Transmit Signal and Bandwidth Optimization in Multiple-Antenna Relay Channels

Chris T. K. Ng and Gerard J. Foschini
Bell Laboratories, Alcatel-Lucent, Holmdel, NJ 07733
Email: {Chris.Ng, Gerard.Foschini}@alcatel-lucent.com

Abstract

Transmit signal and bandwidth optimization is considered in multiple-antenna relay channels. Assuming all terminals have channel state information, the cut-set capacity upper bound and decode-and-forward rate under full-duplex relaying are evaluated by formulating them as convex optimization problems. For half-duplex relays, bandwidth allocation and transmit signals are optimized jointly. Moreover, achievable rates based on the compress-and-forward transmission strategy are presented using rate–distortion and Wyner–Ziv compression schemes. It is observed that when the relay is close to the source, decode-and-forward is almost optimal, whereas compress-and-forward achieves good performance when the relay is close to the destination.

Index Terms

Bandwidth allocation, convex optimization, relay channels, multiple-antenna, transmit covariance matrices.

I. INTRODUCTION

In wireless communications, the transmission rate is fundamentally limited by the channel propagation loss over the range of transmission. Relaying has been proposed as a transmission strategy that can improve the performance of wireless systems. In a relay channel, in addition to the source and the destination, there is also a relay terminal. The relay does not have its own data to send or receive; its intention is to facilitate the transmission between the source and destination. On one hand, the relay brings additional power to the network, as typically, the relay is under its own power source. On the other hand, a relay can also help to shorten the transmission range by enabling communication over two hops, and cooperate with the source to perform joint encoding of the transmit signals. However, if the relay transmission scheme is not designed properly, the relay may also create undesirable interference to the terminal at the destination. In this paper, we investigate the optimization of transmit signals in relay channels. In particular, we consider a multiple-input multiple-output (MIMO) relay channel, where the source, relay, and destination terminal all have multiple antennas. In wireless communications, using multiple transmit and receive antennas has been shown to provide substantial improvement in channel capacity [1], [2]. We also evaluate the MIMO relay channel rates under different network geometry, and investigate the effectiveness of the corresponding relaying schemes.

The three-node relay channel model is proposed in [3]. In [4], a capacity upper bound and achievable coding strategies are presented for the relay channel, but the relay channel capacity remains an open problem. For Gaussian single-antenna relay channels, capacity bounds and power allocation are studied in [5]. Capacity bounds on half-duplex relaying are presented in [6], [7]. Bandwidth and power allocation are considered in [8], [9] for fading orthogonal relay channels, and in [10] for the amplify-and-forward scheme in Gaussian relay networks. Transmitter cooperation versus receiver cooperation in relay channels is compared in [11]. Relay channel coding strategies, with extensions to relay channels with multiple terminals, are given in [12]. For relay channels with multiple-antenna terminals, bounds to the cut-set capacity upper bound and decode-and-forward rate are considered in [13], [14]. In [15], [16], the diversity-multiplexing tradeoff is characterized for full-duplex and half-duplex MIMO relay channels.

In this paper, we consider a multiple-antenna relay channel where the terminals have knowledge of the channel state information (CSI). We consider optimization of the source and relay transmit signals to evaluate the MIMO cut-set capacity upper bound and the decode-and-forward achievable rate, by formulating them as convex optimization problems. In the case of half-duplex relaying, the bandwidth allocation and the multiple-antenna transmit signals are optimized jointly. We also present achievable rates in MIMO relay channels using the compress-and-forward approach, under which the rate–distortion and Wyner–Ziv compression schemes are considered.
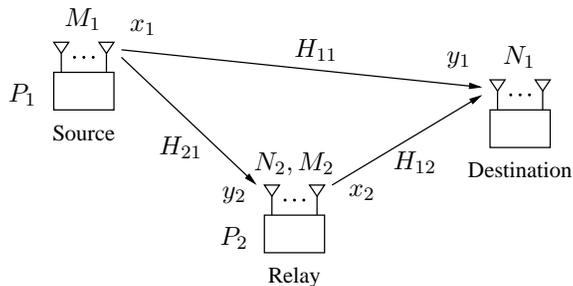
Fig. 1.   Multiple-antenna relay channel.

The remainder of this paper is organized as follows. Section II presents the multiple-antenna relay channel model, and the capacity upper and lower bounds. The cut-set bound and decode-and-forward rate optimization formulations are described in Section III under full-duplex assumptions, while Section IV considers half-duplex relaying where the relay cannot simultaneously transmit and receive in the same frequency band. The compress-and-forward transmission strategy is studied in Section V, and Section VI concludes the paper.

Notation: In this paper, $\mathbb{R}$ ($\mathbb{R}_+$, $\mathbb{R}_{++}$) is the set of real numbers (nonnegative, positive real numbers), $\mathbb{C}$ is the complex field, $\mathbb{1}$ denotes the two-element set $\{0, 1\}$. Dimensions of vectors/matrices are indicated by superscripts. $\mathbb{H}_+^N$ is the set of $N \times N$ positive semidefinite Hermitian matrices. $X \succ Y$ ($X \succeq Y$) means the matrix $X - Y$ is positive (semi)definite. I_N is the $N \times N$ identity matrix. A^T and A^H are the transpose and conjugate transpose, respectively, of a matrix A. The operators $\mathrm{E}[\,\cdot\,]$, $\det$, $\mathbf{tr}$ denote, respectively, expectation, determinant and trace. For random variables, $x \sim \mathcal{CN}(\mu, Q)$, where $x, \mu \in \mathbb{C}^N$, $Q \in \mathbb{H}_+^N$, means that x is a circularly symmetric complex Gaussian random N-vector about mean μ with covariance matrix Q.

II. SYSTEM MODEL

A. Channel Model

Consider a three-node wireless relay channel as illustrated in Fig. 1. The source node wishes to send a message to the destination; the relay node does not have its own message to send, but facilitates the transmission between the source and destination. Suppose the source has M_1 transmit antennas, and the destination has N_1 receive antennas. We assume the relay has M_2 transmit antennas and N_2 receive antennas (for instance, in full-duplex operation the relay may have different sets of transmit and receive antennas). We consider a discrete-time flat-fading channel model, which is described by

$$y_1 = H_{11}x_1 + H_{12}x_2 + z_1 \tag{1}$$

$$y_2 = H_{21}x_1 + z_2 \tag{2}$$

where $x_1 \in \mathbb{C}^{M_1}$, $x_2 \in \mathbb{C}^{M_2}$ are the respective transmit signals of the source and relay; $y_1 \in \mathbb{C}^{N_1}$, $y_2 \in \mathbb{C}^{N_2}$ are the respective receive signals of the destination and relay; and $z_1 \sim \mathcal{CN}(0, I_{N_1}) \in \mathbb{C}^{N_1}$, $z_2 \sim \mathcal{CN}(0, I_{N_2}) \in \mathbb{C}^{N_2}$ are independent zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise at the destination and relay, respectively. The complex baseband channel from the source to destination is $H_{11} \in \mathbb{C}^{N_1 \times M_1}$; from the source to relay is $H_{21} \in \mathbb{C}^{N_2 \times M_1}$; and from the relay to destination is $H_{12} \in \mathbb{C}^{N_1 \times M_2}$.

We consider a block-fading channel model: the channels realize independently according to their distribution at the beginning of each fading block, and they remain unchanged within the duration of the fading block. In this paper, we assume the channel states can be estimated accurately and conveyed timely to all terminals: i.e., we assume channel state information (CSI) is available at all nodes. The source and the relay are under the respective transmit power constraints: $\mathrm{E}[x_i^H x_i] \leq P_i$, $i = 1, 2$, where the expectations are over repeated channel uses within every fading block. Power allocation across fading blocks is not permitted. The transmit signals have zero mean: $\mathrm{E}[x_1] = 0$, $\mathrm{E}[x_2] = 0$.

It is convenient to write (1), (2) in block-matrix form

$$y = Hx + z \tag{3}$$

where

$$y \triangleq \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{C}^N, \qquad H \triangleq \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & 0 \end{bmatrix} \in \mathbb{C}^{N \times M}, \qquad x \triangleq \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^M, \qquad z \triangleq \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \mathbb{C}^N \qquad (4)$$

with $M \triangleq M_1 + M_2$, and $N \triangleq N_1 + N_2$. Moreover, we denote the joint covariance matrix of the transmit signals of the source and relay as

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \triangleq Q \triangleq \mathrm{E}[xx^H] \in \mathbb{H}_+^M \qquad (5)$$

where the conformally partitioned blocks (with respect to x_1, x_2) of Q have dimensions

$$Q_{11} \triangleq \mathrm{E}[x_1 x_1^H] \in \mathbb{H}_+^{M_1}, \qquad Q_{22} \triangleq \mathrm{E}[x_2 x_2^H] \in \mathbb{H}_+^{M_2}, \qquad Q_{12} \triangleq \mathrm{E}[x_1 x_2^H] = Q_{21}^H \in \mathbb{C}^{M_1 \times M_2}. \qquad (6)$$

B. Capacity Bounds and Achievable Rates

The capacity of a relay channel is, in general, an open problem; however, there are known upper and lower bounds. The cut-set bound described in [4], [17] provides an upper bound to the relay channel capacity. Intuitively, the cut-set bound states that, over any possible joint source-relay transmit signals, the relay channel capacity cannot exceed the smaller of i) the maximum rate at which information can flow out of the source, and ii) the maximum rate at which information can flow into the destination.

On the other hand, capacity lower bounds of the relay channel, as achieved by two coding strategies, are given in [4]. In the decode-and-forward strategy [4, Thm. 1], transmission is done in blocks: the relay first fully decodes the message from the source in one block, then in the ensuing block, the relay and the source cooperatively transmit the message to the destination. In the compress-and-forward strategy [4, Thm. 6], the relay does not decode the source's message but sends a compressed version of its observed signal to the destination. The destination then combines the compressed signal with its own receive signal to decode the source's message. A detailed discussion on the relay channel coding strategies can be found in [12].

In the following sections, we evaluate these capacity bounds and achievable rates for full- and half-duplex Gaussian MIMO relay channels. Under half-duplex relaying in Section IV, we impose the constraint that the relay cannot simultaneously transmit and receive in the same frequency band. For performance comparisons, we also consider: direct transmission when the relay is not available (i.e., the capacity of an $M_1 \times N_1$ MIMO channel under transmit power constraint P_1); orthogonal two-hop relaying (Section IV-C); and the scenario when the relay is co-located with the source or destination (Section IV-D).

III. MIMO RELAY CHANNEL CAPACITY BOUNDS

In this section, we present the optimization frameworks for evaluating the MIMO relay channel cut-set capacity upper bound and the decode-and-forward achievable rate. We first adopt the full-duplex assumption where the relay can transmit and receive in the same frequency band at the same time. In practice, full-duplex transmission is difficult to realize. Nevertheless, the full-duplex model provides insight into the design of effective coding strategies for relay channels, and serves as a performance upper bound for half-duplex systems. Half-duplex relaying is considered in Section IV where the relay cannot simultaneously transmit and receive in the same band.

A. Cut-Set Capacity Upper Bound

The cut-set capacity upper bound [4] for the relay channel is given as an optimization in terms of the channel mutual information as follows:

$$R_{\mathrm{CS}} = \max_{p(x_1, x_2)} \min\{I(x_1; y_1, y_2 | x_2), \ I(x_1, x_2; y_1)\} \qquad (7)$$

$$= \max_{p(x)} \min\{I(x_1; y | x_2), \ I(x; y_1)\}. \qquad (8)$$

Gaussian signals are optimal in the cut-set bound and decode-and-forward rate [12, Proposition 2]. We denote the transmit signals by: $x \sim \mathcal{CN}(0, Q)$, where $Q \in \mathbb{H}_+^M$ is the covariance matrix of x. Then the mutual information expressions in (8) evaluate [2] to

$$R_{\mathrm{CS}} = \max_{Q \, : \, \mathbf{tr}\, Q_{ii} \leq P_i, \, i=1,2} \min\big\{\log \det(I_N + H_1 Q_{1|2} H_1^H), \log \det(I_{N_1} + \tilde{H}_1 Q \tilde{H}_1^H)\big\} \tag{9}$$

where H_1 and $\tilde{H}_1$ are, respectively, the first block column and block row of H

$$H_1 \triangleq \begin{bmatrix} H_{11} \\ H_{21} \end{bmatrix} \in \mathbb{C}^{N \times M_1}, \qquad\qquad \tilde{H}_1 \triangleq \begin{bmatrix} H_{11} & H_{12} \end{bmatrix} \in \mathbb{C}^{N_1 \times M} \tag{10}$$

and the conditional covariance matrix $Q_{1|2} \triangleq \mathrm{E}[x_1 x_1^H | x_2]$ is given by the Schur complement of Q_{22} in Q

$$Q_{1|2} = Q_{11} - Q_{12} Q_{22}^{-1} Q_{21} \tag{11}$$

where we assume $Q_{22} \succ 0$. The zero-mean Gaussian signal x is fully characterized by its covariance; therefore, in (9), the sole optimization variable is the joint covariance matrix Q.

The cut-set bound maximization in (9) can be formulated as the following optimization problem:

$$\text{maximize} \quad R_{\mathrm{CS}} \tag{12}$$
$$\text{over} \quad R_{\mathrm{CS}} \in \mathbb{R}_+, \; Q \in \mathbb{H}_+^M, \; Q_{1|2} \in \mathbb{H}_+^{M_1} \tag{13}$$
$$\text{subject to} \quad R_{\mathrm{CS}} \leq \log \det(I_N + H_1 Q_{1|2} H_1^H) \tag{14}$$
$$R_{\mathrm{CS}} \leq \log \det(I_{N_1} + \tilde{H}_1 Q \tilde{H}_1^H) \tag{15}$$
$$\mathbf{tr}(C_1^T Q C_1) \leq P_1 \tag{16}$$
$$\mathbf{tr}(C_2^T Q C_2) \leq P_2 \tag{17}$$
$$Q - C_1 Q_{1|2} C_1^T \succeq 0 \tag{18}$$

where C_1, C_2 are constant matrices defined as

$$C_1 \triangleq \begin{bmatrix} I_{M_1} \\ 0 \end{bmatrix} \in \mathbb{1}^{M \times M_1}, \qquad\qquad C_2 \triangleq \begin{bmatrix} 0 \\ I_{M_2} \end{bmatrix} \in \mathbb{1}^{M \times M_2}. \tag{19}$$

In the optimization, (14), (15) follow from the two terms inside the min expression in (9); and (16), (17) represent the per-node transmit power constraints at the source and relay, respectively. The constraint (18) results from relaxing the equality constraint in (11)

$$Q_{1|2} = Q_{11} - Q_{12} Q_{22}^{-1} Q_{21} \implies Q_{1|2} \preceq Q_{11} - Q_{12} Q_{22}^{-1} Q_{21}. \tag{20}$$

By the semidefiniteness property of Schur complements [18], we have the following identity on the right-hand side of (20):

$$(Q_{11} - Q_{1|2}) - Q_{12} Q_{22}^{-1} Q_{21} \succeq 0 \iff \begin{bmatrix} (Q_{11} - Q_{1|2}) & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \succeq 0 \tag{21}$$

where the right-hand side of (21) is equivalent to $Q - C_1 Q_{1|2} C_1^T \succeq 0$ when written in the block-matrix form as defined in (5). Finally, we show the relaxation in (20) does not increase the optimal value in (12). Suppose given a set of fixed $Q_{11}, Q_{12}, Q_{21}, Q_{22}$, we consider all $X \in \mathbb{H}_+^{M_1}$ such that $X \preceq Q_{11} - Q_{12} Q_{22}^{-1} Q_{21}$. Recalling that the determinant is matrix increasing [19] on the set of positive semidefinite matrices, we get

$$\log \det(I_N + H_1 X H_1^H) \leq \log \det\big(I_N + H_1(Q_{11} - Q_{12} Q_{22}^{-1} Q_{21}) H_1^H\big) \tag{22}$$

which only limits the feasible set in (14).

The maximization in (12)–(18) is a convex optimization problem; in particular, the log-determinant function is concave on positive definite matrices [19]. The solution of (12) can be efficiently computed using standard convex optimization numerical techniques, for instance, by the interior-point method [19], [20]. The above optimization can also be solved by the CVX [21], [22] software package, which uses a successive approximation approach to model the log-determinant inequalities. All optimization formulations presented in this paper are convex problems, unless otherwise noted.

In addition to the per-node power constraints (16), (17), if the source and relay are also under per-antenna power constraints, they can be readily incorporated in the convex optimization formulation. Let the antenna power constraints of the source and relay, respectively, be $p_{1,1}, \ldots, p_{1,M_1}$ and $p_{2,1}, \ldots, p_{1,M_2}$. The per-antenna power constraints are represented by

$$q_{i,i} \leq p_{1,i}, \quad i = 1, \ldots, M_1, \qquad q_{M_1+j,M_1+j} \leq p_{2,j}, \quad j = 1, \ldots, M_2 \qquad (23)$$

where $q_{i,i}$ is the (i,i) entry of the covariance matrix Q, with $i = 1, \ldots, M$.

B. Decode-and-Forward Achievable Rate

The decode-and-forward [4, Thm. 1] relay channel achievable rate is given by

$$R_{\mathrm{DF}} = \max_{p(x_1, x_2)} \min\{ I(x_1; y_2 | x_2), I(x_1, x_2; y_1) \} \qquad (24)$$

$$= \max_{p(x)} \min\{ I(x_1; y_2 | x_2), I(x; y_1) \} \qquad (25)$$

$$= \max_{Q \,:\, \mathbf{tr}\, Q_{ii} \leq P_i, \; i=1,2} \min\{ \log\det(I_{N_2} + H_{21} Q_{1|2} H_{21}^H), \log\det(I_{N_1} + \tilde{H}_1 Q \tilde{H}_1^H) \} \qquad (26)$$

where $Q_{1|2}$ is as given in (11). The decode-and-forward rate can be formulated as the solution to the following convex optimization problem:

$$\begin{aligned} \text{maximize} \quad & R_{\mathrm{DF}} & (27) \\ \text{over} \quad & R_{\mathrm{DF}} \in \mathbb{R}_+, \; Q \in \mathbb{H}_+^M, \; Q_{1|2} \in \mathbb{H}_+^{M_1} & (28) \\ \text{subject to} \quad & R_{\mathrm{DF}} \leq \log\det(I_{N_2} + H_{21} Q_{1|2} H_{21}^H) & (29) \\ & R_{\mathrm{DF}} \leq \log\det(I_{N_1} + \tilde{H}_1 Q \tilde{H}_1^H) & (30) \\ & \mathbf{tr}(C_1^T Q C_1) \leq P_1 & (31) \\ & \mathbf{tr}(C_2^T Q C_2) \leq P_2 & (32) \\ & Q - C_1 Q_{1|2} C_1^T \succeq 0. & (33) \end{aligned}$$

The derivations for the optimization problem formulation parallel those presented in the previous section. Note that the decode-and-forward optimization problem (27)–(33) is similar to the cut-set bound optimization problem (12)–(18). The only difference is the aggregate channel H_1 in (14) versus the source-relay channel H_{21} in (29). Consequently, we expect the decode-and-forward scheme is nearly capacity-achieving when the direct channel H_{11} is weak relative to the source-relay channel H_{21}.

C. Numerical Results

In the numerical examples in this paper, we assume a network geometry as depicted in Fig. 2. In the two-dimensional network, the source is located at coordinates $(0,0)$, the destination is at $(1,0)$, and the relay is at (d_x, d_y). We will use a distance-based path-loss power attenuation exponent $\eta = 4$, combined with independent and identically distributed (i.i.d.) Rayleigh fading for each channel matrix entry. Shadow fading can also be included, but it is omitted here to allow a simple geometric interpretation of the relay network topology. The relay channel matrices are given as

$$H_{11} = H_w^{(1)}, \qquad H_{21} = (d_x^2 + d_y^2)^{-\eta/4} H_w^{(2)}, \qquad H_{12} = \left((1 - d_x)^2 + d_y^2\right)^{-\eta/4} H_w^{(3)} \qquad (34)$$

where $H_w^{(1)} \in \mathbb{C}^{N_1 \times M_1}$, $H_w^{(2)} \in \mathbb{C}^{N_2 \times M_1}$, $H_w^{(3)} \in \mathbb{C}^{N_1 \times M_3}$, with each entry of $H_w^{(1)}$, $H_w^{(2)}$, $H_w^{(3)}$ i.i.d. $\sim \mathcal{CN}(0,1)$. For the numerical experiments presented in this paper, 50 random instances of the channel realizations $H_w^{(1)}$, $H_w^{(2)}$, $H_w^{(3)}$ are generated. Then under each channel realization, the corresponding optimization problems are solved to evaluate the relay channel capacity bounds and achievable rates. For consistent comparison, the same set of channel realizations is used to compute the performance of the different coding schemes under consideration. The convex optimization problems are solved using the barrier interior-point algorithm described in [19, Section 11.3].

The empirical cumulative distribution functions (CDFs) of the cut-set (CS) bound and decode-and-forward (DF) rate of a MIMO relay channel, where the relay is located at $(d_x, d_y) = (1/3, 1/2)$, are shown in Fig. 3 (solid lines).
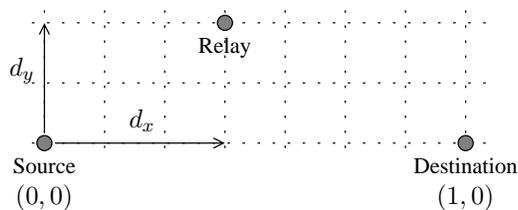
Fig. 2. Geometry of the source, relay, and destination nodes.
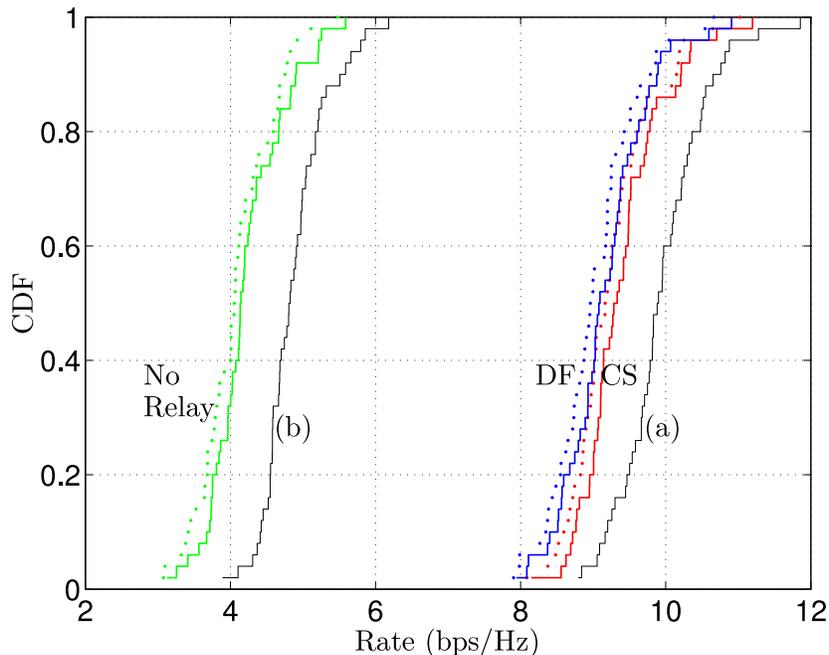

Fig. 3. Full-duplex CDFs of MIMO relay channel cut-set (CS) bound and decode-and-forward (DF) rate. The relay is located at $(1/3, 1/2)$. All terminals have four antennas: $M_1 = N_1 = M_2 = N_2 = 4$, and unit power constraints: $P_1 = P_2 = 0$ dB. The dotted lines represent the rates under per-antenna power constraints. The plots (a) and (b) correspond to the upper and lower bounds, respectively, from [13, Thms. 3.1 and 3.2].

All terminals have four antennas: $M_1 = N_1 = M_2 = N_2 = 4$, and unit power constraints: $P_1 = P_2 = 0$ dB. The dotted lines represent the rates under per-antenna power constraints, where all antennas have equal constraints: $p_{1,1} = \cdots = p_{1,M_1} = P_1/M_1$, and $p_{2,1} = \cdots = p_{2,M_2} = P_2/M_2$. Also shown in the plot is the MIMO capacity of the direct channel H_{11} when the relay node is not available (No Relay). The decode-and-forward achievable rate considerably outperforms the direct channel capacity and is quite close to the cut-set capacity upper bound. For all coding schemes, imposing per-antenna power constraints only slightly reduces the rates as compared to per-node power constraints. For comparison, the upper and lower bounds from [13, Thms. 3.1 and 3.2] are plotted and labeled (a) and (b), respectively. It is observed that the capacity upper and lower bounds can be tightened when the transmit signals of the source and relay are optimized. The upper bound (a) is computed by searching over combinations of $\rho = \{0, 0.05, \ldots, 0.9, 0.95\}$ and $a = \{10^{-1}, 10^{-0.9}, \ldots, 10^{0.9}, 10^1\}$, where the parameters ρ, a are as defined in [13]. For each choice of ρ, a, a convex optimization problem is solved.

IV. HALF-DUPLEX RELAYING

In Section III, we assumed the relay was able to transmit and receive simultaneously in the same band. Such full-duplex radios can be difficult to implement in practice. In this section, we consider a half-duplex relay, where the relay receives in one band and transmits over a different band. In particular, we assume the channel has unit bandwidth, and it is partitioned into sub-channel Band 1 with bandwidth w_1, and another orthogonal sub-channel Band 2 with bandwidth w_2, with $w_1 + w_2 \leq 1$. The relay can only receive in Band 1 and it can only transmit in

Band 2. Hence the channel is described by

$$y_1^{(1)} = H_{11}x_1^{(1)} + z_1^{(1)}, \qquad\qquad y_1^{(2)} = H_{11}x_1^{(2)} + H_{12}x_2^{(2)} + z_1^{(2)} \tag{35}$$
$$y_2^{(1)} = H_{21}x_1^{(1)} + z_2^{(2)}, \qquad\qquad y_2^{(2)} = 0 \tag{36}$$

where the superscripts designate the corresponding bands. The noise powers in the sub-channels are given by

$$\mathrm{E}[z_i^{(1)}(z_i^{(1)})^H] = w_1 I_{N_i}, \qquad\qquad \mathrm{E}[z_i^{(2)}(z_i^{(2)})^H] = w_2 I_{N_i}, \qquad\qquad i = 1, 2. \tag{37}$$

Let $Q_{11}^{(1)}$, $Q^{(2)}$ be the transmit signal covariance matrices in the two bands

$$Q_{11}^{(1)} \triangleq \mathrm{E}[x_1^{(1)}(x_1^{(1)})^H] \in \mathbb{H}_+^{M_1}, \qquad Q^{(2)} \triangleq \mathrm{E}[x^{(2)}(x^{(2)})^H] \in \mathbb{H}_+^{M}, \qquad x^{(2)} \triangleq [x_1^{(2)}\ x_2^{(2)}]^T \in \mathbb{C}^M. \tag{38}$$

We assume the bandwidth allocation and the transmit signal covariances in each band can be optimized with respect to the channel realizations. In the following sections, we consider the cut-set capacity upper bound and the achievable rates under the half-duplex relaying constraint.

A. Half-Duplex Cut-Set Bound

Let the mutual information across the different cut sets be designated as labeled in Fig. 4. The cut set around the source is shown in Fig. 4(a). Let R_1, R_2, respectively, denote the egress information rate out of the source in Band 1 and Band 2. On the other hand, for the cut set around the destination shown in Fig. 4(b), let R_d, R_c be the ingress information rate into the destination in Band 1 and Band 2, respectively. Optimizing over the transmit signals and the bandwidth allocation, the half-duplex cut-set bound is characterized as follows:

$$\text{maximize} \quad R_{\mathrm{hCS}} \tag{39}$$

$$\text{over} \quad R_{\mathrm{hCS}}, R_1, R_2, R_d, R_c, w_1, w_2 \in \mathbb{R}_+,\ Q_{11}^{(1)} \in \mathbb{H}_+^{M_1},\ Q^{(2)} \in \mathbb{H}_+^{M} \tag{40}$$

$$\text{subject to} \quad R_{\mathrm{hCS}} \leq \min(R_1 + R_2,\ R_d + R_c) \tag{41}$$

$$R_1 \leq w_1 \log \det\big(I_N + \tfrac{1}{w_1} H_1 Q_{11}^{(1)} H_1^H\big) \tag{42}$$

$$R_2 \leq w_2 \log \det\big(I_{N_1} + \tfrac{1}{w_2} H_{11} C_1^T Q^{(2)} C_1 H_{11}^H\big) \tag{43}$$

$$R_d \leq w_1 \log \det\big(I_{N_1} + \tfrac{1}{w_1} H_{11} Q_{11}^{(1)} H_{11}^H\big) \tag{44}$$

$$R_c \leq w_2 \log \det\big(I_{N_1} + \tfrac{1}{w_2} \tilde{H}_1 Q^{(2)} \tilde{H}_1^H\big) \tag{45}$$

$$\mathbf{tr}\, Q_{11}^{(1)} + \mathbf{tr}(C_1^T Q^{(2)} C_1) \leq P_1 \tag{46}$$

$$\mathbf{tr}(C_2^T Q^{(2)} C_2) \leq P_2 \tag{47}$$

$$w_1 + w_2 \leq 1 \tag{48}$$

where C_1, C_2 are as defined in (19). By continuity we define: $w \log \det(I + X/w)|_{w=0} \triangleq 0$, for all $X \succeq 0$. The right-hand side of each constraint in (42), (43), (44), (45) is a concave function, being the perspective of the log-determinant function. (Given a function $f(x)$, the perspective of f is defined as the function $g(x,t) = tf(x/t)$, $t \in \mathbb{R}_{++}$, and the perspective operation preserves convexity [19].)

B. Half-Duplex Decode-and-Forward Rate

Fig. 5 depicts the operation of decode-and-forward in the half-duplex mode. In Band 1, the source sends to the relay at rate R_r, of which R_d is decodable at the destination. The relay fully decodes the message from the source, and in Band 2 the source and relay cooperatively send to the destination additional information at rate R_c. The
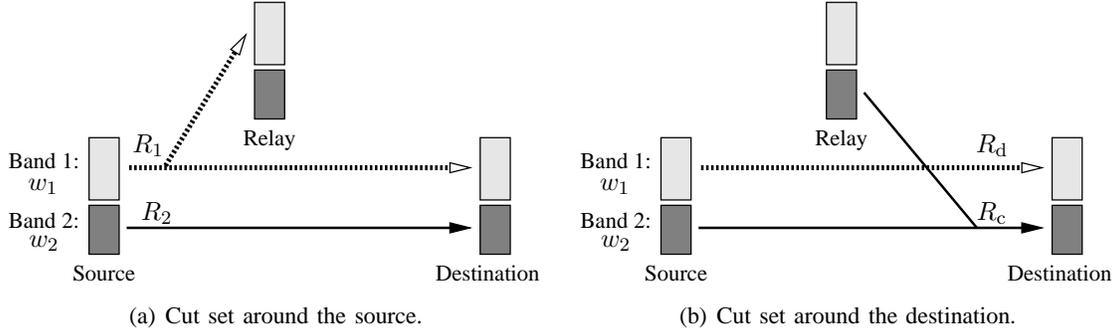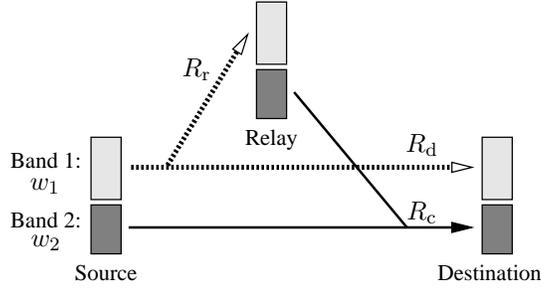
Fig. 4. Half-duplex cut-set bound.

Fig. 5. Half-duplex decode-and-forward.

half-duplex decode-and-forward optimization is given as follows:

$$\text{maximize} \quad R_{\text{hDF}} \tag{49}$$

$$\text{over} \quad R_{\text{hDF}}, R_{\text{r}}, R_{\text{d}}, R_{\text{c}}, w_1, w_2 \in \mathbb{R}_+, \ Q_{11}^{(1)} \in \mathbb{H}_+^{M_1}, \ Q^{(2)} \in \mathbb{H}_+^{M} \tag{50}$$

$$\text{subject to} \quad R_{\text{hDF}} \leq \min(R_{\text{r}}, \ R_{\text{d}} + R_{\text{c}}) \tag{51}$$

$$R_{\text{r}} \leq w_1 \log\det\left(I_{N_2} + \tfrac{1}{w_1} H_{21} Q_{11}^{(1)} H_{21}^H\right) \tag{52}$$

$$R_{\text{d}} \leq w_1 \log\det\left(I_{N_1} + \tfrac{1}{w_1} H_{11} Q_{11}^{(1)} H_{11}^H\right) \tag{53}$$

$$R_{\text{c}} \leq w_2 \log\det\left(I_{N_1} + \tfrac{1}{w_2} \tilde{H}_1 Q^{(2)} \tilde{H}_1^H\right) \tag{54}$$

$$\mathbf{tr}\, Q_{11}^{(1)} + \mathbf{tr}(C_1^T Q^{(2)} C_1) \leq P_1 \tag{55}$$

$$\mathbf{tr}(C_2^T Q^{(2)} C_2) \leq P_2 \tag{56}$$

$$w_1 + w_2 \leq 1. \tag{57}$$

C. Two-Hop Relaying

Two-hop relaying is a simple scheme that imposes relatively small coordination overhead between the source and relay. Its operation is portrayed in Fig. 6. In Band 1, the source transmits to the relay with signal covariance $Q_{11}^{(1)}$. The relay then decodes the message from the source, and re-encodes it to transmit to the destination in Band 2 with covariance $Q_{22}^{(2)} \in \mathbb{H}_+^{M_2}$. The following rate is achievable

$$\text{maximize} \quad R_{\text{2hop}} \tag{58}$$

$$\text{over} \quad R_{\text{2hop}}, R_{\text{sr}}, R_{\text{rd}}, w_1, w_2 \in \mathbb{R}_+, \ Q_{11}^{(1)} \in \mathbb{H}_+^{M_1}, \ Q_{22}^{(2)} \in \mathbb{H}_+^{M_2} \tag{59}$$

$$\text{subject to} \quad R_{\text{2hop}} \leq \min(R_{\text{sr}}, \ R_{\text{rd}}) \tag{60}$$

$$R_{\text{sr}} \leq w_1 \log\det\left(I_{N_2} + \tfrac{1}{w_1} H_{21} Q_{11}^{(1)} H_{21}^H\right) \tag{61}$$

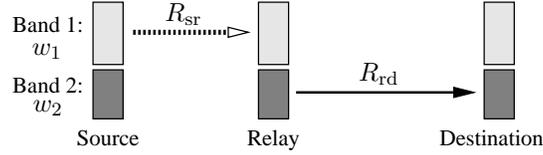$$R_{\text{rd}} \leq w_2 \log\det\left(I_{N_1} + \tfrac{1}{w_2} H_{12} Q_{22}^{(2)} H_{12}^H\right) \tag{62}$$

Fig. 6. Two-hop relaying.

$$\mathbf{tr}\, Q_{11}^{(1)} \le P_1 \tag{63}$$

$$\mathbf{tr}\, Q_{22}^{(2)} \le P_2 \tag{64}$$

$$w_1 + w_2 \le 1 \tag{65}$$

where R_{sr} represents the rate from the source to the relay, and R_{rd} represents the rate from the relay to the destination.

D. Co-Location MIMO Capacity

For comparison, we also consider the performance of the relay channel when the relay is close to the source or the destination. When the relay is co-located with the source, where it can cooperate perfectly with the source without overhead, the resulting channel is equivalent to an $M \times N_1$ MIMO channel. In the MIMO channel, M_1 of the transmit antennas are under a sum power constraint of P_1, and M_2 antennas are under sum power constraint P_2. The MIMO capacity is given by

$$\text{maximize} \quad R_{M \times N_1} \tag{66}$$

$$\text{over} \quad R_{M \times N_1} \in \mathbb{R}_+,\ Q \in \mathbb{H}_+^M \tag{67}$$

$$\text{subject to} \quad R_{M \times N_1} \le \log \det(I_{N_1} + \tilde{H}_1 Q \tilde{H}_1^H) \tag{68}$$

$$\mathbf{tr}(C_1^T Q C_1) \le P_1 \tag{69}$$

$$\mathbf{tr}(C_2^T Q C_2) \le P_2. \tag{70}$$

The above maximization is a convex optimization problem, and can be solved by the software package SDPT3 [23], which directly supports the log-determinant construct in the optimization objective function.

On the other hand, when the relay is co-located with the destination, we assume they can cooperate perfectly without overhead. In this case, the resulting channel is equivalent to an $M_1 \times N$ MIMO channel under a transmit power constraint of P_1. The MIMO channel capacity is

$$R_{M_1 \times N} = \max_{\mathbf{tr}\, Q_{11} \le P_1} \log \det(I_N + H_1 Q_{11} H_1^H) \tag{71}$$

where the solution is given by waterfilling power allocation [24] along the eigenmodes of $H_1^H H_1$.

E. Numerical Results

Fig. 7 shows the empirical CDF of the half-duplex cut-set (hCS) bounds, the half-duplex decode-and-forward (hDF) rates, and the two-hop relaying (2hop) rates for the MIMO relay channel with parameters as described in Section III-C. Again, the dotted lines represent the rates under per-antenna power constraints. For comparison, Fig. 7 also includes the full-duplex cut-set bound, full-duplex decode-and-forward rate, and the MIMO capacity of the direct channel H_{11} without the relay. It is observed that the half-duplex rates fall moderately as compared to the full-duplex rates. Moreover, the gap between the decode-and-forward rate and the cut-set capacity upper bound widens under the half-duplex mode. On the other hand, half-duplex decode-and-forward still provides a sizable capacity gain over direct transmission, while the two-hop relaying scheme achieves only marginally higher rates than when the relay is not available. Similar to the full-duplex case, imposing the per-antenna power constraints reduces the rates only slightly.

Next, we investigate the relay channel capacity bound and achievable rate as a function of the relay position. In the following numerical experiments, we fix $d_y = 1/10$, and vary d_x from $-1/2$ to $1^1/2$; therefore, the relay
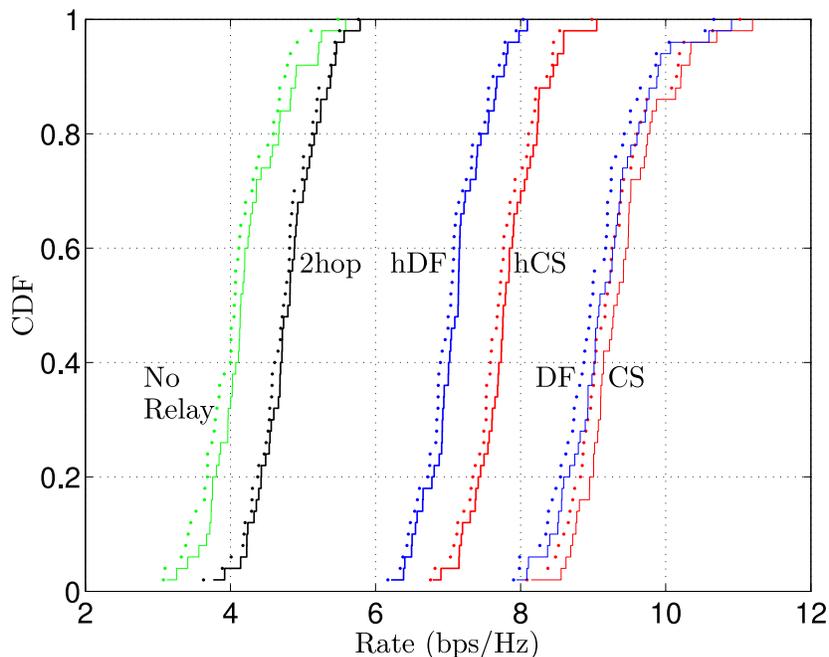
Fig. 7. Half-duplex CDFs of MIMO relay channel cut-set (hCS) bound, decode-and-forward (hDF) rate, and two-hop (2hop) relaying rate. The relay is located at $(1/3, 1/2)$. All terminals have four antennas: $M_1 = N_1 = M_2 = N_2 = 4$, and unit power constraints: $P_1 = P_2 = 0$ dB. The dotted lines represent the rates under per-antenna power constraints.

ranges from being closer to the source, to being closer to the destination. Again, all terminals in the network have four antennas: $M_1 = N_1 = M_2 = N_2 = 4$, and unit power constraints: $P_1 = P_2 = 0$ dB. The average rates for the different full- and half-duplex relaying schemes are plotted in Fig. 8; they are computed over the 50 sets of random channel realizations with distance-based path-loss as given in (34). The 8×4 MIMO capacity given by (66), corresponding to the case where the relay is co-located with the source, is indicated by a circle, whereas the 4×8 MIMO capacity (71), under relay-destination co-location, is indicated by a square. For the half-duplex schemes, the bandwidth allocation in terms of the relay location is shown in Fig. 9.

Under the full-duplex mode of operation, the decode-and-forward relaying scheme offers substantial capacity gain over transmission using only the direct channel. Over a wide range when the relay is close to the source, the decode-and-forward rate almost coincides with the cut-set capacity upper bound, and it is close to the 8×4 MIMO capacity when the relay is at $(d_x, d_y) = (0, 1/10)$. The highest decode-and-forward rate is attained when the relay is located approximately midway between the source and destination. However, as the relay moves from the source and approaches the destination, the decode-and-forward rate begins to deteriorate. In fact, when the relay enters into proximity of the destination, the decode-and-forward rate underperforms direct transmission. This is because the decode-and-forward scheme requires the relay to fully decode the message from the source, and consequently the source-relay channel becomes a performance bottleneck. In practice, the source would typically enlist the relay's help only if it offers a capacity gain over direct transmission; hence the achievable rate may be taken as the maximum of the relaying rate and the direct rate.

The half-duplex decode-and-forward scheme follows similar trends, but exhibits a wider gap from the half-duplex cut-set capacity upper bound, and its maximum capacity gain over the direct channel is less pronounced. In Fig. 9, it is observed that the system bandwidth is disproportionately allocated to Band 1 when the relay is in the vicinity of the destination, which corroborates with the system performance being limited by the source-relay link. The orthogonal two-hop relaying scheme, on the other hand, does not perform as well as decode-and-forward. It only offers moderate capacity gain over direct transmission, where the gain similarly is at its peak when the relay is about equidistant from the source and destination.
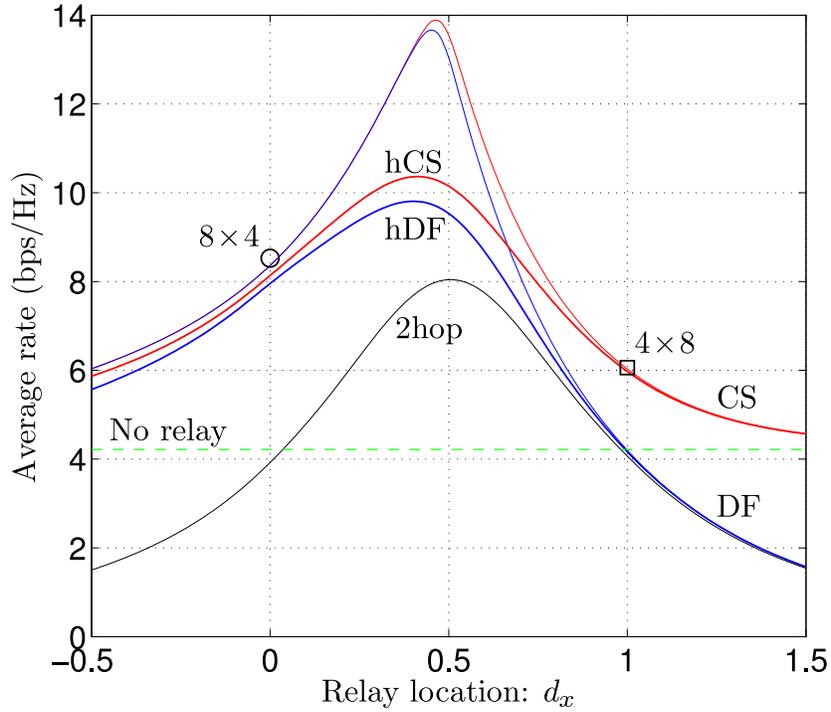
Fig. 8.   Full- and half-duplex cut-set (CS) bound, decode-and-forward (DF) rate, and two-hop (2hop) relaying rate with respect to the relay location.
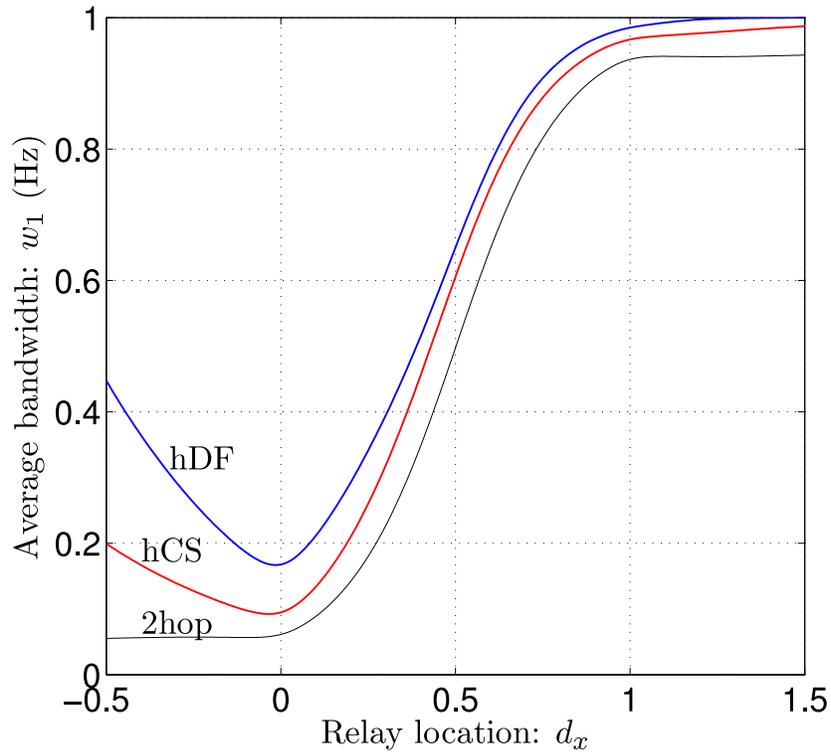

Fig. 9.   Bandwidth allocation in the half-duplex relaying schemes with respect to the relay location.

V. COMPRESS-AND-FORWARD RELAYING

From discussion in the previous section, it is clear that requiring the relay to decode the source's message can become a performance bottleneck when the source-relay channel is weak. In this section, we consider the compress-and-forward [4, Thm. 6] strategy where the relay does not attempt to decode the message from the source. Rather, the relay forwards a compressed version of its observation to the destination. The relay's observation is compressed in the sense that a finite number of bits is used to represent the analog signal (the scheme is sometimes also referred to as quantize-and-forward). Unlike the cut-set bound and decode-and-forward formulations, however, the transmit signal design and bandwidth allocation under compress-and-forward do not appear to be convex problems. In this section, we consider achievable compress-and-forward transmission schemes. We focus on full-duplex transmission; under fixed bandwidth allocation, the compression-and-forward operation readily extends to half-duplex relaying.

A. Compress-and-Forward Transmission

We first describe the general compress-and-forward strategy; specific compression schemes are considered in Sections V-B, V-C. The optimal joint design of the transmit signals and compression rate appears to be intractable; in the following we present suboptimal approaches to consider specific power allocation and compression schemes. We assume the source and the relay use Gaussian signals. Using the capacity-achieving strategy as in a multiple-access channel [17], suppose the destination performs successive interference cancellation to allow simultaneous transmission from the source and relay. In particular, we consider the decode order in which the destination first decodes the relay's message, treating transmission from the source as noise. Then the relay's codeword is subtracted from the observed signal, and the message from the source is decoded. The source-destination transmission is thus interference-free from the relay's signals, and the source optimizes its own transmit signal covariance Q_{11} according to

$$R_{11} = \max_{Q_{11} \,:\, \mathbf{tr}\, Q_{11} \leq P_1} \log \det\big(I_{N_1} + H_{11} Q_{11} H_{11}^H\big) \tag{72}$$

where the solution is given by the waterfilling procedure. Let Q_{11}^* denote the covariance matrix that maximizes (72). Next, the relay optimizes its transmit signal against the interference from the source's transmission

$$R_{12} = \max_{Q_{22} \,:\, \mathbf{tr}\, Q_{22} \leq P_2} \log \det\big(I_{N_1} + \tilde{H}_{12} Q_{22} \tilde{H}_{12}^H\big) \tag{73}$$

where $\tilde{H}_{12}$ is the effective channel from the relay to destination treating interference from the source as noise

$$\tilde{H}_{12} \triangleq (I_{N_1} + H_{11} Q_{11}^* H_{11}^H)^{-1/2} H_{12}. \tag{74}$$

Similarly, the solution in (73) is given by waterfilling against the effective channel $\tilde{H}_{12}$.

In the compress-and-forward approach, the relay sends $\tilde{y}_2 \in \mathbb{C}^{N_2}$ to the destination, which is a compressed version of the relay's receive signal y_2, with compression rate R_{12} as given in (73). The compression schemes considered in this paper can be modeled by

$$\tilde{y}_2 = A y_2 + \tilde{z} \tag{75}$$

where $A \in \mathbb{C}^{N_2 \times N_2}$ is a constant scaling matrix, and $\tilde{z} \sim \mathcal{CN}(0, Z) \in \mathbb{C}^{N_2}$ is independent additive Gaussian compression noise, with $Z \in \mathbb{H}_+^{N_2}$. Upon receiving $\tilde{y}_2$ at the destination, the relay network is equivalent to an $M_1 \times (N_1 + N_2)$ MIMO channel, except that N_2 of its receive antennas are scaled by A and corrupted by compression noise $\tilde{z}$

$$\begin{bmatrix} y_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{bmatrix} H_{11} \\ A H_{21} \end{bmatrix} x_1 + \begin{bmatrix} z_1 \\ A z_2 + \tilde{z} \end{bmatrix}. \tag{76}$$

The corresponding channel capacity is given by

$$R_{\mathrm{CF}} = \log \det\big(I_N + \hat{H}_1 Q_{11}^* \hat{H}_1^H\big) \tag{77}$$

where $\hat{H}_1$ is the effective source-to-destination-and-relay MIMO channel, incorporating the degradation introduced by the compression scheme as given in (76)

$$\hat{H}_1 \triangleq \begin{bmatrix} H_{11} \\ (Z + A A^H)^{-1/2} A H_{21} \end{bmatrix} \in \mathbb{C}^{N \times M_1}. \tag{78}$$

The different compression schemes considered in this paper differ in their respective achieved values of A and Z. They are described in more detail in the next sections.

B. Rate–Distortion Compression

Let $\hat{y}_2 \in \mathbb{C}^{N_2}$ represent the compressed version of the signal y_2 under rate–distortion theory [17]. Let $\hat{z} = y_2 - \hat{y}_2$ be the compression error, where $\hat{z}$ has zero mean. The quality of the compression is characterized by the distortion covariance matrix: $D \in \mathbb{H}_+^{N_2} \triangleq \mathrm{E}[\hat{z}\hat{z}^H]$. In general, we wish D to be small to obtain a high compress-and-forward rate. In the following, we consider a simple approach to model this dependency by considering minimizing: $\mathbf{tr}\, D$, which represents the sum of the quadratic distortion measures in the entries of $\hat{y}_2$. The distortion-rate function prescribes the minimal distortion achievable under the compression rate R_{12}

$$\min_{D\,:\,I(y_2;\hat{y}_2)\leq R_{12}} \mathbf{tr}\, D. \tag{79}$$

After evaluating the mutual information expression in (79) under Gaussian signaling, the minimization in the distortion-rate function can be written as

$$\text{minimize} \quad \mathbf{tr}\, D \tag{80}$$

$$\text{over} \quad D \in \mathbb{H}_+^{N_2} \tag{81}$$

$$\text{subject to} \quad \log\det(S_2 - D) \leq R_{12}. \tag{82}$$

Let D^* denote the optimal distortion covariance matrix in (81); it is computed by the reverse waterfilling [17] procedure along the eigenmodes of S_2. The $\hat{y}_2$ that achieves the minimum sum distortion in (80) has a joint distribution with y_2 described by

$$y_2 = \hat{y}_2 + \hat{z}, \qquad\qquad \hat{y}_2 \sim \mathcal{CN}(0, S_2 - D^*), \qquad\qquad \hat{z} \sim \mathcal{CN}(0, D^*). \tag{83}$$

The joint distribution in (83) can be equated with the form in (75) by setting the corresponding parameters of the compression scheme to be

$$A_{\mathrm{RD}} = (I_{N_2} - D^* S_2^{-1})^{1/2}, \qquad\qquad Z_{\mathrm{RD}} = D^* \tag{84}$$

where the subscripts in $A_{\mathrm{RD}}, Z_{\mathrm{RD}}$ are used to designate the compression scheme under consideration.

C. Wyner–Ziv Compression

In the compress-and-forward strategy in [4, Thm. 6], the transmission scheme also takes advantage of the correlation between the observed signals at the source and relay using Wyner–Ziv compression. In particular, when the destination attempts to reconstruct y_2 form $\tilde{y}_2$, it also has access to its own receive signal y_1, which can be used to improve the performance of the compression process. The Wyner–Ziv compression approach [25] exploits the correlation between y_1 and y_2 as side information at the decoder to achieve a lower compression noise level with the same compression rate R_{12}.

With the transmit signal of the source being as specified in (72), the covariance of the observed signals at the destination (after successive interference cancellation of the relay's signal) and relay, respectively, are given by

$$S_{11} \triangleq \mathrm{E}[y_1 y_1^H | x_2] = I_{N_1} + H_{11} Q_{11}^* H_{11}^H \in \mathbb{H}_+^{N_1} \tag{85}$$

$$S_{22} \triangleq \mathrm{E}[y_2 y_2^H] = I_{N_2} + H_{21} Q_{11}^* H_{21}^H \in \mathbb{H}_+^{N_2}. \tag{86}$$

Moreover, the cross-covariance between y_2 and $y_1 | x_2$ is

$$S_{21} \triangleq \mathrm{E}[y_2 y_1^H | x_2] = H_{21} Q_{11}^* H_{11}^H \in \mathbb{C}^{N_2 \times N_1}. \tag{87}$$

For Gaussian signals under quadratic distortion, the Wyner–Ziv scheme achieves the same rate–distortion tradeoff as if the side information were also present at the encoder [25], [26] (i.e., as if the relay had access to y_1 in the

course of the compression process). Therefore, the Wyner–Ziv compression noise is given by the distortion-rate function of compressing the signal $y_2|y_1$ using rate R_{12}

$$\text{minimize} \quad \mathbf{tr}\,\bar{D} \tag{88}$$

$$\text{over} \quad \bar{D} \in \mathbb{H}_+^{N_2} \tag{89}$$

$$\text{subject to} \quad \log\det(S_{2|1} - \bar{D}) \leq R_{12} \tag{90}$$

where $\bar{D}$ represents the distortion covariance matrix of the Wyner–Ziv compression error, and $S_{2|1}$ is the conditional covariance of y_2 given y_1

$$S_{2|1} \triangleq \mathrm{E}[y_2 y_2^H | y_1] = S_{22} - S_{21}S_{11}^{-1}S_{21}^H \in \mathbb{H}_+^{N_2}. \tag{91}$$

The parameters for the Wyner–Ziv compression scheme are then identified to be

$$A_{\text{WZ}} = (I_{N_2} - \bar{D}^* S_{2|1}^{-1})^{1/2}, \qquad\qquad Z_{\text{WZ}} = \bar{D}^* \tag{92}$$

where $\bar{D}^*$ is the solution to (89) from reverse waterfilling against $S_{2|1}$. Since the side information reduces the compression noise, Wyner–Ziv compression always achieves better performance than the rate–distortion compression scheme. However, rate–distortion compression has lower implementation complexity, since the correlation between the receive signals of the source and relay is not exploited in the compression process.

D. Numerical Results

Fig. 10 shows the compress-and-forward rates under the same channel parameters as those considered in Fig. 8. The Wyner–Ziv (WZ) compress-and-forward rate outperforms the compression-and-forward rate under rate–distortion (RD), which demonstrates the capacity gain from exploiting side information. However, the Wyner–Ziv advantage ceases when the relay is close to the destination: in that regime, the efficiency of the compression scheme has limited impact, as the relay has a strong channel to the destination. Overall, the compress-and-forward rates do not perform as well as the decode-and-forward rates, except when the relay is far from the source and near the destination. Moreover, unlike its decode-and-forward counterpart at $(d_x, d_y) = (0, 1/10)$, the compress-and-forward rates markedly fall short of the 4×8 MIMO capacity when the relay is at $(d_x, d_y) = (1, 1/10)$. Nevertheless, as the relay is under no stipulation to perform any decoding, the compress-and-forward rate is at least as large as that under direct transmission, regardless of network geometry.

VI. Conclusions

We considered the optimization of transmit signals and bandwidth allocation for MIMO relay channels. We assumed that all terminals have channel state information, and we evaluated the cut-set capacity upper bounds and the decode-and-forward rates by formulating them as convex optimization problems. The solutions to the optimization problems can be efficiently computed by numerical convex optimization methods. In the case of half-duplex relaying, where the relay cannot simultaneously transmit and receive in the same frequency band, the bandwidth allocation and the transmit signals are jointly optimized. We also presented achievable relaying rates based on the compress-and-forward strategy, where the relay does not decode the message from the source, but forwards a compressed version of its observation to the destination using the rate–distortion and Wyner–Ziv compression schemes.

When the relay is close to the source, it is observed that the decode-and-forward coding strategy is almost optimal: its achieved rate is near the cut-set capacity upper bound, especially in full-duplex relaying. Moreover, under the half-duplex constraint, decode-and-forward significantly outperforms orthogonal two-hop relaying. For all relaying schemes, the maximum capacity gain over direct transmission is attained when the relay is approximately halfway between the source and destination. On the other hand, when the relay is close to the destination, decode-and-forward underperforms direct transmission as the source-relay link becomes a bottleneck. In this regime good performance is achieved by the compress-and-forward schemes, which always achieve a rate that is equal to or better than the direct transmission rate.
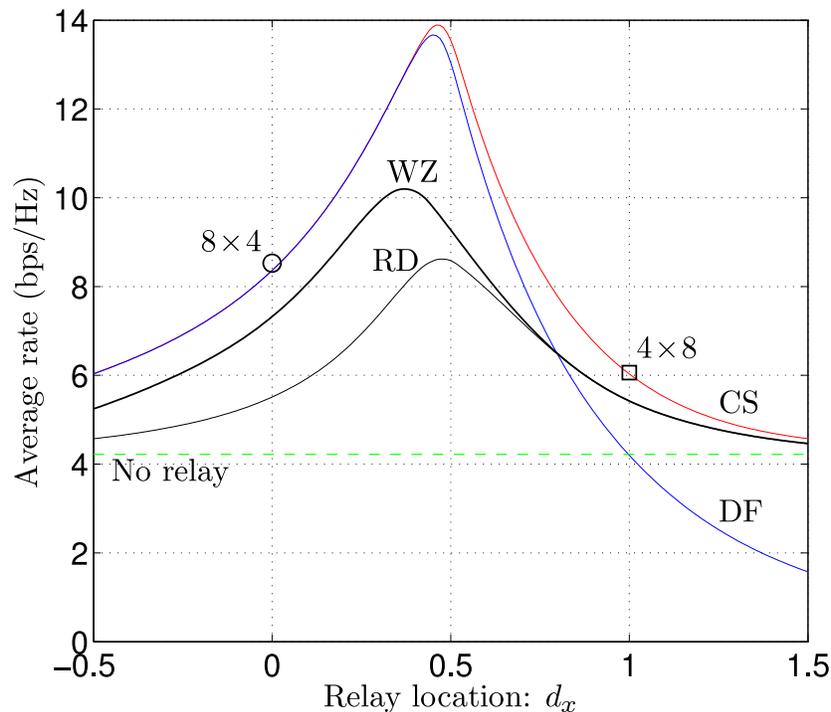
Fig. 10. Compress-and-forward relaying rates under Wyner–Ziv (WZ) compression with side information and the rate–distortion (RD) compression scheme.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[3] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[4] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[5] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[6] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of 'cheap' relay networks," in *Proc. Conf. on Inform. Sciences and Systems*, Baltimore, MD, Mar. 2003.

[7] ——, "On the capacity of Gaussian 'cheap' relay channel," in *Proc. IEEE Globecom Conf.*, San Francisco, CA, Dec. 2003, pp. 1776–1780.

[8] Y. Liang and V. V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3284–3289, Sep. 2005.

[9] Y. Liang, V. V. Veeravalli, and H. V. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3432–3453, Oct. 2007.

[10] I. Maric and R. D. Yates, "Bandwidth and power allocation for cooperative strategies in Gaussian relay networks," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2004, pp. 1907–1911.

[11] C. T. K. Ng and A. J. Goldsmith, "The impact of CSI and power allocation on relay channel capacity and cooperation strategies," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5380–5389, Dec. 2008.

[12] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[13] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[14] C. T. K. Ng, J. N. Laneman, and A. J. Goldsmith, "The role of SNR in achieving MIMO rates in cooperative systems," in *Proc. IEEE Inform. Theory Workshop*, Punta del Este, Uruguay, Mar. 2006, pp. 288–292.

[15] M. Yuksel and E. Erkip, "Diversity-multiplexing tradeoff in half-duplex relay systems," in *Proc. IEEE Internat. Conf. Commun.*, Glasgow, Scotland, Jun. 2007, pp. 689–694.

[16] ——, "Diversity-multiplexing tradeoff in multiple-antenna relay systems," in *Proc. IEEE Int. Symp. on Inform. Theory*, Seattle, WA, Jul. 2006, pp. 1154–1158.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY: Wiley-Interscience, 1991.

[18] F. Zhang, Ed., *The Schur Complement and Its Applications*. New York, NY: Springer, 2005.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[20] J. Renegar, *A Mathematical View of Interior-Point Methods in Convex Optimization*. Philadelphia, PA: MPS-SIAM, 2001.

[21] M. C. Grant and S. P. Boyd, "Graph implementations for nonsmooth convex programs," *Recent Advances in Learning and Control (a tribute to M. Vidyasagar)*, pp. 95–110, 2008, V. Blondel, S. Boyd, and H. Kimura, editors, Lecture Notes in Control and Information Sciences, Springer.

[22] M. Grant and S. Boyd. (2009, Jun.) CVX: Matlab software for disciplined convex programming (web page and software). [Online]. Available: http://stanford.edu/~boyd/cvx

[23] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, vol. 95, no. 2, pp. 189–217, Feb. 2003.

[24] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[25] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder—II: General sources," *Inform. Contr.*, vol. 38, pp. 60–80, Jul. 1978.

[26] R. Zamir, "The rate loss in the Wyner–Ziv problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2073–2084, Nov. 1996.