TECHNISCHE UNIVERSITÄT MÜNCHEN
Fakultät für Elektrotechnik und Informationstechnik
Fachgebiet Methoden der Signalverarbeitung

Capacity Bounds and Achievable Rates for the Gaussian MIMO Relay Channel

Lennart Sebastian Gerdes

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. Dr. rer. nat. Holger Boche
Prüfer der Dissertation:
1. Univ.-Prof. Dr.-Ing. Wolfgang Utschick
2. Univ.-Prof. Dr. sc. techn. Gerhard Kramer

Die Dissertation wurde am 03.08.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 20.06.2016 angenommen.

Abstract

This work considers the performance limits of the relay channel in terms of achievable rates. More specifically, capacity upper bounds and achievable rates for the Gaussian MIMO relay channel and the half-duplex constrained Gaussian MIMO relay channel are studied which are based on the cut-set bound as well as the decode-and-forward and partial decode-and-forward strategies, respectively. To this end, the corresponding rate maximization problems are formulated and, as far as possible, solved by means of efficient convex optimization techniques.

Acknowledgments

The completion of this work would not have been possible without the support, guidance, and encouragement of many people, some of whom deserve special mention here. First and foremost, I would like to thank and express my gratitude to my advisor Professor Wolfgang Utschick, who was instrumental in sparking my interest in signal processing and communication topics and who has supported me ever since I wrote my bachelor's thesis at his institute. He gave me the right amount of freedom to be independent in my research and he was always there to help or offer advice when I needed it. I am also grateful to Professor Gerhard Kramer for his interest in my work and for serving as the second examiner, and I would like to thank Professor Holger Boche for chairing my examination committee.

During my time as a research and teaching assistant at the Fachgebiet Methoden der Signalverarbeitung, I had the pleasure to work with Johannes Brehmer, Peter Breun, Andreas Dotzler, Andreas Gründinger, Stephan Günther, Christian Guthy, Christoph Hellings, Stephan Herrmann, Matthias Hotz, Raphael Hunger, Michael Joham, Daniel Kotzor, Alexander Krebs, Quan Kuang, David Neumann, Michael Newinger, Muhammad Danish Nisar, Max Riemensberger, David Schmidt, Fabian Steiner, Rainer Strobel, Lorenz Weiland, and Thomas Wiese. Not only did they create a friendly and creative working environment, they were always available for open-minded discussions and willing to share their expertise. Specifically, I am very grateful to Max Riemensberger for being a great office mate, for the ideas he contributed, for answering every single one of my numerous mathematical questions, as well as for proofreading this work. Furthermore, I would like to specifically thank Michael Joham for sharing his expert knowledge of MIMO systems and linear algebra and Christoph Hellings and Lorenz Weiland for their contributions to the results presented in this work.

Finally, I am deeply indebted to my family and, in particular, my wife Steffi for their past and continued love and support.

Lennart Gerdes
Munich, July 2015

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

A central aspect of today's and future wireless network standards is the question of how to satisfy the ever increasing demand of high-speed and high-quality service by a steadily growing number of mobile users without an increase of available bandwidth and power. One possibility to improve throughput, spectral efficiency, and reliability is to equip the communication devices with multiple antennas as it is well known that multi-antenna systems offer substantial gains over single-antenna systems. Another means to achieve above goals and to extend coverage is the use of *relays*, which support the communication between source(s) and destination(s), but do not have own information to transmit. In fact, both *multiple-input multiple-output* (MIMO) and relay techniques already play an essential role in current mobile communication systems such as WiMAX and LTE-Advanced [81, 143].

For example, one may improve throughput, spectral efficiency, energy efficiency, reliability, and/or coverage in cellular wireless communication systems by installing relays which can support the communication between base stations and mobile users. Compared to the complex and expensive base stations, such relays can be significantly cheaper since they need not be connected to the backhaul network, but only facilitate the wireless communication between mobile terminals and base stations. Nevertheless, they add (spatial) degrees of freedom to the network that may, among other things, be exploited to increase the diversity and/or the multiplexing gain, particularly if the mobile devices are also equipped with multiple antennas. Furthermore, relays may also facilitate the communication in wireless ad hoc networks and hierarchical wireless sensor networks, cf. [112] and [106], respectively, where the main goals usually are to extend coverage and to preserve the limited battery resources.

In this work, we study the performance limits (in terms of achievable rates) of the smallest relay network, the so-called *relay channel*, where two terminals communicate with the help of one relay. More specifically, we determine capacity upper bounds and achievable rates for the Gaussian MIMO relay channel and the half-duplex constrained

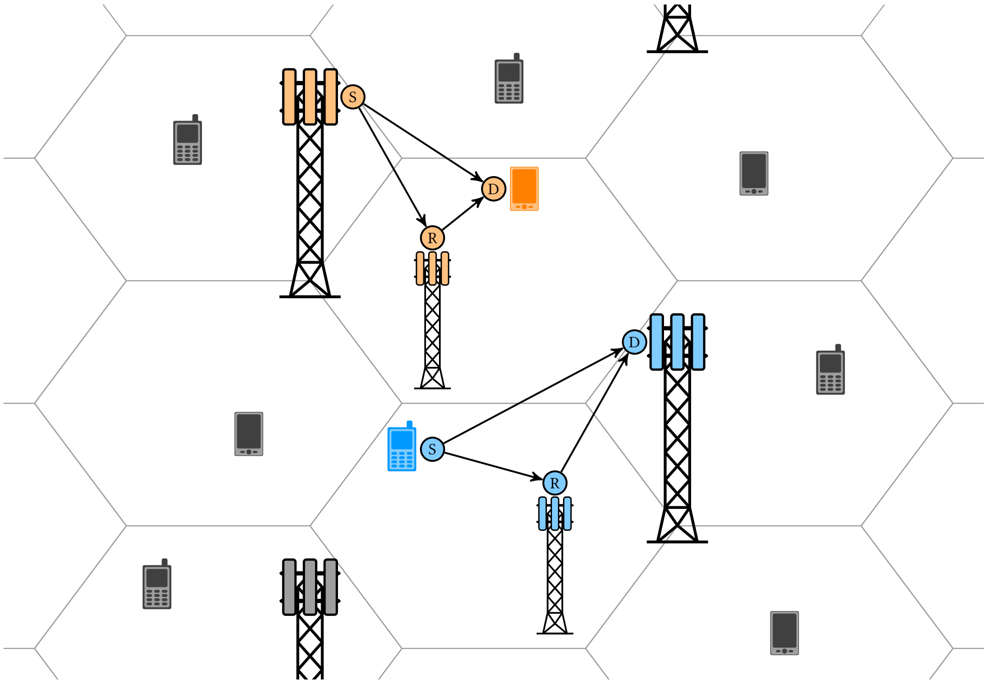


Figure 1.1: Relay-Aided Uplink/Downlink in Cellular Wireless Communication Systems

Gaussian MIMO relay channel, where all three nodes may be equipped with multiple antennas. As illustrated in Figure 1.1, the relay channel can for example be used to model relay-aided uplink and downlink scenarios in cellular wireless communication systems. In addition, the relay channel is an elementary building block of general multi-hop wireless networks. A fundamental understanding of the Gaussian MIMO relay channel and its performance limits may thus help to determine the limits on the performance of various wireless networks.

The reason we study capacity upper bounds and achievable rates is that in contrast to other (multi-user) Gaussian MIMO channels for which characterizations of the capacity (regions) have been found [9], the capacity of the (half-duplex) Gaussian MIMO relay channel is unknown in general. We remark that all capacity upper bounds considered in this work are based on the *cut-set bound* (CSB), which was originally derived in the pioneering work of Cover and El Gamal [21]. Furthermore, the achievable rates we consider are based on the *decode-and-forward* (DF) and *partial decode-and-forward* (PDF) strategies, which go back to Cover and El Gamal as well.

As the name suggests, the DF strategy requires the relay to decode the entire message transmitted by the source node. After being decoded at the relay, the message is then re-encoded and, in cooperation with the source, forwarded to the destination. The PDF strategy, on the other hand, allows to optimize the amount of information the relay must decode. It hence generalizes the DF strategy in that it provides the possibility to tradeoff sending information from the source to the destination via the relay (using DF principles) versus sending it over the direct (source-to-destination) link only. Consequently, the

PDF strategy seems well suited to the Gaussian MIMO relay channel, where the multi-antenna nodes provide spatial degrees of freedom that may be exploited. In fact, this leads to an interesting question we address in this work: Under which conditions on the channel gain matrices can the PDF strategy outperform the DF strategy and/or approach the CSB in the Gaussian MIMO relay channel?

Finally, we remark that throughout this work, perfect channel state information (CSI) is assumed to be available at all nodes. This essentially means that the results we present are applicable to slow fading environments where the channels remain constant long enough such that, first, the channel states can reliably be estimated at the receiving nodes and, second, the CSI can then be conveyed to the other nodes in an accurate and timely manner. Of course, perfect CSI will never be available in practical wireless communication systems, but the results we obtain under this assumption are nevertheless important as they specify the theoretical performance limits. In particular, knowing these limits for example allows to assess the performance degradation due to imperfect or outdated CSI.

1.2 Outline and Contributions

In addition to the introduction and a conclusion (Chapter 9), this work contains seven chapters, which are divided into two parts. In Part I, which consists of Chapters 2–4, we focus on unidirectional communication in the (full-duplex) relay channel, where the relay is able to simultaneously receive and transmit in the same frequency band. Part II, which comprises Chapters 5–8, then considers uni- and bidirectional communication in the half-duplex constrained relay channel, where in contrast to the first part, all nodes can either transmit or receive, but not simultaneously do both, in any frequency band.

Part I: The Relay Channel

In Chapter 2, we introduce the discrete memoryless (full-duplex) relay channel model, and we give a review of information theoretical results on the capacity of this relay channel in so far as they are important to this work. In particular, we discuss the CSB, which is the best known upper bound on the capacity of the relay channel, as well as the rates that can be achieved with the DF and PDF strategies. The chapter concludes with bibliographical notes and an overview of further noteworthy results on the discrete memoryless relay channel that are beyond the scope of this work.

In Chapter 3, we then apply the results described in the preceding chapter to the considered Gaussian MIMO relay channel, which is obtained by applying the *linear MIMO model* to the relay scenario of interest. Using the entropy maximizing property of the Gaussian distribution, we first establish that the CSB C_{CSB} and the maximum achievable DF rate R_{DF} are attained by jointly Gaussian source and relay inputs, and we show that if perfect CSI is available at all nodes, both C_{CSB} and R_{DF} can be evaluated as the solutions of convex optimization problems. Subsequently, we prove that the maximum achievable PDF rate R_{PDF} is attained by jointly Gaussian source and relay inputs as well, and we identify several classes of Gaussian MIMO relay channels for which R_{PDF} can also be determined as the solution of a convex optimization problem. For the general case, we present two different PDF schemes, which both yield suboptimal PDF rates that can be evaluated by standard convex optimization techniques. Finally, bibliographical notes and an overview of further noteworthy results on the Gaussian (MIMO) relay channel round off the chapter.

In Chapter 4, we provide numerical results for various Gaussian MIMO relay channels in which we compare the achievable DF and PDF rates to each other as well as to the CSB. More specifically, we investigate the effects of different antenna configurations and relay positions on the achievable rates and the CSB. As an example scenario, we choose a simple line network where the distance between the source and the destination is normalized to one and where the relay is positioned on the line connecting these two nodes. Moreover, we consider uncorrelated Rayleigh fading, and we assume that the channel gains are determined according to a simplified path loss model. The chapter, and hence the first part of this work, concludes with a discussion of possible practical applications of the results derived in Chapter 3.

Part II: The Relay Channel with Half-Duplex Constraint

The structure of the second part is similar to that of the first one. In Chapter 5, we first consider unidirectional communication in the half-duplex constrained relay channel, to which end we introduce the discrete memoryless half-duplex relay channel model where *time-division duplex* (TDD) is used to separate transmission and reception at the relay. Like for the full-duplex case, we review the information theoretical results on the capacity of this relay channel that are important to this work, i.e., we again discuss the CSB as well as the achievable DF and PDF rates. Subsequently, we extend the half-duplex relay channel model and the discussion of the corresponding information theoretical results to bidirectional communication. Finally, we provide an overview of further noteworthy results and bibliographical notes on uni- and bidirectional communication in the half-duplex constrained relay channel to round off the chapter.

In Chapter 6, we focus on unidirectional communication in the half-duplex constrained Gaussian MIMO relay channel with TDD and two different power constraints, a *per-phase* and an *average* power constraint. First, we establish that the CSB C_{CSB} and the maximum achievable DF rate R_{DF} are attained by jointly Gaussian source and relay inputs for both power constraints, and subsequently, we show that if perfect CSI is available at all nodes, C_{CSB} and R_{DF} can be determined as the solutions of convex optimization problems. In particular, we derive dual decomposition approaches that allow to efficiently solve the corresponding rate maximization problems with respect to the channel inputs and the time-shares of the relay receive and transmit phases in the Lagrangian dual domain. Furthermore, we prove that the maximum achievable PDF rate R_{PDF} is attained by Gaussian channel inputs as well, and we identify classes of half-duplex Gaussian MIMO relay channels for which the dual decomposition approaches can also be applied to evaluate R_{PDF}. For the general case, we again present two approaches that can be used to determine suboptimal PDF rates, and we conclude the chapter with an overview of further noteworthy results and bibliographical notes on the half-duplex Gaussian (MIMO) relay channel.

In Chapter 7, we consider the half-duplex *two-way* Gaussian MIMO relay channel with TDD where the bidirectional communication is *restricted* in the sense that the encoders at the terminals may neither cooperate, nor are they allowed to use previously decoded information to encode their messages. For both the per-phase and the average power constraint, we establish that the *cut-set outer bound* (CSOB) region $\mathcal{C}_{\text{CSB}}$ as well as the achievable DF and PDF rate regions $\mathcal{R}_{\text{DF}}$ and $\mathcal{R}_{\text{PDF}}$, respectively, are attained by Gaussian channel inputs. In addition, we derive convex parameterizations of $\mathcal{C}_{\text{CSB}}$ and $\mathcal{R}_{\text{DF}}$, we show how the dual decomposition approaches presented in Chapter 6 can thus be generalized to determine boundary points of $\mathcal{C}_{\text{CSB}}$ and $\mathcal{R}_{\text{DF}}$, and we provide sufficient conditions under which the same is true for $\mathcal{R}_{\text{PDF}}$. For the general case, we consider two suboptimal PDF rate regions which can also be evaluated by means of dual decomposition, and we again round off the chapter with bibliographical notes and a brief overview of further noteworthy results on bidirectional communication in the half-duplex constrained Gaussian (MIMO) relay channel.

In Chapter 8, we present numerical results for unidirectional communication in the half-duplex constrained Gaussian MIMO relay channel. In particular, we again compare the achievable DF and PDF rates to each other as well as to the CSB, where the focus is on the effects of different antenna configurations, relay positions, and the considered power constraints on the achievable rates and the CSB. The example scenario we choose is the same as for the full-duplex case, i.e., we again consider the aforementioned line network, uncorrelated Rayleigh fading, and the channel gains are determined according to a simplified path loss model. Finally, we discuss how the results that were derived in Chapters 6 and 7 may be helpful in designing practical wireless communication systems to conclude the chapter and the second part of this work.

1.3 Notation

Sets and Topology

$\mathbb{N}$	set of natural numbers
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of nonnegative real numbers
$\mathbb{R}^n, \mathbb{R}_+^n$	set of real/nonnegative real n-vectors
$\mathbb{R}^{m\times n}, \mathbb{R}_+^{m\times n}$	set of real/nonnegative real $m \times n$ matrices
$\mathbb{C}$	set of complex numbers
$\mathbb{C}^n$	set of complex n-vectors
$\mathbb{C}^{m\times n}$	set of complex $m \times n$ matrices

$\emptyset$	empty set
$\mathcal{S}$	an abstract set
$\mathcal{S}^{\mathrm{c}}$	complement of set $\mathcal{S}$
$\lvert\mathcal{S}\rvert$	cardinality of a finite set $\mathcal{S}$
$\mathrm{conv}(\mathcal{S})$	convex hull of set $\mathcal{S}$
$\mathrm{int}(\mathcal{S})$	interior of set $\mathcal{S}$
$\mathcal{S}_1 \subseteq \mathcal{S}_2$	$\mathcal{S}_1$ is a subset of $\mathcal{S}_2$
$\mathcal{S}_1 \subsetneq \mathcal{S}_2$	$\mathcal{S}_1$ is a proper subset of $\mathcal{S}_2$
$\mathcal{V}^{\perp}$	orthogonal complement of $\mathcal{V}$
$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2$	$\mathcal{V}$ is the Minkowski sum of $\mathcal{V}_1$ and $\mathcal{V}_2$, i.e., $\mathcal{V} = \{v = v_1 + v_2 : v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2\}$
$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$	the vector space $\mathcal{V}$ is the internal direct sum of $\mathcal{V}_1$ and $\mathcal{V}_2$, i.e., $\mathcal{V}_1 + \mathcal{V}_2 = \mathcal{V}$ and $\dim(\mathcal{V}_1 \cap \mathcal{V}_2) = 0$
$\mathcal{V} = \mathcal{V}_1 \perp \mathcal{V}_2$	the vector space $\mathcal{V}$ is the orthogonal sum of $\mathcal{V}_1$ and $\mathcal{V}_2$, i.e., $\mathcal{V}_1 \oplus \mathcal{V}_2 = \mathcal{V}$ and $\mathcal{V}_1, \mathcal{V}_2$ are orthogonal subspaces

Vectors and Matrices

$\mathbf{0}$	all-zeros vector/matrix
$\mathbf{1}$	all-ones vector/matrix
$\mathbf{I}$	identity matrix (dimensions indicated by subscripts if necessary)
$\boldsymbol{a}$	a real- or complex-valued column vector
$\boldsymbol{A}$	a real- or complex-valued matrix
$[\boldsymbol{A}]_{k,\ell}$	k-th element in the ℓ-th column of matrix $\boldsymbol{A}$
$\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{A}^{\mathrm{T}}$	transpose of vector $\boldsymbol{a}$/matrix $\boldsymbol{A}$
$\boldsymbol{a}^{\mathrm{H}}, \boldsymbol{A}^{\mathrm{H}}$	Hermitian (complex conjugate transpose) of vector $\boldsymbol{a}$/matrix $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	inverse of a square nonsingular matrix $\boldsymbol{A}$
$\boldsymbol{A}^{+}$	Moore–Penrose pseudoinverse of matrix $\boldsymbol{A}$
$\det(\boldsymbol{A})$	determinant of a square matrix $\boldsymbol{A}$
$\mathrm{null}(\boldsymbol{A})$	kernel (null space) of matrix $\boldsymbol{A}$
$\mathrm{range}(\boldsymbol{A})$	range (column space) of matrix $\boldsymbol{A}$
$\mathrm{rank}(\boldsymbol{A})$	rank of matrix $\boldsymbol{A}$
$\mathrm{row}(\boldsymbol{A})$	row space of matrix $\boldsymbol{A}$
$\mathrm{tr}(\boldsymbol{A})$	trace of a square matrix $\boldsymbol{A}$

Generalized Inequalities

$\boldsymbol{a} \geq \boldsymbol{b}$	componentwise inequality between real-valued vectors $\boldsymbol{a}, \boldsymbol{b}$
$\boldsymbol{a} > \boldsymbol{b}$	strict componentwise inequality between real-valued vectors $\boldsymbol{a}, \boldsymbol{b}$
$\boldsymbol{A} \succcurlyeq \boldsymbol{B}$	$\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite (nonnegative definite)
$\boldsymbol{A} \succ \boldsymbol{B}$	$\boldsymbol{A} - \boldsymbol{B}$ is positive definite

Random Variables and Information Theory

X, Y, Z	discrete random variables
x, y, z	realizations of random variables X, Y, Z
X^n, x^n	length-n sequences: $X^n = (X_1, \ldots, X_n), x^n = (x_1, \ldots, x_n)$
$\mathrm{E}[X]$	expectation of X
$H(X)$	entropy of X
$H(X, Y)$	joint entropy of X and Y
$H(X\|Z)$	conditional entropy of X given Z
$I(X; Y)$	mutual information of X and Y
$I(X; Y\|Z)$	conditional mutual information of X and Y given Z
$X \leftrightarrow Y \leftrightarrow Z$	X, Y, Z form a Markov chain (used in chapters discussing information theoretical results)
$\mathcal{N}_{\mathbb{C}}(\boldsymbol{m}, \boldsymbol{C})$	proper complex Gaussian distribution with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{C}$
$\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$	continuous random vectors or realizations of corresponding random vectors (same notation since bold uppercase letters denote matrices)
$\boldsymbol{x}^n$	length-n sequence: $\boldsymbol{x}^n = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$
$\mathrm{E}[\boldsymbol{x}]$	expectation of $\boldsymbol{x}$
$h(\boldsymbol{x})$	differential entropy of $\boldsymbol{x}$
$h(\boldsymbol{x}, \boldsymbol{y})$	joint differential entropy of $\boldsymbol{x}$ and $\boldsymbol{y}$
$h(\boldsymbol{x}\|\boldsymbol{z})$	conditional differential entropy of $\boldsymbol{x}$ given $\boldsymbol{z}$
$I(\boldsymbol{x}; \boldsymbol{y})$	mutual information of $\boldsymbol{x}$ and $\boldsymbol{y}$
$I(\boldsymbol{x}; \boldsymbol{y}\|\boldsymbol{z})$	conditional mutual information of $\boldsymbol{x}$ and $\boldsymbol{y}$ given $\boldsymbol{z}$
$\boldsymbol{x} \leftrightarrow \boldsymbol{y} \leftrightarrow \boldsymbol{z}$	$\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ form a Markov chain (used in chapters discussing Gaussian MIMO relay channels)

Part I

The Relay Channel

Chapter 2

Information Theoretical Results

The relay channel we consider in this work is a three-terminal network in which one source terminal S transmits information to one destination terminal D with the help of a single relay node R. It is assumed that the relay has no own information to transmit or receive so that its only purpose is to assist the communication from the source to the destination. The type of relaying considered here is therefore also referred to as *supportive* relaying – as opposed to *cooperative* relaying, where nodes mutually help each other in delivering data to the destined sink terminals [28, Chapter 1].

For this system model, our main interest is to investigate how much the relay can improve the performance of the data transfer from the source to the destination in terms of achievable rates. To this end, we first introduce an information theoretical model for the discrete memoryless relay channel in Section 2.1. Subsequently, we give a review of information theoretical results on the capacity of this relay channel in so far as they are important to this work. In particular, we discuss the *cut-set bound* (CSB), which is the best known upper bound on the capacity of the general relay channel, in Section 2.2 as well as the achievable *decode-and-forward* (DF) and *partial decode-and-forward* (PDF) rates in Sections 2.3 and 2.4, respectively. The chapter concludes with bibliographical notes and an overview of further noteworthy results on the relay channel which are beyond the scope of this work in Section 2.5.

2.1 Channel Model

The discrete memoryless relay channel, specified by

$$\left\{\mathcal{X}_\mathrm{S} \times \mathcal{X}_\mathrm{R}, p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R}), \mathcal{Y}_\mathrm{D} \times \mathcal{Y}_\mathrm{R}\right\}, \tag{2.1}$$

consists of four finite sets $\mathcal{X}_\mathrm{S}$, $\mathcal{X}_\mathrm{R}$, $\mathcal{Y}_\mathrm{D}$, $\mathcal{Y}_\mathrm{R}$ and a collection of probability mass functions $p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R})$ on $\mathcal{Y}_\mathrm{D} \times \mathcal{Y}_\mathrm{R}$, one for each $(x_\mathrm{S}, x_\mathrm{R}) \in \mathcal{X}_\mathrm{S} \times \mathcal{X}_\mathrm{R}$.[1] The interpretation is that x_S is the channel input of the source, y_D is the channel output of the destination,

[1]Throughout this work, we shall drop the subscripts of probability mass or density functions, e.g., $p_{X,Y|Z}(x, y|z) = p(x, y|z)$, as they can be inferred by inspection of the arguments.

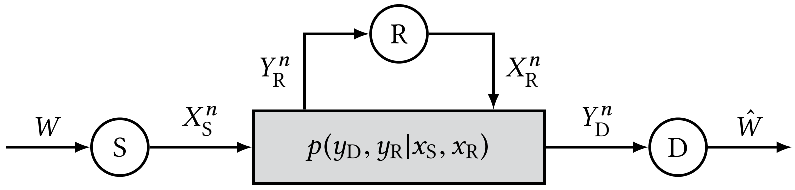


Figure 2.1: Illustration of the Relay Channel

while y_R and x_R denote the relay's observation (channel output) and channel input, respectively, as illustrated in Figure 2.1.

A $(2^{nR}, n)$ code for the relay channel consists of a message set $\mathcal{W} = \{1, 2, \ldots, \lceil 2^{nR} \rceil\}$, a source encoder that assigns a codeword $X_S^n(w) \in \mathcal{X}_S^n$ to each $w \in \mathcal{W}$, a relay encoder that assigns a symbol $X_{R,i}(y_R^{i-1}) \in \mathcal{X}_R$ to each past received sequence $y_R^{i-1} \in \mathcal{Y}_R^{i-1}$ for each time $i \in \{1, \ldots, n\}$, and a decoder (at the destination) that assigns an estimate $\hat{W}(y_D^n) \in \mathcal{W}$ (or possibly an error message) to each received sequence $y_D^n \in \mathcal{Y}_D^n$. For generality, the encoding and decoding functions are allowed to be stochastic.

Note that the definition of the relay encoder includes the condition that the relay input symbols $X_{R,i}$ may depend on the past observations $Y_R^{i-1} = (Y_{R,1}, \ldots, Y_{R,i-1})$ only. This requirement ensures that the relay operates in a *causal* fashion (strictly causal in the terminology of [34, Chapter 16]).

The relay channel is *memoryless* in the sense that the current channel outputs $(Y_{D,i}, Y_{R,i})$ depend on all previous channel inputs (X_S^i, X_R^i) only through the current channel inputs $(X_{S,i}, X_{R,i})$. For any $p(w)$ and choice of the code, the joint probability mass function on $\mathcal{W} \times \mathcal{X}_S^n \times \mathcal{X}_R^n \times \mathcal{Y}_D^n \times \mathcal{Y}_R^n$ hence factors as

$$p(w, x_S^n, x_R^n, y_D^n, y_R^n) = p(w)p(x_S^n|w) \prod_{i=1}^{n} p(x_{R,i}|y_R^{i-1})p(y_{D,i}, y_{R,i}|x_{S,i}, x_{R,i}). \quad (2.2)$$

Moreover, the transmissions are modeled as taking place *synchronously*. More specifically, it is assumed that there is a central clock that governs the operation of all three terminals in the following way [76, Section 9.1]:

- The clock ticks n times, once for each channel use.
- The source and the relay apply their respective inputs $X_{S,i}$ and $X_{R,i}$ to the channel after clock tick $i - 1$ and before clock tick i.
- The destination and the relay see their respective channel outputs $Y_{D,i}$ and $Y_{R,i}$ exactly at clock tick i.

Consequently, there are small delays between transmission and reception, which ensures that the whole network operates in a causal manner.

If W is uniformly distributed over $\mathcal{W}$ and $P_e^{(n)} = \Pr[\hat{W} \neq W]$ denotes the average probability of error, a rate R is said to be *achievable* if there exists a sequence of $(2^{nR}, n)$ codes for which $P_e^{(n)} \to 0$ as $n \to \infty$. Finally, the *capacity* C of the relay channel is defined as the supremum of the set of achievable rates.

Remark 2.1. This is the standard definition of the discrete memoryless relay channel, which was (sometimes with minor differences, e.g., with a restriction to deterministic encoding and decoding functions) already used in the earliest works on the relay channel, cf. [4, 21, 107, 130], and which can today be found in various textbooks on (network) information theory, cf. [24, 34, 76].

Unfortunately, the capacity of the relay channel is still unknown for the general case. However, several upper and lower bounds (achievable rates) on the capacity have been derived, and for some special cases, including the four types of relay channels defined below, these bounds are tight.

Definition 2.1. The relay channel $\left\{\mathcal{X}_{\mathrm{S}} \times \mathcal{X}_{\mathrm{R}}, p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}}), \mathcal{Y}_{\mathrm{D}} \times \mathcal{Y}_{\mathrm{R}}\right\}$

(a) is *degraded* if $p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}})$ can be written in the form

$$p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}}) = p(y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}})p(y_{\mathrm{D}}|y_{\mathrm{R}}, x_{\mathrm{R}}), \tag{2.3}$$

or equivalently, if $X_{\mathrm{S}} \leftrightarrow (X_{\mathrm{R}}, Y_{\mathrm{R}}) \leftrightarrow Y_{\mathrm{D}}$ form a Markov chain;[2]

(b) is *reversely degraded* if $p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}})$ can be written in the form

$$p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}}) = p(y_{\mathrm{D}}|x_{\mathrm{S}}, x_{\mathrm{R}})p(y_{\mathrm{R}}|y_{\mathrm{D}}, x_{\mathrm{R}}), \tag{2.4}$$

or equivalently, if $X_{\mathrm{S}} \leftrightarrow (X_{\mathrm{R}}, Y_{\mathrm{D}}) \leftrightarrow Y_{\mathrm{R}}$ form a Markov chain;

(c) is *semideterministic* if Y_{R} is a function of $(X_{\mathrm{S}}, X_{\mathrm{R}})$, i.e., if $Y_{\mathrm{R}} = y_{\mathrm{R}}(X_{\mathrm{S}}, X_{\mathrm{R}})$;

(d) has *orthogonal sender components* if $\mathcal{X}_{\mathrm{S}} = \mathcal{X}'_{\mathrm{S}} \times \mathcal{X}''_{\mathrm{S}}$ and $p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}})$ can be written in the form

$$p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}}) = p(y_{\mathrm{R}}|x'_{\mathrm{S}}, x_{\mathrm{R}})p(y_{\mathrm{D}}|x''_{\mathrm{S}}, x_{\mathrm{R}}) \tag{2.5}$$

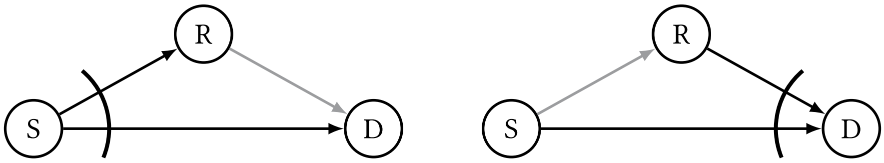
for all $(x'_{\mathrm{S}}, x''_{\mathrm{S}}, x_{\mathrm{R}}, y_{\mathrm{D}}, y_{\mathrm{R}}) \in \mathcal{X}'_{\mathrm{S}} \times \mathcal{X}''_{\mathrm{S}} \times \mathcal{X}_{\mathrm{R}} \times \mathcal{Y}_{\mathrm{D}} \times \mathcal{Y}_{\mathrm{R}}$.

We remark that these four particular relay channels are introduced here because their capacities are achieved by the DF and/or the PDF strategies, which we focus on in this work, and because the converses follow from the CSB. More details on these capacity results are provided in the following sections.

2.2 Cut-Set Bound

As the capacity of the relay channel has not yet been determined for the general case, we can usually only study capacity bounds. The best known upper bound on the capacity of the relay channel is the so-called cut-set bound (CSB). It was originally derived in the pioneering work of Cover and El Gamal [21] and reads as follows:

[2] Note that this type of degradedness is also known as physical degradedness in information theory. Without further specification, the term "degraded" thus means physically degraded throughout this work.



(a) Cut bounding the rate of information transmitted by the source ($\mathcal{S} = \{\mathrm{S}\}$)

(b) Cut bounding the rate of information received by the destination ($\mathcal{S} = \{\mathrm{S}, \mathrm{R}\}$)

Figure 2.2: Illustration of the CSB for the Relay Channel

THEOREM 2.1. *For any relay channel* $\left\{\mathcal{X}_\mathrm{S} \times \mathcal{X}_\mathrm{R}, p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R}), \mathcal{Y}_\mathrm{D} \times \mathcal{Y}_\mathrm{R}\right\}$*, the capacity is upper bounded by*

$$C \le C_\mathrm{CSB} = \max_{p(x_\mathrm{S}, x_\mathrm{R})} \min\left\{I(X_\mathrm{S}; Y_\mathrm{R}, Y_\mathrm{D}|X_\mathrm{R}), I(X_\mathrm{S}, X_\mathrm{R}; Y_\mathrm{D})\right\}. \tag{2.6}$$

Proof. See [21], [24, Theorem 15.7.1], or [34, Theorem 16.1], for example. □

This bound has a nice *max-flow min-cut* interpretation, cf. [24, Section 15.7] or [34, Section 16.2], which is illustrated in Figure 2.2. The first term in (2.6) implies that the source cannot transmit information to the destination at a higher rate than if the relay and the destination cooperate in receiving, cf. Figure 2.2(a). Likewise, the destination cannot receive information at a higher rate than if the source and the relay cooperate in transmitting, cf. Figure 2.2(b), which yields the second term in (2.6).

From today's perspective, the CSB for the relay channel can be viewed as a special case of a more general max-flow min-cut theorem for general multiterminal networks, cf. [24, Section 15.10]. In particular, suppose we have a network that is comprised of a finite node set $\mathcal{N}$, and assume that node A sends information to node B at rate $R_{\mathrm{A},\mathrm{B}}$. If $\mathcal{S} \subseteq \mathcal{N}$ and $\mathcal{S}^\mathrm{c}$ is the complement of $\mathcal{S}$ in $\mathcal{N}$, a *cut* that separates nodes A and B is a partition $(\mathcal{S}, \mathcal{S}^\mathrm{c})$ of $\mathcal{N}$ such that $\mathrm{A} \in \mathcal{S}$ and $\mathrm{B} \in \mathcal{S}^\mathrm{c}$. Letting $X_\mathcal{S} = \{X_\mathrm{A} : \mathrm{A} \in \mathcal{S}\}$ and $Y_{\mathcal{S}^\mathrm{c}} = \{Y_\mathrm{B} : \mathrm{B} \in \mathcal{S}^\mathrm{c}\}$, the following theorem has become an important tool for finding outer bounds on capacity regions in network information theory:

THEOREM 2.2. *If the information rates* $\{R_{\mathrm{A},\mathrm{B}}\}$ *are achievable, there exists some joint probability distribution* $p(x_\mathcal{N})$ *such that*

$$\sum_{\mathrm{A}\in\mathcal{S},\mathrm{B}\in\mathcal{S}^\mathrm{c}} R_{\mathrm{A},\mathrm{B}} \le I(X_\mathcal{S}; Y_{\mathcal{S}^\mathrm{c}}|X_{\mathcal{S}^\mathrm{c}}) \tag{2.7}$$

for all $\mathcal{S} \subseteq \mathcal{N}$*.*

Proof. See [24, Theorem 15.10.1], for example. □

For the relay channel, we have $\mathcal{N} = \{\mathrm{S}, \mathrm{R}, \mathrm{D}\}$ and there is only one rate $R = R_{\mathrm{S},\mathrm{D}}$. Consequently, an upper bound on any achievable rate R is obtained by applying (2.7) to the two cuts specified by $\mathcal{S} = \{\mathrm{S}\}$ and $\mathcal{S} = \{\mathrm{S}, \mathrm{R}\}$, cf. Figure 2.2.

REMARK 2.2. The conditional mutual information terms $I(X_{\mathcal{S}}; Y_{\mathcal{S}^c}|X_{\mathcal{S}^c})$, $\mathcal{S} \subseteq \mathcal{N}$, that upper bound the achievable rates in (2.7) are concave in $p(x_{\mathcal{N}})$, cf. [78, Section 3.7.1].[3] Since the set of joint probability distributions is convex and the intersection (pointwise minimum) of concave functions is concave [11, Section 3.2.3], finding a boundary point of the outer bound region specified by Theorem 2.2 is a convex optimization problem. In particular, this also means that the maximization over $p(x_S, x_R)$ in (2.6) that yields C_{CSB} is a convex optimization problem.

As previously mentioned, the CSB is the best known general upper bound on the capacity of the relay channel, and it is tight in several special cases for which the capacity is known. However, there also exist relay channels for which the CSB has been proved to be loose, cf. Section 2.5.

2.3 Decode-and-Forward

Lower bounds on the capacity of the relay channel are obtained by designing relay protocols/strategies and codes to show that certain rates are achievable, cf. [78, Section 3.7.1]. One approach that suggests itself is to let the relay decode the entire information transmitted by the source and, after re-encoding, forward it to the destination. Due to the decoding and encoding operations performed at the relay, this strategy, which is now commonly referred to as decode-and-forward (DF), belongs to the class of *regenerative* relay protocols, cf. [28, Section 1.6.1]. Like the CSB, the DF strategy was derived by Cover and El Gamal [21], and it gives the following lower bound on the capacity:

THEOREM 2.3. *The capacity of the relay channel* $\left\{\mathcal{X}_S \times \mathcal{X}_R, p(y_D, y_R|x_S, x_R), \mathcal{Y}_D \times \mathcal{Y}_R\right\}$ *is lower bounded by* $C \geq R_{DF}$, *where*

$$R_{DF} = \max_{p(x_S, x_R)} \min\left\{I(X_S; Y_R|X_R), I(X_S, X_R; Y_D)\right\} \tag{2.8}$$

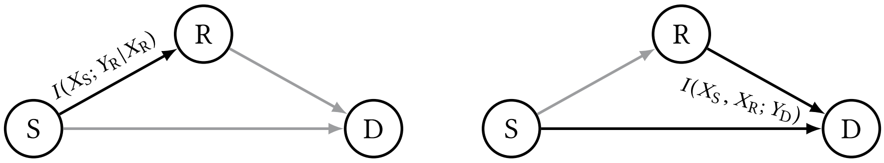
denotes the maximum rate that can be achieved with the decode-and-forward (DF) protocol.

Proof. See [21] or [34, Theorem 16.2], for example. □

The achievable DF rate is limited by two mutual information terms. The first rate bound, $R_{DF} \leq I(X_S; Y_R|X_R)$, is due to the condition that the relay must decode the entire information transmitted by the source, whereas the second one, $R_{DF} \leq I(X_S, X_R; Y_D)$, states that the achievable rate cannot be higher than the combined information transfer from the source and the relay to the destination. An illustration of these two rate bounds is provided in Figure 2.3.

Note that R_{DF} differs from C_{CSB} only in the first mutual information term inside the minimum, where an additional Y_D appears in (2.6) as compared to (2.8). Since the second

[3] In fact, this property is implicitly shown in the proof of [24, Theorem 15.10.1].



(a) $R_{DF} \leq I(X_S; Y_R|X_R)$ (b) $R_{DF} \leq I(X_S, X_R; Y_D)$

Figure 2.3: Illustration of the Rate Bounds for the DF Strategy

term is the same and R_{DF} is also obtained by maximizing over all joint distributions of (X_S, X_R), we can conclude that DF achieves capacity if $I(X_S, X_R; Y_D) \leq I(X_S; Y_R|X_R)$ for all $p(x_S, x_R)$. This is a very stringent condition though, which can only be satisfied if the source-to-relay link is better than both the source-to-destination and relay-to-destination links.

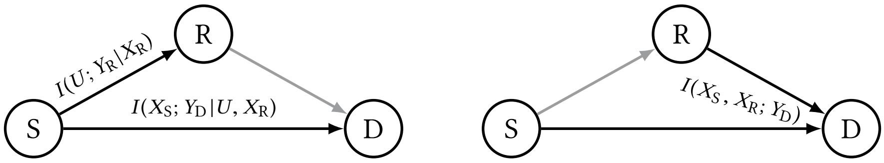
Generally speaking, however, the DF strategy performs quite well when the source-to-relay link is (much) better than the source-to-destination link. One such example is the degraded relay channel, for which DF actually achieves the capacity [21]. This result follows from the fact that $I(X_S; Y_R, Y_D|X_R) = I(X_S; Y_R|X_R)$ if the relay channel is degraded as $X_S \leftrightarrow (X_R, Y_R) \leftrightarrow Y_D$ form a Markov chain in this case.

An interesting and nonobvious property of the DF strategy is that one obtains R_{DF} in (2.8) by maximizing over all joint distributions of (X_S, X_R). This can be explained by having a closer look at the *block Markov superposition encoding* scheme that achieves R_{DF}, cf. [21] or [34, Section 16.4.3], for example. The coding scheme is based on conveying $B - 1$ (independent) messages $W_b \in \mathcal{W} = \{1, 2, \ldots, \lceil 2^{nR} \rceil\}$, $b \in \{1, \ldots, B-1\}$, from the source to the destination in B blocks of transmission, each consisting of n channel uses. As $B \to \infty$ for a fixed n, the rate $\frac{B-1}{B}R$ then is arbitrarily close to R.

The source uses an encoder that has one block memory in the sense that the source codeword in block b depends on w_b as well as w_{b-1}, i.e., in block b, the encoder assigns a codeword $X^n_{S,b}(w_{b-1}, w_b)$ to each tuple $(w_{b-1}, w_b) \in \mathcal{W} \times \mathcal{W}$ (with $w_0 = w_B = 1$). Assuming that the relay has correctly decoded w_b after block b, which is justified if the first rate bound is satisfied, both the source and the relay know w_b. If the relay encoder is deterministic, the source can determine the relay's codeword $x^n_{R,b+1}(w_b)$ sent in block $b + 1$ and generate its own codeword $X^n_{S,b+1}(w_b, w_{b+1})$ from $x^n_{R,b}(w_b)$ via superposition coding so that input distributions of the form $p(x_S|x_R)p(x_R) = p(x_S, x_R)$ are permissible. As a consequence, Remark 2.2 applies here accordingly. For more details on the DF coding scheme, the reader is referred to the above cited references.

2.4 Partial Decode-and-Forward

The DF strategy's requirement that the relay decode the entire information sent by the source is a severe limitation if the source-to-relay link is weak compared to the source-to-destination link. In particular, it can happen that DF is worse than direct transmission, i.e., R_{DF} may be smaller than the maximum of the rates that can be achieved if the relay



(a) $R_{\text{PDF}} \le I(U; Y_{\text{R}}|X_{\text{R}}) + I(X_{\text{S}}; Y_{\text{D}}|U, X_{\text{R}})$ (b) $R_{\text{PDF}} \le I(X_{\text{S}}, X_{\text{R}}; Y_{\text{D}})$

Figure 2.4: Illustration of the Rate Bounds for the PDF Strategy

does not forward any information to the destination, which is given by

$$R_{\text{P2P}} = \max_{x_{\text{R}} \in \mathcal{X}_{\text{R}}} \max_{p(x_{\text{S}})} I(X_{\text{S}}; Y_{\text{D}}|x_{\text{R}}). \tag{2.9}$$

Note that R_{P2P} is achieved when the relay "opens" the source-to-destination channel by transmitting the best constant symbol $x_{\text{R}} \in \mathcal{X}_{\text{R}}$, cf. [21].[4]

Without considering a completely different relay protocol, this problem can be overcome by allowing the relay to *partially* decode the source message. For this purpose, the source message W that is to be conveyed to the destination is split into two independent parts W' and W'' of which the relay is only required to decode W'. By constructing separate codebooks for W' and W'' and using superposition coding at the source, the following lower bound on the capacity of the relay channel can be obtained:

Theorem 2.4. *The capacity of the relay channel* $\left\{\mathcal{X}_{\text{S}} \times \mathcal{X}_{\text{R}}, p(y_{\text{D}}, y_{\text{R}}|x_{\text{S}}, x_{\text{R}}), \mathcal{Y}_{\text{D}} \times \mathcal{Y}_{\text{R}}\right\}$ *is lower bounded by* $C \ge R_{\text{PDF}}$, *where*

$$R_{\text{PDF}} = \max_{p(u, x_{\text{S}}, x_{\text{R}})} \min\left\{I(U; Y_{\text{R}}|X_{\text{R}}) + I(X_{\text{S}}; Y_{\text{D}}|U, X_{\text{R}}), I(X_{\text{S}}, X_{\text{R}}; Y_{\text{D}})\right\} \tag{2.10}$$

denotes the maximum rate that can be achieved with the partial decode-and-forward (PDF) protocol, and where the maximization over $p(u, x_{\text{S}}, x_{\text{R}})$ *is subject to the constraint that* $U \leftrightarrow (X_{\text{S}}, X_{\text{R}}) \leftrightarrow (Y_{\text{D}}, Y_{\text{R}})$ *form a Markov chain.*

Proof. See [21] or [34, Theorem 16.3], for example. □

Here, U is an auxiliary random variable which represents the part of the information the relay must decode. By optimizing over (the distribution of) U, the PDF strategy hence allows to tradeoff sending information from the source to the destination via the relay versus sending it over the direct link. This tradeoff is also reflected in the first mutual information term that upper bounds R_{PDF} in (2.10), cf. Figure 2.4(a). The second term, on the other hand, is the same as for the CSB and the DF rate in (2.6) and (2.8), respectively. It is again given by the rate of the information the source and the relay can cooperatively send to the destination, cf. Figure 2.4(b).

[4] We denote this rate as R_{P2P} because if $p(y_{\text{D}}|x_{\text{S}}, a) = p(y_{\text{D}}|x_{\text{S}}, b)$ for all $a, b \in \mathcal{X}_{\text{R}}$, or if $\mathcal{X}_{\text{R}}$ includes a "quiet" symbol that opens the channel (as is the case for the Gaussian relay channel, cf. Chapter 3), then R_{P2P} is indeed the *point-to-point* (P2P) rate from source to destination as if the relay were not present.

It is easily verified that the PDF strategy includes both the DF strategy and direct transmission as special cases. In particular, if we choose $U = X_\mathrm{S}$ in (2.10), it follows that $I(U; Y_\mathrm{R}|X_\mathrm{R}) + I(X_\mathrm{S}; Y_\mathrm{D}|U, X_\mathrm{R}) = I(X_\mathrm{S}; Y_\mathrm{R}|X_\mathrm{R})$ so that PDF reduces to DF. Similarly, setting $U = \emptyset$ results in $I(U; Y_\mathrm{R}|X_\mathrm{R}) + I(X_\mathrm{S}; Y_\mathrm{D}|U, X_\mathrm{R}) = I(X_\mathrm{S}; Y_\mathrm{D}|X_\mathrm{R}) \le I(X_\mathrm{S}, X_\mathrm{R}; Y_\mathrm{D})$, meaning that PDF reduces to direct transmission. Consequently, we have

$$R_\mathrm{PDF} \ge \max\{R_\mathrm{DF}, R_\mathrm{P2P}\}. \tag{2.11}$$

Furthermore, we remark that PDF is the capacity achieving strategy for several types of relay channels, including the four defined in Section 2.1:

(a) The capacity of the degraded relay channel is given by [21]

$$C = \max_{p(x_\mathrm{S}, x_\mathrm{R})} \min \Big\{ I(X_\mathrm{S}; Y_\mathrm{R}|X_\mathrm{R}), I(X_\mathrm{S}, X_\mathrm{R}; Y_\mathrm{D}) \Big\}, \tag{2.12}$$

and it is achieved by the DF strategy as discussed in the previous section.

(b) If the relay channel is reversely degraded, on the other hand, the capacity is achieved by direct transmission [21], i.e.,

$$C = \max_{x_\mathrm{R} \in \mathcal{X}_\mathrm{R}} \max_{p(x_\mathrm{S})} I(X_\mathrm{S}; Y_\mathrm{D}|x_\mathrm{R}). \tag{2.13}$$

The converse follows from the fact that $I(X_\mathrm{S}; Y_\mathrm{R}, Y_\mathrm{D}|X_\mathrm{R}) = I(X_\mathrm{S}; Y_\mathrm{D}|X_\mathrm{R})$ if the relay channel is reversely degraded since $X_\mathrm{S} \leftrightarrow (X_\mathrm{R}, Y_\mathrm{D}) \leftrightarrow Y_\mathrm{R}$ form a Markov chain in this case.

(c) The capacity of the semideterministic relay channel is equal to [32]

$$C = \max_{p(x_\mathrm{S}, x_\mathrm{R})} \min \Big\{ H(Y_\mathrm{R}|X_\mathrm{R}) + I(X_\mathrm{S}; Y_\mathrm{D}|X_\mathrm{R}, Y_\mathrm{R}), I(X_\mathrm{S}, X_\mathrm{R}; Y_\mathrm{D}) \Big\}. \tag{2.14}$$

Achievability follows from the PDF strategy with $U = Y_\mathrm{R}$, which is feasible as Y_R is a function of $(X_\mathrm{S}, X_\mathrm{R})$, and the converse follows from the CSB.

(d) Finally, if the relay channel has orthogonal sender components, its capacity is given by

$$C = \max_{p(x'_\mathrm{S}, x''_\mathrm{S}, x_\mathrm{R})} \min \Big\{ I(X'_\mathrm{S}; Y_\mathrm{R}|X_\mathrm{R}) + I(X''_\mathrm{S}; Y_\mathrm{D}|X_\mathrm{R}), I(X''_\mathrm{S}, X_\mathrm{R}; Y_\mathrm{D}) \Big\}, \tag{2.15}$$

where the maximum is taken over all joint probability mass functions that factor as $p(x'_\mathrm{S}, x''_\mathrm{S}, x_\mathrm{R}) = p(x'_\mathrm{S}|x_\mathrm{R}) p(x''_\mathrm{S}|x_\mathrm{R}) p(x_\mathrm{R})$ [36]. The proof of achievability uses the PDF strategy with $U = X'_\mathrm{S}$, and the converse again follows from the CSB.

In addition, it was recently shown in [18] that PDF achieves the capacity of a class of degraded semideterministic relay channels, which strictly includes the class of degraded relay channels. Note that for all the relay channels mentioned here, the achievable PDF rate meets the CSB.

2.5 Further Results and Bibliographical Notes

The relay channel was introduced by van der Meulen as one particular instance of a three-terminal communication network [129, 130]. It was subsequently also studied by Aref, Cover, El Gamal, Sato, and Zhang [4, 21, 32, 107, 151] before the research activity died down until the middle of the 1990s, when the interest in relaying began to increase again, cf. [28, Section 1.7].

The most important work on the relay channel is that of Cover and El Gamal [21], who derived both upper and lower bounds on the capacity based on a then new *cut-set bound* (CSB, cf. Section 2.2) and the two fundamental relay strategies that are today known as *decode-and-forward* (DF, cf. Section 2.3) and *compress-and-forward* (CF), respectively. As discussed above, the DF protocol requires the relay to decode the entire source message, which is then re-encoded and, in *cooperation* with the source, transmitted to the destination. When using the CF strategy, on the other hand, the relay does not decode any part of the source message. Rather, it *facilitates* the communication by reliably forwarding an estimate, i.e., a compressed version of its receive signal, to the destination.

In their pioneering work, Cover and El Gamal also proposed a relay protocol that combines the ideas of cooperation (DF) and facilitation (CF) [21, Theorem 7]. If it uses this strategy, the relay decodes only some part of the source message and compresses the remainder. For a long time, the best known lower bound on the capacity of the relay channel was due to this mixed strategy. Recently, however, Chong et al. derived two new coding schemes that are also based on combining DF and CF principles [19, 20]. Since both of these schemes subsume the one by Cover and El Gamal, they possibly enlarge the achievable rate.

The *partial decode-and-forward* (PDF) strategy (cf. Section 2.4) is a special case of Cover and El Gamal's mixed strategy where the relay only forwards information about the part of the source message it has decoded. In this particular form, it first appeared in El Gamal and Aref's work on the semideterministic relay channel [32]. Furthermore, it should be mentioned that PDF is sometimes also referred to as "multipath decode-and-forward", cf. [78, Section 4.2].

The CSB, which was generalized to relay networks (Theorem 2.2) in [4, 31], is still the best general upper bound on the capacity of the relay channel. In addition to those relay channels for which PDF achieves the capacity (cf. Section 2.4), it is tight for a class of *primitive relay channels*, i.e., relay channels having a separate noiseless relay-to-destination link, cf. [73]. More precisely, it could be shown that CF achieves the CSB for primitive relay channels in which the relay's observation Y_R is a deterministic function of the source input X_S and the destination output Y_D [23, 74] as well as for the more general class of semideterministic orthogonal relay channels [18].

All the aforementioned relay channels for which the capacities are known share the common property that they achieve their respective CSBs. However, the CSB is not always tight as demonstrated in [3, 120, 151]. Without characterizing the actual capacity, Zhang established a partial converse for a primitive relay channel which shows that the CSB can be loose [151]. Moreover, Aleksic et al. recently derived the capacity of a class of (primitive) modulo-sum relay channels, which is achieved by the CF strategy and which can be strictly below the CSB [3]. The latter result was then generalized by Tandon and Ulukus, who obtained an upper bound that is strictly tighter than the CSB for a broader subclass of primitive relay channels [120]. Finally, some justifications for why the CSB cannot be tight in all cases were provided by Zahedi [147].

Beyond that, there are several interesting and noteworthy aspects of relaying that are not further elaborated on in this work. For example, the relay channel with different forms of (causal noiseless) *feedback* was considered in [12, 13, 18, 21, 39]. The capacity of the relay channel with feedback from the destination to the relay is achieved by the DF strategy, which follows from the fact that such a relay channel is degraded [21]. In addition, it was shown in [12, 13, 18] that feedback from the relay and/or the destination to the source does not improve the capacity of the semideterministic relay channel. For general relay channels with relay-source or destination-source feedback, achievable rates that improve on the best known lower bounds without feedback were presented in [13, 39], and like for the relay channel without feedback, the best known upper bound on the capacity is the CSB, cf. [13].

Another topic we do not further discuss in the following is the *lookahead* relay channel, where the relay encoder may also use the current observation or even the entire received sequence to generate the relay input. Upper and lower bounds on the capacities of lookahead relay channels were studied in [33], for example. An interesting result is that the classical CSB, which was derived based on the assumption that the relay encoder operates in a causal manner, does not hold for *relay-without-delays*, where the relay inputs may depend on the current and past inputs. More specifically, it was shown in [33] that *instantaneous* (or memoryless) relaying, where the relay input depends on the current observation only, can achieve rates above the classical CSB. The relay-without-delay and instantaneous relaying were subsequently also considered in many other works, cf. [69] and references therein, for example.

To conclude this chapter, we want to remark that excellent surveys on the literature on the relay channel (and more general relay networks) through 2005 and 2007 can be found in [77] and [131], respectively.

Chapter 3

Gaussian MIMO Relay Channel

The information theoretical results described in the previous chapter were all derived for the discrete memoryless relay channel, i.e., a channel model with finite input and output alphabets. However, we focus on the Gaussian relay channel in this work, where the channels connecting the nodes are continuous-alphabet channels in which the receiving nodes are subject to additive Gaussian noise. More precisely, we study the Gaussian *multiple-input multiple-output* (MIMO) relay channel, where all nodes may be equipped with multiple antennas. Before we can apply any of the results presented in Chapter 2 to the Gaussian MIMO relay channel, we thus have to establish that the corresponding theorems remain valid for channels with continuous input and output alphabets. The details are deferred to subsequent sections.

Like for the discrete memoryless case, the capacity of the Gaussian MIMO relay channel remains an open problem.[1] After introducing the system model for the Gaussian MIMO relay channel in Section 3.1, we hence study the same capacity bounds as in the previous chapter. Using the entropy maximizing property of the Gaussian distribution, it is established in Sections 3.2 and 3.3, respectively, that the *cut-set bound* (CSB) C_{CSB} and the maximum achievable *decode-and-forward* (DF) rate R_{DF} are attained by jointly Gaussian source and relay inputs. Furthermore, we show that if the channel gain matrices are known perfectly and instantaneously at all nodes, C_{CSB} and R_{DF} can be determined as the solutions of convex optimization problems.

Rates that can be achieved using the *partial decode-and-forward* (PDF) strategy are discussed in Section 3.4. In particular, we first show that the maximum achievable PDF rate R_{PDF} is also attained by Gaussian channel inputs. The derivation of this result is quite involved (it extends over Sections 3.4.1 and 3.4.2) as the entropy maximizing property cannot directly be applied. Although the optimal channel input distribution is then known, it seems not always possible to (efficiently) evaluate R_{PDF}. In Section 3.4.3, we therefore propose two principally different suboptimal PDF schemes which both result in rate maximization problems that can be solved using convex optimization techniques.

[1] We remark that this is in contrast to several other Gaussian MIMO channels, e.g., the single-user (or point-to-point) MIMO channel, the MIMO *multiple-access channel*, and the MIMO *broadcast channel*, for which characterizations of the capacity (regions) have been found, cf. [9, Chapter 2], for example.

Section 3.4.4 identifies several classes of Gaussian MIMO relay channels for which R_{PDF}, like C_{CSB} and R_{DF}, can be evaluated as the solution of a convex optimization problem, and the chapter concludes in Section 3.5 with an overview of further noteworthy results and bibliographical notes on the Gaussian (MIMO) relay channel.

3.1 System Model

Throughout this work, we use the discrete-time narrowband MIMO channel model that is also considered in [9, Chapter 1] or [50, Chapter 10], for example. That is, we assume frequency-flat fading over the bandwidth of interest, which is justified if the delay spread in the channel is negligible compared to the inverse bandwidth. Using this model, which is also known as the *linear MIMO model*, the input-output characteristic of the channel between nodes A and B is specified by

$$\boldsymbol{y}_{\mathrm{B}} = \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{x}_{\mathrm{A}} + \boldsymbol{n}_{\mathrm{B}}. \tag{3.1}$$

Here, $\boldsymbol{x}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{A}}}$ and $\boldsymbol{y}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{B}}}$ denote the signal vectors transmitted by node A and received at node B, respectively, $\boldsymbol{H}_{\mathrm{AB}} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{A}}}$ is the channel gain matrix such that $[\boldsymbol{H}_{\mathrm{AB}}]_{k,\ell}$ represents the channel gain from transmit antenna ℓ to receive antenna k, and $\boldsymbol{n}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{B}}}$ denotes the additive noise vector received at node B, which is assumed to be independent of the transmit signal $\boldsymbol{x}_{\mathrm{A}}$.

The additive noise is modeled as a zero-mean proper (circularly symmetric) complex Gaussian random vector with covariance Z_{B}, i.e., $\boldsymbol{n}_{\mathrm{B}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, Z_{\mathrm{B}})$.[2] This noise model is reasonable in the sense that if the receiving node is subject to many independent noise sources, their cumulative effect should approach a Gaussian distribution by the central limit theorem. Furthermore, it has been shown that Gaussian noise is the worst-case additive noise [27, 115]. Loosely speaking, this means that if a rate is achievable in a Gaussian channel, it can also be achieved in the channel where the additive Gaussian noise is replaced by any other additive noise of the same covariance.

The model for the Gaussian MIMO relay channel, which is illustrated in Figure 3.1, is obtained by applying the linear MIMO model to the considered relay scenario. The receive signal vectors of the relay and the destination can then be expressed as

$$\begin{aligned} \boldsymbol{y}_{\mathrm{R}} &= \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}, & \boldsymbol{n}_{\mathrm{R}} &\sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, Z_{\mathrm{R}}), \\ \boldsymbol{y}_{\mathrm{D}} &= \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{n}_{\mathrm{D}}, & \boldsymbol{n}_{\mathrm{D}} &\sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, Z_{\mathrm{D}}), \end{aligned} \tag{3.2}$$

where $\boldsymbol{H}_{\mathrm{SR}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{S}}}$, $\boldsymbol{H}_{\mathrm{SD}} \in \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{S}}}$, and $\boldsymbol{H}_{\mathrm{RD}} \in \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{R}}}$ represent the channel gain

[2] A complex random vector $\boldsymbol{x}$ is said to be *circularly symmetric* if its probability distribution is rotationally invariant, i.e., if $\boldsymbol{x}$ and $\boldsymbol{x}\exp(\mathrm{j}\phi)$ have the same distribution for all $\phi \in [0, 2\pi]$. If $\boldsymbol{x}$ is circularly symmetric, it hence follows that $\mathrm{E}[\boldsymbol{x}] = \boldsymbol{0}$ (from $\phi = \pi$) and $\mathrm{E}[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}] = \boldsymbol{0}$ (from $\phi = \pi/2$). Furthermore, $\boldsymbol{x}$ is said to be *proper* if $\mathrm{E}[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}] = \boldsymbol{0}$. As the Gaussian distribution is completely determined by second-order statistics, a complex Gaussian random vector is circularly symmetric if and only if it is zero-mean and proper. For more details on complex random vectors, the reader is referred to [1] or [111].

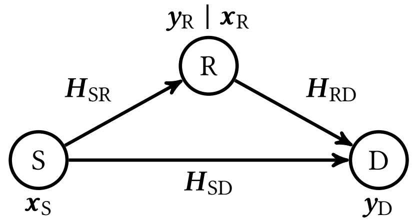


Figure 3.1: Illustration of the Gaussian MIMO Relay Channel

matrices, which we assume to be perfectly and instantaneously known at all nodes. Moreover, $\boldsymbol{n}_{\mathrm{R}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_{\mathrm{R}})$ and $\boldsymbol{n}_{\mathrm{D}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_{\mathrm{D}})$ denote independent zero-mean proper complex Gaussian noise vectors with nonsingular covariance matrices $\boldsymbol{Z}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ and $\boldsymbol{Z}_{\mathrm{D}} \in \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{D}}}$. These noise vectors are also assumed to be independent of the transmit signals $\boldsymbol{x}_{\mathrm{S}} \in \mathbb{C}^{N_{\mathrm{S}}}$ and $\boldsymbol{x}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{R}}}$, and we assume perfectly synchronized transmission and reception between all nodes.

In the following, we often assume the additive Gaussian noise to be white, i.e., $\boldsymbol{Z}_{\mathrm{R}} = \mathbf{I}_{N_{\mathrm{R}}}$ and $\boldsymbol{Z}_{\mathrm{D}} = \mathbf{I}_{N_{\mathrm{D}}}$. Note that this is without loss of generality since any Gaussian MIMO relay channel with nonsingular noise covariances $\boldsymbol{Z}_{\mathrm{R}}$ and $\boldsymbol{Z}_{\mathrm{D}}$ can be transformed into one with additive white Gaussian noise vectors by means of a *noise whitening* operation. In particular, if $\boldsymbol{y}_{\mathrm{R}}$ and $\boldsymbol{y}_{\mathrm{D}}$ are multiplied by $\boldsymbol{Z}_{\mathrm{R}}^{-1/2}$ and $\boldsymbol{Z}_{\mathrm{D}}^{-1/2}$, respectively, we obtain a Gaussian MIMO relay channel with white noise vectors and the effective channel gain matrices $\boldsymbol{Z}_{\mathrm{R}}^{-1/2}\boldsymbol{H}_{\mathrm{SR}}$, $\boldsymbol{Z}_{\mathrm{D}}^{-1/2}\boldsymbol{H}_{\mathrm{SD}}$, and $\boldsymbol{Z}_{\mathrm{D}}^{-1/2}\boldsymbol{H}_{\mathrm{RD}}$, cf. [51].

REMARK 3.1. It is implicitly assumed in (3.2) that the relay is able to completely cancel its self-interference. In addition, the channel outputs $\boldsymbol{y}_{\mathrm{R}}$ and $\boldsymbol{y}_{\mathrm{D}}$ are conditionally independent given the channel inputs $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$ due to the assumption that the noise vectors $\boldsymbol{n}_{\mathrm{R}}$ and $\boldsymbol{n}_{\mathrm{D}}$ are independent. In a corresponding discrete memoryless relay channel, this would mean that the conditional probability mass function $p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}})$ specifying the channel factors as $p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}}) = p(y_{\mathrm{D}}|x_{\mathrm{S}}, x_{\mathrm{R}})p(y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}})$.

Without further conditions on the channel inputs, the capacity of the Gaussian MIMO relay channel is infinite because one can then choose infinite subsets of inputs arbitrarily far apart so that they are distinguishable at the outputs with arbitrarily small probability of error, cf. [24, Chapter 9]. We therefore impose the power constraints

$$\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}] \leq P_{\mathrm{S}}, \qquad \mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}] \leq P_{\mathrm{R}} \tag{3.3}$$

on the channel inputs, where $P_{\mathrm{S}} > 0$ and $P_{\mathrm{R}} > 0$ denote the power budgets available at the source and the relay, respectively.[3] Note that without loss of generality, we can

[3]Adding such power constraints to single-letter mutual information expressions that were obtained for discrete memoryless channels is generally more restrictive than imposing power constraints of the form $\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{x}_i \leq P$ on the $(2^{nR}, n)$ code for the corresponding Gaussian channel. However, both approaches yield the same values for C_{CSB} and R_{DF}, and any rate that is achievable with the approach we consider is also achievable with the standard approach, cf. Appendix B.1.1.

restrict our attention to zero-mean inputs as the optimal $\boldsymbol{x}_\mathrm{S}$ and $\boldsymbol{x}_\mathrm{R}$ are always zero-mean. The reason for this is that channel inputs with nonzero mean consume more transmit power than the corresponding zero-mean signals, but they cannot convey more information since translations do not change the differential entropy of continuous random vectors, cf. [24, Theorem 8.6.3]. As a consequence, the covariance matrices of the source and relay inputs are given by $\boldsymbol{C}_\mathrm{S} = \mathrm{E}[\boldsymbol{x}_\mathrm{S}\boldsymbol{x}_\mathrm{S}^\mathrm{H}]$ and $\boldsymbol{C}_\mathrm{R} = \mathrm{E}[\boldsymbol{x}_\mathrm{R}\boldsymbol{x}_\mathrm{R}^\mathrm{H}]$ so that the transmit power constraints can equivalently be expressed as

$$\mathrm{tr}(\boldsymbol{C}_\mathrm{S}) \leq P_\mathrm{S}, \qquad \mathrm{tr}(\boldsymbol{C}_\mathrm{R}) \leq P_\mathrm{R}. \tag{3.4}$$

When discussing the different capacity bounds in the following sections, we not only have to consider the marginal distributions of $\boldsymbol{x}_\mathrm{S}$ and $\boldsymbol{x}_\mathrm{R}$, but also their joint distribution. In particular, we need the joint covariance matrix of the source and relay inputs, which is equal to

$$\boldsymbol{C} = \mathrm{E}\left[\begin{bmatrix}\boldsymbol{x}_\mathrm{S}\\ \boldsymbol{x}_\mathrm{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_\mathrm{S}\\ \boldsymbol{x}_\mathrm{R}\end{bmatrix}^\mathrm{H}\right] = \begin{bmatrix}\boldsymbol{C}_\mathrm{S} & \boldsymbol{C}_\mathrm{SR}\\ \boldsymbol{C}_\mathrm{SR}^\mathrm{H} & \boldsymbol{C}_\mathrm{R}\end{bmatrix}. \tag{3.5}$$

That is because the covariance matrix completely characterizes the zero-mean proper complex Gaussian distribution, which plays a central role for the CSB and the achievable DF and PDF rates. Finally, we remark that by defining the two selection matrices

$$\boldsymbol{D}_\mathrm{S} = \begin{bmatrix}\mathbf{I}_{N_\mathrm{S}} & \mathbf{0}_{N_\mathrm{S}\times N_\mathrm{R}}\end{bmatrix}, \qquad \boldsymbol{D}_\mathrm{R} = \begin{bmatrix}\mathbf{0}_{N_\mathrm{R}\times N_\mathrm{S}} & \mathbf{I}_{N_\mathrm{R}}\end{bmatrix}, \tag{3.6}$$

both the source and the relay transmit covariance matrices (and hence also the power constraints) can be expressed as linear functions of the joint covariance matrix $\boldsymbol{C}$:

$$\boldsymbol{C}_\mathrm{S} = \boldsymbol{D}_\mathrm{S}\boldsymbol{C}\boldsymbol{D}_\mathrm{S}^\mathrm{H}, \qquad \boldsymbol{C}_\mathrm{R} = \boldsymbol{D}_\mathrm{R}\boldsymbol{C}\boldsymbol{D}_\mathrm{R}^\mathrm{H}. \tag{3.7}$$

3.2 Cut-Set Bound

Since the proof of the cut-set bound (CSB) applies to arbitrary (not necessarily discrete) input and output alphabets, cf. [34, Section 16.5], Theorem 2.1 also holds for continuous-alphabet channels. Taking the source and relay transmit power constraints into account, the capacity of the Gaussian MIMO relay channel is thus upper bounded by

$$C_\mathrm{CSB} = \max_{p(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})} \min\left\{I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{R},\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}), I(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R};\boldsymbol{y}_\mathrm{D})\right\} \quad \text{s.t.} \quad \mathrm{tr}(\boldsymbol{C}_\mathrm{S}) \leq P_\mathrm{S}, \quad \mathrm{tr}(\boldsymbol{C}_\mathrm{R}) \leq P_\mathrm{R}. \tag{3.8}$$

In order to evaluate this bound, we have to determine which probability distribution maximizes the minimum of the two mutual information terms.

Theorem 3.1. *For the Gaussian MIMO relay channel, the cut-set bound (CSB) is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. The power constraints only depend on $\boldsymbol{C}_\mathrm{S}$ and $\boldsymbol{C}_\mathrm{R}$, which are determined by the joint covariance matrix $\boldsymbol{C}$. But for any given $\boldsymbol{C}$, the Gaussian distribution maximizes both differential entropy [24, Theorem 8.6.5] and conditional differential entropy [123] and thus $I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{R},\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) = h(\boldsymbol{y}_\mathrm{R},\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) - h(\boldsymbol{n}_\mathrm{R},\boldsymbol{n}_\mathrm{D})$ and $I(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R};\boldsymbol{y}_\mathrm{D}) = h(\boldsymbol{y}_\mathrm{D}) - h(\boldsymbol{n}_\mathrm{D})$. In particular, the (conditional) differential entropy of a complex random vector is maximized by the zero-mean proper complex Gaussian distribution [94, 122]. □

This result is very useful as it allows us to represent the optimal joint distribution of the source and relay inputs by a single parameter, their joint covariance matrix $\boldsymbol{C}$. Letting $\left[\begin{smallmatrix}\boldsymbol{x}_\mathrm{S}\\ \boldsymbol{x}_\mathrm{R}\end{smallmatrix}\right] \sim \mathcal{N}_\mathbb{C}(\boldsymbol{0},\boldsymbol{C})$, the two mutual information terms defining the CSB simplify to the log-det[4] expressions

$$\begin{aligned} I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{R},\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S|R}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^\mathrm{H}\right), \\ I(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R};\boldsymbol{y}_\mathrm{D}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}\boldsymbol{H}_{\mathrm{\{SR\}D}}^\mathrm{H}\right), \end{aligned} \tag{3.9}$$

where we have assumed that the additive Gaussian noise is white,

$$\boldsymbol{H}_{\mathrm{S\{RD\}}} = \begin{bmatrix}\boldsymbol{H}_\mathrm{SR}\\ \boldsymbol{H}_\mathrm{SD}\end{bmatrix} \in \mathbb{C}^{(N_\mathrm{R}+N_\mathrm{D})\times N_\mathrm{S}} \tag{3.10}$$

denotes the channel gain matrix of the composite channel from the source to the relay and the destination, and

$$\boldsymbol{H}_{\mathrm{\{SR\}D}} = \begin{bmatrix}\boldsymbol{H}_\mathrm{SD} & \boldsymbol{H}_\mathrm{RD}\end{bmatrix} \in \mathbb{C}^{N_\mathrm{D}\times(N_\mathrm{S}+N_\mathrm{R})} \tag{3.11}$$

represents the channel gain matrix of the composite channel from the source and the relay to the destination. Furthermore,

$$\boldsymbol{C}_{\mathrm{S|R}} = \boldsymbol{C}_\mathrm{S} - \boldsymbol{C}_\mathrm{SR}\boldsymbol{C}_\mathrm{R}^{+}\boldsymbol{C}_\mathrm{SR}^\mathrm{H} \tag{3.12}$$

is the conditional covariance matrix of $\boldsymbol{x}_\mathrm{S}$ given $\boldsymbol{x}_\mathrm{R}$ and $\boldsymbol{C}_\mathrm{R}^{+}$ denotes the Moore–Penrose pseudoinverse of $\boldsymbol{C}_\mathrm{R}$, which is equal to $\boldsymbol{C}_\mathrm{R}^{-1}$ if $\boldsymbol{C}_\mathrm{R}$ is nonsingular. Evaluating the CSB consequently requires to solve the following optimization problem:

$$\begin{aligned} C_\mathrm{CSB} = \max_{\boldsymbol{C}} \min \left\{\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S|R}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^\mathrm{H}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}\boldsymbol{H}_{\mathrm{\{SR\}D}}^\mathrm{H}\right)\right\} \\ \text{s.t.}\quad \boldsymbol{C} \succcurlyeq \boldsymbol{0},\quad \operatorname{tr}(\boldsymbol{D}_\mathrm{S}\boldsymbol{C}\boldsymbol{D}_\mathrm{S}^\mathrm{H}) \le P_\mathrm{S},\quad \operatorname{tr}(\boldsymbol{D}_\mathrm{R}\boldsymbol{C}\boldsymbol{D}_\mathrm{R}^\mathrm{H}) \le P_\mathrm{R}. \end{aligned} \tag{3.13}$$

For the actual computation of C_CSB, it would be convenient if this optimization problem were convex since there exist powerful methods to efficiently solve convex

[4]The logarithm function log is assumed to be base 2 unless specified otherwise.

optimization problems, cf. [11], for example. In order to determine whether or not the problem is convex, we need to examine the properties of its objective function and constraints. To this end, first note that both $\log\det(\boldsymbol{C})$ and $\log\det(\boldsymbol{C}_{S|R})$ are concave functions of the joint covariance matrix $\boldsymbol{C}$, cf. [24, Theorems 17.9.1 and 17.10.1], from which it follows that $\log\det\left(\mathbf{I} + \boldsymbol{H}_{\{SR\}D}\boldsymbol{C}\boldsymbol{H}_{\{SR\}D}^{H}\right)$ and $\log\det\left(\mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{S|R}\boldsymbol{H}_{S\{RD\}}^{H}\right)$ are also concave in $\boldsymbol{C} \succcurlyeq \mathbf{0}$ as the arguments of the log-det terms are affine functions of $\boldsymbol{C}$, cf. [11, Section 3.2.2]. Since the power constraints are linear functions of $\boldsymbol{C}$, it can thus be concluded that the optimization problem given in (3.13) is convex.

However, note that as the powerful interior point methods cannot directly be applied to this problem, it is advantageous to reformulate the term involving the conditional covariance matrix $\boldsymbol{C}_{S|R}$ by means of an auxiliary variable. In particular, let

$$\mathbf{0} \preccurlyeq \boldsymbol{C}_{Q} \preccurlyeq \boldsymbol{C}_{S|R} = \boldsymbol{C}_{S} - \boldsymbol{C}_{SR}\boldsymbol{C}_{R}^{+}\boldsymbol{C}_{SR}^{H}. \tag{3.14}$$

Since $\mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{Q}\boldsymbol{H}_{S\{RD\}}^{H} \preccurlyeq \mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{S|R}\boldsymbol{H}_{S\{RD\}}^{H}$ for $\boldsymbol{C}_{Q} \preccurlyeq \boldsymbol{C}_{S|R}$ and since $\det(\boldsymbol{C}_{Q})$ is nondecreasing in $\boldsymbol{C}_{Q} \succcurlyeq \mathbf{0}$, cf. [60, Section 7.7], we then have

$$\log\det\left(\mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{Q}\boldsymbol{H}_{S\{RD\}}^{H}\right) \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{S|R}\boldsymbol{H}_{S\{RD\}}^{H}\right) \tag{3.15}$$

for all $\mathbf{0} \preccurlyeq \boldsymbol{C}_{Q} \preccurlyeq \boldsymbol{C}_{S|R}$, which implies that

$$\max_{\mathbf{0} \preccurlyeq \boldsymbol{C}_{Q} \preccurlyeq \boldsymbol{C}_{S|R}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{Q}\boldsymbol{H}_{S\{RD\}}^{H}\right) = \log\det\left(\mathbf{I} + \boldsymbol{H}_{S\{RD\}}\boldsymbol{C}_{S|R}\boldsymbol{H}_{S\{RD\}}^{H}\right) \tag{3.16}$$

for all conditional covariance matrices $\boldsymbol{C}_{S|R}$. Moreover, note that

$$\boldsymbol{C}_{S|R} \succcurlyeq \boldsymbol{C}_{Q} \quad \Leftrightarrow \quad (\boldsymbol{C}_{S} - \boldsymbol{C}_{Q}) - \boldsymbol{C}_{SR}\boldsymbol{C}_{R}^{+}\boldsymbol{C}_{SR}^{H} \succcurlyeq \mathbf{0}. \tag{3.17}$$

The constraint resulting from the introduction of the auxiliary variable $\boldsymbol{C}_{Q}$ can therefore be reformulated by making use of the following lemma:

Lemma 3.2. *The following conditions are equivalent:*

1. $$\boldsymbol{C}_{R} \succcurlyeq \mathbf{0}, \quad (\boldsymbol{C}_{S} - \boldsymbol{C}_{Q}) - \boldsymbol{C}_{SR}\boldsymbol{C}_{R}^{+}\boldsymbol{C}_{SR}^{H} \succcurlyeq \mathbf{0}, \quad (\mathbf{I} - \boldsymbol{C}_{R}\boldsymbol{C}_{R}^{+})\boldsymbol{C}_{SR}^{H} = \mathbf{0} \tag{3.18}$$
2. $$\begin{bmatrix} \boldsymbol{C}_{S} - \boldsymbol{C}_{Q} & \boldsymbol{C}_{SR} \\ \boldsymbol{C}_{SR}^{H} & \boldsymbol{C}_{R} \end{bmatrix} = \boldsymbol{C} - \boldsymbol{D}_{S}^{H}\boldsymbol{C}_{Q}\boldsymbol{D}_{S} \succcurlyeq \mathbf{0} \tag{3.19}$$

Proof. Note that $(\boldsymbol{C}_{S} - \boldsymbol{C}_{Q}) - \boldsymbol{C}_{SR}\boldsymbol{C}_{R}^{+}\boldsymbol{C}_{SR}^{H}$ is the Schur complement of $\boldsymbol{C}_{R}$ in $\boldsymbol{C} - \boldsymbol{D}_{S}^{H}\boldsymbol{C}_{Q}\boldsymbol{D}_{S}$. The lemma hence follows from the Schur complement condition for positive semidefinite matrices, cf. [2] or [61]. □

The first two conditions of (3.18) are clearly satisfied as $\boldsymbol{C}_{R}$ is itself a covariance matrix and $(\boldsymbol{C}_{S} - \boldsymbol{C}_{Q}) - \boldsymbol{C}_{SR}\boldsymbol{C}_{R}^{+}\boldsymbol{C}_{SR}^{H} \succcurlyeq \mathbf{0}$ is the condition on $\boldsymbol{C}_{Q}$ stated in (3.17). In order to apply Lemma 3.2 to our problem, it remains to verify that $(\mathbf{I} - \boldsymbol{C}_{R}\boldsymbol{C}_{R}^{+})\boldsymbol{C}_{SR}^{H} = \mathbf{0}$ for all possible joint covariance matrices $\boldsymbol{C}$. But since $\boldsymbol{C} \succcurlyeq \mathbf{0}$, this again follows from the Schur complement condition for positive semidefinite matrices. In fact, the desired result can be obtained by setting $\boldsymbol{C}_\mathrm{Q} = \mathbf{0}$ in Lemma 3.2.

With the help of an auxiliary variable $\boldsymbol{C}_\mathrm{Q} \succcurlyeq \mathbf{0}$, the CSB maximization problem given in (3.13) can accordingly be formulated as

$$C_\mathrm{CSB} = \max_{\boldsymbol{C},\boldsymbol{C}_\mathrm{Q}} \min\left\{\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_\mathrm{Q}\boldsymbol{H}_{\mathrm{S\{RD\}}}^\mathrm{H}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}\boldsymbol{H}_{\mathrm{\{SR\}D}}^\mathrm{H}\right)\right\}$$
$$\text{s.t.} \quad \boldsymbol{C} - \boldsymbol{D}_\mathrm{S}^\mathrm{H}\boldsymbol{C}_\mathrm{Q}\boldsymbol{D}_\mathrm{S} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_\mathrm{Q} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{D}_\mathrm{S}\boldsymbol{C}\boldsymbol{D}_\mathrm{S}^\mathrm{H}) \le P_\mathrm{S}, \quad \mathrm{tr}(\boldsymbol{D}_\mathrm{R}\boldsymbol{C}\boldsymbol{D}_\mathrm{R}^\mathrm{H}) \le P_\mathrm{R}. \tag{3.20}$$

Note that the objective function of this optimization problem is concave because it is the pointwise minimum of two functions that are jointly concave in $\boldsymbol{C} \succcurlyeq \mathbf{0}$ and $\boldsymbol{C}_\mathrm{Q} \succcurlyeq \mathbf{0}$, cf. [11, Section 3.2.3]. In addition, all constraints are affine functions of the optimization variables so that we can conclude the following:

Proposition 3.3. *For the Gaussian MIMO relay channel, C_CSB can be determined as the solution of the convex optimization problem given in* (3.20).

Remark 3.2. The constraints $\boldsymbol{C} - \boldsymbol{D}_\mathrm{S}^\mathrm{H}\boldsymbol{C}_\mathrm{Q}\boldsymbol{D}_\mathrm{S} \succcurlyeq \mathbf{0}$ and $\boldsymbol{C}_\mathrm{Q} \succcurlyeq \mathbf{0}$ imply $\boldsymbol{C} \succcurlyeq \mathbf{0}$. While this means that adding the latter constraint is redundant from a theoretical point of view, it might nevertheless be useful (or even required) to explicitly define $\boldsymbol{C}$ as a positive semidefinite matrix when actually solving the problem.

Remark 3.3. Since $\boldsymbol{C}_\mathrm{S}$ and $\boldsymbol{C}_\mathrm{R}$ are linear functions of $p(\boldsymbol{x}_\mathrm{S}, \boldsymbol{x}_\mathrm{R})$, the CSB optimization problem given in (3.8), where the maximization is over all joint distributions satisfying the power constraints, is already convex, cf. Remark 2.2. This section hence reveals that for the Gaussian MIMO relay channel, this still holds if one restricts the joint distribution of $\boldsymbol{x}_\mathrm{S}$ and $\boldsymbol{x}_\mathrm{R}$ to be zero-mean proper complex Gaussian.

3.3 Decode-and-Forward

The original decode-and-forward (DF) coding scheme of Cover and El Gamal [21] uses random encoding and jointly typical decoding. More precisely, their achievability proof is based on *strong typicality*, which can only be used for discrete random variables, cf. [76, Chapter 1]. As pointed out in [77, Remark 28], however, Theorem 2.3 remains valid for continuous-alphabet channels since it can also be derived using *weakly typical* sequences, which apply to continuous random variables as well.

Considering the transmit power constraints, the maximum achievable DF rate for the Gaussian MIMO relay channel is therefore equal to

$$R_\mathrm{DF} = \max_{p(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})} \min\left\{I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R}), I(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R};\boldsymbol{y}_\mathrm{D})\right\}$$
$$\text{s.t.} \quad \mathrm{tr}(\boldsymbol{C}_\mathrm{S}) \le P_\mathrm{S}, \quad \mathrm{tr}(\boldsymbol{C}_\mathrm{R}) \le P_\mathrm{R}. \tag{3.21}$$

As previously discussed in Section 2.3, the only difference between (3.8) and (3.21) is in the first mutual information term of the objective function, where an additional $\boldsymbol{y}_{\text{D}}$ appears in the CSB maximization problem. Apart from that, the optimization problems have the same structure, optimization variables, and constraints. Like for the CSB, we can therefore state the following theorem:

Theorem 3.4. *For the Gaussian MIMO relay channel, the maximum achievable decode-and-forward (DF) rate is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. This result again follows from the entropy maximizing property of the zero-mean proper complex Gaussian distribution, cf. Theorem 3.1. □

Letting $\left[\begin{smallmatrix}\boldsymbol{x}_{\text{S}}\\ \boldsymbol{x}_{\text{R}}\end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C})$ and assuming the Gaussian noise to be white again, the mutual information terms that upper bound the achievable DF rate read as

$$\begin{aligned} I(\boldsymbol{x}_{\text{S}}; \boldsymbol{y}_{\text{R}}|\boldsymbol{x}_{\text{R}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SR}}\boldsymbol{C}_{\text{S}|\text{R}}\boldsymbol{H}_{\text{SR}}^{\text{H}}\right), \\ I(\boldsymbol{x}_{\text{S}}, \boldsymbol{x}_{\text{R}}; \boldsymbol{y}_{\text{D}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\text{SR}\}\text{D}}\boldsymbol{C}\boldsymbol{H}_{\{\text{SR}\}\text{D}}^{\text{H}}\right). \end{aligned} \tag{3.22}$$

As a consequence, the optimization problem that determines the maximum achievable DF rate for the Gaussian MIMO relay channel is given by

$$\begin{aligned} R_{\text{DF}} = \max_{\boldsymbol{C}} \min \Big\{&\log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SR}}\boldsymbol{C}_{\text{S}|\text{R}}\boldsymbol{H}_{\text{SR}}^{\text{H}}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\text{SR}\}\text{D}}\boldsymbol{C}\boldsymbol{H}_{\{\text{SR}\}\text{D}}^{\text{H}}\right)\Big\} \\ &\text{s.t.} \quad \boldsymbol{C} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{D}_{\text{S}}\boldsymbol{C}\boldsymbol{D}_{\text{S}}^{\text{H}}) \le P_{\text{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\text{R}}\boldsymbol{C}\boldsymbol{D}_{\text{R}}^{\text{H}}) \le P_{\text{R}}. \end{aligned} \tag{3.23}$$

Note that except for the fact that the composite channel gain matrix $\boldsymbol{H}_{\text{S}\{\text{RD}\}}$ has been replaced by the source-to-relay channel gain matrix $\boldsymbol{H}_{\text{SR}}$, this rate maximization problem is identical to the CSB maximization problem given in (3.13). This implies that the problem as formulated in (3.23) is convex, but also that interior point methods cannot be directly applied to it. However, by making use of the arguments and results of the previous section, it is straightforward to show that the maximum achievable DF rate can equivalently be evaluated by solving the optimization problem

$$\begin{aligned} R_{\text{DF}} = \max_{\boldsymbol{C}, \boldsymbol{C}_{\text{Q}}} \min \Big\{&\log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SR}}\boldsymbol{C}_{\text{Q}}\boldsymbol{H}_{\text{SR}}^{\text{H}}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\text{SR}\}\text{D}}\boldsymbol{C}\boldsymbol{H}_{\{\text{SR}\}\text{D}}^{\text{H}}\right)\Big\} \\ \text{s.t.} \quad &\boldsymbol{C} - \boldsymbol{D}_{\text{S}}^{\text{H}}\boldsymbol{C}_{\text{Q}}\boldsymbol{D}_{\text{S}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_{\text{Q}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{D}_{\text{S}}\boldsymbol{C}\boldsymbol{D}_{\text{S}}^{\text{H}}) \le P_{\text{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\text{R}}\boldsymbol{C}\boldsymbol{D}_{\text{R}}^{\text{H}}) \le P_{\text{R}}. \end{aligned} \tag{3.24}$$

Since this DF rate maximization problem has obviously the same structure as the CSB maximization problem in (3.20), we can state the following:

Proposition 3.5. *For the Gaussian MIMO relay channel, R_{DF} can be determined as the solution of the convex optimization problem given in* (3.24).

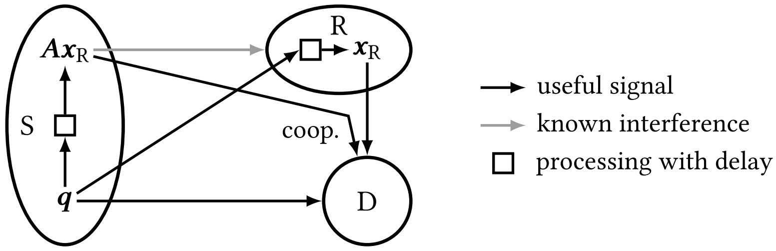


Figure 3.2: Decomposition of the Source Transmit Signal for the DF Strategy

Remark 3.4. Due to the similarity of the problems given in (3.20) and (3.24), we only need to solve one type of optimization problem in order to evaluate C_{CSB} and R_{DF}. That is, if we can evaluate C_{CSB}, then we can also evaluate R_{DF} and vice versa. Accordingly, Remarks 3.2 and 3.3 also apply to the DF rate maximization problem.

We also remark that the auxiliary variable $\boldsymbol{C}_Q$ has a nice interpretation in the context of the DF strategy. Recall that R_{DF} is achieved by a block Markov superposition encoding scheme that uses B blocks of transmission to convey $B-1$ independent messages from the source to the destination. Due to the causality of the relay encoder, the relay's transmit signal $\boldsymbol{x}_R \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_R)$ in block b is a function of the previous message w_{b-1}. Provided that the relay encoding function is deterministic, $\boldsymbol{x}_R$ is then also known to the source. If we let $\boldsymbol{q} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_Q)$ be a function of the current message w_b only (and independent of w_{b-1}), the transmit signal of the source in block b can be expressed as the superposition of $\boldsymbol{q}(w_b)$ and $\boldsymbol{x}_R(w_{b-1})$ given by

$$\boldsymbol{x}_S = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_R, \tag{3.25}$$

where $\boldsymbol{A} \in \mathbb{C}^{N_S \times N_R}$ specifies a linear transformation the source may apply to $\boldsymbol{x}_R$. Since $\boldsymbol{q}$ and $\boldsymbol{x}_R$ are independent, it follows that $\boldsymbol{x}_S \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_Q + \boldsymbol{A}\boldsymbol{C}_R\boldsymbol{A}^H)$.

Of course, both messages w_{b-1} and w_b are ultimately decoded by the relay and, after being re-encoded, forwarded to the destination. In block b, however, w_{b-1} is already known to the relay while w_b has yet to be decoded. Consequently, the source and the relay can *cooperatively* transmit w_{b-1} to the destination by means of $\boldsymbol{A}\boldsymbol{x}_R$ and $\boldsymbol{x}_R$. On the other hand, $\boldsymbol{q}$ contains the current message w_b, which is provided to the relay in order to allow cooperation in the next block, cf. Figure 3.2.

The auxiliary variable $\boldsymbol{C}_Q$ hence specifies the distribution of the signal part $\boldsymbol{q}$, which can be interpreted as the *innovation* the source introduces into the system. Furthermore, if $\boldsymbol{C}^\star$ and $\boldsymbol{C}_Q^\star$ denote the optimizers of the DF rate maximization problem given in (3.24), we have $\boldsymbol{C}_Q^\star = \boldsymbol{C}_{S|R}^\star$, i.e., for the optimal joint distribution of $(\boldsymbol{x}_S, \boldsymbol{x}_R)$, $\boldsymbol{C}_Q$ is equivalent to the conditional covariance matrix $\boldsymbol{C}_{S|R}$. This essentially follows from (3.16) and the fact that $\boldsymbol{C}$ uniquely determines $\boldsymbol{C}_{S|R}$ for proper complex Gaussian signals. Note also that if $\boldsymbol{C}_S = \boldsymbol{C}_Q + \boldsymbol{A}\boldsymbol{C}_R\boldsymbol{A}^H$ with $\boldsymbol{C}_Q = \boldsymbol{C}_{S|R}$, the linear transformation the source applies to $\boldsymbol{x}_R$ must be equal to $\boldsymbol{A} = \boldsymbol{C}_{SR}\boldsymbol{C}_R^+$.

Since the joint covariance matrix $\boldsymbol{C}$ is completely specified by $\boldsymbol{C}_{\mathrm{Q}}$, $\boldsymbol{C}_{\mathrm{R}}$, and $\boldsymbol{A}$, it is also possible to express the achievable DF rate as a function of these three matrices, e.g., to better reflect the structure of $\boldsymbol{x}_{\mathrm{S}}$ given in (3.25). Unfortunately, the objective function of the resulting DF rate maximization problem would be nonconvex as it would contain the product $\boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}$ of the optimization variables $\boldsymbol{A}$ and $\boldsymbol{C}_{\mathrm{R}}$.

However, the influence of the innovative and cooperative parts of the source input on the achievable DF rate can also more easily be identified if the problem in (3.24) is reformulated as follows. Let $\bar{\boldsymbol{C}} = \boldsymbol{C} - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}}$, then

$$R_{\mathrm{DF}} = \max_{\bar{\boldsymbol{C}},\boldsymbol{C}_{\mathrm{Q}}} \min\left\{\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\bar{\boldsymbol{C}}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right)\right\}$$
$$\text{s.t.} \quad \bar{\boldsymbol{C}}, \boldsymbol{C}_{\mathrm{Q}} \succeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{D}_{\mathrm{S}}\bar{\boldsymbol{C}}\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\bar{\boldsymbol{C}}\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \quad (3.26)$$

Here, $\boldsymbol{C}_{\mathrm{Q}}$ again denotes the covariance matrix of the innovative part $\boldsymbol{q}$, whereas $\bar{\boldsymbol{C}}$ can be interpreted as the joint covariance matrix of the cooperative part $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ of $\boldsymbol{x}_{\mathrm{S}}$ and the relay transmit signal $\boldsymbol{x}_{\mathrm{R}}$.

3.4 Partial Decode-and-Forward

In the previous section, it was argued that Theorem 2.3 remains valid for the Gaussian MIMO relay channel because the DF coding scheme can be derived by means of weakly typical sequences, which apply to both discrete and continuous random variables. The same reasoning can be used here to explain why Theorem 2.4 also holds for continuous-alphabet channels. As a consequence, the maximum achievable partial decode-and-forward (PDF) rate for the Gaussian MIMO relay channel is equal to

$$R_{\mathrm{PDF}} = \max_{p(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})} \min\left\{I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}), I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}})\right\}$$
$$\text{s.t.} \quad \boldsymbol{u} \leftrightarrow (\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}}) \leftrightarrow (\boldsymbol{y}_{\mathrm{D}},\boldsymbol{y}_{\mathrm{R}}), \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{S}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}}. \quad (3.27)$$

Similar to the DF coding scheme, R_{PDF} is achieved by a block Markov superposition encoding scheme. However, the PDF strategy is more general in the sense that the relay need not decode the entire source message, i.e., the source may split its message w_b into independent parts w'_b and w''_b and send the second one w''_b directly to the destination via the source-to-destination link. The transmit signal of the source in block b can hence be expressed as the superposition of $\boldsymbol{q}(w'_b)$, $\boldsymbol{x}_{\mathrm{R}}(w'_{b-1})$, and $\boldsymbol{v}(w''_b)$ given by

$$\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{v}, \quad (3.28)$$

where the interpretation of the signal parts is as follows. Like in (3.25), $\boldsymbol{q}$ represents the *innovation to be decoded by the relay* and $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ denotes the *cooperative part*, which allows the source and the relay to cooperatively transmit the message part w'_{b-1} the relay has previously decoded to the destination, cf. Figure 3.3. The new part $\boldsymbol{v}$ represents the

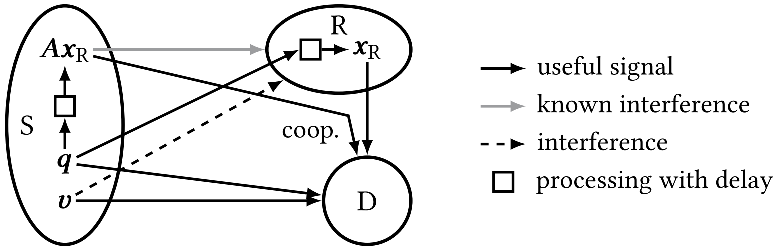


Figure 3.3: Decomposition of the Source Transmit Signal for the PDF Strategy

innovation not to be decoded by the relay, i.e., the part of the source message w_b'' that is conveyed to the destination over the direct link only. Since w_b'' is not supposed to be decoded by the relay, $\boldsymbol{v}$ acts as interference at the relay as illustrated in Figure 3.3.

Note that at this point, it is not clear that the source input $\boldsymbol{x}_{\mathrm{S}}$ can be represented as the superposition of three independent signal parts. For the DF strategy, independence of $\boldsymbol{q}$ and $\boldsymbol{x}_{\mathrm{R}}$ could be assumed because the optimal source and relay inputs are jointly Gaussian, cf. Theorem 3.4. For the PDF strategy, on the other hand, we still need to prove the optimality of Gaussian channel inputs, which is much more involved than proving the result for C_{CSB}, R_{DF}, or R_{P2P}. In particular, we cannot simply invoke the entropy maximizing property of the zero-mean proper complex Gaussian distribution to argue that R_{PDF} is maximized by Gaussian channel inputs. The reason for this is that the entropy maximizing property cannot be applied to the term

$$I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) = h(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) - h(\boldsymbol{n}_{\mathrm{D}}) \\ + h(\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) - h(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}), \tag{3.29}$$

which includes the difference $h(\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) - h(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})$ of two conditional differential entropies involving $\boldsymbol{u}$, $\boldsymbol{x}_{\mathrm{S}}$, and $\boldsymbol{x}_{\mathrm{R}}$.

We remark that for $\boldsymbol{H}_{\mathrm{SR}} = \boldsymbol{H}_{\mathrm{SD}} = \mathbf{I}$, the maximization of such a difference over the conditional probability distribution $p(\boldsymbol{x}_{\mathrm{S}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})$ subject to a shaping constraint on the conditional covariance matrix $\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}]$ was analyzed in [85], where it was proved that the optimal distribution is Gaussian. However, the term $I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}})+I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})$, and hence the difference $h(\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}}+\boldsymbol{n}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})-h(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}}+\boldsymbol{n}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})$, is only one part of the objective function of the PDF rate maximization problem given in (3.27). Therefore, we cannot directly apply [85, Theorem 8] to prove that the achievable PDF rate is maximized by Gaussian channel inputs.

3.4.1 Aligned Gaussian MIMO Relay Channel

As a first step towards proving that the maximum achievable PDF rate for the Gaussian MIMO relay channel is attained by Gaussian channel inputs, we consider the *aligned* Gaussian MIMO relay channel. The results for this special case are then generalized in the following section.

Definition 3.1. The Gaussian MIMO relay channel is said to be *aligned* if $N_\mathrm{S} = N_\mathrm{R} = N_\mathrm{D} = N$ and $\boldsymbol{H}_\mathrm{SR} = \boldsymbol{H}_\mathrm{SD} = \mathbf{I}_N$.

If we consider the aligned Gaussian MIMO relay channel, the general relay channel model given in (3.2) hence reduces to

$$\begin{aligned} \boldsymbol{y}_\mathrm{R} &= \boldsymbol{x}_\mathrm{S} + \boldsymbol{n}_\mathrm{R}, & \boldsymbol{n}_\mathrm{R} &\sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \boldsymbol{Z}_\mathrm{R}), \\ \boldsymbol{y}_\mathrm{D} &= \boldsymbol{x}_\mathrm{S} + \boldsymbol{H}_\mathrm{RD}\boldsymbol{x}_\mathrm{R} + \boldsymbol{n}_\mathrm{D}, & \boldsymbol{n}_\mathrm{D} &\sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \boldsymbol{Z}_\mathrm{D}). \end{aligned} \tag{3.30}$$

For this special case, the theorem below shows that jointly Gaussian source and relay inputs maximize the achievable PDF rate. The proof relies on a *channel enhancement* argument, which, like the idea to first consider an aligned channel, goes back to Weingarten et al. [137].

Theorem 3.6. *For the aligned Gaussian MIMO relay channel, the maximum achievable partial decode-and-forward (PDF) rate is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. Achievability: Let $\boldsymbol{q} \sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \boldsymbol{C}_\mathrm{Q})$, $\boldsymbol{v} \sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \boldsymbol{C}_\mathrm{V})$, $\boldsymbol{x}_\mathrm{R} \sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \boldsymbol{C}_\mathrm{R})$ be independent, $\boldsymbol{u} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_\mathrm{R}$, and $\boldsymbol{x}_\mathrm{S} = \boldsymbol{u} + \boldsymbol{v}$ such that $\boldsymbol{x}_\mathrm{S} \sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \boldsymbol{C}_\mathrm{S})$ with $\boldsymbol{C}_\mathrm{S} = \boldsymbol{C}_\mathrm{Q} + \boldsymbol{A}\boldsymbol{C}_\mathrm{R}\boldsymbol{A}^\mathrm{H} + \boldsymbol{C}_\mathrm{V}$. Then, the mutual information terms characterizing the achievable PDF rate simplify to

$$\begin{aligned} I(\boldsymbol{u}; \boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R}) &= \log\det(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V} + \boldsymbol{Z}_\mathrm{R}) - \log\det(\boldsymbol{C}_\mathrm{V} + \boldsymbol{Z}_\mathrm{R}), \\ I(\boldsymbol{x}_\mathrm{S}; \boldsymbol{y}_\mathrm{D}|\boldsymbol{u}, \boldsymbol{x}_\mathrm{R}) &= \log\det(\boldsymbol{C}_\mathrm{V} + \boldsymbol{Z}_\mathrm{D}) - \log\det(\boldsymbol{Z}_\mathrm{D}), \\ I(\boldsymbol{x}_\mathrm{S}, \boldsymbol{x}_\mathrm{R}; \boldsymbol{y}_\mathrm{D}) &= \log\det\left(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V} + (\boldsymbol{H}_\mathrm{RD} + \boldsymbol{A})\boldsymbol{C}_\mathrm{R}(\boldsymbol{H}_\mathrm{RD} + \boldsymbol{A})^\mathrm{H} + \boldsymbol{Z}_\mathrm{D}\right) \\ &\quad - \log\det(\boldsymbol{Z}_\mathrm{D}). \end{aligned} \tag{3.31}$$

We can therefore conclude that

$$\begin{aligned} R_\mathrm{PDF} \geq R_\mathrm{PDF}^{\mathcal{N}_\mathrm{c}} = &\max_{R, \boldsymbol{C}_\mathrm{Q}, \boldsymbol{C}_\mathrm{V}, \boldsymbol{C}_\mathrm{R}, \boldsymbol{A}} R \\ \text{s.t.}\quad & R \leq \log\det(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V} + \boldsymbol{Z}_\mathrm{R}) - \log\det(\boldsymbol{C}_\mathrm{V} + \boldsymbol{Z}_\mathrm{R}) \\ &\qquad + \log\det(\boldsymbol{C}_\mathrm{V} + \boldsymbol{Z}_\mathrm{D}) - \log\det(\boldsymbol{Z}_\mathrm{D}), \\ & R \leq \log\det\left(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V} + (\boldsymbol{H}_\mathrm{RD} + \boldsymbol{A})\boldsymbol{C}_\mathrm{R}(\boldsymbol{H}_\mathrm{RD} + \boldsymbol{A})^\mathrm{H} + \boldsymbol{Z}_\mathrm{D}\right) - \log\det(\boldsymbol{Z}_\mathrm{D}), \\ & \boldsymbol{C}_\mathrm{Q}, \boldsymbol{C}_\mathrm{V}, \boldsymbol{C}_\mathrm{R} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V} + \boldsymbol{A}\boldsymbol{C}_\mathrm{R}\boldsymbol{A}^\mathrm{H}) \leq P_\mathrm{S}, \quad \mathrm{tr}(\boldsymbol{C}_\mathrm{R}) \leq P_\mathrm{R}, \end{aligned} \tag{3.32}$$

where $R_\mathrm{PDF}^{\mathcal{N}_\mathrm{c}}$ denotes a rate that is achievable by means of the PDF strategy and proper complex Gaussian channel inputs.

To simplify the optimization problem that specifies $R_\mathrm{PDF}^{\mathcal{N}_\mathrm{c}}$, let $\boldsymbol{S} = \boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V}$ be the covariance of the innovative signal $\boldsymbol{q} + \boldsymbol{v}$. Using this definition and applying a primal decomposition approach to the PDF rate maximization problem above, we obtain

$$R_\mathrm{PDF}^{\mathcal{N}_\mathrm{c}} = \max_{\boldsymbol{S}} R_\mathrm{PDF}^{\mathcal{N}_\mathrm{c}}(\boldsymbol{S}) \quad \text{s.t.} \quad \boldsymbol{S} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{S}) \leq P_\mathrm{S} \tag{3.33}$$

with

$$
\begin{aligned}
R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(\boldsymbol{S}) = \max_{R, \boldsymbol{C}_{\mathrm{R}}, \boldsymbol{A}} \; & R \\
\text{s.t.} \quad & R \leq \log\det\left(\boldsymbol{S} + \boldsymbol{Z}_{\mathrm{R}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right) \\
& \qquad + \max_{\boldsymbol{0} \preccurlyeq \boldsymbol{C}_{\mathrm{V}} \preccurlyeq \boldsymbol{S}} \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{R}}\right), \\
& R \leq \log\det\left(\boldsymbol{S} + (\boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{A})\boldsymbol{C}_{\mathrm{R}}(\boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{A})^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right), \\
& \boldsymbol{C}_{\mathrm{R}} \succcurlyeq \boldsymbol{0}, \quad \mathrm{tr}(\boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}) \leq P_{\mathrm{S}} - \mathrm{tr}(\boldsymbol{S}), \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}}.
\end{aligned}
\tag{3.34}
$$

Next, consider the inner maximization problem

$$
\max_{\boldsymbol{C}_{\mathrm{V}}} \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{R}}\right) \quad \text{s.t.} \quad \boldsymbol{0} \preccurlyeq \boldsymbol{C}_{\mathrm{V}} \preccurlyeq \boldsymbol{S}, \tag{3.35}
$$

which up to the additive constant $\log\det\left(\boldsymbol{Z}_{\mathrm{R}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right)$ is mathematically equivalent to the optimization problem that yields the secrecy capacity of the aligned Gaussian MIMO wiretap channel (vector Gaussian wiretap channel) under a shaping constraint, cf. [84, Section II-A].[5] Adopting the considerations from the proof of [84, Theorem 2], the optimal value of the inner maximization problem given in (3.35) can thus be determined as follows.

First, note that the Karush–Kuhn–Tucker (KKT) conditions are necessary conditions for this problem. That is because the Abadie constraint qualification is automatically satisfied if all constraints are linear [6, Section 5.1] and because the KKT conditions readily extend to problems with generalized inequalities such as positive semidefiniteness constraints [11, Section 5.9.2]. As a consequence, any optimizer $\boldsymbol{C}_{\mathrm{V}}^{\star}$ of the inner maximization problem must satisfy

$$
(\boldsymbol{C}_{\mathrm{V}}^{\star} + \boldsymbol{Z}_{\mathrm{D}})^{-1} + \boldsymbol{\Lambda}_1 = (\boldsymbol{C}_{\mathrm{V}}^{\star} + \boldsymbol{Z}_{\mathrm{R}})^{-1} + \boldsymbol{\Lambda}_2, \tag{3.36a}
$$

$$
\boldsymbol{C}_{\mathrm{V}}^{\star}\boldsymbol{\Lambda}_1 = \boldsymbol{0}, \tag{3.36b}
$$

$$
(\boldsymbol{S} - \boldsymbol{C}_{\mathrm{V}}^{\star})\boldsymbol{\Lambda}_2 = \boldsymbol{0}, \tag{3.36c}
$$

where $\boldsymbol{\Lambda}_1 \succcurlyeq \boldsymbol{0}$ and $\boldsymbol{\Lambda}_2 \succcurlyeq \boldsymbol{0}$ are Lagrangian multipliers corresponding to the (generalized) inequality constraints $\boldsymbol{C}_{\mathrm{V}} \succcurlyeq \boldsymbol{0}$ and $\boldsymbol{S} - \boldsymbol{C}_{\mathrm{V}} \succcurlyeq \boldsymbol{0}$, respectively. Now, let $\boldsymbol{Z}$ such that

$$
(\boldsymbol{C}_{\mathrm{V}}^{\star} + \boldsymbol{Z})^{-1} = (\boldsymbol{C}_{\mathrm{V}}^{\star} + \boldsymbol{Z}_{\mathrm{D}})^{-1} + \boldsymbol{\Lambda}_1. \tag{3.37}
$$

It then follows from (3.36b) that an explicit expression for $\boldsymbol{Z}$ (as a function of $\boldsymbol{Z}_{\mathrm{D}}$ and the Lagrangian multiplier $\boldsymbol{\Lambda}_1$) is given by

$$
\boldsymbol{Z} = (\boldsymbol{Z}_{\mathrm{D}}^{-1} + \boldsymbol{\Lambda}_1)^{-1}, \tag{3.38}
$$

[5] Comparing the problems given in (3.35) and [84, eq. (17)], we see that the destination plays the role of the legitimate receiver and the relay that of the eavesdropper.

and because $Z_\mathrm{D} \succ \mathbf{0}$ and $\boldsymbol{\Lambda}_1 \succcurlyeq \mathbf{0}$, we can conclude that $Z \succ \mathbf{0}$. Furthermore, (3.36a) and the definition of Z in (3.37) imply that

$$(C_\mathrm{V}^\star + Z)^{-1} = (C_\mathrm{V}^\star + Z_\mathrm{R})^{-1} + \boldsymbol{\Lambda}_2. \tag{3.39}$$

By means of the variable Z, it is hence possible to determine the optimal value of the inner maximization problem given in (3.35). To this end, note that

$$\begin{aligned}(C_\mathrm{V}^\star + Z)Z^{-1} &= C_\mathrm{V}^\star Z^{-1} + \mathbf{I} \overset{(a)}{=} C_\mathrm{V}^\star (Z_\mathrm{D}^{-1} + \boldsymbol{\Lambda}_1) + \mathbf{I} \\ &\overset{(b)}{=} C_\mathrm{V}^\star Z_\mathrm{D}^{-1} + \mathbf{I} = (C_\mathrm{V}^\star + Z_\mathrm{D})Z_\mathrm{D}^{-1},\end{aligned} \tag{3.40}$$

where (a) follows from (3.38) and (b) is due to (3.36b). As a result,

$$\frac{\det(C_\mathrm{V}^\star + Z_\mathrm{D})}{\det(Z_\mathrm{D})} = \frac{\det(C_\mathrm{V}^\star + Z)}{\det(Z)}. \tag{3.41}$$

Similarly, it holds that

$$\begin{aligned}(S + Z)(C_\mathrm{V}^\star + Z)^{-1} &= (S - C_\mathrm{V}^\star + C_\mathrm{V}^\star + Z)(C_\mathrm{V}^\star + Z)^{-1} \\ &= (S - C_\mathrm{V}^\star)(C_\mathrm{V}^\star + Z)^{-1} + \mathbf{I} \\ &\overset{(c)}{=} (S - C_\mathrm{V}^\star)\left((C_\mathrm{V}^\star + Z_\mathrm{R})^{-1} + \boldsymbol{\Lambda}_2\right) + \mathbf{I} \\ &\overset{(d)}{=} (S - C_\mathrm{V}^\star)(C_\mathrm{V}^\star + Z_\mathrm{R})^{-1} + \mathbf{I} \\ &= (S - C_\mathrm{V}^\star + C_\mathrm{V}^\star + Z_\mathrm{R})(C_\mathrm{V}^\star + Z_\mathrm{R})^{-1} \\ &= (S + Z_\mathrm{R})(C_\mathrm{V}^\star + Z_\mathrm{R})^{-1},\end{aligned} \tag{3.42}$$

where (c) and (d) are due to (3.39) and (3.36c), respectively. Consequently, we obtain

$$\frac{\det(C_\mathrm{V}^\star + Z)}{\det(C_\mathrm{V}^\star + Z_\mathrm{R})} = \frac{\det(S + Z)}{\det(S + Z_\mathrm{R})}. \tag{3.43}$$

The optimal value of the inner problem can therefore be calculated as

$$\begin{aligned}&\log\det(C_\mathrm{V}^\star + Z_\mathrm{D}) - \log\det(C_\mathrm{V}^\star + Z_\mathrm{R}) \\ &\quad= \log\frac{\det(C_\mathrm{V}^\star + Z_\mathrm{D})}{\det(Z_\mathrm{D})} - \log\frac{\det(C_\mathrm{V}^\star + Z_\mathrm{R})}{\det(Z_\mathrm{D})} \\ &\quad\overset{(e)}{=} \log\frac{\det(C_\mathrm{V}^\star + Z)}{\det(Z)} - \log\frac{\det(C_\mathrm{V}^\star + Z_\mathrm{R})}{\det(Z_\mathrm{D})} \\ &\quad= \log\frac{\det(C_\mathrm{V}^\star + Z)}{\det(C_\mathrm{V}^\star + Z_\mathrm{R})} - \log\frac{\det(Z)}{\det(Z_\mathrm{D})} \\ &\quad\overset{(f)}{=} \log\frac{\det(S + Z)}{\det(S + Z_\mathrm{R})} - \log\frac{\det(Z)}{\det(Z_\mathrm{D})} \\ &\quad= \log\det(S + Z) - \log\det(Z) - \left(\log\det(S + Z_\mathrm{R}) - \log\det(Z_\mathrm{D})\right),\end{aligned} \tag{3.44}$$

where (e) follows from (3.41) and (f) is due to (3.43). Using this result, it is straightforward to verify that $R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}(\boldsymbol{S})$ is equal to

$$\begin{aligned} R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}(\boldsymbol{S}) = \max_{R,\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A}} \;& R \\ \text{s.t.} \quad & R \le \log\det(\boldsymbol{S}+\boldsymbol{Z}) - \log\det(\boldsymbol{Z}), \\ & R \le \log\det\left(\boldsymbol{S} + (\boldsymbol{H}_{\mathrm{RD}}+\boldsymbol{A})\boldsymbol{C}_{\mathrm{R}}(\boldsymbol{H}_{\mathrm{RD}}+\boldsymbol{A})^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det(\boldsymbol{Z}_{\mathrm{D}}), \\ & \boldsymbol{C}_{\mathrm{R}} \succcurlyeq \boldsymbol{0}, \quad \mathrm{tr}(\boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}) \le P_{\mathrm{S}} - \mathrm{tr}(\boldsymbol{S}), \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{R}}) \le P_{\mathrm{R}}. \end{aligned} \tag{3.45}$$

Converse: From (3.37), (3.39) and the positive semidefiniteness of the Lagrangian multipliers $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$, it follows that

$$\boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{D}}, \qquad \boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{R}}. \tag{3.46}$$

We can thus use the variable $\boldsymbol{Z}$ to define an *enhanced* aligned Gaussian MIMO relay channel. In particular, let $\tilde{\boldsymbol{Z}}_{\mathrm{R}} = \boldsymbol{Z}$ and

$$\begin{aligned} \tilde{\boldsymbol{y}}_{\mathrm{R}} &= \boldsymbol{x}_{\mathrm{S}} + \tilde{\boldsymbol{n}}_{\mathrm{R}}, & \tilde{\boldsymbol{n}}_{\mathrm{R}} &\sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \tilde{\boldsymbol{Z}}_{\mathrm{R}}), \\ \boldsymbol{y}_{\mathrm{D}} &= \boldsymbol{x}_{\mathrm{S}} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{n}_{\mathrm{D}}, & \boldsymbol{n}_{\mathrm{D}} &\sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{D}}). \end{aligned} \tag{3.47}$$

Since $\boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{R}}$, $\boldsymbol{y}_{\mathrm{R}}$ is a stochastically degraded version of $\tilde{\boldsymbol{y}}_{\mathrm{R}}$, which implies $I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) \le I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})$, cf. Appendix B.2.2. As a result, we obtain

$$\begin{aligned} R_{\mathrm{PDF}} \le \tilde{R}_{\mathrm{PDF}} = \max_{p(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})} & \min\left\{ I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}), I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}}) \right\} \\ \text{s.t.} \quad & \boldsymbol{u} \leftrightarrow (\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}}) \leftrightarrow (\boldsymbol{y}_{\mathrm{D}},\tilde{\boldsymbol{y}}_{\mathrm{R}}), \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{S}}) \le P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{R}}) \le P_{\mathrm{R}}. \end{aligned} \tag{3.48}$$

Moreover, as $\boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{D}}$, it follows that $\boldsymbol{y}_{\mathrm{D}}$ is a stochastically degraded version of $\tilde{\boldsymbol{y}}_{\mathrm{R}}$ as well. Therefore,

$$\begin{aligned} I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) &\le I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \\ &= I(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) = I(\boldsymbol{x}_{\mathrm{S}};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}), \end{aligned} \tag{3.49}$$

where the last equation is due to the Markov condition $\boldsymbol{u} \leftrightarrow (\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}}) \leftrightarrow (\boldsymbol{y}_{\mathrm{D}},\tilde{\boldsymbol{y}}_{\mathrm{R}})$. The optimal PDF strategy for the enhanced relay channel defined in (3.47) is hence equivalent to DF, i.e., $\tilde{R}_{\mathrm{PDF}} = \tilde{R}_{\mathrm{DF}}$ with

$$\begin{aligned} \tilde{R}_{\mathrm{DF}} = \max_{p(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})} & \min\left\{ I(\boldsymbol{x}_{\mathrm{S}};\tilde{\boldsymbol{y}}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}), I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}}) \right\} \\ & \text{s.t.} \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{S}}) \le P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{R}}) \le P_{\mathrm{R}}. \end{aligned} \tag{3.50}$$

However, we know from Theorem 3.4 that the maximum achievable DF rate for any Gaussian MIMO relay channel is attained by proper complex Gaussian channel inputs. Consequently, the achievable PDF rate for the enhanced aligned Gaussian MIMO relay channel is maximized by letting $\boldsymbol{q} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{Q}})$ and $\boldsymbol{x}_{\mathrm{R}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{R}})$ be independent and $\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ such that $\boldsymbol{x}_{\mathrm{S}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{S}})$ with $\boldsymbol{C}_{\mathrm{S}} = \boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}$, i.e.,

$$\begin{aligned} \tilde{R}_{\mathrm{PDF}} = \max_{R, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{R}}, \boldsymbol{A}} \; & R \\ \text{s.t.} \quad & R \leq \log\det(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{Z}) - \log\det(\boldsymbol{Z}), \\ & R \leq \log\det(\boldsymbol{C}_{\mathrm{Q}} + (\boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{A})\boldsymbol{C}_{\mathrm{R}}(\boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{A})^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}) - \log\det(\boldsymbol{Z}_{\mathrm{D}}), \\ & \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{R}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{3.51}$$

If we again apply a primal decomposition approach, this maximization problem can equivalently be written as

$$\tilde{R}_{\mathrm{PDF}} = \max_{\boldsymbol{C}_{\mathrm{Q}}} \tilde{R}_{\mathrm{PDF}}(\boldsymbol{C}_{\mathrm{Q}}) \quad \text{s.t.} \quad \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{Q}}) \leq P_{\mathrm{S}}, \tag{3.52}$$

where

$$\begin{aligned} \tilde{R}_{\mathrm{PDF}}(\boldsymbol{C}_{\mathrm{Q}}) = \max_{R, \boldsymbol{C}_{\mathrm{R}}, \boldsymbol{A}} \; & R \\ \text{s.t.} \quad & R \leq \log\det(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{Z}) - \log\det(\boldsymbol{Z}), \\ & R \leq \log\det(\boldsymbol{C}_{\mathrm{Q}} + (\boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{A})\boldsymbol{C}_{\mathrm{R}}(\boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{A})^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}) - \log\det(\boldsymbol{Z}_{\mathrm{D}}), \\ & \boldsymbol{C}_{\mathrm{R}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}) \leq P_{\mathrm{S}} - \operatorname{tr}(\boldsymbol{C}_{\mathrm{Q}}), \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{3.53}$$

Comparing (3.45) to (3.53), we notice that $\tilde{R}_{\mathrm{PDF}}(\boldsymbol{C}_{\mathrm{Q}}) = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(\boldsymbol{S})$ for $\boldsymbol{C}_{\mathrm{Q}} = \boldsymbol{S}$, from which we can conclude that $\tilde{R}_{\mathrm{PDF}} = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}$ as the constraints in (3.33) and (3.52) are the same. But since $\tilde{R}_{\mathrm{PDF}} \geq R_{\mathrm{PDF}} \geq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}$ in general, this implies $R_{\mathrm{PDF}} = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}$. □

Note that the key to proving Theorem 3.6 is the primal decomposition approach in the achievability part, which we use to split the complicated PDF rate maximization problem into subproblems. That is because this approach leads to an inner optimization problem that is well known from the secrecy capacity of the aligned Gaussian MIMO wiretap channel under a shaping constraint, cf. [84, Section II-A]. In particular, both the optimal value of this inner problem, which enables us to simplify the expression for the achievable rate, and the enhanced channel, which we require for the converse, follow from considerations similar to those in the proof of [84, Theorem 2].

However, there is a difference in how the channel enhancement argument is used in the proofs of Theorem 3.6 and [84, Theorem 2]. Whereas the secrecy capacity of the aligned Gaussian MIMO wiretap channel is derived by enhancing the channel to the legitimate receiver, we enhance the channel from the source to the relay, which plays the role of the eavesdropper in the inner problem. The reason for this is that enhancing the source-to-destination channel would not have yielded the desired converse for our purpose as the second mutual information term inside the minimum of (3.48) would also have increased.

Like for the aligned Gaussian MIMO wiretap channel, on the other hand, we can explain the existence and the properties of the enhanced channel by considering the special case where the noise covariances are diagonal, cf. [84, Section III]. In particular, it follows from [83] that the achievable PDF rate for the *parallel Gaussian relay channel* is maximized if the relay decodes the entire information transmitted over the subchannels for which it receives a better signal than the destination and no information sent over the subchannels for which its receive signal is worse.[6] An enhanced relay channel for which the achievable PDF rate does not increase can therefore be constructed as follows. On any subchannel where the destination receives a better signal than the relay, reduce the noise variance of the relay to that of the destination. Then, on every subchannel of the resulting enhanced parallel Gaussian relay channel, the destination's receive signal is no better than the relay's receive signal. The enhanced relay channel hence belongs to the class of *stochastically degraded relay channels*, for which the optimal PDF strategy reduces to DF, cf. Section 3.4.4.

Following this line of thought, the enhanced aligned Gaussian MIMO relay channel defined in (3.47) can be interpreted as the stochastically degraded relay channel that is obtained by reducing the noise covariance of the relay "just enough" such that R_{PDF} remains the same as for the original aligned relay channel. Because the noise covariances $\boldsymbol{Z}_{\mathrm{R}}$ and $\boldsymbol{Z}_{\mathrm{D}}$ may have different eigendirections, finding this enhanced channel is more involved than for the parallel Gaussian relay channel. However, the fact that such an enhanced channel exists can be concluded from the derivation of [84, Theorem 2].

3.4.2 General Gaussian MIMO Relay Channel

We now extend the result that jointly proper complex Gaussian source and relay inputs maximize the achievable PDF rate to the general Gaussian MIMO relay channel. The main idea for extending the proof from the aligned to the general case is adopted from [137] again. We first argue that any Gaussian MIMO relay channel can be described by a channel model with square channel gain matrices. Then, we use the singular value decomposition (SVD) to enhance $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ by adding small perturbations to their singular values such that the resulting channel gain matrices are invertible. Finally, we show that the maximum achievable PDF rate for the original Gaussian MIMO relay channel can be obtained by a limit process on the maximum achievable PDF rate for the enhanced (perturbed) relay channel.

Theorem 3.7. *For the Gaussian MIMO relay channel, the maximum achievable partial decode-and-forward (PDF) rate is attained by jointly proper complex Gaussian source and relay inputs.*

[6] As pointed out in [83, Remark 2], the parallel relay channel is not a simple combination of independent subchannels since information sent to the relay on one subchannel may be forwarded to the destination on other subchannels. Consequently, it would not be entirely accurate to state that the optimal PDF strategy reduces to DF/P2P transmission on subchannels where the source-to-relay channel is better/worse than the source-to-destination channel.

Proof. Without loss of generality, we may assume that $\boldsymbol{H}_{\mathrm{SR}}, \boldsymbol{H}_{\mathrm{SD}}, \boldsymbol{H}_{\mathrm{RD}} \in \mathbb{C}^{N\times N}$ with $N = \max\{N_{\mathrm{S}}, N_{\mathrm{R}}, N_{\mathrm{D}}\}$. If this is not the case, we can augment the matrices with zeros to obtain an equivalent Gaussian MIMO relay channel with square $N \times N$ channel gain matrices that preserves the achievable PDF rate under the same power constraints. As previously discussed, we may also assume that the additive Gaussian noise is white, i.e., $\boldsymbol{Z}_{\mathrm{R}} = \boldsymbol{Z}_{\mathrm{D}} = \mathbf{I}_N$.

Achievability: Let $\boldsymbol{q} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{Q}})$, $\boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{V}})$, $\boldsymbol{x}_{\mathrm{R}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{R}})$ be independent, $\boldsymbol{u} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$, and $\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{u} + \boldsymbol{v}$ such that $\boldsymbol{x}_{\mathrm{S}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{S}})$ with $\boldsymbol{C}_{\mathrm{S}} = \boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}} + \boldsymbol{C}_{\mathrm{V}}$. Then, the rate

$$R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}} = \max_{\boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{R}}, \boldsymbol{A}} R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}(\boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{R}}, \boldsymbol{A}) \quad \text{s.t.} \quad \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{R}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}} + \boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}) \le P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}}) \le P_{\mathrm{R}} \tag{3.54}$$

is achievable with jointly proper complex Gaussian source and relay inputs, where

$$\begin{aligned} R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}(\boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{R}}, \boldsymbol{A}) = \min\Big\{&\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ &+ \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{H}^{\mathrm{H}}\right)\Big\} \end{aligned} \tag{3.55}$$

and $\boldsymbol{H} = \boldsymbol{H}_{\mathrm{RD}} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{A}$.

Converse: Suppose the SVDs of $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ are given by

$$\boldsymbol{H}_{\mathrm{SR}} = \boldsymbol{U}_{\mathrm{SR}}\boldsymbol{\Sigma}_{\mathrm{SR}}\boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}}, \qquad \boldsymbol{H}_{\mathrm{SD}} = \boldsymbol{U}_{\mathrm{SD}}\boldsymbol{\Sigma}_{\mathrm{SD}}\boldsymbol{V}_{\mathrm{SD}}^{\mathrm{H}}, \tag{3.56}$$

where $\boldsymbol{U}_{\mathrm{SR}}, \boldsymbol{U}_{\mathrm{SD}}, \boldsymbol{V}_{\mathrm{SR}}, \boldsymbol{V}_{\mathrm{SD}} \in \mathbb{C}^{N\times N}$ are the unitary matrices containing the singular vectors, and where the diagonal matrices $\boldsymbol{\Sigma}_{\mathrm{SR}}, \boldsymbol{\Sigma}_{\mathrm{SD}} \in \mathbb{R}_+^{N\times N}$ contain the singular values of $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$, respectively. Moreover, let

$$\bar{\boldsymbol{H}}_{\mathrm{SR}} = \boldsymbol{U}_{\mathrm{SR}}(\boldsymbol{\Sigma}_{\mathrm{SR}} + \varepsilon\mathbf{I})\boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}}, \qquad \bar{\boldsymbol{H}}_{\mathrm{SD}} = \boldsymbol{U}_{\mathrm{SD}}(\boldsymbol{\Sigma}_{\mathrm{SD}} + \varepsilon\mathbf{I})\boldsymbol{V}_{\mathrm{SD}}^{\mathrm{H}} \tag{3.57}$$

for some $\varepsilon > 0$, and consider the following enhanced Gaussian MIMO relay channel:

$$\begin{aligned} \bar{\boldsymbol{y}}_{\mathrm{R}} &= \bar{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}, & \boldsymbol{n}_{\mathrm{R}} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N), \\ \bar{\boldsymbol{y}}_{\mathrm{D}} &= \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{n}_{\mathrm{D}}, & \boldsymbol{n}_{\mathrm{D}} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N). \end{aligned} \tag{3.58}$$

Since both $\bar{\boldsymbol{H}}_{\mathrm{SR}}$ and $\bar{\boldsymbol{H}}_{\mathrm{SD}}$ are invertible, this relay channel is equivalent to an aligned Gaussian MIMO relay channel with $\boldsymbol{Z}_{\mathrm{R}} = (\bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SR}})^{-1}$ and $\boldsymbol{Z}_{\mathrm{D}} = (\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SD}})^{-1}$, for which the achievable PDF rate is maximized by proper complex Gaussian channel inputs, cf. Theorem 3.6. Furthermore, note that

$$\begin{aligned} \boldsymbol{H}_{\mathrm{SR}} &= \boldsymbol{M}_{\mathrm{SR}}\bar{\boldsymbol{H}}_{\mathrm{SR}}, & \boldsymbol{M}_{\mathrm{SR}} &= \boldsymbol{U}_{\mathrm{SR}}\boldsymbol{\Sigma}_{\mathrm{SR}}(\boldsymbol{\Sigma}_{\mathrm{SR}} + \varepsilon\mathbf{I})^{-1}\boldsymbol{U}_{\mathrm{SR}}^{\mathrm{H}} \preccurlyeq \mathbf{I}_N, \\ \boldsymbol{H}_{\mathrm{SD}} &= \boldsymbol{M}_{\mathrm{SD}}\bar{\boldsymbol{H}}_{\mathrm{SD}}, & \boldsymbol{M}_{\mathrm{SD}} &= \boldsymbol{U}_{\mathrm{SD}}\boldsymbol{\Sigma}_{\mathrm{SD}}(\boldsymbol{\Sigma}_{\mathrm{SD}} + \varepsilon\mathbf{I})^{-1}\boldsymbol{U}_{\mathrm{SD}}^{\mathrm{H}} \preccurlyeq \mathbf{I}_N, \end{aligned} \tag{3.59}$$

which according to [114, Lemma 5] is equivalent to

$$\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}} \preccurlyeq \bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SR}}, \qquad \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SD}} \preccurlyeq \bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SD}}. \tag{3.60}$$

However, this means that $\boldsymbol{y}_{\mathrm{R}}$ and $\boldsymbol{y}_{\mathrm{D}}$ are stochastically degraded versions of $\bar{\boldsymbol{y}}_{\mathrm{R}}$ and $\bar{\boldsymbol{y}}_{\mathrm{D}}$, respectively, cf. Appendix B.2.2. As a result, we have

$$\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}} = \bar{R}_{\mathrm{PDF}} \geq R_{\mathrm{PDF}} \geq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}, \tag{3.61}$$

where

$$\begin{aligned} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}} = \max_{\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A}} \ & \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A}) \\ \text{s.t.} \quad & \boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}} \succcurlyeq \boldsymbol{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}} + \boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}} \end{aligned} \tag{3.62}$$

is the maximum achievable PDF rate for the enhanced Gaussian MIMO relay channel defined in (3.58),

$$\begin{aligned} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A}) = \min \Big\{ & \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}})\bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ & + \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\right), \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}})\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}} + \bar{\boldsymbol{H}}\boldsymbol{C}_{\mathrm{R}}\bar{\boldsymbol{H}}^{\mathrm{H}}\right)\Big\}, \end{aligned} \tag{3.63}$$

and $\bar{\boldsymbol{H}} = \boldsymbol{H}_{\mathrm{RD}} + \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{A}$. The proof can hence be completed by showing that $\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}} \to R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}$ as $\varepsilon \to 0$.

To this end, suppose $\boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{R}}$, and $\boldsymbol{A}$ are fixed. Then, $\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A})$ is a continuous function of ε since it is the pointwise minimum of two functions that are continuous in ε. As a consequence,

$$\lim_{\varepsilon\to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A}) = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{R}},\boldsymbol{A}). \tag{3.64}$$

Because (3.64) holds for any $\boldsymbol{A}$ and $\boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{R}} \succcurlyeq \boldsymbol{0}$, it also holds for the maximizers $\boldsymbol{C}_{\mathrm{Q}}^{\star}, \boldsymbol{C}_{\mathrm{V}}^{\star}, \boldsymbol{C}_{\mathrm{R}}^{\star}, \boldsymbol{A}^{\star}$ of the problem given in (3.62). In addition, these maximizers also satisfy the constraints of (3.54), which means that

$$\lim_{\varepsilon\to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}} = \lim_{\varepsilon\to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}}^{\star},\boldsymbol{C}_{\mathrm{V}}^{\star},\boldsymbol{C}_{\mathrm{R}}^{\star},\boldsymbol{A}^{\star}) = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}(\boldsymbol{C}_{\mathrm{Q}}^{\star},\boldsymbol{C}_{\mathrm{V}}^{\star},\boldsymbol{C}_{\mathrm{R}}^{\star},\boldsymbol{A}^{\star}) \leq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}. \tag{3.65}$$

But since $\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}} \geq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}$ in general, this implies $\lim_{\varepsilon\to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}} = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{c}}}$. □

Because the maximum achievable PDF rate is attained by jointly Gaussian source and relay inputs, $\boldsymbol{x}_{\mathrm{S}}$ can indeed be decomposed as $\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{v}$ with $\boldsymbol{q}$, $\boldsymbol{x}_{\mathrm{R}}$, and $\boldsymbol{v}$ being independent. However, like for the DF strategy, it is not convenient to express the correlation of $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$ by means of $\boldsymbol{A}$ and $\boldsymbol{C}_{\mathrm{R}}$ if one actually wants to evaluate R_{PDF}. This is again because the corresponding maximization problem would then contain the product $\boldsymbol{A}\boldsymbol{C}_{\mathrm{R}}\boldsymbol{A}^{\mathrm{H}}$ of two optimization variables.

For the purpose of avoiding this issue, let us define $\boldsymbol{u} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$. In this case, the source input can be represented as

$$\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{u} + \boldsymbol{v} \tag{3.66}$$

with $\boldsymbol{u}$ and $\boldsymbol{x}_{\mathrm{R}}$ being correlated by design (to enable cooperative transmission from the source and the relay to the destination) and independent of $\boldsymbol{v}$. The joint covariance matrix of the source and relay inputs is consequently equal to

$$\boldsymbol{C} = \left[\begin{bmatrix} \boldsymbol{u}+\boldsymbol{v} \\ \boldsymbol{x}_{\mathrm{R}} \end{bmatrix}\begin{bmatrix} \boldsymbol{u}+\boldsymbol{v} \\ \boldsymbol{x}_{\mathrm{R}} \end{bmatrix}^{\mathrm{H}}\right] = \begin{bmatrix} \boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}} & \boldsymbol{C}_{\mathrm{UR}} \\ \boldsymbol{C}_{\mathrm{UR}}^{\mathrm{H}} & \boldsymbol{C}_{\mathrm{R}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_{\mathrm{U}} & \boldsymbol{C}_{\mathrm{UR}} \\ \boldsymbol{C}_{\mathrm{UR}}^{\mathrm{H}} & \boldsymbol{C}_{\mathrm{R}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{C}_{\mathrm{V}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{3.67}$$

where $\boldsymbol{C}_{\mathrm{U}}$ and $\boldsymbol{C}_{\mathrm{V}}$ denote the covariance matrices of $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively, and where $\boldsymbol{C}_{\mathrm{UR}}$ is the cross-covariance matrix that specifies the correlation between $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$. If we further define $\check{\boldsymbol{C}}$ to be the joint covariance matrix of $\boldsymbol{u}$ and $\boldsymbol{x}_{\mathrm{R}}$, it follows that

$$\boldsymbol{C} = \check{\boldsymbol{C}} + \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{D}_{\mathrm{S}}. \tag{3.68}$$

Note also that like for the discrete memoryless relay channel, we then obtain $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$ if $\boldsymbol{u} = \boldsymbol{x}_{\mathrm{S}}$ and $R_{\mathrm{PDF}} = R_{\mathrm{P2P}}$ if $\boldsymbol{u} = \mathbf{0}$, where

$$R_{\mathrm{P2P}} = \max_{p(\boldsymbol{x}_{\mathrm{S}})} I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{x}_{\mathrm{R}} = \mathbf{0}) \quad \text{s.t.} \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}}) \leq P_{\mathrm{S}}. \tag{3.69}$$

Remark 3.5. The auxiliary variable $\boldsymbol{u}$ represents the part of the source message the relay has to decode in the PDF strategy, which means that $\boldsymbol{u}$ must contain $\boldsymbol{q}$ and that it must not contain $\boldsymbol{v}$. It is hence also possible to define $\boldsymbol{u} = \boldsymbol{q}$, in which case the source input can be expressed as $\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{u} + \boldsymbol{w}$ with $\boldsymbol{w} = \boldsymbol{A}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{v}$ and $\boldsymbol{x}_{\mathrm{R}}$ being correlated by design and independent of $\boldsymbol{u}$. But setting $\boldsymbol{u} = \boldsymbol{x}_{\mathrm{S}}$ would then imply that $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$ are independent, and input distributions that factor as $p(\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}) = p(\boldsymbol{x}_{\mathrm{S}})p(\boldsymbol{x}_{\mathrm{R}})$ do not always attain the maximum achievable DF rate. For this reason, we shall mostly prefer the definition where the cooperative part $\boldsymbol{A}\boldsymbol{x}_{\mathrm{R}}$ of the source input belongs to $\boldsymbol{u}$.

Using the decomposition of $\boldsymbol{x}_{\mathrm{S}}$ defined in (3.66) and jointly Gaussian source and relay inputs, i.e., $\boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{V}})$ and $\left[\begin{smallmatrix} \boldsymbol{u} \\ \boldsymbol{x}_{\mathrm{R}} \end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \check{\boldsymbol{C}})$, the mutual information terms which characterize the achievable PDF rate hence simplify to

$$\begin{aligned} I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{U|R}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right), \\ I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\check{\boldsymbol{C}}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \end{aligned} \tag{3.70}$$

where, for simplicity, we have assumed the Gaussian noise to be white again, and where $\boldsymbol{C}_{\mathrm{U|R}} = \boldsymbol{C}_{\mathrm{U}} - \boldsymbol{C}_{\mathrm{UR}}\boldsymbol{C}_{\mathrm{R}}^{+}\boldsymbol{C}_{\mathrm{UR}}^{\mathrm{H}}$ denotes the conditional covariance matrix of $\boldsymbol{u}$ given $\boldsymbol{x}_{\mathrm{R}}$, cf. (3.12). The maximum achievable PDF rate for the Gaussian MIMO relay channel is therefore equal to

$$\begin{aligned} R_{\mathrm{PDF}} = \max_{R,\check{C},C_{\mathrm{V}}} \; & R \\ \text{s.t.} \quad & R \leq \log\det\left(\mathbf{I} + H_{\mathrm{SR}}(C_{\mathrm{U|R}} + C_{\mathrm{V}})H_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + H_{\mathrm{SR}}C_{\mathrm{V}}H_{\mathrm{SR}}^{\mathrm{H}}\right) \\ & \quad + \log\det\left(\mathbf{I} + H_{\mathrm{SD}}C_{\mathrm{V}}H_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq \log\det\left(\mathbf{I} + H_{\mathrm{SD}}C_{\mathrm{V}}H_{\mathrm{SD}}^{\mathrm{H}} + H_{\{\mathrm{SR}\}\mathrm{D}}\check{C}H_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & \check{C}, C_{\mathrm{V}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(C_{\mathrm{V}} + D_{\mathrm{S}}\check{C}D_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(D_{\mathrm{R}}\check{C}D_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{3.71}$$

Note that we can again introduce an auxiliary variable C_{Q} to relax and reformulate the equality constraint on the conditional covariance matrix $C_{\mathrm{U|R}}$, cf. (3.14). It is then straightforward to show that R_{PDF} can equivalently be determined as the solution of the following optimization problem:

$$\begin{aligned} R_{\mathrm{PDF}} = \max_{R,\check{C},C_{\mathrm{Q}},C_{\mathrm{V}}} \; & R \\ \text{s.t.} \quad & R \leq \log\det\left(\mathbf{I} + H_{\mathrm{SR}}(C_{\mathrm{Q}} + C_{\mathrm{V}})H_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + H_{\mathrm{SR}}C_{\mathrm{V}}H_{\mathrm{SR}}^{\mathrm{H}}\right) \\ & \quad + \log\det\left(\mathbf{I} + H_{\mathrm{SD}}C_{\mathrm{V}}H_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq \log\det\left(\mathbf{I} + H_{\mathrm{SD}}C_{\mathrm{V}}H_{\mathrm{SD}}^{\mathrm{H}} + H_{\{\mathrm{SR}\}\mathrm{D}}\check{C}H_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & \check{C} - D_{\mathrm{S}}^{\mathrm{H}}C_{\mathrm{Q}}D_{\mathrm{S}} \succcurlyeq \mathbf{0}, \quad C_{\mathrm{Q}}, C_{\mathrm{V}} \succcurlyeq \mathbf{0}, \\ & \mathrm{tr}(C_{\mathrm{V}} + D_{\mathrm{S}}\check{C}D_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(D_{\mathrm{R}}\check{C}D_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{3.72}$$

Moreover, if one wants to emphasize the influence of the innovative components and the cooperative part of x_{S} on the achievable PDF rate, it is possible to reformulate the problem once more. In particular, if $\bar{C} = \check{C} - D_{\mathrm{S}}^{\mathrm{H}}C_{\mathrm{Q}}D_{\mathrm{S}}$, it follows that

$$\begin{aligned} R_{\mathrm{PDF}} = \max_{R,\bar{C},C_{\mathrm{Q}},C_{\mathrm{V}}} \; & R \\ \text{s.t.} \quad & R \leq \log\det\left(\mathbf{I} + H_{\mathrm{SR}}(C_{\mathrm{Q}} + C_{\mathrm{V}})H_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + H_{\mathrm{SR}}C_{\mathrm{V}}H_{\mathrm{SR}}^{\mathrm{H}}\right) \\ & \quad + \log\det\left(\mathbf{I} + H_{\mathrm{SD}}C_{\mathrm{V}}H_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq \log\det\left(\mathbf{I} + H_{\mathrm{SD}}(C_{\mathrm{Q}} + C_{\mathrm{V}})H_{\mathrm{SD}}^{\mathrm{H}} + H_{\{\mathrm{SR}\}\mathrm{D}}\bar{C}H_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & \bar{C}, C_{\mathrm{Q}}, C_{\mathrm{V}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(C_{\mathrm{Q}} + C_{\mathrm{V}} + D_{\mathrm{S}}\bar{C}D_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(D_{\mathrm{R}}\bar{C}D_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}, \end{aligned} \tag{3.73}$$

where C_{Q} denotes the covariance matrix of the cooperative part to be decoded by the relay, C_{V} represents the covariance matrix of the innovative part not to be decoded by the relay, and where $\bar{C}$ can be interpreted as the joint covariance matrix of the cooperative part Ax_{R} of the source input x_{S} and the relay input x_{R}.[7]

[7] Another equivalent problem could be obtained by defining $\hat{C} = \bar{C} + D_{\mathrm{S}}^{\mathrm{H}}C_{\mathrm{V}}D_{\mathrm{S}}$. In fact, the resulting optimization problem would correspond to the case where $u = q$ such that $x_{\mathrm{S}} = u + w$ with $w = Ax_{\mathrm{R}} + v$ and $\hat{C}$ being equal to the joint covariance matrix of w and x_{R}.

We remark that if $\boldsymbol{C}_{\mathrm{V}} = \mathbf{0}$, the problems given in (3.72) and (3.73) reduce to the DF rate maximization problems given in (3.24) and (3.26), respectively. As opposed to the latter, the optimization problems that determine R_{PDF} are not convex though. This is because of the term $-\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)$ in the first constraint on R, which results from the fact that $\boldsymbol{v}$ must be considered as interference at the relay.

3.4.3 Suboptimal Partial Decode-and-Forward Rates

Since evaluating R_{PDF} would thus require to solve a difficult nonconvex optimization problem, we instead consider two approaches that yield suboptimal solutions to the rate maximization problem given in (3.72). The first one is based on *zero-forcing* (ZF) the interference the relay would suffer from, i.e., the part of the source input the relay is not supposed to decode, which is a common way to deal with interference in wireless communication systems, cf. [9, 50, 128], for example. The second approach uses the so-called *inner approximation algorithm* (IAA), which solves a sequence of approximating convex optimization problems instead of the original nonconvex one [88].

3.4.3.1 Zero-Forcing PDF Approach

In order to suppress the interference the relay would otherwise suffer from, we introduce a ZF receive filter $\boldsymbol{G}$ at the relay and require that all possible realizations of $\boldsymbol{v}$ satisfy $\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{v} = \mathbf{0}$. Because $\boldsymbol{v}$ is a zero-mean proper complex Gaussian random vector, this condition is equivalent to

$$\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} = \mathbf{0}, \tag{3.74}$$

which is linear in $\boldsymbol{C}_{\mathrm{V}}$. By adding the ZF condition to (3.72), the PDF rate maximization problem hence becomes a convex problem provided that $\boldsymbol{G}$ is fixed. Naturally, this raises the question of what is a good choice for $\boldsymbol{G}$. As there is no clear-cut answer from theory, two different types of ZF receive filters are considered below. The first one performs an antenna selection for signal reception at the relay, whereas the second one uses a selection of left singular vectors of the channel gain matrix $\boldsymbol{H}_{\mathrm{SR}}$ to form $\boldsymbol{G}$.

Antenna Selection—Let $\mathcal{A} = \{a_1, \dots, a_{|\mathcal{A}|}\}$ be an index set such that $\emptyset \subseteq \mathcal{A} \subseteq \mathcal{I}_{N_{\mathrm{R}}}$, where $\mathcal{I}_{N_{\mathrm{R}}} = \{1, \dots, N_{\mathrm{R}}\}$, and let $\boldsymbol{e}_i \in \{0,1\}^{N_{\mathrm{R}}}$ denote the i-th canonical unit vector of dimension N_{R}. Then, if we apply the relay receive filter

$$\boldsymbol{G}_{\mathrm{AS}} = \boldsymbol{F}_{\mathcal{A}} = [\boldsymbol{e}_{a_1}, \dots, \boldsymbol{e}_{a_{|\mathcal{A}|}}]^{\mathrm{H}} \in \{0,1\}^{|\mathcal{A}| \times N_{\mathrm{R}}}, \tag{3.75}$$

only a selection $\mathcal{A} \subseteq \mathcal{I}_{N_{\mathrm{R}}}$ of all available antennas is used for signal reception at the relay. For this choice of $\boldsymbol{G}$, the ZF condition reads as

$$\boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}} = \mathbf{0}_{|\mathcal{A}| \times |\mathcal{A}|}, \tag{3.76}$$

or equivalently, $\mathrm{range}(\boldsymbol{C}_{\mathrm{V}}) \subseteq \mathrm{null}(\boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}})$. Since $\boldsymbol{G}_{\mathrm{AS}}\boldsymbol{G}_{\mathrm{AS}}^{\mathrm{H}} = \boldsymbol{F}_{\mathcal{A}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}} = \mathbf{I}_{|\mathcal{A}|}$, it is straightforward to verify that all rates up to

$$
\begin{aligned}
\tilde{R}_{\mathrm{AS}}^{\mathcal{A}} = \max_{R,\check{\boldsymbol{C}},\boldsymbol{C}_{\mathrm{Q}},\boldsymbol{C}_{\mathrm{V}}} \quad & R \\
\text{s.t.} \quad & R \leq \log\det\left(\mathbf{I} + \boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}}\right) + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
& R \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\check{\boldsymbol{C}}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\
& \check{\boldsymbol{C}} - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}} = \mathbf{0}, \\
& \mathrm{tr}(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{D}_{\mathrm{S}}\check{\boldsymbol{C}}\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\check{\boldsymbol{C}}\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}
\end{aligned}
\tag{3.77}
$$

can be achieved by means of the considered ZF PDF scheme with relay receive filter $\boldsymbol{G}_{\mathrm{AS}} = \boldsymbol{F}_{\mathcal{A}}$.

Now, let $\check{\boldsymbol{C}}^{\mathcal{A}}, \boldsymbol{C}_{\mathrm{Q}}^{\mathcal{A}}, \boldsymbol{C}_{\mathrm{V}}^{\mathcal{A}}$ denote the optimizers of the convex optimization problem above. It then follows from the data processing inequality [24, Theorem 2.8.1] that for given channel inputs, the information rate on the source-to-relay link cannot be improved by applying a relay receive filter. Since $\check{\boldsymbol{C}}^{\mathcal{A}}, \boldsymbol{C}_{\mathrm{Q}}^{\mathcal{A}}, \boldsymbol{C}_{\mathrm{V}}^{\mathcal{A}}$ also satisfy the constraints of (3.72), this implies that by dropping the ZF filter, all rates smaller than or equal to

$$
\begin{aligned}
R_{\mathrm{AS}}^{\mathcal{A}} = \min\Big\{ & \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}}^{\mathcal{A}} + \boldsymbol{C}_{\mathrm{V}}^{\mathcal{A}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\
& + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}^{\mathcal{A}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}^{\mathcal{A}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\check{\boldsymbol{C}}^{\mathcal{A}}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right)\Big\}
\end{aligned}
\tag{3.78}
$$

are achievable with the selection $\mathcal{A}$, where $R_{\mathrm{AS}}^{\mathcal{A}} \geq \tilde{R}_{\mathrm{AS}}^{\mathcal{A}}$. The best rate we can achieve using the proposed antenna selection ZF approach is then given by

$$
R_{\mathrm{AS}} = \max_{\emptyset \subseteq \mathcal{A} \subseteq \mathcal{I}_{N_{\mathrm{R}}}} R_{\mathrm{AS}}^{\mathcal{A}}. \tag{3.79}
$$

Because $R_{\mathrm{AS}}^{\mathcal{A}}$ needs to be evaluated for any possible selection $\emptyset \subseteq \mathcal{A} \subseteq \mathcal{I}_{N_{\mathrm{R}}}$, we have to solve $2^{N_{\mathrm{R}}}$ convex optimization problems in order to determine R_{AS}. Note also that $R_{\mathrm{AS}} \geq \max\{R_{\mathrm{DF}}, R_{\mathrm{P2P}}\}$. This is easily verified as R_{P2P} is obtained by letting $\mathcal{A} = \emptyset$ and $\check{\boldsymbol{C}} = \mathbf{0}$ in (3.77), whereas choosing $\mathcal{A} = \mathcal{I}_{N_{\mathrm{R}}}$ and $\boldsymbol{C}_{\mathrm{V}} = \mathbf{0}$ yields R_{DF}.

Remark 3.6. For actually solving the problem given in (3.77), the ZF condition should be rewritten as $\mathrm{tr}(\boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}}) = 0$. Since $\boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}} \succcurlyeq \mathbf{0}$, this is equivalent to $\boldsymbol{F}_{\mathcal{A}}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{F}_{\mathcal{A}}^{\mathrm{H}} = \mathbf{0}$ and avoids specifying redundant constraints.

SVD-Based Zero-Forcing—Obviously, the design of $\boldsymbol{G}_{\mathrm{AS}}$ does not take into account the channel properties. Therefore, we also want to consider a more sophisticated relay receive filter that is based on the SVD of the source-to-relay channel gain matrix. To this end, let the SVD of $\boldsymbol{H}_{\mathrm{SR}}$ be given as

$$
\boldsymbol{H}_{\mathrm{SR}} = \boldsymbol{U}_{\mathrm{SR}}\boldsymbol{\Sigma}_{\mathrm{SR}}\boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}}, \tag{3.80}
$$

where $\boldsymbol{U}_{\mathrm{SR}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ and $\boldsymbol{V}_{\mathrm{SR}} \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$ are the unitary matrices containing the left and right singular vectors of $\boldsymbol{H}_{\mathrm{SR}}$, respectively, and where the diagonal matrix $\boldsymbol{\Sigma}_{\mathrm{SR}} \in \mathbb{R}_{+}^{N_{\mathrm{R}} \times N_{\mathrm{S}}}$ is assumed to contain the singular values of $\boldsymbol{H}_{\mathrm{SR}}$ in descending order. Furthermore, let $\mathcal{B}$ denote an index set such that $\emptyset \subseteq \mathcal{B} \subseteq \mathcal{I}_{\mathrm{rank}(\boldsymbol{H}_{\mathrm{SR}})}$. For a given selection $\mathcal{B}$, we then choose the ZF receive filter based on the left singular vectors of $\boldsymbol{H}_{\mathrm{SR}}$ as

$$\boldsymbol{G}_{\mathrm{SVD}} = \boldsymbol{F}_{\mathcal{B}} \boldsymbol{U}_{\mathrm{SR}}^{\mathrm{H}} \in \mathbb{C}^{|\mathcal{B}| \times N_{\mathrm{R}}}, \tag{3.81}$$

where $\boldsymbol{F}_{\mathcal{B}}$ is defined in the same manner as $\boldsymbol{F}_{\mathcal{A}}$ in (3.75). We remark that receive filters of this type are also considered in [10, 146], where they are used to implement ZF beamforming schemes for the Gaussian MIMO broadcast channel.

For this choice of $\boldsymbol{G}$, the ZF condition requires that $\mathrm{range}(\boldsymbol{C}_{\mathrm{V}}) \subseteq \mathrm{null}(\boldsymbol{F}_{\mathcal{B}} \boldsymbol{U}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{SR}})$, which can equivalently be expressed as

$$\boldsymbol{F}_{\mathcal{B}} \boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{C}_{\mathrm{V}} \boldsymbol{V}_{\mathrm{SR}} \boldsymbol{F}_{\mathcal{B}}^{\mathrm{H}} = \boldsymbol{0}_{|\mathcal{B}| \times |\mathcal{B}|} \tag{3.82}$$

since $\mathrm{null}(\boldsymbol{F}_{\mathcal{B}} \boldsymbol{U}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{SR}}) = \mathrm{null}(\boldsymbol{F}_{\mathcal{B}} \boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}})$ for $\mathcal{B} \subseteq \mathcal{I}_{\mathrm{rank}(\boldsymbol{H}_{\mathrm{SR}})}$. Moreover, note that $\boldsymbol{G}_{\mathrm{SVD}}$ also satisfies $\boldsymbol{G}_{\mathrm{SVD}} \boldsymbol{G}_{\mathrm{SVD}}^{\mathrm{H}} = \boldsymbol{F}_{\mathcal{B}} \boldsymbol{U}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{U}_{\mathrm{SR}} \boldsymbol{F}_{\mathcal{B}}^{\mathrm{H}} = \mathbf{I}_{|\mathcal{B}|}$. By applying the PDF scheme with ZF receive filter $\boldsymbol{G}_{\mathrm{SVD}} = \boldsymbol{F}_{\mathcal{B}} \boldsymbol{U}_{\mathrm{SR}}^{\mathrm{H}}$ at the relay, we can thus achieve all rates $R \le \tilde{R}_{\mathrm{SVD}}^{\mathcal{B}}$, where

$$\begin{aligned}
\tilde{R}_{\mathrm{SVD}}^{\mathcal{B}} = &\max_{R, \check{\boldsymbol{C}}, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}} R \\
&\text{s.t.} \quad R \le \log\det\left(\mathbf{I} + \boldsymbol{F}_{\mathcal{B}} \boldsymbol{\Sigma}_{\mathrm{SR}} \boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{C}_{\mathrm{Q}} \boldsymbol{V}_{\mathrm{SR}} \boldsymbol{\Sigma}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{F}_{\mathcal{B}}^{\mathrm{H}}\right) + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{V}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
&\qquad R \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{V}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}} \check{\boldsymbol{C}} \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\
&\qquad \check{\boldsymbol{C}} - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}} \boldsymbol{C}_{\mathrm{Q}} \boldsymbol{D}_{\mathrm{S}} \succcurlyeq \boldsymbol{0}, \quad \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}} \succcurlyeq \boldsymbol{0}, \quad \boldsymbol{F}_{\mathcal{B}} \boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{C}_{\mathrm{V}} \boldsymbol{V}_{\mathrm{SR}} \boldsymbol{F}_{\mathcal{B}}^{\mathrm{H}} = \boldsymbol{0}, \\
&\qquad \mathrm{tr}(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{D}_{\mathrm{S}} \check{\boldsymbol{C}} \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \le P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}} \check{\boldsymbol{C}} \boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \le P_{\mathrm{R}}.
\end{aligned} \tag{3.83}$$

Let $\check{\boldsymbol{C}}^{\mathcal{B}}, \boldsymbol{C}_{\mathrm{Q}}^{\mathcal{B}}, \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}}$ denote the optimizers of this optimization problem. Using the same arguments as for the antenna selection ZF approach, it follows that all rates up to

$$\begin{aligned}
R_{\mathrm{SVD}}^{\mathcal{B}} = \min\Big\{&\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}}^{\mathcal{B}} + \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}}) \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\
&+ \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}} \check{\boldsymbol{C}}^{\mathcal{B}} \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right)\Big\}
\end{aligned} \tag{3.84}$$

are achievable, where $R_{\mathrm{SVD}}^{\mathcal{B}} \ge \tilde{R}_{\mathrm{SVD}}^{\mathcal{B}}$ due to the data processing inequality. The best rate we can achieve with the SVD-based ZF approach is then obtained by considering all possible selections $\mathcal{B}$, i.e.,

$$R_{\mathrm{SVD}} = \max_{\emptyset \subseteq \mathcal{B} \subseteq \mathcal{I}_{\mathrm{rank}(\boldsymbol{H}_{\mathrm{SR}})}} R_{\mathrm{SVD}}^{\mathcal{B}}. \tag{3.85}$$

While it is easily shown that $R_{\mathrm{SVD}} \ge \max\{R_{\mathrm{DF}}, R_{\mathrm{P2P}}\}$ again, we cannot make a general statement about whether R_{SVD} is smaller or larger than R_{AS}. However, note that choosing $\mathcal{A} = \mathcal{B} = \emptyset$ yields $R_{\mathrm{AS}}^{\mathcal{A}} = R_{\mathrm{SVD}}^{\mathcal{B}} = R_{\mathrm{P2P}}$, whereas $R_{\mathrm{AS}}^{\mathcal{A}} = R_{\mathrm{SVD}}^{\mathcal{B}} \geq R_{\mathrm{DF}}$ if $\mathcal{A} = \mathcal{I}_{N_\mathrm{R}}$ and $\mathcal{B} = \mathcal{I}_{\mathrm{rank}(\boldsymbol{H}_{\mathrm{SR}})}$ are the maximum index sets. Finally, we remark that evaluating R_{SVD} requires to solve at most the same number of convex optimization problems that are necessary to determine R_{AS} because $\mathrm{rank}(\boldsymbol{H}_{\mathrm{SR}}) \leq \min\{N_\mathrm{S}, N_\mathrm{R}\}$.

3.4.3.2 Inner Approximation Algorithm

In contrast to the ZF approach, where the PDF rate maximization problem is treated from the perspective of a communications engineer, we now consider a general mathematical approach to deal with nonconvex optimization problems. In particular, we apply the inner approximation algorithm (IAA), which is explained in more detail in Appendix A.1, to the problem given in (3.72).

For this purpose, we first rewrite the nonconvex PDF rate maximization problem in *hypograph form*, where the compact constraint set is specified by the seven inequality constraint functions $g_1, \ldots, g_7$ such that $g_i(R, \check{\boldsymbol{C}}, \boldsymbol{C}_\mathrm{Q}, \boldsymbol{C}_\mathrm{V}) \geq 0, i \in \{1, \ldots, 7\}$:[8]

$$\begin{aligned} R_{\mathrm{PDF}} = \max_{R, \check{\boldsymbol{C}}, \boldsymbol{C}_\mathrm{Q}, \boldsymbol{C}_\mathrm{V}} \quad & R \\ \text{s.t.} \quad & g_1(\cdot) = \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_\mathrm{V}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ & \qquad + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_\mathrm{V}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) - R \geq 0, \\ & g_2(\cdot) = \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_\mathrm{V}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\check{\boldsymbol{C}}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right) - R \geq 0, \\ & g_3(\cdot) = \check{\boldsymbol{C}} - \boldsymbol{D}_\mathrm{S}^{\mathrm{H}}\boldsymbol{C}_\mathrm{Q}\boldsymbol{D}_\mathrm{S} \succcurlyeq \mathbf{0}, \quad g_4(\cdot) = \boldsymbol{C}_\mathrm{Q} \succcurlyeq \mathbf{0}, \quad g_5(\cdot) = \boldsymbol{C}_\mathrm{V} \succcurlyeq \mathbf{0}, \\ & g_6(\cdot) = P_\mathrm{S} - \mathrm{tr}(\boldsymbol{C}_\mathrm{V} + \boldsymbol{D}_\mathrm{S}\check{\boldsymbol{C}}\boldsymbol{D}_\mathrm{S}^{\mathrm{H}}) \geq 0, \quad g_7(\cdot) = P_\mathrm{R} - \mathrm{tr}(\boldsymbol{D}_\mathrm{R}\check{\boldsymbol{C}}\boldsymbol{D}_\mathrm{R}^{\mathrm{H}}) \geq 0. \end{aligned} \tag{3.86}$$

Clearly, the objective function of this problem is linear and the inequality constraint functions are differentiable. Furthermore, all inequality constraint functions except for g_1 are jointly concave in the optimization variables.

The principle of the IAA is to solve a sequence of convex optimization problems in which the nonconcave inequality constraint functions are approximated by concave ones. In order to apply the IAA to the nonconvex problem given in (3.86), we thus have to approximate g_1. To this end, note that g_1 can be expressed as the sum of a concave and a convex function. In particular, with

$$\begin{aligned} g_{1\mathrm{a}}(R, \boldsymbol{C}_\mathrm{Q}, \boldsymbol{C}_\mathrm{V}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_\mathrm{Q} + \boldsymbol{C}_\mathrm{V})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ &\quad + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_\mathrm{V}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) - R, \\ g_{1\mathrm{b}}(\boldsymbol{C}_\mathrm{V}) &= -\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_\mathrm{V}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right), \end{aligned} \tag{3.87}$$

[8]Note that constraints of the form $\boldsymbol{A} \succcurlyeq \boldsymbol{B}$ can be thought of as generalized inequalities induced by the proper cone of positive semidefinite matrices, cf. [11, Section 2.4]. This does not pose any problems for the application of the IAA as the convex optimization theory the IAA relies on extends from ordinary optimization problems to problems with generalized inequalities, cf. [11, Section 5.9].

it follows that

$$g_1(R, \check{C}, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}) = g_{1\mathrm{a}}(R, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}) + g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}), \tag{3.88}$$

where $g_{1\mathrm{a}}$ is concave and $g_{1\mathrm{b}}$ is convex. This is because $\log\det(\mathbf{I} + \boldsymbol{H}\boldsymbol{C}\boldsymbol{H}^{\mathrm{H}})$ is concave in $\boldsymbol{C} \succcurlyeq \mathbf{0}$ as previously discussed in Section 3.2. Since the sum of concave functions is concave again [11, Section 3.2.1], a concave approximation of g_1 can be obtained by approximating only the convex part $g_{1\mathrm{b}}$.

Now, suppose we have solved the approximating convex optimization problem in iteration $k-1$, and assume that $R_{\mathrm{IAA}}^{(k-1)}, \check{\boldsymbol{C}}^{(k-1)}, \boldsymbol{C}_{\mathrm{Q}}^{(k-1)}, \boldsymbol{C}_{\mathrm{V}}^{(k-1)}$ denote the optimizers. The concave function $\tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)})$ which is used to approximate $g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}})$ in iteration k must then satisfy the following properties, cf. Appendix A.1:

$$g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}) \geq \tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)}), \forall \boldsymbol{C}_{\mathrm{V}} \succcurlyeq \mathbf{0}, \tag{3.89a}$$

$$g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) = \tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)}), \tag{3.89b}$$

$$\nabla g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) = \nabla \tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)}). \tag{3.89c}$$

Therefore, we can choose $\tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)})$ to be the first-order Taylor series of $g_{1\mathrm{b}}$ around $\boldsymbol{C}_{\mathrm{V}}^{(k-1)}$, which is an affine function that by definition satisfies (3.89b) and (3.89c), cf. [90, Section 5.4]. Moreover, condition (3.89a) then holds because the first-order Taylor series is a global underestimator for convex functions, cf. [11, Section 3.1.3].

In order to to determine $\tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)})$, we require the gradient of $g_{1\mathrm{b}}$ at $\boldsymbol{C}_{\mathrm{V}}^{(k-1)}$. If, for simplicity, we assume the logarithm to be the natural logarithm here, i.e., $\log = \log_{\mathrm{e}}$, this gradient is given by [38]

$$\nabla g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) = -\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} \left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}^{(k-1)} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1} \boldsymbol{H}_{\mathrm{SR}} \preccurlyeq \mathbf{0}. \tag{3.90}$$

Consequently, we obtain

$$\begin{aligned} \tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)}) &= g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) + \left\langle \nabla g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}), \boldsymbol{C}_{\mathrm{V}} - \boldsymbol{C}_{\mathrm{V}}^{(k-1)} \right\rangle \\ &= c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) - \mathrm{tr}\Big(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} \left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}^{(k-1)} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1} \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}\Big), \end{aligned} \tag{3.91}$$

where the standard inner product between complex matrices $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{B})$ has been used, cf. [91, Section 5.3], and where

$$\begin{aligned} c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) &= g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) - \left\langle \nabla g_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}), \boldsymbol{C}_{\mathrm{V}}^{(k-1)} \right\rangle \\ &= -\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}^{(k-1)} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) + \mathrm{tr}\Big(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} \left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}^{(k-1)} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1} \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{V}}^{(k-1)}\Big) \end{aligned} \tag{3.92}$$

has been introduced to collect all terms of the Taylor series that do not depend on $\boldsymbol{C}_{\mathrm{V}}$. Accordingly, the concave function we use to approximate the inequality constraint function g_1 in iteration k is given by

$$\begin{aligned}\tilde{g}_1(R, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)}) &= g_{1\mathrm{a}}(R, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}) + \tilde{g}_{1\mathrm{b}}(\boldsymbol{C}_{\mathrm{V}}; \boldsymbol{C}_{\mathrm{V}}^{(k-1)}) \\ &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) \\ &\quad - R + c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) - \mathrm{tr}\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(k-1)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\right),\end{aligned} \tag{3.93}$$

which means that the approximating convex optimization problem that needs to be solved in the k-th iteration of the IAA reads as

$$\begin{aligned}R_{\mathrm{IAA}}^{(k)} = \max_{R, \check{\boldsymbol{C}}, \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}}} \quad & R \\ \text{s.t.} \quad & R \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) \\ & \qquad + c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) - \mathrm{tr}\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(k-1)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\right), \\ & R \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\check{\boldsymbol{C}}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & \check{\boldsymbol{C}} - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_{\mathrm{Q}}, \boldsymbol{C}_{\mathrm{V}} \succcurlyeq \mathbf{0}, \\ & \mathrm{tr}(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{D}_{\mathrm{S}}\check{\boldsymbol{C}}\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\check{\boldsymbol{C}}\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}.\end{aligned} \tag{3.94}$$

Note that the sequence $\left\{R_{\mathrm{IAA}}^{(k)}\right\}$ is nondecreasing in k, i.e., $R_{\mathrm{IAA}}^{(k)} \geq R_{\mathrm{IAA}}^{(k-1)}, \forall k \in \mathbb{N}$, as the optimizers of iteration $k-1$ are also feasible for the approximating convex problem in iteration k. Since any sequence element is upper bounded by C_{CSB}, we can conclude that $\left\{R_{\mathrm{IAA}}^{(k)}\right\}$ converges. Beyond this obvious result, the following can be stated with regard to the convergence of the IAA:

Theorem 3.8. *The IAA stops at a Karush–Kuhn–Tucker (KKT) point of the original nonconvex problem, or the limit of any convergent subsequence is a KKT point.*

Proof. See [88]. □

That is, if the optimization variables of the approximating convex problems converge, the IAA yields a KKT point of the original optimization problem. However, since the original problem is nonconvex, the KKT conditions are not always sufficient for local optimality so that a KKT point is not necessarily a local maximizer.

In addition, note that convergence usually occurs only in the limit as $k \to \infty$, which means that we have to define a practical termination criterion for the IAA. To this end, let $e_s^{(k)} = R_{\mathrm{IAA}}^{(k)} - R_{\mathrm{IAA}}^{(k-s)}, s \in \mathbb{N}$. Although the nonnegative sequence $\left\{e_s^{(k)}\right\}$ need not be decreasing in k, we can choose the absolute rate improvement $e_s^{(k)}$ or the relative rate improvement $e_s^{(k)}/R_{\mathrm{IAA}}^{(k)}$ over the last s iterations as termination criterion. In this case, the IAA stops after iteration $K > s$ if

$$e_s^{(K)} < \delta \qquad \text{or} \qquad \frac{e_s^{(K)}}{R_{\mathrm{IAA}}^{(K)}} < \delta, \tag{3.95}$$

where $\delta > 0$ specifies a predefined accuracy. We remark that choosing $s = 1$ causes the IAA to stop if there is only one iteration with little improvement in the objective value. As $e_s^{(k)} \nleq e_s^{(k-1)}$ in general, better results are usually obtained by choosing $s > 1$.

Now, suppose the absolute or the relative rate improvement is used as termination criterion and the IAA stops after iteration K. Then, the rate

$$R_{\mathrm{IAA}} = \min\Big\{\log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{Q}}^{(K)} + \boldsymbol{C}_{\mathrm{V}}^{(K)})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\Big) - \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(K)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\Big) \\ + \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}^{(K)}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\Big), \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}^{(K)}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\check{\boldsymbol{C}}^{(K)}\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\Big)\Big\} \quad (3.96)$$

can be achieved by means of the PDF strategy, where $R_{\mathrm{PDF}} \geq R_{\mathrm{IAA}} \geq R_{\mathrm{IAA}}^{(K)}$. Note that the last inequality is due to the fact that the approximated constraint is always stricter than the original one. Moreover, since the IAA uses local approximations of the nonconcave inequality constraint function, R_{IAA} strongly depends on the initialization, i.e., on $\boldsymbol{C}_{\mathrm{V}}^{(0)}$. However, by appropriate choices of $\boldsymbol{C}_{\mathrm{V}}^{(0)}$, it can easily be guaranteed that $R_{\mathrm{IAA}} \geq R_{\mathrm{DF}}$ and/or $R_{\mathrm{IAA}} \geq R_{\mathrm{P2P}}$, cf. Section 4.2.2.

3.4.4 Optimal Partial Decode-and-Forward Rates for Special Cases

Since an algorithm to (efficiently) compute the optimal solution of the optimization given in (3.72) has yet to be derived, the maximum achievable PDF rate for the Gaussian MIMO relay channel cannot always be determined. Therefore, the focus of this section is on special cases for which it is possible to evaluate R_{PDF}.

As a first step towards this end, note that C_{CSB}, R_{DF}, and R_{P2P} can all be determined as the solutions of convex optimization problems, cf. Proposition 3.3, Proposition 3.5, and [122], respectively. We can therefore evaluate R_{PDF} whenever the optimal PDF strategy is equivalent to the DF strategy or direct transmission, or if the PDF strategy achieves the CSB. Recall that for the discrete memoryless relay channel, such special cases include the degraded relay channel ($R_{\mathrm{PDF}} = R_{\mathrm{DF}} = C_{\mathrm{CSB}}$), the reversely degraded relay channel ($R_{\mathrm{PDF}} = R_{\mathrm{P2P}} = C_{\mathrm{CSB}}$), the semideterministic relay channel ($R_{\mathrm{PDF}} = C_{\mathrm{CSB}}$), and the relay channel with orthogonal sender components ($R_{\mathrm{PDF}} = C_{\mathrm{CSB}}$), cf. Section 2.4. Unfortunately, only one of these four special cases is also relevant for the Gaussian MIMO relay channel considered in this work. In particular:

(a) The Gaussian MIMO relay channel is never degraded according to the system model defined in Section 3.1. The reason for this is that the Markov chain condition $\boldsymbol{x}_{\mathrm{S}} \leftrightarrow (\boldsymbol{x}_{\mathrm{R}}, \boldsymbol{y}_{\mathrm{R}}) \leftrightarrow \boldsymbol{y}_{\mathrm{D}}$ can never be satisfied if the noise vectors $\boldsymbol{n}_{\mathrm{R}}$ and $\boldsymbol{n}_{\mathrm{D}}$ are independent.

(b) Likewise, the Gaussian MIMO relay channel is never reversely degraded.

(c) The Gaussian MIMO relay channel is never semideterministic according to the system model defined in Section 3.1. This is because the defining condition of the semideterministic relay channel requires that $\boldsymbol{y}_{\mathrm{R}}$ be a function of $(\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}})$, which is not satisfied unless there is no noise at the relay.

(d) The definition of the relay channel with orthogonal sender components can be extended to the Gaussian MIMO relay channel as follows:

Definition 3.2. The Gaussian MIMO relay channel is said to have *orthogonal sender components* if $\operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{row}(\boldsymbol{H}_{SR}) \perp \operatorname{row}(\boldsymbol{H}_{SD})$, i.e., if $\operatorname{row}(\boldsymbol{H}_{S\{RD\}})$ is the direct sum of the two orthogonal subspaces $\operatorname{row}(\boldsymbol{H}_{SR})$ and $\operatorname{row}(\boldsymbol{H}_{SD})$.

Remark 3.7. Here, we assume that $\mathcal{X}_S = \operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{null}^{\perp}(\boldsymbol{H}_{S\{RD\}})$. This is without loss of generality, however, as $\boldsymbol{y}_R = \boldsymbol{n}_R$ and $\boldsymbol{y}_D = \boldsymbol{H}_{RD}\boldsymbol{x}_R + \boldsymbol{n}_D$ for any $\boldsymbol{x}_S \in \operatorname{null}(\boldsymbol{H}_{S\{RD\}})$. Similarly, we may also assume that $\mathcal{X}'_S = \operatorname{row}(\boldsymbol{H}_{SR})$ and $\mathcal{X}''_S = \operatorname{row}(\boldsymbol{H}_{SD})$.

Note that according to this definition, the Gaussian MIMO relay channel has orthogonal sender components if and only if $\operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{row}(\boldsymbol{H}_{SR}) + \operatorname{row}(\boldsymbol{H}_{SD})$ such that $\operatorname{row}(\boldsymbol{H}_{SR}) \subseteq \operatorname{null}(\boldsymbol{H}_{SD})$ and $\operatorname{row}(\boldsymbol{H}_{SD}) \subseteq \operatorname{null}(\boldsymbol{H}_{SR})$. Moreover, this condition is equivalent to $p(\boldsymbol{y}_D, \boldsymbol{y}_R | \boldsymbol{x}_S, \boldsymbol{x}_R) = p(\boldsymbol{y}_R | \boldsymbol{x}'_S, \boldsymbol{x}_R) p(\boldsymbol{y}_D | \boldsymbol{x}''_S, \boldsymbol{x}_R)$ for all $(\boldsymbol{x}'_S, \boldsymbol{x}''_S, \boldsymbol{x}_R, \boldsymbol{y}_D, \boldsymbol{y}_R) \in \mathcal{X}'_S \times \mathcal{X}''_S \times \mathcal{X}_R \times \mathcal{Y}_D \times \mathcal{Y}_R$, where $\mathcal{X}'_S = \operatorname{row}(\boldsymbol{H}_{SR})$ and $\mathcal{X}''_S = \operatorname{row}(\boldsymbol{H}_{SD})$ constitute an orthogonal decomposition of the source alphabet $\mathcal{X}_S = \operatorname{row}(\boldsymbol{H}_{S\{RD\}})$, cf. Definition 2.1 for the discrete memoryless relay channel. While this shows that Definition 3.2 provides an appropriate characterization of the Gaussian MIMO relay channel with orthogonal sender components, the following proposition states an equivalent condition that is much easier to work with:

Proposition 3.9. *The Gaussian MIMO relay channel has orthogonal sender components if and only if* $\boldsymbol{H}_{SR}\boldsymbol{H}_{SD}^{H} = \boldsymbol{0}$.

Proof. This result follows from the facts that $\operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{row}(\boldsymbol{H}_{SR}) + \operatorname{row}(\boldsymbol{H}_{SD})$ is always satisfied and that both $\operatorname{row}(\boldsymbol{H}_{SR}) = \operatorname{null}^{\perp}(\boldsymbol{H}_{SR}) = \operatorname{range}(\boldsymbol{H}_{SR}^{H}) \subseteq \operatorname{null}(\boldsymbol{H}_{SD})$ and $\operatorname{row}(\boldsymbol{H}_{SD}) = \operatorname{null}^{\perp}(\boldsymbol{H}_{SD}) = \operatorname{range}(\boldsymbol{H}_{SD}^{H}) \subseteq \operatorname{null}(\boldsymbol{H}_{SR})$ are equivalent to the condition stated in the proposition. □

Clearly, this is a very stringent condition on the channel gain matrices $\boldsymbol{H}_{SR}$ and $\boldsymbol{H}_{SD}$, which will not be satisfied unless one assigns nonoverlapping frequency bands to the channels from the source to the relay and the destination.[9] However, the relay then operates in half-duplex mode, which is not what we focus on in this first part of the work. Nevertheless, it can be shown that like for the discrete memoryless case, the PDF strategy achieves the CSB for the Gaussian MIMO relay channel with orthogonal sender components.

Theorem 3.10. *For the Gaussian MIMO relay channel with orthogonal sender components, the partial decode-and-forward (PDF) strategy achieves the cut-set bound (CSB), i.e.,* $R_{PDF} = C_{CSB}$.

[9]The Gaussian relay channel where the channel from the source to the relay uses a different frequency band is referred to as *sender frequency-division* Gaussian relay channel in [34, Section 16.6.3].

Proof. With $\mathcal{X}_{\mathrm{S}} = \operatorname{row}(\boldsymbol{H}_{\mathrm{S\{RD\}}})$, $\mathcal{X}'_{\mathrm{S}} = \operatorname{row}(\boldsymbol{H}_{\mathrm{SR}})$, and $\mathcal{X}''_{\mathrm{S}} = \operatorname{row}(\boldsymbol{H}_{\mathrm{SD}})$, it follows from Definition 3.2 that $\mathcal{X}_{\mathrm{S}} = \mathcal{X}'_{\mathrm{S}} \oplus \mathcal{X}''_{\mathrm{S}}$. For any $\boldsymbol{x}_{\mathrm{S}} \in \mathcal{X}_{\mathrm{S}}$, there hence exist unique $\boldsymbol{x}'_{\mathrm{S}} \in \mathcal{X}'_{\mathrm{S}}$ and $\boldsymbol{x}''_{\mathrm{S}} \in \mathcal{X}''_{\mathrm{S}}$ such that $\boldsymbol{x}_{\mathrm{S}} = \boldsymbol{x}'_{\mathrm{S}} + \boldsymbol{x}''_{\mathrm{S}}$ [93, Lemma 4.10.1], which means that $\boldsymbol{x}_{\mathrm{S}} \equiv (\boldsymbol{x}'_{\mathrm{S}}, \boldsymbol{x}''_{\mathrm{S}})$. As a result, we can follow the steps for the discrete memoryless relay channel with orthogonal sender components [36] to show that the two mutual information terms characterizing the CSB simplify to

$$\begin{aligned} I(\boldsymbol{x}_{\mathrm{S}}; \boldsymbol{y}_{\mathrm{R}}, \boldsymbol{y}_{\mathrm{D}} | \boldsymbol{x}_{\mathrm{R}}) &= I(\boldsymbol{x}'_{\mathrm{S}}; \boldsymbol{y}_{\mathrm{R}} | \boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}''_{\mathrm{S}}; \boldsymbol{y}_{\mathrm{D}} | \boldsymbol{x}_{\mathrm{R}}), \\ I(\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}; \boldsymbol{y}_{\mathrm{D}}) &= I(\boldsymbol{x}''_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}; \boldsymbol{y}_{\mathrm{D}}). \end{aligned} \tag{3.97}$$

Moreover, $\operatorname{tr}(\boldsymbol{C}_{\mathrm{S}}) = \operatorname{tr}(\boldsymbol{C}'_{\mathrm{S}} + \boldsymbol{C}''_{\mathrm{S}})$ since $\boldsymbol{x}'^{\mathrm{H}}_{\mathrm{S}} \boldsymbol{x}''_{\mathrm{S}} = 0$ for any $(\boldsymbol{x}'_{\mathrm{S}}, \boldsymbol{x}''_{\mathrm{S}}) \in \mathcal{X}' \times \mathcal{X}''_{\mathrm{S}}$. For the Gaussian MIMO relay channel with orthogonal sender components, the CSB is thus equal to

$$C_{\mathrm{CSB}} = \max_{p(\boldsymbol{x}'_{\mathrm{S}}, \boldsymbol{x}''_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}})} \min \left\{ I(\boldsymbol{x}'_{\mathrm{S}}; \boldsymbol{y}_{\mathrm{R}} | \boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}''_{\mathrm{S}}; \boldsymbol{y}_{\mathrm{D}} | \boldsymbol{x}_{\mathrm{R}}), I(\boldsymbol{x}''_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}; \boldsymbol{y}_{\mathrm{D}}) \right\}$$
$$\text{s.t.} \quad \operatorname{tr}(\boldsymbol{C}'_{\mathrm{S}} + \boldsymbol{C}''_{\mathrm{S}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}}, \tag{3.98}$$

where it suffices to maximize over all joint probability density functions of the form $p(\boldsymbol{x}'_{\mathrm{S}}, \boldsymbol{x}''_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}) = p(\boldsymbol{x}'_{\mathrm{S}} | \boldsymbol{x}_{\mathrm{R}}) p(\boldsymbol{x}''_{\mathrm{S}} | \boldsymbol{x}_{\mathrm{R}}) p(\boldsymbol{x}_{\mathrm{R}})$. The rate specified by this problem can be achieved using the PDF strategy with $\boldsymbol{u} = \boldsymbol{q} = \boldsymbol{x}'_{\mathrm{S}}$, i.e., if the auxiliary variable $\boldsymbol{u}$ is defined to only contain the innovative part of $\boldsymbol{x}_{\mathrm{S}}$ the relay must decode. □

For the Gaussian MIMO relay channel with orthogonal sender components, we can therefore conclude from Proposition 3.3 that R_{PDF} can be determined as the solution of a convex optimization problem. As previously stated, however, the Gaussian MIMO relay channel with orthogonal sender components is characterized by the very restrictive condition that the row spaces of $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ be orthogonal.

In the following, we identify two more relevant classes of relay channels for which convex optimization techniques can be used to evaluate R_{PDF}, the *stochastically degraded* and the *reversely stochastically degraded* Gaussian relay channel, where either the source-to-relay or the source-to-destination channel is better than the other one.

Definition 3.3. The Gaussian MIMO relay channel is said to be *stochastically degraded* if $\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SR}} \boldsymbol{Z}^{-1}_{\mathrm{R}} \boldsymbol{H}_{\mathrm{SR}} \succcurlyeq \boldsymbol{H}^{\mathrm{H}}_{\mathrm{SD}} \boldsymbol{Z}^{-1}_{\mathrm{D}} \boldsymbol{H}_{\mathrm{SD}}$. Similarly, the Gaussian MIMO relay channel is said to be *reversely stochastically degraded* if $\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SR}} \boldsymbol{Z}^{-1}_{\mathrm{R}} \boldsymbol{H}_{\mathrm{SR}} \preccurlyeq \boldsymbol{H}^{\mathrm{H}}_{\mathrm{SD}} \boldsymbol{Z}^{-1}_{\mathrm{D}} \boldsymbol{H}_{\mathrm{SD}}$. If the Gaussian MIMO relay channel is either stochastically or reversely stochastically degraded, it is said to be of *stochastically degraded nature.*

For stochastically degraded and reversely stochastically degraded Gaussian relay channels, we show below that $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$ and $R_{\mathrm{PDF}} = R_{\mathrm{P2P}}$, respectively. Of course, this immediately implies that the maximum achievable PDF rate for any Gaussian relay channel of stochastically degraded nature can be determined as the solution of a convex optimization problem, cf. Proposition 3.5 and [122]. In order to prove the two results, which are formalized in Theorems 3.12 and 3.13, we make use of the following lemma (which was already applied in the derivation of Theorem 3.7):

LEMMA 3.11. *Let* $\boldsymbol{H}_1 \in \mathbb{C}^{N_1 \times M}$ *and* $\boldsymbol{H}_2 \in \mathbb{C}^{N_2 \times M}$*, then* $\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{H}_1 \succcurlyeq \boldsymbol{H}_2^{\mathrm{H}}\boldsymbol{H}_2$ *if and only if there exists an* $\boldsymbol{M} \in \mathbb{C}^{N_2 \times N_1}$ *such that* $\boldsymbol{H}_2 = \boldsymbol{M}\boldsymbol{H}_1$ *and* $\boldsymbol{M}\boldsymbol{M}^{\mathrm{H}} \preccurlyeq \mathbf{I}_{N_2}$*.*

Proof. See [114, Lemma 5]. □

THEOREM 3.12. *If the Gaussian MIMO relay channel is stochastically degraded, the optimal partial decode-and-forward (PDF) strategy is equivalent to the decode-and-forward (DF) strategy, i.e.,* $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$*.*

Proof. Applying Lemma 3.11 to $\boldsymbol{H}_1 = \boldsymbol{Z}_{\mathrm{R}}^{-1/2}\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_2 = \boldsymbol{Z}_{\mathrm{D}}^{-1/2}\boldsymbol{H}_{\mathrm{SD}}$, it follows that the Gaussian MIMO relay channel is stochastically degraded if and only if there exists an $\boldsymbol{M} \in \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{R}}}$ such that $\boldsymbol{H}_{\mathrm{SD}} = \boldsymbol{M}\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{M}\boldsymbol{Z}_{\mathrm{R}}\boldsymbol{M}^{\mathrm{H}} \preccurlyeq \boldsymbol{Z}_{\mathrm{D}}$. We thus have

$$\begin{aligned}
&I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) = I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{n}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \\
&= I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) = I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{M}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \\
&\overset{(a)}{=} I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}) + \tilde{\boldsymbol{n}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \overset{(b)}{\leq} I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}})|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \\
&\overset{(c)}{\leq} I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) = I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}),
\end{aligned} \tag{3.99}$$

where in (a) we have defined $\tilde{\boldsymbol{n}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_{\mathrm{D}} - \boldsymbol{M}\boldsymbol{Z}_{\mathrm{R}}\boldsymbol{M}^{\mathrm{H}})$ to be independent of $\boldsymbol{x}_{\mathrm{S}}$, $\boldsymbol{x}_{\mathrm{R}}$, and $\boldsymbol{n}_{\mathrm{R}}$, (b) is due to the fact that $\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}}+\boldsymbol{n}_{\mathrm{R}})+\tilde{\boldsymbol{n}}$ is a degraded version of $\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}}+\boldsymbol{n}_{\mathrm{R}})$, and (c) follows from the data processing inequality [24, Theorem 2.8.1] and the Markov chain $\boldsymbol{x}_{\mathrm{S}} \leftrightarrow \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}} \leftrightarrow \boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S}} + \boldsymbol{n}_{\mathrm{R}})$. As a consequence,

$$\begin{aligned}
I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) &\leq I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \\
&= I(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) = I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}),
\end{aligned} \tag{3.100}$$

where the last equation is due to the Markov relationship $\boldsymbol{u} \leftrightarrow (\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}) \leftrightarrow (\boldsymbol{y}_{\mathrm{R}}, \boldsymbol{y}_{\mathrm{D}})$. Now, let $\mathcal{P} = \{p(\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}) : \mathrm{tr}(\boldsymbol{C}_{\mathrm{S}}) \leq P_{\mathrm{S}}, \mathrm{tr}(\boldsymbol{C}_{\mathrm{R}}) \leq P_{\mathrm{R}}\}$ denote the set of feasible channel input distributions, then

$$R_{\mathrm{PDF}} \leq \max_{p(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}}) \in \mathcal{P}} \min\left\{I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}), I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}})\right\} = R_{\mathrm{DF}}. \tag{3.101}$$

But since $R_{\mathrm{PDF}} \geq R_{\mathrm{DF}}$ in general, it follows that $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$. □

Note that the proof of Theorem 3.12 does not rely on the optimal channel input distribution. Rather, it is based on the fact that the stochastic degradedness condition $\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{SR}} \succcurlyeq \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{D}}^{-1}\boldsymbol{H}_{\mathrm{SD}}$ implies $I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}}) \leq I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})$ for *all* feasible joint distributions $p(\boldsymbol{u}, \boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}})$, cf. [113]. As a consequence, Theorem 3.12 can easily be generalized to discrete memoryless relay channels.

Remark 3.8. If $I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{u},\boldsymbol{x}_\mathrm{R}) \leq I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{R}|\boldsymbol{u},\boldsymbol{x}_\mathrm{R})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})$, the relay is *conditionally more capable* than the destination given $\boldsymbol{u}$ and $\boldsymbol{x}_\mathrm{R}$. Here, the auxiliary variable $\boldsymbol{u}$ and the channel input of the relay $\boldsymbol{x}_\mathrm{R}$ can be interpreted as a time-sharing variable and a channel state, respectively, if we only consider these two mutual information terms. Since the set of feasible joint distributions $p(\boldsymbol{u},\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})$ is convex, the time-sharing variable $\boldsymbol{u}$ is obsolete in the more capable condition above. That is, the above condition is equivalent to the condition that $I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) \leq I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R})$ for all feasible $p(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})$, meaning that the relay is conditionally more capable than the destination given $\boldsymbol{x}_\mathrm{R}$. For Gaussian channels, the fact that the relay is more capable than the destination given a channel state is equivalent to the fact that $\boldsymbol{y}_\mathrm{D}$ is a stochastically degraded version of $\boldsymbol{y}_\mathrm{R}$ given the same channel state, cf. Appendix B.2.2. For discrete memoryless channels, on the other hand, the more capable condition is less strict than the stochastically degraded condition, cf. Appendix B.2.1, suggesting that Theorem 3.12 should in general be formulated using the weaker more capable condition. However, for the Gaussian MIMO relay channel considered here, it is more convenient to use the stochastically degraded condition because it is equivalent to and much easier to work with than the more capable condition.

Theorem 3.13. *If the Gaussian MIMO relay channel is reversely stochastically degraded, the optimal partial decode-and-forward (PDF) strategy is equivalent to direct transmission, i.e.,* $R_\mathrm{PDF} = R_\mathrm{P2P}$.

Proof. It again follows from Lemma 3.11 that the Gaussian MIMO relay channel is reversely stochastically degraded if and only if there exists an $\boldsymbol{M} \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{D}}$ such that $\boldsymbol{H}_\mathrm{SR} = \boldsymbol{M}\boldsymbol{H}_\mathrm{SD}$ and $\boldsymbol{M}\boldsymbol{Z}_\mathrm{D}\boldsymbol{M}^\mathrm{H} \preccurlyeq \boldsymbol{Z}_\mathrm{R}$. Similar to the proof of Theorem 3.12, we can therefore construct a chain of inequalities to show that $I(\boldsymbol{u};\boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R}) \leq I(\boldsymbol{u};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})$, which in turn implies that

$$\begin{aligned} I(\boldsymbol{u};\boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R}) + I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{u},\boldsymbol{x}_\mathrm{R}) &\leq I(\boldsymbol{u};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) + I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{u},\boldsymbol{x}_\mathrm{R}) \\ &= I(\boldsymbol{u},\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) = I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}). \end{aligned} \tag{3.102}$$

If $\mathcal{P}$ again denotes the set of feasible channel input distributions, it hence follows that

$$\begin{aligned} R_\mathrm{PDF} &\leq \max_{p(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})\in\mathcal{P}} \min\left\{I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}), I(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R};\boldsymbol{y}_\mathrm{D})\right\} \\ &= \max_{p(\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})\in\mathcal{P}} I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R}) = R_\mathrm{P2P}, \end{aligned} \tag{3.103}$$

and since $R_\mathrm{PDF} \geq R_\mathrm{P2P}$ in general, we can conclude that $R_\mathrm{PDF} = R_\mathrm{P2P}$. □

Remark 3.9. If $I(\boldsymbol{u};\boldsymbol{y}_\mathrm{R}|\boldsymbol{x}_\mathrm{R}) \leq I(\boldsymbol{u};\boldsymbol{y}_\mathrm{D}|\boldsymbol{x}_\mathrm{R})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_\mathrm{S},\boldsymbol{x}_\mathrm{R})$, the destination is *conditionally less noisy* than the relay given $\boldsymbol{x}_\mathrm{R}$, cf. [124, Definition 6]. Here, $\boldsymbol{x}_\mathrm{R}$ again plays the role of a channel state if we only consider these two mutual information terms. For Gaussian channels, the fact that the destination is less noisy than the relay given a channel state is equivalent to the fact that $\boldsymbol{y}_{\mathrm{R}}$ is a stochastically degraded version of $\boldsymbol{y}_{\mathrm{D}}$ given the same channel state, cf. Appendix B.2.2. But for discrete memoryless channels, the less noisy condition is less strict than the stochastically degraded condition, cf. Appendix B.2.1, so Theorem 3.13 should in general be formulated using the weaker less noisy condition.

Corollary 3.14. *If $N_{\mathrm{S}} = 1$, i.e., if the source is equipped with one antenna, the Gaussian MIMO relay channel is of stochastically degraded nature so that $R_{\mathrm{PDF}} = \max\{R_{\mathrm{DF}}, R_{\mathrm{P2P}}\}$.*

Proof. For $N_{\mathrm{S}} = 1$, the two channel gain matrices $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ become channel vectors, i.e., $\boldsymbol{H}_{\mathrm{SR}} = \boldsymbol{h}_{\mathrm{SR}} \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$ and $\boldsymbol{H}_{\mathrm{SD}} = \boldsymbol{h}_{\mathrm{SD}} \in \mathbb{C}^{N_{\mathrm{D}} \times 1}$. As a consequence, the conditions for stochastic degradedness and reversely stochastic degradedness reduce to the scalar inequalities $\boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{Z}_{\mathrm{R}}^{-1} \boldsymbol{h}_{\mathrm{SR}} \geq \boldsymbol{h}_{\mathrm{SD}}^{\mathrm{H}} \boldsymbol{Z}_{\mathrm{D}}^{-1} \boldsymbol{h}_{\mathrm{SD}}$ and $\boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{Z}_{\mathrm{R}}^{-1} \boldsymbol{h}_{\mathrm{SR}} \leq \boldsymbol{h}_{\mathrm{SD}}^{\mathrm{H}} \boldsymbol{Z}_{\mathrm{D}}^{-1} \boldsymbol{h}_{\mathrm{SD}}$, respectively, of which one is always satisfied. □

While the Gaussian MIMO relay channel is not always of stochastically degraded nature, Corollary 3.14 reveals that this is the case whenever the source is equipped with one antenna. Moreover, we remark that Theorems 3.12 and 3.13 and Corollary 3.14 generalize a result presented in [35], where for the Gaussian relay channel with single-antenna nodes and unit noise variances it was shown that

$$R_{\mathrm{PDF}} = \begin{cases} R_{\mathrm{DF}} & \text{if } |h_{\mathrm{SR}}|^2 \geq |h_{\mathrm{SD}}|^2, \\ R_{\mathrm{P2P}} & \text{if } |h_{\mathrm{SR}}|^2 < |h_{\mathrm{SD}}|^2. \end{cases} \tag{3.104}$$

3.5 Further Results and Bibliographical Notes

The single-antenna Gaussian relay channel (with real-valued channel coefficients) was already considered in the seminal work of Cover and El Gamal [21]. In particular, they established that both the *cut-set bound* (CSB) and the achievable *decode-and-forward* (DF) rate are maximized by jointly Gaussian source and relay inputs. Using this fact, they also formulated the maximization problems that yield C_{CSB} and R_{DF}, where the only optimization variable is the correlation coefficient ρ of the source and relay inputs x_{S} and x_{R}. Later, closed-form expressions for C_{CSB} and R_{DF} were derived by El Gamal et al. [35], who also showed that the optimal *partial decode-and-forward* (PDF) strategy always reduces to DF or direct transmission if all nodes are equipped with a single antenna, cf. (3.104).

The maximum achievable *compress-and-forward* (CF) rate R_{CF} for the Gaussian relay channel has yet to be determined because the optimal distribution of the source input x_{S}, the relay input x_{R}, and the relay quantization $\hat{y}_{\mathrm{R}}$ is unknown. For jointly Gaussian $(x_{\mathrm{S}}, x_{\mathrm{R}}, \hat{y}_{\mathrm{R}})$, on the other hand, the achievable CF rate was for example given by Høst-Madsen and Zhang in [63].[10] Furthermore, Cover and El Gamal's mixed strategy

[10]Unlike for the CSB and the achievable DF and PDF rates, one needs to be careful when applying the information theoretical results for the CF strategy to the Gaussian relay channel. The reason for this is

that combines DF and CF principles was recently applied to the Gaussian relay channel by Luo et al. [87]. They showed that if the signals are restricted to be jointly Gaussian, the maximum rate that can be achieved using this mixed strategy can be strictly greater than the maximum of the achievable DF and CF rates.

Another relay strategy that has often been considered for the Gaussian relay channel is *amplify-and-forward* (AF), which has the relay forward a linear function of its past receive signal, cf. [35, 80, 108], for example. The AF strategy is a special case of *linear relaying*, where the relay input is restricted to be a linear combination of its past receive symbols. Despite its simplicity, linear relaying can sometimes outperform the more sophisticated (P)DF and CF strategies in the Gaussian relay channel [35].

We remark that most of the abovementioned results for the single-antenna Gaussian relay channel can nowadays be found in standard textbooks on network information theory or cooperative communications, cf. [34, Chapter 16], [76, Chapter 9], and [78, Section 4.2], for example.

Bounds on the capacity of the Gaussian MIMO relay channel with perfect channel state information (CSI) were first studied by Wang et al. [132], who exploited matrix inequalities to establish a generally loose upper bound to the CSB. This upper bound can be evaluated by solving a maximization problem over the covariance matrices $\boldsymbol{C}_{\mathrm{S}}, \boldsymbol{C}_{\mathrm{R}}$ and a scalar parameter that is related to the cross correlation of $\boldsymbol{x}_{\mathrm{S}}$ and $\boldsymbol{x}_{\mathrm{R}}$, and it is equal to the CSB if C_{CSB} is achieved by independent source and relay inputs, cf. [46]. In addition to an upper bound on the capacity, several achievable rates based on suboptimal DF schemes were also derived in [132].

The CSB and the maximum achievable DF rate for the Gaussian MIMO relay channel were subsequently also considered by Simoens et al. [116]. However, they replaced the conditional covariance matrix $\boldsymbol{C}_{\mathrm{S|R}}$ by $\boldsymbol{C}_{\mathrm{S}}$ in the CSB and the DF rate maximization problems so that the convex optimization problems resulting from their considerations only yield upper bounds to C_{CSB} and R_{DF}. Ng and Foschini then were the first to show that if perfect CSI is available at all nodes and if the covariance matrix of the relay input is positive definite, C_{CSB} and R_{DF} can be determined as the solutions of convex optimization problems [95]. The key to proving this was to introduce an auxiliary variable and to apply the Schur complement condition for positive definite matrices, which allows to relax and reformulate the equality constraint on the conditional covariance matrix $\boldsymbol{C}_{\mathrm{S|R}}$. The same result was later independently derived by Gerdes and Utschick for the more general case where $\boldsymbol{C}_{\mathrm{R}}$ may also be singular [46].

Employing PDF in the Gaussian MIMO relay channel was first considered by Lo et al. [86], who termed the strategy "transmit-side message splitting". More specifically, they formulated the PDF rate maximization problem for jointly proper complex Gaussian source and relay inputs, but they did not solve the resulting nonconvex optimization problem. Moreover, no attempt was made to characterize the input distribution that

that the derivation of the achievable CF rate for the discrete memoryless relay channel uses the *Markov lemma* [24, Lemma 15.8.1], which requires strong typicality. For Gaussian input distributions, however, the Markov lemma can be generalized, cf. [77, Remark 30].

maximizes the achievable PDF rate. Rather, proper complex Gaussian channel inputs were assumed as part of the system model.

The results presented in Section 3.4 are all due to Gerdes, Hellings, Riemensberger, Weiland, and Utschick [40, 47, 48, 57, 136]. It should be noted that the order in which the results are presented in this work differs from that in which they were derived. In particular, the two suboptimal PDF schemes (cf. Section 3.4.3) were actually derived first, i.e., at a time when the optimal input distribution for the PDF strategy was not yet known. In [48] and [136], proper complex Gaussian channel inputs were hence chosen to obtain tractable expressions and to guarantee that the resulting achievable PDF rates are no smaller than the maximum of R_{DF} and R_{P2P}. Subsequently, the maximum achievable PDF rate R_{PDF} was established for stochastically degraded and reversely stochastically degraded Gaussian relay channels (cf. Section 3.4.4) in [47]. The first step towards finding the input distribution that maximizes the achievable PDF rate for the general Gaussian MIMO relay channel was made in [57], where it was shown that jointly proper source and relay inputs are optimal if the channel inputs are restricted to be complex Gaussian. Furthermore, we remark that the primal decomposition approach, which turned out to be a key ingredient for proving Theorem 3.6, was first used in [57]. Finally, the fact that the maximum achievable PDF rate for the Gaussian MIMO relay channel is attained by jointly proper complex Gaussian source and relay inputs is the most recent result, which was proved in [40].

A suboptimal CF scheme for the Gaussian MIMO relay channel was proposed by Ng and Foschini [95]. Since the joint design of the channel inputs and the quantization at the relay $\hat{\boldsymbol{y}}_{\mathrm{R}}$ seems intractable, they assumed $(\boldsymbol{x}_{\mathrm{S}}, \boldsymbol{x}_{\mathrm{R}}, \hat{\boldsymbol{y}}_{\mathrm{R}})$ to be jointly Gaussian and took a greedy approach which works as follows. The destination performs successive decoding with the relay's message, i.e., the encoded relay quantization, being decoded first. Then, the source covariance $\boldsymbol{C}_{\mathrm{S}}$ and the relay covariance $\boldsymbol{C}_{\mathrm{R}}$ are successively chosen as to maximize the information rates on the source-to-destination channel and the resulting effective relay-to-destination channel, respectively. Eventually, the quantization at the relay is determined based on rate distortion compression.

Compared to the few works on achievable DF, PDF, and CF rates, there is a large number of publications on AF schemes for the Gaussian MIMO relay channel, cf. [103, 105, 121] and references therein, for example. Like for the CF strategy, the problem of determining the maximum achievable AF rate seems intractable. Even if Gaussian channel inputs are assumed, a difficult nonconvex optimization problem is obtained since the AF strategy suffers from noise amplification induced by the relay. Accordingly, various suboptimal AF schemes have been proposed in the literature.

In addition to the topics discussed so far, there are many more interesting aspects of the Gaussian (MIMO) relay channel that go beyond the scope of this work. Two of these, we nevertheless want to mention in the following. First, note that an important assumption of our system model (cf. Section 3.1) is that the noise vectors received at the relay and the destination are independent. While this assumption is often reasonable, it does not account for scenarios where the relay and the destination are interfered with by common random sources. Therefore, Zhang et al. [149] studied the effect of noise correlation on the (single-antenna) Gaussian relay channel.

What is more, another important assumption we make is that the relay is able to perfectly cancel its self-interference. From a practical perspective, a major problem with this assumption is that even if the self-interference signal is assumed to be perfectly known, the transmit signal of the relay may drown out its desired receive signal due to the limited dynamic range of the relay's receiver circuitry. This aspect of full-duplex relaying was for example paid attention to by Day et al. [25], who explicitly modeled the dynamic range limitations of the receiving and the transmitting nodes in the Gaussian MIMO relay channel. Beyond that, we remark that the implementation of in-band full-duplex wireless communication devices has generally received much attention during the last few years. For recent discussions about the challenges and opportunities of full-duplex operation with different self-interference cancellation (mitigation) techniques, we refer to [59, 104] and references therein.

Chapter 4

Numerical Results and Discussion

In this chapter, we provide numerical results for various Gaussian MIMO relay channels in which we compare the achievable *decode-and-forward* (DF) and *partial decode-and-forward* (PDF) rates to each other as well as to the *cut-set bound* (CSB). In particular, the effects of different antenna configurations and relay positions on the achievable rates and the CSB are investigated. Furthermore, we point out practical applications for the results derived in Chapter 3, which were obtained based on the assumptions that perfect channel state information (CSI) is available at all nodes and that the relay is able to completely cancel its self-interference.

After introducing the example scenario we use to generate the channel gain matrices in Section 4.1, numerical results are presented in Section 4.2. More specifically, we first compare the CSB C_{CSB} to the maximum achievable DF rate R_{DF} in Section 4.2.1. This is in order to identify channel conditions for which the DF strategy either performs very well ($R_{\mathrm{DF}} \approx C_{\mathrm{CSB}}$) or very poorly ($R_{\mathrm{DF}} \ll C_{\mathrm{CSB}}$). Whenever the first case occurs, i.e., if R_{DF} closely approaches C_{CSB}, we (approximately) know the capacity of the Gaussian MIMO relay channel so that there is no need to consider other relay strategies. The second case is interesting because if the gap between R_{DF} and C_{CSB} is large, the PDF strategy can potentially achieve much higher rates than the DF strategy.

In Section 4.2.2, we compare rates that can be achieved by means of the PDF strategy to C_{CSB} and R_{DF}. Thereby, we distinguish between antenna configurations for which the Gaussian MIMO relay channel has *disjoint sender components* and those for which it does not. This is possible because for the considered example scenario, there is a simple condition in terms of N_{S}, N_{R}, and N_{D} which almost surely characterizes Gaussian MIMO relay channels with disjoint sender components. In addition, we derive conditions on N_{S}, N_{R}, and N_{D} which almost surely are necessary for the Gaussian MIMO relay channel to be stochastically degraded or reversely stochastically degraded and which show that Gaussian MIMO relay channels with disjoint sender components are not of stochastically degraded nature. For channel conditions where we cannot evaluate R_{PDF}, we employ the two suboptimal approaches presented in Section 3.4.3, the *zero-forcing* (ZF) scheme and the *inner approximation algorithm* (IAA), to obtain achievable PDF rates.

Finally, we conclude this chapter, and hence the first part of this work, in Section 4.3 with a discussion of possible practical applications for the results derived in the previous chapter. That is, we shed some light on how these results can be used in a meaningful way although they are only valid for the idealistic case where perfect CSI is available at all nodes and where the relay, working in full-duplex mode, is able to completely cancel its self-interference.

4.1 Example Scenario

As an example scenario for our simulations, we choose the line network that is depicted in Figure 4.1. The line network is a simple but commonly used network model in which the distance between the source and the destination d_{SD} is normalized to one and where the relay is positioned on the line connecting the source and the destination such that $d_{\mathrm{SR}} = d$ and $d_{\mathrm{RD}} = 1 - d$ for some $d \in (0, 1)$. The advantage of considering this simple scenario is that the geometry of our three-terminal relay network is parameterized by a single scalar parameter.

Moreover, we assume uncorrelated Rayleigh fading, which means that the entries of the channel gain matrices are *independent and identically distributed* (i.i.d.) zero-mean proper (circularly symmetric) complex Gaussian random variables, cf. [17]. The variances of the channel gains are determined according to a simplified path loss model, cf. [50, Section 2.6]. More precisely, if d_{AB} denotes the distance between nodes A and B, the corresponding channel gain matrix is given by

$$\boldsymbol{H}_{\mathrm{AB}} = d_{\mathrm{AB}}^{-\alpha/2} \tilde{\boldsymbol{H}}_{\mathrm{AB}}, \tag{4.1}$$

where α is the pass loss exponent and the elements of $\tilde{\boldsymbol{H}}_{\mathrm{AB}}$ are i.i.d. standard complex Gaussian random variables, i.e., $[\tilde{\boldsymbol{H}}_{\mathrm{AB}}]_{k,\ell} \sim \mathcal{N}_{\mathbb{C}}(0, 1)$, $k \in \{1, \ldots, N_{\mathrm{B}}\}$, $\ell \in \{1, \ldots, N_{\mathrm{A}}\}$. Finally, the additive Gaussian noise is assumed to be white for simplicity, i.e., $\boldsymbol{Z}_{\mathrm{R}} = \mathbf{I}_{N_{\mathrm{R}}}$ and $\boldsymbol{Z}_{\mathrm{D}} = \mathbf{I}_{N_{\mathrm{D}}}$.

We remark that many real wireless propagation environments are not accurately described using the (uncorrelated) Rayleigh fading model, cf. [26, 49] and references therein, for example. However, the numerical results presented in the following section are obtained based on the assumption that perfect CSI is available at all nodes, and our focus is not on how specific MIMO channel models affect the achievable DF or PDF rates for the Gaussian MIMO relay channel. Rather, we are interested in how the geometry and the numbers of antennas the nodes are equipped with affect the performance of the DF and PDF strategies as compared to the CSB and *point-to-point* (P2P) transmission from source to destination. Since the model that is used to generate the channel gain matrices is of secondary importance for this purpose, we assume uncorrelated Rayleigh fading mainly because it is a simple and well-understood fading model. Beyond that, a convenient property of uncorrelated Rayleigh fading is that the channel gain matrices are drawn from a continuous distribution, by which we mean the following:

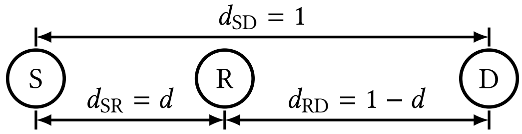


Figure 4.1: Line Network

Definition 4.1. The channel gain matrices are said to be *drawn from a continuous distribution* if the joint distribution of $(\boldsymbol{H}_{\mathrm{SR}}, \boldsymbol{H}_{\mathrm{SD}}, \boldsymbol{H}_{\mathrm{RD}})$ is absolutely continuous with respect to the Lebesgue measure on the probability space $\mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{S}}} \times \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{S}}} \times \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{R}}}$.

Note that if the joint distribution of $(\boldsymbol{H}_{\mathrm{SR}}, \boldsymbol{H}_{\mathrm{SD}}, \boldsymbol{H}_{\mathrm{RD}})$ is absolutely continuous, all channel gain matrices have full rank with probability one. This fact is very convenient since it almost surely allows us to identify Gaussian MIMO relay channels with disjoint sender components by means of a condition that only depends on N_{S}, N_{R}, and N_{D}, cf. Section 4.2.2. Furthermore, it also enables us to derive necessary conditions in terms of N_{S}, N_{R}, and N_{D} which almost surely have to be satisfied for the Gaussian MIMO relay channel to be stochastically degraded or reversely stochastically degraded.

4.2 Numerical Results

As previously mentioned, we want to examine how the geometry of our three-terminal relay network and the numbers of source, relay, and destination antennas affect the achievable DF and PDF rates as compared to the CSB and the capacity of the source-to-destination channel. To this end, we consider eight different antenna configurations, for each of which we vary d between 0.1 (the relay is very close to the source) and 0.9 (the relay is very close to the destination). All further parameters are fixed and remain constant for all simulations. In particular:

- The path loss exponent is set to $\alpha = 4$, which is a typical value for urban macrocell environments or multi-level office buildings [50, Table 2.2].
- The power budgets available at the source and the relay are assumed to be equal and given by $P_{\mathrm{S}} = P_{\mathrm{R}} = 10$, which is a reasonable assumption in wireless ad hoc networks, for example.
- All numerical results are based on the same 1000 independent realizations of $(\tilde{\boldsymbol{H}}_{\mathrm{SR}}, \tilde{\boldsymbol{H}}_{\mathrm{SD}}, \tilde{\boldsymbol{H}}_{\mathrm{RD}})$. For each of these realizations and every considered value of d, we determine the corresponding channel gain matrices $\boldsymbol{H}_{\mathrm{SR}}$, $\boldsymbol{H}_{\mathrm{SD}}$, and $\boldsymbol{H}_{\mathrm{RD}}$ by scaling $\tilde{\boldsymbol{H}}_{\mathrm{SR}}$, $\tilde{\boldsymbol{H}}_{\mathrm{SD}}$, and $\tilde{\boldsymbol{H}}_{\mathrm{RD}}$ according to (4.1), respectively.
- For any value of d, the rate values presented below are therefore averages over 1000 independent channel realizations.

In order to evaluate C_{CSB}, R_{DF}, R_{P2P}, and the achievable PDF rates, the optimization problems yielding these values were solved using CVX [52, 53], a package for specifying and solving convex programs, with either SeDuMi [119] or SDPT3 [126, 127].

Remark 4.1. For some of the generated realizations of $(\tilde{\boldsymbol{H}}_{\mathrm{SR}}, \tilde{\boldsymbol{H}}_{\mathrm{SD}}, \tilde{\boldsymbol{H}}_{\mathrm{RD}})$, it turned out that one or several of the problems could not be solved reliably, e.g., because neither SeDuMi nor SDPT3 could achieve the desired accuracy. In this case, the realization was discarded and a new one was generated.

4.2.1 Cut-Set Bound and Decode-and-Forward

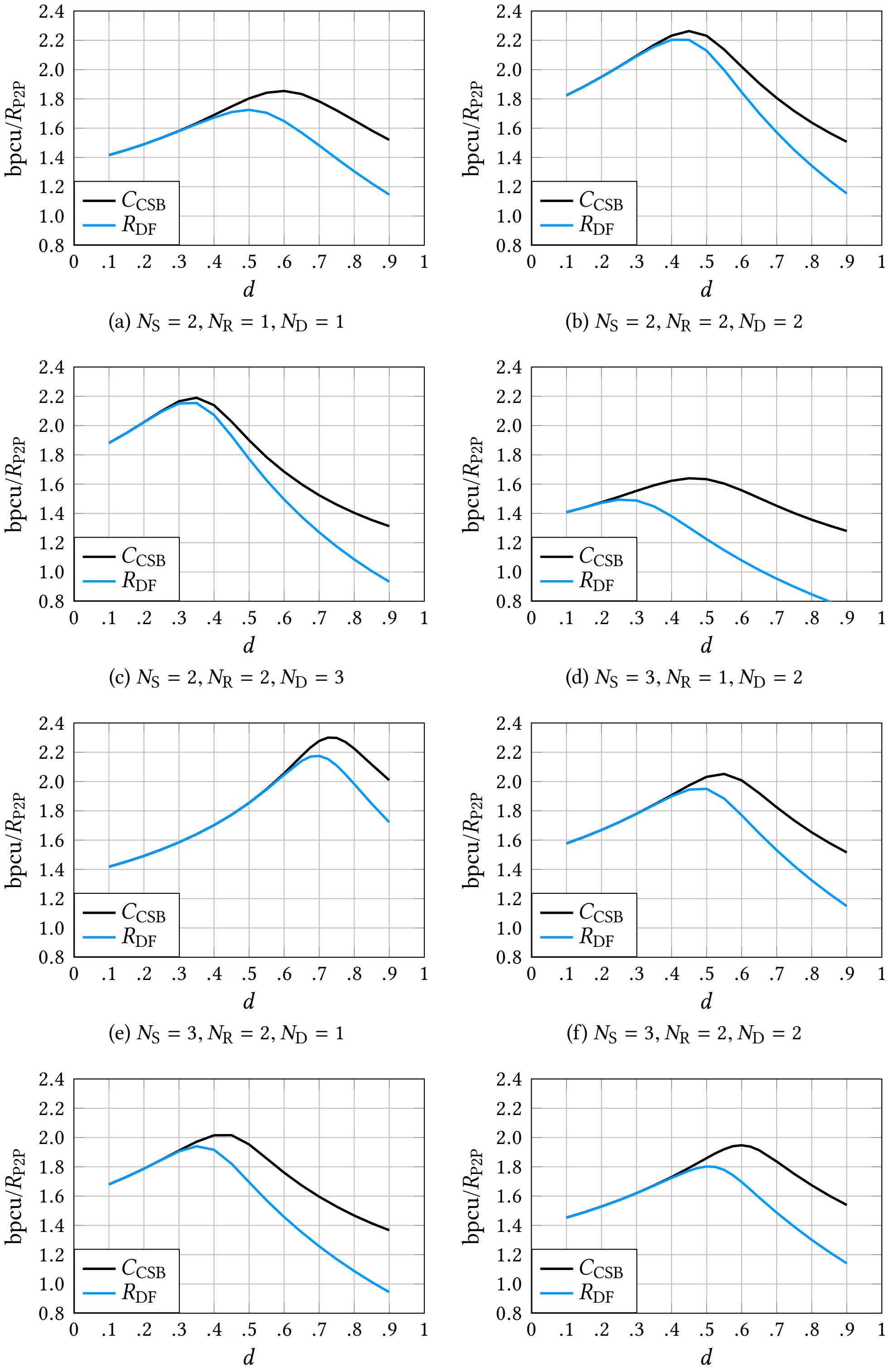
In Figure 4.2, we compare the maximum achievable DF rate to both the CSB and the capacity of the source-to-destination channel for eight different antenna configurations.[1] Note that the results for C_{CSB} and R_{DF} are normalized with respect to R_{P2P}, i.e., $R_{\mathrm{DF}} = 1.0$ means that the maximum achievable DF rate is equal to the capacity of the source-to-destination channel, for example.

The most obvious observation from Figure 4.2 is that R_{DF} closely approaches C_{CSB} as long as d is small enough, i.e., if the relay is close enough to the source. This result is not surprising as the DF strategy is known to perform well whenever the source-to-relay link is strong compared to the source-to-destination and relay-to-destination links. In particular, it has been shown that R_{DF} approaches the capacity of the single-antenna Gaussian relay channel if the relay is close enough to the source [77]. However, what "close enough" means for the Gaussian MIMO relay channel strongly depends on N_{S}, N_{R}, and N_{D} (as well as P_{S}, P_{R}, and α). For example, if $N_{\mathrm{R}} = N_{\mathrm{D}}$, close enough roughly means $d \leq d_{\max} \approx 0.4$ for the parameters we have chosen, where $d_{\max}$ is almost independent of N_{S}, cf. Figures 4.2(a), 4.2(b), 4.2(f), and 4.2(h). If $N_{\mathrm{R}} < N_{\mathrm{D}}$, on the other hand, the relay must be closer to the source, i.e., $d_{\max}$ decreases, cf. Figures 4.2(c), 4.2(d), and 4.2(g), whereas it increases if $N_{\mathrm{R}} > N_{\mathrm{D}}$, cf. Figure 4.2(e).

Another notable observation from Figure 4.2 is that the source-to-relay channel eventually becomes the bottleneck if $N_{\mathrm{R}} < N_{\mathrm{D}}$. That is, the DF strategy performs worse than P2P transmission from source to destination if the relay has fewer antennas than the destination and d becomes too large, with the meaning of "too large" depending on N_{S}, N_{R}, and N_{D} again, cf. Figures 4.2(c), 4.2(d), and 4.2(g). In this case, it is clear that PDF improves on DF since $R_{\mathrm{PDF}} \geq \max\{R_{\mathrm{DF}}, R_{\mathrm{P2P}}\}$ in general. However, note that C_{CSB} decreases for all considered antenna configurations if d becomes too large, i.e., the potential rate gains due to the relay are anyway limited if the relay is too close to the destination.

We can also see from Figure 4.2 that both the maximum potential rate gain, i.e., the maximum of C_{CSB} over all values of d, and the corresponding optimal relay position strongly depend on the antenna configuration. Not surprisingly, the more antennas the relay is equipped with as compared to N_{S} and N_{D}, the larger the maximum potential rate gain. The optimal distance between the source and the relay decreases with the

[1] We do not consider antenna configurations with $N_{\mathrm{S}} = 1$ here because $R_{\mathrm{PDF}} = \max\{R_{\mathrm{DF}}, R_{\mathrm{P2P}}\}$ if the source is equipped with a single antenna, cf. Corollary 3.14. Since PDF can then improve on DF only if the source-to-relay channel is worse than the source-to-destination channel, this case is not of particular interest to us.



(a) $N_S = 2, N_R = 1, N_D = 1$

(b) $N_S = 2, N_R = 2, N_D = 2$

(c) $N_S = 2, N_R = 2, N_D = 3$

(d) $N_S = 3, N_R = 1, N_D = 2$

(e) $N_S = 3, N_R = 2, N_D = 1$

(f) $N_S = 3, N_R = 2, N_D = 2$

(g) $N_S = 3, N_R = 2, N_D = 3$

(h) $N_S = 4, N_R = 2, N_D = 2$

Figure 4.2: Comparison of C_{CSB} and R_{DF} for Gaussian MIMO Relay Channels: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (results averaged over 1000 independent channel realizations)

ratio of $(1 + N_R/N_D)$ and $(1 + N_R/N_S)$, which can be explained as follows. If the first term is small compared to the second one, the potential rate gain due to cooperative transmission of the source and the relay tends to be larger than that due to cooperative reception of the relay and the destination. Hence, the optimal relay position tends to be closer to the source in order to facilitate cooperative transmission. If the first term is large compared to the second one, the opposite is true and the optimal relay position tends to be closer to the destination.

Finally, the most interesting result concerning the question under which conditions the PDF strategy outperforms the DF strategy is the following. The maximum value of R_{DF} over all relay positions is close to the maximum potential rate gain if $N_S \leq N_R$, cf. Figures 4.2(b) and 4.2(c). But for the other six considered antenna configurations where $N_S > N_R$, the gaps between the maximum values of R_{DF} and C_{CSB} are quite large. This suggests that PDF can potentially achieve much higher rates than DF if the source is equipped with more antennas than the relay.

4.2.2 Partial Decode-and-Forward

In this section, we distinguish between antenna configurations for which the Gaussian MIMO relay channel has disjoint sender components and those for which it does not. This is possible because for the considered example scenario with uncorrelated Rayleigh fading, we can identify a simple condition in terms of N_S, N_R, and N_D that almost surely characterizes Gaussian MIMO relay channels with disjoint sender components.

4.2.2.1 Gaussian MIMO Relay Channels with Disjoint Sender Components

Recall from Section 3.4.4 that Gaussian MIMO relay channels with orthogonal sender components are characterized by the very restrictive condition that the row spaces of $\boldsymbol{H}_{SR}$ and $\boldsymbol{H}_{SD}$ be orthogonal. The more general class of Gaussian MIMO relay channels with disjoint sender components is obtained by relaxing this condition as follows:

Definition 4.2. The Gaussian MIMO relay channel is said to have *disjoint sender components* if $\operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{row}(\boldsymbol{H}_{SR}) \oplus \operatorname{row}(\boldsymbol{H}_{SD})$.

Proposition 4.1. *The Gaussian MIMO relay channel has disjoint sender components if and only if* $\operatorname{rank}(\boldsymbol{H}_{SR}) + \operatorname{rank}(\boldsymbol{H}_{SD}) = \operatorname{rank}(\boldsymbol{H}_{S\{RD\}})$.

Proof. Note that by definition, $\operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{row}(\boldsymbol{H}_{SR}) \oplus \operatorname{row}(\boldsymbol{H}_{SD})$ is equivalent to $\operatorname{row}(\boldsymbol{H}_{S\{RD\}}) = \operatorname{row}(\boldsymbol{H}_{SR}) + \operatorname{row}(\boldsymbol{H}_{SD})$, $\dim(\operatorname{row}(\boldsymbol{H}_{SR}) \cap \operatorname{row}(\boldsymbol{H}_{SD})) = 0$. However, the first of these two conditions is always satisfied while the second one is equivalent to $\operatorname{rank}(\boldsymbol{H}_{SR}) + \operatorname{rank}(\boldsymbol{H}_{SD}) = \operatorname{rank}(\boldsymbol{H}_{S\{RD\}})$ [89]. □

The proposition below shows that whenever the channel gain matrices are drawn from a continuous distribution, the Gaussian MIMO relay channel almost surely has disjoint sender components if and only if the source has at least as many antennas as the relay and the destination combined. In addition, note that since Gaussian MIMO relay channels with orthogonal sender components also belong to the class of Gaussian MIMO relay channels with disjoint sender components, it immediately follows that this condition is also almost surely necessary for a Gaussian MIMO relay channel to have orthogonal sender components. Most importantly, however, this proposition, together with Proposition 4.3, implies that Gaussian MIMO relay channels with disjoint sender components are not of stochastically degraded nature with probability one.

Proposition 4.2. *If the channel gain matrices are drawn from a continuous distribution, the probability that the Gaussian MIMO relay channel has disjoint sender components is one if* $N_\mathrm{S} \geq N_\mathrm{R} + N_\mathrm{D}$ *and zero otherwise.*

Proof. Suppose the channel gain matrices are drawn from a continuous distribution. Then, $\boldsymbol{H}_\mathrm{SR}$, $\boldsymbol{H}_\mathrm{SD}$, and $\boldsymbol{H}_\mathrm{S\{RD\}}$ almost surely have full rank, i.e.,

$$\operatorname{rank}(\boldsymbol{H}_\mathrm{SR}) = \min\{N_\mathrm{S}, N_\mathrm{R}\}, \tag{4.2a}$$

$$\operatorname{rank}(\boldsymbol{H}_\mathrm{SD}) = \min\{N_\mathrm{S}, N_\mathrm{D}\}, \tag{4.2b}$$

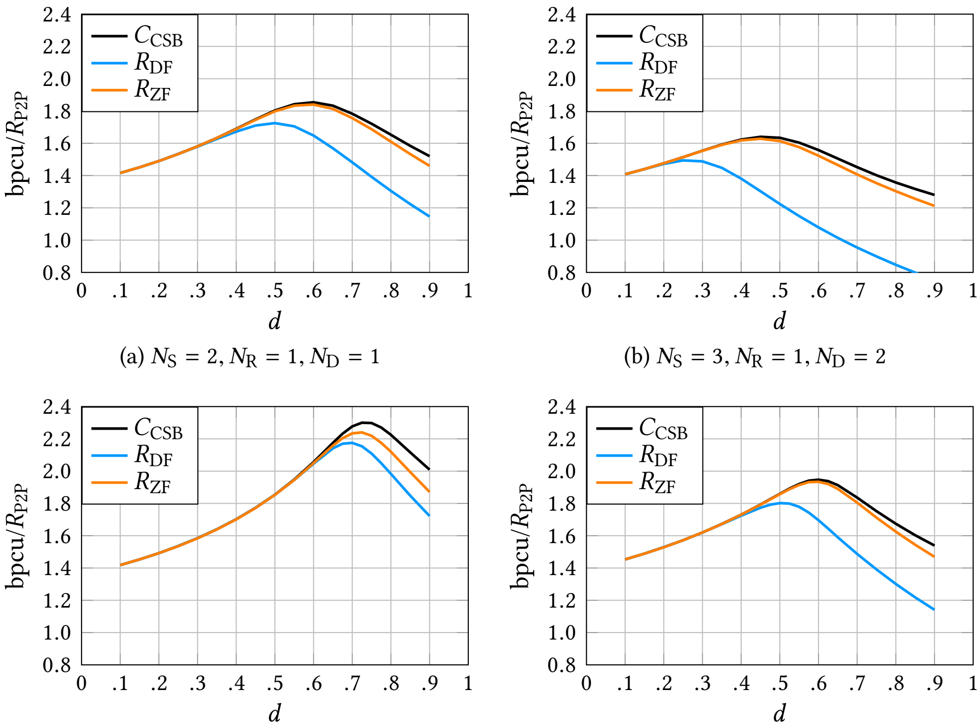
$$\operatorname{rank}(\boldsymbol{H}_\mathrm{S\{RD\}}) = \min\{N_\mathrm{S}, N_\mathrm{R} + N_\mathrm{D}\} \tag{4.2c}$$

with probability one. In this case, the condition $\operatorname{rank}(\boldsymbol{H}_\mathrm{SR}) + \operatorname{rank}(\boldsymbol{H}_\mathrm{SD}) = \operatorname{rank}(\boldsymbol{H}_\mathrm{S\{RD\}})$, which according to Proposition 4.1 characterizes Gaussian MIMO relay channels with disjoint sender components, is satisfied if and only if $\min\{N_\mathrm{S}, N_\mathrm{R} + N_\mathrm{D}\} = N_\mathrm{R} + N_\mathrm{D}$, or equivalently, if $N_\mathrm{S} \geq N_\mathrm{R} + N_\mathrm{D}$. □

Note that four of the eight antenna configurations considered in Figure 4.2 satisfy this condition (with equality). For the corresponding Gaussian MIMO relay channels, which thus almost surely have disjoint sender components and are not of stochastically degraded nature, we cannot evaluate the maximum achievable PDF rates. In Figure 4.3, we hence compare R_ZF, which is achieved by means of the ZF PDF approach with $\boldsymbol{G} = \mathbf{I}_{N_\mathrm{R}}$, to both R_DF and C_CSB for these four antenna configurations, where the results are normalized with respect to R_P2P again.[2]

It can be observed from Figure 4.3 that a suboptimal PDF scheme outperforms the optimal DF scheme whenever R_DF does not already approach C_CSB. Moreover, for three of the four antenna configurations, the basic ZF PDF scheme almost achieves the maximum potential rate gain and the gaps between R_PDF and C_CSB are very small for all considered relay positions, cf. Figures 4.3(a), 4.3(b), and 4.3(d). We can hence conclude that PDF is almost capacity achieving in these three cases. For the fourth antenna configuration, the suboptimal PDF scheme also improves on the DF strategy,

[2]The reason we only consider this basic ZF PDF approach here is that for any Gaussian MIMO relay channel with disjoint sender components, choosing the ZF filter $\boldsymbol{G} = \mathbf{I}_{N_\mathrm{R}}$ does not restrict the information transfer from source to relay while still allowing $\boldsymbol{C}_\mathrm{V}$ to have rank up to $N_\mathrm{D} = \operatorname{rank}(\boldsymbol{H}_\mathrm{SD})$. Nevertheless, we remark that this ZF PDF scheme does not attain R_PDF for all Gaussian MIMO relay channels with disjoint sender components.



(a) $N_S = 2, N_R = 1, N_D = 1$

(b) $N_S = 3, N_R = 1, N_D = 2$

(c) $N_S = 3, N_R = 2, N_D = 1$

(d) $N_S = 4, N_R = 2, N_D = 2$

Figure 4.3: Comparison of C_{CSB}, R_{DF}, and R_{ZF} for Gaussian MIMO Relay Channels with Disjoint Sender Components: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (results averaged over 1000 independent channel realizations)

but there are still considerable gaps between R_{PDF} and C_{CSB} for all relay positions where R_{DF} does not approach C_{CSB}, cf. Figure 4.3(c). A possible explanation for this result is that $(N_D + N_R)/N_D = 3 \gg 1.67 = (N_S + N_R)/N_S$ if $N_S = 3$, $N_R = 2$, and $N_D = 1$, which suggests that cooperative reception of the relay and the destination may be more beneficial than cooperative transmission of the source and the relay in this case. Consequently, a relay strategy like *compress-and-forward* (CF), which achieves gains related to cooperative reception, may perform better than the DF and PDF strategies, which achieve gains related to cooperative transmission, cf. [77].

Coming back to our main question, the results in Figure 4.3 clearly show that even a suboptimal PDF scheme can achieve much higher rates than the optimal DF scheme if $R_{DF} \ll C_{CSB}$ and $N_S \geq N_R + N_D$ (or $N_S = N_R + N_D$, to be more precise). Furthermore, it has been illustrated that this suboptimal PDF scheme almost achieves the maximum potential rate gain and generally achieves rates close to the CSB for at least some of these antenna configurations. An important question that remains open, however, is whether this is only because the corresponding Gaussian MIMO relay channels have disjoint sender components, or if the same holds for antenna configurations where $N_R < N_S < N_R + N_D$.

4.2.2.2 Gaussian MIMO Relay Channels without Disjoint Sender Components

If $N_\mathrm{S} < N_\mathrm{R} + N_\mathrm{D}$, it follows from Proposition 4.2 that the corresponding Gaussian MIMO relay channel almost surely does not have disjoint sender components. In this case, we can evaluate R_PDF whenever the Gaussian MIMO relay channel is of stochastically degraded nature.

From Corollary 3.14, we know that every Gaussian relay channel is either stochastically degraded or reversely stochastically degraded if the source is equipped with a single antenna. Unfortunately, there are no necessary and sufficient conditions in terms of N_S, N_R, and N_D that almost surely characterize stochastically degraded or reversely stochastically degraded Gaussian MIMO relay channels if $N_\mathrm{S} > 1$. We only have the following necessary conditions the antenna configuration must satisfy for the corresponding Gaussian MIMO relay channel to be stochastically degraded or reversely stochastically degraded:

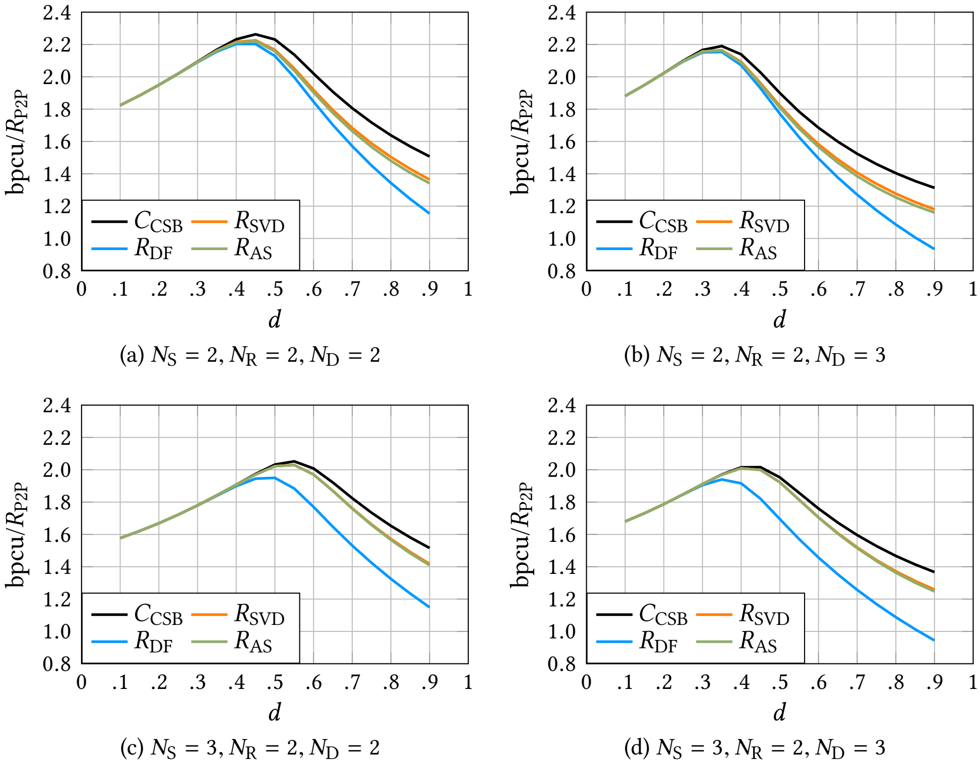
Proposition 4.3. *If the channel gain matrices are drawn from a continuous distribution, the Gaussian MIMO relay channel almost surely is not stochastically degraded or reversely stochastically degraded if $N_\mathrm{S} > N_\mathrm{R}$ or $N_\mathrm{S} > N_\mathrm{D}$, respectively.*

Proof. Suppose that $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{H}_\mathrm{SR} \succcurlyeq \boldsymbol{H}_\mathrm{SD}^\mathrm{H}\boldsymbol{H}_\mathrm{SD}$, which according to Definition 3.3 means that the Gaussian MIMO relay channel is stochastically degraded. Then, by Lemma 3.11, $\exists \boldsymbol{M} \in \mathbb{C}^{N_\mathrm{D}\times N_\mathrm{R}}$ such that $\boldsymbol{H}_\mathrm{SD}^\mathrm{H} = \boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{M}^\mathrm{H}$. That is, every column of $\boldsymbol{H}_\mathrm{SD}^\mathrm{H}$ lies in the column space of $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}$, which in turn means that $\mathrm{rank}([\boldsymbol{H}_\mathrm{SR}^\mathrm{H}, \boldsymbol{H}_\mathrm{SD}^\mathrm{H}]) = \mathrm{rank}(\boldsymbol{H}_\mathrm{S\{RD\}}^\mathrm{H}) = \mathrm{rank}(\boldsymbol{H}_\mathrm{SR}^\mathrm{H})$. But $\mathrm{rank}(\boldsymbol{H}_\mathrm{S\{RD\}}^\mathrm{H}) = \min\{N_\mathrm{S}, N_\mathrm{R} + N_\mathrm{D}\}$ and $\mathrm{rank}(\boldsymbol{H}_\mathrm{SR}^\mathrm{H}) = \min\{N_\mathrm{S}, N_\mathrm{R}\}$ with probability one if the channel gain matrices are drawn from a continuous distribution, which implies $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{H}_\mathrm{SR} \not\succcurlyeq \boldsymbol{H}_\mathrm{SD}^\mathrm{H}\boldsymbol{H}_\mathrm{SD}$ almost surely if $N_\mathrm{S} > N_\mathrm{R}$.

The proof for reversely stochastic degradedness proceeds along the same lines after switching the roles of the relay and the destination. □

Consequently, it is not always possible to conclude from the antenna configuration whether or not a Gaussian MIMO relay channel is of stochastically degraded nature. If $N_\mathrm{S} > \max\{N_\mathrm{R}, N_\mathrm{D}\}$, Proposition 4.3 implies that the Gaussian MIMO relay channel almost surely is neither stochastically degraded nor reversely stochastically degraded. But if $N_\mathrm{S} \le N_\mathrm{R}$ and/or $N_\mathrm{S} \le N_\mathrm{D}$, we need to check if the source-to-relay and source-to-destination channel gain matrices satisfy $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{H}_\mathrm{SR} \succcurlyeq \boldsymbol{H}_\mathrm{SD}^\mathrm{H}\boldsymbol{H}_\mathrm{SD}$ or $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{H}_\mathrm{SR} \preccurlyeq \boldsymbol{H}_\mathrm{SD}^\mathrm{H}\boldsymbol{H}_\mathrm{SD}$. In these two cases, it follows from Theorems 3.12 and 3.13, respectively, that $R_\mathrm{PDF} = R_\mathrm{DF}$ or $R_\mathrm{PDF} = R_\mathrm{P2P}$. If neither of the two conditions is satisfied, on the other hand, the Gaussian MIMO relay channel is not of stochastically degraded nature and we employ the zero-forcing (ZF) approach and the inner approximation algorithm (IAA) to obtain achievable PDF rates.

ZF—For the four antenna configurations where $N_\mathrm{S} < N_\mathrm{R} + N_\mathrm{D}$, we first compare the PDF rates that can be achieved using the ZF PDF approach to the CSB and the maximum



(a) $N_S = 2, N_R = 2, N_D = 2$

(b) $N_S = 2, N_R = 2, N_D = 3$

(c) $N_S = 3, N_R = 2, N_D = 2$

(d) $N_S = 3, N_R = 2, N_D = 3$

Figure 4.4: Comparison of C_{CSB}, R_{DF}, R_{AS}, and R_{SVD} for Gaussian MIMO Relay Channels without Disjoint Sender Components: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (results averaged over 1000 independent channel realizations)

achievable DF rate in Figure 4.4. Like in the previous figures, the results are normalized with respect to R_{P2P}. Note that $R_{AS} = R_{SVD} = R_{DF}$ or $R_{AS} = R_{SVD} = R_{P2P}$ for channel realizations for which the Gaussian MIMO relay channel is stochastically degraded or reversely stochastically degraded, respectively, since $R_{PDF} \geq R_{AS} \geq \max\{R_{DF}, R_{P2P}\}$ and $R_{PDF} \geq R_{SVD} \geq \max\{R_{DF}, R_{P2P}\}$ in general. That is, if the Gaussian MIMO relay channel is of stochastically degraded nature, both the ZF approach based on antenna selection and that based on the singular value decomposition (SVD) of $\boldsymbol{H}_{SR}$ achieve R_{PDF}. Furthermore, we can directly conclude that the winning selections

$$\mathcal{A}^{\text{win}} = \underset{\emptyset \subseteq \mathcal{A} \subseteq \mathcal{I}_{N_R}}{\arg\max}\, R_{AS}^{\mathcal{A}}, \qquad \mathcal{B}^{\text{win}} = \underset{\emptyset \subseteq \mathcal{B} \subseteq \mathcal{I}_{\text{rank}(\boldsymbol{H}_{SR})}}{\arg\max}\, R_{SVD}^{\mathcal{B}} \tag{4.3}$$

are given by $\mathcal{A}^{\text{win}} = \mathcal{I}_{N_R}$ and $\mathcal{B}^{\text{win}} = \mathcal{I}_{\text{rank}(\boldsymbol{H}_{SR})}$ if the relay channel is stochastically degraded and by $\mathcal{A}^{\text{win}} = \mathcal{B}^{\text{win}} = \emptyset$ if it is reversely stochastically degraded. Otherwise, R_{AS} and R_{SVD} are obtained by considering all possible selections, cf. (3.79) and (3.85). For two of the four antenna configurations considered in Figure 4.4 and some selected values of d, the percentages of the winning selections $\mathcal{A}^{\text{win}}$ and $\mathcal{B}^{\text{win}}$ are presented in Tables 4.1 and 4.2.

Table 4.1: Percentages of Winning Selections for Zero-Forcing PDF Schemes: $N_S = 2$, $N_R = 2$, $N_D = 2$ (results for 1000 independent channel realizations)

(a) $d = 0.5$

sel.	$\mathcal{A}^{\text{win}}$	$\mathcal{B}^{\text{win}}$
$\emptyset$	0.1	0.1
$\{1\}$	7.8	21.2
$\{2\}$	9.3	0
$\{1,2\}$	82.8	78.7

(b) $d = 0.7$

sel.	$\mathcal{A}^{\text{win}}$	$\mathcal{B}^{\text{win}}$
$\emptyset$	0.6	0.4
$\{1\}$	21.3	49.7
$\{2\}$	21.7	0.1
$\{1,2\}$	56.4	49.8

(c) $d = 0.9$

sel.	$\mathcal{A}^{\text{win}}$	$\mathcal{B}^{\text{win}}$
$\emptyset$	4.9	4.4
$\{1\}$	33.8	72.8
$\{2\}$	36.3	0.5
$\{1,2\}$	25.0	22.3

Table 4.2: Percentages of Winning Selections for Zero-Forcing PDF Schemes: $N_S = 3$, $N_R = 2$, $N_D = 2$ (results for 1000 independent channel realizations)

(a) $d = 0.5$

sel.	$\mathcal{A}^{\text{win}}$	$\mathcal{B}^{\text{win}}$
$\emptyset$	0	0
$\{1\}$	0.4	1.1
$\{2\}$	0.2	0
$\{1,2\}$	99.4	98.9

(b) $d = 0.7$

sel.	$\mathcal{A}^{\text{win}}$	$\mathcal{B}^{\text{win}}$
$\emptyset$	0	0
$\{1\}$	3.3	9.9
$\{2\}$	3.7	0
$\{1,2\}$	93.0	90.1

(c) $d = 0.9$

sel.	$\mathcal{A}^{\text{win}}$	$\mathcal{B}^{\text{win}}$
$\emptyset$	0.5	0.2
$\{1\}$	13.0	33.3
$\{2\}$	13.7	0.2
$\{1,2\}$	72.8	66.3

The most interesting observation from Figure 4.4 is that for the two cases where $N_S > N_R$, cf. Figures 4.4(c) and 4.4(d), the results look similar to those for the Gaussian MIMO relay channels with disjoint sender components in Figures 4.3(a), 4.3(b), and 4.3(d). In particular, the suboptimal ZF PDF scheme can achieve much higher rates than the DF strategy if $R_{DF} \ll C_{CSB}$, and R_{AS} and R_{SVD} closely approach the maximum potential rate gain. The only minor difference is that for larger values of d, the gaps between R_{AS} or R_{SVD} and C_{CSB} in Figures 4.4(c) and 4.4(d) are not quite as small as those between R_{ZF} and C_{CSB} in Figures 4.3(a), 4.3(b), and 4.3(d). We can therefore conclude that the PDF strategy is able to outperform the DF strategy whenever the source is equipped with more antennas than the relay.

Note that this conclusion is also supported by the results in Table 4.2. For $N_S = 3$, $N_R = N_D = 2$, and $d = 0.5$, we see from Figure 4.4(c) that $R_{DF} \ll R_{AS} = R_{SVD} \approx C_{CSB}$ while Table 4.2(a) shows that the maximum index set $\{1,2\}$ is the optimal ZF selection for about 99% of the generated channel realizations. This means that the differences between R_{DF} and R_{AS} or R_{SVD} are almost entirely due to the fact that because $N_S > N_R$, the condition $\text{range}(\boldsymbol{C}_V) \subseteq \text{null}(\boldsymbol{H}_{SR})$ does not imply $\boldsymbol{C}_V = \boldsymbol{0}$. Even for larger values of d, this is the main reason why PDF is able to outperform DF if $N_S > N_R$ as the maximum index set remains the optimal ZF selection for the vast majority of the generated channel realizations, cf. Tables 4.2(b) and 4.2(c).

Table 4.3: Percentages of Channel Realizations for which the Gaussian MIMO Relay Channel is Stochastically Degraded (P_{sto}) or Reversely Stochastically Degraded (P_{rev}): $N_{\mathrm{S}} = 2$, $N_{\mathrm{R}} = 2$, $N_{\mathrm{D}} = 2$ (results for 1000 independent channel realizations)

d	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
P_{sto}	100	99.5	96.8	89.4	75.8	61.6	40.9	24.7	13.0
P_{rev}	0	0	0	0	0.1	0.2	0.3	0.5	2.3

For the two antenna configuration where $N_{\mathrm{S}} \leq N_{\mathrm{R}}$, on the other hand, Figures 4.4(a) and 4.4(b) show that the suboptimal ZF PDF approach cannot significantly outperform the DF strategy. In particular, the maximum values of R_{AS} and R_{SVD} are only marginally greater than the maximum values of R_{DF}, which can partly be explained by the fact that the latter already come close to the maximum potential rate gains. What is more, the differences between R_{DF} and R_{AS} or R_{SVD} are smaller than those between C_{CSB} and R_{AS} or R_{SVD}, respectively, as long as the relay is not very close to the destination.

The main reason for this result is revealed in Table 4.3, which for $N_{\mathrm{S}} = N_{\mathrm{R}} = N_{\mathrm{D}} = 2$ shows how likely the Gaussian MIMO relay channel turned out to be stochastically degraded or reversely stochastically degraded. For $d \leq 0.3$, it can be observed from Table 4.3 that the Gaussian MIMO relay channel was stochastically degraded for at least 96% of the generated channel realizations, and even for $d = 0.5$, more than 75% of the channel realizations satisfied the stochastically degradedness condition. In this case, however, we already know that PDF cannot outperform DF as both strategies achieve exactly the same rates, i.e., $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$.

Beyond that, Table 4.1 shows that the percentages of the channel realizations for which $\mathcal{A}^{\mathrm{win}} = \{1, 2\}$ or $\mathcal{B}^{\mathrm{win}} = \{1, 2\}$ are even higher. In contrast to the cases where $N_{\mathrm{S}} > N_{\mathrm{R}}$, the ZF condition $\mathrm{range}(\boldsymbol{C}_{\mathrm{V}}) \subseteq \mathrm{null}(\boldsymbol{H}_{\mathrm{SR}})$ almost surely implies $\boldsymbol{C}_{\mathrm{V}} = \boldsymbol{0}$ for the considered example scenario if $N_{\mathrm{S}} \leq N_{\mathrm{R}}$. As a result, $R_{\mathrm{AS}} = R_{\mathrm{DF}}$ and $R_{\mathrm{SVD}} = R_{\mathrm{DF}}$ if $\mathcal{A}^{\mathrm{win}} = \{1, 2\}$ and $\mathcal{B}^{\mathrm{win}} = \{1, 2\}$, respectively. When the relay is placed exactly in the middle between the source and the destination, this applies for about 80% of the channel realizations, cf. Table 4.1(a), and for $d = 0.7$, this percentage is still about 50%, cf. Table 4.1(b). On the other hand, the empty selection, which results in $R_{\mathrm{AS}} = R_{\mathrm{SVD}} = R_{\mathrm{P2P}}$, really plays a role only if the relay is very close to the destination, cf. Table 4.1(c). This is reasonable as Table 4.3 shows that $\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{SR}} \preccurlyeq \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{SD}}$ for more than 0.5% of the generated channel realizations only if $d > 0.8$.

Finally, we see from Figure 4.4 that $R_{\mathrm{SVD}} > R_{\mathrm{AS}}$ on average, as one might expect since the ZF filter design based on antenna selection completely neglects the channel properties. However, the difference is not remarkable for any antenna configuration, and for some channel realizations, it actually turned out that $R_{\mathrm{AS}} > R_{\mathrm{SVD}}$. On the one hand, this may indicate that we should not expect to obtain significantly higher PDF rates with more sophisticated ZF filter designs, but on the other, it demonstrates that the SVD-based ZF filter design is certainly not optimal. The latter is also supported by the fact that $\mathcal{B}^{\text{win}} = \{2\}$ for some channel realizations, cf. Table 4.1(c), for example. This means that the singular vector that belongs to the smaller singular value of $\boldsymbol{H}_{\text{SR}}$ can be a better ZF filter for our suboptimal PDF approach than that belonging to the stronger one although it is clearly worse if only the information transfer from the source to the relay is considered.

More sophisticated ZF filter designs should of course be based on both the source-to-relay and the source-to-destination channels. To this end, one could for example use the generalized singular value decomposition (GSVD) of $\boldsymbol{H}_{\text{SR}}$ and $\boldsymbol{H}_{\text{SD}}$ to form the ZF filter $\boldsymbol{G}$ as follows. Suppose the channel gain matrices $\boldsymbol{H}_{\text{SR}}$ and $\boldsymbol{H}_{\text{SD}}$ are given as in the first theorem stated in [99], and assume that $\alpha_i/\beta_i \geq 1$ for $i \in \{1, \ldots, \ell'\}$ and $\alpha_i/\beta_i < 1$ for $i \in \{\ell' + 1, \ldots, \ell\}$. Then, let

$$\boldsymbol{G}_{\text{GSVD}} = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k+\ell'}]^{\text{H}} \boldsymbol{U}_{\text{SR}}^{\prime\,\text{H}} \in \mathbb{C}^{(k+\ell')\times N_{\text{R}}}, \tag{4.4}$$

where $\boldsymbol{e}_i \in \{0, 1\}^{N_{\text{R}}}$ denotes the i-th canonical unit vector of dimension N_{R}. For this choice of $\boldsymbol{G}$, it follows that the source signal part $\boldsymbol{v}$, which the relay is not supposed to decode, must not be transmitted along those directions of the composite channel $\boldsymbol{H}_{\text{S\{RD\}}}$ for which the relay receives a stronger signal than the destination. However, we remark that it is not clear whether $\boldsymbol{G}_{\text{GSVD}}$ is the optimal ZF filter for the (generally suboptimal) ZF PDF approach. In addition, the results for the PDF rates that can be obtained by means of the IAA, which are presented below, suggest that any improvement over R_{SVD} due to more sophisticated ZF filter designs is likely to be marginal.

IAA—For the four antenna configurations considered in Figure 4.4, we now want to compare C_{CSB}, R_{DF}, and R_{SVD} to the achievable PDF rates that can be obtained by means of the IAA.[3] To this end, first recall from Section 3.4.3 that the principle of the IAA is to locally approximate the nonconcave inequality constraint functions, which means that the results may strongly depend on $\boldsymbol{C}_{\text{V}}^{(0)}$. Therefore, we choose the following two different initializations:

1. $\boldsymbol{C}_{\text{V}}^{(0)} = \boldsymbol{0} \ \rightarrow\ R_{\text{IAA1}}$
   Because setting $\boldsymbol{C}_{\text{V}} = \boldsymbol{0}$ in the PDF rate maximization problem yields R_{DF} and the sequence of achievable PDF rates that is generated by the IAA is nondecreasing, this choice of $\boldsymbol{C}_{\text{V}}^{(0)}$ guarantees that $R_{\text{IAA1}} \geq R_{\text{DF}}$. Moreover, setting $\boldsymbol{C}_{\text{V}}^{(0)} = \boldsymbol{0}$ means that the IAA starts at a point where there is no interference signal at the relay, and if the achievable PDF rate can be increased by moving away from this point, the initial covariance matrix $\boldsymbol{C}_{\text{V}}$ has no predefined structure.
2. $\boldsymbol{C}_{\text{V}}^{(0)} = \boldsymbol{C}_{\text{V}}^{\mathcal{B}^{\text{win}}} \rightarrow R_{\text{IAA2}}$
   With this choice of $\boldsymbol{C}_{\text{V}}^{(0)}$, we ensure that $R_{\text{IAA2}} \geq R_{\text{SVD}} \geq \max\{R_{\text{DF}}, R_{\text{P2P}}\}$. Furthermore, using this initialization allows us to assess the results obtained by means of

[3] R_{AS} is omitted here since the SVD-based ZF approach yields better results on average.

the ZF PDF approach. In particular, if the differences between R_{IAA2} and R_{SVD} are small, we can reasonably assume that the SVD-based ZF approach yields good suboptimal solutions to the PDF rate maximization problem.

Note that for stochastically degraded Gaussian MIMO relay channels, it can immediately be concluded that $R_{\mathrm{IAA1}} = R_{\mathrm{IAA2}} = R_{\mathrm{DF}}$, whereas $R_{\mathrm{IAA2}} = R_{\mathrm{P2P}}$ if the Gaussian MIMO relay channel is reversely stochastically degraded. Consequently, it is not necessary to run the IAA in order to determine R_{IAA1} or R_{IAA2} in these cases.

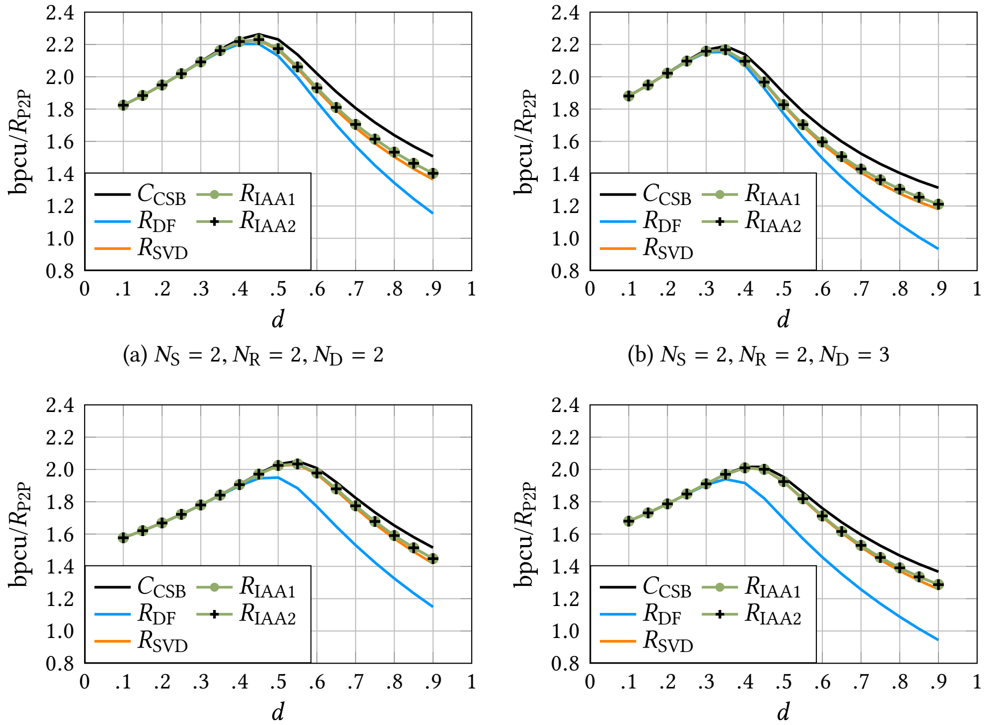
In addition to the initialization, the achievable PDF rates we obtain from the IAA of course depend on the termination criterion. For the results presented in Figure 4.5, where we compare R_{IAA1} and R_{IAA2} to C_{CSB}, R_{DF}, and R_{SVD}, the termination criterion for the IAA was based on the relative rate improvement over the last three iterations. More precisely, the IAA was stopped after iteration K if $K > 3$ and

$$\frac{R_{\mathrm{IAA}}^{(K)} - R_{\mathrm{IAA}}^{(K-3)}}{R_{\mathrm{IAA}}^{(K)}} < 10^{-2}. \tag{4.5}$$

The first notable observation from Figure 4.5 is that $R_{\mathrm{IAA1}} = R_{\mathrm{IAA2}}$ for all considered antenna configurations and relay positions, i.e., there is no difference between the achievable PDF rates we obtain with the two different initializations of the IAA. This result suggests that for both of the initializations we have chosen, the IAA always finds good local solutions to the nonconvex PDF rate maximization problem as it would be unlikely that $R_{\mathrm{IAA1}} = R_{\mathrm{IAA2}}$ if this were not the case. In particular, since the structure of $\boldsymbol{C}_{\mathrm{V}}$ is not predefined if $\boldsymbol{C}_{\mathrm{V}}^{(0)} = \boldsymbol{0}$, one would expect that $R_{\mathrm{IAA1}} > R_{\mathrm{IAA2}}$ for at least some channel realizations (and thus also on average) if the IAA converged to bad local solutions with $\boldsymbol{C}_{\mathrm{V}}^{(0)} = \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}^{\mathrm{win}}}$. Likewise, one would reasonably expect that $R_{\mathrm{IAA1}} < R_{\mathrm{IAA2}}$ if the IAA did not always converge to good local solutions with $\boldsymbol{C}_{\mathrm{V}}^{(0)} = \boldsymbol{0}$.

The second notable observation from Figure 4.5 is that the differences between R_{SVD} and R_{IAA1}, R_{IAA2} are very small for all four antenna configurations and the considered relay positions. As a consequence, we can conclude that the SVD-based ZF approach already yields good suboptimal solutions to the PDF rate maximization problem, which in turn implies that we are not likely to achieve much higher PDF rates with more sophisticated ZF filter designs.

Finally, we also want to compare the computational effort that needs to be spent in order to evaluate R_{IAA1}, R_{IAA2}, and R_{SVD}. To this end, first note that if the Gaussian MIMO relay channel is not of stochastically degraded nature, the evaluation of R_{SVD} for one particular channel realization requires to solve $2^{\min\{N_{\mathrm{S}},N_{\mathrm{R}}\}} = 4$ convex optimization problems for each of the antenna configurations considered in Figure 4.5. Furthermore, we remark that this is also the minimum number of iterations for the IAA with the termination criterion specified in (4.5). Because the Gaussian MIMO relay channel is almost surely neither stochastically degraded nor reversely stochastically degraded if $N_{\mathrm{S}} > \max\{N_{\mathrm{R}}, N_{\mathrm{D}}\}$, we here consider the computational effort for evaluating R_{IAA1} and R_{IAA2} using the antenna configuration $N_{\mathrm{S}} = 3$, $N_{\mathrm{R}} = 2$, and $N_{\mathrm{D}} = 2$ as an example. For



(a) $N_S = 2, N_R = 2, N_D = 2$ (b) $N_S = 2, N_R = 2, N_D = 3$

(c) $N_S = 3, N_R = 2, N_D = 2$ (d) $N_S = 3, N_R = 2, N_D = 3$

Figure 4.5: Comparison of C_{CSB}, R_{DF}, R_{SVD}, and R_{IAA} for Gaussian MIMO Relay Channels without Disjoint Sender Components: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (results averaged over 1000 independent channel realizations)

this case, Table 4.4 shows the average, median, and maximum numbers of the iterations the IAA required to compute R_{IAA1} and R_{IAA2}, where the average iteration numbers have been rounded to the nearest hundredth.

Recall that in each iteration of the IAA, one approximating convex optimization problem must be solved, which dominates the computational cost of the iteration. Moreover, the computational complexity of such an approximating problem is comparable to that of one of the four convex optimization problems which need to be solved in order to evaluate R_{SVD}. We therefore simply measure the computational effort for the IAA in terms of the iteration numbers.

From Table 4.4, we can observe that both the average and the median numbers of the iterations the IAA required to compute R_{IAA1} and R_{IAA2} are small. In particular, except for the case where $d = 0.9$, the median numbers are equal to the minimum iteration number for the chosen termination criterion, and even for $d = 0.9$, the median numbers only increase by one. The average iteration numbers are also equal or at least very close to the minimum iteration number as long as d does not become too large. Whereas the median numbers are identical for both initializations, the average iteration numbers for computing R_{IAA1} are slightly higher than those for R_{IAA2} if the relay is closer to

Table 4.4: Number of Iterations for Computation of R_{IAA}: $N_{\mathrm{S}} = 3$, $N_{\mathrm{R}} = 2$, $N_{\mathrm{D}} = 2$ (results for 1000 independent channel realizations)

(a) R_{IAA1} ($\boldsymbol{C}_{\mathrm{V}}^{(0)} = \mathbf{0}$)

d	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
average	4.00	4.00	4.00	4.00	4.08	4.39	4.83	5.67	6.86
median	4	4	4	4	4	4	4	4	5
maximum	5	4	4	5	18	20	20	21	22

(b) R_{IAA2} ($\boldsymbol{C}_{\mathrm{V}}^{(0)} = \boldsymbol{C}_{\mathrm{V}}^{\mathcal{B}^{\mathrm{win}}}$)

d	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
average	4.00	4.00	4.00	4.00	4.02	4.19	4.51	5.09	5.88
median	4	4	4	4	4	4	4	4	5
maximum	5	4	4	5	8	18	34	29	32

the destination. However, the difference is less than one iteration on average for all considered relay positions.

Consequently, we can conclude that for most of the channel realizations considered here, the computational effort for the two suboptimal PDF approaches (ZF and IAA) is comparable although the maximum iteration numbers in Table 4.4 reveal that the computational cost of computing R_{IAA1} or R_{IAA2} can sometimes be significantly higher than that of R_{SVD}. Here, it should of course be noted that we first need to evaluate R_{SVD} before R_{IAA2} can be computed and that the complexity of the SVD-based ZF approach increases exponentially with $\min\{N_{\mathrm{S}}, N_{\mathrm{R}}\}$. For Gaussian MIMO relay channels that are not of stochastically degraded nature, the IAA with initialization $\boldsymbol{C}_{\mathrm{V}}^{(0)} = \mathbf{0}$ hence does not only obtain the highest PDF rates, but it usually also has the best performance-complexity tradeoff of the considered suboptimal PDF approaches.

4.3 Practical Applications

To conclude the first part of this work, we want to discuss the practical usability of the theoretical results derived in Chapter 3 and the corresponding numerical results presented in this chapter. Clearly, as these results are based on the idealistic assumptions that perfect CSI is available at all nodes and that the relay operates in full-duplex mode with perfect self-interference cancellation, they cannot directly be applied to practical communication systems. However, there are several aspects which may help to design future relay-aided wireless communication systems.

Using the results of Chapter 3, it is possible to efficiently determine upper and lower bounds on the capacity of the Gaussian MIMO relay channel. In particular, we showed that evaluating the CSB or the maximum achievable DF rate only requires to solve one convex optimization problem, and for certain types of Gaussian MIMO relay channels, the same is true for the maximum achievable PDF rate. Otherwise, convex optimization techniques can still be employed to obtain good suboptimal PDF rates with reasonable complexity. The evaluated upper and lower capacity bounds can serve as benchmarks when studying the performance of other relay strategies, the influence of the MIMO channel model on the achievable rates, or the performance degradation due to channel estimation errors, imperfect self-interference cancellation at the relay, or imperfect CSI at the transmitting nodes.

For the considered scenario with uncorrelated Rayleigh fading, it could for example be concluded that using the PDF strategy instead of the simpler DF strategy is really beneficial only if the source is equipped with more antennas than the relay. Moreover, for such antenna configurations, the results showed that a rather simple ZF scheme can achieve PDF rates that closely approach the CSB. While our focus was on how the achievable and potential rate gains are affected by the antenna configurations and relay positions, one can just as well examine the impact of different transmit power budgets and/or path loss coefficients on the rate gains, or how the results change for different MIMO channel models.

Numerical results we obtain under the assumptions of perfect CSI and perfect self-interference cancellation at the relay may also help to answer further interesting questions about (MIMO) relaying. For example, an important property of the DF and PDF strategies is that the source and the relay can cooperatively transmit to the destination. However, this only holds if both the absolute values and the phases of all channel gains are known at the source and the relay. In particular, for *phase fading*, i.e., if only the absolute values of the channel gains are known at the transmitting nodes and the phases are assumed to vary uniformly over $[0, 2\pi]$, the maximum achievable ergodic DF rate is attained by $\boldsymbol{C}_\mathrm{S} = \frac{P_\mathrm{S}}{N_\mathrm{S}}\mathbf{I}_{N_\mathrm{S}}$, $\boldsymbol{C}_\mathrm{R} = \frac{P_\mathrm{R}}{N_\mathrm{R}}\mathbf{I}_{N_\mathrm{R}}$, and $\boldsymbol{C}_\mathrm{SR} = \mathbf{0}$, cf. [77]. Furthermore, even if the channel phases are known, one may require the source and relay inputs to be independent for practical reasons. By evaluating the achievable DF and PDF rates for these channel inputs and comparing them to R_DF and $R_\mathrm{PDF}/R_\mathrm{ZF}/R_\mathrm{SVD}/R_\mathrm{IAA1}$, respectively, it can be assessed how critical the knowledge of phase information at the transmitting nodes and the cooperation between the source and the relay are to the performance of the DF and PDF strategies.

Beyond that, we believe that a thorough understanding of the Gaussian MIMO relay channel with perfect CSI can generally yield valuable insights for the analysis of ergodic or outage rates in more realistic scenarios where the transmitting nodes only have imperfect or statistical CSI. Likewise, we can reasonably assume that a thorough understanding of the Gaussian MIMO relay channel with perfect self-interference cancellation at the relay is helpful if one wants to examine what happens if the relay is not able to completely remove its self-interference.

The self-interference problem can be avoided if the relay operates in half-duplex mode, with the additional benefit that half-duplex devices are easier to build and thus cheaper than full-duplex devices. From a theoretical point of view, on the other hand, evaluating the CSB or the maximum achievable DF and PDF rates for the half-duplex case is more difficult than for the full-duplex case (assuming perfect self-interference cancellation at the relay). This is because in addition to the channel inputs, we have to optimize the durations of the time slots or the bandwidths of the frequency bands that are used for reception and transmission at the relay.

As far as the optimization of the channel inputs is concerned, however, the problems we face for the half-duplex case are the same as those for the full-duplex case. Therefore, many results for the relay channel straightforwardly carry over to the half-duplex relay channel, which we discuss in the second part of this work. In particular, this means that we can directly make use of the results derived in Chapter 3 to tackle the corresponding rate maximization problems for the half-duplex Gaussian MIMO relay channel.

Part II

The Relay Channel with Half-Duplex Constraint

Chapter 5

Information Theoretical Results

In the second part of this work, we again consider the relay channel where one source terminal S wants to convey information to one destination terminal D with the help of a single relay node R. In contrast to the first part, however, we now impose a *half-duplex* constraint on the relay, which means that it can either transmit or receive, but not simultaneously transmit and receive, in the same frequency band. Consequently, orthogonal resources have to be assigned for reception and transmission at the relay. For this purpose, we assume that reception and transmission at the relay are separated in time, i.e., we consider a *time-division duplex* (TDD) protocol where the relay and the destination listen to the source in the first phase and where the destination listens to the source and the relay in the second phase.

For this system model, we again want to investigate how much the relay can improve the data transfer from the source to the destination in terms of achievable rates. To this end, we first introduce an information theoretical model for the discrete memoryless half-duplex relay channel in Section 5.1. Subsequently, we review the information theoretical results on the capacity of the half-duplex relay channel that are important to this work. More specifically, we discuss the *cut-set bound* (CSB) as well as the achievable *decode-and-forward* (DF) and *partial decode-and-forward* (PDF) rates for the half-duplex relay channel in Sections 5.2, 5.3, and 5.4, respectively. The system model and the results presented in these four sections can be viewed as the natural extension of the system model and the results presented in Sections 2.1–2.4 to the half-duplex relay channel where TDD is used to separate transmission and reception at the relay.

Furthermore, we also want to study the half-duplex *two-way* relay channel in the second part of this work, which models the more general scenario where two terminals exchange information with the help of a single relay node. In Section 5.5, the half-duplex relay channel model and the corresponding information theoretical results are therefore generalized to bidirectional communication with all three nodes being subject to a half-duplex constraint. The chapter concludes in Section 5.6 with an overview of further noteworthy results and some bibliographical notes on uni- and bidirectional communication in the half-duplex constrained relay channel.

5.1 Channel Model

The discrete memoryless half-duplex relay channel, specified by

$$\left\{\mathcal{X}_\mathrm{S} \times \mathcal{X}_\mathrm{R} \times \mathcal{M}, p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R}, m), \mathcal{Y}_\mathrm{D} \times \mathcal{Y}_\mathrm{R}\right\}, \tag{5.1}$$

consists of five finite sets $\mathcal{X}_\mathrm{S}$, $\mathcal{X}_\mathrm{R}$, $\mathcal{M}$, $\mathcal{Y}_\mathrm{D}$, $\mathcal{Y}_\mathrm{R}$ and a collection of probability mass functions $p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R}, m)$, one for each $(x_\mathrm{S}, x_\mathrm{R}, m) \in \mathcal{X}_\mathrm{S} \times \mathcal{X}_\mathrm{R} \times \mathcal{M}$. Like for the full-duplex case, the interpretation is that x_S and x_R are the channel inputs of the source and the relay, respectively, whereas y_R and y_D denote the channel outputs of the relay and the destination, cf. Figure 5.1. Without loss of generality, it is assumed here that the transmit alphabet of the relay $\mathcal{X}_\mathrm{R}$ contains a "quiet" symbol q and that the receive alphabet of the relay $\mathcal{Y}_\mathrm{R}$ contains an "erasure" symbol e.

In order to account for the half-duplex constraint, the half-duplex relay channel model contains an additional random variable $M \in \mathcal{M} = \{1, 2\}$, which models the state/mode of the relay, i.e., whether the relay receives or transmits.[1] As the relay has to receive information from the source before it can forward this information to the destination, the relay is defined to be in receive mode if $M = 1$ and in transmit mode if $M = 2$. If we further define

$$X_{\mathrm{A}m} = X_\mathrm{A}|\{M = m\}, \qquad Y_{\mathrm{B}m} = Y_\mathrm{B}|\{M = m\}, \tag{5.2}$$

the constraint that the relay cannot simultaneously transmit and receive in the same frequency band can be expressed in terms of the conditional probability mass function that specifies the half-duplex relay channel as

$$p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R}, m) = \begin{cases} p(y_\mathrm{D1}, y_\mathrm{R1}|x_\mathrm{S1})\delta_q(x_\mathrm{R1}) & \text{if } m = 1, \\ p(y_\mathrm{D2}|x_\mathrm{S2}, x_\mathrm{R2})\delta_e(y_\mathrm{R2}) & \text{if } m = 2, \end{cases} \tag{5.3}$$

where

$$\delta_a(b) = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4}$$

In particular, this also means that the channel input distribution $p(x_\mathrm{S}, x_\mathrm{R}, m)$ takes on the following form:

$$p(x_\mathrm{S}, x_\mathrm{R}, m) = \begin{cases} p(x_\mathrm{S1})\delta_q(x_\mathrm{R1}) & \text{if } m = 1, \\ p(x_\mathrm{S2}, x_\mathrm{R2}) & \text{if } m = 2. \end{cases} \tag{5.5}$$

[1]Note that if we consider unidirectional communication, the source and the destination are not affected by a half-duplex constraint. This is because the source does not want to receive any information and thus may transmit the whole time, whereas the destination does not have any information to transmit and hence may listen the whole time.

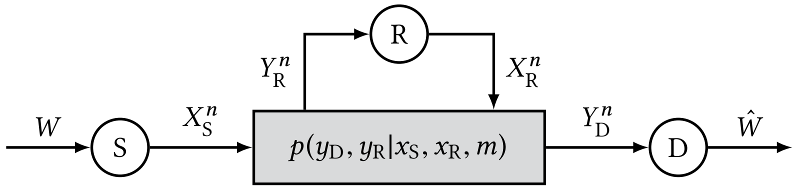


Figure 5.1: Illustration of the Half-Duplex Relay Channel

REMARK 5.1. Our notation differs from that of Khojastepour et al. [66–68], who simply treat M as a state that is known to all nodes. Their notation hence follows Wolfowitz' notation for finite-state channels with state calculable by both sender and receiver [138, Section 3.5]. We treat M as a channel input, however, since the distribution of M can be chosen like the distribution of (X_S, X_R).

REMARK 5.2. Although we treat M as a channel input in our notation, we assume that it is noncausally known to all three nodes. That is, we assume that the source and the destination always know in advance which mode (receive or transmit) the relay is using. Therefore, M cannot be used to convey information from the source to the destination, as would be the case if the destination did not know M, cf. [75].

Let $m^n \in \mathcal{M}^n$ be a sequence of relay modes that is noncausally known to all three nodes. Then, a $(2^{nR}, n)$ code for the half-duplex relay channel consists of a message set $\mathcal{W} = \{1, 2, \ldots, \lceil 2^{nR} \rceil\}$, an encoder (at the source) that assigns a codeword $X_S^n(w) \in \mathcal{X}_S^n$ to each message $w \in \mathcal{W}$, a relay encoder that assigns a symbol $X_{R,i}(y_R^{i-1}) \in \mathcal{X}_R$ to each past received sequence $y_R^{i-1} \in \mathcal{Y}_R^{i-1}$ for each $i \in \{1, \ldots, n\}$ with $X_{R,i}(y_R^{i-1}) = q$ if $m_i = 1$, and a decoder (at the destination) that assigns an estimate $\hat{W}(y_D^n) \in \mathcal{W}$ (or possibly an error message) to each received sequence $y_D^n \in \mathcal{Y}_D^n$. For generality, the encoding and decoding functions are allowed to be stochastic.

Note that the definition of the relay encoder includes two important conditions. The first one is that the relay's transmit symbol $X_{R,i}$ may depend on its past observations $Y_R^{i-1} = (Y_{R,1}, \ldots, Y_{R,i-1})$ only, which ensures that the relay operates in a *causal* manner. Beyond that, the relay encoder must select the quiet symbol q whenever the relay is in receive mode, i.e., $X_{R,i} = q$ if $m_i = 1$, to satisfy the half-duplex constraint.

The half-duplex relay channel is *memoryless* in the sense that the current channel outputs $(Y_{D,i}, Y_{R,i})$ depend on (X_S^i, X_R^i, M^i) only through $(X_{S,i}, X_{R,i}, M_i)$. For any $p(w)$ and choice of the code, the joint probability mass function on $\mathcal{W} \times \mathcal{X}_S^n \times \mathcal{X}_R^n \times \mathcal{M}^n \times \mathcal{Y}_D^n \times \mathcal{Y}_R^n$ hence factors as

$$p(w, x_S^n, x_R^n, m^n, y_D^n, y_R^n) = p(w)p(x_S^n|w) \prod_{i=1}^{n} p(x_{R,i}|y_R^{i-1})p(y_{D,i}, y_{R,i}|x_{S,i}, x_{R,i}, m_i), \tag{5.6}$$

where $p(y_{D,i}, y_{R,i}|x_{S,i}, x_{R,i}, m_i)$ satisfies (5.3) for any $i \in \{1, \ldots, n\}$ and $p(x_{R,i}|y_R^{i-1}) =$

$\delta_q(x_{R,i})$ for any $i \in \{1, \dots, n\}$ such that $m_i = 1$. Furthermore, the transmissions are modeled as taking place *synchronously* again, cf. Section 2.1 for a detailed explanation of what this means.

If W is uniformly distributed over $\mathcal{W}$ and $P_e^{(n)} = \Pr[\hat{W} \neq W]$ denotes the average probability of error, a rate R is said to be *achievable* if there exists a sequence of $(2^{nR}, n)$ codes for which $P_e^{(n)} \to 0$ as $n \to \infty$. The *capacity* C of the half-duplex relay channel is then defined as the supremum of the set of achievable rates.

We remark that if $n_m = \sum_{i=1}^{n} \delta_m(m_i)$, it follows that

$$t_m = \frac{n_m}{n} \to p(m) \in [0, 1] \quad \text{as} \quad n \to \infty. \tag{5.7}$$

That is, as the number of channel uses becomes large, the (normalized) durations of the relay receive and transmit phases t_1 and t_2, which we also refer to as the *time-shares* of the relay receive and transmit phases in the following, can be interpreted as the probabilities that the relay is in receive or in transmit mode, respectively. For our half-duplex relay channel model, where M is assumed to be known to all three nodes, any maximization with respect to the time-shares t_1 and t_2 thus implicitly corresponds to a maximization with respect to $p(m)$.

Like for the full-duplex case, the capacity of the half-duplex relay channel is still unknown for the general case. However, for the four types of relay channels defined in Definition 2.1, the definitions and the corresponding capacity results that are stated in Section 2.4 can easily be extended to the half-duplex case.

Definition 5.1. The discrete memoryless half-duplex relay channel $\left\{\mathcal{X}_S \times \mathcal{X}_R \times \mathcal{M}, p(y_D, y_R|x_S, x_R, m), \mathcal{Y}_D \times \mathcal{Y}_R\right\}$

(a) is *degraded* if $p(y_{D1}, y_{R1}|x_{S1})$ can be written in the form

$$p(y_{D1}, y_{R1}|x_{S1}) = p(y_{R1}|x_{S1})p(y_{D1}|y_{R1}), \tag{5.8}$$

or equivalently, if $X_{S1} \leftrightarrow Y_{R1} \leftrightarrow Y_{D1}$ form a Markov chain;

(b) is *reversely degraded* if $p(y_{D1}, y_{R1}|x_{S1})$ can be written in the form

$$p(y_{D1}, y_{R1}|x_{S1}) = p(y_{D1}|x_{S1})p(y_{R1}|y_{D1}), \tag{5.9}$$

or equivalently, if $X_{S1} \leftrightarrow Y_{D1} \leftrightarrow Y_{R1}$ form a Markov chain;

(c) is *semideterministic* if Y_{R1} is a function of X_{S1}, i.e., if $Y_{R1} = y_{R1}(X_{S1})$;

(d) has *orthogonal sender components* if $\mathcal{X}_S = \mathcal{X}_S' \times \mathcal{X}_S''$ and $p(y_{D1}, y_{R1}|x_{S1})$ can be written in the form

$$p(y_{D1}, y_{R1}|x_{S1}) = p(y_{R1}|x_{S1}')p(y_{D1}|x_{S1}'') \tag{5.10}$$

for all $(x_{S1}', x_{S1}'', y_{D1}, y_{R1}) \in \mathcal{X}_S' \times \mathcal{X}_S'' \times \mathcal{Y}_D \times \mathcal{Y}_R$.

Note that the relay transmit mode is irrelevant to Definition 5.1, i.e., whether a half-duplex relay channel is degraded, reversely degraded, semideterministic, or whether it has orthogonal sender components only depends on the properties of the channel in relay receive mode. Since the capacities of these four half-duplex relay channels are achieved by the DF strategy and/or the PDF strategy, the corresponding capacity results are stated at the end of Section 5.4.

5.2 Cut-Set Bound

As the capacity of the half-duplex relay channel is unknown for the general case, we usually have to study capacity bounds again. To this end, we first introduce the cut-set bound (CSB), which is the best known upper bound on the capacity of the half-duplex relay channel.

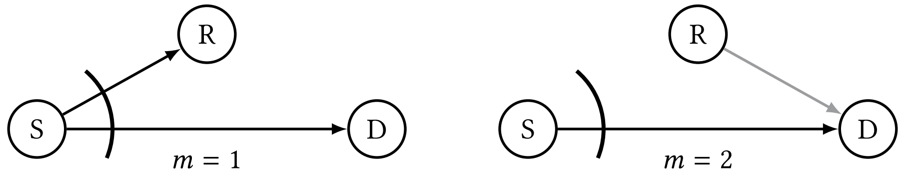
Theorem 5.1. *For any half-duplex relay channel* $\Big\{\mathcal{X}_\mathrm{S} \times \mathcal{X}_\mathrm{R} \times \mathcal{M}, p(y_\mathrm{D}, y_\mathrm{R}|x_\mathrm{S}, x_\mathrm{R}, m), \mathcal{Y}_\mathrm{D} \times \mathcal{Y}_\mathrm{R}\Big\}$*, the capacity is upper bounded by* $C \leq C_\mathrm{CSB}$*, where*

$$C_\mathrm{CSB} = \max_{\substack{R, t_1, t_2 \\ p(x_\mathrm{S1}) p(x_\mathrm{S2}, x_\mathrm{R2})}} R \quad \text{s.t.} \quad \begin{aligned} & R \leq t_1 I(X_\mathrm{S1}; Y_\mathrm{R1}, Y_\mathrm{D1}) + t_2 I(X_\mathrm{S2}; Y_\mathrm{D2}|X_\mathrm{R2}), \\ & R \leq t_1 I(X_\mathrm{S1}; Y_\mathrm{D1}) + t_2 I(X_\mathrm{S2}, X_\mathrm{R2}; Y_\mathrm{D2}), \\ & t_1, t_2 \geq 0, \ t_1 + t_2 = 1. \end{aligned} \tag{5.11}$$

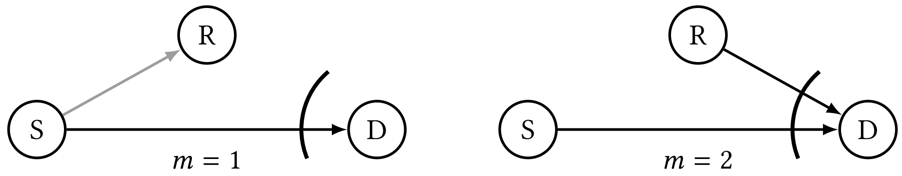
Proof. See [66, Section 3.2], for example. □

This result is obtained by extending the arguments of Cover and El Gamal [21] to the considered half-duplex relay channel model. Like for the full-duplex case, the CSB thus has a nice *max-flow min-cut* interpretation, which is illustrated in Figure 5.2. The first constraint on R in (5.11) implies that the source cannot transmit information to the destination at a higher rate than if the relay and the destination fully cooperate in receiving, cf. Figure 5.2(a). Of course, the relay and the destination can only cooperate in receiving if the relay is in receive mode, i.e., if $m = 1$. If $m = 2$, on the other hand, the relay is transmitting and cannot receive any information due to the half-duplex constraint. The second constraint on R in (5.11) results from the fact that the destination cannot receive information at a higher rate than if the source and the relay cooperate in transmitting, cf. Figure 5.2(b), which, however, is only possible if the relay is in transmit mode, i.e., if $m = 2$.

We remark that the CSB for the half-duplex relay channel can also be viewed as a special case of a more general max-flow min-cut theorem for general multiterminal networks with a finite number of states and the state being known noncausally to all nodes, cf. [66]. In particular, suppose we have a network that is comprised of a finite set of nodes $\mathcal{N}$, and assume that $M \in \mathcal{M}$ denotes the state of the network, where $\mathcal{M}$ is finite. Furthermore, assume that node A sends information to node B at rate $R_\mathrm{A,B}$. If $\mathcal{S} \subseteq \mathcal{N}$ and $\mathcal{S}^\mathrm{c}$ is the complement of $\mathcal{S}$ in $\mathcal{N}$, a *cut* that separates nodes A and B is a partition



(a) Cut bounding the rate of information transmitted by the source ($\mathcal{S} = \{\mathrm{S}\}$)



(b) Cut bounding the rate of information received by the destination ($\mathcal{S} = \{\mathrm{S}, \mathrm{R}\}$)

Figure 5.2: Illustration of the CSB for the Half-Duplex Relay Channel

$(\mathcal{S}, \mathcal{S}^{\mathrm{c}})$ of $\mathcal{N}$ such that $\mathrm{A} \in \mathcal{S}$ and $\mathrm{B} \in \mathcal{S}^{\mathrm{c}}$. Letting $X_{\mathcal{S}m} = \{X_{\mathrm{A}} : \mathrm{A} \in \mathcal{S}\}|\{M = m\}$ and $Y_{\mathcal{S}m} = \{Y_{\mathrm{B}} : \mathrm{B} \in \mathcal{S}\}|\{M = m\}$, we have the following result:

Theorem 5.2. *Let the state of the network M be known noncausally to all nodes in the network. Then, if the information rates $\{R_{\mathrm{A,B}}\}$ are achievable, there exists $\prod_{m\in\mathcal{M}} p(x_{\mathcal{N}m})$ such that*

$$\sum_{\mathrm{A}\in\mathcal{S},\mathrm{B}\in\mathcal{S}^{\mathrm{c}}} R_{\mathrm{A,B}} \le \max_{t_m} \min_{\mathcal{S}} \sum_{m\in\mathcal{M}} t_m I(X_{\mathcal{S}m}; Y_{\mathcal{S}^{\mathrm{c}}m} | X_{\mathcal{S}^{\mathrm{c}}m}), \tag{5.12}$$

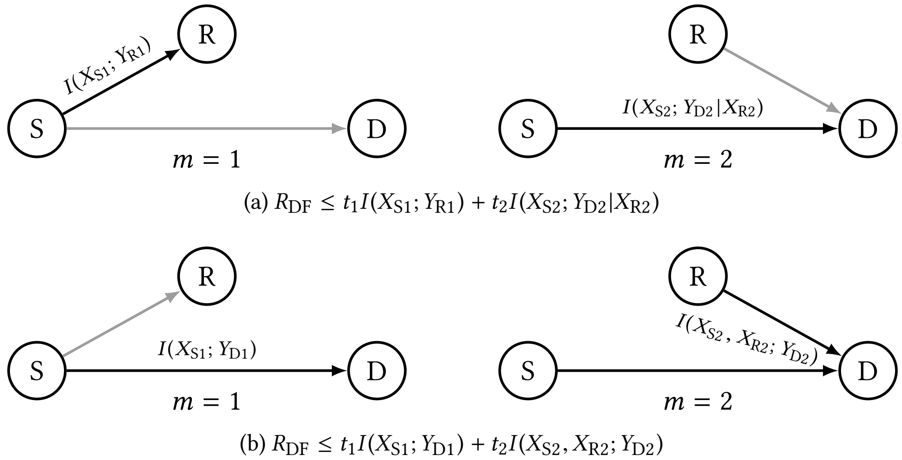
where the minimum is taken over all subsets $\mathcal{S} \subseteq \mathcal{N}$ and the maximum is with respect to all $t_m \ge 0$ such that $\sum_{m\in\mathcal{M}} t_m = 1$.

Proof. See [66, Corollary 2]. □

For the half-duplex relay channel, the node set is given by $\mathcal{N} = \{\mathrm{S}, \mathrm{R}, \mathrm{D}\}$ and there is only one rate $R = R_{\mathrm{S,D}}$. Therefore, we obtain C_{CSB} in (5.11) by applying (5.12) to the two cuts specified by $\mathcal{S} = \{\mathrm{S}\}$ and $\mathcal{S} = \{\mathrm{S}, \mathrm{R}\}$, cf. Figure 5.2. To conclude this section about the CSB for the half-duplex relay channel, we remark that C_{CSB} does not depend on the exact sequence of relay modes m^n, but only on the portions of the time the relay is used in receive and transmit mode, i.e., on the time-shares t_1 and t_2.

5.3 Decode-and-Forward

The first lower bound on the capacity of the half-duplex relay channel we consider is the achievable decode-and-forward (DF) rate. If the relay uses the DF strategy in the half-duplex relay channel, it is required to decode the entire information transmitted by the source *while the relay is in receive mode.* That is, the source may send additional ("new") information to the destination in the relay transmit phase, which the relay is of



(a) $R_{DF} \le t_1 I(X_{S1};Y_{R1}) + t_2 I(X_{S2};Y_{D2}|X_{R2})$

(b) $R_{DF} \le t_1 I(X_{S1};Y_{D1}) + t_2 I(X_{S2},X_{R2};Y_{D2})$

Figure 5.3: Illustration of the Rate Bounds for the Half-Duplex DF Strategy

course not able to decode as it cannot receive any information when it is transmitting. Since the relay does not have to decode the entire information the source communicates to the destination, this strategy is sometimes already referred to as a partial decode-and-forward (PDF) strategy, cf. [116], for example. Nevertheless, we consider this strategy as a DF strategy because the relay decodes everything it can receive.

THEOREM 5.3. *The capacity of the half-duplex relay channel specified by* $\left\{\mathcal{X}_S \times \mathcal{X}_R \times \mathcal{M}, p(y_D, y_R|x_S, x_R, m), \mathcal{Y}_D \times \mathcal{Y}_R\right\}$ *is lower bounded by* $C \ge R_{DF}$, *where*

$$R_{DF} = \max_{\substack{R,t_1,t_2\\ p(x_{S1})p(x_{S2},x_{R2})}} R \quad \text{s.t.} \quad \begin{aligned} &R \le t_1 I(X_{S1};Y_{R1}) + t_2 I(X_{S2};Y_{D2}|X_{R2}),\\ &R \le t_1 I(X_{S1};Y_{D1}) + t_2 I(X_{S2},X_{R2};Y_{D2}),\\ &t_1, t_2 \ge 0,\ t_1 + t_2 = 1 \end{aligned} \tag{5.13}$$

denotes the maximum rate that can be achieved with the decode-and-forward (DF) protocol.

Proof. See [68, Section II], for example. □

Similar to the CSB, this result is obtained by extending the DF strategy derived by Cover and El Gamal [21] to the considered half-duplex relay channel model, and like for the full-duplex case, the achievable DF rate is limited by two mutual information terms. The first rate bound, $R_{DF} \le t_1 I(X_{S1};Y_{R1}) + t_2 I(X_{S2};Y_{D2}|X_{R2})$, is due to the condition that the relay must decode the entire information transmitted by the source in the relay receive phase and the fact that the source may transmit new information to the destination in the relay transmit phase. The second rate bound, $R_{DF} \le t_1 I(X_{S1};Y_{D1}) + t_2 I(X_{S2},X_{R2};Y_{D2})$, again states that the achievable DF rate cannot be higher than the combined information transfer from the source and the relay to the destination, with the restriction that the relay does not transmit when it is receiving, i.e., if $m = 1$. An illustration of these two rate bounds is given in Figure 5.3.

Note also that the optimization problems that determine R_{DF} and C_{CSB} only differ in the first mutual information term of the first rate bound, where an additional Y_{D1} appears in (5.11) as compared to (5.13). Since the second rate bound, the optimization variables, and the constraints are exactly the same, the DF strategy performs quite well whenever the source-to-relay link is (much) better than the source-to-destination link. For example, DF is the capacity achieving strategy for the degraded half-duplex relay channel, cf. Section 5.4.

In contrast to the full-duplex case, where R_{DF} is achieved by a block Markov superposition encoding scheme that is based on conveying $B-1$ messages from the source to the destination in B blocks of transmission, the DF coding scheme for the half-duplex relay channel operates on only one block. This one block consists of n channel uses, where the relay is in receive mode for the first n_1 channel uses and in transmit mode for the remaining $n_2 = n - n_1$ channel uses.

More specifically, the source splits its message W into two independent parts W_1 and W_2, and the source encoder then assigns two codewords $X_{\mathrm{S1}}^{n_1}(w_1)$ and $X_{\mathrm{S2}}^{n_2}(w_1, w_2)$ to each $(w_1, w_2) = w \in \mathcal{W}$. Assuming that the relay has correctly decoded w_1 after the relay receive phase, which is justified if the rate associated with W_1 is no greater than $t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{R1}})$, both the source and the relay know w_1. If the relay encoder is deterministic, the source can determine the relay's codeword $x_{\mathrm{R2}}^{n_2}(w_1)$ sent in the relay transmit phase, which implies that the source can generate $X_{\mathrm{S2}}^{n_2}(w_1, w_2)$ from $x_{\mathrm{R2}}^{n_2}(w_1)$ via superposition coding. As a result, input distributions of the form $p(x_{\mathrm{S2}}|x_{\mathrm{R2}})p(x_{\mathrm{R2}}) = p(x_{\mathrm{S2}}, x_{\mathrm{R2}})$ are permissible for the relay transmit phase.

We remark that in this DF coding scheme, W_2 represents the new information the source may communicate to the destination in the relay transmit phase. If $W_2 = \emptyset$, or equivalently, if $W_1 = W$, it follows that $I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|X_{\mathrm{R2}}) = 0$ and we obtain the DF strategy that is referred to as "full DF" in [116]. Like for the full-duplex case, it may happen that the best rate that can be achieved with this strategy is smaller than

$$R_{\mathrm{P2P}} = \max_{x_{\mathrm{R2}} \in \mathcal{X}_{\mathrm{R}}} \max_{p(x_{\mathrm{S2}})} I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|x_{\mathrm{R2}}). \tag{5.14}$$

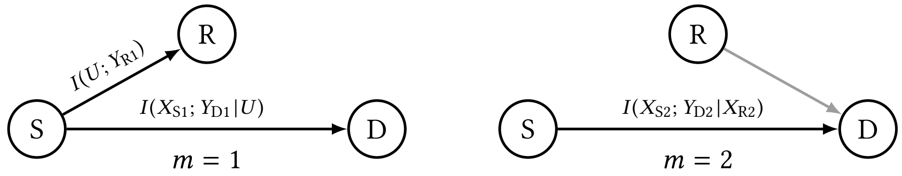
However, the maximum achievable DF rate R_{DF} as defined in Theorem 5.3 is never smaller than R_{P2P}, i.e.,

$$R_{\mathrm{DF}} \geq R_{\mathrm{P2P}}, \tag{5.15}$$

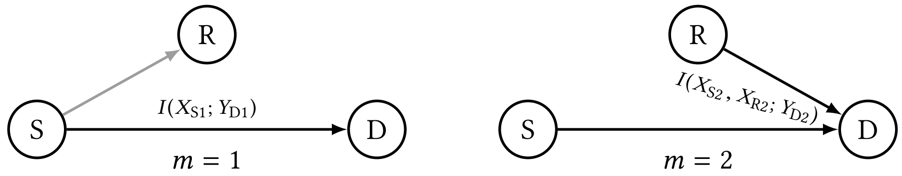
since R_{P2P} can always be achieved by setting $t_1 = 0$ and $t_2 = 1$ in (5.13).

5.4 Partial Decode-and-Forward

While the considered half-duplex DF strategy never performs worse than direct transmission, higher rates can usually be achieved if the constraint that the relay decode the entire information the source transmits during the relay receive phase is relaxed. To



(a) $R_{\mathrm{PDF}} \leq t_1 I(U; Y_{\mathrm{R1}}) + t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}|U) + t_2 I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|X_{\mathrm{R2}})$



(b) $R_{\mathrm{PDF}} \leq t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}) + t_2 I(X_{\mathrm{S2}}, X_{\mathrm{R2}}; Y_{\mathrm{D2}})$

Figure 5.4: Illustration of the Rate Bounds for the Half-Duplex PDF Strategy

this end, we consider the partial decode-and-forward (PDF) strategy, where the part W_1 of the source message $W = (W_1, W_2)$ the source transmits during the relay receive phase is further decomposed into two independent parts W_1' and W_1'' of which the relay is only required to decode W_1'. By constructing separate codebooks for W_1' and W_1'' and using superposition coding at the source to generate $X_{\mathrm{S1}}^{n_1}(w_1', w_1'')$, the following lower bound on the capacity of the half-duplex relay channel can be obtained:

Theorem 5.4. *The capacity of the half-duplex relay channel specified by* $\left\{\mathcal{X}_{\mathrm{S}} \times \mathcal{X}_{\mathrm{R}} \times \mathcal{M},\ p(y_{\mathrm{D}}, y_{\mathrm{R}}|x_{\mathrm{S}}, x_{\mathrm{R}}, m), \mathcal{Y}_{\mathrm{D}} \times \mathcal{Y}_{\mathrm{R}}\right\}$ *is lower bounded by* $C \geq R_{\mathrm{PDF}}$*, where*

$$R_{\mathrm{PDF}} = \max_{\substack{R, t_1, t_2 \\ p(u, x_{\mathrm{S1}}) p(x_{\mathrm{S2}}, x_{\mathrm{R2}})}} R \quad \text{s.t.} \quad \begin{aligned} R &\leq t_1 I(U; Y_{\mathrm{R1}}) + t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}|U) \\ &\quad + t_2 I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|X_{\mathrm{R2}}), \\ R &\leq t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}) + t_2 I(X_{\mathrm{S2}}, X_{\mathrm{R2}}; Y_{\mathrm{D2}}), \\ t_1, t_2 &\geq 0,\ t_1 + t_2 = 1 \end{aligned} \tag{5.16}$$

denotes the maximum rate that can be achieved with the partial decode-and-forward (PDF) protocol, and where the maximization over $p(u, x_{\mathrm{S1}})$ *is subject to the constraint that* $U \leftrightarrow X_{\mathrm{S1}} \leftrightarrow (Y_{\mathrm{D1}}, Y_{\mathrm{R1}})$ *form a Markov chain.*

Proof. The achievability of R_{PDF} can be proved by extending the half-duplex DF coding scheme as explained above. A rigorous proof, which follows along the same lines as the proof of Theorem 5.3 in [68, Section II], is given in [118]. □

Similar to the full-duplex case, U is an auxiliary random variable which represents the part of the information the relay must decode (after the relay receive phase). By optimizing over (the distribution of) U, the PDF strategy therefore allows to tradeoff sending information to the destination via the relay versus sending it over the direct link (in the relay receive phase). This tradeoff is also reflected in the first two mutual information terms of the first rate bound on R_{PDF} in (5.16), cf. Figure 5.4(a). The second rate bound, on the other hand, is again the same as for the CSB and the DF rate in (5.11) and (5.13), respectively. It is given by the rate of the information the source and the relay can cooperatively send to the destination, with the restriction that the relay does not transmit when it is receiving, cf. Figure 5.4(b).

Like for the full-duplex case, it is also easy to verify that the PDF strategy includes the DF strategy as a special case. In particular, note that if we choose $U = X_{\mathrm{S1}}$ in (5.16), it follows that $t_1 I(U; Y_{\mathrm{R1}}) + t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}|U) = t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{R1}})$ so that PDF reduces to DF. Moreover, as the DF strategy already includes direct transmission as a special case for the half-duplex relay channel, cf. (5.15), we can conclude that

$$R_{\mathrm{PDF}} \geq R_{\mathrm{DF}} \geq R_{\mathrm{P2P}}. \tag{5.17}$$

Beyond that, it is easily shown that PDF is the capacity achieving strategy for the four half-duplex relay channels defined in Section 5.1:

(a) The capacity of the degraded half-duplex relay channel is given by [68]

$$C = \max_{\substack{R, t_1, t_2 \\ p(x_{\mathrm{S1}})p(x_{\mathrm{S2}}, x_{\mathrm{R2}})}} R \quad \text{s.t.} \quad \begin{aligned} & R \leq t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{R1}}) + t_2 I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|X_{\mathrm{R2}}), \\ & R \leq t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}) + t_2 I(X_{\mathrm{S2}}, X_{\mathrm{R2}}; Y_{\mathrm{D2}}), \\ & t_1, t_2 \geq 0, \ t_1 + t_2 = 1. \end{aligned} \tag{5.18}$$

It is achieved by the DF strategy, and the converse follows from the fact that $I(X_{\mathrm{S1}}; Y_{\mathrm{R1}}, Y_{\mathrm{D1}}) = I(X_{\mathrm{S1}}; Y_{\mathrm{R1}})$ if the half-duplex relay channel is degraded since $X_{\mathrm{S1}} \leftrightarrow Y_{\mathrm{R1}} \leftrightarrow Y_{\mathrm{D1}}$ form a Markov chain in this case.

(b) If the half-duplex relay channel is reversely degraded, on the other hand, the capacity is achieved by direct transmission, i.e.,

$$C = \max_{x_{\mathrm{R2}} \in \mathcal{X}_{\mathrm{R}}} \max_{p(x_{\mathrm{S2}})} I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|x_{\mathrm{R2}}). \tag{5.19}$$

The converse follows from the fact that $I(X_{\mathrm{S1}}; Y_{\mathrm{R1}}, Y_{\mathrm{D1}}) = I(X_{\mathrm{S1}}; Y_{\mathrm{D1}})$ if the half-duplex relay channel is reversely degraded since $X_{\mathrm{S1}} \leftrightarrow Y_{\mathrm{D1}} \leftrightarrow Y_{\mathrm{R1}}$ form a Markov chain in this case.

(c) The capacity of the semideterministic half-duplex relay channel is equal to

$$C = \max_{\substack{R, t_1, t_2 \\ p(x_{\mathrm{S1}})p(x_{\mathrm{S2}}, x_{\mathrm{R2}})}} R \quad \text{s.t.} \quad \begin{aligned} & R \leq t_1 H(Y_{\mathrm{R1}}) + t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}|Y_{\mathrm{R1}}) \\ & \qquad + t_2 I(X_{\mathrm{S2}}; Y_{\mathrm{D2}}|X_{\mathrm{R2}}), \\ & R \leq t_1 I(X_{\mathrm{S1}}; Y_{\mathrm{D1}}) + t_2 I(X_{\mathrm{S2}}, X_{\mathrm{R2}}; Y_{\mathrm{D2}}), \\ & t_1, t_2 \geq 0, \ t_1 + t_2 = 1. \end{aligned} \tag{5.20}$$

Achievability follows from the PDF strategy with $U = Y_{\mathrm{R1}}$, which is feasible as Y_{R1} is a function of X_{S1}, and the converse follows from the CSB.

(d) Finally, if the half-duplex relay channel has orthogonal sender components, its capacity is given by

$$C = \max_{\substack{R, t_1, t_2 \\ p(x'_{S1}, x''_{S1}) p(x_{S2}, x_{R2})}} R \quad \text{s.t.} \quad \begin{aligned} R &\le t_1 I(X'_{S1}; Y_{R1}) + t_1 I(X''_{S1}; Y_{D1}) \\ &\quad + t_2 I(X_{S2}; Y_{D2} | X_{R2}), \\ R &\le t_1 I(X''_{S1}; Y_{D1}) + t_2 I(X_{S2}, X_{R2}; Y_{D2}), \\ t_1&, t_2 \ge 0,\ t_1 + t_2 = 1, \end{aligned} \tag{5.21}$$

where it is sufficient to maximize with respect to joint probability mass functions $p(x'_{S1}, x''_{S1})$ that factor as $p(x'_{S1})p(x''_{S1})$. The proof of achievability uses the PDF strategy with $U = X'_{S1}$, and the converse again follows from the CSB.

5.5 Bidirectional Communication

In this section, we extend the half-duplex relay channel model and the corresponding information theoretical results to the half-duplex *two-way* relay channel, which models the more general scenario where two terminals exchange information with the help of a relay. More specifically, we consider the *restricted* half-duplex two-way relay channel in this work. That is, we assume that the bidirectional communication is restricted in the sense that the encoders at the two terminals may neither cooperate, nor are they allowed to use previously decoded information to encode their messages.

The discrete memoryless half-duplex two-way relay channel, specified by

$$\left\{ \mathcal{X}_A \times \mathcal{X}_B \times \mathcal{X}_R \times \mathcal{M}, p(y_A, y_B, y_R | x_A, x_B, x_R, m), \mathcal{Y}_A \times \mathcal{Y}_B \times \mathcal{Y}_R \right\}, \tag{5.22}$$

consists of seven finite sets $\mathcal{X}_A$, $\mathcal{X}_B$, $\mathcal{X}_R$, $\mathcal{M}$, $\mathcal{Y}_A$, $\mathcal{Y}_B$, $\mathcal{Y}_R$ and a collection of probability mass functions $p(y_A, y_B, y_R | x_A, x_B, x_R, m)$, one for each $(x_A, x_B, x_R, m) \in \mathcal{X}_A \times \mathcal{X}_B \times \mathcal{X}_R \times \mathcal{M}$. The interpretation is that x_A, x_B, x_R and y_A, y_B, y_R are the channel inputs and outputs, respectively, of terminal A, terminal B, and the relay, cf. Figure 5.5. Without loss of generality, it is assumed here that the three transmit alphabets $\mathcal{X}_A$, $\mathcal{X}_B$, and $\mathcal{X}_R$ contain a "quiet" symbol q and that the three receive alphabets $\mathcal{Y}_A$, $\mathcal{Y}_B$, and $\mathcal{Y}_R$ contain an "erasure" symbol e.

In order to account for the half-duplex constraint, the half-duplex two-way relay channel model contains the random variable $M \in \mathcal{M} = \{1, \dots, 6\}$, which models the state/mode of the three-terminal relay network, i.e., which nodes transmit and which nodes receive. Note that $|\mathcal{M}| = 6$ since there are six states/modes where either one or two nodes transmit. Evidently, no information can be conveyed when all nodes are silent or when all nodes transmit simultaneously, where the latter is due to the half-duplex constraint imposed on all nodes. We define these six states as follows:

$M = 1$: Terminal A transmits to the relay and terminal B.

$M = 2$: Terminal B transmits to the relay and terminal A.

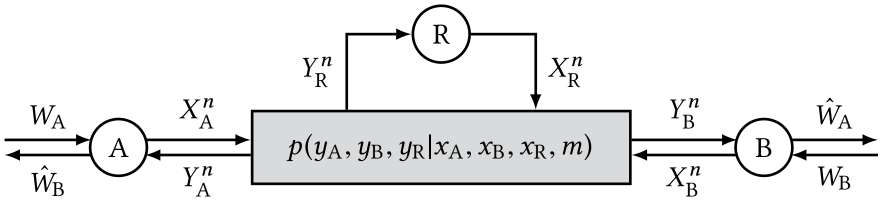


Figure 5.5: Illustration of the Half-Duplex Two-Way Relay Channel

$M = 3$: Terminals A and B transmit to the relay.
$M = 4$: The relay transmits to terminals A and B.
$M = 5$: Terminal B and the relay transmit to terminal A.
$M = 6$: Terminal A and the relay transmit to terminal B.

Using (5.2), the constraint that the nodes cannot simultaneously transmit and receive in the same frequency band can be expressed in terms of the conditional probability mass function that specifies the half-duplex two-way relay channel as

$$p(y_A, y_B, y_R|x_A, x_B, x_R, m) = \begin{cases} p(y_{B1}, y_{R1}|x_{A1})\delta_q(x_{B1})\delta_q(x_{R1})\delta_e(y_{A1}) & \text{if } m = 1, \\ p(y_{A2}, y_{R2}|x_{B2})\delta_q(x_{A2})\delta_q(x_{R2})\delta_e(y_{B2}) & \text{if } m = 2, \\ p(y_{R3}|x_{A3}, x_{B3})\delta_q(x_{R3})\delta_e(y_{A3})\delta_e(y_{B3}) & \text{if } m = 3, \\ p(y_{A4}, y_{B4}|x_{R4})\delta_q(x_{A4})\delta_q(x_{B4})\delta_e(y_{R4}) & \text{if } m = 4, \\ p(y_{A5}|x_{B5}, x_{R5})\delta_q(x_{A5})\delta_e(y_{B5})\delta_e(y_{R5}) & \text{if } m = 5, \\ p(y_{B6}|x_{A6}, x_{R6})\delta_q(x_{B6})\delta_e(y_{A6})\delta_e(y_{R6}) & \text{if } m = 6. \end{cases} \tag{5.23}$$

Like for the half-duplex relay channel, we assume that all nodes know the state of the network noncausally, cf. Remark 5.2, which implies that M cannot be used to exchange information between terminals A and B.

Let $m^n \in \mathcal{M}^n$ be a sequence of states that is noncausally known to all three nodes. Then, a $(2^{nR_A}, 2^{nR_B}, n)$ code for the restricted half-duplex two-way relay channel consists of two message sets $\mathcal{W}_A = \{1, 2, \ldots, \lceil 2^{nR_A} \rceil\}$ and $\mathcal{W}_B = \{1, 2, \ldots, \lceil 2^{nR_B} \rceil\}$, two encoders (one at each terminal) that assign codewords $X_A^n(w_A) \in \mathcal{X}_A^n$ and $X_B^n(w_B) \in \mathcal{X}_B^n$ to each $w_A \in \mathcal{W}_A$ and $w_B \in \mathcal{W}_B$, respectively, a relay encoder that assigns a symbol $X_{R,i}(y_R^{i-1}) \in \mathcal{X}_R$ to each past received sequence $y_R^{i-1} \in \mathcal{Y}_R^{i-1}$ for each $i \in \{1, \ldots, n\}$, and two decoders (one at each terminal again) that assign estimates $\hat{W}_A(y_B^n) \in \mathcal{W}_A$ and $\hat{W}_B(y_A^n) \in \mathcal{W}_B$ (or possibly an error message) to each received sequence $y_B^n \in \mathcal{Y}_B^n$ and $y_A^n \in \mathcal{Y}_A^n$, respectively. For generality, the encoding and decoding functions are allowed to be stochastic again.

The definition of the relay encoder already includes the condition that the relay's transmit symbol $X_{R,i}$ may depend on its past observations $Y_R^{i-1} = (Y_{R,1}, \ldots, Y_{R,i-1})$ only, which ensures that the relay operates in a *causal* manner. Similarly, the terminal encoders are restricted in the sense that the codewords X_A^n and X_B^n may depend on the messages w_{A} and w_{B} only. Beyond that, however, all nodes must satisfy the half-duplex constraint. To this end, we require the encoders at the terminals and the relay to select the quiet symbol whenever the corresponding node is in receive mode, i.e., for any $i \in \{1, \dots, n\}$, we require

$$\begin{aligned} X_{\mathrm{A},i}(w_{\mathrm{A}}) &= q \quad \text{if } m_i \in \{2,4,5\}, \\ X_{\mathrm{B},i}(w_{\mathrm{B}}) &= q \quad \text{if } m_i \in \{1,4,6\}, \\ X_{\mathrm{R},i}(y_{\mathrm{R}}^{i-1}) &= q \quad \text{if } m_i \in \{1,2,3\}. \end{aligned} \tag{5.24}$$

The half-duplex two-way relay channel is *memoryless* in the sense that the current channel outputs $(Y_{\mathrm{A},i}, Y_{\mathrm{B},i}, Y_{\mathrm{R},i})$ depend on all previous channel inputs and states $(X_{\mathrm{A}}^i, X_{\mathrm{B}}^i, X_{\mathrm{R}}^i, M^i)$ only through $(X_{\mathrm{A},i}, X_{\mathrm{B},i}, X_{\mathrm{R},i}, M_i)$. For any $p(w_{\mathrm{A}})$, $p(w_{\mathrm{B}})$, and choice of the code, the joint probability mass function on $\mathcal{W}_{\mathrm{A}} \times \mathcal{W}_{\mathrm{B}} \times \mathcal{X}_{\mathrm{A}}^n \times \mathcal{X}_{\mathrm{B}}^n \times \mathcal{X}_{\mathrm{R}}^n \times \mathcal{M}^n \times \mathcal{Y}_{\mathrm{A}}^n \times \mathcal{Y}_{\mathrm{B}}^n \times \mathcal{Y}_{\mathrm{R}}^n$ hence factors as

$$\begin{aligned} p(w_{\mathrm{A}}, w_{\mathrm{B}}, x_{\mathrm{A}}^n, x_{\mathrm{B}}^n, x_{\mathrm{R}}^n, m^n, y_{\mathrm{A}}^n, y_{\mathrm{B}}^n, y_{\mathrm{R}}^n) &= p(w_{\mathrm{A}})p(w_{\mathrm{B}})p(x_{\mathrm{A}}^n|w_{\mathrm{A}})p(x_{\mathrm{B}}^n|w_{\mathrm{B}}) \cdot \\ &\prod_{i=1}^{n} p(x_{\mathrm{R},i}|y_{\mathrm{R}}^{i-1}) p(y_{\mathrm{A},i}, y_{\mathrm{B},i}, y_{\mathrm{R},i}|x_{\mathrm{A},i}, x_{\mathrm{B},i}, x_{\mathrm{R},i}, m_i), \end{aligned} \tag{5.25}$$

where $p(y_{\mathrm{A},i}, y_{\mathrm{B},i}, y_{\mathrm{R},i}|x_{\mathrm{A},i}, x_{\mathrm{B},i}, x_{\mathrm{R},i}, m_i)$ and $p(x_{\mathrm{R},i}|y_{\mathrm{R}}^{i-1})$, $p(x_{\mathrm{A},i}|w_{\mathrm{A}})$, $p(x_{\mathrm{B},i}|w_{\mathrm{B}})$ must satisfy (5.23) and (5.24), respectively, for any $i \in \{1, \dots, n\}$. Finally, the transmissions are modeled as taking place *synchronously* again.

If W_{A} and W_{B} are uniformly distributed over $\mathcal{W}_{\mathrm{A}}$ and $\mathcal{W}_{\mathrm{B}}$, respectively, and if $P_e^{(n)} = \Pr[\hat{W}_{\mathrm{A}} \neq W_{\mathrm{A}} \vee \hat{W}_{\mathrm{B}} \neq W_{\mathrm{B}}]$ denotes the average probability of error, a rate pair $(R_{\mathrm{A}}, R_{\mathrm{B}})$ is said to be *achievable* if there exists a sequence of $(2^{nR_{\mathrm{A}}}, 2^{nR_{\mathrm{B}}}, n)$ codes for which $P_e^{(n)} \to 0$ as $n \to \infty$. The *capacity region* $\mathcal{C}$ of the restricted half-duplex two-way relay channel is defined as the closure of the set of achievable rate pairs.

Theorem 5.5. *Suppose the rate pair* $(R_{\mathrm{A}}, R_{\mathrm{B}})$ *is achievable in the half-duplex two-way relay channel, then* $(R_{\mathrm{A}}, R_{\mathrm{B}}) \in \mathcal{C}_{\mathrm{CSB}}$ *with*

$$\begin{aligned} \mathcal{C}_{\mathrm{CSB}} = \bigcup_{\prod_{m=1}^{6} p(x_{\mathrm{A}m}, x_{\mathrm{B}m}, x_{\mathrm{R}m})} \Big\{ &(R_{\mathrm{A}}, R_{\mathrm{B}}) \in \mathbb{R}_+^2 : t_1, \dots, t_6 \geq 0, \textstyle\sum_{m=1}^{6} t_m = 1, \\ &R_{\mathrm{A}} \leq t_1 I(X_{\mathrm{A}1}; Y_{\mathrm{R}1}, Y_{\mathrm{B}1}) + t_3 I(X_{\mathrm{A}3}; Y_{\mathrm{R}3}|X_{\mathrm{B}3}) + t_6 I(X_{\mathrm{A}6}; Y_{\mathrm{B}6}|X_{\mathrm{R}6}), \\ &R_{\mathrm{A}} \leq t_1 I(X_{\mathrm{A}1}; Y_{\mathrm{B}1}) + t_4 I(X_{\mathrm{R}4}; Y_{\mathrm{B}4}) + t_6 I(X_{\mathrm{A}6}, X_{\mathrm{R}6}; Y_{\mathrm{B}6}), \\ &R_{\mathrm{B}} \leq t_2 I(X_{\mathrm{B}2}; Y_{\mathrm{R}2}, Y_{\mathrm{A}2}) + t_3 I(X_{\mathrm{B}3}; Y_{\mathrm{R}3}|X_{\mathrm{A}3}) + t_5 I(X_{\mathrm{B}5}; Y_{\mathrm{A}5}|X_{\mathrm{R}5}), \\ &R_{\mathrm{B}} \leq t_2 I(X_{\mathrm{B}2}; Y_{\mathrm{A}2}) + t_4 I(X_{\mathrm{R}4}; Y_{\mathrm{A}4}) + t_5 I(X_{\mathrm{B}5}, X_{\mathrm{R}5}; Y_{\mathrm{A}5}) \Big\}. \end{aligned} \tag{5.26}$$

Proof. This result is a direct consequence of Theorem 5.2. In particular, the two rate bounds on $R_{\mathrm{A}} = R_{\mathrm{A,B}}$ are obtained from the cuts $\mathcal{S} = \{\mathrm{A}\}$ and $\mathcal{S} = \{\mathrm{A}, \mathrm{R}\}$, whereas the two rate bounds on $R_{\mathrm{B}} = R_{\mathrm{B,A}}$ follow from the cuts $\mathcal{S} = \{\mathrm{B}\}$ and $\mathcal{S} = \{\mathrm{B}, \mathrm{R}\}$. □

The outer bound on the capacity region defined in Theorem 5.5 is referred to as the *cut-set outer bound* (CSOB) in the following. We remark that the CSOB is not only valid for the restricted half-duplex two-way relay channel, but also for half-duplex two-way relay channels where the terminal encoders may cooperate. This is because input distributions of the form $p(x_{A3}, x_{B3})$ are permissible in Theorem 5.5. For any half-duplex two-way Gaussian relay channel, however, we show in Chapter 7 that $\mathcal{C}_{CSB}$ is attained if the joint distribution factors as $p(x_{A3}, x_{B3}) = p(x_{A3})p(x_{B3})$.

Whereas the CSOB region $\mathcal{C}_{CSB}$ does not depend on the sequence of states m^n, but only on the time-shares $t_m = n_m/n$, the coding schemes that yield the achievable DF and PDF rate regions given below are based on a specific state sequence. In particular, the DF and PDF coding schemes for the restricted half-duplex two-way relay channel operate on one block of n channel uses where $m = 1$ for the first n_1 channel uses, $m = 2$ for the next n_2 channel uses, and so forth. Therefore, we commonly use the term *phase m* when we refer to the n_m consecutive channel uses the half-duplex two-way relay channel is used in state m.

If the relay uses the DF strategy in the half-duplex two-way relay channel, it is required to decode the entire information transmitted by terminals A and B *while the relay is in receive mode*, i.e., in phases 1, 2, and 3. Similar to the case of unidirectional communication in the half-duplex relay channel, this means that in phases 5 and 6, where one terminal and the relay transmit to the respective other terminal, the two terminals may exchange information which the relay does not decode.

Theorem 5.6. *If $p(x_{A3}, x_{B3}) = p(x_{A3})p(x_{B3})$, the following rate region is achievable for the restricted half-duplex two-way relay channel:*

$$\begin{aligned}
\mathcal{R}_{DF} = \bigcup_{\prod_{m=1}^{6} p(x_{Am}, x_{Bm}, x_{Rm})} \Big\{ & (R_A, R_B) \in \mathbb{R}_+^2 : t_1, \dots, t_6 \geq 0, \textstyle\sum_{m=1}^{6} t_m = 1, \\
R_A &\leq t_1 I(X_{A1}; Y_{R1}) + t_3 I(X_{A3}; Y_{R3}|X_{B3}) + t_6 I(X_{A6}; Y_{B6}|X_{R6}), \\
R_A &\leq t_1 I(X_{A1}; Y_{B1}) + t_4 I(X_{R4}; Y_{B4}) + t_6 I(X_{A6}, X_{R6}; Y_{B6}), \\
R_B &\leq t_2 I(X_{B2}; Y_{R2}) + t_3 I(X_{B3}; Y_{R3}|X_{A3}) + t_5 I(X_{B5}; Y_{A5}|X_{R5}), \\
R_B &\leq t_2 I(X_{B2}; Y_{A2}) + t_4 I(X_{R4}; Y_{A4}) + t_5 I(X_{B5}, X_{R5}; Y_{A5}), \\
R_A + R_B &\leq t_1 I(X_{A1}; Y_{R1}) + t_2 I(X_{B2}; Y_{R2}) + t_3 I(X_{A3}, X_{B3}; Y_{R3}) \\
&\quad + t_5 I(X_{B5}; Y_{A5}|X_{R5}) + t_6 I(X_{A6}; Y_{B6}|X_{R6}) \Big\}.
\end{aligned} \tag{5.27}$$

Proof. See [118] for a rigorous proof. An outline of the coding scheme that achieves $\mathcal{R}_{DF}$ is given in Appendix B.3.1. □

Note that the achievable DF rate region $\mathcal{R}_{DF}$ and the CSOB region $\mathcal{C}_{CSB}$ are defined by similar rate constraints. In particular, the second rate bounds on R_A and R_B in (5.27) are exactly the same as in (5.26), and the first rate bounds only differ in the first mutual information terms, where (Y_{R1}, Y_{B1}) and (Y_{R2}, Y_{A2}) in (5.26) are replaced by Y_{R1} and Y_{R2}

in (5.27), respectively. The main difference between $\mathcal{R}_{\mathrm{DF}}$ and $\mathcal{C}_{\mathrm{CSB}}$ is the additional constraint on the sum rate in (5.27). This constraint results from the fact that the relay must decode the entire information transmitted by terminal A and terminal B in phase 3 when it uses DF.

A better inner bound on the capacity region $\mathcal{C}$ can again be obtained if the relay uses the PDF strategy, where the constraint that the relay decode the entire information transmitted by the terminals in phases 1 and 2 is relaxed.[2] More specifically, if the relay is only required to decode some parts of the messages transmitted by terminals A and B during phases 1 and 2, respectively, the following rate region can be achieved:

Theorem 5.7. *If $p(x_{\mathrm{A3}}, x_{\mathrm{B3}}) = p(x_{\mathrm{A3}})p(x_{\mathrm{B3}})$ and $p(u_{\mathrm{A}}, x_{\mathrm{A1}})$ and $p(u_{\mathrm{B}}, x_{\mathrm{B2}})$ are such that $U_{\mathrm{A}} \leftrightarrow X_{\mathrm{A1}} \leftrightarrow (Y_{\mathrm{B1}}, Y_{\mathrm{R1}})$ and $U_{\mathrm{B}} \leftrightarrow X_{\mathrm{B2}} \leftrightarrow (Y_{\mathrm{A2}}, Y_{\mathrm{R2}})$ form Markov chains, respectively, the following rate region is achievable for the restricted half-duplex two-way relay channel:*

$$\begin{aligned}\mathcal{R}_{\mathrm{PDF}} = \bigcup_{\substack{p(u_{\mathrm{A}},x_{\mathrm{A1}})p(u_{\mathrm{B}},x_{\mathrm{B2}})\\ \prod_{m=3}^{6} p(x_{\mathrm{A}m},x_{\mathrm{B}m},x_{\mathrm{R}m})}} &\Big\{(R_{\mathrm{A}}, R_{\mathrm{B}}) \in \mathbb{R}_+^2 : t_1, \ldots, t_6 \geq 0, \textstyle\sum_{m=1}^{6} t_m = 1,\\
R_{\mathrm{A}} &\leq t_1 I(U_{\mathrm{A}}; Y_{\mathrm{R1}}) + t_1 I(X_{\mathrm{A1}}; Y_{\mathrm{B1}}|U_{\mathrm{A}}) + t_3 I(X_{\mathrm{A3}}; Y_{\mathrm{R3}}|X_{\mathrm{B3}}) + t_6 I(X_{\mathrm{A6}}; Y_{\mathrm{B6}}|X_{\mathrm{R6}}),\\
R_{\mathrm{A}} &\leq t_1 I(X_{\mathrm{A1}}; Y_{\mathrm{B1}}) + t_4 I(X_{\mathrm{R4}}; Y_{\mathrm{B4}}) + t_6 I(X_{\mathrm{A6}}, X_{\mathrm{R6}}; Y_{\mathrm{B6}}),\\
R_{\mathrm{B}} &\leq t_2 I(U_{\mathrm{B}}; Y_{\mathrm{R2}}) + t_2 I(X_{\mathrm{B2}}; Y_{\mathrm{A2}}|U_{\mathrm{B}}) + t_3 I(X_{\mathrm{B3}}; Y_{\mathrm{R3}}|X_{\mathrm{A3}}) + t_5 I(X_{\mathrm{B5}}; Y_{\mathrm{A5}}|X_{\mathrm{R5}}),\\
R_{\mathrm{B}} &\leq t_2 I(X_{\mathrm{B2}}; Y_{\mathrm{A2}}) + t_4 I(X_{\mathrm{R4}}; Y_{\mathrm{A4}}) + t_5 I(X_{\mathrm{B5}}, X_{\mathrm{R5}}; Y_{\mathrm{A5}}),\\
R_{\mathrm{A}} + R_{\mathrm{B}} &\leq t_1 I(U_{\mathrm{A}}; Y_{\mathrm{R1}}) + t_1 I(X_{\mathrm{A1}}; Y_{\mathrm{B1}}|U_{\mathrm{A}}) + t_2 I(U_{\mathrm{B}}; Y_{\mathrm{R2}}) + t_2 I(X_{\mathrm{B2}}; Y_{\mathrm{A2}}|U_{\mathrm{B}})\\
&\quad + t_3 I(X_{\mathrm{A3}}, X_{\mathrm{B3}}; Y_{\mathrm{R3}}) + t_5 I(X_{\mathrm{B5}}; Y_{\mathrm{A5}}|X_{\mathrm{R5}}) + t_6 I(X_{\mathrm{A6}}; Y_{\mathrm{B6}}|X_{\mathrm{R6}})\Big\}.\end{aligned} \tag{5.28}$$

Proof. This result is a rather straightforward generalization of Theorem 5.6, cf. [118]. An outline of the coding scheme that achieves $\mathcal{R}_{\mathrm{PDF}}$ is given in Appendix B.3.2. □

Here, U_{A} and U_{B} are again auxiliary random variables which represent the parts of the information the relay must decode. Obviously, if we choose $U_{\mathrm{A}} = X_{\mathrm{A1}}$ and $U_{\mathrm{B}} = X_{\mathrm{B2}}$ in (5.28), $\mathcal{R}_{\mathrm{PDF}}$ reduces to $\mathcal{R}_{\mathrm{DF}}$ so that we have

$$\mathcal{R}_{\mathrm{DF}} \subseteq \mathcal{R}_{\mathrm{PDF}} \subseteq \mathcal{C} \subseteq \mathcal{C}_{\mathrm{CSB}}. \tag{5.29}$$

Furthermore, note that $\mathcal{R}_{\mathrm{DF}}$ and $\mathcal{R}_{\mathrm{PDF}}$ of course include unidirectional DF and PDF transmission, respectively, from terminal A to terminal B ($t_2 = t_3 = t_4 = t_5 = 0$), from terminal B to terminal A ($t_1 = t_3 = t_4 = t_6 = 0$), and time-sharing between unidirectional transmission in both directions ($t_3 = t_4 = 0$) as special cases. Similarly, the CSOB region $\mathcal{C}_{\mathrm{CSB}}$ reduces to the CSB if only unidirectional transmission is considered.

[2] If the relay does not decode the entire information transmitted by one of the terminals in phase 3, this information never arrives at the respective other terminal. This is because neither terminal listens in phase 3 and since the relay cannot forward any information it has not previously decoded if it uses the PDF strategy. Without loss of generality, we hence assume that the relay decodes the entire information transmitted by both terminals in phase 3.

To conclude this section, we remark that the definitions of and the capacity results for the degraded, the reversely degraded, and the semideterministic half-duplex relay channels as well as the half-duplex relay channel with orthogonal sender components can easily be extended to the restricted half-duplex two-way relay channel. In particular, it can be shown that PDF achieves the capacity regions of these types of half-duplex constrained two-way relay channels.

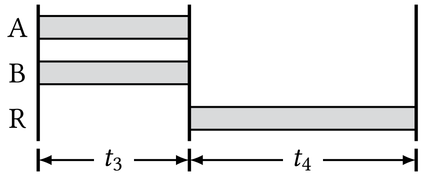
5.6 Further Results and Bibliographical Notes

The *cut-set bound* (CSB, cf. Section 5.2) on the capacity of the half-duplex relay channel with *time-division duplex* (TDD) was established by Khojastepour et al. [66] as a special case of the max-flow min-cut theorem for general multiterminal networks with a finite number of states and the state being known to all nodes (Theorem 5.2). In [68], they also derived the maximum achievable *decode-and-forward* (DF) rate (cf. Section 5.3) for the half-duplex ("cheap") relay channel, which they used to establish the capacity of the degraded half-duplex relay channel. The corresponding DF coding scheme was later generalized by Stein [118], who obtained the expression for the maximum achievable *partial decode-and-forward* (PDF) rate (cf. Section 5.4).

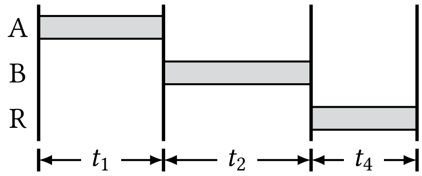
Achievable *compress-and-forward* (CF) rates for the half-duplex relay channel with TDD were independently derived by Stein [118] and Yao et al. [144]. In addition, Cover and El Gamal's mixed strategy [21, Theorem 7] was also generalized to the half-duplex relay channel in [118]. Like for DF and PDF, the corresponding half-duplex coding schemes operate on one block of n channel uses where the relay is in receive mode for the first n_1 channel uses and in transmit mode for the remaining n_2 channel uses.

Throughout this work, we assume that all three nodes noncausally know whether the relay is in receive or in transmit mode, which is a reasonable assumption for communication protocols where a TDD schedule is defined and communicated to all nodes a priori. In [75], Kramer considered a random sleep-listen-or-talk (SLoT) strategy for the relay, where the destination does not know the relay mode in advance. In particular, he showed that strategies with deterministic schedules are generally suboptimal as the relay may also forward information to the destination through its choice of operating mode if it uses a random schedule. However, the potential rate gain is limited to 1 bpcu since $|\mathcal{M}| = 2$, and there are several challenges for implementing random schedules as discussed in [78, Section 4.3].

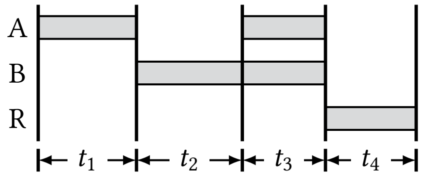
Instead of separating reception and transmission at the relay in time, it is of course also possible to consider *frequency-division duplex* (FDD) protocols, where the relay uses different frequency bands for transmission and reception. If the channel from the source to the relay is assigned a separate frequency band, which is referred to as *sender frequency-division* in [34, Chapter 16], the relay channel has orthogonal sender components. Its capacity is thus achieved by the PDF strategy, cf. Section 2.4. On the other hand, the capacity of the relay channel with *receiver frequency-division*, where the



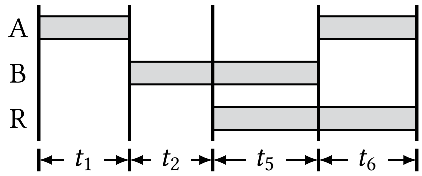
(a) Multiple-access broadcast (MABC) protocol



(b) Time-division broadcast (TDBC) protocol



(c) Hybrid broadcast (HBC) protocol



(d) One-way time-sharing (OWTS) protocol

Figure 5.6: Timing Diagrams of MABC, TDBC, HBC, and OWTS Protocols: Gray areas denote transmissions by the respective nodes, where it is assumed that the nodes listen when they are not transmitting.

relay-to-destination channel is assigned a separate frequency band, remains unknown. For this relay channel with *orthogonal receiver components* as well as for the half-duplex relay channel with a more general FDD protocol in which the source may utilize the entire spectrum for transmission, upper bounds on the capacity and achievable rates can be derived in the same way as for the half-duplex relay channel with TDD.

The (restricted) half-duplex two-way relay channel model was introduced by Rankov and Wittneben [101, 102], who noted that a significant portion of the loss in spectral efficiency suffered due to the half-duplex constraint can be compensated when bidirectional communication is considered. Various achievable rate regions for the half-duplex two-way relay channel have since been derived, most of them based on a two-phase TDD communication protocol in which the terminals first transmit their messages to the relay in a multiple-access phase before the relay then broadcasts its signal to both terminals, cf. [70–72, 96, 100, 109, 110, 141] and references therein, for example. This protocol, which is illustrated in Figure 5.6(a), is commonly referred to as *multiple-access broadcast* (MABC) protocol.

However, note that all information is sent via the relay in the MABC protocol since the terminals cannot overhear each other's transmissions during the multiple-access phase due to the half-duplex constraint. As a result, this protocol is only optimal if the direct link between the two terminals can be neglected, e.g., because of shadowing effects. Otherwise, communication protocols that also utilize the direct link yield larger achievable rate regions in general, cf. [70–72], for example. Protocols of this kind having been considered in the literature include the *time-division broadcast* (TDBC) protocol (cf. Figure 5.6(b)), the *hybrid broadcast* (HBC) protocol (cf. Figure 5.6(c)), as well as the *one-way time-sharing* (OWTS) protocol (cf. Figure 5.6(d)).[3]

[3]The protocol names are due to [72] (MABC, TDBC, HBC) and [41] (OWTS).

We remark that the six-phase protocol that was used to derive the achievable rate regions $\mathcal{R}_{\text{DF}}$ and $\mathcal{R}_{\text{PDF}}$ (cf. Theorems 5.6 and 5.7) includes all four of these protocols as special cases. In particular, the MABC protocol (consisting of phases 3 and 4), the TDBC protocol (phases 1, 2, 4), the HBC protocol (phases 1, 2, 3, 4), and the OWTS protocol (phases 1, 2, 5, 6) can all be obtained from the considered six-phase protocol by setting some of the time-shares $t_1, \ldots, t_6$ to zero. Consequently, any rate region that is achievable with the MABC/TDBC/HBC/OWTS protocol and the DF or PDF strategy is a subset of $\mathcal{R}_{\text{DF}}$ or $\mathcal{R}_{\text{PDF}}$, respectively. Furthermore, an outer bound on the largest rate region that can be achieved with some specific protocol can be obtained by setting the appropriate time-shares in the CSOB region (cf. Theorem 5.5) to zero.

Chapter 6

Half-Duplex Gaussian MIMO Relay Channel

After having discussed various information theoretical results for the discrete memoryless half-duplex relay channel in the previous chapter, we now apply these results to the half-duplex constrained Gaussian *multiple-input multiple-output* (MIMO) relay channel. To this end, we first have to ensure that the achievable rates and the capacity upper bounds that were derived for a channel model with finite input and output alphabets are also valid for continuous-alphabet channels. However, since the same arguments as for the full-duplex case can be used to argue why this is the case, we do not elaborate on the details. In addition, note that this chapter focuses on unidirectional communication in the half-duplex constrained Gaussian MIMO relay channel, whereas bidirectional communication is subsequently addressed in Chapter 7. The structure of this chapter is therefore identical to that of Chapter 3.

The system model for the half-duplex Gaussian MIMO relay channel is introduced in Section 6.1, where we remark that two different power constraints are considered, a *per-phase* power constraint and an *average* power constraint, cf. Sections 6.1.1 and 6.1.2. Using the entropy maximizing property of the Gaussian distribution, we then establish in Sections 6.2 and 6.3, respectively, that the *cut-set bound* (CSB) C_{CSB} and the maximum achievable *decode-and-forward* (DF) rate R_{DF} are attained by jointly Gaussian source and relay inputs for both power constraints. Furthermore, we show that if the channel gain matrices are known perfectly and instantaneously at all nodes, C_{CSB} and R_{DF} can be determined as the solutions of convex optimization problems. More specifically, we derive dual decomposition approaches that allow to efficiently solve the corresponding rate maximization problems with respect to the channel inputs and the time-shares of the relay receive and transmit phases in the Lagrangian dual domain.

Rates that can be achieved using the *partial decode-and-forward* (PDF) strategy are subsequently discussed in Section 6.4, where we first show that the maximum achievable PDF rate R_{PDF} is also attained by Gaussian channel inputs. The derivation proceeds along the same lines as for the full-duplex case: In Section 6.4.1, we use a channel enhancement argument to prove that jointly Gaussian source and relay inputs maximize the achievable PDF rate for the *aligned* half-duplex Gaussian MIMO relay channel, and in Section 6.4.2, we extend this result to the general half-duplex Gaussian MIMO relay channel. However, because an algorithm to (efficiently) evaluate R_{PDF} for the general case has yet to be derived, Section 6.4.3 again considers the two approaches that were introduced in Section 3.4.3 to obtain suboptimal solutions to the PDF rate maximization problem. Using either of these approaches, suboptimal PDF rates can be determined by means of convex optimization techniques. In Section 6.4.4, we show that convex optimization techniques can also be used to evaluate R_{PDF} for half-duplex Gaussian relay channels of *stochastically degraded nature*, and the chapter concludes in Section 6.5 with an overview of further noteworthy results and bibliographical notes on the half-duplex Gaussian (MIMO) relay channel.

6.1 System Model

Note that like for the full-duplex case, our half-duplex Gaussian MIMO relay channel model, which is illustrated in Figure 6.1, is obtained by applying the *linear MIMO model* to the considered relay scenario. More precisely, let $\boldsymbol{H}_{\text{SR}} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{S}}}$, $\boldsymbol{H}_{\text{SD}} \in \mathbb{C}^{N_{\text{D}} \times N_{\text{S}}}$, and $\boldsymbol{H}_{\text{RD}} \in \mathbb{C}^{N_{\text{D}} \times N_{\text{R}}}$ again denote the channel gain matrices, which are assumed to be perfectly and instantaneously known at all nodes. Then, if $\boldsymbol{x}_{\text{A}m}$ and $\boldsymbol{y}_{\text{B}m}$ represent the transmit signal of node A and the receive signal of node B during phase m, respectively, the receive signals of the relay and the destination during the relay receive phase (1) and the relay transmit phase (2) can be expressed as follows:

(1) The source transmits to the relay and the destination:

$$\begin{aligned} \boldsymbol{y}_{\text{R1}} &= \boldsymbol{H}_{\text{SR}}\boldsymbol{x}_{\text{S1}} + \boldsymbol{n}_{\text{R1}}, \qquad && \boldsymbol{n}_{\text{R1}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{R}}), \\ \boldsymbol{y}_{\text{D1}} &= \boldsymbol{H}_{\text{SD}}\boldsymbol{x}_{\text{S1}} + \boldsymbol{n}_{\text{D1}}, \qquad && \boldsymbol{n}_{\text{D1}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{D}}). \end{aligned} \tag{6.1}$$

(2) The source and the relay transmit to the destination:

$$\boldsymbol{y}_{\text{D2}} = \boldsymbol{H}_{\text{SD}}\boldsymbol{x}_{\text{S2}} + \boldsymbol{H}_{\text{RD}}\boldsymbol{x}_{\text{R2}} + \boldsymbol{n}_{\text{D2}}, \qquad \boldsymbol{n}_{\text{D2}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{D}}). \tag{6.2}$$

Here, the channel from the source to the destination is assumed to be the same for both phases. For the results presented in this chapter, this is without loss of generality as we anyhow require all channel gain matrices to be perfectly known at all nodes. Moreover, $\boldsymbol{n}_{\text{R1}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{R}})$, $\boldsymbol{n}_{\text{D1}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{D}})$, and $\boldsymbol{n}_{\text{D2}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{D}})$ denote zero-mean proper complex Gaussian noise vectors with nonsingular covariance matrices $\boldsymbol{Z}_{\text{R}} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{R}}}$ and $\boldsymbol{Z}_{\text{D}} \in \mathbb{C}^{N_{\text{D}} \times N_{\text{D}}}$. These noise vectors are assumed to be independent of each other as well as of the transmit signals. In particular, the additive Gaussian noise vectors $\boldsymbol{n}_{\text{D1}}$ and $\boldsymbol{n}_{\text{D2}}$, which are received at the destination during the relay receive and transmit phases, respectively, are assumed to be independent. Finally, we assume perfectly synchronized transmission and reception between all nodes.

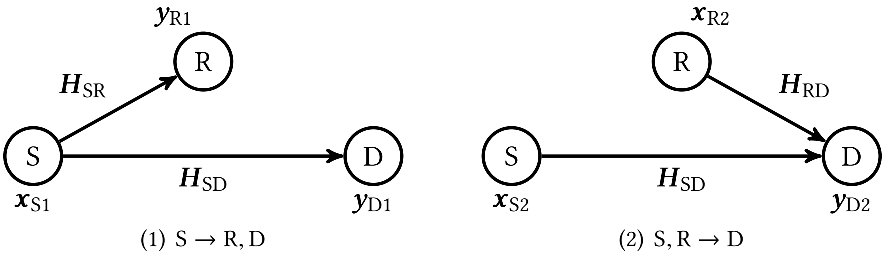


Figure 6.1: Illustration of the Half-Duplex Gaussian MIMO Relay Channel

Following the same reasoning as in Section 3.1, it is clear that the capacity of the half-duplex Gaussian MIMO relay channel is infinite without further restrictions on $\boldsymbol{x}_{S1}$, $\boldsymbol{x}_{S2}$, and $\boldsymbol{x}_{R2}$. Like for the full-duplex case, we therefore impose transmit power constraints on the channel inputs of the source and the relay. However, for the half-duplex case, we actually consider two different power constraints.

6.1.1 Per-Phase Power Constraint

First, suppose that if node A transmits during phase $m \in \mathcal{M}$, then it is subject to the power constraint

$$\mathrm{E}[\boldsymbol{x}_{Am}^{H}\boldsymbol{x}_{Am}] \leq P_{Am}, \tag{6.3}$$

where $P_{Am} > 0$ represents the power budget that is available to node A in phase m. This constraint is referred to as *per-phase* power constraint in the following. We remark that with this constraint, the power node A may consume during phase m is independent of the duration of this phase, i.e., independent of t_m.

Note that without loss of generality, we can again assume that all channel inputs are zero-mean because channel inputs with nonzero mean consume more transmit power to convey the same amount of information, cf. Section 3.1. For the half-duplex Gaussian MIMO relay channel, the per-phase power constraint hence requires that

$$\mathrm{tr}(\boldsymbol{C}_{S1}) \leq P_S, \qquad \mathrm{tr}(\boldsymbol{C}_{S2}) \leq P_S, \qquad \mathrm{tr}(\boldsymbol{C}_{R2}) \leq P_R, \tag{6.4}$$

where $\boldsymbol{C}_{S1}$, $\boldsymbol{C}_{S2}$, and $\boldsymbol{C}_{R2}$ denote the covariance matrices of $\boldsymbol{x}_{S1}$, $\boldsymbol{x}_{S2}$, and $\boldsymbol{x}_{R2}$, respectively, and where we have assumed that $P_{S1} = P_{S2} = P_S$ and $P_{R2} = P_R$ for simplicity.

6.1.2 Average Power Constraint

The second power constraint we consider is an *average* power constraint, which requires the channel inputs of any node A that is part of the network to satisfy

$$\sum_{m \in \mathcal{M}} t_m \, \mathrm{E}[\boldsymbol{x}_{Am}^{H}\boldsymbol{x}_{Am}] \leq P_A. \tag{6.5}$$

Here, $P_\mathrm{A} > 0$ is the total power budget that is available to node A and $\mathcal{M}$ represents the set of all states/modes of the network. In contrast to the per-phase power constraint, the powers node A may consume during the different phases strongly depend on the durations of the phases, i.e., on the time shares $t_m, m \in \mathcal{M}$.

For the half-duplex Gaussian MIMO relay channel, the average power constraint requires the channel inputs of the source and the relay to satisfy

$$t_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}1}) + t_2 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}2}) \leq P_\mathrm{S}, \qquad t_2 \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}2}) \leq P_\mathrm{R}. \tag{6.6}$$

We remark that if the values of P_S and P_R are the same as for the per-phase power constraint in (6.4), the average power constraint is less restrictive. That is, any set of covariance matrices $(\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_{\mathrm{S}2}, \boldsymbol{C}_{\mathrm{R}2})$ that satisfies (6.4) also satisfies (6.6) because the time-shares $t_1, t_2 \geq 0$ are normalized such that $t_1 + t_2 = 1$.[1]

6.2 Cut-Set Bound

The derivation of the cut-set bound (CSB) for the half-duplex relay channel applies to arbitrary input and output alphabets so that Theorem 5.1 also holds for continuous-alphabet channels. Consequently, the capacity of the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint is upper bounded by

$$\begin{aligned} C_\mathrm{CSB} = \max_{\substack{R, t_1, t_2 \\ p(\boldsymbol{x}_{\mathrm{S}1}) p(\boldsymbol{x}_{\mathrm{S}2}, \boldsymbol{x}_{\mathrm{R}2})}} R \quad \text{s.t.} \quad & R \leq t_1 I(\boldsymbol{x}_{\mathrm{S}1}; \boldsymbol{y}_{\mathrm{R}1}, \boldsymbol{y}_{\mathrm{D}1}) + t_2 I(\boldsymbol{x}_{\mathrm{S}2}; \boldsymbol{y}_{\mathrm{D}2} | \boldsymbol{x}_{\mathrm{R}2}), \\ & R \leq t_1 I(\boldsymbol{x}_{\mathrm{S}1}; \boldsymbol{y}_{\mathrm{D}1}) + t_2 I(\boldsymbol{x}_{\mathrm{S}2}, \boldsymbol{x}_{\mathrm{R}2}; \boldsymbol{y}_{\mathrm{D}2}), \\ & t_1, t_2 \geq 0, \ t_1 + t_2 = 1, \\ & (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_{\mathrm{S}2}, \boldsymbol{C}_{\mathrm{R}2}) \in \mathcal{P}, \end{aligned} \tag{6.7}$$

where the set $\mathcal{P}$ is given by

$$\mathcal{P} = \left\{ (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_{\mathrm{S}2}, \boldsymbol{C}_{\mathrm{R}2}) : \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}1}) \leq P_\mathrm{S}, \ \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}2}) \leq P_\mathrm{S}, \ \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}2}) \leq P_\mathrm{R} \right\} \tag{6.8}$$

for the per-phase power constraint and by

$$\mathcal{P} = \left\{ (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_{\mathrm{S}2}, \boldsymbol{C}_{\mathrm{R}2}) : t_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}1}) + t_2 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S}2}) \leq P_\mathrm{S}, \ t_2 \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}2}) \leq P_\mathrm{R} \right\} \tag{6.9}$$

for the average power constraint. Now, the first step to evaluate this bound for either power constraint is to determine the probability distribution that attains C_CSB. To this end, the entropy maximizing property of the Gaussian distribution can be used to show that like for the full-duplex case, C_CSB is attained by proper complex Gaussian channel inputs for both the per-phase and the average power constraint.

[1] Adding the average power constraints to the single-letter mutual information expressions that were obtained for the discrete memoryless half-duplex relay channel yields the same values for the CSB and the maximum achievable DF rate as imposing standard power constraints of the form $\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x}_i^\mathrm{H} \boldsymbol{x}_i \leq P$ on the $(2^{nR}, n)$ code for the half-duplex Gaussian MIMO relay channel, cf. Appendix B.1.2.

Theorem 6.1. *For the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint, the cut-set bound (CSB) is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. Suppose t_1 and t_2 are fixed, and let $\boldsymbol{C}_2$ denote the joint covariance matrix of $\boldsymbol{x}_{\mathrm{S}2}$ and $\boldsymbol{x}_{\mathrm{R}2}$. Then, both transmit power constraints only depend on $\boldsymbol{C}_{\mathrm{S}1}$ and $\boldsymbol{C}_2$ as $\boldsymbol{C}_{\mathrm{S}2}$ and $\boldsymbol{C}_{\mathrm{R}2}$ are determined by $\boldsymbol{C}_2$. For any t_1, t_2 and $\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_2$, however, the Gaussian distribution simultaneously maximizes all mutual information terms in (6.7) since it maximizes both differential entropy [24, Theorem 8.6.5] and conditional differential entropy [123]. More precisely, the (conditional) differential entropy of a complex random vector is maximized by the zero-mean proper complex Gaussian distribution [94, 122]. □

The optimal distributions of the channel inputs for the relay receive and transmit phases can hence be represented by the covariance matrix of $\boldsymbol{x}_{\mathrm{S}1}$ and the joint covariance matrix of $\boldsymbol{x}_{\mathrm{S}2}$ and $\boldsymbol{x}_{\mathrm{R}2}$. In particular, let $\boldsymbol{x}_{\mathrm{S}1} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_{\mathrm{S}1})$ and $\left[\begin{smallmatrix}\boldsymbol{x}_{\mathrm{S}2}\\ \boldsymbol{x}_{\mathrm{R}2}\end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_2)$, then the mutual information terms that define the CSB simplify to the log-det expressions

$$\begin{aligned}
I(\boldsymbol{x}_{\mathrm{S}1}; \boldsymbol{y}_{\mathrm{R}1}, \boldsymbol{y}_{\mathrm{D}1}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S}\{\mathrm{RD}\}}\boldsymbol{C}_{\mathrm{S}1}\boldsymbol{H}_{\mathrm{S}\{\mathrm{RD}\}}^{\mathrm{H}}\right), \\
I(\boldsymbol{x}_{\mathrm{S}1}; \boldsymbol{y}_{\mathrm{D}1}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S}1}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
I(\boldsymbol{x}_{\mathrm{S}2}; \boldsymbol{y}_{\mathrm{D}2} | \boldsymbol{x}_{\mathrm{R}2}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S}|\mathrm{R}2}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
I(\boldsymbol{x}_{\mathrm{S}2}, \boldsymbol{x}_{\mathrm{R}2}; \boldsymbol{y}_{\mathrm{D}2}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right).
\end{aligned} \tag{6.10}$$

Here, we have assumed that the additive Gaussian noise is white, $\boldsymbol{H}_{\mathrm{S}\{\mathrm{RD}\}}$ denotes the channel gain matrix of the composite channel from the source to the relay and the destination, cf. (3.10), and $\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}$ represents the channel gain matrix of the composite channel from the source and the relay to the destination, cf. (3.11). Moreover, $\boldsymbol{C}_{\mathrm{S}|\mathrm{R}2}$ is the conditional covariance matrix of $\boldsymbol{x}_{\mathrm{S}2}$ given $\boldsymbol{x}_{\mathrm{R}2}$, which is equal to the Schur complement of $\boldsymbol{C}_{\mathrm{R}2}$ in $\boldsymbol{C}_2$. In order to evaluate the CSB for the half-duplex Gaussian MIMO relay channel, we thus have to solve the optimization problem

$$\begin{aligned}
C_{\mathrm{CSB}} = \max_{R, t_1, t_2, \boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_2} \; & R \\
\text{s.t.} \quad & R \le t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S}\{\mathrm{RD}\}}\boldsymbol{C}_{\mathrm{S}1}\boldsymbol{H}_{\mathrm{S}\{\mathrm{RD}\}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S}|\mathrm{R}2}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
& R \le t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S}1}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\
& t_1, t_2 \ge 0, \; t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_2 \succcurlyeq \boldsymbol{0}, \quad (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P},
\end{aligned} \tag{6.11}$$

where $\mathcal{P}$ is given by (6.8) or (6.9) depending on the considered power constraint, and where $\boldsymbol{D}_{\mathrm{S}}$ and $\boldsymbol{D}_{\mathrm{R}}$ are the selection matrices that were introduced in (3.6).

In the following, we show that C_{CSB} can be determined as the solution of a convex optimization problem for both the per-phase and the average power constraint. More specifically, we derive dual decomposition approaches that allow to efficiently solve the corresponding optimization problems with respect to the covariance matrices and the time-shares of the relay receive and transmit phases in the Lagrangian dual domain. We remark that the average power constraint is considerably more difficult to deal with, both from a theoretical and a practical point of view. This is because we need to introduce more dual variables as for the per-phase power constraint and because the constraint sets of the subproblems encountered in the dual domain are unbounded only for the average power constraint. The latter means that several additional mathematical details have to be taken into account in order to ensure the correctness of the optimization strategy. On the other hand, the per-phase power constraint can easily be incorporated into the framework for the average power constraint. Since the converse is not true, the dual decomposition approach for the average power constraint is more general.

6.2.1 Per-Phase Power Constraint

First, suppose the source and the relay are subject to the per-phase power constraint. In this case, the CSB for the half-duplex Gaussian MIMO relay channel is equal to

$$\begin{aligned} C_{\mathrm{CSB}} = \max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{S1}},\boldsymbol{C}_2} \; & R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S|R2}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_2\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\; t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2 \succcurlyeq \mathbf{0}, \\ & \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.12}$$

Note that for fixed time-shares, it can be shown that the optimization problem given in (6.12) is convex in the remaining optimization variables. However, similar to the full-duplex case, it is advantageous to reformulate the first rate bound of this optimization problem by introducing an auxiliary variable $\boldsymbol{C}_{\mathrm{Q}} = \boldsymbol{C}_{\mathrm{S|R2}} \succcurlyeq \mathbf{0}$, relaxing the equality constraint to $\mathbf{0} \preccurlyeq \boldsymbol{C}_{\mathrm{Q}} \preccurlyeq \boldsymbol{C}_{\mathrm{S|R2}}$, and applying Lemma 3.2, which eventually yields

$$\begin{aligned} C_{\mathrm{CSB}} = \max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{S1}},\boldsymbol{C}_2,\boldsymbol{C}_{\mathrm{Q}}} \; & R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_2\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\; t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.13}$$

Since $\log\det(\mathbf{I} + \boldsymbol{H}\boldsymbol{C}\boldsymbol{H}^{\mathrm{H}})$ is concave in $\boldsymbol{C} \succcurlyeq \mathbf{0}$, cf. Section 3.2, the nonnegative weighted sum of concave functions is concave [11, Section 3.2.1], and the remaining constraints are linear in $\boldsymbol{C}_{\mathrm{S1}}$, $\boldsymbol{C}_2$, $\boldsymbol{C}_{\mathrm{Q}}$, this maximization problem is convex in $(R, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2, \boldsymbol{C}_{\mathrm{Q}})$ for given $t_1, t_2 \geq 0$. However, it is not jointly convex in all optimization variables, so further reformulation steps are needed to obtain a convex optimization problem.

One way to convexify the maximization problem given in (6.13) is to apply a change of variables. In particular, let

$$\boldsymbol{K}_{\mathrm{S1}} = t_1 \boldsymbol{C}_{\mathrm{S1}}, \qquad \boldsymbol{K}_2 = t_2 \boldsymbol{C}_2, \qquad \boldsymbol{K}_{\mathrm{Q}} = t_2 \boldsymbol{C}_{\mathrm{Q}}, \tag{6.14}$$

then the optimization problem that determines C_{CSB} can equivalently be expressed as

$$\begin{aligned} C_{\mathrm{CSB}} = & \max_{R, t_1, t_2, \boldsymbol{K}_{\mathrm{S1}}, \boldsymbol{K}_2, \boldsymbol{K}_{\mathrm{Q}}} R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}(\boldsymbol{K}_{\mathrm{S1}}/t_1)\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{K}_{\mathrm{Q}}/t_2)\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{K}_{\mathrm{S1}}/t_1)\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}(\boldsymbol{K}_2/t_2)\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0, \; t_1 + t_2 = 1, \qquad \boldsymbol{K}_{\mathrm{S1}}, \boldsymbol{K}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{K}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{K}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & \mathrm{tr}(\boldsymbol{K}_{\mathrm{S1}}) \leq t_1 P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{K}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq t_2 P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{K}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq t_2 P_{\mathrm{R}}. \end{aligned} \tag{6.15}$$

Note that $g : (\boldsymbol{C}, t) \mapsto t \log\det(\mathbf{I} + \boldsymbol{H}(\boldsymbol{C}/t)\boldsymbol{H}^{\mathrm{H}})$ with $\mathrm{dom}(g) = \{(\boldsymbol{C}, t) : \boldsymbol{C} \succcurlyeq \mathbf{0}, t > 0\}$ is the *perspective* of the function $f : \boldsymbol{C} \mapsto \log\det(\mathbf{I} + \boldsymbol{H}\boldsymbol{C}\boldsymbol{H}^{\mathrm{H}})$, which is concave on $\mathrm{dom}(f) = \{\boldsymbol{C} : \boldsymbol{C} \succcurlyeq \mathbf{0}\}$. It hence follows from the convexity preserving property of the perspective operation [11, Section 3.2.6] that g is jointly concave in $\boldsymbol{C}$ and t on its entire domain, which implies that the two rate constraints define a convex set. In addition, all other constraints are linear functions of the optimization variables so that we have the following result:

Proposition 6.2. *For the half-duplex Gaussian MIMO relay channel with the per-phase power constraint, C_{CSB} can be determined as the solution of the convex optimization problem given in* (6.15).

Remark 6.1. The perspective function is defined only for positive t. However, if either $t_1 = 0, t_2 = 1$ or $t_1 = 1, t_2 = 0$ were optimal in (6.13), the CSB would be equal to the capacity of the source-to-destination link, i.e.,

$$C_{\mathrm{CSB}} = R_{\mathrm{P2P}} = \max_{\boldsymbol{C}_{\mathrm{S}}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) \quad \text{s.t.} \quad \boldsymbol{C}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{S}}) \leq P_{\mathrm{S}}. \tag{6.16}$$

But unless the relay cannot help the source convey information to the destination at all, which happens if and only if $\boldsymbol{H}_{\mathrm{SR}} = \mathbf{0}$ and/or $\boldsymbol{H}_{\mathrm{RD}} = \mathbf{0}$, we must have $C_{\mathrm{CSB}} > R_{\mathrm{P2P}}$. For all relevant scenarios, we therefore only need to apply the change of variables and the perspective operation for positive t_1, t_2. Furthermore, continuity arguments can be used to show that the first and second terms of both rate bounds in (6.15) vanish if $t_1 = 0, t_2 = 1$ and $t_1 = 1, t_2 = 0$, respectively, so that the constraints need not explicitly exclude these cases.

While the optimization problem given in (6.15) is convex, we remark that it does not satisfy the ruleset of *disciplined convex programming* (DCP) [54]. This is important as problems adhering to the DCP ruleset can automatically be converted to solvable form, which allows to directly apply standard semidefinite program (SDP) solvers to such problems. Unfortunately, the same does not hold for optimization problems that violate the DCP ruleset, even if they are convex. Consequently, other solution methods like subgradient algorithms need to be considered in order to determine C_{CSB} from the rate maximization problem given in (6.15).

Instead of developing such an algorithm, we propose an alternative approach to efficiently evaluate C_{CSB} based on dual decomposition. To this end, let us first define the two rate regions

$$\begin{aligned}\mathcal{S}_{\mathrm{CSB1}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ \boldsymbol{C}_{\mathrm{S1}} &\succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{S1}}) \leq P_{\mathrm{S}} \Big\}\end{aligned} \tag{6.17}$$

and

$$\begin{aligned}\mathcal{S}_{\mathrm{CSB2}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_{2}\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ \boldsymbol{C}_{\mathrm{Q}} &\succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) &\leq P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}} \Big\},\end{aligned} \tag{6.18}$$

which are associated with the contributions of the relay receive and transmit phases to the CSB, respectively. Note that it is straightforward to show that both $\mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{CSB2}}$ are compact (i.e., closed and bounded since $\mathcal{S}_{\mathrm{CSB1}}, \mathcal{S}_{\mathrm{CSB2}} \subseteq \mathbb{R}^2$) and convex. By means of these two rate regions, the maximization problem given in (6.13) can now be rewritten as

$$\begin{aligned}C_{\mathrm{CSB}} = \max_{R,t_1,t_2,\boldsymbol{s}_1,\boldsymbol{s}_2} R \quad \text{s.t.} \quad & R\mathbf{1} \leq t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \\ & \boldsymbol{s}_1 \in \mathcal{S}_{\mathrm{CSB1}},\ \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{CSB2}}, \quad t_1, t_2 \geq 0,\ t_1 + t_2 = 1.\end{aligned} \tag{6.19}$$

Next, let us define

$$\begin{aligned}\mathcal{S}_{\mathrm{CSB}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : \boldsymbol{s} &= t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \\ &\boldsymbol{s}_1 \in \mathcal{S}_{\mathrm{CSB1}},\ \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{CSB2}}, \quad t_1, t_2 \geq 0,\ t_1 + t_2 = 1 \Big\},\end{aligned} \tag{6.20}$$

which is a compact convex set as well since the convex sum of convex sets is convex, cf. [11, Section 2.3.2]. Using this definition, we can rewrite (6.19) once more to obtain

$$C_{\mathrm{CSB}} = \max_{R,\boldsymbol{s}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}. \tag{6.21}$$

Because $\mathcal{S}_{\mathrm{CSB}}$ is a convex set with nonempty interior, this is a convex optimization problem for which *strong duality* holds [11, Section 5.3.2], which means that we can equivalently solve the problem in the Lagrangian dual domain.

Derivation of the Dual Problem—In the approach considered here, the constraint $R\mathbf{1} \leq \boldsymbol{s}$ is incorporated into the objective function using the Lagrangian multiplier $\boldsymbol{\lambda}$. This leads to a dual problem in which the relay receive and transmit phases decouple – hence the term "dual decomposition". The corresponding Lagrangian function reads as

$$L(R, \boldsymbol{s}, \boldsymbol{\lambda}) = R - \boldsymbol{\lambda}^{\mathrm{T}} (R\mathbf{1} - \boldsymbol{s}), \tag{6.22}$$

and the resulting dual function is given by

$$\Theta(\boldsymbol{\lambda}) = \sup_{R, \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}} L(R, \boldsymbol{s}, \boldsymbol{\lambda}) = \begin{cases} \max_{\boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s} & \text{if } \mathbf{1}^{\mathrm{T}} \boldsymbol{\lambda} = 1, \\ +\infty & \text{otherwise} \end{cases} \tag{6.23}$$

as $\mathcal{S}_{\mathrm{CSB}}$ is compact. In order to determine C_{CSB}, we therefore have to solve the dual problem

$$C_{\mathrm{CSB}} = \min_{\boldsymbol{\lambda}} \max_{\boldsymbol{s}} \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s} \quad \text{s.t.} \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}, \quad \boldsymbol{\lambda} \geq \mathbf{0},\ \mathbf{1}^{\mathrm{T}} \boldsymbol{\lambda} = 1. \tag{6.24}$$

In the following, we explain how this can be done by means of the *cutting-plane* method, an outer-approximation method in which the dual function $\Theta(\boldsymbol{\lambda})$ is approximated and iteratively refined by a set of linear inequalities.

Solution by means of the Cutting-Plane Algorithm—Note that an equivalent formulation of the dual problem is given by

$$C_{\mathrm{CSB}} = \min_{z, \boldsymbol{\lambda}} z \quad \text{s.t.} \quad z \geq \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s}, \forall \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}, \quad \boldsymbol{\lambda} \geq \mathbf{0},\ \mathbf{1}^{\mathrm{T}} \boldsymbol{\lambda} = 1. \tag{6.25}$$

If the cutting-plane algorithm is applied to this problem, the set $\mathcal{S}_{\mathrm{CSB}}$ is approximated by a finite set $\{\boldsymbol{s}^{(0)}, \dots, \boldsymbol{s}^{(k-1)}\}$ of feasible rate vectors, i.e., $\boldsymbol{s}^{(\ell)} \in \mathcal{S}_{\mathrm{CSB}}, \forall \ell \in \{0, \dots, k-1\}$. More precisely, in the k-th iteration of the cutting-plane algorithm, the approximated dual problem

$$\min_{z, \boldsymbol{\lambda}} z \quad \text{s.t.} \quad z \geq \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s}^{(\ell)}, \forall \ell \in \{0, \dots, k-1\}, \quad \boldsymbol{\lambda} \geq \mathbf{0},\ \mathbf{1}^{\mathrm{T}} \boldsymbol{\lambda} = 1 \tag{6.26}$$

is solved instead of the original dual problem. This approximated dual problem is a linear program, and it is called the *master program* of the cutting-plane algorithm.

Let $z^{(k)}$ and $\boldsymbol{\lambda}^{(k)}$ denote the optimizers of the master program in the k-th iteration. In order to check whether $(z^{(k)}, \boldsymbol{\lambda}^{(k)})$ is an optimal solution to the original dual problem given in (6.25), we need to solve the *Lagrangian subproblem*

$$\begin{aligned} \Theta(\boldsymbol{\lambda}^{(k)}) &= \max_{\boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{(k),\mathrm{T}} \boldsymbol{s} = \max_{\substack{\boldsymbol{s}_1 \in \mathcal{S}_{\mathrm{CSB}1}, \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{CSB}2} \\ t_1, t_2 \geq 0,\ t_1 + t_2 = 1}} t_1 \boldsymbol{\lambda}^{(k),\mathrm{T}} \boldsymbol{s}_1 + t_2 \boldsymbol{\lambda}^{(k),\mathrm{T}} \boldsymbol{s}_2 \\ &= \max_{m=1,2} \max_{\boldsymbol{s}_m \in \mathcal{S}_{\mathrm{CSB}m}} \boldsymbol{\lambda}^{(k),\mathrm{T}} \boldsymbol{s}_m, \end{aligned} \tag{6.27}$$

i.e., we need to evaluate the dual function at $\boldsymbol{\lambda}^{(k)}$. As can be seen from (6.27), the Lagrangian subproblem decomposes into two independent subproblems. In particular, the evaluation of $\Theta(\boldsymbol{\lambda}^{(k)})$ requires to solve two *weighted sum rate* (WSR) maximization problems, one over $\mathcal{S}_{\mathrm{CSB1}}$ and one over $\mathcal{S}_{\mathrm{CSB2}}$. We remark that for solving these problems, standard SDP solvers like SDPT3 [126, 127] or LogdetPPA [133] which are capable of dealing with log-det terms in the objective function can be used.

Note that if $z^{(k)} \geq \Theta(\boldsymbol{\lambda}^{(k)})$, it follows that $(z^{(k)}, \boldsymbol{\lambda}^{(k)})$ is an optimal solution to the original dual problem given in (6.25). If $z^{(k)} < \Theta(\boldsymbol{\lambda}^{(k)})$, on the other hand, $(z^{(k)}, \boldsymbol{\lambda}^{(k)})$ is no valid solution to the original dual problem since the constraint $z \geq \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}$ is violated for $\boldsymbol{s}^{(k)} = \arg\max_{\boldsymbol{s}\in\mathcal{S}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}$, where

$$\boldsymbol{s}^{(k)} = \begin{cases} \boldsymbol{s}_1^{(k)} & \text{if } \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}_1^{(k)} \geq \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}_2^{(k)}, \\ \boldsymbol{s}_2^{(k)} & \text{if } \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}_1^{(k)} < \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}_2^{(k)} \end{cases} \tag{6.28}$$

is the maximizer of the Lagrangian subproblem in iteration k and

$$\boldsymbol{s}_m^{(k)} = \underset{\boldsymbol{s}_m\in\mathcal{S}_{\mathrm{CSB}m}}{\arg\max}\ \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}_m. \tag{6.29}$$

In that case, the approximation of the dual problem is refined by imposing the additional constraint $z \geq \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(k)}$ on the master program of the next iteration, i.e., by adding $\boldsymbol{s}^{(k)}$ to the set of rate vectors that define the approximation of the constraint set $\mathcal{S}_{\mathrm{CSB}}$.

In each iteration, one more constraint is thus added to the master program, which implies that the sequence $\{z^{(k)}\}$ is nondecreasing. Furthermore, note that C_{CSB}, which is the optimal value of the dual problem, is bounded by $z^{(k)} \leq C_{\mathrm{CSB}} \leq \min_{\ell\in\{1,\ldots,k\}} \Theta(\boldsymbol{\lambda}^{(\ell)})$ with the upper bound being nonincreasing in k. Hence, we stop the algorithm after iteration K and set $C_{\mathrm{CSB}} = z^{(K)}$ if $\min_{\ell\in\{1,\ldots,K\}} \Theta(\boldsymbol{\lambda}^{(\ell)}) - z^{(K)} < \delta$, where $\delta > 0$ specifies the desired accuracy of the optimal value. For a convergence analysis and further details on the cutting-plane method, we refer to Appendix A.2.

Primal Recovery—The proposed dual decomposition approach allows to evaluate C_{CSB} without determining the optimal time-shares $t_1^\star$ and $t_2^\star$ of the relay receive and transmit phases. But we are usually also interested in these values, e.g., for designing resource allocation protocols. In order to determine $t_1^\star$ and $t_2^\star$, we need to recover the optimal primal variables from the sequences of primal and dual variables generated by the cutting-plane algorithm, a process called *primal recovery*.

Suppose the cutting-plane algorithm has converged after iteration K, and consider the dual problem of the corresponding master program, which reads as

$$\max_{R',\tau_0,\ldots,\tau_{K-1}} R' \quad \text{s.t.} \quad R'\mathbf{1} \leq \sum_{\ell=0}^{K-1} \tau_\ell \boldsymbol{s}^{(\ell)}, \quad \tau_\ell \geq 0, \forall \ell \in \{0,\ldots,K-1\}, \quad \sum_{\ell=0}^{K-1} \tau_\ell = 1. \tag{6.30}$$

Here, R' and τ_ℓ are Lagrangian multipliers associated with the constraints $\mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1$ and $z \geq \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(\ell)}$ of the master program, respectively. Moreover, we remark that this problem is an approximation of the primal problem given in (6.21), where $\mathcal{S}_{\mathrm{CSB}}$ is replaced by $\operatorname{conv}(\{\boldsymbol{s}^{(0)},\dots,\boldsymbol{s}^{(K-1)}\}) \subseteq \mathcal{S}_{\mathrm{CSB}}$, i.e., by a convex combination of feasible points. Now, let $\mathcal{I}_1 = \{\ell : \boldsymbol{s}^{(\ell)} \in \mathcal{S}_{\mathrm{CSB1}}\} \subseteq \{0,\dots,K-1\}$ and $\mathcal{I}_2 = \{\ell : \boldsymbol{s}^{(\ell)} \in \mathcal{S}_{\mathrm{CSB2}}\}\backslash\mathcal{I}_1 \subseteq \{0,\dots,K-1\}$, then we can rewrite the problem given in (6.30) as

$$\begin{aligned}\max_{R',\tau_0,\dots,\tau_{K-1}} \ R' \quad \text{s.t.} \quad & R'\mathbf{1} \leq \sum_{\ell\in\mathcal{I}_1}\tau_\ell\boldsymbol{s}^{(\ell)} + \sum_{\ell\in\mathcal{I}_2}\tau_\ell\boldsymbol{s}^{(\ell)}, \\ & \tau_\ell \geq 0, \forall\ell\in\{0,\dots,K-1\}, \quad \sum_{\ell=0}^{K-1}\tau_\ell = 1.\end{aligned} \tag{6.31}$$

Furthermore, because $\mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{CSB2}}$ are convex, it follows that

$$\sum_{\ell\in\mathcal{I}_m}\tau_\ell\boldsymbol{s}^{(\ell)} = \sum_{k\in\mathcal{I}_m}\tau_k\sum_{\ell\in\mathcal{I}_m}\frac{\tau_\ell}{\sum_{k\in\mathcal{I}_m}\tau_k}\boldsymbol{s}^{(\ell)} = \Bigg(\sum_{k\in\mathcal{I}_m}\tau_k\Bigg)\boldsymbol{s}_m \tag{6.32}$$

for some $\boldsymbol{s}_m \in \mathcal{S}_{\mathrm{CSB}m}$, $m \in \{1,2\}$. Finally, if we insert (6.32) into (6.31) and compare the result to the original optimization problem given in (6.19), we can conclude that

$$t_m^\star = \sum_{\ell\in\mathcal{I}_m}\tau_\ell. \tag{6.33}$$

That is, the optimal time-shares $t_1^\star$ and $t_2^\star$ can easily be obtained from the Lagrangian dual variables that correspond to the constraints $z \geq \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(\ell)}$ in the master program of the final cutting-plane iteration (cf. also Appendix A.2).

Remark 6.2. Since $\mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{CSB2}}$ are convex, time-sharing within either of the two rate regions that correspond to the relay receive and transmit phases, respectively, is not necessary. As a result, there is usually only one $\ell \in \mathcal{I}_m$ such that $\tau_\ell > 0$ for $m = 1$ and $m = 2$.

6.2.2 Average Power Constraint

Now, suppose the source and the relay are subject to the average power constraint. The CSB for the half-duplex Gaussian MIMO relay channel is then given by

$$\begin{aligned} C_{\mathrm{CSB,av}} = & \max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{S1}},\boldsymbol{C}_2} R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) + t_2\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S|R2}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_2\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0, \ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2 \succcurlyeq \mathbf{0}, \\ & t_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) + t_2\operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad t_2\operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.34}$$

Like the corresponding problem with the per-phase power constraint (6.12), this optimization problem is jointly convex in $(R, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2)$ if the time-shares $t_1, t_2 \geq 0$ are fixed. Furthermore, following the steps in Section 6.2.1, we can reformulate the problem given in (6.34) by means of an auxiliary variable $\boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}$ such that

$$\begin{aligned} C_{\mathrm{CSB,av}} = &\max_{R, t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2, \boldsymbol{C}_{\mathrm{Q}}} R \\ \text{s.t.}\quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_2\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & t_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) + t_2 \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad t_2 \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.35}$$

This reformulated problem is again convex in $(R, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2, \boldsymbol{C}_{\mathrm{Q}})$ for given $t_1, t_2 \geq 0$, but not jointly convex in all optimization variables. Moreover, by applying the same change of variables as for the per-phase power constraint, cf. (6.14), it can easily be shown that the maximization problem given in (6.35) is equivalent to

$$\begin{aligned} C_{\mathrm{CSB,av}} = &\max_{R, t_1, t_2, \boldsymbol{K}_{\mathrm{S1}}, \boldsymbol{K}_2, \boldsymbol{K}_{\mathrm{Q}}} R \\ \text{s.t.}\quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}(\boldsymbol{K}_{\mathrm{S1}}/t_1)\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{K}_{\mathrm{Q}}/t_2)\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{K}_{\mathrm{S1}}/t_1)\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}(\boldsymbol{K}_2/t_2)\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{K}_{\mathrm{S1}}, \boldsymbol{K}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{K}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{K}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & \operatorname{tr}(\boldsymbol{K}_{\mathrm{S1}}) + \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{K}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{K}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.36}$$

Note that the rate constraints in (6.36) are the same as in the corresponding optimization problem with the per-phase power constraint, cf. (6.15), where they were shown to define a convex set. In addition, all other constraints are linear functions of the optimization variables so that we can conclude the following:

Proposition 6.3. *For the half-duplex Gaussian MIMO relay channel with the average power constraint,* $C_{\mathrm{CSB,av}}$ *can be determined as the solution of the convex optimization problem given in* (6.36).

Like for the per-phase power constraint, this convex optimization problem does not satisfy the DCP ruleset, so we again propose a dual decomposition approach that allows to efficiently evaluate $C_{\mathrm{CSB,av}}$ in the Lagrangian dual domain. For this purpose, we first define the two rate-power regions

$$\begin{aligned} \mathcal{Z}_{\mathrm{CSB1}} = \Big\{(\boldsymbol{s}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : {}& s_1 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right), \\ & s_2 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & p_1 = \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}),\ p_2 = 0, \quad \boldsymbol{C}_{\mathrm{S1}} \succcurlyeq \mathbf{0}\Big\} \end{aligned} \tag{6.37}$$

and

$$\begin{aligned}\mathcal{Z}_{\mathrm{CSB2}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : {} & s_1 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & s_2 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & p_1 = \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}),\ p_2 = \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}), \\ & \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}\Big\},\end{aligned} \tag{6.38}$$

which specify the contributions of the relay receive and transmit phases to the CSB, respectively, both in terms of rate and power consumption. Note that $\mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{CSB2}}$ are closed and convex, but not compact, because neither the rates nor the transmit powers are bounded above (except for p_2 in $\mathcal{Z}_{\mathrm{CSB1}}$). Moreover, with $\boldsymbol{p}_{\mathrm{tx}} = [P_{\mathrm{S}}, P_{\mathrm{R}}]^{\mathrm{T}}$, the maximization problem given in (6.35) can be rewritten as

$$\begin{aligned}C_{\mathrm{CSB,av}} = \max_{R,t_1,t_2,\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{p}_1,\boldsymbol{p}_2} R \quad \text{s.t.} \quad & R\mathbf{1} \leq t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \quad t_1\boldsymbol{p}_1 + t_2\boldsymbol{p}_2 \leq \boldsymbol{p}_{\mathrm{tx}}, \\ & (\boldsymbol{s}_1,\boldsymbol{p}_1) \in \mathcal{Z}_{\mathrm{CSB1}},\ (\boldsymbol{s}_2,\boldsymbol{p}_2) \in \mathcal{Z}_{\mathrm{CSB2}}, \\ & t_1, t_2 \geq 0,\ t_1 + t_2 = 1.\end{aligned} \tag{6.39}$$

For the next reformulation step, let us define

$$\begin{aligned}\mathcal{Z}_{\mathrm{CSB}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : {} & \boldsymbol{s} = t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \quad \boldsymbol{p} = t_1\boldsymbol{p}_1 + t_2\boldsymbol{p}_2, \\ & (\boldsymbol{s}_1,\boldsymbol{p}_1) \in \mathcal{Z}_{\mathrm{CSB1}},\ (\boldsymbol{s}_2,\boldsymbol{p}_2) \in \mathcal{Z}_{\mathrm{CSB2}}, \quad t_1, t_2 \geq 0,\ t_1 + t_2 = 1\Big\},\end{aligned} \tag{6.40}$$

which is also closed and convex since the convex sum of convex sets is convex, cf. [11, Section 2.3.2]. Using this definition, we can rewrite (6.39) as

$$C_{\mathrm{CSB,av}} = \max_{R,\boldsymbol{s},\boldsymbol{p}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{p} \leq \boldsymbol{p}_{\mathrm{tx}}, \quad (\boldsymbol{s},\boldsymbol{p}) \in \mathcal{Z}_{\mathrm{CSB}}. \tag{6.41}$$

Because $\mathcal{Z}_{\mathrm{CSB}}$ is a convex set with nonempty interior, this convex optimization problem can again be solved in the Lagrangian dual domain, cf. [11, Section 5.3.2].

Derivation of the Dual Problem—If we incorporate the constraints $R\mathbf{1} \leq \boldsymbol{s}$ and $\boldsymbol{p} \leq \boldsymbol{p}_{\mathrm{tx}}$ into the objective function using the Lagrangian multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, respectively, the Lagrangian function reads as

$$L(R,\boldsymbol{s},\boldsymbol{p},\boldsymbol{\lambda},\boldsymbol{\mu}) = R - \boldsymbol{\lambda}^{\mathrm{T}}(R\mathbf{1} - \boldsymbol{s}) - \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{p} - \boldsymbol{p}_{\mathrm{tx}}), \tag{6.42}$$

and the corresponding dual function is given by

$$\begin{aligned}\Theta(\boldsymbol{\lambda},\boldsymbol{\mu}) &= \sup_{R,(\boldsymbol{s},\boldsymbol{p})\in\mathcal{Z}_{\mathrm{CSB}}} L(R,\boldsymbol{s},\boldsymbol{p},\boldsymbol{\lambda},\boldsymbol{\mu}) \\ &= \begin{cases}\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_{\mathrm{tx}} + \sup_{(\boldsymbol{s},\boldsymbol{p})\in\mathcal{Z}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p} & \text{if } \mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1, \\ +\infty & \text{otherwise.}\end{cases}\end{aligned} \tag{6.43}$$

Proposition 6.4. *For any $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $\mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1$ and $\lambda_2 > 0$, the dual function $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is finite if and only if $\boldsymbol{\mu} > \mathbf{0}$.*

Proof. First, note that since

$$\begin{aligned}\sup_{(\boldsymbol{s},\boldsymbol{p})\in\mathcal{Z}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p} &= \sup_{\substack{(\boldsymbol{s}_1,\boldsymbol{p}_1)\in\mathcal{Z}_{\mathrm{CSB1}},(\boldsymbol{s}_2,\boldsymbol{p}_2)\in\mathcal{Z}_{\mathrm{CSB2}}\\ t_1,t_2\geq 0,\ t_1+t_2=1}} t_1(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_1 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_1) + t_2(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_2 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_2)\\ &= \sup_{m=1,2}\ \sup_{(\boldsymbol{s}_m,\boldsymbol{p}_m)\in\mathcal{Z}_{\mathrm{CSB}m}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_m - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_m,\end{aligned} \tag{6.44}$$

it follows that $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < +\infty$ if and only if $\sup_{(\boldsymbol{s}_m,\boldsymbol{p}_m)\in\mathcal{Z}_{\mathrm{CSB}m}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_m - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_m < +\infty$ for both $m = 1$ and $m = 2$.

In order to prove the "if" part of the proposition, assume that $\boldsymbol{\mu} > \mathbf{0}$. For any $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $\mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1$, we then have

$$\begin{aligned}\sup_{(\boldsymbol{s}_1,\boldsymbol{p}_1)\in\mathcal{Z}_{\mathrm{CSB1}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_1 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_1 &\leq \sup_{\boldsymbol{C}_{\mathrm{S1}}\succcurlyeq\mathbf{0}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right) - \mu_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}})\\ &= \sup_{\boldsymbol{C}_{\mathrm{S1}}\succcurlyeq\mathbf{0}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\right) - \mu_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}),\end{aligned} \tag{6.45}$$

where the inequality in the first line is due to the fact that $\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\right)$ for any $\boldsymbol{C}_{\mathrm{S1}} \succcurlyeq \mathbf{0}$. Now, let $\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{S\{RD\}}} = \boldsymbol{U}\boldsymbol{\Phi}\boldsymbol{U}^{\mathrm{H}}$ with $\boldsymbol{\Phi} = \operatorname{diag}(\phi_1, \ldots, \phi_{N_{\mathrm{S}}}) \succcurlyeq \mathbf{0}$ and $\boldsymbol{C}_{\mathrm{S1}} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{H}}$ with $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_{N_{\mathrm{S}}}) \succcurlyeq \mathbf{0}$ denote the eigenvalue decompositions of $\boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{S\{RD\}}}$ and $\boldsymbol{C}_{\mathrm{S1}}$, respectively. Then, the trace of $\boldsymbol{C}_{\mathrm{S1}}$ is independent of the modal matrix $\boldsymbol{V}$ and equal to the sum of its nonnegative eigenvalues $\sigma_1, \ldots, \sigma_{N_{\mathrm{S}}}$. Furthermore, if $\mu_1 > 0$, Hadamard's inequality [24, Theorem 17.9.2] can be used to show that the supremum is attained for $\boldsymbol{V} = \boldsymbol{U}$ and $\sigma_i = \max\left\{\frac{1}{\mu_1} - \frac{1}{\phi_i}, 0\right\}$, $i \in \{1, \ldots, N_{\mathrm{S}}\}$, i.e.,

$$\begin{aligned}&\sup_{\boldsymbol{C}_{\mathrm{S1}}\succcurlyeq\mathbf{0}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{S\{RD\}}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{S\{RD\}}}\boldsymbol{C}_{\mathrm{S1}}\right) - \mu_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}})\\ &\qquad = \sum_{i=1}^{N_{\mathrm{S}}} \max_{\sigma_i\geq 0} \log\left(1 + \phi_i\sigma_i\right) - \mu_1\sigma_i = \sum_{i\in\mathcal{I}} \log\left(\frac{\phi_i}{\mu_1}\right) - \left(1 - \frac{\mu_1}{\phi_i}\right) < +\infty,\end{aligned} \tag{6.46}$$

where $\mathcal{I} = \{i : \phi_i > \mu_1\} \subseteq \{1, \ldots, N_{\mathrm{S}}\}$.

Similarly, for any $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $\mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1$, we have

$$\begin{aligned}\sup_{(\boldsymbol{s}_2,\boldsymbol{p}_2)\in\mathcal{Z}_{\mathrm{CSB2}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_2 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_2 &\leq \sup_{\boldsymbol{C}_2\succcurlyeq\mathbf{0}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_2\boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\right) - \nu \operatorname{tr}(\boldsymbol{C}_2)\\ &= \sup_{\boldsymbol{C}_2\succcurlyeq\mathbf{0}} \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{\{SR\}D}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{\{SR\}D}}\boldsymbol{C}_2\right) - \nu \operatorname{tr}(\boldsymbol{C}_2)\end{aligned} \tag{6.47}$$

with $\nu = \min\{\mu_1, \mu_2\} > 0$ if $\boldsymbol{\mu} > \mathbf{0}$. Using the same arguments as above, it can hence be shown that $\sup_{(\boldsymbol{s}_2,\boldsymbol{p}_2)\in\mathcal{Z}_{\mathrm{CSB2}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_2 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_2 < +\infty$ whenever $\boldsymbol{\mu} > \mathbf{0}$.

Conversely, assume that $\boldsymbol{\mu} \ngtr \mathbf{0}$. In particular, if $\mu_1 \leq 0$, we have

$$\sup_{(\boldsymbol{s}_1,\boldsymbol{p}_1)\in\mathcal{Z}_{\mathrm{CSB1}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_1 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_1 \geq \sup_{a\geq 0} \log\det\left(\mathbf{I} + a\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + |\mu_1|aN_{\mathrm{S}} = +\infty \tag{6.48}$$

for any $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $\mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1$. Likewise, if $\mu_2 \leq 0$ and $\boldsymbol{\lambda} \geq \mathbf{0}$ is such that $\lambda_2 > 0$, it follows that

$$\sup_{(\boldsymbol{s}_2,\boldsymbol{p}_2)\in\mathcal{Z}_{\mathrm{CSB2}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_2 - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_2 \geq \sup_{b\geq 0} \lambda_2\log\det\left(\mathbf{I} + b\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\right) + |\mu_2|bN_{\mathrm{R}} = +\infty, \tag{6.49}$$

which completes the proof. □

The meaning of Proposition 6.4 is as follows. Note that for the two subproblems $\sup_{(\boldsymbol{s}_m,\boldsymbol{p}_m)\in\mathcal{Z}_{\mathrm{CSB}m}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}_m - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_m$, $m \in \{1,2\}$, the Lagrangian multipliers μ_1 and μ_2 can be understood as *prices* for the transmit powers p_{1m} and p_{2m} of the source and the relay in phase m, respectively. If the prices are positive, both subproblems have finite optimal solutions because the transmit power cost $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_m$ is a linear function of $\boldsymbol{p}_m$, whereas the rates only increase logarithmically with the powers. If one of the prices is zero (or negative), on the other hand, the transmit power of the corresponding node and the associated rate value(s) can be increased to infinity without incurring any costs.

Remark 6.3. For $\boldsymbol{\lambda} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$, the dual function is also finite if $\mu_2 = 0$ and $\mu_1 > 0$. However, whether or not the relay transmits, and hence the entire relay transmit phase, has no effect on $\Theta(\boldsymbol{\lambda},\boldsymbol{\mu})$ if $\lambda_2 = 0$ since the only rate expression that depends on C_{R2} (s_{2m}) is weighted with zero in this case. If $\boldsymbol{\lambda} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ were an optimal solution to the dual problem, it would thus follow that $C_{\mathrm{CSB,av}} = R_{\mathrm{P2P}}$. But this cannot happen unless $\boldsymbol{H}_{\mathrm{SR}} = \mathbf{0}$ and/or $\boldsymbol{H}_{\mathrm{RD}} = \mathbf{0}$, cf. Remark 6.1, so for all relevant scenarios, we can conclude that $\lambda_2^\star > 0$, which in turn implies $\boldsymbol{\mu}^\star > \mathbf{0}$.

For any $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $\lambda_2 > 0$, it follows from Proposition 6.4 that the dual function is equal to

$$\Theta(\boldsymbol{\lambda},\boldsymbol{\mu}) = \begin{cases} \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_{\mathrm{tx}} + \max_{(\boldsymbol{s},\boldsymbol{p})\in\mathcal{Z}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s} - \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p} & \text{if } \mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1, \boldsymbol{\mu} > \mathbf{0}, \\ +\infty & \text{otherwise.} \end{cases} \tag{6.50}$$

In order to determine the CSB for the half-duplex Gaussian MIMO relay channel with the average power constraint, we therefore have to solve the dual problem

$$C_{\mathrm{CSB,av}} = \min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \max_{\boldsymbol{s},\boldsymbol{p}} \ \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{p}_{\mathrm{tx}} - \boldsymbol{p}) + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s} \quad \text{s.t.} \quad (\boldsymbol{s},\boldsymbol{p}) \in \mathcal{Z}_{\mathrm{CSB}}, \quad \boldsymbol{\lambda} \geq \mathbf{0},\ \mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1, \quad \boldsymbol{\mu} > \mathbf{0}. \tag{6.51}$$

Like for the corresponding dual problem with the per-phase power constraint, cf. (6.24), this can for example be done by means of the cutting-plane method.

Solution by means of the Cutting-Plane Algorithm—Suppose we apply the cutting-plane algorithm to the dual problem given in (6.51). The master program of the k-th iteration then reads as

$$C_{\text{CSB,av}} = \min_{z,\boldsymbol{\lambda},\boldsymbol{\mu}} z \quad \text{s.t.} \quad \begin{aligned} & z \geq \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{p}_{\text{tx}} - \boldsymbol{p}^{(\ell)}) + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(\ell)}, \forall \ell \in \{0, \ldots, k-1\}, \\ & \boldsymbol{\lambda} \geq \mathbf{0},\ \mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1, \quad \boldsymbol{\mu} > \mathbf{0}, \end{aligned} \tag{6.52}$$

where $(\boldsymbol{s}^{(\ell)}, \boldsymbol{p}^{(\ell)}) \in \mathcal{Z}_{\text{CSB}}, \forall \ell \in \{0, \ldots, k-1\}$. Furthermore, we can conclude from (6.44) that the Lagrangian subproblem in iteration k, i.e., the evaluation of the dual function at $(\boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}^{(k)})$, requires to solve two independent convex optimization problems, one over each of the two rate-power regions $\mathcal{Z}_{\text{CSB1}}$ and $\mathcal{Z}_{\text{CSB2}}$ that are associated with the relay receive and the relay transmit phase, respectively. For this purpose, we can again use SDP solvers like SDPT3 [126, 127] or LogdetPPA [133] that are capable of dealing with log-det terms in the objective function.

Remark 6.4. For the cutting-plane algorithm to work in practice, the constraint $\boldsymbol{\mu} > \mathbf{0}$ in the master program must be replaced by $\boldsymbol{\mu} \geq \mathbf{0}$. This does not change the optimal solution of the dual problem given in (6.51) because we know that $\boldsymbol{\mu}^\star > \mathbf{0}$. However, if $z^{(k)}, \boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}^{(k)}$ are the optimizers of the master program in iteration k, the initialization of the cutting-plane algorithm must ensure that $z^{(k)}$ is finite and that $\boldsymbol{\mu}^{(k)} > \mathbf{0}$ (or $\mu_2^{(k)} = 0$ and $\lambda_2^{(k)} = 0$) for all $k \in \mathbb{N}$. Otherwise, the algorithm runs into problems when the dual function is evaluated.

Remark 6.5. Because $\boldsymbol{\mu}^\star > \mathbf{0}$, the *complementary slackness* condition $\boldsymbol{\mu}^{\star,\mathrm{T}}(\boldsymbol{p}^\star - \boldsymbol{p}_{\text{tx}}) = 0$ of the primal problem given in (6.41) implies that $\boldsymbol{p}^\star = \boldsymbol{p}_{\text{tx}}$, which means that both the source and the relay exhaust their available transmit power budgets.

Primal Recovery—If the cutting-plane algorithm is used to solve the dual problem, the optimal time-shares $t_1^\star$ and $t_2^\star$ of the relay receive and transmit phases can be recovered in the same way as for the per-phase power constraint. In particular, assume that the cutting-plane algorithm has converged after iteration K, and consider the dual problem of the corresponding master program, which reads as

$$\max_{R',\tau_0,\ldots,\tau_{K-1}} R' \quad \text{s.t.} \quad \begin{aligned} & R'\mathbf{1} \leq \sum_{\ell=0}^{K-1} \tau_\ell \boldsymbol{s}^{(\ell)}, \quad \sum_{\ell=0}^{K-1} \tau_\ell \boldsymbol{p}^{(\ell)} \leq \boldsymbol{p}_{\text{tx}}, \\ & \tau_\ell \geq 0, \forall \ell \in \{0, \ldots, K-1\}, \quad \sum_{\ell=0}^{K-1} \tau_\ell = 1. \end{aligned} \tag{6.53}$$

Here, R' and τ_ℓ again represent Lagrangian multipliers that are associated with the constraints $\mathbf{1}^{\mathrm{T}}\boldsymbol{\lambda} = 1$ and $z \geq \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{p}_{\text{tx}} - \boldsymbol{p}^{(\ell)}) + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(\ell)}$ of the master program, respectively. Moreover, this problem is an approximation of the primal problem given in (6.41), where the set $\mathcal{Z}_{\text{CSB}}$ is replaced by $\operatorname{conv}\left(\{(\boldsymbol{s}^{(0)}, \boldsymbol{p}^{(0)}), \ldots, (\boldsymbol{s}^{(K-1)}, \boldsymbol{p}^{(K-1)})\}\right) \subseteq \mathcal{Z}_{\text{CSB}}$. Letting

$\mathcal{I}_1 = \{\ell : (\boldsymbol{s}^{(\ell)}, \boldsymbol{p}^{(\ell)}) \in \mathcal{Z}_{\mathrm{CSB1}}\} \subseteq \{0, \ldots, K-1\}$ and $\mathcal{I}_2 = \{\ell : (\boldsymbol{s}^{(\ell)}, \boldsymbol{p}^{(\ell)}) \in \mathcal{Z}_{\mathrm{CSB2}}\} \setminus \mathcal{I}_1 \subseteq \{0, \ldots, K-1\}$, it follows that we can rewrite the problem given in (6.53) as

$$\max_{R', \tau_0, \ldots, \tau_{K-1}} R' \quad \text{s.t.} \quad R'\mathbf{1} \leq \sum_{\ell \in \mathcal{I}_1} \tau_\ell \boldsymbol{s}^{(\ell)} + \sum_{\ell \in \mathcal{I}_2} \tau_\ell \boldsymbol{s}^{(\ell)}, \quad \sum_{\ell \in \mathcal{I}_1} \tau_\ell \boldsymbol{p}^{(\ell)} + \sum_{\ell \in \mathcal{I}_2} \tau_\ell \boldsymbol{p}^{(\ell)} \leq \boldsymbol{p}_{\mathrm{tx}},$$
$$\tau_\ell \geq 0, \forall \ell \in \{0, \ldots, K-1\}, \quad \sum_{\ell=0}^{K-1} \tau_\ell = 1. \tag{6.54}$$

Furthermore, note that since $\mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{CSB2}}$ are convex, it can again be shown that $\sum_{\ell \in \mathcal{I}_m} \tau_\ell (\boldsymbol{s}^{(\ell)}, \boldsymbol{p}^{(\ell)}) = (\sum_{k \in \mathcal{I}_m} \tau_k)(\boldsymbol{s}_m, \boldsymbol{p}_m)$ for some $(\boldsymbol{s}_m, \boldsymbol{p}_m) \in \mathcal{Z}_{\mathrm{CSB}m}$, $m \in \{1, 2\}$. If we plug these expressions into (6.54) and compare the result to the original optimization problem given in (6.39), we can eventually conclude that

$$t_m^\star = \sum_{\ell \in \mathcal{I}_m} \tau_\ell. \tag{6.55}$$

This means that the optimal time-shares can again be obtained from the Lagrangian dual variables that correspond to the constraints $z \geq \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{p}_{\mathrm{tx}} - \boldsymbol{p}^{(\ell)}) + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s}^{(\ell)}$ in the master program of the final cutting-plane iteration. Beyond that, Remark 6.2 applies here as well since both rate-power regions $\mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{CSB2}}$ are convex.

To conclude this section, we remark that the average and the per-phase power constraints can also be dealt with simultaneously.[2] To this end, we simply need to add the per-phase transmit power constraints we want to impose on the source and the relay to the definitions of the rate-power regions $\mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{CSB2}}$. Since these sets are then bounded, Proposition 6.4 becomes obsolete as no condition on $\boldsymbol{\mu}$ is required for the dual function to be finite in this case. Consequently, the per-phase power constraint can easily be incorporated into the dual decomposition approach presented in this section. Because the converse is not true, the approach for the average power constraint is more general than that for the per-phase power constraint.

6.3 Decode-and-Forward

Since Theorem 5.3 can be derived using weakly or strongly typical sequences, it is also valid for continuous-alphabet channels. Therefore, the maximum achievable decode-and-forward (DF) rate for the half-duplex Gaussian MIMO relay channel is equal to

$$\begin{aligned} R_{\mathrm{DF}} = \max_{\substack{R, t_1, t_2 \\ p(\boldsymbol{x}_{\mathrm{S1}}) p(\boldsymbol{x}_{\mathrm{S2}}, \boldsymbol{x}_{\mathrm{R2}})}} R \quad \text{s.t.} \quad & R \leq t_1 I(\boldsymbol{x}_{\mathrm{S1}}; \boldsymbol{y}_{\mathrm{R1}}) + t_2 I(\boldsymbol{x}_{\mathrm{S2}}; \boldsymbol{y}_{\mathrm{D2}} | \boldsymbol{x}_{\mathrm{R2}}), \\ & R \leq t_1 I(\boldsymbol{x}_{\mathrm{S1}}; \boldsymbol{y}_{\mathrm{D1}}) + t_2 I(\boldsymbol{x}_{\mathrm{S2}}, \boldsymbol{x}_{\mathrm{R2}}; \boldsymbol{y}_{\mathrm{D2}}), \\ & t_1, t_2 \geq 0, \ t_1 + t_2 = 1, \\ & (\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{S2}}, \boldsymbol{C}_{\mathrm{R2}}) \in \mathcal{P}, \end{aligned} \tag{6.56}$$

[2]Note that P_{S} and P_{R} in (6.4) should then be larger than the corresponding values in (6.6) because otherwise the per-phase power constraint dominates the average power constraint, cf. Section 6.1.2.

where $\mathcal{P}$ is given by (6.8) or (6.9) depending on the considered power constraint. As already discussed in Section 5.3, the only difference between the optimization problems given in (6.7) and (6.56) is in the first mutual information term of the first rate bound, where an additional $\boldsymbol{y}_{\mathrm{D1}}$ appears in in (6.7) as compared to (6.56). Apart from that, both problems have the same structure, optimization variables, and constraints. Like for the CSB (and the full-duplex case), we can hence use the entropy maximizing property of the Gaussian distribution to show that the maximum achievable DF rate is attained by proper complex Gaussian channel inputs for both the per-phase and the average power constraint.

Theorem 6.5. *For the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint, the maximum achievable decode-and-forward (DF) rate is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. The proof of this result follows exactly the same lines as the proof of Theorem 6.1. □

As a consequence, the optimal distributions of the source and relay inputs can again be represented by the covariance matrix of $\boldsymbol{x}_{\mathrm{S1}}$ and the joint covariance matrix of $\boldsymbol{x}_{\mathrm{S2}}$ and $\boldsymbol{x}_{\mathrm{R2}}$. With $\boldsymbol{x}_{\mathrm{S1}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{S1}})$, $\left[\begin{smallmatrix}\boldsymbol{x}_{\mathrm{S2}}\\ \boldsymbol{x}_{\mathrm{R2}}\end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_2)$, and assuming the Gaussian noise to be white, the mutual information terms that upper bound R_{DF} are given by

$$\begin{aligned}
I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right), \\
I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
I(\boldsymbol{x}_{\mathrm{S2}};\boldsymbol{y}_{\mathrm{D2}}|\boldsymbol{x}_{\mathrm{R2}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S|R2}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
I(\boldsymbol{x}_{\mathrm{S2}},\boldsymbol{x}_{\mathrm{R2}};\boldsymbol{y}_{\mathrm{D2}}) &= \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right).
\end{aligned} \tag{6.57}$$

In order to determine the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel with the power constraint specified by $\mathcal{P}$, we thus have to solve the optimization problem

$$\begin{aligned}
R_{\mathrm{DF}} = \max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{S1}},\boldsymbol{C}_2,\boldsymbol{C}_{\mathrm{Q}}} \quad & R \\
\text{s.t.} \quad & R \le t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
& R \le t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\
& t_1, t_2 \ge 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\
& (\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P}.
\end{aligned} \tag{6.58}$$

Note that to obtain this formulation of the DF rate maximization problem, we have already reformulated the constraint involving the conditional covariance matrix $\boldsymbol{C}_{\mathrm{S|R2}}$ by introducing an auxiliary variable $\boldsymbol{C}_{\mathrm{Q}} = \boldsymbol{C}_{\mathrm{S|R2}} \succcurlyeq \mathbf{0}$, relaxing the equality constraint to $\mathbf{0} \preccurlyeq \boldsymbol{C}_{\mathrm{Q}} \preccurlyeq \boldsymbol{C}_{\mathrm{S|R2}}$, and applying Lemma 3.2.

While the problem given in (6.58) is only convex if the time-shares $t_1, t_2 \geq 0$ are fixed, we could again apply a change of variables to obtain a problem that is jointly convex in all optimization variables. For both the per-phase and the average power constraint, however, the resulting convex optimization problems would not satisfy the DCP ruleset. So instead of elaborating on a change of variables, we discuss how the corresponding maximum achievable DF rates can also be evaluated using the dual decomposition approaches presented in Sections 6.2.1 and 6.2.2.

If the source and the relay are subject to the per-phase power constraint, the DF rate maximization problem is given by

$$\begin{aligned} R_{\mathrm{DF}} = \max_{R, t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2, \boldsymbol{C}_{\mathrm{Q}}} \; & R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{S1}} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{Q}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{S1}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}} \boldsymbol{C}_2 \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\; t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}} \boldsymbol{C}_{\mathrm{Q}} \boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}} \boldsymbol{C}_2 \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}} \boldsymbol{C}_2 \boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.59}$$

Now, we first define the two rate regions

$$\begin{aligned} \mathcal{S}_{\mathrm{DF1}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : \; & s_1 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}} \boldsymbol{C}_{\mathrm{S1}} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right), \\ & s_2 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{S1}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & \boldsymbol{C}_{\mathrm{S1}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) \leq P_{\mathrm{S}} \Big\} \end{aligned} \tag{6.60}$$

and

$$\begin{aligned} \mathcal{S}_{\mathrm{DF2}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : \; & s_1 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C}_{\mathrm{Q}} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & s_2 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}} \boldsymbol{C}_2 \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}} \boldsymbol{C}_{\mathrm{Q}} \boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}} \boldsymbol{C}_2 \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}} \boldsymbol{C}_2 \boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}} \Big\}, \end{aligned} \tag{6.61}$$

where $\mathcal{S}_{\mathrm{DF1}} \subseteq \mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{DF2}} = \mathcal{S}_{\mathrm{CSB2}}$ are both compact and convex. By means of these two rate regions, which specify the contributions of the relay receive and the relay transmit phase to R_{DF}, respectively, (6.59) can be rewritten as

$$\begin{aligned} R_{\mathrm{DF}} = \max_{R, t_1, t_2, \boldsymbol{s}_1, \boldsymbol{s}_2} R \quad \text{s.t.} \quad & R\mathbf{1} \leq t_1 \boldsymbol{s}_1 + t_2 \boldsymbol{s}_2, \\ & \boldsymbol{s}_1 \in \mathcal{S}_{\mathrm{DF1}},\; \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{DF2}}, \quad t_1, t_2 \geq 0,\; t_1 + t_2 = 1. \end{aligned} \tag{6.62}$$

Furthermore, if we define

$$\begin{aligned} \mathcal{S}_{\mathrm{DF}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : \; & \boldsymbol{s} = t_1 \boldsymbol{s}_1 + t_2 \boldsymbol{s}_2, \\ & \boldsymbol{s}_1 \in \mathcal{S}_{\mathrm{DF1}},\; \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{DF2}}, \quad t_1, t_2 \geq 0,\; t_1 + t_2 = 1 \Big\}, \end{aligned} \tag{6.63}$$

we can rewrite (6.62) once more to obtain

$$R_{\mathrm{DF}} = \max_{R,\boldsymbol{s}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{DF}}. \tag{6.64}$$

Finally, since $\mathcal{S}_{\mathrm{DF}}$ is a convex set with nonempty interior, this is a convex optimization problem that can equivalently be solved in the Lagrangian dual domain. In particular, note that for this purpose, we can apply the dual decomposition approach presented in Section 6.2.1, where we simply need to replace $\mathcal{S}_{\mathrm{CSB}}$ with $\mathcal{S}_{\mathrm{DF}}$, or equivalently, replace $\mathcal{S}_{\mathrm{CSB1}}$ with $\mathcal{S}_{\mathrm{DF1}}$ as $\mathcal{S}_{\mathrm{DF2}} = \mathcal{S}_{\mathrm{CSB2}}$.

Similarly, if the source and the relay are subject to the average power constraint, the DF rate maximization problem reads as

$$\begin{aligned} R_{\mathrm{DF,av}} = \max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{S1}},\boldsymbol{C}_2,\boldsymbol{C}_{\mathrm{Q}}} & R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & t_1 \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}) + t_2 \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}}, \quad t_2 \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{6.65}$$

If we define the two rate-power regions

$$\begin{aligned} \mathcal{Z}_{\mathrm{DF1}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ p_1 &= \operatorname{tr}(\boldsymbol{C}_{\mathrm{S1}}),\ p_2 = 0, \quad \boldsymbol{C}_{\mathrm{S1}} \succcurlyeq \mathbf{0}\Big\} \end{aligned} \tag{6.66}$$

and

$$\begin{aligned} \mathcal{Z}_{\mathrm{DF2}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ p_1 &= \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}),\ p_2 = \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}), \\ \boldsymbol{C}_{\mathrm{Q}} &\succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}\Big\}, \end{aligned} \tag{6.67}$$

which specify the contributions (in terms of rate and power) of the relay receive and the relay transmit phase to $R_{\mathrm{DF,av}}$, respectively, the DF rate maximization problem given in (6.65) can be rewritten as

$$\begin{aligned} R_{\mathrm{DF,av}} = \max_{R,t_1,t_2,\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{p}_1,\boldsymbol{p}_2} R \quad \text{s.t.} \quad & R\mathbf{1} \leq t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \quad t_1\boldsymbol{p}_1 + t_2\boldsymbol{p}_2 \leq \boldsymbol{p}_{\mathrm{tx}}, \\ & (\boldsymbol{s}_1,\boldsymbol{p}_1) \in \mathcal{Z}_{\mathrm{DF1}},\ (\boldsymbol{s}_2,\boldsymbol{p}_2) \in \mathcal{Z}_{\mathrm{DF2}}, \\ & t_1, t_2 \geq 0,\ t_1 + t_2 = 1. \end{aligned} \tag{6.68}$$

Note that $\mathcal{Z}_{\mathrm{DF1}} \subseteq \mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{DF2}} = \mathcal{Z}_{\mathrm{CSB2}}$ are closed and convex. They are not compact,

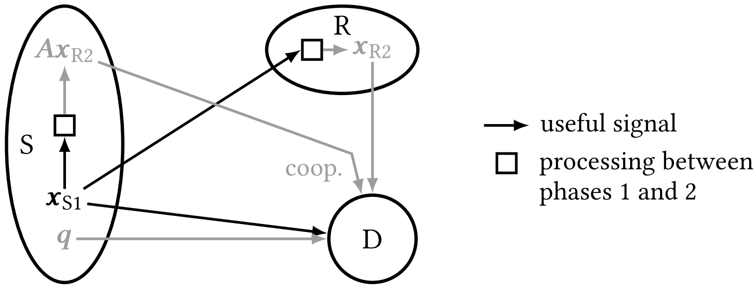


Figure 6.2: Decomposition of the Source Transmit Signals for the Half-Duplex DF Strategy (signals for relay transmit phase in gray)

however, because neither the rates nor the transmit powers are bounded above (except for p_2 in $\mathcal{Z}_{\mathrm{DF1}}$). If we further define

$$\mathcal{Z}_{\mathrm{DF}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : \boldsymbol{s} = t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \quad \boldsymbol{p} = t_1\boldsymbol{p}_1 + t_2\boldsymbol{p}_2, \\ (\boldsymbol{s}_1,\boldsymbol{p}_1) \in \mathcal{Z}_{\mathrm{DF1}},\ (\boldsymbol{s}_2,\boldsymbol{p}_2) \in \mathcal{Z}_{\mathrm{DF2}}, \quad t_1, t_2 \geq 0,\ t_1 + t_2 = 1\Big\}, \tag{6.69}$$

which is also closed and convex, we can rewrite (6.68) again to obtain

$$R_{\mathrm{DF,av}} = \max_{R,\boldsymbol{s},\boldsymbol{p}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{p} \leq \boldsymbol{p}_{\mathrm{tx}}, \quad (\boldsymbol{s},\boldsymbol{p}) \in \mathcal{Z}_{\mathrm{DF}}. \tag{6.70}$$

Like for the per-phase power constraint, this optimization problem is convex and can equivalently be solved in the Lagrangian dual domain. To this end, we can apply the dual decomposition approach presented in Section 6.2.2 if $\mathcal{Z}_{\mathrm{CSB}}$ is replaced with $\mathcal{Z}_{\mathrm{DF}}$, or equivalently, if $\mathcal{Z}_{\mathrm{CSB1}}$ is replaced with $\mathcal{Z}_{\mathrm{DF1}}$ since $\mathcal{Z}_{\mathrm{DF2}} = \mathcal{Z}_{\mathrm{CSB2}}$.

The quintessence of this section is that due to the similarity of the optimization problems that determine the CSB and the maximum achievable DF rate, we only need to solve one type of problem in order to evaluate $C_{\mathrm{CSB}}/C_{\mathrm{CSB,av}}$ and $R_{\mathrm{DF}}/R_{\mathrm{DF,av}}$. Like for the full-duplex case, this means that if we can evaluate $C_{\mathrm{CSB}}/C_{\mathrm{CSB,av}}$, then we can also evaluate $R_{\mathrm{DF}}/R_{\mathrm{DF,av}}$ and vice versa, cf. Remark 3.4.

Before we conclude this section, however, we would also like to point out that the auxiliary variable $\boldsymbol{C}_{\mathrm{Q}}$ in the DF rate maximization problems given in (6.59) and (6.65) has a nice interpretation again. Recall that in the half-duplex DF coding scheme, the source splits its message w into two independent parts w_1, w_2 and the source encoder generates the two codewords $\boldsymbol{x}_{\mathrm{S1}}(w_1)$ and $\boldsymbol{x}_{\mathrm{S2}}(w_1, w_2)$. Due to the causality of the relay encoder, the relay's transmit signal $\boldsymbol{x}_{\mathrm{R2}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{R2}})$ must be a function of the message part w_1 only, and provided that the relay encoding function is deterministic, $\boldsymbol{x}_{\mathrm{R2}}$ is also known to the source. If we let $\boldsymbol{q} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{Q}})$ be a function of w_2 only, the transmit signal of the source in the relay transmit phase can be expressed as the superposition of $\boldsymbol{q}(w_2)$ and $\boldsymbol{x}_{\mathrm{R2}}(w_1)$ given by

$$\boldsymbol{x}_{\mathrm{S2}} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R2}}, \tag{6.71}$$

where $\boldsymbol{A} \in \mathbb{C}^{N_S \times N_R}$ specifies a linear transformation the source may apply to $\boldsymbol{x}_{R2}$. Since $\boldsymbol{q}$ and $\boldsymbol{x}_{R2}$ are independent, it follows that $\boldsymbol{x}_{S2} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_Q + \boldsymbol{A}\boldsymbol{C}_{R2}\boldsymbol{A}^H)$.

Using this form of superposition coding, the source and the relay can *cooperatively* transmit w_1 to the destination in the relay transmit phase by means of $\boldsymbol{A}\boldsymbol{x}_{R2}$ and $\boldsymbol{x}_{R2}$, whereas the signal part $\boldsymbol{q}$, which contains the message w_2, represents the information the relay never decodes as it cannot receive when it is transmitting, cf. Figure 6.2. The auxiliary variable $\boldsymbol{C}_Q$ hence specifies the distribution of the signal part that contains the *new* information the source may communicate to the destination in the relay transmit phase, cf. Section 5.3.

6.4 Partial Decode-and-Forward

Theorem 5.4 can be derived using weakly or strongly typical sequences, so it is also valid for continuous-alphabet channels. Consequently, the maximum achievable partial decode-and-forward (PDF) rate for the half-duplex Gaussian MIMO relay channel is determined by the optimization problem

$$
\begin{aligned}
R_{\mathrm{PDF}} = \max_{\substack{R, t_1, t_2 \\ p(\boldsymbol{u}, \boldsymbol{x}_{S1}) p(\boldsymbol{x}_{S2}, \boldsymbol{x}_{R2})}} R \quad \text{s.t.} \quad & R \leq t_1 I(\boldsymbol{u}; \boldsymbol{y}_{R1}) + t_1 I(\boldsymbol{x}_{S1}; \boldsymbol{y}_{D1} | \boldsymbol{u}) \\
& \quad + t_2 I(\boldsymbol{x}_{S2}; \boldsymbol{y}_{D2} | \boldsymbol{x}_{R2}), \\
& R \leq t_1 I(\boldsymbol{x}_{S1}; \boldsymbol{y}_{D1}) + t_2 I(\boldsymbol{x}_{S2}, \boldsymbol{x}_{R2}; \boldsymbol{y}_{D2}), \\
& t_1, t_2 \geq 0, \; t_1 + t_2 = 1, \\
& \boldsymbol{u} \leftrightarrow \boldsymbol{x}_{S1} \leftrightarrow (\boldsymbol{y}_{D1}, \boldsymbol{y}_{R1}), \quad (\boldsymbol{C}_{S1}, \boldsymbol{C}_{S2}, \boldsymbol{C}_{R2}) \in \mathcal{P},
\end{aligned} \tag{6.72}
$$

where the set $\mathcal{P}$ is again given by (6.8) for the per-phase power constraint or by (6.9) for the average power constraint.

Recall that the half-duplex PDF strategy generalizes the half-duplex DF strategy in the sense that the relay need not decode the entire message the source transmits during the relay receive phase. That is, the source may further split w_1 into two independent parts w_1' and w_1'' and send the second one w_1'' directly to the destination via the source-to-destination link. The transmit signal of the source in the relay receive phase can thus be expressed as the superposition of $\boldsymbol{u}(w_1')$ and $\boldsymbol{v}(w_1'')$ given by

$$\boldsymbol{x}_{S1} = \boldsymbol{u} + \boldsymbol{v}, \tag{6.73}$$

where $\boldsymbol{u}$ contains the part of the information *the relay must decode* and $\boldsymbol{v}$ represents the information *not to be decoded by the relay*, i.e., the part of w_1 that is conveyed to the destination over the direct link only. Note that since w_1'' is not supposed to be decoded by the relay, $\boldsymbol{v}$ acts as interference at the relay as illustrated in Figure 6.3. Furthermore, we remark that the transmit signal of the source in the relay transmit phase can again be expressed as the superposition of $\boldsymbol{q}(w_2)$ and $\boldsymbol{x}_{R2}(w_1')$, cf. (6.71), with the only difference being that $\boldsymbol{x}_{R2}$ is a function of w_1' instead of w_1.

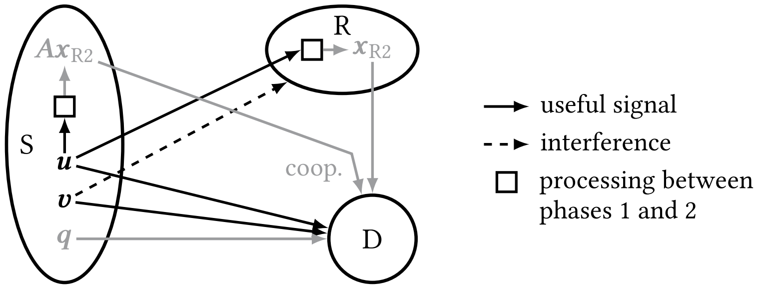


Figure 6.3: Decomposition of the Source Transmit Signals for the Half-Duplex PDF Strategy (signals for relay transmit phase in gray)

However, it is not yet clear that the source inputs $\boldsymbol{x}_{S1}$ and $\boldsymbol{x}_{S2}$ can be represented as the superpositions of independent signal parts for the PDF strategy. In particular, independence of $\boldsymbol{q}$ and $\boldsymbol{x}_{R2}$ for the DF strategy could be assumed because the optimal source and relay inputs are jointly Gaussian, cf. Theorem 6.5. For the PDF strategy, on the other hand, the optimality of Gaussian channel inputs still needs to be proved. Since we cannot simply invoke the entropy maximizing property of the zero-mean proper complex Gaussian distribution for this purpose, we proceed like for the full-duplex case. That is, we first prove that jointly Gaussian source and relay inputs maximize the achievable PDF rate for the *aligned* half-duplex Gaussian MIMO relay channel, and subsequently, we use a limiting argument to extend this result from the aligned to the general half-duplex Gaussian MIMO relay channel.

6.4.1 Aligned Half-Duplex Gaussian MIMO Relay Channel

Definition 6.1. The half-duplex Gaussian MIMO relay channel is said to be *aligned* if $N_S = N_R = N_D = N$ and $\boldsymbol{H}_{SR} = \boldsymbol{H}_{SD} = \mathbf{I}_N$.

If we consider the aligned half-duplex Gaussian MIMO relay channel, the general half-duplex relay channel model specified in (6.1) and (6.2) therefore reduces to

$$\begin{aligned} \boldsymbol{y}_{R1} &= \boldsymbol{x}_{S1} + \boldsymbol{n}_{R1}, & \boldsymbol{n}_{R1} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_R), \\ \boldsymbol{y}_{D1} &= \boldsymbol{x}_{S1} + \boldsymbol{n}_{D1}, & \boldsymbol{n}_{D1} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_D), \\ \boldsymbol{y}_{D2} &= \boldsymbol{x}_{S2} + \boldsymbol{H}_{RD}\boldsymbol{x}_{R2} + \boldsymbol{n}_{D2}, & \boldsymbol{n}_{D2} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_D). \end{aligned} \tag{6.74}$$

For this special case, the theorem below reveals that jointly Gaussian source and relay inputs maximize the achievable PDF rate. We remark that like for the full-duplex case, the proof of this result is based on a *channel enhancement* argument, which goes back to Weingarten et al. [137].

Theorem 6.6. *For the aligned half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint, the maximum achievable partial decode-and-forward (PDF) rate is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. Achievability: Let $\boldsymbol{u} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{U}})$ and $\boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{V}})$ be independent such that $\boldsymbol{x}_{\mathrm{S1}} = \boldsymbol{u} + \boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{S1}})$ with $\boldsymbol{C}_{\mathrm{S1}} = \boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}}$, and let $\left[\begin{smallmatrix} \boldsymbol{x}_{\mathrm{S2}} \\ \boldsymbol{x}_{\mathrm{R2}} \end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_2)$. The mutual information terms that characterize the achievable PDF rate then simplify to

$$
\begin{aligned}
I(\boldsymbol{u}; \boldsymbol{y}_{\mathrm{R1}}) &= \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}_{\mathrm{R}}\right) - \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{R}}\right), \\
I(\boldsymbol{x}_{\mathrm{S1}}; \boldsymbol{y}_{\mathrm{D1}} | \boldsymbol{u}) &= \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right), \\
I(\boldsymbol{x}_{\mathrm{S1}}; \boldsymbol{y}_{\mathrm{D1}}) &= \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right), \\
I(\boldsymbol{x}_{\mathrm{S2}}; \boldsymbol{y}_{\mathrm{D2}} | \boldsymbol{x}_{\mathrm{R2}}) &= \log\det\left(\boldsymbol{C}_{\mathrm{S|R2}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right), \\
I(\boldsymbol{x}_{\mathrm{S2}}, \boldsymbol{x}_{\mathrm{R2}}; \boldsymbol{y}_{\mathrm{D2}}) &= \log\det\left(\boldsymbol{H}\boldsymbol{C}_2\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right),
\end{aligned}
\tag{6.75}
$$

where $\boldsymbol{H} = [\mathbf{I}, \boldsymbol{H}_{\mathrm{RD}}]$ and where $\boldsymbol{C}_{\mathrm{S|R2}}$ again denotes the conditional covariance matrix of $\boldsymbol{x}_{\mathrm{S2}}$ given $\boldsymbol{x}_{\mathrm{R2}}$. It hence follows that

$$
\begin{aligned}
R_{\mathrm{PDF}} \geq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}} = & \max_{R, t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2} R \\
\text{s.t.} \quad & R \leq t_1 \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}_{\mathrm{R}}\right) + t_2 \log\det\left(\boldsymbol{C}_{\mathrm{S|R2}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right) \\
& \qquad + t_1 \max_{\mathbf{0} \preccurlyeq \boldsymbol{C}_{\mathrm{V}} \preccurlyeq \boldsymbol{C}_{\mathrm{S1}}} \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{R}}\right), \\
& R \leq t_1 \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}_{\mathrm{D}}\right) + t_2 \log\det\left(\boldsymbol{H}\boldsymbol{C}_2\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right), \\
& t_1, t_2 \geq 0, \ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2 \succcurlyeq \mathbf{0}, \quad (\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P},
\end{aligned}
\tag{6.76}
$$

where $R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}$ is achievable by means of the PDF strategy and proper complex Gaussian channel inputs. Note that the inner maximization problem

$$
\max_{\boldsymbol{C}_{\mathrm{V}}} \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{C}_{\mathrm{V}} + \boldsymbol{Z}_{\mathrm{R}}\right) \quad \text{s.t.} \quad \mathbf{0} \preccurlyeq \boldsymbol{C}_{\mathrm{V}} \preccurlyeq \boldsymbol{C}_{\mathrm{S1}} \tag{6.77}
$$

is identical to the optimization problem given in (3.35). Consequently, we can follow the steps in the proof of Theorem 3.6 to show that the optimal value of this inner problem is equal to

$$
\begin{aligned}
&\log\det(\boldsymbol{C}_{\mathrm{V}}^{\star} + \boldsymbol{Z}_{\mathrm{D}}) - \log\det(\boldsymbol{C}_{\mathrm{V}}^{\star} + \boldsymbol{Z}_{\mathrm{R}}) = \\
&\qquad \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}\right) - \log\det\left(\boldsymbol{Z}\right) - \Big(\log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}_{\mathrm{R}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right)\Big)
\end{aligned}
\tag{6.78}
$$

for some $\boldsymbol{Z} \succ \mathbf{0}$ such that $\boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{R}}$ and $\boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{D}}$. Using this result, it is straightforward to verify that

$$
\begin{aligned}
R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}} = & \max_{R, t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2} R \\
\text{s.t.} \quad & R \leq t_1 \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}\right) - t_1 \log\det\left(\boldsymbol{Z}\right) \\
& \qquad + t_2 \log\det\left(\boldsymbol{C}_{\mathrm{S|R2}} + \boldsymbol{Z}_{\mathrm{D}}\right) - t_2 \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right) \\
& R \leq t_1 \log\det\left(\boldsymbol{C}_{\mathrm{S1}} + \boldsymbol{Z}_{\mathrm{D}}\right) + t_2 \log\det\left(\boldsymbol{H}\boldsymbol{C}_2\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}\right) - \log\det\left(\boldsymbol{Z}_{\mathrm{D}}\right), \\
& t_1, t_2 \geq 0, \ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_2 \succcurlyeq \mathbf{0}, \quad (\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P}.
\end{aligned}
\tag{6.79}
$$

Converse: Since $\mathbf{0} \prec \boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{R}}$, we can use $\boldsymbol{Z}$ to define an *enhanced* aligned half-duplex Gaussian MIMO relay channel. In particular, if $\tilde{\boldsymbol{Z}}_{\mathrm{R}} = \boldsymbol{Z}$ and

$$\begin{aligned} \tilde{\boldsymbol{y}}_{\mathrm{R}1} &= \boldsymbol{x}_{\mathrm{S}1} + \tilde{\boldsymbol{n}}_{\mathrm{R}1}, & \tilde{\boldsymbol{n}}_{\mathrm{R}1} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tilde{\boldsymbol{Z}}_{\mathrm{R}}), \\ \boldsymbol{y}_{\mathrm{D}1} &= \boldsymbol{x}_{\mathrm{S}1} + \boldsymbol{n}_{\mathrm{D}1}, & \boldsymbol{n}_{\mathrm{D}1} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_{\mathrm{D}}), \\ \boldsymbol{y}_{\mathrm{D}2} &= \boldsymbol{x}_{\mathrm{S}2} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}2} + \boldsymbol{n}_{\mathrm{D}2}, & \boldsymbol{n}_{\mathrm{D}2} &\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_{\mathrm{D}}), \end{aligned} \tag{6.80}$$

$\boldsymbol{y}_{\mathrm{R}1}$ is a stochastically degraded version of $\tilde{\boldsymbol{y}}_{\mathrm{R}1}$, which implies that $I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}1}) \le I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}1})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}1})$, cf. Appendix B.2.2. As a result, we have $R_{\mathrm{PDF}} \le \tilde{R}_{\mathrm{PDF}}$ with

$$\begin{aligned} \tilde{R}_{\mathrm{PDF}} = \max_{\substack{R,t_1,t_2 \\ p(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}1})p(\boldsymbol{x}_{\mathrm{S}2},\boldsymbol{x}_{\mathrm{R}2})}} R \quad \text{s.t.} \quad & R \le t_1 I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}1}) + t_1 I(\boldsymbol{x}_{\mathrm{S}1};\boldsymbol{y}_{\mathrm{D}1}|\boldsymbol{u}) \\ & \qquad + t_2 I(\boldsymbol{x}_{\mathrm{S}2};\boldsymbol{y}_{\mathrm{D}2}|\boldsymbol{x}_{\mathrm{R}2}), \\ & R \le t_1 I(\boldsymbol{x}_{\mathrm{S}1};\boldsymbol{y}_{\mathrm{D}1}) + t_2 I(\boldsymbol{x}_{\mathrm{S}2},\boldsymbol{x}_{\mathrm{R}2};\boldsymbol{y}_{\mathrm{D}2}), \\ & t_1, t_2 \ge 0,\ t_1 + t_2 = 1, \\ & \boldsymbol{u} \leftrightarrow \boldsymbol{x}_{\mathrm{S}1} \leftrightarrow (\boldsymbol{y}_{\mathrm{D}1}, \tilde{\boldsymbol{y}}_{\mathrm{R}1}), \quad (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_{\mathrm{S}2}, \boldsymbol{C}_{\mathrm{R}2}) \in \mathcal{P}. \end{aligned} \tag{6.81}$$

Moreover, note that because $\boldsymbol{Z} \preccurlyeq \boldsymbol{Z}_{\mathrm{D}}$, $\boldsymbol{y}_{\mathrm{D}1}$ is a stochastically degraded version of $\tilde{\boldsymbol{y}}_{\mathrm{R}1}$ as well, which means that

$$I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}1}) + I(\boldsymbol{x}_{\mathrm{S}1};\boldsymbol{y}_{\mathrm{D}1}|\boldsymbol{u}) \le I(\boldsymbol{u};\tilde{\boldsymbol{y}}_{\mathrm{R}1}) + I(\boldsymbol{x}_{\mathrm{S}1};\tilde{\boldsymbol{y}}_{\mathrm{R}1}|\boldsymbol{u}) = I(\boldsymbol{x}_{\mathrm{S}1};\tilde{\boldsymbol{y}}_{\mathrm{R}1}). \tag{6.82}$$

The optimal PDF strategy for the enhanced half-duplex relay channel defined in (6.80) is therefore equivalent to DF, i.e., $\tilde{R}_{\mathrm{PDF}} = \tilde{R}_{\mathrm{DF}}$ with

$$\begin{aligned} \tilde{R}_{\mathrm{DF}} = \max_{\substack{R,t_1,t_2 \\ p(\boldsymbol{x}_{\mathrm{S}1})p(\boldsymbol{x}_{\mathrm{S}2},\boldsymbol{x}_{\mathrm{R}2})}} R \quad \text{s.t.} \quad & R \le t_1 I(\boldsymbol{x}_{\mathrm{S}1};\tilde{\boldsymbol{y}}_{\mathrm{R}1}) + t_2 I(\boldsymbol{x}_{\mathrm{S}2};\boldsymbol{y}_{\mathrm{D}2}|\boldsymbol{x}_{\mathrm{R}2}), \\ & R \le t_1 I(\boldsymbol{x}_{\mathrm{S}1};\boldsymbol{y}_{\mathrm{D}1}) + t_2 I(\boldsymbol{x}_{\mathrm{S}2},\boldsymbol{x}_{\mathrm{R}2};\boldsymbol{y}_{\mathrm{D}2}), \\ & t_1, t_2 \ge 0,\ t_1 + t_2 = 1, \\ & (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_{\mathrm{S}2}, \boldsymbol{C}_{\mathrm{R}2}) \in \mathcal{P}. \end{aligned} \tag{6.83}$$

However, the maximum achievable DF rate for any half-duplex Gaussian MIMO relay channel is attained by proper complex Gaussian channel inputs, cf. Theorem 6.5. As a consequence, $\tilde{R}_{\mathrm{PDF}}$ is attained by $\boldsymbol{u} = \boldsymbol{x}_{\mathrm{S}1} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{S}1})$ and $\left[\begin{smallmatrix}\boldsymbol{x}_{\mathrm{S}2} \\ \boldsymbol{x}_{\mathrm{R}2}\end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_2)$, i.e.,

$$\begin{aligned} \tilde{R}_{\mathrm{PDF}} = & \max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{S}1},\boldsymbol{C}_2} R \\ \text{s.t.} \quad & R \le t_1 \log\det(\boldsymbol{C}_{\mathrm{S}1} + \boldsymbol{Z}) - t_1 \log\det(\boldsymbol{Z}) \\ & \qquad + t_2 \log\det(\boldsymbol{C}_{\mathrm{S}|\mathrm{R}2} + \boldsymbol{Z}_{\mathrm{D}}) - t_2 \log\det(\boldsymbol{Z}_{\mathrm{D}}) \\ & R \le t_1 \log\det(\boldsymbol{C}_{\mathrm{S}1} + \boldsymbol{Z}_{\mathrm{D}}) + t_2 \log\det(\boldsymbol{H}\boldsymbol{C}_2\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{Z}_{\mathrm{D}}) - \log\det(\boldsymbol{Z}_{\mathrm{D}}), \\ & t_1, t_2 \ge 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{C}_2 \succcurlyeq \mathbf{0}, \quad (\boldsymbol{C}_{\mathrm{S}1}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P}. \end{aligned} \tag{6.84}$$

The optimization problems given in (6.79) and (6.84) are obviously the same so that $\tilde{R}_{\mathrm{PDF}} = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}$. But since $\tilde{R}_{\mathrm{PDF}} \ge R_{\mathrm{PDF}} \ge R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}$ in general, this implies $R_{\mathrm{PDF}} = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}$. □

We remark that like for Theorem 3.6, the key to proving Theorem 6.6 is the inner maximization problem, which up to an additive constant is mathematically equivalent to the optimization problem that yields the secrecy capacity of the aligned Gaussian MIMO wiretap channel under a shaping constraint, cf. [84, Section II-A]. In particular, recall that both the optimal value of this inner problem and the enhanced channel we require for the converse follow from considerations similar to those in the proof of [84, Theorem 2]. Furthermore, note that like for the full-duplex case, the enhanced half-duplex relay channel used in the converse belongs to the class of *stochastically degraded half-duplex relay channels*, for which the optimal PDF strategy always reduces to DF as shown in Section 6.4.4.

6.4.2 General Half-Duplex Gaussian MIMO Relay Channel

We now generalize the result that the maximum achievable PDF rate is attained by jointly proper complex Gaussian source and relay inputs. The main idea for extending the proof from the aligned to the general half-duplex Gaussian MIMO relay channel is exactly the same as for the full-duplex case, cf. Section 3.4.2.

Theorem 6.7. *For the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint, the maximum achievable partial decode-and-forward (PDF) rate is attained by jointly proper complex Gaussian source and relay inputs.*

Proof. Without loss of generality, we may assume that $\boldsymbol{H}_{\text{SR}}, \boldsymbol{H}_{\text{SD}}, \boldsymbol{H}_{\text{RD}} \in \mathbb{C}^{N \times N}$ with $N = \max\{N_{\text{S}}, N_{\text{R}}, N_{\text{D}}\}$. If this is not the case, we can augment the matrices with zeros to obtain an equivalent half-duplex Gaussian MIMO relay channel with square $N \times N$ channel gain matrices that preserves the achievable PDF rate under the same power constraints. Beyond that, we may also assume the additive Gaussian noise to be white again, i.e., $\boldsymbol{Z}_{\text{R}} = \boldsymbol{Z}_{\text{D}} = \mathbf{I}_N$.

Achievability: Let $\boldsymbol{u} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\text{U}})$ and $\boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\text{V}})$ be independent such that $\boldsymbol{x}_{\text{S1}} = \boldsymbol{u} + \boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\text{S1}})$ with $\boldsymbol{C}_{\text{S1}} = \boldsymbol{C}_{\text{U}} + \boldsymbol{C}_{\text{V}}$, and let $\left[\begin{smallmatrix} \boldsymbol{x}_{\text{S2}} \\ \boldsymbol{x}_{\text{R2}} \end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_2)$. Then,

$$
\begin{aligned}
R_{\text{PDF}}^{\mathcal{N}_{\text{C}}} = \max_{t_1, t_2, \boldsymbol{C}_{\text{S1}}, \boldsymbol{C}_{\text{V}}, \boldsymbol{C}_2} \; & R_{\text{PDF}}^{\mathcal{N}_{\text{C}}}(t_1, t_2, \boldsymbol{C}_{\text{S1}}, \boldsymbol{C}_{\text{V}}, \boldsymbol{C}_2) \\
\text{s.t.} \quad & t_1, t_2 \geq 0, \; t_1 + t_2 = 1, \qquad \mathbf{0} \preccurlyeq \boldsymbol{C}_{\text{V}} \preccurlyeq \boldsymbol{C}_{\text{S1}}, \quad \boldsymbol{C}_2 \succcurlyeq \mathbf{0}, \\
& (\boldsymbol{C}_{\text{S1}}, \boldsymbol{D}_{\text{S}} \boldsymbol{C}_2 \boldsymbol{D}_{\text{S}}^{\text{H}}, \boldsymbol{D}_{\text{R}} \boldsymbol{C}_2 \boldsymbol{D}_{\text{R}}^{\text{H}}) \in \mathcal{P}
\end{aligned}
\tag{6.85}
$$

is achievable with jointly proper complex Gaussian source and relay inputs, where

$$
\begin{aligned}
R_{\text{PDF}}^{\mathcal{N}_{\text{C}}}(t_1, t_2, \boldsymbol{C}_{\text{S1}}, \boldsymbol{C}_{\text{V}}, \boldsymbol{C}_2) = \min\Big\{ & \\
& t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SR}} \boldsymbol{C}_{\text{S1}} \boldsymbol{H}_{\text{SR}}^{\text{H}}\right) - t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SR}} \boldsymbol{C}_{\text{V}} \boldsymbol{H}_{\text{SR}}^{\text{H}}\right) \\
& \quad + t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SD}} \boldsymbol{C}_{\text{V}} \boldsymbol{H}_{\text{SD}}^{\text{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SD}} \boldsymbol{C}_{\text{S|R2}} \boldsymbol{H}_{\text{SD}}^{\text{H}}\right), \\
& t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{SD}} \boldsymbol{C}_{\text{S1}} \boldsymbol{H}_{\text{SD}}^{\text{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\text{SR}\}\text{D}} \boldsymbol{C}_2 \boldsymbol{H}_{\{\text{SR}\}\text{D}}^{\text{H}}\right)\Big\}.
\end{aligned}
\tag{6.86}
$$

Converse: Suppose the singular value decompositions (SVDs) of the channel gain matrices $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ are given by

$$\boldsymbol{H}_{\mathrm{SR}} = \boldsymbol{U}_{\mathrm{SR}}\boldsymbol{\Sigma}_{\mathrm{SR}}\boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}}, \qquad \boldsymbol{H}_{\mathrm{SD}} = \boldsymbol{U}_{\mathrm{SD}}\boldsymbol{\Sigma}_{\mathrm{SD}}\boldsymbol{V}_{\mathrm{SD}}^{\mathrm{H}}, \tag{6.87}$$

where $\boldsymbol{U}_{\mathrm{SR}}, \boldsymbol{U}_{\mathrm{SD}}, \boldsymbol{V}_{\mathrm{SR}}, \boldsymbol{V}_{\mathrm{SD}} \in \mathbb{C}^{N\times N}$ are the unitary matrices containing the singular vectors, and where the diagonal matrices $\boldsymbol{\Sigma}_{\mathrm{SR}}, \boldsymbol{\Sigma}_{\mathrm{SD}} \in \mathbb{R}_{+}^{N\times N}$ contain the singular values of $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$, respectively. For some $\varepsilon > 0$, let

$$\bar{\boldsymbol{H}}_{\mathrm{SR}} = \boldsymbol{U}_{\mathrm{SR}}(\boldsymbol{\Sigma}_{\mathrm{SR}} + \varepsilon\mathbf{I})\boldsymbol{V}_{\mathrm{SR}}^{\mathrm{H}}, \qquad \bar{\boldsymbol{H}}_{\mathrm{SD}} = \boldsymbol{U}_{\mathrm{SD}}(\boldsymbol{\Sigma}_{\mathrm{SD}} + \varepsilon\mathbf{I})\boldsymbol{V}_{\mathrm{SD}}^{\mathrm{H}}, \tag{6.88}$$

and consider the following enhanced half-duplex Gaussian MIMO relay channel:

$$\begin{aligned}
\bar{\boldsymbol{y}}_{\mathrm{R1}} &= \bar{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{R1}}, && \boldsymbol{n}_{\mathrm{R1}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N),\\
\bar{\boldsymbol{y}}_{\mathrm{D1}} &= \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{D1}}, && \boldsymbol{n}_{\mathrm{D1}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N),\\
\bar{\boldsymbol{y}}_{\mathrm{D2}} &= \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S2}} + \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R2}} + \boldsymbol{n}_{\mathrm{D2}}, && \boldsymbol{n}_{\mathrm{D2}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N).
\end{aligned} \tag{6.89}$$

Note that since $\bar{\boldsymbol{H}}_{\mathrm{SR}}$ and $\bar{\boldsymbol{H}}_{\mathrm{SD}}$ are invertible, this half-duplex relay channel is equivalent to an aligned half-duplex Gaussian MIMO relay channel with $\boldsymbol{Z}_{\mathrm{R}} = (\bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SR}})^{-1}$ and $\boldsymbol{Z}_{\mathrm{D}} = (\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SD}})^{-1}$, for which we know from Theorem 6.6 that the achievable PDF rate is maximized by proper complex Gaussian channel inputs. Moreover, $\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}} \preccurlyeq \bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SD}} \preccurlyeq \bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\bar{\boldsymbol{H}}_{\mathrm{SD}}$, which means that $\boldsymbol{y}_{\mathrm{R1}}$ and $\boldsymbol{y}_{\mathrm{D1}}$ are stochastically degraded versions of $\bar{\boldsymbol{y}}_{\mathrm{R1}}$ and $\bar{\boldsymbol{y}}_{\mathrm{D1}}$, respectively, cf. Appendix B.2.2. As a result, it follows that

$$\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}} = \bar{R}_{\mathrm{PDF}} \geq R_{\mathrm{PDF}} \geq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}, \tag{6.90}$$

where

$$\begin{aligned}
\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}} = \max_{t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2}\ & \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}(t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2)\\
\text{s.t.}\quad & t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \qquad \mathbf{0} \preccurlyeq \boldsymbol{C}_{\mathrm{V}} \preccurlyeq \boldsymbol{C}_{\mathrm{S1}}, \quad \boldsymbol{C}_2 \succcurlyeq \mathbf{0},\\
& (\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P}
\end{aligned} \tag{6.91}$$

is the maximum achievable PDF rate for the enhanced half-duplex Gaussian MIMO relay channel defined in (6.89),

$$\begin{aligned}
\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}(t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2) = \min\Big\{ &\\
& t_1 \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{S1}}\bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\right) - t_1 \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\bar{\boldsymbol{H}}_{\mathrm{SR}}^{\mathrm{H}}\right)\\
& \quad + t_1 \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S|R2}}\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\right),\\
& t_1 \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{S1}}\bar{\boldsymbol{H}}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \bar{\boldsymbol{H}}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\bar{\boldsymbol{H}}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right)\Big\},
\end{aligned} \tag{6.92}$$

and $\bar{\boldsymbol{H}}_{\{\mathrm{SR}\}\mathrm{D}} = [\bar{\boldsymbol{H}}_{\mathrm{SD}}, \boldsymbol{H}_{\mathrm{RD}}]$. Like for the full-duplex case, we can therefore complete the proof by showing that $\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}} \to R_{\mathrm{PDF}}^{\mathcal{N}_{\mathbb{C}}}$ as $\varepsilon \to 0$.

To this end, suppose t_1, t_2 and $\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2$ are fixed. Then, $\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2)$ is a continuous function of ε since it is the pointwise minimum of two functions that are continuous in ε. As a consequence, we have

$$\lim_{\varepsilon \to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2) = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(t_1, t_2, \boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2). \tag{6.93}$$

Because (6.93) holds for any t_1, t_2 and $\boldsymbol{C}_{\mathrm{S1}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2 \succcurlyeq \mathbf{0}$, it also holds for the maximizers $t_1^\star, t_2^\star, \boldsymbol{C}_{\mathrm{S1}}^\star, \boldsymbol{C}_{\mathrm{V}}^\star, \boldsymbol{C}_2^\star$ of the problem given in (6.91). In addition, these maximizers also satisfy the constraints of (6.85), which means that

$$\lim_{\varepsilon \to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}} = \lim_{\varepsilon \to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(t_1^\star, t_2^\star, \boldsymbol{C}_{\mathrm{S1}}^\star, \boldsymbol{C}_{\mathrm{V}}^\star, \boldsymbol{C}_2^\star) = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}(t_1^\star, t_2^\star, \boldsymbol{C}_{\mathrm{S1}}^\star, \boldsymbol{C}_{\mathrm{V}}^\star, \boldsymbol{C}_2^\star) \leq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}. \tag{6.94}$$

But since $\bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}} \geq R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}$ in general, it follows that $\lim_{\varepsilon \to 0} \bar{R}_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}} = R_{\mathrm{PDF}}^{\mathcal{N}_{\mathrm{C}}}$. □

Because the maximum achievable PDF rate is attained by jointly Gaussian source and relay inputs, $\boldsymbol{x}_{\mathrm{S1}}$ and $\boldsymbol{x}_{\mathrm{S2}}$ can indeed be decomposed as $\boldsymbol{x}_{\mathrm{S1}} = \boldsymbol{u} + \boldsymbol{v}$ and $\boldsymbol{x}_{\mathrm{S2}} = \boldsymbol{q} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{R2}}$ with $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{q}$, and $\boldsymbol{x}_{\mathrm{R2}}$ being independent. However, like for the DF strategy (as well as the full-duplex case), it is not convenient to express the correlation of $\boldsymbol{x}_{\mathrm{S2}}$ and $\boldsymbol{x}_{\mathrm{R2}}$ by means of $\boldsymbol{A}$ and $\boldsymbol{C}_{\mathrm{R2}}$ if one actually wants to evaluate R_{PDF}. The reason for this is that the PDF rate maximization problem would then contain the product $\boldsymbol{A}\boldsymbol{C}_{\mathrm{R2}}\boldsymbol{A}^{\mathrm{H}}$ of two optimization variables.

In order to avoid this issue, we represent the optimal source and relay inputs as in the proofs of Theorems 6.6 and 6.7. In particular, let $\boldsymbol{u} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{U}})$ and $\boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{V}})$ be independent such that $\boldsymbol{x}_{\mathrm{S1}} = \boldsymbol{u} + \boldsymbol{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})$, and let $\left[\begin{smallmatrix} \boldsymbol{x}_{\mathrm{S2}} \\ \boldsymbol{x}_{\mathrm{R2}} \end{smallmatrix}\right] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_2)$. Then, the optimization problem that specifies the maximum achievable PDF rate for the half-duplex Gaussian MIMO relay channel can be expressed as

$$\begin{aligned} R_{\mathrm{PDF}} = & \max_{R, t_1, t_2, \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2, \boldsymbol{C}_{\mathrm{Q}}} R \\ \text{s.t.} \quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ & \qquad + t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ & t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ & (\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P} \end{aligned} \tag{6.95}$$

with the set $\mathcal{P}$ given by (6.8) or (6.9) depending on the considered power constraint. We remark that to obtain this formulation of the PDF rate maximization problem, we have again reformulated the term involving the conditional covariance matrix $\boldsymbol{C}_{\mathrm{S|R2}}$ by introducing the auxiliary variable $\boldsymbol{C}_{\mathrm{Q}} = \boldsymbol{C}_{\mathrm{S|R2}} \succcurlyeq \mathbf{0}$, relaxing the equality constraint to $\mathbf{0} \preccurlyeq \boldsymbol{C}_{\mathrm{Q}} \preccurlyeq \boldsymbol{C}_{\mathrm{S|R2}}$, and applying Lemma 3.2. Furthermore, note that we have chosen $\boldsymbol{C}_{\mathrm{U}}$ and $\boldsymbol{C}_{\mathrm{V}}$ as optimization variables in (6.95) (instead of $\boldsymbol{C}_{\mathrm{S1}}$ and $\boldsymbol{C}_{\mathrm{V}}$) since this is more convenient for the following considerations.

The PDF rate maximization problem reduces to the DF rate maximization problem given in (6.58) if $\boldsymbol{C}_{\mathrm{V}} = \mathbf{0}$, i.e., if $\boldsymbol{u} = \boldsymbol{x}_{\mathrm{S1}}$. However, note that even for fixed time-shares $t_1, t_2 \geq 0$, the general PDF rate maximization problem is nonconvex due to the term $-t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)$ in the first rate constraint, which results from the fact that $\boldsymbol{v}$ must be considered as interference at the relay, cf. Figure 6.3. Like for the full-duplex case, this means that evaluating the maximum achievable PDF rate for the half-duplex Gaussian MIMO relay channel would require to solve a difficult nonconvex optimization problem (for both the per-phase and the average power constraint).

6.4.3 Suboptimal Partial Decode-and-Forward Rates

Instead of trying to evaluate R_{PDF} for the general case, we thus focus on approaches that yield suboptimal solutions to the PDF rate maximization problem. In particular, we again apply the *zero-forcing* (ZF) PDF approach and the *inner approximation algorithm* (IAA), which were already considered in Section 3.4.3, to the problem given in (6.95).

6.4.3.1 Zero-Forcing PDF Approach

Recall that the ZF approach is based on suppressing the interference the relay would suffer from, i.e., the part of the source input the relay is not supposed to decode. To this end, we introduce a ZF receive filter $\boldsymbol{G}$ at the relay and require that all possible realizations of $\boldsymbol{v}$ satisfy $\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{v} = \mathbf{0}$, which is equivalent to

$$\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} = \mathbf{0} \tag{6.96}$$

if $\boldsymbol{v}$ is a zero-mean proper complex Gaussian random vector. By adding the relay receive filter and the corresponding ZF constraint to (6.95), we obtain

$$\begin{aligned}
\tilde{R}_{\mathrm{ZF}} = &\max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{U}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_2,\boldsymbol{C}_{\mathrm{Q}}} R \\
\text{s.t.}\quad & R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{U}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) + t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) \\
&\qquad + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\
& R \leq t_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\
& t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \qquad \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\
& (\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P}, \quad \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} = \mathbf{0},
\end{aligned} \tag{6.97}$$

where, without loss of generality, we have assumed that the ZF receive filter is chosen such that $\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}} = \mathbf{I}_M$ for some $M \in \{1, \ldots, N_{\mathrm{R}}\}$. Since the ZF condition is linear in $\boldsymbol{C}_{\mathrm{V}}$, this problem is convex in $(R, \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}}, \boldsymbol{C}_2, \boldsymbol{C}_{\mathrm{Q}})$ for given time-shares $t_1, t_2 \geq 0$. For any fixed ZF filter $\boldsymbol{G}$, the best rates that can be achieved using the ZF PDF approach can therefore again be determined by means of the dual decomposition approaches that were proposed in Sections 6.2.1 and 6.2.2.

In particular, if the source and the relay are subject to the per-phase power constraint, we define

$$\begin{aligned}\mathcal{S}_{\mathrm{ZF1}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{U}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) \\ &\quad + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}} &\succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}}) \leq P_{\mathrm{S}}, \quad \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} = \mathbf{0} \Big\}\end{aligned} \tag{6.98}$$

and

$$\begin{aligned}\mathcal{S}_{\mathrm{ZF2}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ \boldsymbol{C}_{\mathrm{Q}} &\succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) &\leq P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}} \Big\},\end{aligned} \tag{6.99}$$

where the compact convex rate regions $\mathcal{S}_{\mathrm{ZF1}} \subseteq \mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{ZF2}} = \mathcal{S}_{\mathrm{DF2}} = \mathcal{S}_{\mathrm{CSB2}}$ specify the contributions of the relay receive and transmit phases to $\tilde{R}_{\mathrm{ZF}}$, respectively. Furthermore, let

$$\begin{aligned}\mathcal{S}_{\mathrm{ZF}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : \boldsymbol{s} &= t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \\ \boldsymbol{s}_1 &\in \mathcal{S}_{\mathrm{ZF1}},\ \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{ZF2}}, \quad t_1, t_2 \geq 0,\ t_1 + t_2 = 1 \Big\},\end{aligned} \tag{6.100}$$

then for the per-phase power constraint, the optimization problem given in (6.97) can be rewritten as

$$\tilde{R}_{\mathrm{ZF}} = \max_{R, \boldsymbol{s}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{ZF}}. \tag{6.101}$$

Since $\mathcal{S}_{\mathrm{ZF}}$ is a convex set with nonempty interior, this is a convex optimization problem which can equivalently be solved in the Lagrangian dual domain. For this purpose, we can again apply the dual decomposition approach presented in Section 6.2.1, where we simply need to replace $\mathcal{S}_{\mathrm{CSB}}$ with $\mathcal{S}_{\mathrm{ZF}}$, or equivalently, replace $\mathcal{S}_{\mathrm{CSB1}}$ with $\mathcal{S}_{\mathrm{ZF1}}$ as $\mathcal{S}_{\mathrm{ZF2}} = \mathcal{S}_{\mathrm{CSB2}}$.

Similarly, if the source and the relay are subject to the average power constraint, we define the two closed convex rate-power regions

$$\begin{aligned}\mathcal{Z}_{\mathrm{ZF1}} = \Big\{ (\boldsymbol{s}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{U}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) \\ &\quad + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ p_1 &= \mathrm{tr}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}}),\ p_2 = 0, \\ \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}} &\succcurlyeq \mathbf{0}, \quad \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} = \mathbf{0} \Big\}\end{aligned} \tag{6.102}$$

and

$$\begin{aligned}\mathcal{Z}_{\mathrm{ZF2}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ p_1 &= \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}),\ p_2 = \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}), \\ \boldsymbol{C}_{\mathrm{Q}} &\succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}\Big\},\end{aligned} \tag{6.103}$$

where $\mathcal{Z}_{\mathrm{ZF1}} \subseteq \mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{ZF2}} = \mathcal{Z}_{\mathrm{DF2}} = \mathcal{Z}_{\mathrm{CSB2}}$. If we further define

$$\begin{aligned}\mathcal{Z}_{\mathrm{ZF}} = \Big\{(\boldsymbol{s},\boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : \boldsymbol{s} = t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \quad \boldsymbol{p} = t_1\boldsymbol{p}_1 + t_2\boldsymbol{p}_2, \\ (\boldsymbol{s}_1,\boldsymbol{p}_1) \in \mathcal{Z}_{\mathrm{ZF1}},\ (\boldsymbol{s}_2,\boldsymbol{p}_2) \in \mathcal{Z}_{\mathrm{ZF2}}, \quad t_1, t_2 \geq 0,\ t_1 + t_2 = 1\Big\},\end{aligned} \tag{6.104}$$

which is also closed and convex, we can rewrite (6.97) as

$$\tilde{R}_{\mathrm{ZF,av}} = \max_{R,\boldsymbol{s},\boldsymbol{p}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{p} \leq \boldsymbol{p}_{\mathrm{tx}}, \quad (\boldsymbol{s},\boldsymbol{p}) \in \mathcal{Z}_{\mathrm{ZF}} \tag{6.105}$$

for the average power constraint. This problem is convex again, and because the interior of $\mathcal{Z}_{\mathrm{ZF}}$ is nonempty, it can be solved in the Lagrangian dual domain. To this end, we can apply the dual decomposition approach presented in Section 6.2.2 if $\mathcal{Z}_{\mathrm{CSB}}$ is replaced with $\mathcal{Z}_{\mathrm{ZF}}$, or equivalently, if $\mathcal{Z}_{\mathrm{CSB1}}$ is replaced with $\mathcal{Z}_{\mathrm{ZF1}}$ since $\mathcal{Z}_{\mathrm{ZF2}} = \mathcal{Z}_{\mathrm{CSB2}}$.

Because the rate(-power) regions $\mathcal{S}_{\mathrm{ZF1}}/\mathcal{Z}_{\mathrm{ZF1}}$ and $\mathcal{S}_{\mathrm{ZF2}}/\mathcal{Z}_{\mathrm{ZF2}}$ are convex, time-sharing within the relay receive or the relay transmit phase is again not necessary, cf. Remark 6.2. That is, we usually have $t_m^\star = \tau_\ell$ and $\boldsymbol{s}_m^\star = \boldsymbol{s}^{(\ell)} \in \mathcal{S}_{\mathrm{ZF}m}/(\boldsymbol{s}_m^\star, \boldsymbol{p}_m^\star) = (\boldsymbol{s}^{(\ell)}, \boldsymbol{p}^{(\ell)}) \in \mathcal{Z}_{\mathrm{ZF}m}$, $m \in \{1,2\}$, for one particular $\ell \in \{1,\ldots,K\}$ if the cutting-plane algorithm has converged after K iterations. In this case, the optimal primal variables $\boldsymbol{C}_{\mathrm{U}}^\star, \boldsymbol{C}_{\mathrm{V}}^\star, \boldsymbol{C}_2^\star, \boldsymbol{C}_{\mathrm{Q}}^\star$ can easily be recovered as well, and by dropping the ZF filter, we can then conclude that all rates smaller than or equal to

$$\begin{aligned}\min\Big\{&t_1^\star \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{U}}^\star + \boldsymbol{C}_{\mathrm{V}}^\star)\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) - t_1^\star \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^\star\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ &\quad + t_1^\star \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}^\star\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2^\star \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}^\star\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ &t_1^\star \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}}^\star + \boldsymbol{C}_{\mathrm{V}}^\star)\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + t_2^\star \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2^\star\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right)\Big\}\end{aligned} \tag{6.106}$$

are achievable using the ZF PDF approach with relay receive filter $\boldsymbol{G}$. Moreover, if R_{ZF} and $R_{\mathrm{ZF,av}}$ denote the values of (6.106) for the per-phase and the average power constraint, respectively, it follows from the data processing inequality [24, Theorem 2.8.1] that $R_{\mathrm{ZF}} \geq \tilde{R}_{\mathrm{ZF}}$ and $R_{\mathrm{ZF,av}} \geq \tilde{R}_{\mathrm{ZF,av}}$.

Note also that like for the full-duplex case, there is no clear-cut answer from theory as to what is the best choice for the ZF receive filter. In general, we therefore have to use heuristics to choose $\boldsymbol{G}$. Reasonable ZF receive filters can for example be obtained using antenna selection or the SVD of $\boldsymbol{H}_{\mathrm{SR}}$, cf. Section 3.4.3, or from the generalized singular value decomposition (GSVD) of $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ as discussed in Section 4.2.2.

6.4.3.2 Inner Approximation Algorithm

In contrast to the ZF approach, which is specific to the PDF rate maximization problem, the IAA is a general mathematical approach to deal with nonconvex optimization problems. In particular, the idea of the IAA is to solve a sequence of approximating convex problems instead of the original nonconvex one, cf. Appendix A.1. If we want to apply the IAA to the nonconvex PDF rate maximization problem given in (6.95), then like for the full-duplex case, we need to approximate the first rate constraint. In iteration k, we therefore replace $-\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)$ by its first-order Taylor series around $\boldsymbol{C}_{\mathrm{V}}^{(k-1)}$, which is equal to[3]

$$c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)})-\operatorname{tr}\!\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(k-1)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\right), \tag{6.107}$$

cf. (3.91), and where $\boldsymbol{C}_{\mathrm{V}}^{(k-1)}$ is an optimizer of the approximating convex optimization problem in iteration $k-1$. The approximating convex optimization problem that needs to be solved in the k-th iteration of the IAA hence reads as

$$\begin{aligned}
R_{\mathrm{IAA}}^{(k)}=\;&\max_{R,t_1,t_2,\boldsymbol{C}_{\mathrm{U}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_2,\boldsymbol{C}_{\mathrm{Q}}}\; R\\
\text{s.t.}\quad & R\le t_1\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)+t_1\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right)\\
&\quad+t_1c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)})-t_1\operatorname{tr}\!\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(k-1)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\right)\\
&\quad+t_2\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right),\\
& R\le t_1\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right)+t_2\log\det\left(\mathbf{I}+\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right),\\
& t_1,t_2\ge 0,\;\; t_1+t_2=1,\qquad \boldsymbol{C}_{\mathrm{U}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_{\mathrm{Q}}\succcurlyeq\mathbf{0},\quad \boldsymbol{C}_2-\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}}\succcurlyeq\mathbf{0},\\
& (\boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}},\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}},\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}})\in\mathcal{P}
\end{aligned} \tag{6.108}$$

with $\mathcal{P}$ depending on the considered power constraint. Note that if $t_1,t_2\ge 0$ are fixed, this problem is convex in $(R,\boldsymbol{C}_{\mathrm{U}},\boldsymbol{C}_{\mathrm{V}},\boldsymbol{C}_2,\boldsymbol{C}_{\mathrm{Q}})$ for both the per-phase and the average power constraint. Consequently, the optimal solutions can be obtained using the dual decomposition approaches presented in Sections 6.2.1 and 6.2.2.

In particular, suppose the source and the relay are subject to the per-phase power constraint. If we define

$$\begin{aligned}
\mathcal{S}_{\mathrm{IAA1}}^{(k)}=\Big\{\boldsymbol{s}\in\mathbb{R}_+^2:\; & s_1\le\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)\\
&\quad+\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right)+c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)})\\
&\quad-\operatorname{tr}\!\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(k-1)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\right),\\
& s_2\le\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right),\\
& \boldsymbol{C}_{\mathrm{U}},\boldsymbol{C}_{\mathrm{V}}\succcurlyeq\mathbf{0},\quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{U}}+\boldsymbol{C}_{\mathrm{V}})\le P_{\mathrm{S}}\Big\}
\end{aligned} \tag{6.109}$$

[3] We remark that $c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)})$ simply collects all terms of the Taylor series that do not depend on $\boldsymbol{C}_{\mathrm{V}}$, cf. (3.92), and like for the full-duplex case, $\log=\log_{\mathrm{e}}$ is assumed here for simplicity.

and

$$\mathcal{S}_{\mathrm{IAA2}}^{(k)} = \Big\{ s \in \mathbb{R}_+^2 : s_1 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0}, \\ \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \le P_{\mathrm{S}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \le P_{\mathrm{R}} \Big\}, \tag{6.110}$$

the rate regions $\mathcal{S}_{\mathrm{IAA1}}^{(k)} \subseteq \mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{IAA2}}^{(k)} = \mathcal{S}_{\mathrm{DF2}} = \mathcal{S}_{\mathrm{CSB2}}$, $k \in \mathbb{N}$, which specify the contributions of the relay receive and transmit phases to $R_{\mathrm{IAA}}^{(k)}$, respectively, are compact and convex. Moreover, if

$$\mathcal{S}_{\mathrm{IAA}}^{(k)} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^2 : \boldsymbol{s} = t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \\ \boldsymbol{s}_1 \in \mathcal{S}_{\mathrm{IAA1}}^{(k)}, \ \boldsymbol{s}_2 \in \mathcal{S}_{\mathrm{IAA2}}^{(k)}, \quad t_1, t_2 \ge 0, \ t_1 + t_2 = 1 \Big\}, \tag{6.111}$$

then for the per-phase power constraint, the optimization problem given in (6.108) can be rewritten as

$$R_{\mathrm{IAA}}^{(k)} = \max_{R, \boldsymbol{s}} R \quad \text{s.t.} \quad R\mathbf{1} \le \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{IAA}}^{(k)}. \tag{6.112}$$

Similarly, if the source and the relay are subject to the average power constraint, we can define the two closed convex rate-power regions

$$\mathcal{Z}_{\mathrm{IAA1}}^{(k)} = \Big\{ (\boldsymbol{s}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : s_1 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right) \\ + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right) + c(\boldsymbol{C}_{\mathrm{V}}^{(k-1)}) \\ - \mathrm{tr}\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}^{(k-1)}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\right), \\ s_2 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}})\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ p_1 = \mathrm{tr}(\boldsymbol{C}_{\mathrm{U}} + \boldsymbol{C}_{\mathrm{V}}), \ p_2 = 0, \quad \boldsymbol{C}_{\mathrm{U}}, \boldsymbol{C}_{\mathrm{V}} \succcurlyeq \mathbf{0} \Big\} \tag{6.113}$$

and

$$\mathcal{Z}_{\mathrm{IAA2}}^{(k)} = \Big\{ (\boldsymbol{s}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : s_1 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\right), \\ s_2 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}_2\boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}^{\mathrm{H}}\right), \\ p_1 = \mathrm{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}), \ p_2 = \mathrm{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}), \\ \boldsymbol{C}_{\mathrm{Q}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_2 - \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q}}\boldsymbol{D}_{\mathrm{S}} \succcurlyeq \mathbf{0} \Big\}, \tag{6.114}$$

where $\mathcal{Z}_{\mathrm{IAA1}}^{(k)} \subseteq \mathcal{Z}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{IAA2}}^{(k)} = \mathcal{Z}_{\mathrm{DF2}} = \mathcal{Z}_{\mathrm{CSB2}}$ for all $k \in \mathbb{N}$. Furthermore, if

$$\mathcal{Z}_{\mathrm{IAA}}^{(k)} = \Big\{ (\boldsymbol{s}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 : \boldsymbol{s} = t_1\boldsymbol{s}_1 + t_2\boldsymbol{s}_2, \quad \boldsymbol{p} = t_1\boldsymbol{p}_1 + t_2\boldsymbol{p}_2, \\ (\boldsymbol{s}_1, \boldsymbol{p}_1) \in \mathcal{Z}_{\mathrm{IAA1}}^{(k)}, \ (\boldsymbol{s}_2, \boldsymbol{p}_2) \in \mathcal{Z}_{\mathrm{IAA2}}^{(k)}, \quad t_1, t_2 \ge 0, \ t_1 + t_2 = 1 \Big\}, \tag{6.115}$$

then for the average power constraint, we can rewrite (6.108) as

$$R^{(k)}_{\mathrm{IAA,av}} = \max_{R,\boldsymbol{s},\boldsymbol{p}} R \quad \text{s.t.} \quad R\mathbf{1} \leq \boldsymbol{s}, \quad \boldsymbol{p} \leq \boldsymbol{p}_{\mathrm{tx}}, \quad (\boldsymbol{s},\boldsymbol{p}) \in \mathcal{Z}^{(k)}_{\mathrm{IAA}}. \tag{6.116}$$

Since $\mathcal{S}^{(k)}_{\mathrm{IAA}}$ and $\mathcal{Z}^{(k)}_{\mathrm{IAA}}$ are convex sets with nonempty interiors, the problems given in (6.112) and (6.116) are convex and can also be solved in the Lagrangian dual domain. For this purpose, we can again apply the dual decomposition approaches proposed in Sections 6.2.1 and 6.2.2 if $\mathcal{S}_{\mathrm{CSB}}$ and $\mathcal{Z}_{\mathrm{CSB}}$ are replaced with $\mathcal{S}^{(k)}_{\mathrm{IAA}}$ and $\mathcal{Z}^{(k)}_{\mathrm{IAA}}$, respectively. We remark that this is equivalent to replacing $\mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{Z}_{\mathrm{CSB1}}$ with $\mathcal{S}^{(k)}_{\mathrm{IAA1}}$ and $\mathcal{Z}^{(k)}_{\mathrm{IAA1}}$ as $\mathcal{S}^{(k)}_{\mathrm{IAA2}} = \mathcal{S}_{\mathrm{CSB2}}$ and $\mathcal{Z}^{(k)}_{\mathrm{IAA2}} = \mathcal{Z}_{\mathrm{CSB2}}$ for all $k \in \mathbb{N}$.

Note also that the sequences $\left\{R^{(k)}_{\mathrm{IAA}}\right\}$ and $\left\{R^{(k)}_{\mathrm{IAA,av}}\right\}$ are nondecreasing in k, i.e., $R^{(k)}_{\mathrm{IAA}} \geq R^{(k-1)}_{\mathrm{IAA}}$ and $R^{(k)}_{\mathrm{IAA,av}} \geq R^{(k-1)}_{\mathrm{IAA,av}}$ for all $k \in \mathbb{N}$. This is because the optimizers of iteration $k-1$ are also feasible for the approximating convex problems in iteration k. As the sequence elements $R^{(k)}_{\mathrm{IAA}}$ and $R^{(k)}_{\mathrm{IAA,av}}$ are upper bounded by C_{CSB} and $C_{\mathrm{CSB,av}}$, respectively, we can conclude that both sequences converge. In addition, it follows from Theorem A.1 that the IAA stops at a Karush–Kuhn–Tucker (KKT) point of the original nonconvex problem, or the limit of any convergent subsequence is a KKT point. However, since convergence usually occurs only in the limit as $k \to \infty$, we have to define a practical termination criterion. As discussed in Section 3.4.3, this criterion may for example be based on the absolute or the relative rate improvement over the last iteration(s) .

What is more, because the rate(-power) regions $\mathcal{S}^{(k)}_{\mathrm{IAA1}}/\mathcal{Z}^{(k)}_{\mathrm{IAA1}}$ and $\mathcal{S}^{(k)}_{\mathrm{IAA2}}/\mathcal{Z}^{(k)}_{\mathrm{IAA2}}$ are convex for all $k \in \mathbb{N}$, time-sharing within the relay receive and transmit phases is not necessary, cf. Remark 6.2. Like for the ZF PDF approach, this means that we can easily recover all primal optimal variables $t^{(k)}_1, t^{(k)}_2, \boldsymbol{C}^{(k)}_{\mathrm{U}}, \boldsymbol{C}^{(k)}_{\mathrm{V}}, \boldsymbol{C}^{(k)}_2, \boldsymbol{C}^{(k)}_{\mathrm{Q}}$ of the approximating optimization problems. If the IAA has terminated after K iterations, it follows that all rates smaller than or equal to

$$\begin{aligned}\min \Big\{ & t^{(K)}_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}(\boldsymbol{C}^{(K)}_{\mathrm{U}} + \boldsymbol{C}^{(K)}_{\mathrm{V}})\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SR}}\right) - t^{(K)}_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}^{(K)}_{\mathrm{V}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SR}}\right) \\ & + t^{(K)}_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}^{(K)}_{\mathrm{V}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SD}}\right) + t^{(K)}_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}^{(K)}_{\mathrm{Q}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SD}}\right), \\ & t^{(K)}_1 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{SD}}(\boldsymbol{C}^{(K)}_{\mathrm{U}} + \boldsymbol{C}^{(K)}_{\mathrm{V}})\boldsymbol{H}^{\mathrm{H}}_{\mathrm{SD}}\right) + t^{(K)}_2 \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{SR}\}\mathrm{D}}\boldsymbol{C}^{(K)}_2\boldsymbol{H}^{\mathrm{H}}_{\{\mathrm{SR}\}\mathrm{D}}\right)\Big\}\end{aligned} \tag{6.117}$$

are achievable by means of the PDF strategy. Furthermore, if R_{IAA} and $R_{\mathrm{IAA,av}}$ denote the values of (6.117) for the per-phase and the average power constraint, respectively, we have $R_{\mathrm{PDF}} \geq R_{\mathrm{IAA}} \geq R^{(K)}_{\mathrm{IAA}}$ and $R_{\mathrm{PDF,av}} \geq R_{\mathrm{IAA,av}} \geq R^{(K)}_{\mathrm{IAA,av}}$ since the approximated rate constraint is stricter than the original one. Finally, we remark that R_{IAA} and $R_{\mathrm{IAA,av}}$ strongly depend on how the IAA is initialized. However, with appropriate initializations, e.g., $\boldsymbol{C}^{(0)}_{\mathrm{V}} = \mathbf{0}$, we can ensure that $R_{\mathrm{IAA}} \geq R_{\mathrm{DF}}$ and $R_{\mathrm{IAA,av}} \geq R_{\mathrm{DF,av}}$.

6.4.4 Optimal Partial Decode-and-Forward Rates for Special Cases

Recall that for the discrete memoryless half-duplex relay channel, special cases for which the maximum achievable PDF rate can be determined include the degraded half-duplex relay channel, the reversely degraded half-duplex relay channel, the semideterministic half-duplex relay channel, and the half-duplex relay channel with orthogonal sender components. In particular, it was shown in Section 5.4 that the PDF strategy achieves the CSB for these four types of discrete memoryless half-duplex relay channels. Like for the full-duplex case, however, only one of these special cases is also relevant for the half-duplex Gaussian MIMO relay channel considered in this work.

(a) The half-duplex Gaussian MIMO relay channel is never degraded according to the system model defined in Section 6.1. The reason for this is that the Markov chain condition $\boldsymbol{x}_{\mathrm{S1}} \leftrightarrow \boldsymbol{y}_{\mathrm{R1}} \leftrightarrow \boldsymbol{y}_{\mathrm{D1}}$ can never be satisfied if the noise vectors $\boldsymbol{n}_{\mathrm{R1}}$ and $\boldsymbol{n}_{\mathrm{D1}}$ are independent.
(b) Likewise, the half-duplex Gaussian MIMO relay channel is never reversely degraded.
(c) The half-duplex Gaussian MIMO relay channel is never semideterministic according to the system model defined in Section 6.1. This is because the defining condition of the semideterministic half-duplex relay channel requires that $\boldsymbol{y}_{\mathrm{R1}}$ be a function of $\boldsymbol{x}_{\mathrm{S1}}$, which is not satisfied unless there is no noise at the relay.
(d) For the half-duplex relay channel with orthogonal sender components, we can extend the definition to the half-duplex Gaussian MIMO relay channel as follows:

Definition 6.2. The half-duplex Gaussian MIMO relay channel has *orthogonal sender components* if $\mathrm{row}(\boldsymbol{H}_{\mathrm{S\{RD\}}}) = \mathrm{row}(\boldsymbol{H}_{\mathrm{SR}}) \perp \mathrm{row}(\boldsymbol{H}_{\mathrm{SD}})$, i.e., if $\mathrm{row}(\boldsymbol{H}_{\mathrm{S\{RD\}}})$ is the direct sum of the two orthogonal subspaces $\mathrm{row}(\boldsymbol{H}_{\mathrm{SR}})$ and $\mathrm{row}(\boldsymbol{H}_{\mathrm{SD}})$.

Remark 6.6. Like for the full-duplex case, it is assumed here that $\mathcal{X}_{\mathrm{S}} = \mathrm{row}(\boldsymbol{H}_{\mathrm{S\{RD\}}}) = \mathrm{null}^{\perp}(\boldsymbol{H}_{\mathrm{S\{RD\}}})$. This is without loss of generality, however, as $\boldsymbol{y}_{\mathrm{R1}} = \boldsymbol{n}_{\mathrm{R1}}$ and $\boldsymbol{y}_{\mathrm{D1}} = \boldsymbol{n}_{\mathrm{D1}}$ for any $\boldsymbol{x}_{\mathrm{S1}} \in \mathrm{null}(\boldsymbol{H}_{\mathrm{S\{RD\}}})$. Similarly, we may also assume that $\mathcal{X}'_{\mathrm{S}} = \mathrm{row}(\boldsymbol{H}_{\mathrm{SR}})$ and $\mathcal{X}''_{\mathrm{S}} = \mathrm{row}(\boldsymbol{H}_{\mathrm{SD}})$.

According to this definition, it follows that if the half-duplex Gaussian MIMO relay channel has orthogonal sender components, then $p(\boldsymbol{y}_{\mathrm{D1}}, \boldsymbol{y}_{\mathrm{R1}}|\boldsymbol{x}_{\mathrm{S1}}) = p(\boldsymbol{y}_{\mathrm{R1}}|\boldsymbol{x}'_{\mathrm{S1}})p(\boldsymbol{y}_{\mathrm{D1}}|\boldsymbol{x}''_{\mathrm{S1}})$ for all $(\boldsymbol{x}'_{\mathrm{S1}}, \boldsymbol{x}''_{\mathrm{S1}}, \boldsymbol{y}_{\mathrm{D1}}, \boldsymbol{y}_{\mathrm{R1}}) \in \mathcal{X}'_{\mathrm{S}} \times \mathcal{X}''_{\mathrm{S}} \times \mathcal{Y}_{\mathrm{D}} \times \mathcal{Y}_{\mathrm{R}}$, where $\mathcal{X}'_{\mathrm{S}} = \mathrm{row}(\boldsymbol{H}_{\mathrm{SR}})$ and $\mathcal{X}''_{\mathrm{S}} = \mathrm{row}(\boldsymbol{H}_{\mathrm{SD}})$ form an orthogonal decomposition of $\mathcal{X}_{\mathrm{S}} = \mathrm{row}(\boldsymbol{H}_{\mathrm{S\{RD\}}})$, cf. Definition 5.1. While this shows that Definition 6.2 provides an appropriate characterization of the half-duplex Gaussian MIMO relay channel with orthogonal sender components, there is an equivalent condition that is much easier to work with.

Proposition 6.8. *The half-duplex Gaussian MIMO relay channel has orthogonal sender components if and only if* $\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} = \boldsymbol{0}$.

Proof. Since the condition that defines the half-duplex Gaussian MIMO relay channel with orthogonal sender components is the same as for the full-duplex case, the proof of this result is identical to the proof of Proposition 3.9. □

Now, this condition on $\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_{\mathrm{SD}}$ will not be satisfied unless we assign nonoverlapping frequency bands to the source-to-relay and the source-to-destination channels. But this case is mostly of theoretical interest because if we use *frequency-division duplex* (FDD) to orthogonalize the channels, we can also use FDD instead of *time-division duplex* (TDD) to separate transmission and reception at the relay. Nevertheless, it can be shown that the PDF strategy achieves the CSB for the considered half-duplex Gaussian MIMO relay channel with orthogonal sender components.

Theorem 6.9. *For the half-duplex Gaussian MIMO relay channel with orthogonal sender components, the partial decode-and-forward (PDF) strategy achieves the cut-set bound (CSB), i.e.,* $R_{\mathrm{PDF}} = C_{\mathrm{CSB}}$ *and* $R_{\mathrm{PDF,av}} = C_{\mathrm{CSB,av}}$.

Proof. If the half-duplex Gaussian MIMO relay channel has orthogonal sender components, we can follow the steps in the proof of Theorem 3.10 to show that the CSB is equal to

$$C_{\mathrm{CSB}} = \max_{\substack{R,t_1,t_2 \\ p(\boldsymbol{x}'_{\mathrm{S1}},\boldsymbol{x}''_{\mathrm{S1}})p(\boldsymbol{x}_{\mathrm{S2}},\boldsymbol{x}_{\mathrm{R2}})}} R \quad \text{s.t.} \quad \begin{aligned} &R \leq t_1 I(\boldsymbol{x}'_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}) + t_1 I(\boldsymbol{x}''_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}) \\ &\qquad + t_2 I(\boldsymbol{x}_{\mathrm{S2}};\boldsymbol{y}_{\mathrm{D2}}|\boldsymbol{x}_{\mathrm{R2}}), \\ &R \leq t_1 I(\boldsymbol{x}''_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}) + t_2 I(\boldsymbol{x}_{\mathrm{S2}},\boldsymbol{x}_{\mathrm{R2}};\boldsymbol{y}_{\mathrm{D2}}), \\ &t_1, t_2 \geq 0,\ t_1 + t_2 = 1, \\ &(\boldsymbol{C}'_{\mathrm{S1}} + \boldsymbol{C}''_{\mathrm{S1}}, \boldsymbol{D}_{\mathrm{S}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}, \boldsymbol{D}_{\mathrm{R}}\boldsymbol{C}_2\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \in \mathcal{P}, \end{aligned} \tag{6.118}$$

where it suffices to maximize over all joint probability density functions that factor as $p(\boldsymbol{x}'_{\mathrm{S1}},\boldsymbol{x}''_{\mathrm{S1}}) = p(\boldsymbol{x}'_{\mathrm{S1}})p(\boldsymbol{x}''_{\mathrm{S1}})$. The upper bound specified by this problem can be achieved using the PDF strategy with $\boldsymbol{u} = \boldsymbol{x}'_{\mathrm{S1}}$. □

Consequently, the maximum achievable PDF rate for a half-duplex Gaussian MIMO relay channel with orthogonal sender components can be evaluated based on dual decomposition, cf. Section 5.2. As noted above, however, this result is not of great practical relevance.

In the following, we show that the proposed dual decomposition approaches can also be used to evaluate R_{PDF} and $R_{\mathrm{PDF,av}}$ for the more relevant *stochastically degraded* and *reversely stochastically degraded* half-duplex Gaussian relay channels, which are characterized by the fact that the source-to-relay channel is better or worse (in a probabilistic sense) than the source-to-destination channel.

Definition 6.3. The half-duplex Gaussian MIMO relay channel is *stochastically degraded* if $\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{SR}} \succcurlyeq \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{D}}^{-1}\boldsymbol{H}_{\mathrm{SD}}$. Similarly, the half-duplex Gaussian MIMO relay channel is *reversely stochastically degraded* if $\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{SR}} \preccurlyeq \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{D}}^{-1}\boldsymbol{H}_{\mathrm{SD}}$. If the half-duplex Gaussian MIMO relay channel is either stochastically or reversely stochastically degraded, it is said to be of *stochastically degraded nature.*

Like for the full-duplex case, we can show that if the half-duplex Gaussian relay channel is stochastically degraded or reversely stochastically degraded, the optimal PDF strategy reduces to DF or direct transmission. For any half-duplex Gaussian relay channel of stochastically degraded nature, the maximum achievable PDF rates for the per-phase and the average power constraint can thus be determined using the dual decomposition approaches proposed in Sections 6.2.1 and 6.2.2, respectively.

Theorem 6.10. *If the half-duplex Gaussian MIMO relay channel is stochastically degraded, the optimal partial decode-and-forward (PDF) strategy is equivalent to the decode-and-forward (DF) strategy, i.e.,* $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$ *and* $R_{\mathrm{PDF,av}} = R_{\mathrm{DF,av}}$.

Proof. Applying Lemma 3.11 to $\boldsymbol{H}_1 = \boldsymbol{Z}_{\mathrm{R}}^{-1/2}\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{H}_2 = \boldsymbol{Z}_{\mathrm{D}}^{-1/2}\boldsymbol{H}_{\mathrm{SD}}$, it follows that the half-duplex Gaussian MIMO relay channel is stochastically degraded if and only if there exists an $\boldsymbol{M} \in \mathbb{C}^{N_{\mathrm{D}}\times N_{\mathrm{R}}}$ such that $\boldsymbol{H}_{\mathrm{SD}} = \boldsymbol{M}\boldsymbol{H}_{\mathrm{SR}}$ and $\boldsymbol{M}\boldsymbol{Z}_{\mathrm{R}}\boldsymbol{M}^{\mathrm{H}} \preccurlyeq \boldsymbol{Z}_{\mathrm{D}}$. Therefore,

$$\begin{aligned} I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}|\boldsymbol{u}) &= I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{H}_{\mathrm{SD}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{D1}}|\boldsymbol{u}) = I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{M}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{D1}}|\boldsymbol{u}) \\ &\overset{(a)}{=} I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{R1}}) + \tilde{\boldsymbol{n}}_1|\boldsymbol{u}) \overset{(b)}{\leq} I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{R1}})|\boldsymbol{u}) \\ &\overset{(c)}{\leq} I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{R1}}|\boldsymbol{u}) = I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}|\boldsymbol{u}), \end{aligned} \tag{6.119}$$

where in (a) we have defined $\tilde{\boldsymbol{n}}_1 \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{D}} - \boldsymbol{M}\boldsymbol{Z}_{\mathrm{R}}\boldsymbol{M}^{\mathrm{H}})$ to be independent of $\boldsymbol{x}_{\mathrm{S1}}$ and $\boldsymbol{n}_{\mathrm{R1}}$, (b) is due to the fact that $\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}}+\boldsymbol{n}_{\mathrm{R1}})+\tilde{\boldsymbol{n}}_1$ is a degraded version of $\boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}}+\boldsymbol{n}_{\mathrm{R1}})$, and (c) follows from the data processing inequality [24, Theorem 2.8.1] and the Markov chain $\boldsymbol{x}_{\mathrm{S1}} \leftrightarrow \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{R1}} \leftrightarrow \boldsymbol{M}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x}_{\mathrm{S1}} + \boldsymbol{n}_{\mathrm{R1}})$. As a consequence,

$$\begin{aligned} I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R1}}) + I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}|\boldsymbol{u}) &\leq I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R1}}) + I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}|\boldsymbol{u}) \\ &= I(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}) = I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}), \end{aligned} \tag{6.120}$$

where the last equation is due to the Markov relationship $\boldsymbol{u} \leftrightarrow \boldsymbol{x}_{\mathrm{S1}} \leftrightarrow (\boldsymbol{y}_{\mathrm{R1}}, \boldsymbol{y}_{\mathrm{D1}})$. The maximum achievable PDF rate is hence attained by choosing $\boldsymbol{u} = \boldsymbol{x}_{\mathrm{S1}}$, which implies that the optimal PDF strategy is equivalent to the DF strategy. □

Note that like for Theorem 3.12, the proof of Theorem 6.10 does not rely on the optimal channel input distribution. Rather, it is based on the fact that the condition $\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{SR}} \succcurlyeq \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{D}}^{-1}\boldsymbol{H}_{\mathrm{SD}}$ implies $I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}|\boldsymbol{u}) \leq I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}|\boldsymbol{u})$ for *all* feasible joint distributions $p(\boldsymbol{u}, \boldsymbol{x}_{\mathrm{S1}})$, so Theorem 6.10 can easily be generalized to discrete memoryless half-duplex relay channels.

Remark 6.7. If $I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}|\boldsymbol{u}) \leq I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}}|\boldsymbol{u})$ for all feasible $p(\boldsymbol{u}, \boldsymbol{x}_{\mathrm{S1}})$, the relay is *conditionally more capable* than the destination given $\boldsymbol{u}$, where the auxiliary variable $\boldsymbol{u}$ can be interpreted as a time-sharing variable if we only consider these two mutual information terms. Since the set of feasible joint distributions $p(\boldsymbol{u}, \boldsymbol{x}_{\mathrm{S1}})$ is convex, the time-sharing variable $\boldsymbol{u}$ is obsolete in the condition above. That is, the above condition is equivalent to the condition that $I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{D1}}) \leq I(\boldsymbol{x}_{\mathrm{S1}};\boldsymbol{y}_{\mathrm{R1}})$ for all feasible $p(\boldsymbol{x}_{\mathrm{S1}})$, meaning that the relay is more capable than the destination. For Gaussian channels, the fact that the relay is more capable than the destination is equivalent to the fact that $\boldsymbol{y}_\mathrm{D}$ is a stochastically degraded version of $\boldsymbol{y}_\mathrm{R}$, cf. Appendix B.2.2. For discrete memoryless channels, on the other hand, the more capable condition is less strict than the stochastically degraded condition, cf. Appendix B.2.1, so Theorem 6.10 should in general be formulated using the weaker more capable condition. In the context of Gaussian MIMO relay channels, however, it is more convenient to use the stochastically degraded condition as it is equivalent to and much easier to work with than the more capable condition, cf. Remark 3.8.

Theorem 6.11. *If the half-duplex Gaussian MIMO relay channel is reversely stochastically degraded, the optimal partial decode-and-forward (PDF) strategy is equivalent to direct transmission from source to destination, i.e.,* $R_\mathrm{PDF} = R_\mathrm{PDF,av} = R_\mathrm{P2P}$.

Proof. It again follows from Lemma 3.11 that the half-duplex Gaussian MIMO relay channel is reversely stochastically degraded if and only if there exists an $\boldsymbol{M} \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{D}}$ such that $\boldsymbol{H}_\mathrm{SR} = \boldsymbol{M}\boldsymbol{H}_\mathrm{SD}$ and $\boldsymbol{M}\boldsymbol{Z}_\mathrm{D}\boldsymbol{M}^\mathrm{H} \preccurlyeq \boldsymbol{Z}_\mathrm{R}$. Similar to the proof of Theorem 6.10, we can therefore construct a chain of inequalities to show that $I(\boldsymbol{u};\boldsymbol{y}_\mathrm{R1}) \leq I(\boldsymbol{u};\boldsymbol{y}_\mathrm{D1})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_\mathrm{S1})$, which in turn implies that

$$I(\boldsymbol{u};\boldsymbol{y}_\mathrm{R1}) + I(\boldsymbol{x}_\mathrm{S1};\boldsymbol{y}_\mathrm{D1}|\boldsymbol{u}) \leq I(\boldsymbol{u};\boldsymbol{y}_\mathrm{D1}) + I(\boldsymbol{x}_\mathrm{S};\boldsymbol{y}_\mathrm{D1}|\boldsymbol{u}) = I(\boldsymbol{x}_\mathrm{S1};\boldsymbol{y}_\mathrm{D1}). \tag{6.121}$$

Consequently, the maximum achievable PDF rate is attained by choosing $\boldsymbol{u} = \boldsymbol{0}$, i.e., the optimal PDF strategy reduces to direct transmission from source to destination. □

Remark 6.8. If $I(\boldsymbol{u};\boldsymbol{y}_\mathrm{R1}) \leq I(\boldsymbol{u};\boldsymbol{y}_\mathrm{D1})$ for all feasible $p(\boldsymbol{u},\boldsymbol{x}_\mathrm{S1})$, the destination is *less noisy* than the relay. For Gaussian channels, the fact that the destination is less noisy than the relay is equivalent to the fact that $\boldsymbol{y}_\mathrm{R}$ is a stochastically degraded version of $\boldsymbol{y}_\mathrm{D}$, cf. Appendix B.2.2. But for discrete memoryless channels, the less noisy condition is less strict than the stochastically degraded condition, cf. Appendix B.2.1. Therefore, Theorem 6.11 should in general be formulated using the weaker less noisy condition, cf. Remark 3.9.

Remark 6.9. Because direct transmission is a special case of the DF strategy for the half-duplex relay channel, it immediately follows from Theorem 6.11 that $R_\mathrm{PDF} = R_\mathrm{DF}$ and $R_\mathrm{PDF,av} = R_\mathrm{DF,av}$ if the half-duplex Gaussian MIMO relay channel is reversely stochastically degraded.

Obviously, the half-duplex Gaussian MIMO relay channel is not always of stochastically degraded nature. However, like for the full-duplex case, this is the case whenever the source is equipped with a single antenna.

Corollary 6.12. *If* $N_\mathrm{S} = 1$*, i.e., if the source is equipped with only one antenna, the half-duplex Gaussian MIMO relay channel is of stochastically degraded nature so that* $R_\mathrm{PDF} = R_\mathrm{DF}$ *and* $R_\mathrm{PDF,av} = R_\mathrm{DF,av}$.

Proof. Since the conditions that define the stochastically and reversely stochastically degraded half-duplex Gaussian MIMO relay channels are the same as for the full-duplex case, the proof of this result is exactly the same as the proof of Corollary 3.14. □

6.5 Further Results and Bibliographical Notes

Early on, the half-duplex constrained single-antenna Gaussian relay channel with *time-division duplex* (TDD) was considered by Høst-Madsen [62]. Assuming the source and the relay to be subject to average power constraints, he stated the *cut-set bound* (CSB) and the maximum achievable *decode-and-forward* (DF) rate for the half-duplex Gaussian relay channel. For fixed time-shares, he also derived closed-form expressions for the optimal correlations of x_{S2} and x_{R2}. An achievable *compress-and-forward* (CF) rate for the same half-duplex relay channel model was proposed by Høst-Madsen and Zhang [63]. Because the optimal joint distribution of the channel inputs x_{S1}, x_{S2}, x_{R2} and the relay quantization $\hat{y}_{R1}$ is unknown, they assumed $(x_{S1}, x_{S2}, x_{R2}, \hat{y}_{R1})$ to be jointly Gaussian. For fixed time-shares, the achievable CF rate with Gaussian signals can then be computed in closed form, but like for the CSB and the maximum achievable DF rate, the optimal time-shares must generally be found by numerical search methods.

For the single-antenna half-duplex Gaussian relay channel with TDD and per-phase power constraints imposed on the source and the relay, the CSB and the maximum achievable DF rate were derived by Khojastepour et al. [67]. Moreover, Yao et al. [144] stated an achievable CF rate for this half-duplex relay channel model. Their CF strategy is also based on choosing $(x_{S1}, x_{S2}, x_{R2}, \hat{y}_{R1})$ to be jointly Gaussian, and for fixed time-shares, the corresponding CF rate can again be computed in closed form. For jointly Gaussian channel inputs, which were not yet known to be optimal, the maximum achievable *partial decode-and-forward* (PDF) rate for the half-duplex Gaussian relay channel with TDD and per-phase power constraints was stated by Stein in [118]. There, it was already observed that PDF cannot improve on DF if $|h_{SR}|^2 \geq |h_{SD}|^2$ and that both PDF and DF reduce to direct transmission if $|h_{SR}|^2 < |h_{SD}|^2$ (assuming equal noise variances). Beyond that, Cover and El Gamal's mixed strategy [21, Theorem 7] was also applied to the half-duplex single-antenna Gaussian relay channel in [118], where jointly Gaussian $(x_{S1}, x_{S2}, x_{R2}, \hat{y}_{R1})$ were chosen again.

The *receiver frequency-division* single-antenna Gaussian relay channel, in which the relay-to-destination channel is assigned a separate frequency band, was considered by El Gamal et al. [35] and by Liang and Veeravalli [82]. More specifically, for the average power constraint and equal frequency-shares, closed-form expressions for the CSB, the maximum achievable DF rate, an achievable CF rate with jointly Gaussian channel inputs and relay quantization, and an achievable *amplify-and-forward* (AF) rate were derived in [35]. Furthermore, it was shown that PDF always reduces to DF or direct transmission and that *linear relaying* can outperform the (P)DF and CF strategies for certain channel conditions. In [82], the properties of the optimal frequency allocation for the DF strategy in the receiver frequency-division Gaussian relay channel with the average power constraint were analyzed, and it was proved that DF achieves the CSB if the relay-to-destination channel is weak compared to the other channels.

For the half-duplex Gaussian MIMO relay channel with the average power constraint and the more general *frequency-division duplex* (FDD) protocol where the source may utilize the entire spectrum for transmission, Ng and Foschini [95] showed that the CSB and the maximum achievable DF rate can be determined as the solutions of convex optimization problems. The key to proving this was the convexity preserving property of the perspective operation. In combination with the change of variables discussed in Section 6.2.1, this property was also exploited by Simoens et al. [116], who considered the CSB and the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel with TDD and the per-phase power constraint. But like for the full-duplex case, they simply replaced the conditional covariance matrix $\boldsymbol{C}_{\mathrm{S|R2}}$ by $\boldsymbol{C}_{\mathrm{S2}}$ in the first rate constraints so that the convex optimization problems resulting from their considerations only yield upper bounds to C_{CSB} and R_{DF}.

The fact that both the CSB and the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel with TDD and the per-phase power constraint can indeed be determined as the solutions of convex optimization problems was later proved by Gerdes and Utschick [45]. In order to obtain the convex rate maximization problems, the terms involving $\boldsymbol{C}_{\mathrm{S|R2}}$ were first reformulated by means of an auxiliary variable before the change of variables and the perspective operation were applied, cf. Section 6.2.1. Moreover, it was also noted in [45] that the convex optimization problems that yield C_{CSB} and R_{DF} do not satisfy the ruleset of *disciplined convex programming* (DCP), which means that it is not possible to directly apply standard semidefinite program (SDP) solvers to these problems.

The dual decomposition approaches that can be used to evaluate $C_{\mathrm{CSB}}/C_{\mathrm{CSB,av}}$ and $R_{\mathrm{DF}}/R_{\mathrm{DF,av}}$ in the Lagrangian dual domain, e.g., by means of the cutting-plane method, are due to Gerdes, Riemensberger, and Utschick [41, 43]. In particular, the approach to determine the CSB and the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel with TDD in the dual domain was first derived for the per-phase power constraint in [41] and then generalized to the average power constraint in [43]. The beauty of the proposed dual decomposition approaches is that the phases of the TDD communication protocol decouple in the dual problem. In addition, evaluating the dual function only requires to solve one convex optimization problem for each phase of the protocol so that the complexity scales linearly with the number of protocol phases. Therefore, the dual decomposition approaches can easily be extended to the half-duplex two-way Gaussian MIMO relay channel, cf. Chapter 7.

The results concerning achievable PDF rates for the half-duplex Gaussian MIMO relay channel presented in this chapter had not explicitly been reported so far, with the exception of the zero-forcing (ZF) PDF approach for the per-phase power constraint, cf. [135]. However, we remark that the results of Sections 6.4.1, 6.4.2 and Section 6.4.4 are rather straightforward extensions of the corresponding results for the full-duplex case, cf. Sections 3.4.1, 3.4.2 and Section 3.4.4, respectively, which were originally derived in [40] and [47]. Furthermore, for the per-phase power constraint, it was already pointed out in [48] and [136] that the ZF PDF approach and the inner approximation algorithm (IAA) can of course also be applied to the half-duplex Gaussian MIMO relay channel and that the resulting rate maximization problems can again be tackled by the dual decomposition approach derived in [41].

A suboptimal CF scheme for the half-duplex Gaussian MIMO relay channel with TDD and per-phase power constraints imposed on the source and the relay was proposed by Simoens et al. [117], who assumed $(\boldsymbol{x}_{\mathrm{S}1}, \boldsymbol{x}_{\mathrm{S}2}, \boldsymbol{x}_{\mathrm{R}2}, \hat{\boldsymbol{y}}_{\mathrm{R}1})$ to be jointly Gaussian and took a greedy alternating optimization approach that works as follows. In the relay receive phase, the destination performs successive decoding with the relay's message, i.e., the encoded relay quantization, being decoded first. In the first and second steps of the algorithm, $\boldsymbol{C}_{\mathrm{S}2}$ and $\boldsymbol{C}_{\mathrm{R}2}$ are then chosen as to maximize the information rates on the source-to-destination channel and the resulting effective relay-to-destination channel, respectively. The third step consists of an outer and an inner loop. In the outer loop, the achievable CF rate is maximized with respect to the time-shares t_1 and t_2 while the inner loop alternates between optimizing the achievable rate with respect to $\boldsymbol{C}_{\mathrm{S}1}$ and the relay quantization.

Note that this greedy optimization approach is similar to that proposed by Ng and Foschini for the full-duplex case [95]. However, an important difference between the two suboptimal CF schemes is that Ng and Foschini used rate distortion quantization with squared-error distortion to determine the relay quantization, whereas in [117], the relay quantization is chosen as to maximize $I(\boldsymbol{x}_{\mathrm{S}1}; \hat{\boldsymbol{y}}_{\mathrm{R}1} | \boldsymbol{y}_{\mathrm{D}1})$, i.e., the information $\hat{\boldsymbol{y}}_{\mathrm{R}1}$ contains about $\boldsymbol{x}_{\mathrm{S}1}$ given $\boldsymbol{y}_{\mathrm{D}1}$. Although not mentioned in [117], this means that the relay quantization is determined according to the *information bottleneck method* [14, 56, 125], which implicitly uses information divergence (relative entropy) as distortion measure. Consequently, it could be shown in [117] that this approach achieves higher CF rates than classical rate distortion quantization with squared-error distortion.

Like for the full-duplex case, there are also many works on AF schemes for the half-duplex constrained Gaussian MIMO relay channel. However, it again does not seem possible to determine the maximum achievable AF rate since the AF strategy suffers from noise amplification induced by the relay, which means that difficult nonconvex optimization problems have to be solved even if Gaussian channel inputs are assumed. For an overview of suboptimal AF schemes for the half-duplex Gaussian MIMO relay channel, we refer to [105] and references therein.

Finally, the results that have so far been mentioned in this chapter are based on the assumption that all three nodes noncausally know which resources the relay uses for transmission and reception. That is, all nodes know the time-shares of the relay receive and transmit phases or the frequency-shares of the source-to-relay and the relay-to-destination channels in advance. On the other hand, Kramer considered a random sleep-listen-or-talk (SLoT) strategy for the half-duplex relay channel with TDD, where the destination does not a priori know when the relay transmits [75], cf. Section 5.6. In this work, Kramer not only showed that fixed TDD schedules are generally suboptimal, but also that Gaussian source and relay inputs do not always attain the highest achievable rates for the half-duplex Gaussian relay channel.

Chapter 7

Half-Duplex Two-Way Gaussian MIMO Relay Channel

In this chapter, we consider the half-duplex *two-way* Gaussian MIMO relay channel, which models the scenario where two terminals exchange information with the help of a relay. More specifically, we study bounds on the capacity region of the *restricted* half-duplex two-way Gaussian MIMO relay channel, where the bidirectional communication is restricted in the sense that the encoders at the terminals may neither cooperate, nor are they allowed to use previously decoded information to encode their messages. To this end, we apply the information theoretical results of Section 5.5, which also hold for continuous-alphabet channels, to the Gaussian MIMO relay channel model. Note that since many results presented here are direct extensions of the results for unidirectional communication, this chapter is less detailed than the previous one.

The system model for the restricted half-duplex two-way Gaussian MIMO relay channel is introduced in Section 7.1. We remark that like in Chapter 6, two different power constraints are considered, the *per-phase* and the *average* power constraint. In Sections 7.2 and 7.3, respectively, we then prove that the *cut-set outer bound* (CSOB) region $\mathcal{C}_{\text{CSB}}$ and the achievable *decode-and-forward* (DF) rate region $\mathcal{R}_{\text{DF}}$ are attained by jointly Gaussian channel inputs for both power constraints. These results again follow from the entropy maximizing property of the Gaussian distribution. In addition, we exemplarily show how the dual decomposition approach that was derived in Section 6.2.1 can be extended to the half-duplex two-way Gaussian MIMO relay channel in order to evaluate $\mathcal{C}_{\text{CSB}}$ and $\mathcal{R}_{\text{DF}}$ for the per-phase power constraint.

The achievable *partial decode-and-forward* (PDF) rate region $\mathcal{R}_{\text{PDF}}$ is discussed in Section 7.4. Even though the entropy maximizing property of the Gaussian distribution cannot directly be applied, we can again show that $\mathcal{R}_{\text{PDF}}$ is also attained by jointly Gaussian channel inputs. As opposed to $\mathcal{C}_{\text{CSB}}$ and $\mathcal{R}_{\text{DF}}$, however, $\mathcal{R}_{\text{PDF}}$ can in general not be evaluated using convex optimization techniques. In Section 7.4.1, we therefore consider the suboptimal PDF rate regions that are obtained by applying the *zero-forcing* (ZF) PDF approach or the *inner approximation algorithm* (IAA) to the half-duplex two-way

Gaussian MIMO relay channel. These rate regions can always be determined by means of the dual decomposition approaches proposed in Section 6.2. Special cases for which it is possible to evaluate $\mathcal{R}_{\text{PDF}}$ based on dual decomposition are studied in Section 7.4.2, and the chapter eventually concludes in Section 7.5 with bibliographical notes and an overview of further noteworthy results on bidirectional communication in the half-duplex constrained Gaussian (MIMO) relay channel.

7.1 System Model

Our half-duplex two-way Gaussian MIMO relay channel model is again obtained by applying the *linear MIMO model* to the considered relay scenario. In particular, let $\boldsymbol{H}_{\text{KL}} \in \mathbb{C}^{N_{\text{L}} \times N_{\text{K}}}$ denote the channel gain matrix of the channel between nodes K and L, where $\text{K}, \text{L} \in \{\text{A}, \text{B}, \text{R}\}$, and let $\boldsymbol{x}_{\text{K}m}$ and $\boldsymbol{y}_{\text{L}m}$ represent the input and output signals of nodes K and L during phase m, respectively. The six phases of the considered communication protocol, which are illustrated in Figure 7.1, can then be described as follows:

(1) Terminal A transmits to the relay and terminal B:

$$\begin{aligned} \boldsymbol{y}_{\text{R1}} &= \boldsymbol{H}_{\text{AR}}\boldsymbol{x}_{\text{A1}} + \boldsymbol{n}_{\text{R1}}, \qquad \boldsymbol{n}_{\text{R1}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{R}}), \\ \boldsymbol{y}_{\text{B1}} &= \boldsymbol{H}_{\text{AB}}\boldsymbol{x}_{\text{A1}} + \boldsymbol{n}_{\text{B1}}, \qquad \boldsymbol{n}_{\text{B1}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{B}}). \end{aligned} \tag{7.1}$$

(2) Terminal B transmits to the relay and terminal A:

$$\begin{aligned} \boldsymbol{y}_{\text{R2}} &= \boldsymbol{H}_{\text{BR}}\boldsymbol{x}_{\text{B2}} + \boldsymbol{n}_{\text{R2}}, \qquad \boldsymbol{n}_{\text{R2}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{R}}), \\ \boldsymbol{y}_{\text{A2}} &= \boldsymbol{H}_{\text{BA}}\boldsymbol{x}_{\text{B2}} + \boldsymbol{n}_{\text{A2}}, \qquad \boldsymbol{n}_{\text{A2}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{A}}). \end{aligned} \tag{7.2}$$

(3) Terminals A and B transmit to the relay:

$$\boldsymbol{y}_{\text{R3}} = \boldsymbol{H}_{\text{AR}}\boldsymbol{x}_{\text{A3}} + \boldsymbol{H}_{\text{BR}}\boldsymbol{x}_{\text{B3}} + \boldsymbol{n}_{\text{R3}}, \qquad \boldsymbol{n}_{\text{R3}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{R}}). \tag{7.3}$$

(4) The relay transmits to terminals A and B:

$$\begin{aligned} \boldsymbol{y}_{\text{A4}} &= \boldsymbol{H}_{\text{RA}}\boldsymbol{x}_{\text{R4}} + \boldsymbol{n}_{\text{A4}}, \qquad \boldsymbol{n}_{\text{A4}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{A}}), \\ \boldsymbol{y}_{\text{B4}} &= \boldsymbol{H}_{\text{RB}}\boldsymbol{x}_{\text{R4}} + \boldsymbol{n}_{\text{B4}}, \qquad \boldsymbol{n}_{\text{B4}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{B}}). \end{aligned} \tag{7.4}$$

(5) Terminal B and the relay transmit to terminal A:

$$\boldsymbol{y}_{\text{A5}} = \boldsymbol{H}_{\text{BA}}\boldsymbol{x}_{\text{B5}} + \boldsymbol{H}_{\text{RA}}\boldsymbol{x}_{\text{R5}} + \boldsymbol{n}_{\text{A5}}, \qquad \boldsymbol{n}_{\text{A5}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{A}}). \tag{7.5}$$

(6) Terminal A and the relay transmit to terminal B:

$$\boldsymbol{y}_{\text{B6}} = \boldsymbol{H}_{\text{AB}}\boldsymbol{x}_{\text{A6}} + \boldsymbol{H}_{\text{RB}}\boldsymbol{x}_{\text{R6}} + \boldsymbol{n}_{\text{B6}}, \qquad \boldsymbol{n}_{\text{B6}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\text{B}}). \tag{7.6}$$

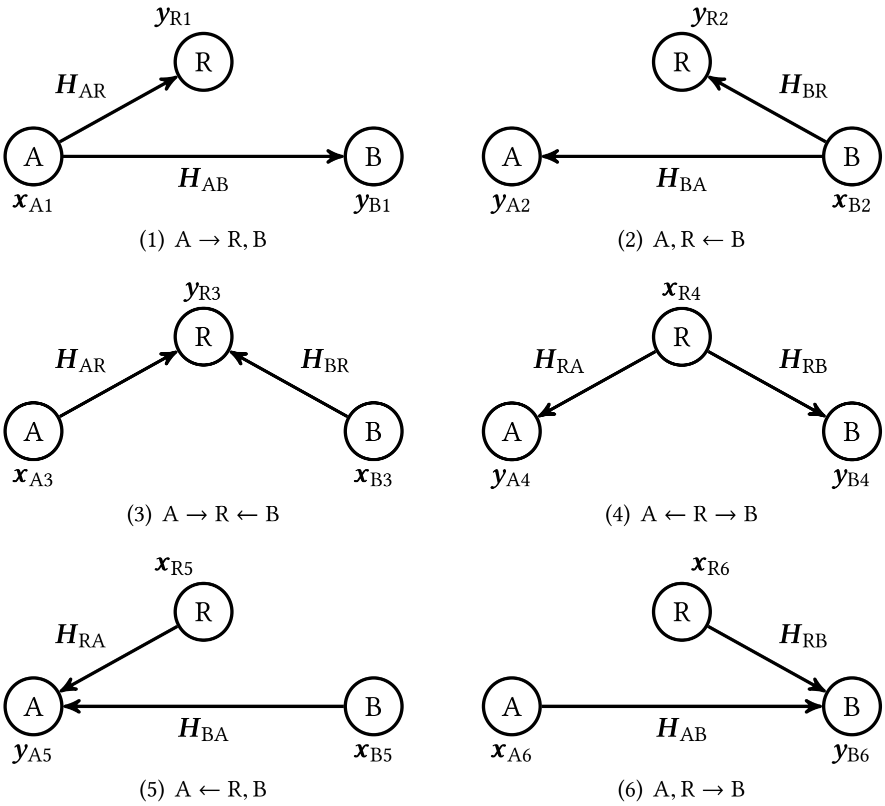


Figure 7.1: Illustration of the Half-Duplex Two-Way Gaussian MIMO Relay Channel

Here, the channels are assumed to remain constant for all phases, which for the results presented in this chapter is without loss of generality because we anyhow require all channel gain matrices to be perfectly known at all nodes. Furthermore, $\boldsymbol{n}_{Am} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, Z_A)$, $\boldsymbol{n}_{Bm} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, Z_B)$, and $\boldsymbol{n}_{Rm} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, Z_R)$ denote the zero-mean proper complex Gaussian noise vectors that are received at terminal A, terminal B, and the relay during phase m, respectively. The noise vectors are assumed to be independent of each other as well as of the transmit signals. More precisely, the additive Gaussian noise $\boldsymbol{n}_{Km}$ received at node K during phase m is assumed to be independent of the noise $\boldsymbol{n}_{Lm'}$ received at another node L for all phases $m' \in \{1, \ldots, 6\}$ and independent of $\boldsymbol{n}_{Km'}$ for all $m' \neq m$. Finally, we assume the noise covariance matrices $Z_A \in \mathbb{C}^{N_A \times N_A}$, $Z_B \in \mathbb{C}^{N_B \times N_B}$, and $Z_R \in \mathbb{C}^{N_R \times N_R}$ to be nonsingular.

Since we consider the restricted half-duplex two-way Gaussian MIMO relay channel, it follows that the channel input distribution of phase 3 must factor as $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) = p(\boldsymbol{x}_{A3})p(\boldsymbol{x}_{B3})$. In addition, it is clear that the capacity region is unbounded if we do not impose further restrictions on the channel inputs, cf. Section 3.1. Like in Chapter 6, we thus require the channel inputs to satisfy a *per-phase* power constraint or an *average* power constraint. Without loss of generality, we may again assume that all channel inputs are zero-mean because channel inputs with nonzero mean consume more transmit power to convey the same amount of information, cf. Section 3.1.

The per-phase power constraint hence imposes the conditions

$$\operatorname{tr}(\boldsymbol{C}_{\mathrm{A}m}) \leq P_{\mathrm{A}}, \qquad \operatorname{tr}(\boldsymbol{C}_{\mathrm{B}m}) \leq P_{\mathrm{B}}, \qquad \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}m}) \leq P_{\mathrm{R}}, \qquad \forall m \in \{1,\dots,6\} \tag{7.7}$$

on the channel inputs $\boldsymbol{x}_{\mathrm{A}m}$, $\boldsymbol{x}_{\mathrm{B}m}$, and $\boldsymbol{x}_{\mathrm{R}m}$ if $\boldsymbol{C}_{\mathrm{A}m}$, $\boldsymbol{C}_{\mathrm{B}m}$, and $\boldsymbol{C}_{\mathrm{R}m}$ denote the respective covariance matrices. Note that for simplicity, we have assumed here that the power budgets that are available to the nodes in a particular phase do not change over the different phases. On the other hand, the average power constraint requires the channel inputs of terminal A, terminal B, and the relay to satisfy

$$\sum_{m=1}^{6} t_m \operatorname{tr}(\boldsymbol{C}_{\mathrm{A}m}) \leq P_{\mathrm{A}}, \qquad \sum_{m=1}^{6} t_m \operatorname{tr}(\boldsymbol{C}_{\mathrm{B}m}) \leq P_{\mathrm{B}}, \qquad \sum_{m=1}^{6} t_m \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}m}) \leq P_{\mathrm{R}}. \tag{7.8}$$

We again remark that the average power constraint is dominated by the per-phase power constraint if the values of $P_{\mathrm{A}} > 0$, $P_{\mathrm{B}} > 0$, and $P_{\mathrm{R}} > 0$ in (7.7) and (7.8) are the same. That is, if the covariance matrices $(\boldsymbol{C}_{\mathrm{A}m}, \boldsymbol{C}_{\mathrm{B}m}, \boldsymbol{C}_{\mathrm{R}m})$, $m \in \{1,\dots,6\}$, satisfy (7.7), they also satisfy (7.8) since the time-shares $t_m \geq 0$ are normalized such that $\sum_{m=1}^{6} t_m = 1$.

7.2 Cut-Set Outer Bound

For any half-duplex two-way Gaussian MIMO relay channel, the capacity region is a subset of the cut-set outer bound (CSOB) region

$$\begin{aligned}
\mathcal{C}_{\mathrm{CSB}} = \bigcup_{\prod_{m=1}^{6} p(\boldsymbol{x}_{\mathrm{A}m}, \boldsymbol{x}_{\mathrm{B}m}, \boldsymbol{x}_{\mathrm{R}m})} \Big\{ &(R_{\mathrm{A}}, R_{\mathrm{B}}) \in \mathbb{R}_{+}^{2} : \\
& t_1, \dots, t_6 \geq 0,\ \textstyle\sum_{m=1}^{6} t_m = 1,\ (\boldsymbol{C}_{\mathrm{A}m}, \boldsymbol{C}_{\mathrm{B}m}, \boldsymbol{C}_{\mathrm{R}m})_{m \in \{1,\dots,6\}} \in \mathcal{P}, \\
& R_{\mathrm{A}} \leq t_1 I(\boldsymbol{x}_{\mathrm{A}1}; \boldsymbol{y}_{\mathrm{R}1}, \boldsymbol{y}_{\mathrm{B}1}) + t_3 I(\boldsymbol{x}_{\mathrm{A}3}; \boldsymbol{y}_{\mathrm{R}3} | \boldsymbol{x}_{\mathrm{B}3}) + t_6 I(\boldsymbol{x}_{\mathrm{A}6}; \boldsymbol{y}_{\mathrm{B}6} | \boldsymbol{x}_{\mathrm{R}6}), \\
& R_{\mathrm{A}} \leq t_1 I(\boldsymbol{x}_{\mathrm{A}1}; \boldsymbol{y}_{\mathrm{B}1}) + t_4 I(\boldsymbol{x}_{\mathrm{R}4}; \boldsymbol{y}_{\mathrm{B}4}) + t_6 I(\boldsymbol{x}_{\mathrm{A}6}, \boldsymbol{x}_{\mathrm{R}6}; \boldsymbol{y}_{\mathrm{B}6}), \\
& R_{\mathrm{B}} \leq t_2 I(\boldsymbol{x}_{\mathrm{B}2}; \boldsymbol{y}_{\mathrm{R}2}, \boldsymbol{y}_{\mathrm{A}2}) + t_3 I(\boldsymbol{x}_{\mathrm{B}3}; \boldsymbol{y}_{\mathrm{R}3} | \boldsymbol{x}_{\mathrm{A}3}) + t_5 I(\boldsymbol{x}_{\mathrm{B}5}; \boldsymbol{y}_{\mathrm{A}5} | \boldsymbol{x}_{\mathrm{R}5}), \\
& R_{\mathrm{B}} \leq t_2 I(\boldsymbol{x}_{\mathrm{B}2}; \boldsymbol{y}_{\mathrm{A}2}) + t_4 I(\boldsymbol{x}_{\mathrm{R}4}; \boldsymbol{y}_{\mathrm{A}4}) + t_5 I(\boldsymbol{x}_{\mathrm{B}5}, \boldsymbol{x}_{\mathrm{R}5}; \boldsymbol{y}_{\mathrm{A}5}) \Big\},
\end{aligned} \tag{7.9}$$

where for the per-phase power constraint, the set $\mathcal{P}$ is given by

$$\begin{aligned}
\mathcal{P} = \Big\{ &(\boldsymbol{C}_{\mathrm{A}m}, \boldsymbol{C}_{\mathrm{B}m}, \boldsymbol{C}_{\mathrm{R}m})_{m \in \{1,\dots,6\}} : \\
& \operatorname{tr}(\boldsymbol{C}_{\mathrm{A}m}) \leq P_{\mathrm{A}},\ \operatorname{tr}(\boldsymbol{C}_{\mathrm{B}m}) \leq P_{\mathrm{B}},\ \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}m}) \leq P_{\mathrm{R}},\ \forall m \in \{1,\dots,6\} \Big\},
\end{aligned} \tag{7.10}$$

and where for the average power constraint, $\mathcal{P}$ is equal to

$$\begin{aligned}
\mathcal{P} = \Big\{ &(\boldsymbol{C}_{\mathrm{A}m}, \boldsymbol{C}_{\mathrm{B}m}, \boldsymbol{C}_{\mathrm{R}m})_{m \in \{1,\dots,6\}} : \\
& \textstyle\sum_{m=1}^{6} t_m \operatorname{tr}(\boldsymbol{C}_{\mathrm{A}m}) \leq P_{\mathrm{A}},\ \sum_{m=1}^{6} t_m \operatorname{tr}(\boldsymbol{C}_{\mathrm{B}m}) \leq P_{\mathrm{B}},\ \sum_{m=1}^{6} t_m \operatorname{tr}(\boldsymbol{C}_{\mathrm{R}m}) \leq P_{\mathrm{R}} \Big\}.
\end{aligned} \tag{7.11}$$

Now, we can once again invoke the entropy maximizing property of the Gaussian distribution to prove that the CSOB region is attained by jointly proper complex Gaussian channel inputs for both the per-phase and the average power constraint. Beyond that, the theorem below reveals that we only need to consider channel input distributions for which $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3})$ factors as $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) = p(\boldsymbol{x}_{A3})p(\boldsymbol{x}_{B3})$. As a result, we can conclude that $\mathcal{C}_{\text{CSB}}$ is the tightest outer bound on the capacity region of the restricted half-duplex two-way Gaussian MIMO relay channel that can be obtained by means of standard cut-set arguments.

Theorem 7.1. *For the half-duplex two-way Gaussian MIMO relay channel with the per-phase or the average power constraint, the cut-set outer bound (CSOB) region $\mathcal{C}_{\text{CSB}}$ is attained by jointly proper complex Gaussian channel inputs such that $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3})$ factors as $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) = p(\boldsymbol{x}_{A3})p(\boldsymbol{x}_{B3})$.*

Proof. The fact that $\mathcal{C}_{\text{CSB}}$ is attained by jointly proper complex Gaussian channel inputs can be proved by following the steps in the proof of Theorem 6.1. In particular, for fixed time-shares t_m and (joint) covariance matrices $\boldsymbol{C}_m$, $m \in \{1, \ldots, 6\}$, the Gaussian distribution simultaneously maximizes all mutual information terms in (7.9) because it maximizes both differential entropy [24, Theorem 8.6.5] and conditional differential entropy [123]. More precisely, the (conditional) differential entropy of a complex random vector is maximized by the zero-mean proper (circularly symmetric) complex Gaussian distribution [94, 122].

In order to prove that it suffices to consider input distributions $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3})$ that factor as $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) = p(\boldsymbol{x}_{A3})p(\boldsymbol{x}_{B3})$, note that the relay's receive signal in phase 3 is given by $\boldsymbol{y}_{R3} = \boldsymbol{H}_{AR}\boldsymbol{x}_{A3} + \boldsymbol{H}_{BR}\boldsymbol{x}_{B3} + \boldsymbol{n}_{R3}$, where the additive Gaussian noise is independent of both $\boldsymbol{x}_{A3}$ and $\boldsymbol{x}_{B3}$. The two mutual information terms associated with phase 3 in (7.9) can therefore be upper bounded by

$$\begin{aligned} I(\boldsymbol{x}_{A3}; \boldsymbol{y}_{R3}|\boldsymbol{x}_{B3}) &= h(\boldsymbol{y}_{R3}|\boldsymbol{x}_{B3}) - h(\boldsymbol{y}_{R3}|\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) \\ &= h(\boldsymbol{H}_{AR}\boldsymbol{x}_{A3} + \boldsymbol{n}_{R3}|\boldsymbol{x}_{B3}) - h(\boldsymbol{n}_{R3}) \qquad (7.12) \\ &\leq h(\boldsymbol{H}_{AR}\boldsymbol{x}_{A3} + \boldsymbol{n}_{R3}) - h(\boldsymbol{n}_{R3}), \\ I(\boldsymbol{x}_{B3}; \boldsymbol{y}_{R3}|\boldsymbol{x}_{A3}) &= h(\boldsymbol{y}_{R3}|\boldsymbol{x}_{A3}) - h(\boldsymbol{y}_{R3}|\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) \\ &= h(\boldsymbol{H}_{BR}\boldsymbol{x}_{B3} + \boldsymbol{n}_{R3}|\boldsymbol{x}_{A3}) - h(\boldsymbol{n}_{R3}) \qquad (7.13) \\ &\leq h(\boldsymbol{H}_{BR}\boldsymbol{x}_{B3} + \boldsymbol{n}_{R3}) - h(\boldsymbol{n}_{R3}), \end{aligned}$$

with equality in (7.12) and (7.13) if and only if $\boldsymbol{x}_{A3}$ and $\boldsymbol{x}_{B3}$ are independent [24, Corollary to Theorem 8.6.1]. □

As a consequence, the optimal channel input distributions can be represented by the (joint) covariance matrices $\boldsymbol{C}_m$, $m \in \{1, \ldots, 6\}$, and all mutual information terms in (7.9) reduce to log-det expressions. In the following, we make use of this result to show how $\mathcal{C}_{\text{CSB}}$ can be evaluated if the relay and both terminals are subject to the per-phase power constraint. To this end, first note that we can determine the boundary of the CSOB region $\mathcal{C}_{\text{CSB}}$ by solving WSR maximization problems over $\mathcal{C}_{\text{CSB}}$ for different weight vectors $\boldsymbol{w} \in \mathbb{R}_+^2$. More specifically, the boundary of $\mathcal{C}_{\text{CSB}}$ can be determined with arbitrary precision by varying the ratio of the weights $\frac{w_1}{w_2}$ from zero to infinity. For a given weight vector $\boldsymbol{w}$, the optimization problem we then need to solve reads as

$$\max_{\boldsymbol{r}} \ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{C}_{\text{CSB}}. \tag{7.14}$$

However, this formulation of the WSR maximization problem is not very convenient if we actually want to perform the optimization, so we seek a parameterization that is more suitable to the problem. For this purpose, we first find a convex parameterization of the CSOB region $\mathcal{C}_{\text{CSB}}$. Because the objective function is linear, we then obtain a convex optimization problem for which strong duality holds and which can thus be solved in the Lagrangian dual domain. In particular, this means that the WSR maximization problem given in (7.14) can be solved by an extension of the dual decomposition approach that was derived in Section 6.2.1.

Convex Parameterization of Cut-Set Outer Bound Region $\mathcal{C}_{\text{CSB}}$—For the per-phase power constraint and additive white Gaussian noise vectors, let us define the rate regions

$$\begin{aligned}\mathcal{S}_{\text{CSB1}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [s_1, s_2, 0, 0]^{\mathrm{T}}, &\\ s_1 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{A\{RB\}}}\boldsymbol{C}_{\text{A1}}\boldsymbol{H}_{\text{A\{RB\}}}^{\mathrm{H}}\right), &\\ s_2 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{AB}}\boldsymbol{C}_{\text{A1}}\boldsymbol{H}_{\text{AB}}^{\mathrm{H}}\right), &\\ \boldsymbol{C}_{\text{A1}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\text{A1}}) \le P_{\text{A}} \Big\}, &\end{aligned} \tag{7.15}$$

$$\begin{aligned}\mathcal{S}_{\text{CSB2}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [0, 0, s_3, s_4]^{\mathrm{T}}, &\\ s_3 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{B\{RA\}}}\boldsymbol{C}_{\text{B2}}\boldsymbol{H}_{\text{B\{RA\}}}^{\mathrm{H}}\right), &\\ s_4 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{BA}}\boldsymbol{C}_{\text{B2}}\boldsymbol{H}_{\text{BA}}^{\mathrm{H}}\right), &\\ \boldsymbol{C}_{\text{B2}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\text{B2}}) \le P_{\text{B}} \Big\}, &\end{aligned} \tag{7.16}$$

$$\begin{aligned}\mathcal{S}_{\text{CSB3}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [s_1, 0, s_3, 0]^{\mathrm{T}}, &\\ s_1 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{AR}}\boldsymbol{C}_{\text{A3}}\boldsymbol{H}_{\text{AR}}^{\mathrm{H}}\right), &\\ s_3 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{BR}}\boldsymbol{C}_{\text{B3}}\boldsymbol{H}_{\text{BR}}^{\mathrm{H}}\right), &\\ \boldsymbol{C}_{\text{A3}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\text{A3}}) \le P_{\text{A}}, &\\ \boldsymbol{C}_{\text{B3}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\text{B3}}) \le P_{\text{B}} \Big\}, &\end{aligned} \tag{7.17}$$

$$\begin{aligned}\mathcal{S}_{\text{CSB4}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [0, s_2, 0, s_4]^{\mathrm{T}}, &\\ s_2 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{RB}}\boldsymbol{C}_{\text{R4}}\boldsymbol{H}_{\text{RB}}^{\mathrm{H}}\right), &\\ s_4 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\text{RA}}\boldsymbol{C}_{\text{R4}}\boldsymbol{H}_{\text{RA}}^{\mathrm{H}}\right), &\\ \boldsymbol{C}_{\text{R4}} \succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\text{R4}}) \le P_{\text{R}} \Big\}, &\end{aligned} \tag{7.18}$$

$$\begin{aligned}\mathcal{S}_{\mathrm{CSB5}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [0, 0, s_3, s_4]^{\mathrm{T}},\\ s_3 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BA}}\boldsymbol{C}_{\mathrm{Q5}}\boldsymbol{H}_{\mathrm{BA}}^{\mathrm{H}}\right),\\ s_4 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{BR}\}\mathrm{A}}\boldsymbol{C}_5\boldsymbol{H}_{\{\mathrm{BR}\}\mathrm{A}}^{\mathrm{H}}\right),\\ \boldsymbol{C}_{\mathrm{Q5}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_5 - \boldsymbol{D}_{\mathrm{B}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q5}}\boldsymbol{D}_{\mathrm{B}} \succcurlyeq \mathbf{0},\\ \mathrm{tr}(\boldsymbol{D}_{\mathrm{B}}\boldsymbol{C}_5\boldsymbol{D}_{\mathrm{B}}^{\mathrm{H}}) \le P_{\mathrm{B}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{RB}}\boldsymbol{C}_5\boldsymbol{D}_{\mathrm{RB}}^{\mathrm{H}}) \le P_{\mathrm{R}}\Big\},\end{aligned} \tag{7.19}$$

$$\begin{aligned}\mathcal{S}_{\mathrm{CSB6}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [s_1, s_2, 0, 0]^{\mathrm{T}},\\ s_1 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{Q6}}\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right),\\ s_2 \le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\{\mathrm{AR}\}\mathrm{B}}\boldsymbol{C}_6\boldsymbol{H}_{\{\mathrm{AR}\}\mathrm{B}}^{\mathrm{H}}\right),\\ \boldsymbol{C}_{\mathrm{Q6}} \succcurlyeq \mathbf{0}, \quad \boldsymbol{C}_6 - \boldsymbol{D}_{\mathrm{A}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{Q6}}\boldsymbol{D}_{\mathrm{A}} \succcurlyeq \mathbf{0},\\ \mathrm{tr}(\boldsymbol{D}_{\mathrm{A}}\boldsymbol{C}_6\boldsymbol{D}_{\mathrm{A}}^{\mathrm{H}}) \le P_{\mathrm{A}}, \quad \mathrm{tr}(\boldsymbol{D}_{\mathrm{RA}}\boldsymbol{C}_6\boldsymbol{D}_{\mathrm{RA}}^{\mathrm{H}}) \le P_{\mathrm{R}}\Big\},\end{aligned} \tag{7.20}$$

where $\boldsymbol{D}_{\mathrm{A}}$, $\boldsymbol{D}_{\mathrm{B}}$, $\boldsymbol{D}_{\mathrm{RA}}$, $\boldsymbol{D}_{\mathrm{RB}}$ are selection matrices defined as

$$\begin{aligned}\boldsymbol{D}_{\mathrm{A}} &= \begin{bmatrix}\mathbf{I}_{N_{\mathrm{A}}} & \mathbf{0}_{N_{\mathrm{A}}\times N_{\mathrm{R}}}\end{bmatrix}, \qquad \boldsymbol{D}_{\mathrm{RA}} = \begin{bmatrix}\mathbf{0}_{N_{\mathrm{R}}\times N_{\mathrm{A}}} & \mathbf{I}_{N_{\mathrm{R}}}\end{bmatrix},\\ \boldsymbol{D}_{\mathrm{B}} &= \begin{bmatrix}\mathbf{I}_{N_{\mathrm{B}}} & \mathbf{0}_{N_{\mathrm{B}}\times N_{\mathrm{R}}}\end{bmatrix}, \qquad \boldsymbol{D}_{\mathrm{RB}} = \begin{bmatrix}\mathbf{0}_{N_{\mathrm{R}}\times N_{\mathrm{B}}} & \mathbf{I}_{N_{\mathrm{R}}}\end{bmatrix},\end{aligned} \tag{7.21}$$

and where $\boldsymbol{H}_{\mathrm{A}\{\mathrm{RB}\}}$, $\boldsymbol{H}_{\mathrm{B}\{\mathrm{RA}\}}$, $\boldsymbol{H}_{\{\mathrm{AR}\}\mathrm{B}}$, $\boldsymbol{H}_{\{\mathrm{BR}\}\mathrm{A}}$ denote the composite channel gain matrices given by

$$\begin{aligned}\boldsymbol{H}_{\mathrm{A}\{\mathrm{RB}\}} &= \begin{bmatrix}\boldsymbol{H}_{\mathrm{AR}}\\ \boldsymbol{H}_{\mathrm{AB}}\end{bmatrix}, \qquad \boldsymbol{H}_{\{\mathrm{AR}\}\mathrm{B}} = \begin{bmatrix}\boldsymbol{H}_{\mathrm{AB}} & \boldsymbol{H}_{\mathrm{RB}}\end{bmatrix},\\ \boldsymbol{H}_{\mathrm{B}\{\mathrm{RA}\}} &= \begin{bmatrix}\boldsymbol{H}_{\mathrm{BR}}\\ \boldsymbol{H}_{\mathrm{BA}}\end{bmatrix}, \qquad \boldsymbol{H}_{\{\mathrm{BR}\}\mathrm{A}} = \begin{bmatrix}\boldsymbol{H}_{\mathrm{BA}} & \boldsymbol{H}_{\mathrm{RA}}\end{bmatrix}.\end{aligned} \tag{7.22}$$

Note that $\mathcal{S}_{\mathrm{CSB1}}, \ldots, \mathcal{S}_{\mathrm{CSB6}}$, which are easily shown to be compact and convex, specify the contributions of the six phases to the four rate bounds of the CSOB region $\mathcal{C}_{\mathrm{CSB}}$, cf. (7.9). Furthermore, we remark that in $\mathcal{S}_{\mathrm{CSB5}}$ and $\mathcal{S}_{\mathrm{CSB6}}$, respectively, the constraints on s_3 and s_1 have been reformulated by introducing the auxiliary variables $\boldsymbol{C}_{\mathrm{Q5}} = \boldsymbol{C}_{\mathrm{B|R5}} \succcurlyeq \mathbf{0}$ and $\boldsymbol{C}_{\mathrm{Q6}} = \boldsymbol{C}_{\mathrm{A|R6}} \succcurlyeq \mathbf{0}$, relaxing these equality constraints, and applying Lemma 3.2.

By means of these six rate regions, the WSR maximization problem given in (7.14) can be rewritten as

$$\begin{aligned}\max_{\boldsymbol{r}, t_m, \boldsymbol{s}_m} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \le \sum_{m=1}^{6} t_m \boldsymbol{s}_m, \quad \sum_{m=1}^{6} t_m = 1,\\ \boldsymbol{s}_m \in \mathcal{S}_{\mathrm{CSB}m}, \ t_m \ge 0, \ \forall m \in \{1, \ldots, 6\},\end{aligned} \tag{7.23}$$

where each row of $\boldsymbol{A} = \left[\begin{smallmatrix}1 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\end{smallmatrix}\right]^{\mathrm{T}}$ selects one of the four rate constraints. This optimization problem is convex for given time shares $t_1, \ldots, t_6$, but it is not jointly convex in all optimization variables since $t_m \boldsymbol{s}_m$ is not jointly concave in t_m and $\boldsymbol{s}_m$.

Consequently, another reformulation step is necessary. For this purpose, we further define the set

$$\mathcal{S}_{\mathrm{CSB}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = \textstyle\sum_{m=1}^{6} t_m \boldsymbol{s}_m, \quad \sum_{m=1}^{6} t_m = 1, \\ \boldsymbol{s}_m \in \mathcal{S}_{\mathrm{CSB}m},\ t_m \geq 0,\ \forall m \in \{1, \ldots, 6\} \Big\}, \tag{7.24}$$

which is the convex sum of the six convex rate regions $\mathcal{S}_{\mathrm{CSB}1}, \ldots, \mathcal{S}_{\mathrm{CSB}6}$ and thus convex as well, cf. [11, Section 2.3.2]. Using this definition, it follows that we can rewrite (7.23) once more to obtain

$$\max_{\boldsymbol{r},\boldsymbol{s}} \ \boldsymbol{w}^{\mathrm{T}} \boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}. \tag{7.25}$$

Since $\mathcal{S}_{\mathrm{CSB}}$ is a convex set with nonempty interior, this is a convex optimization problem for which *strong duality* holds [11, Section 5.3.2]. In particular, the constraints of this reformulated WSR maximization problem specify a convex set, which means that a convex parameterization of the CSOB region is given by

$$\mathcal{C}_{\mathrm{CSB}} = \Big\{ \boldsymbol{r} \in \mathbb{R}_+^2 : \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s},\ \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}} \Big\}. \tag{7.26}$$

Derivation of the Dual Problem—As strong duality holds for the WSR maximization problem given in (7.25), it can equivalently be solved in the Lagrangian dual domain. To this end, we incorporate the constraint $\boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s}$ into the objective function using the Lagrangian multiplier $\boldsymbol{\lambda}$. The resulting Lagrangian function is then equal to

$$L(\boldsymbol{r}, \boldsymbol{s}, \boldsymbol{\lambda}) = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{r} - \boldsymbol{\lambda}^{\mathrm{T}} \left(\boldsymbol{A}\boldsymbol{r} - \boldsymbol{s} \right), \tag{7.27}$$

and since $\mathcal{S}_{\mathrm{CSB}}$ is also compact, the corresponding dual function is given by

$$\Theta(\boldsymbol{\lambda}) = \sup_{\boldsymbol{r}, \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}} L(\boldsymbol{r}, \boldsymbol{s}, \boldsymbol{\lambda}) = \begin{cases} \max_{\boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s} & \text{if } \boldsymbol{A}^{\mathrm{T}} \boldsymbol{\lambda} = \boldsymbol{w}, \\ +\infty & \text{otherwise.} \end{cases} \tag{7.28}$$

As a consequence, it follows that the Lagrangian dual problem of the WSR maximization problem given in (7.25) reads as

$$\min_{\boldsymbol{\lambda}} \max_{\boldsymbol{s}} \ \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s} \quad \text{s.t.} \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{CSB}}, \quad \boldsymbol{\lambda} \geq \boldsymbol{0},\ \boldsymbol{A}^{\mathrm{T}} \boldsymbol{\lambda} = \boldsymbol{w}. \tag{7.29}$$

Solution by means of the Cutting-Plane Algorithm—Like for unidirectional communication, this dual problem can for example be solved by means of the cutting-plane algorithm, an outer-approximation method in which the dual function is approximated and iteratively refined by a set of linear inequalities. In particular, recall from Section 6.2 that in each iteration of the cutting-plane algorithm, we must solve an approximated dual problem, the so-called *master program*, and the *Lagrangian subproblem*, i.e., we have to evaluate the dual function.

If we apply the cutting-plane algorithm to the dual problem given in (7.29), the master program in the k-th iteration reads as

$$\min_{z,\boldsymbol{\lambda}} z \quad \text{s.t.} \quad z \geq \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(\ell)}, \forall \ell \in \{0,\ldots,k-1\}, \quad \boldsymbol{\lambda} \geq \mathbf{0},\ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{w}, \tag{7.30}$$

where for all $\ell \in \{0,\ldots,k-1\}$, we have $\boldsymbol{s}^{(\ell)} \in \mathcal{S}_{\mathrm{CSB}m}$ for some $m \in \{1,\ldots,6\}$. Furthermore, note that since

$$\Theta(\boldsymbol{\lambda}^{(k)}) = \max_{\boldsymbol{s}\in\mathcal{S}_{\mathrm{CSB}}} \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s} = \max_{m=1,\ldots,6} \max_{\boldsymbol{s}_m\in\mathcal{S}_{\mathrm{CSB}m}} \boldsymbol{\lambda}^{(k),\mathrm{T}}\boldsymbol{s}_m, \tag{7.31}$$

evaluating the dual function at $\boldsymbol{\lambda}^{(k)}$ requires to solve six independent WSR maximization problems, one over each of the compact convex rate regions $\mathcal{S}_{\mathrm{CSB1}},\ldots,\mathcal{S}_{\mathrm{CSB6}}$ that are associated with the six phases of the communication protocol. We again remark that for this purpose, standard SDP solvers like SDPT3 [126, 127] or LogdetPPA [133] which are capable of dealing with log-det terms in the objective function can be used.

Primal Recovery—The proposed dual decomposition approach allows to determine the optimal value of the primal problem given in (7.25) without explicitly optimizing the rate vector that attains the maximum WSR or the corresponding optimal time-shares which need to be allocated to the six phases of the communication protocol. In order to determine $\boldsymbol{r}^\star$ and $t_1^\star,\ldots,t_6^\star$, we again have to perform a *primal recovery*, i.e., we must recover the optimal primal variables from the sequences of primal and dual variables generated by the cutting-plane algorithm.

To this end, suppose the cutting-plane algorithm has converged after iteration K, and consider the dual problem of the corresponding master program, which reads as

$$\max_{\boldsymbol{r}',\tau_0,\ldots,\tau_{K-1}} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r}' \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r}' \leq \sum_{\ell=0}^{K-1}\tau_\ell \boldsymbol{s}^{(\ell)}, \quad \tau_\ell \geq 0, \forall \ell \in \{0,\ldots,K-1\}, \quad \sum_{\ell=0}^{K-1}\tau_\ell = 1. \tag{7.32}$$

Here, $\boldsymbol{r}'$ and τ_ℓ are Lagrangian multipliers associated with the constraints $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{w}$ and $z \geq \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{s}^{(\ell)}$ of the master program, respectively. Moreover, we remark that this problem is an approximation of the primal problem given in (7.25), where $\mathcal{S}_{\mathrm{CSB}}$ is replaced by $\operatorname{conv}(\{\boldsymbol{s}^{(0)},\ldots,\boldsymbol{s}^{(K-1)}\}) \subseteq \mathcal{S}_{\mathrm{CSB}}$, i.e., by a convex combination of feasible points. Now, if we let $\mathcal{I}_m = \{\ell : \boldsymbol{s}^{(\ell)} \in \mathcal{S}_{\mathrm{CSB}m}\} \setminus \bigcup_{j=1}^{m-1}\mathcal{I}_j \subseteq \{0,\ldots,K-1\}$, $m \in \{1,\ldots,6\}$, we can equivalently express the problem given in (7.32) as

$$\begin{aligned} \max_{\boldsymbol{r}',\tau_0,\ldots,\tau_{K-1}} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r}' \quad \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{r}' \leq \sum_{m=1}^{6}\sum_{\ell\in\mathcal{I}_m}\tau_\ell \boldsymbol{s}^{(\ell)}, \\ & \tau_\ell \geq 0, \forall \ell \in \{0,\ldots,K-1\}, \quad \sum_{\ell=0}^{K-1}\tau_\ell = 1. \end{aligned} \tag{7.33}$$

In addition, note that because the rate regions $\mathcal{S}_{\mathrm{CSB}m}$ are convex, it follows that

$$\sum_{\ell \in \mathcal{I}_m} \tau_\ell \boldsymbol{s}^{(\ell)} = \sum_{k \in \mathcal{I}_m} \tau_k \sum_{\ell \in \mathcal{I}_m} \frac{\tau_\ell}{\sum_{k \in \mathcal{I}_m} \tau_k} \boldsymbol{s}^{(\ell)} = \left(\sum_{k \in \mathcal{I}_m} \tau_k \right) \boldsymbol{s}_m \tag{7.34}$$

for some $\boldsymbol{s}_m \in \mathcal{S}_{\mathrm{CSB}m}$, $m \in \{1, \dots, 6\}$. Finally, if we insert (7.34) into (7.33) and compare the result to the optimization problem given in (7.23), we can conclude that

$$\boldsymbol{r}^\star = \boldsymbol{r}', \qquad t_m^\star = \sum_{\ell \in \mathcal{I}_m} \tau_\ell. \tag{7.35}$$

That is, the optimal rate vector $\boldsymbol{r}^\star \in \partial \mathcal{C}_{\mathrm{CSB}}$ and the optimal time-shares $t_m^\star$ can easily be obtained from the Lagrangian dual variables that correspond to the constraints $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{\lambda} = \boldsymbol{w}$ and $z \geq \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{s}^{(\ell)}$ in the master program of the final cutting-plane iteration.

REMARK 7.1. Since all $\mathcal{S}_{\mathrm{CSB}m}$ are convex, time-sharing within the rate regions associated with the six phases of the communication protocol is not necessary. As a result, there is usually only one $\ell \in \mathcal{I}_m$ such that $\tau_\ell > 0$ for any $m \in \{1, \dots, 6\}$, cf. Remark 6.2.

Compared to the evaluation of the CSB for the half-duplex Gaussian MIMO relay channel with the per-phase power constraint, cf. Section 6.2.1, the main difference is that six independent WSR maximization problems need to be solved in each iteration of the cutting-plane method instead of just two. Accordingly, we can also extend the dual decomposition approach presented in Section 6.2.2 to the half-duplex two-way Gaussian MIMO relay channel if we want to determine the CSOB region for the average power constraint, cf. [43].

7.3 Decode-and-Forward

Provided that $p(\boldsymbol{x}_{\mathrm{A3}}, \boldsymbol{x}_{\mathrm{B3}})$ factors as $p(\boldsymbol{x}_{\mathrm{A3}}, \boldsymbol{x}_{\mathrm{B3}}) = p(\boldsymbol{x}_{\mathrm{A3}}) p(\boldsymbol{x}_{\mathrm{B3}})$, the achievable decode-and-forward (DF) rate region $\mathcal{R}_{\mathrm{DF}}$ for the restricted half-duplex two-way Gaussian MIMO relay channel is specified by

$$\begin{aligned} \mathcal{R}_{\mathrm{DF}} = \bigcup_{\prod_{m=1}^{6} p(\boldsymbol{x}_{\mathrm{A}m}, \boldsymbol{x}_{\mathrm{B}m}, \boldsymbol{x}_{\mathrm{R}m})} \Big\{ & (R_{\mathrm{A}}, R_{\mathrm{B}}) \in \mathbb{R}_+^2 : \\ & t_1, \dots, t_6 \geq 0,\ \textstyle\sum_{m=1}^{6} t_m = 1,\ (C_{\mathrm{A}m}, C_{\mathrm{B}m}, C_{\mathrm{R}m})_{m \in \{1,\dots,6\}} \in \mathcal{P}, \\ & R_{\mathrm{A}} \leq t_1 I(\boldsymbol{x}_{\mathrm{A1}}; \boldsymbol{y}_{\mathrm{R1}}) + t_3 I(\boldsymbol{x}_{\mathrm{A3}}; \boldsymbol{y}_{\mathrm{R3}} | \boldsymbol{x}_{\mathrm{B3}}) + t_6 I(\boldsymbol{x}_{\mathrm{A6}}; \boldsymbol{y}_{\mathrm{B6}} | \boldsymbol{x}_{\mathrm{R6}}), \\ & R_{\mathrm{A}} \leq t_1 I(\boldsymbol{x}_{\mathrm{A1}}; \boldsymbol{y}_{\mathrm{B1}}) + t_4 I(\boldsymbol{x}_{\mathrm{R4}}; \boldsymbol{y}_{\mathrm{B4}}) + t_6 I(\boldsymbol{x}_{\mathrm{A6}}, \boldsymbol{x}_{\mathrm{R6}}; \boldsymbol{y}_{\mathrm{B6}}), \\ & R_{\mathrm{B}} \leq t_2 I(\boldsymbol{x}_{\mathrm{B2}}; \boldsymbol{y}_{\mathrm{R2}}) + t_3 I(\boldsymbol{x}_{\mathrm{B3}}; \boldsymbol{y}_{\mathrm{R3}} | \boldsymbol{x}_{\mathrm{A3}}) + t_5 I(\boldsymbol{x}_{\mathrm{B5}}; \boldsymbol{y}_{\mathrm{A5}} | \boldsymbol{x}_{\mathrm{R5}}), \\ & R_{\mathrm{B}} \leq t_2 I(\boldsymbol{x}_{\mathrm{B2}}; \boldsymbol{y}_{\mathrm{A2}}) + t_4 I(\boldsymbol{x}_{\mathrm{R4}}; \boldsymbol{y}_{\mathrm{A4}}) + t_5 I(\boldsymbol{x}_{\mathrm{B5}}, \boldsymbol{x}_{\mathrm{R5}}; \boldsymbol{y}_{\mathrm{A5}}), \\ R_{\mathrm{A}} + {} & R_{\mathrm{B}} \leq t_1 I(\boldsymbol{x}_{\mathrm{A1}}; \boldsymbol{y}_{\mathrm{R1}}) + t_2 I(\boldsymbol{x}_{\mathrm{B2}}; \boldsymbol{y}_{\mathrm{R2}}) + t_3 I(\boldsymbol{x}_{\mathrm{A3}}, \boldsymbol{x}_{\mathrm{B3}}; \boldsymbol{y}_{\mathrm{R3}}) \\ & \quad + t_5 I(\boldsymbol{x}_{\mathrm{B5}}; \boldsymbol{y}_{\mathrm{A5}} | \boldsymbol{x}_{\mathrm{R5}}) + t_6 I(\boldsymbol{x}_{\mathrm{A6}}; \boldsymbol{y}_{\mathrm{B6}} | \boldsymbol{x}_{\mathrm{R6}}) \Big\}, \end{aligned} \tag{7.36}$$

with $\mathcal{P}$ depending on the considered power constraint, cf (7.10) and (7.11). Like for the CSOB region, it is easily shown that the achievable DF rate region is attained by jointly proper complex Gaussian channel inputs.

THEOREM 7.2. *For the restricted half-duplex two-way Gaussian MIMO relay channel with the per-phase or the average power constraint, the achievable decode-and-forward (DF) rate region $\mathcal{R}_{\mathrm{DF}}$ is attained by jointly proper complex Gaussian channel inputs.*

Proof. The proof of this result follows exactly the same lines as the first part of the proof of Theorem 7.1. □

Consequently, the optimal channel input distributions can again be represented by the (joint) covariance matrices $\boldsymbol{C}_m$, $m \in \{1, \ldots, 6\}$, and all mutual information terms in (7.36) reduce to log-det expressions. Beyond that, the boundary of $\mathcal{R}_{\mathrm{DF}}$ can also be determined by solving WSR maximization problems of the form

$$\max_{\boldsymbol{r}} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{R}_{\mathrm{DF}} \tag{7.37}$$

if the weight vector $\boldsymbol{w} \in \mathbb{R}_+^2$ is varied appropriately. As $\mathcal{R}_{\mathrm{DF}}$ and $\mathcal{C}_{\mathrm{CSB}}$ are similar in structure, any such WSR maximization problem can again be solved based on dual decomposition. In particular, we can find convex parameterizations of $\mathcal{R}_{\mathrm{DF}}$ for both the per-phase and the average power constraint so that strong duality holds for the resulting convex optimization problems. These problems can then be solved in the dual domain using the cutting-plane method, and the optimal rate vectors and the optimal time-shares can eventually be determined by means of primal recovery.

Below, we exemplarily derive a suitable convex parameterization of $\mathcal{R}_{\mathrm{DF}}$ for the per-phase power constraint. For the extension of the dual decomposition approach to the average power constraint, we again refer to [43].

Convex Parameterization of Achievable DF Rate Region $\mathcal{R}_{\mathrm{DF}}$—Assuming the Gaussian noise to be white again, let us define the six compact convex rate regions

$$\begin{aligned}\mathcal{S}_{\mathrm{DF1}} = \Big\{\boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [s_1, s_2, 0, 0]^{\mathrm{T}},\\ s_1 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{A1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right),\\ s_2 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{A1}}\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right),\\ \boldsymbol{C}_{\mathrm{A1}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{A1}}) \leq P_{\mathrm{A}}\Big\},\end{aligned} \tag{7.38}$$

$$\begin{aligned}\mathcal{S}_{\mathrm{DF2}} = \Big\{\boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [0, 0, s_3, s_4]^{\mathrm{T}},\\ s_3 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{B2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right),\\ s_4 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BA}}\boldsymbol{C}_{\mathrm{B2}}\boldsymbol{H}_{\mathrm{BA}}^{\mathrm{H}}\right),\\ \boldsymbol{C}_{\mathrm{B2}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{B2}}) \leq P_{\mathrm{B}}\Big\},\end{aligned} \tag{7.39}$$

$$\begin{aligned}\mathcal{S}_{\mathrm{DF3}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [s_1, 0, s_3, 0]^{\mathrm{T}}, \\ s_1 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{A3}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right), \\ s_3 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{B3}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right), \\ s_1 + s_3 \leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{A3}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{B3}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right), \\ \boldsymbol{C}_{\mathrm{A3}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{A3}}) \leq P_{\mathrm{A}}, \\ \boldsymbol{C}_{\mathrm{B3}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{B3}}) \leq P_{\mathrm{B}} \Big\},\end{aligned} \tag{7.40}$$

and $\mathcal{S}_{\mathrm{DF}m} = \mathcal{S}_{\mathrm{CSB}m}$ for $m \in \{4, 5, 6\}$. Note that $\mathcal{S}_{\mathrm{DF}m} \subseteq \mathcal{S}_{\mathrm{CSB}m}$ for $m \in \{1, 2, 3\}$, where the main difference between the CSOB region and the achievable DF rate region is the additional sum rate constraint in $\mathcal{S}_{\mathrm{DF3}}$ as compared to $\mathcal{S}_{\mathrm{CSB3}}$. Obviously, this constraint comes from the third phase of the communication protocol, a multiple-access phase in which both terminals transmit to the relay. The sum rate constraint in $\mathcal{S}_{\mathrm{DF3}}$ is due to the fact that the relay must decode the messages from both terminals when it uses DF. Moreover, similar to the relay receive phase for unidirectional communication in the half-duplex constrained Gaussian MIMO relay channel, the composite channel gain matrices $\boldsymbol{H}_{\mathrm{A\{RB\}}}$ and $\boldsymbol{H}_{\mathrm{B\{RA\}}}$ have been replaced by $\boldsymbol{H}_{\mathrm{AR}}$ and $\boldsymbol{H}_{\mathrm{BR}}$ in $\mathcal{S}_{\mathrm{DF1}}$ and $\mathcal{S}_{\mathrm{DF2}}$ as compared to $\mathcal{S}_{\mathrm{CSB1}}$ and $\mathcal{S}_{\mathrm{CSB2}}$, respectively.

On the other hand, the rate regions associated with the phases in which the relay transmits are the same for the CSOB region and the DF rate region. We remark that phase 4 corresponds to a bidirectional broadcast channel, whose capacity region was derived in [96] (cf. also [79, 139]). Beyond that, phases 5 and 6 are already known from the relay transmit phase for unidirectional communication. They correspond to multiple-access channels with correlated sources [22], where the information that serves as the source of the relay is completely known at terminal B (phase 5) or terminal A (phase 6) since the relay has no own information to transmit.

Having defined the six rate regions $\mathcal{S}_{\mathrm{DF1}}, \ldots, \mathcal{S}_{\mathrm{DF6}}$, we can now express the WSR maximization problem that yields a point on the boundary of $\mathcal{R}_{\mathrm{DF}}$ as follows:

$$\begin{aligned}\max_{\boldsymbol{r}, t_m, \boldsymbol{s}_m} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \sum_{m=1}^{6} t_m \boldsymbol{s}_m, \quad \sum_{m=1}^{6} t_m = 1, \\ \boldsymbol{s}_m \in \mathcal{S}_{\mathrm{DF}m}, \ t_m \geq 0, \ \forall m \in \{1, \ldots, 6\}.\end{aligned} \tag{7.41}$$

Note that the sum rate constraint is not explicitly expressed by a fifth constraint like in (7.36), but rather it is implicitly accounted for in the definition of $\mathcal{S}_{\mathrm{DF3}}$. Moreover, (7.41) and (7.23) have exactly the same structure, so the final step to obtain a convex parameterization of $\mathcal{R}_{\mathrm{DF}}$ is the same as for the CSOB region. In particular, we define $\mathcal{S}_{\mathrm{DF}}$ to be the convex sum of the six convex rate regions $\mathcal{S}_{\mathrm{DF1}}, \ldots, \mathcal{S}_{\mathrm{DF6}}$, i.e.,

$$\begin{aligned}\mathcal{S}_{\mathrm{DF}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = \textstyle\sum_{m=1}^{6} t_m \boldsymbol{s}_m, \quad \textstyle\sum_{m=1}^{6} t_m = 1, \\ \boldsymbol{s}_m \in \mathcal{S}_{\mathrm{DF}m}, \ t_m \geq 0, \ \forall m \in \{1, \ldots, 6\} \Big\},\end{aligned} \tag{7.42}$$

which is also convex, cf. [11, Section 2.3.2] again. Using this definition, it follows that we can rewrite (7.41) once more to obtain

$$\max_{\boldsymbol{r},\boldsymbol{s}} \ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{DF}}. \tag{7.43}$$

Since the constraints of this problem define a convex set, we can conclude that a convex parameterization of the achievable DF rate region $\mathcal{R}_{\mathrm{DF}}$ is given by

$$\mathcal{R}_{\mathrm{DF}} = \left\{ \boldsymbol{r} \in \mathbb{R}_+^2 : \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s},\ \boldsymbol{s} \in \mathcal{S}_{\mathrm{DF}} \right\}. \tag{7.44}$$

Furthermore, the interior of $\mathcal{S}_{\mathrm{DF}}$ is nonempty, so the reformulated WSR maximization problem can again be solved in the Lagrangian dual domain by means of the dual decomposition approach that was presented in the previous section. After performing the primal recovery, the optimal time-shares $t_1^\star, \dots, t_6^\star$ then eventually reveal which of the six protocol phases should be used (and for how long) to attain $\boldsymbol{r}^\star \in \partial\mathcal{R}_{\mathrm{DF}}$.

Remark 7.2. Because all $\mathcal{S}_{\mathrm{DF}m}$, $m \in \{1, \dots, 6\}$, are convex, Remark 7.1 applies here as well. In addition, note that $\mathcal{S}_{\mathrm{DF}} = \operatorname{conv}\left(\bigcup_{m=1}^{6} \mathcal{S}_{\mathrm{DF}m}\right)$ and that $\bigcup_{m=1}^{6} \mathcal{S}_{\mathrm{DF}m}$ is connected. The latter is due to $\boldsymbol{0} \in \bigcap_{m=1}^{6} \mathcal{S}_{\mathrm{DF}m}$ and the fact that the union of connected sets with nonempty intersection is connected. It hence follows from the strengthened version of Carathéodory's theorem [30, Theorem 18(ii)] that $\boldsymbol{s}^\star \in \mathcal{S}_{\mathrm{DF}} \subseteq \mathbb{R}_+^4$ can be represented as a convex combination of no more than four points in $\bigcup_{m=1}^{6} \mathcal{S}_{\mathrm{DF}m}$, which means that at most four of the six protocol phases need to be used in order to attain $\boldsymbol{r}^\star \in \partial\mathcal{R}_{\mathrm{DF}}$.

7.4 Partial Decode-and-Forward

Let $\mathcal{P}$ again denote the set of feasible covariance matrices for the per-phase or the average power constraint, cf. (7.10) and (7.11). The corresponding achievable partial decode-and-forward (PDF) rate region $\mathcal{R}_{\mathrm{PDF}}$ for the restricted half-duplex two-way Gaussian MIMO relay channel is then given by

$$\begin{aligned}
\mathcal{R}_{\mathrm{PDF}} = \bigcup_{\substack{p(\boldsymbol{u}_{\mathrm{A}},\boldsymbol{x}_{\mathrm{A1}})p(\boldsymbol{u}_{\mathrm{B}},\boldsymbol{x}_{\mathrm{B2}}) \\ \prod_{m=3}^{6} p(\boldsymbol{x}_{\mathrm{A}m},\boldsymbol{x}_{\mathrm{B}m},\boldsymbol{x}_{\mathrm{R}m})}} \Big\{ & (R_{\mathrm{A}}, R_{\mathrm{B}}) \in \mathbb{R}_+^2 : \\
& t_1, \dots, t_6 \geq 0,\ \textstyle\sum_{m=1}^{6} t_m = 1,\ (\boldsymbol{C}_{\mathrm{A}m}, \boldsymbol{C}_{\mathrm{B}m}, \boldsymbol{C}_{\mathrm{R}m})_{m\in\{1,\dots,6\}} \in \mathcal{P}, \\
R_{\mathrm{A}} \leq\ & t_1 I(\boldsymbol{u}_{\mathrm{A}}; \boldsymbol{y}_{\mathrm{R1}}) + t_1 I(\boldsymbol{x}_{\mathrm{A1}}; \boldsymbol{y}_{\mathrm{B1}} | \boldsymbol{u}_{\mathrm{A}}) + t_3 I(\boldsymbol{x}_{\mathrm{A3}}; \boldsymbol{y}_{\mathrm{R3}} | \boldsymbol{x}_{\mathrm{B3}}) + t_6 I(\boldsymbol{x}_{\mathrm{A6}}; \boldsymbol{y}_{\mathrm{B6}} | \boldsymbol{x}_{\mathrm{R6}}), \\
R_{\mathrm{A}} \leq\ & t_1 I(\boldsymbol{x}_{\mathrm{A1}}; \boldsymbol{y}_{\mathrm{B1}}) + t_4 I(\boldsymbol{x}_{\mathrm{R4}}; \boldsymbol{y}_{\mathrm{B4}}) + t_6 I(\boldsymbol{x}_{\mathrm{A6}}, \boldsymbol{x}_{\mathrm{R6}}; \boldsymbol{y}_{\mathrm{B6}}), \\
R_{\mathrm{B}} \leq\ & t_2 I(\boldsymbol{u}_{\mathrm{B}}; \boldsymbol{y}_{\mathrm{R2}}) + t_2 I(\boldsymbol{x}_{\mathrm{B2}}; \boldsymbol{y}_{\mathrm{A2}} | \boldsymbol{u}_{\mathrm{B}}) + t_3 I(\boldsymbol{x}_{\mathrm{B3}}; \boldsymbol{y}_{\mathrm{R3}} | \boldsymbol{x}_{\mathrm{A3}}) + t_5 I(\boldsymbol{x}_{\mathrm{B5}}; \boldsymbol{y}_{\mathrm{A5}} | \boldsymbol{x}_{\mathrm{R5}}), \\
R_{\mathrm{B}} \leq\ & t_2 I(\boldsymbol{x}_{\mathrm{B2}}; \boldsymbol{y}_{\mathrm{A2}}) + t_4 I(\boldsymbol{x}_{\mathrm{R4}}; \boldsymbol{y}_{\mathrm{A4}}) + t_5 I(\boldsymbol{x}_{\mathrm{B5}}, \boldsymbol{x}_{\mathrm{R5}}; \boldsymbol{y}_{\mathrm{A5}}), \\
R_{\mathrm{A}} + R_{\mathrm{B}} \leq\ & t_1 I(\boldsymbol{u}_{\mathrm{A}}; \boldsymbol{y}_{\mathrm{R1}}) + t_1 I(\boldsymbol{x}_{\mathrm{A1}}; \boldsymbol{y}_{\mathrm{B1}} | \boldsymbol{u}_{\mathrm{A}}) + t_2 I(\boldsymbol{u}_{\mathrm{B}}; \boldsymbol{y}_{\mathrm{R2}}) + t_2 I(\boldsymbol{x}_{\mathrm{B2}}; \boldsymbol{y}_{\mathrm{A2}} | \boldsymbol{u}_{\mathrm{B}}) \\
& + t_3 I(\boldsymbol{x}_{\mathrm{A3}}, \boldsymbol{x}_{\mathrm{B3}}; \boldsymbol{y}_{\mathrm{R3}}) + t_5 I(\boldsymbol{x}_{\mathrm{B5}}; \boldsymbol{y}_{\mathrm{A5}} | \boldsymbol{x}_{\mathrm{R5}}) + t_6 I(\boldsymbol{x}_{\mathrm{A6}}; \boldsymbol{y}_{\mathrm{B6}} | \boldsymbol{x}_{\mathrm{R6}}) \Big\},
\end{aligned} \tag{7.45}$$

where the union is over all input distributions such that $p(\boldsymbol{x}_{A3}, \boldsymbol{x}_{B3}) = p(\boldsymbol{x}_{A3})p(\boldsymbol{x}_{B3})$, and where the auxiliary variables $\boldsymbol{u}_A$ and $\boldsymbol{u}_B$ must be chosen such that $\boldsymbol{u}_A \leftrightarrow \boldsymbol{x}_{A1} \leftrightarrow (\boldsymbol{y}_{B1}, \boldsymbol{y}_{R1})$ and $\boldsymbol{u}_B \leftrightarrow \boldsymbol{x}_{B2} \leftrightarrow (\boldsymbol{y}_{A2}, \boldsymbol{y}_{R2})$ form Markov chains, respectively. Unlike for $\mathcal{C}_{CSB}$ or $\mathcal{R}_{DF}$, we cannot simply invoke the entropy maximizing property of the Gaussian distribution to argue that $\mathcal{R}_{PDF}$ is attained by jointly Gaussian channel inputs. However, we can again use a channel enhancement argument to prove that the optimal channel inputs for the PDF strategy are jointly Gaussian as well.

Theorem 7.3. *For the restricted half-duplex two-way Gaussian MIMO relay channel with the per-phase or the average power constraint, the achievable partial decode-and-forward (PDF) rate region $\mathcal{R}_{PDF}$ is attained by jointly proper complex Gaussian channel inputs.*

Proof. First, consider the *aligned* half-duplex two-way Gaussian MIMO relay channel, which is characterized by $N_A = N_B = N_R = N$ and $\boldsymbol{H}_{AR} = \boldsymbol{H}_{AB} = \boldsymbol{H}_{BR} = \boldsymbol{H}_{BA} = \mathbf{I}_N$. Following the steps in the proof of Theorem 6.6, it can be shown that $\mathcal{R}_{PDF}^{\mathcal{N}_C} = \tilde{\mathcal{R}}_{PDF}$, where $\mathcal{R}_{PDF}^{\mathcal{N}_C}$ denotes a PDF rate region for the aligned relay channel that can be achieved with jointly proper complex Gaussian channel inputs, and where $\tilde{\mathcal{R}}_{PDF}$ is the achievable PDF rate region of an *enhanced* aligned half-duplex two-way Gaussian MIMO relay channel. This enhanced aligned relay channel is obtained by replacing $\boldsymbol{n}_{R1} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_R)$ and $\boldsymbol{n}_{R2} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_R)$ with $\tilde{\boldsymbol{n}}_{R1} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_1)$ and $\tilde{\boldsymbol{n}}_{R2} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Z}_2)$, respectively, where $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are such that $\mathbf{0} \prec \boldsymbol{Z}_1, \boldsymbol{Z}_2 \preccurlyeq \boldsymbol{Z}_R, \boldsymbol{Z}_D$. Since $\mathcal{R}_{PDF}^{\mathcal{N}_C} \subseteq \mathcal{R}_{PDF} \subseteq \tilde{\mathcal{R}}_{PDF}$ in general, it follows that jointly proper complex Gaussian channel inputs attain $\mathcal{R}_{PDF}$ if the half-duplex two-way Gaussian MIMO relay channel is aligned.

The theorem can therefore be proved by extending the above result to the general half-duplex two-way Gaussian MIMO relay channel. To this end, we can follow the steps in the proof of Theorem 6.7. In particular, note that any half-duplex two-way Gaussian MIMO relay channel can be described by a channel model with square channel gain matrices, and we can enhance $\boldsymbol{H}_{AR}$, $\boldsymbol{H}_{AB}$, $\boldsymbol{H}_{BR}$, and $\boldsymbol{H}_{BA}$ by adding small perturbations to their singular values such that the resulting channel gain matrices are invertible. Furthermore, we can show that the achievable PDF rate region $\mathcal{R}_{PDF}$ for the original relay channel can be obtained by a limit process on the achievable PDF rate region for the enhanced (perturbed) relay channel. □

The fact that $\mathcal{R}_{PDF}$ is attained by jointly proper complex Gaussian channel inputs implies that we may again represent the channel input distributions by the (joint) covariance matrices $\boldsymbol{C}_m$, $m \in \{1, \ldots, 6\}$. Moreover, we may decompose $\boldsymbol{x}_{A1}$ and $\boldsymbol{x}_{B2}$ as

$$\boldsymbol{x}_{A1} = \boldsymbol{u}_A + \boldsymbol{v}_A, \tag{7.46}$$

$$\boldsymbol{x}_{B2} = \boldsymbol{u}_B + \boldsymbol{v}_B \tag{7.47}$$

with $\boldsymbol{u}_A \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{U1})$ and $\boldsymbol{u}_B \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{U2})$ being independent of $\boldsymbol{v}_A \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{V1})$ and $\boldsymbol{v}_B \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{V2})$, respectively. Assuming the nodes are subject to the per-phase power constraint and the additive Gaussian noise is white, a parameterization of the achievable PDF rate region $\mathcal{R}_{\mathrm{PDF}}$ for the restricted half-duplex two-way Gaussian MIMO relay channel is thus given by

$$\mathcal{R}_{\mathrm{PDF}} = \left\{ \boldsymbol{r} \in \mathbb{R}_+^2 : \boldsymbol{A}\boldsymbol{r} \le \boldsymbol{s},\ \boldsymbol{s} \in \mathcal{S}_{\mathrm{PDF}} \right\}, \tag{7.48}$$

where

$$\begin{aligned} \mathcal{S}_{\mathrm{PDF}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = \textstyle\sum_{m=1}^{6} t_m \boldsymbol{s}_m, \quad \sum_{m=1}^{6} t_m = 1, \\ \boldsymbol{s}_m \in \mathcal{S}_{\mathrm{PDF}m},\ t_m \ge 0,\ \forall m \in \{1, \dots, 6\} \Big\} \end{aligned} \tag{7.49}$$

and

$$\begin{aligned} \mathcal{S}_{\mathrm{PDF1}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [s_1, s_2, 0, 0]^{\mathrm{T}}, \\ s_1 &\le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}(\boldsymbol{C}_{\mathrm{U1}} + \boldsymbol{C}_{\mathrm{V1}})\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right) \\ &\quad - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right) \\ &\quad + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right), \\ s_2 &\le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}(\boldsymbol{C}_{\mathrm{U1}} + \boldsymbol{C}_{\mathrm{V1}})\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right), \\ &\boldsymbol{C}_{\mathrm{U1}}, \boldsymbol{C}_{\mathrm{V1}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U1}} + \boldsymbol{C}_{\mathrm{V1}}) \le P_{\mathrm{A}} \Big\}, \end{aligned} \tag{7.50}$$

$$\begin{aligned} \mathcal{S}_{\mathrm{PDF2}} = \Big\{ \boldsymbol{s} \in \mathbb{R}_+^4 : \boldsymbol{s} = [0, 0, s_3, s_4]^{\mathrm{T}}, \\ s_3 &\le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}(\boldsymbol{C}_{\mathrm{U2}} + \boldsymbol{C}_{\mathrm{V2}})\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right) \\ &\quad - \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right) \\ &\quad + \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BA}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BA}}^{\mathrm{H}}\right), \\ s_4 &\le \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BA}}(\boldsymbol{C}_{\mathrm{U2}} + \boldsymbol{C}_{\mathrm{V2}})\boldsymbol{H}_{\mathrm{BA}}^{\mathrm{H}}\right), \\ &\boldsymbol{C}_{\mathrm{U2}}, \boldsymbol{C}_{\mathrm{V2}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U2}} + \boldsymbol{C}_{\mathrm{V2}}) \le P_{\mathrm{B}} \Big\}, \end{aligned} \tag{7.51}$$

$\mathcal{S}_{\mathrm{PDF3}} = \mathcal{S}_{\mathrm{DF3}} \subseteq \mathcal{S}_{\mathrm{CSB3}}$, and $\mathcal{S}_{\mathrm{PDF}m} = \mathcal{S}_{\mathrm{DF}m} = \mathcal{S}_{\mathrm{CSB}m}$ for $m \in \{4, 5, 6\}$. However, we remark that this parameterization of $\mathcal{R}_{\mathrm{PDF}}$ is not a convex one because $\mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}_{\mathrm{PDF2}}$ are generally nonconvex, which in turn implies that $\mathcal{S}_{\mathrm{PDF}}$ need not be convex. In contrast to the CSOB region $\mathcal{C}_{\mathrm{CSB}}$ and the achievable DF rate region $\mathcal{R}_{\mathrm{DF}}$, the achievable PDF rate region $\mathcal{R}_{\mathrm{PDF}}$ can therefore not always be evaluated by means of the dual decomposition approach presented in Section 7.2.

7.4.1 Suboptimal Partial Decode-and-Forward Rate Regions

Instead of trying to evaluate $\mathcal{R}_{\mathrm{PDF}}$ for the general case, we hence focus on approaches that allow to determine suboptimal achievable PDF rate regions for the restricted half-duplex two-way Gaussian MIMO relay channel based on dual decomposition. For this purpose, we again apply the zero-forcing (ZF) PDF approach and the inner approximation algorithm (IAA), which were originally considered in Section 3.4.3.

7.4.1.1 Zero-Forcing PDF Approach

Note that the two rate regions $\mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}_{\mathrm{PDF2}}$ are generally not convex because of the terms $-\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right)$ and $-\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right)$ in the constraints on s_1 and s_3, respectively. These terms are due to the fact that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ must be considered as interference at the relay. In order to suppress the interference the relay would suffer from during phases 1 and 2, we introduce two ZF receive filters $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ at the relay and require that all possible realizations of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ satisfy $\boldsymbol{G}_1\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{v}_1=\mathbf{0}$ and $\boldsymbol{G}_2\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{v}_2=\mathbf{0}$, which is equivalent to

$$\boldsymbol{G}_1\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{G}_1^{\mathrm{H}}=\mathbf{0},\qquad \boldsymbol{G}_2\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\boldsymbol{G}_2^{\mathrm{H}}=\mathbf{0} \tag{7.52}$$

for $\boldsymbol{v}_1\sim\mathcal{N}_{\mathbb{C}}(\mathbf{0},\boldsymbol{C}_{\mathrm{V1}})$ and $\boldsymbol{v}_2\sim\mathcal{N}_{\mathbb{C}}(\mathbf{0},\boldsymbol{C}_{\mathrm{V2}})$. If the first condition is added to $\mathcal{S}_{\mathrm{PDF1}}$ and the second one to $\mathcal{S}_{\mathrm{PDF2}}$, we then obtain the two compact convex rate regions

$$\begin{aligned}\mathcal{S}_{\mathrm{ZF1}}=\Big\{\boldsymbol{s}\in\mathbb{R}_+^4:\ &\boldsymbol{s}=[s_1,s_2,0,0]^{\mathrm{T}},\\ &s_1\le\log\det\left(\mathbf{I}+\boldsymbol{G}_1\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{U1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{G}_1^{\mathrm{H}}\right)\\ &\qquad+\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right),\\ &s_2\le\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{AB}}(\boldsymbol{C}_{\mathrm{U1}}+\boldsymbol{C}_{\mathrm{V1}})\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right),\\ &\boldsymbol{C}_{\mathrm{U1}},\boldsymbol{C}_{\mathrm{V1}}\succcurlyeq\mathbf{0},\quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U1}}+\boldsymbol{C}_{\mathrm{V1}})\le P_{\mathrm{A}},\\ &\boldsymbol{G}_1\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{G}_1^{\mathrm{H}}=\mathbf{0}\Big\},\end{aligned} \tag{7.53}$$

$$\begin{aligned}\mathcal{S}_{\mathrm{ZF2}}=\Big\{\boldsymbol{s}\in\mathbb{R}_+^4:\ &\boldsymbol{s}=[0,0,s_3,s_4]^{\mathrm{T}},\\ &s_3\le\log\det\left(\mathbf{I}+\boldsymbol{G}_2\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{U2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\boldsymbol{G}_2^{\mathrm{H}}\right)\\ &\qquad+\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{BA}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BA}}^{\mathrm{H}}\right),\\ &s_4\le\log\det\left(\mathbf{I}+\boldsymbol{H}_{\mathrm{BA}}(\boldsymbol{C}_{\mathrm{U2}}+\boldsymbol{C}_{\mathrm{V2}})\boldsymbol{H}_{\mathrm{BA}}^{\mathrm{H}}\right),\\ &\boldsymbol{C}_{\mathrm{U2}},\boldsymbol{C}_{\mathrm{V2}}\succcurlyeq\mathbf{0},\quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U2}}+\boldsymbol{C}_{\mathrm{V2}})\le P_{\mathrm{B}},\\ &\boldsymbol{G}_2\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\boldsymbol{G}_2^{\mathrm{H}}=\mathbf{0}\Big\},\end{aligned} \tag{7.54}$$

where, without loss of generality, we have assumed that $\boldsymbol{G}_1\boldsymbol{G}_1^{\mathrm{H}}=\mathbf{I}_{M_1}$ and $\boldsymbol{G}_2\boldsymbol{G}_2^{\mathrm{H}}=\mathbf{I}_{M_2}$ for some $M_1,M_2\in\{1,\ldots,N_{\mathrm{R}}\}$.

Now, if we replace $\mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}_{\mathrm{PDF2}}$ with $\mathcal{S}_{\mathrm{ZF1}}$ and $\mathcal{S}_{\mathrm{ZF2}}$ in (7.49), respectively, the resulting set, which we term $\mathcal{S}_{\mathrm{ZF}}$, is convex. This is because $\mathcal{S}_{\mathrm{PDF3}},\ldots,\mathcal{S}_{\mathrm{PDF6}}$ are convex and because the convex sum of convex sets is convex, cf. [11, Section 2.3.2] again. As a consequence, we can conclude that

$$\mathcal{R}_{\mathrm{ZF}}=\left\{\boldsymbol{r}\in\mathbb{R}_+^2:\boldsymbol{A}\boldsymbol{r}\le\boldsymbol{s},\ \boldsymbol{s}\in\mathcal{S}_{\mathrm{ZF}}\right\} \tag{7.55}$$

is a convex rate region which can be achieved by means of the ZF PDF approach with relay receive filters $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$. In addition, it also follows that like for $\mathcal{C}_{\mathrm{CSB}}$ and $\mathcal{R}_{\mathrm{DF}}$, the boundary of this achievable rate region can be determined by solving WSR maximization problems of the form

$$\max_{\boldsymbol{r},\boldsymbol{s}} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{ZF}} \tag{7.56}$$

if the weight vector $\boldsymbol{w} \in \mathbb{R}_+^2$ is varied appropriately. Finally, since the interior of $\mathcal{S}_{\mathrm{ZF}}$ is nonempty, any such WSR maximization problem can again be solved using the dual decomposition approach presented in Section 7.2, and it is straightforward to verify that Remark 7.2 also applies to the ZF PDF scheme.

7.4.1.2 Inner Approximation Algorithm

Recall that the IAA is a general mathematical approach to tackle nonconvex optimization problems which is based on solving a sequence of approximating convex problems instead of the original nonconvex one. In particular, in every iteration of the IAA, the nonconvex inequality constraints are approximated by convex ones such that the constraint set of the approximating convex problem is contained in the original nonconvex constraint set, cf. Appendix A.1.

Now, suppose we would like to determine a point on the boundary of the achievable PDF rate region $\mathcal{R}_{\mathrm{PDF}}$ by solving the WSR rate maximization problem

$$\max_{\boldsymbol{r},\boldsymbol{s}} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}_{\mathrm{PDF}} \tag{7.57}$$

for some $\boldsymbol{w} \in \mathbb{R}_+^2$, but as discussed above, $\mathcal{S}_{\mathrm{PDF}}$ is nonconvex so that this optimization problem is nonconvex. Like for unidirectional communication in the (half-duplex constrained) Gaussian MIMO relay channel, however, we can apply the IAA to obtain a suboptimal solution to above problem. To this end, we approximate the convex functions $-\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right)$ and $-\log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right)$ in $\mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}_{\mathrm{PDF2}}$ by their first-order Taylor series around $\boldsymbol{C}_{\mathrm{V1}}$ and $\boldsymbol{C}_{\mathrm{V2}}$, respectively. More precisely, in iteration k of the IAA, we replace these two functions by their first-order Taylor series around $\boldsymbol{C}_{\mathrm{V1}}^{(k-1)}$ and $\boldsymbol{C}_{\mathrm{V2}}^{(k-1)}$, which are equal to[1]

$$c_1(\boldsymbol{C}_{\mathrm{V1}}^{(k-1)}) - \mathrm{tr}\Big(\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}^{(k-1)}\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\Big), \tag{7.58}$$

$$c_2(\boldsymbol{C}_{\mathrm{V2}}^{(k-1)}) - \mathrm{tr}\Big(\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}^{(k-1)}\boldsymbol{H}_{\mathrm{BR}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\Big), \tag{7.59}$$

cf. (3.91), and where $\boldsymbol{C}_{\mathrm{V1}}^{(k-1)}$ and $\boldsymbol{C}_{\mathrm{V2}}^{(k-1)}$ are the optimizers of the approximating convex optimization problem in iteration $k-1$.

Assuming all three nodes are subject to the per-phase power constraint, this means that the two nonconvex rate regions $\mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}_{\mathrm{PDF2}}$, which specify the contributions of phases 1 and 2 to the achievable PDF rate region $\mathcal{R}_{\mathrm{PDF}}$, respectively, are approximated

[1] We remark that $c_1(\boldsymbol{C}_{\mathrm{V1}}^{(k-1)})$ and $c_2(\boldsymbol{C}_{\mathrm{V2}}^{(k-1)})$, respectively, collect all terms of the Taylor series that do not depend on $\boldsymbol{C}_{\mathrm{V1}}$ and $\boldsymbol{C}_{\mathrm{V2}}$, cf. (3.92). Besides, we again assume $\log = \log_{\mathrm{e}}$ for simplicity.

by the two convex rate regions

$$\begin{aligned}\mathcal{S}^{(k)}_{\mathrm{IAA1}} = \Big\{ & \boldsymbol{s} \in \mathbb{R}^4_+ : \boldsymbol{s} = [s_1, s_2, 0, 0]^{\mathrm{T}}, \\ & s_1 \le \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}(\boldsymbol{C}_{\mathrm{U1}} + \boldsymbol{C}_{\mathrm{V1}})\boldsymbol{H}^{\mathrm{H}}_{\mathrm{AR}}\Big) \\ & \quad + \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{V1}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{AB}}\Big) + c(\boldsymbol{C}^{(k-1)}_{\mathrm{V1}}) \\ & \quad - \mathrm{tr}\Big(\boldsymbol{H}^{\mathrm{H}}_{\mathrm{AR}}\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}^{(k-1)}_{\mathrm{V1}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{AR}}\Big)^{-1}\boldsymbol{H}_{\mathrm{AR}}\boldsymbol{C}_{\mathrm{V1}}\Big), \\ & s_2 \le \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}(\boldsymbol{C}_{\mathrm{U1}} + \boldsymbol{C}_{\mathrm{V1}})\boldsymbol{H}^{\mathrm{H}}_{\mathrm{AB}}\Big), \\ & \boldsymbol{C}_{\mathrm{U1}}, \boldsymbol{C}_{\mathrm{V1}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U1}} + \boldsymbol{C}_{\mathrm{V1}}) \le P_{\mathrm{A}} \Big\},\end{aligned} \tag{7.60}$$

$$\begin{aligned}\mathcal{S}^{(k)}_{\mathrm{IAA2}} = \Big\{ & \boldsymbol{s} \in \mathbb{R}^4_+ : \boldsymbol{s} = [0, 0, s_3, s_4]^{\mathrm{T}}, \\ & s_3 \le \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}(\boldsymbol{C}_{\mathrm{U2}} + \boldsymbol{C}_{\mathrm{V2}})\boldsymbol{H}^{\mathrm{H}}_{\mathrm{BR}}\Big) \\ & \quad + \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{BA}}\boldsymbol{C}_{\mathrm{V2}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{BA}}\Big) + c(\boldsymbol{C}^{(k-1)}_{\mathrm{V2}}) \\ & \quad - \mathrm{tr}\Big(\boldsymbol{H}^{\mathrm{H}}_{\mathrm{BR}}\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}^{(k-1)}_{\mathrm{V2}}\boldsymbol{H}^{\mathrm{H}}_{\mathrm{BR}}\Big)^{-1}\boldsymbol{H}_{\mathrm{BR}}\boldsymbol{C}_{\mathrm{V2}}\Big), \\ & s_4 \le \log\det\Big(\mathbf{I} + \boldsymbol{H}_{\mathrm{BA}}(\boldsymbol{C}_{\mathrm{U2}} + \boldsymbol{C}_{\mathrm{V2}})\boldsymbol{H}^{\mathrm{H}}_{\mathrm{BA}}\Big), \\ & \boldsymbol{C}_{\mathrm{U2}}, \boldsymbol{C}_{\mathrm{V2}} \succcurlyeq \mathbf{0}, \quad \mathrm{tr}(\boldsymbol{C}_{\mathrm{U2}} + \boldsymbol{C}_{\mathrm{V2}}) \le P_{\mathrm{B}} \Big\}.\end{aligned} \tag{7.61}$$

Furthermore, note that because the first-order Taylor series is a global underestimator for convex functions, cf. [11, Section 3.1.3], we have $\mathcal{S}^{(k)}_{\mathrm{IAA1}} \subseteq \mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}^{(k)}_{\mathrm{IAA2}} \subseteq \mathcal{S}_{\mathrm{PDF2}}$ for all $k \in \mathbb{N}$. The constraint set of the approximating convex optimization problem in iteration k is thus obtained as

$$\mathcal{R}^{(k)}_{\mathrm{IAA}} = \Big\{\boldsymbol{r} \in \mathbb{R}^2_+ : \boldsymbol{A}\boldsymbol{r} \le \boldsymbol{s},\ \boldsymbol{s} \in \mathcal{S}^{(k)}_{\mathrm{IAA}}\Big\} \subseteq \mathcal{R}_{\mathrm{PDF}}, \tag{7.62}$$

where

$$\begin{aligned}\mathcal{S}^{(k)}_{\mathrm{IAA}} = \Big\{\boldsymbol{s} \in \mathbb{R}^4_+ : \boldsymbol{s} = \textstyle\sum_{m=1}^{6} t_m \boldsymbol{s}_m, \quad \sum_{m=1}^{6} t_m = 1, \\ \boldsymbol{s}_m \in \mathcal{S}^{(k)}_{\mathrm{IAA}m},\ t_m \ge 0,\ \forall m \in \{1, \dots, 6\}\Big\} \subseteq \mathcal{S}_{\mathrm{PDF}}\end{aligned} \tag{7.63}$$

and $\mathcal{S}^{(k)}_{\mathrm{IAA}m} = \mathcal{S}_{\mathrm{PDF}m} = \mathcal{S}_{\mathrm{DF}m}$, $m \in \{3, \dots, 6\}$, for all $k \in \mathbb{N}$. As $\mathcal{S}^{(k)}_{\mathrm{IAA}}$ is a convex set with nonempty interior, it again follows that the approximating convex problem

$$\max_{\boldsymbol{r},\boldsymbol{s}}\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \le \boldsymbol{s}, \quad \boldsymbol{s} \in \mathcal{S}^{(k)}_{\mathrm{IAA}} \tag{7.64}$$

can be solved using the dual decomposition approach presented in Section 7.2.

Note that if $\boldsymbol{r}^{(k)}$ denotes an optimizer of this approximating convex problem, we have $\boldsymbol{r}^{(k)} \in \partial\mathcal{R}^{(k)}_{\mathrm{IAA}} \subseteq \mathcal{R}_{\mathrm{PDF}}$. Moreover, the sequence $\{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{r}^{(k)}\}$ is nondecreasing in k as $\boldsymbol{r}^{(k)} \in \mathcal{R}^{(k+1)}_{\mathrm{IAA}}$ for all $k \in \mathbb{N}$, which in turn implies that the sequence converges because $\mathcal{R}_{\mathrm{PDF}}$ is compact. Beyond that, it follows from Theorem A.1 that the IAA stops at a Karush–Kuhn–Tucker (KKT) point of the original nonconvex problem given in (7.57), or the limit of any convergent subsequence is a KKT point.

Since all rate vectors $\boldsymbol{r}^{(k)}$, $k \in \mathbb{N}$, generated by the IAA are element of $\mathcal{R}_{\mathrm{PDF}}$, the IAA allows to determine a suboptimal achievable PDF rate region by means of convex programming techniques. The obtained rate region of course depends on the initialization of the IAA as well as on the termination criterion. However, with an appropriate initialization of the IAA, e.g., by choosing $\boldsymbol{C}_{\mathrm{V1}}^{(0)} = \mathbf{0}$ and $\boldsymbol{C}_{\mathrm{V2}}^{(0)} = \mathbf{0}$ for any $\boldsymbol{w} \in \mathbb{R}_{+}^{2}$, we can ensure that the obtained PDF rate region contains $\mathcal{R}_{\mathrm{DF}}$.

7.4.2 Optimal Partial Decode-and-Forward Rate Regions for Special Cases

Like for unidirectional communication in the half-duplex constrained Gaussian MIMO relay channel, there exist some special cases for which it is possible to evaluate $\mathcal{R}_{\mathrm{PDF}}$ based on dual decomposition. In particular, note that whenever both $\mathcal{S}_{\mathrm{PDF1}}$ and $\mathcal{S}_{\mathrm{PDF2}}$ can be shown to be convex, it follows that $\mathcal{S}_{\mathrm{PDF}}$ and $\mathcal{R}_{\mathrm{PDF}}$ are convex. The WSR maximization problem given in (7.57) can then be solved using the dual decomposition approach derived in Section 7.2. In the following propositions, we state such sufficient conditions under which the set $\mathcal{S}_{\mathrm{PDF1}}$ is convex. Obviously, the results also apply to $\mathcal{S}_{\mathrm{PDF2}}$ if the roles of terminals A and B are reversed.

Proposition 7.4. *If $\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{AR}} \succcurlyeq \boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{B}}^{-1}\boldsymbol{H}_{\mathrm{AB}}$, then $\mathcal{S}_{\mathrm{PDF1}} = \mathcal{S}_{\mathrm{DF1}}$.*

Proof. This result follows from the fact that $I(\boldsymbol{u}_{\mathrm{A1}};\boldsymbol{y}_{\mathrm{R1}}) + I(\boldsymbol{x}_{\mathrm{A1}};\boldsymbol{y}_{\mathrm{B1}}|\boldsymbol{u}_{\mathrm{A1}}) \leq I(\boldsymbol{x}_{\mathrm{A1}};\boldsymbol{y}_{\mathrm{R1}})$ for all feasible $p(\boldsymbol{u}_{\mathrm{A1}}, \boldsymbol{x}_{\mathrm{A1}})$ if $\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{AR}} \succcurlyeq \boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{B}}^{-1}\boldsymbol{H}_{\mathrm{AB}}$, i.e., if $\boldsymbol{y}_{\mathrm{B1}}$ is a stochastically degraded version of $\boldsymbol{y}_{\mathrm{R1}}$, cf. Theorem 6.10. □

Proposition 7.5. *If $\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{AR}} \preccurlyeq \boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{B}}^{-1}\boldsymbol{H}_{\mathrm{AB}}$, then $\mathcal{S}_{\mathrm{PDF1}} = \mathcal{S}_{\mathrm{P2P1}}$ with*

$$\begin{aligned}\mathcal{S}_{\mathrm{P2P1}} = \Big\{\boldsymbol{s} \in \mathbb{R}_{+}^{4} : \boldsymbol{s} = [s_1, s_2, 0, 0]^{\mathrm{T}}&,\\ s_1 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{A1}}\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right),\\ s_2 &\leq \log\det\left(\mathbf{I} + \boldsymbol{H}_{\mathrm{AB}}\boldsymbol{C}_{\mathrm{A1}}\boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\right),\\ \boldsymbol{C}_{\mathrm{A1}} &\succcurlyeq \mathbf{0}, \quad \operatorname{tr}(\boldsymbol{C}_{\mathrm{A1}}) \leq P_{\mathrm{A}}\Big\}.\end{aligned} \tag{7.65}$$

Proof. Similarly, this result follows from the fact that $I(\boldsymbol{u}_{\mathrm{A1}};\boldsymbol{y}_{\mathrm{R1}}) + I(\boldsymbol{x}_{\mathrm{A1}};\boldsymbol{y}_{\mathrm{B1}}|\boldsymbol{u}_{\mathrm{A1}}) \leq I(\boldsymbol{x}_{\mathrm{A1}};\boldsymbol{y}_{\mathrm{B1}})$ for all feasible $p(\boldsymbol{u}_{\mathrm{A1}}, \boldsymbol{x}_{\mathrm{A1}})$ if $\boldsymbol{H}_{\mathrm{AR}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{R}}^{-1}\boldsymbol{H}_{\mathrm{AR}} \preccurlyeq \boldsymbol{H}_{\mathrm{AB}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{B}}^{-1}\boldsymbol{H}_{\mathrm{AB}}$, i.e., if $\boldsymbol{y}_{\mathrm{R1}}$ is a stochastically degraded version of $\boldsymbol{y}_{\mathrm{B1}}$, cf. Theorem 6.11. □

Remark 7.3. Note that $\mathcal{S}_{\mathrm{P2P1}} \subseteq \mathcal{S}_{\mathrm{PDF6}}$. That is, if the channel from terminal A to the relay is worse than the channel from terminal A to terminal B, phase 1 will never be used because its contribution to the achievable PDF rate region is always smaller than or equal to that of phase 6.

Remark 7.4. If we only consider unidirectional communication from terminal A to terminal B, Propositions 7.4 and 7.5 characterize the stochastically degraded and the reversely stochastically degraded half-duplex Gaussian MIMO relay channels.

Loosely speaking, we can thus state that a sufficient condition for $\mathcal{R}_{\mathrm{PDF}}$ to be convex is that for both directions of information transfer, the half-duplex constrained Gaussian MIMO relay channel is of stochastically degraded nature. Note that in this case, Remark 7.2 also applies to $\mathcal{R}_{\mathrm{PDF}}$. To conclude this section, we remark that like for unidirectional communication, the ZF PDF approach, the approach using the IAA, and Propositions 7.4 and 7.5 can be generalized to the case where the nodes are subject to the average power constraint or a combination of the per-phase and average power constraints.

7.5 Further Results and Bibliographical Notes

The restricted half-duplex two-way Gaussian relay channel with single-antenna nodes has been considered in many works, of which we only want to mention a few here. For example, Rankov and Wittneben [101, 102] initially studied the sum rate that can be achieved with the MABC protocol (cf. Figure 5.6(a)) and the relay using *decode-and-forward* (DF) or *amplify-and-forward* (AF). Furthermore, Wyrembelski et al. [141] considered the achievable DF rate region for the MABC protocol, and Kim et al. [70–72] compared various achievable rate regions for the MABC protocol, the TDBC protocol (cf. Figure 5.6(b)), and the HBC protocol (cf. Figure 5.6(c)) that are based on the relay using DF, AF, or *compress-and-forward* (CF). In [71], they also introduced a mixed forward scheme for the TDBC protocol, where the relay uses DF for one direction and CF for the other one, and a lattice DF scheme for the MABC protocol, where the relay decodes the sum of the terminals' codewords after the multiple-access phase. Moreover, outer bounds on the rate regions that are achievable with the MABC, TDBC, and HBC protocols were derived in [72].

The general *cut-set outer bound* (CSOB) region $\mathcal{C}_{\mathrm{CSB}}$ and the achievable DF rate region $\mathcal{R}_{\mathrm{DF}}$ for the half-duplex two-way single-antenna Gaussian relay channel were independently derived by Stein [118] and by Ashar et al. [5]. In addition, the latter also noted that at most four of the six optimal time-shares $t_1^\star, \dots, t_6^\star$ need to be positive for any point on the boundary of $\mathcal{C}_{\mathrm{CSB}}$. However, they used a different argument than we did in Remark 7.2 to arrive at this conclusion.

The vast majority of publications on the half-duplex two-way Gaussian MIMO relay channel has focused on the MABC protocol so far. Early on, rate regions that can be achieved with the MABC protocol and the relay using DF were for example considered by Hammerström et al. [55], and the same topic was later extensively studied by Bjelaković, Oechtering, Schnurr, Wyrembelski, and Boche [97, 98, 140, 141], with emphasis on the properties of the optimal channel inputs for the (bidirectional) broadcast phase. Many works have also addressed achievable AF rate regions for the MABC protocol. However, even for Gaussian channel inputs, difficult nonconvex optimization problems would have to be solved in order to jointly optimize the channel inputs of the relay and both terminals if all nodes are equipped with multiple antennas. As a consequence, the optimal AF strategy for the MABC protocol has only been obtained for single-antenna terminals [150], whereas for the general case, only suboptimal solutions have been derived, cf. [134, 142], for example.

The results on the general CSOB region $\mathcal{C}_{\mathrm{CSB}}$ and the achievable DF rate region $\mathcal{R}_{\mathrm{DF}}$ for the half-duplex two-way Gaussian MIMO relay channel are due to Gerdes, Riemensberger, and Utschick [41, 43]. In particular, like for unidirectional communication, the approach to evaluate $\mathcal{C}_{\mathrm{CSB}}$ and $\mathcal{R}_{\mathrm{DF}}$ in the dual domain was first derived for the per-phase power constraint in [41] and then generalized to the average power constraint in [43]. Beyond that, it was also shown by Gerdes, Riemensberger, and Utschick [42] how the convex parameterization of $\mathcal{R}_{\mathrm{DF}}$ and the dual decomposition approach can be used to efficiently solve utility maximization problems over $\mathcal{R}_{\mathrm{DF}}$ if the utility function is nondecreasing and concave. Note that in addition to the (weighted) sum rate, utility functions that satisfy these conditions include the utilities associated with max-min fairness [7], proportional fairness [65], and α-fairness [92], which is a generalization of the first two fairness measures.

Finally, we remark that none of the results for the *partial decode-and-forward* (PDF) strategy presented in Section 7.4 had been published so far. However, it is clear that these results are rather straightforward extensions of the corresponding results for unidirectional communication in the half-duplex constrained Gaussian MIMO relay channel, cf. Section 6.4.

Chapter 8

Numerical Results and Discussion

In this chapter, we provide numerical results for various half-duplex Gaussian MIMO relay channels in which we compare the achievable *decode-and-forward* (DF) and *partial decode-and-forward* (PDF) rates to each other as well as to the *cut-set bound* (CSB).[1] In particular, we investigate the effects of different antenna configurations and relay positions on the achievable rates and the CSB for both the per-phase and the average power constraint. Beyond that, we point out some practical applications for the results derived in Chapters 6 and 7, which were obtained based on the assumption that perfect channel state information (CSI) is available at all nodes.

After introducing the example scenario we use to generate the channel gain matrices in Section 8.1, numerical results are presented in Section 8.2. Like for the full-duplex case, we first compare C_{CSB} and $C_{\text{CSB,av}}$ to the maximum achievable DF rates R_{DF} and $R_{\text{DF,av}}$ in Section 8.2.1, respectively. The reason for this is again to identify channel conditions for which the DF strategy performs very well or very poorly. If $R_{\text{DF}} \approx C_{\text{CSB}}$ or $R_{\text{DF,av}} \approx C_{\text{CSB,av}}$, we (approximately) know the capacity of the corresponding half-duplex Gaussian MIMO relay channel, in which case there is no need to consider other relay strategies. On the other hand, the PDF strategy can potentially achieve much higher rates than the DF strategy if $R_{\text{DF}} \ll C_{\text{CSB}}$ or $R_{\text{DF,av}} \ll C_{\text{CSB,av}}$.

In Section 8.2.2, we then compare rates that can be achieved by means of the PDF strategy to $C_{\text{CSB}}/C_{\text{CSB,av}}$ and $R_{\text{DF}}/R_{\text{DF,av}}$, where we again distinguish between antenna configurations for which the half-duplex Gaussian MIMO relay channel has *disjoint sender components* and those for which it does not. Like for the full-duplex case, this is possible because for the considered example scenario, there is a necessary and sufficient condition in terms of N_{S}, N_{R}, and N_{D} that almost surely characterizes half-duplex Gaussian MIMO relay channels with disjoint sender components. What is more, we also derive conditions on N_{S}, N_{R}, and N_{D} which almost surely are necessary for the half-duplex Gaussian MIMO relay channel to be stochastically degraded or reversely stochastically degraded and which show that half-duplex Gaussian MIMO relay channels

[1] We remark that numerical results for the half-duplex two-way Gaussian MIMO relay channel are not presented in this work.

with disjoint sender components are not of stochastically degraded nature with probability one. For half-duplex Gaussian MIMO relay channels that are not of stochastically degraded nature, we employ the *zero-forcing* (ZF) approach presented in Section 6.4.3 to obtain achievable PDF rates.[2]

This chapter and the second part of this work eventually conclude in Section 8.3 with a discussion of practical applications for the results derived in Chapters 6 and 7. More specifically, we discuss how these results may be helpful in designing practical wireless communication systems although they were derived and are valid only for the idealistic case where perfect CSI is available at all nodes.

8.1 Example Scenario

The example scenario for the simulation results presented in this chapter is the same as for the full-duplex case. That is, we consider the line network that is depicted in Figure 4.1, where the distance between the source and the destination d_{SD} is normalized to one and where the relay is positioned on the line connecting the source and the destination such that $d_{\mathrm{SR}} = d$ and $d_{\mathrm{RD}} = 1 - d$ for some $d \in (0, 1)$. Furthermore, we assume uncorrelated Rayleigh fading, which means that the entries of the channel gain matrices are *independent and identically distributed* (i.i.d.) zero-mean proper (circularly symmetric) complex Gaussian random variables, and the variances of the channel gains are determined according to a simplified path loss model, cf. (4.1). Finally, we again assume that the additive white Gaussian noise is white, i.e., $\boldsymbol{Z}_{\mathrm{R}} = \mathbf{I}_{N_{\mathrm{R}}}$ and $\boldsymbol{Z}_{\mathrm{D}} = \mathbf{I}_{N_{\mathrm{D}}}$.

8.2 Numerical Results

As mentioned in the introduction to this chapter, we want to examine how the geometry of our three-terminal relay network and the numbers of source, relay, and destination antennas affect the achievable DF and PDF rates as compared to the CSB. To this end, we consider the same eight antenna configurations as for the full-duplex case, and for each of these antenna configurations, we again vary the distance between the source and the relay $d_{\mathrm{SR}} = d$ from 0.1 to 0.9. All further parameters remain constant and are chosen as follows, cf. Section 4.2:

- The path loss exponent is set to $\alpha = 4$, which is a typical value for urban macrocell environments or multi-level office buildings [50, Table 2.2].
- The source and the relay have equal power budgets given by $P_{\mathrm{S}} = P_{\mathrm{R}} = 10$, which is a reasonable assumption in wireless ad hoc networks, for example.
- All results are based on the same realizations of $(\tilde{\boldsymbol{H}}_{\mathrm{SR}}, \tilde{\boldsymbol{H}}_{\mathrm{SD}}, \tilde{\boldsymbol{H}}_{\mathrm{RD}})$ as for the full-duplex case. For each realization and every considered value of d, we determine the channel gain matrices by scaling $\tilde{\boldsymbol{H}}_{\mathrm{SR}}$, $\tilde{\boldsymbol{H}}_{\mathrm{SD}}$, and $\tilde{\boldsymbol{H}}_{\mathrm{RD}}$ according to (4.1).

[2] The *inner approximation algorithm* (IAA) is not considered in this chapter.

- For any given d, the rate values presented below are therefore averages over 1000 independent channel realizations again.

In order to evaluate $C_{\text{CSB}}/C_{\text{CSB,av}}$, $R_{\text{DF}}/R_{\text{DF,av}}$, and the corresponding achievable PDF rates, the optimization problems yielding these values were solved by means of the dual decomposition approaches proposed in Chapter 6. More precisely, the cutting-plane algorithm was applied to solve the dual problems (with $\delta = 10^{-2}$), where the master programs were solved using CVX [52, 53] with either SeDuMi [119] or SDPT3 [126, 127] and where the Lagrangian subproblems were solved using SDPT3.

8.2.1 Cut-Set Bound and Decode-and-Forward

8.2.1.1 Per-Phase Power Constraint

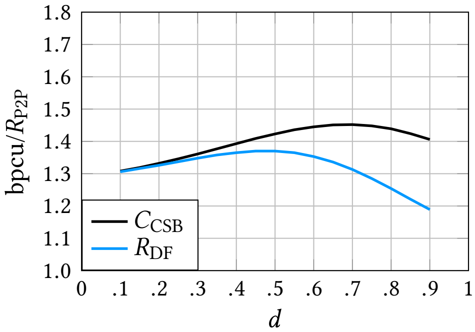
In Figure 8.1, we first compare the maximum achievable DF rate to the CSB and the capacity of the source-to-destination channel for the per-phase power constraint. We remark that like for the full-duplex case, the results for C_{CSB} and R_{DF} are normalized with respect to R_{P2P}, i.e., $R_{\text{DF}} = 1.0$ means that the maximum achievable DF rate is equal to the capacity of the source-to-destination channel, for example.

The most obvious difference between the results for the half-duplex Gaussian MIMO relay channel in Figure 8.1 and the corresponding results for the full-duplex case in Figure 4.2 is the range of values for C_{CSB} and R_{DF}. In particular, note that Figure 8.1 only shows values between 1.0 and 1.8, whereas in Figure 4.2, the range is exactly twice as large (0.8 to 2.4). The lower limit on the values in Figure 8.1 is easily explained by the fact that $R_{\text{DF}} \geq R_{\text{P2P}}$ for the half-duplex Gaussian MIMO relay channel, cf. (5.15). Furthermore, the reason why the largest values in Figure 8.1 are much smaller than in Figure 4.2 is because the (potential) gains of relaying are much smaller if the relay is forced to operate in half-duplex mode.[3]

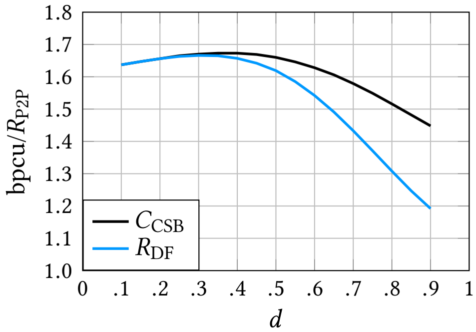
Apart from these quantitative differences, the results in Figure 8.1 and Figure 4.2 share several qualitative similarities. First, note that like for the full-duplex case, the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel with the per-phase power constraint approaches the CSB if d is sufficiently small, i.e., if the source-to-relay link is strong compared to the source-to-destination and relay-to-destination links. Of course, the meaning of "sufficiently small" again depends on N_{S}, N_{R}, and N_{D} (in addition to the fixed parameters P_{S}, P_{R}, and α), and compared to the full-duplex case, the relay must be closer to the source for R_{DF} to approach C_{CSB}. However, another similarity between the half- and full-duplex cases is that the source-to-relay channel eventually becomes the bottleneck if $N_{\text{R}} < N_{\text{D}}$, which for the half-duplex Gaussian MIMO relay channel means that $R_{\text{DF}}/R_{\text{P2P}} \to 1$ as $d \to 1$.

From Figure 8.1, we can also observe that both the maximum potential rate gain, i.e., the maximum of C_{CSB} over all considered values of d, and the corresponding optimal

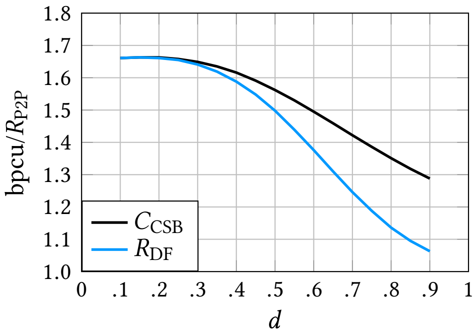
[3] Recall that for the results presented in Figure 4.2, it was assumed that the relay is able to perfectly cancel its self-interference.



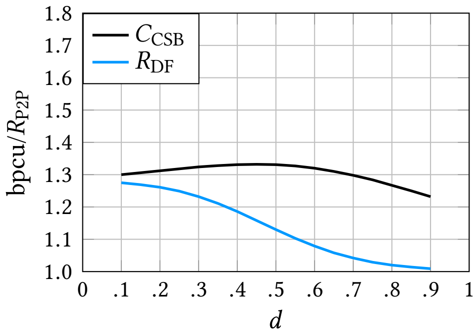
(a) $N_S = 2, N_R = 1, N_D = 1$



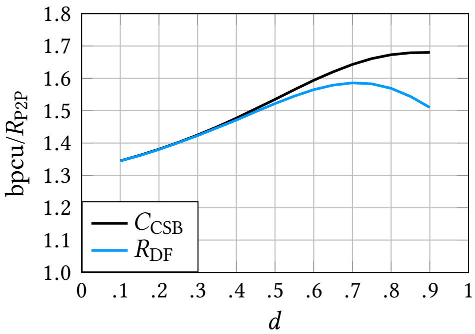
(b) $N_S = 2, N_R = 2, N_D = 2$



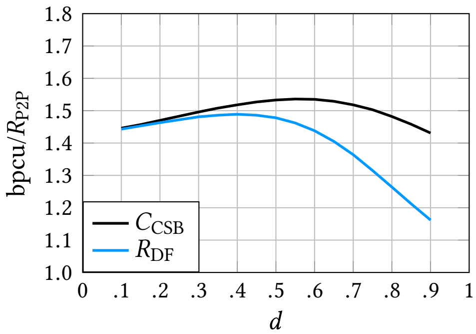
(c) $N_S = 2, N_R = 2, N_D = 3$



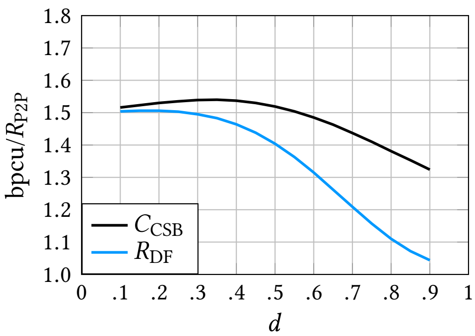
(d) $N_S = 3, N_R = 1, N_D = 2$



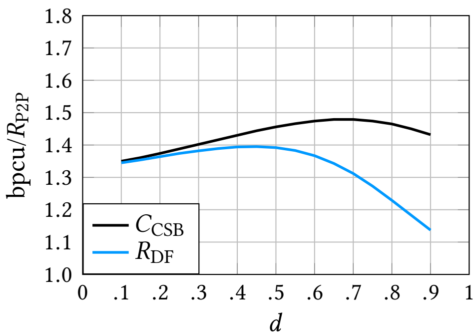
(e) $N_S = 3, N_R = 2, N_D = 1$



(f) $N_S = 3, N_R = 2, N_D = 2$



(g) $N_S = 3, N_R = 2, N_D = 3$



(h) $N_S = 4, N_R = 2, N_D = 2$

Figure 8.1: Comparison of C_{CSB} and R_{DF} for Half-Duplex Gaussian MIMO Relay Channels: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (per-phase power constraint, results averaged over 1000 independent channel realizations)

relay position strongly depend on the antenna configuration again. Obviously, the more antennas the relay is equipped with as compared to N_S and N_D, the larger the maximum potential rate gain. Moreover, like for the full-duplex case, the optimal distance between the source and the relay decreases with the ratio of $(1 + N_R/N_D)$ and $(1 + N_R/N_S)$ (cf. Section 4.2.1 for an explanation).

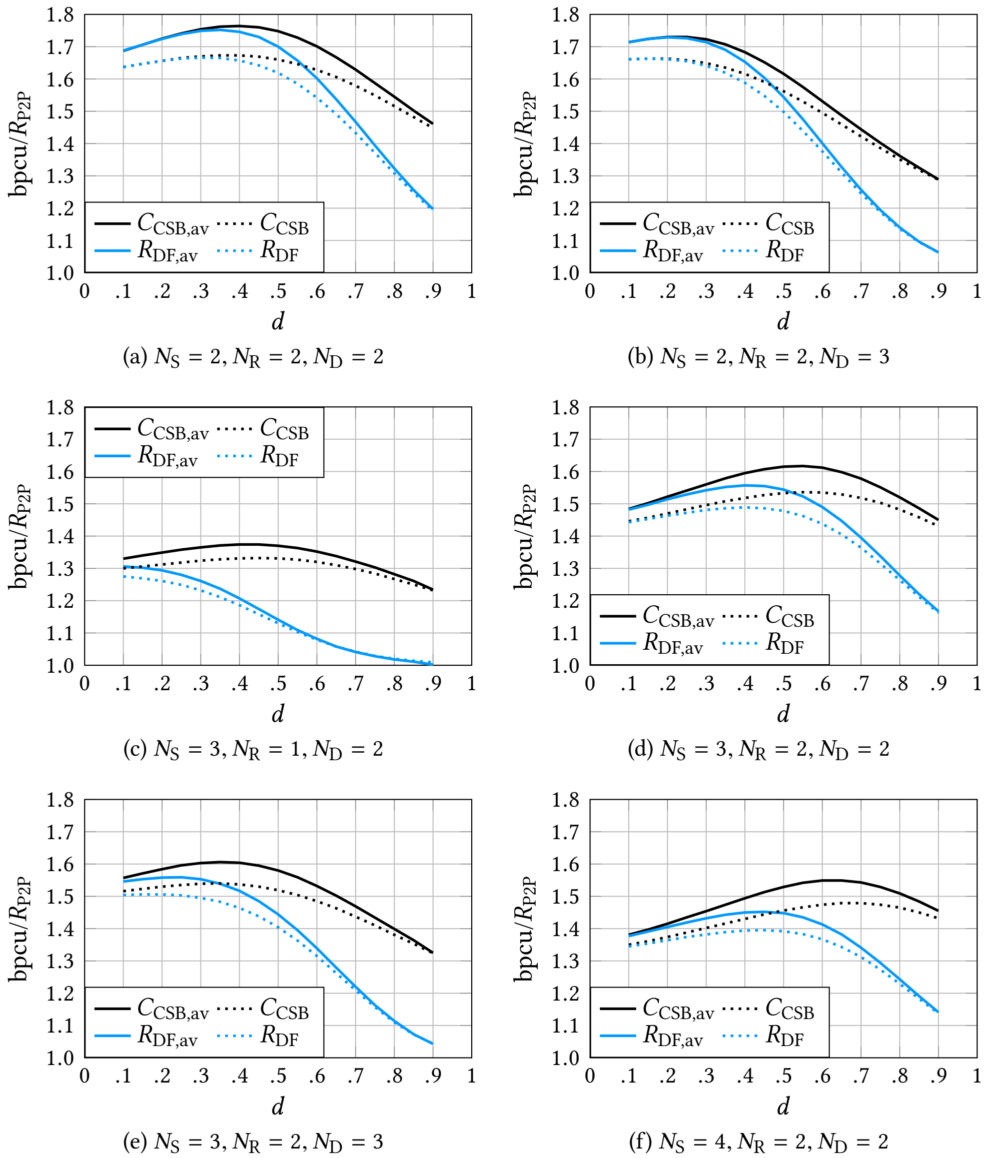
However, the most interesting similarity between the results presented in Figure 8.1 and Figure 4.2 is that if $N_S \leq N_R$, the maximum value of R_{DF} over all relay positions is also close to the maximum potential rate gain for the half-duplex Gaussian MIMO relay channel, cf. Figures 8.1(b) and 8.1(c). On the other hand, the gaps between the maximum values of R_{DF} and C_{CSB} are again quite large for the other six antenna configurations where $N_S > N_R$. If the source is equipped with more antennas than the relay, we can therefore reasonably expect that the PDF strategy can also achieve much higher rates than the DF strategy for the half-duplex Gaussian MIMO relay channel.

8.2.1.2 Average Power Constraint

In Figure 8.2, the maximum achievable DF rate is compared to the CSB and the capacity of the source-to-destination channel for the average power constraint. We remark that only six of the eight antenna configurations are considered for the average power constraint because the algorithms to compute $C_{CSB,av}$ and $R_{DF,av}$ proved to be numerically unstable for the two antenna configurations with $N_D = 1$. Moreover, note that for comparison, we have also plotted the CSB and the maximum achievable DF rate for the per-phase power constraint in Figure 8.2.

Since the per-phase power constraint is more restrictive than the average power constraint, cf. Section 6.1, it can be observed from Figure 8.2 that $C_{CSB,av} \geq C_{CSB}$ and $R_{DF,av} \geq R_{DF}$ for all antenna configurations and relay positions. The additional gains due to the less restrictive power constraint are up to 10% of R_{P2P}, and they are most pronounced around the relay positions where $C_{CSB,av}$ and $R_{DF,av}$ attain their respective maximum values. If the relay is placed close to the destination, however, the gaps between $C_{CSB,av}/R_{DF,av}$ and C_{CSB}/R_{DF} vanish. This is because the source-to-relay link, and thus the relay-receive phase, increasingly becomes the bottleneck of the information transfer as d approaches one. The optimal time-share of the relay-receive phase $t_1^\star$ then also approaches one so that the relay transmit power constraint has almost no effect on the optimal solution. Furthermore, the average transmit power constraint imposed on the source becomes $t_1^\star \operatorname{tr}(\boldsymbol{C}_{S1}) + t_2^\star \operatorname{tr}(\boldsymbol{C}_{S2}) \approx \operatorname{tr}(\boldsymbol{C}_{S1}) \leq P_S$ in this case, i.e., it basically amounts to a per-phase power constraint for the relay-receive phase.

Beyond that, we see from Figure 8.2 that the qualitative behavior of the results for the two different power constraints is remarkably similar. In particular, if the relay is close enough to the source, the maximum achievable DF rate approaches the CSB, regardless of whether we consider the per-phase or the average power constraint. Moreover, the relay positions at which the CSB and the maximum achievable DF rate attain their respective maximum values are almost the same for both power constraints, and like



(a) $N_S = 2, N_R = 2, N_D = 2$
(b) $N_S = 2, N_R = 2, N_D = 3$
(c) $N_S = 3, N_R = 1, N_D = 2$
(d) $N_S = 3, N_R = 2, N_D = 2$
(e) $N_S = 3, N_R = 2, N_D = 3$
(f) $N_S = 4, N_R = 2, N_D = 2$

Figure 8.2: Comparison of $C_{CSB,av}$ and $R_{DF,av}$ for Half-Duplex Gaussian MIMO Relay Channels: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (average power constraint, results averaged over 1000 independent channel realizations)

Table 8.1: Average Numbers of Cutting-Plane Iterations for Computing CSBs and Maximum Achievable DF Rates ($\delta = 10^{-2}$, results averaged over all relay positions and 1000 independent channel realizations)

	C_{CSB}	R_{DF}	$C_{\mathrm{CSB,av}}$	$R_{\mathrm{DF,av}}$
$N_{\mathrm{S}} = 2, N_{\mathrm{R}} = 2, N_{\mathrm{D}} = 2$	4.59	4.63	12.02	11.70
$N_{\mathrm{S}} = 2, N_{\mathrm{R}} = 2, N_{\mathrm{D}} = 3$	4.10	4.06	11.99	11.58
$N_{\mathrm{S}} = 3, N_{\mathrm{R}} = 1, N_{\mathrm{D}} = 2$	5.04	4.51	12.28	11.58
$N_{\mathrm{S}} = 3, N_{\mathrm{R}} = 2, N_{\mathrm{D}} = 2$	5.23	5.16	12.58	11.99
$N_{\mathrm{S}} = 3, N_{\mathrm{R}} = 2, N_{\mathrm{D}} = 3$	4.83	4.89	12.64	12.15
$N_{\mathrm{S}} = 4, N_{\mathrm{R}} = 2, N_{\mathrm{D}} = 2$	5.19	5.45	12.84	12.47

for the per-phase power constraint, there are considerable gaps between the maximum values of $C_{\mathrm{CSB,av}}$ and $R_{\mathrm{DF,av}}$ for all antenna configurations with $N_{\mathrm{S}} > N_{\mathrm{R}}$.

On the other hand, a major difference between the per-phase and the average power constraint is the computational effort that needs to be spent in order to evaluate the CSB $C_{\mathrm{CSB}}/C_{\mathrm{CSB,av}}$ or the maximum achievable DF rate $R_{\mathrm{DF}}/R_{\mathrm{DF,av}}$. In fact, Table 8.1 reveals that for all considered antenna configurations, the average numbers of iterations the cutting-plane algorithm required until convergence were more than twice as large for the average power constraint. This is because compared to the per-phase power constraint, two additional (scalar) dual variables are needed to formulate the dual problem if the nodes are subject to the average power constraint.

However, we remark that the numbers of required iterations are reasonably small for both power constraints. Consequently, the results in Table 8.1 also confirm that the dual decomposition approaches that were derived in Chapter 6 indeed allow to efficiently evaluate the CSB and the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint.

8.2.2 Partial Decode-and-Forward

Like for the full-duplex case, we can distinguish between antenna configurations for which the half-duplex Gaussian MIMO relay channel has disjoint sender components and those for which it does not. This is again possible because there exists a simple condition in terms of N_{S}, N_{R}, and N_{D} which almost surely characterizes half-duplex Gaussian MIMO relay channels with disjoint sender components for the considered example scenario with uncorrelated Rayleigh fading.

Definition 8.1. The half-duplex Gaussian MIMO relay channel is said to have *disjoint sender components* if $\mathrm{row}(\boldsymbol{H}_{\mathrm{S\{RD\}}}) = \mathrm{row}(\boldsymbol{H}_{\mathrm{SR}}) \oplus \mathrm{row}(\boldsymbol{H}_{\mathrm{SD}})$.

Proposition 8.1. *The half-duplex Gaussian MIMO relay channel has disjoint sender components if and only if* $\mathrm{rank}(\boldsymbol{H}_{\mathrm{SR}}) + \mathrm{rank}(\boldsymbol{H}_{\mathrm{SD}}) = \mathrm{rank}(\boldsymbol{H}_{\mathrm{S\{RD\}}})$.

Proof. Because the condition that defines the half-duplex Gaussian MIMO relay channel with disjoint sender components is the same as for the full-duplex case, the proof of this result is the same as the proof of Proposition 4.1. □

Proposition 8.2. *If the channel gain matrices are drawn from a continuous distribution, the probability that the half-duplex Gaussian MIMO relay channel has disjoint sender components is one if* $N_\mathrm{S} \geq N_\mathrm{R} + N_\mathrm{D}$ *and zero otherwise.*

Proof. Since the condition that defines the half-duplex Gaussian MIMO relay channel with disjoint sender components is the same as for the full-duplex case, the proof of this result is identical to the proof of Proposition 4.2. □

We remark that since the class of half-duplex Gaussian MIMO relay channels with disjoint sender components properly contains the class of half-duplex Gaussian MIMO relay channels with orthogonal sender components, Proposition 8.2 also implies that the source must have at least as many antennas as the relay and the destination combined for a half-duplex Gaussian MIMO relay channel to have orthogonal sender components. More importantly, however, this proposition, together with Proposition 8.3, implies that half-duplex Gaussian MIMO relay channels with disjoint sender components are almost surely not of stochastically degraded nature.

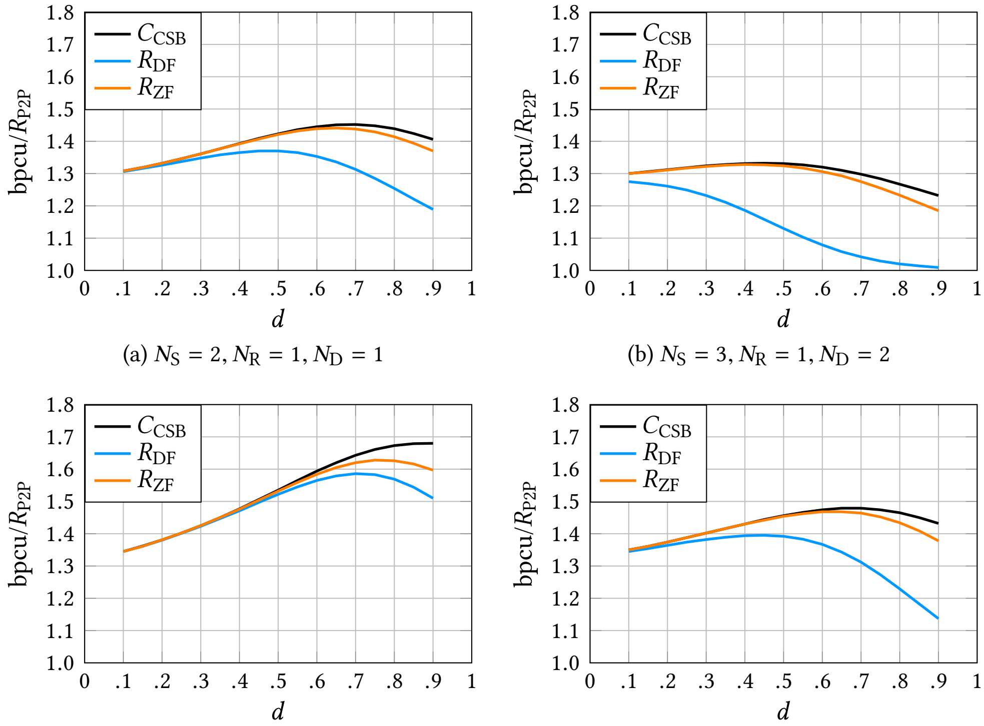
Proposition 8.3. *If the channel gain matrices are drawn from a continuous distribution, the half-duplex Gaussian MIMO relay channel almost surely is not stochastically degraded or reversely stochastically degraded if* $N_\mathrm{S} > N_\mathrm{R}$ *or* $N_\mathrm{S} > N_\mathrm{D}$*, respectively.*

Proof. Because the conditions that define the stochastically and the reversely stochastically degraded half-duplex Gaussian MIMO relay channels are the same as for the full-duplex case, the proof of this result is identical to the proof of Proposition 4.3. □

If the source is equipped with more antennas than the relay and the destination, it clearly follows from Proposition 8.3 that the corresponding half-duplex Gaussian MIMO relay channel is almost surely not of stochastically degraded nature. On the other hand, if $N_\mathrm{S} \leq N_\mathrm{R}$ and/or $N_\mathrm{S} \leq N_\mathrm{D}$, we need to check if the source-to-relay and source-to-destination channel gain matrices satisfy $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{H}_\mathrm{SR} \succcurlyeq \boldsymbol{H}_\mathrm{SD}^\mathrm{H}\boldsymbol{H}_\mathrm{SD}$ or $\boldsymbol{H}_\mathrm{SR}^\mathrm{H}\boldsymbol{H}_\mathrm{SR} \preccurlyeq \boldsymbol{H}_\mathrm{SD}^\mathrm{H}\boldsymbol{H}_\mathrm{SD}$. In these two cases, it follows from Theorems 6.10 and 6.11, respectively, that $R_\mathrm{PDF} = R_\mathrm{DF}$ and $R_\mathrm{PDF,av} = R_\mathrm{DF,av}$ or that $R_\mathrm{PDF} = R_\mathrm{PDF,av} = R_\mathrm{P2P}$. If neither of the two conditions is satisfied, however, the half-duplex Gaussian MIMO relay channel is not of stochastically degraded nature and we employ the zero-forcing (ZF) approach proposed in Section 6.4.3 to obtain achievable PDF rates.

8.2.2.1 Per-Phase Power Constraint

Let us first consider the per-phase power constraint and the antenna configurations where $N_\mathrm{S} \geq N_\mathrm{R}+N_\mathrm{D}$. For the corresponding half-duplex Gaussian MIMO relay channels,



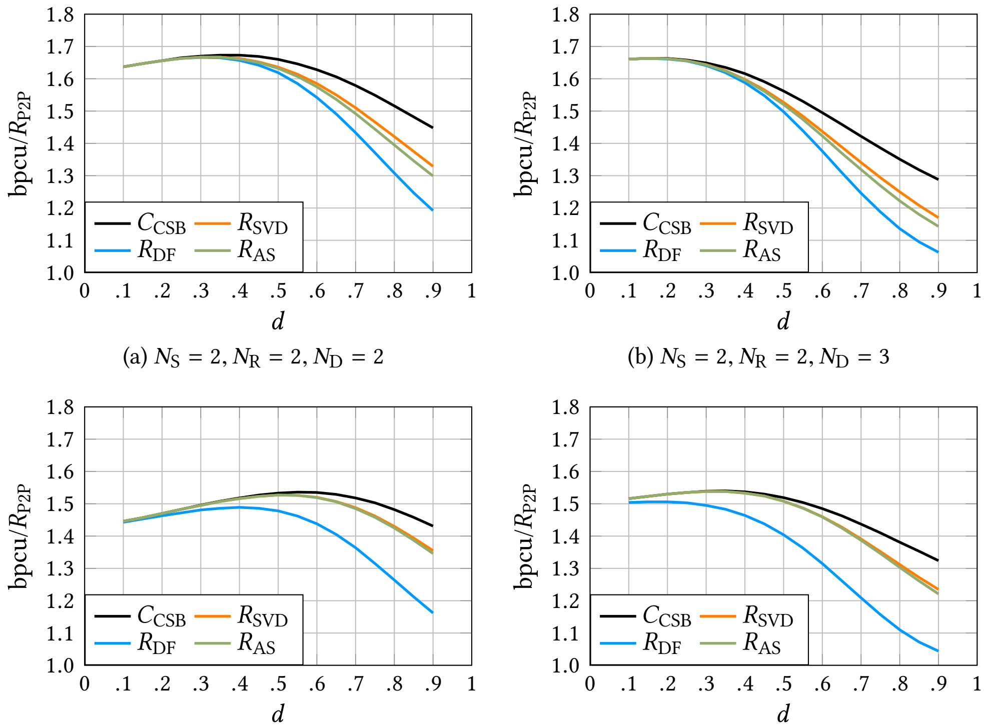
(a) $N_S = 2, N_R = 1, N_D = 1$ (b) $N_S = 3, N_R = 1, N_D = 2$

(c) $N_S = 3, N_R = 2, N_D = 1$ (d) $N_S = 4, N_R = 2, N_D = 2$

Figure 8.3: Comparison of C_{CSB}, R_{DF}, and R_{ZF} for Half-Duplex Gaussian MIMO Relay Channels with Disjoint Sender Components: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (per-phase power constraint, results averaged over 1000 independent channel realizations)

which almost surely have disjoint sender components and are not of stochastically degraded nature, we cannot evaluate the maximum achievable PDF rates. In Figure 8.3, we hence compare R_{ZF}, which is achieved by means of the ZF PDF approach with $\boldsymbol{G} = \mathbf{I}_{N_R}$, to R_{DF} and C_{CSB}, where the results are normalized with respect to R_{P2P}.

Note that like for the full-duplex case, cf. Figure 4.3, we can observe from Figure 8.3 that the suboptimal PDF scheme clearly outperforms the optimal DF scheme whenever R_{DF} does not already approach C_{CSB}. In fact, for three of the four considered antenna configurations, the basic ZF PDF scheme almost achieves the maximum potential rate gain, cf. Figures 8.3(a), 8.3(b), and 8.3(d). On the other hand, while the basic ZF PDF scheme also improves on the DF strategy for the fourth antenna configuration, cf. Figure 8.3(c), there are still considerable gaps between R_{ZF} and C_{CSB} for all relay positions where $R_{DF} \ll C_{CSB}$ (cf. Section 4.2.2 for a possible explanation). However, the results in Figure 8.3 confirm that much higher rates can be achieved by using PDF instead of DF for scenarios where $R_{DF} \ll C_{CSB}$ and $N_S \geq N_R + N_D$. Furthermore, the results show that the basic ZF PDF scheme almost achieves the maximum potential rate gain and that R_{ZF} generally comes close to C_{CSB} for at least some half-duplex Gaussian MIMO relay channels with disjoint sender components.



(a) $N_{\text{S}} = 2, N_{\text{R}} = 2, N_{\text{D}} = 2$

(b) $N_{\text{S}} = 2, N_{\text{R}} = 2, N_{\text{D}} = 3$

(c) $N_{\text{S}} = 3, N_{\text{R}} = 2, N_{\text{D}} = 2$

(d) $N_{\text{S}} = 3, N_{\text{R}} = 2, N_{\text{D}} = 3$

Figure 8.4: Comparison of C_{CSB}, R_{DF}, R_{AS}, and R_{SVD} for Half-Duplex Gaussian MIMO Relay Channels without Disjoint Sender Components: $P_{\text{S}} = 10$, $P_{\text{R}} = 10$, $\alpha = 4$ (per-phase power constraint, results averaged over 1000 independent channel realizations)

Now, in order to examine whether this only holds for half-duplex Gaussian MIMO relay channels with disjoint sender components, let us turn our attention to the antenna configurations where $N_{\text{S}} < N_{\text{R}} + N_{\text{D}}$. Since we do not know how to determine the maximum achievable PDF rates for all channel realizations, Figure 8.4 compares PDF rates that can be achieved using the ZF approach to C_{CSB} and R_{DF}. Like for the full-duplex case, the ZF filters we consider are based on antenna selection and the singular value decomposition (SVD) of $\boldsymbol{H}_{\text{SR}}$, cf. Section 3.4.3.

We remark that because $R_{\text{PDF}} \geq R_{\text{AS}}, R_{\text{SVD}} \geq R_{\text{DF}} \geq R_{\text{P2P}}$ in general, it follows that $R_{\text{AS}} = R_{\text{SVD}} = R_{\text{DF}}$ or $R_{\text{AS}} = R_{\text{SVD}} = R_{\text{DF}} = R_{\text{P2P}}$ for channel realizations for which the half-duplex Gaussian MIMO relay channel is stochastically degraded or reversely stochastically degraded, respectively. That is, if the half-duplex Gaussian MIMO relay channel is of stochastically degraded nature, both the ZF approach based on antenna selection and that based on the SVD of $\boldsymbol{H}_{\text{SR}}$ achieve $R_{\text{PDF}} = R_{\text{DF}}$. Moreover, we can again conclude that $\mathcal{A}^{\text{win}} = \mathcal{I}_{N_{\text{R}}}$ and $\mathcal{B}^{\text{win}} = \mathcal{I}_{\text{rank}(\boldsymbol{H}_{\text{SR}})}$ if the half-duplex relay channel is stochastically degraded, cf. Section 4.2.2. If it is reversely stochastically degraded, on the other hand, we have $R_{\text{PDF}} = R_{\text{P2P}}$, which means that R_{PDF} is attained by $t_1^\star = 0, t_2^\star = 1$. In this case, we define $\mathcal{A}^{\text{win}} = \mathcal{B}^{\text{win}} = \emptyset$ since R_{PDF} is achieved by not letting the relay decode any information that is conveyed from the source to the destination. If the half-duplex Gaussian MIMO relay channel is not of stochastically degraded nature, R_{AS} and R_{SVD} are obtained by considering all possible selections $\mathcal{A}$ and $\mathcal{B}$, cf. (3.79) and (3.85), respectively. Finally, note that for two of the four antenna configurations considered in Figure 8.4 and some selected values of d, the percentages of the winning selections $\mathcal{A}^{\mathrm{win}}$ and $\mathcal{B}^{\mathrm{win}}$ are presented in Tables 8.2 and 8.3.

Like for the full-duplex case, the most interesting observation from Figure 8.4 is that for the antenna configurations where $N_{\mathrm{S}} > N_{\mathrm{R}}$, cf. Figures 8.4(c) and 8.4(d), the results look similar to those for the half-duplex Gaussian MIMO relay channels with disjoint sender components in Figures 8.3(a), 8.3(b), and 8.3(d). More specifically, we see that the suboptimal ZF PDF scheme can again achieve much higher rates than the DF strategy if $R_{\mathrm{DF}} \ll C_{\mathrm{CSB}}$ and that both R_{AS} and R_{SVD} closely approach the maximum potential rate gain. Although for larger values of d the gaps between R_{AS} or R_{SVD} and C_{CSB} in Figures 8.4(c) and 8.4(d) are not quite as small as those between R_{ZF} and C_{CSB} in Figures 8.3(a), 8.3(b), and 8.3(d), we can thus conclude that the half-duplex Gaussian MIMO relay channel need not have disjoint sender components for the PDF strategy being able to outperform the DF strategy. Rather, it again suffices that the source is equipped with more antennas than the relay.

Furthermore, we remark that this conclusion is also supported by the results shown in Table 8.3. In particular, for $N_{\mathrm{S}} = 3$, $N_{\mathrm{R}} = N_{\mathrm{D}} = 2$, and $d = 0.5$, we see from Figure 8.4(c) that $R_{\mathrm{DF}} \ll R_{\mathrm{AS}} = R_{\mathrm{SVD}} \approx C_{\mathrm{CSB}}$ while Table 8.3(a) reveals that the maximum index set $\{1, 2\}$ is the optimal ZF selection for about 99% of the generated channel realizations. Hence, the differences between R_{DF} and R_{AS} or R_{SVD} are again almost entirely due to the fact that the condition $\mathrm{range}(\boldsymbol{C}_{\mathrm{V}}) \subseteq \mathrm{null}(\boldsymbol{H}_{\mathrm{SR}})$ does not imply $\boldsymbol{C}_{\mathrm{V}} = \boldsymbol{0}$ if $N_{\mathrm{S}} > N_{\mathrm{R}}$. Note that even for relay positions closer to the destination, this is the main reason why the suboptimal PDF scheme is able to outperform the DF strategy as the maximum index set remains the optimal ZF selection for the vast majority of the generated channel realizations, cf. Tables 8.3(b) and 8.3(c).

For the two antenna configuration where $N_{\mathrm{S}} \leq N_{\mathrm{R}}$, on the other hand, Figures 8.4(a) and 8.4(b) show that the ZF PDF approach can again not significantly outperform the DF strategy. Part of the reason for this is that R_{DF} already approaches C_{CSB} if d is small. Beyond that, however, we can see that the differences between R_{DF} and R_{AS} or R_{SVD} are smaller than those between C_{CSB} and R_{AS} or R_{SVD}, respectively, if $R_{\mathrm{DF}} < C_{\mathrm{CSB}}$. The main reason for this result is that the half-duplex Gaussian MIMO relay channel is very likely to be stochastically degraded if $N_{\mathrm{S}} \leq N_{\mathrm{R}}$. For example, Table 8.4 shows that for $N_{\mathrm{S}} = N_{\mathrm{R}} = N_{\mathrm{D}} = 2$ and $d = 0.5$, more than 75% of the generated channel realizations satisfied the stochastically degradedness condition.[4] In this case, however, we already know that $R_{\mathrm{PDF}} = R_{\mathrm{DF}}$, i.e., we know that PDF cannot outperform DF as both strategies achieve exactly the same rates.

[4]Note that the results in Table 8.4 are identical to those in Table 4.3 since the results for the full- and half-duplex Gaussian MIMO relay channels are based on the same 1000 channel realizations.

Table 8.2: Percentages of Winning Selections for Zero-Forcing PDF Schemes: $N_S = 2$, $N_R = 2$, $N_D = 2$ (results for per-phase power constraint, 1000 independent channel realizations)

(a) $d = 0.5$

sel.	$\mathcal{A}^{win}$	$\mathcal{B}^{win}$
$\emptyset$	0.1	0.1
$\{1\}$	7.8	21.8
$\{2\}$	9.2	0
$\{1, 2\}$	82.9	78.1

(b) $d = 0.7$

sel.	$\mathcal{A}^{win}$	$\mathcal{B}^{win}$
$\emptyset$	0.7	0.4
$\{1\}$	20.8	50.5
$\{2\}$	21.7	0.1
$\{1, 2\}$	56.8	49.0

(c) $d = 0.9$

sel.	$\mathcal{A}^{win}$	$\mathcal{B}^{win}$
$\emptyset$	5.4	3.9
$\{1\}$	33.3	73.5
$\{2\}$	34.9	0.5
$\{1, 2\}$	26.4	22.1

Table 8.3: Percentages of Winning Selections for Zero-Forcing PDF Schemes: $N_S = 3$, $N_R = 2$, $N_D = 2$ (results for per-phase power constraint, 1000 independent channel realizations)

(a) $d = 0.5$

sel.	$\mathcal{A}^{win}$	$\mathcal{B}^{win}$
$\emptyset$	0	0
$\{1\}$	0.4	1.3
$\{2\}$	0.1	0
$\{1, 2\}$	99.5	98.7

(b) $d = 0.7$

sel.	$\mathcal{A}^{win}$	$\mathcal{B}^{win}$
$\emptyset$	0	0
$\{1\}$	3.3	10.7
$\{2\}$	4.1	0
$\{1, 2\}$	92.6	89.3

(c) $d = 0.9$

sel.	$\mathcal{A}^{win}$	$\mathcal{B}^{win}$
$\emptyset$	0.6	0.2
$\{1\}$	12.9	35.0
$\{2\}$	13.6	0.2
$\{1, 2\}$	72.9	64.6

In addition, Table 8.2 shows that the percentages of the channel realizations for which $\mathcal{A}^{win} = \{1, 2\}$ or $\mathcal{B}^{win} = \{1, 2\}$ are even higher than P_{sto}. Recall that in contrast to the antenna configurations where $N_S > N_R$, the ZF condition range$(\boldsymbol{C}_V) \subseteq$ null$(\boldsymbol{H}_{SR})$ almost surely implies $\boldsymbol{C}_V = \mathbf{0}$ for the considered example scenario if $N_S \leq N_R$. As a result, $R_{AS} = R_{DF}$ and $R_{SVD} = R_{DF}$ if $\mathcal{A}^{win} = \{1, 2\}$ and $\mathcal{B}^{win} = \{1, 2\}$, respectively. Like for the full-duplex case, this applies for about 80% of the channel realizations when the relay is exactly in the middle between the source and the relay, cf. Table 8.2(a), and for $d = 0.7$, this percentage is still about 50%, cf. Table 8.2(b).

Table 8.4: Percentages of Channel Realizations for which the Half-Duplex Gaussian MIMO Relay Channel is Stochastically Degraded (P_{sto}) or Reversely Stochastically Degraded (P_{rev}): $N_S = 2$, $N_R = 2$, $N_D = 2$ (results for 1000 independent channel realizations)

d	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
P_{sto}	100	99.5	96.8	89.4	75.8	61.6	40.9	24.7	13.0
P_{rev}	0	0	0	0	0.1	0.2	0.3	0.5	2.3

Of course, the ZF PDF approach does not generally attain the maximum achievable PDF rate for the half-duplex Gaussian MIMO relay channel, and achievable PDF rates greater than R_{AS} and R_{SVD} may for example be obtained by more sophisticated ZF filter designs or the application of the IAA, cf. Section 4.2.2. However, the results presented in this section strongly suggest that for the considered Rayleigh fading environment and the per-phase power constraint, the PDF strategy can achieve significant gains as compared to the DF strategy if and only if $N_{\mathrm{S}} > N_{\mathrm{R}}$.

8.2.2.2 Average Power Constraint

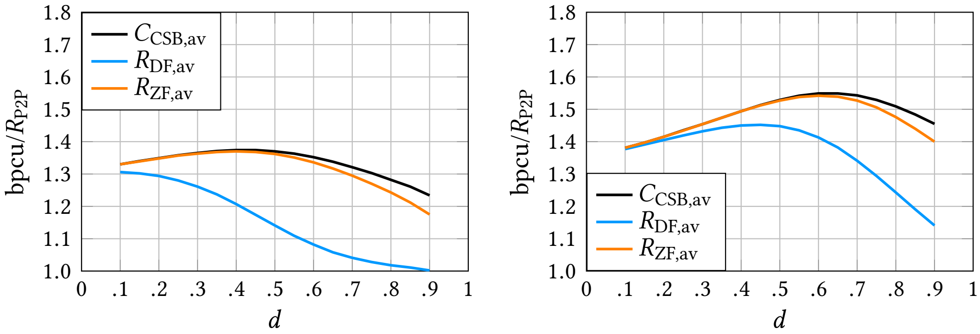
For the average power constraint and the antenna configurations where $N_{\mathrm{S}} \geq N_{\mathrm{R}} + N_{\mathrm{D}}$ and $N_{\mathrm{D}} > 1$, we compare $R_{\mathrm{ZF,av}}$ to $R_{\mathrm{DF,av}}$ and $C_{\mathrm{CSB,av}}$ in Figure 8.5. The results are normalized with respect to R_{P2P} again, and the two antenna configurations with $N_{\mathrm{D}} = 1$ are not considered here because the algorithms to compute $C_{\mathrm{CSB,av}}$, $R_{\mathrm{DF,av}}$, and $R_{\mathrm{ZF,av}}$ proved to be numerically unstable for this case.

Note that qualitatively the results in Figures 8.5(a) and 8.5(b) are very similar to the corresponding results for the per-phase power constraint in Figures 8.3(b) and 8.3(d), respectively. In particular, the basic ZF PDF scheme almost achieves the maximum potential rate gain, and $R_{\mathrm{ZF,av}}$ generally comes close to $C_{\mathrm{CSB,av}}$ for all considered relay positions so that the PDF strategy clearly outperforms the DF strategy whenever $R_{\mathrm{DF,av}} \ll C_{\mathrm{CSB,av}}$. For half-duplex Gaussian MIMO relay channels with disjoint sender components, the conclusions we can draw from these results hence are the same as those we were able to draw for the per-phase power constraint in the previous subsection.

Next, let us turn our attention to achievable PDF rates for half-duplex Gaussian MIMO relay channels that do not have disjoint sender components and where the nodes are subject to the average power constraint. To this end, we compare $R_{\mathrm{ZF,av}}$ to $R_{\mathrm{DF,av}}$ and $C_{\mathrm{CSB,av}}$ in Figure 8.6. We remark that Figure 8.6 only shows results for the antenna configurations where $N_{\mathrm{R}} < N_{\mathrm{S}} < N_{\mathrm{R}} + N_{\mathrm{D}}$ because if $N_{\mathrm{S}} \leq N_{\mathrm{R}}$, the ZF condition $\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} = \mathbf{0}$ implies $\boldsymbol{C}_{\mathrm{V}} = \mathbf{0}$ and thus $R_{\mathrm{ZF,av}} = R_{\mathrm{DF,av}}$ for the considered example scenario.

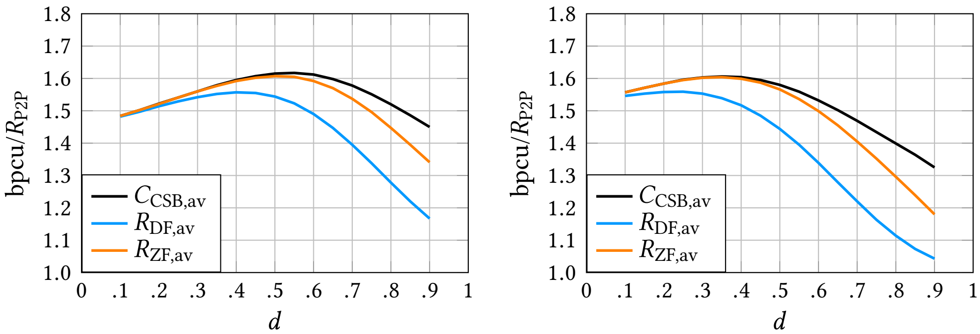
The results in Figure 8.6 again reveal that the PDF strategy can significantly outperform the DF strategy if the source is equipped with more antennas than the relay. In fact, we can again see that the basic ZF PDF scheme almost achieves the maximum potential rate gain, and for all relay positions where $R_{\mathrm{DF,av}} < C_{\mathrm{CSB,av}}$, the gaps between $R_{\mathrm{ZF,av}}$ and $C_{\mathrm{CSB,av}}$ are smaller than those between $R_{\mathrm{ZF,av}}$ and $R_{\mathrm{DF,av}}$. Since $R_{\mathrm{ZF,av}}$ is achieved by only considering the ZF filter $\boldsymbol{G} = \mathbf{I}_{N_{\mathrm{R}}}$, which corresponds to $\mathcal{A} = \mathcal{B} = \{1, 2\}$, the gains in comparison to $R_{\mathrm{DF,av}}$ are entirely due to the fact that $N_{\mathrm{S}} > N_{\mathrm{R}}$ as the ZF condition $\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{C}_{\mathrm{V}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} = \mathbf{0}$ does not imply $\boldsymbol{C}_{\mathrm{V}} = \mathbf{0}$ in this case.

Finally, we compare the computational effort that needs to be spent in order to evaluate the achievable PDF rates for the per-phase and the average power constraint in Table 8.5. Note that for simplicity, the results for R_{AS} and R_{SVD} are based only on the results for the winning selections $\mathcal{A}^{\mathrm{win}}$ and $\mathcal{B}^{\mathrm{win}}$, respectively. Like for the CSB and



(a) $N_S = 3, N_R = 1, N_D = 2$ (b) $N_S = 4, N_R = 2, N_D = 2$

Figure 8.5: Comparison of $C_{CSB,av}$, $R_{DF,av}$, and $R_{ZF,av}$ for Half-Duplex Gaussian MIMO Relay Channels with Disjoint Sender Components: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (average power constraint, results averaged over 1000 independent channel realizations)



(a) $N_S = 3, N_R = 2, N_D = 2$ (b) $N_S = 3, N_R = 2, N_D = 3$

Figure 8.6: Comparison of $C_{CSB,av}$, $R_{DF,av}$, and $R_{ZF,av}$ for Half-Duplex Gaussian MIMO Relay Channels without Disjoint Sender Components: $P_S = 10$, $P_R = 10$, $\alpha = 4$ (average power constraint, results averaged over 1000 independent channel realizations)

Table 8.5: Average Numbers of Cutting-Plane Iterations for Computing Achievable PDF Rates ($\delta = 10^{-2}$, results averaged over all relay positions and 1000 independent channel realizations)

	R_{ZF}	R_{SVD}	R_{AS}	$R_{ZF,av}$
$N_S = 2, N_R = 2, N_D = 2$	—	4.95	4.89	—
$N_S = 2, N_R = 2, N_D = 3$	—	4.59	4.49	—
$N_S = 3, N_R = 1, N_D = 2$	5.66	—	—	13.77
$N_S = 3, N_R = 2, N_D = 2$	—	5.19	5.18	13.39
$N_S = 3, N_R = 2, N_D = 3$	—	5.10	5.07	13.34
$N_S = 4, N_R = 2, N_D = 2$	5.55	—	—	14.01

the maximum achievable DF rate, cf. Table 8.1, we see that for all considered antenna configurations, the average numbers of iterations the cutting-plane algorithm required until convergence were more than twice as large for the average power constraint. However, the iteration numbers are again reasonably small for both power constraints, so the dual decomposition approaches that were derived in Chapter 6 also allow to efficiently evaluate achievable PDF rates for the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint.

8.3 Practical Applications

To round off the second part of this work, we want to discuss the practical usability of the theoretical results derived in Chapters 6 and 7 and the numerical results presented in this chapter. Since these results are based on the idealistic assumption that perfect CSI is available at all nodes, they can of course not directly be applied in practice. Nevertheless, there are several aspects that may be helpful in designing future relay-aided wireless communication systems.

8.3.1 Unidirectional Communication

By means of the dual decomposition approaches derived in Chapter 6, it is possible to efficiently determine upper and lower bounds on the capacity of the half-duplex Gaussian MIMO relay channel with the per-phase power constraint, the average power constraint, or a combination of both. In particular, we showed that convex optimization techniques can be used to evaluate the CSB and the maximum achievable DF rate for any half-duplex Gaussian MIMO relay channel and that for certain classes of half-duplex Gaussian MIMO relay channels, the same holds for the maximum achievable PDF rate. In addition, we showed how convex optimization techniques can always be employed to determine suboptimal PDF rates.

The most immediate application of these results is to obtain benchmarks for assessing the performance of other relay strategies, the influence of different MIMO channel models on the achievable rates, or the performance degradation due to imperfect CSI. From the numerical results presented in this chapter, we could for example conclude that for uncorrelated Rayleigh fading, using the PDF strategy instead of the simpler DF strategy is really beneficial only if the source is equipped with more antennas than the relay. Moreover, we observed that like for the full-duplex case, a rather simple ZF scheme can then achieve PDF rates close to the CSB. While we focused on how the achievable and potential rate gains are affected by different antenna configurations and relay positions, there are many other aspects that could be examined as well, e.g., the impact of different transmit power budgets on the rate gains or how critical the knowledge of phase information and the cooperation between the source and the relay are to the performance of the DF and PDF strategies.

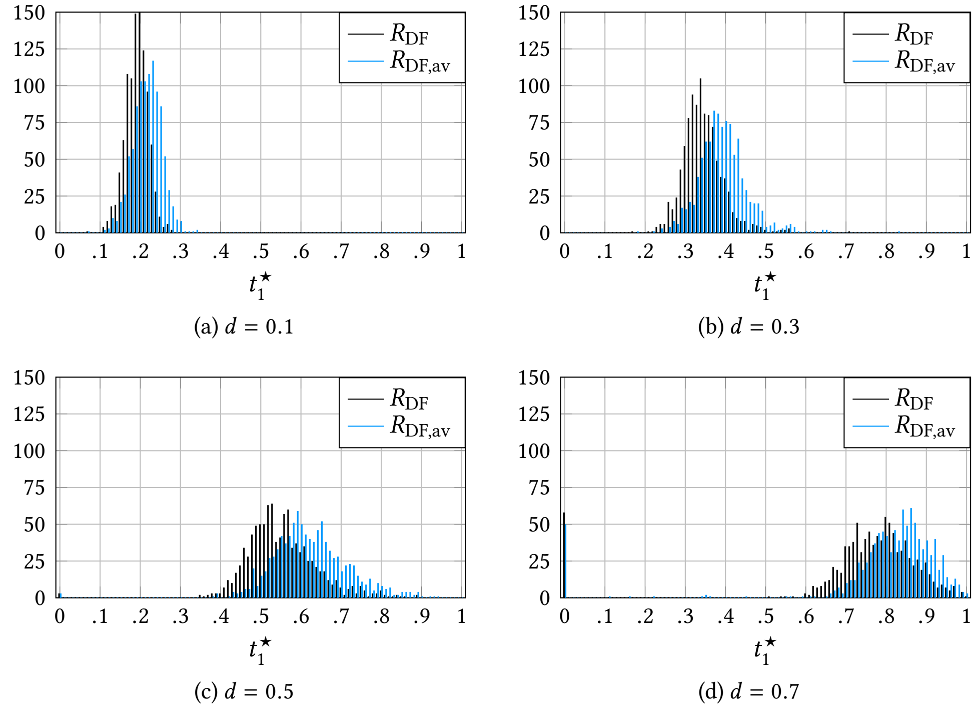


Figure 8.7: Histograms of Optimal Time-Shares for Maximum Achievable DF Rates: $N_S = 2$, $N_R = 2$, $N_D = 2$, $P_S = 10$, $P_R = 10$, $\alpha = 4$ (results for 1000 independent channel realizations)

For half-duplex relay channels where *time-division duplex* (TDD) is used to separate transmission and reception at the relay, a particularly interesting question is how long the relay should receive and transmit. In practical communication systems, the time-shares may not assume any value between zero and one, but rather a fixed number of bits will be used to represent the possible durations of the relay receive and transmit phases. In order to determine a suitable quantization, however, it would be very helpful if the distribution of the optimal time-shares were known. Using the results of Chapter 6, we can for example determine the empirical distribution of the optimal time-shares if the relay employs the DF strategy. For $N_S = N_R = N_D = 2$ and four different relay positions, the histograms of the time-shares that attain the maximum achievable DF rates R_{DF} and $R_{DF,av}$ are depicted in Figure 8.7. Even though the results are based on only 1000 channel realizations, they clearly show that the range and the distribution of the optimal time-shares strongly depend on the relay position/channel statistics and, to a lesser degree, on the considered power constraint.

Furthermore, note that once a quantization has been chosen, the results of Chapter 6 can also be used to quantify the rate loss due to the quantized time-shares. In particular, recall that for any given $t_1, t_2 \geq 0$, evaluating the achievable DF rates for the half-duplex Gaussian MIMO relay channel with the per-phase or the average power constraint only requires to solve one standard convex optimization problem, cf. (6.58) and (6.65), respectively. As a result, we can easily determine the DF rates that can be achieved with quantized time-shares and compare them to R_{DF} and $R_{\text{DF,av}}$.

8.3.2 Bidirectional Communication

The most immediate application of the results we derived for the case of bidirectional communication in the half-duplex constrained Gaussian MIMO relay channel is the same as for unidirectional communication. By means of the dual decomposition approaches proposed in Chapter 7, it is possible to determine both upper and lower bounds on the capacity region of the restricted half-duplex Gaussian MIMO relay channel. In particular, the *cut-set outer bound* (CSOB) region $\mathcal{C}_{\text{CSB}}$, the achievable DF rate region $\mathcal{R}_{\text{DF}}$, and the achievable PDF rate regions $\mathcal{R}_{\text{PDF}}/\mathcal{R}_{\text{ZF}}/\mathcal{R}_{\text{IAA}}$ we can evaluate may serve as benchmarks when studying other relay strategies, the influence of different MIMO channel models on the achievable rates, the performance degradation due to imperfect CSI, etc.

Beyond that, the results of Chapter 7 may be helpful in deciding which protocol(s) should be used for bidirectional communication in half-duplex constrained relay channels with TDD. For example, suppose the relay employs the DF strategy and we are interested in the maximum sum rate that can be achieved in the restricted half-duplex two-way Gaussian MIMO relay channel. From Remark 7.2, we know that at most four of the six phases illustrated in Figure 7.1 are needed to attain the maximum sum rate, but we do not a priori know which ones. However, this is not at all necessary as the solution to the general sum rate maximization problem (the WSR problem given in (7.41) with $\boldsymbol{w} = \mathbf{1}$), which can be obtained by means of the dual decomposition approach presented in Section 7.3, yields this information.

What is more, this dual decomposition may be applied to any utility maximization problem over $\mathcal{R}_{\text{DF}}$ with nondecreasing and concave utility function, and by setting some of the time-shares t_m, $m \in \{1, \dots, 6\}$, to zero, it may also be used to evaluate the performance of specific communication protocols like the MABC, TDBC, HBC, and OWTS protocols that have been considered in the literature. As a consequence, the results of Chapter 7 may for example be helpful in designing resource allocation protocols for a DF relay. In particular, an interesting question that could be addressed is whether there exist specific protocols using less than all six phases that perform well for a wide variety of scenarios.

Chapter 9

Conclusion

In this work, we examined the information theoretical performance limits of uni- and bidirectional communication in the (half-duplex constrained) Gaussian *multiple-input multiple-output* (MIMO) relay channel. As opposed to several other (multi-user) Gaussian MIMO channels for which characterizations of the capacity (regions) have been found, the capacity of the (half-duplex) Gaussian MIMO relay channel remains unknown in general. Therefore, we studied capacity upper bounds and achievable rates on the basis of the *cut-set bound* (CSB) and the *decode-and-forward* (DF) and *partial decode-and-forward* (PDF) strategies, respectively.

For the Gaussian MIMO relay channel, we showed in Chapter 3 that the CSB C_{CSB} and the maximum achievable DF rate R_{DF} can be determined as the solutions of convex optimization problems if perfect channel state information (CSI) is available at all nodes. To arrive at these results, we first made use of the fact that C_{CSB} and R_{DF} are attained by jointly proper complex Gaussian channel inputs and then exploited the Schur complement condition for positive semidefinite matrices. In Chapters 6 and 7, these results were also generalized to uni- and bidirectional communication in the half-duplex constrained Gaussian MIMO relay channel where the nodes use *time-division duplex* (TDD) to separate transmission and reception. In particular, we derived dual decomposition approaches that allow to efficiently solve the corresponding optimization problems with respect to the channel inputs and the time-shares allocated to the different phases of the TDD communication protocols in the Lagrangian dual domain.

Another important result of this work is that the optimal channel input distribution for the PDF strategy is proper complex Gaussian as well. In particular, for the Gaussian MIMO relay channel, we proved in Chapter 3 that the maximum achievable PDF rate R_{PDF} is attained by jointly proper complex Gaussian source and relay inputs. We remark that unlike for C_{CSB} and R_{DF}, the proof of this result was quite involved because the entropy maximizing property of the Gaussian distribution could not directly be applied. Moreover, even though R_{PDF} is attained by Gaussian channel inputs, the general PDF rate maximization problem as formulated in (3.72) or (3.73) is nonconvex. Rather than trying to solve this nonconvex problem, we considered two approaches which both yield suboptimal PDF rates that can be evaluated using standard convex optimization techniques. More specifically, the first approach was based on *zero-forcing* (ZF) the interference the relay would suffer from in the PDF strategy, whereas the second one used the so-called *inner approximation algorithm* (IAA). After all, we showed that for certain classes of relay channels, namely *stochastically degraded* and *reversely stochastically degraded* Gaussian relay channels, the maximum achievable PDF rate R_{PDF} can be determined as the solution of a convex rate maximization problem. Finally, like for the CSB and the DF strategy, Chapters 6 and 7 generalized these results to uni- and bidirectional communication in the half-duplex constrained Gaussian MIMO relay channel, respectively.

The most immediate practical application of the results we derived in Chapters 3, 6, and 7 certainly is to obtain benchmarks for assessing the performance of other relay strategies, the influence of different MIMO channel models on the achievable rates, the performance degradation due to imperfect or outdated CSI, etc. For example, from the numerical results presented in Chapters 4 and 8, we could conclude that the PDF strategy can achieve significant gains as compared to the DF strategy if and only if the source is equipped with more antennas than the relay. It must of course be mentioned that we only considered a scenario with uncorrelated Rayleigh fading. However, the results nevertheless suggest that using the PDF strategy instead of the simpler DF strategy may considerably improve the downlink performance of relay-aided cellular wireless communication systems (source ≡ base station), but not so much the uplink performance (source ≡ mobile terminal).

An interesting question that was not addressed in this work is how the *compress-and-forward* (CF) strategy compares to the DF and PDF strategies. As pointed out in Sections 3.5 and 6.5, the maximum achievable CF rate for the (half-duplex) Gaussian MIMO relay channel has yet to be determined because the optimal joint distribution of the channel inputs and the relay quantization is unknown. However, one could compare the achievable DF and PDF rates to the rates that are achievable with the suboptimal CF scheme proposed by Simoens et al. [117], which is based on choosing the channel inputs and the relay quantization to be jointly Gaussian and on determining the relay quantization according to the *information bottleneck method*, cf. Section 6.5. Since the CF strategy generally performs well if the relay-to-destination link is strong compared to the source-to-relay and source-to-destination links, it is reasonable to expect that this suboptimal CF scheme can at least outperform the (suboptimal) PDF scheme(s) if the relay is placed close to the destination.

Apart from that, it would of course be desirable to be able to evaluate the maximum achievable PDF rate R_{PDF} for any (half-duplex) Gaussian MIMO relay channel. However, as mentioned above, we did not attempt to solve the general PDF rate maximization problem in this work. In particular, a question we did not address is whether or not the primal decomposition approach that was the key to proving Theorems 3.6 and 6.6 may also be helpful for deriving an algorithm which is able to compute R_{PDF} in cases where the (half-duplex) Gaussian MIMO relay channel is not of stochastically degraded nature. To conclude this work, we want to briefly explain why there is a good reason to believe that the answer to this question is "yes".

If we define $\boldsymbol{S} = \boldsymbol{C}_{\mathrm{Q}} + \boldsymbol{C}_{\mathrm{V}}$ and apply the primal decomposition approach to the PDF rate maximization problem given in (3.73), then one of the resulting subproblems is a standard convex optimization problem. Furthermore, as noted in [57], the second subproblem we obtain is mathematically equivalent to the optimization problem that yields the sum capacity of the 2-user Gaussian MIMO broadcast channel under a shaping constraint. According to the minimax duality for linear conic constraints derived in [29], this sum capacity can equivalently be determined from a dual problem that is convex, cf. [57]. For any given shaping matrix $\boldsymbol{S} \succcurlyeq \mathbf{0}$, it hence follows that both subproblems, and thus the inner optimization problem, can efficiently be solved. If one can also find an efficient method to solve the outer problem, i.e., the maximization with respect to the shaping matrix $\boldsymbol{S}$, it will therefore be possible to solve the general PDF rate maximization problem by means of primal decomposition.

Appendix A

Algorithms

A.1 Inner Approximation Algorithm

The inner approximation algorithm (IAA) described in [88] is a general mathematical approach to deal with nonconvex optimization problems. More specifically, the IAA solves a sequence of approximating convex problems in order to locate a Karush–Kuhn–Tucker (KKT) point of the original nonconvex problem.

To explain how the IAA works, consider the following nonconvex optimization problem in epigraph form:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad g_i(\boldsymbol{x}) \leq 0, \forall i \in \mathcal{I}, \quad g_j(\boldsymbol{x}) \leq 0, \forall j \in \mathcal{J}. \tag{A.1}$$

Here, $f : \mathbb{C}^n \to \mathbb{R}$ is a differentiable objective function, and $\mathcal{I}$ and $\mathcal{J}$ are finite index sets such that $g_i : \mathbb{C}^n \to \mathbb{R}$, $i \in \mathcal{I}$, are differentiable convex functions and $g_j : \mathbb{C}^n \to \mathbb{R}$, $j \in \mathcal{J}$, are differentiable nonconvex functions. Furthermore, the constraint set

$$\mathcal{X} = \left\{\boldsymbol{x} \in \mathbb{C}^n : g_i(\boldsymbol{x}) \leq 0, \forall i \in \mathcal{I},\ g_j(\boldsymbol{x}) \leq 0, \forall j \in \mathcal{J}\right\} \tag{A.2}$$

is assumed to be compact (i.e., closed and bounded since $\mathcal{X} \subseteq \mathbb{C}^n$). In every iteration of the IAA, each nonconvex inequality constraint function is approximated by a convex function such that the constraint set of the approximating convex problem is contained in the original constraint set $\mathcal{X}$ (hence the name *inner* approximation). The detailed steps of the IAA are given in Algorithm 1.

Theorem A.1. *The IAA stops at a KKT point of the original nonconvex optimization problem given in* (A.1)*, or the limit of any convergent subsequence is a KKT point.*

Proof. See [88]. □

The key to the IAA and Theorem A.1 lies in how the nonconvex inequality constraint functions are approximated. In particular, since $g_j(\boldsymbol{x}) \leq \tilde{g}_j(\boldsymbol{x}; \boldsymbol{x}^{(k-1)})$ for any $\boldsymbol{x} \in \mathcal{X}^{(k-1)}$ and $j \in \mathcal{J}$, it follows that $\boldsymbol{x}^{(k)} \in \mathcal{X}^{(k-1)} \subseteq \mathcal{X}$, i.e., all points generated by the IAA are

Algorithm 1 Inner Approximation Algorithm (IAA)

Require: initial point $\boldsymbol{x}^{(0)} \in \mathcal{X}$

Set $p^{(0)} = f(\boldsymbol{x}^{(0)})$ and define $\mathcal{S}^{(0)} = \{\boldsymbol{x} \in \mathcal{X} : f(\boldsymbol{x}) = p^{(0)}\}$.

Repeat

Step 1: In iteration $k \in \mathbb{N}$, replace each inequality constraint $g_j(\boldsymbol{x}) \leq 0, j \in \mathcal{J}$, by a constraint $\tilde{g}_j(\boldsymbol{x}; \boldsymbol{x}^{(k-1)}) \leq 0$, where $\boldsymbol{x}^{(k-1)} \in \mathcal{S}^{(k-1)}$ and $\tilde{g}_j(\boldsymbol{x}; \boldsymbol{x}^{(k-1)})$ is a differentiable convex function such that

(i) $g_j(\boldsymbol{x}) \leq \tilde{g}_j(\boldsymbol{x}; \boldsymbol{x}^{(k-1)}), \forall \boldsymbol{x} \in \mathcal{X}^{(k-1)}$,

(ii) $g_j(\boldsymbol{x}^{(k-1)}) = \tilde{g}_j(\boldsymbol{x}^{(k-1)}; \boldsymbol{x}^{(k-1)})$,

(iii) $\nabla g_j(\boldsymbol{x}^{(k-1)}) = \nabla \tilde{g}_j(\boldsymbol{x}^{(k-1)}; \boldsymbol{x}^{(k-1)})$.

The constraint set $\mathcal{X}^{(k-1)} = \left\{\boldsymbol{x} \in \mathbb{C}^n : g_i(\boldsymbol{x}) \leq 0, \forall i \in \mathcal{I},\ \tilde{g}_j(\boldsymbol{x}; \boldsymbol{x}^{(k-1)}) \leq 0, \forall j \in \mathcal{J}\right\}$ of the approximating convex problem, which is a subset of $\mathcal{X}$ due to condition (i), must satisfy Slater's constraint qualification for convex optimization problems, cf. [6, Section 5.2] or [11, Section 5.2.3].

Step 2: Solve the approximating convex problem

$$p^{(k)} = \min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{X}^{(k-1)}$$

and let $\mathcal{S}^{(k)} = \{\boldsymbol{x} \in \mathcal{X} : f(\boldsymbol{x}) = p^{(k)}\}$.

until $p^{(k)} = p^{(k-1)}$.

Return $\boldsymbol{x}^{(k-1)}$, which is a KKT point of the original nonconvex optimization problem given in (A.1).

contained in the compact set $\mathcal{X}$. If $\boldsymbol{x}^{(k)}$ does not satisfy the termination criterion, then $p^{(k)} < p^{(k-1)}$ because $\boldsymbol{x}^{(k-1)} \in \mathcal{X}^{(k-1)}$. On the other hand, $p^{(k)} = p^{(k-1)}$ implies that $\boldsymbol{x}^{(k-1)}$ is a KKT point of the original nonconvex problem. To see this, note that $\boldsymbol{x}^{(k-1)}$ is a global optimizer of the k-th approximating convex problem in this case. As it is assumed that Slater's constraint qualification holds, this means that $\boldsymbol{x}^{(k-1)}$ must satisfy the KKT conditions of the approximating problem. But because $g_j(\boldsymbol{x}^{(k-1)}) = \tilde{g}_j(\boldsymbol{x}^{(k-1)}; \boldsymbol{x}^{(k-1)})$ and $\nabla g_j(\boldsymbol{x}^{(k-1)}) = \nabla \tilde{g}_j(\boldsymbol{x}^{(k-1)}; \boldsymbol{x}^{(k-1)})$, it then follows that $\boldsymbol{x}^{(k-1)}$ also satisfies the KKT conditions of the original nonconvex problem. Finally, Theorem A.1 is established by noting that Zangwill's convergence theorem [148, Convergence Theorem A] can be applied to the IAA.

Since convergence usually occurs only in the limit as $k \to \infty$, practical rules for terminating the IAA have to be defined in general, cf. [6, Section 7.2]. Such termination criteria may for example be based on the total or the relative improvement of the objective function over several iterations. To conclude this section, we remark that the IAA is explicitly mentioned in [16, Section 5.3.4] as one method to deal with nonconvex utility maximization problems in wireless communication sytems.

A.2 Cutting-Plane Method

A cutting-plane algorithm for solving certain convex optimization problems was first derived by Kelley [64] and Cheney and Goldstein [15]. The attractiveness of the cutting-plane method lies in the fact that it does not require any line searches and that the subproblem to be solved at each iteration is a linear program and thus simple to solve. Furthermore, we remark that various cutting-plane algorithms have been developed, cf. [37] and references therein, for example, and that the cutting-plane method belongs to the more general class of bundle methods, cf. [58, Chapter XV].

In order to explain the dual cutting-plane algorithm used in this work, consider the inequality constrained primal optimization problem

$$p^\star = \min_{\boldsymbol{x}} \; f(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}, \quad \boldsymbol{x} \in \mathcal{S}, \tag{A.3}$$

where $f : \mathbb{C}^n \to \mathbb{R}$ is the objective function, $\boldsymbol{g} : \mathbb{C}^n \to \mathbb{R}^m$ specifies m scalar inequality constraints, and $\mathcal{S} \subseteq \mathbb{C}^n$ denotes an abstract constraint set, which we assume to be closed. The (Lagrangian) dual problem associated with this primal problem is

$$d^\star = \max_{\boldsymbol{u}} \; \Theta(\boldsymbol{u}) \quad \text{s.t.} \quad \boldsymbol{u} \geq \boldsymbol{0}, \tag{A.4}$$

where

$$\Theta(\boldsymbol{u}) = \inf_{\boldsymbol{x} \in \mathcal{S}} \; L(\boldsymbol{x}, \boldsymbol{u}) \tag{A.5}$$

denotes the (Lagrangian) dual function and

$$L(\boldsymbol{x}, \boldsymbol{u}) = f(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}) \tag{A.6}$$

is the Lagrangian function. Any $\boldsymbol{u}$ for which $\Theta(\boldsymbol{u}) = -\infty$ is certainly not an optimizer of the dual problem, so we can equivalently express the dual problem as

$$d^\star = \max_{\boldsymbol{u}} \; \Theta(\boldsymbol{u}) \quad \text{s.t.} \quad \boldsymbol{u} \in \mathcal{U} = \{\boldsymbol{u} \in \mathbb{R}^m : \boldsymbol{u} \geq \boldsymbol{0}, \; \Theta(\boldsymbol{u}) > -\infty\}. \tag{A.7}$$

Using the fact that the dual function $\Theta : \mathbb{R}^m \to \mathbb{R}$ is concave [8, Proposition 6.2.1], it is straightforward to show that the set $\mathcal{U}$ is convex. Beyond that, we assume $\mathcal{U}$ to be closed in the following.[1]

Note that from the definitions of the dual function in (A.5) and the Lagrangian function in (A.6), it follows that

$$\Theta(\boldsymbol{u}) \leq f(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{S}. \tag{A.8}$$

[1] This assumption is not really a restriction. It is for example true if the set $\mathcal{S}$ is compact and the functions f and $\boldsymbol{g}$ are continuous, in which case $\Theta(\boldsymbol{u})$ is real-valued for all $\boldsymbol{u} \in \mathbb{R}^m$ and $\mathcal{U} = \{\boldsymbol{u} \in \mathbb{R}^m : \boldsymbol{u} \geq \boldsymbol{0}\} = \mathbb{R}_+^m$, cf. [8, Section 8.3].

By means of introducing an auxiliary variable z, the dual problem given in (A.7) can therefore be reformulated as

$$d^\star = \max_{z,\boldsymbol{u}} \; z \quad \text{s.t.} \quad z \le f(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{S}, \quad \boldsymbol{u} \in \mathcal{U}, \tag{A.9}$$

which is a linear program in the variables z and $\boldsymbol{u}$. However, this problem has infinitely many constraints (unless $\mathcal{S}$ is finite), and the constraints are not explicitly known.

The cutting-plane algorithm we use in this work is based on approximating the reformulated dual problem given in (A.9) and successively refining this approximation as follows. In particular, suppose we have the points $\boldsymbol{x}^{(0)}, \ldots, \boldsymbol{x}^{(k-1)} \in \mathcal{S}$. In the k-th iteration, we then solve the approximated dual problem

$$\max_{z,\boldsymbol{u}} \; z \quad \text{s.t.} \quad z \le f(\boldsymbol{x}^{(\ell)}) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}^{(\ell)}), \forall \ell \in \{0, \ldots, k-1\}, \quad \boldsymbol{u} \in \mathcal{U}, \tag{A.10}$$

which is also known as the *master program* of the cutting-plane algorithm. It is a linear program with a finite number of constraints, and it can for example be solved by means of the simplex algorithm [6, Section 2.7].

Now, let $(z^{(k)}, \boldsymbol{u}^{(k)})$ be an optimal solution to the master program in the k-th iteration. Then, we have $z^{(k)} \ge d^\star$, where $d^\star$ is the optimal value of the dual problem, because the approximated dual problem has less restrictive constraints than the original dual problem. In order to check whether $(z^{(k)}, \boldsymbol{u}^{(k)})$ is also an optimal solution to the original dual problem, we need to solve the *Lagrangian subproblem*

$$\min_{\boldsymbol{x}} \; f(\boldsymbol{x}) + \boldsymbol{u}^{(k),\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{S}, \tag{A.11}$$

i.e., we have to evaluate the dual function Θ at $\boldsymbol{u}^{(k)} \in \mathcal{U}$.[2] Let $\boldsymbol{x}^{(k)}$ be an optimal solution to above subproblem so that $\Theta(\boldsymbol{u}^{(k)}) = f(\boldsymbol{x}^{(k)}) + \boldsymbol{u}^{(k),\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}^{(k)})$. If $z^{(k)} \le \Theta(\boldsymbol{u}^{(k)})$, it follows that $(z^{(k)}, \boldsymbol{u}^{(k)})$ is an optimal solution to the (reformulated) dual problem given in (A.9). Otherwise, $(z^{(k)}, \boldsymbol{u}^{(k)})$ is no valid solution to the dual problem as it violates the constraint $z \le f(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})$ for $\boldsymbol{x} = \boldsymbol{x}^{(k)} \in \mathcal{S}$. In that case, we refine the approximation of the dual problem by imposing the additional constraint $z \le f(\boldsymbol{x}^{(k)}) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}^{(k)})$ on the master program of the next iteration, which is done by adding $\boldsymbol{x}^{(k)}$ to the current collection of points $\boldsymbol{x}^{(0)}, \ldots, \boldsymbol{x}^{(k-1)}$. The current optimizer of the master program $(z^{(k)}, \boldsymbol{u}^{(k)})$ does not fulfill the added constraint and is thus cut away from the feasible set, hence the name *cutting-plane* method.

Remark A.1. The initial point $\boldsymbol{x}^{(0)} \in \mathcal{S}$ should be strictly feasible, i.e., $\boldsymbol{g}(\boldsymbol{x}^{(0)}) < \boldsymbol{0}$. This is because otherwise the master program in the first iteration may become unbounded if $\boldsymbol{g}(\boldsymbol{x}^{(0)})$ has a positive component, e.g., when $\mathcal{U} = \mathbb{R}_+^m$, or any $\boldsymbol{u} \in \mathcal{U}$ is optimal if $\boldsymbol{g}(\boldsymbol{x}^{(0)}) = \boldsymbol{0}$. In both cases, the cutting-plane algorithm will not work properly.

[2] Assuming that f and $\boldsymbol{g}$ are continuous, we may replace inf by min in (A.11) as $\mathcal{S}$ was assumed to be closed and $\Theta(\boldsymbol{u}) > -\infty$ for any $\boldsymbol{u} \in \mathcal{U}$.

Algorithm 2 Cutting-Plane Algorithm

Require: initial point $\boldsymbol{x}^{(0)} \in \mathcal{S}$ such that $\mathbf{g}(\boldsymbol{x}^{(0)}) < \mathbf{0}$

Repeat

Step 1: In iteration $k \in \mathbb{N}$, solve the *master program*

$$z^{(k)} = \max_{z,\boldsymbol{u}} z \quad \text{s.t.} \quad z \le f(\boldsymbol{x}^{(\ell)}) + \boldsymbol{u}^{\mathrm{T}}\mathbf{g}(\boldsymbol{x}^{(\ell)}), \forall \ell \in \{0,\dots,k-1\}, \quad \boldsymbol{u} \in \mathcal{U}.$$

Step 2: Solve the *Lagrangian subproblem*

$$\Theta(\boldsymbol{u}^{(k)}) = \min_{\boldsymbol{x}} f(\boldsymbol{x}) + \boldsymbol{u}^{(k),\mathrm{T}}\mathbf{g}(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{S}.$$

until $z^{(k)} = \Theta(\boldsymbol{u}^{(k)})$.

Return $d^\star = z^{(k)}$ and $\boldsymbol{u}^{(k)}$.

The cutting-plane method can also be understood as follows. For any $\boldsymbol{x} \in \mathcal{S}$, the hyperplane given by

$$\mathcal{H}(\boldsymbol{x}) = \{(z,\boldsymbol{u}) : \boldsymbol{u} \in \mathcal{U},\ z = f(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}}\mathbf{g}(\boldsymbol{x})\} \tag{A.12}$$

bounds the dual function Θ from above. Moreover, the master program in iteration k is equivalent to the problem

$$\max_{\boldsymbol{u}} \hat{\Theta}_k(\boldsymbol{u}) \quad \text{s.t.} \quad \boldsymbol{u} \in \mathcal{U}, \tag{A.13}$$

where

$$\hat{\Theta}_k(\boldsymbol{u}) = \min_{\ell=0,\dots,k-1} f(\boldsymbol{x}^{(\ell)}) + \boldsymbol{u}^{\mathrm{T}}\mathbf{g}(\boldsymbol{x}^{(\ell)}) \tag{A.14}$$

is a piecewise linear function that provides an outer approximation, or more precisely, an outer linearization, of the concave dual function Θ by considering only k bounding hyperplanes (instead of infinitely many if $\mathcal{S}$ is not finite). In the k-th iteration, the master program of the cutting-plane algorithm hence uses k hyperplanes to approximate the dual function. With each additional hyperplane, the approximation becomes tighter until it is good enough so that the master program yields an optimal solution to the dual problem, cf. Algorithm 2.

In practice, however, this may only occur in the limit as $k \to \infty$ unless the dual function Θ is polyhedral and can be put into the form

$$\Theta(\boldsymbol{u}) = \min_{i\in\mathcal{I}} \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{u} + b_i, \tag{A.15}$$

where $\mathcal{I}$ is a finite index set and where $\boldsymbol{a}_i \in \mathbb{R}^m$ and $b_i \in \mathbb{R}$ are given vectors and scalars, respectively. The convergence properties of the cutting-plane algorithm are established in the following theorem:

Theorem A.2.

(a) Assume that $\{\boldsymbol{g}(\boldsymbol{x}^{(k)})\}$ is a bounded sequence. Then, every limit point of a sequence of dual variables $\{\boldsymbol{u}^{(k)}\}$ generated by the cutting-plane algorithm is a dual optimal solution.

(b) Assume that the dual function Θ is polyhedral of the form (A.15). *Then, the cutting-plane algorithm terminates finitely, i.e., for some $k \in \mathbb{N}$, $\boldsymbol{u}^{(k)}$ is a dual optimal solution.*

Proof. See [8, Proposition 8.3.1]. □

We remark that the boundedness assumption in Theorem A.2(a) can be replaced by the assumption that $\Theta(\boldsymbol{u})$ is real-valued for all $\boldsymbol{u} \in \mathbb{R}^m$, which can be ascertained if $\mathcal{S}$ is a finite set, or alternatively, if f and $\boldsymbol{g}$ are continuous and $\mathcal{S}$ is compact [8, Section 8.3].

If the dual function is not polyhedral, convergence usually occurs only in the limit as $k \to \infty$, which means that we need some practical termination criterion. To this end, note that $\{z^{(k)}\}$ is a nonincreasing sequence that upper bounds $d^\star$. The reason for this is that $\max_{\boldsymbol{u}\in\mathcal{U}} \Theta(\boldsymbol{u}) \leq \max_{\boldsymbol{u}\in\mathcal{U}} \hat{\Theta}_k(\boldsymbol{u}) \leq \max_{\boldsymbol{u}\in\mathcal{U}} \hat{\Theta}_{k-1}(\boldsymbol{u}), \forall k \in \mathbb{N}$, because any $\hat{\Theta}_k$ is an outer approximation of Θ and the approximations become tighter with increasing k. A lower bound on $d^\star$ is given by the nondecreasing sequence $\{y^{(k)}\}$, where $y^{(k)} = \max_{\ell=1,\ldots,k} \Theta(\boldsymbol{u}^{(\ell)})$. For any $k \in \mathbb{N}$, the optimal dual objective value is thus sandwiched between $y^{(k)}$ and $z^{(k)}$, i.e., we have

$$y^{(k)} \leq d^\star \leq z^{(k)}, \forall k \in \mathbb{N}, \tag{A.16}$$

and the difference $z^{(k)} - y^{(k)}$ is nonincreasing in k. Moreover, the sequence of dual function values $\{\Theta(\boldsymbol{u}^{(k)})\}$ converges to $d^\star$ so that $\lim_{k\to\infty} y^{(k)} = d^\star$, and it can be shown that $\lim_{k\to\infty} z^{(k)} = d^\star$ [58, Chapter XII, Theorem 4.2.3]. Therefore, we can terminate the cutting-plane algorithm after the k-th iteration if $z^{(k)} - y^{(k)} < \delta$, where $\delta > 0$ specifies a predefined accuracy.

A.2.1 Primal Recovery

So far, we have only explained a cutting-plane algorithm for solving a dual problem. However, the main concern usually is to find an optimal solution to the primal problem, not just solving the dual one. Consequently, we also need to discuss how the optimal primal variables can be obtained from the sequences of primal and dual variables that are generated by the cutting-plane algorithm, a process known as *primal recovery*. Note that for this purpose, we need to assume that *strong duality* holds, which means that the solutions to the primal problem and its dual problem are equivalent in the sense that their optimal objective values are equal.

In particular, suppose the functions $f : \mathbb{C}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{C}^n \to \mathbb{R}^m$ are convex and the abstract constraint set $\mathcal{S}$ is convex and closed. Let $\mathcal{X} = \{\boldsymbol{x} \in \mathcal{S} : \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}\}$ denote the feasible set of the primal optimization problem given in (A.3), which we assume to be nonempty, and let $\mathcal{X}^\star = \{\boldsymbol{x}^\star \in \mathcal{X} : \boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})\}$ be the set of primal optimal solutions, which is closed and convex. Finally, we assume that the interior of $\mathcal{X}$ is nonempty, i.e., $\operatorname{int}(\{\boldsymbol{x} \in \operatorname{int}(\mathcal{S}) : \boldsymbol{g}(\boldsymbol{x}) < \boldsymbol{0}\}) \neq \emptyset$. Then, Slater's constraint qualification is satisfied and strong duality holds, i.e., $p^\star = d^\star$ [11, Section 5.3.2].

Now, suppose the cutting-plane algorithm is applied to the dual problem given in (A.7). If it is properly initialized, i.e., if $\boldsymbol{x}^{(0)} \in \mathcal{S}$ such that $\boldsymbol{g}(\boldsymbol{x}^{(0)}) < \boldsymbol{0}$, the cutting-plane algorithm generates a sequence $\{\boldsymbol{u}^{(k)}\}$ of dual feasible points as well as a corresponding sequence $\{\boldsymbol{x}^{(k)}\}$ of primal variables. At iteration k, we can then obtain a primal feasible point $\hat{\boldsymbol{x}}^{(k)} \in \operatorname{conv}(\{\boldsymbol{x}^{(0)}, \ldots, \boldsymbol{x}^{(k-1)}\})$ by solving the following linear program, which is closely related to the master program of the k-th iteration:

$$\begin{aligned} \min_{\lambda_0,\ldots,\lambda_{k-1}} \sum_{\ell=0}^{k-1} \lambda_\ell f(\boldsymbol{x}^{(\ell)}) \quad \text{s.t.} \quad & \sum_{\ell=0}^{k-1} \lambda_\ell \boldsymbol{g}(\boldsymbol{x}^{(\ell)}) \leq \boldsymbol{0}, \\ & \sum_{\ell=0}^{k-1} \lambda_\ell = 1, \quad \lambda_\ell \geq 0, \forall \ell \in \{0,\ldots,k-1\}. \end{aligned} \tag{A.17}$$

In fact, this linear program is the dual problem of the master program given in (A.10), as the following considerations show. Dualizing the constraints $z \leq f(\boldsymbol{x}^{(\ell)}) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}^{(\ell)})$, $\ell \in \{0,\ldots,k-1\}$, of the master program yields the dual function

$$\begin{aligned} \Theta_{\mathrm{MP}}(\boldsymbol{\lambda}) &= \sup_{z,\boldsymbol{u}\geq\boldsymbol{0}} z - \sum_{\ell=0}^{k-1} \lambda_\ell \left(z - f(\boldsymbol{x}^{(\ell)}) - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}^{(\ell)})\right) \\ &= \begin{cases} \sum_{\ell=0}^{k-1} \lambda_\ell f(\boldsymbol{x}^{(\ell)}) & \text{if } \sum_{\ell=0}^{k-1} \lambda_\ell = 1,\ \sum_{\ell=0}^{k-1} \lambda_\ell \boldsymbol{g}(\boldsymbol{x}^{(\ell)}) \leq \boldsymbol{0}, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned} \tag{A.18}$$

Therefore, the dual problem of the master program is equivalent to the problem given in (A.17), whose optimal value is thus equal to $z^{(k)}$.

Theorem A.3. *Let $\boldsymbol{x}^{(0)} \in \mathcal{S}$ such that $\boldsymbol{g}(\boldsymbol{x}^{(0)}) < \boldsymbol{0}$, and let $\{\boldsymbol{u}^{(k)}\}$ and $\{\boldsymbol{x}^{(k)}\}$ be the sequences of dual and primal variables, respectively, that are generated by the cutting-plane algorithm with $\boldsymbol{x}^{(k)} \in \{\bar{\boldsymbol{x}} \in \mathcal{S} : \bar{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathcal{S}} f(\boldsymbol{x}) + \boldsymbol{u}^{(k),\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})\}$. Moreover, let $(\hat{\lambda}_0,\ldots,\hat{\lambda}_{k-1})$ be an optimal solution to the problem given in* (A.17)*, and define*

$$\hat{\boldsymbol{x}}^{(k)} = \sum_{\ell=0}^{k-1} \hat{\lambda}_\ell \boldsymbol{x}^{(\ell)}. \tag{A.19}$$

Then, $\hat{\boldsymbol{x}}^{(k)} \in \mathcal{X}$, i.e., $\hat{\boldsymbol{x}}^{(k)}$ is a feasible point for the primal problem given in (A.3)*. Furthermore, if $z^{(k)} - \Theta(\boldsymbol{u}) \leq \delta$ for some $\boldsymbol{u} \in \mathcal{U}$, it follows that $f(\hat{\boldsymbol{x}}^{(k)}) \leq p^\star + \delta$.*

Proof. See [6, Theorem 6.5.2]. □

At each iteration of the cutting-plane algorithm, a primal feasible point can thus be obtained by solving the linear program given in (A.17). Beyond that, we do not even need to solve this problem in practice if we use a primal-dual linear program solver for the master program which solves both the primal and the dual linear program at the same time. In this case, we simply get $(\hat{\lambda}_0, \dots, \hat{\lambda}_{k-1})$ as the optimal dual variables that are associated with the inequality constraints $z \leq f(\boldsymbol{x}^{(\ell)}) + \boldsymbol{u}^\top \boldsymbol{g}(\boldsymbol{x}^{(\ell)})$, $\ell \in \{0, \dots, k-1\}$, of the master program, meaning that the computation of the primal feasible point $\hat{\boldsymbol{x}}^{(k)}$ is basically for free.

What is more, recall that the cutting-plane algorithm refines the approximation of the dual function as long as $z^{(k)} > \Theta(\boldsymbol{u}^{(k)})$ and that it stops if $z^{(k)} = \Theta(\boldsymbol{u}^{(k)})$. Because $z^{(k)} - \Theta(\boldsymbol{u}) \leq 0$ for $\boldsymbol{u} = \boldsymbol{u}^{(k)}$ if the termination criterion is satisfied, it follows from Theorem A.3 that $f(\hat{\boldsymbol{x}}^{(k)}) \leq p^\star$ in this case. However, note that since $\hat{\boldsymbol{x}}^{(k)} \in \mathcal{X}$, we also have $f(\hat{\boldsymbol{x}}^{(k)}) \geq p^\star = \min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$. As a result, we can conclude that $f(\hat{\boldsymbol{x}}^{(k)}) = p^\star$ if $z^{(k)} = \Theta(\boldsymbol{u}^{(k)})$, i.e., $\hat{\boldsymbol{x}}^{(k)}$ is an optimal solution to the primal problem. In addition, note that $y^{(k)} = \max_{\ell=1,\dots,k} \Theta(\boldsymbol{u}^{(\ell)})$ implies $y^{(k)} = \Theta(\boldsymbol{u}^{(\ell)})$ for some $\ell \in \{1, \dots, k\}$. If for some $\delta > 0$ the practical termination criterion $z^{(k)} - y^{(k)} < \delta$ is satisfied, it hence follows that there exists some $\boldsymbol{u} \in \mathcal{U}$ for which $z^{(k)} - \Theta(\boldsymbol{u}) < \delta$ so that $f(\hat{\boldsymbol{x}}^{(k)}) < p^\star + \delta$ by Theorem A.3. As a consequence, the practical termination criterion $z^{(k)} - y^{(k)} < \delta$ guarantees that $f(\hat{\boldsymbol{x}}^{(k)}) - p^\star < \delta$, which means that the cutting-plane algorithm obtains a near-optimal primal feasible solution $\hat{\boldsymbol{x}}^{(k)}$ provided that δ is small.

Appendix B

Information Theoretical Background

B.1 Standard Power Constraints for the Gaussian MIMO Relay Channel

B.1.1 Full-Duplex

A $(2^{nR}, n)$ code for the (discrete-time) Gaussian MIMO relay channel consists of a message set $\mathcal{W} = \{1, 2, \ldots, \lceil 2^{nR} \rceil\}$, a source encoder that assigns a codeword $\boldsymbol{x}_\mathrm{S}^n(w) \in \mathcal{X}_\mathrm{S}^n = \prod_{i=1}^n \mathbb{C}^{N_\mathrm{S}}$ to each $w \in \mathcal{W}$, a relay encoder that assigns an $\boldsymbol{x}_{\mathrm{R},i}(\boldsymbol{y}_\mathrm{R}^{i-1}) \in \mathcal{X}_\mathrm{R} = \mathbb{C}^{N_\mathrm{R}}$ to each past received sequence $\boldsymbol{y}_\mathrm{R}^{i-1} \in \mathcal{Y}_\mathrm{R}^{i-1} = \prod_{j=1}^{i-1} \mathbb{C}^{N_\mathrm{R}}$ for each $i \in \{1, \ldots, n\}$, and a decoder (at the destination) that assigns an estimate $\hat{W}(\boldsymbol{y}_\mathrm{D}^n) \in \mathcal{W}$ (or possibly an error message) to each received sequence $\boldsymbol{y}_\mathrm{D}^n \in \mathcal{Y}_\mathrm{D}^n = \prod_{i=1}^n \mathbb{C}^{N_\mathrm{D}}$.

The Gaussian MIMO relay channel we consider is *memoryless* in the sense that the current channel outputs $(\boldsymbol{y}_{\mathrm{D},i}, \boldsymbol{y}_{\mathrm{R},i})$ depend on all previous channel inputs $(\boldsymbol{x}_\mathrm{S}^i, \boldsymbol{x}_\mathrm{R}^i)$ only through the current channel inputs $(\boldsymbol{x}_{\mathrm{S},i}, \boldsymbol{x}_{\mathrm{R},i})$. For any $p(w)$ and choice of the code, $p(w, \boldsymbol{x}_\mathrm{S}^n, \boldsymbol{x}_\mathrm{R}^n, \boldsymbol{y}_\mathrm{D}^n, \boldsymbol{y}_\mathrm{R}^n)$ hence factors as

$$p(w, \boldsymbol{x}_\mathrm{S}^n, \boldsymbol{x}_\mathrm{R}^n, \boldsymbol{y}_\mathrm{D}^n, \boldsymbol{y}_\mathrm{R}^n) = p(w)p(\boldsymbol{x}_\mathrm{S}^n|w) \prod_{i=1}^n p(\boldsymbol{x}_{\mathrm{R},i}|\boldsymbol{y}_\mathrm{R}^{i-1}) p(\boldsymbol{y}_{\mathrm{D},i}, \boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i}, \boldsymbol{x}_{\mathrm{R},i}). \tag{B.1}$$

Moreover, the transmissions are modeled as taking place *synchronously*, and since the additive Gaussian noise at the relay and the destination is assumed to be independent, $p(\boldsymbol{y}_{\mathrm{D},i}, \boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i}, \boldsymbol{x}_{\mathrm{R},i})$ factors as

$$p(\boldsymbol{y}_{\mathrm{D},i}, \boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i}, \boldsymbol{x}_{\mathrm{R},i}) = p(\boldsymbol{y}_{\mathrm{D},i}|\boldsymbol{x}_{\mathrm{S},i}, \boldsymbol{x}_{\mathrm{R},i}) p(\boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i}, \boldsymbol{x}_{\mathrm{R},i}). \tag{B.2}$$

Note that except for the last condition, the definition of a code for the Gaussian MIMO relay channel is so far identical to the definition of a code for the discrete memoryless relay channel, cf. Section 2.1.

As mentioned in Section 3.1, however, we need to constrain the channel inputs for continuous-alphabet channels because otherwise the capacity is unbounded. The most common input constraint is an *average* power constraint over the n channel uses of the code. If we impose this standard power constraint on the channel inputs of the source and the relay, we obtain

$$\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{x}_{\mathrm{S},i}(w)\|_2^2 \le P_{\mathrm{S}}, \quad \forall w \in \mathcal{W} \tag{B.3}$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{x}_{\mathrm{R},i}(\boldsymbol{y}_{\mathrm{R},1},\ldots,\boldsymbol{y}_{\mathrm{R},i-1})\|_2^2 \le P_{\mathrm{R}}, \quad \forall(\boldsymbol{y}_{\mathrm{R},1},\ldots,\boldsymbol{y}_{\mathrm{R},n}) \in \mathcal{Y}_{\mathrm{R}}^n. \tag{B.4}$$

Now, let Q be a mixing or *time-sharing* random variable that is uniformly distributed on $\{1,\ldots,n\}$ and that is independent of $(\boldsymbol{x}_{\mathrm{S},1},\ldots,\boldsymbol{x}_{\mathrm{S},n})$ and $(\boldsymbol{x}_{\mathrm{R},1},\ldots,\boldsymbol{x}_{\mathrm{R},n})$. Furthermore, let us define $\boldsymbol{x}_{\mathrm{A},i} = \boldsymbol{x}_{\mathrm{A}}|\{Q=i\}$ and $\boldsymbol{y}_{\mathrm{B},i} = \boldsymbol{y}_{\mathrm{B}}|\{Q=i\}$. Then,

$$\begin{aligned}\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}] &= \mathrm{E}\Big[\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}|Q]\Big]\\ &= \sum_{i=1}^{n}\Pr[Q=i]\,\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}|Q=i] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[\boldsymbol{x}_{\mathrm{S},i}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S},i}] \le P_{\mathrm{S}}\end{aligned} \tag{B.5}$$

and

$$\begin{aligned}\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}] &= \mathrm{E}\Big[\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}|Q]\Big]\\ &= \sum_{i=1}^{n}\Pr[Q=i]\,\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}|Q=i] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[\boldsymbol{x}_{\mathrm{R},i}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R},i}] \le P_{\mathrm{R}}.\end{aligned} \tag{B.6}$$

In addition, following the steps of Cover and El Gamal [21] while taking into account that we are dealing with a continuous-alphabet channel, cf. [145, Chapter 11], it can be shown that

$$\begin{aligned}R &\le \min\left\{\frac{1}{n}\sum_{i=1}^{n}I(\boldsymbol{x}_{\mathrm{S},i};\boldsymbol{y}_{\mathrm{R},i},\boldsymbol{y}_{\mathrm{D},i}|\boldsymbol{x}_{\mathrm{R},i}),\ \frac{1}{n}\sum_{i=1}^{n}I(\boldsymbol{x}_{\mathrm{S},i},\boldsymbol{x}_{\mathrm{R},i};\boldsymbol{y}_{\mathrm{D},i})\right\} + \varepsilon_n\\ &= \min\Big\{I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}},\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{x}_{\mathrm{R}},Q),\ I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}}|Q)\Big\} + \varepsilon_n\\ &\le \min\Big\{I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}},\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{x}_{\mathrm{R}}),\ I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}})\Big\} + \varepsilon_n\end{aligned} \tag{B.7}$$

if R is an achievable rate for the Gaussian MIMO relay channel, where the last inequality is due to the concavity of the mutual information terms with respect to the joint input distribution $p(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})$, cf. [78, Section 3.7.1].

Consequently, the upper bound on the capacity of the Gaussian MIMO relay channel we obtain if the source and relay inputs are subject to the standard power constraint is equivalent to C_{CSB} given in (3.8). The reason for this is that time-sharing over the n channel uses of the code is not necessary to achieve the cut-set bound (CSB) since both $I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}},\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{x}_{\mathrm{R}})$ and $I(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}};\boldsymbol{y}_{\mathrm{D}})$ are concave in $p(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})$. Likewise, the maximum decode-and-forward (DF) rate that can be achieved with the standard power constraint is equivalent to R_{DF} given in (3.21) because $I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}})$ is also concave in $p(\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})$. On the other hand, R_{PDF} given in (3.27) is in general smaller than the maximum partial decode-and-forward (PDF) rate that is achievable with the standard power constraint as $I(\boldsymbol{u};\boldsymbol{y}_{\mathrm{R}}|\boldsymbol{x}_{\mathrm{R}}) + I(\boldsymbol{x}_{\mathrm{S}};\boldsymbol{y}_{\mathrm{D}}|\boldsymbol{u},\boldsymbol{x}_{\mathrm{R}})$ is not concave in $p(\boldsymbol{u},\boldsymbol{x}_{\mathrm{S}},\boldsymbol{x}_{\mathrm{R}})$. However, whether or not time-sharing in the code can actually improve the achievable PDF rate for the Gaussian MIMO relay channel is not considered in this work, so for simplicity, we simply add the power constraints given in (3.3) to the single-letter mutual information expressions that were obtained for the discrete memoryless relay channel.

B.1.2 Half-Duplex

Let $m^n \in \mathcal{M}^n = \{1,2\}^n$ be a sequence of relay modes that is noncausally known to all three nodes, where $m=1$ means that the relay is in receive mode and $m=2$ means that the relay is in transmit mode. Then, a $(2^{nR}, n)$ code for the half-duplex Gaussian MIMO relay channel consists of a message set $\mathcal{W} = \{1,2,\dots,\lceil 2^{nR}\rceil\}$, an encoder (at the source) that assigns a codeword $\boldsymbol{x}_{\mathrm{S}}^n(w) \in \mathcal{X}_{\mathrm{S}}^n = \prod_{i=1}^n \mathbb{C}^{N_{\mathrm{S}}}$ to each message $w \in \mathcal{W}$, a relay encoder that assigns an $\boldsymbol{x}_{\mathrm{R},i}(\boldsymbol{y}_{\mathrm{R}}^{i-1}) \in \mathcal{X}_{\mathrm{R}} = \mathbb{C}^{N_{\mathrm{R}}}$ to each past received sequence $\boldsymbol{y}_{\mathrm{R}}^{i-1} \in \mathcal{Y}_{\mathrm{R}}^{i-1} = \prod_{j=1}^{i-1} \mathbb{C}^{N_{\mathrm{R}}}$ for each $i \in \{1,\dots,n\}$ with $\boldsymbol{x}_{\mathrm{R},i}(\boldsymbol{y}_{\mathrm{R}}^{i-1}) = \boldsymbol{0}$ if $m_i = 1$, and a decoder (at the destination) that assigns an estimate $\hat{W}(\boldsymbol{y}_{\mathrm{D}}^n) \in \mathcal{W}$ (or possibly an error message) to each received sequence $\boldsymbol{y}_{\mathrm{D}}^n \in \mathcal{Y}_{\mathrm{D}}^n = \prod_{i=1}^n \mathbb{C}^{N_{\mathrm{D}}}$. Note that the relay encoder must select $\boldsymbol{x}_{\mathrm{R},i} = \boldsymbol{0}$ whenever the relay is in receive mode in order to satisfy the half-duplex constraint.

The half-duplex Gaussian MIMO relay channel we consider is *memoryless* in the sense that the current channel outputs $(\boldsymbol{y}_{\mathrm{D},i},\boldsymbol{y}_{\mathrm{R},i})$ depend on $(\boldsymbol{x}_{\mathrm{S}}^i,\boldsymbol{x}_{\mathrm{R}}^i,M^i)$ only through $(\boldsymbol{x}_{\mathrm{S},i},\boldsymbol{x}_{\mathrm{R},i},M_i)$ so that for any $p(w)$ and choice of the code, $p(w,\boldsymbol{x}_{\mathrm{S}}^n,\boldsymbol{x}_{\mathrm{R}}^n,m^n,\boldsymbol{y}_{\mathrm{D}}^n,\boldsymbol{y}_{\mathrm{R}}^n)$ factors as

$$p(w,\boldsymbol{x}_{\mathrm{S}}^n,\boldsymbol{x}_{\mathrm{R}}^n,m^n,\boldsymbol{y}_{\mathrm{D}}^n,\boldsymbol{y}_{\mathrm{R}}^n) = p(w)p(\boldsymbol{x}_{\mathrm{S}}^n|w)\prod_{i=1}^n p(\boldsymbol{x}_{\mathrm{R},i}|\boldsymbol{y}_{\mathrm{R}}^{i-1})p(\boldsymbol{y}_{\mathrm{D},i},\boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i},\boldsymbol{x}_{\mathrm{R},i},m_i). \tag{B.8}$$

Moreover, if we define $\boldsymbol{x}_{\mathrm{A}m,i} = \boldsymbol{x}_{\mathrm{A},i}|\{M=m\}$ and $\boldsymbol{y}_{\mathrm{B}m,i} = \boldsymbol{y}_{\mathrm{B},i}|\{M=m\}$, the conditional probability density function $p(\boldsymbol{y}_{\mathrm{D},i},\boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i},\boldsymbol{x}_{\mathrm{R},i},m_i)$ satisfies

$$p(\boldsymbol{y}_{\mathrm{D},i},\boldsymbol{y}_{\mathrm{R},i}|\boldsymbol{x}_{\mathrm{S},i},\boldsymbol{x}_{\mathrm{R},i},m_i) = \begin{cases} p(\boldsymbol{y}_{\mathrm{D}1,i},\boldsymbol{y}_{\mathrm{R}1,i}|\boldsymbol{x}_{\mathrm{S}1,i}) & \text{if } m_i = 1, \boldsymbol{x}_{\mathrm{R}1,i} = \boldsymbol{0},\\ p(\boldsymbol{y}_{\mathrm{D}2,i}|\boldsymbol{x}_{\mathrm{S}2,i},\boldsymbol{x}_{\mathrm{R}2,i}) & \text{if } m_i = 2, \boldsymbol{y}_{\mathrm{R}2,i} = \boldsymbol{0},\\ 0 & \text{otherwise} \end{cases} \tag{B.9}$$

for any $i \in \{1,\dots,n\}$, and $p(\boldsymbol{x}_{\mathrm{R},i}|\boldsymbol{y}_{\mathrm{R}}^{i-1}) = 0$ for any $i \in \{1,\dots,n\}$ such that $m_i = 1$ and $\boldsymbol{x}_{\mathrm{R},i} \neq \boldsymbol{0}$. Finally, like for the full-duplex case, the transmissions are modeled as taking place *synchronously*.

Now, let us impose the standard power constraints given in (B.3) and (B.4) on the source and relay inputs, respectively. Furthermore, suppose $m_i = 1$ for n_1 and $m_i = 2$ for $n_2 = n - n_1$ channel uses, let $t_1 = \frac{n_1}{n}$ and $t_2 = \frac{n-n_1}{n} = \frac{n_2}{n}$, and let Q again denote a *time-sharing* random variable that is uniformly distributed on $\{1, \ldots, n\}$ and that is independent of $(\boldsymbol{x}_{\mathrm{S},1}, \ldots, \boldsymbol{x}_{\mathrm{S},n})$ and $(\boldsymbol{x}_{\mathrm{R},1}, \ldots, \boldsymbol{x}_{\mathrm{R},n})$. With $\boldsymbol{x}_{\mathrm{A}m} = \boldsymbol{x}_{\mathrm{A}}|\{M = m\}$ and $\boldsymbol{x}_{\mathrm{A},i} = \boldsymbol{x}_{\mathrm{A}}|\{Q = i\}$, it then follows that

$$\begin{aligned} t_1\,\mathrm{E}[\boldsymbol{x}_{\mathrm{S}1}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}1}] + t_2\,\mathrm{E}[\boldsymbol{x}_{\mathrm{S}2}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}2}] &= \sum_{m=1}^{2}\Pr[M=m]\,\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}|M=m] \\ = \mathrm{E}\Big[\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}|M]\Big] &= \mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}] = \mathrm{E}\Big[\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}|Q]\Big] \\ = \sum_{i=1}^{n}\Pr[Q=i]\,\mathrm{E}[\boldsymbol{x}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S}}|Q=i] &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[\boldsymbol{x}_{\mathrm{S},i}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{S},i}] \le P_{\mathrm{S}} \end{aligned} \tag{B.10}$$

and

$$\begin{aligned} t_2\,\mathrm{E}[\boldsymbol{x}_{\mathrm{R}2}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}2}] &= \sum_{m=1}^{2}\Pr[M=m]\,\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}|M=m] \\ = \mathrm{E}\Big[\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}|M]\Big] &= \mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}] = \mathrm{E}\Big[\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}|Q]\Big] \\ = \sum_{i=1}^{n}\Pr[Q=i]\,\mathrm{E}[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}}|Q=i] &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[\boldsymbol{x}_{\mathrm{R},i}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R},i}] \le P_{\mathrm{R}} \end{aligned} \tag{B.11}$$

as $\mathrm{E}[\boldsymbol{x}_{\mathrm{R}1}^{\mathrm{H}}\boldsymbol{x}_{\mathrm{R}1}] = 0$. Beyond that, all mutual information terms that characterize the CSB and the maximum achievable DF rate for the half-duplex Gaussian MIMO relay channel are concave in the channel inputs. The corresponding values we obtain if the source and the relay are subject to the standard power constraint are therefore equivalent to $C_{\mathrm{CSB,av}}$ and $R_{\mathrm{DF,av}}$ given in (6.34) and (6.65), respectively. Like for the full-duplex case, however, the maximum PDF rate that can be achieved with the standard power constraint is generally larger than $R_{\mathrm{PDF,av}}$ considered in Section 6.4.

To conclude this discussion about power constraints for the half-duplex Gaussian MIMO relay channel, we remark that the per-phase power constraints defined in (6.4) are similarly related to the source and relay input constraints

$$\frac{1}{n_1}\sum_{i:m_i=1}\|\boldsymbol{x}_{\mathrm{S},i}(w)\|_2^2 \le P_{\mathrm{S}}, \quad \frac{1}{n_2}\sum_{i:m_i=2}\|\boldsymbol{x}_{\mathrm{S},i}(w)\|_2^2 \le P_{\mathrm{S}}, \quad \forall w \in \mathcal{W} \tag{B.12}$$

and

$$\frac{1}{n_2}\sum_{i:m_i=2}\|\boldsymbol{x}_{\mathrm{R},i}(\boldsymbol{y}_{\mathrm{R},1}, \ldots, \boldsymbol{y}_{\mathrm{R},i-1})\|_2^2 \le P_{\mathrm{R}}, \quad \forall(\boldsymbol{y}_{\mathrm{R},1}, \ldots, \boldsymbol{y}_{\mathrm{R},n}) \in \mathcal{Y}_{\mathrm{R}}^n. \tag{B.13}$$

Furthermore, the results straightforwardly carry over to bidirectional communication in the half-duplex constrained Gaussian MIMO relay channel.

B.2 Degraded Broadcast Channels

The broadcast channel models a multi-user communication system where one source wants to convey (different) information to several destinations. While the capacity region of the general broadcast channel has yet to be determined, there are several special cases for which it is known. In particular, it has been shown that *superposition coding* achieves the (private message) capacity region of the *(stochastically) degraded* broadcast channel. Furthermore, superposition coding is also optimal for the *less noisy* and the *more capable* 2-user broadcast channels.

In this section, we discuss how the conditions that characterize these classes of 2-user broadcast channels (where the source-to-destination channels can be ordered according to different criteria) are related to each other and how these relations differ for the discrete memoryless and the Gaussian MIMO broadcast channels.

B.2.1 Discrete Memoryless Broadcast Channel

The 2-user discrete memoryless broadcast channel, which is specified by

$$\left\{\mathcal{X}, p(y_\mathrm{A}, y_\mathrm{B}|x), \mathcal{Y}_\mathrm{A} \times \mathcal{Y}_\mathrm{B}\right\}, \tag{B.14}$$

consists of three finite sets $\mathcal{X}$, $\mathcal{Y}_\mathrm{A}$, $\mathcal{Y}_\mathrm{B}$ and a collection of probability mass functions $p(y_\mathrm{A}, y_\mathrm{B}|x)$, one for each $x \in \mathcal{X}$. The interpretation is that x is the channel input of the source, whereas y_A and y_B denote the channel outputs of the destinations A and B, respectively, as illustrated in Figure B.1.

Assuming that no common message is conveyed from the source to both destinations, a $(2^{nR_\mathrm{A}}, 2^{nR_\mathrm{B}}, n)$ code for the 2-user broadcast channel consists of two message sets $\mathcal{W}_\mathrm{A} = \{1, 2, \ldots, \lceil 2^{nR_\mathrm{A}} \rceil\}$ and $\mathcal{W}_\mathrm{B} = \{1, 2, \ldots, \lceil 2^{nR_\mathrm{B}} \rceil\}$, a source encoder that assigns a codeword $X^n(w_\mathrm{A}, w_\mathrm{B})$ to each message tuple $(w_\mathrm{A}, w_\mathrm{B}) \in \mathcal{W}_\mathrm{A} \times \mathcal{W}_\mathrm{B}$, and two decoders (one at each destination) that assign estimates $\hat{W}_\mathrm{A}(y_\mathrm{A}^n)$ and $\hat{W}_\mathrm{B}(y_\mathrm{B}^n)$ (or possibly an error message) to each received sequence $y_\mathrm{A}^n \in \mathcal{Y}_\mathrm{A}^n$ and $y_\mathrm{B}^n \in \mathcal{Y}_\mathrm{B}^n$, respectively.

The broadcast channel is *memoryless* in the sense that the current channel outputs $(Y_{\mathrm{A},i}, Y_{\mathrm{B},i})$ depend on all previous channel inputs X^i only through the current channel input X_i. For any $p(w_\mathrm{A}, w_\mathrm{B})$ and choice of the code, the joint probability mass function

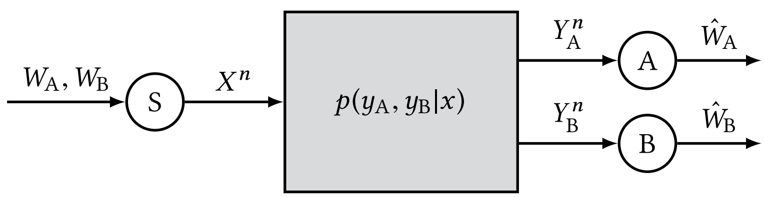


Figure B.1: Illustration of the 2-User Broadcast Channel

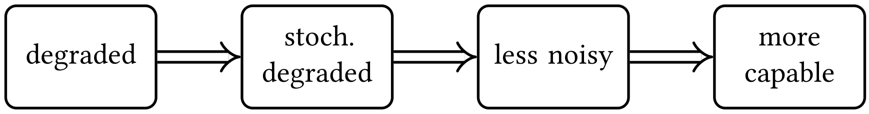


Figure B.2: Degraded Discrete Memoryless Broadcast Channels

on $\mathcal{W}_\mathrm{A} \times \mathcal{W}_\mathrm{B} \times \mathcal{X}^n \times \mathcal{Y}_\mathrm{A}^n \times \mathcal{Y}_\mathrm{B}^n$ hence factors as

$$p(w_\mathrm{A}, w_\mathrm{B}, x^n, y_\mathrm{A}^n, y_\mathrm{B}^n) = p(w_\mathrm{A}, w_\mathrm{B}) p(x^n | w_\mathrm{A}, w_\mathrm{B}) \prod_{i=1}^{n} p(y_{\mathrm{A},i}, y_{\mathrm{B},i} | x_i). \tag{B.15}$$

If W_A and W_B are uniformly distributed over $\mathcal{W}_\mathrm{A}$ and $\mathcal{W}_\mathrm{B}$, respectively, and if $P_e^{(n)} = \Pr[\hat{W}_\mathrm{A} \neq W_\mathrm{A} \vee \hat{W}_\mathrm{B} \neq W_\mathrm{B}]$ denotes the average probability of error, a rate tuple $(R_\mathrm{A}, R_\mathrm{B})$ is said to be *achievable* if there exists a sequence of $(2^{nR_\mathrm{A}}, 2^{nR_\mathrm{B}}, n)$ codes for which $P_e^{(n)} \to 0$ as $n \to \infty$. The (private message) *capacity region* $\mathcal{C}_\mathrm{BC}$ of the 2-user broadcast channel is defined as the closure of the set of achievable rate tuples $(R_\mathrm{A}, R_\mathrm{B})$. Finally, we remark that an important property of the broadcast channel is that $\mathcal{C}_\mathrm{BC}$ depends on the conditional probability mass function $p(y_\mathrm{A}, y_\mathrm{B}|x)$ only through the conditional marginals $p(y_\mathrm{A}|x)$ and $p(y_\mathrm{B}|x)$ [34, Lemma 5.1].

The capacity region $\mathcal{C}_\mathrm{BC}$ of the discrete memoryless broadcast channel is not known in general. However, for the following classes of broadcast channels where the channel from the source to destination A is better than the channel to destination B, the (private message) capacity region is known and can be achieved using superposition coding, cf. [34, Chapter 5]:

Definition B.1. The 2-user broadcast channel $\left\{\mathcal{X}, p(y_\mathrm{A}, y_\mathrm{B}|x), \mathcal{Y}_\mathrm{A} \times \mathcal{Y}_\mathrm{B}\right\}$ is said to be

(a) *degraded* if $X \leftrightarrow Y_\mathrm{A} \leftrightarrow Y_\mathrm{B}$ form a Markov chain;

(b) *stochastically degraded* if there exists a $\tilde{Y}_\mathrm{A}$ with the same conditional distribution as Y_A (given X) such that $X \leftrightarrow \tilde{Y}_\mathrm{A} \leftrightarrow Y_\mathrm{B}$ form a Markov chain;

(c) *less noisy* if $I(U; Y_\mathrm{A}) \geq I(U; Y_\mathrm{B})$ for all $p(u, x)$ such that $U \leftrightarrow X \leftrightarrow (Y_\mathrm{A}, Y_\mathrm{B})$ form a Markov chain;

(d) *more capable* if $I(X; Y_\mathrm{A}) \geq I(X; Y_\mathrm{B})$ for all $p(x)$.

Obviously, a degraded broadcast channel is also stochastically degraded. Moreover, it can be shown that if a discrete memoryless broadcast channel is stochastically degraded, then it is less noisy, and if it is less noisy, it is more capable. However, the converse to each of these statements is not true, cf. [34, Section 5.6]. The implications between these four classes of discrete memoryless broadcast channels are illustrated in Figure B.2.

Note that the capacity region of a stochastically degraded broadcast channel is the same as that of its corresponding degraded broadcast channel because $\mathcal{C}_\mathrm{BC}$ only depends on $p(y_\mathrm{A}|x)$ and $p(y_\mathrm{B}|x)$. Furthermore, the less noisy condition guarantees that if the source uses superposition coding, destination A can always recover the message for destination B. To conclude the discussion of the discrete memoryless broadcast channel, we remark that the capacity region is not known for the less noisy K-user discrete memoryless broadcast channel with $K > 3$ destinations or for the more capable K-user discrete memoryless broadcast channel with $K > 2$ destinations [34, Remark 5.15].

B.2.2 Gaussian MIMO Broadcast Channel

The Gaussian MIMO broadcast channel models a wireless communication system in which the source and each destination can have multiple antennas and where the destinations are subject to additive Gaussian noise. Using the discrete-time narrowband MIMO channel model, cf. [9, Chapter 1] and [50, Chapter 10], for example, the 2-user Gaussian MIMO broadcast channel is specified by

$$\begin{aligned} \boldsymbol{y}_{\mathrm{A}} &= \boldsymbol{H}_{\mathrm{A}}\boldsymbol{x} + \boldsymbol{n}_{\mathrm{A}}, \qquad & \boldsymbol{n}_{\mathrm{A}} &\sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{A}}), \\ \boldsymbol{y}_{\mathrm{B}} &= \boldsymbol{H}_{\mathrm{B}}\boldsymbol{x} + \boldsymbol{n}_{\mathrm{B}}, \qquad & \boldsymbol{n}_{\mathrm{B}} &\sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{B}}). \end{aligned} \tag{B.16}$$

Here, $\boldsymbol{x} \in \mathbb{C}^{M}$ is the transmit signal of the source and $\boldsymbol{y}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{A}}}$ and $\boldsymbol{y}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{B}}}$ denote the signal vectors that are received by destinations A and B, respectively. Moreover, $\boldsymbol{H}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{A}} \times M}$ and $\boldsymbol{H}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{B}} \times M}$ represent the channel gain matrices (of appropriate dimensions) while $\boldsymbol{n}_{\mathrm{A}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{A}})$ and $\boldsymbol{n}_{\mathrm{B}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{B}})$ denote independent zero-mean proper (circularly symmetric) complex Gaussian noise vectors, not necessarily white, with nonsingular covariance matrices $\boldsymbol{Z}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{A}} \times N_{\mathrm{A}}}$ and $\boldsymbol{Z}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{B}}}$. The noise vectors are also assumed to be independent of the transmit signal $\boldsymbol{x}$.

The degraded, stochastically degraded, less noisy, and more capable Gaussian MIMO broadcast channels can be defined in the same way as the corresponding discrete memoryless broadcast channels, cf. Definition B.1. However, the Gaussian MIMO broadcast channel is never degraded according to the system model since $\boldsymbol{n}_{\mathrm{A}}$ and $\boldsymbol{n}_{\mathrm{B}}$ are assumed to be independent. Beyond that, the stochastically degraded Gaussian MIMO broadcast channel can be defined in a more concise manner by means of the channel gain matrices $\boldsymbol{H}_{\mathrm{A}}, \boldsymbol{H}_{\mathrm{B}}$ and the noise covariances $\boldsymbol{Z}_{\mathrm{A}}, \boldsymbol{Z}_{\mathrm{B}}$.

Definition B.2. The 2-user Gaussian MIMO broadcast channel is said to be *stochastically degraded* if $\boldsymbol{H}_{\mathrm{A}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{A}}^{-1}\boldsymbol{H}_{\mathrm{A}} \succcurlyeq \boldsymbol{H}_{\mathrm{B}}^{\mathrm{H}}\boldsymbol{Z}_{\mathrm{B}}^{-1}\boldsymbol{H}_{\mathrm{B}}$.

Note that according to this definition, the Gaussian MIMO broadcast channel is stochastically degraded if and only if there exists an $\boldsymbol{M} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{A}}}$ such that $\boldsymbol{H}_{\mathrm{B}} = \boldsymbol{M}\boldsymbol{H}_{\mathrm{A}}$ and $\boldsymbol{M}\boldsymbol{Z}_{\mathrm{A}}\boldsymbol{M}^{\mathrm{H}} \preccurlyeq \boldsymbol{Z}_{\mathrm{B}}$, cf. [114, Lemma 5]. But if the latter condition holds, there exists a random vector $\tilde{\boldsymbol{y}}_{\mathrm{A}}$ with $\tilde{\boldsymbol{y}}_{\mathrm{A}}|\boldsymbol{x} \sim \boldsymbol{y}_{\mathrm{A}}|\boldsymbol{x}$ such that $\boldsymbol{x} \leftrightarrow \tilde{\boldsymbol{y}}_{\mathrm{A}} \leftrightarrow \boldsymbol{y}_{\mathrm{B}}$ form a Markov chain. In particular, let $\tilde{\boldsymbol{n}}_{\mathrm{A}}$ be a random vector that follows the same distribution as $\boldsymbol{n}_{\mathrm{A}}$, i.e., $\tilde{\boldsymbol{n}}_{\mathrm{A}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{A}})$, and let $\tilde{\boldsymbol{n}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{B}} - \boldsymbol{M}\boldsymbol{Z}_{\mathrm{A}}\boldsymbol{M}^{\mathrm{H}})$ be independent of $\tilde{\boldsymbol{n}}_{\mathrm{A}}$ such that $\boldsymbol{n}_{\mathrm{B}} = \boldsymbol{M}\tilde{\boldsymbol{n}}_{\mathrm{A}} + \tilde{\boldsymbol{n}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{Z}_{\mathrm{B}})$. If we then define $\tilde{\boldsymbol{y}}_{\mathrm{A}} = \boldsymbol{H}_{\mathrm{A}}\boldsymbol{x} + \tilde{\boldsymbol{n}}_{\mathrm{A}}$, it follows that

$$\boldsymbol{y}_{\mathrm{B}} = \boldsymbol{H}_{\mathrm{B}}\boldsymbol{x} + \boldsymbol{n}_{\mathrm{B}} = \boldsymbol{M}\boldsymbol{H}_{\mathrm{A}}\boldsymbol{x} + \boldsymbol{n}_{\mathrm{B}} = \boldsymbol{M}\left(\boldsymbol{H}_{\mathrm{A}}\boldsymbol{x} + \tilde{\boldsymbol{n}}_{\mathrm{A}}\right) + \tilde{\boldsymbol{n}} = \boldsymbol{M}\tilde{\boldsymbol{y}}_{\mathrm{A}} + \tilde{\boldsymbol{n}}. \tag{B.17}$$

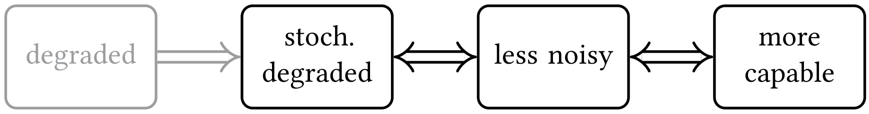


Figure B.3: Degraded Gaussian MIMO Broadcast Channels

While this only shows that the condition of Definition B.2 implies that of Definition B.1, the proof of Theorem B.1 given below reveals that the two conditions are equivalent. Hence, Definition B.2 indeed provides an appropriate characterization of the stochastically degraded 2-user Gaussian MIMO broadcast channel.

THEOREM B.1. *The conditions defining the stochastically degraded, less noisy, and more capable Gaussian MIMO broadcast channels are equivalent.*

Proof. Like for the discrete memoryless case, it can be shown that if a Gaussian MIMO broadcast channel is stochastically degraded, then it is less noisy, and if it is less noisy, then it is also more capable. The theorem can therefore be proved by showing that a more capable Gaussian MIMO broadcast channel is stochastically degraded.

To this end, let $\boldsymbol{x} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C})$. Since $I(\boldsymbol{x};\boldsymbol{y}_\mathrm{A}) \geq I(\boldsymbol{x};\boldsymbol{y}_\mathrm{B})$ for all $p(\boldsymbol{x})$ if destination A is more capable than destination B, we must have

$$\log\det\left(\mathbf{I} + \boldsymbol{H}_\mathrm{A}\boldsymbol{C}\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\right) \geq \log\det\left(\mathbf{I} + \boldsymbol{H}_\mathrm{B}\boldsymbol{C}\boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\right), \forall \boldsymbol{C} \succcurlyeq \mathbf{0}. \tag{B.18}$$

Now, suppose $\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\boldsymbol{H}_\mathrm{A} \not\succcurlyeq \boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\boldsymbol{H}_\mathrm{B}$, i.e., $\exists \boldsymbol{a} \in \mathbb{C}^M : \boldsymbol{a}^\mathrm{H}\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\boldsymbol{H}_\mathrm{A}\boldsymbol{a} < \boldsymbol{a}^\mathrm{H}\boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\boldsymbol{H}_\mathrm{B}\boldsymbol{a}$. But since $\log\det(\mathbf{I} + \boldsymbol{A}\boldsymbol{B}) = \log\det(\mathbf{I} + \boldsymbol{B}\boldsymbol{A})$ and the logarithm is an increasing function, this would imply

$$\begin{aligned} &\log\det\left(\mathbf{I} + \boldsymbol{H}_\mathrm{A}\boldsymbol{a}\boldsymbol{a}^\mathrm{H}\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\right) = \log\left(1 + \boldsymbol{a}^\mathrm{H}\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\boldsymbol{H}_\mathrm{A}\boldsymbol{a}\right) \\ <\ &\log\left(1 + \boldsymbol{a}^\mathrm{H}\boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\boldsymbol{H}_\mathrm{B}\boldsymbol{a}\right) = \log\det\left(\mathbf{I} + \boldsymbol{H}_\mathrm{B}\boldsymbol{a}\boldsymbol{a}^\mathrm{H}\boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\right), \end{aligned} \tag{B.19}$$

which contradicts (B.18) as $\boldsymbol{C} = \boldsymbol{a}\boldsymbol{a}^\mathrm{H} \succcurlyeq \mathbf{0}$ is a valid covariance matrix. If the Gaussian MIMO broadcast channel is more capable, it hence follows that $\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\boldsymbol{H}_\mathrm{A} \succcurlyeq \boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\boldsymbol{H}_\mathrm{B}$, i.e., it is also stochastically degraded. □

As opposed to the discrete memoryless case, for which there is a hierarchy of four distinct classes of broadcast channels with one destination having a better channel than the other, Theorem B.1 reveals that there is basically only one such class of Gaussian MIMO broadcast channels. In particular, for the Gaussian MIMO broadcast channel, it holds that $\boldsymbol{y}_\mathrm{B}$ is a stochastically degraded version of $\boldsymbol{y}_\mathrm{A}$ if and only if destination A is both more capable and less noisy than destination B, cf. Figure B.3. Furthermore, we can test for this case by checking whether or not the simple condition $\boldsymbol{H}_\mathrm{A}^\mathrm{H}\boldsymbol{Z}_\mathrm{A}^{-1}\boldsymbol{H}_\mathrm{A} \succcurlyeq \boldsymbol{H}_\mathrm{B}^\mathrm{H}\boldsymbol{Z}_\mathrm{B}^{-1}\boldsymbol{H}_\mathrm{B}$ of Definition B.2 is satisfied. The equivalence of the stochastically degraded, less noisy, and more capable Gaussian MIMO broadcast channels was first established in [44].

B.3 Coding Schemes for the Half-Duplex Two-Way Relay Channel

B.3.1 Decode-and-Forward

Let $W_{\text{A}} \in \{1, 2, \dots, \lceil 2^{nR_{\text{A}}} \rceil\}$ denote the message to be transmitted from node A to node B, and let $W_{\text{B}} \in \{1, 2, \dots, \lceil 2^{nR_{\text{B}}} \rceil\}$ denote the message to be sent from terminal B to terminal A. Both messages are split into six independent parts, i.e., $W_{\text{A}} = (W_{\text{A}1}, \dots, W_{\text{A}6})$ and $W_{\text{B}} = (W_{\text{B}1}, \dots, W_{\text{B}6})$ such that $W_{\text{A}k} \in \{1, 2, \dots, \lceil 2^{nR_{\text{A}k}} \rceil\}$, $W_{\text{B}k} \in \{1, 2, \dots, \lceil 2^{nR_{\text{B}k}} \rceil\}$, $k \in \{1, \dots, 6\}$. The messages are then conveyed to the respective other terminal as follows:

Phase 1: Node A transmits $X_{\text{A}1}(W_{\text{A}1}, W_{\text{A}2}, W_{\text{A}3})$.

Phase 2: Node B transmits $X_{\text{B}2}(W_{\text{B}1}, W_{\text{B}2}, W_{\text{B}3})$.

Phase 3: Node A transmits $X_{\text{A}3}(W_{\text{A}4}, W_{\text{A}5})$ and node B sends $X_{\text{B}3}(W_{\text{B}4}, W_{\text{B}5})$. The two codewords are independent.

After phase 3, the relay reliably decodes the message parts $(W_{\text{A}1}, \dots, W_{\text{A}5})$ and $(W_{\text{B}1}, \dots, W_{\text{B}5})$, which requires

$$\begin{aligned}
R_{\text{A}1} + R_{\text{A}2} + R_{\text{A}3} &< t_1 I(X_{\text{A}1}; Y_{\text{R}1}), \\
R_{\text{B}1} + R_{\text{B}2} + R_{\text{B}3} &< t_2 I(X_{\text{B}2}; Y_{\text{R}2}), \\
R_{\text{A}4} + R_{\text{A}5} &< t_3 I(X_{\text{A}3}; Y_{\text{R}3} | X_{\text{B}3}), \\
R_{\text{B}4} + R_{\text{B}5} &< t_3 I(X_{\text{B}3}; Y_{\text{R}3} | X_{\text{A}3}), \\
R_{\text{A}4} + R_{\text{A}5} + R_{\text{B}4} + R_{\text{B}5} &< t_3 I(X_{\text{A}3}, X_{\text{B}3}; Y_{\text{R}3}).
\end{aligned} \tag{B.20}$$

Phase 4: The relay transmits $X_{\text{R}4}(W_{\text{A}1}, W_{\text{A}4}, W_{\text{B}1}, W_{\text{B}4})$.

Phase 5: The relay sends $X_{\text{R}5}(W_{\text{B}2}, W_{\text{B}5})$ and node B transmits $X_{\text{B}5}(W_{\text{B}2}, W_{\text{B}5}, W_{\text{B}6})$. Note that the two codewords are not independent, but correlated by design in general.

Phase 6: The relay sends $X_{\text{R}6}(W_{\text{A}2}, W_{\text{A}5})$ and node A transmits $X_{\text{A}6}(W_{\text{A}2}, W_{\text{A}5}, W_{\text{A}6})$. Again, the two codewords are not independent, but correlated by design.

After phase 6, each terminal reliably decodes all parts of the message transmitted by the respective other terminal. Reliable decoding at terminal A imposes the conditions

$$\begin{aligned}
R_{\text{B}1} + R_{\text{B}4} &< t_4 I(X_{\text{R}4}; Y_{\text{A}4}), \\
R_{\text{B}2} + R_{\text{B}5} &< t_5 I(X_{\text{R}5}; Y_{\text{A}5}), \\
R_{\text{B}6} &< t_5 I(X_{\text{B}5}; Y_{\text{A}5} | X_{\text{R}5}), \\
R_{\text{B}3} &< t_2 I(X_{\text{B}2}; Y_{\text{A}2}),
\end{aligned} \tag{B.21}$$

whereas reliable decoding at terminal B requires

$$\begin{aligned} R_{A1} + R_{A4} &< t_4 I(X_{R4}; Y_{B4}), \\ R_{A2} + R_{A5} &< t_6 I(X_{R6}; Y_{B6}), \\ R_{A6} &< t_6 I(X_{A6}; Y_{B6}|X_{R6}), \\ R_{A3} &< t_1 I(X_{A1}; Y_{B1}). \end{aligned} \tag{B.22}$$

Noting that $R_A = \sum_{k=1}^{6} R_{Ak}$ and $R_B = \sum_{k=1}^{6} R_{Bk}$, putting all constraints together, and taking the closure of the resulting achievable rate region yields $\mathcal{R}_{DF}$.

B.3.2 Partial Decode-and-Forward

Let $W_A \in \{1, 2, \ldots, \lceil 2^{nR_A} \rceil\}$ denote the message to be transmitted from node A to node B, and let $W_B \in \{1, 2, \ldots, \lceil 2^{nR_B} \rceil\}$ denote the message to be sent from terminal B to terminal A. Both messages are split into seven independent parts, i.e., $W_A = (W_{A1}, \ldots, W_{A7})$ and $W_B = (W_{B1}, \ldots, W_{B7})$ such that $W_{Ak} \in \{1, 2, \ldots, \lceil 2^{nR_{Ak}} \rceil\}$ and $W_{Bk} \in \{1, 2, \ldots, \lceil 2^{nR_{Bk}} \rceil\}$, $k \in \{1, \ldots, 7\}$. The messages are then conveyed to the respective other terminal as follows:

Phase 1: Terminal A first chooses $U_A(W_{A1}, W_{A2}, W_{A3})$ and subsequently generates $X_{A1}(W_{A1}, W_{A2}, W_{A3}, W_{A4})$ using superposition coding. It then transmits X_{A1}.

Phase 2: Terminal B first chooses $U_B(W_{B1}, W_{B2}, W_{B3})$ and subsequently generates $X_{B2}(W_{B1}, W_{B2}, W_{B3}, W_{B4})$ using superposition coding. It then transmits X_{B2}.

Phase 3: Node A transmits $X_{A3}(W_{A5}, W_{A6})$ and node B sends $X_{B3}(W_{B5}, W_{B6})$. The two codewords are independent.

After phase 3, the relay reliably decodes the messages $(W_{A1}, W_{A2}, W_{A3}, W_{A5}, W_{A6})$ and $(W_{B1}, W_{B2}, W_{B3}, W_{B5}, W_{B6})$, which requires

$$\begin{aligned} R_{A1} + R_{A2} + R_{A3} &< t_1 I(U_A; Y_{R1}), \\ R_{B1} + R_{B2} + R_{B3} &< t_2 I(U_B; Y_{R2}), \\ R_{A5} + R_{A6} &< t_3 I(X_{A3}; Y_{R3}|X_{B3}), \\ R_{B5} + R_{B6} &< t_3 I(X_{B3}; Y_{R3}|X_{A3}), \\ R_{A5} + R_{A6} + R_{B5} + R_{B6} &< t_3 I(X_{A3}, X_{B3}; Y_{R3}). \end{aligned} \tag{B.23}$$

Phase 4: The relay transmits $X_{R4}(W_{A1}, W_{A5}, W_{B1}, W_{B5})$.

Phase 5: The relay sends $X_{R5}(W_{B2}, W_{B6})$ and node B transmits $X_{B5}(W_{B2}, W_{B6}, W_{B7})$. Note that the two codewords are not independent, but correlated by design in general.

Phase 6: The relay sends $X_{R6}(W_{A2}, W_{A6})$ and node A transmits $X_{A6}(W_{A2}, W_{A6}, W_{A7})$. Again, the two codewords are not independent, but correlated by design.

After phase 6, each terminal reliably decodes all parts of the message transmitted by the respective other terminal. Reliable decoding at terminal A imposes the conditions

$$\begin{aligned} R_{\mathrm{B}1} + R_{\mathrm{B}5} &< t_4 I(X_{\mathrm{R}4}; Y_{\mathrm{A}4}), \\ R_{\mathrm{B}2} + R_{\mathrm{B}6} &< t_5 I(X_{\mathrm{R}5}; Y_{\mathrm{A}5}), \\ R_{\mathrm{B}7} &< t_5 I(X_{\mathrm{B}5}; Y_{\mathrm{A}5} | X_{\mathrm{R}5}), \\ R_{\mathrm{B}3} &< t_2 I(U_{\mathrm{B}}; Y_{\mathrm{A}2}), \\ R_{\mathrm{B}4} &< t_2 I(X_{\mathrm{B}2}; Y_{\mathrm{A}2} | U_{\mathrm{B}}), \end{aligned} \tag{B.24}$$

whereas reliable decoding at terminal B requires

$$\begin{aligned} R_{\mathrm{A}1} + R_{\mathrm{A}5} &< t_4 I(X_{\mathrm{R}4}; Y_{\mathrm{B}4}), \\ R_{\mathrm{A}2} + R_{\mathrm{A}6} &< t_6 I(X_{\mathrm{R}6}; Y_{\mathrm{B}6}), \\ R_{\mathrm{A}7} &< t_6 I(X_{\mathrm{A}6}; Y_{\mathrm{B}6} | X_{\mathrm{R}6}), \\ R_{\mathrm{A}3} &< t_1 I(U_{\mathrm{A}}; Y_{\mathrm{B}1}), \\ R_{\mathrm{A}4} &< t_1 I(X_{\mathrm{A}1}; Y_{\mathrm{B}1} | U_{\mathrm{A}}). \end{aligned} \tag{B.25}$$

Noting that $R_{\mathrm{A}} = \sum_{k=1}^{7} R_{\mathrm{A}k}$ and $R_{\mathrm{B}} = \sum_{k=1}^{7} R_{\mathrm{B}k}$, putting all constraints together, and taking the closure of the resulting achievable rate region yields $\mathcal{R}_{\mathrm{PDF}}$.

Appendix C

Abbreviations and Acronyms

AF	amplify-and-forward
bpcu	bits per channel use
CF	compress-and-forward
CSB	cut-set bound
CSI	channel state information
CSOB	cut-set outer bound
DCP	disciplined convex programming
DF	decode-and-forward
FDD	frequency-division duplex
GSVD	generalized singular value decomposition
HBC	hybrid broadcast
IAA	inner approximation algorithm
i.i.d.	independent and identically distributed
KKT	Karush–Kuhn–Tucker
MABC	multiple-access broadcast
MIMO	multiple-input multiple-output
OWTS	one-way time-sharing
P2P	point-to-point
PDF	partial decode-and-forward
SDP	semidefinite program
SVD	singular value decomposition
s.t.	subject to
TDBC	time-division broadcast
TDD	time-division duplex
WSR	weighted sum rate
ZF	zero-forcing

Bibliography

[1] T. Adalı, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101–5125, November 2011.

[2] A. Albert, "Conditions for positive and nonnegative definiteness in terms of pseudoinverses," *SIAM Journal on Applied Mathematics*, vol. 17, no. 2, pp. 434–440, 1969.

[3] 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.

[4] M. R. Aref, "Information flow in relay networks," Ph.D. dissertation, Stanford University, Stanford, October 1980.

[5] K. I. Ashar, V. Prathyusha, S. Bhashyam, and A. Thangaraj, "Outer bounds for the capacity region of a Gaussian two-way relay channel," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, October 2012, pp. 1645–1652.

[6] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. Hoboken: Wiley-Interscience, 2006.

[7] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Englewood Cliffs: Prentice Hall, 1992.

[8] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont: Athena Scientific, 2003.

[9] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. New York: Cambridge University Press, 2007.

[10] F. Boccardi and H. Huang, "A near-optimum technique using linear precoding for the MIMO broadcast channel," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2007, pp. III–17–III–20.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization.* New York: Cambridge University Press, 2004.

[12] S. I. Bross, "A converse result for the discrete memoryless relay channel with relay–transmitter feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3789–3792, August 2006.

[13] S. I. Bross and M. A. Wigger, "On the relay channel with receiver–transmitter feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 275–291, January 2009.

[14] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, "Information bottleneck for Gaussian variables," *Journal of Machine Learning Research*, vol. 6, pp. 165–188, January 2005.

[15] E. W. Cheney and A. A. Goldstein, "Newton's method for convex programming and Tchebycheff approximation," *Numerische Mathematik*, vol. 1, no. 1, pp. 253–268, 1959.

[16] M. Chiang, "Nonconvex optimization for communication networks," in *Advances in Applied Mathematics and Global Optimization*, ser. Advances in Mechanics and Mathematics, D. Y. Gao and H. D. Sherali, Eds. New York: Springer, 2009, vol. 17, ch. 5, pp. 137–196.

[17] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2363–2371, October 2003.

[18] H.-F. Chong and M. Motani, "The capacity of several new classes of semi-deterministic relay channels," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6397–6404, October 2011.

[19] H.-F. Chong and M. Motani, "On achievable rates for the general relay channel," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1249–1266, March 2011.

[20] H.-F. Chong, M. Motani, and H. K. Garg, "Generalized backward decoding strategies for the relay channel," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 394–401, January 2007.

[21] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, September 1979.

[22] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Transactions on Information Theory*, vol. 26, no. 6, pp. 648–657, November 1980.

[23] T. M. Cover and Y.-H. Kim, "Capacity of a class of deterministic relay channels," in *Proceedings of IEEE International Symposium on Information Theory*, June 2007, pp. 591–595.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken: Wiley-Interscience, 2006.

[25] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1541–1553, September 2012.

[26] M. Debbah and R. R. Müller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.

[27] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3072–3081, November 2001.

[28] M. Dohler and Y. Li, *Cooperative Communications: Hardware, Channel & PHY*. West Sussex: Wiley-Interscience, 2010.

[29] A. Dotzler, M. Riemensberger, and W. Utschick, "Minimax duality for MIMO interference networks," *Information*, vol. 7, no. 2, pp. 1–42, June 2016.

[30] H. G. Eggleston, *Convexity.* Cambridge: Cambridge University Press, 1963.

[31] A. El Gamal, "On information flow in relay networks," in *Proceedings of National Telecommunications Conference*, vol. 2, November–December 1981, pp. D4.1.1–D4.1.4.

[32] A. El Gamal and M. R. Aref, "The capacity of the semideterministic relay channel," *IEEE Transactions on Information Theory*, vol. 28, no. 3, p. 536, May 1982.

[33] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3413–3431, October 2007.

[34] A. El Gamal and Y.-H. Kim, *Network Information Theory.* New York: Cambridge University Press, 2011.

[35] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1545–1561, April 2006.

[36] A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

[37] J. Elzinga and T. G. Moore, "A central cutting plane algorithm for the convex programming problem," *Mathematical Programming*, vol. 8, no. 1, pp. 134–145, 1975.

[38] A. Feiten, S. Hanly, and R. Mathar, "Derivatives of mutual information in Gaussian vector channels with applications," in *Proceedings of IEEE International Symposium on Information Theory*, June 2007, pp. 2296–2300.

[39] Y. Gabbai and S. I. Bross, "Achievable rates for the discrete memoryless relay channel with partial feedback configurations," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4989–5007, November 2006.

[40] L. Gerdes, C. Hellings, L. Weiland, and W. Utschick, "On the maximum achievable partial decode-and-forward rate for the Gaussian MIMO relay channel," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6751–6758, December 2015.

[41] L. Gerdes, M. Riemensberger, and W. Utschick, "On achievable rate regions for half-duplex Gaussian MIMO relay channels: A decomposition approach," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1319–1330, September 2012.

[42] L. Gerdes, M. Riemensberger, and W. Utschick, "Utility maximization in the half-duplex two-way Gaussian MIMO relay channel," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2012, pp. 2897–2900.

[43] L. Gerdes, M. Riemensberger, and W. Utschick, "Bounds on the capacity regions of half-duplex Gaussian MIMO relay channels," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 43, pp. 1–17, March 2013.

[44] L. Gerdes, M. Riemensberger, and W. Utschick, "On the equivalence of degraded Gaussian MIMO broadcast channels," in *Proceedings of International ITG Workshop on Smart Antennas*, March 2015.

[45] L. Gerdes and W. Utschick, "Optimized capacity bounds for the half-duplex Gaussian MIMO relay channel," in *Proceedings of International ITG Workshop on Smart Antennas*, February 2011.

[46] L. Gerdes and W. Utschick, "Optimized capacity bounds for the MIMO relay channel," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2011, pp. 3336–3339.

[47] L. Gerdes, L. Weiland, M. Riemensberger, and W. Utschick, "Optimal partial decode-and-forward rates for stochastically degraded Gaussian relay channels," in *Proceedings of Conference on Information Sciences and Systems*, March 2014.

[48] L. Gerdes, L. Weiland, and W. Utschick, “A zero-forcing partial decode-and-forward scheme for the Gaussian MIMO relay channel,” in *Proceedings of IEEE International Conference on Communications*, June 2013, pp. 3349–3354.

[49] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, “Outdoor MIMO wireless channels: Models and performance prediction,” *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 1926–1934, December 2002.

[50] A. Goldsmith, *Wireless Communications.* New York: Cambridge University Press, 2005.

[51] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[52] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Berlin: Springer, 2008, vol. 371, pp. 95–110.

[53] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.0 beta,” http://cvxr.com/cvx, September 2013.

[54] M. Grant, S. Boyd, and Y. Ye, “Disciplined convex programming,” in *Global Optimization*, ser. Nonconvex Optimization and Its Applications, P. Pardalos, L. Liberti, and N. Maculan, Eds. New York: Springer, 2006, vol. 84, pp. 155–210.

[55] I. Hammerström, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, “MIMO two-way relaying with transmit CSI at the relay,” in *Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications*, June 2007.

[56] P. Harremoës and N. Tishby, “The information bottleneck revisited or how to choose a good distortion measure,” in *Proceedings of IEEE International Symposium on Information Theory*, June 2007, pp. 566–570.

[57] C. Hellings, L. Gerdes, L. Weiland, and W. Utschick, “On optimal Gaussian signaling in MIMO relay channels with partial decode-and-forward,” *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3153–3164, June 2014.

[58] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms II*, ser. A Series of Comprehensive Studies in Mathematics. Berlin: Springer, 1996, 2nd corrected printing.

[59] S. Hong, J. Brand, J. I. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, “Applications of self-interference cancellation in 5G and beyond,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, February 2014.

[60] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge: Cambridge University Press, 1985.

[61] R. A. Horn and F. Zhang, “Basic properties of the Schur complement,” in *The Schur Complement and Its Applications*, ser. Numerical Methods and Algorithms, F. Zhang, Ed. New York: Springer, 2005, vol. 4, ch. 1, pp. 17–46.

[62] A. Høst-Madsen, “On the capacity of wireless relaying,” in *Proceedings of IEEE Vehicular Technology Conference*, September 2002, pp. 1333–1337.

[63] A. Høst-Madsen and J. Zhang, “Capacity bounds and power allocation for wireless relay channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.

[64] J. E. Kelley, Jr., “The cutting-plane method for solving convex programs,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 4, pp. 703–712, December 1960.

[65] F. Kelly, “Charging and rate control for elastic traffic,” *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, January–February 1997.

[66] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, “Bounds on achievable rates for general multi-terminal networks with practical constraints,” in *Information Processing in Sensor Networks*, ser. Lecture Notes in Computer Science, F. Zhao and L. Guibas, Eds. Berlin: Springer, 2003, vol. 2634, pp. 146–161.

[67] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, “On capacity of Gaussian ’cheap’ relay channel,” in *Proceedings of IEEE Global Telecommunications Conference*, vol. 3, December 2003, pp. 1776–1780.

[68] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, “On the capacity of ‘cheap’ relay networks,” in *Proceedings of Conference on Information Sciences and Systems*, March 2003.

[69] M. N. Khormuji, “Coding schemes for relay networks,” Ph.D. dissertation, KTH Royal Institute of Technology, Stockholm, September 2011.

[70] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, “Comparison of bi-directional relaying protocols,” in *Proceedings of IEEE Sarnoff Symposium*, April 2008.

[71] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, “Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols,” *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6405–6418, October 2011.

[72] S. J. Kim, P. Mitran, and V. Tarokh, “Performance bounds for bidirectional coded cooperation protocols,” *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5235–5241, November 2008.

[73] Y.-H. Kim, "Coding techniques for primitive relay channels," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, September 2007, pp. 129–135.

[74] Y.-H. Kim, "Capacity of a class of deterministic relay channels," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 1328–1329, March 2008.

[75] G. Kramer, "Models and theory for relay channels with receive constraints," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, September–October 2004, pp. 1312–1321.

[76] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 4-5, pp. 265–444, 2007.

[77] 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.

[78] G. Kramer, I. Marić, and R. D. Yates, "Cooperative communications," *Foundations and Trends in Networking*, vol. 1, no. 3-4, pp. 271–425, 2006.

[79] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *Proceedings of IEEE Information Theory Workshop*, September 2007, pp. 313–318.

[80] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[81] Q. Li, G. Li, W. Lee, M.-i. Lee, D. Mazzarese, B. Clerckx, and Z. Li, "MIMO techniques in WiMAX and LTE: A feature overview," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 86–92, May 2010.

[82] Y. Liang and V. V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3284–3289, September 2005.

[83] Y. Liang, V. V. Veeravalli, and H. V. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3432–3453, October 2007.

[84] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, June 2009.

[85] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Transactions on Information Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[86] C. K. Lo, S. Vishwanath, and R. W. Heath, Jr., "Rate bounds for MIMO relay channels," *Journal of Communications and Networks*, vol. 10, no. 2, pp. 194–203, June 2008.

[87] K. Luo, R. H. Gohary, and H. Yanikomeroglu, "Analysis of the generalized DF-CF for Gaussian relay channels: Decode or compress?" *IEEE Transactions on Communications*, vol. 61, no. 5, pp. 1810–1821, May 2013.

[88] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, July–August 1978.

[89] G. Marsaglia and G. P. H. Styan, "Equalities and inequalities for ranks of matrices," *Linear and Multilinear Algebra*, vol. 2, no. 3, pp. 269–292, 1974.

[90] K. Meyberg and P. Vachenauer, *Höhere Mathematik 1*, 6th ed. Berlin: Springer, 2001.

[91] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra.* Philadelphia: SIAM, 2000.

[92] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, October 2000.

[93] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science*, ser. Applied Mathematical Sciences, J. E. Marsden and L. Sirovich, Eds. New York: Springer, 2000, vol. 40.

[94] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[95] C. T. K. Ng and G. J. Foschini, "Transmit signal and bandwidth optimization in multiple-antenna relay channels," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 2987–2992, November 2011.

[96] T. J. Oechtering, C. Schnurr, I. Bjelaković, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 454–458, January 2008.

[97] T. J. Oechtering, R. F. Wyrembelski, and H. Boche, "Multiantenna bidirectional broadcast channels—optimal transmit strategies," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1948–1958, May 2009.

[98] T. J. Oechtering, R. F. Wyrembelski, and H. Boche, "On the optimal transmission for the MIMO bidirectional broadcast channel," in *Proceedings of IEEE International Conference on Communications*, June 2009.

[99] C. C. Paige and M. A. Saunders, "Towards a generalized singular value decomposition," *SIAM Journal on Numerical Analysis*, vol. 18, no. 3, pp. 398–405, June 1981.

[100] P. Popovski and T. Koike-Akino, "Coded bidirectional relaying in wireless networks," in *New Directions in Wireless Communications Research*, V. Tarokh, Ed. New York: Springer, 2009, ch. 11, pp. 291–316.

[101] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, October–November 2005, pp. 1066–1071.

[102] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 379–389, February 2007.

[103] Y. Rong and F. Gao, "Optimal beamforming for non-regenerative MIMO relays with direct link," *IEEE Communications Letters*, vol. 13, no. 12, pp. 926–928, December 2009.

[104] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, September 2014.

[105] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on the optimization of amplify-and-forward MIMO relay systems," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1331–1346, September 2012.

[106] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Hierarchical sensor networks: Capacity bounds and cooperative strategies using the multiple-access relay channel models," in *IEEE International Conference on Sensor and Ad Hoc Communications and Networks*, October 2004, pp. 191–199.

[107] H. Sato, "Information transmission through a channel with relay," University of Hawaii, Honolulu, Tech. Rep. B76-7, The Aloha System, March 1976.

[108] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proceedings of IEEE International Symposium on Information Theory*, June 2000, p. 22.

[109] C. Schnurr, T. J. Oechtering, and S. Stanczak, "Achievable rates for the restricted half-duplex two-way relay channel," in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, November 2007, pp. 1468–1472.

[110] C. Schnurr, S. Stanczak, and T. J. Oechtering, “Achievable rates for the restricted half-duplex two-way relay channel under a partial-decode-and-forward protocol,” in *Proceedings of IEEE Information Theory Workshop*, May 2008, pp. 134–138.

[111] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals.* Cambridge University Press, 2010.

[112] S. Serbetli and A. Yener, “Relay assisted F/TDMA ad hoc networks: Node classification, power allocation and relaying strategies,” *IEEE Transactions on Communications*, vol. 56, no. 6, pp. 937–947, June 2008.

[113] X. Shang, B. Chen, G. Kramer, and H. V. Poor, “MIMO Z-interference channels: Capacity under strong and noisy interference,” in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, November 2009, pp. 667–671.

[114] X. Shang and H. V. Poor, “Noisy-interference sum-rate capacity for vector Gaussian interference channels,” *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 132–153, January 2013.

[115] I. Shomorony and A. S. Avestimehr, “Worst-case additive noise in wireless networks,” *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3833–3847, June 2013.

[116] S. Simoens, O. Muñoz-Medina, J. Vidal, and A. del Coso, “On the Gaussian MIMO relay channel with full channel state information,” *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3588–3599, September 2009.

[117] S. Simoens, O. Muñoz-Medina, J. Vidal, and A. del Coso, “Compress-and-forward cooperative MIMO relaying with full channel state information,” *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 781–791, February 2010.

[118] M. Stein, “Two-way relay channel with half-duplex constraint,” Diploma thesis, Fachgebiet Methoden der Signalverarbeitung, Technische Universität München, Munich, September 2010.

[119] J. F. Sturm, “Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones,” *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999, current version and user’s guide available at http://sedumi.ie.lehigh.edu.

[120] R. Tandon and S. Ulukus, “A new upper bound on the capacity of a class of primitive relay channels,” in *Proceedings of Allerton Conference on Communication, Control, and Computing*, September 2008, pp. 1562–1569.

[121] X. Tang and Y. Hua, “Optimal design of non-regenerative MIMO wireless relays,” *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1398–1407, April 2007.

[122] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November–December 1999.

[123] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Transactions on Information Theory*, vol. 33, no. 5, pp. 711–716, September 1987.

[124] R. Timo, T. J. Oechtering, and M. A. Wigger, "Source coding problems with conditionally less noisy side information," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5516–5532, September 2014.

[125] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, September 1999, pp. 368–377.

[126] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3 — a Matlab software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 545–581, 1999, current version and user's guide available at http://www.math.nus.edu.sg/~mattohkc/sdpt3.html.

[127] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "On the implementation and usage of SDPT3 – a Matlab software package for semidefinite-quadratic-linear programming, version 4.0," in *Handbook on Semidefinite, Conic and Polynomial Optimization*, ser. International Series in Operations Research & Management Science, M. F. Anjos and J. B. Lasserre, Eds. New York: Springer, 2012, vol. 166, ch. 25, pp. 715–754.

[128] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* New York: Cambridge University Press, 2005.

[129] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel." Ph.D. dissertation, Department of Statistics, University of California, Berkeley, February 1968.

[130] E. C. van der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[131] E. C. van der Meulen, "A survey of the relay channel," in *Multiple Access Channels: Theory and Practice*, E. Biglieri and L. Györfi, Eds. Boston: IOS Press, 2007, vol. 10, pp. 73–96.

[132] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 29–43, January 2005.

[133] C. Wang, D. Sun, and K.-C. Toh, "Solving log-determinant optimization problems by a Newton-CG primal proximal point algorithm," *SIAM Journal on Optimization*,

vol. 20, no. 6, pp. 2994–3013, 2010, current version and user's guide available at http://www.math.nus.edu.sg/~mattohkc/LogdetPPA.html.

[134] R. Wang and M. Tao, "Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1352–1365, March 2012.

[135] L. Weiland, "Achievable rates for the Gaussian multiple-input multiple-output relay channel using the partial decode-and-forward scheme," Master's thesis, Fachgebiet Methoden der Signalverarbeitung, Technische Universität München, Munich, December 2012.

[136] L. Weiland, L. Gerdes, and W. Utschick, "Partial decode-and-forward rates for the Gaussian MIMO relay channel: Inner approximation of non-convex rate constraints," in *Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications*, June 2013, pp. 540–544.

[137] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, September 2006.

[138] J. Wolfowitz, *Coding Theorems of Information Theory*, 3rd ed., ser. A Series of Modern Surveys in Mathematics. Berlin: Springer, 1978.

[139] Y. Wu, "Broadcasting when receivers know some messages a priori," in *Proceedings of IEEE International Symposium on Information Theory*, June–July 2007, pp. 1141–1145.

[140] R. F. Wyrembelski, T. J. Oechtering, I. Bjelaković, C. Schnurr, and H. Boche, "Capacity of Gaussian MIMO bidirectional broadcast channels," in *Proceedings of IEEE International Symposium on Information Theory*, July 2008, pp. 584–588.

[141] R. F. Wyrembelski, T. J. Oechtering, and H. Boche, "Decode-and-forward strategies for bidirectional relaying," in *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2008.

[142] S. Xu and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way MIMO relay system," *IEEE Transactions on Wireless Communications*, vol. 10, no. 5, pp. 1645–1655, May 2011.

[143] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMAX and LTE-Advanced mobile systems," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 100–105, October 2009.

[144] S. Yao, M. Skoglund, T. T. Kim, and H. V. Poor, "Half-duplex relaying based on quantize-and-forward," in *Proceedings of IEEE International Symposium on Information Theory*, July 2011, pp. 2447–2451.

[145] R. W. Yeung, *Information Theory and Network Coding*. New York: Springer, 2008.

[146] T. Yoo and A. 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.

[147] S. Zahedi, "On reliable communication over relay channels," Ph.D. dissertation, Stanford University, Stanford, August 2005.

[148] W. I. Zangwill, *Nonlinear Programming: A Unified Approach*. Englewood Cliffs: Prentice Hall, 1969.

[149] L. Zhang, J. Jiang, A. 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.

[150] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 5, pp. 699–712, June 2009.

[151] Z. Zhang, "Partial converse for a relay channel," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1106–1110, September 1988.