

On the Gaussian MIMO Relay Channel With Full Channel State Information

Sebastien Simoens, Olga Muñoz-Medina, Josep Vidal, *Member, IEEE*, and Aitor del Coso, *Member, IEEE*

Abstract—This paper addresses the problem of source and relay transmit covariance optimization on the Gaussian MIMO relay channel with full channel state information (CSI), i.e., assuming perfect knowledge of all channels. For full-duplex relaying, we show that the cut-set bound on capacity can be computed as the solution of a convex problem, thus providing a tighter bound than previously published. For time division duplex (TDD) relaying, both upper and lower bounds on capacity are derived, and the transmit covariance matrices are optimized for decode-and-forward (DF) strategies with either partial or full decoding at the relay. A generic procedure is introduced to formulate these problems into a standard convex form, and to solve them efficiently. Suboptimum precoders are also proposed which have a specific matrix structure that either leads to a closed-form expression or at least reduces the dimension of the optimization problem. Practical aspects related to transmit power constraints and CSI availability are then discussed. Finally, simulations in a cellular downlink scenario show that the partial DF strategy can achieve a rate very close to capacity for realistic values of the source to relay SNR, and that the rate loss due to suboptimum precoder structures remains small for typical antenna configurations.

Index Terms—Cooperative, relay, channel state information (CSI), MIMO.

I. INTRODUCTION

ALTHOUGH capacity bounds for the relay channel have been available for years [1], the research activity in this domain is still thriving. This interest is supported by the recent standardization activities in, e.g., the IEEE802.16 j and m Task Groups [2], where cooperative [3] relaying strategies are proposed. Some promising research topics include the multiple-relay channel (e.g., [4]), the two-way relay channel [5], [6] and the multiuser relay channel (e.g., [7]). Moreover, it is now possible to integrate multiple antennas not only in infrastructure devices (e.g., base stations, fixed relay stations) but also in mobile devices (e.g., handsets), and to exploit MIMO channel

state information (CSI) not only at the receiving node (CSIR) but also at the transmitting node (CSIT). Point-to-point MIMO with full¹ CSI is now a well investigated topic: the transmit covariance that attains the Gaussian MIMO channel capacity is derived by Telatar in [8], while the maximization of various other cost functions is performed in [9]. However, several aspects of the MIMO relay channel with full CSI are still unsolved even for the one-way single relay channel and this is the domain in which our paper aims at contributing. We therefore consider a single Relay (R) which cooperates with a Source (S) and a Destination (D) to maximize the information rate from S to D. The number of antennas at S, R and D are respectively denoted N_S , N_R and N_D and each can be greater than 1.

In [1], the capacity of the full-duplex single-antenna single-relay channel is studied: the DF (Theorem 1) and compress-and-forward (CF, Theorem 6) coding strategies are introduced and an upper-bound on the relay channel capacity, known as the cut-set bound (CSB) is provided. With DF, the relay decodes at least a part of the message transmitted by the source, whereas with CF the relay observation is compressed and forwarded to the destination as a new codeword. Neither DF nor CF is capacity-achieving in the general relay channel. However, DF (respectively, CF) meets the CSB and becomes capacity-achieving when the capacity of the S-R link (respectively, the R-D link) gets infinitely large (see, e.g., [7]). In this paper full-duplex relaying is only briefly addressed and more focus is put on half-duplex relaying, which lends itself more easily to practical implementation by preventing the transmit chain of the relay terminal to interfere with its receive chain. We consider a time division duplex (TDD)² protocol with a relay-receive slot of duration denoted by $t \in [0; 1]$ followed by a relay-transmit slot of duration $(1 - t)$. Using the same conventions as [11] we assume a “fixed-dynamic” TDD protocol, i.e., t is not random but can be optimized for a given channel realization.

In [12], the CSB and the achievable rates of partial DF and CF strategies are expressed in the case of Gaussian TDD relaying, assuming single antenna devices. In [13], an upper-bound and several lower bounds are derived for the full-duplex MIMO relay channel with full CSI. As explained in Section II of this paper the upper-bound in [13] is in general larger than the CSB. Two achievable rates are also derived in [13]. The first rate (Theorem 3.2) is simply the minimum between the capacity of the source-destination point-to-point MIMO channel and that of a cascaded channel in which the destination treats

Manuscript received September 08, 2008; accepted March 10, 2009. First published April 10, 2009; current version published August 12, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Walid Hachem. This work was supported in part by the ICT ROCKET Project of the 7th European Research Framework, by the Spanish/Catalan Science and Technology Commissions, and by FEDER funds from the EC: TEC2006-06481/TCM and TEC2004-04526.

S. Simoens was with Motorola Labs Paris, Parc Les Algorithmes, 91193, Gif-sur-Yvette Cedex, France. He is now with Thales Avionics, Valence 26000, France (e-mail: sebastien.simoens@fr.thalesgroup.com).

O. Muñoz-Medina and J. Vidal are with the Signal Theory and Communications Department, Technical University of Catalonia, 08034 Barcelona, Spain (e-mail: olga.munoz@upc.edu; josep.vidal@upc.edu).

A. del Coso is with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), 08860, Castelldefels, Spain (e-mail: aitor.delcoso@cttc.es).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2009.2020744

¹In this paper, full CSI refers to the assumption that perfect knowledge of all channels is available at each node.

²Note that frequency division duplex (FDD) relaying as proposed in, e.g., [10] also assumes that the transmissions from S to R and from R to D take place in different time slots.

the signal received from the source as noise. The second achievable rate (derived in the Appendix of [13]) corresponds to a Block-Markov coding strategy (see, e.g., [7]). As the authors point out, it is expected to yield a tighter bound than Theorem 3.2. The problem of maximizing the rate w.r.t. the source and relay transmit covariance matrices for this strategy is formulated but not solved since only suboptimum transmit covariance matrices are proposed. In [14], two strategies for the full-duplex MIMO relay channel are introduced based on either superposition coding or dirty paper coding at the source and both are proven to achieve a higher rate than [13, Theorem 3.2]. However, once again the achievable rate maximization problem is formulated but not solved for the general case. In [11] and [15], a Diversity Multiplexing Tradeoff analysis of the full and half duplex MIMO relay channel is carried on. In [16], [17], Source and Relay precoders are derived for linear relaying (LR) with full MIMO CSI. Although LR—which generalizes amplify-and-forward—is interesting thanks to its processing simplicity, it presents the drawback in single-relay TDD case that the time-sharing is constrained to $t = 1/2$, which severely limits the achievable rate, therefore LR will not be considered in the following.

The present paper contributes to the study of MIMO relaying with full CSI as follows:

- The CSB is formulated as a convex optimization problem for both full-duplex and TDD relaying. In the full-duplex case, this upper-bound on capacity is tighter than the one proposed in [13]. In the TDD case, the formulation is obtained by exploiting the convexity-preserving property of perspective function [18] (Section II-A).
- A convex formulation of the achievable rates maximization problem is obtained for DF strategies with either partial or full decoding at the relay for TDD relaying (Section III).
- Two generic procedures are proposed to efficiently solve the above problems and compute the upper and lower bounds. For this purpose, several tools are exploited from optimization theory (e.g., duality, epigraph [18]), optimization algorithms (e.g., gradient projection [19], barrier method [18], [19]) and differentiation techniques (e.g., patterned matrix derivatives [21], [22]). (Section II-B).
- Suboptimum source and relay precoder structures are proposed for the DF strategy with full decoding. In this case, either analytical expressions can be derived from KKT conditions (Section III-B–1) or at least the problem dimensions can be reduced. (Section III-B–2).
- The extension of the optimization procedures to handle more realistic constraints such as multicarrier transmission and per-antenna power constraint is discussed (Section III-C) and the upper and lower bounds are finally benchmarked by simulations in a cellular downlink scenario (Section IV).

Note that parts of this work were published in [23]. The present paper extends [23] on many aspects, including the convex formulation in TDD case, the use of patterned derivatives, a discussion on implementation constraints and additional simulation results.

Notation. We use boldface capital font for matrices and boldface lower-case for column vectors. Given a scalar $x \in \mathbb{R}$ we

denote $x^+ \triangleq \max(x, 0)$ and \mathbf{x}^+ its component-wise extension to vectors $\mathbf{x} \in \mathbb{R}^N$. $(\cdot)^*$ denotes the component-wise complex-conjugate operator, $(\cdot)^T$ is the transpose and $(\cdot)^H$ the Hermitian transpose. The notation $\text{vec}(\cdot)$ stands for the operator that stacks the columns of a matrix into a column vector. Likewise, $\mathbf{X} = \text{mat}(\mathbf{x})$ denotes the $N \times N$ matrix obtained from an $N^2 \times 1$ vector \mathbf{x} such that $\text{vec}(\text{mat}(\mathbf{x})) = \mathbf{x}$. The diagonal matrix whose diagonal terms are given by vector \mathbf{x} is denoted $\text{diag}(\mathbf{x})$. The identity matrix of size $N \times N$ is denoted \mathbf{I}_N and the null matrix is denoted $\mathbf{0}_N$. As in [18], the cone of Positive semidefinite (PSD) matrices of size $k \times k$ is denoted by \mathcal{S}_+^k and the set of positive (respectively, strictly positive) real numbers is \mathbb{R}_+ (respectively, \mathbb{R}_{++}). We write $\mathbf{X} \succeq \mathbf{0}$ for a PSD matrix and $\mathbf{x} \geq 0$ for a vector with components in \mathbb{R}_+ . The mutual information between vectors \mathbf{x} and \mathbf{y} is written $I(\mathbf{x}; \mathbf{y})$. Given a complex matrix \mathbf{H} of size $M \times N$ and a complex PSD matrix $\mathbf{X} \in \mathcal{S}_+^M$, the following quantity is defined: $\mathcal{C}(\mathbf{X}, \mathbf{H}) \triangleq \log_2 |\mathbf{I}_M + \mathbf{H}\mathbf{X}\mathbf{H}^H|$. For a given \mathbf{H} , the function $f : \mathbf{X} \rightarrow \mathcal{C}(\mathbf{X}, \mathbf{H})$ is concave on \mathcal{S}_+^M .

II. THE CUT-SET BOUND WITH FULL MIMO CSI

In this section, we show that the computation of the CSB can be formulated as a convex optimization problem in the full-duplex and TDD relaying cases. Procedures to solve this problem are then proposed. These procedures will be directly applicable to DF coding strategies in Section III.

A. Formulating the CSB as a Convex Problem

1) *Full-Duplex Relaying Case:* In this paper the channels from S to D, S to R and R to D are all assumed static and are denoted, respectively $\mathbf{H}_0, \mathbf{H}_1$ and \mathbf{H}_2 . As stated earlier, full CSI is assumed. Moreover, unless explicitly stated otherwise, each node is subject to a maximum transmit power constraint, denoted by P_S and P_R for the source and relay respectively. The CSI and power assumptions will be discussed in more details in Section III-C. In full-duplex MIMO relaying, the signals received at the relay and destination can be written as in [13]

$$\begin{aligned} \mathbf{y}_R &= \mathbf{H}_1 \mathbf{x}_S + \mathbf{n}_R \\ \mathbf{y}_D &= \mathbf{H}_0 \mathbf{x}_S + \mathbf{H}_2 \mathbf{x}_R + \mathbf{n}_D \end{aligned} \quad (1)$$

where circularly symmetric complex white Gaussian noise of unit variance is assumed³ at the relay and destination, i.e., $\mathbf{n}_R \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_R})$ and $\mathbf{n}_D \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_D})$. The capacity C_{FD} of the full-duplex relay channel is upper-bounded by the cut-set bound $C_{\text{CSB,FD}}$ whose expression is given by [13, eq. (3)]

$$C_{\text{FD}} \leq C_{\text{CSB,FD}} = \max_{p(\mathbf{x}_S, \mathbf{x}_R)} \min(I(\mathbf{x}_S; \mathbf{y}_D, \mathbf{y}_R | \mathbf{x}_R), I(\mathbf{x}_S, \mathbf{x}_R; \mathbf{y}_D)) \quad (2)$$

where the maximization is performed over the joint distribution of the source and relay codebooks $p(\mathbf{x}_S, \mathbf{x}_R)$. The authors in [13] show that the optimum $p(\mathbf{x}_S, \mathbf{x}_R)$ is Gaussian and conclude

³Note that correlated noise of covariance Γ_R and Γ_D can be modeled by replacing $\mathbf{H}_0, \mathbf{H}_1$ and \mathbf{H}_2 by respectively, $\tilde{\mathbf{H}}_0 \triangleq \Gamma_D^{-1/2} \mathbf{H}_0$, $\tilde{\mathbf{H}}_1 \triangleq \Gamma_R^{-1/2} \mathbf{H}_1$ and $\tilde{\mathbf{H}}_2 \triangleq \Gamma_D^{-1/2} \mathbf{H}_2$ in all the equations of this paper.

that the optimization of (2) must be carried on w.r.t. three matrices $\mathbf{R}_S \triangleq E[\mathbf{x}_S \mathbf{x}_S^H]$, $\mathbf{R}_R \triangleq E[\mathbf{x}_R \mathbf{x}_R^H]$ and the cross correlation $E[\mathbf{x}_S \mathbf{x}_R^H]$. This optimization seems highly nontrivial and nonconvex. Therefore, the authors exploit matrix inequalities and introduce a scalar parameter ρ that captures the cross correlation. They finally obtain (see [13, Theorem 3.1]) an upper-bound which involves a maximization only over \mathbf{R}_S , \mathbf{R}_R , and ρ

$$\begin{aligned} C_{FD} &\leq C_{CSB,FD} \leq \max_{\rho \in [0;1], \mathbf{R}_S \succeq 0, \mathbf{R}_R \succeq 0} \min(C_A, C_B) \\ C_A &\triangleq \mathcal{C} \left(\mathbf{R}_S, \sqrt{1-\rho^2} \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \right) \\ C_B &\triangleq \inf_{a>0} \mathcal{C} \left(\begin{bmatrix} \mathbf{R}_S & \mathbf{0}_{N_S \times N_R} \\ \mathbf{0}_{N_R \times N_S} & \mathbf{R}_R \end{bmatrix}, \right. \\ &\quad \left. \begin{bmatrix} \sqrt{1+\frac{\rho^2}{a}} \mathbf{H}_0 & \sqrt{1+a} \mathbf{H}_2 \end{bmatrix} \right) \\ \text{s.t.} \quad &\text{tr}(\mathbf{R}_S) \leq P_S \quad \text{tr}(\mathbf{R}_R) \leq P_R. \end{aligned} \quad (3)$$

Although its derivation is very elegant, this bound unfortunately suffers several restrictions: *i*) It is in the general case strictly larger than the CSB (e.g., equality with the CSB requires $N_S \leq N_R$) *ii*) Although C_A and C_B are concave in \mathbf{R}_S and \mathbf{R}_R for a fixed ρ , the problem is not convex in $(\mathbf{R}_S, \mathbf{R}_R, \rho)$ and thus the proposed algorithm in [13] includes a nonconvex one-dimensional optimization over ρ . *iii*) Its numerical evaluation is computationally intensive. Indeed, it is not possible to obtain a closed-form expression for the partial derivatives of C_B w.r.t. \mathbf{R}_S and \mathbf{R}_R . We therefore have to resort to numerical differentiation with respect to each component of these two matrices, which requires a number of evaluations of C_B that is proportional to N_S^2 (respectively, N_R^2), and each evaluation of C_B requires by definition to solve an optimization problem.

The aforementioned limitations can be overcome by considering the joint covariance matrix:

$$\mathbf{R}_{SR} \triangleq \begin{bmatrix} \mathbf{R}_S & E[\mathbf{x}_S \mathbf{x}_R^H] \\ E[\mathbf{x}_R \mathbf{x}_S^H] & \mathbf{R}_R \end{bmatrix}. \quad (4)$$

Let also define the following matrices:

$$\mathbf{D}_S \triangleq [\mathbf{I}_{N_S} \quad \mathbf{0}_{N_S \times N_R}] \quad \text{and} \quad \mathbf{D}_R \triangleq [\mathbf{0}_{N_R \times N_S} \quad \mathbf{I}_{N_R}]. \quad (5)$$

From (4) and (5), the following relationships hold:

$$\mathbf{R}_S = \mathbf{D}_S \mathbf{R}_{SR} \mathbf{D}_S^H \quad \text{and} \quad \mathbf{R}_R = \mathbf{D}_R \mathbf{R}_{SR} \mathbf{D}_R^H. \quad (6)$$

Note that if \mathbf{R}_{SR} is PSD, then \mathbf{R}_S and \mathbf{R}_R are PSD too

$$\mathbf{R}_{SR} \succeq 0 \Rightarrow \mathbf{R}_S \succeq 0 \quad \text{and} \quad \mathbf{R}_R \succeq 0. \quad (7)$$

Indeed, for any two vectors $\mathbf{x} \in \mathbb{C}^{N_S}$ and $\mathbf{y} \in \mathbb{C}^{N_R}$, defining $\tilde{\mathbf{x}} \triangleq \mathbf{D}_S^H \mathbf{x}$ and $\tilde{\mathbf{y}} \triangleq \mathbf{D}_R^H \mathbf{y}$, the following holds:

$$\begin{aligned} \mathbf{x}^H \mathbf{R}_S \mathbf{x} &\stackrel{(a)}{=} \mathbf{x}^H \mathbf{D}_S \mathbf{R}_{SR} \mathbf{D}_S^H \mathbf{x} = \tilde{\mathbf{x}}^H \mathbf{R}_{SR} \tilde{\mathbf{x}} \stackrel{(b)}{\geq} 0 \\ \mathbf{y}^H \mathbf{R}_R \mathbf{y} &\stackrel{(a)}{=} \mathbf{y}^H \mathbf{D}_R \mathbf{R}_{SR} \mathbf{D}_R^H \mathbf{y} = \tilde{\mathbf{y}}^H \mathbf{R}_{SR} \tilde{\mathbf{y}} \stackrel{(b)}{\geq} 0 \end{aligned} \quad (8)$$

where (a) comes from (6) and (b) from the positive semidefiniteness of \mathbf{R}_{SR} .

The cut-set bound (2) can, therefore, be expressed as

$$\begin{aligned} C_{CSB,FD} &= \max_{\mathbf{R}_{SR} \succeq 0} \min \left\{ \mathcal{C} \left(\mathbf{R}_{SR}, \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \mathbf{D}_S \right), \right. \\ &\quad \left. \mathcal{C}(\mathbf{R}_{SR}, [\mathbf{H}_0 \quad \mathbf{H}_2]) \right\} \\ \text{s.t.} \quad &\text{tr}(\mathbf{D}_S \mathbf{R}_{SR} \mathbf{D}_S^H) \leq P_S \quad \text{and} \\ &\text{tr}(\mathbf{D}_R \mathbf{R}_{SR} \mathbf{D}_R^H) \leq P_R. \end{aligned} \quad (9)$$

The objective is concave on the PSD cone, because it is the pointwise minimum of two concave functions [18]. The constraints are affine. Therefore, the problem is convex. Thus, any locally optimal point is globally optimal and we can rely on the convex optimization literature [18], [19] to solve the problem efficiently.

2) *TDD Relaying Case:* As in [20] we can define three TDD relaying protocols. In Protocol I, S transmits to both R and D during the first slot. Then during the second slot both S and R transmit to D. Protocol II is a variant of Protocol I in which only the Relay transmits during the second slot. Protocol III is also a variant of Protocol I in which D only receives signals transmitted during the second slot. In this section we consider the more general and more complex Protocol I, from which Protocols II and III can be easily derived. The CSB can be expressed as follows (see [12, eq. (77)]):

$$\begin{aligned} C_{CSB,TDD} &= \max_{t \in [0;1], p(\mathbf{x}_S^{(1)}, \mathbf{x}_S^{(2)}, \mathbf{x}_R^{(2)})} \min\{C_A, C_B\} \\ C_A &\triangleq tI(\mathbf{x}_S^{(1)}; \mathbf{y}_R^{(1)}, \mathbf{y}_D^{(1)} | \mathbf{x}_R^{(1)} = \mathbf{0}) \\ &\quad + (1-t)I(\mathbf{x}_S^{(2)}; \mathbf{y}_D^{(2)} | \mathbf{x}_R^{(2)}) \\ C_B &\triangleq tI(\mathbf{x}_S^{(1)}; \mathbf{y}_D^{(1)} | \mathbf{x}_R^{(1)} = \mathbf{0}) \\ &\quad + (1-t)I(\mathbf{x}_S^{(2)}, \mathbf{x}_R^{(2)}; \mathbf{y}_D^{(2)}) \end{aligned} \quad (10)$$

where the superscript (i) indicates the slot during which the signal was transmitted or received. Using similar arguments as in the full-duplex case, the cut-set bound for the TDD MIMO relay channel can be expressed as

$$\begin{aligned} C_{CSB,TDD} &= \max_{t \in [0;1], \mathbf{R}_S^{(1)} \succeq 0, \mathbf{R}_{SR}^{(2)} \succeq 0} \min\{C_A, C_B\} \\ C_A &\triangleq t\mathcal{C} \left(\mathbf{R}_S^{(1)}, \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \right) + (1-t)\mathcal{C}(\mathbf{R}_{SR}^{(2)}, \mathbf{H}_0 \mathbf{D}_S) \\ C_B &\triangleq t\mathcal{C}(\mathbf{R}_S^{(1)}, \mathbf{H}_0) + (1-t)\mathcal{C}(\mathbf{R}_{SR}^{(2)}, [\mathbf{H}_0 \quad \mathbf{H}_2]) \\ \text{s.t.} \quad &\text{tr}(\mathbf{R}_S^{(1)}) \leq P_S; \text{tr}(\mathbf{D}_S \mathbf{R}_{SR}^{(2)} \mathbf{D}_S^H) \\ &\leq P_S; \text{tr}(\mathbf{D}_R \mathbf{R}_{SR}^{(2)} \mathbf{D}_R^H) \leq P_R. \end{aligned} \quad (11)$$

Equation (11) is convex in $(\mathbf{R}_S^{(1)}, \mathbf{R}_{SR}^{(2)})$ for a given t , but convexity in $(\mathbf{R}_S^{(1)}, \mathbf{R}_{SR}^{(2)}, t)$ cannot be claimed at this stage. However, let us introduce the following changes of variables into

(11): $t_1 \triangleq t, t_2 \triangleq 1 - t, \mathbf{Q}_S^{(1)} \triangleq t_1 \mathbf{R}_S^{(1)}$ and $\mathbf{Q}_{SR}^{(2)} \triangleq t_2 \mathbf{R}_{SR}^{(2)}$. The problem now reads

$$\begin{aligned} C_{\text{CSB,TDD}} &= \max_{t_1 > 0, t_2 > 0, \mathbf{Q}_S^{(1)} \succeq 0, \mathbf{Q}_{SR}^{(2)} \succeq 0} \min\{C_A, C_B\} \\ C_A &\triangleq t_1 \mathcal{C}\left(\mathbf{Q}_S^{(1)}/t_1, \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix}\right) + t_2 \mathcal{C}\left(\mathbf{Q}_{SR}^{(2)}/t_2, \mathbf{H}_0 \mathbf{D}_S\right) \\ C_B &\triangleq t_1 \mathcal{C}\left(\mathbf{Q}_S^{(1)}/t_1, \mathbf{H}_0\right) + t_2 \mathcal{C}\left(\mathbf{Q}_{SR}^{(2)}/t_2, [\mathbf{H}_0 \quad \mathbf{H}_2]\right) \\ \text{s.t.} & \\ \text{tr}\left(\mathbf{Q}_S^{(1)}\right) - t_1 P_S &\leq 0; \text{tr}\left(\mathbf{D}_S \mathbf{Q}_{SR}^{(2)} \mathbf{D}_S^H\right) - t_2 P_S \leq 0; \\ \text{tr}\left(\mathbf{D}_R \mathbf{Q}_{SR}^{(2)} \mathbf{D}_R^H\right) - t_2 P_R &\leq 0; t_1 + t_2 - 1 \leq 0 \end{aligned} \quad (12)$$

where the trivial cases $t = 0$ and $t = 1$ were excluded from the domain. The function $g : (\mathbf{X}, t) \mapsto t f(\mathbf{X}/t)$ defined on $\mathcal{S}_+^M \times \mathbb{R}_{++}$ is the perspective (cf. [18, Sec. 3.2.6]) of the function $f : \mathbf{X} \mapsto \mathcal{C}(\mathbf{X}, \mathbf{H})$ which is concave on \mathcal{S}_+^M , therefore, g is concave on $\mathcal{S}_+^M \times \mathbb{R}_{++}$. Equation (12) is, therefore, convex in standard form.

B. Computing the Cut-Set Bound

The previous section has shown that the CSB can be expressed as the solution of a convex problem in the full-duplex and TDD cases. Since it does not seem possible to derive a closed-form expression of the solution of (9) and (12), efficient procedures are sought hereafter to solve them numerically. For this purpose, the convex optimization literature is exploited. We focus on solving (12) from which the solution of the simpler problem (9) will be straightforward.

We start by writing the equivalent epigraph [18] form of (12) such that closed-form expressions of partial derivatives can be found for the objective and the constraints

$$\begin{aligned} C_{\text{CSB,TDD}} &= - \min_{(\varepsilon, t_1, t_2, \mathbf{Q}_S^{(1)}, \mathbf{Q}_{SR}^{(2)}) \in \mathcal{X}} -\varepsilon \\ \text{s.t.} & \\ \varepsilon - C_A &\leq 0; \varepsilon - C_B \leq 0 \\ \text{tr}\left(\mathbf{Q}_S^{(1)}\right) - t_1 P_S &\leq 0; \text{tr}\left(\mathbf{D}_S \mathbf{Q}_{SR}^{(2)} \mathbf{D}_S^H\right) - t_2 P_S \leq 0; \\ \text{tr}\left(\mathbf{D}_R \mathbf{Q}_{SR}^{(2)} \mathbf{D}_R^H\right) - t_2 P_R &\leq 0; t_1 + t_2 - 1 \leq 0 \end{aligned} \quad (13)$$

where

$$\mathcal{X} \triangleq \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_{++} \times \mathcal{S}_+^{N_S} \times \mathcal{S}_+^{N_S + N_R}. \quad (14)$$

The optimization problem in (13) needs to be carried on with respect to three real-valued variables and two PSD matrices. In [22], a framework is presented for computing the derivatives of functions of complex-valued matrices which are patterned, i.e., which have a special structure (e.g., Toeplitz, Hermitian). In our problem, the Hermitian-symmetric structure of matrices $\mathbf{Q}_S^{(1)}$ and $\mathbf{Q}_{SR}^{(2)}$ is exploited. The computation of closed-form expressions for the partial derivatives and gradients involved in this paper is detailed in the Appendix. Following (60), let define the following parameterization of matrices $\mathbf{Q}_S^{(1)}$ and $\mathbf{Q}_{SR}^{(2)}$

$$\mathbf{Q}_S^{(1)} \triangleq \mathcal{F}\left(\mathbf{r}_S^{(1)}, \mathbf{c}_S^{(1)}, \left(\mathbf{c}_S^{(1)}\right)^*\right) \quad (15)$$

$$\mathbf{Q}_{SR}^{(2)} \triangleq \mathcal{F}\left(\mathbf{r}_{SR}^{(2)}, \mathbf{c}_{SR}^{(2)}, \left(\mathbf{c}_{SR}^{(2)}\right)^*\right). \quad (16)$$

In order to simplify the notations, all the variables in the problem are stacked into the following vector:

$$\mathbf{v} \triangleq \left(\varepsilon, t_1, t_2, \left(\mathbf{r}_S^{(1)}\right)^T, \left(\mathbf{c}_S^{(1)}\right)^T, \left(\mathbf{r}_{SR}^{(2)}\right)^T, \left(\mathbf{c}_{SR}^{(2)}\right)^T\right)^T \quad (17)$$

and the set \mathcal{V} is defined such that

$$\mathbf{v} \in \mathcal{V} \Leftrightarrow \left(\varepsilon, t_1, t_2, \mathbf{Q}_S^{(1)}, \mathbf{Q}_{SR}^{(2)}\right) \in \mathcal{X}. \quad (18)$$

Moreover, the inequality constraints in (13) are denoted as $f_j(\mathbf{v}) \leq 0, j = 1, \dots, J$ and stacked into a vector-valued function \mathbf{f} as follows:

$$\mathbf{f}(\mathbf{v}) \triangleq (f_1(\mathbf{v}), \dots, f_J(\mathbf{v}))^T. \quad (19)$$

Using the notations (17)–(19) and the gradient expressions of the Appendix, two alternative ways of solving (13) are now presented.

1) *Dual Method*: A first method to compute the cut-set bound consists in solving the dual problem

$$\max_{\boldsymbol{\mu} \geq 0} g(\boldsymbol{\mu}) \quad \text{where } g(\boldsymbol{\mu}) \triangleq \inf_{\mathbf{v} \in \mathcal{V}} \left(\underbrace{-\varepsilon + \boldsymbol{\mu}^T \mathbf{f}(\mathbf{v})}_{\triangleq \mathcal{L}(\mathbf{v}, \boldsymbol{\mu})} \right) \quad (20)$$

where $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu})$ denotes the Lagrangian and $g(\boldsymbol{\mu})$ is the dual function. If the problem is strongly dual (see [18, Sec. 5.2.3]) then there is no duality gap and solving (20) yields the cut-set bound. In order to prove strong duality, we show that Slater's condition holds. Let denote by $\tilde{\mathcal{V}}$ (respectively, $\tilde{\mathcal{X}}$) the subset of \mathcal{V} (respectively, \mathcal{X}) over which all the constraints are inactive, i.e., $f_j(\mathbf{v}) < 0 \forall j \in \{1, \dots, J\}$. This set is called the interior set of \mathcal{V} (respectively, \mathcal{X}) and Slater's condition is satisfied if $\tilde{\mathcal{V}}$ (respectively, $\tilde{\mathcal{X}}$) is nonempty. Note that for instance the following point is strictly feasible:

$$\mathbf{x}_0 \triangleq \left(0, \frac{1}{4}, \frac{1}{4}, \frac{P_S}{8N_S} \mathbf{I}_{N_S}, \frac{\min(P_S, P_R)}{8(N_S + N_R)} \mathbf{I}_{N_S + N_R}\right) \in \tilde{\mathcal{X}}. \quad (21)$$

Therefore, Slater's condition is satisfied and strong duality holds. The computation of the dual function $g(\boldsymbol{\mu})$ at a given $\boldsymbol{\mu}_0$ requires to minimize the Lagrangian $\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}_0)$ over \mathcal{V} . The gradient of the Lagrangian $\nabla \mathcal{L}(\mathbf{v}, \boldsymbol{\mu}_0)$ can be computed in closed-form as described in the Appendix. However, a straightforward application of gradient descent methods cannot guarantee that the sequence of points belong to \mathcal{V} . We therefore resort to the gradient projection method (GPM) which is described in [19, Sec. 2.3]. The GPM is an iterative algorithm which computes at step k the following points:

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \alpha^{(k)} \left(\bar{\mathbf{v}}^{(k)} - \mathbf{v}^{(k)}\right) \quad (22)$$

$$\bar{\mathbf{v}}^{(k)} = \mathcal{P}_{\mathcal{V}}\left(\mathbf{v}^{(k)} - s^{(k)} \nabla \mathcal{L}\Big|_{\mathbf{v}^{(k)}}\right) \quad (23)$$

where $s^{(k)}$ is a positive scalar, $\alpha^{(k)} \in (0, 1]$ is the step size and $\mathcal{P}_{\mathcal{V}}$ denotes the projection on the set \mathcal{V} . The GPM is selected here because the projection operator turns out to be simple. The projection of t_1 and t_2 onto \mathbb{R}_{++} can be practically handled by

restricting their domain to a closed interval $[\eta; +\infty)$ where $\eta > 0$. In this case their projections are, respectively, $\max(t_1, \eta)$ and $\max(t_2, \eta)$. By selecting a small enough η , the error introduced on the final solution can be made arbitrarily small. Finally, a projection from the set of Hermitian matrices onto the PSD cone is needed. Let consider the eigenvalue decomposition of an $N \times N$ Hermitian matrix: $\mathbf{M} = \mathbf{U}\text{diag}(\boldsymbol{\lambda})\mathbf{U}^H$. The projection of \mathbf{M} onto \mathcal{S}_+^N is [24]

$$\mathcal{P}_{\mathcal{S}_+^N}(\mathbf{M}) = \mathbf{U}\text{diag}(\boldsymbol{\lambda}^+)\mathbf{U}^H. \quad (24)$$

There exist various methods to select parameters $s^{(k)}$ and $\alpha^{(k)}$ in order to ensure convergence. In our simulations we used the Armijo Rule along the feasible direction ([19, Sec. 2.3]).

The previous paragraph presented a procedure to compute the dual function by minimizing the Lagrangian. In order to solve the dual problem (20), it remains to maximize the dual function with respect to $\boldsymbol{\mu}$. From (20), the dual function may not be differentiable. However, as shown below, a closed-form expression of a subgradient can be found (see also [19, Sec. 6.3]). Since the dual function is concave in $\boldsymbol{\mu}$, a vector \mathbf{h} is a subgradient of g at $\boldsymbol{\mu}_0$ if for all $\boldsymbol{\mu}_1$

$$g(\boldsymbol{\mu}_1) \leq g(\boldsymbol{\mu}_0) + \mathbf{h}^T(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0). \quad (25)$$

Let $\hat{\mathbf{v}}_0$ and $\hat{\mathbf{v}}_1$ be minimizers of the Lagrangian at, respectively, $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$. Then

$$\begin{aligned} g(\boldsymbol{\mu}_1) &\triangleq \mathcal{L}(\hat{\mathbf{v}}_1, \boldsymbol{\mu}_1) \leq \mathcal{L}(\hat{\mathbf{v}}_0, \boldsymbol{\mu}_1) \\ \Rightarrow g(\boldsymbol{\mu}_1) - g(\boldsymbol{\mu}_0) &\leq \mathcal{L}(\hat{\mathbf{v}}_0, \boldsymbol{\mu}_1) - \mathcal{L}(\hat{\mathbf{v}}_0, \boldsymbol{\mu}_0) \\ &= \mathbf{f}(\hat{\mathbf{v}}_0)^T(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0). \end{aligned} \quad (26)$$

From (26) and (25) it can be concluded that $\mathbf{f}(\hat{\mathbf{v}}_0)$ is a subgradient of g at $\boldsymbol{\mu}_0$, and the dual can be solved by subgradient methods. The subgradient method generates a sequence of dual-feasible points according to the following iteration:

$$\boldsymbol{\mu}^{(k+1)} = \mathcal{P}_{\mathcal{M}}\left(\boldsymbol{\mu}^{(k)} + s^{(k)}\mathbf{h}\right) \quad (27)$$

where \mathbf{h} is the subgradient, $s^{(k)}$ is a positive scalar step size and $\mathcal{P}_{\mathcal{M}}$ is the projection on the set \mathcal{M} of dual-feasible points. We refer the reader to [19, Sec. 6.3] in which several step size selection strategies are proposed and in particular to Proposition 6.3.1 which states that for sufficiently small step size, the distance to the optimum $\hat{\boldsymbol{\mu}}$ is reduced at each iteration.

2) *Barrier Method*: The previous paragraph described a procedure that computes the cut-set bound by solving the dual problem. An alternate approach is to solve the primal problem using an interior-point algorithm. The barrier method (see [18, Sec. 11.3.1.]) is the selected algorithm in this paragraph. Given two fixed real parameters $\alpha > 0$ and $\beta > 1$, the barrier method is an iterative procedure which solves at the i th iteration the following unconstrained minimization problem:

$$\min_{\mathbf{v} \in \tilde{\mathcal{V}}} \left\{ -\varepsilon + \frac{1}{m(i)} \sum_{j=1}^J \phi_j(\mathbf{v}) \right\} \quad (28)$$

$$\phi_j(\mathbf{v}) \triangleq -\log(-f_j(\mathbf{v})) \quad (29)$$

where $m(1) = \alpha$ and $m(i) = \beta m(i-1)$. The function ϕ_j is the logarithmic barrier associated to the j th inequality constraint. This function tends to $+\infty$ when $f_j(\mathbf{v}) \rightarrow 0$. The expression of barrier functions is straightforward for the constraints $-t_1 < 0$ and $-t_2 < 0$. Moreover, as pointed out in [25, Theorem 5.1], $-\log|\mathbf{Q}|$ is a barrier function for the positive definiteness constraint $\mathbf{Q} \succ 0$. Therefore, the minimization of (28) can be carried on without the need for any projection as long as the descent starts from an initial point \mathbf{v}_0 in $\tilde{\mathcal{V}}$. For instance the point in (21) is a possible starting point. In our simulations we used a steepest descent with backtracking line search [18] to provide the step size. For sufficiently small step size, the candidate next point is guaranteed to lie within the interior set.

The two procedures presented above are expected to provide an efficient numerical computation of the cut-set bound in the full-duplex and TDD cases. Unfortunately, it is not easy to assess their complexity for a given accuracy. For instance, as stated in [19, Sec. 6.3], step size selection rules are empirical. When maximizing the dual function in (27), we applied the step size updating rule of [19, eq. (6.28)]

$$s^{(k)} = \alpha^{(k)} \frac{(\tilde{g} - g(\boldsymbol{\mu}^{(k)}))}{\|\mathbf{h}^{(k)}\|^2} \quad (30)$$

where $\alpha^{(k)} = (1+m)/(k+m)$ with m a fixed positive integer and \tilde{g} is an upper-bound on the optimum dual. Any primal-feasible point $\tilde{\mathbf{v}}$ obtained when minimizing the Lagrangian at previous iterations gives an upper-bound $\tilde{g} = f(\tilde{\mathbf{v}})$, but not all such points are primal-feasible and therefore \tilde{g} may be updated infrequently. As a result we observed that in our simulations solving the primal problem with an interior point method was more computationally efficient than solving the dual problem.

Having addressed the computation of the cut-set bound in this section, we will now investigate in the next section the problem of precoder and resource allocation optimization for specific coding strategies.

III. DECODE-AND-FORWARD (DF) WITH FULL MIMO CSI

This section focuses on DF relaying strategies for the TDD MIMO relay channel. The partial DF coding strategy is addressed in Section III-A before considering simpler coding and optimization strategies in Section III-B.

A. Partial DF

In partial DF relaying [12], S transmits a first message ω_0 at a rate R_0 using a signal $\mathbf{x}_S^{(1)}(\omega_0)$ during the first slot. The relay (R) transmits $\mathbf{x}_R^{(2)}(\omega_0)$ during the second slot while S transmits $\mathbf{x}_S^{(2)}(\omega_0, \omega_1) = \mathbf{x}_{S,0}^{(2)}(\omega_0) + \mathbf{x}_{S,1}^{(2)}(\omega_1)$. This strategy is called partial DF because R only decodes the part ω_0 of the message (ω_0, ω_1) in order to cooperate with S during the second slot. Because we assume a synchronized scenario, the signals $\mathbf{x}_R^{(2)}(\omega_0)$ and $\mathbf{x}_{S,0}^{(2)}(\omega_0)$ are correlated, whereas ω_1 is mapped onto an independent signal $\mathbf{x}_{S,1}^{(2)}(\omega_1)$ transmitted at rate R_1 using superposition coding. The destination successively decodes $\hat{\omega}_0$ and $\hat{\omega}_1$.

The derivation of the achievable rate is a straightforward extension of [12, Proof of Prop. 2]

$$R_{\text{PDF}} = \max_{t \in [0;1], \mathbf{P}(\mathbf{x}_S^{(1)}, \mathbf{x}_{S,0}^{(2)}, \mathbf{x}_{S,1}^{(2)}, \mathbf{x}_R^{(2)})} R_0 + R_1 \quad (31)$$

$$= \max_{t \in [0;1], \mathbf{R}_S^{(1)} \succeq 0, \mathbf{R}_{S,1}^{(2)} \succeq 0, \mathbf{R}_{SR,0}^{(2)} \succeq 0} \min(R_A, R_B) \quad (32)$$

$$\begin{aligned} R_A &\triangleq t\mathcal{C}(\mathbf{R}_S^{(1)}, \mathbf{H}_1) + (1-t)\mathcal{C}(\mathbf{R}_{S,1}^{(2)}, \mathbf{H}_0 \mathbf{D}_S) \\ R_B &\triangleq t\mathcal{C}(\mathbf{R}_S^{(1)}, \mathbf{H}_0) \\ &\quad + (1-t)\mathcal{C} \left(\begin{bmatrix} \mathbf{R}_{S,1}^{(2)} & \mathbf{0}_{N_S \times (N_S + N_R)} \\ \mathbf{0}_{(N_S + N_R) \times N_S} & \mathbf{R}_{SR,0}^{(2)} \end{bmatrix}, \right. \\ &\quad \left. \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_0 & \mathbf{H}_2 \end{bmatrix} \right) \end{aligned} \quad (33)$$

where

$$\begin{aligned} \mathbf{R}_S^{(1)} &\triangleq E \left[\mathbf{x}_S^{(1)} \left(\mathbf{x}_S^{(1)} \right)^H \right] \\ \mathbf{R}_{S,1}^{(2)} &\triangleq E \left[\mathbf{x}_{S,1}^{(2)} \left(\mathbf{x}_{S,1}^{(2)} \right)^H \right] \quad \text{and} \\ \mathbf{R}_{SR,0}^{(2)} &\triangleq E \left[\begin{bmatrix} \left(\mathbf{x}_{S,0}^{(2)} \right)^T & \left(\mathbf{x}_R^{(2)} \right)^T \end{bmatrix}^T \right. \\ &\quad \left. \times \begin{bmatrix} \left(\mathbf{x}_{S,0}^{(2)} \right)^H & \left(\mathbf{x}_R^{(2)} \right)^H \end{bmatrix} \right] \end{aligned}$$

and the transmit power constraints can be written as

$$\text{tr}(\mathbf{R}_S^{(1)}) \leq P_S; \text{tr}(\mathbf{D}_R \mathbf{R}_{SR,0}^{(2)} \mathbf{D}_R^H) \leq P_R \quad (34)$$

$$\text{tr}(\mathbf{R}_{S,1}^{(2)}) + \text{tr}(\mathbf{D}_S \mathbf{R}_{SR,0}^{(2)} \mathbf{D}_S^H) \leq P_S. \quad (35)$$

Equation (31) with the constraints (34) and (35) is similar to (11) and can be turned into a convex problem in standard form just like (13). The Source and Relay precoders during the first and second slot and the time sharing variable which maximize the achievable rate of the partial DF strategy can be computed by exactly the same procedures as for the cut-set bound in the previous section.

B. Full DF

The partial DF strategy requires to implement superposition coding at the source and successive interference cancellation at the destination. In this section, we consider a strategy with reduced implementation complexity at the expense of a lower achievable rate. We call it Full DF (FDF) and define it as a variant of partial DF in which the relay decodes the whole message and the source does not superimpose a new message during

the second slot (i.e., $R_1 = 0$). In this case, the achievable rate simplifies as

$$R_{\text{FDF}} = \max_{t \in [0;1], \mathbf{R}_S^{(1)} \succeq 0, \mathbf{R}_{SR}^{(2)} \succeq 0} \min(tR_A, tR_{B,1} + (1-t)R_{B,2}) \quad (36)$$

$$\text{s.t. } R_A \triangleq \mathcal{C}(\mathbf{R}_S^{(1)}, \mathbf{H}_1) \quad (37)$$

$$R_{B,1} \triangleq \mathcal{C}(\mathbf{R}_S^{(1)}, \mathbf{H}_0) \quad (38)$$

$$R_{B,2} \triangleq \mathcal{C}(\mathbf{R}_{SR}^{(2)}, [\mathbf{H}_0 \ \mathbf{H}_2]) \quad (39)$$

$$\begin{aligned} \text{tr}(\mathbf{R}_S^{(1)}) &\leq P_S; \text{tr}(\mathbf{D}_R \mathbf{R}_{SR}^{(2)} \mathbf{D}_R^H) \leq P_R; \\ \text{tr}(\mathbf{D}_S \mathbf{R}_{SR}^{(2)} \mathbf{D}_S^H) &\leq P_S. \end{aligned} \quad (40)$$

For any fixed $t \in [0; 1]$, it can be noticed that R_{FDF} is a nondecreasing function of $R_{B,2}$, which only depends on $\mathbf{R}_{SR}^{(2)}$. Therefore, the optimization (36) can be carried on in two steps [see (41) and (42) at the bottom of the page], where the trivial case $t = 0$ was excluded from the domain.

Note that problem (42) is defined for TDD protocol I. The computation of the rates $R_{\text{FDF,II}}$ and $R_{\text{FDF,III}}$ of the FDF strategy for protocols II and III is somewhat simpler. The rate $R_{\text{FDF,II}}$ is obtained by solving (42) in which $\hat{R}_{B,2}$ is replaced by the capacity $\hat{\mathcal{C}}(\mathbf{H}_2, P_R)$ of the point-to-point MIMO link between R and D

$$\hat{\mathcal{C}}(\mathbf{H}, P) \triangleq \max_{\mathbf{R} \succeq 0} \mathcal{C}(\mathbf{R}, \mathbf{H}) \text{ s.t. } \text{tr}(\mathbf{R}) \leq P \quad (43)$$

where problem (43) is the MIMO channel capacity with CSIT as solved by Telatar in [8]. The rate $R_{\text{FDF,III}}$ is given by

$$\begin{aligned} R_{\text{FDF,III}} &= \max_{t \in [0;1]} \min(t\hat{\mathcal{C}}(\mathbf{H}_1, P_S), (1-t)\hat{R}_{B,2}) \\ &= \frac{\hat{\mathcal{C}}(\mathbf{H}_1, P_S)\hat{R}_{B,2}}{\hat{\mathcal{C}}(\mathbf{H}_1, P_S) + \hat{R}_{B,2}}. \end{aligned} \quad (44)$$

Let also define the noncooperative DF (NCDF) strategy as a special case of FDF for Protocol II in which the destination only receives the signal during the second slot. This strategy is well known in the literature and standard bodies under various other names such as regenerative multi-hopping or packet forwarding. It will be used as a benchmark to highlight the gains due to cooperation in the performance evaluation Section IV of this paper. The achievable rate for NCDF is

$$R_{\text{NC,DF}} \triangleq \max_{t \in [0;1], \mathbf{R}_S^{(1)} \succeq 0, \mathbf{R}_R^{(2)} \succeq 0} \min \left(t\mathcal{C}(\mathbf{R}_S^{(1)}, \mathbf{H}_1), (1-t)\mathcal{C}(\mathbf{R}_R^{(2)}, \mathbf{H}_2) \right) \quad (45)$$

$$\hat{R}_{B,2} \triangleq \max_{\mathbf{R}_{SR}^{(2)} \succeq 0} R_{B,2} \quad (41)$$

$$R_{\text{FDF}} = \max_{t \in [0;1], \mathbf{Q}_S^{(1)} \succeq 0} \min \left\{ t\mathcal{C}(\mathbf{Q}_S^{(1)}/t, \mathbf{H}_1), t\mathcal{C}(\mathbf{Q}_S^{(1)}/t, \mathbf{H}_0) + (1-t)\hat{R}_{B,2} \right\} \quad (42)$$

$$\begin{aligned}
& \text{s.t. } \text{tr}(\mathbf{R}_S^{(1)}) \leq P_S; \text{tr}(\mathbf{R}_R^{(2)}) \leq P_R \\
R_{\text{NC,DF}} &= \max_{t \in [0;1]} (t\hat{C}(\mathbf{H}_1, P_S), (1-t)\hat{C}(\mathbf{H}_2, P_R)) \\
&= \frac{\hat{C}(\mathbf{H}_1, P_S)\hat{C}(\mathbf{H}_2, P_R)}{\hat{C}(\mathbf{H}_1, P_S) + \hat{C}(\mathbf{H}_2, P_R)}. \quad (46)
\end{aligned}$$

From (46) it is clear that the achievable rate of NCDF is upper-bounded by the minimum between the capacity of the first hop link and that of the second hop link

$$R_{\text{NC,DF}} \leq \min(\hat{C}(\mathbf{H}_1, P_S), \hat{C}(\mathbf{H}_2, P_R)). \quad (47)$$

Note that when N_R grows to infinity while the average SNR per receive antenna remains constant on each link, then (11) reduces to (45), i.e., NCDF becomes capacity achieving. This means that the achievable rate gain to be expected from advanced cooperative coding strategies such as partial DF tends to reduce as the number of antennas at the relay becomes large. Coming back to the more complex cooperative case, it can first be noticed that problems (41) and (42) can be solved using the convex optimization procedures described previously. However, we decide to evaluate suboptimum precoders at the Source and Relay with a structure that further reduces the optimization complexity.

1) *Suboptimum Source Precoder During 1st Slot:* Let first consider the problem (42) in which the Source precoder during the first time slot is optimized. Let introduce the SVD of \mathbf{H}_0 and \mathbf{H}_1

$$\mathbf{H}_0 = \mathbf{U}_0 \text{diag}(\boldsymbol{\lambda}_0) \mathbf{V}_0^H \quad \text{and} \quad \mathbf{H}_1 = \mathbf{U}_1 \text{diag}(\boldsymbol{\lambda}_1) \mathbf{V}_1^H \quad (48)$$

We arbitrarily impose the following structure to the source covariance matrix:

$$\mathbf{R}_S^{(1)} = \underbrace{\mathbf{V}_0 \text{diag}(\mathbf{p}_0) \mathbf{V}_0^H}_{\triangleq \mathbf{R}_0} + \underbrace{\mathbf{V}_1 \text{diag}(\mathbf{p}_1) \mathbf{V}_1^H}_{\triangleq \mathbf{R}_1}. \quad (49)$$

The structure (49) stems from the intuition that the source shall transmit part of its power on the eigenmodes of the channel to the relay and the rest on the eigenmodes of the channel to the destination (note the similarity with the precoder optimization in [6]). Let L_0 and L_1 denote the number of nonzero singular values of \mathbf{H}_0 and \mathbf{H}_1 . Inserting (48) and (49) into (37) and (38) gives

$$\begin{aligned}
R_A &= \log_2 |\mathbf{I}_{N_R} + \mathbf{H}_1 \mathbf{R}_0 \mathbf{H}_1^H + \mathbf{H}_1 \mathbf{R}_1 \mathbf{H}_1^H| \\
&\stackrel{(a)}{\geq} \log_2 |\mathbf{I}_{N_R} + \mathbf{H}_1 \mathbf{R}_1 \mathbf{H}_1^H| \\
&\stackrel{(b)}{=} \sum_{i=1}^{L_1} \log_2 (1 + \lambda_{1,i}^2 p_{1,i}) \triangleq J_1 \\
R_{B,1} &\stackrel{(a)}{\geq} \log_2 |\mathbf{I}_{N_D} + \mathbf{H}_0 \mathbf{R}_0 \mathbf{H}_0^H| \\
&\stackrel{(b)}{=} \sum_{i=1}^{L_0} \log_2 (1 + \lambda_{0,i}^2 p_{0,i}) \triangleq J_0 \quad (50)
\end{aligned}$$

where (a) comes from the Minkowski determinant inequality and (b) comes from (49). Inserting these lower bounds on R_A

and $R_{B,1}$ into the epigraph form of (36) yields the following lower-bound on R_{FDF} :

$$R_{\text{FDF}} \geq -\min_{\varepsilon, t \in [0;1], \mathbf{p}_0 \geq \mathbf{0}_{L_0}, \mathbf{p}_1 \geq \mathbf{0}_{L_1}} (-\varepsilon) \quad (51)$$

$$\text{s.t. } \varepsilon - tJ_1 \leq 0 \quad (52)$$

$$\varepsilon - tJ_0 - (1-t)\hat{R}_{B,2} \leq 0 \quad (53)$$

$$\mathbf{1}_{L_0}^T \mathbf{p}_0 + \mathbf{1}_{L_1}^T \mathbf{p}_1 - P_S \leq 0. \quad (54)$$

Before solving the above-defined problem, it can first be noticed that when the source-relay and source-destination channels are orthogonal, equality occurs in (50). If \mathbf{H}_0 and \mathbf{H}_1 have i.i.d. complex Gaussian components, let consider the distribution of the angle ϕ between any two rows \mathbf{h}_0 and \mathbf{h}_1 of respectively \mathbf{H}_0 and \mathbf{H}_1 . For $N_S > 1$ the quantity $\xi \triangleq |\mathbf{h}_1^* \mathbf{h}_0|^2 / (\|\mathbf{h}_0\|_2^2 \|\mathbf{h}_1\|_2^2) = \cos^2(\phi)$ is Beta-distributed with parameters 1 and $N_S - 1$ [26]. This distribution concentrates around 0 as $N_S \rightarrow +\infty$. Therefore the source precoder that solves (51) becomes optimum for (42) as N_S grows.

We now derive a procedure for solving (51). Introducing the perspective function as in (12) allows to turn (51) into a convex problem which has a reduced number of dimensions compared to (42) leading to a reduction of the computational complexity. Unfortunately, writing the KKT conditions for this problem does not seem to lead to a closed-form expression. However, for a fixed t in (51), the optimization w.r.t. $(\varepsilon, \mathbf{p}_1, \mathbf{p}_2)$ is also a convex problem for which the KKT conditions lead to

$$\hat{p}_{0,i} = (\alpha - 1/\lambda_{0,i}^2)^+; \hat{p}_{1,i} = (\beta - 1/\lambda_{1,i}^2)^+. \quad (55)$$

The solution (55) can be obtained by a water-filling algorithm with two water levels $\alpha > 0$ and $\beta > 0$ which are not independent due to the total source power constraint (54). Let define $P_1 \triangleq (\mathbf{1}_{L_1}^T \mathbf{p}_1) / P_S$ the fraction of Source power transmitted on the source-relay channel eigenmodes, while the rest $1 - P_1$ is transmitted on the eigenmodes of the source-destination channel. Finding α and β amounts to finding the optimum $\hat{P}_1 \in [0; 1]$. Let first assume that both constraints (52) and (53) are active at the optimum, which yields

$$tJ_1 = tJ_0 - (1-t)\hat{R}_{B,2}. \quad (56)$$

It can be checked that $J_1 = 0$ at $P_1 = 0$ and J_1 is nondecreasing with P_1 . Likewise, J_0 is nonincreasing with P_1 and equals 0 at $P_1 = 1$. Therefore, the optimum \hat{P}_1 is found by solving (56) under the condition that $(1-t)\hat{R}_{B,2} \leq tJ_1$ at $P_1 = 1$. When this constraint is not satisfied, either (52) or (53) is not active and the solution is trivial (i.e., $\hat{P}_1 = 0$ or $\hat{P}_1 = 1$). In order to solve (51), it remains to perform a one-dimensional optimization with respect to the variable t . Fortunately, it can be checked numerically that the solution $\hat{\varepsilon}(t)$ of (51) for a given t turns out to be a unimodal function of t over the interval $[0; 1]$, i.e., a function that is either strictly increasing or strictly decreasing. Therefore, efficient one-dimensional search techniques such as the Golden Section Search (see [19, Appendix C.3]) can be employed to find the optimum \hat{t} .

To summarize, a suboptimum approach to source precoder optimization with reduced complexity is proposed in this section. It consists in transmitting a fraction of the source power

on the eigenmodes of the channel to the relay and the rest on the eigenmodes of the channel to the destination. The power assignment is provided by a water-filling algorithm with two water levels that are related by the total source power constraint. This precoder tends to become optimum as the number of antennas at the source becomes large.

2) Suboptimum Source and Relay Precoder During 2nd Slot:

Let now consider problem (41) in which the source and relay precoders are optimized during the second time slot under an individual power constraint. Introducing the SVD $[\mathbf{H}_2\mathbf{H}_0] = \mathbf{U}\text{diag}(\boldsymbol{\lambda})\mathbf{V}^H$ and the change of variable $\mathbf{R} \triangleq \mathbf{V}^H\mathbf{R}_{SR}^{(2)}\mathbf{V}$, (41) can be rewritten as

$$\begin{aligned} \hat{R}_{B,2} &= \max_{\mathbf{R} \succeq 0} \log_2 |\mathbf{I}_{N_D} + \text{diag}(\boldsymbol{\lambda})\mathbf{R}(\text{diag}(\boldsymbol{\lambda}))^H| \\ \text{s.t. } \text{tr}(\mathbf{D}_S\mathbf{V}\mathbf{R}\mathbf{V}^H\mathbf{D}_S^H) &\leq P_S; \text{tr}(\mathbf{D}_R\mathbf{V}\mathbf{R}\mathbf{V}^H\mathbf{D}_R^H) \leq P_R. \end{aligned} \quad (57)$$

As in [13, Appendix B], we arbitrarily enforce a diagonal structure $\mathbf{R} = \text{diag}(\mathbf{p}_2)$. This turns the matrix optimization problem (57) into the following vector optimization:

$$\begin{aligned} \max_{\mathbf{p}_2 \geq 0} \sum_{i=1}^L \log_2(1 + \lambda_i^2 p_{2,i}) \\ \text{s.t. } \mathbf{a}^T \mathbf{p}_2 \leq P_S; \mathbf{b}^T \mathbf{p}_2 \leq P_R \end{aligned} \quad (58)$$

where L is the number of nonzero eigenvalues of $[\mathbf{H}_2\mathbf{H}_0]$ and the vectors \mathbf{a} and \mathbf{b} have their components defined by $a_i \triangleq \sum_{j=1}^{N_S} |V_{N_R+j,i}|^2$ and $b_i \triangleq \sum_{j=1}^{N_R} |V_{j,i}|^2$. The problem (58) can then be solved at a much lower computational cost than (41). Note that if we replace the individual power constraints in (57) by a sum-power constraint $\text{tr}(\mathbf{R}_{SR}^{(2)}) = \text{tr}(\mathbf{R}) \leq P_S + P_R$, then Hadamard determinant inequality can be applied as in [8] to show that the optimum \mathbf{R} in (57) is diagonal. In other words, under a total transmit power constraint at the source and relay, transmitting on the eigenmodes of the joint channel $[\mathbf{H}_2\mathbf{H}_0]$ is the optimum precoding strategy.

C. Accounting for Actual System Constraints

Due to practical system implementation constraints, several key assumptions made in the previous paragraphs can be questioned. We now review them and discuss how more realistic constraints can be accounted for.

1) *Per-Antenna Power Constraint*: A multiple-antenna transmitter radio architecture typically includes one power amplifier per antenna with an individual maximum output power constraint. In this case, the sum-power constraints $\text{tr}(\mathbf{R}) \leq P$ on transmit covariance matrices can be replaced by per-antenna power constraints $R_{i,i} \leq P/N$ where P is the total device power and N is the number of antennas. The problems addressed in this paper remain convex and the procedures proposed to solve them remain applicable.

2) *MIMO-Orthogonal Frequency Division Multiplexing (OFDM) Transmission*: If the system is wideband and employs OFDM, the equations must be modified to account for the parallel transmission on multiple subcarriers. A term $\mathcal{C}(\mathbf{R}, \mathbf{H})$ shall be replaced by a sum $\sum_{i=1}^{N_C} \mathcal{C}(\mathbf{R}_i, \mathbf{H}_i)$ where N_C is the number of subcarriers, and \mathbf{R}_i and \mathbf{H}_i are the signal covariance and MIMO channel matrices on the i th subcarrier. Likewise a

transmit power constraint shall now read $\sum_{i=1}^{N_C} \text{tr}(\mathbf{R}_i) \leq P$. Standards for wideband systems (e.g., [27]) typically impose a spectral mask that is almost flat in order to make the interference generated by the system as white as possible. This constraint can be translated into a per-subcarrier power constraint. Assuming a perfectly flat mask gives: $\text{tr}(\mathbf{R}_i) \leq P/N_C$. Per-antenna and per-subcarrier power constraints can also be applied simultaneously. Again, all these constraints result in convex problems that can be solved as presented in this paper.

3) *Imperfect CSI*: State-of-the-art TDD systems often rely on uplink/downlink channel reciprocity to enable CSIT exploitation without having to explicitly feed back the estimated channel coefficients. In this case, CSIT imperfection comes from the estimation noise and the variations of the channel between the moment it is estimated in one direction and the moment it is applied in the other direction. However, in cooperative relaying it is hard if not impossible to avoid explicit signaling because for instance the source cannot estimate the relay-destination channel by means of reciprocity. Quantized codebooks for MIMO systems [28] allow to reduce the amount of feedback. Therefore, in an actual system CSIT will be degraded by errors due to estimation, channel variations and limited-rate precoder and/or CSI feedback. Addressing these aspects in the half-duplex MIMO relay channel is out of the scope of this paper, but the expressions derived here can be taken as a performance upper-bound when evaluating the degradation introduced by these unavoidable imperfections.

IV. SIMULATION RESULTS

In this section, simulation results are presented for the TDD MIMO relay channel. The upper and lower bounds on capacity are evaluated and the suboptimality of the precoder optimization procedures in Section III-B-1) and Section III-B-2) is discussed.

The simulations below assume $N_S = 4$ antennas at the source and $N_R = N_D = 2$ antennas at the relay and destination. Such an antenna configuration is well suited to a cellular downlink scenario. The source and relay are subject to the same power constraint $P_S = P_R = 1$. The MIMO channel on the S-D, S-R, and R-D links is modeled by i.i.d. complex Gaussian components of respective variance γ_0, γ_1 and γ_2 . Therefore, under the above-defined power constraints, γ_0, γ_1 and γ_2 represent the average SNR on the S-D, S-R, and R-D links. The destination is far from both S and R, with $\gamma_0 = \gamma_2 = 0$ dB. The average SNR γ_1 on the S-R link is varied from 0 to 30 dB.

On Fig. 1, the average achievable rates of the partial and full DF strategies are plotted. For comparison purpose, the average capacity of the S-D link is also plotted with (solid line) or without (dashed line) CSIT. In this last case, the source covariance matrix that maximizes the ergodic capacity is [7] $\mathbf{R}_S = (P_S/N_S)\mathbf{I}_{N_S}$. The average capacity gain provided by CSIT can be decomposed into an array gain and a waterfilling gain [29]. Here since $N_S = 2N_D$ there is a 3 dB array gain plus a large waterfilling gain because γ_0 is low. It can be observed that single-hop transmission always outperforms noncooperative DF. This is obvious from inequality (47) and the fact that the capacity of the S-D link is larger than that of the R-D link. Both partial and full DF achieve a large rate increase over noncooperative DF and single-hop transmission,

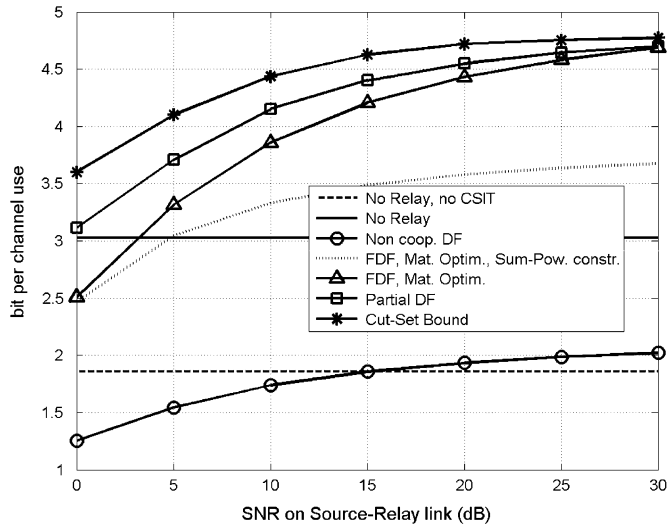


Fig. 1. Upper and lower bounds on TDD MIMO relaying channel capacity with Source and Relay precoder optimization in a $4 \times 2 \times 2$ antenna configuration.

and are less than 0.5 bit (per channel use) away from the cut-set bound for $\gamma_1 \geq 15$ dB. Partial DF outperforms FDF only at low γ_1 , when the rate becomes limited by the S-R link capacity. It can be argued that under the above simulation assumptions, the comparison with noncooperative DF and single-hop transmission is unfair since the total transmit power is larger for cooperative protocols during the second slot. Therefore, we also plot (dotted curve) the achievable rate of the FDF strategy when the sum-power is constrained to remain lower than P_S during the second slot. The figure shows that even in this case FDF outperforms noncooperative protocols at $\gamma_1 \geq 5$ dB.

On Fig. 2, various precoder optimization strategies for FDF are compared. The highest rate is achieved by matrix convex optimization of the source and relay precoders during both slots, using one of the procedures described in Section II-B. As stated in Section III-B-1), the vector optimization of the source precoder becomes optimum when N_S is large, but here it can be observed that the incurred loss is already small at $N_S = 4$ (only 0.1 bit). An additional rate penalty occurs at high γ_1 when the suboptimum source and relay precoder structure of Section III-B-2) is enforced during the second slot. Overall, the degradation due to suboptimum precoding is lower than 0.5 bit over the whole SNR range. Finally, the dashed and dotted curves illustrate the large rate loss when precoders are not optimized during the first slot (i.e., $\mathbf{R}_S^{(1)} = (P_S/N_S)\mathbf{I}_{N_S}$) or during both slots ($\mathbf{R}_S^{(2)} = \mathbf{R}_S^{(1)} = (P_S/N_S)\mathbf{I}_{N_S}$, $\mathbf{R}_R^{(2)} = (P_R/N_R)\mathbf{I}_{N_R}$).

On Fig. 3, the optimum power fraction \hat{P}_1 and the optimum time-sharing \hat{t} obtained by the suboptimum procedure of Section III-B-1) are plotted vs. γ_1 . As expected, when the capacity of the S-R link becomes much larger than that of the other links, most of the source transmit power during the first slot is assigned to the eigenmodes of \mathbf{H}_0 , and most of the time resource is assigned to the second slot in order to maximize R_{FDF} .

On Fig. 4, the achievable rate performance of the three TDD protocols defined in Section II-A-2) is compared. The SNR γ_2 on the R-D link is varied from 0 to 30 dB in an $N_S = N_R = N_D = 2$ antenna configuration with $\gamma_0 = 5$ and $\gamma_1 = 20$ dB. It can be observed that at high γ_2 , protocol III only provides a

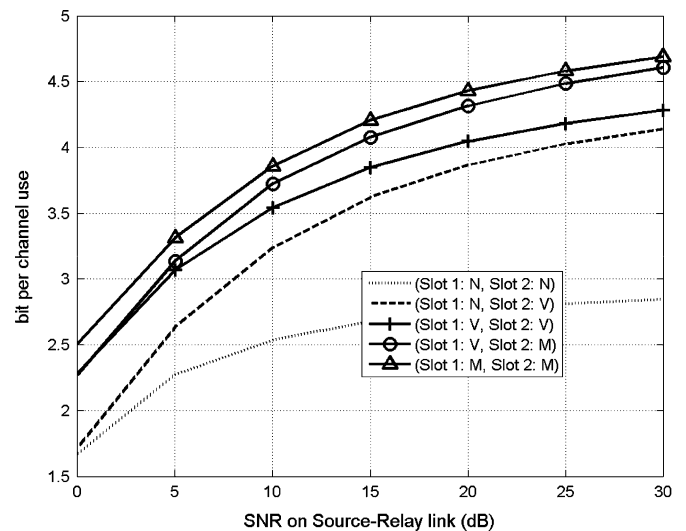


Fig. 2. Optimum vs. Sub-optimum precoder optimization for full DF strategy in a $4 \times 2 \times 2$ antenna configuration. (Legend: N = No optimization, V = Vector Optimization, M = Matrix Optimization).

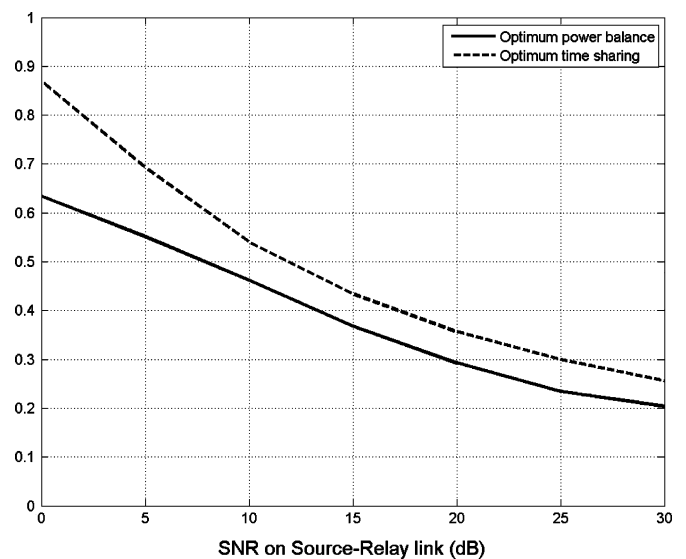


Fig. 3. Suboptimum source precoder during the first slot for full DF strategy: variation of the time-sharing and power balancing versus SNR on the Source-Relay link.

marginal gain w.r.t. NCDF. Indeed, from (44) and (46) it can be checked that when $\hat{C}(\mathbf{H}_2, P_R) \gg \hat{C}(\mathbf{H}_1, P_S)$ and $\hat{R}_{B,2} \gg \hat{C}(\mathbf{H}_1, P_S)$ then $R_{\text{FDF,III}} \approx R_{\text{NC,DF}} \approx \hat{C}(\mathbf{H}_1, P_S)$. At high γ_2 protocol III is outperformed by protocol II because the duration t of the first slot required for the relay to correctly decode the message cannot be made arbitrarily small due to the finite capacity of the S-R link, and only protocol II allows the destination to receive some information during this first slot. Protocol I enables the implementation of the partial DF strategy which yields a large rate increase over both protocols II and III.

Note that although in the above simulations a spatially uncorrelated channel model is assumed, the capacity bounds presented in this paper are derived without making any assumption on the channel statistics. It is however worth wondering how they are affected by spatial correlation. For a point-to-point

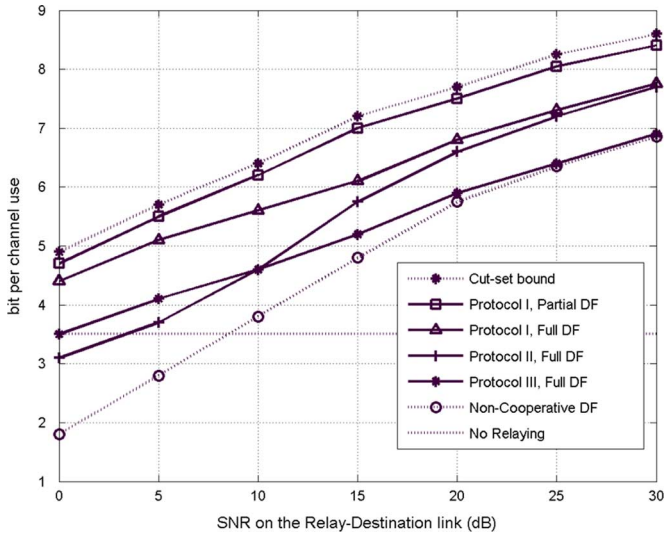


Fig. 4. Comparison of TDD Protocols I, II and III in a $2 \times 2 \times 2$ antenna configuration: achievable rate versus SNR on the R-D link for a 20-dB SNR on the S-R link and 5-dB SNR on the S-D link.

MIMO link, it is well known [30] that for a given SNR the capacity reduces as the spatial correlation increases. The same phenomenon occurs on the relay channel: for instance in (9) the left-hand term in the $\min \{ \}$ decreases as \mathbf{H}_0 and \mathbf{H}_1 become spatially correlated and the right-hand term decreases as \mathbf{H}_0 and \mathbf{H}_2 become spatially correlated.

V. CONCLUSION

We presented a generic methodology to maximize capacity upper and lower bounds on the MIMO relay channel with full CSI in the full-duplex and TDD relaying cases. The optimum source and relay transmit covariance matrices and TDD time-sharing parameter can be derived by convex optimization, and the gap between the achievable rate and the capacity can be quantified for various DF strategies. In particular, we verified that for realistic antenna configurations and SNR ranges this gap can actually be quite small. Our optimization procedure illustrates the application of several mathematical tools borrowed from convex optimization theory, nonlinear programming, as well as from complex matrix differentiation to the practical problem of precoding for the MIMO relay channel. As pointed out in this paper, the bounds computed here can serve as a benchmark when studying relaying strategies accounting for CSI imperfections.

APPENDIX

A. Gradient of a Real-Valued Function of Hermitian-Symmetric Complex Matrices

The differentiation of the objective and constraints w.r.t. the real variable ε is straightforward. The differentiation w.r.t. the real variables t_1 and t_2 requires to compute

$$\frac{\partial (t\mathcal{C}(\mathbf{Q}/t, \mathbf{H}))}{\partial t} = \mathcal{C}(\mathbf{Q}/t) - \frac{1}{t} \text{Tr} \left(\mathbf{H} \left(\mathbf{I}_N + \frac{1}{t} \mathbf{Q} \mathbf{H}^H \mathbf{H} \right)^{-1} \mathbf{Q} \mathbf{H}^H \right). \quad (59)$$

The remainder of this Appendix addresses the differentiation w.r.t. matrices $\mathbf{Q}_S^{(1)}$ and $\mathbf{Q}_{SR}^{(2)}$. Using similar notations as [22, example 5], an $N \times N$ Hermitian-symmetric matrix \mathbf{X} can be generated by the following so-called ‘‘pattern producing’’ function:

$$\mathcal{F} : \mathbb{R}^N \times \mathbb{C}^{N(N-1)/2} \times \mathbb{C}^{N(N-1)/2} \rightarrow \mathbb{C}^{N \times N} \quad (60)$$

$$(\mathbf{r}, \mathbf{c}, \mathbf{c}^*) \mapsto \text{mat}(\mathbf{L}_d \mathbf{r} + \mathbf{L}_l \mathbf{c} + \mathbf{L}_u \mathbf{c}^*) \triangleq \mathbf{X}$$

where \mathbf{L}_d is an $N^2 \times N$ matrix that maps the N independent components of the real vector \mathbf{r} onto the diagonal of \mathbf{X} , \mathbf{L}_l (respectively, \mathbf{L}_u) is an $N^2 \times (N(N-1)/2)$ matrix that maps the $N(N-1)/2$ independent components of \mathbf{c} (respectively, \mathbf{c}^*) onto the lower-triangular (respectively, upper-triangular) part of \mathbf{X} . A convenient way of writing partial derivatives of matrix-valued functions of matrices is introduced in [21]. With this notation, the partial derivatives of \mathcal{F} read

$$\mathcal{D}_r \mathcal{F} = \mathbf{L}_d; \mathcal{D}_c \mathcal{F} = \mathbf{L}_l; \mathcal{D}_{c^*} \mathcal{F} = \mathbf{L}_u. \quad (61)$$

The partial derivatives involved in our optimization problem can now be computed using [22, eqs. (15), (16), and (17)]:

$$\mathcal{D}_r \mathcal{C}(\mathbf{X}, \mathbf{H}) = \mathcal{D}_{\tilde{\mathbf{X}}} \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})|_{\tilde{\mathbf{X}}=\mathbf{X}} \mathcal{D}_r \mathcal{F}(\mathbf{r}, \mathbf{c}, \mathbf{c}^*) + \mathcal{D}_{\tilde{\mathbf{X}}^*} \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})|_{\tilde{\mathbf{X}}=\mathbf{X}} \mathcal{D}_r \mathcal{F}(\mathbf{r}, \mathbf{c}, \mathbf{c}^*) \quad (62)$$

where $\tilde{\mathbf{X}}$ is the extension of \mathbf{X} to the set of unpatterned matrices. For unpatterned matrices, the partial derivatives of \mathcal{C} are well known [21] and equal to

$$\frac{\partial \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})}{\partial \tilde{\mathbf{X}}} = (\mathbf{H}^H (\mathbf{I}_N + \mathbf{H} \tilde{\mathbf{X}} \mathbf{H}^H)^{-1} \mathbf{H})^T$$

$$\frac{\partial \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})}{\partial \tilde{\mathbf{X}}^*} = \mathbf{0}_{N \times N}. \quad (63)$$

Therefore

$$\mathcal{D}_{\tilde{\mathbf{X}}} \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})|_{\tilde{\mathbf{X}}=\mathbf{X}} \triangleq \left(\text{vec} \left(\frac{\partial \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})}{\partial \tilde{\mathbf{X}}} \Big|_{\tilde{\mathbf{X}}=\mathbf{X}} \right) \right)^T = (\text{vec}((\mathbf{H}^H (\mathbf{I} + \mathbf{H} \mathbf{X} \mathbf{H}^H)^{-1} \mathbf{H})^T))^T \quad (64)$$

and

$$\mathcal{D}_{\tilde{\mathbf{X}}^*} \mathcal{C}(\tilde{\mathbf{X}}, \mathbf{H})|_{\tilde{\mathbf{X}}=\mathbf{X}} = \mathbf{0}_{1 \times N^2}. \quad (65)$$

Inserting (61), (64) and (65) into (62) yields

$$\mathcal{D}_r \mathcal{C}(\mathbf{X}, \mathbf{H}) = (\text{vec}((\mathbf{H}^H (\mathbf{I} + \mathbf{H} \mathbf{X} \mathbf{H}^H)^{-1} \mathbf{H})^T))^T \mathbf{L}_d. \quad (66)$$

Likewise

$$\mathcal{D}_{c^*} \mathcal{C}(\mathbf{X}, \mathbf{H}) = (\text{vec}((\mathbf{H}^H (\mathbf{I} + \mathbf{H} \mathbf{X} \mathbf{H}^H)^{-1} \mathbf{H})^T))^T \mathbf{L}_u. \quad (67)$$

Finally, the gradient of \mathcal{C} is given by [22, Theorem 2]

$$\nabla \mathcal{C} = [\mathcal{D}_r \mathcal{C}(\mathbf{X}, \mathbf{H}) \quad 2\mathcal{D}_{c^*} \mathcal{C}(\mathbf{X}, \mathbf{H})]^T. \quad (68)$$

The partial derivatives of the trace functions that arise from transmit power constraints in our problems can be computed

as in (62)–(67) using the well-known [21] unpatterned partial derivatives

$$\frac{\partial \text{Tr}(\mathbf{A}\tilde{\mathbf{X}}\mathbf{B})}{\partial \tilde{\mathbf{X}}} = (\mathbf{A}\mathbf{B})^T. \quad (69)$$

Note that because the mapping \mathcal{F} is linear, it preserves convexity. However, in (9) and (12), matrices are not only Hermitian but also PSD. In [22], the Cholesky decomposition $\mathbf{Q} = \mathbf{L}\mathbf{L}^H$ is proposed to parameterize PSD matrices. Unfortunately in this case, the mapping is nonlinear and a function which is convex in \mathbf{Q} may be nonconvex in \mathbf{L} . This is the reason why in the algorithms proposed in this paper the Hermitian symmetry is guaranteed by the pattern-producing function, but the positive semidefiniteness is enforced by either gradient projection or a barrier function.

ACKNOWLEDGMENT

The authors would like to thank A. Hjørungnes and D. Palomar for fruitful discussions on differentiation techniques and E. Calvo for suggesting the use of the perspective function.

REFERENCES

- [1] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, Sep. 1979.
- [2] *Air Interface for Fixed and Mobile Broadband Wireless Access Systems—Advanced Air Interface*, IEEE 802.16m Task Group, PAR approved by the IEEE-SA Standards Board, Dec. 6, 2006.
- [3] Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, Oct. 2004.
- [4] F. Xue and S. Sandhu, "Cooperation in a half-duplex Gaussian diamond relay channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, Oct. 2007.
- [5] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun. (JSAC)*, vol. 25, no. 2, Feb. 2007.
- [6] I. Hammerstrom, M. Kuhn, C. Esli, Z. Jihan, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2007.
- [7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, Sep. 2005.
- [8] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, Nov./Dec. 1999.
- [9] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, Sep. 2003.
- [10] W. Tong *et al.*, "Duplex and multiplex configurations for OFDMA in-band relay," IEEE 802.16 Doc nb. C802.16mmr-05/011 MMR Study Group Sep. 2005.
- [11] M. Yuksel and E. Erkip, "Multiple-antenna cooperative wireless systems: A diversity-multiplexing tradeoff perspective," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, Oct. 2007.
- [12] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.
- [13] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, Jan. 2005.
- [14] C. K. Lo, S. Vishwanath, and R. W. Heath, Jr., "Rate bounds for MIMO relay channels using precoding," *Proc. IEEE Globecom'05* Nov. 2005.
- [15] K. Azarian, H. ElGamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, Dec. 2005.
- [16] O. Muñoz, J. Vidal, and A. Agustin, "Linear transceiver design in non-regenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, Jun. 2007.

- [17] I. Hammerstrom and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2798–2802, Aug. 2007.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004 [Online]. Available: <http://www.stanford.edu/~boyd/cvxbook/>, Avail.
- [19] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. New York: Athena Scientific, 1999.
- [20] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun. (JSAC)*, vol. 22, no. 6, Aug. 2004.
- [21] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pt. I, pp. 2740–2746, Jun. 2007.
- [22] A. Hjørungnes and D. Palomar, "Patterned complex-valued matrix derivatives," in *Proc. Fifth IEEE Workshop on Sensor Array Multi-Channel Signal Process.*, Darmstadt, Germany, Jul. 2008 [Online]. Available: <http://brage.unik.no/personer/arehj/>
- [23] S. Simoens, O. Muñoz, J. Vidal, and A. Del Coso, "Capacity bounds for Gaussian MIMO relay channel with full channel state information," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Recife, Brazil, Jul. 2008 [Online]. Available: http://sebastien.simoens.free.fr/publications_simoens.html
- [24] S. Ye and R. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, Nov. 2003.
- [25] M. J. Todd, "Semidefinite Optimization," in *Acta Numerica*. Cambridge, U.K.: Cambridge Univ. Press, 2001, vol. 10, pp. 515–560.
- [26] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, Nov. 2006.
- [27] *Amendment to IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems—Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE Std 802.16e-2005, Feb. 2006.
- [28] D. J. Love and R. W. Heath, Jr., "Multimode precoding for MIMO wireless systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, Oct. 2005.
- [29] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [30] D. Gesbert *et al.*, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, Apr. 2003.



Sebastien Simoens graduated from Ecole Nationale Supérieure des Télécommunications (ENST) Paris, France, in 1998.

Until 2008, he was with Motorola Labs Paris where he conducted research on signal processing for broadband wireless communications. He was involved in several European projects including IST-FIREWORKS and ICT-ROCKET. Since September 2005, he has been pursuing the Ph.D. degree with the Signal Theory and Communications Department, Technical University of Catalonia (UPC), Barcelona, Spain. Since November 2008, he has been with Thales Aerospace Division, Valence, France, working on signal processing for inertial navigation sensors.



Olga Muñoz-Medina received the M.S. degree in 1993 and the Ph.D. degree in 1998, both in electrical engineering, from the Universitat Politècnica de Catalunya (UPC), Spain.

In 1994, she joined the Department of Signal Theory and Communications of the same University and became an Associate Professor in 2001. She teaches graduate and undergraduate courses related to signal processing. She has participated in industrial projects as Spread Spectrum Link for the Automatic Transfer Vehicle (ATV) with Alcatel Espacio S.A. for European Spatial Agency, and in European Commission-funded projects as TSUNAMI, ROMANTIK, FIREWORKS, and ROCKET. Her current research interests include physical layer aware MAC techniques, cooperative transmission for multihop systems, and multibase coordination.



Josep Vidal (M'91) received the Telecommunication Engineering and Ph.D. degrees from the Universitat Politècnica de Catalunya (UPC), Barcelona.

From 1989 to 1990, he joined the LTS of the Ecole Polytechnique de Lausanne as an Associate Researcher. He was awarded Premio Extraordinario de Doctorado in 1996. Since 1993, he has been teaching different topics in Telecommunication Engineering at UPC. His current research interests are in statistical signal processing, information and communication theory, where he has authored more than 100 journal and conference papers. Since 2000, has led UPC participation in the EC-funded projects, SATURN, ROMANTIK, FIREWORKS and ROCKET, as project coordinator in some of them. He has held research appointments with INP Toulouse and the University of Hawaii.



Aitor del Coso (M'08) received the M.Sc. degree in telecommunications from the Universidad Politécnica de Madrid (UPM), Spain, and the Ph.D. degree in signal theory and communications from the Universitat Politècnica de Catalunya (UPC), in 2003 and 2008, respectively. He conducted his Ph.D. studies at the Access Technologies area of the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain, from 2004 to 2008.

He also held visiting positions with the Politecnico de Milano, Italy, in 2005; New Jersey Institute of Technology (NJIT), in 2006; and Motorola Research Labs, France, in 2007. Currently, he is with the Multimedia Telecommunications Systems Group of Thales Alenia Space Espana, Madrid. His research interests lie within the fields of wireless and satellite communications, communication theory, signal processing, and information theory.