RESEARCH   **Open Access**

Bounds on the capacity regions of half-duplex Gaussian MIMO relay channels

Lennart Gerdes[*], Maximilian Riemensberger and Wolfgang Utschick

Abstract

This article considers uni- and bidirectional communication in the half-duplex Gaussian multiple-input multiple-output (MIMO) relay channel. Assuming perfect channel state information at all nodes and the use of time division duplex communications protocol to separate transmissions and receptions at all nodes, we propose a dual decomposition approach to efficiently determine upper and lower bounds on the capacity and the capacity region of the half-duplex relay channel and the restricted half-duplex two-way relay channel, respectively. Our approach allows to quantify the fundamental limits of the considered relay networks, and the obtained results may serve as benchmarks when studying different and/or suboptimal relay strategies or the impact of channel estimation errors. Furthermore, we discuss how our dual decomposition approach may be used for designing optimal resource allocation protocols.

Keywords: Relay channel, Two-way relay channel, MIMO, Half-duplex, Fundamental limits, Cut-set outer bound, Decode-and-forward

1 Introduction

A central aspect of today's and future wireless network standards is the question of how to provide high-speed and high-quality service to a steadily growing number of mobile users without an increase of available bandwidth. One means 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 [1,2]. 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 usually do not have own information to transmit. The concept of relaying goes back as far as 1971 when van der Meulen introduced the relay channel model [3]. In contrast to point-to-point channels, the capacity of the relay channel remains unknown in general, but upper and lower bounds have of course been derived [4].

In this study, we consider the combination of multiple-antenna systems and the concept of relaying. In particular, we determine upper and lower bounds on the capacity and the capacity region of the Gaussian multiple-input multiple-output (MIMO) relay channel and the

Gaussian MIMO two-way relay channel with a half-duplex constraint[a]. While this topic is interesting and relevant in itself, note also that both the relay channel and the two-way relay channel are elementary building blocks of general multi-hop wireless networks. A fundamental understanding of these two small networks and their performance limits can thus help to determine the limits on the performances of larger communication networks, e.g., by decomposing a larger network into subgraphs whose performances can be more easily specified.

In their pioneering study on the relay channel, Cover and El Gamal derived a capacity upper bound and achievable rates based on a then new cut-set bound (CSB) and two coding schemes that are now referred to as *decode-and-forward* (DF) and *compress-and-forward* (CF), respectively. In [5,6], the cut-set bound and the DF scheme are used to derive bounds on the capacity of the half-duplex relay channel. For Gaussian single-antenna channels, corresponding bounds are presented in [7,8]. A (generally loose) upper bound to the CSB of the full-duplex Gaussian MIMO relay channel is provided in [9]. Achievable rates for this channel based on point-to-point transmission, the cascaded relay channel, and a suboptimal DF scheme are given, too. In [10], it is shown that the cut-set bound and the maximum achievable DF rate

*Correspondence: gerdes@tum.de
Associate Institute for Signal Processing, Technische Universität München, 80290 München, Germany

for this MIMO relay channel can be obtained as the solutions of convex optimization problems, which also holds if a half-duplex constraint is imposed and frequency division duplex (FDD) with an average power constraint is considered. For the full-duplex case, the same result was later independently derived in [11] and then extended to the half-duplex relay channel with a time division duplex (TDD) protocol and per protocol phase transmit power constraints imposed on source and relay [12]. A similar study for both the full-duplex and the half-duplex case with TDD is presented in [13]. However, it can be verified that the expressions resulting from those derivations are only upper bounds to the optimal solutions.

The two-way relay channel models the more common and important scenario where two terminals want to exchange information with the aid of a relay. It was introduced in [14], where the authors showed that a significant portion of the loss in spectral efficiency suffered in the one-way relay channel due to the half-duplex constraint can be compensated when bidirectional communication is considered. Most scientific articles have analyzed the half-duplex two-way relay channel in combination with a communication protocol consisting of two phases, a multiple access (MAC) phase and a broadcast (BC) phase [14-18]. In the MAC phase, the terminals transmit their messages to the relay, and subsequently, in the BC phase, the relay broadcasts its message to the terminals. With this protocol, however, all information is sent via the relay since the terminals cannot overhear each other's transmissions due to the half-duplex constraint. As a result, protocols composed of more than two phases that utilize the direct link between the terminals can yield larger achievable rate regions in general [19-21].

The contributions of this article are as follows. We present a dual decomposition approach that allows to evaluate upper and lower bounds on the capacity and the capacity region of the half-duplex Gaussian MIMO relay channel and the restricted half-duplex Gaussian MIMO two-way relay channel, respectively. To this end, perfect channel state information (CSI) at all nodes and the use of TDD protocols to separate transmissions and receptions at all nodes are assumed. We show how the proposed dual decomposition approach can be applied to efficiently tackle the joint optimization of input signals and time allocation that needs to be solved in order to obtain the desired results. In the dual domain, the problem decomposes into subproblems that are easier to solve and for which standard convex optimization tools can be used. With our optimization approach, it is hence possible to efficiently obtain numerical results that quantify the fundamental limits of uni- and bidirectional communication in the half-duplex Gaussian MIMO relay channel. These results can then serve as benchmarks when studying different and/or suboptimal relay strategies or the impact

of channel estimation errors on the performance of the considered relay networks. Moreover, our dual decomposition approach may be used for designing optimal resource allocation protocols, as discussed later in this article.

The approach proposed here is a nontrivial extension of a similar dual decomposition approach presented in [21]. There, we considered bounds on achievable rate regions for the same relay networks, but the transmit powers of all nodes were assumed to be bounded above by some finite value for every protocol phase. In this study, we modify this approach such that it can handle the average transmit power constraints under which the information theoretic capacity bounds (cut-set bound and achievable DF rate) we are interested in here were derived. We remark that the problems we need to solve become considerably more difficult due to the average transmit power constraints, both from a theoretical and practical point of view. This is because we need to introduce more dual variables and because the constraint sets of the subproblems encountered in the dual domain become unbounded. The latter means that several additional mathematical details have to be taken into account in order to ensure correctness of the optimization strategy. What is more, the power constraints considered in [21] can easily be incorporated into the optimization framework presented in this article, which is not the case vice versa. In this sense, the optimization approach presented here is more general than that of [21].

The remainder of this article is organized as follows. Section 2 introduces the system model for the restricted half-duplex Gaussian MIMO two-way relay channel. It should be mentioned here that our analysis focuses on the half-duplex two-way relay channel since it includes the half-duplex relay channel as a special case. In Section 3, we derive an outer bound on the capacity region of the restricted half-duplex Gaussian MIMO two-way relay channel and show how it can numerically be evaluated by means of the aforementioned dual decomposition approach. An inner bound on the capacity region is given by the rate region that can be achieved when the relay uses the decode-and-forward scheme. This achievable rate region and how it can be evaluated is discussed in Section 4. Numerical results for both uni- and bidirectional communication in the half-duplex Gaussian MIMO relay channel are presented in Section 5, and Section 6 concludes the article.

Notation: $\mathbb{R}_+$ stands for the set of nonnegative real numbers. Matrices are denoted by bold capital letters, vectors by bold lowercase characters. The identity matrix, the zero matrix/vector, and the all-ones vector are specified by $\mathbf{I}$, $\mathbf{0}$, and $\mathbf{1}$, respectively, where the dimensions are indicated by subscripts if necessary. $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^{\dagger}$, $\boldsymbol{A}^{\mathrm{T}}$, $\boldsymbol{A}^{\mathrm{H}}$, and $\mathrm{tr}(\boldsymbol{A})$ denote the inverse, Moore-Penrose pseudoinverse,

transpose, conjugate transpose, and trace of a matrix $\boldsymbol{A}$, while $\boldsymbol{A} \succeq \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite. $E[\cdot]$ is the expectation operator and $\boldsymbol{x} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{C})$ means that $\boldsymbol{x}$ is a circularly symmetric complex Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$. Finally, $I(X; Y|Z)$ denotes the conditional mutual information of random variables X and Y given Z and $h(X|Y)$ is the differential entropy of X given Y.

2 System model

In the one-way relay channel, one source transmits information to one destination with the help of a relay. This simple unidirectional relay network is obviously only a special case of the two-way relay channel, where two terminals exchange information with the aid of the relay. Therefore, our analysis focuses on the half-duplex Gaussian MIMO two-way relay channel. More specifically, we consider the *restricted* two-way relay channel, i.e., the bidirectional communication is restricted in the sense that the encoders at the two terminals can neither cooperate, nor are they able to use previously decoded information to encode their messages. The most general communication protocol for this channel model is composed of all six phases (network states) where either one or two nodes transmit, as first noted in [22]. Evidently, no information can be conveyed when all nodes are silent or when all nodes transmit at the same time, where the latter is due to the half-duplex constraint imposed on all nodes. The six different phases are illustrated in Figure 1, where nodes 1 and 2 represent the two terminals and R is the relay.

Let N_A and N_B be the number of antennas at node A and node B, let $\boldsymbol{x}_A^{(i)} \in \mathbb{C}^{N_A}$ and $\boldsymbol{y}_B^{(i)} \in \mathbb{C}^{N_B}$ denote the transmit signal of node A and the receive signal of node B during phase i, respectively, and let $\boldsymbol{H}_{AB} \in \mathbb{C}^{N_B \times N_A}$ denote the channel gain matrix between nodes A and B for all $i \in \{1, \ldots, 6\}$. Then, the phases are characterized as follows:

(1) Node 1 transmits to node 2 and the relay:

$$\boldsymbol{y}_R^{(1)} = \boldsymbol{H}_{1R}\boldsymbol{x}_1^{(1)} + \boldsymbol{n}_R^{(1)}, \quad \boldsymbol{n}_R^{(1)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_R}),$$
$$\boldsymbol{y}_2^{(1)} = \boldsymbol{H}_{12}\boldsymbol{x}_1^{(1)} + \boldsymbol{n}_2^{(1)}, \quad \boldsymbol{n}_2^{(1)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_2}).$$

(2) Node 2 transmits to node 1 and the relay:

$$\boldsymbol{y}_R^{(2)} = \boldsymbol{H}_{2R}\boldsymbol{x}_2^{(2)} + \boldsymbol{n}_R^{(2)}, \quad \boldsymbol{n}_R^{(2)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_R}),$$
$$\boldsymbol{y}_1^{(2)} = \boldsymbol{H}_{21}\boldsymbol{x}_2^{(2)} + \boldsymbol{n}_1^{(2)}, \quad \boldsymbol{n}_1^{(2)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_1}).$$

(3) Node 1 and node 2 transmit to the relay:

$$\boldsymbol{y}_R^{(3)} = \boldsymbol{H}_{1R}\boldsymbol{x}_1^{(3)} + \boldsymbol{H}_{2R}\boldsymbol{x}_2^{(3)} + \boldsymbol{n}_R^{(3)}, \quad \boldsymbol{n}_R^{(3)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_R}).$$

(4) The relay transmits to node 1 and node 2:

$$\boldsymbol{y}_1^{(4)} = \boldsymbol{H}_{R1}\boldsymbol{x}_R^{(4)} + \boldsymbol{n}_1^{(4)}, \quad \boldsymbol{n}_1^{(4)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_1}),$$
$$\boldsymbol{y}_2^{(4)} = \boldsymbol{H}_{R2}\boldsymbol{x}_R^{(4)} + \boldsymbol{n}_2^{(4)}, \quad \boldsymbol{n}_2^{(4)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_2}).$$

(5) The relay and node 2 transmit to node 1:

$$\boldsymbol{y}_1^{(5)} = \boldsymbol{H}_{R1}\boldsymbol{x}_R^{(5)} + \boldsymbol{H}_{21}\boldsymbol{x}_2^{(5)} + \boldsymbol{n}_1^{(5)}, \quad \boldsymbol{n}_1^{(5)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_1}).$$

(6) The relay and node 1 transmit to node 2:

$$\boldsymbol{y}_2^{(6)} = \boldsymbol{H}_{R2}\boldsymbol{x}_R^{(6)} + \boldsymbol{H}_{12}\boldsymbol{x}_1^{(6)} + \boldsymbol{n}_2^{(6)}, \quad \boldsymbol{n}_2^{(6)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I}_{N_2}).$$

Here, we have assumed that the channels are the same for all network states in order to simplify the notation. This is without loss of generality, however, since we anyhow require all channels to be perfectly known at all nodes for the discussions below. Moreover, the additive white Gaussian noise $\boldsymbol{n}_A^{(i)}$ received at node A during phase i is assumed to be independent of the noise $\boldsymbol{n}_B^{(j)}$ received at another node B for all phases $j \in \{1, \ldots, 6\}$ and independent of $\boldsymbol{n}_A^{(j)}$ for all $j \neq i$.

With each node A that transmits in the ith phase a transmit covariance matrix

$$\boldsymbol{R}_A^{(i)} = E\left[\boldsymbol{x}_A^{(i)}\boldsymbol{x}_A^{(i),H}\right] \tag{1}$$

is associated, and the average transmit power consumed by the node during this phase is given by $p_A^{(i)} = \operatorname{tr}(\boldsymbol{R}_A^{(i)})$. Furthermore, if the two nodes A and B transmit simultaneously during phase i, we have a joint transmit covariance matrix

$$\boldsymbol{R}^{(i)} = E\left[\begin{bmatrix}\boldsymbol{x}_A^{(i)}\\\boldsymbol{x}_B^{(i)}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_A^{(i)}\\\boldsymbol{x}_B^{(i)}\end{bmatrix}^H\right] = \begin{bmatrix}\boldsymbol{R}_A^{(i)} & \boldsymbol{R}_{AB}^{(i)}\\\boldsymbol{R}_{AB}^{(i),H} & \boldsymbol{R}_B^{(i)}\end{bmatrix} \tag{2}$$

for this phase. By defining the selection matrices

$$\boldsymbol{D}_A^{(i)} = \begin{bmatrix}\mathbf{I}_{N_A} & \mathbf{0}_{N_A \times N_B}\end{bmatrix}, \quad \boldsymbol{D}_B^{(i)} = \begin{bmatrix}\mathbf{0}_{N_B \times N_A} & \mathbf{I}_{N_B}\end{bmatrix}, \tag{3}$$

the transmit covariance matrices $\boldsymbol{R}_A^{(i)}$ and $\boldsymbol{R}_B^{(i)}$ of the two transmitting nodes can be expressed as linear functions of the joint transmit covariance matrix $\boldsymbol{R}^{(i)}$:

$$\boldsymbol{R}_A^{(i)} = \boldsymbol{D}_A^{(i)}\boldsymbol{R}^{(i)}\boldsymbol{D}_A^{(i),H}, \quad \boldsymbol{R}_B^{(i)} = \boldsymbol{D}_B^{(i)}\boldsymbol{R}^{(i)}\boldsymbol{D}_B^{(i),H}. \tag{4}$$

3 Outer bound on capacity region

In this section, we establish an outer bound on the capacity region of the restricted half-duplex Gaussian MIMO two-way relay channel and, as our main contribution, propose an efficient method to evaluate it. The outer bound region is obtained by applying the cut-set bound, which was originally derived for the one-way relay channel in [4], to the information flow from node 1 to node 2 as well as to the information flow from node 2 to node 1. In particular, we first consider the cut-set outer bound for the general half-duplex two-way relay channel and then show that, for Gaussian channels, it is equivalent to the cut-set outer bound for the restricted half-duplex two-way relay channel. While it is not known
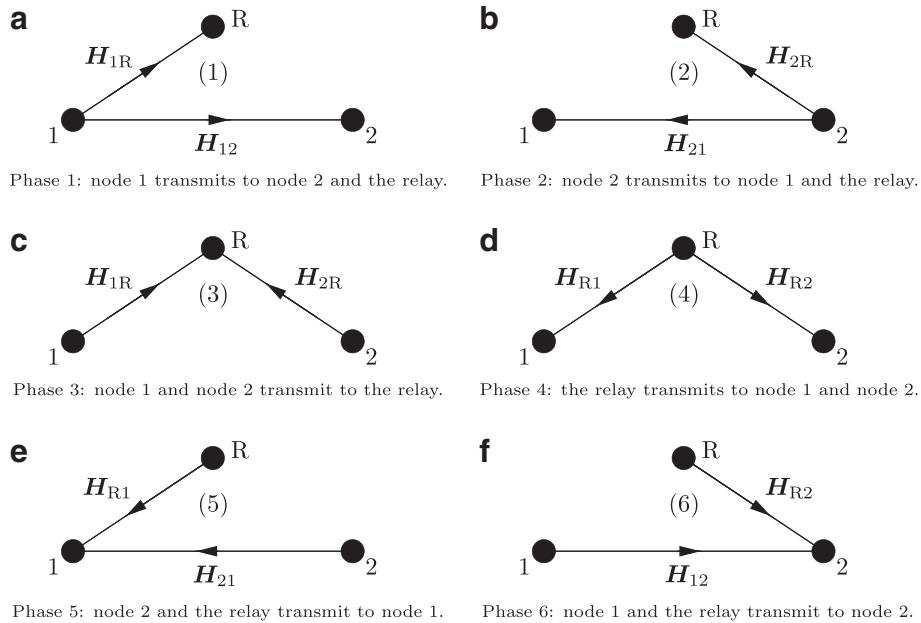
Figure 1 Half-duplex MIMO two-way relay channel. (a) Phase 1: node 1 transmits to node 2 and relay. **(b)** Phase 2: node 2 transmits to node 1 and relay. **(c)** Phase 3: node 1 and node 2 transmit to relay. **(d)** Phase 4: relay transmits to node 1 and node 2. **(e)** Phase 5: relay and node 2 transmit to node 1. **(f)** Phase 6: relay and node 1 transmit to node 2.
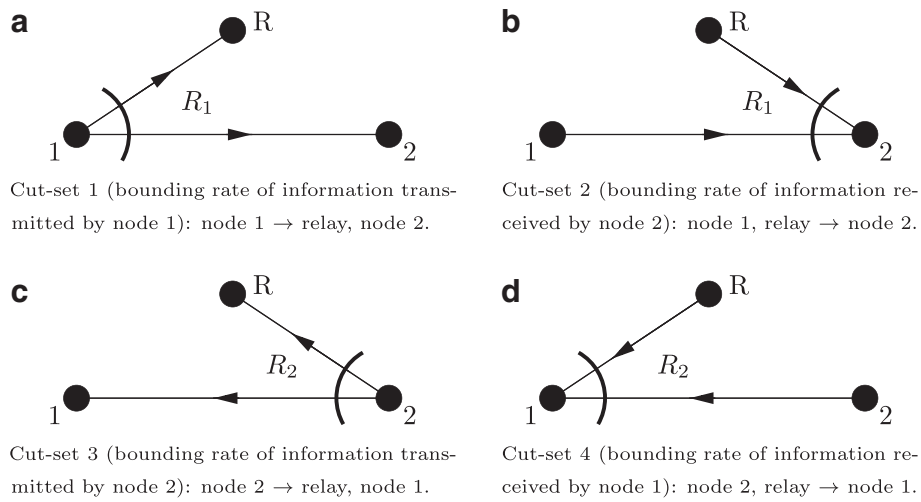
whether the cut-set bound is tight in general, there is no known tighter bound for the relay channel. What is more, it is tight for all classes of relay channels for which the capacity is known. These include the physically degraded and the reversely degraded relay channel [4], the semideterministic relay channel [23], and the relay channel with orthogonal components [24].

Theorem 1. *Suppose (R_1, R_2) is an achievable rate pair for the half-duplex two-way relay channel, where R_1 is associated with the rate of the information sent from node 1 to node 2 and R_2 with that of the reverse direction. Then,*

where $X_A^{(i)}$ and $Y_B^{(i)}$ represent the channel input of node A *and the channel output of node* B *during phase i, respectively, and the duration of the ith phase is denoted by τ_i.*

Proof. The result directly follows from ([5], Thm. 1) by considering all six network states (TDD phases) and both directions of data transmission. In particular, the rate bounds originate from the four cut-sets depicted in Figure 2; the first two cut-sets (shown in Figures 2a,b) yield the upper bounds on R_1, whereas the bounds on R_2 are determined by the third (Figure 2c) and fourth

$$(R_1, R_2) \in \mathcal{C}_{OB} = \bigcup_{\prod_{i=1}^{6} p_{X_1^{(i)} X_2^{(i)} X_R^{(i)}}} \left\{ (C_1, C_2) \in \mathbb{R}_+^2 : \tau_i \geq 0, \forall i \in \{1, \ldots, 6\}, \sum_{i=1}^{6} \tau_i = 1, \right.$$

$$C_1 \leq \tau_1 I(X_1^{(1)}; Y_R^{(1)} Y_2^{(1)}) + \tau_3 I(X_1^{(3)}; Y_R^{(3)}|X_2^{(3)}) + \tau_6 I(X_1^{(6)}; Y_2^{(6)}|X_R^{(6)}),$$

$$C_1 \leq \tau_1 I(X_1^{(1)}; Y_2^{(1)}) + \tau_4 I(X_R^{(4)}; Y_2^{(4)}) + \tau_6 I(X_1^{(6)} X_R^{(6)}; Y_2^{(6)}),$$

$$C_2 \leq \tau_2 I(X_2^{(2)}; Y_R^{(2)} Y_1^{(2)}) + \tau_3 I(X_2^{(3)}; Y_R^{(3)}|X_1^{(3)}) + \tau_5 I(X_2^{(5)}; Y_1^{(5)}|X_R^{(5)}),$$

$$\left. C_2 \leq \tau_2 I(X_2^{(2)}; Y_1^{(2)}) + \tau_4 I(X_R^{(4)}; Y_1^{(4)}) + \tau_5 I(X_2^{(5)} X_R^{(5)}; Y_1^{(5)}) \right\},$$

(5)

Cut-set 1 (bounding rate of information transmitted by node 1): node 1 → relay, node 2.

Cut-set 2 (bounding rate of information received by node 2): node 1, relay → node 2.

Cut-set 3 (bounding rate of information transmitted by node 2): node 2 → relay, node 1.

Cut-set 4 (bounding rate of information received by node 1): node 2, relay → node 1.

Figure 2 Cut-sets that determine the outer bound region $\mathcal{C}_{\text{OB}}$. (a) Cut-set 1 (bounding rate of information transmitted by node 1): node 1 → relay, node 2. **(b)** Cut-set 2 (bounding rate of information received by node 2): node 1, relay → node 2. **(c)** Cut-set 3 (bounding rate of information transmitted by node 2): node 2 → relay, node 1. **(d)** Cut-set 4 (bounding rate of information received by node 1): node 2, relay → node 1.

(Figure 2d) cut-sets. After having identified which of the six protocol phases need to be considered for which cut-set, e.g., phases 1, 3, 6 for the first one, straightforward application of ([5], Thm. 1) gives the constraints specified in (5). □

We remark that the order of the phases in the transmission protocol is irrelevant if we only consider this outer bound region $\mathcal{C}_{\text{OB}}$; only the portion of the time τ_i that phase i is used matters. While it is clear that the optimal joint input distribution factors as $\prod_{i=1}^{6} p_{X_1^{(i)} X_2^{(i)} X_R^{(i)}}$, the following proposition additionally shows that $p_{X_1^{(3)} X_2^{(3)}} = p_{X_1^{(3)}} p_{X_2^{(3)}}$ maximizes $\mathcal{C}_{\text{OB}}$ for the Gaussian relay channel.

Proposition 2. *The input distribution for phase 3 that maximizes $\mathcal{C}_{\text{OB}}$ for the half-duplex Gaussian two-way relay channel factors as $p_{X_1^{(3)} X_2^{(3)}} = p_{X_1^{(3)}} p_{X_2^{(3)}}$.*

Proof. In the third phase, both node 1 and node 2 transmit to the relay so that the input-output characteristic for the Gaussian relay channel is generally specified by $Y_R^{(3)} = f_1(X_1^{(3)}) + f_2(X_2^{(3)}) + N_R^{(3)}$, where f_1 and f_2 are deterministic functions that represent the transformations of the input signals induced by the channel gains, and where $N_R^{(3)}$ denotes the Gaussian noise received at the relay, which is independent of the signals $X_1^{(3)}$ and $X_2^{(3)}$. There are two mutual information terms associated with phase 3 in (5): $I(X_1^{(3)}; Y_R^{(3)} | X_2^{(3)})$ in the first condition and $I(X_2^{(3)}; Y_R^{(3)} | X_1^{(3)})$ in the third one. Both of them are

maximized if $X_1^{(3)}$ and $X_2^{(3)}$ are independent, as shown by the following chain of inequalities:

$$\begin{aligned} I(X_1^{(3)}; Y_R^{(3)} | X_2^{(3)}) &= h(Y_R^{(3)} | X_2^{(3)}) - h(Y_R^{(3)} | X_1^{(3)} X_2^{(3)}) \\ &= h(f_1(X_1^{(3)}) + N_R^{(3)} | X_2^{(3)}) - h(N_R^{(3)}) \\ &\leq h(f_1(X_1^{(3)}) + N_R^{(3)}) - h(N_R^{(3)}), \end{aligned}$$

with equality if and only if $X_1^{(3)}$ and $X_2^{(3)}$ are independent ([25], Cor. to Thm. 8.6.1). The same of course holds if the roles of $X_1^{(3)}$ and $X_2^{(3)}$ are reversed, which proves the proposition. □

Proposition 2 hence implies that $\mathcal{C}_{\text{OB}}$ is also the cut-set outer bound for the restricted half-duplex Gaussian two-way relay channel, which requires that $p_{X_1^{(3)} X_2^{(3)}} = p_{X_1^{(3)}} p_{X_2^{(3)}}$ as the terminals must not cooperate in encoding their messages. Moreover, it can be shown that Gaussian inputs are optimal for each phase ([26], Prop. 2). Since a Gaussian distribution is completely determined by its mean and covariance, the optimal zero mean input for phase i is specified by $\mathbf{R}^{(i)}$, where $\mathbf{R}_{12}^{(3)} = \mathbf{0}_{N_1 \times N_2}$ holds for the optimal $\mathbf{R}^{(3)}$ as a consequence of Proposition 2. Note also that the cut-set bound was derived under the assumption of average transmit power constraints on every node, i.e., $\sum_{i=1}^{6} \tau_i E\left[\mathbf{x}_A^{(i),\text{H}} \mathbf{x}_A^{(i)}\right] = \sum_{i=1}^{6} \tau_i \text{tr}(\mathbf{R}_A^{(i)}) \leq P_A$ if P_A denotes the available transmit power node A may consume on average.

Now, let us turn to the main subject of this section and the entire article, which is to evaluate the outer bound region $\mathcal{C}_{\text{OB}}$ for the Gaussian MIMO relay channel.

One way of achieving this, and the one we choose here, is to determine its boundary by solving weighted sum rate (WSR) maximization problems over $\mathcal{C}_{OB}$ for different weight vectors $\boldsymbol{w} \in \mathbb{R}_+^2$. In particular, the boundary of $\mathcal{C}_{OB}$ can be determined with arbitrary precision by varying the ratio of the weights $\frac{w_1}{w_2}$ from zero to infinity[b]. For a given weight vector, the weighted sum rate maximization we then need to solve reads as

$$\max_{\boldsymbol{r}} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{r} \in \mathcal{C}_{OB}. \qquad (6)$$

We remark that the maximum of problem (6) is well-defined and that a maximizer $\boldsymbol{r}^{\star} \in \mathcal{C}_{OB}$ exists. This is because $\mathcal{C}_{OB}$ is closed and bounded (and thus compact) if the transmit powers P_1, P_2, and P_R the nodes may consume on average are finite, which we of course assume below. Hence, Weierstrass' theorem ([27], Thm. 2.3.1) guarantees that problem (6) attains its maximum.

For the purpose of solving such a WSR maximization problem, we take an approach that is similar to that chosen in [21] and which can be summarized as follows. Since the formulation of (6) is not very convenient if we actually want to perform the optimization, we seek a parameterization that is more suitable to the problem. As a first step towards this end, we find a convex parameterization of the outer bound region $\mathcal{C}_{OB}$ in Section 3.1. Since the objective function is linear, we obtain a convex optimization problem for which strong duality holds so that it can equivalently be solved in the dual domain. The corresponding dual problem is derived in Section 3.2. We then choose to solve this dual problem by means of the cutting plane algorithm, which is discussed in Section 3.3. Finally, we need to recover the optimal primal solution from the optimal solution to the dual problem. How this so-called primal reconstruction works for the considered weighted sum rate maximization problem is explained in Section 3.4.

3.1 Convex parameterization of outer bound region $\mathcal{C}_{OB}$

As a first step towards a convex parameterization of the outer bound region $\mathcal{C}_{OB}$, we define six rate-power regions $\mathcal{S}_1, \ldots, \mathcal{S}_6$, one for each phase of the transmission protocol. Basically, $\mathcal{S}_i$ specifies the contribution of protocol phase i to the outer bound region, both in terms of rates and power consumption. For the Gaussian MIMO relay channel with the optimal Gaussian inputs, the mutual information terms specifying the rates boil down to the well-known log-det expressions. Consequently, the six rate-power regions are given by[c]

$$\mathcal{S}_1 = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \leq \log \det\big(\mathbf{I}_{N_2+N_R} + \boldsymbol{H}_1 \boldsymbol{R}^{(1)} \boldsymbol{H}_1^{\mathrm{H}}\big),$$
$$r_2 \leq \log \det\big(\mathbf{I}_{N_2} + \boldsymbol{H}_{12} \boldsymbol{R}^{(1)} \boldsymbol{H}_{12}^{\mathrm{H}}\big),$$
$$p_1 = \mathrm{tr}\big(\boldsymbol{R}^{(1)}\big), p_2 = 0, p_3 = 0,$$
$$\boldsymbol{R}^{(1)} \succeq \boldsymbol{0} \Big\}, \qquad (7)$$

$$\mathcal{S}_2 = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \leq \log \det\big(\mathbf{I}_{N_1+N_R} + \boldsymbol{H}_2 \boldsymbol{R}^{(2)} \boldsymbol{H}_2^{\mathrm{H}}\big),$$
$$r_2 \leq \log \det\big(\mathbf{I}_{N_1} + \boldsymbol{H}_{21} \boldsymbol{R}^{(2)} \boldsymbol{H}_{21}^{\mathrm{H}}\big),$$
$$p_1 = 0, p_2 = \mathrm{tr}\big(\boldsymbol{R}^{(2)}\big), p_3 = 0,$$
$$\boldsymbol{R}^{(2)} \succeq \boldsymbol{0} \Big\}, \qquad (8)$$

$$\mathcal{S}_3 = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \leq \log \det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{1R} \boldsymbol{R}_1^{(3)} \boldsymbol{H}_{1R}^{\mathrm{H}}\big),$$
$$r_2 \leq \log \det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{2R} \boldsymbol{R}_2^{(3)} \boldsymbol{H}_{2R}^{\mathrm{H}}\big),$$
$$p_1 = \mathrm{tr}\big(\boldsymbol{R}_1^{(3)}\big), p_2 = \mathrm{tr}\big(\boldsymbol{R}_2^{(3)}\big), p_3 = 0,$$
$$\boldsymbol{R}_1^{(3)} \succeq \boldsymbol{0}, \boldsymbol{R}_2^{(3)} \succeq \boldsymbol{0} \Big\}, \qquad (9)$$

$$\mathcal{S}_4 = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \leq \log \det\big(\mathbf{I}_{N_2} + \boldsymbol{H}_{R2} \boldsymbol{R}^{(4)} \boldsymbol{H}_{R2}^{\mathrm{H}}\big),$$
$$r_2 \leq \log \det\big(\mathbf{I}_{N_1} + \boldsymbol{H}_{R1} \boldsymbol{R}^{(4)} \boldsymbol{H}_{R1}^{\mathrm{H}}\big),$$
$$p_1 = 0, p_2 = 0, p_3 = \mathrm{tr}\big(\boldsymbol{R}^{(4)}\big),$$
$$\boldsymbol{R}^{(4)} \succeq \boldsymbol{0} \Big\}, \qquad (10)$$

$$\mathcal{S}_5 = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \leq \log \det\big(\mathbf{I}_{N_1} + \boldsymbol{H}_{21} \boldsymbol{Q}^{(5)} \boldsymbol{H}_{21}^{\mathrm{H}}\big),$$
$$r_2 \leq \log \det\big(\mathbf{I}_{N_1} + \boldsymbol{H}_5 \boldsymbol{R}^{(5)} \boldsymbol{H}_5^{\mathrm{H}}\big),$$
$$p_1 = 0, p_2 = \mathrm{tr}\big(\boldsymbol{D}_2^{(5)} \boldsymbol{R}^{(5)} \boldsymbol{D}_2^{(5),\mathrm{H}}\big),$$
$$p_3 = \mathrm{tr}\big(\boldsymbol{D}_R^{(5)} \boldsymbol{R}^{(5)} \boldsymbol{D}_R^{(5),\mathrm{H}}\big),$$
$$\boldsymbol{Q}^{(5)} \succeq \boldsymbol{0}, \boldsymbol{R}^{(5)} - \boldsymbol{D}_2^{(5),\mathrm{H}} \boldsymbol{Q}^{(5)} \boldsymbol{D}_2^{(5)} \succeq \boldsymbol{0} \Big\} \qquad (11)$$

$$\mathcal{S}_6 = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \leq \log \det\big(\mathbf{I}_{N_2} + \boldsymbol{H}_{12} \boldsymbol{Q}^{(6)} \boldsymbol{H}_{12}^{\mathrm{H}}\big),$$
$$r_2 \leq \log \det\big(\mathbf{I}_{N_2} + \boldsymbol{H}_6 \boldsymbol{R}^{(6)} \boldsymbol{H}_6^{\mathrm{H}}\big),$$
$$p_1 = \mathrm{tr}\big(\boldsymbol{D}_1^{(6)} \boldsymbol{R}^{(6)} \boldsymbol{D}_1^{(6),\mathrm{H}}\big), p_2 = 0,$$
$$p_3 = \mathrm{tr}\big(\boldsymbol{D}_R^{(6)} \boldsymbol{R}^{(6)} \boldsymbol{D}_R^{(6),\mathrm{H}}\big),$$
$$\boldsymbol{Q}^{(6)} \succeq \boldsymbol{0}, \boldsymbol{R}^{(6)} - \boldsymbol{D}_1^{(6),\mathrm{H}} \boldsymbol{Q}^{(6)} \boldsymbol{D}_1^{(6)} \succeq \boldsymbol{0} \Big\}, \qquad (12)$$

with $\boldsymbol{H}_1 = \big[\boldsymbol{H}_{1R}^{\mathrm{H}} \ \boldsymbol{H}_{12}^{\mathrm{H}}\big]^{\mathrm{H}}$, $\boldsymbol{H}_2 = \big[\boldsymbol{H}_{2R}^{\mathrm{H}} \ \boldsymbol{H}_{21}^{\mathrm{H}}\big]^{\mathrm{H}}$, $\boldsymbol{H}_5 = \big[\boldsymbol{H}_{21} \ \boldsymbol{H}_{R1}\big]$, $\boldsymbol{H}_6 = \big[\boldsymbol{H}_{12} \ \boldsymbol{H}_{R2}\big]$ and $\boldsymbol{D}_2^{(5)}, \boldsymbol{D}_R^{(5)}, \boldsymbol{D}_1^{(6)}, \boldsymbol{D}_R^{(6)}$ being appropriate selection matrices as defined in (3). It is straightforward to verify that $\mathcal{S}_1, \ldots, \mathcal{S}_6$ are convex sets which are parameterized by means of the (joint) transmit covariance matrices $\boldsymbol{R}^{(1)}, \ldots, \boldsymbol{R}^{(6)}$, respectively. They are not compact, however, because neither the rates nor the transmit powers are bounded above. In fact, this is the main difference to the problem considered in [21], where the average transmit powers for each phase and thus also

the rate regions associated with each phase are bounded. As a result, the derivation of the dual problem and its solution by means of the cutting plane algorithm become considerably more difficult, as discussed in Sections 3.2 and 3.3.

Remark 1. In order to arrive at above formulations for $\mathcal{S}_5$ and $\mathcal{S}_6$, the corresponding constraints on r_1 have to be reformulated. This is done by introducing the auxiliary variables $Q^{(5)}$ and $Q^{(6)}$ to relax the equality constraints on the conditional covariance matrices $R_{2|R}^{(5)} = R_2^{(5)} - R_{2R}^{(5)} R_R^{(5),\dagger} R_{2R}^{(5),H}$ and $R_{1|R}^{(6)} = R_1^{(6)} - R_{1R}^{(6)} R_R^{(6),\dagger} R_{1R}^{(6),H}$, respectively, before applying the (generalized) Schur complement condition. For more details, we refer the reader to [10], where this reformulation was first presented assuming that $R_R^{(5),-1}$ exists, or to [11], where the same result was later independently derived for the more general case when $R_R^{(5)}$ need not have full rank.

Suppose that P_1, P_2, and P_R denote the finite transmit powers that terminal 1, terminal 2, and the relay may consume on average, respectively, and let $\boldsymbol{p}_{\text{Tx}} = \begin{bmatrix} P_1 & P_2 & P_R \end{bmatrix}^{\text{T}}$. Having defined the six rate-power regions and the vector $\boldsymbol{p}_{\text{Tx}}$, we can now rewrite problem (6) as follows:

$$\max_{\boldsymbol{r},\tau_i,\boldsymbol{r}_i,\boldsymbol{p}_i} \boldsymbol{w}^{\text{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \sum_{i=1}^{6} \tau_i \boldsymbol{B}_i \boldsymbol{r}_i, \sum_{i=1}^{6} \tau_i \boldsymbol{p}_i \leq \boldsymbol{p}_{\text{Tx}}, \sum_{i=1}^{6} \tau_i = 1,$$
$$\tau_i \geq 0, (\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i, \forall i \in \{1, \dots, 6\}.$$
$$(13)$$

Like in [21], each row of $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}^{\text{T}}$ selects one of the four rate constraints as defined in the outer bound region $\mathcal{C}_{\text{OB}}$ (cf. (5)), and the corresponding rows of the matrices $\boldsymbol{B}_i \in \{0,1\}^{4 \times 2}$ specify the structures of these constraints with regard to the sets $\mathcal{S}_i$: $\boldsymbol{B}_1 = \boldsymbol{B}_6 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}^{\text{T}}$, $\boldsymbol{B}_2 = \boldsymbol{B}_5 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{\text{T}}$, $\boldsymbol{B}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}^{\text{T}}$, $\boldsymbol{B}_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{\text{T}}$. Furthermore, the fact that the three nodes are subject to average transmit power constraints is reflected in the term $\sum_{i=1}^{6} \tau_i \boldsymbol{p}_i \leq \boldsymbol{p}_{\text{Tx}}$, where $\boldsymbol{p}_i = \begin{bmatrix} p_1^{(i)} & p_2^{(i)} & p_R^{(i)} \end{bmatrix}^{\text{T}}$ is the vector of average transmit powers consumed by the three nodes during phase i.

Remark 2. The optimization problem (13) would be convex for fixed $\tau_1, \dots, \tau_6$. The reason it is a nonconvex parameterization of (6) if the time shares are optimization variables is that the functions $\tau_i \boldsymbol{B}_i \boldsymbol{r}_i$ and $\tau_i \boldsymbol{p}_i$ are not jointly concave in $\tau_i, \boldsymbol{r}_i$ and jointly convex in $\tau_i, \boldsymbol{p}_i$, respectively.

Consequently, another reformulation step is required, and for this purpose, we define the set

$$\mathcal{S} = \left\{ (\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}_+^4 \times \mathbb{R}_+^3 : \boldsymbol{y} = \sum_{i=1}^{6} \tau_i \boldsymbol{B}_i \boldsymbol{r}_i, \boldsymbol{z} = \sum_{i=1}^{6} \tau_i \boldsymbol{p}_i, \sum_{i=1}^{6} \tau_i = 1, \right.$$
$$\left. \tau_i \geq 0, (\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i, \forall i \in \{1, \dots, 6\} \right\}.$$
$$(14)$$

Proposition 3. $\mathcal{S}$ *is a convex set.*

Proof. See Appendix 1. □

Using this definition of $\mathcal{S}$, the weighted sum rate maximization problem (6) is equivalently expressed as

$$\max_{\boldsymbol{r},\boldsymbol{y},\boldsymbol{z}} \boldsymbol{w}^{\text{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{y}, \boldsymbol{z} \leq \boldsymbol{p}_{\text{Tx}}, (\boldsymbol{y},\boldsymbol{z}) \in \mathcal{S}. \quad (15)$$

Because $\mathcal{S}$ is a convex set with nonempty relative interior, (15) is a convex optimization problem for which strong duality holds ([28], Sec. 5.3.2). In particular, the constraints of problem (15) specify a convex set, which means that a convex parameterization of the outer bound region $\mathcal{C}_{\text{OB}}$ is given by

$$\mathcal{C}_{\text{OB}} = \left\{ \boldsymbol{r} \in \mathbb{R}_+^2 : \boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{y}, \boldsymbol{z} \leq \boldsymbol{p}_{\text{Tx}}, (\boldsymbol{y},\boldsymbol{z}) \in \mathcal{S} \right\}. \quad (16)$$

3.2 Derivation of the dual function

Since we have strong duality for problem (15), we can equivalently solve it in the dual domain. In the approach considered here, the constraints $\boldsymbol{A}\boldsymbol{r} \leq \boldsymbol{y}$ and $\boldsymbol{z} \leq \boldsymbol{p}_{\text{Tx}}$ are incorporated into the objective function using the Lagrangian multipliers $\boldsymbol{\lambda} \in \mathbb{R}^4$ and $\boldsymbol{\mu} \in \mathbb{R}^3$. This leads to a dual problem where the six phases are decoupled. In particular, it will show that this approach allows to solve (15) without explicitly optimizing the time allocation parameters $\tau_1, \dots \tau_6$. The Lagrangian function reads as

$$L(\boldsymbol{r},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{w}^{\text{T}}\boldsymbol{r} - \boldsymbol{\lambda}^{\text{T}}(\boldsymbol{A}\boldsymbol{r} - \boldsymbol{y}) - \boldsymbol{\mu}^{\text{T}}(\boldsymbol{z} - \boldsymbol{p}_{\text{Tx}}),$$
$$(17)$$

and the resulting dual function is given by

$$\Theta(\boldsymbol{\lambda},\boldsymbol{\mu}) = \sup_{\boldsymbol{r},(\boldsymbol{y},\boldsymbol{z})\in\mathcal{S}} L(\boldsymbol{r},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\lambda},\boldsymbol{\mu})$$
$$= \begin{cases} \boldsymbol{\mu}^{\text{T}}\boldsymbol{p}_{\text{Tx}} + \sup_{(\boldsymbol{y},\boldsymbol{z})\in\mathcal{S}} \left\{ \boldsymbol{\lambda}^{\text{T}}\boldsymbol{y} - \boldsymbol{\mu}^{\text{T}}\boldsymbol{z} \right\} & \text{if } \boldsymbol{A}^{\text{T}}\boldsymbol{\lambda} = \boldsymbol{w}, \\ +\infty & \text{otherwise.} \end{cases}$$
$$(18)$$

Applying the definition of $\mathcal{S}$ yields

$$\sup_{(\boldsymbol{y},\boldsymbol{z})\in\mathcal{S}}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{y}-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{z}\right\}=\max_{i=1,\ldots,6}\left(\sup_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i\right\}\right).$$

(19)

If none of the channel gain matrices between the two terminals or between one of the terminals and the relay is a zero matrix, we have the following proposition[d].

Proposition 4. *For any $\boldsymbol{\mu} \geq \boldsymbol{0}$ and any $\boldsymbol{\lambda} \geq \boldsymbol{0}$ that satisfies $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{w}$, the value of the dual function $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is finite if and only if the following three conditions hold:*

1. $\mu_1 > 0$ *or* $\mu_1 = 0, \lambda_1 = \lambda_2 = 0$,
2. $\mu_2 > 0$ *or* $\mu_2 = 0, \lambda_3 = \lambda_4 = 0$,
3. $\mu_3 > 0$ *or* $\mu_3 = 0, \lambda_2 = \lambda_4 = 0$.

Proof. See Appendix 2. □

The meaning of Proposition 4 is as follows. For the subproblems

$$\sup_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{T}\boldsymbol{p}_i\right\}, \quad i \in \{1,\ldots,6\}, \tag{20}$$

the Lagrangian multipliers μ_1, μ_2, and μ_3 can be understood as prices associated with the powers p_1, p_2, and p_{R} consumed by node 1, node 2, and the relay, respectively. If all prices are positive, each of the subproblems is guaranteed to have a finite optimal solution because the cost of power $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i$ is a linear function of $p_1^{(i)}$, $p_2^{(i)}$, $p_{\mathrm{R}}^{(i)}$, whereas $\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i$ increases only logarithmically with the powers. If one of the prices is zero, however, the transmit power of the corresponding node and the associated transmit data rates can be increased to infinity without incurring any costs. Consequently, the subproblems for all phases $i \in \{1,\ldots,6\}$ in which this node transmits take the value infinity unless all the entries of the $\boldsymbol{r}_i$'s to which transmissions by the node contribute are weighted with zero.

Remark 3. Note that $\lambda_3 = \lambda_4 = 0$ $(\lambda_1 = \lambda_2 = 0)$ may result in $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$ only if $w_2 = 0$ $(w_1 = 0)$ because otherwise $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} \neq \boldsymbol{w}$. If $w_2 = 0$ $(w_1 = 0)$, however, the WSR maximization over $\mathcal{C}_{\mathrm{OB}}$ (6) reduces to maximizing the cut-set bound for the one-way relay channel with terminal 1 (terminal 2) being the source and terminal 2

(terminal 1) being the destination. In particular, $w_2 = 0$ yields $\lambda_3^{\star} = \lambda_4^{\star} = 0$, which in turn implies $\mu_2^{\star} = 0$ and

$$\max_{i=1,\ldots,6}\max_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i\right\}$$

$$=\max_{i=1,6}\max_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i\right\} \tag{21}$$

That is, only phases 1 and 6 of our 6-phase protocol need to be considered for the optimal solution, which is equivalent to setting $\tau_2 = \tau_3 = \tau_4 = \tau_5 = 0$ in (13). The de facto communication protocol for this case is therefore consistent with that used for the half-duplex one-way relay channel if terminal 1 is the source and terminal 2 is the destination [12,21]. Similarly, $w_1 = 0$ implies $\lambda_1^{\star} = \lambda_2^{\star} = 0$, $\mu_1^{\star} = 0$, and the optimal solution involves only phases 2 and 5. If $\boldsymbol{w} > \boldsymbol{0}$, on the other hand, we can conclude from Proposition 4 that $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$ requires $\mu_1 > 0$ and $\mu_2 > 0$.

Remark 4. For $\mu_3 = 0$, it follows from Proposition 4 that $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$ only if $\lambda_2 = \lambda_4 = 0$. But $\lambda_2 = \lambda_4 = 0$ means that transmissions by the relay have no effect on the dual function since the corresponding rates are all weighted with zero. This is independent of P_{R} and the channel gain matrices, and as a result, we have

$$\max_{i=1,\ldots,6}\max_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i\right\}$$

$$=\max_{i=1,2,3}\max_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i\right\} \tag{22}$$

in this case. Moreover, it is clear that phase 3 contributes nothing to the bidirectional communication if the relay cannot forward the information it previously received. Hence, the optimal solution could only involve phases 1 and 2 if $\mu_3 = 0$, meaning that only the direct link between the terminals would be utilized, and $\boldsymbol{\lambda} = \begin{bmatrix} w_1 & 0 & w_2 & 0 \end{bmatrix}^{\mathrm{T}}$ would be the optimizer of the dual problem. But for this $\boldsymbol{\lambda}$ the primal feasibility and complementary slackness conditions of the primal problem (15) would only be satisfied simultaneously if $X_2^{(2)} - Y_1^{(2)} - Y_{\mathrm{R}}^{(2)}$ and $X_1^{(1)} - Y_2^{(1)} - Y_{\mathrm{R}}^{(1)}$ formed Markov chains. This is an academic special case that our system model does not permit. Consequently, $\mu_3 > 0$ if $\boldsymbol{\lambda} \geq \boldsymbol{0}$ and $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$.

From Proposition 4 and the two subsequent remarks, it follows that for $\boldsymbol{\lambda} \geq \boldsymbol{0}$ and positive weight vectors $\boldsymbol{w} > \boldsymbol{0}$ the dual function is equal to

$$\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{cases} \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_{\mathrm{Tx}} + \max_{i=1,\ldots,6}\left(\max_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i}\left\{\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{r}_i-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{p}_i\right\}\right) & \text{if } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{w},\ \boldsymbol{\mu} > \boldsymbol{0}, \\ +\infty & \text{otherwise.} \end{cases} \tag{23}$$

In order to determine an optimal solution to the original weighted sum rate maximization problem (6), we thus have to solve the dual problem

$$\min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{p}_{\mathrm{Tx}} + \max_{i=1,\dots,6} \max_{(\boldsymbol{r}_i,\boldsymbol{p}_i)\in\mathcal{S}_i} \left\{ \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{B}_i \boldsymbol{r}_i - \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{p}_i \right\}$$

$$\text{s.t.} \quad \boldsymbol{\lambda} \geq \boldsymbol{0},\, \boldsymbol{A}^{\mathrm{T}} \boldsymbol{\lambda} = \boldsymbol{w},\, \boldsymbol{\mu} > \boldsymbol{0}. \tag{24}$$

Remark 5. Because $\boldsymbol{\mu} > \boldsymbol{0}$, the constraint set of this dual problem is not closed so that the existence of a minimizing solution cannot be guaranteed by Weierstrass' theorem. However, since the maximum of (6) is well-defined and strong duality holds, the minimum of (24) is also well-defined.

3.3 Solution by means of cutting plane algorithm

A simple yet efficient algorithm that can be used to solve the dual problem (24) is the cutting plane algorithm ([27], Sec. 6.4), an outer-approximation method where the feasible set of the problem is approximated by a finite number of feasible points and iteratively refined by a set of linear inequalities. In each iteration of the cutting plane algorithm, a linear program, the so-called *master program*, must be solved and the dual function $\Theta(\boldsymbol{\lambda},\boldsymbol{\mu})$ must be evaluated. In the ℓth iteration, the master program reads as

$$\min_{\alpha,\boldsymbol{\lambda},\boldsymbol{\mu}} \alpha \ \text{s.t.}\ \alpha \geq \boldsymbol{\mu}^{\mathrm{T}} \left(\boldsymbol{p}_{\mathrm{Tx}} - \boldsymbol{p}^{(k)} \right) + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{v}^{(k)}, \forall k \in \{1,\dots,\ell\},$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0},\, \boldsymbol{A}^{\mathrm{T}} \boldsymbol{\lambda} = \boldsymbol{w},\, \boldsymbol{\mu} > \boldsymbol{0}, \tag{25}$$

where, for all $k \in \{1,\dots,\ell\}$, we have $(\boldsymbol{r}^{(k)},\boldsymbol{p}^{(k)}) \in \mathcal{S}_i$ for some $i \in \{1,\dots,6\}$ and $\boldsymbol{v}^{(k)} = \boldsymbol{B}_i \boldsymbol{r}^{(k)}$. As can be seen from (23), evaluating the dual function requires to solve six independent convex optimization problems, one over each of the sets $\mathcal{S}_i$ associated with the six phases of the communication protocol. For this purpose, standard semidefinite program (SDP) solvers like SDPT3 [29] that are capable of dealing with the weighted sum of log-det terms in the objective function can be applied. For a convergence analysis and more details on the cutting plane method, we refer the reader to ([27], Sec. 6.4).

Remark 6. In order for the cutting plane algorithm to work in practice, we replace the constraint $\boldsymbol{\mu} > \boldsymbol{0}$ by $\boldsymbol{\mu} \geq \boldsymbol{0}$. This does not change the optimal solution of the dual problem (24) because we know that the optimizer satisfies $\boldsymbol{\mu}^{\star} > \boldsymbol{0}$. However, proper initialization of the cutting plane method then has to be ensured. In particular, if $\alpha^{(1)},\boldsymbol{\lambda}^{(1)},\boldsymbol{\mu}^{(1)}$ are the optimizers of the master program in the first iteration, we actually have to choose *several* initial points $(\boldsymbol{r}^{(k)},\boldsymbol{p}^{(k)}) \in \mathcal{S}_i$ for some $i \in \{1,\dots,6\}$ such that $\alpha^{(1)}$ is finite and $\boldsymbol{\mu}^{(1)} > \boldsymbol{0}$ is guaranteed. Otherwise, the algorithm runs into problems when the dual function is evaluated.

Remark 7. Since $\boldsymbol{\mu}^{\star} > \boldsymbol{0}$ for $\boldsymbol{w} > \boldsymbol{0}$, it follows from the complimentary slackness condition of the primal problem (15) that $\boldsymbol{z}^{\star} = \boldsymbol{p}_{\mathrm{Tx}}$, which means that the three nodes use all their available transmit power. This in turn implies $\boldsymbol{r}^{\star} > \boldsymbol{0}$ whenever $\boldsymbol{w} > \boldsymbol{0}$, i.e., the tangents to the boundary of $\mathcal{C}_{\mathrm{OB}}$ at the optimal unidirectional points $(C_{1,\max},0)$ and $(0,C_{2,\max})$ are orthogonal to the axes.

3.4 Primal reconstruction

As previously mentioned, the proposed dual decomposition approach allows to determine the optimal value of (6) without explicitly optimizing the time shares allocated to the six phases of the communication protocol. On the one hand, the decoupling of the phases considerably simplifies the optimization, but on the other, we want to know the optimal rate vector $\boldsymbol{r}^{\star}$ and possibly the optimal time shares $\tau_i^{\star}$, e.g., for the purpose of designing resource allocation protocols. To this end, we need to generate the optimal primal solution from the optimal solution to the dual problem, a process that is generally referred to as *primal reconstruction* or *primal recovery*. Since we apply the cutting plane algorithm to solve the dual problem, the primal recovery scheme to obtain the optimal rate vector $\boldsymbol{r}^{\star}$ and the optimal time shares $\tau_i^{\star}$ is fairly simple. Assume the cutting plane algorithm has converged to the optimal solution of the dual problem after L iterations, and consider the dual problem of the corresponding master program (25) given by

$$\max_{\boldsymbol{x},\boldsymbol{u}} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} \leq \sum_{k=1}^{L} u_k \boldsymbol{v}^{(k)},\, \sum_{k=1}^{L} u_k \boldsymbol{p}^{(k)} \leq \boldsymbol{p}_{\mathrm{Tx}},$$

$$\sum_{k=1}^{L} u_k = 1,\, u_k \geq 0, \forall k \in \{1,\dots,L\}. \tag{26}$$

We remark that this problem is an approximation of the primal problem (15) where the set $\mathcal{S}$ is replaced by a convex combination of feasible points $\left\{ (\boldsymbol{v}^{(1)},\boldsymbol{p}^{(1)}),\dots,(\boldsymbol{v}^{(L)},\boldsymbol{p}^{(L)}) \right\} \subset \mathcal{S}$ and where $\boldsymbol{x} \in \mathbb{R}^2$ and u_k denote the Lagrangian multipliers associated with the constraints $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{\lambda} = \boldsymbol{w}$ and $\alpha \geq \boldsymbol{\mu}^{\mathrm{T}} \left(\boldsymbol{p}_{\mathrm{Tx}} - \boldsymbol{p}^{(k)} \right) + \boldsymbol{\lambda}^{T} \boldsymbol{v}^{(k)}$ of the master program, respectively. Letting $\mathcal{K}_i = \left\{ k : (\boldsymbol{r}^{(k)},\boldsymbol{p}^{(k)}) \in \mathcal{S}_i,\, \boldsymbol{v}^{(k)} = \boldsymbol{B}_i \boldsymbol{r}^{(k)} \right\}^{\mathrm{e}}$, we can rewrite (26) as

$$\max_{\boldsymbol{x},\boldsymbol{u}} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} \leq \sum_{i=1}^{6} \sum_{k\in\mathcal{K}_i} u_k \boldsymbol{B}_i \boldsymbol{r}^{(k)},\, \sum_{i=1}^{6} \sum_{k\in\mathcal{K}_i} u_k \boldsymbol{p}^{(k)} \leq \boldsymbol{p}_{\mathrm{Tx}},$$

$$\sum_{k=1}^{L} u_k = 1,\, u_k \geq 0, \forall k \in \{1,\dots,L\}. \tag{27}$$

Furthermore, it can be shown that $\sum_{k\in\mathcal{K}_i} u_k \boldsymbol{B}_i \boldsymbol{r}^{(k)} = \left(\sum_{k\in\mathcal{K}_i} u_k \right) \boldsymbol{B}_i \tilde{\boldsymbol{r}}_i$ and $\sum_{k\in\mathcal{K}_i} u_k \boldsymbol{p}^{(k)} = \left(\sum_{k\in\mathcal{K}_i} u_k \right) \tilde{\boldsymbol{p}}_i$ for

some $(\tilde{r}_i, \tilde{p}_i) \in \mathcal{S}_i$ since $\mathcal{S}_i$ is a convex set for all $i \in \{1, \dots, 6\}$. If we insert these expressions in (26) and compare the result to (13), we can conclude that

$$\tau_i^\star = \sum_{k \in \mathcal{K}_i} u_k^\star, \ \forall i \in \{1, \dots, 6\}. \tag{28}$$

The optimal time shares $\tau_i^\star$ are therefore easily obtained from the optimal Lagrangian dual variables $u_k^\star$, $k \in \{1, \dots, L\}$, that correspond to the constraints $\alpha \geq \boldsymbol{\mu}^{\mathrm{T}}\left(\boldsymbol{p}_{\mathrm{Tx}} - \boldsymbol{p}^{(k)}\right) + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\nu}^{(k)}$ in the master program. Moreover, it is clear that $\boldsymbol{x}^\star$, which denotes the vector of optimal dual variables corresponding to the equality constraints $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{w}$, yields the optimal rate vector $\boldsymbol{r}^\star$.

Remark 8. Since all $\mathcal{S}_i$ are convex, time sharing *within* any of the six phases of the communication protocol is not necessary. As a result, there will be no more than one $k \in \mathcal{K}_i$ with $u_k^\star > 0$ for every $i \in \{1, \dots, 6\}$.

4 An achievable rate region using the DF scheme
To obtain an inner bound on the capacity region of the restricted half-duplex Gaussian MIMO two-way relay channel, we consider the rate region that is achievable with the decode-and-forward coding scheme in this section. Like the cut-set bound, the DF coding scheme is due to Cover and El Gamal [4]. Requiring the relay to decode the source message can be a severe constraint so that other relaying strategies like compress-and-forward or amplify-and-forward can achieve higher rates for certain channel conditions. For single-antenna nodes, this is for example illustrated in [8,26]. Nevertheless, we consider only the DF strategy in this article because the corresponding achievable rate region $\mathcal{R}_{\mathrm{DF}}$ is very similar in structure to $\mathcal{C}_{\mathrm{OB}}$ and can thus be evaluated using the same methodology as described in the previous section.

Theorem 5. *If the relay uses the decode-and-forward coding scheme, the following rate region is achievable for the restricted half-duplex two-way relay channel:*

Proof. This result is derived in [22] by adapting the DF coding scheme to the 6-phase communication protocol introduced in Section 2 (with the phases performed in exactly that order) and applying it to both directions of information transfer. A brief outline of the coding scheme that achieves $\mathcal{R}_{\mathrm{DF}}$ is given in Appendix 3. $\square$

In theory, a different ordering of the phases may increase the achievable rate region $\mathcal{R}_{\mathrm{DF}}$. To the best of our knowledge, however, the 6-phase protocol we use is the most general protocol for the half-duplex two-way relay channel that has been considered in the literature so far. In particular, it includes the 2-phase *multiple access broadcast protocol* (MABC: consisting of phases 3, 4), the 3-phase *time division broadcast protocol* (TDBC: 1, 2, 4), and the 4-phase *hybrid broadcast protocol* (HBC: 1, 2, 3, 4) used in [14-20], for example[f]. In addition, it also covers the approach of using time sharing between the one-way relay channels in both directions to exchange information between the terminals, which we termed *one-way time sharing* (OWTS: 1, 2, 5, 6) in [21].

Like for the outer bound region $\mathcal{C}_{\mathrm{OB}}$, the optimal joint input distribution factors as $\prod_{i=1}^{6} p_{X_1^{(i)} X_2^{(i)} X_{\mathrm{R}}^{(i)}}$, where $p_{X_1^{(3)} X_2^{(3)}} = p_{X_1^{(3)}} p_{X_2^{(3)}}$ must be fulfilled due to the assumption of the restricted half-duplex two-way relay channel, which prohibits the nodes from cooperating in encoding their messages. Furthermore, the optimal input distribution for each phase $i \in \{1, \dots, 6\}$ can be shown to be Gaussian again.

Note that, as in Section 3, our main objective is again to evaluate the achievable rate region $\mathcal{R}_{\mathrm{DF}}$ for the Gaussian MIMO relay channel. Clearly, the boundary of the achievable rate region $\mathcal{R}_{\mathrm{DF}}$ can also be determined by means of solving WSR maximization problems with different weight vectors. As $\mathcal{R}_{\mathrm{DF}}$ and $\mathcal{C}_{\mathrm{OB}}$ are very similar in structure, the approach we use to solve one such problem is essentially the same as for the outer bound region. First, we find a convex parameterization for $\mathcal{R}_{\mathrm{DF}}$. Subsequently,

$$
\begin{aligned}
\mathcal{R}_{\mathrm{DF}} = \bigcup_{\prod_{i=1}^{6} p_{X_1^{(i)} X_2^{(i)} X_{\mathrm{R}}^{(i)}}} & \Bigg\{ (R_1, R_2) \in \mathbb{R}_+^2 : \tau_i \geq 0, \forall i \in \{1, \dots, 6\}, \ \sum_{i=1}^{6} \tau_i = 1, \\
& R_1 \leq \tau_1 I(X_1^{(1)}; Y_{\mathrm{R}}^{(1)}) + \tau_3 I(X_1^{(3)}; Y_{\mathrm{R}}^{(3)} | X_2^{(3)}) + \tau_6 I(X_1^{(6)}; Y_2^{(6)} | X_{\mathrm{R}}^{(6)}), \\
& R_1 \leq \tau_1 I(X_1^{(1)}; Y_2^{(1)}) + \tau_4 I(X_{\mathrm{R}}^{(4)}; Y_2^{(4)}) + \tau_6 I(X_1^{(6)} X_{\mathrm{R}}^{(6)}; Y_2^{(6)}), \\
& R_2 \leq \tau_2 I(X_2^{(2)}; Y_{\mathrm{R}}^{(2)}) + \tau_3 I(X_2^{(3)}; Y_{\mathrm{R}}^{(3)} | X_1^{(3)}) + \tau_5 I(X_2^{(5)}; Y_1^{(5)} | X_{\mathrm{R}}^{(5)}), \\
& R_2 \leq \tau_2 I(X_2^{(2)}; Y_1^{(2)}) + \tau_4 I(X_{\mathrm{R}}^{(4)}; Y_1^{(4)}) + \tau_5 I(X_2^{(5)} X_{\mathrm{R}}^{(5)}; Y_1^{(5)}), \\
R_1 + R_2 \leq {} & \tau_1 I(X_1^{(1)}; Y_{\mathrm{R}}^{(1)}) + \tau_2 I(X_2^{(2)}; Y_{\mathrm{R}}^{(2)}) + \tau_3 I(X_1^{(3)} X_2^{(3)}; Y_{\mathrm{R}}^{(3)}) \\
& + \tau_5 I(X_2^{(5)}; Y_1^{(5)} | X_{\mathrm{R}}^{(5)}) + \tau_6 I(X_1^{(6)}; Y_2^{(6)} | X_{\mathrm{R}}^{(6)}) \Bigg\}.
\end{aligned}
\tag{29}
$$

we solve the problem in the dual domain by means of the cutting plane algorithm, and finally, we perform the primal reconstruction.

For the purpose of deriving a convex parameterization for $\mathcal{R}_{\text{DF}}$, let

$$\mathcal{S}_1' = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \le \log\det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{1R}\boldsymbol{R}^{(1)}\boldsymbol{H}_{1R}^{\text{H}}\big),$$
$$r_2 \le \log\det\big(\mathbf{I}_{N_2} + \boldsymbol{H}_{12}\boldsymbol{R}^{(1)}\boldsymbol{H}_{12}^{\text{H}}\big),$$
$$p_1 = \text{tr}\big(\boldsymbol{R}^{(1)}\big),\, p_2 = 0,\, p_3 = 0,$$
$$\boldsymbol{R}^{(1)} \succeq \mathbf{0} \Big\}, \tag{30}$$

$$\mathcal{S}_2' = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \le \log\det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{2R}\boldsymbol{R}^{(2)}\boldsymbol{H}_{2R}^{\text{H}}\big),$$
$$r_2 \le \log\det\big(\mathbf{I}_{N_1} + \boldsymbol{H}_{21}\boldsymbol{R}^{(2)}\boldsymbol{H}_{21}^{\text{H}}\big),$$
$$p_1 = 0,\, p_2 = \text{tr}\big(\boldsymbol{R}^{(2)}\big),\, p_3 = 0,$$
$$\boldsymbol{R}^{(2)} \succeq \mathbf{0} \Big\}, \tag{31}$$

$$\mathcal{S}_3' = \Big\{ (\boldsymbol{r}, \boldsymbol{p}) \in \mathbb{R}_+^2 \times \mathbb{R}_+^3 : r_1 \le \log\det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{1R}\boldsymbol{R}_1^{(3)}\boldsymbol{H}_{1R}^{\text{H}}\big),$$
$$r_2 \le \log\det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{2R}\boldsymbol{R}_2^{(3)}\boldsymbol{H}_{2R}^{\text{H}}\big),$$
$$r_1 + r_2 \le \log\det\big(\mathbf{I}_{N_R} + \boldsymbol{H}_{1R}\boldsymbol{R}_1^{(3)}\boldsymbol{H}_{1R}^{\text{H}}$$
$$+ \boldsymbol{H}_{2R}\boldsymbol{R}_2^{(3)}\boldsymbol{H}_{2R}^{\text{H}}\big),$$
$$p_1 = \text{tr}\big(\boldsymbol{R}_1^{(3)}\big),\, p_2 = \text{tr}\big(\boldsymbol{R}_2^{(3)}\big),\, p_3 = 0,$$
$$\boldsymbol{R}_1^{(3)} \succeq \mathbf{0},\, \boldsymbol{R}_2^{(3)} \succeq \mathbf{0} \Big\}, \tag{32}$$

and $\mathcal{S}_i' = \mathcal{S}_i$ for $i \in \{4, 5, 6\}$. Like $\mathcal{S}_i$ defined in the previous section, every $\mathcal{S}_i'$ is a convex set that is parameterized by means of the (joint) transmit covariance matrix $\boldsymbol{R}^{(i)}$ and that specifies the contribution of phase i to $\mathcal{R}_{\text{DF}}$. Having defined these unbounded convex sets $\mathcal{S}_i'$, we can now express the weighted sum rate maximization problem that yields a point on the boundary of $\mathcal{R}_{\text{DF}}$ as follows:

$$\max_{\boldsymbol{r},\tau_i,\boldsymbol{r}_i,\boldsymbol{p}_i} \boldsymbol{w}^{\text{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}'\boldsymbol{r} \le \sum_{i=1}^{6} \tau_i\boldsymbol{B}_i'\boldsymbol{r}_i, \sum_{i=1}^{6}\tau_i\boldsymbol{p}_i \le \boldsymbol{p}_{\text{Tx}}, \sum_{i=1}^{6}\tau_i = 1,$$
$$\tau_i \ge 0,\, (\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i',\, \forall i \in \{1, \dots, 6\}. \tag{33}$$

Observe that the main difference compared to (13) is the additional constraint on the sum rate $R_1 + R_2$ in $\mathcal{R}_{\text{DF}}$ so that $\boldsymbol{A}' = \left[\begin{smallmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{smallmatrix}\right]^{\text{T}}$ and $\boldsymbol{B}_1' = \boldsymbol{B}_6' = \left[\begin{smallmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{smallmatrix}\right]^{\text{T}}$, $\boldsymbol{B}_2' = \boldsymbol{B}_5' = \left[\begin{smallmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{smallmatrix}\right]^{\text{T}}$, $\boldsymbol{B}_3' = \left[\begin{smallmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{smallmatrix}\right]^{\text{T}}$, $\boldsymbol{B}_4' = \left[\begin{smallmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{smallmatrix}\right]^{\text{T}}$. This constraint comes from the third phase of the communication protocol, a multiple access phase where both terminals transmit to the relay. The sum rate constraint in $\mathcal{R}_{\text{DF}}$ occurs because the relay must decode the messages from node 1 and node 2 when it uses DF.

Since all results from Section 3 apply here accordingly[g], the remaining steps of the optimization follow along the same lines as for the outer bound region $\mathcal{C}_{\text{OB}}$. First, we define the convex set

$$\mathcal{S}' = \Big\{ (\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}_+^5 \times \mathbb{R}_+^3 :\, \boldsymbol{y} = \sum_{i=1}^{6}\tau_i\boldsymbol{B}_i'\boldsymbol{r}_i,\, \boldsymbol{z} = \sum_{i=1}^{6}\tau_i\boldsymbol{p}_i,$$
$$\sum_{i=1}^{6}\tau_i = 1, \tau_i \ge 0,\, (\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i',\, \forall i \in \{1, \dots, 6\} \Big\} \tag{34}$$

and reformulate (33) as

$$\max_{\boldsymbol{r},\boldsymbol{y},\boldsymbol{z}} \boldsymbol{w}^{\text{T}}\boldsymbol{r} \quad \text{s.t.} \quad \boldsymbol{A}'\boldsymbol{r} \le \boldsymbol{y}, \boldsymbol{z} \le \boldsymbol{p}_{\text{Tx}}, (\boldsymbol{y}, \boldsymbol{z}) \in \mathcal{S}'. \tag{35}$$

Then, we use the dual decomposition approach in combination with the cutting plane method to obtain an optimal solution to this convex optimization problem.

Note that, after having obtained the solution, the optimal time shares $\tau_i^\star$, $i \in \{1, \dots, 6\}$, i.e., the optimal durations of the six protocol phases, tell us which of these phases are part of the optimal transmission protocol for a given weight vector $\boldsymbol{w}$. In particular, the optimal protocol includes phase i if and only if $\tau_i^\star > 0$. Furthermore, our dual decomposition approach cannot only be applied to WSR maximization problems, but to any convex optimization problem for which strong duality holds. As a result, this approach may be used for the design of resource allocation protocols, e.g., by considering utility maximization problems with concave utility functions.

5 Numerical results

In this section, numerical results yielding bounds on the capacity of the half-duplex Gaussian relay channel as well as numerical results giving bounds on the capacity region of the restricted half-duplex Gaussian two-relay channel are presented. More specifically, we evaluate and compare the outer bound region $\mathcal{C}_{\text{OB}}$ and the rate region $\mathcal{R}_{\text{DF}}$ that can be achieved with the relay using the decode-and-forward scheme for different scenarios in the two-way case. For unidirectional communication, these regions reduce to the cut-set bound C_{OB} and the achievable rate R_{DF}, which give upper and lower bounds on the capacity of the half-duplex Gaussian one-way relay channel.

As an example scenario, let us consider the line network depicted in Figure 3. This is a simple but commonly used geometry (cf. [20,26]) where the distance $d_{12} = 1$ between the terminals is fixed and the relay is positioned on the line connecting the two terminals such that $d_{1R} = |d|$ and $d_{2R} = |1 - d|$. Furthermore, it is assumed that each node may consume the same transmit power

Figure 3 Line network: the relay is positioned on the line connecting the two terminals that want to exchange information.
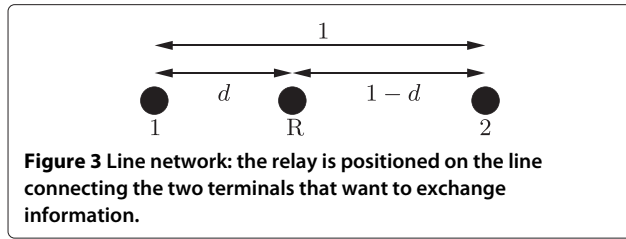
$P_1 = P_2 = P_R = 10$ on average, which for instance is a reasonable assumption in ad hoc networks. Finally, we assume that the path loss exponent is equal to $\alpha = 4$, which is a typical value for urban macrocell environments or multi-level office buildings (cf. [30], Table 2.2), and that all channel coefficients are perfectly known at all nodes.

Within this framework, two different relay network configurations are considered. In the first one, all nodes have a single antenna and the real-valued scalar channel coefficients are specified by $h_{AB} = d_{AB}^{-\alpha/2}$, which of course implies $h_{AB} = h_{BA}$. Note that, due to the assumption of real-valued channels, all rate vectors obtained with the presented optimization framework have to be divided by 2 since the rates are specified by $\frac{1}{2}\log(\cdot)$ in this case as opposed to $\log(\cdot)$ for complex-valued channels. In the second configuration, all nodes are equipped with two antennas. The channel gain matrices are then assumed to be complex random and independent, where the entries of $\boldsymbol{H}_{AB}$ are independent and identically distributed complex Gaussian random variables with zero mean and variance $d_{AB}^{-\alpha}$. In addition, we assume that the channels are reciprocal, i.e., $\boldsymbol{H}_{AB} = \boldsymbol{H}_{BA}^{\mathrm{T}}$.
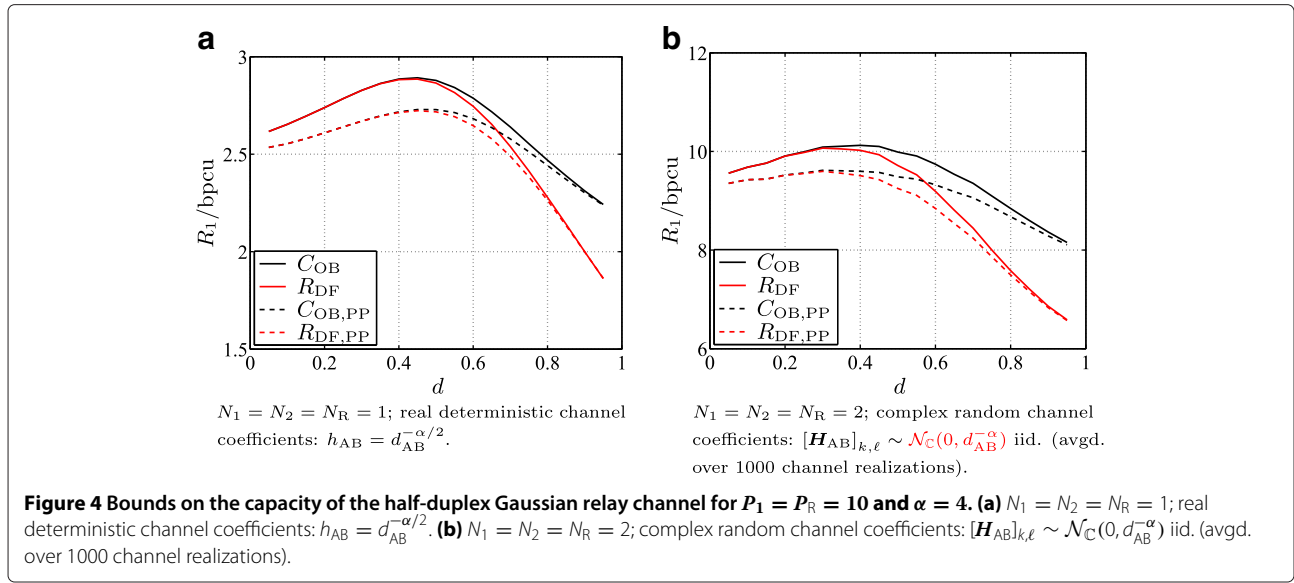
For both the single- and the multi-antenna scenario, Figure 4 shows the cut-set outer bound C_{OB} and the achievable DF rate R_{DF} for the half-duplex one-way relay channel over the distance $d = d_{1R}$ between terminal 1 and the relay. Here, we have assumed that terminal 1 is the source and that terminal 2 is the destination of the unidirectional communication, which means that only phases 1 and 6 of the 6-phase communication protocol are used. We remark that the results for the multi-antenna case are averaged over 1000 independent channel realizations. For comparison, the best outer bound $C_{OB,PP}$ and the best achievable DF rate $R_{DF,PP}$ that can be obtained if the source and the relay are subject to *per protocol phase* transmit power constraints of the form $\mathrm{tr}(\boldsymbol{R}_A^{(i)}) \leq P_A, i \in \{1, 6\}$, are plotted as well. Note that this condition is more restrictive than the average transmit power constraint $\sum_{i=1,6} \tau_i \mathrm{tr}(\boldsymbol{R}_A^{(i)}) \leq P_A$ with $\tau_1, \tau_6 \geq 0, \tau_1 + \tau_6 = 1$ so that $C_{OB} \geq C_{OB,PP}$ and $R_{DF} \geq R_{DF,PP}$.

It can be observed from Figure 4 that the decode-and-forward strategy achieves capacity if the relay is close enough to the source, which is a well-known fact that has previously been noted for the full-duplex case, e.g., in [26]. We also see that the optimal relay positions lie in

the range $0.3 \leq d \leq 0.5$, with the optimal values of d being almost the same for both power constraints. These observations are to be interpreted with caution, however, as the optimal relay position heavily depends on the path loss coefficient as well as the available transmit powers. Another non-surprising observation is that, although a factor of 2 is due to the fact that we use real-valued channels for the single-antenna configuration, substantial rate gains can be achieved without increasing P_1 or P_R when multiple antennas are used at each node. More interestingly, the gap between C_{OB} and $C_{OB,PP}$ as well as that between R_{DF} and $R_{DF,PP}$ vanishes when the relay is moved closer to the destination. This can be explained as follows. The source-relay link, and thus the phase in which the relay listens to the source, increasingly becomes the bottleneck of the information transfer with increasing d. As d approaches $d_{12} = 1$, the optimal time share $\tau_1^\star$ of phase 1 also approaches 1. Hence, the relay power and the transmit power constraint imposed on the relay have no effect on the optimal solution. Furthermore, the average transmit power constraint imposed on the source becomes $\tau_1^\star \mathrm{tr}(\boldsymbol{R}_1^{(1)}) + \tau_6^\star \mathrm{tr}(\boldsymbol{R}_1^{(6)}) \approx \mathrm{tr}(\boldsymbol{R}_1^{(1)}) \leq P_1$, i.e., it basically amounts to a per phase power constraint for phase 1.

For the bidirectional communication in the half-duplex two-way relay channel, we consider three different relay positions: (a) the relay is exactly in the middle between the two terminals ($d = 0.5$); (b) the relay is placed near terminal 1 ($d = 0.25$); (c) the relay is very close to terminal 1 ($d = 0.1$). For each of these scenarios, Figures 5 (single-antenna) and 6 (multi-antenna, results for one particular random channel realization) show the achievable DF rate regions $\mathcal{R}_{DF}$ and the outer bound regions $\mathcal{C}_{OB}$. For comparison, the best achievable DF rate regions $\mathcal{R}_{DF,PP}$ and the best outer bound regions $\mathcal{C}_{OB,PP}$ that can be obtained with per phase power constraints imposed on all nodes are also illustrated. Like for unidirectional transmission, we observe that $\mathcal{C}_{OB} \supset \mathcal{C}_{OB,PP}$ and $\mathcal{R}_{DF} \supset \mathcal{R}_{DF,PP}$ for all scenarios since the average power constraint $\sum_{i=1}^{6} \tau_i \mathrm{tr}(\boldsymbol{R}_A^{(i)}) \leq P_A$ with $\tau_i \geq 0, \forall i \in \{1, \ldots, 6\}$, and $\sum_{i=1}^{6} \tau_i = 1$ is less restrictive than the per phase power constraint $\mathrm{tr}(\boldsymbol{R}_A^{(i)}) \leq P_A, \forall i \in \{1, \ldots, 6\}$.
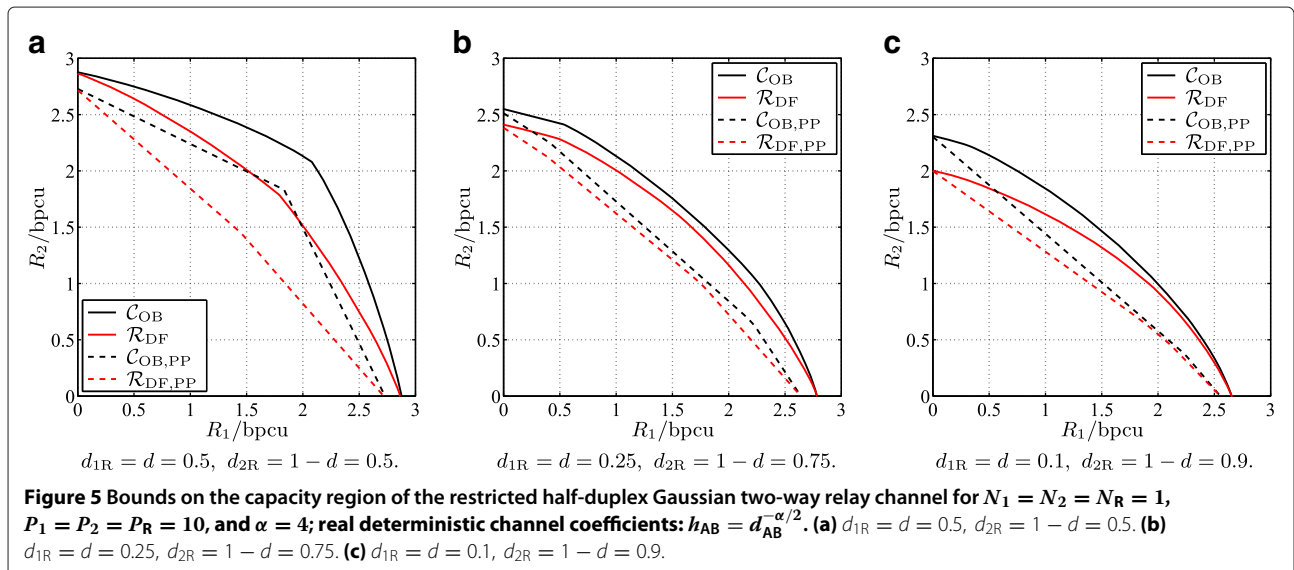
First of all, note that the results shown in Figures 5 and 6 allow to draw the same conclusions as for the one-way case: If the relay is close enough to terminal 1, the decode-and-forward scheme achieves the cut-set bound for the unidirectional communication from terminal 1 to terminal 2, i.e., $R_{1,max} = C_{1,max}$, regardless of whether we consider average or per phase transmit power constraints. Furthermore, the same $R_{2,max}$ (or $C_{2,max}$) is obtained for both types of power constraints when d_{2R} approaches 1. Beyond that, a noteworthy observation is that the greatest benefit of the less restrictive average transmit power constraints is obtained if we are interested in the sum rate
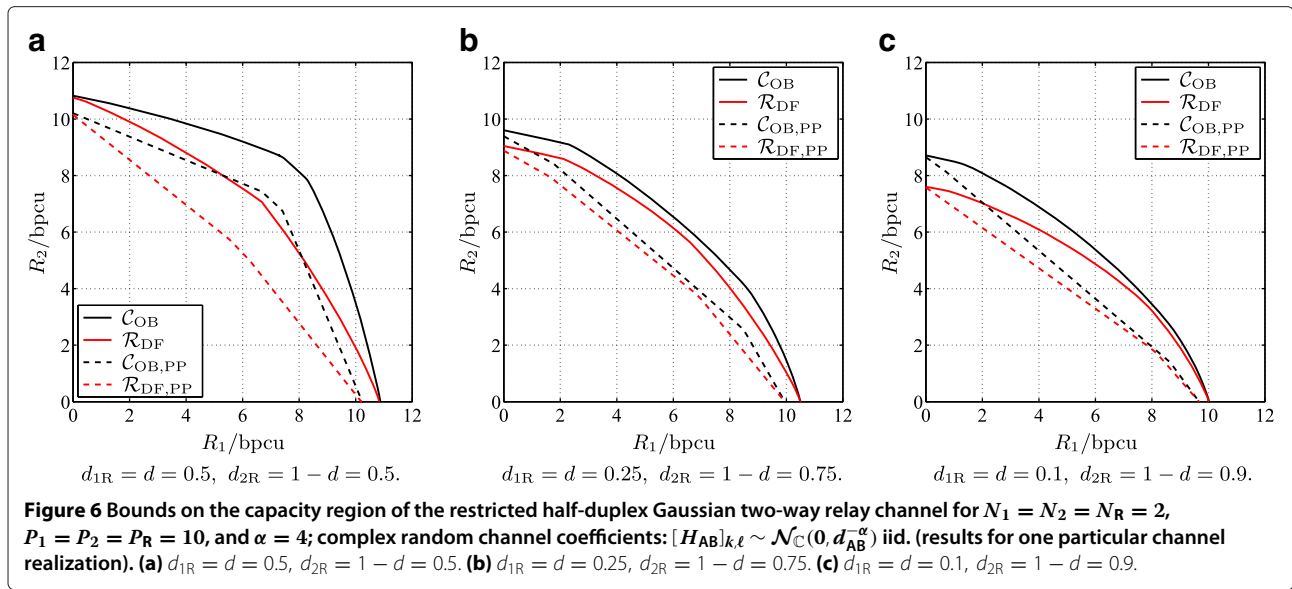
a

b

$N_1 = N_2 = N_R = 1$; real deterministic channel coefficients: $h_{AB} = d_{AB}^{-\alpha/2}$.

$N_1 = N_2 = N_R = 2$; complex random channel coefficients: $[H_{AB}]_{k,\ell} \sim \mathcal{N}_{\mathbb{C}}(0, d_{AB}^{-\alpha})$ iid. (avgd. over 1000 channel realizations).

Figure 4 Bounds on the capacity of the half-duplex Gaussian relay channel for $P_1 = P_R = 10$ and $\alpha = 4$. (a) $N_1 = N_2 = N_R = 1$; real deterministic channel coefficients: $h_{AB} = d_{AB}^{-\alpha/2}$. **(b)** $N_1 = N_2 = N_R = 2$; complex random channel coefficients: $[H_{AB}]_{k,\ell} \sim \mathcal{N}_{\mathbb{C}}(0, d_{AB}^{-\alpha})$ iid. (avgd. over 1000 channel realizations).

$R_1 + R_2$, whereas the performance improvement is less pronounced for asymmetric rate requirements. Finally, observe that the gaps between the boundaries of $\mathcal{C}_{OB}$ and $\mathcal{C}_{OB,PP}$ are like the gaps between the boundaries of $\mathcal{R}_{DF}$ and $\mathcal{R}_{DF,PP}$ for all ratios $\frac{R_1}{R_2}$ and all scenarios considered here.

In order to assess the complexity of determining the achievable rate regions and outer bound regions, Table 1 illustrates the average number of iterations the cutting plane algorithm needs per weighted sum rate maximization problem until it converges for the different scenarios in the multi-antenna case[h]. Here, the parameter ε that specifies the absolute accuracy of the optimal value was set to 10^{-2}. Note that the number of required iterations

is very small if we consider the per protocol phase transmit power constraint. Unfortunately, the numbers roughly triple with the average power constraint that yields the information theoretic bounds on the capacity and the capacity region of the half-duplex Gaussian relay channel and the half-duplex Gaussian two-way relay channel, respectively. The main reason for this is that we need more dual variables to formulate the dual problem in the latter case. Since the number of required iterations remains reasonably small, however, these results confirm that the proposed dual decomposition approach indeed allows to efficiently evaluate achievable rate regions and corresponding outer bounds for the considered half-duplex Gaussian relay networks. Assuming knowledge of all channel gain

a

b

c

$d_{1R} = d = 0.5,\ d_{2R} = 1 - d = 0.5.$

$d_{1R} = d = 0.25,\ d_{2R} = 1 - d = 0.75.$

$d_{1R} = d = 0.1,\ d_{2R} = 1 - d = 0.9.$

Figure 5 Bounds on the capacity region of the restricted half-duplex Gaussian two-way relay channel for $N_1 = N_2 = N_R = 1$, $P_1 = P_2 = P_R = 10$, and $\alpha = 4$; real deterministic channel coefficients: $h_{AB} = d_{AB}^{-\alpha/2}$. (a) $d_{1R} = d = 0.5,\ d_{2R} = 1 - d = 0.5.$ **(b)** $d_{1R} = d = 0.25,\ d_{2R} = 1 - d = 0.75.$ **(c)** $d_{1R} = d = 0.1,\ d_{2R} = 1 - d = 0.9.$

Figure 6 Bounds on the capacity region of the restricted half-duplex Gaussian two-way relay channel for $N_1 = N_2 = N_R = 2$, $P_1 = P_2 = P_R = 10$, and $\alpha = 4$; complex random channel coefficients: $[H_{AB}]_{k,\ell} \sim \mathcal{N}_\mathbb{C}(0, d_{AB}^{-\alpha})$ iid. (results for one particular channel realization). (a) $d_{1R} = d = 0.5$, $d_{2R} = 1 - d = 0.5$. **(b)** $d_{1R} = d = 0.25$, $d_{2R} = 1 - d = 0.75$. **(c)** $d_{1R} = d = 0.1$, $d_{2R} = 1 - d = 0.9$.

matrices, it is hence possible to numerically evaluate their fundamental limits.

6 Conclusion

In this article, we presented a generic method that allows to determine the fundamental limits of uni- and bidirectional communication in the half-duplex Gaussian MIMO relay channel. More specifically, we proposed a dual decomposition approach to evaluate upper and lower bounds on the capacity or the capacity region of the considered MIMO relay channels, for which perfect channel state information (CSI) was assumed. To this end, we modified the approach that was previously proposed in [21] such that the average transmit power constraints under which the cut-set outer bound and the achievable decode-and-forward (DF) rates were derived can be handled. It was shown that the joint optimization of input signals and time allocation decomposes into subproblems that are easier to solve in the dual domain, and we gave an example of how to solve the resulting dual problem by means of the cutting plane algorithm. The beauty of the proposed approach lies in the fact that the phases of the respective communication protocol decouple in the

dual problem. As a result, evaluating the dual function only requires to solve one convex problem for each phase of the communication protocol, which can be done by applying standard semidefinite program (SDP) tools like SDPT3. It is this property that makes dual decomposition so attractive here, especially since the cutting plane algorithm converges after a reasonably small number of iterations.

Furthermore, we remark that our results may be used for protocol design with DF relays in the future. For the one-way case, we can determine what fraction of time the relay should listen to the source and how long it should transmit. For the two-way case, the benefit of our approach is even greater. By not restricting ourselves to any specific protocol from the outset, we let an optimization problem determine which protocol phases should be used and for what fraction of time they should be active to obtain the best performance. At the same time, the approach allows to evaluate any specific communication protocol. All we need to do is set the time shares of the phases that shall not be part of the considered protocol to zero.

Finally, note that average and per phase transmit power constraints can easily be combined using the framework presented in this article. For this purpose, we simply need to add the per phase transmit power constraints to the definitions of the sets $\mathcal{S}_i$ and $\mathcal{S}_i'$ that specify the contributions of the different protocol phases to the outer bound region and the achievable rate region, respectively. Since the sets are then bounded, Proposition 4 becomes obsolete as we do not need a condition on the dual variables $\boldsymbol{\mu}$ for the dual function to be finite. The per phase power constraints considered in [21] can therefore easily be incorporated into the optimization framework

Table 1 Average number of cutting plane iterations needed per weighted sum rate maximization problem ($N_1 = N_2 = N_R = 2$ and $\varepsilon = 10^{-2}$)

	one-way	two-way
$\mathcal{C}_{OB}$	15.1	18.2
$\mathcal{R}_{DF}$	14.4	17.8
$\mathcal{C}_{OB,PP}$	4.6	5.2
$\mathcal{R}_{DF,PP}$	4.6	5.3

presented in this article. Since the converse is not true, the optimization approach presented here generalizes that of [21].

Endnotes

[a]In contrast to full-duplex devices, half-duplex nodes cannot transmit and receive simultaneously in the same frequency band, which means that they require orthogonal resources (time, frequency) for transmission and reception.
[b]Another option to determine points on the boundary of the outer bound region would be to solve rate balancing problems over $\mathcal{C}_{\text{OB}}$ for different ratios of the two rates.
[c]Note that r_1 and r_2 only denote two entries of the sets $\mathcal{S}_1, \ldots, \mathcal{S}_6$. They are not to be confused with R_1 and R_2, which specify the rates of the information exchanged by nodes 1 and 2.
[d]Note that this assumption is not really a restriction. If the relay is not connected to both terminals, it cannot help the communication between the terminals. And while the direct channel between the terminals may be very weak, e.g., due to high path loss, it is still reasonable to assume it supports rates strictly greater than zero.
[e]If there exists an $(\boldsymbol{r}^{(k)}, \boldsymbol{p}^{(k)})$ such that $(\boldsymbol{r}^{(k)}, \boldsymbol{p}^{(k)}) \in \mathcal{S}_i$ and $\boldsymbol{v}^{(k)} = \boldsymbol{B}_i \boldsymbol{r}^{(k)}$ for more than one $i \in \{1, \ldots, 6\}$, we assign the index k to only one set $\mathcal{K}_i$ so that $\mathcal{K}_i \cap \mathcal{K}_j = \emptyset$ for $i \neq j$.
[f]The protocol names are due to [19,20], which are the only two articles among references [14-20] that do not only consider the multiple access broadcast (MABC) protocol.
[g]The reasoning why $\mu_3 > 0$ must hold for $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$ is more complicated in this case since $\lambda_2 = \lambda_4 = 0$ does not imply (22). However, the final conclusion remains the same.
[h]In order to obtain the results for the one-way case, we simply let $\boldsymbol{w} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and considered only phases 1 and 6 in the evaluation of the dual function as explained in Remark 3.

Appendix 1
Proof of Proposition 3
Let $(\boldsymbol{y}, \boldsymbol{z}), (\boldsymbol{y}', \boldsymbol{z}') \in \mathcal{S}$ and $\lambda \in [0, 1]$. Moreover, define $\alpha_i = \lambda \tau_i$ and $\beta_i = (1 - \lambda) \tau_i'$. Then,

$$
\begin{aligned}
\lambda \boldsymbol{y} + (1 - \lambda) \boldsymbol{y}' &= \lambda \sum_{i=1}^{6} \tau_i \boldsymbol{B}_i \boldsymbol{r}_i + (1 - \lambda) \sum_{i=1}^{6} \tau_i' \boldsymbol{B}_i \boldsymbol{r}_i' \\
&= \sum_{i=1}^{6} \boldsymbol{B}_i (\alpha_i \boldsymbol{r}_i + \beta_i \boldsymbol{r}_i') \\
&= \sum_{i=1}^{6} (\alpha_i + \beta_i) \boldsymbol{B}_i \left(\frac{\alpha_i}{\alpha_i + \beta_i} \boldsymbol{r}_i + \frac{\beta_i}{\alpha_i + \beta_i} \boldsymbol{r}_i' \right)
\end{aligned}
$$

and

$$
\begin{aligned}
\lambda \boldsymbol{z} + (1 - \lambda) \boldsymbol{z}' &= \lambda \sum_{i=1}^{6} \tau_i \boldsymbol{p}_i + (1 - \lambda) \sum_{i=1}^{6} \tau_i' \boldsymbol{p}_i' = \sum_{i=1}^{6} \alpha_i \boldsymbol{p}_i + \beta_i \boldsymbol{p}_i' \\
&= \sum_{i=1}^{6} (\alpha_i + \beta_i) \left(\frac{\alpha_i}{\alpha_i + \beta_i} \boldsymbol{p}_i + \frac{\beta_i}{\alpha_i + \beta_i} \boldsymbol{p}_i' \right).
\end{aligned}
$$

Since $\alpha_i, \beta_i \geq 0$, $(\boldsymbol{r}_i, \boldsymbol{p}_i), (\boldsymbol{r}_i', \boldsymbol{p}_i') \in \mathcal{S}_i$, and $\mathcal{S}_i$ is convex, it follows that $\frac{\alpha_i}{\alpha_i + \beta_i} (\boldsymbol{r}_i, \boldsymbol{p}_i) + \frac{\beta_i}{\alpha_i + \beta_i} (\boldsymbol{r}_i', \boldsymbol{p}_i') \in \mathcal{S}_i$, i.e.,

$$
\lambda \boldsymbol{y} + (1 - \lambda) \boldsymbol{y}' = \sum_{i=1}^{6} (\alpha_i + \beta_i) \boldsymbol{B}_i \tilde{\boldsymbol{r}}_i,
$$

$$
\lambda \boldsymbol{z} + (1 - \lambda) \boldsymbol{z}' = \sum_{i=1}^{6} (\alpha_i + \beta_i) \tilde{\boldsymbol{p}}_i, \quad \text{where} \quad (\tilde{\boldsymbol{r}}_i, \tilde{\boldsymbol{p}}_i) \in \mathcal{S}_i.
$$

Furthermore, $0 \leq \alpha_i + \beta_i \leq 1$, $\forall i \in \{1, \ldots, 6\}$, and $\sum_{i=1}^{6} \alpha_i + \beta_i = \sum_{i=1}^{6} (\lambda \tau_i + (1 - \lambda) \tau_i') = 1$, which means that $\lambda(\boldsymbol{y}, \boldsymbol{z}) + (1 - \lambda)(\boldsymbol{y}', \boldsymbol{z}') \in \mathcal{S}$. This proves the proposition. $\qquad\square$

Appendix 2
Proof of Proposition 4
For any $\boldsymbol{\lambda} \geq \boldsymbol{0}$ such that $\boldsymbol{A}^{\text{T}} \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 + \lambda_2 \\ \lambda_3 + \lambda_4 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \boldsymbol{w}$, note that $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$ is equivalent to $\sup_{(\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i} \{\boldsymbol{\lambda}^{\text{T}} \boldsymbol{B}_i \boldsymbol{r}_i - \boldsymbol{\mu}^{\text{T}} \boldsymbol{p}_i\} = \max_{(\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i} \{\boldsymbol{\lambda}^{\text{T}} \boldsymbol{B}_i \boldsymbol{r}_i - \boldsymbol{\mu}^{\text{T}} \boldsymbol{p}_i\} < \infty$ for all $i \in \{1, \ldots, 6\}$. For $\boldsymbol{\mu} > \boldsymbol{0}$, we hence prove the "if" part of the proposition by exemplarily showing that $\sup_{(\boldsymbol{r}_i, \boldsymbol{p}_i) \in \mathcal{S}_i} \{\boldsymbol{\lambda}^T \boldsymbol{B}_i \boldsymbol{r}_i - \boldsymbol{\mu}^T \boldsymbol{p}_i\} < \infty$ for $i = 1$ as corresponding statements for $i = 2, \ldots, 6$ follow along the same lines.

With $\boldsymbol{B}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}^{\text{T}}$ and only terminal 1 transmitting during phase 1, we have

$$
\begin{aligned}
&\max_{(\boldsymbol{r}, \boldsymbol{p}) \in \mathcal{S}_1} \left\{ \boldsymbol{\lambda}^{\text{T}} \boldsymbol{B}_1 \boldsymbol{r} - \boldsymbol{\mu}^{\text{T}} \boldsymbol{p} \right\} \\
&= \max_{(\boldsymbol{r}, \boldsymbol{p}) \in \mathcal{S}_1} \{\lambda_1 r_1 + \lambda_2 r_2 - \mu_1 p_1\} \leq \max_{(\boldsymbol{r}, \boldsymbol{p}) \in \mathcal{S}_1} \{(\lambda_1 + \lambda_2) r_1 - \mu_1 p_1\} \\
&= \max_{\boldsymbol{R}^{(1)} \succeq \boldsymbol{0}} \left\{ (\lambda_1 + \lambda_2) \log \det(\boldsymbol{I}_{N_2 + N_{\text{R}}} + \boldsymbol{H}_1 \boldsymbol{R}^{(1)} \boldsymbol{H}_1^{\text{H}}) - \mu_1 \text{tr}(\boldsymbol{R}^{(1)}) \right\} \\
&= \max_{\boldsymbol{R}^{(1)} \succeq \boldsymbol{0}} \left\{ (\lambda_1 + \lambda_2) \log \det(\boldsymbol{I}_{N_1} + \boldsymbol{H}_1^{\text{H}} \boldsymbol{H}_1 \boldsymbol{R}^{(1)}) - \mu_1 \text{tr}(\boldsymbol{R}^{(1)}) \right\},
\end{aligned}
$$

where the inequality is due to the fact that $I(X_1^{(1)}; Y_{\text{R}}^{(1)} Y_2^{(1)}) = I(X_1^{(1)}; Y_2^{(1)}) + I(X_1^{(1)}; Y_{\text{R}}^{(1)} | Y_2^{(1)}) \geq I(X_1^{(1)}; Y_2^{(1)})$, which follows from the nonnegativity of and the chain rule for mutual information ([25], Chap. 2). Now, suppose $\boldsymbol{H}_1^{\text{H}} \boldsymbol{H}_1 = \boldsymbol{V} \boldsymbol{\Phi} \boldsymbol{V}^{\text{H}}$ with $\boldsymbol{\Phi} = diag(\varphi_1, \ldots, \varphi_{N_1})$ is the eigenvalue decomposition of $\boldsymbol{H}_1^{\text{H}} \boldsymbol{H}_1$, and let us also express $\boldsymbol{R}^{(1)}$ by means of its eigenvalue decomposition $\boldsymbol{R}^{(1)} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{U}^H$. Then, the trace of $\boldsymbol{R}^{(1)}$ is independent of the modal matrix $\boldsymbol{U}$ and equal to the sum of its nonnegative eigenvalues $\sigma_1, \ldots, \sigma_{N_1}$. Moreover, Hadamard's

inequality ([25], Thm. 17.9.2) can be used to show that, with $\boldsymbol{U} = \boldsymbol{V}$,

$$\max_{\boldsymbol{R}^{(1)} \geq \boldsymbol{0}} \left\{ (\lambda_1 + \lambda_2) \log \det (\mathbf{I}_{N_1} + \boldsymbol{H}_1^{\mathrm{H}} \boldsymbol{H}_1 \boldsymbol{R}^{(1)}) - \mu_1 \mathrm{tr}(\boldsymbol{R}^{(1)}) \right\}$$

$$= \max_{\sigma_k \geq 0} \sum_{k=1}^{N_1} (\lambda_1 + \lambda_2) \log (1 + \varphi_k \sigma_k) - \mu_1 \sigma_k.$$

For $\mu_1 > 0$, the right-hand side of above equality has a waterfilling type solution given by

$$\sigma_k^{\star} = \max \left\{ \frac{\lambda_1 + \lambda_2}{\mu_1} - \frac{1}{\varphi_k}, 0 \right\},$$

which implies that $0 \leq \sigma_k^{\star} \leq \frac{\lambda_1 + \lambda_2}{\mu_1} < \infty$ for all $k \in \{1, \dots, N_1\}$, and consequently,

$$\max_{(\boldsymbol{r}, \boldsymbol{p}) \in \mathcal{S}_1} \left\{ \boldsymbol{\lambda}^T \boldsymbol{B}_1 \boldsymbol{r} - \boldsymbol{\mu}^T \boldsymbol{p} \right\} \leq \sum_{k=1}^{N_1} (\lambda_1 + \lambda_2) \log \left(1 + \varphi_k \sigma_k^{\star} \right)$$

$$- \mu_1 \sigma_k^{\star} < \infty.$$

The proofs of the converse and the "if" part of the proposition for $\boldsymbol{\mu} \not> \boldsymbol{0}$ are omitted because they directly follow from the necessary and sufficient conditions for $\Theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) < \infty$ if $\mu_k = 0$, $k \in \{1, 2, 3\}$. □

Appendix 3

Outline of coding scheme that achieves $\mathcal{R}_{\mathrm{DF}}$

The achievability of $\mathcal{R}_{\mathrm{DF}}$ is proved in [22] for a discrete memoryless channel (DMC) without feedback. The coding scheme uses random encoding and jointly typical decoding on the nth extension of the DMC (see [25], Sec. 7.5 for a definition), meaning that the data transmission is performed with n channel uses. Furthermore, it is assumed that TDD phase i is used n_i times, where $\frac{n_i}{n} \to \tau_i \in [0, 1]$ as n grows large.

The message $W_1 \in \{1, \dots, 2^{nR_1}\}$ is to be transmitted from node 1 to node 2, whereas $W_2 \in \{1, \dots, 2^{nR_2}\}$ denotes the message to be sent from terminal 2 to terminal 1 that is independent of W_1. Both messages are split into six parts: $W_1 = (W_{11}, \dots, W_{16})$ and $W_2 = (W_{21}, \dots, W_{26})$ such that $W_{1a} \in \{1, \dots, 2^{nR_{1a}}\}$ and $W_{2b} \in \{1, \dots, 2^{nR_{2b}}\}$, $a, b \in \{1, \dots, 6\}$. The messages are then conveyed to the other terminal as follows:

Phase 1: Node 1 transmits a codeword $X_1^{(1)}(W_{11}, W_{12}, W_{13})$.

Phase 2: Node 2 transmits a codeword $X_2^{(2)}(W_{21}, W_{22}, W_{23})$.

Phase 3: Node 1 transmits a codeword $X_1^{(3)}(W_{14}, W_{15})$ and node 2 sends $X_2^{(3)}(W_{24}, W_{25})$. The two codewords are independent!

After phase 3, the relay reliably decodes the messages $(W_{11}, \dots, W_{15})$ and $(W_{21}, \dots, W_{25})$, which requires

$$R_{11} + R_{12} + R_{13} < \tau_1 I(X_1^{(1)}; Y_{\mathrm{R}}^{(1)}),$$

$$R_{21} + R_{22} + R_{23} < \tau_2 I(X_2^{(2)}; Y_{\mathrm{R}}^{(2)}),$$

$$R_{14} + R_{15} < \tau_3 I(X_1^{(3)}; Y_{\mathrm{R}}^{(3)} | X_2^{(3)}),$$

$$R_{24} + R_{25} < \tau_3 I(X_2^{(3)}; Y_{\mathrm{R}}^{(3)} | X_1^{(3)}),$$

$$R_{14} + R_{15} + R_{24} + R_{25} < \tau_3 I(X_1^{(3)} X_2^{(3)}; Y_{\mathrm{R}}^{(3)}).$$

Phase 4: The relay transmits a codeword $X_{\mathrm{R}}^{(4)}(W_{11}, W_{14}, W_{21}, W_{24})$.

Phase 5: The relay sends a codeword $X_{\mathrm{R}}^{(5)}(W_{22}, W_{25})$, whereas node 2 transmits $X_2^{(5)}(W_{22}, W_{25}, W_{26})$. Note that the two codewords are not independent, but correlated by design in general!

Phase 6: The relay sends a codeword $X_{\mathrm{R}}^{(6)}(W_{12}, W_{15})$ and node 1 transmits $X_1^{(6)}(W_{12}, W_{15}, W_{16})$. Again, note that the two codewords are not independent, but correlated by design in general!

After phase 6, each terminal reliably decodes all parts of the message transmitted by the respective other terminal. Reliable decoding at terminal 1 imposes the conditions

$$R_{21} + R_{24} < \tau_4 I(X_{\mathrm{R}}^{(4)}; Y_1^{(4)}),$$

$$R_{22} + R_{25} < \tau_5 I(X_{\mathrm{R}}^{(5)}; Y_1^{(5)}),$$

$$R_{26} < \tau_5 I(X_2^{(5)}; Y_1^{(5)} | X_{\mathrm{R}}^{(5)}),$$

$$R_{23} < \tau_2 I(X_2^{(2)}; Y_1^{(2)}),$$

whereas reliable decoding at terminal 2 requires

$$R_{11} + R_{14} < \tau_4 I(X_{\mathrm{R}}^{(4)}; Y_2^{(4)}),$$

$$R_{12} + R_{15} < \tau_6 I(X_{\mathrm{R}}^{(6)}; Y_2^{(6)}),$$

$$R_{16} < \tau_6 I(X_1^{(6)}; Y_2^{(6)} | X_{\mathrm{R}}^{(6)}),$$

$$R_{13} < \tau_1 I(X_1^{(1)}; Y_2^{(1)}).$$

Noting that $R_1 = \sum_{a=1}^{6} R_{1a}$, $R_2 = \sum_{b=1}^{6} R_{2b}$, putting all constraints together, and taking the closure of the resulting achievable rate region yields $\mathcal{R}_{\mathrm{DF}}$.

While the achievable rate region $\mathcal{R}_{\mathrm{DF}}$ was derived for a DMC, we remark that Theorem 5 remains valid for channel models with continuous random variables. This is because the decode-and-forward strategy can be derived by means of weakly typical sequences and since the concept of weak typicality applies to continuous random variables as well (cf. [26], Rem. 28).

References

1. GJ Foschini, MJ Gans, On limits of wireless communications in a fading environment when using multiple antennas. Wirel. Personal Commun. **6**, 311–335 (1998)
2. IE Telatar, Capacity of multi-antenna Gaussian channels. Europ. Trans. Telecommun. **10**, 585–595 (1999)
3. EC van der Meulen, Three-terminal communication channels. Adv. Appl. Probab. **3**, 120–154 (1971)
4. TM Cover, A EL Gamal, Capacity theorems for the relay channel. IEEE Trans. Inf. Theory. **25**(5), 572–584 (1979)
5. MA Khojastepour, A Sabharwal, B Aazhang, in *Information Processing in Sensor Networks, Volume 2634 of* Lecture Notes in Computer Science, ed. by F Zhao, L Guibas. Bounds on achievable rates for general multi-terminal networks with practical constraints (Springer, Berlin, 2003), pp. 146–161
6. MA Khojastepour, A Sabharwal, B Aazhang, in *37th Annual Conference on Information Sciences and Systems (CISS)*. On the capacity of 'Cheap' relay networks, (Baltimore 2003)
7. MA Khojastepour, A Sabharwal, B Aazhang, in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 3. On capacity of Gaussian 'Cheap' relay channel, (San Francisco, 2003), pp. 1776–1780
8. A Host-Madsen, J Zhang, Capacity bounds and power allocation for wireless relay channels. IEEE Trans. Inf. Theory. **51**(6), 2020–2040 (2005)
9. B Wang, J Zhang, A Host-Madsen, On the capacity of MIMO relay channels. IEEE Trans. Inf. Theory. **51**, 29–43 (2005)
10. CTK Ng, GJ Foschini, Transmit signal and bandwidth optimization in multiple-antenna relay channels. IEEE Trans. Commun. **59**(11), 2987–2992 (2011)
11. L Gerdes, W Utschick, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Optimized capacity bounds for the MIMO relay channel, Prague, 2011), pp. 3336–3339
12. L Gerdes, W Utschick, in *International ITG Workshop on Smart Antennas (WSA)*. Optimized capacity bounds for the half-duplex Gaussian MIMO relay channel, (Aachen, 2011)
13. S Simoens, O Muñoz-Medina, J Vidal, A del Coso, On the Gaussian MIMO relay channel with full channel state information. IEEE Trans. Signal Process. **57**(9), 3588–3599 (2009)
14. B Rankov, A Wittneben, in *39th Asilomar Conference on Signals, Systems and Computers*. Spectral efficient signaling for half-duplex relay channels, (Monterey, 2005), pp. 1066–1071
15. I Hammerstroem, M Kuhn, C Esli, J Zhao, A Wittneben, G Bauch, in *IEEE Workshop on Signal Process. Advances in Wireless Communications (SPAWC) (Helsinki)*. MIMO two-way relaying with transmit CSI at the relay, (2007)
16. TJ Oechtering, C Schnurr, I Bjelakovic, H Boche, Broadcast capacity region of two-phase bidirectional relaying. IEEE Trans. Inf. Theory. **54**, 454–458 (2008)
17. C Schnurr, TJ Oechtering, S Stanczak, in *41st Asilomar Conference on Signals, Systems and Computers*. Achievable rates for the restricted half-duplex two-way relay channel, (Monterey, 2007), pp. 1468–1472
18. TJ Oechtering, RF Wyrembelski, H Boche, in *IEEE International Conference on Communications (ICC)*. On the optimal transmission for the MIMO bidirectional broadcast channel, (Dresden, 2009)
19. SJ Kim, P Mitran, V Tarokh, Performance bounds for bidirectional coded cooperation protocols. IEEE Trans. Inf. Theory. **54**(11), 5235–5241 (2008)
20. SJ Kim, N Devroye, P Mitran, V Tarokh, in *IEEE Sarnoff Symposium*. Comparison of bi-directional relaying protocols, (Princeton 2008)
21. L Gerdes, M Riemensberger, W Utschick, On achievable rate regions for half-duplex Gaussian MIMO relay channels: a decomposition approach. IEEE J. Sel. Areas Commun. **30**(8), 1319–1330 (2012)
22. M Stein, Towards optimal schemes for the half-duplex two-way relay channel. Submitted to IEEE J. Sel. Areas Commun (2011). http://arxiv.org/abs/1101.3198
23. A El Gamal, M Aref, The capacity of the semideterministic relay channel. IEEE Trans. Inf. Theory. **28**(3), 536 (1982)
24. A El Gamal, S Zahedi, Capacity of a class of relay channels with orthogonal components. IEEE Trans. Inf. Theory. **51**(5), 1815–1817 (2005)
25. TM Cover, JA Thomas, *Elements of Information Theory*, 2nd edn. (John Wiley & Sons, Hoboken, 2006)
26. G Kramer, M Gastpar, P Gupta, Cooperative strategies and capacity theorems for relay networks. IEEE Trans. Inf. Theory. **51**(9), 3037–3063 (2005)
27. MS Bazaraa, HD Sherali, CM Shetty, *Nonlinear Programming*, 3rd edn. (John Wiley & Sons, Hoboken, 2006)
28. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, New York, 2004)
29. KC Toh, MJ Todd. RH Tutuncu, On the implementation and usage of SDPT3—a MATLAB software package for semidefinite-quadratic-linear programming, version 4.0 (2010). http://hdl.handle.net/1813/15133
30. A Goldsmith, *Wireless Communications*. (Cambridge University Press, New York, 2005)