is equivalent to the case of no-cooperation. Hence, in this regime it is always gDoF-optimal to operate the channel as a non-cooperative IC [12] and HD unilateral cooperation does not help in managing interference. It is very surprising that in this regime, no matter how strong the cooperation link is, unilateral causal cooperation cannot beat the performance of the non-cooperative system. In other words, $\mathsf{d}(\alpha, 0) = \mathsf{d}(\alpha, \beta) = \mathsf{d}(\alpha, \infty)$ for $\alpha \in [2/3, 2)$. For $\alpha \in [2/3, 1)$, an optimal scheme for the LDA only uses b_{3c}, b_{3p} at PTx and b_{4c}, b_{4p} at CTx [104]; for $\alpha \in [1, 2)$, b_{3c} at the PTx and b_{4c} at the CTx, both of normalized length $\alpha/2$, are optimal [104].

From the LDA to the Additive White Gaussian Noise (AWGN): In Appendix 5.E, we show how the LDA schemes described above can be 'translated' into schemes for the Gaussian HD-CCIC that are to within a constant gap from the upper bound in (5.5). The 'translation' is as follows: (i) the different pieces of information conveyed through the b-vectors in the LDA correspond to independent Gaussian codewords which are summed together and sent through the Gaussian HD-CCIC; (ii) the position, from top to bottom, of a b-vector within the transmit signal vector in the LDA corresponds to the transmit power of the corresponding Gaussian codeword in the Gaussian HD-CCIC; the higher the position of the b-vector, the larger the power of the corresponding Gaussian codeword; (iii) the length of a b-vector in the LDA corresponds to the rate of the corresponding Gaussian codeword in the Gaussian HD-CCIC; the longer the b-vector, the higher the rate of the corresponding Gaussian codeword; (iv) the transmission of the sum of two b-vectors in the LDA corresponds to a Gaussian codeword being DPC-precoded against known interference in the Gaussian HD-CCIC; (v) at the receiver side, stripping decoding is used with the Gaussian codeword corresponding to the top-most not-yet-decoded b-vector in the received LDA signal being decoded while treating the other signals as noise. Therefore, the LDA schemes described above tell us exactly: (a) how many Gaussian codewords must be superposed, with which power and at what rate, (b) if DPC is needed and if so against which interfering codeword, and (c) the decoding order at the receivers. With this, the achievable scheme is completely specified and the achievable rate can be computed.

Comparisons: In Figure 5.1, the different operating regimes are numbered from 1 to 10, each discussed in Appendix 5.B. We conclude the analysis of the symmetric Gaussian HD-CCIC with few comments:

- Everywhere, except in regions 3 (i.e., $\min\{\alpha, \beta\} \geq 2$), 8 and 10 (i.e., $\alpha \leq 2/3$ and $\beta \geq \max\{2-2\alpha, 2\alpha\}$) in Figure 5.1, HD unilateral cooperation might not be worth implementing since the same gDoF is achieved without cooperation [12].

- The symmetric Gaussian HD-CCIC attains the same gDoF of the non-causal Gaussian CIC [64] in regions 4 and 5 (i.e., $2/3 \leq \alpha \leq 2$) in Figure 5.1. Thus, in these two regions, the performance of the system, in terms of gDoF, is not worsened by allowing causal learning at CTx.

- In regions 1, 4, 5 and 6 of Figure 5.1 the gDoF equals that of the equivalent FD channel analyzed in Section 4.4 and is equal to the non-cooperative case. Since the FD channel is an outer bound for the HD channel and no-cooperative strategies are possible under the HD constraint, we conclude that in these regimes the same gap results found for the FD case in Section 4.4 hold in the case of HD source cooperation. In this case, gDoF-wise, there is no loss in having a HD CTx compared to a more powerful FD CTx.

- All the achievable schemes use successive decoding at the receivers, which, in practice, is simpler than joint decoding. Thus our proposed schemes, which are optimal to within a constant gap, may be used as guidelines to deploy practical cognitive radio systems.

- The computed gap is quite large; possible ways to reduce the gap may be: (i) apply joint decoding at the receivers; (ii) develop block Markov coding schemes instead of taking inspiration by the LDA; (iii) design achievable strategies that exploit the randomness into the switch to convey further useful information; (iv) derive tighter upper bounds than those used in this chapter.

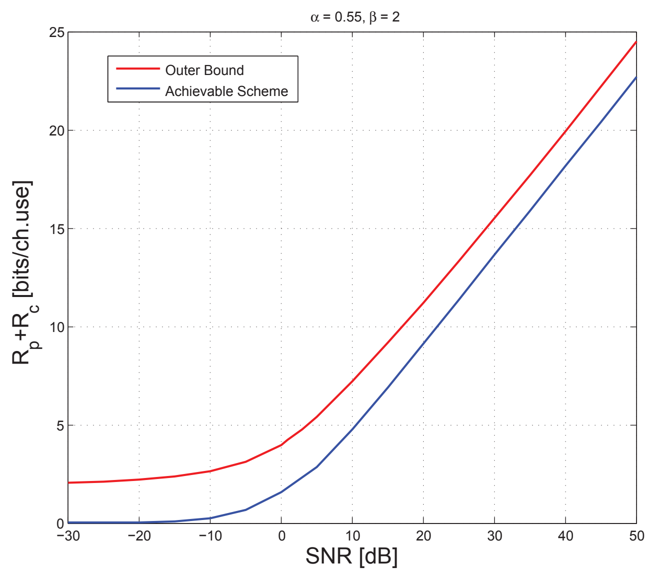


Figure 5.6: Numerical evaluation of the gap for the symmetric Gaussian HD-CCIC with $\alpha = 0.55$ and $\beta = 2$ (Region 8 in Figure 5.1).

On numerical evaluation of the gap: The gap in Theorem 12 is pessimistic and it is due to the crude bounding of the upper and lower bounds, which seems to be necessary to obtain expressions that can be easily handled and compared analytically. In order to illustrate this point, in Figure 5.6 we numerically evaluate the outer bound in (5.5) and the lower bound obtained from the scheme in Appendix 5.E.3 when the channel parameters fall into region 8 in Figure 5.1, where the gap is the largest. By numerically optimizing all the optimization variables, i.e., power splits, correlation coefficients, fraction of time the CTx listens, we observe from Figure 5.6 that the gap is of around 1.2 bits/user, i.e., more than 3 bits/user less than the analytical one. Although we can claim this gap reduction only for the simulated set of channel gains, we believe that this is a more general result.

5.5 Sum-capacity to within a constant gap for the Gaussian HD symmetric Z-channel

In this section, we prove Theorem 13, i.e., we characterize the sum-capacity to within a constant gap for the symmetric Gaussian Z-channel defined by $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_p} = 0$. Figure 5.7 shows the gDoF $\mathsf{d}^{\mathrm{Z}}(\alpha, \beta)$ (abbreviated with d) and the gap (per user) for the symmetric Gaussian Z-channel for the different regions in the (α, β) plane, where the whole set of parameters has been partitioned into multiple sub-regions depending upon different levels of cooperation (β) and interference (α) strengths.

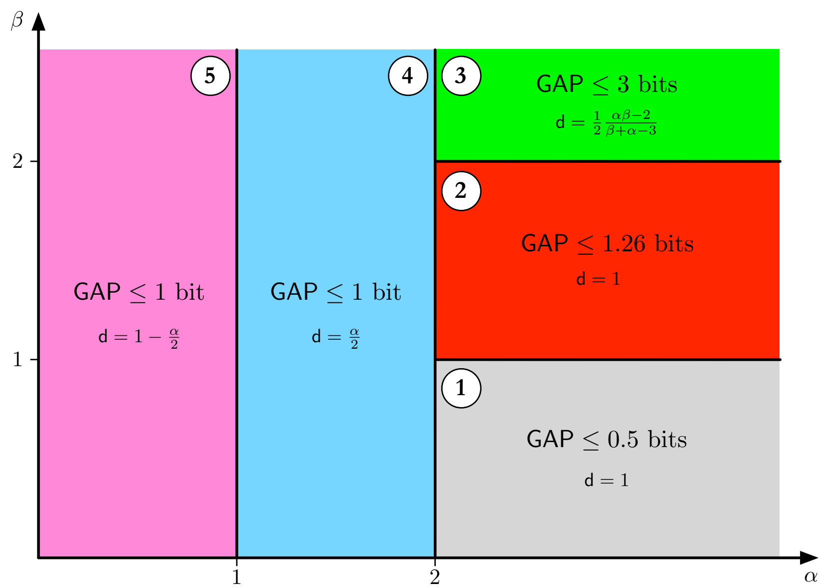


Figure 5.7: Optimal gDoF and constant gap for the Z-channel in the different regimes of (α, β).

The proof of the constant gap result can be found in Appendix 5.C. This (approximate) sum-capacity characterization implies that the gDoF can be obtained from Definition 7 with (5.5) (evaluated for $\mathsf{I_p} = 0$) and equals

$$\begin{aligned}\mathsf{d}^{\mathrm{Z}} &= \frac{1}{2} \max_{\gamma\in[0,1]} \min\left\{\gamma \max\left\{1, \beta\right\} + (1-\gamma)\, 2,\right.\\ &\qquad\qquad \left.\gamma + (1-\gamma)\left(\max\left\{1, \alpha\right\} + [1-\alpha]^{+}\right)\right\}\\ &= \begin{cases} \max\left\{1-\frac{1}{2}\alpha, \frac{1}{2}\alpha\right\} & \alpha \in [0, 2) \\ 1 + \frac{1}{2}\frac{[\beta-2]^{+}\ (\alpha-2)}{\beta+\alpha-3} & \alpha \in [2, \infty) \end{cases}. \end{aligned} \tag{5.9}$$

For future reference, for the non-cooperative Z-channel from [103] we have

$$\mathsf{d}^{\mathrm{Z}}\left(\alpha, 0\right) = \min\{\max\{1-\alpha/2, \alpha/2\},\ 1\},$$

and for the non-causal cognitive Z-channel from [64] we have

$$\mathsf{d}^{\mathrm{Z}}\left(\alpha, \infty\right) = \max\{1-\alpha/2, \alpha/2\}.$$

It hence follows that cooperation can improve the gDoF only in very strong interference, i.e., $\alpha > 2$.

The interpretation of the gDoF in (5.9) is similar to that of the interference-symmetric case in (5.6d). In particular, if the channel has weak or strong interference, i.e., $\alpha \leq 2$, the gDoF is the same as the one of the non-cooperative Z-channel [103]; in this regime it might not be worth to engage in unilateral cooperation. In very strong interference, i.e., $\alpha > 2$, unilateral cooperation gives larger gDoF than in the case of no-cooperation only when $\beta > 2\mathsf{d}^{\mathsf{Z}}(\alpha, 0) = 2$. An achievable scheme for the LDA in this regime is exactly the same developed for the corresponding interference-symmetric channel in Figure 5.2 and Figure 5.5, with the only difference that now the signal $X_{\mathsf{p}}[2]$ is not received at $Y_{\mathsf{c}}[2]$ since $\mathsf{I}_{\mathsf{p}} = 0$.

We conclude this section with few comments:

- In regions 1, 4 and 5 of Figure 5.7 the gDoF of the HD channel is as that in FD analyzed in Section 4.5 and so the same gap results found for the FD case hold in the HD case. Moreover, in region 2 in Figure 5.7 the gDoF equals that of the non-cooperative Z-channel [103]. Hence, in regions 1, 2, 4 and 5 cooperation might not be worth implementing since the same gDoF is attained without cooperation.

- The symmetric Z-channel achieves the same gDoF of the non-causal cognitive symmetric Z-channel everywhere except for $\alpha > 2$ (regions 1, 2 and 3 in Figure 5.7), i.e., for $\alpha \leq 2$, causal cognition attains the ultimate performance of the ideal non-causal cognitive Z-channel.

- By comparing Figure 5.1 and Figure 5.7, we observe that the gDoF of the Z-channel is always greater than or equal to that of the interference-symmetric channel. This is because, as already remarked for the FD case in Section 4.5, the PTx does not cooperate in sending the message of the CTx, i.e., by removing the link between PTx and CRx we rid CRx of only interfering signals. We observe that the Z-channel outperforms the interference-symmetric Gaussian HD-CCIC when $0 \leq \alpha \leq 2/3$.

5.6 Sum-capacity to within a constant gap for the Gaussian HD symmetric S-channel

In this section, we prove Theorem 14, i.e., we characterize the sum-capacity to within a constant gap for the symmetric Gaussian S-channel defined by

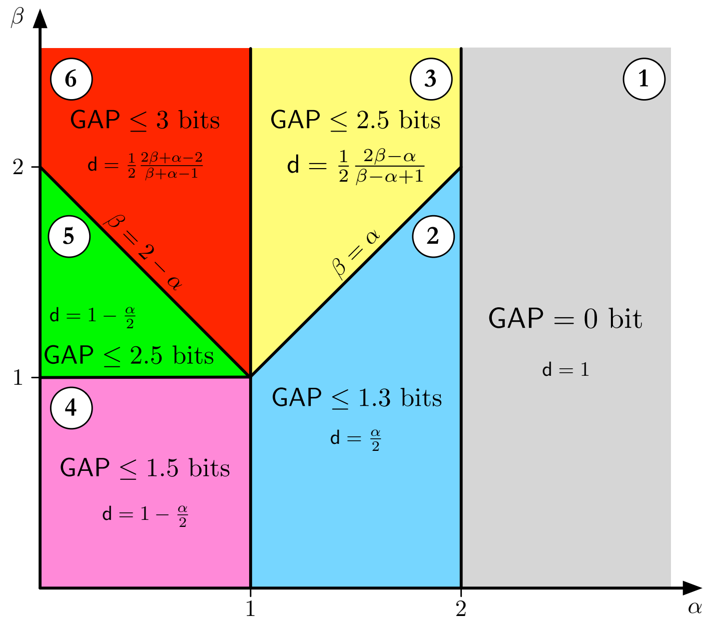


Figure 5.8: Optimal gDoF and constant gap for the S-channel in the different regimes of (α, β).

$\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{I_c} = 0$. Figure 5.8 shows the gDoF $\mathsf{d}^{\mathrm{S}}(\alpha, \beta)$ (abbreviated with d) and the gap (per user) for the symmetric Gaussian S-channel for the different regions in the (α, β) plane, where the whole set of parameters has been partitioned into multiple sub-regions depending upon different levels of cooperation (β) and interference (α) strengths.

The proof of the constant gap result can be found in Appendix 5.D. This (approximate) sum-capacity characterization implies that the gDoF can be obtained from Definition 7 with (5.5) (evaluated for $\mathsf{I_c} = 0$) and equals

$$
\begin{aligned}
\mathsf{d}^{\mathrm{S}} &= \frac{1}{2} \max_{\gamma \in [0,1]} \min \left\{ \gamma + 2\left(1-\gamma\right), \right. \\
&\qquad \left. \gamma \max\left\{\beta, \alpha, 1\right\} + \left(1-\gamma\right)\left(\max\left\{1, \alpha\right\} + [1-\alpha]^{+}\right) \right\} \\
&= \begin{cases} 1 - \frac{1}{2}\alpha + \frac{1}{2}\frac{\alpha[\alpha+\beta-2]^{+}}{\beta+\alpha-1} & \alpha \in [0,1) \\ \frac{1}{2}\alpha + \frac{1}{2}\frac{(2-\alpha)[\beta-\alpha]^{+}}{\beta-\alpha+1} & \alpha \in [1,2) \\ 1 & \alpha \in [2,\infty) \end{cases} .
\end{aligned} \tag{5.10}
$$

For future reference, for the non-cooperative S-channel from [103] we

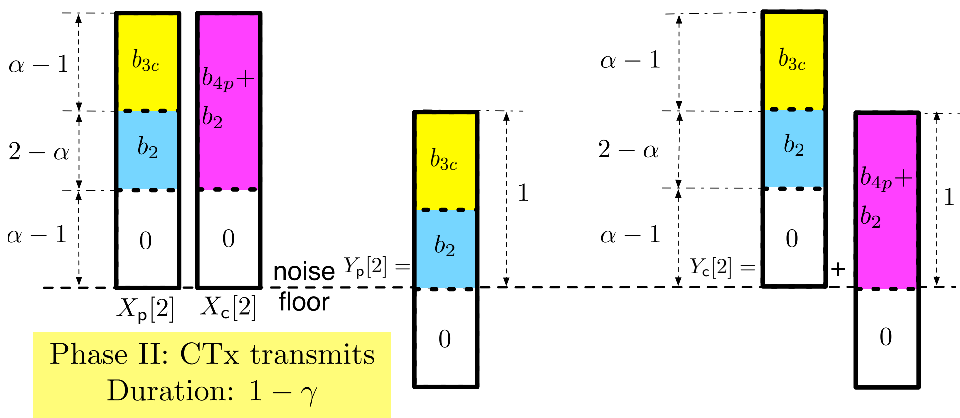


Figure 5.9: Phase II ($M_{\mathsf{c}} = 1$) for $\alpha \leq 1 \leq 2$ and $\beta > \alpha$.

have

$$\mathsf{d}^{\mathrm{S}}(\alpha, 0) = \min\{\max\{1 - \alpha/2, \alpha/2\},\ 1\},$$

and for the non-causal cognitive S-channel from [64] we have

$$\mathsf{d}^{\mathrm{S}}(\alpha, \infty) = 1.$$

It hence follows that cooperation can improve the performance only if the channel is not in very strong interference, i.e., $\alpha < 2$. It is interesting to note the different behavior of the Z- and S-channel: for the Z-channel HD unilateral cooperation is useful only in very strong interference, while for the S-channel only when not in very strong interference. Also in this case the interpretation of the gDoF in (5.10) is similar to that of the interference-symmetric case in (5.6d). In particular,

- If the channel has very strong interference, i.e., $\alpha > 2$, the gDoF is the same as for the non-cooperative S-channel [103]; in this regime it might not be worth to engage in unilateral cooperation.

- In weak and strong interference, i.e., $\alpha \leq 2$, unilateral cooperation gives larger gDoF than in the case of no-cooperation only when $\beta > 2\mathsf{d}^{\mathrm{S}}(\alpha, 0) = 2\max\{1-\alpha/2, \alpha/2\}$. A representation of the LDA schemes used for $\alpha < 2$ and $\beta > 2\max\{1 - \alpha/2, \alpha/2\}$ is given in Figure 5.2, Figure 5.9 and Figure 5.10 which can be interpreted as done for the interference-symmetric case in Section 5.4. It is worth noting that the CTx uses only private messages since it does not create interference at the PRx.

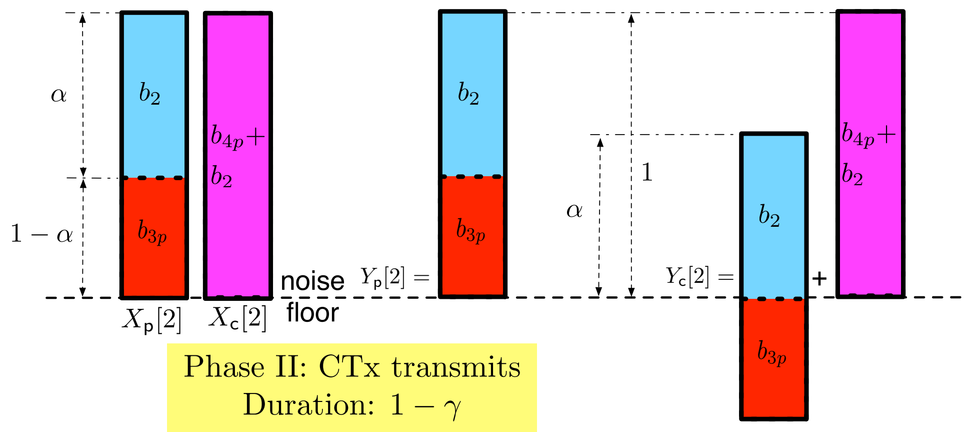


Figure 5.10: Phase II ($M_{\mathsf{c}} = 1$) for $\alpha < 1$ and $\beta > 2-\alpha$.

We conclude the analysis of the S-channel with few comments:

- There are some regimes (regions 1, 2 and 4 in Figure 5.8) in which the gDoF of the HD channel is as that in FD. In these regions, the same additive gap results found for the FD case in Section 4.6 hold in HD. Moreover, in region 5 in Figure 5.8 the gDoF equals that of the non-cooperative S-channel [103].

- The S-channel achieves the same gDoF of the non-causal cognitive S-channel [64] for $\alpha \geq 2$ (region 1 in Figure 5.8). Thus, in this region the S-channel attains the ultimate performance of the ideal non-causal cognitive S-channel.

- The S-channel outperforms the interference-symmetric Gaussian HD-CCIC when either $0 \leq \alpha \leq 2/3$ or when $\alpha \leq 2$ and $\beta \geq \max\{2-\alpha, \alpha\}$ (regions 3 and 6, and parts of regions 4 and 5 in Figure 5.8). On the other hand, the interference-symmetric Gaussian HD-CCIC outperforms the S-channel in very strong interference and strong cooperation, i.e., $\min\{\alpha, \beta\} \geq 2$. This is so because, as already remarked for the FD case in Section 4.6, in the very strong interference and cooperation regime, the system performance is enhanced by allowing the CTx to help the PTx to convey the information to the PRx, but this is not possible since $\mathsf{I}_{\mathsf{c}} = 0$.

- When $\alpha \geq 2$ (region 1 in Figure 5.8) we have an exact sum-capacity result, i.e., the gap between the sum-capacity outer bound and inner bound is equal to zero.

5.7 Conclusions and future directions

In this chapter we studied the CCIC where, differently from Chapter 4, the cognitive source, who assists the primary source in the transmission, is constrained to operate in HD mode. We analyzed both the interference-symmetric and interference-asymmetric (Z- and S-) channels, which correspond to different network deployments. For each topology we determined the sum-capacity to within a constant gap and hence the gDoF. This was accomplished by adapting the upper bounds on the sum-capacity of Chapter 4 to the HD case and by designing transmission strategies based on the LDA of the Gaussian noise channel at high SNR. In particular, the various schemes exploit binning and superposition encoding, PDF relaying and successive decoding. Moreover, by using the LDA model, we obtained a closed-form expression for the different optimization variables (e.g., schedule, power splits, coding schemes and corresponding decoding orders, etc.); this result sheds light on how to design the HD cognitive source, which is an important practical task for future wireless networks. Finally, we compared the interference-symmetric and interference-asymmetric models by highlighting the regimes where the gDoF is as that of the classical IC without cooperation and by identifying the regimes where the system attains the ultimate limits predicted by the ideal non-causal cognitive model. Moreover, we also showed that there are some regimes where no losses (in terms of gDoF) incur by considering a HD, rather than a FD, CTx.

With respect to the results presented in this chapter, interesting future research directions may include: (i) characterizing the whole capacity region (not only the sum-capacity) as done for the FD case in Chapter 4, (ii) extending the gap result to the general Gaussian channel, which is defined by 5 different channel gains and (iii) designing transmission strategies which exploit the randomness inside the switch to convey further information; this last point is critical and might lead to much smaller gap results compared to those presented in this chapter.

Appendix

5.A Derivation of the sum-capacity outer bounds and evaluation for the Gaussian noise channel

In (4.9), by following the approach of [18], we incorporate the HD constraints by substituting X_{c} with the pair $(X_{\mathsf{c}}, M_{\mathsf{c}})$. Moreover, for any triplet of random variables (A, B, C) we bound $I(A, X_{\mathsf{c}}, M_{\mathsf{c}}; B|C) \leq H(M_{\mathsf{c}}) + I(A, X_{\mathsf{c}}; B|C, M_{\mathsf{c}})$ since, for a binary-valued random variable M_{c}, we have $I(M_{\mathsf{c}}; B|C) \leq H(M_{\mathsf{c}})$.

From (4.9a) we obtain

$$\begin{aligned} R_{\mathsf{p}} &\leq I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|X_{\mathsf{c}}, M_{\mathsf{c}}\right) \\ &= \gamma I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|X_{\mathsf{c}}, M_{\mathsf{c}}=0\right) + (1-\gamma)\, I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|X_{\mathsf{c}}, M_{\mathsf{c}}=1\right), \end{aligned}$$

which is exactly (5.3a).

From (4.9b) we obtain

$$\begin{aligned} R_{\mathsf{p}} &\leq I\left(X_{\mathsf{p}}, X_{\mathsf{c}}, M_{\mathsf{c}}; Y_{\mathsf{p}}\right) \\ &= I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{p}}|M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{p}}\right) \\ &\leq H\left(M_{\mathsf{c}}\right) + \gamma I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{p}}|M_{\mathsf{c}}=0\right) + (1-\gamma)\, I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{p}}|M_{\mathsf{c}}=1\right), \end{aligned}$$

which is exactly (5.3b).

From (4.9c) we obtain

$$\begin{aligned} R_{\mathsf{c}} &\leq I\left(X_{\mathsf{c}}, M_{\mathsf{c}}; Y_{\mathsf{c}}|X_{\mathsf{p}}\right) \\ &= I\left(X_{\mathsf{c}}; Y_{\mathsf{c}}|X_{\mathsf{p}}, M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{c}}|X_{\mathsf{p}}\right) \\ &\leq H\left(M_{\mathsf{c}}\right) + \gamma I\left(X_{\mathsf{c}}; Y_{\mathsf{c}}|X_{\mathsf{p}}, M_{\mathsf{c}}=0\right) + (1-\gamma)\, I\left(X_{\mathsf{c}}; Y_{\mathsf{c}}|X_{\mathsf{p}}, M_{\mathsf{c}}=1\right), \end{aligned}$$

which is exactly (5.3c).

From (4.9d) we obtain

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} &\leq I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|Y_{\mathsf{c}}, X_{\mathsf{c}}, M_{\mathsf{c}}\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}, M_{\mathsf{c}}; Y_{\mathsf{c}}\right) \\ &= I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|Y_{\mathsf{c}}, X_{\mathsf{c}}, M_{\mathsf{c}}\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{c}}|M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{c}}\right) \\ &\leq H\left(M_{\mathsf{c}}\right) + \gamma \left[I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|Y_{\mathsf{c}}, X_{\mathsf{c}}, M_{\mathsf{c}}=0\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{c}}|M_{\mathsf{c}}=0\right)\right] \\ &\quad + (1-\gamma)\left[I\left(X_{\mathsf{p}}; Y_{\mathsf{p}}, T_{\mathsf{f}}|Y_{\mathsf{c}}, X_{\mathsf{c}}, M_{\mathsf{c}}=1\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{c}}|M_{\mathsf{c}}=1\right)\right], \end{aligned}$$

which is exactly (5.3d).

From (4.9e) we obtain

$$\begin{aligned} R_{\mathsf{p}} + R_{\mathsf{c}} &\leq I\left(X_{\mathsf{c}}, M_{\mathsf{c}}; Y_{\mathsf{c}}|Y_{\mathsf{p}}, X_{\mathsf{p}}\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}, M_{\mathsf{c}}; Y_{\mathsf{p}}\right) \\ &= I\left(X_{\mathsf{c}}; Y_{\mathsf{c}}|Y_{\mathsf{p}}, X_{\mathsf{p}}, M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{c}}|Y_{\mathsf{p}}, X_{\mathsf{p}}\right) \\ &\quad + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{p}}|M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{p}}\right) \\ &\leq 2H\left(M_{\mathsf{c}}\right) + \gamma \left[I\left(X_{\mathsf{c}}; Y_{\mathsf{c}}|Y_{\mathsf{p}}, X_{\mathsf{p}}, M_{\mathsf{c}}=0\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{p}}|M_{\mathsf{c}}=0\right)\right] \\ &\quad + (1-\gamma)\left[I\left(X_{\mathsf{c}}; Y_{\mathsf{c}}|Y_{\mathsf{p}}, X_{\mathsf{p}}, M_{\mathsf{c}}=1\right) + I\left(X_{\mathsf{p}}, X_{\mathsf{c}}; Y_{\mathsf{p}}|M_{\mathsf{c}}=1\right)\right], \end{aligned}$$

which is exactly (5.3e).

From (4.9f) we obtain

$$
\begin{aligned}
&R_{\mathsf{p}} + R_{\mathsf{c}} \le I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}}, M_{\mathsf{c}} | T_{\mathsf{p}}, T_{\mathsf{f}}\right) + I\left(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}}, M_{\mathsf{c}} | T_{\mathsf{c}}\right) \\
&= I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{p}}, T_{\mathsf{f}}, M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{p}} | T_{\mathsf{p}}, T_{\mathsf{f}}\right) \\
&\quad + I\left(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{c}}, M_{\mathsf{c}}\right) + I\left(M_{\mathsf{c}}; Y_{\mathsf{c}}, T_{\mathsf{f}} | T_{\mathsf{c}}\right) \\
&\le 2H\left(M_{\mathsf{c}}\right) + \gamma\left[I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{p}}, T_{\mathsf{f}}, M_{\mathsf{c}}=0\right) + I\left(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{c}}, M_{\mathsf{c}}=0\right)\right] \\
&\quad + (1-\gamma)\left[I\left(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{p}}, T_{\mathsf{f}}, M_{\mathsf{c}}=1\right) + I\left(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}} | T_{\mathsf{c}}, M_{\mathsf{c}}=1\right)\right],
\end{aligned}
$$

which is exactly (5.3f).

We now evaluate the sum-capacity upper bounds above for the Gaussian HD-CCIC in (5.1). We let $\begin{bmatrix} P_{\mathsf{p},\ell} & \rho_\ell \sqrt{P_{\mathsf{p},\ell} P_{\mathsf{c},\ell}} \\ \rho_\ell^* \sqrt{P_{\mathsf{p},\ell} P_{\mathsf{c},\ell}} & P_{\mathsf{c},\ell} \end{bmatrix}$ for $|\rho_\ell| \le 1$, $\ell \in [0:1]$ and $(P_{\mathsf{p},0}, P_{\mathsf{p},1}, P_{\mathsf{c},0}, P_{\mathsf{c},1}) \in \mathbb{R}_+^4$ satisfying the power constraint $\gamma P_{u,0} + (1-\gamma) P_{u,1} \le 1$, $u \in \{\mathsf{p}, \mathsf{c}\}$. In particular, since the PTx always transmits we define, for some $\tau \in [0,1]$,

$$
P_{\mathsf{p},0} = \frac{\tau}{\gamma}, \quad P_{\mathsf{p},1} = \frac{1-\tau}{1-\gamma},
$$

while, since the CTx's transmission only affects the receiver outputs when $M_{\mathsf{c}} = 1$, we let

$$
P_{\mathsf{c},0} = 0, \quad P_{\mathsf{c},1} = \frac{1}{1-\gamma}.
$$

It is also easy to see that $|\rho_0| = 0$ is the optimal choice since when the CTx is in receiving mode, everything is independent of X_{c}. For all the sum-capacity outer bounds we let $H\left(M_{\mathsf{c}}\right) = \mathcal{H}(\gamma) \le \log(2)$. All the mutual information terms below, evaluated in Gaussian noise, are already maximized with respect to $|\rho_1| \in [0,1]$; the steps of this optimization are not reported since they are similar to those in Appendix 4.D.

From (5.3a), we have

$$
\begin{aligned}
R_{\mathsf{p}} &\le \gamma \log\left(1 + \left(\mathsf{C} + \mathsf{S_p}\right) P_{\mathsf{p},0}\right) + (1-\gamma)\log\left(1 + \mathsf{S_p} P_{\mathsf{p},1}\right) \\
&= \gamma \log\left(1 + \left(\mathsf{C} + \mathsf{S_p}\right) \frac{\tau}{\gamma}\right) + (1-\gamma)\log\left(1 + \mathsf{S_p} \frac{1-\tau}{1-\gamma}\right) \\
&= \mathcal{H}(\gamma) + \gamma \log\left(\gamma + \left(\mathsf{C} + \mathsf{S_p}\right)\tau\right) + (1-\gamma)\log\left(1 - \gamma + \mathsf{S_p}(1-\tau)\right) \\
&\le \log(2) + \gamma \log\left(1 + \mathsf{C} + \mathsf{S_p}\right) + (1-\gamma)\log\left(1 + \mathsf{S_p}\right),
\end{aligned}
$$

which is exactly (5.4a).

From (5.3b), we have

$$
\begin{aligned}
R_{\mathsf{p}} &\leq \mathcal{H}(\gamma) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}} P_{\mathsf{p},0}\right) + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}} P_{\mathsf{p},1}} + \sqrt{\mathsf{I}_{\mathsf{c}} P_{\mathsf{c},1}}\right)^2\right) \\
&= \mathcal{H}(\gamma) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}} \frac{\tau}{\gamma}\right) + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}} \frac{1-\tau}{1-\gamma}} + \sqrt{\mathsf{I}_{\mathsf{c}} \frac{1}{1-\gamma}}\right)^2\right) \\
&= 2\mathcal{H}(\gamma) + \gamma \log\left(\gamma + \mathsf{S}_{\mathsf{p}} \tau\right) + (1-\gamma)\log\left(1-\gamma + \left(\sqrt{\mathsf{S}_{\mathsf{p}}(1-\tau)} + \sqrt{\mathsf{I}_{\mathsf{c}}}\right)^2\right) \\
&\leq 2\log(2) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}}\right) + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}}} + \sqrt{\mathsf{I}_{\mathsf{c}}}\right)^2\right),
\end{aligned}
$$

which is exactly (5.4b).

From (5.3c), we have

$$
\begin{aligned}
R_{\mathsf{c}} &\leq \mathcal{H}(\gamma) + (1-\gamma)\log\left(1 + \mathsf{S}_{\mathsf{c}} P_{\mathsf{c},1}\right) \\
&= \mathcal{H}(\gamma) + (1-\gamma)\log\left(1 + \mathsf{S}_{\mathsf{c}} \frac{1}{1-\gamma}\right) \\
&= \mathcal{H}(\gamma) - (1-\gamma)\log(1-\gamma) + (1-\gamma)\log\left(1-\gamma + \mathsf{S}_{\mathsf{c}}\right) \\
&\leq 1.52\log(2) + (1-\gamma)\log\left(1 + \mathsf{S}_{\mathsf{c}}\right),
\end{aligned}
$$

which is exactly (5.4c).

From (5.3d), we have

$$
\begin{aligned}
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq \mathcal{H}(\gamma) + \gamma \log\left(1 + (\mathsf{S}_{\mathsf{p}} + \mathsf{C} + \mathsf{I}_{\mathsf{p}}) P_{\mathsf{p},0}\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{p}} P_{\mathsf{p},1}}{1 + \mathsf{I}_{\mathsf{p}} P_{\mathsf{p},1}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{c}} P_{\mathsf{c},1}} + \sqrt{\mathsf{I}_{\mathsf{p}} P_{\mathsf{p},1}}\right)^2\right) \\
&= \mathcal{H}(\gamma) + \gamma \log\left(1 + (\mathsf{S}_{\mathsf{p}} + \mathsf{C} + \mathsf{I}_{\mathsf{p}}) \frac{\tau}{\gamma}\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{p}} \frac{1-\tau}{1-\gamma}}{1 + \mathsf{I}_{\mathsf{p}} \frac{1-\tau}{1-\gamma}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{c}} \frac{1}{1-\gamma}} + \sqrt{\mathsf{I}_{\mathsf{p}} \frac{1-\tau}{1-\gamma}}\right)^2\right) \\
&= 2\mathcal{H}(\gamma) + \gamma \log\left(\gamma + (\mathsf{S}_{\mathsf{p}} + \mathsf{C} + \mathsf{I}_{\mathsf{p}}) \tau\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{p}}(1-\tau)}{1-\gamma + \mathsf{I}_{\mathsf{p}}(1-\tau)}\right) \\
&\quad + (1-\gamma)\log\left(1-\gamma + \left(\sqrt{\mathsf{S}_{\mathsf{c}}} + \sqrt{\mathsf{I}_{\mathsf{p}}(1-\tau)}\right)^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq 2\log(2) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}} + \mathsf{C} + \mathsf{I}_{\mathsf{p}}\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{p}}}{\mathsf{I}_{\mathsf{p}}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{c}}} + \sqrt{\mathsf{I}_{\mathsf{p}}}\right)^2\right),
\end{aligned}
$$

which is exactly (5.4d).

From (5.3e), we have

$$
\begin{aligned}
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq 2\mathcal{H}(\gamma) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}} P_{\mathsf{p},0}\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{c}} P_{\mathsf{c},1}}{1 + \mathsf{I}_{\mathsf{c}} P_{\mathsf{c},1}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}} P_{\mathsf{p},1}} + \sqrt{\mathsf{I}_{\mathsf{c}} P_{\mathsf{c},1}}\right)^2\right) \\
&= 2\mathcal{H}(\gamma) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}} \frac{\tau}{\gamma}\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{c}} \frac{1}{1-\gamma}}{1 + \mathsf{I}_{\mathsf{c}} \frac{1}{1-\gamma}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}} \frac{1-\tau}{1-\gamma}} + \sqrt{\mathsf{I}_{\mathsf{c}} \frac{1}{1-\gamma}}\right)^2\right) \\
&= 3\mathcal{H}(\gamma) + \gamma \log\left(\gamma + \mathsf{S}_{\mathsf{p}} \tau\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{c}}}{1-\gamma+\mathsf{I}_{\mathsf{c}}}\right) \\
&\quad + (1-\gamma)\log\left(1 - \gamma + \left(\sqrt{\mathsf{S}_{\mathsf{p}}(1-\tau)} + \sqrt{\mathsf{I}_{\mathsf{c}}}\right)^2\right) \\
&\leq 3\log(2) + \gamma \log\left(1 + \mathsf{S}_{\mathsf{p}}\right) + (1-\gamma)\log\left(1 + \frac{\mathsf{S}_{\mathsf{c}}}{\mathsf{I}_{\mathsf{c}}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \left(\sqrt{\mathsf{S}_{\mathsf{p}}} + \sqrt{\mathsf{I}_{\mathsf{c}}}\right)^2\right),
\end{aligned}
$$

which is exactly (5.4e).

From (5.3f), we have

$$
\begin{aligned}
R_{\mathsf{p}} + R_{\mathsf{c}} &\leq 2\mathcal{H}(\gamma) + \gamma \log\left(1 + \left(\mathsf{I}_{\mathsf{p}} + \mathsf{C} + \mathsf{S}_{\mathsf{p}}\right) P_{\mathsf{p},0}\right) \\
&\quad + (1-\gamma)\log\left(1 + \mathsf{I}_{\mathsf{c}} P_{\mathsf{c},1} + \frac{\mathsf{S}_{\mathsf{p}} P_{\mathsf{p},1}}{1 + \mathsf{I}_{\mathsf{p}} P_{\mathsf{p},1}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \mathsf{I}_{\mathsf{p}} P_{\mathsf{p},1} + \frac{\mathsf{S}_{\mathsf{c}} P_{\mathsf{c},1}}{1 + \mathsf{I}_{\mathsf{c}} P_{\mathsf{c},1}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= 2\mathcal{H}(\gamma) + \gamma \log\left(1 + (\mathsf{I_p} + \mathsf{C} + \mathsf{S_p})\,\frac{\tau}{\gamma}\right) \\
&\quad + (1-\gamma)\log\left(1 + \mathsf{I_c}\frac{1}{1-\gamma} + \frac{\mathsf{S_p}\frac{1-\tau}{1-\gamma}}{1+\mathsf{I_p}\frac{1-\tau}{1-\gamma}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \mathsf{I_p}\frac{1-\tau}{1-\gamma} + \frac{\mathsf{S_c}\frac{1}{1-\gamma}}{1+\mathsf{I_c}\frac{1}{1-\gamma}}\right) \\
&= 3\mathcal{H}(\gamma) - (1-\gamma)\log(1-\gamma) + \gamma\log\left(\gamma + (\mathsf{I_p} + \mathsf{C} + \mathsf{S_p})\,\tau\right) \\
&\quad + (1-\gamma)\log\left(1 - \gamma + \mathsf{I_c} + \frac{\mathsf{S_p}(1-\tau)}{1+\mathsf{I_p}\frac{1-\tau}{1-\gamma}}\right) \\
&\quad + (1-\gamma)\log\left(1 - \gamma + \mathsf{I_p}(1-\tau) + \frac{\mathsf{S_c}}{1+\mathsf{I_c}\frac{1}{1-\gamma}}\right) \\
&\leq 3.51\log(2) + \gamma\log\left(1 + \mathsf{I_p} + \mathsf{C} + \mathsf{S_p}\right) \\
&\quad + (1-\gamma)\log\left(1 + \mathsf{I_c} + \frac{\mathsf{S_p}}{1+\mathsf{I_p}}\right) \\
&\quad + (1-\gamma)\log\left(1 + \mathsf{I_p} + \frac{\mathsf{S_c}}{1+\mathsf{I_c}}\right),
\end{aligned}
$$

which is exactly (5.4f).

5.B Proof of Theorem 12

Let $\mathsf{I_p} = \mathsf{I_c} = \mathsf{I} = \mathsf{S}^{\alpha}$, $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ and $\mathsf{d}(\alpha, \beta) = \mathsf{d}$ for brevity. We analyze the different regimes in Figure 5.1.

Regime 1 ($\alpha \geq 2$, $\beta \leq 1$) / Very Strong Interference 1: in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and $\mathsf{C} \leq \mathsf{S}$ and we have $\mathsf{d} \leq 1$ as in the FD case. Thus, we have

$$
\mathsf{GAP} = \frac{(\text{eq}(4.16\text{a}) + \text{eq}(4.16\text{b})) - (\text{eq}(4.15\text{a}) + \text{eq}(4.15\text{b}))}{2} \leq 0.5 \text{ bits/user.}
$$

Regime 2 ($\alpha \geq 2$, $1 < \beta \leq 2$) / Very Strong Interference 2: in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and $\mathsf{S} < \mathsf{C} \leq \mathsf{S}(1+\mathsf{S})$ and we have $\mathsf{d} \leq 1$, achieved with $\gamma = 0$, as in the classical non-cooperative IC [12]. We hence use the same transmission strategy as in the non-cooperative IC (see also Section 4.4.1), whose achievable sum-rate is given by

$$
(R_{\mathsf{p}} + R_{\mathsf{c}})^{\text{IB}} \leq 2\log\left(1 + \mathsf{S}\right).
$$

The tightest sum-capacity outer bound in (5.5) is

$$
\begin{aligned}
(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} &\leq \mathrm{eq}(5.4\mathrm{a}) + \mathrm{eq}(5.4\mathrm{c}) \\
&\overset{\mathsf{C} \leq \mathsf{S}(1+\mathsf{S})}{\leq} 2\log(1+\mathsf{S}) + 2.52\log(2).
\end{aligned}
$$

Thus,

$$
\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}}{2} \leq 1.26 \text{ bits/user}.
$$

Regime 3 ($\min\{\alpha, \beta\} \geq 2$**) / Very Strong Interference 3:** in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and $\mathsf{C} > \mathsf{S}(1+\mathsf{S})$ and we have

$$
2\mathsf{d} \leq \max_{\gamma} \min\left\{\gamma\beta + 2(1-\gamma), \gamma + (1-\gamma)\alpha\right\}.
$$

In this expression the first term is increasing in γ while the second one is decreasing in γ. Thus the optimal γ can be found by equating the two terms and is given by $\gamma^\star = \frac{\alpha-2}{\beta+\alpha-3}$ as in (5.8), which leads to

$$
\mathsf{d} \leq \frac{1}{2}\frac{\beta\alpha - 2}{\beta + \alpha - 3}.
$$

In this regime we use the transmission strategy described in Appendix 5.E.1 and Appendix 5.E.2 that, since the interference is very strong, uses common messages only. The sum-rate achieved with this scheme is

$$
\begin{aligned}
(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} &= \gamma' \log(1+\mathsf{S}) + \gamma' \log\left(1 + \frac{\mathsf{C}}{1+\mathsf{S}}\right) - \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{S}}\right) \\
&\quad + 2\left(1-\gamma'\right)\log(1+\mathsf{S}),
\end{aligned} \tag{5.11}
$$

where (see Appendix 5.E.2 for the details) $\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x} \overset{\mathsf{SNR}\gg 1}{\to} \gamma^\star$ as in (5.8) for $x := \log\left(1 + \frac{\mathsf{I}}{(1+\mathsf{S})^2}\right) \overset{\mathsf{SNR}\gg 1}{\to} \alpha - 2$ (recall that $\log\left(1 + \frac{\mathsf{C}}{1+\mathsf{S}}\right) \overset{\mathsf{SNR}\gg 1}{\to} \beta - 1$). The tightest sum-capacity outer bounds in (5.5) for this regime are eq(5.4a) + eq(5.4c)

$$
(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} \leq \gamma \log(1 + \mathsf{C} + \mathsf{S}) + 2(1-\gamma)\log(1+\mathsf{S}) + 2.52\log(2),
$$

which gives (note that we evaluate $(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}$ for the optimal γ in the upper bound; this is a possible suboptimal choice)

$$
\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{2} \leq 1.76 \text{ bits/user},
$$

and (5.4e)

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} \le \gamma\log(1+\mathsf{S}) + (1-\gamma)\log\left(1+\frac{\mathsf{S}}{\mathsf{I}}\right)$$
$$+ (1-\gamma)\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) + 3\log(2),$$

which gives

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \le 3 \text{ bits/user}.$$

Thus, for this regime we have $\mathsf{GAP} \le 3$ bits/user.

Regime 4 ($1 \le \alpha < 2$) / Strong Interference: in this regime $\mathsf{S} \le \mathsf{I} < \mathsf{S}(1+\mathsf{S})$ and we have $\mathsf{d} \le \frac{\alpha}{2}$ as in the FD case and in the classical IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative IC in strong interference (common messages only), i.e.,

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})} = \log(1+\mathsf{S}+\mathsf{I}).$$

As upper bound on the sum-capacity for this regime we use the one in (4.14e), valid for the FD case, which can be further upper bounded as

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} \le \log(1+\mathsf{S}+\mathsf{I}) + 2\log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \le 1 \text{ bit/user}.$$

Regime 5 ($2/3 \le \alpha < 1$) / Moderately Weak Interference: in this regime $\mathsf{I} < \mathsf{S}$ and $\mathsf{S}(\mathsf{S}+1) \le \mathsf{I}(\mathsf{I}+1)^2$ and we have $\mathsf{d} \le 1-\frac{\alpha}{2}$ as in the FD case and in the classical IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative IC in moderately weak interference (common and private messages with the power split of [12]), i.e.,

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})} = \log(1+\mathsf{S}+\mathsf{I}) + \log\left(\frac{1+\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2).$$

As upper bound on the sum-capacity for this regime we use the one in (4.14e), valid for the FD case, which can be further upper bounded as

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} \le \log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log(1+\mathsf{S}+\mathsf{I}) + \log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})}-(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 2 \text{ bits/user}.$$

Regime 6 ($1/2 \leq \alpha < 2/3$, $\beta \leq 2\alpha - 1$) / Weak Interference 1: in this regime $\mathsf{S}(\mathsf{S}+1) > \mathsf{I}(\mathsf{I}+1)^2$, $\mathsf{S} \leq \mathsf{I}\,(1+\mathsf{I})$ and $\mathsf{C} \leq \frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}}$ and we have $\mathsf{d} \leq \alpha$ as in the FD case and in the classical IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative IC in weak interference (common and private messages with the power split of [12]), i.e.,

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})} = 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2).$$

As upper bound on the sum-capacity for this regime we use the one in (4.14f), valid for the FD case, which can be further upper bounded as

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} \leq 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + 4\log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})}-(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 3 \text{ bits/user}.$$

Regime 7 ($1/2 \leq \alpha < 2/3$, $2\alpha - 1 < \beta \leq 2\alpha$) / Weak Interference 2: in this regime $\mathsf{S}(\mathsf{S}+1) > \mathsf{I}(\mathsf{I}+1)^2$, $\mathsf{S} \leq \mathsf{I}\,(1+\mathsf{I})$ and $\frac{\mathsf{I}(1+\mathsf{I})}{1+\mathsf{S}} < \mathsf{C} \leq \mathsf{I}^2$ and we have $\mathsf{d} \leq \alpha$, achieved with $\gamma = 0$, as in the classical non-cooperative IC [12]. We hence use the same transmission strategy as in the non-cooperative IC, whose achievable sum-rate is given by

$$(R_\mathsf{p}+R_\mathsf{c})^{\mathrm{IB}} \leq 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\log(2).$$

The tightest sum-capacity outer bound in (5.5) is

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{\mathrm{OB}} &\leq \text{eq}(5.4\text{f}) \\ &\overset{\mathsf{C}\leq\mathsf{I}^2}{\leq} 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + 3.51\log(2).\end{aligned}$$

Thus,

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{\mathrm{OB}}-(R_\mathsf{p}+R_\mathsf{c})^{\mathrm{IB}}}{2} \leq 2.8 \text{ bits/user}.$$

Regime 8 ($1/2 \leq \alpha < 2/3, \beta > 2\alpha$) / Weak Interference 3: in this regime $\mathsf{S}(\mathsf{S}+1) > \mathsf{I}(\mathsf{I}+1)^2$, $\mathsf{S} \leq \mathsf{I}\,(1+\mathsf{I})$ and $\mathsf{C} > \mathsf{I}^2$ and we have

$$2\mathsf{d} \leq \max_{\gamma} \min\left\{\gamma + (1-\gamma)(2-\alpha), \gamma\beta + 2(1-\gamma)\alpha\right\}.$$

In this expression the first term is decreasing in γ while the second one is increasing in γ. Thus the optimal γ can be found by equating the two terms and is given by $\gamma^\star = \frac{2-3\alpha}{\beta-3\alpha+1}$ in (5.7), which leads to $\mathsf{d} \leq \frac{1}{2}\frac{2\beta-\alpha\beta-2\alpha}{\beta-3\alpha+1}$. In this regime we use the transmission strategy described in Appendix 5.E.1 and Appendix 5.E.3 that, since the interference is weak, uses both common and private messages. The sum-rate achieved with this scheme is

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})} &= \gamma' \log(1+\mathsf{S}) - \gamma' \log\left(1+\frac{\mathsf{S}}{1+\mathsf{S}}\right) + \gamma' \log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)\\ &+ 2\left(1-\gamma'\right)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - 2\left(1-\gamma'\right)\log\left(1+\frac{\mathsf{I}}{1+\mathsf{I}}\right)\\ &+ \left(1-\gamma'\right)\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}+\frac{\mathsf{I}}{1+\mathsf{I}}\right)\\ &- \left(1-\gamma'\right)\log\left(1+\frac{\mathsf{S}\mathsf{I}+\mathsf{I}+\mathsf{I}^2}{(1+\mathsf{I})^2}+\frac{\mathsf{S}}{1+\mathsf{I}}\right),\end{aligned} \tag{5.12}$$

where (see Appendix 5.E.3 for the details) $\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x} \overset{\mathsf{S}\gg 1}{\rightarrow} \gamma^\star$ as in (5.7) for $x := \log\left(1+\frac{\mathsf{S}^2}{(1+\mathsf{I})^3+\mathsf{S}+\mathsf{S}\mathsf{I}}\right) \overset{\mathsf{SNR}\gg 1}{\rightarrow} 2-3\alpha$. The tightest sum-capacity outer bounds in (5.5) for this regime are (5.4e)

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} &\leq \gamma \log(1+\mathsf{S}) + (1-\gamma)\log\left(1+\frac{\mathsf{S}}{\mathsf{I}}\right)\\ &+ (1-\gamma)\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) + 3\log(2),\end{aligned}$$

which gives (note that we evaluate $(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}$ for the optimal γ in the upper bound; this is a possible suboptimal choice)

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 5 \text{ bits/user},$$

and (5.4f)

$$\begin{aligned}(R_{\mathsf{p}}+R_{\mathsf{c}})^{(\mathrm{OB})} \leq\ & \gamma \log\left(1+\mathsf{I}+\mathsf{C}+\mathsf{S}\right)\\ &+(1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)\\ &+(1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+3.51\log(2),\end{aligned}$$

which gives

$$\mathsf{GAP}=\frac{(R_{\mathsf{p}}+R_{\mathsf{c}})^{(\mathrm{OB})}-(R_{\mathsf{p}}+R_{\mathsf{c}})^{(\mathrm{IB})}}{2}\leq 3.3\text{ bits/user}.$$

Thus, for this regime we have $\mathsf{GAP}\leq 5$ bits/user.

Regime 9 ($\alpha<1/2,\beta\leq 2-2\alpha$) / Weak Interference 4: in this regime $\mathsf{I}(\mathsf{I}+1)<\mathsf{S}$ and $\mathsf{C}\leq\frac{\mathsf{S}^2}{\mathsf{I}^2}$ and we have $\mathsf{d}\leq 1-\alpha$, achieved with $\gamma=0$, as in the classical non-cooperative IC [12]. We hence use the same transmission strategy as in the classical IC, whose achievable sum-rate is given by

$$(R_{\mathsf{p}}+R_{\mathsf{c}})^{\mathrm{IB}}\leq 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)-2\log(2).$$

The tightest sum-capacity outer bound in (5.5) is

$$\begin{aligned}(R_{\mathsf{p}}+R_{\mathsf{c}})^{\mathrm{OB}} &\leq \mathrm{eq}(5.4\mathrm{f})\\ &\overset{\mathsf{C}\leq\frac{\mathsf{S}^2}{\mathsf{I}^2}}{\leq} 2\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)+3.51\log(2).\end{aligned}$$

Thus,

$$\mathsf{GAP}=\frac{(R_{\mathsf{p}}+R_{\mathsf{c}})^{\mathrm{OB}}-(R_{\mathsf{p}}+R_{\mathsf{c}})^{\mathrm{IB}}}{2}\leq 2.8\text{ bits/user}.$$

Regime 10 ($\alpha<1/2,\beta>2-2\alpha$) / Weak Interference 5: in this regime $\mathsf{I}(\mathsf{I}+1)<\mathsf{S}$ and $\mathsf{C}>\frac{\mathsf{S}^2}{\mathsf{I}^2}$ and we have

$$2\mathsf{d}\leq\max_{\gamma}\min\left\{\gamma+(1-\gamma)(2-\alpha),\gamma\beta+2(1-\gamma)(1-\alpha)\right\}.$$

In this expression the first term is decreasing in γ while the second term is increasing in γ. So the optimal γ can be found by equating the two terms and is given by $\gamma^\star = \frac{\alpha}{\beta+\alpha-1}$ in (5.7), which leads to $\mathsf{d} \leq \frac{1}{2}\frac{2\beta+2\alpha-\alpha\beta-2}{\beta+\alpha-1}$. In this regime we use the transmission strategy described in Appendix 5.E.1 and Appendix 5.E.4 that, since the interference is weak, uses both common and private messages. The sum-rate achieved with this scheme is

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})} &= \gamma' \log(1+\mathsf{S}) - \gamma' \log\left(1+\frac{\mathsf{S}}{1+\mathsf{S}}\right) + \gamma' \log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right) \\ &+ \left(1-\gamma'\right)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - \left(1-\gamma'\right)\log(1+\mathsf{I}) \\ &+ \left(1-\gamma'\right)\log\left(1+\frac{\mathsf{I}}{1+\mathsf{I}}+\mathsf{S}\right) - \left(1-\gamma'\right)\log\left(1+\frac{\mathsf{I}}{1+\mathsf{I}}\right), \quad (5.13)\end{aligned}$$

where (see Appendix 5.E.4 for the details) $\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x} \overset{\mathsf{SNR}\gg 1}{\rightarrow} \gamma^\star$ as in (5.7) for $x := \log\left(1+\frac{\mathsf{SI}}{(1+\mathsf{I})^2+\mathsf{S}}\right) \overset{\mathsf{SNR}\gg 1}{\rightarrow} \alpha$. The tightest sum-capacity outer bounds in (5.5) for this regime are (5.4e)

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} &\leq \gamma\log(1+\mathsf{S}) + (1-\gamma)\log\left(1+\frac{\mathsf{S}}{\mathsf{I}}\right) \\ &+ (1-\gamma)\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) + 3\log(2),\end{aligned}$$

which gives (note that we evaluate $(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}$ for the optimal γ in the upper bound; this is a possible suboptimal choice)

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 3.5 \text{ bits/user},$$

and (5.4f)

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} &\leq \gamma\log(1+\mathsf{I}+\mathsf{C}+\mathsf{S}) \\ &+ (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) \\ &+ (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) + 3.51\log(2),\end{aligned}$$

which gives

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 2.8 \text{ bits/user}.$$

Thus, for this regime we have $\mathsf{GAP} \leq 3.5$ bits/user.

5.C Proof of Theorem 13

Let $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$, $\mathsf{I_c} = \mathsf{I}$, $\mathsf{I_p} = 0$ and $\mathsf{d}^{\mathrm{Z}} = \mathsf{d}$ for brevity. We analyze the different regimes in Figure 5.7.

Regime 1 ($\alpha \geq 2$, $\beta \leq 1$) / Very Strong Interference 1: in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and $\mathsf{C} \leq \mathsf{S}$ and we have $\mathsf{d} \leq 1$ as in the FD case and in the classical IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative Z-IC in very strong interference [103, Theorem 2], i.e.,

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} = 2\log\left(1+\mathsf{S}\right).$$

As upper bound on the sum-capacity for this regime we use eq(4.29a)+eq(4.29c), valid for the FD case, which can be further upper bounded as

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} \leq 2\log\left(1+\mathsf{S}\right) + \log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{2} \leq 0.5 \text{ bits/user}.$$

Regime 2 ($\alpha \geq 2$, $1 < \beta \leq 2$) / Very Strong Interference 2: in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and $\mathsf{S} < \mathsf{C} \leq \mathsf{S}(1+\mathsf{S})$ and we have $\mathsf{d} \leq 1$, achieved with $\gamma = 0$, as in the classical non-cooperative Z-channel [103]. We hence use the same transmission strategy as in the non-cooperative Z-channel (see [103, Theorem 2]), whose achievable sum-rate is given by

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}} \leq 2\log\left(1+\mathsf{S}\right).$$

The tightest sum-capacity outer bound in (5.5) is

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} \leq \text{eq(5.4a)} + \text{eq(5.4c)} \overset{\mathsf{C} \leq \mathsf{S}(1+\mathsf{S})}{\leq} 2\log\left(1+\mathsf{S}\right) + 2.52\log(2).$$

Thus,

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}}{2} \leq 1.26 \text{ bits/user}.$$

Regime 3 ($\min\{\alpha,\beta\} \geq 2$) / Very Strong Interference 3: in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and $\mathsf{C} > \mathsf{S}(1+\mathsf{S})$ and we have

$$2\mathsf{d} \leq \max_{\gamma} \min\left\{\gamma\beta + 2(1-\gamma), \gamma + (1-\gamma)\alpha\right\}.$$

In this expression the first term is increasing in γ while the second one is decreasing in γ. Thus the optimal γ can be found by equating the two terms and is given by $\gamma^\star = \frac{\alpha-2}{\beta+\alpha-3}$ as in (5.8), which leads to

$$\mathsf{d} \leq \frac{1}{2}\frac{\beta\alpha-2}{\beta+\alpha-3}.$$

In this regime we use the transmission strategy described in Appendix 5.E.1 and Appendix 5.E.2 that, since the interference is very strong, uses common messages only. Recall that, since $\mathsf{I}_\mathsf{p} = 0$, the signal $X_\mathsf{p}[2]$ is not received at the CRx. The sum-rate achieved with this scheme is

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})} &= \gamma' \log(1+\mathsf{S}) + \gamma' \log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right) - \log\left(1+\frac{\mathsf{S}}{1+\mathsf{S}}\right) \\ &\quad + 2\left(1-\gamma'\right)\log(1+\mathsf{S}),\end{aligned} \tag{5.14}$$

where (see Appendix 5.E.2 for the details) $\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x} \overset{\mathsf{SNR}\gg 1}{\to} \gamma^\star$ as in (5.8) for $x := \log\left(1+\frac{\mathsf{I}}{(1+\mathsf{S})^2}\right) \overset{\mathsf{SNR}\gg 1}{\to} \alpha-2$ (recall that $\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right) \overset{\mathsf{SNR}\gg 1}{\to} \beta-1$). The tightest sum-capacity outer bounds in (5.5) for this regime are eq(5.4a) + eq(5.4c)

$$(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} \leq \gamma\log(1+\mathsf{C}+\mathsf{S}) + 2(1-\gamma)\log(1+\mathsf{S}) + 2.52\log(2),$$

which gives (note that we evaluate $(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}$ for the optimal γ in the upper bound; this is a possible suboptimal choice)

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 1.76 \text{ bits/user},$$

and (5.4e)

$$\begin{aligned}(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} &\leq \gamma\log(1+\mathsf{S}) + (1-\gamma)\log\left(1+\frac{\mathsf{S}}{\mathsf{I}}\right) \\ &\quad + (1-\gamma)\log\left(1+\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)^2\right) + 3\log(2),\end{aligned}$$

which gives

$$\mathsf{GAP} = \frac{(R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p}+R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 3 \text{ bits/user}.$$

Thus, for this regime we have $\mathsf{GAP} \leq 3$ bits/user.

Regime 4 ($1 \leq \alpha < 2$) / Strong Interference: in this regime $\mathsf{S} \leq \mathsf{I} < \mathsf{S}(1+\mathsf{S})$ and we have $\mathsf{d} \leq \frac{\alpha}{2}$ as in the FD case and in the classical Z-IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative Z-IC in strong interference [103, Theorem 2], i.e.,

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} = \log\left(1 + \mathsf{S} + \mathsf{I}\right).$$

As upper bound on the sum-capacity for this regime we use (4.29d), valid for the FD case, which can be further upper bounded as

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} \leq \log\left(1 + \mathsf{S} + \mathsf{I}\right) + 2\log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{2} \leq 1 \text{ bit/user}.$$

Regime 5 ($\alpha < 1$) / Weak Interference: in this regime $\mathsf{I} < \mathsf{S}$ and we have $\mathsf{d} \leq 1 - \frac{\alpha}{2}$ as in the FD case and in the classical Z-IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative Z-IC in weak interference [103, Theorem 2], i.e.,

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} = \log\left(1 + \mathsf{S}\right) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right).$$

As upper bound on the sum-capacity for this regime we use (4.29d), valid for the FD case, which can be further upper bounded as

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} \leq \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \log\left(1 + \mathsf{S}\right) + 2\log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{2} \leq 1 \text{ bit/user}.$$

5.D Proof of Theorem 14

Let $\mathsf{S}_{\mathsf{p}} = \mathsf{S}_{\mathsf{c}} = \mathsf{S}$, $\mathsf{I}_{\mathsf{p}} = \mathsf{I}$, $\mathsf{I}_{\mathsf{c}} = 0$ and $\mathsf{d}^{\mathrm{S}} = \mathsf{d}$ for brevity. We analyze the different regimes in Figure 5.8.

Regime 1 ($\alpha \geq 2$) / Very Strong Interference: in this regime $\mathsf{I} \geq \mathsf{S}(1+\mathsf{S})$ and we have $\mathsf{d} \leq 1$ as in the FD case and in the classical S-IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative S-IC in very strong interference [103, Theorem 2], i.e.,

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} = 2\log\left(1+\mathsf{S}\right).$$

As upper bound on the sum-capacity for this regime we use eq(4.36a)+(4.36b), valid for the FD case, i.e.

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} \leq 2\log\left(1+\mathsf{S}\right).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{2} \leq 0 \text{ bit/user},$$

i.e., for this regime the sum-capacity is known exactly.

Regime 2 ($1 \leq \alpha < 2$, $\beta \leq \alpha$) / Strong Interference 1: in this regime $\mathsf{S} \leq \mathsf{I} < \mathsf{S}(1+\mathsf{S})$ and $\mathsf{C} \leq \mathsf{I}$ and we have $\mathsf{d} \leq \frac{\alpha}{2}$ as in the FD case and in the classical non-cooperative S-IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative S-IC in strong interference [103, Theorem 2], i.e.,

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} = \log\left(1+\mathsf{S}+\mathsf{I}\right).$$

As upper bound on the sum-capacity for this regime we use (4.36c), valid for the FD case, which can be further upper bounded as

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} \leq \log\left(1+\mathsf{S}+\mathsf{I}\right) + \log(2) + \log(3).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})}}{2} \leq 1.3 \text{ bits/user}.$$

Regime 3 ($1 \leq \alpha < 2$, $\beta > \alpha$) / Strong Interference 2: in this regime $\mathsf{S} \leq \mathsf{I} < \mathsf{S}(1+\mathsf{S})$ and $\mathsf{C} > \mathsf{I}$ and we have

$$2\mathsf{d} \leq \max_{\gamma} \min\left\{\gamma + 2\left(1-\gamma\right), \gamma\beta + \left(1-\gamma\right)\alpha\right\}.$$

In this expression the first term is decreasing in γ while the second term is increasing in γ. Thus the optimal γ can be found by equating the two terms and is given by $\gamma^\star = \frac{2-\alpha}{\beta-\alpha+1}$, which leads to $\mathsf{d} \leq \frac{1}{2}\frac{2\beta-\alpha}{\beta-\alpha+1}$. In this regime we use the transmission strategy described in Appendix 5.E.1 and Appendix 5.E.3, but with different power splits. In particular

- In the transmitted signal $X_{\mathsf{p}}[2]$ we set $\delta_3 = 0$, i.e., the non-cooperative message of the PTx is common since we are in strong cooperation. We further choose $\delta_2 = 1 - \delta_1 = \frac{\mathsf{S}}{1+\mathsf{I}}$.
- In the transmitted signal $X_{\mathsf{c}}[2]$ we set $\delta_4 = 1$, i.e., private message only for the CTx; this is because since $\mathsf{I}_{\mathsf{c}} = 0$, there is no interference at the PRx.

With these choices we obtain that Phase I is successful if

$$R_{b1} \leq \text{eq(5.25)}, \tag{5.15}$$
$$R_{b2} \leq \text{eq(5.26)}, \tag{5.16}$$

while Phase II is successful if

$$R_{b2} \leq (1-\gamma)\log\left(1 + \frac{\mathsf{S}^2}{1+\mathsf{I}}\right), \tag{5.17}$$
$$R_{b3c} \leq (1-\gamma)\log\left(1+\mathsf{S}+\mathsf{I}\right) - (1-\gamma)\log\left(1 + \mathsf{S}\frac{\mathsf{I}}{1+\mathsf{I}} + \mathsf{S}\right), \tag{5.18}$$
$$R_{b4p} \leq (1-\gamma)\log\left(1+\mathsf{S}\right). \tag{5.19}$$

By imposing that R_{b_2} is the same in both phases, i.e., that (5.26) and (5.17) are equal, we get that γ should be chosen equal to

$$\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right) + x} \overset{\mathsf{SNR}\gg 1}{\to} \gamma^\star \quad x := \log\left(1+\frac{\mathsf{S}^2}{1+\mathsf{I}}\right) \overset{\mathsf{SNR}\gg 1}{\to} 2-\alpha.$$

Therefore the total sum-rate decoded at the PRx and at the CRx through the two phases is

$$\begin{aligned}
(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} &= R_{b_1} + R_{b_2} + R_{b_{3c}} + R_{b_{4p}} \\
&= \gamma'\log\left(1+\mathsf{S}\right) - \gamma'\log\left(1+\frac{\mathsf{S}}{1+\mathsf{S}}\right) + \gamma'\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right) \\
&+ \left(1-\gamma'\right)\log\left(1+\mathsf{S}+\mathsf{I}\right) - \left(1-\gamma'\right)\log\left(1+\mathsf{S}\frac{\mathsf{I}}{1+\mathsf{I}}+\mathsf{S}\right) \\
&+ (1-\gamma)\log\left(1+\mathsf{S}\right).
\end{aligned}$$

The tightest sum-capacity upper bounds for this regime are (5.4d) and (5.4e) (evaluated for $\mathsf{I_c} = 0$), which can be respectively further upper bounded as

$$\begin{aligned}(R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{OB})} &\leq \gamma \log(1+\mathsf{C}) + (1-\gamma)\log(1+\mathsf{S}+\mathsf{I}) \\ &\quad + 2\log(2) + \gamma\log(3) + 2(1-\gamma)\log(2), \\ (R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{OB})} &\leq \gamma\log(1+\mathsf{S}) + 2(1-\gamma)\log(1+\mathsf{S}) + 3\log(2).\end{aligned}$$

Both outer bounds lead to

$$\mathsf{GAP} = \frac{(R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 2.5 \text{ bits/user}.$$

Regime 4 ($\alpha < 1$, $\beta \leq 1$) / Weak Interference 1: in this regime $\mathsf{I} < \mathsf{S}$ and $\mathsf{C} \leq \mathsf{S}$ and we have $\mathsf{d} \leq 1 - \frac{\alpha}{2}$ as in the FD case and in the classical non-cooperative S-IC. We use as lower bound on the sum-capacity the sum-rate achieved by the non-cooperative S-IC in weak interference [103, Theorem 2], i.e.,

$$(R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{IB})} = \log(1+\mathsf{S}) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right).$$

As upper bound on the sum-capacity for this regime we use (4.36c), valid for the FD case, which can be further upper bounded as

$$(R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{OB})} \leq \log(1+\mathsf{S}+\mathsf{I}) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + 2\log(2).$$

Thus, for this regime we have

$$\mathsf{GAP} = \frac{(R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{OB})} - (R_\mathsf{p} + R_\mathsf{c})^{(\mathrm{IB})}}{2} \leq 1.5 \text{ bits/user},$$

Regime 5 ($\alpha < 1$, $1 < \beta \leq 2 - \alpha$) / Weak Interference 2: in this regime $\mathsf{I} < \mathsf{S}$ and $\mathsf{S} < \mathsf{C} \leq \frac{\mathsf{S}^2}{1+\mathsf{I}}$ and we have $\mathsf{d} \leq 1 - \frac{\alpha}{2}$, achieved with $\gamma = 0$, as in the classical non-cooperative S-channel [103]. We hence use the same transmission strategy as in the non-cooperative S-channel (see [103, Theorem 2]), whose achievable sum-rate is given by

$$(R_\mathsf{p} + R_\mathsf{c})^{\mathrm{IB}} \leq \log(1+\mathsf{S}) + \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right).$$

The tightest sum-capacity outer bound in (5.5) is (5.4d), which can be further upper bounded as

$$(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} \leq \log\left(1 + \frac{\mathsf{S}}{\mathsf{I}}\right) + \log\left(1 + \mathsf{S} + \mathsf{I}\right) + 3\log(2).$$

Thus,

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}}{2} \leq 2.5 \text{ bits/user}.$$

Regime 6 ($\alpha < 1$, $\beta > 2 - \alpha$) / Weak Interference 3: in this regime $\mathsf{I} < \mathsf{S}$ and $\mathsf{C} > \frac{\mathsf{S}^2}{1+\mathsf{I}}$ and we have

$$2\mathsf{d} \leq \max_{\gamma} \min\left\{\gamma + 2\left(1-\gamma\right), \gamma\beta + \left(1-\gamma\right)\left(2-\alpha\right)\right\}.$$

In this expression the first term is decreasing in γ while the second term is increasing in γ. Thus the optimal γ can be found by equating the two terms and is given by $\gamma^{\star} = \frac{\alpha}{\beta+\alpha-1}$, which leads to $\mathsf{d} \leq \frac{1}{2}\frac{2\beta+\alpha-2}{\beta+\alpha-1}$. In this regime we use the transmission strategy described in Appendix 5.E.1 and Appendix 5.E.4, adapted to the S-channel. We obtain that Phase I is successful if

$$R_{b1} \leq \mathrm{eq}(5.25), \tag{5.20}$$
$$R_{b2} \leq \mathrm{eq}(5.26), \tag{5.21}$$

and that Phase II is successful if

$$R_{b2} \leq (1-\gamma)\log\left(1+\mathsf{S}\right) - (1-\gamma)\log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right), \tag{5.22}$$
$$R_{b3p} \leq (1-\gamma)\log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right), \tag{5.23}$$
$$R_{b4p} \leq \mathrm{eq}(5.41). \tag{5.24}$$

By imposing that R_{b_2} is the same in both phases, i.e., that (5.26) and (5.22) are equal, we get that γ should be chosen equal to

$$\gamma' = \frac{x}{\log\left(1 + \frac{\mathsf{C}}{1+\mathsf{S}}\right) + x} \overset{\mathsf{SNR}\gg 1}{\rightarrow} \gamma^{\star}$$
$$x := \log\left(1 + \frac{\mathsf{SI}}{1+\mathsf{S}+\mathsf{I}}\right) \overset{\mathsf{SNR}\gg 1}{\rightarrow} \min\{1, \alpha\} = \alpha.$$

Therefore the total sum-rate decoded at the PRx and at the CRx through the two phases is

$$\begin{aligned}(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{IB})} &= R_{b_1} + R_{b_2} + R_{b_{3p}} + R_{b_{4p}} \\ &= \gamma' \log\left(1+\mathsf{S}\right) - \gamma' \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{S}}\right) + \gamma' \log\left(1 + \frac{\mathsf{C}}{1+\mathsf{S}}\right) \\ &+ \left(1-\gamma'\right) \log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) + \left(1-\gamma'\right) \log\left(1 + \mathsf{S} + \frac{\mathsf{I}}{1+\mathsf{I}}\right) \\ &- \left(1-\gamma'\right) \log\left(1 + \frac{\mathsf{I}}{1+\mathsf{I}}\right).\end{aligned}$$

The tightest sum-capacity upper bounds for this regime are (5.4d) and (5.4e) (evaluated for $\mathsf{I}_{\mathsf{c}} = 0$), which can be respectively further upper bounded as

$$\begin{aligned}(R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} &\leq \gamma \log\left(1+\mathsf{C}\right) + (1-\gamma)\log\left(1+\mathsf{S}+\mathsf{I}\right) + (1-\gamma)\log\left(1+\frac{\mathsf{S}}{\mathsf{I}}\right) \\ &\quad + 2\log(2) + \gamma\log(3) + (1-\gamma)\log(2), \\ (R_{\mathsf{p}} + R_{\mathsf{c}})^{(\mathrm{OB})} &\leq \gamma \log\left(1+\mathsf{S}\right) + 2\left(1-\gamma\right)\log\left(1+\mathsf{S}\right) + 3\log(2).\end{aligned}$$

With the first outer bound we obtain

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}}{2} \leq 3 \text{ bits/user},$$

while with the second one we have

$$\mathsf{GAP} = \frac{(R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{OB}} - (R_{\mathsf{p}} + R_{\mathsf{c}})^{\mathrm{IB}}}{2} \leq 2 \text{ bits/user}.$$

Thus, for this regime we have $\mathsf{GAP} \leq 3$ bits/user.

5.E Transmission strategies

Here we develop achievable schemes inspired by the LDA-based transmission strategies. In the following all signals X_{b_j} for some subscript j, are independent proper-complex Gaussian random variables with zero mean and unit variance and represent codebooks used to convey the bits in b_j.

5.E.1 Phase I of duration $\gamma \in [0, 1]$ (see also Figure 5.2)

The transmitted signals are

$$\begin{aligned}X_{\mathsf{p}}[1] &= \sqrt{1-\eta}X_{b_1} + \sqrt{\eta}X_{b_2}, \\ X_{\mathsf{c}}[1] &= 0.\end{aligned}$$

Thus, the received signals at the CTx and at the PRx are

$$T_{\mathsf{f}}[1] = \sqrt{\mathsf{C}}\left(\sqrt{1-\eta}X_{b_1} + \sqrt{\eta}X_{b_2}\right) + Z_{\mathsf{f}}[1],$$
$$Y_{\mathsf{p}}[1] = \sqrt{\mathsf{S}}\left(\sqrt{1-\eta}X_{b_1} + \sqrt{\eta}X_{b_2}\right) + Z_{\mathsf{p}}[1].$$

The CTx applies successive decoding of X_{b_1} followed by X_{b_2} from $T_{\mathsf{f}}[1]$ which is possible if

$$R_{b_1} \leq \gamma \log\left(1+\mathsf{C}\right) - \gamma\log\left(1+\mathsf{C}\eta\right),$$
$$R_{b_2} \leq \gamma \log\left(1+\mathsf{C}\eta\right).$$

The PRx decodes X_{b_1} treating X_{b_2} as noise from $Y_{\mathsf{p}}[1]$, which is possible if

$$R_{b_1} \leq \gamma \log\left(1+\mathsf{S}\right) - \gamma\log\left(1+\mathsf{S}\eta\right).$$

Since $\mathsf{C} \geq \mathsf{S}$ and motivated by the observation in [12] that all the terms that appear as noise should be at most at the level of the noise, we choose $\eta = \frac{1}{1+\mathsf{S}}$. With this we have that Phase I is successful if

$$R_{b_1} \leq \gamma \log\left(1+\mathsf{S}\right) - \gamma\log\left(1+\frac{\mathsf{S}}{1+\mathsf{S}}\right), \tag{5.25}$$
$$R_{b_2} \leq \gamma \log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right). \tag{5.26}$$

5.E.2 Phase II of duration $(1-\gamma)$ for Region 3 in Figure 5.1 (see also Figure 5.5)

The transmitted signals are

$$X_{\mathsf{p}}[2] = X_{b_{3c}},$$
$$X_{\mathsf{c}}[2] = \sqrt{\eta}X_{b_2} + \sqrt{1-\eta}X_{b_{4c}},$$

where we choose $\eta = \frac{1}{1+\mathsf{S}}$. Thus, the received signal at the PRx and at the CRx are

$$Y_{\mathsf{p}}[2] = \sqrt{\mathsf{S}}X_{b_{3c}} + \sqrt{\mathsf{I}}e^{j\theta_{\mathsf{c}}}\left(\sqrt{\frac{1}{1+\mathsf{S}}}X_{b_2} + \sqrt{\frac{\mathsf{S}}{1+\mathsf{S}}}X_{b_{4c}}\right) + Z_{\mathsf{p}}[2],$$
$$Y_{\mathsf{c}}[2] = \sqrt{\mathsf{I}}e^{j\theta_{\mathsf{p}}}X_{b_{3c}} + \sqrt{\mathsf{S}}\left(\sqrt{\frac{1}{1+\mathsf{S}}}X_{b_2} + \sqrt{\frac{\mathsf{S}}{1+\mathsf{S}}}X_{b_{4c}}\right) + Z_{\mathsf{c}}[2].$$

The PRx applies successive decoding from $Y_{\mathsf{p}}[2]$ as follows: $X_{b_{4c}}$, X_{b_2}, $X_{b_{3c}}$, which is possible if

$$R_{b_{4c}} \le (1-\gamma)\log\left(1+\mathsf{S}+\mathsf{I}\right) - (1-\gamma)\log\left(1+\mathsf{S}+\frac{\mathsf{I}}{1+\mathsf{S}}\right), \tag{5.27}$$

$$R_{b_2} \le (1-\gamma)\log\left(1+\mathsf{S}+\frac{\mathsf{I}}{1+\mathsf{S}}\right) - (1-\gamma)\log\left(1+\mathsf{S}\right), \tag{5.28}$$

$$R_{b_{3c}} \le (1-\gamma)\log\left(1+\mathsf{S}\right). \tag{5.29}$$

The CRx successively decodes $X_{b_{3c}}$ and $X_{b_{4c}}$ (treating X_{b_2} as noise) from $Y_{\mathsf{c}}[2]$, which is possible if

$$R_{b_{3c}} \le (1-\gamma)\log\left(1+\mathsf{I}+\mathsf{S}\right) - (1-\gamma)\log\left(1+\mathsf{S}\right), \tag{5.30}$$

$$R_{b_{4c}} \le (1-\gamma)\log\left(1+\mathsf{S}\right) - (1-\gamma)\log\left(1+\frac{\mathsf{S}}{1+\mathsf{S}}\right). \tag{5.31}$$

Thus, Phase II is successful if

$$\begin{aligned} R_{b_{3c}} &\le \min\left\{\text{eq.}(5.29), \text{eq.}(5.30)\right\} \overset{\mathsf{I}\ge\mathsf{S}(1+\mathsf{S})}{=} \text{eq.}(5.29), \\ R_{b_{4c}} &\le \min\left\{\text{eq.}(5.27), \text{eq.}(5.31)\right\} \overset{\mathsf{I}\ge\mathsf{S}(1+\mathsf{S})}{=} \text{eq.}(5.31), \\ R_{b_2} &\le \text{eq}(5.28). \end{aligned}$$

By imposing that R_{b_2} is the same in both phases, i.e., that (5.26) and (5.28) are equal, we get that γ should be chosen equal to

$$\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x}, \quad x := \log\left(1+\frac{\mathsf{I}}{(1+\mathsf{S})^2}\right).$$

Therefore the total sum-rate decoded at the PRx and at the CRx through the two phases is

$$(R_{\mathsf{p}}+R_{\mathsf{c}})^{(\mathrm{IB})} = R_{b_1}+R_{b_2}+R_{b_{3c}}+R_{b_{4c}},$$

as given in (5.11).

5.E.3 Phase II of duration $(1-\gamma)$ for Region 8 in Figure 5.1 (see also Figure 5.3)

The transmitted signals are

$$\begin{aligned} X_{\mathsf{p}}[2] &= \sqrt{\delta_1}X_{b_{3c}} + \sqrt{\delta_2}X_{b_2} + \sqrt{\delta_3}X_{b_{3p}}, \\ X_{\mathsf{c}}[2] &= \sqrt{\delta_4}X_{b_{4p}} + \sqrt{1-\delta_4}X_{b_{4c}}, \end{aligned}$$

where we choose $\delta_1 = 1 - \delta_2 - \delta_3$, $\quad \delta_2 = \frac{\mathsf{S}}{(1+\mathsf{I})^2}$, $\quad \delta_3 = \delta_4 = \frac{1}{1+\mathsf{I}}$, and where $X_{b_{4p}}$ is DPC-ed against X_{b_2} at $Y_{\mathsf{c}}[2]$. Thus, the received signal at the PRx and at the CRx are

$$\begin{aligned} Y_{\mathsf{p}}[2] &= \sqrt{\mathsf{S}}\left(\sqrt{\delta_1}X_{b_{3c}} + \sqrt{\delta_2}X_{b_2} + \sqrt{\delta_3}X_{b_{3p}}\right) \\ &\quad + \sqrt{\mathsf{I}}e^{\mathsf{j}\theta_{\mathsf{c}}}\left(\sqrt{\delta_4}X_{b_{4p}} + \sqrt{1-\delta_4}X_{b_{4c}}\right) + Z_{\mathsf{p}}[2], \\ Y_{\mathsf{c}}[2] &= \sqrt{\mathsf{I}}e^{\mathsf{j}\theta_{\mathsf{p}}}\left(\sqrt{\delta_1}X_{b_{3c}} + \sqrt{\delta_2}X_{b_2} + \sqrt{\delta_3}X_{b_{3p}}\right) \\ &\quad + \sqrt{\mathsf{S}}\left(\sqrt{\delta_4}X_{b_{4p}} + \sqrt{1-\delta_4}X_{b_{4c}}\right) + Z_{\mathsf{c}}[2]. \end{aligned}$$

The PRx applies successive decoding from $Y_{\mathsf{p}}[2]$ as follows: $X_{b_{3c}}$, X_{b_2}, $X_{b_{4c}}$ and $X_{b_{3p}}$ (by treating $X_{b_{4p}}$ as noise), which is possible if

$$R_{b_{3c}} \le (1-\gamma)\log(1+\mathsf{S}+\mathsf{I}) - (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}^2+\mathsf{S}+\mathsf{SI}}{(1+\mathsf{I})^2}\right), \quad (5.32)$$

$$R_{b_2} \le (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}^2+\mathsf{S}+\mathsf{SI}}{(1+\mathsf{I})^2}\right) - (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right), \quad (5.33)$$

$$R_{b_{4c}} \le (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) - (1-\gamma)\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}+\frac{\mathsf{I}}{1+\mathsf{I}}\right), \quad (5.34)$$

$$R_{b_{3p}} \le (1-\gamma)\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}+\frac{\mathsf{I}}{1+\mathsf{I}}\right) - (1-\gamma)\log\left(1+\frac{\mathsf{I}}{1+\mathsf{I}}\right). \quad (5.35)$$

The CRx applies successive decoding from $Y_{\mathsf{c}}[2]$ as follows: $X_{b_{4c}}$, $X_{b_{3c}}$ and $X_{b_{4p}}$, which is possible if

$$R_{b_{4c}} \le (1-\gamma)\log(1+\mathsf{S}+\mathsf{I}) - (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right), \quad (5.36)$$

$$\begin{aligned} R_{b_{3c}} &\le (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right) \\ &\quad - (1-\gamma)\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}+\frac{\mathsf{IS}+\mathsf{I}+\mathsf{I}^2}{(1+\mathsf{I})^2}\right), \end{aligned} \quad (5.37)$$

$$R_{b_{4p}} \le (1-\gamma)\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}+\frac{\mathsf{I}}{1+\mathsf{I}}\right) - (1-\gamma)\log\left(1+\frac{\mathsf{I}}{1+\mathsf{I}}\right). \quad (5.38)$$

Thus, Phase II is successful if

$$\begin{aligned}
R_{b_{3c}} &\leq \min\{\text{eq.}(5.32), \text{eq.}(5.37)\} \overset{\mathsf{S}(\mathsf{S}+1)\geq\mathsf{I}(\mathsf{I}+1)^2}{=} \text{eq.}(5.37),\\
R_{b_{4c}} &\leq \min\{\text{eq.}(5.34), \text{eq.}(5.36)\} \overset{\mathsf{S}(\mathsf{S}+1)\geq\mathsf{I}(\mathsf{I}+1)^2}{=} \text{eq.}(5.34),\\
R_{b_2} &\leq \text{eq}(5.33),\\
R_{b_{3p}} &\leq \text{eq}(5.35),\\
R_{b_{4p}} &\leq \text{eq}(5.38).
\end{aligned}$$

By imposing that R_{b_2} is the same in both phases, i.e., that (5.26) and (5.33) are equal, we get that γ should be chosen equal to

$$\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x}, \quad x := \log\left(1+\frac{\mathsf{S}^2}{(1+\mathsf{I})^3+\mathsf{S}+\mathsf{SI}}\right).$$

Therefore the total sum-rate decoded at the PRx and at the CRx through the two phases is

$$(R_{\mathsf{p}}+R_{\mathsf{c}})^{(\mathrm{IB})} = R_{b_1}+R_{b_2}+R_{b_{3c}}+R_{b_{3p}}+R_{b_{4c}}+R_{b_{4p}},$$

as given in (5.12).

5.E.4 Phase II of duration $(1-\gamma)$ for Region 10 in Figure 5.1 (see also Figure 5.4)

The transmitted signals are

$$\begin{aligned}
X_{\mathsf{p}}[2] &= \sqrt{1-\delta}X_{b_2}+\sqrt{\delta}X_{b_{3p}},\\
X_{\mathsf{c}}[2] &= X_{b_{4p}},
\end{aligned}$$

where we choose $\delta = \frac{1}{1+\mathsf{I}}$, and where $X_{b_{4p}}$ is DPC-ed against X_{b_2} at $Y_{\mathsf{c}}[2]$. Thus, the received signal at the PRx and at the CRx are

$$\begin{aligned}
Y_{\mathsf{p}}[2] &= \sqrt{\mathsf{S}}\left(\sqrt{1-\delta}X_{b_2}+\sqrt{\delta}X_{b_{3p}}\right)+\sqrt{\mathsf{I}}e^{j\theta_{\mathsf{c}}}X_{b_{4p}}+Z_{\mathsf{p}}[2],\\
Y_{\mathsf{c}}[2] &= \sqrt{\mathsf{I}}e^{j\theta_{\mathsf{p}}}\left(\sqrt{1-\delta}X_{b_2}+\sqrt{\delta}X_{b_{3p}}\right)+\sqrt{\mathsf{S}}X_{b_{4p}}+Z_{\mathsf{c}}[2].
\end{aligned}$$

The PRx applies successive decoding from $Y_{\mathsf{p}}[2]$ as follows: X_{b_2} and $X_{b_{3p}}$ (by treating $X_{b_{4p}}$ as noise), that is possible if

$$R_{b_2} \leq (1-\gamma)\log(1+\mathsf{S}+\mathsf{I})-(1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right), \tag{5.39}$$

$$R_{b_{3p}} \leq (1-\gamma)\log\left(1+\mathsf{I}+\frac{\mathsf{S}}{1+\mathsf{I}}\right)-(1-\gamma)\log(1+\mathsf{I}). \tag{5.40}$$

The CRx decodes $X_{b_{4p}}$ (by treating $X_{b_{3p}}$ as noise) from $Y_{\mathsf{c}}[2]$, which is possible if

$$R_{b_{4p}} \le (1-\gamma)\log\left(1+\mathsf{S}+\frac{\mathsf{I}}{1+\mathsf{I}}\right) - (1-\gamma)\log\left(1+\frac{\mathsf{I}}{1+\mathsf{I}}\right). \tag{5.41}$$

Thus Phase 2 is successful if

$$\begin{aligned} R_{b_2} &\le \text{eq}(5.39), \\ R_{b_{3p}} &\le \text{eq}(5.40), \\ R_{b_{4p}} &\le \text{eq}(5.41). \end{aligned}$$

By imposing that R_{b_2} is the same in both phases, i.e., that (5.26) and (5.39) are equal, we get that γ should be chosen equal to

$$\gamma' = \frac{x}{\log\left(1+\frac{\mathsf{C}}{1+\mathsf{S}}\right)+x}, \quad x := \log\left(1+\frac{\mathsf{S}\mathsf{I}}{(1+\mathsf{I})^2+\mathsf{S}}\right).$$

Therefore the total sum-rate decoded at the PRx and at the CRx through the two phases is

$$(R_{\mathsf{p}}+R_{\mathsf{c}})^{(\mathrm{IB})} = R_{b_1}+R_{b_2}+R_{b_{3p}}+R_{b_{4p}},$$

as given in (5.13).

Chapter 6

Conclusions

In this thesis we conducted an information theoretic study on two practically relevant cooperative wireless networks, namely the HD relay network and the CCIC, or the IC with unilateral source cooperation.

The first part of the thesis was dedicated to the study of the HD relay network, where the communication between a source and a destination is assisted by N relay stations operating in HD mode. In particular, in **Chapter 2**, we analyzed the case $N = 1$, i.e., the classical relay channel. For this system we first determined the exact capacity of the LDA channel, by showing that random switch and correlated non-uniform input bits at the relay are optimal. We then showed that, for the Gaussian noise case, the cut-set outer bound is achievable to within a constant gap by PDF and CF, evaluated both with deterministic and random switch. This constant gap result implies the exact knowledge of the gDoF, which was derived in closed form. We finally designed an 'optimal to within a constant gap' scheme inspired by the LDA of the Gaussian noise channel at high SNR, which is based on superposition encoding, PDF relaying and successive decoding. Publications related to this chapter are [65–67]. In **Chapter 3**, we analyzed the general case of N relays. We first showed that, for the Gaussian noise case, the cut-set outer bound is achievable to within a constant gap (which only depends on N) by NNC; we also extended this constant gap result to the case of multi-antenna nodes, by showing that the gap only depends on the total number of antennas in the system. We then proved that, for any memoryless HD N-relay network with independent noises and for which the cut-set outer bound is achievable to within a constant gap under certain assumptions, the (approximately) optimal schedule has at most $N+1$ states, out of the 2^N possible ones, with a strictly positive probability; the Gaussian noise network with single-antenna nodes is a practically relevant model where this result holds; interestingly, it was shown that this result holds for the case of Gaussian noise networks with multi-antenna nodes as well where the antennas at the relays are switched between transmit and receive modes independently of one another; in other words, the (approximately) optimal schedule has at most $N+1$ active states, independently of the total number of antennas in the system. We also proved that the gDoF of the Gaussian noise network is the solution of a LP, where the coefficients of the linear inequality constraints are the solution of several LPs referred to as the assignment / MWBM problem; beyond its application to Gaussian relay networks, this technique was also showed to be useful to derive the gDoF of Gaussian broadcast networks and to solve user scheduling problems. Finally, through two simple network examples we highlighted under which channel conditions a best-relay selection scheme is strictly suboptimal in terms of gDoF and we showed that independently switching the antennas at the relays not only achieves in general strictly higher rates compared to using the antennas for the same purpose, but can actually provide a strictly larger multiplexing gain. Publications related to this chapter are [71–76].

In the second part of the thesis we studied the two-user CCIC, an IC where one capable source, i.e., the cognitive source CTx, cooperates with / assists the other source, i.e., the primary source PTx, to convey information. In contrast to the original overlay cognitive paradigm, where the CTx a priori knows the message of the PTx, in the CCIC the CTx causally learns the primary's data through a noisy in-band link. In particular, in **Chapter 4**, we assumed a FD mode of operation at the cognitive source, i.e., the CTx receives and transmits over the same time-frequency-space resources. For this system, we first derived two novel outer bounds of the type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ and $R_{\mathsf{p}}+2R_{\mathsf{c}}$ on the capacity region of the class of ISD CCICs where the noises at the different source-destination pairs are independent. We then derived an achievable rate region based on Gelfand-Pinsker binning, superposition coding, PDF relaying at the CTx and simultaneous decoding at the receivers. We specialized the outer and lower bounds on the capacity to the practically relevant Gaussian noise case and we proved that these bounds are a constant number of bits apart from one another for the symmetric case (i.e., the two direct links and the two interfering links are of the same strength) and for two asymmetric scenarios, namely the Z-channel (i.e., the link between the

PTx and the CRx is absent) and the S-channel (i.e., the link between the CTx and the PRx is absent). We showed that the capacity regions of the Z-channel and S-channel are only characterized by constraints on the single rates and on the sum-rate, while the one of the symmetric channel has also bounds of the type $2R_{\mathsf{p}} + R_{\mathsf{c}}$ and $R_{\mathsf{p}} + 2R_{\mathsf{c}}$ which are active in weak interference when the cooperation link is weaker than the direct link, i.e., in this regime unilateral cooperation is too weak to allow for a full utilization of the channel resources. We finally identified the set of parameters where causal cooperation achieves the same gDoF of the non-cooperative IC, i.e., regimes where cooperation might not be worth implementing, and of the ideal non-causal CIC, i.e., regimes where the performance is not worsened by allowing causal learning at the cognitive source. Publications related to this chapter are [77–83]. In **Chapter 5**, we constrained the cognitive source to operate in HD mode, i.e., at each time instant the CTx can either listen or transmit but not both, and we conducted an analysis similar to the FD case. In particular, our main contribution was the characterization to within a constant gap of the sum-capacity for the symmetric Z-, S- and fully-connected channels. This was accomplished by adapting to the HD case (by properly accounting for random switch) the sum-capacity outer bounds of Chapter 4 and by designing novel transmission strategies based on the LDA of the Gaussian noise channel at high SNR. Similarly to the FD case in Chapter 4, also for the HD case we highlighted the parameter regimes where the gDoF equals those of the non-cooperative IC and of the non-causal CIC. We finally identified the regimes where no losses (in terms of gDoF) incur by assuming HD mode of operation at the CTx with respect of employing a CTx with FD capabilities. Publications related to this chapter are [84, 85].

Chapter 7

Résumé [Français]

7.1 Introduction

La prochaine majeure innovation des réseaux cellulaires de quatrième génération consistera en un déploiement massif de l'infrastructure radio, c'est-à-dire, des stations de base et des stations relais. Ces nœuds seront déployés sous différentes formes, caractérisés principalement par leur largeur de bande disponible et le nombre de canaux de fréquence simultanées sur lesquels ils peuvent opérer au même temps (agrégation du spectre), le type de liens vers le réseau principal de l'opérateur (par exemple, sans-fil, par câble haut débit/faible délai), leur capacité de collaborer avec d'autres nœuds similaires et leur zone de couverture et tolérance à l'interférence. La coopération au niveau de la couche physique est considérée comme un ingrédient clé des réseaux cellulaires du futur. La transmission sans-fil permet ainsi au même signal émis d'être entendu par de multiples noeuds, permettant que les nœuds s'aident à retransmettre leur message vers la destination. La coopération promet aussi d'offrir des solutions intelligentes afin de faire face et de gérer l'interférence, garantir une qualité de service uniforme pour l'utilisateur mobile à l'intérieur de la cellule et permettre une utilisation distribuée et agressive du spectre. Tous ces facteurs sont d'une extrême importance et il devient donc critique de comprendre comment correctement concevoir de tels réseaux coopératifs.

À partir du travail de référence de Shannon "Théorie mathématique des communications", la théorie de l'information a joué un rôle essentiel dans l'évolution de systèmes de communication sans fil. Le cœur de la théorie de l'information pour les réseaux sans fil est de fournir des aperçus fondamentaux pour plusieurs problèmes clés (comme l'interférence), tout en déterminant les performances limites de ces systèmes. Ceci motive depuis des nombreuses années des chercheurs du domaine à concevoir des techniques et des stratégies de transmission s'approchant aussi près que possible de ces limites.

Dans cette thèse, nous conduisons une étude du point de vue de la théorie de l'information sur deux groupes pertinents dans la pratique de systèmes sans fil coopératifs, où les différents nœuds radio de l'infrastructure (stations de base et relais), en exploitant la nature transmissive du moyen sans fil, coopèrent entre eux afin d'augmenter les prestations du réseau (par exemple, le débit moyen, la couverture, la robustesse). En particulier, nous nous concentrons sur les réseaux multi-relais semi-duplex et sur le canal d'interférence causal-cognitif (CCIC), ou le canal d'interférence (IC) avec coopération unilatérale à la source.

Le réseau multi-relais est un exemple fondamental d'un système sans fil coopératif [2], où plusieurs relais aident la communication d'une station source (connectée à une infrastructure de réseau) à un utilisateur mobile. L'ajout de stations relais aux infrastructures cellulaires d'aujourd'hui promet de stimuler la performance de réseau en termes de couverture, débit, et robustesse. En réalité, les nœuds relais fournissent des couvertures améliorées dans des zones cibles, offrant une façon par laquelle la station de base peut communiquer avec des utilisateurs situés aux bords de la cellule. De plus, l'utilisation de nœuds relais peut offrir une alternative moins chère et avec une consommation d'énergie inférieure par rapport à l'installation de nouvelles stations de base, particulièrement pour des régions où le déploiement de solutions en fibre optique est impossible. Selon le mode de fonctionnement, les relais sont classifiés en deux catégories : plein-duplex (FD) et semi-duplex (HD). On dit qu'un relais fonctionne en mode FD s'il peut recevoir et transmettre simultanément sur les mêmes ressources de temps-fréquence-espace, et qu'il est en mode HD autrement. Bien que des performances plus hautes puissent être atteintes avec des relais FD, dans les réseaux sans fil commerciaux l'hypothèse du modèle HD est à présent plus pratique que celui de FD. En effet, des restrictions pratiques existent quand un nœud peut simultanément transmettre et recevoir, comme par exemple l'efficacité avec laquelle l'auto-interférence peut être éliminée, ce qui rend son implémentation difficile [3–5]. Il est donc plus réaliste de supposer que les stations relais opèrent en mode HD, soit en duplex avec division de fréquence (FDD) soit en duplex avec division de temps (TDD). En FDD, les relais utilisent une bande de fréquence pour transmettre et une autre pour recevoir, tandis qu'en TDD, les relais écoutent pour une fraction de temps et transmettent ensuite dans le temps restant. Nous analysons d'abord le cas de relais simple, c'est-à-dire, le canal relais classique pour lequel nous cherchons à dériver la performance maximale en termes de capacité, dans l'esprit de [6]. Nous donnerons de l'intuition intéressante sur la conception d'une station relais HD, ce qui est un élément critique pour les réseaux mobiles du futur. Nous considérons alors un nombre général N de stations relais HD. Pour un tel réseau il y a 2^N états possibles de configuration écoute-transmission dont la probabilité doit être optimisée. En raison de la complexité prohibitive de ce problème d'optimisation (c'est-à-dire, exponentielle dans le nombre de relais N), il est critique d'identifier, le cas échéant, les propriétés structurelles de tels réseaux qui peuvent être exploitées pour trouver des solutions optimales ayant une complexité limitée. En utilisant les propriétés des fonctions sous modulaires et des programmes linéaires (LPs), nous montrerons qu'une classe de réseaux multi-relais HD pratiquement pertinente possède des propriétés structurelles qui permettent une remarquable réduction de la complexité (d'exponentiel en N à linéaire en N).

Le CCIC, ou l'IC avec la coopération à la source unilatérale, représente un aspect particulier des réseaux sans fil du futur, à savoir, une application pratique du paradigme de revêtement cognitif [7]. Il consiste d'une source *primaire* PTx (l'émetteur principal) et une source CTx *cognitif / capable* (l'émetteur cognitif) qui ont pour but de communiquer de façon fiable avec deux destinataires différents, à savoir le PRx (le récepteur primaire) et le CRx (le récepteur cognitif), via un canal commun. Différemment du IC classique non-coopératif, dans le CCIC le CTx (grâce aux capacités radio avancées) peut entendre le PTx par un lien sans-fil partagé bruyant; le CTx peut donc exploiter ces informations pour améliorer le taux de performance des deux systèmes (principal et cognitif). La caractéristique majeure et nouvelle du CCIC est le concept de *connaissance causale / coopération à la source*, qui représente tant un outil de contrôle de l'interférence qu'un modèle pratique pour la technologie radio cognitive. En réalité, la coopération unilatérale à la source offre un moyen de gérer et faire face à l'interférence 'avec intelligence'. Dans les systèmes sans fil d'aujourd'hui, l'approche générale pour traiter l'interférence consiste soit à l'éviter, en essayant de rendre orthogonale (dans le temps / la fréquence / l'espace) la transmission des utilisateurs, soit à simplement la traiter comme du bruit. Cependant, ces approches peuvent sévèrement limiter la capacité du système puisqu'une orthogonalization parfaite au niveau de l'utilisateur n'est pas possible dans la pratique [1]. Au contraire, dans le CCIC le CTx, qui peut causalement apprendre les données du primaire par un lien bruyant, peut protéger soit sa propre information (en pré-codant contre un peu d'interférence connue) soit celle du primaire (en répartissant certaines de ses ressources de transmission pour aider le PTx à transmettre des données au PRx) de l'interférence. Ainsi, les techniques de transmission conçues pour le CCIC ont pour but de démultiplier la structure de l'interférence, au lieu de simplement la traiter comme le bruit. Le CCIC représente aussi un modèle plus pertinent dans la pratique pour le paradigme de revêtement cognitif, comparé au cas où le CTx est assumé à a priori (avant que la transmission ne commence) connaître le message du PTx [11], qui n'est adapté qu'a certains scénarios limités. Au contraire, dans le CCIC, le CTx apprend causalement les données du PTx par un lien bruyant. Ainsi, les techniques de transmission conçues pour le CCIC tiennent en compte le temps dont le CTx nécessite pour décoder et de (possible) pertes ultérieures dans le décodage du message du PTx. Nous étudions des configurations de déploiement différentes, qui correspondent à différents scénarios d'interférence. Dans le scénario d'interférence symétrique les deux destinations sont dans la zone de couverture des deux sources; ceci implique que les deux destinations souffrent de l'interférence. Dans le scénario d'interférence asymétrique, une destination ne souffre pas de l'interférence; dans ce cas un des liens d'interférence est absent. En raison de l'asymétrie en coopération, on doit considérer deux scénarios asymétriques d'interférence : le canal Z, où le lien du PTx au CRx est inexistant (c'est-à-dire, le CRx est hors de la portée du PTx) et le canal S, où le lien du CTx au PRx est inexistant (c'est-à-dire, le PRx est hors de la portée du CTx). Nous assumons de plus deux modes de fonctionnement différents au CTx, à savoir FD (c'est-à-dire que le CTx peut simultanément recevoir et transmettre sur

[1] Un exemple célèbre de comment 'traiter l'interférence comme du bruit' limite sévèrement la capacité du système est donné par un réseau ad hoc où n paires aléatoirement placés ont pour but de communiquer. Dans [8] les auteurs ont montré que, si chaque nœud décode seulement le signal du voisin le plus proche (en traitant tous les autres signaux comme du bruit) le débit par paire destination-source converges vers zéro en $O\left(1/\sqrt{n}\right)$ pour des réseaux *denses* (c'est-à-dire, quand l'aire est fixée et $n \to \infty$). Dans [9], il est montré que cette loi de convergence est optimale d'un point de vue de la théorie de l'information pour des réseaux *étendus* (quand n est fixé et l'aire grandit linéairement avec n) dans la haute atténuation (c'est-à-dire, pour un exposant d'atténuation $\alpha > 4$). Dans [10], Özgür et al. ont montré une nouvelle loi de convergence pour des réseaux denses et des réseaux étendus à faible atténuation: si les nœuds peuvent coopérer, la capacité totale du réseau augmente avec n, c'est-à-dire, le débit de chaque paire destination-source n'est pas détérioré avec l'augmentation de n. Ceci a été accompli par une nouvelle architecture de coopération hiérarchique, où les nœuds dans le même cluster coopèrent dans la livraison des messages à leurs destinations.

la même ressource de temps, fréquence, espace) et HD TDD (c'est-à-dire, dans chaque créneau horaire, le CTx écoute pour une fraction du temps et transmet ensuite pour le temps restant). Pour chaque topologie nous étudions la performance ultime en termes de capacité dans l'esprit de [6, 12], en dérivant de nouvelles bornes supérieures pour la région de capacité et en concevant des stratégies de transmission permettant d'atteindre ces limites dans le cas du canal avec bruit Gaussien.

7.2 Contributions de cette dissertation

Dans cette thèse nous analysons deux modèles de canal sans fil pertinents dans la pratique avec coopération entre les nœuds, à savoir le réseau avec relais HD et le CCIC, ou l'IC avec coopération unilatérale au niveau de la source. Ces deux scénarios sont étudiés dans deux parties différentes, à savoir la Partie I et la Partie II, respectivement. En particulier, notre analyse se sert d'outils de la théorie de l'information et de la théorie des graphes. Les propriétés des fonctions sous-modulaires et de la programmation linéaire sont aussi utilisées.

Cette thèse a abouti à 13 publications de conférence et 6 articles de journal, tous actuellement en phase de soumission ou déjà publié par IEEE. Des parties de ces œuvres sont réimprimées ensuite avec la permission d'IEEE.

7.2.1 Partie I

Dans la Partie I, nous étudions le réseau de relais HD où N stations relais aident la communication entre une source et une destination fonctionnant en HD. Particulièrement,

Chapitre 2. Dans le Chapitre 2, nous analysons le cas avec bruit Gaussien pour $N = 1$, c'est-à-dire, le canal relais Gaussien, dont la capacité exacte $C^{(\mathrm{HD-RC})}$ est inconnue. Nous faisons des progrès vers la détermination de sa capacité en caractérisant son gDoF (degrés-de-liberté généralisés) $\mathsf{d}^{(\mathrm{HD-RC})}$ (voir la Définition 1) analytiquement et en prouvant un résultat à écart constant (voir la Définition 2). Nous proposons aussi une stratégie de transmission inspirée par le LDA (approximation linéaire déterministe), qui est de façon prouvable asymptotiquement optimal.

Définition 1. *Le gDoF du canal relais Gaussien HD est defini comme*

$$\mathsf{d}^{(\mathrm{HD-RC})} := \lim_{\mathsf{SNR}\to+\infty} \frac{C^{(\mathrm{HD-RC})}}{\log(1+\mathsf{SNR})},$$

où SNR *est le rapport signal sur bruit.*

Définition 2. *On dit que la capacité* $C^{(\mathrm{HD-RC})}$ *est connue avec un écart de* GAP *bits, si on peut montrer des débits réalisables* $R^{(\mathrm{in})}$ *et une limite supérieure* $R^{(\mathrm{out})}$ *tels que:*

$$R^{(\mathrm{in})} \leq C^{(\mathrm{HD-RC})} \leq R^{(\mathrm{out})} \leq R^{(\mathrm{in})} + \mathsf{GAP}.$$

Nos propres contributions principales peuvent être résumés comme suit:

1. Nous déterminons la capacité exacte du canal LDA: nous montrons qu'un commutateur aléatoire et des entrées non uniformes corrélés au niveau du relais sont optimaux. Nous montrons aussi qu'un commutateur déterministe induit une perte maximale de 1 bit. Particulièrement, nous démonstrons le théorème suivant,

   Théorème 1. *La capacité du LDA est donnée par:*

$$C^{(\mathrm{HD})} = \begin{cases} \beta_{\mathrm{sd}} & \textit{si } \mathsf{C1} \\ \beta_{\mathrm{sd}} + \max_{\gamma\in[0,1]} \min\left\{\mathsf{A}(\gamma), \gamma(\beta_{\mathrm{sr}} - \beta_{\mathrm{sd}})\right\} & \textit{autrement} \end{cases}, \tag{7.1}$$

   où

$$\begin{aligned} &\mathsf{C1} : \beta_{\mathrm{sd}} \leq \max\{\beta_{\mathrm{sr}}, \beta_{\mathrm{rd}}\}, \\ &\mathsf{A}(\gamma) := (1-\theta^*(\gamma)) \log \frac{1}{1-\theta^*(\gamma)} + \theta^*(\gamma) \log \frac{L-1}{\theta^*(\gamma)}, \\ &\theta^*(\gamma) := 1 - \max\left\{\frac{1}{L}, \gamma\right\}, \ L := 2^{(\beta_{\mathrm{rd}} - \beta_{\mathrm{sd}})}, \end{aligned}$$

   et où $\beta_{\mathrm{sr}}, \beta_{\mathrm{sd}}, \beta_{\mathrm{rd}}$ *sont des entiers non négatifs avec le sens suivant:* β_{sr} *est le nombre de bits envoyés par la source et observés au relais,* β_{sd} *est le nombre de bits envoyés par la source et observés à la destination et* β_{rd} *est le nombre de bits envoyés par le relais et observés à la destination.* $C^{(\mathrm{HD})}$ *est atteinte avec un commutateur aléatoire et des entrées non uniformes corrélés au niveau du relais. De plus, une stratégie avec un commutateur déterministe et des bits indépendant et identiquement distribués selon Bernoulli*(1/2) *au relais est au maximum 1 bit de la capacité dans* (7.1).

2. Nous dérivons le gDoF pour le canal relais Gaussien sous forme analytique: nous montrons que PDF (décoder-et-retransmettre partiel)

ainsi que CF (compresser-et-retransmettre) sont gDoF optimales, tant avec un commutateur déterministe qu'avec un commutateur aléatoire au relais. Nous montrons aussi qu'une technique inspirée par le LDA avec un commutateur déterministe est gDoF optimal. Particulièrement, nous prouvons le théorème suivant,

Théorème 2. *Le gDoF* $\mathsf{d}^{(\mathrm{HD-RC})}$ *du canal relais Gaussien HD est*

$$\mathsf{d}^{(\mathrm{HD-RC})} = \begin{cases} \beta_{\mathrm{sd}} + \frac{(\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}})(\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}})}{(\beta_{\mathrm{rd}}-\beta_{\mathrm{sd}})+(\beta_{\mathrm{sr}}-\beta_{\mathrm{sd}})} & \textit{pour } \beta_{\mathrm{sr}} > \beta_{\mathrm{sd}},\ \beta_{\mathrm{rd}} > \beta_{\mathrm{sd}} \\ \beta_{\mathrm{sd}} & \textit{autrement} \end{cases},$$

où: β_{sr} *est l'exposant SNR du lien source-relais,* β_{sd} *est l'exposant SNR du lien source-destination et* β_{rd} *est l'exposant SNR du lien relais-destination.*

3. Pour le cas de bruit Gaussien, nous montrons que les stratégies de transmission ci-dessus sont optimales à l'intérieur d'un intervalle constant, uniformément sur tous les paramètres du canal. En particulier, PDF est optimale à 1 bit, CF à 1.61 bits et la stratégie inspirée par le LDA à 3 bits. Dans tous les cas, l'écart est plus petit que celui de 5 bits existant dans la littérature pour le cas d'un relais [28]. Nos résultats d'écart constant sont résumés dans le tableau suivant,

Stratégie de transmission	LDAi	CF	PDF
Écart Analytique	3 bits	1.61 bits	1 bit
Écart Numérique	1.32 bits	1.16 bits	1 bit

où LDAi est une stratégie réalisable inspirée par le LDA.

4. Pour les trois systèmes de codage, nous obtenons une expression analytique pour la durée des phases de transmission et réception au relais avec interrupteur déterministe. Ce résultat met en lumière la conception d'un nœud HD de relais dans les réseaux sans fil de demain.

5. Nous prouvons que PDF avec un commutateur aléatoire est exactement optimale pour un réseau des relais placés sur une ligne sans mémoire générale, soit, lorsque le lien direct entre la source et la destination est absent. Une expression analytique pour la distribution d'entrée optimale avec la politique de commutateur aléatoire n'est cependant pas disponible.

Le travail présenté dans ce chapitre a donné lieu aux publications suivantes:

- [65] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “Gaussian half-duplex relay channels: generalized degrees of freedom and constant gap result”, in 2013 IEEE International Conference on Communications (ICC 2013), Budapest (Hungary), June 2013.
- [66] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “The capacity to within a constant gap of the Gaussian half-duplex relay channel”, in 2013 IEEE International Symposium on Information Theory (ISIT 2013), Istanbul (Turkey), July 2013.
- [67] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “On the Gaussian half-duplex relay channel”, in IEEE Transactions on Information Theory, Volume 60, Issue n.5, May 2014, Pages 2542-2562.

Une mise en œuvre concrète de la stratégie de transmission inspirée par le LDA peut être trouvée dans

- [68] R. Thomas, M. Cardone, R. Knopp, D. Tuninetti, B. T. Maharaja, “An LTE implementation of a novel strategy for the Gaussian half-duplex relay channel”, to appear in 2015 IEEE International Conference on Communications (ICC 2015), London (United Kingdom), June 2015.

Chapitre 3. Dans le Chapitre 3, nous étudions le réseau relais HD avec un certain nombre de relais N générale, en suivant l'approche proposée dans [18]. Nos principales contributions peuvent être résumées comme suit:

1. Pour le cas du bruit Gaussien pertinent dans la pratique, nous prouvons que NNC (codage de réseau avec bruit) avec un interrupteur déterministe atteint la limite *cut-set* (correctement évaluée pour tenir compte de l'interrupteur aléatoire) à moins de $1.96(N+2)$ bits. Cet écart est plus petit que l'écart de $5N$ bits disponible dans la littérature [28]. Notre résultat d'écart pour un réseau HD de relais est obtenu comme un cas particulier d'un résultat plus général pour un réseau Gaussien multidiffusion HD, qui étend l'écart de 1.26 bits au nœud pour le cas FD [20] à un écart au nœud de 1.96 bits pour le cas de HD. Nous montrons également que ce résultat d'intervalle peut être

étendu au cas des nœuds multi-antennes et est de $1,96$ bits par utilisation du canal par antenne. En particulier, nous prouvons le théorème suivant

Théorème 3. *La cut-set borne supérieure de la capacité du réseau relais Gaussien HD avec unique antenne et N relais est atteinte par NNC avec interrupteur déterministe à moins de*

$$\mathsf{GAP} \leq 1.96(N+2) \text{ bits.} \tag{7.2}$$

2. Afin de déterminer le gDoF du canal Gaussien, nous avons besoin de trouver une approximation précise à SNR élevé pour les différents termes d'information mutuelle impliqués dans la borne supérieure *cut-set*. En conséquence d'un intérêt indépendant, au-delà de son application au réseau relais Gaussien étudié dans ce chapitre, nous montrons que telles approximations précises peuvent être trouvées comme solution de problèmes d'appariement maximum bipartite pondéré (MWBM), ou problèmes d'affectation [69], pour lesquels des algorithmes efficaces à temps polynomial, tels que l'algorithme Hongrois [70], existent. À titre d'exemple, on montre que cette technique est utile pour dériver le gDoF de réseaux Gaussiens diffusés avec et sans relais et pour résoudre des problèmes de planification de l'utilisateur. En particulier, nous démontrons le théorème suivant

   Théorème 4. *Le gDoF* $\mathsf{d}^{(\mathrm{HD-RN})}$ *du réseau multi-relais Gaussien HD est la solution du problème de programmation linéaire suivant:*

$$\text{maximizer } \{\mathbf{f}^T \ \mathbf{x}\} \tag{7.3}$$

$$\text{sujet } \begin{bmatrix} -\mathbf{A} & \mathbf{1}_{2^N} \\ \mathbf{1}_{2^N}^T & 0 \end{bmatrix} \mathbf{x} \leq \mathbf{f}, \quad \mathbf{x} \geq 0, \tag{7.4}$$

   où $\mathbf{x}^T := [\lambda_{\mathrm{vect}}, \mathsf{d}^{(\mathrm{HD-RN})}]$ *avec* $\lambda_{\mathrm{vect}} := [\lambda_s] \in \mathbb{R}_+^{1\times 2^N}$, $\mathbf{f}^T := [\mathbf{0}_{2^N}^T, 1]$ *et où les éléments de la matrice non négative* $\mathbf{A} \in \mathbb{R}^{2^N \times 2^N}$ *peuvent être calculés comme solutions de* $2^{N-1}\left(2^N+1\right)$ *problèmes d'affectation indépendants.*

3. Nous prouvons la conjecture de Brahma et al. [33] au-delà des réseaux Gaussiens avec une topologie de diamant. En particulier, nous montrons que pour tout réseau HD avec N relais, avec des bruits indépendants et pour lesquels la borne supérieure *cut-set* est approximativement optimale à une constante sous certaines hypothèses, la politique

de commutation de relais (approximativement) optimale est simple, à savoir, au plus $N+1$ états (sur les 2^N possibles en total) ont une probabilité strictement positive. L'idée principale est d'utiliser l'extension de Lovàsz et l'algorithme glouton pour les polyèdres sous-modulaires pour mettre en évidence les propriétés structurelles du minimum d'une fonction sous-modulaire. Puis, en utilisant la propriété de point col des problèmes de min-max et l'existence de solutions faisables basiques optimales pour la programmation linéaire, une politique de relais (approximativement) optimale avec le nombre d'états actifs déclaré peut être démontrée. Réseaux à relais avec bruit Gaussien satisfont à toutes les hypothèses et donc admettent une programmation simple. Plus important encore, lorsque les nœuds sont équipés de plusieurs antennes et les antennes aux relais peuvent être commutées entre mode d'émission et de réception indépendamment les unes des autres, la stratégie de commutation a au plus $N+1$ états actifs (comme dans le cas mono-antenne), quel que soit le nombre total d'antennes dans le système. En particulier, nous prouvons le théorème suivant

Théorème 5. *En général, pour chaque réseau relais HD sans mémoire, pour lequel:*

3.1. entrées indépendantes sont approximativement (à l'intérieur d'un intervalle constant) optimales dans la borne supérieure cut-set, c'est-à-dire s'il existe une distribution d'entrée en forme de produit

$$\mathbb{P}_{X_{[1:N+1]}|S_{[1:N]}} = \prod_{i\in[1:N+1]} \mathbb{P}_{X_i|S_{[1:N]}} \tag{7.5}$$

pour laquelle nous pouvons limiter la capacité $C^{(\mathrm{HD-RN})}$ comme

$$\mathsf{C}' - \mathsf{G}_1 \le C^{(\mathrm{HD-RN})} \le \mathsf{C}' + \mathsf{G}_2, : \; \mathsf{C}' := \max_{\mathbb{P}_{S_{[1:N]}}} \min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(fix)}, \tag{7.6}$$

où G_1 et G_2 sont des constantes non-négatives qui peuvent être dépendantes en N mais pas de la probabilité de transition du canal, et où

$$I_{\mathcal{A}}^{(fix)} := I\left(X_{N+1}, X_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{[1:N]}\right) \tag{7.7}$$

$$= \sum_{s\in[0:1]^N} \lambda_s \; f_s(\mathcal{A}), \tag{7.8}$$

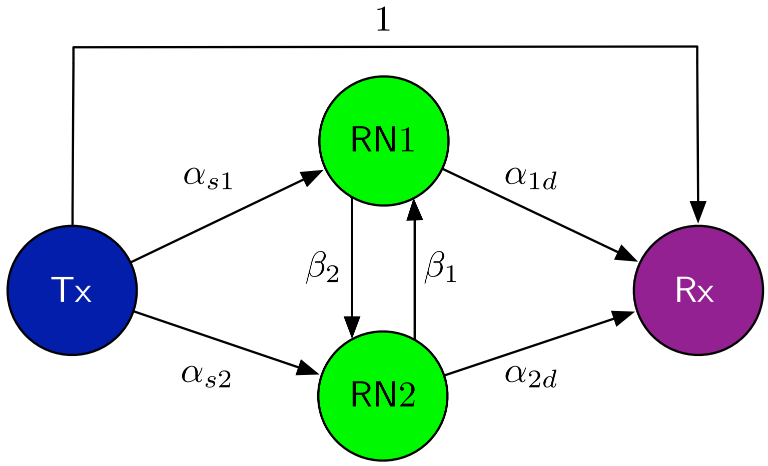


Figure 7.1: Example d'un reseau avec $N = 2$ relais et avec nœuds mono-antenne.

avec

$$\lambda_s := \mathbb{P}[S_{[1:N]} = s] \in [0,1] : \sum_{s\in[0:1]^N} \lambda_s = 1, \tag{7.9}$$

$$f_s(\mathcal{A}) := I\left(X_{N+1}, X_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}} | X_{\mathcal{A}}, S_{[1:N]} = s\right), \tag{7.10}$$

3.2. les "bruits sont indépendants", c'est-à-dire

$$\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]},S_{[1:N]}} = \prod_{i\in[1:N+1]} \mathbb{P}_{Y_i|X_{[1:N+1]},S_{[1:N]}}, \tag{7.11}$$

3.3. les fonctions en (7.10) *ne sont pas en fonction de* $\{\lambda_s, s \in [0 : 1]^N\}$*, soit elles peuvent dépendre de l'état* s *mais pas de* $\{\lambda_s, s \in [0 : 1]^N\}$*,*

donc des politiques de relais simples sont (environ) optimales en (7.6)*, c'est-à-dire, la fonction de probabilité de masse (environ) optimale* $\mathbb{P}_{S_{[1:N]}}$ *a au plus* $N + 1$ *entrées non nulles / états actifs.*

4. Nous considérons enfin deux exemples de réseaux: pour le premier scénario dans la Figure 7.1, constitué de $N = 2$ relais à antenne unique, nous mettons en évidence les conditions de canal dans lesquelles une

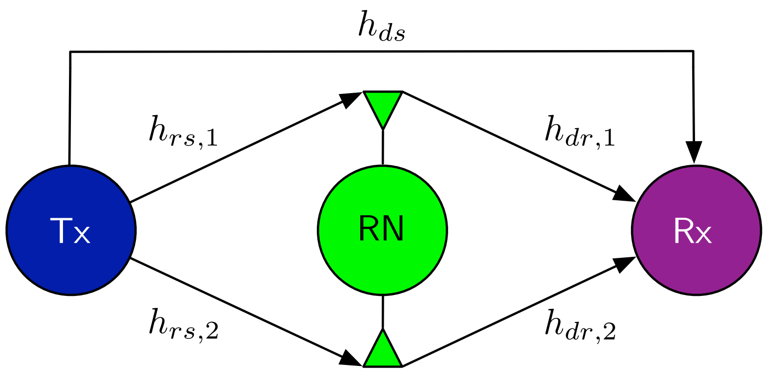


Figure 7.2: Example d'un reseau avec $N = 1$ relais avec $m_r = 2$ antennes, et source et destination mono-antenne.

stratégie de sélection du meilleur relais est strictement sous-optimale en termes de gDoF et nous gagnons un aperçu de la nature du gain de débit atteignable dans des réseaux avec plusieurs relais; pour le deuxième scénario dans la Figure 7.2, consistant en $N = 1$ relais équipé de 2 antennes, nous montrons que la commutation des deux antennes indépendamment au niveau du relais non seulement atteint en général des débits strictement plus élevés par rapport à l'utilisation des antennes pour le même but, mais peut fournir effectivement un facteur de pré-logarithme strictement plus grand.

Le travail présenté dans ce chapitre a donné lieu aux publications suivantes:

- [71] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “Gaussian half-duplex relay networks: improved gap and a connection with the assignment problem”, in 2013 IEEE Information Theory Workshop (ITW 2013), Seville (Spain), September 2013.

- [72] M. Cardone, D. Tuninetti, R. Knopp, “On user scheduling for maximum throughput in K-user MISO broadcast channels”, to appear in 2015 IEEE International Conference on Communications (ICC 2015), London (United Kingdom), June 2015.

- [73] M. Cardone, D. Tuninetti, R. Knopp, “The approximate optimality of simple schedules for half-duplex multi-relay networks”, to appear in 2015 IEEE Information Theory Workshop (ITW 2015), Jerusalem (Israel), May 2015.

- [74] M. Cardone, D. Tuninetti, R. Knopp, "Gaussian MIMO half-duplex relay networks: approximate optimality of simple schedules", to appear in 2015 IEEE International Symposium on Information Theory (ISIT 2015), Hong Kong, June 2015.

- [75] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "Gaussian half-duplex relay networks: improved constant gap and connections with the assignment problem", in IEEE Transactions on Information Theory, Volume 60, Issue n.6, June 2014, Pages 3559-3575.

- [76] M. Cardone, D. Tuninetti, R. Knopp, "On the optimality of simple schedules for networks with multiple half-duplex relays", submitted to IEEE Transactions on Information Theory, December 2014.

7.2.2 Partie II

Dans la Partie II, nous étudions le CCIC, ou l'IC avec coopération de la source unilatérale, qui se compose de deux paires de source-destination partageant le même canal et où le CTx espionne le PTx par un lien de communication avec perte et peut donc consacrer une partie de ses ressources de transmission pour aider la communication du paire primaire. En particulier,

Chapitre 4. Dans le Chapitre 4, nous considérons le mode de fonctionnement FD à la source cognitive, à savoir le CTx peut recevoir et transmettre simultanément sur les mêmes ressources de fréquence-espace-temps. Nos principales contributions peuvent être résumées comme suit:

1. Nous développons un cadre général pour calculer les bornes supérieures du type $2R_{\mathsf{p}}+R_{\mathsf{c}}$ et $R_{\mathsf{p}}+2R_{\mathsf{c}}$ (où R_{p}, respectivement R_{c}, est le débit de transmission du PTx, respectivement du CTx) sur la capacité du CCIC général ISD (injectif semi-déterministe) lorsque les bruits aux différents couples source-destination sont indépendants; ce cadre comprend par exemple les retours d'information de la destination prévue. Comme cas particulier, nous retrouvons et renforçons les limites dérivées dans [47, 57]. L'ingrédient technique clé est la preuve de deux chaînes de Markov. En particulier, nous prouvons le théorème suivant

   Théorème 6. *Pour le ISD CCIC de Figure 7.3 satisfaisant*

$$\mathbb{P}_{Y_{\mathsf{Fc}},Y_{\mathsf{p}},Y_{\mathsf{c}}|X_{\mathsf{p}},X_{\mathsf{c}}} = \mathbb{P}_{Y_{\mathsf{p}}|X_{\mathsf{p}},X_{\mathsf{c}}}\mathbb{P}_{Y_{\mathsf{Fc}},Y_{\mathsf{c}}|X_{\mathsf{p}},X_{\mathsf{c}}}$$

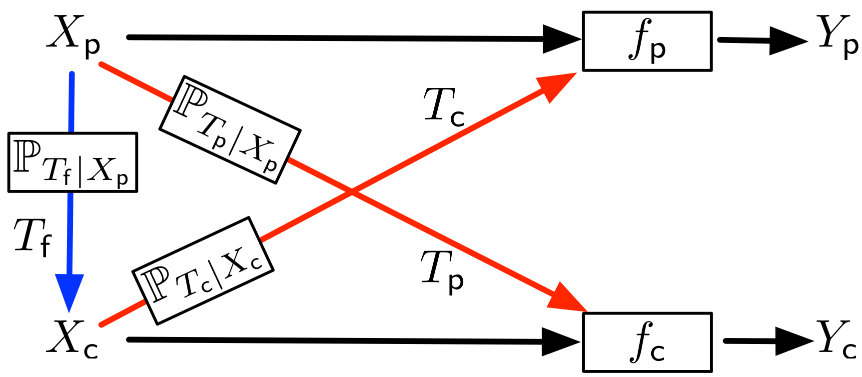


Figure 7.3: Le CCIC général ISD, où $Y_{\mathsf{p}} = f_{\mathsf{p}}(X_{\mathsf{p}}, T_{\mathsf{c}})$, $Y_{\mathsf{c}} = f_{\mathsf{c}}(X_{\mathsf{c}}, T_{\mathsf{p}})$ et $Y_{\mathsf{Fc}} = f_{\mathsf{f}}(X_{\mathsf{c}}, T_{\mathsf{f}})$, où f_u, $u \in \{\mathsf{p}, \mathsf{c}\}$, est une fonction déterministe et invertible donnée X_u et f_{f} est une fonction déterministe et invertible donnée X_{c}.

la région de capacité est bornée supérieurement par

$$2R_{\mathsf{p}}+R_{\mathsf{c}} \leq I(Y_{\mathsf{p}}; X_{\mathsf{p}}, X_{\mathsf{c}})+I(Y_{\mathsf{p}}; X_{\mathsf{p}}|Y_{\mathsf{c}}, T_{\mathsf{f}}, X_{\mathsf{c}})+I(Y_{\mathsf{c}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}}|T_{\mathsf{c}}),$$
$$R_{\mathsf{p}}+2R_{\mathsf{c}} \leq I(Y_{\mathsf{c}}; X_{\mathsf{p}}, X_{\mathsf{c}})+I(Y_{\mathsf{c}}; X_{\mathsf{c}}|Y_{\mathsf{p}}, T_{\mathsf{f}}, X_{\mathsf{p}})+I(Y_{\mathsf{p}}, T_{\mathsf{f}}; X_{\mathsf{p}}, X_{\mathsf{c}}|T_{\mathsf{p}}),$$

pour certaines distributions d'entrée $\mathbb{P}_{X_{\mathsf{p}}, X_{\mathsf{c}}}$.

2. Nous concevons une stratégie de transmission pour le CCIC général sans mémoire et nous en dérivons la région de débit atteignable. La stratégie proposée utilise la superposition et l'encodage *binning*, transmission PDF et le décodage simultané sur les récepteurs. Dès que le CCIC partage des caractéristiques communes avec l'IC classique non coopératif, des messages à la fois communes (décodés également au niveau du récepteur non-voulu) et privés (traités comme du bruit au niveau du récepteur non-voulu) sont utilisés. En outre, nous utilisons des messages coopératifs (envoyés en coopération avec le CTx) et non-coopératifs (directement envoyés au PRx sans l'aide du CTx) pour le PTx, tandis que les messages du CTx sont seulement non coopératifs.

3. Nous évaluons la limite externe et les régions de débit atteignable pour le canal de bruit Gaussien pertinent dans la pratique dans la Figure 7.4.

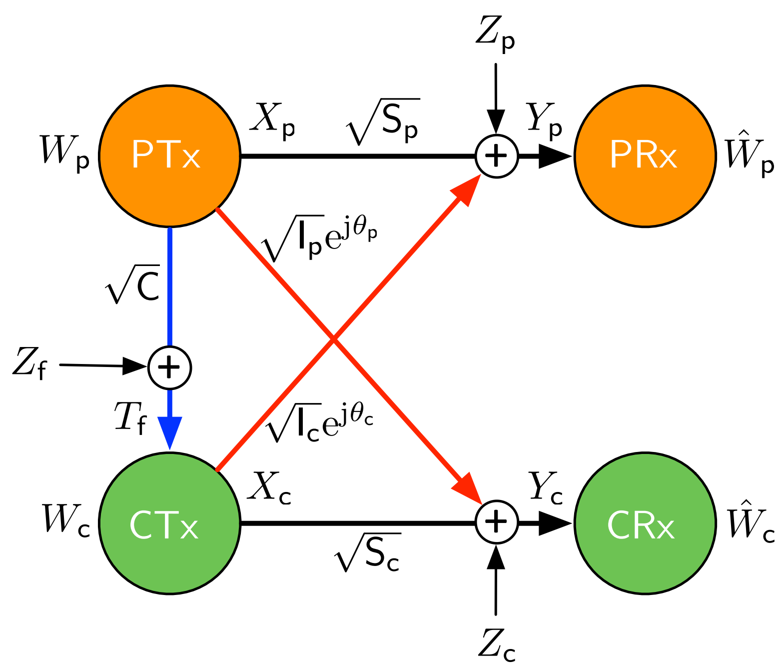


Figure 7.4: Le CCIC de bruit Gaussien.

Nous prouvons que pour le cas symétrique, à savoir, lorsque les deux liens directs et les deux liens croisés interférents sont de la même force, pour le canal Z, à savoir, lorsque le lien du PTx au CRx est absent, et pour le canal S, soit, lorsque le lien du CTx au PRx est absent, la région atteignable est à un nombre constant de bits (uniformément sur tous les gains de canal) de la région de la borne extérieure. En particulier, nos résultats d'écart constant sont indiqués dans les trois théorèmes suivants.

Théorème 7. *La limite extérieure de la région de capacité du CCIC Gaussien symétrique (c'est-à-dire avec référence à la Figure 7.4 où* $S_p = S_c = S$ *et* $I_p = I_c = I$*) est atteignable à moins de* 5 *bits. Particulièrement,*

3.1. quand $I \geq S$*, alors* $\mathsf{GAP} \leq 1$ *bit,*

3.2. quand $I < S$ *et* $C \leq S$*, alors* $\mathsf{GAP} \leq 5$ *bits,*

3.3. quand $I < S$ *et* $S < C$*, alors* $\mathsf{GAP} \leq 2$ *bits.*

Ces résultats d'écart constant sont également reportés sur la Figure

7.5, où α est l'exposant SNR des liens d'interférence, à savoir $\mathsf{I} = \mathsf{S}^{\alpha}$ et β est l'exposant SNR de la liaison de coopération, à savoir, $\mathsf{C} = \mathsf{S}^{\beta}$.

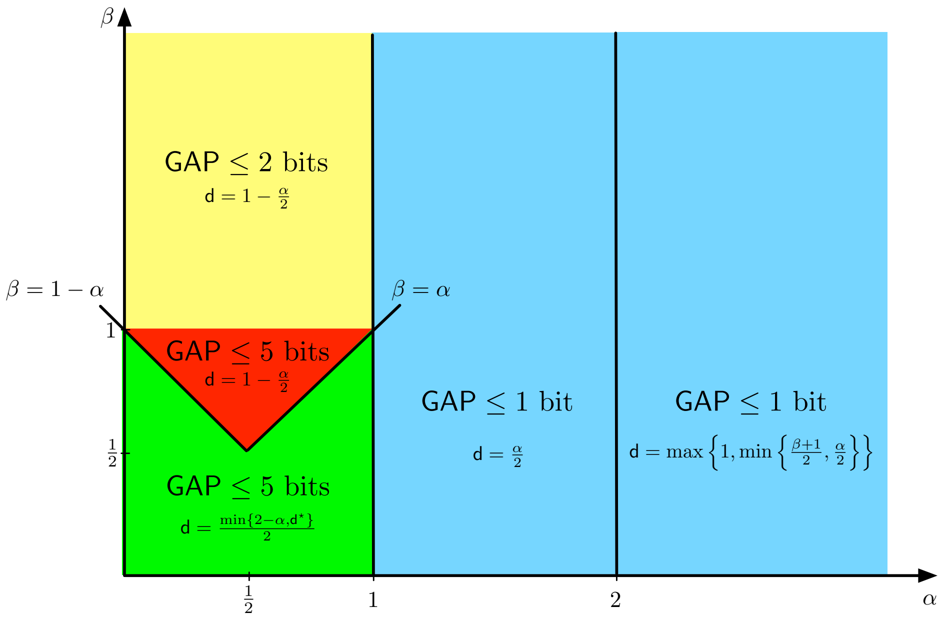


Figure 7.5: Différent régimes dependant de valeurs α et β, avec $\mathsf{d}^{\star} := \max\{\alpha, 1-\alpha\} + \max\{\alpha, 1+\beta-\max\{\alpha,\beta\}\}$.

Théorème 8. *La limite extérieure de la région de capacité du canal Z (à savoir, avec référence à la Figure 7.4* $\mathsf{I_p} = 0$*, le lien PTx → CRx est non-existant) est caractérisée sous* 2 *bits. En particulier,*

3.1. quand $\mathsf{C} \leq \mathsf{S_p}$*, alors* $\mathsf{GAP} \leq 2$ *bits,*

3.2. quand $\mathsf{C} > \mathsf{S_p}$ *et* $\mathsf{S_c} \leq \mathsf{I_c}$*, alors* $\mathsf{GAP} \leq 1.5$ *bits,*

3.3. quand $\mathsf{C} > \mathsf{S_p}$ *et* $\mathsf{S_c} > \mathsf{I_c}$*, alors* $\mathsf{GAP} \leq 1$ *bit.*

Ces résultats d'écart constant sont également représentés sur la Figure 7.6 pour le cas de liens directs aussi fortes, soit, $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}^1$ *et où* α *est l'exposant SNR de la liaison d'interférence, soit* $\mathsf{I_c} = \mathsf{S}^{\alpha}$ *et* β *est l'exposant SNR du lien de coopération, à savoir,* $\mathsf{C} = \mathsf{S}^{\beta}$.

Théorème 9. *La limite extérieure de la région de capacité du canal S (à savoir, avec référence à la Figure 7.4* $\mathsf{I_c} = 0$*, le lien CTx → PRx est non-existant) est atteignable à moins de* 3 *bits. En particulier,*

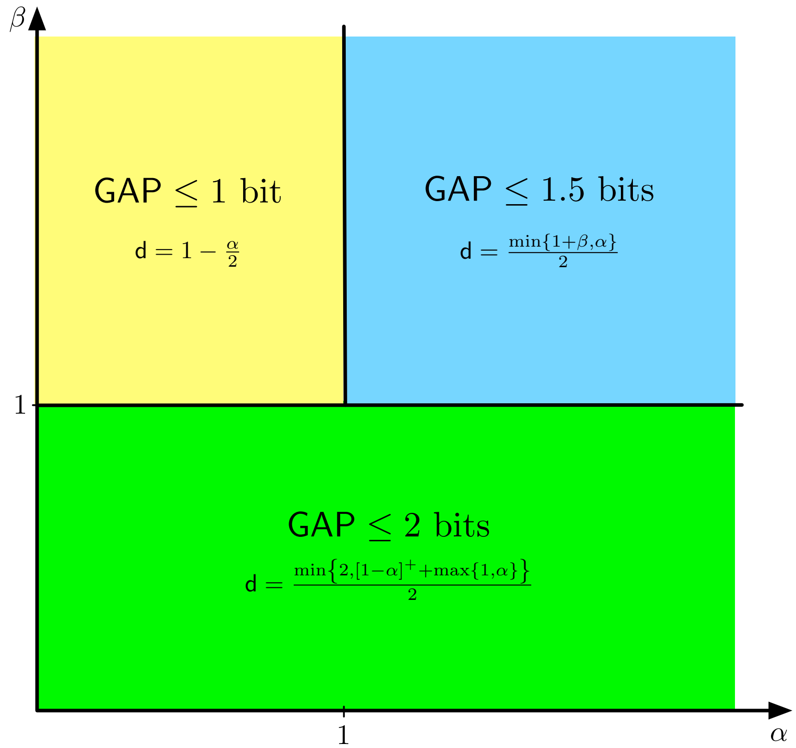


Figure 7.6: gDoF optimals et écart constant pour le canal Z dans les différents régimes dans le plan (α, β).

3.1. quand $\mathsf{C} \leq \max\{\mathsf{S_p}, \mathsf{I_p}\}$*, alors* $\mathsf{GAP} \leq 2.5$ *bits,*

3.2. quand $\max\{\mathsf{S_p}, \mathsf{I_p}\} < \mathsf{C} \leq \mathsf{I_p}\mathsf{S_p}$*, alors* $\mathsf{GAP} \leq 3$ *bits,*

3.3. quand $\mathsf{C} > \mathsf{I_p}\mathsf{S_p}$*, alors* $\mathsf{GAP} \leq 1$ *bit.*

Ces résultats d'écart constant sont également rapportés sur la Figure 7.7 pour le cas de liens directs aussi forts $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}^1$ *et où* α *est l'exposant SNR de la liaison d'interférence, à savoir,* $\mathsf{I_p} = \mathsf{S}^\alpha$ *et* β *est l'exposant SNR du lien de coopération, à savoir,* $\mathsf{C} = \mathsf{S}^\beta$*.*

Fait intéressant, nous montrons que les régions de capacité des deux scénarios asymétriques (c'est-à-dire le canal Z et le canal S) ne possèdent pas de limites du type $2R_\mathsf{p} + R_\mathsf{c}$ et $R_\mathsf{p} + 2R_\mathsf{c}$, à savoir, la coopération unilatérale permet une pleine utilisation des ressources du canal. D'autre part, nous prouvons que les deux nouvelles limites extérieures du type $2R_\mathsf{p} + R_\mathsf{c}$ et $R_\mathsf{p} + 2R_\mathsf{c}$ sont actives pour le canal

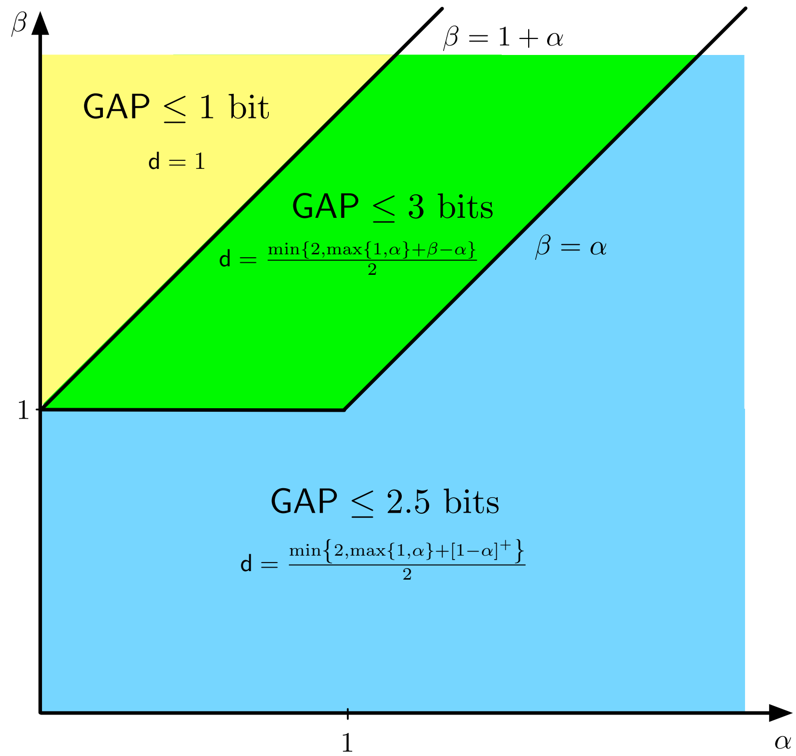


Figure 7.7: gDoF optimals et écart constant pour le canal S dans les différents régimes dans le plan (α, β).

symétrique avec interférences faible et lorsque le lien de coopération est plus faible que le lien direct, ce qui signifie que pour ce régime la coopération unilatérale est trop faible et laisse des ressources de système sous-utilisées.

4. Les résultats d'écart constant impliquent la connaissance exacte du gDoF pour les canaux Z, S et symétrique. Pour chaque configuration, le gDoF est rapporté sur la Figure 7.5 (pour le canal symétrique), sur la Figure 7.6 (pour le canal Z) et sur la Figure 7.7 (pour le canal S). Nous identifions les régimes de paramètres où le CCIC Gaussien (à la fois avec des configurations symétriques et asymétriques) est équivalent en termes de gDoF au IC Gaussien non-coopératif [12] (c'est-à-dire, l'implémentation de la coopération unilatérale pourrait ne pas être intéressant dans les systèmes pratiques) et au CIC Gaussien non-causal [64] (en d'autres mots, la coopération causale unilatérale atteint

la performance ultime de la technologie radio cognitive). Ces comparaisons mettent en lumières les régimes de paramètres et les topologies de réseau qui pourraient fournir dans la pratique un gain de débit illimité par rapport à des technologies actuellement disponibles (non-cognitives).

Le travail présenté dans ce chapitre a donné lieu aux publications suivantes:

- [77] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “Approximate sum-capacity of full- and half-duplex asymmetric interference channels with unilateral source cooperation”, in 2013 Information Theory and Applications Workshop (ITA 2013), San Diego (USA), February 2013.
- [78] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “On the interference channel with causal cognition”, in 2013 IEEE International Conference on Communications (ICC 2013), Budapest (Hungary), June 2013.
- [79] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “On the Gaussian interference channel with unilateral generalized feedback”, in 6th International Symposium on Communications, Control and Signal Processing (ISCCSP 2014), Athens (Greece), May 2014.
- [80] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “On the capacity of full-duplex causal cognitive interference channels to within a constant gap”, in 2014 IEEE International Conference on Communications (ICC 2014), Sydney (Australia), June 2014.
- [81] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “New outer bounds for the interference channel with unilateral source cooperation”, in 2014 IEEE International Symposium on Information Theory (ISIT 2014), Honolulu (Hawaii), July 2014.
- [82] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, “On the capacity of the two-user Gaussian causal cognitive interference channel”, in IEEE Transactions on Information Theory, Volume 60, Issue n.5, May 2014, Pages 2512-2541.
- [83] M. Cardone, D. Tuninetti, R. Knopp, “The two-user causal cognitive interference channel: novel outer bounds and constant gap result for the symmetric Gaussian noise channel in weak Interference”, submitted to IEEE Transactions on Information Theory, March 2014.

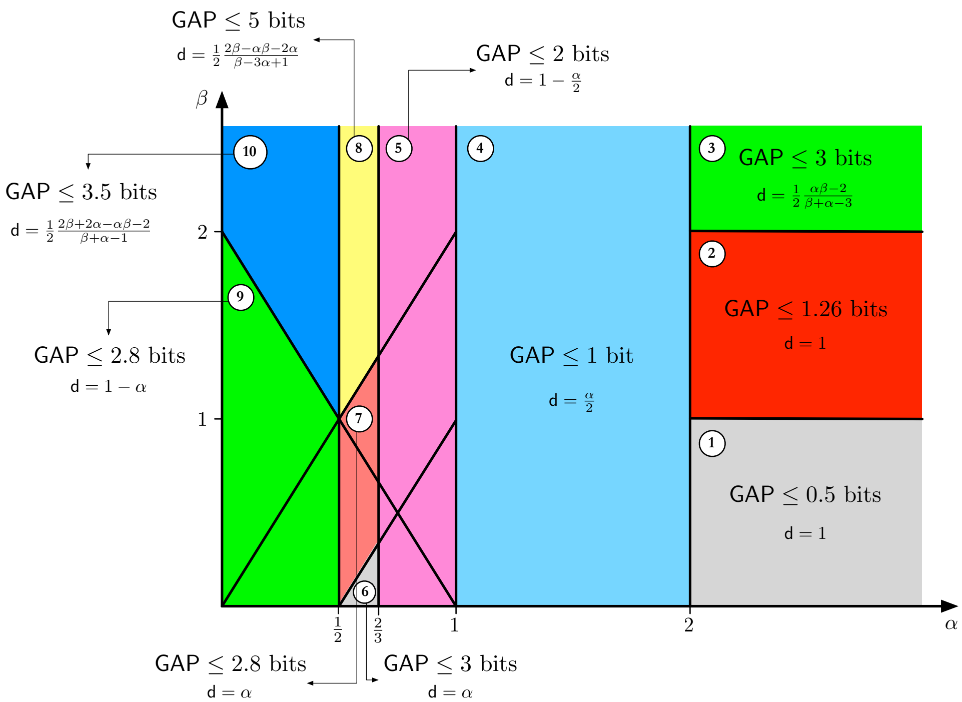


Figure 7.8: Différent régimes dependant de valeurs α et β.

Chapitre 5. Dans le Chapitre 5, nous considérons le mode de fonctionnement HD à la source cognitive, à savoir, dans chaque tranche de temps le CTx écoute pour une fraction du temps et puis transmet dans le temps restant. Nos principales contributions peuvent être résumées comme suit:

1. Nous caractérisons la somme des capacités à l'intérieur d'un intervalle constant pour le canal Z Gaussien symétrique, le canal S Gaussien symétrique et le HD-CCIC Gaussien symétrique entièrement connecté; cela est accompli en adaptant les bornes supérieurs de la somme des capacités pour la coopération FD unilatérale dans le Chapitre 4 au cas de la coopération unilatérale HD en utilisant le cadre théorique de [18], à savoir, en tenant correctement en compte d'un commutateur aléatoire au CTx, et en concevant des nouvelles stratégies de transmission inspirées par le LDA du canal avec bruit Gaussien à SNR élevé. En particulier, l'intervalle est de 5 bits/utilisateur pour le cas symétrique et de 3 bits/utilisateur pour le canal Z symétrique et le canal S symétrique. Nous remarquons que ces résultats d'écart limité, différemment de [55], sont dérivés en tenant correctement en compte

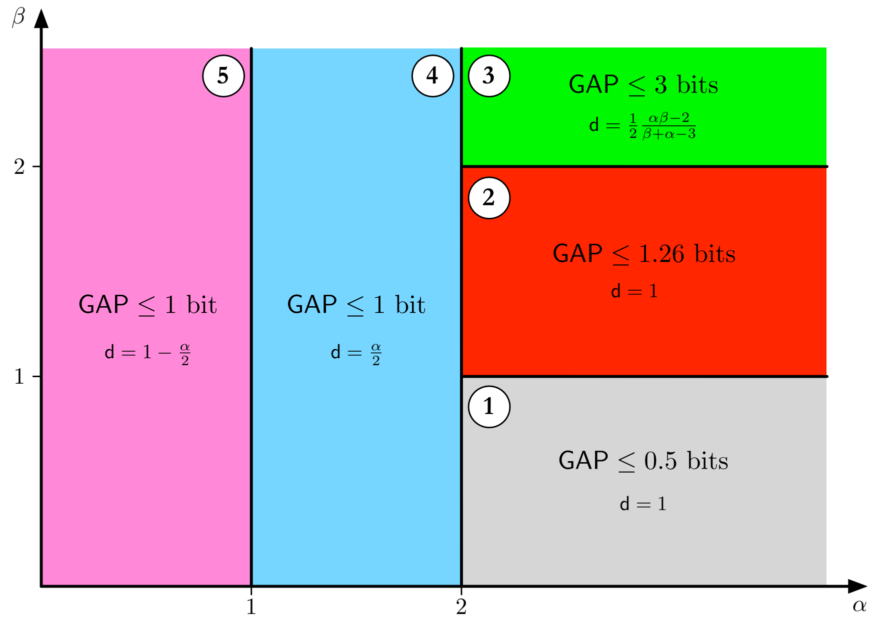


Figure 7.9: gDoF optimals et écartement constant pour le canal Z dans les différents régimes dans le plan (α, β).

l'utilisation d'un commutateur aléatoire au CTx, et ils sont aussi plus petits que ceux dérivés en [55]. En particulier, nos résultats d'écart constants sont indiqués dans les trois théorèmes suivants.

Théorème 10. *La somme des capacités du HD - CCIC Gaussien symétrique (à savoir, avec référence à la Figure 7.4 où* $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ *et* $\mathsf{I_p} = \mathsf{I_c} = \mathsf{I}$*) est atteignable à moins de* 5 *bits/utilisateur.*

La Figure 7.8 montre l'écart (par utilisateur) pour le HD - CCIC symétrique Gaussien pour les différentes régions dans le plan (α, β)*, où l'ensemble entier des paramètres a été partitionné en plusieurs sous-régions en fonction de différents niveaux de coopération (*β*, avec* $\mathsf{C} = \mathsf{S}^\beta$*) et d'interférence (*$\alpha$*, avec* $\mathsf{I} = \mathsf{S}^\alpha$*).*

Théorème 11. *La somme des capacités du canal Z symétrique (c'est-à-dire, avec référence à la Figure 7.4 où* $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ *et* $\mathsf{I_p} = 0$*, donc le lien PTx* $\to$ *CRx est non-existant) est caractérisée à moins de* 3 *bits/utilisateur.*

Le diagramme dans la Figure 7.9 montre l'écart (par utilisateur) pour le canal Z symétrique Gaussien pour les différentes régions dans le

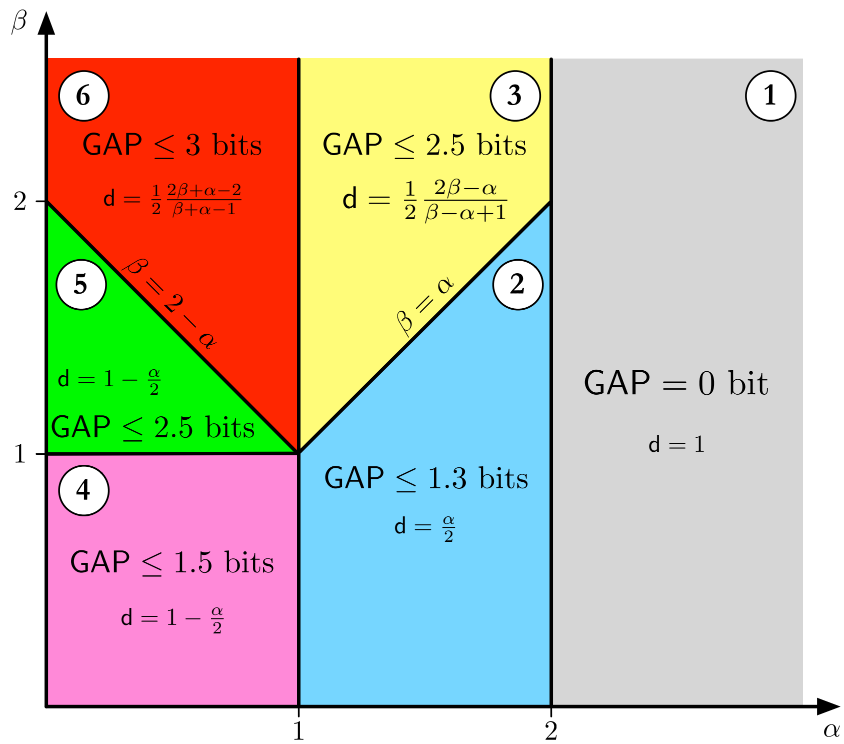


Figure 7.10: gDoF optimals et écart constant pour le canal S dans les différents régimes dans le plan (α, β).

plan (α, β), où l'ensemble entier des paramètres a été partitionné en plusieurs sous-régions en fonction des différents niveaux de coopération (β, avec $\mathsf{C} = \mathsf{S}^{\beta}$) et d'interférence ($\alpha$, avec $\mathsf{I_c} = \mathsf{S}^{\alpha}$).

Théorème 12. *La somme des capacités du canal S symétrique (c'est-à-dire, avec référence à la Figure 7.4 où $\mathsf{S_p} = \mathsf{S_c} = \mathsf{S}$ et $\mathsf{I_c} = 0$, donc le lien CTx → PRx est non-existant) est atteignable à moins de 3 bits/utilisateur.*

Le diagramme dans Figure 7.10 montre l'écart (par utilisateur) pour le canal S symétrique Gaussien pour les différentes régions dans le plan (α, β), où l'ensemble des paramètres a été partitionné en plusieurs sous-régions en fonction de différents niveaux de coopération (β, avec $\mathsf{C} = \mathsf{S}^{\beta}$) et d'interférence ($\alpha$, avec $\mathsf{I_p} = \mathsf{S}^{\alpha}$).

2. En utilisant le modèle LDA, nous obtenons une expression analytique

pour le gDoF et pour les différentes variables d'optimisation (par exemple, la stratégie, les allocations de puissance, les systèmes de codage et ordres de décodage correspondants, etc.). Ce résultat met en lumière la façon dont la conception du CTx HD doit être correctement effectuée, ce qui est une tâche pratique importante pour les réseaux sans fil de demain.

3. Comme pour le cas FD au Chapitre 4, nous comparons le gDoF du HD-CCIC Gaussien avec celui de: (i) le IC classique non coopératif, à savoir, où il n'y a pas de coopération entre les nœuds [12], et (ii) le CIC non-causal, à savoir, où le CTx a une connaissance non-causale du message du PTx [64]. En particulier, nous trouvons les régimes de paramètres où la coopération unilatérale HD ne donne pas d'avantages par rapport à l'IC non coopératif [12], et ceux où il atteint les limites de performances ultimes du CIC non-causal [64]. Fait intéressant, nous montrons que dans les régimes où le HD-CCIC Gaussien surpasse l'IC non coopératif le lien de coopération doit être capable de transmettre de manière fiable un débit supérieur à la somme des capacités de l'IC non coopératif correspondant. Pour chaque configuration, le gDoF est montré sur la Figure 7.8 (pour le canal symétrique), sur la Figure 7.9 (pour le canal Z) et sur la Figure 7.10 (pour le canal S).

4. Nous identifions enfin les régimes où une perte, en termes de gDoF, est subie en utilisant le mode de fonctionnement HD au CTx par rapport au cas FD analysé au Chapitre 4. Ces pertes pourraient motiver (dans ces régimes) l'utilisation d'un CTx plus cher avec des capacités FD dans les futurs réseaux sans fil.

Le travail présenté dans ce chapitre a donné lieu aux publications suivantes:

- [84] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "The symmetric sum-capacity of the Gaussian half-duplex causal cognitive interference channel to within a constant gap", in 2013 IEEE International Symposium on Information Theory (ISIT 2013), Istanbul (Turkey), July 2013.

- [85] M. Cardone, D. Tuninetti, R. Knopp, U. Salim, "On the Gaussian interference channel with half-duplex causal cognition", in IEEE Journal on Selected Areas in Communications, Volume 32, Issue n.11, November 2014, Pages 2177-2189.

Bibliography

[1] S. Brahma, C. Fragouli, and A. Özgür, "On the structure of approximately optimal schedules for half-duplex diamond networks," in *51st Annual Allerton Conference on Communication, Control, and Computing*, October 2013, pp. 1561–1566.

[2] LTE-A, *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (EUTRA).* 3GPP TR 36.806 V9.0.0, 2010.

[3] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, November 2010, pp. 1558–1562.

[4] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers*, November 2011, pp. 2002–2006.

[5] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, December 2012.

[6] A. S. Avestimehr, "Wireless network information flow: a deterministic approach," Ph.D. dissertation, EECS Department, University of California, Berkeley, October 2008. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-128.html

[7] A. Goldsmith, S. A. Jafar, I. Marić, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894 –914, May 2009.

[8] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388 –404, March 2000.

[9] L.-L. Xie and P. R. Kumar, "On the path-loss attenuation regime for positive cost and linear scaling of transport capacity in wireless networks," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2313–2328, June 2006.

[10] A. Özgür, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3549 –3572, October 2007.

[11] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1813 – 1827, May 2006.

[12] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534 –5562, December 2008.

[13] E. C. van der Meulen, "Three-terminal communication channel," *Advances in Applied Probability*, vol. 3, pp. 120–154, 1971.

[14] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572 – 584, September 1979.

[15] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulo-sum relay channels," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 921 –930, March 2009.

[16] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037 – 3063, September 2005.

[17] A. Host-Madsen, "On the capacity of wireless relaying," in *IEEE 56th Vehicular Technology Conference (VTC)*, vol. 3, September 2002, pp. 1333 – 1337.

[18] G. Kramer, "Models and theory for relay channels with receive constraints," in *42nd Annual Allerton Conference on Commununication, Control, and Computing*, October 2004, pp. 1312–1321.

[19] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872 –1905, April 2011.

[20] S. H. Lim, Y.-H. Kim, A. E. Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132 –3152, May 2011.

[21] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *IEEE International Symposium on Information Theory (ISIT)*, June 2000, p. 22.

[22] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 3008–3023, July 2008.

[23] L. Ong, S. J. Johnson, and C. M. Kellett, "On the capacity of the binary-symmetric parallel-relay network," *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 3, pp. 217–230, February 2014.

[24] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 765–779, March 2005.

[25] B. Chern and A. Özgür, "Achieving the capacity of the n-relay Gaussian diamond network within log(n) bits," in *IEEE Information Theory Workshop (ITW)*, September 2012, pp. 377–380.

[26] A. Sengupta, I.-H. Wang, and C. Fragouli, "Optimizing quantize-map-and-forward relaying for Gaussian diamond networks," in *IEEE Information Theory Workshop (ITW)*, September 2012, pp. 381–385.

[27] U. Niesen and S. N. Diggavi, "The approximate capacity of the Gaussian n-relay diamond network," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 845–859, February 2013.

[28] A. Özgür and S. N. Diggavi, "Approximately achieving Gaussian relay network capacity with lattice-based QMF codes," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8275–8294, December 2013.

[29] T. Lutz, C. Hausl, and R. Kotter, "Bits through deterministic relay cascades with half-duplex constraint," *IEEE Transactions on Information Theory*, vol. 58, no. 1, pp. 369–381, January 2012.

[30] T. Lutz, "Coding for half-duplex relay networks and the trapdoor channel," Ph.D. dissertation, Technische Universität Müchen (TUM), Munich, Germany, May 2014. [Online]. Available: http://mediatum.ub.tum.de/node?id=1231171

[31] H. Bagheri, A. S. Motahari, and A. K. Khandani, "On the capacity of the half-duplex diamond channel," in *IEEE International Symposium on Information Theory (ISIT)*, June 2010, pp. 649 –653.

[32] F. Xue and S. Sandhu, "Cooperation in a half-duplex Gaussian diamond relay channel," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3806 –3814, October 2007.

[33] S. Brahma, A. Özgür, and C. Fragouli, "Simple schedules for half-duplex networks," in *IEEE International Symposium on Information Theory (ISIT)*, July 2012, pp. 1112 –1116.

[34] S. Brahma and C. Fragouli, "Structure of optimal schedules in diamond networks," in *IEEE International Symposium on Information Theory (ISIT)*, June 2014, pp. 641–645.

[35] L. Ong, M. Motani, and S. J. Johnson, "On capacity and optimal scheduling for the half-duplex multiple-relay channel," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5770 –5784, September 2012.

[36] R. H. Etkin, F. Parvaresh, I. Shomorony, and A. S. Avestimehr, "Computing half-duplex schedules in Gaussian relay networks via min-cut approximations," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7204–7220, November 2014.

[37] A. Carleial, "A case where interference does not reduce capacity (corresp.)," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, September 1975.

[38] H. Sato, "On the capacity region of a discrete two-user channel for strong interference (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 377–379, May 1978.

[39] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49 – 60, January 1981.

[40] A. E. Gamal and M. Costa, "The capacity region of a class of deterministic interference channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 343–346, March 1982.

[41] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1522 –1544, April 2006.

[42] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 581–586, March 2004.

[43] S. Gelfand and M. Pinsker, "Coding for channel with random parameters," *Problems of control and information theory*, vol. 9, no. 1, pp. 19 – 31, 1980.

[44] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439 – 441, May 1983.

[45] D. Tuninetti, "An outer bound region for interference channels with generalized feedback," in *Information Theory and Applications Workshop (ITA)*, February 2010, pp. 1–5.

[46] ——, "An outer bound for the memoryless two-user interference channel with general cooperation," in *IEEE Information Theory Workshop (ITW)*, September 2012, pp. 217–221.

[47] V. M. Prabhakaran and P. Viswanath, "Interference channels with source cooperation," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 156 –186, January 2011.

[48] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *IEEE International Symposium on Information Theory (ISIT)*, June 2007, pp. 2871–2874.

[49] R. Tandon and S. Ulukus, "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4063 –4086, July 2011.

[50] A. Hekstra and F. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 44 –53, January 1989.

[51] S. Yang and D. Tuninetti, "Interference channel with generalized feedback (a.k.a. with source cooperation): Part i: Achievable region," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2686 –2710, May 2011.

[52] ——, "Interference channels with source cooperation in the strong cooperation regime: Symmetric capacity to within 2 bits/s/hz with dirty paper coding," in *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers*, November 2011, pp. 2140 –2144.

[53] Z. Wu and M. Vu, "Partial decode-forward binning schemes for the causal cognitive relay channels," *CoRR*, vol. abs/1111.3966, 2011.

[54] M. Mirmohseni, B. Akhbari, and M. Aref, "On the capacity of interference channel with causal and noncausal generalized feedback at the cognitive transmitter," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2813–2837, May 2012.

[55] R. Wu, "Interference channels with half-duplex source cooperation," Master's thesis, Electrical and Computer Engineering Department, University of Illinois at Urbana-Champaign, May 2011. [Online]. Available: http://hdl.handle.net/2142/24136

[56] C. Suh and D. N. C. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

[57] A. Sahai, V. Aggarwal, M. Yuksel, and A. Sabharwal, "Capacity of all nine models of channel output feedback for the two-user interference channel," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 6957–6979, November 2013.

[58] S.-Q. Le, R. Tandon, M. Motani, and H. Poor, "The capacity region of the symmetric linear deterministic interference channel with partial feedback," in *50th Annual Allerton Conference on Communication, Control, and Computing*, October 2012, pp. 1864–1871.

[59] ——, "Approximate capacity region for the symmetric Gaussian interference channel with noisy feedback," *submitted to IEEE Transactions on Information Theory*, December 2012.

[60] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4391–4399, November 2007.

[61] A. Jovičić and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3945–3958, September 2009.

[62] I. Marić, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3536–3548, October 2007.

[63] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the memoryless cognitive interference channel and some new capacity results," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4087–4109, July 2011.

[64] ——, "Inner and outer bounds for the Gaussian cognitive interference channel and new capacity results," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 820 –848, February 2012.

[65] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "Gaussian half-duplex relay channels: Generalized degrees of freedom and constant gap result," in *IEEE International Conference on Communications (ICC)*, June 2013, pp. 3355–3359.

[66] ——, "The capacity to within a constant gap of the Gaussian half-duplex relay channel," in *IEEE International Symposium on Information Theory (ISIT)*, July 2013, pp. 749–753.

[67] ——, "On the Gaussian half-duplex relay channel," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2542–2562, May 2014.

[68] R. Thomas, M. Cardone, R. Knopp, D. Tuninetti, and B. T. Maharaja, "An LTE implementation of a novel strategy for the Gaussian half-duplex relay channel," *to appear in IEEE International Conference on Communications (ICC) 2015.*

[69] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems.* SIAM, 2009.

[70] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, pp. 83–97, 1955.

[71] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "Gaussian half-duplex relay networks: Improved gap and a connection with the assignment problem," in *IEEE Information Theory Workshop (ITW)*, September 2013, pp. 1–5.

[72] M. Cardone, D. Tuninetti, and R. Knopp, "On user scheduling for maximum throughput in K-user MISO broadcast channels," *to appear in IEEE International Conference on Communications (ICC) 2015.*

[73] ——, "The approximate optimality of simple schedules for half-duplex multi-relay networks," *to appear in IEEE Information Theory Workshop (ITW) 2015.*

[74] ——, "Gaussian MIMO half-duplex relay networks: approximate optimality of simple schedules," *to appear in IEEE International Symposium on Information Theory (ISIT) 2015.*

[75] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "Gaussian half-duplex relay networks: Improved constant gap and connections with the assignment problem," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3559–3575, June 2014.

[76] M. Cardone, D. Tuninetti, and R. Knopp, "On the optimality of simple schedules for networks with multiple half-duplex relays," *submitted to IEEE Transactions on Information Theory*, December 2014.

[77] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "Approximate sum-capacity of full- and half-duplex asymmetric interference channels with unilateral source cooperation," in *Information Theory and Applications Workshop (ITA)*, February 2013, pp. 1–9.

[78] ——, "On the interference channel with causal cognition," in *IEEE International Conference on Communications (ICC)*, June 2013, pp. 3225–3229.

[79] ——, "On the Gaussian interference channel with unilateral generalized feedback," in *6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, May 2014, pp. 178–181.

[80] ——, "On the capacity of full-duplex causal cognitive interference channels to within a constant gap," in *IEEE International Conference on Communications (ICC)*, June 2014, pp. 1965–1970.

[81] ——, "New outer bounds for the interference channel with unilateral source cooperation," in *IEEE International Symposium on Information Theory (ISIT)*, June 2014, pp. 1426–1430.

[82] ——, "On the capacity of the two-user Gaussian causal cognitive interference channel," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2512–2541, May 2014.

[83] M. Cardone, D. Tuninetti, and R. Knopp, "The two-user causal cognitive interference channel: Novel outer bounds and constant gap result for the symmetric Gaussian noise channel in weak interference," *submitted to IEEE Transactions on Information Theory*, March 2014.

[84] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "The symmetric sum-capacity of the Gaussian half-duplex causal cognitive interference channel to within a constant gap," in *IEEE International Symposium on Information Theory (ISIT)*, July 2013, pp. 754–758.

[85] ——, "On the Gaussian interference channel with half-duplex causal cognition," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 11, pp. 2177–2189, November 2014.

[86] L. Zhang, J. Jiang, A. J. Goldsmith, and S. Cui, "Study of Gaussian relay channels with correlated noises," *IEEE Transactions on Communications*, vol. 59, no. 3, pp. 863 –876, March 2011.

[87] A. E. Gamal and Y.-H. Kim, *Network Information Theory.* Cambridge U.K.: Cambridge University Press, 2011.

[88] F. Bach, "Learning with sub modular functions: a convex optimization perspective," *Foundations and Trends®in Machine Learning*, vol. 6, no. 2-3, pp. 145–373, December 2013.

[89] V. Chvátal, *Linear Programming.* W.H.Freeman, 1983.

[90] L. Lovász and M. D. Plummer, *Matching Theory.* American Mathematical Society, 2009.

[91] P. Rost and G. Fettweis, "Achievable rates in Gaussian half-duplex multiple relay networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, November 2008, pp. 1–6.

[92] F. Parvaresh and R. Etkin, "Efficient capacity computation and power optimization for relay networks," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1782–1792, March 2014.

[93] I.-H. Wang and D. N. C. Tse, "Interference mitigation through limited transmitter cooperation," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2941–2965, May 2011.

[94] S. Kannan, A. Raja, and P. Viswanath, "Approximately optimal wireless broadcasting," *IEEE Transactions on Information Theory*, vol. 58, no. 12, pp. 7154–7167, December 2012.

[95] H. Sato, "An outer bound to the capacity region of broadcast channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 374–377, May 1978.

[96] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[97] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, October 2003.

[98] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, September 2004.

[99] T. Yoo and A. J. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, March 2006.

[100] J. G. Broida and S. G. Williamson, *A comprehensive introduction to linear algebra*. Addison-Wesley, 1989.

[101] J. M. Steele, *The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities*. Cambridge University Press, 2004.

[102] G. Kramer, "Directed information for channel with feedback," Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, 1998.

[103] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1345 – 1356, June 2004.

[104] G. Bresler and D. Tse, "The two-user Gaussian interference channel: a deterministic view," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 333–354, June 2008.