

# Converse Results for the Downlink Multicell Processing With Finite Backhaul Capacity

Tianyu Yang, Nan Liu<sup>id</sup>, Wei Kang<sup>id</sup>, and Shlomo Shamai (Shitz)<sup>id</sup>

**Abstract**—In this paper, we study outer bounds on the capacity region of the downlink multicell processing model with finite backhaul capacity for the simple case of two base stations and two mobile users. It is modeled as a two-user multiple access diamond channel. It consists of a first hop from the central processor to the base stations via orthogonal links of finite capacity and the second hop from the base stations to the mobile users via a Gaussian interference channel. The outer bound is derived using the converse tools of the multiple access diamond channel and that of the Gaussian MIMO broadcast channel. Through numerical results, it is shown that our outer bound improves upon the existing outer bounds greatly in the medium backhaul capacity range, and as a result, the gap between the outer bounds and the rate of the time-sharing of the known achievable schemes is significantly reduced.

**Index Terms**—Multi-cell processing, diamond channel, converse, MIMO broadcast channel.

## I. INTRODUCTION

THE multi-cell processing system, as reviewed in [1], has been used to increase the throughput and to cope with the inter-cell interference. The downlink multi-cell processing system, when first considered, consists of different base stations linked to the central processor via backhaul links of *unlimited* capacity, and therefore, the amount of cooperation among the different base stations is unbounded. This network can be modeled by a MIMO broadcast channel and the sum-rate characterization was found in [2]. Later on, due to the impracticality of unlimited capacity backhaul links, Simeone *et al.* [3], Park *et al.* [4], Hong and Caire [5],

Manuscript received June 9, 2017; revised July 23, 2018; accepted September 2, 2018. Date of publication September 13, 2018; date of current version December 19, 2018. This work was supported in part by the National Natural Science Foundation of China under Grants 61571123, 61571122, and 61521061, in part by the Six Talent Peaks Project in Jiangsu Province, in part by the National Major Project under Grant 2017ZX03001002-004, and in part by the Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant 2018A03. S. Shamai (Shitz) was supported by the European Union's Horizon 2020 Research and Innovation Programme under Grant 694630. This paper was presented in part at the 2017 IEEE International Symposium on Information Theory.

T. Yang and W. Kang are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: tianyu@seu.edu.cn; wkang@seu.edu.cn).

N. Liu is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: nanliu@seu.edu.cn).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion Israel Institute of Technology, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

Communicated by M. Costa, Associate Editor for Shannon Theory.

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.2018.2869881

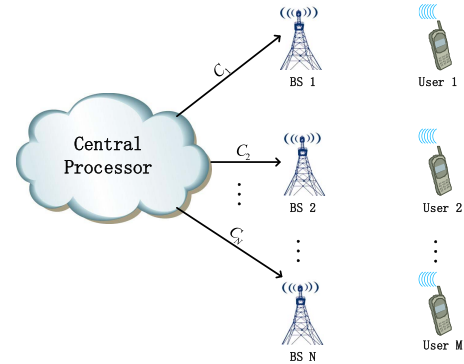


Fig. 1. The multicell processing system with finite backhaul capacity.

Liu and Kang [6], and Yi and Liu [7] studied the problem of finding the capacity region of the downlink multicell processing system when the capacities of the backhaul links are *finite*, see Fig. 1. The papers proposed various achievable schemes to efficiently utilize the finite capacity backhaul links. More specifically, in [3], a compressed dirty-paper coding scheme is proposed, where the base stations are treated as the antennas of the central processor and the dirty-paper coding codewords for each antenna are compressed and transmitted on the backhaul links. The scheme is improved in [4] by allowing the quantization noise of the base stations be correlated. The scheme of reverse compute-and-forward was proposed in [5] where linear precoding is performed at the central processor and the backhaul links are used to transmit linear combinations of the messages over a finite field. Such linear precoding transforms the channel seen at each mobile user into a point-to-point channel where integer-valued interference is eliminated by precoding and the remaining noninteger residual interference is treated as noise. By regarding the network model as a multi-user diamond channel, an achievability scheme is proposed in [6] and [7] by combining Marton's achievability for the broadcast channel [8] and the achievability of sending correlated codewords over a multiple access diamond channel [9], [10].

The outer bound on the capacity region for this network is unknown except for the simple cut-set bound [11] on the sum rate, which is the minimum of the capacity between the first hop from the central processor to the base stations and that of the second hop from the base stations to the mobile users. Utilizing the simple cut-set bound, some constant-gap results for the multicell processing system with finite backhaul capacity

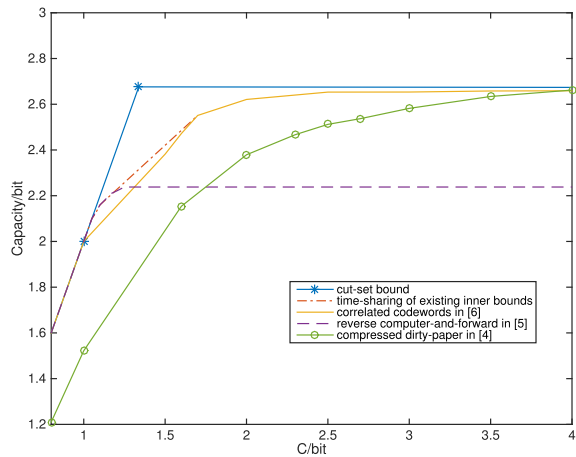


Fig. 2. Existing upper and lower bounds on the sum capacity for the symmetric case.

have recently been obtained [12]–[14]. When the capacity of the backhaul links are relatively large, the achievable sum rate of the scheme of compressed dirty-paper coding approaches that of the simple cut-set bound. On the other hand, when the capacity of the backhaul links are relatively small, the scheme of reverse compute-and-forward reaches the simple cut-set bound [6]. In the medium capacity region, there is still a relatively large gap between the simple cut-set upper bound and the sum rate of the time-sharing of the known achievable schemes, see Fig. 2. So it is unknown how well the proposed achievable schemes are and whether further efforts are needed in proposing better achievable schemes than existing ones for the downlink multicell processing system.

In this paper, we derive a novel outer bound on the capacity region of the downlink multicell processing network consisting of two base stations and two users. Similar to [6], we regard the network as a 2-user multiple access diamond channel. As a result, the converse is derived using the converse tools of the multiple access diamond channel used in [15] and [16] and that of the Gaussian MIMO broadcast channel used in [17]. The derived outer bound is expressed in terms of the capacity region of the Gaussian MIMO broadcast channel given input covariance constraint, which has been found in [17]–[19].

Comparing numerically the proposed outer bound, the cut-set bound and the performance of various achievable schemes for the multicell processing system in terms of the sum-rate, we see that our outer bound improves upon the cut-set bound greatly in the medium backhaul capacity range, and as a result, the gap between the outer bounds and the time-sharing of the known achievable schemes is significantly reduced.

The remainder of this paper is organized as follows. In Section II, we provide the system model. In Section III, we derive a new outer bound for the capacity region as well as an upper bound on the sum capacity. The proof of the main result is provided in Section IV. Numerical results are provided in Section V, and this is followed by conclusions in Section VI.

## II. SYSTEM MODEL

In this paper, we consider the downlink multicell processing system with two base stations and two users. This network

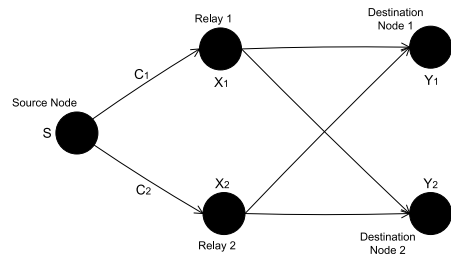


Fig. 3. The 2-user multiple access diamond channel.

model can be seen as the 2-user Gaussian multiple access diamond channel [6], see Fig. 3. The source node (central processor) can communicate with Relays (base stations) 1 and 2 via backhaul links of capacities  $C_1$  and  $C_2$ , respectively. The channel between the two relay nodes and the two destination nodes (mobile users) is characterized by

$$Y_1 = X_1 + aX_2 + U_1, \quad (1)$$

$$Y_2 = bX_1 + X_2 + U_2, \quad (2)$$

where  $Y_1$  and  $Y_2$  are the received signal at Users 1 and 2, respectively,  $X_1$  and  $X_2$  are the input signals from Relays 1 and 2, respectively,  $U_1$ ,  $U_2$  are two independent zero-mean unit-variance Gaussian random variables that are independent to  $(X_1, X_2)$ , and  $a, b \in \mathbb{R}$  are the channel gains from Relay 1 to Destination 2 and Relay 2 to Destination 1, respectively. The transmitted signals at the two relays must satisfy the average power constraints: for any  $x_k^n$  that Relay  $k$  sends into the channel, it must satisfy

$$\frac{1}{n} \sum_{i=1}^n x_{ki}^2 \leq P_k, \quad k = 1, 2.$$

Let  $W_1$  and  $W_2$  be two independent messages that the source node would like to transmit to Destination nodes (mobile users) 1 and 2, respectively. Assume that  $W_k$  is uniformly distributed on  $\{1, 2, \dots, M_k\}$ ,  $k = 1, 2$ .

An  $(M_1, M_2, n, \epsilon_n)$  code consists an encoding function at source node:

$$f^n : \{1, 2, \dots, M_1\} \times \{1, 2, \dots, M_2\} \\ \rightarrow \{1, 2, \dots, 2^{nC_1}\} \times \{1, 2, \dots, 2^{nC_2}\},$$

two encoding functions at the relay nodes:

$$f_k^n : \{1, 2, \dots, 2^{nC_k}\} \rightarrow \mathcal{X}_k, \quad k = 1, 2,$$

and two decoding functions at the destination nodes:

$$g_k^n : \mathcal{Y}_k \rightarrow \{1, 2, \dots, M_k\}, \quad k = 1, 2.$$

The average probability of error is defined as

$$\epsilon_n = \sum_{w_1=1}^{M_1} \sum_{w_2=1}^{M_2} \frac{1}{M_1 M_2} \Pr[g_1^n(Y_1^n) \neq w_1 \text{ or} \\ g_2^n(Y_2^n) \neq w_2 | W_1 = w_1, W_2 = w_2].$$

Rate pair  $(R_1, R_2)$  is said to be achievable if there exists a sequence of  $(2^{nR_1}, 2^{nR_2}, n, \epsilon_n)$  code such that  $\epsilon_n \rightarrow 0$

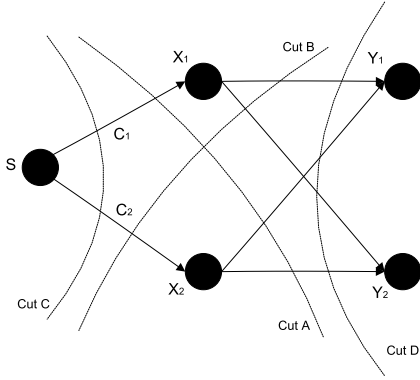


Fig. 4. Cuts for the 2-user Gaussian multiple access diamond channel.

as  $n \rightarrow \infty$ . The capacity region of the 2-user Gaussian multiple access diamond channel is the closure of the set of all achievable rates pairs.

### III. AN OUTER BOUND FOR THE CAPACITY REGION OF THE 2-USER GAUSSIAN MULTIPLE ACCESS DIAMOND CHANNEL

The existing *simple* cut-set upper bound on the sum capacity is

$$R_1 + R_2 \leq \min\{C_1 + C_2, \max_{\rho \in [-1, 1]} C_{\text{MIMO}}^{\text{sum}}(\rho)\}, \quad (3)$$

where  $C_{\text{MIMO}}^{\text{sum}}(\rho) \triangleq \max_{(R_1, R_2) \in \mathcal{C}_{\text{MIMO}}(\rho)} R_1 + R_2$ , and  $\mathcal{C}_{\text{MIMO}}(\rho)$  denotes the capacity region of the broadcast channel described in (1) and (2) under a given covariance constraint, where  $\bar{\mathbf{X}} \triangleq [X_1, X_2]^T$  is the transmitted signal of the 2 antennas of the transmitter, and  $Y_1$  and  $Y_2$  are the received signals of the single-antenna Receivers 1 and 2, respectively. The input of the transmitter must satisfy a covariance constraint, i.e.,

$$E[\bar{\mathbf{X}}\bar{\mathbf{X}}^T] \preceq \begin{bmatrix} P_1 & \rho\sqrt{P_1P_2} \\ \rho\sqrt{P_1P_2} & P_2 \end{bmatrix} \triangleq \mathbf{K}. \quad (4)$$

where  $\mathbf{A} \preceq \mathbf{B}$  means that the matrix  $\mathbf{B} - \mathbf{A}$  is positive semi-definite. The capacity region of the MIMO broadcast channel, i.e.,  $\mathcal{C}_{\text{MIMO}}(\rho)$ , has been found in [17]–[19]. Note that the existing simple cut-set bound, i.e., (3), is the minimum of the capacity of Cuts C and D of Fig. 4.

To improve upon the existing simple cut-set bound, our proposed outer bound on the capacity region of the 2-user Gaussian multiple access diamond channel needs the following definitions:

- 1) Since we are considering the entire capacity region and not just the sum capacity, we generalize  $C_{\text{MIMO}}^{\text{sum}}(\rho)$  to the definitions of  $C_{\text{MIMO}12}^v(\rho)$  and  $C_{\text{MIMO}21}^v(\rho)$  as: for  $\rho \in [-1, 1]$  and  $v \in [1, \infty)$ , define

$$C_{\text{MIMO}12}^v(\rho) \triangleq \max_{(R_1, R_2) \in \mathcal{C}_{\text{MIMO}}(\rho)} R_1 + vR_2,$$

$$C_{\text{MIMO}21}^v(\rho) \triangleq \max_{(R_1, R_2) \in \mathcal{C}_{\text{MIMO}}(\rho)} R_2 + vR_1,$$

By varying over  $v$ ,  $C_{\text{MIMO}12}^v(\rho)$  and  $C_{\text{MIMO}21}^v(\rho)$  trace the capacity region of the 2-user broadcast channel

described in (1) and (2) under the given covariance constraint in (4), where  $\bar{\mathbf{X}} \triangleq [X_1, X_2]^T$  is the transmitted signal of the 2 antennas of the transmitter.

- 2) Since we are considering the cross-cuts A and B of Fig. 4, as well as Cuts C and D in the existing simple cut-set bound, we define for  $\rho \in [-1, 1]$

$$f_A(\rho) \triangleq C_1 + \frac{1}{2} \log \left( 1 + \max\{a^2, 1\}(1 - \rho^2)P_2 \right),$$

$$f_B(\rho) \triangleq C_2 + \frac{1}{2} \log \left( 1 + \max\{b^2, 1\}(1 - \rho^2)P_1 \right),$$

where  $f_A(\rho)$  and  $f_B(\rho)$  can be intuitively seen as the capacity of Cuts A and B of Fig. 4, respectively, when the transmitted signal of the two base stations are jointly Gaussian with correlation  $\rho$ .

- 3) Define  $f_C(\rho)$  as

$$f_C(\rho) \triangleq C_1 + C_2 - \frac{1}{2} \log \frac{1}{1 - \rho^2},$$

Note that the cut-set bound for Cut C is  $C_1 + C_2 = f_C(0)$ . We would like to improve upon this bound by proving that when the transmitted signals of the second hop becomes correlated with correlation coefficient  $\rho$ , the rate of the first cut decreases from  $C_1 + C_2$  to  $C_1 + C_2 - \frac{1}{2} \log \frac{1}{1 - \rho^2} = f_C(\rho)$ . The intuition is that when independent information is sent via the two orthogonal links of the first hop, the rate sent over the first hop is  $C_1 + C_2$ . To increase the rate sent over the second hop, we would like the two relays to send correlated codewords through the Gaussian channel. In order to do this, correlated information has to be sent via the two orthogonal links of the first hop. As a result, the information rate sent through the first hop can not be as high as  $C_1 + C_2$ , and we will show that instead, the rate sent over the first hop decreases from  $f_C(0)$  to  $f_C(\rho)$ . We are able to prove the decrease from  $f_C(0)$  to  $f_C(\rho)$  only for small enough  $\rho$  that satisfy  $\rho \in \mathcal{A}_x$ , where

$$\mathcal{A}_x = \begin{cases} [0, \rho_x] & \text{if } x \geq 0 \\ [\rho_x, 0] & \text{if } x < 0, \end{cases} \quad x = a, b.$$

and  $\rho_x$  is defined as

$$\rho_x = \begin{cases} \text{sgn}(x) \left( \sqrt{1 + \frac{1}{4x^2 P_1 P_2}} - \sqrt{\frac{1}{4x^2 P_1 P_2}} \right), & x \neq 0 \\ 0, & x = 0 \end{cases} \quad (5)$$

where  $\text{sgn}(\cdot)$  is the sign function of  $\cdot$ .

With the above definitions, we are ready to state the main result of this paper.

*Theorem 1: The rate pair  $(R_1, R_2)$  is achievable only if it satisfies for  $\mu \geq 1$ ,*

$$\begin{aligned} & R_1 + \mu R_2 \\ & \leq \max \left\{ \max_{\rho \in \mathcal{A}_b} \min\{T_A^{12}(\rho), T_B^{12}(\rho)\}, \max_{\rho \in [-1, 1] \cap \mathcal{A}_a^c} T_A^{12}(\rho) \right\} \end{aligned} \quad (6)$$

and

$$R_2 + \mu R_1 \leq \max \left\{ \max_{\rho \in \mathcal{A}_a} \min \{T_A^{21}(\rho), T_B^{21}(\rho)\}, \max_{\rho \in [-1,1] \cap \mathcal{A}_a^c} T_A^{21}(\rho) \right\}, \quad (7)$$

where

$$T_A^m(\rho) \triangleq \min_{\kappa \in [0,1]} \left[ \kappa \min \{f_A(\rho), f_B(\rho), f_C(0)\} + (1-\kappa) C_{MIMO m}^{\frac{\mu-\kappa}{1-\kappa}}(\rho) \right], \quad m = 12, 21, \quad (8)$$

$$T_B^m(\rho) \triangleq \min_{\substack{\lambda \in [0, \min\{\frac{\mu}{2}, 1\}] \\ \lambda \neq 1}} \left[ \lambda f_C(\rho) + (1-\lambda) C_{MIMO m}^{\frac{\mu-\lambda}{1-\lambda}}(\rho) \right], \quad m = 12, 21.$$

*Proof:* The proof of Theorem 1 is provided in Section IV. ■

As can be seen, the outer bound on the capacity region of the 2-user multiple access diamond channel in Theorem 1 is expressed in terms of the capacity region of the Gaussian MIMO broadcast channel with input covariance constraint. For completeness, we write the expressions for  $C_{MIMO12}^v(\rho)$  and  $C_{MIMO21}^v(\rho)$  here [17]–[19]:

$$C_{MIMO12}^v(\rho) = \max_{\mathbf{0} \leq \mathbf{K}_1 \leq \mathbf{K}} \left[ \frac{1}{2} \log(\mathbf{G}_1^T \mathbf{K}_1 \mathbf{G}_1 + 1) + \frac{v}{2} \log \frac{\mathbf{G}_2^T \mathbf{K} \mathbf{G}_2 + 1}{\mathbf{G}_2^T \mathbf{K}_1 \mathbf{G}_2 + 1} \right]$$

$$C_{MIMO21}^v(\rho) = \max_{\mathbf{0} \leq \mathbf{K}_2 \leq \mathbf{K}} \left[ \frac{v}{2} \log \frac{\mathbf{G}_1^T \mathbf{K} \mathbf{G}_1 + 1}{\mathbf{G}_1^T \mathbf{K}_2 \mathbf{G}_1 + 1} + \frac{1}{2} \log(\mathbf{G}_2^T \mathbf{K}_2 \mathbf{G}_2 + 1) \right]$$

which is achieved by dirty-paper coding. Here,  $\mathbf{G}_1^T = [1 \ a]$ ,  $\mathbf{G}_2^T = [b \ 1]$  and  $\mathbf{K}$  is defined in (4).

*Remarks:*

- 1) In Theorem 1, the upper bound of  $T_A^m(\rho)$ ,  $m = 12, 21$ , is proved using the cut-set bound from the four cuts, i.e., Cuts A, B, C and D of Fig. 4. The more difficult part is to prove that when  $\rho \in \mathcal{A}_b$  ( $\rho \in \mathcal{A}_a$ ), we also have the upper bound of  $T_B^m(\rho)$ ,  $m = 12, 21$ . This upper bound is mainly the improvement from  $f_C(0) = C_1 + C_2$  to  $f_C(\rho) = C_1 + C_2 - \frac{1}{2} \log \frac{1}{1-\rho^2}$ , which indicate that when correlated signals of correlation  $\rho$  are sent through the first hop, the rate sent across the first hop is reduced from  $C_1 + C_2$  to  $C_1 + C_2 - \frac{1}{2} \log \frac{1}{1-\rho^2}$ . The converse techniques we use to prove this include 1) the bounding of the correlation between the transmitted signals of the two relays via an auxiliary random variable [15], [16], which was inspired by Ozarow in solving the Gaussian multiple description problem [20]; 2) the single-letterization technique from [21, pp. 314, eq. (3.34)]; 3) a generalization of the entropy power inequality (EPI) [19, Corollary 4]; and 4) the derivation of the capacity region of the two-user Gaussian

MIMO broadcast channel with private messages in [17, Sec. III.A].

- 2) When  $C_1$  and  $C_2$  goes to infinity,  $f_A(\rho)$ ,  $f_B(\rho)$ ,  $f_C(\rho)$  which include the terms  $C_1$  and/or  $C_2$  also go to infinity. Therefore, the minimum over  $\kappa$  for  $T_A^{12}(\rho)$  is achieved at  $\kappa = 0$ . Similarly, the minimum over  $\lambda$  for  $T_B^{12}(\rho)$  is achieved at  $\lambda = 0$ . Hence,  $T_A^{12}(\rho) = T_B^{12}(\rho) = C_{MIMO12}^\mu(\rho)$ . Thus, the upper bound in (6) becomes

$$R_1 + \mu R_2 \leq \max_{\rho \in [-1,1]} C_{MIMO12}^\mu(\rho),$$

and similarly, we have

$$R_2 + \mu R_1 \leq \max_{\rho \in [-1,1]} C_{MIMO21}^\mu(\rho).$$

Hence, Theorem 1 coincides with the result of the capacity region of the MIMO broadcast channel where Relays 1 and 2 are the antennas of the MIMO transmitter, with individual antenna power constraints of  $P_1$  and  $P_2$ , respectively.

From Theorem 1, we may obtain the following corollary which characterizes an upper bound for the *sum capacity* of the 2-user Gaussian multiple access diamond channel.

*Corollary 1:* An upper bound for the sum capacity of the 2-user Gaussian multiple access diamond channel is

$$R_1 + R_2 \leq \max \left\{ \max_{\rho \in [-1,1] \cap \mathcal{A}_x^c} T_1(\rho), \max_{\rho \in \mathcal{A}_x} T_2(\rho) \right\}, \quad \text{for } x = a \text{ and } x = b, \quad (9)$$

where  $T_1(\rho)$  and  $T_2(\rho)$  are defined as

$$T_1(\rho) = \min \{f_A(\rho), f_B(\rho), f_C(0), C_{MIMO}^{\text{sum}}(\rho)\},$$

$$T_2(\rho) = \min \left\{ T_1(\rho), \frac{1}{2} (f_C(\rho) + C_{MIMO}^{\text{sum}}(\rho)) \right\}.$$

*Proof:* In Theorem 1, take  $\mu = 1$ . We have

$$T_A^m(\rho) \leq C_{MIMO m}^1(\rho), \quad m = 12, 21$$

when we take  $\kappa = 0$  in the minimum of  $\kappa$  in (8). When we take  $\kappa \rightarrow 1$ , we have

$$T_A^m(\rho) \leq \lim_{\kappa \rightarrow 1} \left[ \kappa \min \{f_A(\rho), f_B(\rho), f_C(0)\} + (1-\kappa) C_{MIMO m}^1(\rho) \right]$$

$$= \min \{f_A(\rho), f_B(\rho), f_C(0)\}, \quad m = 12, 21.$$

Taking  $\lambda = \frac{1}{2}$ , we have

$$T_B^m(\rho) \leq \frac{1}{2} f_C(\rho) + \frac{1}{2} C_{MIMO m}^1(\rho), \quad m = 12, 21,$$

Noting that  $C_{MIMO m}^1(\rho) = C_{MIMO}^{\text{sum}}(\rho)$ ,  $m = 12, 21$ , we have proved Corollary 1. ■

Comparing the result of Corollary 1 to the existing simple cut-set bound in (3), we see that Corollary 1 implies that the upper bound of

$$R_1 + R_2 \leq \max_{\rho \in [-1,1]} T_1(\rho) \quad (10)$$

is true since we have  $T_1(\rho) \geq T_2(\rho)$  for  $\forall \rho \in [-1, 1]$ . The upper bound of (10) is tighter than the existing simple cut-set bound of (3), as it further considers the capacities of

Cuts A and B. Furthermore, the result of Corollary 1, i.e., (9), is strictly tighter than the cut-set bound in (10) because when  $\rho \in \mathcal{A}_x$ ,  $x = a, b$ , the upper bound of  $T_2(\rho)$  exists, which is smaller than  $T_1(\rho)$ . Thus, Corollary 1 provides a novel upper bound that is tighter than the existing simple cut-set bound of (3) on the sum-capacity of the channel.

To conclude this section, we make some discussions about the generalization of our upper bound of Theorem 1. By utilizing the same techniques as those used in the proof of Theorem 1, one can generalize the result of Theorem 1 to the case of 2 users and  $N$  relays,  $N \geq 3$ . The improvement over the simple cut-set bound is still two folds: 1) improve upon the simple cut-set bound by considering all possible cuts, and 2) improve the cut-set bound of the first-hop from  $\sum_{i=1}^N C_i$  to  $\sum_{i=1}^N C_i - \left[ I(X_1, X_2) + \sum_{i=3}^N I(X_i; X_1, X_2, \dots, X_{i-1}) \right]$ , with the mutual informations evaluated for jointly Gaussian random variables  $X_1, X_2, \dots, X_N$ . Note the similarity of the  $N$ -relay case,  $N \geq 3$  and the 2-relay case where the cut-set bound of the first hop is improved from  $C_1 + C_2$  to  $C_1 + C_2 - \frac{1}{2} \log \frac{1}{1-\rho^2} = C_1 + C_2 - I(X_1; X_2)$ , where the mutual information is evaluated for jointly Gaussian random variables  $X_1$  and  $X_2$ . We expect the improvement from  $\sum_{i=1}^N C_i$  to  $\sum_{i=1}^N C_i - \left[ I(X_1, X_2) + \sum_{i=3}^N I(X_i; X_1, X_2, \dots, X_{i-1}) \right]$  to scale linearly with the number of antennas, as  $\left[ I(X_1, X_2) + \sum_{i=3}^N I(X_i; X_1, X_2, \dots, X_{i-1}) \right]$  evaluated for the jointly Gaussian distribution scales linearly with  $N$ .

The idea and techniques of improving the rate upper bound of the first-hop from  $\sum_{i=1}^N C_i$  to  $\sum_{i=1}^N C_i - \left[ I(X_1, X_2) + \sum_{i=3}^N I(X_i; X_1, X_2, \dots, X_{i-1}) \right]$  can be used in other related multi-cell processing scenarios. For example, the recent proposition of F-RAN [22] over C-RAN is to save the fronthaul usage by enabling caching and signal processing capabilities at the base station. The technique presented in this paper may provide a better characterization on the upper bound of the capacity of the F-RAN system. Another example is the wiretapped diamond channel problem [23], which considers a multicell processing problem such as the one depicted in Fig. 3 except that Destination node 2 is an eavesdropper, eavesdropping on the message requested by Destination node 1, and does not require any message of its own from the base stations. We believe the technique in this work will be useful in deriving an upper bound on the secret capacity in the wiretapped diamond channel problem.

For the more general case of  $N$  relays and  $M$  users, where  $M \geq 3$ , to generalize the proof and results of Theorem 1, we need the derivation of [17, Sec. III.A] to hold for the  $M$ -user Gaussian MIMO broadcast channel,  $M \geq 3$ . Since [17, Sec. III.A] is valid for two users only, a converse that is tighter than the cut-set bound is still open for the case of  $N$  relays and  $M$  users,  $M \geq 3$ .

#### IV. PROOF OF THEOREM 1

Without loss of generality, we may assume  $a > 0$ . Since the case of  $a < 0$  can easily be transformed into the case of  $a > 0$  by defining  $X'_2 = -X_2$  and  $Y'_2 = -Y_2$ . This results in

the following equivalent channel

$$Y_1 = X_1 - aX'_2 + U_1, \quad (11)$$

$$Y'_2 = -bX_1 + X'_2 + U_2. \quad (12)$$

where  $-a$  is positive. The channel in (11) and (12) is equivalent to the channel in (1) and (2) because the transformation is a one-to-one mapping.

For any sequence of  $(2^{nR_1}, 2^{nR_2}, n, \epsilon_n)$  code, let  $X_k^n$  denote the input of Relay  $k$  into the  $n$  uses of the channel  $p(y_1, y_2 | x_1, x_2)$ , and  $Y_k^n$  denote the corresponding output received at Receiver  $k$ ,  $k = 1, 2$ . Due to the power constraint, we have

$$\frac{1}{n} \sum_{i=1}^n E[X_{ki}^2] \leq P_k, \quad k = 1, 2 \quad (13)$$

Define a random variable  $Q$  that is independent of everything else and uniformly distributed on  $\{1, 2, \dots, n\}$ , further define

$$X_1 \triangleq X_{1Q}, \quad X_2 \triangleq X_{2Q}, \quad Y_1 \triangleq Y_{1Q}, \quad Y_2 \triangleq Y_{2Q}. \quad (14)$$

Define the correlation coefficient between  $X_1$  and  $X_2$  as

$$\rho \triangleq \frac{E[X_1 X_2]}{\sqrt{E[X_1^2] E[X_2^2]}}.$$

Note that  $\rho \in [-1, 1]$ . Further define

$$\bar{P}_k \triangleq E[X_k^2], \quad k = 1, 2.$$

From (13) and (14), we have

$$\begin{aligned} \bar{P}_k &= E[X_k^2] = E[X_{kQ}^2] = E_Q \left[ E_{X|Q} \left[ X_{kQ}^2 | Q = i \right] \right] \\ &= \frac{1}{n} \sum_{i=1}^n E[X_{ki}^2] \leq P_k, \quad k = 1, 2. \end{aligned} \quad (15)$$

Define

$$\rho^* \triangleq \frac{\sqrt{\bar{P}_1 \bar{P}_2}}{\sqrt{P_1 P_2}} \rho. \quad (16)$$

Based on (15), we have

$$|\rho^*| \leq |\rho|. \quad (17)$$

Hence,  $\rho^* \in [-1, 1]$ . Define  $\mathbf{X} \triangleq [X_1 \ X_2]^T$ , and further define  $\mathbf{K}$  as

$$\mathbf{K} \triangleq \begin{bmatrix} P_1 & \rho^* \sqrt{P_1 P_2} \\ \rho^* \sqrt{P_1 P_2} & P_2 \end{bmatrix}.$$

We can see that

$$\begin{aligned} E[\mathbf{X}\mathbf{X}^T] &= \begin{bmatrix} \bar{P}_1 & \rho \sqrt{\bar{P}_1 \bar{P}_2} \\ \rho \sqrt{\bar{P}_1 \bar{P}_2} & \bar{P}_2 \end{bmatrix} \\ &\leq \begin{bmatrix} P_1 & \rho \sqrt{P_1 P_2} \\ \rho \sqrt{P_1 P_2} & P_2 \end{bmatrix} \\ &= \begin{bmatrix} P_1 & \rho^* \sqrt{P_1 P_2} \\ \rho^* \sqrt{P_1 P_2} & P_2 \end{bmatrix} = \mathbf{K}. \end{aligned} \quad (18)$$

We first derive a modified cut-set bound that considers the cross-cuts, i.e., Cuts A and B of Fig. 4, as well as the cut of the first and second hop, i.e., Cuts C and D. We give this result in the following Lemma.

*Lemma 1:* When  $\rho^*$  satisfies  $\rho^* \in [-1, 1]$ , we have

$$R_1 + R_2 \leq \min \{f_A(\rho^*), f_B(\rho^*), f_C(0)\} \quad (19)$$

The proof of Lemma 1 is provided in Appendix A, where converse techniques from the physically degraded broadcast channel [11] are used.

Next, we will use (19) to derive an upper bound on  $R_1 + \mu R_2$ , for  $\mu \geq 1$ . The case of  $\mu \leq 1$  can be derived similarly by swapping the indices 1 and 2 and the numbers  $b$  and  $a$ .

For any  $\kappa \in [0, 1]$ , we have

$$\begin{aligned} R_1 + \mu R_2 &= \kappa(R_1 + R_2) + (1 - \kappa) \left( R_1 + \frac{\mu - \kappa}{1 - \kappa} R_2 \right) \\ &\leq \kappa(R_1 + R_2) + (1 - \kappa) C_{\text{MIMO12}}^{\frac{\mu - \kappa}{1 - \kappa}}(\rho^*), \end{aligned}$$

where the last step follows because  $C_{\text{MIMO12}}^{\frac{\mu - \kappa}{1 - \kappa}}(\rho^*)$  corresponds to the upper bound on  $R_1 + \frac{\mu - \kappa}{1 - \kappa} R_2$  when  $C_1, C_2 \rightarrow \infty$ . Using the upper bound for sum-rate  $R_1 + R_2$  in (19), we have

$$\begin{aligned} R_1 + \mu R_2 &\leq \min_{\kappa \in [0, 1]} \left[ \kappa \min \{f_A(\rho^*), f_B(\rho^*), f_C(0)\} \right. \\ &\quad \left. + (1 - \kappa) C_{\text{MIMO12}}^{\frac{\mu - \kappa}{1 - \kappa}}(\rho^*) \right] = T_A^{12}(\rho^*). \quad (20) \end{aligned}$$

Note that (20) is valid when  $\rho^*$  satisfies  $\rho^* \in [-1, 1]$ . We remark here that we did not consider Cut D of Fig. 4 in our derivations because this cut would give us the upper bound of  $C_{\text{MIMO12}}^\mu(\rho^*)$ , which we would obtain by setting  $\kappa = 0$  in (20). Hence, it is redundant to consider this cut.

We now proceed to derive another upper bound on  $R_1 + \mu R_2$  which is valid when  $\rho^*$  satisfies  $\rho^* \in \mathcal{A}_b$ . Using Fano's inequality, we have

$$\begin{aligned} n(R_1 + R_2) &\leq I(W_1, W_2; Y_1^n, Y_2^n) + n\epsilon_n \\ &\leq I(X_1^n, X_2^n; Y_1^n, Y_2^n) + n\epsilon_n \quad (21) \\ &\leq nC_1 + nC_2 - I(X_1^n; X_2^n) + n\epsilon_n, \quad (22) \end{aligned}$$

where (21) follows from the Markov Chain  $(W_1, W_2) \rightarrow (X_1^n, X_2^n) \rightarrow (Y_1^n, Y_2^n)$ , and (22) follows from [15, eq. (33)]. Using Fano's inequality, we further have

$$\begin{aligned} n(R_1 + R_2) &\leq H(W_1, W_2) \\ &= H(W_2) + H(W_1|W_2) \\ &\leq I(W_2; Y_2^n) + I(W_1; Y_1^n|W_2) + H(W_2|Y_2^n) \\ &\quad + H(W_1|Y_1^n, W_2) \\ &\leq I(W_2; Y_2^n) + I(W_1; Y_1^n|W_2) + 2n\epsilon_n \\ &\leq I(W_2; Y_2^n) + I(X_1^n, X_2^n; Y_1^n|W_2) + 2n\epsilon_n, \quad (23) \end{aligned}$$

and

$$\begin{aligned} nR_2 &= H(W_2) \\ &= I(W_2; Y_2^n) + H(W_2|Y_2^n) \\ &\leq I(W_2; Y_2^n) + n\epsilon_n. \quad (24) \end{aligned}$$

Thus, omitting the  $\epsilon_n$  term which will go to zero and  $n \rightarrow \infty$ , from (22), (23) and (24), we have that for any

$\lambda \in [0, \min(\frac{\mu}{2}, 1)]$  and  $\lambda \neq 1$ ,

$$\begin{aligned} &n((1 - \lambda)R_1 + (\mu - \lambda)R_2) \\ &= n((1 - \lambda)(R_1 + R_2) + (\mu - 1)R_2) \\ &\leq (1 - \lambda)I(X_1^n, X_2^n; Y_1^n|W_2) + (\mu - \lambda)I(W_2; Y_2^n) \quad (25) \\ &= (1 - \lambda)I(X_1^n, X_2^n; Y_1^n|W_2) + (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n) \\ &\quad - (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n|W_2), \quad (26) \end{aligned}$$

where (25) follows from (23), (24) and the fact that  $\lambda \leq 1$  and  $\mu \geq 1$ , and (26) is because of  $W_2 \rightarrow (X_1^n, X_2^n) \rightarrow (Y_1^n, Y_2^n)$  forms a Markov chain. Furthermore, we have

$$\begin{aligned} &n(R_1 + \mu R_2) \\ &= n\lambda(R_1 + R_2) + n(1 - \lambda)R_1 + n(\mu - \lambda)R_2 \\ &\leq \lambda(nC_1 + nC_2 - I(X_1^n; X_2^n)) \\ &\quad + n(1 - \lambda)R_1 + n(\mu - \lambda)R_2 \quad (27) \\ &\leq n\lambda(C_1 + C_2) - \lambda I(X_1^n; X_2^n) \\ &\quad + (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n) + (1 - \lambda)I(X_1^n, X_2^n; Y_1^n|W_2) \\ &\quad - (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n|W_2) \quad (28) \\ &\leq (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n) - \lambda I(X_1^n, X_2^n; Z^n) \\ &\quad + \lambda I(X_1^n; Z^n|X_2^n) + \lambda I(X_2^n; Z^n|X_1^n) \\ &\quad + n\lambda(C_1 + C_2) + (1 - \lambda)I(X_1^n, X_2^n; Y_1^n|W_2) \\ &\quad - (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n|W_2), \quad (29) \end{aligned}$$

where (27) follows from (22) and the fact that  $\lambda \geq 0$ , (28) follows from (26), and (29) follows by introducing a sequence of auxiliary random variables  $Z^n$  and utilizing the fact that [15, eq. (34)]

$$I(X_1^n; X_2^n) = I(X_1^n, X_2^n; Z^n) - I(X_2^n; Z^n|X_1^n) - I(X_1^n; Z^n|X_2^n).$$

where  $Z^n$  is the output of the following memoryless Gaussian channel with  $Y_2^n$  being the input:

$$Z = Y_2 + U_3, \quad (30)$$

and  $U_3$  is a Gaussian random variable with zero mean and variance  $N_3$ . Further define  $Z \triangleq Z_Q$ .

We single-letterize (29) by single-letterizing three terms using the following Lemma.

*Lemma 2:* We have the following single-letterization:

$$\begin{aligned} &(\mu - \lambda)I(X_1^n, X_2^n; Y_2^n) - \lambda I(X_1^n, X_2^n; Z^n) \\ &\leq n(\mu - 2\lambda)I(X_1, X_2; Y_2) + n\lambda I(X_1, X_2; Y_2|Z), \end{aligned}$$

and

$$\begin{aligned} &\lambda(I(X_1^n; Z^n|X_2^n) + I(X_2^n; Z^n|X_1^n)) \\ &\leq n\lambda(I(X_1; Z|X_2) + I(X_2; Z|X_1)) \end{aligned}$$

and

$$\begin{aligned} &I(X_1^n, X_2^n; Y_1^n|W_2) - \frac{\mu - \lambda}{1 - \lambda} I(X_1^n, X_2^n; Y_2^n|W_2) \\ &\leq n \left[ I(X_1, X_2; Y_1|U) - \frac{\mu - \lambda}{1 - \lambda} I(X_1, X_2; Y_2|U) \right]. \end{aligned}$$

where

$$U \triangleq (W_2, Y_1^{Q-1}, Y_2^{Q+1}, Q). \quad (31)$$

The proof of Lemma 2 is provided in Appendix B, where techniques from [21, pp. 314, eq. (3.34)] is used.

Then from (29), Lemma 2, and the fact that  $\lambda \leq \min(\frac{\mu}{2}, 1)$ , we obtain the following single-letterization:

$$\begin{aligned} R_1 + \mu R_2 &\leq \lambda(C_1 + C_2) + \lambda I(X_1; Z|X_2) + \lambda I(X_2; Z|X_1) \\ &\quad + (1 - \lambda)I(X_1, X_2; Y_1|U) - (\mu - \lambda)I(X_1, X_2; Y_2|U) \\ &\quad + (\mu - 2\lambda)I(X_1, X_2; Y_2) + \lambda I(X_1, X_2; Y_2|Z), \quad (32) \end{aligned}$$

where the mutual informations are evaluated using the joint distribution of defined random variables  $(X_1, X_2, Y_1, Y_2, Z, U)$  which satisfy

$$\begin{aligned} p(x_1, x_2, y_1, y_2, z, u) \\ = p(x_1, x_2)p(u|x_1, x_2)p(y_1, y_2|x_1, x_2)p(z|y_2). \quad (33) \end{aligned}$$

Next, we derive an upper bound on (32) by using the fact that the term  $p(y_1, y_2|x_1, x_2)$  in (33), which refers to the channel in (1) and (2), and the term  $p(z|y_2)$  in (33), which refers to the channel in (30), are Gaussian channels. To derive an upper bound on (32), we provide an upper bound for the following three terms:

- 1)  $(\mu - 2\lambda)I(X_1, X_2; Y_2) + \lambda I(X_1, X_2; Y_2|Z)$ ;
- 2)  $I(X_1; Z|X_2) + I(X_2; Z|X_1)$ ;
- 3)  $I(X_1, X_2; Y_1|U) - \frac{\mu - \lambda}{1 - \lambda}I(X_1, X_2; Y_2|U)$ .

The upper bounds for the first two terms make use of standard Gaussian techniques, and are given by the following Lemma.

*Lemma 3: We have*

$$\begin{aligned} &(\mu - 2\lambda)I(X_1, X_2; Y_2) + \lambda I(X_1, X_2; Y_2|Z) \\ &\leq \frac{\mu - \lambda}{2} \log \left( b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 \right) \\ &\quad - \frac{\lambda}{2} \log \left( \frac{b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 + N_3}{1 + N_3} \right), \end{aligned}$$

and

$$\begin{aligned} &I(X_1; Z|X_2) + I(X_2; Z|X_1) \\ &\leq \frac{1}{2} \log \frac{(1 - \rho^{*2})b^2 P_1 + 1 + N_3}{1 + N_3} \\ &\quad + \frac{1}{2} \log \frac{(1 - \rho^{*2})P_2 + 1 + N_3}{1 + N_3}. \end{aligned}$$

The proof of Lemma 3 is provided in Appendix C, where the extremal inequality of [19, Corollary 4] is used.

As for the third term  $I(X_1, X_2; Y_1|U) - \frac{\mu - \lambda}{1 - \lambda}I(X_1, X_2; Y_2|U)$ , since  $\frac{\mu - \lambda}{1 - \lambda} \geq 1$ , we have:

$$\begin{aligned} &I(X_1, X_2; Y_1|U) - \frac{\mu - \lambda}{1 - \lambda}I(X_1, X_2; Y_2|U) \\ &\leq \sup_{\substack{p(u, x_1, x_2): \\ E[\mathbf{X}\mathbf{X}^T] \leq \mathbf{K}}} \left( I(X_1, X_2; Y_1|U) - \frac{\mu - \lambda}{1 - \lambda}I(X_1, X_2; Y_2|U) \right) \quad (34) \end{aligned}$$

$$= C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*) - \frac{\mu - \lambda}{1 - \lambda} \max_{p(x_1, x_2): E[\mathbf{X}\mathbf{X}^T] \leq \mathbf{K}} I(\mathbf{X}; Y_2) \quad (35)$$

$$\begin{aligned} &= C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*) \\ &\quad - \frac{\mu - \lambda}{2(1 - \lambda)} \log(b^2 P_1 + P_2 + 1 + 2b\rho^* \sqrt{P_1 P_2}), \quad (36) \end{aligned}$$

where (34) follows because  $U, X_1$  and  $X_2$  defined in (14) and (31) satisfy the constraint of the optimization in (34) due to (18), and according to [17, Sec. III.A], we have (35).

From (32), Lemma 3, and (36), we have:

$$\begin{aligned} R_1 + \mu R_2 &\leq \lambda(C_1 + C_2) + \frac{\lambda}{2} \log \frac{(1 - \rho^{*2})b^2 P_1 + 1 + N_3}{1 + N_3} \\ &\quad + \frac{\lambda}{2} \log \frac{(1 - \rho^{*2})P_2 + 1 + N_3}{1 + N_3} + (1 - \lambda)C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*) \\ &\quad - \frac{\lambda}{2} \log \left( \frac{b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 + N_3}{1 + N_3} \right) \\ &\quad + \frac{\mu - \lambda}{2} \log \left( b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 \right) \\ &\quad - \frac{\mu - \lambda}{2} \log(b^2 P_1 + P_2 + 1 + 2b\rho^* \sqrt{P_1 P_2}) \\ &= \frac{\lambda}{2} \log \frac{((1 - \rho^{*2})b^2 P_1 + 1 + N_3)((1 - \rho^{*2})P_2 + 1 + N_3)}{(1 + N_3)(b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 + N_3)} \\ &\quad + \lambda(C_1 + C_2) + (1 - \lambda)C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*) \quad (37) \end{aligned}$$

The above is true for any  $N_3 \geq 0$ . Take  $N_3$  as

$$N_3 = b\sqrt{P_1 P_2} \left( \frac{1}{\rho^*} - \rho^* \right) - 1. \quad (38)$$

It can be seen that the value of  $N_3$  in (38) is non-negative because we consider the case where  $\rho^*$  satisfies  $\rho^* \in \mathcal{A}_b$ . Plugging (38) into (37), we obtain

$$\begin{aligned} R_1 + \mu R_2 &\leq \lambda \left( C_1 + C_2 - \frac{1}{2} \log \frac{1}{1 - \rho^{*2}} \right) + (1 - \lambda)C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*) \\ &= \lambda f_C(\rho^*) + (1 - \lambda)C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*). \end{aligned}$$

The above is true for any  $\lambda \in [0, \min(\frac{\mu}{2}, 1)]$  and  $\lambda \neq 1$ . Thus, we have

$$\begin{aligned} R_1 + \mu R_2 &\leq \min_{\substack{\lambda \in [0, \min(\frac{\mu}{2}, 1)] \\ \lambda \neq 1}} \left[ \lambda f_C(\rho^*) + (1 - \lambda)C_{\text{MIMO12}}^{\frac{\mu - \lambda}{1 - \lambda}}(\rho^*) \right] \\ &= T_B^{12}(\rho^*) \quad (39) \end{aligned}$$

when  $\rho^*$  satisfies  $\rho^* \in \mathcal{A}_b$ . From (20) and (39), we have proved (6) of Theorem 1. The result of (7) of Theorem 1 can be derived similarly by swapping the indices 1 and 2 and changing the number  $b$  to  $a$ .

## V. NUMERICAL RESULTS

To illustrate the tightness of the derived upper bound in Theorem 1, we plot and compare the upper and lower bounds on the sum capacity. More specifically, we plot the existing simple cut-set upper bound on the sum capacity in (3), the new cut-set upper bound of (10) implied by our result in Corollary 1, the new upper bound of Corollary 1, and the achievable sum rates of existing schemes for the 2-user Gaussian multiple access diamond channel.

The results are shown in Fig. 5 for the symmetric case of  $a = b = 0.9$ ,  $P_1 = P_2 = 10$  and  $C_1 = C_2 = C$ . We only plot the region of  $C \in [1, 3]$ , since this is the

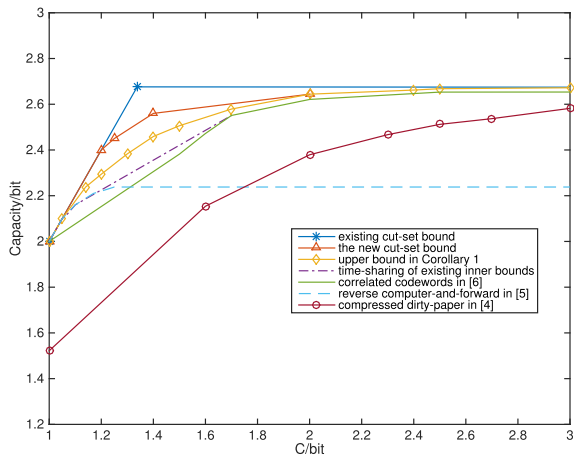


Fig. 5. Upper and lower bounds on the sum capacity for the case of  $a = b = 0.9$ ,  $P_1 = P_2 = 10$  and  $C_1 = C_2 = C$ .

interesting case where the existing simple cut-set upper bound and the existing lower bounds on the sum capacity do not meet. As can be seen, in the region of  $C \in [1.2, 2.55]$ , the new cut-set bound of (10) improves upon the existing simple cut-set bound of (3), which means that in this region, it is beneficial to consider the cross-cuts in the cut-set bound, i.e., Cuts A and B of Fig. 4. In the region of  $C \in [1.05, 2]$ , the upper bound of Corollary 1 improves upon the new cut-set bound of (10), which means that in this region, the derived upper bound of  $\frac{1}{2}(f_C(\rho) + C_{\text{MIMO}}^{\text{sum}}(\rho))$  for  $\rho \in \mathcal{A}_x$ ,  $x = a, b$  is useful. Overall, in the region of  $C \in [1.05, 2.55]$ , our new upper bound of Corollary 1 improves upon the existing simple cut-set upper bound strictly. Furthermore, in the region of  $C \in [1.05, 2]$ , the improvement is rather significant.

The sum rate achieved by the achievable schemes of sending correlated codewords by the relays [6], the compressed dirty-paper coding allowing correlated quantization noise [4] and the reverse compute-and-forward scheme [5] are denoted by the solid, circled, and dashed lines, respectively. Furthermore, the sum rate of the time-sharing of all the existing achievable schemes, which is the largest known lower bound for the sum capacity, is denoted by the dot-dashed line. In the gap between the derived upper bound in Corollary 1, i.e., the diamond line, and the largest known lower bound for the sum capacity, i.e., the dot-dashed line, lies the sum capacity of the 2-user Gaussian multiple access diamond channel for this symmetric case. The gap is greatly reduced due to the outer bound derived in this paper, for example, at  $C = 1.4$ , the new gap using Corollary 1 of this paper is only 1/3 of the gap between existing inner and outer bounds. Since the gap between the existing inner bound and the *new* outer bound is not large, we may conclude that the existing achievable schemes perform reasonably well in this scenario.

In the case of  $a = b = 0.3$ ,  $P_1 = P_2 = 10$  and  $C_1 = C_2 = C$ , the results are shown in Fig. 6. Though the new cut-set bound in (10) still improves upon the simple existing cut-set bound, the improvement of the new upper bound in Corollary 1 over the new cut-set bound is insignificant. This is because the reduction from (10) to Corollary 1 depends on the set  $\mathcal{A}_x$ ,  $x = a, b$ , and for smaller  $a, b$  values,  $\mathcal{A}_x$  is

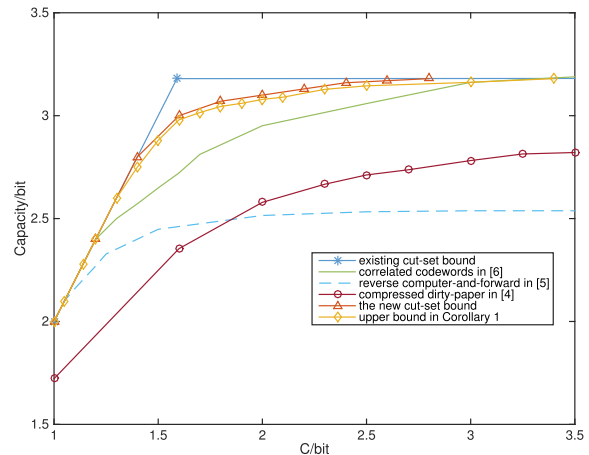


Fig. 6. Upper and lower bounds on the sum capacity for the case of  $a = b = 0.3$ ,  $P_1 = P_2 = 10$  and  $C_1 = C_2 = C$ .

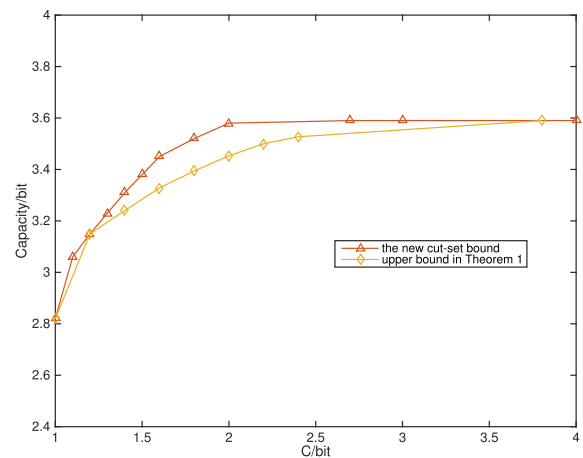


Fig. 7. Upper bounds on  $R_1 + 2R_2$  for the case of  $a = b = 0.9$ ,  $P_1 = P_2 = 10$  and  $C_1 = C_2 = C$ .

a smaller region around 0, where the improvement of  $f_C(\rho)$  over  $f_C(0)$  is small. However, note that multicell processing is more likely used for base stations that have strong links to all the users. Hence, the values of larger  $a, b$  is of more practical interest.

Finally, we plot the new cut-set upper bound and the result of Theorem 1 for the rate  $R_1 + 2R_2$ , for the case of  $a = b = 0.9$ ,  $P_1 = P_2 = 10$  and  $C_1 = C_2 = C$  in Fig. 7. The existing cut-set bound and the achievable schemes characterizes the sum rate only, and thus, are not plotted. Hence, we conclude that Theorem 1 offers a first converse result for arbitrary linear combinations of  $R_1$  and  $R_2$ . The triangle line depicts the upper bound on  $R_1 + 2R_2$  based on new cut-set bound, i.e., considering  $T_A^m(\rho)$ ,  $m = 12, 21$  only, and the diamond line depicts the upper bound in Theorem 1, i.e., considering  $T_B^m(\rho)$  as well as  $T_A^m(\rho)$ ,  $m = 12, 21$ . The improvement of the diamond line over the triangle line illustrates the need to characterize the rate reduction of the first hop due to correlated codewords sent by the relays.

## VI. CONCLUSION

In this paper, we derive a novel outer bound for the capacity of the 2-user Gaussian multiple access



diamond channel. Through numerical results, we show that the derived outer bound greatly reduces the gap between known inner and outer bounds when the capacities of the backhaul links are in the medium range. Future research directions include deriving converse results for the *multi*-user Gaussian multiple access diamond channel, the case where there are receiver side information at one or more of the receivers, and the uplink multicell processing system by possibly exploiting uplink-downlink duality.

APPENDIX A  
PROOF OF LEMMA 1

Based on the four cuts demonstrated in Fig. 4, we have the following cut-set upper bound on the sum capacity, i.e.,  $R_1 + R_2$ :

1) Considering Cut C, we have

$$R_1 + R_2 \leq C_1 + C_2 = f_C(0). \quad (40)$$

2) Considering Cut B, we have two cases:

a) For the case of  $|b| \leq 1$ ,

$$\begin{aligned} n(R_1 + R_2) &= H(W_1, W_2) \\ &= H(W_1, W_2, X_2^n) \end{aligned} \quad (41)$$

$$= H(X_2^n) + H(W_1, W_2 | X_2^n) \quad (42)$$

$$\begin{aligned} &= H(X_2^n) + H(W_1 | X_2^n, W_2) + H(W_2 | X_2^n) \\ &\leq nC_2 + H(W_1 | X_2^n, W_2) + H(W_2 | X_2^n) \\ &= nC_2 + I(W_1; Y_1^n | X_2^n, W_2) \\ &\quad + H(W_1 | Y_1^n, X_2^n, W_2) + I(W_2; Y_2^n | X_2^n) + H(W_2 | Y_2^n, X_2^n) \\ &\leq nC_2 + I(W_1; Y_1^n | X_2^n, W_2) + I(W_2; Y_2^n | X_2^n) + 2n\epsilon_n \end{aligned} \quad (43)$$

$$= nC_2 + I(X_1^n; Y_1^n | X_2^n, W_2) + I(W_2; Y_2^n | X_2^n) + 2n\epsilon_n, \quad (44)$$

where (41) follows from the fact that without loss of generality, we consider deterministic encoding at the source node, i.e.,  $(X_1^n, X_2^n)$  is a deterministic function of  $(W_1, W_2)$ , (43) follows from Fano's inequality, and (44) follows from the fact that we consider deterministic encoders and the Markov Chain  $W_1 \rightarrow (X_1^n, X_2^n, W_2) \rightarrow Y_1^n$ . Define  $\tilde{Y}_2^n$  as the following channel

$$\tilde{Y}_2^n = Y_1^n + \left(\frac{1}{b} - a\right) X_2^n + \tilde{U}^n,$$

where  $\tilde{U}^n$  is an i.i.d. sequence of Gaussian random variables with zero mean and variance  $\frac{1}{b^2} - 1$ , and it is independent of everything else. Note that given  $X_2^n$ ,  $\tilde{Y}_2^n$  is a physically degraded version of  $Y_1^n$ . Furthermore, note the similarity between

$$\tilde{Y}_2^n = X_1^n + \frac{1}{b} X_2^n + U_1^n + \tilde{U}^n$$

and

$$Y_2^n = bX_1^n + X_2^n + U_2^n,$$

which means that we have

$$I(W_2; Y_2^n | X_2^n) = I(W_2; \tilde{Y}_2^n | X_2^n).$$

Thus, from (44), we continue to write as follows while for the simplicity of presentation, we have dropped the  $2n\epsilon_n$  term,

$$\begin{aligned} n(R_1 + R_2) &\leq nC_2 + I(X_1^n; Y_1^n | X_2^n, W_2) + I(W_2; \tilde{Y}_2^n | X_2^n) \\ &= nC_2 + \sum_{i=1}^n [I(X_{1i}; Y_{1i} | X_{2i}^n, W_2, Y_1^{i-1}) \\ &\quad + I(W_2; \tilde{Y}_{2i} | X_{2i}^n, \tilde{Y}_2^{i-1})] \\ &\leq nC_2 + \sum_{i=1}^n [I(X_{1i}; Y_{1i} | X_{2i}^n, W_2, \tilde{Y}_2^{i-1}) \\ &\quad + I(W_2; \tilde{Y}_{2i} | X_{2i}^n, \tilde{Y}_2^{i-1})] \end{aligned} \quad (45)$$

$$\begin{aligned} &\leq nC_2 + \sum_{i=1}^n [I(X_{1i}; Y_{1i} | X_{2i}^n, W_2, \tilde{Y}_2^{i-1}) \\ &\quad + I(W_2, X_{2\{i\}^c}, \tilde{Y}_2^{i-1}; \tilde{Y}_{2i} | X_{2i}),] \end{aligned} \quad (46)$$

where (45) follows from the fact that given  $X_2^n$ ,  $\tilde{Y}_2^n$  is a physically degraded version of  $Y_1^n$ . Define auxiliary random variables

$$V_i = (W_2, X_{2\{i\}^c}, \tilde{Y}_2^{i-1}).$$

Following from (46), we have

$$\begin{aligned} n(R_1 + R_2) - nC_2 &\leq \sum_{i=1}^n [I(X_{1i}; Y_{1i} | X_{2i}, V_i) + I(V_i; \tilde{Y}_{2i} | X_{2i})] \\ &= n(I(X_{1Q}; Y_{1Q} | X_{2Q}, V_Q, Q) + I(V_Q; \tilde{Y}_{2Q} | X_{2Q}, Q)) \\ &\leq n(I(X_{1Q}; Y_{1Q} | X_{2Q}, V_Q, Q) + I(V_Q, Q; \tilde{Y}_{2Q} | X_{2Q})) \\ &= n(I(X_1; Y_1 | X_2, V) + I(V; \tilde{Y}_2 | X_2)) \quad (47) \\ &\leq nE_{X_2} [I(X_1; Y_1 | X_2 = x_2, V) + I(V; \tilde{Y}_2 | X_2 = x_2)], \end{aligned} \quad (48)$$

where (47) follows from the definition in (14) and

$$V \triangleq (V_Q, Q), \quad \tilde{Y}_2 \triangleq \tilde{Y}_{2Q}. \quad (49)$$

Note that the sum capacity of the degraded broadcast channel where the input of the channel is  $X_1$  given  $X_2 = x_2$  and the outputs of the channel is  $Y_1$  and  $\tilde{Y}_2$ , respectively, is given by [11]

$$\begin{aligned} &\max_{p(v, x_1)} \left[ I(X_1; Y_1 | X_2 = x_2, V) + I(V; \tilde{Y}_2 | X_2 = x_2) \right] \\ &= \frac{1}{2} \log \left( 1 + E[(X_1 - E[X_1 | X_2 = x_2])^2 | X_2 = x_2] \right). \end{aligned}$$

Hence, for the particular  $p(v, x_1)$  as defined by the codebook and (49), we have

$$\begin{aligned} I(X_1; Y_1 | X_2 = x_2, V) + I(V; \tilde{Y}_2 | X_2 = x_2) \\ \leq \frac{1}{2} \log \left( 1 + E[(X_1 - E[X_1 | X_2 = x_2])^2 | X_2 = x_2] \right). \end{aligned} \quad (50)$$

Hence, following from (48) and (50), we have

$$\begin{aligned} & R_1 + R_2 - C_2 \\ & \leq E_{X_2} \left[ \frac{1}{2} \log \left( 1 + E[(X_1 - E[X_1|X_2 = x_2])^2 | X_2 = x_2] \right) \right] \\ & \leq \frac{1}{2} \log \left( 1 + E_{X_2} \left[ E[(X_1 - E[X_1|X_2 = x_2])^2 | X_2 = x_2] \right] \right) \end{aligned} \quad (51)$$

$$\leq \frac{1}{2} \log \left( 1 + E[X_1^2] - \frac{E^2[X_1 X_2]}{E[X_2^2]} \right) \quad (52)$$

$$\begin{aligned} & = \frac{1}{2} \log \left( 1 + (1 - \rho^2) \bar{P}_1 \right) \\ & \leq \frac{1}{2} \log \left( 1 + (1 - \rho^{*2}) P_1 \right), \end{aligned} \quad (53)$$

where (51) follows from the convexity of the  $\log(\cdot)$  function, (52) follows from the fact that the mean-squared error (MSE) of the optimal Bayes least square (BLS) estimator is smaller than that of the linear least squared (LLS) estimator, and (53) follows from (15) and (17).

b) Similarly, for the case of  $|b| > 1$ , following from (42), we have

$$\begin{aligned} & n(R_1 + R_2) \\ & = H(X_2^n) + H(W_1, W_2 | X_2^n) \\ & = H(X_2^n) + H(W_2 | X_2^n, W_1) + H(W_1 | X_2^n) \\ & \leq nC_2 + I(X_1^n; Y_2^n | X_2^n, W_1) + I(W_1; Y_1^n | X_2^n) + 2n\epsilon_n. \end{aligned}$$

By following similar steps as (44) to (53), we may conclude that

$$R_1 + R_2 \leq C_2 + \frac{1}{2} \log \left( 1 + b^2(1 - \rho^{*2}) P_1 \right). \quad (54)$$

Combining (53) and (54), we have

$$\begin{aligned} R_1 + R_2 & \leq C_2 + \frac{1}{2} \log \left( 1 + \max\{b^2, 1\}(1 - \rho^{*2}) P_1 \right) \\ & = f_B(\rho^*). \end{aligned} \quad (55)$$

3) Due to symmetry, when we consider Cut A, we obtain

$$\begin{aligned} R_1 + R_2 & \leq C_1 + \frac{1}{2} \log \left( 1 + \max\{a^2, 1\}(1 - \rho^{*2}) P_2 \right) \\ & = f_A(\rho^*). \end{aligned} \quad (56)$$

Thus, combining (40), (55) and (56), we have proved Lemma 1.

#### APPENDIX B PROOF OF LEMMA 2

$$1) (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n) - \lambda I(X_1^n, X_2^n; Z^n)$$

First, we derive the following single-letterization.

$$\begin{aligned} & I(X_1^n, X_2^n; Y_2^n) - I(X_1^n, X_2^n; Z^n) \\ & \leq I(X_1^n, X_2^n; Y_2^n, Z^n) - I(X_1^n, X_2^n; Z^n) \\ & = I(X_1^n, X_2^n; Y_2^n | Z^n) \\ & = \sum_{i=1}^n I(X_1^n, X_2^n; Y_{2i} | Y_2^{i-1}, Z^n) \\ & \leq \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{2i} | Z_i) \end{aligned} \quad (57)$$

$$\begin{aligned} & = nI(X_1, X_2; Y_2 | Z, Q) \\ & \leq nI(X_1, X_2; Y_2 | Z), \end{aligned} \quad (58)$$

where (57) follows from the Markov chain  $(X_1^n, X_2^n, Z^n, Y_2^{i-1}) \rightarrow (X_{1i}, X_{2i}, Z_i) \rightarrow Y_{2i}$ , and (58) comes from conditioning reduces entropy and the Markov Chain  $Q \rightarrow (Z, X_1, X_2) \rightarrow Y_2$ . Then, using the above results, we have

$$\begin{aligned} & (\mu - \lambda)I(X_1^n, X_2^n; Y_2^n) - \lambda I(X_1^n, X_2^n; Z^n) \\ & = (\mu - 2\lambda)I(X_1^n, X_2^n; Y_2^n) + \lambda I(X_1^n, X_2^n; Y_2^n) \\ & \quad - \lambda I(X_1^n, X_2^n; Z^n). \\ & \leq n(\mu - 2\lambda)I(X_1, X_2; Y_2) + n\lambda I(X_1, X_2; Y_2 | Z), \end{aligned} \quad (59)$$

where (59) follows from (58).

$$2) \lambda I(X_1^n; Z^n | X_2^n) + \lambda I(X_2^n; Z^n | X_1^n)$$

Based on (30), we have

$$\begin{aligned} & \lambda(I(X_1^n; Z^n | X_2^n) + I(X_2^n; Z^n | X_1^n)) \\ & \leq \lambda \left( \sum_{i=1}^n I(X_{1i}; Z_i | X_{2i}) + I(X_{2i}; Z_i | X_{1i}) \right) \\ & = n\lambda(I(X_1; Z | X_2, Q) + I(X_2; Z | X_1, Q)) \\ & \leq n\lambda(I(X_1; Z | X_2) + I(X_2; Z | X_1)), \end{aligned} \quad (60)$$

where (60) follows from the Markov chain  $(X_1^n, X_2^n, Z^{i-1}) \rightarrow (X_{1i}, X_{2i}) \rightarrow Z_i$ .

$$3) I(X_1^n, X_2^n; Y_1^n | W_2) - \frac{\mu - \lambda}{1 - \lambda} I(X_1^n, X_2^n; Y_2^n | W_2)$$

First, we derive the following single-letterization to be used later. From [21, pp. 314, eq. (3.34)], we have

$$\begin{aligned} H(Y_1^n | W_2) - H(Y_2^n | W_2) & = \sum_{i=1}^n \left[ H(Y_{1i} | W_2, Y_1^{i-1}, Y_{2(i+1)}^n) \right. \\ & \quad \left. - H(Y_{2i} | W_2, Y_1^{i-1}, Y_{2(i+1)}^n) \right]. \end{aligned}$$

Let us define  $T_i \triangleq (W_2, Y_1^{i-1}, Y_{2(i+1)}^n)$ , and further define

$$T \triangleq T_Q. \quad (62)$$

Note that the auxiliary random variables thus defined satisfy

$$T \rightarrow (X_1, X_2) \rightarrow (Y_1, Y_2, Z). \quad (63)$$

Based on the definition of the random variables in (14) and (62), we have

$$\begin{aligned} H(Y_1^n|W_2) - H(Y_2^n|W_2) &= n(H(Y_1|T, Q) - H(Y_2|T, Q)), \\ nH(Y_1|T, Q) &= \sum_{i=1}^n H(Y_{1i}|W_2, Y_1^{i-1}, Y_{2(i+1)}^n) \\ &\leq \sum_{i=1}^n H(Y_{1i}|W_2, Y_1^{i-1}) = H(Y_1^n|W_2). \end{aligned}$$

Thus, there exist a  $\gamma \geq 0$  such that

$$\begin{aligned} H(Y_1^n|W_2) &= n(H(Y_1|T, Q) + \gamma), \\ H(Y_2^n|W_2) &= n(H(Y_2|T, Q) + \gamma). \end{aligned} \quad (64)$$

Then, using the above results, we have

$$\begin{aligned} I(X_1^n, X_2^n; Y_1^n|W_2) - \frac{\mu - \lambda}{1 - \lambda} I(X_1^n, X_2^n; Y_2^n|W_2) &= H(Y_1^n|W_2) - H(Y_1^n|W_2, X_1^n, X_2^n) - \frac{\mu - \lambda}{1 - \lambda} H(Y_2^n|W_2) \\ &\quad + \frac{\mu - \lambda}{1 - \lambda} H(Y_2^n|X_1^n, X_2^n, W_2) \\ &= n \left[ H(Y_1|T, Q) + \gamma - H(Y_1|X_1, X_2, Q) - \frac{\mu - \lambda}{1 - \lambda} \gamma \right. \\ &\quad \left. - \frac{\mu - \lambda}{1 - \lambda} H(Y_2|T, Q) + \frac{\mu - \lambda}{1 - \lambda} H(Y_2|X_1, X_2, Q) \right] \quad (65) \end{aligned}$$

$$\begin{aligned} &= n \left[ H(Y_1|T, Q) - H(Y_1|X_1, X_2, T, Q) + \left(1 - \frac{\mu - \lambda}{1 - \lambda}\right) \gamma \right. \\ &\quad \left. - \frac{\mu - \lambda}{1 - \lambda} H(Y_2|T, Q) + \frac{\mu - \lambda}{1 - \lambda} H(Y_2|X_1, X_2, T, Q) \right] \quad (66) \end{aligned}$$

$$\begin{aligned} &= n \left[ I(X_1, X_2; Y_1|T, Q) - \frac{\mu - \lambda}{1 - \lambda} I(X_1, X_2; Y_2|T, Q) \right. \\ &\quad \left. + \left(1 - \frac{\mu - \lambda}{1 - \lambda}\right) \gamma \right] \\ &\leq n \left[ I(X_1, X_2; Y_1|T, Q) - \frac{\mu - \lambda}{1 - \lambda} I(X_1, X_2; Y_2|T, Q) \right] \quad (67) \end{aligned}$$

$$= n \left[ I(X_1, X_2; Y_1|U) - \frac{\mu - \lambda}{1 - \lambda} I(X_1, X_2; Y_2|U) \right], \quad (68)$$

where (65) follows from (64), (66) follows from (63), (67) follows because  $\mu \geq 1$ ,  $\lambda < 1$  and  $\gamma \geq 0$ , and (68) follows from the definition of the auxiliary random variable  $U$  in (31).

#### APPENDIX C PROOF OF LEMMA 3

1)  $(\mu - 2\lambda)I(X_1, X_2; Y_2) + \lambda I(X_1, X_2; Y_2|Z)$

First, we provide the upper bound on  $I(X_1, X_2; Y_2|Z)$ . We have

$$\begin{aligned} I(X_1, X_2; Y_2|Z) &= I(X_1, X_2; Y_2) - I(X_1, X_2; Z) \\ &= (H(Y_2) - H(Z)) - \frac{1}{2} \log(2\pi e) + \frac{1}{2} \log(2\pi e)(1 + N_3) \\ &\leq \frac{1}{2} \log \left( b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 \right) \\ &\quad - \frac{1}{2} \log \left( \frac{b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 + N_3}{1 + N_3} \right), \quad (70) \end{aligned}$$

where (69) follows from the distribution of (33), and (70) follows from [19, Corollary 4] and the fact that

$$\begin{aligned} E[(bX_1 + X_2)^2] &= b^2 E[X_1^2] + E[X_2^2] + 2b\rho \sqrt{E[X_1^2]E[X_2^2]} \\ &\leq b^2 P_1 + P_2 + 2b\rho \sqrt{\bar{P}_1 \bar{P}_2} \quad (71) \\ &= b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2}, \quad (72) \end{aligned}$$

where in (71), we have used (15), and (72) follows from the definition of  $\rho^*$  in (16).

Then, using the above result, we have

$$\begin{aligned} (\mu - 2\lambda)I(X_1, X_2; Y_2) + \lambda I(X_1, X_2; Y_2|Z) &\leq \frac{\mu - \lambda}{2} \log \left( b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 \right) \\ &\quad - \frac{\lambda}{2} \log \left( \frac{b^2 P_1 + P_2 + 2b\rho^* \sqrt{P_1 P_2} + 1 + N_3}{1 + N_3} \right), \quad (73) \end{aligned}$$

where (73) follows from (72), the fact that given variance, the Gaussian distribution maximizes the differential entropy, (70), and the fact that  $0 \leq \lambda \leq \frac{\mu}{2}$ .

2)  $I(X_1; Z|X_2) + I(X_2; Z|X_1)$

Let us first calculate

$$\begin{aligned} h(Z|X_2 = x_2) &\leq \frac{1}{2} \log(2\pi e) E[(Z - E[Z|X_2 = x_2])^2 | X_2 = x_2] \quad (74) \end{aligned}$$

$$\leq \frac{1}{2} \log(2\pi e) \left( E[Z^2] - \frac{E^2[Z X_2]}{E[X_2^2]} \right) \quad (75)$$

$$= \frac{1}{2} \log(2\pi e) (b^2(1 - \rho^2)E[X_1^2] + 1 + N_3)$$

$$\leq \frac{1}{2} \log(2\pi e) (b^2(1 - \rho^2)P_1 + 1 + N_3) \quad (76)$$

$$\leq \frac{1}{2} \log(2\pi e) (b^2(1 - \rho^{*2})P_1 + 1 + N_3), \quad (77)$$

where (74) follows from the fact that given the covariance, the Gaussian distribution maximizes the differential entropy, (75) follows from the fact that the mean-squared error (MSE) of the optimal Bayes least square (BLS) estimator is smaller than that of the linear least squared (LLS) estimator, (76) follows from (15), and (77) follows from (17). Then, we have

$$\begin{aligned} h(Z|X_2) &= \int_{\mathbb{R}} h(Z|X_2 = x_2) f(x_2) dx_2 \\ &\leq \frac{1}{2} \log(2\pi e) (b^2(1 - \rho^{*2})P_1 + 1 + N_3), \quad (78) \end{aligned}$$

where (78) follows from (77), and finally, we have

$$I(X_1; Z|X_2) = h(Z|X_2) - h(Z|X_1, X_2) \leq \frac{1}{2} \log \frac{(1 - \rho^{*2})b^2 P_1 + 1 + N_3}{1 + N_3}. \quad (79)$$

Similarly, we have

$$I(X_2; Z|X_1) \leq \frac{1}{2} \log \frac{(1 - \rho^{*2})P_2 + 1 + N_3}{1 + N_3}. \quad (80)$$

Thus, from (79) and (80), we have

$$I(X_1; Z|X_2) + I(X_2; Z|X_1) \leq \frac{1}{2} \log \frac{(1 - \rho^{*2})b^2 P_1 + 1 + N_3}{1 + N_3} + \frac{1}{2} \log \frac{(1 - \rho^{*2})P_2 + 1 + N_3}{1 + N_3}. \quad (81)$$

REFERENCES

[1] O. Simeone *et al.*, “Cooperative wireless cellular systems: An information-theoretic view,” *Found. Trends Commun. Inf. Theory*, vol. 8, nos. 1–2, pp. 1–177, 2012, doi: 10.1561/01000000048.

[2] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), “Sum rate characterization of joint multiple cell-site processing,” *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4473–4497, Dec. 2007.

[3] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), “Downlink multicell processing with limited-backhaul capacity,” *EURASIP J. Adv. Signal Process.*, vol. 2009, Jun. 2009, Art. no. 3.

[4] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), “Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks,” *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5646–5658, Nov. 2013.

[5] S.-N. Hong and G. Caire, “Compute-and-forward strategies for cooperative distributed antenna systems,” *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5227–5243, Sep. 2013.

[6] N. Liu and W. Kang, “A new achievability scheme for downlink multicell processing with finite backhaul capacity,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2014, pp. 1006–1010.

[7] X. Yi and N. Liu, “An achievability scheme for downlink multicell processing with finite backhaul capacity: The general case,” in *Proc. Int. Conf. Wireless Commun. Signal Process.*, Nanjing, China, Oct. 2015, pp. 1–5.

[8] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inf. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.

[9] R. Ahlswede and T. S. Han, “On source coding with side information via a multiple-access channel and related problems in multi-user information theory,” *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 396–412, May 1983.

[10] D. Traskov and G. Kramer, “Reliable communication in networks with multi-access interference,” in *Proc. IEEE Inf. Theory Workshop (ITW)*, Lake Tahoe, CA, USA, Sep. 2007, pp. 343–348.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1991.

[12] S. Ganguly and Y.-H. Kim, “On the capacity of cloud radio access networks,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 2063–2067.

[13] P. Patil and W. Yu. (2018). “Generalized compression strategy for the downlink cloud radio access network.” [Online]. Available: <https://arxiv.org/abs/1801.00394>

[14] S. H. Lim, K. T. Kim, and Y.-H. Kim, “Distributed decode-forward for relay networks,” *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4103–4118, Jul. 2017.

[15] W. Kang and N. Liu, “The Gaussian multiple access diamond channel,” *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6049–6059, Nov. 2015.

[16] S. S. Bidokhti and G. Kramer, “Capacity bounds for diamond networks with an orthogonal broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7103–7122, Dec. 2016.

[17] Y. Geng and C. Nair, “The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages,” *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2087–2104, Apr. 2014.

[18] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), “The capacity region of the Gaussian multiple-input multiple-output broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[19] T. Liu and P. Viswanath, “An extremal inequality motivated by multiterminal information-theoretic problems,” *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[20] L. Ozarow, “On a source-coding problem with two channels and three receivers,” *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.

[21] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York, NY, USA: Academic, 1981.

[22] S.-H. Park, O. Simeone, and S. Shamai (Shitz), “Joint optimization of cloud and edge processing for fog radio access networks,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7621–7632, Nov. 2016.

[23] S. Lee, W. Zhao, and A. Khisti, “Secure degrees of freedom of the Gaussian diamond-wiretap channel,” *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 496–508, Jan. 2017.

**Tianyu Yang** received the B.Eng. degree in information engineering from Southeast University, Nanjing, P. R. China in 2014 and the M.S. degree in information security from Southeast University, Nanjing, P. R. China in 2017. He is now with Tencent Technology, Shanghai. His research interests are in network information theory for wireless networks, wireless communication and computer science.

**Nan Liu** received the B.Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, P. R. China in 2001, and the Ph.D. degree in electrical and computer engineering from University of Maryland, College Park, MD in 2007.

From 2007–2008, she was a postdoctoral scholar in the Wireless Systems Lab, Department of Electrical Engineering, Stanford University. In 2009, she became a professor in the National Mobile Communications Research Laboratory, School of Information Science and Engineering in Southeast University, Nanjing, China. Her research interests are in network information theory for wireless networks, SON algorithms for next generation cellular networks and energy-efficient communications. She is an associate editor for *China Communications*.

**Wei Kang** received the B.Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2001, the M.Eng. degree in electrical engineering from McGill University, Montreal, Canada, in 2003, and the Ph.D. degree in electrical engineering from University of Maryland, College Park, in 2008. He joined the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2009, where he is currently an Associate Professor. His research mainly focuses on network information theory.

**Shlomo Shamai (Shitz)** received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975–1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now a Technion Distinguished Professor, and holds the William Fondiller Chair of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow, an URSI Fellow, a member of the Israeli Academy of Sciences and Humanities and a foreign member of the US National Academy of Engineering. He is the recipient of the 2011 Claude E. Shannon Award, the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering and the 2017 IEEE Richard W. Hamming Medal.

He has been awarded the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and is a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 and 2015 European Commission FP7, Network of Excellence in Wireless Communications (NEWCOM++, NEWCOM#) Best Paper Awards, the 2010 Thomson Reuters Award for International Excellence in Scientific Research, the 2014 EURASIP Best Paper Award (for the EURASIP Journal on Wireless Communications and Networking), and the 2015 IEEE Communications Society Best Tutorial Paper Award. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served twice on the Board of Governors of the Information Theory Society. He has also served on the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY and on the IEEE Information Theory Society Nominations and Appointments Committee.