The Approximate Capacity of the MIMO Relay Channel

Xianglan Jin, *Member, IEEE*, and Young-Han Kim, *Fellow, IEEE*

***Abstract*— Capacity bounds are studied for the multiple-antenna complex Gaussian relay channel with t_1 transmitting antennas at the sender, r_2 receiving and t_2 transmitting antennas at the relay, and r_3 receiving antennas at the receiver. It is shown that the partial decode–forward coding scheme achieves within $\min(t_1, r_2)$ bits from the cutset bound and at least one half of the cutset bound, establishing a good approximate characterization of the capacity. A similar additive gap of $\min(t_1 + t_2, r_3) + r_2$ bits is shown to be achieved by the compress–forward coding scheme. The corresponding results for time-division half-duplex relay channels are also established.**

***Index Terms*— Additive gap, channel capacity, compress–forward, cutset bound, multiplicative gap, partial decode–forward.**

I. INTRODUCTION

THE relay channel, whereby point-to-point communication between a sender and a receiver is aided by a relay, is a canonical building block for cooperative wireless communication. Introduced by van der Meulen [1], this channel model has been studied extensively in the literature, including the now classical paper by Cover and El Gamal [2]. The problem of characterizing the capacity in a computable expression, however, remains open even for simple channel models, and consequently a large body of the literature has been devoted to the study of upper and lower bounds on the capacity. Reminiscent of the max-flow min-cut theorem [3], the cutset bound was established by Cover and El Gamal [2], which sets an intuitive upper limit on the capacity. On the other direction, there are a myriad of coding schemes, typically referred to as "*–forward" [4], each establishing a lower bound on the capacity. Among these, the two most versatile coding schemes are partial decode–forward [2, Th. 7] and compress–forward [2, Th. 6], which are complementary to each other (providing digital-to-digital and analog-to-digital relays, respectively) and have been successfully extended to general relay networks for unicast, multicast, broadcast, and multiple access [5]–[9].

The Gaussian relay channel, whereby the signals from the sender and the relay are corrupted by additive white Gaussian noise, is one of the most basic channel models studied in the literature. The capacity of the Gaussian relay channel, however, is again unknown for any nondegenerate channel parameters. Instead, the following results have been established for single-antenna Gaussian relay channels.

- Partial decode–forward, which is a superposition of decode–forward and direct transmission, reduces to the better of the two [10].
- Partial decode–forward achieves within one bit from the cutset bound [6], [11] (a similar 1-bit gap result can be also obtained using compress–forward [12]).

These results establish simple approximate expressions of the capacity, which are particularly useful at high signal-to-noise ratio (SNR). A natural question arises on how these results can be extended to multiple-antenna (also known as multiple-input multiple-output or MIMO) Gaussian relay channels.

Capacity bounds for MIMO relay channels have been studied in several papers. By convex optimization techniques [13], Wang *et al.* [14] derived upper and lower bounds based on looser versions of the cutset bound and the decode–forward bound. These results have been improved by more advanced coding schemes (partial decode–forward and compress–forward) with suboptimal decoding rules by Simoens *et al.* [15] and Ng and Foschini [16]. The usual focus of this line of work, however, has been on the optimization of resources (power and bandwidth) for practical implementations and on numerical computation of resulting capacity bounds. In the same vein, a recent study by Gerdes *et al.* [17] established the optimal input distribution of the partial decode–forward lower bound for the Gaussian MIMO relay channel.

The most relevant results to our main question are established in [6], [9], and [18]. In the award-winning paper, Avestimehr *et al.* [6] proposed a powerful approach to studying wireless networks, which, inter alia, establishes an approximate capacity of the Gaussian MIMO relay network using the quantize–map–forward scheme. Among several extensions of this seminal work, one of the tightest results to date for MIMO relay networks has been established by Kolte *et al.* [18], who carefully compared the noisy network coding lower bound for the general relay network [8] with the cutset bound. In an alternative direction, the distributed decode–forward scheme [9], which generalizes partial decode–forward to general relay networks, can be used to approximate the capacity of the

Manuscript received September 7, 2015; revised May 12, 2016; accepted July 25, 2016. Date of publication November 23, 2016; date of current version January 18, 2017. This work was supported in part by the Korean National Research Foundation under the Basic Science Research Program under Grant NRF-2014R1A1A2053589 and in part by the National Science Foundation under Grant CCF-1320895. This work was presented at the 2014 IEEE International Symposium on Information Theory, Honolulu, Hawaii.

X. Jin is with the Department of Electrical and Computer Engineering, Pusan National University, Busan 46241, South Korea (e-mail: jinxl77@pusan.ac.kr).

Y.-H. Kim is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407 USA (e-mail: yhk@ucsd.edu).

Communicated by S. Avestimehr, Associate Editor for Communications.

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/TIT.2016.2632625

Gaussian MIMO relay network. These results [6], [9], [18] can be readily specialized to the 3-node relay channel and yield constant gap results in terms of the number of antennas.

This paper provides more direct and tighter answers to our main question through an elementary yet careful analysis of the partial decode–forward and compress–forward lower bounds for the MIMO relay channel. The main contributions are summarized as follows.

- For the complex Gaussian relay channel with t_1 transmitting antennas at the sender, r_2 receiving and t_2 transmitting antennas at the relay, and r_3 receiving antennas at the receiver, we show that the partial decode–forward achieves within $\min(t_1, r_2)$ bits of the cutset bound (Theorem 1).
- This gap is somewhat relaxed when noncoherent transmission is employed (Proposition 4).
- Unlike the single-antenna counterpart, partial decode–forward can achieve rates arbitrarily higher than the better of decode–forward and direct transmission in MIMO relay channels (Proposition 5).
- To complement the additive gap result, we show that both coherent and noncoherent partial decode–forward coding schemes achieve at least half the cutset bound (Theorem 2).
- We show that compress–forward achieves $\min(t_1 + t_2, r_3) + r_2$ bits within the cutset bound (Theorem 3).

In conclusion, the paper establishes simple approximate expressions of the capacity, which are particularly useful at high and low SNR. Beyond these analytical results, we also discuss how these expressions can be computed efficiently.

Several models for *half-duplex* relay channels have been studied in the literature, in which the relay or relay antenna cannot transmit and receive simultaneously; see, for example, [6], [10], [16], [19]–[23]. Among these, arguably the most interesting model was proposed by Kramer [21], whereby the relay state is switched between the transmitting and receiving modes based on the channel output at the relay. As an extension, Cardone *et al.* [23] introduced an individual antenna switching strategy at the relay for MIMO relay networks and established the constant gap result of 1.96 bits per antenna using the noisy network coding scheme. This result extends their earlier work on the single-antenna relay channel case [22] that shows partial decode–forward achieves the capacity within 1 bit and compress–forward achieves the capacity within 1.61 bits. Our main results for the full-duplex case can be easily adapted to the half-duplex MIMO relay channel with t_1 transmitting antennas at the sender, a_2 half-duplex antennas at the relay, and r_3 receiving antennas at the receiver as follows.

- Partial decode–forward achieves within $\min(t_1, a_2)$ bits of the cutset bound as well as at least half the cutset bound (Proposition 13).
- Compress–forward achieves within $\min(t_1 + 2a_2, r_3 + 1.58a_2)$ bits of the cutset bound (Proposition 14).

The rest of the paper is organized as follows. In the next section, we formally define the channel model and review the cutset upper bound, the partial decode–forward lower bound, and the compress–forward lower bound on the capacity. The main results on additive and multiplicative gaps for partial decode–forward and compress–forward are also stated therein. The proofs of these results are given in Sections III and IV. Section V is devoted to the computational aspects of our results, namely, how the capacity bounds can be computed efficiently via appropriate convex optimization formulations. Using these computational tools, the main results are verified by numerical simulations. In Section VI, half-duplex MIMO relay channels are discussed.

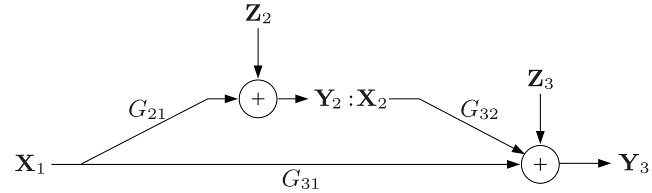

Fig. 1. The MIMO relay channel.

Throughout the paper, we use the following notation. The superscript $(\cdot)^{\mathsf{H}}$ denotes the complex conjugate transpose of a (complex) matrix; $\mathrm{tr}(\cdot)$ denotes the trace of a matrix; I_n denotes the $n \times n$ identity matrix (where the subscript n is omitted when it is irrelevant or clear from the context); $\mathbb{C}^{n\times m}$ denotes a set of $n \times m$ complex matrices; $A \succeq B$ denotes that $A - B$ is hermitian and positive semidefinite for hermitian matrices A and B, and $\mathsf{E}(\cdot)$ denotes the expectation with respect to the random variables in the argument.

II. Problem Setup and Main Results

We model the point-to-point communication system with a relay as a MIMO relay channel with sender node 1, relay node 2, and receiver node 3; see Fig. 1. Throughout the paper, we assume the complex signal model, but corresponding results for the real case can be easily obtained; see the conference version [24] of the current paper for some results on the real model. The relay and the receiver have r_2 and r_3 receiving antennas with respective channel outputs

$$\begin{aligned}\mathbf{Y}_2 &= G_{21}\mathbf{X}_1 + \mathbf{Z}_2,\\ \mathbf{Y}_3 &= G_{31}\mathbf{X}_1 + G_{32}\mathbf{X}_2 + \mathbf{Z}_3,\end{aligned} \tag{1}$$

where $G_{21} \in \mathbb{C}^{r_2\times t_1}$, $G_{31} \in \mathbb{C}^{r_3\times t_1}$, and $G_{32} \in \mathbb{C}^{r_3\times t_2}$ are complex channel gain matrices, $\mathbf{X}_1 \in \mathbb{C}^{t_1}$ and $\mathbf{X}_2 \in \mathbb{C}^{t_2}$ are the respective inputs at the sender and the relay, and $\mathbf{Z}_2 \sim \mathrm{CN}(0, I_{r_2})$ and $\mathbf{Z}_3 \sim \mathrm{CN}(0, I_{r_3})$ are independent complex Gaussian noise components. For simplicity, we will often use the shorthand notation

$$G_{3*} = \begin{bmatrix} G_{31} & G_{32}\end{bmatrix} \quad \text{and} \quad G_{*1} = \begin{bmatrix} G_{21} \\ G_{31}\end{bmatrix}.$$

We assume average power constraints P across the t_1 and t_2 transmitting antennas at the sender and the relay, respectively. As in the standard relay channel model [2], the encoder is defined by $\mathbf{x}_1^n(m)$, the relay encoder is defined by $\mathbf{x}_{2i}(\mathbf{y}_2^{i-1})$, $i = 1, \ldots, n$, and the decoder is defined by $\hat{m}(\mathbf{y}_3^n)$. We assume that the message M is uniformly distributed over the message set. The average probability of error is defined as $P_e^{(n)} = \mathsf{P}\{\hat{M} \neq M\}$. A rate R is said to be achievable for the relay channel if there exists a sequence of $(2^{nR}, n)$ codes such that $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity C of the relay channel is the supremum of all achievable rates.

The following upper bound on the capacity is well known.

Proposition 1 (Cutset Bound [2, Th. 4]): The capacity C of the MIMO relay channel is upper bounded by

$$\begin{aligned} R_{\text{CS}} &= \sup_{F(\mathbf{x}_1,\mathbf{x}_2)} \min\{I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3), I(\mathbf{X}_1;\mathbf{Y}_2,\mathbf{Y}_3|\mathbf{X}_2)\} \quad (2)\\ &= \max_K \min\{\log|I_{r_3}+G_{3*}KG_{3*}^{\mathsf{H}}|,\\ &\qquad \log|I_{r_2+r_3}+G_{*1}K_{1|2}G_{*1}^{\mathsf{H}}|\} \quad (3)\\ &= \max_K \min\{\log|I_{r_3}+[G_{31}\ G_{32}]K[G_{31}\ G_{32}]^{\mathsf{H}}|,\\ &\qquad \log|I_{t_1}+(G_{21}^{\mathsf{H}}G_{21}+G_{31}^{\mathsf{H}}G_{31})K_{1|2}|\} \quad (4)\end{aligned}$$

where the supremum in (2) *is over all joint distributions $F(\mathbf{x}_1,\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j) \le P,\ j=1,2$, the maxima in* (3) *and* (4) *are over all $(t_1+t_2)\times(t_1+t_2)$ matrices*

$$K = \begin{bmatrix} K_1 & K_{12}\\ K_{12}^{\mathsf{H}} & K_2\end{bmatrix} \succeq 0 \quad (5)$$

such that $\operatorname{tr}(K_j) \le P,\ j=1,2$, *and*

$$K_{1|2} = K_1 - K_{12}K_2^{-1}K_{12}^{\mathsf{H}}.$$

The equality in (4) is justified by the following simple fact that will be used repeatedly throughout the paper.

Lemma 1: For $\gamma \in [0,1]$, a $r\times t$ matrix G, and a $t\times t$ matrix $K \succeq 0$,

$$\begin{aligned}|I_r+\gamma GKG^{\mathsf{H}}| &= |I_t+\gamma G^{\mathsf{H}}GK|\\ &\ge \gamma^{\min(t,r)}|I_r+GKG^{\mathsf{H}}|. \quad (6)\end{aligned}$$

We compare the cutset bound with two lower bounds on the capacity. The first lower bound is based on the partial decode–forward coding scheme, in which the relay recovers part of the message and forwards it.

Proposition 2 (Partial Decode–Forward Bound [2, Th. 7]): The capacity C of the MIMO relay channel is lower bounded by

$$\begin{aligned}R_{\text{PDF}} &= \sup\min\{I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3),\\ &\qquad I(\mathbf{U};\mathbf{Y}_2|\mathbf{X}_2)+I(\mathbf{X}_1;\mathbf{Y}_3|\mathbf{X}_2,\mathbf{U})\} \quad (7)\\ &= \sup\min\{I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3),\\ &\qquad I(\mathbf{X}_1;\mathbf{U},\mathbf{Y}_3|\mathbf{X}_2)-I(\mathbf{X}_1;\mathbf{U}|\mathbf{X}_2,\mathbf{Y}_2)\} \quad (8)\end{aligned}$$

where the suprema are over all joint distributions $F(\mathbf{u},\mathbf{x}_1,\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j) \le P,\ j=1,2$.

Remark 1: The equivalence between (7) and (8) is due to the fact that $\mathbf{U}\to\mathbf{X}_1\to\mathbf{Y}_2$ form a Markov chain. It can be readily checked that the partial decode–forward lower bound does not increase by (coded) time sharing.

The partial decode–forward lower bound can be relaxed in several directions. First, by limiting the input distribution to a more practical product form, we obtain the *noncoherent* partial decode–forward lower bound:

$$\begin{aligned}R_{\text{NPDF}} &= \sup\min\{I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3),\\ &\qquad I(\mathbf{U};\mathbf{Y}_2|\mathbf{X}_2)+I(\mathbf{X}_1;\mathbf{Y}_3|\mathbf{X}_2,\mathbf{U})\} \quad (9)\\ &= \sup\min\{I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3),\\ &\qquad I(\mathbf{X}_1;\mathbf{U},\mathbf{Y}_3|\mathbf{X}_2)-I(\mathbf{X}_1;\mathbf{U}|\mathbf{X}_2,\mathbf{Y}_2)\} \quad (10)\end{aligned}$$

where the suprema are over all *product* distributions $F(\mathbf{u},\mathbf{x}_1)F(\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j) \le P,\ j=1,2$. Second, by setting $\mathbf{U}=\mathbf{X}_1$, which is equivalent to having the relay recover the entire message, we obtain the decode–forward lower bound:

$$\begin{aligned}R_{\text{DF}} &= \sup\min\{I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3), I(\mathbf{X}_1;\mathbf{Y}_2|\mathbf{X}_2)\} \quad (11)\\ &= \max_K\min\{\log|I_{r_3}+G_{3*}KG_{3*}^{\mathsf{H}}|,\\ &\qquad \log|I_{r_2}+G_{21}K_{1|2}G_{21}^{\mathsf{H}}|\} \quad (12)\end{aligned}$$

where the supremum in (11) is over all distributions $F(\mathbf{x}_1,\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j) \le P,\ j=1,2$, and the maximum in (12) is over all $(t_1+t_2)\times(t_1+t_2)$ matrices $K\succeq 0$ of the form (5) such that $\operatorname{tr}(K_j)\le P,\ j=1,2$. Third, by setting $\mathbf{U}=\emptyset$ and $\mathbf{X}_2=0$, we obtain the direct-transmission lower bound:

$$\begin{aligned}R_{\text{DT}} &= \sup I(\mathbf{X}_1;\mathbf{Y}_3)\\ &= \max_{K_1}\log|I_{r_3}+G_{31}K_1G_{31}^{\mathsf{H}}| \quad (13)\end{aligned}$$

where the supremum is over all distributions $F(\mathbf{x}_1)$ such that $\mathsf{E}(\mathbf{X}_1^{\mathsf{H}}\mathbf{X}_1)\le P$ and the maximum is over all $t_1\times t_1$ matrices $K_1\succeq 0$ such that $\operatorname{tr}(K_1)\le P$.

Remark 2: Since decode–forward and direct transmission schemes are two special cases of partial decode–forward, we have in general

$$R_{\text{PDF}} \ge \max(R_{\text{DF}}, R_{\text{DT}}). \quad (14)$$

Next, we present another important lower bound, in which the relay compresses its noisy observation instead of recovering the message.

Proposition 3 (Compress–Forward Bound [2, Th. 6], [10]): The capacity C of the MIMO relay channel is lower bounded by

$$R_{\text{CF}} = \sup I(\mathbf{X}_1;\hat{\mathbf{Y}}_2,\mathbf{Y}_3|\mathbf{X}_2) \quad (15)$$

where the supremum is over all conditional distributions $F(\mathbf{x}_1)F(\mathbf{x}_2)F(\hat{\mathbf{y}}_2|\mathbf{y}_2,\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j)\le P,\ j=1,2$ and

$$I(\mathbf{X}_2;\mathbf{Y}_3) \ge I(\mathbf{Y}_2;\hat{\mathbf{Y}}_2|\mathbf{X}_2,\mathbf{Y}_3).$$

This lower bound can be expressed equivalently as

$$\begin{aligned}R_{\text{CF}} = \sup\min\{&I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3)-I(\mathbf{Y}_2;\hat{\mathbf{Y}}_2|\mathbf{X}_1,\mathbf{X}_2,\mathbf{Y}_3),\\ &I(\mathbf{X}_1;\hat{\mathbf{Y}}_2,\mathbf{Y}_3|\mathbf{X}_2)\} \quad (16)\end{aligned}$$

where the supremum is over all conditional distributions $F(\mathbf{x}_1)F(\mathbf{x}_2)F(\hat{\mathbf{y}}_2|\mathbf{y}_2,\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j)\le P,\ j=1,2$.

Remark 3: The compress–forward lower bound before taking the supremum in (15) or (16) is not a convex function of the conditional distribution $F(\mathbf{x}_1)F(\mathbf{x}_2)F(\hat{\mathbf{y}}_2|\mathbf{y}_2,\mathbf{x}_2)$ in general and can be potentially improved by (coded) time sharing [25, Remark 16.4].

Remark 4: By setting $\hat{\mathbf{Y}}_2 = \emptyset$, compress–forward reduces to direct transmission and thus $R_{\text{CF}} \geq R_{\text{DT}}$.

Remark 5: It is worthwhile to compare the cutset bound in (2), the partial decode–forward lower bound in (8), and the compress–forward lower bound in (16). The first term in the partial decode–forward lower bound is identical to the cooperative multiple access channel (MAC) bound $I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_3)$, but the second term differs from the cooperative broadcast channel (BC) bound $I(\mathbf{X}_1; \mathbf{Y}_2, \mathbf{Y}_3|\mathbf{X}_2)$ in that $\mathbf{U}$ replaces $\mathbf{Y}_2$ and there is a rate loss of $I(\mathbf{X}_1; \mathbf{U}|\mathbf{X}_2, \mathbf{Y}_2)$. In comparison, the second term of the compress–forward lower bound is of the same form to the cooperative BC bound (except for $\hat{\mathbf{Y}}_2$ in place of $\mathbf{Y}_2$), but the first term differs from the cooperative MAC bound in that there is a rate loss of $I(\mathbf{Y}_2; \hat{\mathbf{Y}}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_3)$ in addition to the loss from using the noncoherent input distribution.

We are now ready to state the main results of the paper.

Theorem 1: For every G_{21}, G_{31}, G_{32}, and P,

$$\Delta_{\text{PDF}} := R_{\text{CS}} - R_{\text{PDF}} \leq \min(t_1, r_2). \tag{17}$$

This result improves upon the constant gap of $t_1 + t_2 + r_2 + r_3$, which is achieved by distributed decode–forward [9, Remark 5] when specialized to the 3-node relay channel. Theorem 1 will be proved in Subsection III-A.

As a supplement to the additive gap result in Theorem 1, which is useful in approximating the capacity at high SNR, we establish the following multiplicative gap to provide a tighter approximation in low SNR, which will be proved in Subsection III-C.

Theorem 2: For every G_{21}, G_{31}, G_{32}, and P,

$$\frac{R_{\text{CS}}}{R_{\text{PDF}}} \leq 2. \tag{18}$$

In other words, partial decode–forward always achieves at least half the capacity.

The above results can be relaxed by using the noncoherent partial decode–forward.

Proposition 4: For every G_{21}, G_{31}, G_{32}, and P,

$$\begin{aligned}\Delta_{\text{NPDF}} &:= R_{\text{CS}} - R_{\text{NPDF}} \\ &\leq \max\big[\min(t_1, r_2),\ \min(t_1 + t_2,\ r_3)\big]\end{aligned} \tag{19}$$

and

$$\frac{R_{\text{CS}}}{R_{\text{NPDF}}} \leq 2. \tag{20}$$

Proposition 4 will be proved in Subsections III-B and III-C.

For the single-antenna case, the partial decode–forward lower bound can be shown [10, Sec. II] to be *equal* to the maximum of the decode–forward and direct–transmission lower bounds; cf. (14). With multiple antennas, however, partial decode–forward is in general much richer than decode–forward and direct transmission.

Proposition 5: If $t_1, t_2, r_2, r_3 \geq 2$,

$$\sup_{G_{21}, G_{31}, G_{32}, P} \big[R_{\text{PDF}} - \max(R_{\text{DF}}, R_{\text{DT}})\big] = \infty.$$

Proposition 5 will be proved in Subsection III-D.

As mentioned earlier, several generalizations of the compress–forward coding scheme have been shown to achieve the capacity of general MIMO relay networks within a finite number of bits, which can be then specialized back to the 3-node relay channel. In particular, the quantize–map–forward scheme by Avestimehr *et al.* [6] achieves within $12(r_2 + r_3) + 3(t_1 + t_2)$ bits from the cutset bound. This bound can be improved by specializing a recent result [18, Th. 1] for general MIMO relay networks and making it channel-independent.

Proposition 6 (Kolte et al. [18]): For every G_{21}, G_{31}, G_{32}, and P,

$$\begin{aligned}\Delta_{\text{CF}} &:= R_{\text{CS}} - R_{\text{CF}} \\ &\leq \text{DOF} \cdot \log\left(1 + \frac{t_1 + t_2}{\text{DOF}}\right) \\ &\quad + \min_{\sigma^2}\left[\frac{r_2 + r_3}{\sigma^2}\log e + \text{DOF} \cdot \log(1 + \sigma^2)\right]\end{aligned} \tag{21}$$

where $\text{DOF} = \max[\min(t_1 + t_2,\ r_3), \min(t_1,\ r_2 + r_3)]$.

In this paper, we tighten this result further as follows.

Theorem 3: For every G_{21}, G_{31}, G_{32}, and P,

$$\begin{aligned}\Delta_{\text{CF}} &\leq \min_{\sigma^2}\max\big[\min(t_1 + t_2,\ r_3) + r_2\log(1 + 1/\sigma^2), \\ &\qquad \min(t_1,\ r_2)\log(1 + \sigma^2)\big] \qquad (22)\\ &\leq \min(t_1 + t_2,\ r_3) + r_2. \qquad (23)\end{aligned}$$

This theorem will be proved in Section IV.

No multiplicative gap is known between the compress–forward lower bound and the cutset bound. This follows partly from the fact that the optimal distribution for (15) or (16) is rather difficult to characterize. It can be shown, however, that when restricted to Gaussian distributions, the compress–forward lower bound (even with time sharing) may have an unbounded multiplicative gap from the cutset bound. As a compromise, we state the following simple consequence of Remark 4 and the proof of Theorem 2.

Proposition 7: For every G_{21}, G_{31}, G_{32}, and P,

$$\frac{R_{\text{CS}}}{\max(R_{\text{DF}}, R_{\text{CF}})} \leq 2. \tag{24}$$

III. Partial Decode–Forward

In this section, we establish the results on partial decode–forward stated in the previous section (Theorem 1, Theorem 2, Proposition 4, and Proposition 5).

A. Partial Decode–Forward (Proof of Theorem 1)

We evaluate the partial decode–forward lower bound in (7) with $(\mathbf{X}_1, \mathbf{X}_2) \sim \text{CN}(0, K)$, where $K \succeq 0$ is of the form in (5), and

$$\mathbf{U} = G_{21}\mathbf{X}_1 + \mathbf{Z}_2', \tag{25}$$

where $\mathbf{Z}_2' \sim \text{CN}(0, I_{r_2})$ is independent of $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Z}_2, \mathbf{Z}_3)$. Note that $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{U}, \mathbf{Y}_3)$ has the same distribution as $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_2, \mathbf{Y}_3)$. The first term of the minimum in (7) is

$$I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_3) = \log\big|I_{r_3} + G_{3*}KG_{3*}^{\mathsf{H}}\big|. \tag{26}$$

For the second term, since

$$\begin{aligned}\mathrm{Cov}(\mathbf{X}_1|\mathbf{U},\mathbf{X}_2) &= \mathrm{Cov}(\mathbf{X}_1|\mathbf{Y}_2,\mathbf{X}_2)\\ &= K_{1|2} - K_{1|2}G_{21}^{\mathsf{H}}(I_{r_2}+G_{21}K_{1|2}G_{21}^{\mathsf{H}})^{-1}G_{21}K_{1|2}\\ &= K_{1|2}\Big(I_{t_1}+G_{21}^{\mathsf{H}}G_{21}K_{1|2}\Big)^{-1},\end{aligned}$$

we have

$$\begin{aligned}&I(\mathbf{U};\mathbf{Y}_2|\mathbf{X}_2)+I(\mathbf{X}_1;\mathbf{Y}_3|\mathbf{X}_2,\mathbf{U})\\ &= \log\frac{|I_{r_3}+G_{31}\,\mathrm{Cov}(\mathbf{X}_1|\mathbf{U},\mathbf{X}_2)G_{31}^{\mathsf{H}}|}{|I_{r_2}+G_{21}\,\mathrm{Cov}(\mathbf{X}_1|\mathbf{U},\mathbf{X}_2)G_{21}^{\mathsf{H}}|} + \log|I_{r_2}+G_{21}K_{1|2}G_{21}^{\mathsf{H}}|\\ &= \log|I_{t_1}+(G_{21}^{\mathsf{H}}G_{21}+G_{31}^{\mathsf{H}}G_{31})K_{1|2}| + \log\frac{|I_{t_1}+G_{21}^{\mathsf{H}}G_{21}K_{1|2}|}{|I_{t_1}+2G_{21}^{\mathsf{H}}G_{21}K_{1|2}|} \qquad (27)\\ &\ge \log|I_{t_1}+(G_{21}^{\mathsf{H}}G_{21}+G_{31}^{\mathsf{H}}G_{31})K_{1|2}| - \min(t_1,r_2) \qquad (28)\end{aligned}$$

where the last inequality follows by Lemma 1. Since $\min(a,b)-\min(c,d)\le\max(a-c,\,b-d)$, comparing (26) and (28) with the cutset bound in (4) completes the proof of Theorem 1.

We can prove Theorem 1 alternatively using the following result that is applicable to a more general class of relay channels and follows by setting $p(u|x_1,x_2)=p_{Y_2|X_1,X_2}(u|x_1,x_2)$ in the second form of the partial decode–forward lower bound in (8).

Proposition 8: For a discrete memoryless relay channel $p(y_2,y_3|x_1,x_2)=p(y_2|x_1,x_2)p(y_3|x_1,x_2)$,

$$\begin{aligned}\Delta_{\mathrm{PDF}} &:= R_{\mathrm{CS}}-R_{\mathrm{PDF}}\\ &\le \max_{p(x_1,x_2)} I(X_1;U|X_2,Y_2)\end{aligned}$$

where $p(u|x_1,x_2)=p_{Y_2|X_1,X_2}(u|x_1,x_2)$.

Now, we apply Proposition 8 to the MIMO relay channel by setting $\mathbf{U}$ as the form in (25). Then,

$$\begin{aligned}\Delta_{\mathrm{PDF}} &\le \sup_{F(\mathbf{x}_1,\mathbf{x}_2)} I(\mathbf{X}_1;\mathbf{U}|\mathbf{X}_2,\mathbf{Y}_2)\\ &= \sup_{F(\mathbf{x}_1,\mathbf{x}_2)} h(\mathbf{U}|\mathbf{X}_2,\mathbf{Y}_2)-h(\mathbf{Z}_2')\\ &= \max_K \log|I_{r_2}+G_{21}\,\mathrm{Cov}(\mathbf{X}_1|\mathbf{Y}_2,\mathbf{X}_2)G_{21}^{\mathsf{H}}| \qquad (29)\\ &= \max_K \log|I_{t_1}+G_{21}^{\mathsf{H}}G_{21}K_{1|2}(I_{t_1}+G_{21}^{\mathsf{H}}G_{21}K_{1|2})^{-1}|\\ &= \max_K \log\frac{|I_{t_1}+2G_{21}^{\mathsf{H}}G_{21}K_{1|2}|}{|I_{t_1}+G_{21}^{\mathsf{H}}G_{21}K_{1|2}|}\\ &\le \min(t_1,r_2),\end{aligned}$$

where the suprema are over all joint cdfs $F(\mathbf{x}_1,\mathbf{x}_2)$ such that $\mathsf{E}(\mathbf{X}_j^{\mathsf{H}}\mathbf{X}_j)\le P,\ j=1,2$, and the maxima are over all jointly Gaussian pairs $(\mathbf{X}_1,\mathbf{X}_2)$ and their covariance matrices K of the form (5). Here the equality in (29) is due to the maximum differential entropy lemma (see, for example, [25, Sec. 2.2]).

B. Noncoherent Partial Decode–Forward (Proof of the First Statement of Proposition 4)

We use the following fact.

Lemma 2: Let $K\succeq 0$ be of the form in (5). Then, for every G_{31} and G_{32}, we have

$$G_{31}K_1G_{31}^{\mathsf{H}}+G_{32}K_2G_{32}^{\mathsf{H}} \succeq G_{32}K_{12}^{\mathsf{H}}G_{31}^{\mathsf{H}}+G_{31}K_{12}G_{32}^{\mathsf{H}}. \qquad (30)$$

Proof: Consider

$$\begin{bmatrix}G_{31} & G_{32}\end{bmatrix}\begin{bmatrix}K_1 & -K_{12}\\ -K_{12}^{\mathsf{H}} & K_2\end{bmatrix}\begin{bmatrix}G_{31} & G_{32}\end{bmatrix}^{\mathsf{H}}\succeq 0.$$

■

To prove Proposition 4, let $K\succeq 0$ be of the form in (5). Let $\mathbf{X}_1\sim\mathrm{CN}(0,K_1)$ and $\mathbf{X}_2\sim\mathrm{CN}(0,K_2)$ be independent of each other, and define $\mathbf{U}$ as in (25). Then, by Lemmas 1 and 2, the first term of the minimum in the noncoherent partial decode–forward lower bound in (9) is

$$\begin{aligned}&I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_3)\\ &= \log|I_{r_3}+G_{31}K_1G_{31}^{\mathsf{H}}+G_{32}K_2G_{32}^{\mathsf{H}}| \qquad (31)\\ &\ge \log\Big|I_{r_3}+\frac{1}{2}(G_{31}K_1G_{31}^{\mathsf{H}}+G_{32}K_2G_{32}^{\mathsf{H}} + G_{32}K_{12}^{\mathsf{H}}G_{31}^{\mathsf{H}}+G_{31}K_{12}G_{32}^{\mathsf{H}})\Big| \qquad (32)\\ &\ge \log|I_{r_3}+G_{3*}KG_{3*}^{\mathsf{H}}|-\min(t_1+t_2,\,r_3). \qquad (33)\end{aligned}$$

Following steps similar to the coherent case in Subsection III-A, we have

$$\mathrm{Cov}(\mathbf{X}_1|\mathbf{U},\mathbf{X}_2)=\mathrm{Cov}(\mathbf{X}_1|\mathbf{U})=K_1(I_{t_1}+G_{21}^{\mathsf{H}}G_{21}K_1)^{-1}$$

and

$$\begin{aligned}&I(\mathbf{U};\mathbf{Y}_2|\mathbf{X}_2)+I(\mathbf{X}_1;\mathbf{Y}_3|\mathbf{X}_2,\mathbf{U})\\ &= \log\frac{|I_{r_3}+G_{31}\,\mathrm{Cov}(\mathbf{X}_1|\mathbf{U})G_{31}^{\mathsf{H}}|}{|I_{r_2}+G_{21}\,\mathrm{Cov}(\mathbf{X}_1|\mathbf{U})G_{21}^{\mathsf{H}}|} + \log|I_{r_2}+G_{21}K_1G_{21}^{\mathsf{H}}|\\ &= \log|I_{t_1}+(G_{21}^{\mathsf{H}}G_{21}+G_{31}^{\mathsf{H}}G_{31})K_1| + \log\frac{|I_{r_2}+G_{21}K_1G_{21}^{\mathsf{H}}|}{|I_{t_1}+2G_{21}^{\mathsf{H}}G_{21}K_1|}\\ &\ge \log|I_{t_1}+(G_{21}^{\mathsf{H}}G_{21}+G_{31}^{\mathsf{H}}G_{31})K_1|-\min(t_1,r_2). \qquad (34)\end{aligned}$$

Recalling $K_1\succeq K_{1|2}$ and comparing (33) and (34) with the cutset bound in (4) completes the proof.

C. Multiplicative Gap (Proofs of Theorem 2 and the Second Statement of Proposition 4)

We establish the factor-of-two gap of noncoherent partial decode–forward, which in turn implies the factor-of-two gap of (coherent) partial decode–forward in Theorem 2. By setting $\mathbf{U}=\mathbf{X}_1$ or $\emptyset$ in (9) and specializing (12) to independent (X_1,X_2), it can be readily checked that R_{NPDF} and $\max(R_{\mathrm{DF}},R_{\mathrm{DT}})$ are simultaneously lower bounded by

$$\begin{aligned}&\max\Big\{\max_{K_1,K_2}\min\big(\log|I_{r_3}+G_{31}K_1G_{31}^{\mathsf{H}}+G_{32}K_2G_{32}^{\mathsf{H}}|,\ \log|I_{r_2}+G_{21}K_1G_{21}^{\mathsf{H}}|\big),\ \max_{K_1}\log|I_{r_3}+G_{31}K_1G_{31}^{\mathsf{H}}|\Big\}\\ &= \max_{K_1,K_2}\min\Big\{\log|I_{r_3}+G_{31}K_1G_{31}^{\mathsf{H}}+G_{32}K_2G_{32}^{\mathsf{H}}|,\ \max\big(\log|I_{r_2}+G_{21}K_1G_{21}^{\mathsf{H}}|,\ \log|I_{r_3}+G_{31}K_1G_{31}^{\mathsf{H}}|\big)\Big\}. \qquad (35)\end{aligned}$$

We further bound each term in (35) from below. By (32) and the fact that $|I + A/2|^2 = |I + A + A^2/4| \geq |I + A|$ for any $A \succeq 0$, we have

$$\begin{aligned}\log|I_{r_3} &+ G_{31}K_1G_{31}^{\mathsf{H}} + G_{32}K_2G_{32}^{\mathsf{H}}| \\ &\geq \log\Big|I_{r_3} + \frac{1}{2}(G_{31}K_1G_{31}^{\mathsf{H}} + G_{32}K_2G_{32}^{\mathsf{H}} \\ &\qquad + G_{32}K_{12}^{\mathsf{H}}G_{31}^{\mathsf{H}} + G_{31}K_{12}G_{32}^{\mathsf{H}})\Big| \\ &\geq \frac{1}{2}\log|I_{r_3} + [G_{31}\ G_{32}]K[G_{31}\ G_{32}]^{\mathsf{H}}|. \end{aligned} \tag{36}$$

Similarly, since $|I+A|\cdot|I+B| \geq |I+A+B|$ for any $A, B \succeq 0$ (see [26]),

$$\begin{aligned}\max\big\{&\log|I_{t_1} + G_{21}^{\mathsf{H}}G_{21}K_1|, \log|I_{t_1} + G_{31}^{\mathsf{H}}G_{31}K_1|\big\} \\ &\geq \frac{1}{2}\Big(\log|I_{t_1} + G_{21}^{\mathsf{H}}G_{21}K_1| + \log|I_{t_1} + G_{31}^{\mathsf{H}}G_{31}K_1|\Big) \\ &\geq \frac{1}{2}\log\big|I_{t_1} + (G_{21}^{\mathsf{H}}G_{21} + G_{31}^{\mathsf{H}}G_{31})K_1\big|. \end{aligned} \tag{37}$$

Comparing (36) and (37) with the cutset bound in (4) establishes that

$$\begin{aligned} R_{\mathrm{PDF}} &\geq \max(R_{\mathrm{NPDF}}, R_{\mathrm{DF}}, R_{\mathrm{DT}}) \\ &\geq \min\{R_{\mathrm{NPDF}}, \max(R_{\mathrm{DF}}, R_{\mathrm{DT}})\} \geq \frac{1}{2}R_{\mathrm{CS}}. \end{aligned} \tag{38}$$

We can establish Theorem 2 directly based on the following result that is applicable to a more general class of relay channels.

Proposition 9: For a discrete memoryless relay channel $p(y_2, y_3|x_1, x_2) = p(y_2|x_1, x_2)p(y_3|x_1, x_2)$,

$$\frac{R_{\mathrm{CS}}}{R_{\mathrm{PDF}}} \leq 2.$$

Proof: Consider

$$\begin{aligned} I(X_1; Y_2, Y_3|X_2) &= I(X_1; Y_2|X_2) + I(X_1; Y_3|X_2, Y_2) \\ &\leq I(X_1; Y_2|X_2) + I(X_1; Y_3|X_2) \\ &\leq 2\max\big(I(X_1; Y_2|X_2),\ I(X_1; Y_3|X_2)\big), \end{aligned} \tag{39}$$

where the first inequality follows since $Y_2 \to (X_1, X_2) \to Y_3$ form a Markov chain. By setting $U = X_1$ and $U = \emptyset$ in (7), and using the inequality in (39), we have

$$\begin{aligned} R_{\mathrm{PDF}} &\geq \max_{p(x_1,x_2)} \max\big\{\min\big(I(X_1, X_2; Y_3), I(X_1; Y_2|X_2)\big), \\ &\qquad I(X_1; Y_3|X_2)\big\} \\ &= \max_{p(x_1,x_2)} \min\big\{I(X_1, X_2; Y_3), \\ &\qquad \max\big(I(X_1; Y_2|X_2),\ I(X_1; Y_3|X_2)\big)\big\} \\ &\geq \frac{1}{2}\max_{p(x_1,x_2)} \min\big\{I(X_1, X_2; Y_3), \\ &\qquad 2\max\big(I(X_1; Y_2|X_2),\ I(X_1; Y_3|X_2)\big)\big\} \\ &\geq \frac{1}{2}\max_{p(x_1,x_2)} \min\big(I(X_1, X_2; Y_3),\ I(X_1; Y_2, Y_3|X_2)\big) \\ &= \frac{1}{2}R_{\mathrm{CS}}. \end{aligned}$$

■

D. Decode–Forward and Direct Transmission (Proof of Proposition 5)

Consider the MIMO relay channel with $G_{31} = \mathrm{diag}(g, 1)$, $G_{21} = \mathrm{diag}(1, g)$, $G_{32} = \mathrm{diag}(g, g)$, $g > 1$, which is equivalent to a product of two mismatched single-antenna relay channels, one with the direct channel stronger than the sender-to-relay channel and the other in the opposite direction. Setting $K_{1|2} = K_1 = K_2 = (P/2)I_2$ in (26) and (28), we have

$$\begin{aligned} R_{\mathrm{PDF}} &\geq \min\Big\{\log\Big(1 + g^2P\Big)\Big(1 + (1+g^2)\frac{P}{2}\Big), \\ &\qquad \log\Big(1 + (1+g^2)\frac{P}{2}\Big)^2 - 2\Big\} \\ &= \log\Big(1 + (1+g^2)\frac{P}{2}\Big)^2 - 2. \end{aligned} \tag{40}$$

In comparison,

$$\begin{aligned} R_{\mathrm{DF}} = R_{\mathrm{DT}} &= \max_{P_1+P_2\leq P} \log(1 + P_1)(1 + g^2P_2) \\ &\leq \log(1 + P)(1 + g^2P). \end{aligned} \tag{41}$$

Therefore, we have

$$R_{\mathrm{PDF}} - \max(R_{\mathrm{DF}}, R_{\mathrm{DT}}) \geq \log\frac{\big(1 + (1+g^2)\frac{P}{2}\big)^2}{(1+P)(1+g^2P)} - 2$$

which tends to infinity as $g \to \infty$. Intuitively, partial decode–forward can choose different operations per antenna, the flexibility of which is missing in decode–forward and direct transmission. Based on this example, more examples of larger dimensions can be constructed.

IV. Compress–Forward

We prove Theorem 3. Let $K \succeq 0$ be of the form in (5). Let $\mathbf{X}_1 \sim \mathrm{CN}(0, K_1)$ and $\mathbf{X}_2 \sim \mathrm{CN}(0, K_2)$ be independent of each other, and

$$\hat{\mathbf{Y}}_2 = \mathbf{Y}_2 + \hat{\mathbf{Z}}_2 \tag{42}$$

where $\hat{\mathbf{Z}}_2 \sim \mathrm{CN}(0, \sigma^2 I_{r_2})$ is independent of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$. Then,

$$\begin{aligned} I(&\mathbf{Y}_2; \hat{\mathbf{Y}}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_3) \\ &= h(\hat{\mathbf{Y}}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_3) - h(\hat{\mathbf{Y}}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_2, \mathbf{Y}_3) \\ &= r_2\log(1 + 1/\sigma^2) \end{aligned} \tag{43}$$

and

$$\begin{aligned} I(&\mathbf{X}_1; \hat{\mathbf{Y}}_2, \mathbf{Y}_3|\mathbf{X}_2) \\ &= \log\frac{\left|\begin{bmatrix}(1+\sigma^2)I_{r_2} & 0\\ 0 & I_{r_3}\end{bmatrix} + G_{*1}K_1G_{*1}^{\mathsf{H}}\right|}{\left|\begin{bmatrix}(1+\sigma^2)I_{r_2} & 0\\ 0 & I_{r_3}\end{bmatrix}\right|} \\ &= \log\left|I_{r_2+r_3} + \begin{bmatrix}\frac{1}{\sqrt{1+\sigma^2}}G_{21}\\ G_{31}\end{bmatrix}K_1\begin{bmatrix}\frac{1}{\sqrt{1+\sigma^2}}G_{21}\\ G_{31}\end{bmatrix}^{\mathsf{H}}\right| \\ &= \log\left|I_{r_2+r_3} + \begin{bmatrix}\frac{1}{1+\sigma^2}A & \frac{1}{\sqrt{1+\sigma^2}}B\\ \frac{1}{\sqrt{1+\sigma^2}}B^{\mathsf{H}} & D\end{bmatrix}\right| \\ &= \log|I_{r_3} + D| \\ &\quad + \log\left|I_{r_2} + \frac{1}{1+\sigma^2}\big(A - B(I_{r_3} + D)^{-1}B^{\mathsf{H}}\big)\right| \end{aligned} \tag{44}$$

$$\geq \log |I_{r_3} + D| + \log\left|I_{r_2} + A - B(I_{r_3} + D)^{-1} B^{\mathsf{H}}\right| - \log(1 + \sigma^2)^{\min(t_1, r_2)} \quad (45)$$

$$= \log\left|I_{r_2+r_3} + \begin{bmatrix} A & B \\ B^{\mathsf{H}} & D \end{bmatrix}\right| - \min(t_1, r_2) \log(1 + \sigma^2) \quad (46)$$

$$= \log |I_{t_1} + (G_{21}^{\mathsf{H}} G_{21} + G_{31}^{\mathsf{H}} G_{31}) K_1| - \min(t_1, r_2) \log(1 + \sigma^2), \quad (47)$$

where $A = G_{21} K_1 G_{21}^{\mathsf{H}}$, $B = G_{21} K_1 G_{31}^{\mathsf{H}}$, $D = G_{31} K_1 G_{31}^{\mathsf{H}}$, (44) and (46) are due to

$$\left|\begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix}\right| = |\tilde{D}| \cdot |\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C}|,$$

and (45) follows by Lemma 1. The first statement of Theorem 3 is now established by substituting (33), (43), and (47) in the compress–forward lower bound in (16) and comparing it with the cutset bound in (4). Setting $\sigma^2 = 1$ in (22) yields the second statement in (23).

V. Computation of the Capacity Bounds

A. Formulations of Optimization Problems

1) Cutset Bound: Computing the cutset upper bound in (3) can be formulated as the following convex optimization problem [16]:

$$\begin{aligned} \text{maximize} \quad & R_{\mathrm{CS}} \\ \text{over} \quad & R_{\mathrm{CS}} \geq 0, K \succeq 0, K_{1|2} \succeq 0 \\ \text{subject to} \quad & R_{\mathrm{CS}} \leq \log\left|I_{r_3} + G_{3*} K G_{3*}^{\mathsf{H}}\right| \\ & R_{\mathrm{CS}} \leq \log\left|I_{r_2+r_3} + G_{*1} K_{1|2} G_{*1}^{\mathsf{H}}\right| \\ & \mathrm{tr}(A_1^{\mathsf{H}} K A_1) \leq P, \mathrm{tr}(A_2^{\mathsf{H}} K A_2) \leq P \\ & K - A_1 K_{1|2} A_1^{\mathsf{H}} \succeq 0 \end{aligned} \quad (48)$$

where

$$A_1 = \begin{bmatrix} I_{t_1} \\ 0_{t_2 \times t_1} \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0_{t_1 \times t_2} \\ I_{t_2} \end{bmatrix}.$$

The optimization problem in (48) can be solved by standard convex optimization techniques or packages, e.g., [27].

2) Partial Decode–Forward Lower Bound: Since direct computation of (7) or (8) is intractable, we instead consider three lower bounds on R_{PDF}, namely, R_{DF}, R_{DT}, and the special case of R_{PDF} evaluated by (25), and take the maximum of the three. Note that all three lower bounds can be viewed as the partial decode–forward lower bound evaluated by (25) with a more general choice of $\mathbf{Z}_2' \sim \mathrm{CN}(0, \sigma^2 I_{r_2})$, where $\sigma^2 = 0, \infty$, and 1, respectively. Considering more values of σ^2 can further improve the bound at the cost of complexity.

As with the cutset bound, both R_{DF} and R_{DT} can be computed efficiently as a convex optimization problem. The third bound, characterized by (26) and (27), is nonconvex. Thus, we evaluate the bound with the optimal solution to the convex optimization problem defined by (26) and (28). A similar approach can be taken for computation of R_{NPDF}.

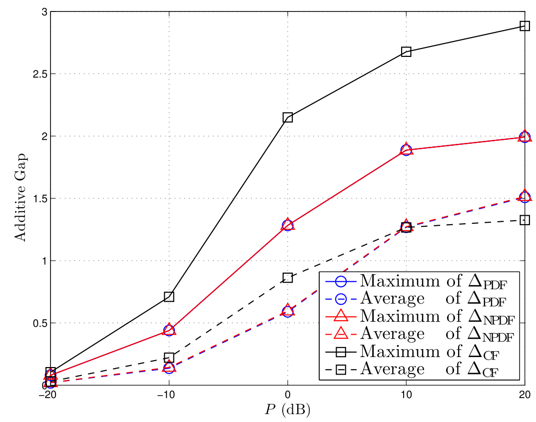

Fig. 2. The additive gaps between the cutset bound and the partial decode–forward and compress–forward lower bounds for randomly generated 2×2 MIMO relay channels.

3) Compress–Forward Lower Bound: We consider two convex lower bounds on R_{CF}, namely, the special case of R_{CF} evaluated by (42) with $\sigma^2 = 1$, namely,

$$\max_{K_1, K_2} \min \left\{ \log |I_{r_3} + G_{31} K_1 G_{31}^{\mathsf{H}} + G_{32} K_2 G_{32}^{\mathsf{H}}| - r_2, \log\left|I_{t_1} + \left(\frac{1}{2} G_{21}^{\mathsf{H}} G_{21} + G_{31}^{\mathsf{H}} G_{31}\right) K_1\right| \right\}$$

and R_{DT} (which corresponds to $\sigma^2 = \infty$). As in the case of partial decode–forward, considering more values of σ^2 can further improve the bound at the cost of complexity.

B. Numerical Results

We consider the additive and multiplicative gaps on 2000 2×2 MIMO relay channels with random channel gains independently distributed according to $\mathrm{CN}(0, 1)$. The gaps are evaluated by relaxed bounds discussed in the previous subsection. The maximum and average of the additive gaps plotted against the power constraint P at the sender and the relay are shown in Fig. 2 and similar multiplicative gaps are shown in Fig. 3. The simulation results are consistent with the theoretical predictions in Theorems 1, 2, and 3, and Proposition 4.

VI. Half-Duplex MIMO Relay Channels

Half-duplex relay channel models are often investigated to study wireless communication systems in which relays cannot send and receive in the same time slot or frequency band. In this section, we consider the time-division half-duplex relay channel model proposed by Kramer [21], which has been recently extended to MIMO relay networks by Cardone *et al.* [23], with minor modifications. In this model, the relay has a total of a_2 antennas, each of which can either transmit or receive at a given time. Let $S_i \in \{0, 1\}$, $i = 1, \ldots, a_2$, denote the operation of antenna i at the relay, say, $S_i = 0$ means the antenna is in the receiving mode and $S_i = 1$ means the antenna is in the transmitting mode. Then, the channel outputs at the relay and the receiver can be expressed as

$$\begin{aligned} \mathbf{Y}_2 &= \bar{D}(\mathbf{S}) G_{21} \mathbf{X}_1 + \mathbf{Z}_2, \\ \mathbf{Y}_3 &= G_{31} \mathbf{X}_1 + G_{32} D(\mathbf{S}) \mathbf{X}_2 + \mathbf{Z}_3, \end{aligned} \quad (49)$$

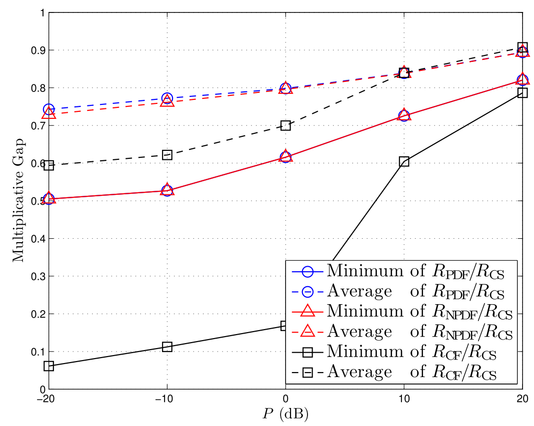

Fig. 3. The multiplicative gaps between the cutset bound and the partial decode–forward and compress–forward lower bounds for randomly generated 2×2 MIMO relay channels.

where $G_{21} \in \mathbb{C}^{a_2 \times t_1}$, $G_{31} \in \mathbb{C}^{r_3 \times t_1}$, and $G_{32} \in \mathbb{C}^{r_3 \times a_2}$ are channel gain matrices, $\mathbf{X}_1 \in \mathbb{C}^{t_1}$ and $\mathbf{X}_2 \in \mathbb{C}^{a_2}$ are the respective inputs at the sender and the relay, and $\mathbf{Z}_2 \sim \mathrm{CN}(0, I_{a_2})$ and $\mathbf{Z}_3 \sim \mathrm{CN}(0, I_{r_3})$ are independent complex Gaussian noise components. Here and henceforth, we use the shorthand notation $D(\mathbf{S}) = \mathrm{diag}(\mathbf{S})$ and $\bar{D}(\mathbf{S}) = I_{a_2} - D(\mathbf{S})$ to represent the antenna mode vector $\mathbf{S} = (S_1, \ldots, S_{a_2})$ and its complement in matrix forms.

Assume that the relay can dynamically choose the antenna modes based on its channel outputs. Then this model can be viewed as a special case of the general relay channel with the equivalent channel input $(\mathbf{X}_2, \mathbf{S})$ at the relay. This observation leads to the following upper bound on the capacity.

Proposition 10 (Cutset Bound [21]): The capacity C of the time-division half-duplex MIMO relay channel in (49) *is upper bounded by*

$$\begin{aligned} R_{\mathrm{CS}} &= \sup \min\big\{I(\mathbf{X}_1, \mathbf{X}_2, \mathbf{S}; \mathbf{Y}_3),\ I(\mathbf{X}_1; \mathbf{Y}_2, \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{S})\big\} \\ &= \sup \min\big\{I(\mathbf{S}; \mathbf{Y}_3) + I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_3 | \mathbf{S}), \\ &\qquad I(\mathbf{X}_1; \mathbf{Y}_2, \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{S})\big\}, \end{aligned} \tag{50}$$

where the suprema are over all pmfs $p(\mathbf{s})$ and conditional cdfs $F(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{s})$ such that $\mathsf{E}(\mathbf{X}_1^{\mathsf{H}} \mathbf{X}_1) \le P$ and $\mathsf{E}(\mathbf{X}_2^{\mathsf{H}} D(\mathbf{S}) \mathbf{X}_2) \le P$. This bound can be further upper bounded by

$$\sup_{p(\mathbf{s}), K(\mathbf{s})} \min(a_2 + J_1,\ J_2), \tag{51}$$

where the supremum is over all pmfs $p(\mathbf{s})$ and $(t_1 + a_2) \times (t_1 + a_2)$ matrix-valued functions $K(\mathbf{s})$ such that

$$K(\mathbf{s}) = \begin{bmatrix} K_1(\mathbf{s}) & K_{12}(\mathbf{s}) \\ K_{12}^{\mathsf{H}}(\mathbf{s}) & K_2(\mathbf{s}) \end{bmatrix} \succeq 0, \quad \mathbf{s} \in \{0,1\}^{a_2}, \tag{52}$$

$$\mathsf{E}[\mathrm{tr}(K_1(\mathbf{S}))] = \sum_{\mathbf{s}} p(\mathbf{s}) \, \mathrm{tr}(K_1(\mathbf{s})) \le P, \tag{53}$$

$$\mathsf{E}[\mathrm{tr}(D(\mathbf{S}) K_2(\mathbf{S}))] = \sum_{\mathbf{s}} p(\mathbf{s}) \, \mathrm{tr}(D(\mathbf{s}) K_2(\mathbf{s})) \le P. \tag{54}$$

Here,

$$\begin{aligned} J_1 &= \mathsf{E}\big(\log |I_{r_3} + [G_{31}\ G_{32} D(\mathbf{S})] K(\mathbf{S}) [G_{31}\ G_{32} D(\mathbf{S})]^{\mathsf{H}}|\big), \\ J_2 &= \mathsf{E}\big(\log |I_{t_1} + (G_{21}^{\mathsf{H}} \bar{D}(\mathbf{S}) G_{21} + G_{31}^{\mathsf{H}} G_{31}) K_{1|2}(\mathbf{S})|\big), \end{aligned}$$

and $K_{1|2}(\mathbf{S}) = K_1(\mathbf{S}) - K_{12}(\mathbf{S}) K_2^{-1}(\mathbf{S}) K_{12}^{\mathsf{H}}(\mathbf{S})$.

The partial decode–forward coding scheme can achieve the following lower bound.

Proposition 11 (Partial Decode–Forward Lower Bound [21]): The capacity C of the time-division half-duplex MIMO relay channel is lower bounded by

$$\begin{aligned} &R_{\mathrm{PDF}} \\ &\quad = \sup \min\{I(\mathbf{X}_1, \mathbf{X}_2, \mathbf{S}; \mathbf{Y}_3), \\ &\qquad\qquad I(\mathbf{U}; \mathbf{Y}_2 | \mathbf{X}_2, \mathbf{S}) + I(\mathbf{X}_1; \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{U}, \mathbf{S})\} \\ &\quad = \sup \min\big\{I(\mathbf{X}_1, \mathbf{X}_2, \mathbf{S}; \mathbf{Y}_3), \\ &\qquad\qquad I(\mathbf{X}_1; \mathbf{U}, \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{S}) - I(\mathbf{X}_1; \mathbf{U} | \mathbf{X}_2, \mathbf{Y}_2, \mathbf{S})\big\} \end{aligned}$$

where the suprema are over all pmfs $p(\mathbf{s})$ and conditional cdfs $F(\mathbf{u}, \mathbf{x}_1, \mathbf{x}_2 | \mathbf{s})$ such that $\mathsf{E}(\mathbf{X}_1^{\mathsf{H}} \mathbf{X}_1) \le P$ and $\mathsf{E}(\mathbf{X}_2^{\mathsf{H}} D(\mathbf{S}) \mathbf{X}_2) \le P$.

Similarly, the compress–forward coding scheme can achieve the following lower bound.

Proposition 12 (Compress–Forward Lower Bound [22]): The capacity C of the time-division half-duplex MIMO relay channel is lower bounded by

$$\begin{aligned} &R_{\mathrm{CF}} \\ &\quad = \sup \min\{I(\mathbf{X}_1, \mathbf{X}_2, \mathbf{S}; \mathbf{Y}_3) - I(\mathbf{Y}_2; \hat{\mathbf{Y}}_2 | \mathbf{X}_1, \mathbf{X}_2, \mathbf{S}, \mathbf{Y}_3), \\ &\qquad\qquad I(\mathbf{X}_1; \hat{\mathbf{Y}}_2, \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{S})\} \\ &\quad \ge \sup \min\{I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_3 | \mathbf{S}) - I(\mathbf{Y}_2; \hat{\mathbf{Y}}_2 | \mathbf{X}_1, \mathbf{X}_2, \mathbf{S}, \mathbf{Y}_3), \\ &\qquad\qquad I(\mathbf{X}_1; \hat{\mathbf{Y}}_2, \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{S})\}, \end{aligned} \tag{55}$$

where the suprema are over all pmfs $p(\mathbf{s})$ and conditional cdfs $F(\mathbf{x}_1|\mathbf{s}) F(\mathbf{x}_2|\mathbf{s}) F(\hat{\mathbf{y}}_2|\mathbf{y}_2, \mathbf{x}_2, \mathbf{s})$ such that $\mathsf{E}(\mathbf{X}_1^{\mathsf{H}} \mathbf{X}_1) \le P$ and $\mathsf{E}(\mathbf{X}_2^{\mathsf{H}} D(\mathbf{S}) \mathbf{X}_2) \le P$.

Comparing the cutset upper bound and the partial decode–forward lower bound, we can establish the following gap result.

Proposition 13: For every G_{21}, G_{31}, G_{32}, and P,

$$\Delta_{\mathrm{PDF}} := C - R_{\mathrm{PDF}} \le \min(t_1, a_2) \tag{56}$$

and

$$\frac{R_{\mathrm{CS}}}{R_{\mathrm{PDF}}} \le 2. \tag{57}$$

Proof: The multiplicative gap in (57) follows immediately from Proposition 9. For the additive gap in (56), we adapt Proposition 8 by setting $\mathbf{U} = \bar{D}(\mathbf{S}) G_{21} \mathbf{X}_1 + \mathbf{Z}_2'$, where $\mathbf{Z}_2' \sim \mathrm{CN}(0, I_{a_2})$ is independent of $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{S}, \mathbf{Z}_2, \mathbf{Z}_3)$. Then,

$$\begin{aligned} &\Delta_{\mathrm{PDF}} \\ &\quad \le \sup_{p(\mathbf{s}) F(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{s})} I(\mathbf{X}_1; \mathbf{U} | \mathbf{X}_2, \mathbf{Y}_2, \mathbf{S}) \\ &\quad = \sup_{p(\mathbf{s}) F(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{s})} h(\mathbf{U} | \mathbf{X}_2, \mathbf{Y}_2, \mathbf{S}) - h(\mathbf{U} | \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_2, \mathbf{S}) \\ &\quad = \sup_{p(\mathbf{s}) F(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{s})} h(\mathbf{U} | \mathbf{X}_2, \mathbf{Y}_2, \mathbf{S}) - h(\mathbf{Z}_2') \\ &\quad = \sup_{p(\mathbf{s}) F(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{s})} \sum_{\mathbf{s}} p(\mathbf{s}) (h(\mathbf{U} | \mathbf{X}_2, \mathbf{Y}_2, \mathbf{S} = \mathbf{s}) - h(\mathbf{Z}_2')). \end{aligned}$$

Now following arguments similar to those in Subsection III-A, we have for each $\mathbf{s}$

$$h(\mathbf{U}|\mathbf{X}_2, \mathbf{Y}_2, \mathbf{S} = \mathbf{s}) - h(\mathbf{Z}_2')$$
$$\le \sup_{K(\mathbf{s})} \log \frac{|I_{t_1} + 2G_{21}^{\mathsf{H}} \bar{D}(\mathbf{s}) G_{21} K_{1|2}(\mathbf{s})|}{|I_{t_1} + G_{21}^{\mathsf{H}} \bar{D}(\mathbf{s}) G_{21} K_{1|2}(\mathbf{s})|}$$
$$\le \min(t_1, a_2),$$

where the supremum is over all covariance matrices $K(\mathbf{s})$ of the form (52). This completes the proof of (56). ∎

An additive gap that is somewhat weaker than (56) can be established by comparing the cutset upper bound and the compress–forward lower bound, which still improves upon the existing gap result of $1.96(t_1 + a_2 + r_3)$ in [23].

Proposition 14: For every G_{21}, G_{31}, G_{32}, and P,

$$\Delta_{\mathrm{CF}} = C - R_{\mathrm{CF}}$$
$$\le \min_{\sigma^2} \max\Big[\min\Big(t_1 + \max\Big[0,\ a_2 \log \frac{2\sigma^2}{1+\sigma^2}\Big],\ r_3\Big) + a_2 \log(2(1 + 1/\sigma^2)),$$
$$\min(t_1, a_2) \log(1 + \sigma^2)\Big] \tag{58}$$
$$\le \min(t_1 + 2a_2,\ a_2 \log 3 + r_3). \tag{59}$$

Proof: Let $\hat{\mathbf{Y}}_2 = \bar{D}(\mathbf{S})\mathbf{Y}_2 + \hat{\mathbf{Z}}_2$, where $\hat{\mathbf{Z}}_2 \sim \mathrm{CN}(0, \sigma^2 I_{a_2})$ is independent of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{S}$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$. Consider any $(p(\mathbf{s}), K(\mathbf{s}))$ satisfying (52)–(54). Given $\{\mathbf{S} = \mathbf{s}\}$, let $\mathbf{X}_1 \sim \mathrm{CN}(0, K_1(\mathbf{s}))$ and $\mathbf{X}_2 \sim \mathrm{CN}(0, K_2(\mathbf{s}))$ be conditionally independent. Then, we have

$$I(\mathbf{X}_1; \hat{\mathbf{Y}}_2, \mathbf{Y}_3 | \mathbf{X}_2, \mathbf{S} = \mathbf{s})$$
$$= \log\left|I_{r_2+r_3} + \begin{bmatrix} \frac{\bar{D}(\mathbf{s})G_{21}}{\sqrt{1+\sigma^2}} \\ G_{31} \end{bmatrix} K_1(\mathbf{s}) \begin{bmatrix} \frac{\bar{D}(\mathbf{s})G_{21}}{\sqrt{1+\sigma^2}} \\ G_{31} \end{bmatrix}^{\mathsf{H}}\right|$$
$$\ge \log|I_{t_1} + (G_{21}^{\mathsf{H}} \bar{D}(\mathbf{s}) G_{21} + G_{31}^{\mathsf{H}} G_{31}) K_1(\mathbf{s})|$$
$$- \min(t_1, r_2) \log(1 + \sigma^2), \tag{60}$$

where $t_2 = \sum_{i=1}^{a_2} s_i$ denotes the number of transmitting antennas and $r_2 = a_2 - t_2$ denotes the number of receiving antennas when $\mathbf{s} \in \{0,1\}^{a_2}$ is the antenna mode vector. Similarly,

$$I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_3 | \mathbf{S} = \mathbf{s})$$
$$= \log|I_{r_3} + G_{31} K_1(\mathbf{s}) G_{31}^{\mathsf{H}} + G_{32} D(\mathbf{s}) K_2(\mathbf{s}) D(\mathbf{s}) G_{32}^{\mathsf{H}}|$$
$$\ge \log|I_{r_3} + [G_{31}\ \ G_{32}D(\mathbf{s})] K(\mathbf{s}) [G_{31}\ \ G_{32}D(\mathbf{s})]^{\mathsf{H}}|$$
$$- \min(t_1 + t_2, r_3), \tag{61}$$

and

$$I(\mathbf{Y}_2; \hat{\mathbf{Y}}_2 | \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_3, \mathbf{S} = \mathbf{s}) = r_2 \log(1 + 1/\sigma^2). \tag{62}$$

Subtracting (62) from (61) and comparing it with $a_2 + J_1$ in (51) before taking the expectation, we can obtain the gap

$$\min(t_1 + t_2,\ r_3) + r_2 \log(1 + 1/\sigma^2) + a_2$$
$$\le \min\Big(t_1 + t_2 \log \frac{2\sigma^2}{1+\sigma^2},\ r_3\Big) + a_2 \log(2(1 + 1/\sigma^2))$$
$$\le \min\Big(t_1 + \max\Big[0,\ a_2 \log \frac{2\sigma^2}{1+\sigma^2}\Big],\ r_3\Big)$$
$$+ a_2 \log(2(1 + 1/\sigma^2)). \tag{63}$$

Similarly, comparing (60) with J_2 in (51), we obtain the gap

$$\min(t_1, a_2) \log(1 + \sigma^2). \tag{64}$$

Combining (63) and (64) and considering all possible σ^2, we can establish (58). Finally, setting $\sigma^2 = 2$ in (58) yields (59). ∎

As a final remark, we note that similar additive and multiplicative gap results can be established for fixed time-division and frequency-division half-duplex relay channel models [6], [10], [25, Secs. 16.6.3 and 16.8]. We refer the reader to [6, Th. 8.3] and an earlier conference version [24] of this paper for such results.

ACKNOWLEDGMENTS

The authors would like to thank the Associate Editor and the anonymous reviewers for many valuable comments that helped improve the presentation of this paper.

REFERENCES

[1] E. C. van der Meulen, “The discrete memoryless channel with two senders and one receiver,” in *Proc. 2nd Int. Symp. Inf. Theory*, Tsahkadsor, Armenian, 1971, pp. 103–135.
[2] 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.
[3] L. R. Ford, Jr., and D. R. Fulkerson, “Maximal flow through a network,” *Canad. J. Math.*, vol. 8, no. 3, pp. 399–404, 1956.
[4] Y.-H. Kim. (2014). *Dictionary of Relaying Schemes*. [Online]. Available: http://circuit.ucsd.edu/yhk/relaying.html
[5] 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.
[6] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
[7] M. H. Yassaee and M. R. Aref, “Slepian–Wolf coding over cooperative relay networks,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3462–3482, Jun. 2011.
[8] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, “Noisy network coding,” *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.
[9] S. H. Lim, K. T. Kim, and Y.-H. Kim, “Distributed decode–forward for relay networks,” *IEEE Trans. Inf. Theory*. [Online]. Available: http://arxiv.org/pdf/1510.00832v1.pdf
[10] A. El Gamal, M. Mohseni, and S. Zahedi, “Bounds on capacity and minimum energy-per-bit for AWGN relay channels,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.
[11] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, “A deterministic approach to wireless relay networks,” in *Proc. 45th Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Sep. 2007, pp. 26–28.
[12] W. Chang, S.-Y. Chung, and Y. H. Lee. (Nov. 2010). “Gaussian relay channel capacity to within a fixed number of bits.” [Online]. Available: https://arxiv.org/abs/1011.5065
[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[14] B. Wang, J. Zhang, and A. Høst-Madsen, “On the capacity of MIMO relay channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.
[15] S. Simoens, O. Munoz-Medina, J. Vidal, and A. del Coso, “On the Gaussian MIMO relay channel with full channel state information,” *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3588–3599, Sep. 2009.
[16] C. T. K. Ng and G. J. Foschini, “Transmit signal and bandwidth optimization in multiple-antenna relay channels,” *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2987–2992, Nov. 2011.

[17] L. Gerdes, C. Hellings, L. Weiland, and W. Utschick, "On the maximum achievable partial decode-and-forward rate for the Gaussian MIMO relay channel," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6751–6758, Dec. 2015.

[18] R. Kolte, A. Özgür, and A. E. Gamal, "Capacity approximations for Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4721–4734, Jun. 2015.

[19] A. Host-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.

[20] 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.

[21] G. Kramer, "Models and theory for relay channels with receive constraints," in *Proc. 42nd Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Sep./Oct. 2004, pp. 1312–1321.

[22] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "On the Gaussian half-duplex relay channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2542–2562, May 2014.

[23] M. Cardone, D. Tuninetti, and R. Knopp, "Gaussian MIMO half-duplex relay networks: Approximate optimality of simple schedules," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 1267–1271.

[24] X. Jin and Y.-H. Kim, "The approximate capacity of the MIMO relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 2102–2106.

[25] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[26] M. Fiedler, "Bounds for the determinant of the sum of hermitian matrices," *Proc. Amer. Math. Soc.*, vol. 30, no. 1, pp. 27–31, 1971.

[27] M. Grant and S. Boyd. (Sep. 2013). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.0 Beta*. [Online]. Available: http://cvxr.com/cvx

Xianglan Jin (M'12) received her B.S. and M.S. degrees in telecommunications engineering from Beijing University of Posts and Telecommunications, China, in 1999 and 2002, respectively, and her Ph.D. degree in electrical engineering and computer science from Seoul National University, Korea, in 2008. She was an engineer at Beijing Samsung Telecommunications R&D center from April 2002 to June 2004. From 2008 to 2010, she was a postdoctoral scholar in the Department of Electrical Engineering and Computer Science, Seoul National University. From 2010 to 2014, she was a Research Assistant Professor in the Department of Information and Communication Engineering at Dongguk University, Korea. She visited the Department of Electrical and Computer Engineering at the University of California, San Diego from 2015 to 2016. Currently, she is a Research Professor in the Department of Electrical and Computer Engineering at Pusan National University, Korea. Her research interests include MIMO, cooperative communication, and network information theory.

Young-Han Kim (S'99–M'06–SM'12–F'15) received his B.S. degree with honors in electrical engineering from Seoul National University, Korea, in 1996 and his M.S. degrees in electrical engineering and in statistics, and his Ph.D. degree in electrical engineering from Stanford University in 2001, 2006, and 2006, respectively. In 2006, he joined the University of California, San Diego, where he is currently an Associate Professor in the Department of Electrical and Computer Engineering. His research interests are in information theory, communication engineering, and data science. He coauthored the book Network Information Theory (Cambridge University Press 2011). He is a recipient of the 2008 NSF Faculty Early Career Development (CAREER) Award, the 2009 US-Israel Binational Science Foundation Bergmann Memorial Award, the 2012 IEEE Information Theory Paper Award, and the 2015 IEEE Information Theory Society James L. Massey Research and Teaching Award for Young Scholars. He served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and a Distinguished Lecturer for the IEEE Information Theory Society.