Interference Mitigation Through Limited Receiver Cooperation

I-Hsiang Wang, *Student Member, IEEE*, and David N. C. Tse, *Fellow, IEEE*

Abstract*—Interference is a major issue limiting the performance in wireless networks. Cooperation among receivers can help mitigate interference by forming distributed MIMO systems. The rate at which receivers cooperate, however, is limited in most scenarios. How much interference can one bit of receiver cooperation mitigate? In this paper, we study the two-user Gaussian interference channel with conferencing decoders to answer this question in a simple setting. We identify two regions regarding the gain from receiver cooperation: linear and saturation regions. In the linear region, receiver cooperation is efficient and provides a** *degrees-of-freedom* **gain, which is either** *one cooperation bit buys one over-the-air bit* **or** *two cooperation bits buy one over-the-air bit. In the saturation region, receiver cooperation is inefficient and provides a** *power* **gain, which is bounded regardless of the rate at which receivers cooperate. The conclusion is drawn from the characterization of capacity region to within two bits/s/Hz, regardless of channel parameters. The proposed strategy consists of two parts: 1) the transmission scheme, where superposition encoding with a simple power split is employed and 2) the cooperative protocol, where one receiver quantize-bin-and-forwards its received signal and the other after receiving the side information decode-bin-and-forwards its received signal.**

***Index Terms*—Capacity to within a bounded gap, distributed MIMO system, interference management, receiver cooperation.**

I. Introduction

In modern communication systems, interference is one of the fundamental factors that limit performance. The simplest information theoretic model for studying this issue is the *two-user interference channel*. Characterizing its capacity region is a long-standing open problem, except for several special cases (e.g., the strong interference regime [1]). The largest achievable region to date is reported by Han and Kobayashi [2] and the core of the scheme is a superposition coding strategy. Recent progress has been made on both inner bounds and outer bounds: Etkin, Tse, and Wang characterized the capacity region of the two-user Gaussian interference channel to within one bit [3] by using a superposition coding scheme with a simple power-split configuration and by providing new upper bounds. The bounded gap-to-optimality result [3] leads to an uniform approximation of the capacity region and provides a strong guarantee on the performance of the proposed scheme. Later, Motahari and Khandani [4], Shang, Kramer, and Chen [5], and Annapureddy and Veeravalli [6] independently improve the outer bounds and characterize the sum capacity in a very weak interference regime and a mixed interference regime.

In the above interference channel setup, transmitters or receivers are not allowed to communicate with one another and each user has to combat interference on its own. In various scenarios, however, nodes are not isolated and transmitters/receivers can exchange certain amount of information. Cooperation among transmitters/receivers can help mitigate interference by forming distributed MIMO systems which provide two kinds of gains: *degrees-of-freedom* gain and *power* gain. The rate at which they cooperate, however, is limited, due to physical constraints. Therefore, one of the fundamental questions is, how much *interference* can limited *transmitter/receiver cooperation* mitigate? How much gain can it provide?

In this paper, we consider a two-user Gaussian interference channel with *conferencing decoders* to answer this question regarding receiver cooperation. Transmitter cooperation is addressed in [32]. Conferencing among encoders/decoders has been studied in [7]–[12]. Our model is similar to those in [11] and [12] but in an interference channel setup. The work in [11] characterizes the capacity region of the compound multiple access channel (MAC) with unidirectional conferencing between decoders. For general setup (i.e., bidirectional conferencing), it provides achievable rates and finds the maximum achievable individual rate to within a bounded gap, but is not able to establish an uniform approximation result on the capacity region. The work in [12] considers one-sided Gaussian interference channels with unidirectional conferencing between decoders and characterizes the capacity region in strong interference regimes and the asymptotic sum capacity at high SNR. For general receiver cooperation, works including [13] and [14], investigate cooperation in interference channels with a setup where the cooperative links are in the same band as the links in the interference channel. In particular, [14] characterizes the sum capacity of Gaussian interference channels with symmetric in-band receiver cooperation to within a bounded gap. Our work, on the other hand, is focused on the Gaussian interference channel with out-of-band (orthogonal) receiver cooperation and studies its entire capacity region. Works on interference channels with additional relays [15]–[17] and two-hop interference-relay networks [18] are also related to our problem, since the receivers also serve as relays in our setup.

We propose a strategy achieving the capacity region universally to within 2 bits/s/Hz per user, regardless of channel parameters. The two-bit gap is the worst case gap which can be loose in some regimes, and it is vanishingly small at high

Manuscript received November 11, 2009; revised September 07, 2010; accepted January 13, 2011. Date of current version April 20, 2011. This work was supported by the National Science Foundation under Grant CCF-0830796. The material in this paper was presented in part at the IEEE Information Theory Workshop, Taormina, Italy, October 2009.

The authors are with the Wireless Foundations, Department of EECS, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: ihsiang@eecs.berkeley.edu; dtse@eecs.berkeley.edu).

Communicated by E. Erkip, 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.2011.2120290

0018-9448/$26.00 © 2011 IEEE

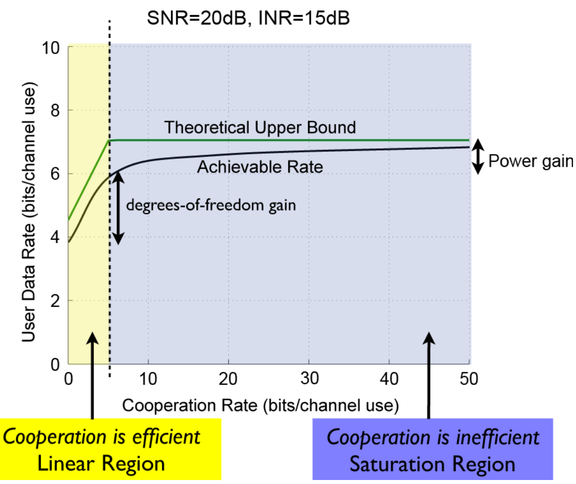


Fig. 1. The gain from limited receiver cooperation.

SNR when compared to the capacity. The strategy consists of two parts: 1) the transmission scheme, describing how transmitters encode their messages and 2) the cooperative protocol, describing how receivers exchange information and decode messages. For transmission, both transmitters use superposition coding [2] with the same common-private power split as in the case without cooperation [3]. For the cooperative protocol, it is appealing to apply the decode-forward or compress-forward schemes, originally proposed in [19] for the relay channel, like most works dealing with more complicated networks, including [10]–[13], [20], etc. It turns out neither conventional compress-forward nor decode-forward achieves capacity to within a bounded gap for the problem at hand. On the other hand, [21]–[25] observe that the conventional compress-forward scheme [19] may be improved by the destination directly decoding the sender's message instead of requiring to first decode the quantized signal of the relay. We use such an improved compress-forward scheme as part of our cooperative protocol. One of the receivers quantizes its received signal at an appropriate distortion, bins the quantization codeword and sends the bin index to the other receiver. The other receiver then decodes its own information based on its own received signal and the received bin index. After decoding, it bin-and-forwards the decoded common messages back to the former receiver and helps it decode. Note that although an arbitrary number of rounds is allowed in the conferencing formulation, it turns out that two rounds are sufficient to achieve within 2 bits of the capacity.

We identify two regions regarding the gain from receiver cooperation: linear and saturation regions, as illustrated through a numerical example in Fig. 1. In the plot we fix the signal-to-noise ratios (SNR) and interference-to-noise ratios (INR) to be 20 dB and 15 dB respectively and we plot the user data rate versus the cooperation rate. In the linear region, receiver cooperation is *efficient*, in the sense that the growth of each user's "over-the-air" data rate is roughly linear with respect to the capacity of receiver-cooperative links. The gain in this region is the *degrees-of-freedom* gain that distributed MIMO systems provide. On the other hand, in the saturation region, receiver cooperation is *inefficient* in the sense that the growth of each user's over-the-air data rate becomes saturated as one increases the rate in receiver-cooperative links. The gain is the *power* gain which is bounded regardless of the cooperation rate. We will focus on the system performance in the linear region, because not only that in most scenarios the rate at which receivers can cooperate is limited, but also that the gain from cooperation is more significant.

With the bounded gap-to-optimality result, we find that the fundamental gain from cooperation in the linear region as follows: either *one cooperation bit buys one over-the-air bit* or *two cooperation bits buy one over-the-air bit* until saturation, depending on channel parameters. In the symmetric setup, at high SNR, when INR is below 50% of SNR in dB scale, one-bit cooperation per direction buys roughly one-bit gain per user until full receiver cooperation performance is reached, while when INR is between 67% and 200% of SNR in dB scale, one-bit cooperation per direction buys roughly half-bit gain per user. (The example in Fig. 1 falls in the latter case, and as can be seen, the slope of the linear region is about 0.5.) In the weak interference regime, for a given pair of (SNR, INR), when the receiver-cooperative link capacity $C^B > \log \mathsf{INR}$, cooperation between receivers can get a close-to-interference-free (that is, within a bounded gap) performance. In the strong interference regime, in contrast to that without cooperation, system performance can be boost *beyond* interference-free performance, by utilizing receiver-cooperative links not only for interference mitigation but also for forwarding desired information, since the cross link is stronger than the direct link.

The rest of this paper is organized as follows. In Section II, we introduce the channel model and formulate the problem. In Section III, we provide intuitive discussions about achievability and motivate our two-round strategy. Then we give examples to illustrate why it is not a good idea to use cooperative protocols based on conventional compress-forward or decode-forward. In Section IV, we describe the strategy concretely and derive its achievable rates and in Section V we show that the achievable rate region is within 2 bits per user to the outer bounds we provide. In addition, we characterize the capacity region of the compound MAC with conferencing decoders to within 1 bit, as a by-product. In Section VII, focusing on the symmetric setup, we illustrate the fundamental gain from receiver cooperation by deriving the optimal number of *generalized degrees of freedom* (g.d.o.f.) and compare it with the achievable ones of suboptimal schemes.

II. Problem Formulation

A. Channel Model

The two-user Gaussian interference channel with conferencing decoders is depicted in Fig. 2.

Transmitter-Receiver Links: The transmitter-receiver links are modeled as the *normalized* Gaussian interference channel

$$y_1 = h_{11}x_1 + h_{12}x_2 + z_1$$
$$y_2 = h_{21}x_1 + h_{22}x_2 + z_2,$$

where the additive noise processes $\{z_i[n]\}$, $(i = 1, 2)$, are independent $\mathcal{CN}(0,1)$, i.i.d. over time. In this paper, we use [.] to denote time indices. Transmitter i intends to convey message m_i

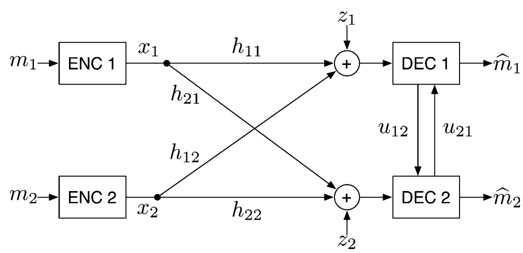


Fig. 2. Channel model.

to receiver i by encoding it into a block codeword $\{x_i[n]\}_{n=1}^{N}$, with transmit power constraints

$$\frac{1}{N}\sum_{n=1}^{N}\left|x_i[n]\right|^2 \le 1,\ i=1,2,$$

for arbitrary block length N. Note that the outcome of each encoder depends solely on its own message. Messages m_1, m_2 are independent. Define channel parameters

$$\mathsf{SNR}_i := |h_{ii}|^2,\ \mathsf{INR}_i := |h_{ij}|^2,\ i,j=1,2,\ i\neq j.$$

Receiver-Cooperative Links: For $(i,j)=(1,2),(2,1)$, the receiver-cooperative links are noiseless with capacity $\mathsf{C}_{ij}^{\mathsf{B}}$ from receiver i to j. Encoding must satisfy causality constraints: for any time index $n=1,2,\ldots,N$, the cooperation signal from receiver 2 to 1, $u_{21}[n]$, is only a function of $\{y_2[1],\ldots,y_2[n-1],u_{12}[1],\ldots,u_{12}[n-1]\}$ and the cooperation signal from receiver 1 to 2, $u_{12}[n]$, is only a function of $\{y_1[1],\ldots,y_1[n-1],u_{21}[1],\ldots,u_{21}[n-1]\}$.

In the rest of this paper, we use v^n to denote the sequence $\{v[1],\ldots,v[n]\}$.

B. Strategies, Rates, and Capacity Region

We give the basic definitions for the coding strategies, achievable rates of the strategy, and the capacity region of the channel.

Definition 2.1 (Strategy and Average Probability of Error): An (M_1,M_2,N)-strategy consists of the following: for $i,j=1,2,\ i\neq j$,

1) message set $\mathcal{M}_i := \{1,2,\ldots,M_i\}$ for user i;
2) encoding function $e_i^{(N)}:\mathcal{M}_i\to\mathbb{C}^N,\ m_i\mapsto x_i^N$ at transmitter i;
3) set of relay functions $\{r_i^{(n)}\}_{n=1}^{N}$ such that $u_{ij}[n]=r_i^{(n)}(y_i^{n-1},u_{ji}^{n-1})\in\{1,2,\ldots,2^{\mathsf{C}_{ij}^{\mathsf{B}}}\},\ \forall n=1,2,\ldots,N$ at receiver i;
4) decoding function $d_i^{(N)}:\mathbb{C}^N\times\{1,2,\ldots,2^{N\mathsf{C}_{ji}^{\mathsf{B}}}\}\to\mathcal{M}_i,\ (y_i^N,u_{ji}^N)\mapsto\hat{m}_i$ at receiver i.

The average probability of error

$$P_e^{(N)} := \frac{1}{M_1M_2}\sum_{\substack{m_1\in\mathcal{M}_1\\ m_2\in\mathcal{M}_2}}\Pr\left\{\begin{array}{ll} d_1^{(N)}(y_1^N,u_{21}^N)\neq m_1 \text{ or} & m_1,m_2\\ d_2^{(N)}(y_2^N,u_{12}^N)\neq m_2 & \text{are sent}\end{array}\right\}.$$

Definition 2.2 (Achievable Rates and Capacity Region): A rate tuple (R_1,R_2) is achievable if for any $\epsilon>0$ and for all sufficiently large N, there exists an (M_1,M_2,N) strategy with $M_i\ge 2^{NR_i}$, for $i=1,2$, such that $P_e^{(N)}<\epsilon$. The capacity region $\mathscr{C}$ is the collections of all achievable (R_1,R_2).

C. Notations

We summarize below the notations used in the rest of this paper.

- For a real number a, $(a)^+ := \max(a,0)$ denotes its positive part.
- For sets $A,B\subseteq\mathbb{R}^k$ in an k-dimensional space, $A\oplus B := \{a+b : a\in A, b\in B\}$ denotes the direct sum of A and B. $\mathrm{conv}\{A\}$ denotes the convex hull of the set A.
- With a little abuse of notations, for $x,y\in\mathbb{F}_q$, $x\oplus y$ denotes the modulo-q sum of x and y.
- Unless specified, all the logarithms $\log(.)$ are of base 2.

III. Motivation of Strategies

Before introducing our main result, we first provide intuitive discussions about achievability and motivate our two-round strategy (to be described in detail in Section IV) from a high-level perspective. Then we give examples to illustrate why cooperative protocols based on conventional compress-forward or decode-forward may not be good for cooperation between receivers to mitigate interference. Throughout the discussion in this section, we will make use of the *linear deterministic model* proposed in [25], [26].

The linear deterministic model is a tool for studying Gaussian networks so that an uniform approximation of the capacity can be found. It is also used for the two-user interference channel [27]. The model captures the signal interaction in the original Gaussian scenario to some extent and is useful for illustrating some subtle facts which are not easy to be uncovered in the Gaussian scenario. Throughout this paper, all discussions involving the linear deterministic model are either aimed to elucidate a certain phenomenon that arises in the Gaussian scenario, or to provide an intuitive argument for a certain claim without rigorously proving it.

A. Optimal Strategy in the Linear Deterministic Channel

First, consider the following symmetric channel: $\mathsf{SNR}_1=\mathsf{SNR}_2=\mathsf{SNR}$, $\mathsf{INR}_1=\mathsf{INR}_2=\mathsf{INR}$ and $\mathsf{C}_{12}^{\mathsf{B}}=\mathsf{C}_{21}^{\mathsf{B}}=\mathsf{C}^{\mathsf{B}}$. Set INR to be 2/3 of SNR in dB scale, that is, $\log\mathsf{INR}=\frac{2}{3}\log\mathsf{SNR}$. Set $\mathsf{C}^{\mathsf{B}}=\frac{1}{3}\log\mathsf{SNR}$. The corresponding linear deterministic channel (LDC) is depicted in Fig. 3. The bits at the levels of transmitters/receivers can be thought of as chunks of binary expansions of the transmitted/received signals. Note that in this example, one bit in the LDC corresponds to $\frac{1}{3}\log\mathsf{SNR}$ bits in the Gaussian channel. Because $\mathsf{INR}<\mathsf{SNR}$, the least significant bit (LSB) of each transmitter appears below noise level at the other receiver and is invisible.

In the discussions below, bit $a_k\in\mathbb{F}_2$ denotes the bit sent at the kth level from the most significant bit (MSB) at transmitter 1 and similarly $b_k\in\mathbb{F}_2$ denotes the bit sent at the kth level at transmitter 2.

We begin with the baseline where two receivers are not allowed to cooperate. The transmitted signals are naturally broken down into two parts: 1) the common levels, which appear at both receivers and 2) the private levels, which only appear at its own receiver. Each transmitter splits its message into common and

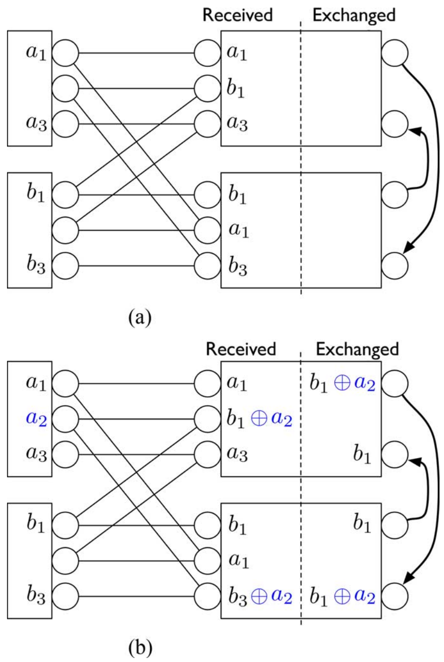


Fig. 3. An example channel. (a) Without cooperation. (b) With cooperation.

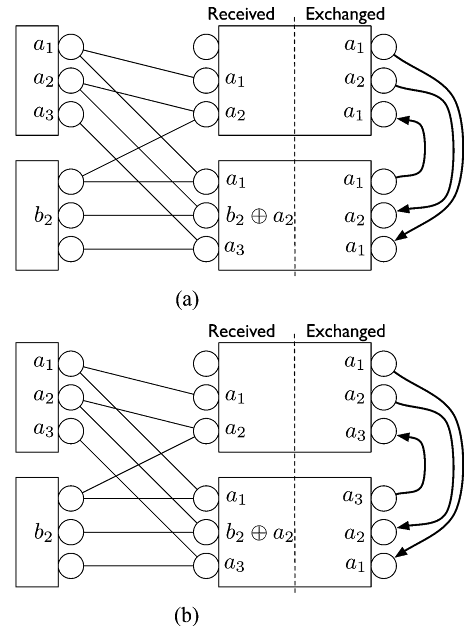


Fig. 4. An asymmetric example. (a) Suboptimal scheme. (b) Optimal scheme.

private parts, which are linearly modulated onto the common and private levels of the signal respectively. Each receiver then decodes both user's common messages and its own private message by solving the linear equations it received. This is shown to be optimal in the two-user interference channel [27]. In this example [Fig. 3(a)], bits a_1 and b_1 are common, while a_3 and b_3 are private. The sum capacity without cooperation is 4 bits. One cannot turn on the bit a_2 (or b_2) since the number of variables (bits) to be solved at the receiver 1, that is, $\{a_1, a_3, b_1\}$, has already met the maximum number of equations it has.

With receiver cooperation, the natural split of transmitted signals does not change. This suggests that the encoding procedure and the aim of each decoder remain the same. Each receiver with the help from the other receiver, however, is able to decode more information because it has additional linear equations. Since each user's private message is not of interest to the other receiver, a natural scheme for receiver cooperation is to exchange linear combinations formed by the signals *above* the private signal level so that the undesired signal does not pollute the cooperative information. In this example, as illustrated in Fig. 3(b), with one-bit cooperation in each direction in the LDC, the optimal sum rate is 5 bits, achieved by turning on one over-the-air bit a_2. This causes collisions at the second level at receiver 1 and at the third level at receiver 2, but they can be resolved with cooperation: receiver 1 sends $b_1 \oplus a_2$ to receiver 2 and receiver 2 sends b_1 to receiver 1. Now receiver 1 can solve (a_1, a_2, a_3, b_1) and receiver 2 can solve (b_1, b_3, a_1, a_2). In fact, the exchanged linear combinations are not unique. For example, receiver 1 can send $(b_1 \oplus a_2) \oplus a_1$ and receiver 2 can send $b_1 \oplus a_1$ and this again achieves the same rates. As long as receiver 1 does not send a linear combination containing the private bit a_3 and the sent linear combination is linearly independent of the signals at receiver 2 (and *vice versa* for the linear combination sent from receiver 2 to receiver 1), the scheme is optimal for this example channel. The above discussion regarding the scheme in the LDC naturally leads to an implementable one-round scheme in the Gaussian channel, where both receivers quantize and bin their received signals at their own private signal level.

In the above example, it is optimal that each receiver sends to each other linear combinations formed by its received signals above its private signal level. Is this optimal in general? The answer is no. Consider the following asymmetric example: $\mathsf{SNR}_2 = \mathsf{INR}_2$, SNR_1 is 2/3 of SNR_2 in dB and INR_1 is 1/3 of SNR_2 in dB. $\mathsf{C}^{\mathsf{B}}_{12} = \frac{2}{3}\log\mathsf{SNR}_2$ and $\mathsf{C}^{\mathsf{B}}_{21} = \frac{1}{3}\log\mathsf{SNR}_2$. The corresponding LDC is depicted in Fig. 4, where one bit in the LDC corresponds to $\frac{1}{3}\log\mathsf{SNR}_2$ in the Gaussian channel. First consider the same scheme as that in the previous exmaple. Note that if receiver 2 just forwards signals above its private signal level, it can only forward a_1 to receiver 1 and achieves R_1 up to 2 bits. On the other hand, if receiver 2 forwards a_3 to receiver 1, which is below user 2's private signal level, it achieves $R_1 = 3$ bits. From this example, we see that once there is "useful" information (which should not be polluted by the receiver's own private bits) which lies *at or below* the private signal level (in this example, the bit a_3), the one-round scheme described in the previous example is suboptimal. To extract the useful information at or below the private signal level, one of the receivers (in this example, receiver 2) can first decode and then form linear combinations using (decoded) common messages *only*.

It turns out that without loss of generality, the above situation (where there is useful information for the other receiver lies at or below the private signal level) only happens at most at one receiver. In other words, there exists a receiver where no useful information (for the other receiver) lies at or below the private signal level. The reason is the following:

1) It is straightforward to see that the capacity region is convex and hence if a scheme can achieve $\max_{(R_1,R_2)\in\mathscr{C}} \{\mu_1 R_1 + \mu_2 R_2\}$ for all $\mu_1, \mu_2 \geq 0$, it is optimal.

2) If $\mu_1 \geq \mu_2$, we weigh user 1's rate more. Since the private bits are cheaper to support in the sense that they do not cause interference at receiver 2, user 1 should be transmitting at its full private rate, which is equal to the number of levels at or below the private signal level at receiver 1. Therefore, all levels at or below the private signal level are occupied by user 1's private bits and there is no useful information at receiver 1 for receiver 2.
3) Similarly if $\mu_1 \leq \mu_2$, there is no useful information at receiver 2 for receiver 1 at or below the private signal level.

Hence, the following two-round strategy turns out to be optimal in the LDC (the proof is omitted here): if $\mu_1 \geq \mu_2$, receiver 1 forms a certain number (no more than the cooperative link capacity) of linear combinations composed of the signals above its private signal level and sends them to receiver 2. After receiver 2 decodes, it forms a certain number of linear combinations composed of the decoded common bits and sends them to receiver 1. If $\mu_1 \leq \mu_2$, the roles of receiver 1 and 2 are exchanged. Note that depending on the operating point in the capacity region, we use different configurations, implying that time-sharing is needed to achieve the full capacity region.

From the above discussion, a natural and implementable two-round strategy for Gaussian channels emerges. For the transmission, we use a superposition Gaussian random coding scheme with a simple power-split configuration, as described in [3]. For the cooperative protocol, one of the receivers quantize-and-bins its received signal at its private signal level and forwards the bin index; after the other receiver decodes with the helping side information, it bin-and-forwards the decoded common messages back to the first receiver and helps it decode. In Section V, we shall prove that this strategy achieves the capacity region universally to within 2 bits per user.

B. Conventional Compress-Forward and Decode-Forward

We have motivated the two-round strategy to be proposed formally in the next section from a high level perspective. Below we shall illustrate why conventional compress-forward (CF) and decode-forward (DF) are not good in certain regimes.

It is a standard approach to evaluate achievable rates of Gaussian relay networks using conventional compress-forward with Gaussian vector quantization (VQ) assuming joint Gaussianity of the received signals at relays and destination in the literature, including [10]–[13], [20], etc. What if we replace the quantize-binning part in the two-round strategy proposed above by the conventional compress-forward with Gaussian VQ, as in [10], [11], and [28], where the two-round idea is also used?

Let us consider another symmetric channel with $\log \mathsf{INR} = \frac{3}{5}\log \mathsf{SNR}$ and $\mathsf{C}^{\mathsf{B}} = \frac{1}{5}\log \mathsf{SNR}$. From its corresponding LDC in Fig. 5, one can see that the two received signals of the Gaussian channel, (y_1, y_2), are not jointly Gaussian. The reason is that, supposing they are jointly Gaussian, the conditional distribution of y_2 given y_1 should be marginally Gaussian. As Fig. 5 suggests, however, conditioning on receiver 1's signal results in a hole at the third level of receiver 2's signal, which was occupied by a_1. Therefore, transmitter 2's common codebook is not dense enough to make the conditional distribution of y_2 given y_1 marginally Gaussian. The incorrect assumption results in larger quantization distortions, as depicted in Fig. 5(b) [1]. The information sent from receiver 1 to receiver 2, a_1, is *redundant* and cannot help mitigate interference a_2. Hence, the achievable sum rate is 7 bits (4 bits for user 1 and 3 bit for user 2), while the one-round scheme in Fig. 5(a) achieves 8 bits. Recall that 1 bit in the LDC corresponds to $\frac{1}{5}\log \mathsf{SNR}$ in the Gaussian channel, therefore the performance loss is unbounded as $\mathsf{SNR} \to \infty$. The main reason why conventional compress-forward does not work well is that the scheme does not well utilize the dependency between the two received signals.

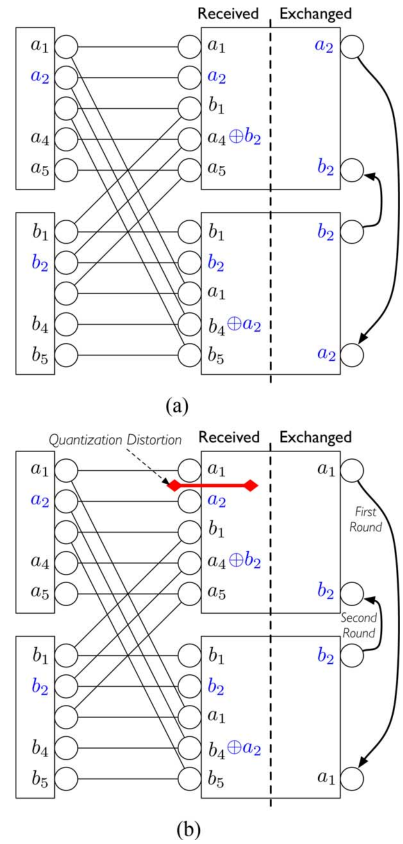


Fig. 5. Another example channel. (a) Optimal scheme (one round). (b) Conventional compress-forward in first round.

Another standard approach is to use decode-forward for the two receivers to cooperate. Let us go back to the first example and consider the channel in Fig. 3. Note that there is no gain if we require both common messages to be decoded at one of the receivers at the first stage without cooperation. By symmetry we can assume that, without loss of generality, each receiver first decodes its own common message and then bin-and-forwards the decoded information to the other receiver. At the second stage, it then decodes the other user's common message with the help from cooperation and decodes its own private message. In the corresponding LDC, the common bit a_2 cannot be decoded at the first stage and hence the total throughput using this strategy is at most 4 bits, which is again the same as that without

[1]If we view the received signals as vectors of bits rather than binary expansions of Gaussian signals, we are not restricted to send the MSB a_1 to receiver 2 and a_2 can be sent instead. However, this kind of scheme cannot be implemented in the Gaussian scenario using conventional compress-forward with Gaussian VQ.

cooperation. The reason why decode-forward is not good for the two receivers to cooperate is that, it is too costly to decode users' own common message at the first stage without the help from cooperation.

IV. A Two-Round Strategy

In this section we describe the two-round strategy and derive its achievable rate region. The strategy consists of two parts: 1) the transmission scheme and 2) the cooperative protocol.

A. Transmission Scheme

We use a simple superposition coding scheme with Gaussian random codebooks. For each transmitter, it splits its own message into common and private (sub-)messages. Each common message is aimed at both receivers, while each private one is aimed at its own receiver. Each message is encoded into a codeword drawn from a Gaussian random codebook with a certain power. For transmitter i, the power for its private and common codes are Q_{ip} and $Q_{ic} = 1 - Q_{ip}$ respectively, for $i = 1, 2$. As [3] points out, since the private signal is undesired at the unintended receiver, a reasonable configuration is to make the private interference at or below the noise level so that it does not cause much damage and can still convey additional information in the direct link if it is stronger than the cross link. When the interference is stronger than the desired signal, simply set the whole message to be common. In a word, for $(i, j) = (1, 2)$ or $(2, 1)$, $Q_{ip} = \min\left\{\frac{1}{\mathsf{INR}_j}, 1\right\}$ if $\mathsf{SNR}_i > \mathsf{INR}_j$ and $Q_{ip} = 0$ otherwise.

B. Cooperative Protocol

The cooperative protocol is two-round. We briefly describe it as follows: for $(i, j) = (1, 2)$ or $(2, 1)$, at the first round, receiver j quantizes its received signal and sends out the bin index (the procedure is described in detail below). At the second round, receiver i receives this side information, decodes its desired messages (both users' common messages and its own private message) with the decoder described in detail below, randomly bins the decoded common messages and sends the bin indices to receiver j. Finally receiver j decodes with the help from the receiver-cooperative link. We call this a two-round strategy $\mathrm{STG}_{j \to i \to j}$, meaning that the processing order is: receiver j quantize-and-bins, receiver i decode-and-bins and receiver j decodes. Its achievable rate region is denoted by $\mathscr{R}_{j \to i \to j}$. By time-sharing, we can obtain achievable rate region $\mathscr{R} := \mathrm{conv}\left\{\mathscr{R}_{2 \to 1 \to 2} \cup \mathscr{R}_{1 \to 2 \to 1}\right\}$, convex hull of the union of two rate regions.

Remark 4.1 (Engineering Interpretation): There is a simple way to understand the strategy from an engineering perspective. To achieve $\max_{(R_1, R_2) \in \mathscr{R}} \{\mu_1 R_1 + \mu_2 R_2\}$ for some nonnegative (μ_1, μ_2), the processing configuration can be easily determined: strategy $\mathrm{STG}_{j \to i \to j}$ should be used, where $i = \arg\min_{l=1,2}\{\mu_l\}$ and $j = \arg\max_{l=1,2}\{\mu_l\}$. In a word, the receiver which decodes last is the one we favor most. This is the high-level intuition we obtained from the discussion in the LDC in Section III-A.

In the following, we describe each component in detail, including quantize-binning, decode-binning, and their corresponding decoders. For simplicity, we consider strategy $\mathrm{STG}_{2 \to 1 \to 2}$.

Quantize-Binning (Receiver 2): Upon receiving its signal from the transmitter-receiver link, receiver 2 does not decode messages immediately. Instead, serving as a relay, it first quantizes its signal by a pregenerated Gaussian quantization codebook with certain distortion and then sends out a bin index determined by a pregenerated binning function. How should we set the distortion? As discussed in the previous section, note that both its own (user 2's) private signal and the noise it encounters are not of interest to receiver 1. Therefore, a natural configuration is to set the distortion level equal to *the aggregate power level of the noise and user 2's private signal*.

Decoder at Receiver 1: After retrieving the receiver-cooperative side information, that is, the bin index, receiver 1 decodes two common messages and its own private message, by searching in transmitters' codebooks for a codeword triple (indexed by user 1 and user 2's common messages and user 1's own private message) that is jointly typical with its received signal and some quantization point (codeword) in the given bin. If there is no such unique codeword triple, it declares an error.

Decode-Binning (Receiver 1): After receiver 1 decodes, it uses two pregenerated binning functions to bin the two common messages and sends out these two bin indices to receiver 2.

Decoder at Receiver 2: After receiving these two bin indices, receiver 2 decodes two common messages and its own private message, by searching in the corresponding bins (containing common messages) and user 2's private codebook for a codeword triple that is jointly typical with its received signal.

Remark 4.2 (Difference From the Conventional CF): The action of receiver 2 as a relay is very similar to that of the relay in the conventional compress-forward with Gaussian vector quantization. Note that the main difference from the conventional compress-forward with Gaussian vector quantization lies in the *decoding* procedure (at receiver 1) and the chosen distortion. In the conventional Gaussian compress-forward, the decoder first searches in the bin for one quantization codeword that is jointly typical with its received signal from its own transmitter *only*, assuming that the two received signals are jointly Gaussian. This may not be true since a single user may not transmit at the capacity in its own link, which results in "holes" in signal space. As a consequence, this scheme may not utilize the dependency of two received signals well and cause larger distortions. Our scheme, on the other hand, utilizes the dependency in a better way by *jointly* deciding the quantization codeword and the message triple, consequently allows smaller distortions and is able to reveal the beneficial side information to the other receiver. Quantize-binning and its corresponding decoding part of our scheme is very similar to *extended hash-and-forward* proposed in [22], in which it was pointed out that the scheme has no advantage over conventional compress-forward in a single-source single-relay setting. In the Gaussian single-relay channel (with orthogonal noise-free relay-destination link), the received signal at the relay and the destination are indeed jointly Gaussian when communicating at the quantize-map-and-forward achievable rate and hence the performances of the two schemes are the same. Due to the above mentioned issues, however, we recognize in our problem where the channel consists of two source-destination pairs and two relays, the scheme has an unbounded advantage over the conventional compress-forward in certain regimes. Such phenomena are also observed in single-source single-destination Gaussian relay networks [25], [29] and interference-relay channels [17], [29].

C. Achievable Rates

The following theorem establishes the achievable rates of strategy $\text{STG}_{2\longrightarrow 1\longrightarrow 2}$. Let R_{ic} and R_{ip} denote the rates for user i's common message and private message respectively, for $i=1,2$.

Theorem 4.3 (Achievable Rate Region for $\text{STG}_{2\longrightarrow 1\longrightarrow 2}$): The rate tuple $(R_{1c}, R_{2c}, R_{1p}, R_{2p})$ satisfying the following constraints is achievable:

Constraints at receiver 1: At the bottom of the page, where

$$\xi_1 = I\left(\hat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1\right).$$

For $i=1,2$, $x_{ic}\sim\mathcal{CN}\left(0,Q_{ic}\right)$ is the common codebook generating random variable. $x_1 = x_{1p}+x_{1c}$ is the superposition codebook generating variable, where $x_{1p}\sim\mathcal{CN}\left(0,Q_{1p}\right)$ is independent of x_{1c}. $\hat{y}_2 \overset{d}{=} y_2+\hat{z}_2$ is the quantization codebook generating random variable and $\hat{z}_2\sim\mathcal{CN}\left(0,\Delta_2\right)$, independent of everything else. Δ_2 is the quantization distortion at receiver 2.

Constraints at receiver 2:

$$\begin{aligned} R_{2p} &\le I\left(x_2; y_2 | x_{2c}, x_{1c}\right) \\ R_{1c}+R_{2p} &\le I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c}+R_{2p} &\le I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c}+R_{1c}+R_{2p} &\le I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}}, \end{aligned}$$

where $x_2 = x_{2p}+x_{2c}$ is the superposition codebook generating variable and $x_{2p}\sim\mathcal{CN}\left(0,Q_{2p}\right)$ is independent of x_{2c}.

Proof: For details, see Appendix A. Here we give some high-level comments on these rate constraints. First, unlike interference channels without cooperation, here receiver 1 is required to decode m_{2c} correctly so that it can help receiver 2. This additional requirement gives the rate constraint (2) on R_{2c}.

Second, in the set of constraints at receiver 1, on the right-hand side they are all minimum of two terms. The second term corresponds to the case when the receiver-cooperative link is strong enough to convey the quantized $\hat{y}_2^N$ correctly. The first term corresponds to the case when receiver 1 can only figure out a set of candidates of quantized $\hat{y}_2^N$. Regarding the "rate loss" term ξ_1, in Section III we see that in the LDC as long as the quantization level is chosen such that no private signals pollute the cooperative information, there is no such penalty. In fact, $\xi_1 = I\left(\hat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1\right)$ corresponds to the number of private bits polluting the cooperative linear combinations in the LDC if one chooses the quantization distortion to be too small. In the Gaussian channel, however, due to the carry-over of real additions, the private part will always "leak" into the levels above the quantization level and hence there is always at least a bounded rate loss even if we choose the quantization distortion properly.

Finally, in the set of constraints at receiver 2, since receiver 1 only helps receiver 2 decode m_{1c} and m_{2c}, there is no enhancement in R_{2p}. ■

We shall use the following shorthand notations throughout the rest of the paper: for $(i,j)=(1,2),(2,1)$,

$$\begin{aligned} \mathsf{SNR}_{ip} &:= |h_{ii}|^2 Q_{ip} = \mathsf{SNR}_i \cdot Q_{ip}, \\ \mathsf{INR}_{ip} &:= |h_{ij}|^2 Q_{jp} = \mathsf{INR}_i \cdot Q_{jp}. \end{aligned}$$

Next, we quantify the "rate loss" term ξ_1 in the set of rate constraints at receiver 1, in terms of distortions Δ_2:

$$\begin{aligned} \xi_1 &= I(\hat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1) \\ &= h(\hat{y}_2 | x_{1c}, x_1, x_{2c}, y_1) - h(\hat{y}_2 | x_{1c}, x_1, x_{2c}, y_1, y_2) \\ &= h\left(h_{22}x_{2p}+z_2+\hat{z}_2 | h_{12}x_{2p}+z_1\right) - h\left(\hat{z}_2\right) \\ &= \log\left(\frac{1+\Delta_2}{\Delta_2} + \frac{\mathsf{SNR}_{2p}}{(1+\mathsf{INR}_{1p})\Delta_2}\right) \\ &\le \log\left(\frac{1+\Delta_2+\mathsf{SNR}_{2p}}{\Delta_2}\right). \end{aligned} \tag{14}$$

Below we shall see why the intuition of quantizing at the private signal level works. By choosing $\Delta_2 = 1+\mathsf{SNR}_{2p}$, the "rate loss" ξ_1 is upper bounded by 1. In particular, when $\mathsf{SNR}_2 \le \mathsf{INR}_1$, we have $\mathsf{SNR}_{2p}=0$ and hence $\xi_1 = 1$. On the other hand, note that for receiver 1 the unwanted signal power

$$R_{1p} \le \min\left\{ I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + (\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+, I\left(x_1; y_1, \hat{y}_2 | x_{1c}, x_{2c}\right) \right\} \tag{1}$$

$$R_{2c} \le \min\left\{ I\left(x_{2c}; y_1 | x_1\right) + (\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+, I\left(x_{2c}; y_1, \hat{y}_2 | x_1\right) \right\} \tag{2}$$

$$\begin{aligned} R_{2c}+R_{1p} &\le \min\left\{ I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+, I\left(x_{2c}, x_1; y_1, \hat{y}_2 | x_{1c}\right) \right\} \\ R_{1c}+R_{1p} &\le \min\left\{ I\left(x_1; y_1 | x_{2c}\right) + (\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+, I\left(x_1; y_1, \hat{y}_2 | x_{2c}\right) \right\} \\ R_{1c}+R_{2c}+R_{1p} &\le \min\left\{ I\left(x_1, x_{2c}; y_1\right) + (\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+, I\left(x_1, x_{2c}; y_1, \hat{y}_2\right) \right\}, \end{aligned} \tag{3}$$

level in y_2 is exactly $1+\mathsf{SNR}_{2p}$ and receiver 1 treats the unwanted signals as noise anyway. Hence, replacing $\hat{y}_2$ by y_2 only increases the rate by a bounded gain.

Remark 4.4: The above configuration of the distortion may not be optimal. The achievable rates can be further improved if we optimize over all possible distortions. For example, if the cooperative link capacity is large, one could lower the distortion level to yield a finer description of received signals. With the above simple configuration, however, we are able to show that it achieves the capacity region to within a bounded gap. Also note that in this paper, we generate the quantization codebook in a slightly different way than that in conventional lossy source coding, where instead a "test channel" $y_2 = \hat{y}_2 + \hat{z}_2$ is used. With this choice the rate loss ξ_1 can be made smaller, while the calculations become more complicated.

V. Characterization of the Capacity Region to Within 2 Bits

The main result in this section is the characterization of the capacity region to within 2 bits per user universally, regardless of channel parameters. To prove it, first we provide outer bounds of the capacity region. Ideas about how to prove them are outlined and details are left in appendices. Then we make use of Theorem 4.3 to evaluate the achievable rate region and show that it is within 2 bits per user to the proposed outer bounds.

A. Outer Bounds

To prove the outer bounds, the main idea is the following: first, upper bound the rates by mutual informations via Fano's inequality and data processing inequality; second, decompose them into two parts: 1) terms which are similar to those in Gaussian interference channels without cooperation and 2) terms which correspond to the enhancement from cooperation. We use the genie-aided techniques in [3] to upper bound the first part and obtain namely the Z-channel bound (where the genie gives interfering symbols x_j^N to receiver i, $i \neq j$) and ETW-bound (where the genie gives the interference term caused by user i at receiver j, $s_i^N := h_{ji}x_i^N + z_j^N$ to receiver i). For the second part, we make use of the fact that u_{12}^N and u_{21}^N are both functions of (y_1^N, y_2^N) and other straightforward bounding techniques. The results are summarized in the following lemma.

Lemma 5.1: $\mathscr{C} \subseteq \bar{\mathscr{C}}$, where $\bar{\mathscr{C}}$ consists of nonnegative rate tuples (R_1, R_2) satisfying the inequalities (4)–(13) at the bottom of the page.

Proof: Details are left in Appendix B. Below we give a short outline and intuitions. First of all, bounds (4), (5), and (9) are straightforward cut-set upper bounds of individual rates and sum rate respectively.

Bound (6) corresponds to the ETW-bound in Gaussian interference channels without cooperation. In the genie-aided

$$R_1 \le \log(1+\mathsf{SNR}_1) + \min\left\{\mathsf{C}_{21}^{\mathsf{B}}, \log\left(1+\frac{\mathsf{INR}_2}{1+\mathsf{SNR}_1}\right)\right\} \tag{4}$$

$$R_2 \le \log(1+\mathsf{SNR}_2) + \min\left\{\mathsf{C}_{12}^{\mathsf{B}}, \log\left(1+\frac{\mathsf{INR}_1}{1+\mathsf{SNR}_2}\right)\right\} \tag{5}$$

$$R_1+R_2 \le \log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \mathsf{C}_{21}^{\mathsf{B}} + \mathsf{C}_{12}^{\mathsf{B}} \tag{6}$$

$$R_1+R_2 \le \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) + \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \mathsf{C}_{12}^{\mathsf{B}} \tag{7}$$

$$R_1+R_2 \le \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \mathsf{C}_{21}^{\mathsf{B}} \tag{8}$$

$$R_1+R_2 \le \log\left(1+\mathsf{SNR}_1+\mathsf{SNR}_2+\mathsf{INR}_1+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right) \tag{9}$$

$$\begin{aligned} 2R_1+R_2 \le{}& \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) \\ &+ \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \mathsf{C}_{21}^{\mathsf{B}} + \mathsf{C}_{12}^{\mathsf{B}} \end{aligned} \tag{10}$$

$$\begin{aligned} R_1+2R_2 \le{}& \log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) \\ &+ \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \end{aligned} \tag{11}$$

$$\begin{aligned} 2R_1+R_2 \le{}& \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2+\mathsf{SNR}_1+\frac{\mathsf{INR}_1}{1+\mathsf{INR}_1}+\frac{|h_{11}h_{22}-h_{12}h_{21}|^2}{1+\mathsf{INR}_1}\right) \\ &+ \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \mathsf{C}_{21}^{\mathsf{B}} \end{aligned} \tag{12}$$

$$\begin{aligned} R_1+2R_2 \le{}& \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1+\mathsf{SNR}_2+\frac{\mathsf{INR}_2}{1+\mathsf{INR}_2}+\frac{|h_{11}h_{22}-h_{12}h_{21}|^2}{1+\mathsf{INR}_2}\right) \\ &+ \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) + \mathsf{C}_{12}^{\mathsf{B}} \end{aligned} \tag{13}$$

channel, we upper bound the gain from receiver cooperation by $\mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}}$, that is, in both directions each bit is useful.

Bounds (7) and (8) correspond to the Z-channel bounds. In the genie-aided channel, since the genie gives interfering symbols x_j^N to receiver i, $i \neq j$, there is no interference at receiver i. Intuitively, the cooperation from receiver j to i is now providing only the power gain and the genie can provide y_j^N to receiver i to upper bound this power gain. The gain from the cooperation from receiver i to j is upper bounded by $\mathsf{C}_{ij}^{\mathsf{B}}$.

Bounds (10) and (11) on $R_i + 2R_j$ are derived by giving side information s_i^N to receiver i and side information x_i^N and y_i^N to one of the receiver j's. In the genie-aided channel there is an underlying Z-channel structure and hence the gain from one direction of the cooperation is absorbed into a power gain. The rest is upper bounded by $\mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}}$.

Bounds (12) and (13) on $R_i + 2R_j$ are derived by giving side information y_j^N and $\tilde{s}_i^N := h_{ji}x_i^N + \tilde{z}_j^N$, where $\tilde{z}_j \sim \mathcal{CN}(0,1)$ and independent of everything else, to receiver i and side information y_i^N to one of the receiver j's. In the genie-aided channel, there is an underlying point-to-point MIMO channel and hence the gain from both directions of cooperation is absorbed into the MIMO system. The rest is upper bounded by $\mathsf{C}_{ij}^{\mathsf{B}}$.

Note that the derivation of all bounds works for all INR's and SNR's. ■

We make the following observations:

Remark 5.2 (Dependence on Phases): The sum-rate cut-set bound (9) not only depends on SNR's and INR's but also on the phases of channel coefficients, due to the term $|h_{11}h_{22} - h_{12}h_{21}|^2$. In particular, when the receiver-cooperative link capacities C^{B}'s are large, the two receivers become near-fully cooperated and the system performance is constrained by that of the SIMO MAC; that is, it enters the saturation region. Therefore, this bound becomes active and the outer bound depends on phases.

Remark 5.3 (Strong Interference Regime): When $\mathsf{SNR}_1 \leq \mathsf{INR}_2$ and $\mathsf{SNR}_2 \leq \mathsf{INR}_1$, unlike the Gaussian interference channel of which the capacity region is equal to that of a compound MAC in the strong interference regime [1], here we cannot apply Sato's argument. Recall that when there is no cooperation, once user i's own message is decoded successfully at receiver i, it can produce $\tilde{y}_j^N$ which has the same distribution as y_j^N. Since the error probability for decoding user j's message at receiver j only depends on the *marginal* distribution of y_j^N, it can be concluded that at receiver i one can achieve the same performance for decoding user j's message by using the same decoder as that in receiver j and hence receiver i can decode user j's message successfully as well. When there is cooperation, however, the error probability for decoding user j's message at receiver j depends on the *joint* distribution of (y_j^N, u_{ij}^N). Note that the additive noise terms in $\tilde{y}_j^N$ and y_j^N have different correlations with the noise term z_i^N and u_{ij}^N can be highly correlated with z_i^N. As a consequence, the joint distributions of (y_j^N, u_{ij}^N) and $(\tilde{y}_j^N, u_{ij}^N)$ are not guaranteed to be the same and receiver i may not be able to achieve the same performance for decoding user j's message by using the same decoder as that in receiver j. Therefore, we cannot claim that the capacity region under strong interference condition is the same as that of compound MAC with conferencing decoders (CMAC-CD). Instead, we take the Z-channel bounds (7) and (8), which are within 1 bit to the sum rate cut-set bound of CMAC-CD in strong interference regimes. This will be discussed in the last part of this section.

B. Capacity Region to Within 2 Bits

Subsequently we investigate three qualitatively different cases, namely, weak interference, mixed interference, and strong interference,[2] in the rest of this section. We summarize the main achievability result in the following theorem (recall that $\bar{\mathscr{C}}$ is the outer bound region defined in Lemma 5.1):

Theorem 5.4 (Within Two-Bit Gap to Capacity Region):

$$\mathscr{R} \subseteq \mathscr{C} \subseteq \bar{\mathscr{C}} \subseteq \mathscr{R} \oplus \left([0,2] \times [0,2]\right),$$

Proof: Proved by Lemma 5.5, 5.8, and 5.11 in the rest of this section. ■

C. Weak Interference

In the case $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 > \mathsf{INR}_1$, the configuration of superposition coding is to split message m_i into m_{ic} and m_{ip}, for both users $i = 1, 2$. We first consider $\mathrm{STG}_{2 \longrightarrow 1 \longrightarrow 2}$: referring to Theorem 4.3, we obtain the set of achievable rates $(R_{1c}, R_{2c}, R_{1p}, R_{2p})$. The term $\xi_1 \leq 1$ bit, due to (14) in Section IV-C and the chosen distortion $\Delta_2 = 1 + \mathsf{SNR}_{2p}$.

To simplify calculations, note that the right-hand-side of (1)–(3) are at most a bounded gap from their lower bounds $I(x_1; y_1|x_{1c}, x_{2c})$, $I(x_{2c}; y_1|x_1)$ and $I(x_1; y_1|x_{2c})$ respectively. Therefore, we replace these three constraints by

$$\begin{aligned} R_{1p} &\leq I(x_1; y_1|x_{1c}, x_{2c}), \\ R_{2c} &\leq I(x_{2c}; y_1|x_1), \\ R_{1c} + R_{1p} &\leq I(x_1; y_1|x_{2c}) \end{aligned}$$

in the following calculations. Next, rewriting $R_{ip} = R_i - R_{ic}$ for $i = 1, 2$, applying Fourier-Motzkin algorithm to eliminate R_{1c} and R_{2c} and removing redundant terms (details omitted here), we obtain an achievable $\mathscr{R}_{2 \longrightarrow 1 \longrightarrow 2}$, which consists of nonnegative (R_1, R_2) satisfying

$$R_1 \leq \min \left\{ \begin{array}{ll} I(x_1; y_1|x_{2c}), & I(x_1; y_1|x_{1c}, x_{2c}) \\ +I(x_{1c}, x_2; y_2|x_{2c}) & + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\}$$

$$R_2 \leq \min \left\{ \begin{array}{l} I(x_2; y_2|x_{1c}) + \mathsf{C}_{12}^{\mathsf{B}}, \\ I(x_{2c}; y_1|x_1) + I(x_2; y_2|x_{1c}, x_{2c}) \end{array} \right\}$$

$$R_1 + R_2 \leq \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1) + I(x_2; y_2|x_{1c}, x_{2c}) \\ + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \quad (15)$$

$$R_1 + R_2 \leq I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2|x_{1c}, x_{2c}) \quad (16)$$

$$R_1 + R_2 \leq \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1|x_{1c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ +I(x_{1c}, x_2; y_2|x_{2c}) + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \quad (17)$$

$$R_1 + R_2 \leq \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2|x_{1c}) \\ +I(x_{1c}, x_2; y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \quad (18)$$

$$R_1 + R_2 \leq I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}_{12}^{\mathsf{B}} \quad (19)$$

[2]The definitions of these cases are the following: (1) weak interference, where $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 > \mathsf{INR}_1$; (2) mixed interference, where $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 \leq \mathsf{INR}_1$; (3) strong interference, where $\mathsf{SNR}_1 \leq \mathsf{INR}_2$ and $\mathsf{SNR}_2 \leq \mathsf{INR}_1$.

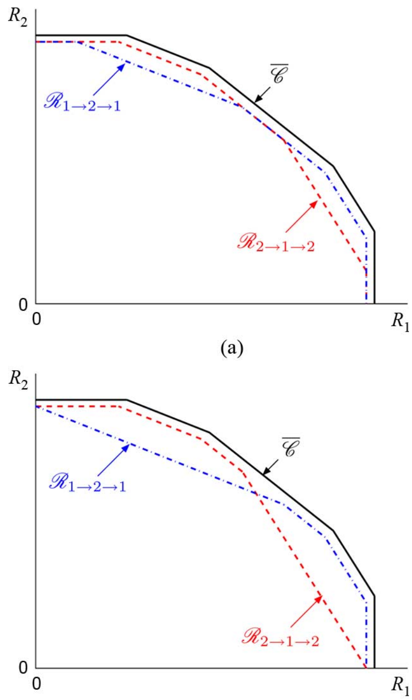


Fig. 6. Time-sharing to achieve approximate capacity region. (a) Taking union is required, while time-sharing is not. (b) Time-sharing is required.

$$R_1 + R_2 \le \left\{ \begin{array}{l} I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{2c}; y_1|x_1) \\ +I(x_{1c}, x_2; y_2|x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\} \quad (20)$$

$$2R_1 + R_2 \le \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1) + I(x_1; y_1|x_{1c}, x_{2c}) \\ +I(x_{1c}, x_2; y_2|x_{2c}) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \\ +\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\}$$

$$2R_1 + R_2 \le \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1|x_{1c}, x_{2c}) \\ +I(x_{1c}, x_2; y_2|x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\} \quad (21)$$

$$R_1 + 2R_2 \le \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1|x_{1c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}^{\mathsf{B}}_{12} \\ +I(x_2; y_2|x_{1c}, x_{2c}) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \end{array} \right\}$$

$$R_1 + 2R_2 \le \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1|x_{1c}) + I(x_{2c}; y_1|x_1) \\ +I(x_{1c}, x_2; y_2|x_{2c}) + I(x_2; y_2|x_{1c}, x_{2c}) \\ +\mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \end{array} \right\}$$

$$R_1 + 2R_2 \le \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2|x_{1c}) + I(x_{1c}, x_2; y_2) \\ +I(x_2; y_2|x_{1c}, x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\}$$

$$R_1 + 2R_2 \le \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2|x_{1c}) + I(x_{2c}; y_1|x_1) \\ +I(x_{1c}, x_2; y_2|x_{2c}) + I(x_2; y_2|x_{1c}, x_{2c}) \\ +\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\}.$$

We will show that except (21), all bounds are within a bounded gap from the corresponding outer bounds in Lemma 5.1. By symmetry, however, one can write down $\mathscr{R}_{1\longrightarrow 2\longrightarrow 1}$ and see that the troublesome constraint (21) can be compensated by time-sharing with rate points in $\mathscr{R}_{1\longrightarrow 2\longrightarrow 1}$. Therefore, the resulting $\mathscr{R} := \mathrm{conv}\left\{\mathscr{R}_{2\longrightarrow 1\longrightarrow 2} \cup \mathscr{R}_{1\longrightarrow 2\longrightarrow 1}\right\}$ is within a bounded gap from the outer bounds in Lemma 5.1. An illustration is provided in Fig. 6.

We give the following lemma.

Lemma 5.5 (Rate Region in the Weak Interference Regime):

$$\mathscr{R} \subseteq \mathscr{C} \subseteq \overline{\mathscr{C}} \subseteq \mathscr{R} \oplus \left([0,2] \times [0,2]\right),$$

in the weak interference regime.

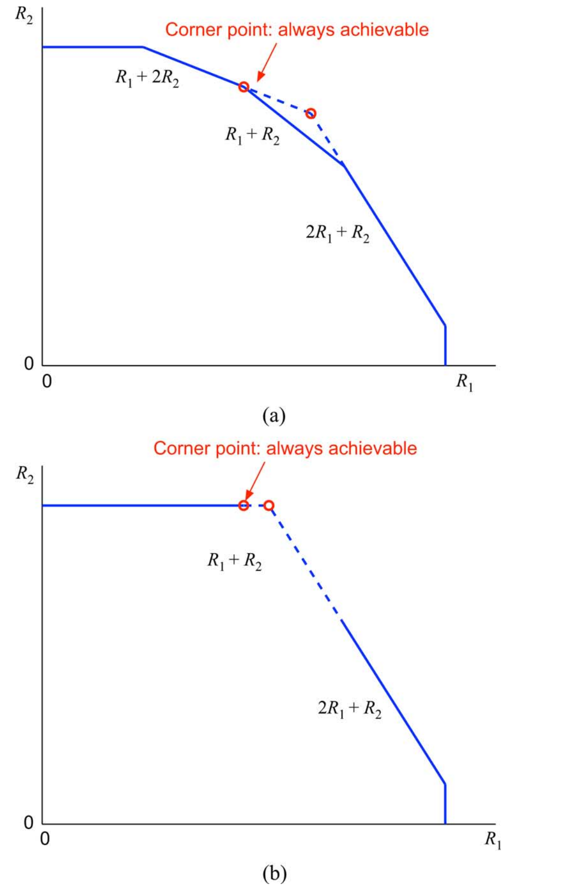


Fig. 7. Situations in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$. (a) $R_1 + 2R_2$ bound is active. (b) $R_1 + 2R_2$ bound is not active.

Proof: We need the following claims:

Claim: In $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$, whenever the $2R_1 + R_2$ bound (21) is active,

a) if $R_1 + 2R_2$ bounds are active, the corner point where $R_1 + R_2$ bound and $R_1 + 2R_2$ bound intersect can be achieved;

b) if $R_1 + 2R_2$ bounds are not active, the corner point where $R_1 + R_2$ bound and R_2 bound intersect can be achieved. Above two situations are illustrated in Fig. 7.

Proof: In both situations, we will argue that the value of $R_1 + R_2$ at the intersection of the dashed lines are always greater than the value of $R_1 + R_2$ at the desired corner point. Details are left in Appendix C. ∎

Therefore, the $2R_1 + R_2$ bound (21) and, by symmetry, its corresponding $R_1 + 2R_2$ bound in $\mathscr{R}_{1\longrightarrow 2\longrightarrow 1}$ do not show up in $\mathscr{R} = \mathrm{conv}\{\mathscr{R}_{2\longrightarrow 1\longrightarrow 2} \cup \mathscr{R}_{1\longrightarrow 2\longrightarrow 1}\}$ and $\mathscr{R}$ is within 2 bits per user to the outer bounds in Lemma 5.1. To show this, we first look at the bounds in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ except (21). We claim that

Claim 5.7: The bounds in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ except (21) satisfy:

- R_1 bound is within 2 bits to outer bounds;
- R_2 bound is within 2 bits to outer bounds;
- $R_1 + R_2$ bound is within 3 bits to outer bounds;
- $2R_1 + R_2$ bound is within 4 bits to outer bounds;
- $R_1 + 2R_2$ bound is within 5 bits to outer bounds.

Proof: See Appendix C. ■

By symmetry, we obtain similar results for $\mathscr{R}_{1\longrightarrow 2\longrightarrow 1}$ and hence conclude that the bounds in $\mathscr{R}$ satisfies (1) both R_1 and R_2 bounds are within 2 bits; (2) R_1+R_2 bound is within 3 bits; (3) both $2R_1+R_2$ and R_1+2R_2 bound are within 5 bits to their corresponding outer bounds. This completes the proof. ■

D. Mixed Interference

In the case $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 \le \mathsf{INR}_1$, the configuration of superposition coding is to split message m_1 into m_{1c} and m_{1p}, while making the whole m_2 common. We first consider $\mathrm{STG}_{2\longrightarrow 1\longrightarrow 2}$: by Theorem 4.3, rates satisfying those at the bottom of the page, and the following are achievable:

$$\begin{aligned} R_{1c} &\le I(x_{1c};y_2|x_2)+\mathsf{C}^{\mathsf{B}}_{12}\\ R_2 &\le I(x_2;y_2|x_{1c})+\mathsf{C}^{\mathsf{B}}_{12}\\ R_2+R_{1c} &\le I(x_2,x_{1c};y_2)+\mathsf{C}^{\mathsf{B}}_{12}, \end{aligned}$$

where $\xi_1 = 1$ since $\mathsf{SNR}_2 \le \mathsf{INR}_1$.

Again to simplify calculations, note that the right-hand-side of (22)–(24) are at most a bounded gap from their lower bounds $I(x_1;y_1|x_{1c},x_2)$, $I(x_2;y_1|x_1)$ and $I(x_1;y_1|x_2)$ respectively. Therefore, we replace these three constraints by

$$\begin{aligned} R_{1p} &\le I(x_1;y_1|x_{1c},x_2),\\ R_2 &\le I(x_2;y_1|x_1),\\ R_{1c}+R_{1p} &\le I(x_1;y_1|x_2) \end{aligned}$$

in the following calculations. Next, rewriting $R_{1p} = R_1 - R_{1c}$, applying Fourier-Motzkin algorithm to eliminate R_{1c} and removing redundant terms (details omitted here), we obtain an achievable $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$, consists of nonnegative (R_1,R_2) satisfying

$$\begin{aligned} R_1 &\le \min\left\{\begin{array}{l} I(x_1;y_1|x_2),I(x_1;y_1|x_{1c},x_2)\\ +I(x_{1c};y_2|x_2)+\mathsf{C}^{\mathsf{B}}_{12}\end{array}\right\}\\ R_2 &\le \min\left\{I(x_2;y_1|x_1),I(x_2;y_2|x_{1c})+\mathsf{C}^{\mathsf{B}}_{12}\right\}\\ R_1+R_2 &\le I(x_1,x_2;y_1)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+\\ R_1+R_2 &\le I(x_1,x_2;y_1,\hat{y}_2)\\ R_1+R_2 &\le I(x_1;y_1|x_{1c},x_2)+I(x_{1c},x_2;y_2)+\mathsf{C}^{\mathsf{B}}_{12}\\ R_1+R_2 &\le \left\{\begin{array}{l} I(x_1,x_2;y_1|x_{1c})+I(x_{1c};y_2|x_2)\\ +\mathsf{C}^{\mathsf{B}}_{12}+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+\end{array}\right\}\\ R_1+R_2 &\le I(x_1,x_2;y_1,\hat{y}_2|x_{1c})+I(x_{1c};y_2|x_2)+\mathsf{C}^{\mathsf{B}}_{12}\\ R_1+2R_2 &\le \left\{\begin{array}{l} I(x_1,x_2;y_1|x_{1c})+I(x_{1c},x_2;y_2)\\ +\mathsf{C}^{\mathsf{B}}_{12}+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+\end{array}\right\}\\ R_1+2R_2 &\le I(x_1,x_2;y_1,\hat{y}_2|x_{1c})+I(x_{1c},x_2;y_2)+\mathsf{C}^{\mathsf{B}}_{12}. \end{aligned}$$

Comparing $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ with the outer bounds in Lemma 5.1, one can easily conclude that:

Lemma 5.8 (Mixed Interference Rate Region):

$$\mathscr{R}_{2\longrightarrow 1\longrightarrow 2} \subseteq \mathscr{C} \subseteq \overline{\mathscr{C}} \subseteq \mathscr{R}_{2\longrightarrow 1\longrightarrow 2} \oplus \left([0,1.5]\times[0,1.5]\right),$$

in the mixed interference regime. Besides, $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2} \subseteq \mathscr{R}$.

Proof: We investigate the bounds in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ and claim that:

Claim 5.9: The bounds in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ satisfy
- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- R_1+R_2 bound is within 3 bits to outer bounds;
- R_1+2R_2 bound is within 3 bits to outer bounds.

Proof: See Appendix C ■

This completes the proof. ■

E. Strong Interference

In the case $\mathsf{SNR}_1 \le \mathsf{INR}_2$ and $\mathsf{SNR}_2 \le \mathsf{INR}_1$, it turns out that a one-round strategy $\mathrm{STG}_{\mathrm{OneRound}}$ described below suffices to achieve capacity to within a bounded gap. The transmission scheme is the same as that described in Section IV-A. The difference is that, both receivers quantize-and-bins their received signals and decode with the help from the side information, as described in Section IV-B. It is called one-round since both receivers decode after one-round exchange of informaion. Below is the coding theorem for this strategy:

Theorem 5.10: The rate tuple $(R_{1c},R_{2c},R_{1p},R_{2p})$ satisfying the following constraints are achievable for $\mathrm{STG}_{\mathrm{OneRound}}$:

Constraints at receiver 1: See equation at the bottom of the next page.

Constraints at receiver 2: Above constraints with index "1" and "2" exchanged.

Proof: The proof follows the same line as the proof of Theorem 4.3. There is no rate constraint for R_{jc} at receiver i for $(i,j) = (1,2)$ or $(2,1)$, since decoding m_{jc} incorrectly at receiver i does not account for an error. ■

$$R_{1p} \le \min\left\{I(x_1;y_1|x_{1c},x_2)+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+, I(x_1;y_1,\hat{y}_2|x_{1c},x_2)\right\} \quad (22)$$

$$R_2 \le \min\left\{I(x_2;y_1|x_1)+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+, I(x_2;y_1,\hat{y}_2|x_1)\right\} \quad (23)$$

$$R_2+R_{1p} \le \min\left\{I(x_2,x_1;y_1|x_{1c})+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+, I(x_2,x_1;y_1,\hat{y}_2|x_{1c})\right\}$$

$$R_{1c}+R_{1p} \le \min\left\{I(x_1;y_1|x_2)+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+, I(x_1;y_1,\hat{y}_2|x_2)\right\} \quad (24)$$

$$R_{1c}+R_2+R_{1p} \le \min\left\{I(x_1,x_2;y_1)+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+, I(x_1,x_2;y_1,\hat{y}_2)\right\}$$

Now, in the strong interference regime, the configuration of superposition coding is to make the whole message m_i common for both users $i = 1, 2$; in a word, there is no superposition eventually. One-round strategy $\text{STG}_{\text{OneRound}}$ yields achievable rate region $\mathscr{R}_{\text{OneRound}}$, which consists of nonnegative (R_1, R_2) satisfying

$$\begin{aligned} R_2 &\le \min\left\{\begin{array}{l} I(x_2; y_1|x_1) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+, \\ I(x_2; y_1, \hat{y}_2|x_1) \end{array}\right\} \\ R_1 &\le \min\left\{\begin{array}{l} I(x_1; y_1|x_2) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+ \\ , I(x_1; y_1, \hat{y}_2|x_2) \end{array}\right\} \\ R_1 + R_2 &\le \min\left\{\begin{array}{l} I(x_1, x_2; y_1) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+ \\ , I(x_1, x_2; y_1, \hat{y}_2) \end{array}\right\} \\ R_1 &\le \min\left\{\begin{array}{l} I(x_1; y_2|x_2) + (\mathsf{C}_{12}^{\mathsf{B}} - \xi_2)^+ \\ , I(x_1; y_2, \hat{y}_1|x_2) \end{array}\right\} \\ R_2 &\le \min\left\{\begin{array}{l} I(x_2; y_2|x_1) + (\mathsf{C}_{12}^{\mathsf{B}} - \xi_2)^+ \\ , I(x_2; y_2, \hat{y}_1|x_1) \end{array}\right\} \\ R_2 + R_1 &\le \min\left\{\begin{array}{l} I(x_2, x_1; y_2) + (\mathsf{C}_{12}^{\mathsf{B}} - \xi_2)^+ \\ , I(x_2, x_1; y_2, \hat{y}_1) \end{array}\right\}, \end{aligned} \quad (25)$$

where $\xi_i = 1$, for both $i = 1, 2$.

Comparing $\mathscr{R}_{\text{OneRound}}$ with the outer bounds in Lemma 5.1, one can easily conclude that

Lemma 5.11 (Strong Interference Rate Region):

$$\mathscr{R}_{\text{OneRound}} \subseteq \mathscr{C} \subseteq \bar{\mathscr{C}} \subseteq \mathscr{R}_{\text{OneRound}} \oplus \big([0,1] \times [0,1]\big),$$

in the strong interference regime. Besides, $\mathscr{R}_{\text{OneRound}} \subseteq \mathscr{R}$.

Proof: We investigate the bounds in $\mathscr{R}_{\text{OneRound}}$ and claim that:

Claim 5.12: The bounds in $\mathscr{R}_{\text{OneRound}}$ satisfy:
- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- $R_1 + R_2$ bound is within 2 bits to outer bounds.

Proof: See Appendix C. ■

This completes the proof. ■

F. Approximate Capacity of Compound MAC With Conferencing Decoders

As a side product of this work, we characterize the capacity region of the *compound multiple access channel with conferencing decoders* (CMAC-CD) to within 1 bit. The channel is defined as follows.

Definition 5.13: A compound multiple access channel with conferencing decoders (CMAC-CD), is a channel with the same setup as depicted in Fig. 2, while both receivers aim to decode both m_1 and m_2.

We give straightforward cut-set upper bounds as follows:

Lemma 5.14: If (R_1, R_2) is achievable, it must satisfy the following constraints:

$$\begin{aligned} R_1 &\le \min\left\{\begin{array}{l} \log(1+\mathsf{SNR}_1) + \mathsf{C}_{21}^{\mathsf{B}}, \\ \log(1+\mathsf{INR}_2) + \mathsf{C}_{12}^{\mathsf{B}}, \\ \log(1+\mathsf{SNR}_1+\mathsf{INR}_2) \end{array}\right\} \\ R_2 &\le \min\left\{\begin{array}{l} \log(1+\mathsf{SNR}_2) + \mathsf{C}_{12}^{\mathsf{B}}, \\ \log(1+\mathsf{INR}_1) + \mathsf{C}_{21}^{\mathsf{B}}, \\ \log(1+\mathsf{SNR}_2+\mathsf{INR}_1) \end{array}\right\} \\ R_1 + R_2 &\le \log(1+\mathsf{SNR}_1+\mathsf{INR}_1) + \mathsf{C}_{21}^{\mathsf{B}} \\ R_1 + R_2 &\le \log(1+\mathsf{SNR}_2+\mathsf{INR}_2) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_1 + R_2 &\le \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2 \\ +|h_{11}h_{22}-h_{12}h_{21}|^2 \end{array}\right). \end{aligned}$$

Proof: These are straightforward cut-set bounds. We omit the details here. ■

For achievability, we adapt the one-round scheme proposed above with no superposition coding at transmitters. Therefore, the rate region is exactly the same as (25). Hence, we conclude that:

Theorem 5.15 (Within 1 Bit to CMAC-CD Capacity Region): The scheme achieves the capacity of compound MAC with conferencing decoders to within 1 bit.

Proof: Following the same line in the proof of Lemma 5.11, we can conclude that the bounds in $\mathscr{R}_{\text{OneRound}}$ satisfy:
- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- $R_1 + R_2$ bound is within 1 bit to outer bounds.

This completes the proof. ■

This result implies that for the Gaussian compound MAC with conferencing decoders, a simple one-round strategy suffices to achieve the capacity region to within 1 bit universally, regardless of channel parameters.

VI. One-Round Strategy Versus Two-Round Strategy

In Section V we show that for the two-user Gaussian interference channel with conferencing decoders, the two-round strategy proposed in Section IV along with time-sharing achieves the capacity region to within 2 bits universally. One of the drawbacks of the two-round strategy, however, is the decoding latency. The quantize-binning receiver cannot proceed to decoding until the other receiver decodes and forwards the

---

<u>*Constraints at receiver* 1</u>:

$$\begin{aligned} R_{1p} &\le \min\left\{I(x_1; y_1|x_{1c}, x_{2c}) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+, I(x_1; y_1, \hat{y}_2|x_{1c}, x_{2c})\right\} \\ R_{2c} + R_{1p} &\le \min\left\{I(x_{2c}, x_1; y_1|x_{1c}) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+, I(x_{2c}, x_1; y_1, \hat{y}_2|x_{1c})\right\} \\ R_{1c} + R_{1p} &\le \min\left\{I(x_1; y_1|x_{2c}) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+, I(x_1; y_1, \hat{y}_2|x_{2c})\right\} \\ R_{1c} + R_{2c} + R_{1p} &\le \min\left\{I(x_1, x_{2c}; y_1) + (\mathsf{C}_{21}^{\mathsf{B}} - \xi_1)^+, I(x_1, x_{2c}; y_1, \hat{y}_2)\right\}. \end{aligned}$$

bin indices back. The latency is two times the block length, which can be large. To avoid such large delay, fortunately in some cases, the one-round strategy $\mathrm{STG}_{\mathrm{OneRound}}$ described in Section V-E suffices. One of such cases is the strong interference regime. This can be easily justified in the corresponding linear deterministic channel (LDC). At strong interference, all transmitted signals in the LDC are common. There is no useful information lies below the noise level since the signal is corrupted by the noise. Hence, quantize-binning at the noise level is sufficient to convey the useful information.

Another such cases is the symmetric setup, where $\mathsf{SNR} = \mathsf{SNR}_1 = \mathsf{SNR}_2$, $\mathsf{INR} = \mathsf{INR}_1 = \mathsf{INR}_2$ and $\mathsf{C}^\mathsf{B} = \mathsf{C}^\mathsf{B}_{12} = \mathsf{C}^\mathsf{B}_{21}$.

For the symmetric setup, a natural performance measure is the symmetric capacity, defined as follows:

Definition 6.1 (Symmetric Capacity):

$$C_{\mathrm{sym}} := \sup\left\{R : (R,R) \in \mathscr{C}\right\}.$$

It turns out that the one-round strategy suffices to achieve C_{sym} to within a bounded gap.

Theorem 6.2 (Bounded Gap to the Symmetric Capacity): The one-round strategy $\mathrm{STG}_{\mathrm{OneRound}}$ can achieve the symmtric capacity to within 3 bits.

Proof: See Appendix D. ■

The justification in the corresponding LDC is again simple. Since the performance measure in which we are interested is the symmetric capacity, we can without loss of generality assume that both transmitters are transmitting at full private rate, that is, the entropy of each user's private signals is equal to the number of levels below the private signal level. Therefore, at each receiver, there is no useful information below the private signal level and quantize-binning at the private signal level suffices to convey the useful information.

VII. Generalized Degrees of Freedom Characterization

With the characterization of the capacity region to within a bounded gap, we attempt to answer the original fundamental question: how much interference can one bit of receiver cooperation mitigate? For simplicity, we consider the symmetric setup.

By Lemma 5.1 and Theorem 5.4, we have the characterization of the symmetric capacity to within 2 bits:

Corollary 7.1 (Approximate Symmetric Capacity): Let $\bar{C}_{\mathrm{sym}}$ be the minimum of the below four terms:

$$\log(1+\mathsf{SNR}) + \min\left\{\mathsf{C}^\mathsf{B}, \log\left(1+\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right)\right\},$$
$$\log\left(1+\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)+\mathsf{C}^\mathsf{B},$$
$$\frac{1}{2}\log\left(1+\mathsf{SNR}+\mathsf{INR}\right)+\frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)+\frac{1}{2}\mathsf{C}^\mathsf{B},$$
$$\frac{1}{2}\log\left(1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2\right).$$

Then, $\bar{C}_{\mathrm{sym}} - 2 \le C_{\mathrm{sym}} \le \bar{C}_{\mathrm{sym}}$.

A. Generalized Degrees of Freedom

To study the behavior of the system performance in the linear region, we use the notion of *generalized degrees of freedom* (g.d.o.f.), which is originally proposed in [3]. A natural extension from the definition in [3] would be the following: let

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log\mathsf{INR}}{\log\mathsf{SNR}} = \alpha;\ \lim_{\mathsf{SNR}\to\infty}\frac{\mathsf{C}^\mathsf{B}}{\log\mathsf{SNR}} = \kappa,$$

and define the number of generalized degrees of freedom per user as

$$d := \lim_{\substack{\text{fix } \alpha,\kappa \\ \mathsf{SNR}\to\infty}} \frac{C_{\mathrm{sym}}}{\log\mathsf{SNR}}, \tag{26}$$

if the limit exists. With fixed α and κ, however, there are certain channel realizations under which (26) has different values and hence the limit does not exist. This happens when $\alpha = 1$, where the phases of the channel gains matter both in inner and outer bounds. In particular, its value can depend on whether the system MIMO matrix is well-conditioned or not.

From the above discussion we see that the limit does not exist, since for different channel phases and different INR settings the value of (26) may be different. The reason is that, the original notion proposed in [3] cannot capture the impact of *phases* in MIMO situations, while from Lemma 5.1 and Theorem 5.4, or Corollary 7.1, we see that our results depend on phases heavily, if the receiver-cooperative link capacity C^B is so large that MIMO sum-rate cut-set bound becomes active. Therefore, instead of claiming that the limit (26) exists for *all* channel realizations, we pose a reasonable distribution, namely, i.i.d. uniform distribution, on the phases, show that the limit exists *almost surely* and define the limit to be the number of *generalized degrees of freedom* per user.

Lemma 7.2: Let

$$|h_{ij}| = g_{ij},\ \angle h_{ij} = \Theta_{ij},\ \forall i,j \in \{1,2\},$$

where g_{ij}'s are deterministic and Θ_{ij}'s are i.i.d. uniformly distributed over $[0,2\pi]$. Then the limit (26) exists almost surely and is defined as the number of generalized degrees of freedom (per user) in the system.

Proof: We leave the proof in Appendix E. ■

Now that the number of g.d.o.f. is well-defined, we can give the following theorem:

Theorem 7.3 (Generalized Degrees of Freedom Per User): We have a direct consequence from Corollary 7.1:

For $0 \le \alpha < 1$,

$$d = \min\left\{1,\ \max(\alpha, 1-\alpha)+\kappa,\ 1-\frac{\alpha}{2}+\frac{\kappa}{2}\right\}.$$

For $\alpha \ge 1$,

$$d = \min\left\{\alpha,\ 1+\kappa,\ \frac{\alpha}{2}+\frac{\kappa}{2}\right\}.$$

Numerical plots for g.d.o.f. are given in Fig. 8. We observe that at different values of α, the gain from cooperation varies. By investigating the g.d.o.f., we conclude that at high SNR, when INR is below 50% of SNR in dB scale, one-bit cooperation per

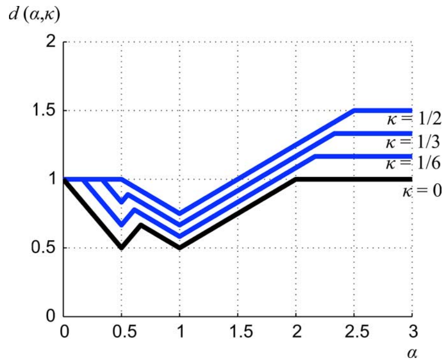


Fig. 8. Generalized degrees of freedom.

direction buys roughly one-bit gain per user until full receiver cooperation performance is reached, while when INR is between 67% and 200% of SNR in dB scale, one-bit cooperation per direction buys roughly half-bit gain per user until saturation.

B. Gain From Limited Receiver Cooperation

The fundamental behavior of the gain from receiver cooperation is explained in the rest of this section, by looking at two particular points: $\alpha = \frac{1}{2}$ and $\alpha = \frac{2}{3}$. Furthermore, we use the linear deterministic channel (LDC) for illustration.

At $\alpha = \frac{1}{2}$, the plot of d versus κ is given in Fig. 9(a). The slope is 1 until full receiver cooperation performance is reached, implying that one-bit cooperation buys one over-the-air bit per user. We look at a particular point $\kappa = \frac{1}{4}$ and use its corresponding LDC [Fig. 9(b)] to provide insights. Note that 1 bit in the LDC corresponds to $\frac{1}{4}\log\mathsf{SNR}$ in the Gaussian channel and since $\mathsf{C}^{\mathsf{B}} \approx \frac{1}{4}\log\mathsf{SNR}$, in the corresponding LDC each receiver is able to sent one-bit information to the other. Without cooperation, the optimal way is to turn on bits not causing interference, that is, the *private* bits a_3, a_4, b_3, b_4. We cannot turn on more bits without cooperation since it causes collisions, for example, at the fourth level of receiver 2 if we turn on a_2 bit. Now with receiver cooperation, we want to support two more bits a_2, b_2. Note that prior to turning on a_2, b_2, there are "holes" left in receiver signal spaces and turning on each of these bits only causes one collision at one receiver. Therefore, we need 1 bit in each direction to resolve the collision at each receiver. We can achieve 3 bits per user in the corresponding LDC and $d = \frac{3}{4}$ in the Gaussian channel. We cannot turn on more bits in the LDC since it causes collisions while no cooperation capability is left.

At $\alpha = \frac{2}{3}$, the plot of d versus κ is given in Fig. 9(c). The slope is $\frac{1}{2}$ until full receiver cooperation performance is reached, implying that two-bit cooperation buys one over-the-air bit per user. We look at a particular point $\kappa = \frac{1}{3}$ and use its corresponding LDC (Fig. 9(d)) to provide insights. Note that now 1 bit in the LDC corresponds to $\frac{1}{3}\log\mathsf{SNR}$ in the Gaussian channel and since $\mathsf{C}^{\mathsf{B}} \approx \frac{1}{3}\log\mathsf{SNR}$, in the corresponding LDC each receiver is able to sent one-bit information to the other. Without cooperation, the optimal way is to turn on bits a_1, a_3, b_1, b_3. We cannot turn on more bits without cooperation since it causes collisions, for example, at the second level of receiver 2 if we turn on a_2 bit. Now with receiver cooperation,

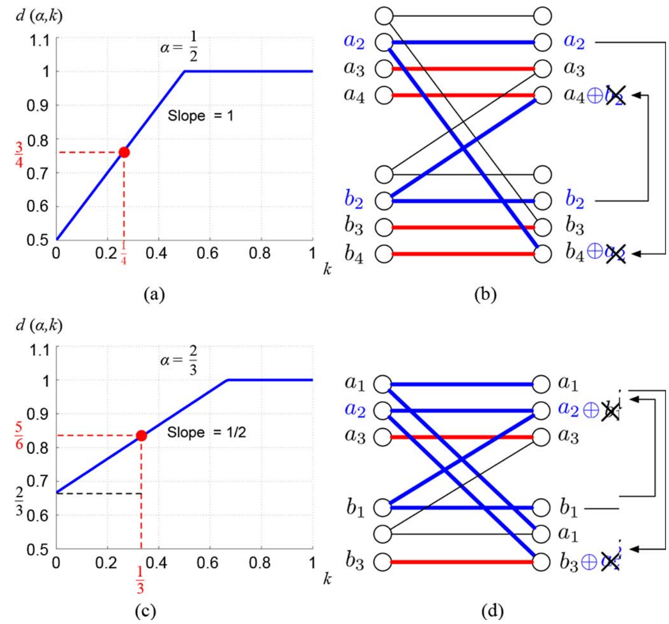


Fig. 9. Gain from cooperation.

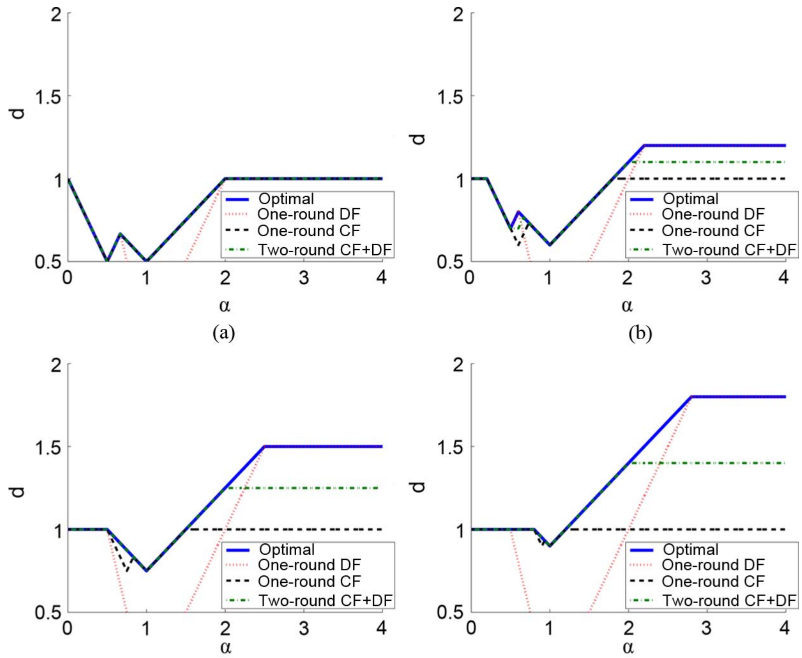
we want to support one over-the-air bit a_2. Note that prior to turning on a_2, there are no "holes" left in receiver signal spaces and turning on a_2 causes collisions at *both* receivers. Therefore, we need 2 bits in total to resolve collisions at both receivers. We can achieve 5 bits in total in the corresponding LDC and $d = \frac{5}{6}$ in the Gaussian channel. We cannot turn on more bits in the LDC since it causes collision while no cooperation capability is left.

From above examples and illustrations, we see that whether *one cooperation bit buys one more bit* or *two cooperation bits buy one more bit* depends on whether there are "holes" in receiver signal spaces before increasing data rates. The "holes" play a central role not only in why conventional compress-forward is suboptimal in certain regimes, as mentioned in Section III-B, but also in the fundamental behavior of the gain from receiver cooperation. We notice that in [14], there is a similar behavior about the gain from in-band receiver cooperation as discussed in Section III-B of [14]. We conjecture that the behavior can be explained via the concept of "holes" as well.

C. Comparison With Suboptimal Strategies

Pointed out by the motivating examples in Section III-B, conventional compress-forward and decode-forward are not good for receiver cooperation to mitigate interference in certain regimes, which are used in [11] and [12]. These suboptimal schemes include:

1) One-round compress-forward (CF) strategy: the conventional compress-forward is used for the two receivers to first exchange information and then decode.
2) One-round decode-forward (DF) strategy: at the first stage both receivers decode one of the common messages with stronger signal strength without help from the receiver-cooperative links, by treating other signals as noise. Both then bin-and-forward the decoded information to each other. At the second stage, both receivers



(a) (b) (c) (d)

Fig. 10. Number of generalized degrees of freedom. (a) $\kappa = 0$. (b) $\kappa = 0.2$. (c) $\kappa = 0.5$. (d) $\kappa = 0.8$.

make use of the bin index send over receiver-cooperative links to decode and enhance the rate.

3) Two-round CF+DF strategy: at the first stage one of the receivers, say, receiver 1, compresses its received signal and forwards it to the other receiver. At the second stage, receiver 2 decodes with the side information received at the first round and then bin-and-forwards the decoded information to receiver 1. Then at the third stage receiver 1 decodes with the help from receiver-cooperative links.

Comparisons of these strategies in terms of the number of generalized degrees of freedom for different scaling exponents α of $\log \mathsf{INR}$ and κ of C^{B} are depicted in Fig. 10. None of them achieves the optimal g.d.o.f. universally. Note that although the two-round CF+DF strategy outperforms one-round CF/DF strategies, it cannot achieve the optimal number of g.d.o.f. for all α's and κ's. By Theorem 6.2, the one-round strategy based on our cooperative protocol, on the other hand, is sufficient to achieve the symmetric capacity to within 3 bits universally and hence achieves the optimal number of g.d.o.f. for all α's and κ's.

Appendix A
Proof of Theorem 4.3

We will first describe the strategy in detail and analyze the error probability rigorously.

A) Description of the Strategy: In the following, consider all $i,j \in \{1,2\}$ and $i \neq j$.

Codebook Generation: Transmitter i splits its message $m_i \longrightarrow (m_{ic}, m_{ip})$. Consider block length-$N$ encoding. First we generate $2^{NR_{ic}}$ common codewords $\{x_{ic}^N(m_{ic}),\ 1 \leq m_{ic} \leq 2^{NR_{ic}}\}$, according to distribution $p\left(x_{ic}^N\right) = \prod_{n=1}^N p\big(x_{ic}[n]\big)$ with $x_{ic}[n] \sim \mathcal{CN}(0, Q_{ic})$ for all n. Then for each common codeword $x_{ic}^N(m_{ic})$ serving as a *cloud center*, we generate $2^{NR_{ip}}$ codewords $\{x_i^N(m_{ic}, m_{ip}),\ 1 \leq m_{ip} \leq 2^{NR_{ip}}\}$, according to conditional distribution $p\left(x_i^N | x_{ic}^N\right) = \prod_{n=1}^N p\big(x_i[n] | x_{ic}[n]\big)$ such that for all n, $x_i[n] = x_{ic}[n] + x_{ip}[n]$, where $x_{ip}[n] \sim \mathcal{CN}(0, Q_{ip})$ and independent of everything else. The power split configuration is such that $Q_{ip} + Q_{ic} = 1$, $\mathsf{INR}_{jp} := Q_{ip}|h_{ji}|^2 \leq 1$ if $\mathsf{SNR}_i > \mathsf{INR}_j$ and no such split if $\mathsf{SNR}_i \leq \mathsf{INR}_j$. Hence, $Q_{ip} = \min\left\{1, \frac{1}{\mathsf{INR}_j}\right\}$ if $\mathsf{SNR}_i > \mathsf{INR}_j$ and $Q_{ip} = 0$ otherwise.

For receiver 2 serving as relay, it generates a quantization codebook $\hat{\mathscr{Y}}_2$, of size $\left|\hat{\mathscr{Y}}_2\right| = 2^{N\hat{R}_2}$, randomly according to marginal distribution $p(\hat{y}_2^N)$, marginalized over joint distribution $p\left(y_2^N, x_{1c}^N, x_1^N, x_{2c}^N\right) p(\hat{y}_2^N | y_2^N, x_{1c}^N, x_1^N, x_{2c}^N)$, where

$$p(\hat{y}_2^N | y_2^N, x_{1c}^N, x_1^N, x_{2c}^N) = \prod_{n=1}^N p\Big(\hat{y}_2[n] \big| y_2[n], x_{1c}[n], x_1[n], x_{2c}[n]\Big).$$

The conditional distribution is such that for all n, $\hat{y}_2[n] = y_2[n] + \hat{z}_2[n]$, where $\hat{z}_2[n] \sim \mathcal{CN}(0, \Delta_2)$, independent of everything else. Parameters $\hat{R}_2$ and Δ_2 are to be specified later. For each element in codebook $\hat{\mathscr{Y}}_2$, map it into $\{1, \ldots, 2^{N\mathsf{C}_{21}^{\mathsf{B}}}\}$ through a uniformly generated random mapping $b_2 : \hat{\mathscr{Y}}_2 \longrightarrow \{1, \ldots, 2^{N\mathsf{C}_{21}^{\mathsf{B}}}\}, \hat{y}_2^N \mapsto l_{21}$ *(binning)*.

For receiver 1 serving as relay, it generates two binning functions $b_1^{(1c)}$ and $b_1^{(2c)}$ independently according to uniform distributions, such that the message set $\{1 \leq m_{ic} \leq 2^{NR_{ic}}\}$ is partitioned into $2^{\lambda_1^{(ic)} N\mathsf{C}_{12}^{\mathsf{B}}}$ bins, for $i = 1, 2$, where $0 \leq \lambda_1^{(ic)} \leq 1$, $\lambda_1^{(1c)} + \lambda_1^{(2c)} = 1$ and

$$b_1^{(ic)} : \{1, \ldots, 2^{NR_{ic}}\} \longrightarrow \{1, \ldots, 2^{\lambda_1^{(ic)} N\mathsf{C}_{12}^{\mathsf{B}}}\},$$
$$m_{ic} \mapsto l_{12}^{(ic)} \in \{1, \ldots, 2^{\lambda_1^{(ic)} N\mathsf{C}_{12}^{\mathsf{B}}}\}.$$

The superscript notation "(ic)" denotes which message set is partitioned into bins, while the subscript "1" denotes the binning procedure is at receiver 1.

Encoding: Transmitter i sends out signals according to its message and the codebook. Receiver 2, serving as relay, chooses the quantization codeword which is jointly typical with y_2^N (if there is more than one, it chooses the one with the smallest index) and then sends out the bin index l_{21} for the quantization codeword. After decoding (m_{1c}, m_{1p}, m_{2c}) (to be specified below), receiver 1 sends out bin indices $\left(l_{12}^{(1c)}, l_{12}^{(2c)}\right)$ according to binning functions $\left(b_1^{(1c)}, b_1^{(2c)}\right)$.

Decoding At Receiver 1: To draw comparison with the decoding procedure in the conventional compress-forward, the above decoding can be interpreted as a two-stage procedure as follows. It first constructs a *list* of message triples (both users' common messages and its own private message), each element of which indices a codeword triple that is jointly typical with its received signal from the transmitter-receiver link. Then, for each message triple in this list, it constructs an *ambiguity set* of quantization codewords, each element of which is jointly typical with the codeword triple and the received signal. Finally, it searches through all ambiguity sets and finds one that contains a quantization codeword with the same bin index it received. If there is no such unique ambiguity set, it declares an error. The two-stage interpretation is illustrated in Fig. 11.

To be specific, upon receiving signal y_1 and receiver-cooperative side information l_{21}, receiver i constructs a list of candidates $L_i(y_i^N)$, defined at the bottom of this page, where $A_\epsilon^{(N)}$ denotes the set of jointly ϵ-typical N-sequences [30].

For each element $\underline{m} \in L(y_1^N)$, construct an ambiguity set of quantization codewords $B(\underline{m})$, defined as shown in the equation at the bottom of the page. Declare the transmitted message is $\underline{\hat{m}}$ if there exists an unique $\underline{\hat{m}}$ such that $\exists\ \hat{y}_2^N \in B(\underline{\hat{m}})$ with $b_2(\hat{y}_2^N) = l_{21}$. Otherwise, declare an error.

Decoding at Receiver 2: After receiving bin indices $\left(l_{12}^{(1c)}, l_{12}^{(2c)}\right)$, receiver 2 searches for an unique message triple (m_{2c}, m_{2p}, m_{1c}) such that $\left(x_{2c}^N(m_{2c}), x_2^N(m_{2c}, m_{2p}), x_{1c}^N(m_{1c}), y_2^N\right) \in A_\epsilon^{(N)}$ and $b_1^{(ic)}(m_{ic}) = l_{12}^{(ic)}$, for $i = 1, 2$. If there is no such unique triple, it declares an error.

B) Analysis:

Error Probability Analysis at Receiver 1: Without loss of generality, assume that all transmitted messages are 1's. For simplicity, we first focus on the case where receiver 1 aims to decode while receiver 2 serves as a relay to help it decode.

At receiver 1, due to law of large numbers, the probability that the truly transmitted $\underline{1} := (m_{1c} = 1, m_{2c} = 1, m_{1p} = 1) \notin L(y_1^N)$ goes to zero as $N \to \infty$. Besides, the probability that $B(\underline{1})$ does not contain the truly selected $\hat{y}_2^N$ is also negligible when N is sufficiently large. Consider the following error events:

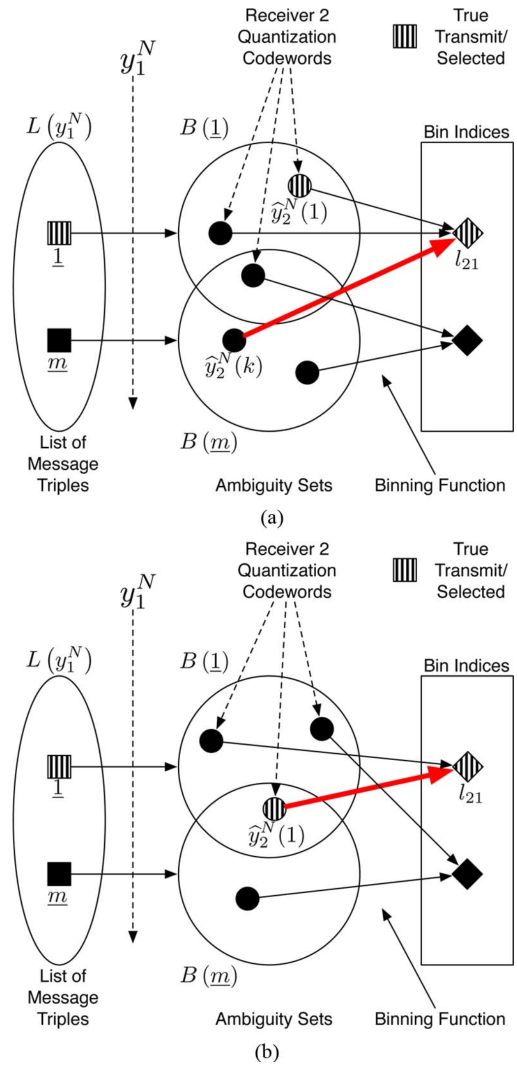


Fig. 11. Decoding at receiver 1 and error events. (a) Error event (1). (b) Error event (2).

First, there is no quantization codeword jointly typical with received signals. This probability goes to zero as $N \to \infty$ if $\hat{R}_2 \geq I(\hat{y}_2; y_2)$, which is a known result in the source coding literature.

Second, there exists $\underline{m} \neq \underline{1}$ such that both of them are in the candidate list $L(y_1^N)$ and the ambiguity set $B(\underline{m})$ contains some quantization codeword $\hat{y}_2^N$ with bin index $b_2(\hat{y}_2^N) = l_{21}$.

$$L(y_1^N) := \left\{ \underline{m} := (m_{1c}, m_{1p}, m_{2c}) \middle| \left(x_{1c}^N(m_{1c}), x_1^N(m_{1c}, m_{1p}), x_{2c}^N(m_{2c}), y_1^N\right) \in A_\epsilon^{(N)} \right\},$$
$$B(\underline{m}) := \left\{ \hat{y}_2^N \in \hat{\mathscr{Y}}_2 \middle| \left(\hat{y}_2^N, x_{1c}^N(m_{1c}), x_1^N(m_{1c}, m_{1p}), x_{2c}^N(m_{2c}), y_i^N\right) \in A_\epsilon^{(N)} \right\}.$$

This event can further be distinguished into two cases: First, this $\hat{y}_2^N \in B(\underline{m})$ is not the actual selected quantization codeword [illustrated in Fig. 11(a)]; second, this $\hat{y}_2^N \in B(\underline{m})$ is indeed the selected quantization codeword [illustrated in Fig. 11(b)]. In the following we analyze the error probability of these two typical error events.

Again, refer to Fig. 11. for illustration. Define error events as follows: for any nonempty $S \subseteq \{1c, 1p, 2c\}$,

$E_S^{(1)}$:= the event that there exists some $\underline{m} \neq \underline{1}$, (where $m_s \neq 1, \forall s \in S$ and $m_s = 1, \forall s \notin S$), such that $\underline{m} \in L(y_1^N)$ and $B(\underline{m})$ contains some $\hat{y}_2^N(k)$, $k \in \{1, 2, \ldots, 2^{N\hat{R}_2}\}$ with $b_2(\hat{y}_2^N(k)) = l_{21}$. Note: this $\hat{y}_2^N(k)$ is not the truly selected quantization codeword $\hat{y}_2^N(1)$.

$E_S^{(2)}$:= the event that there exists some $\underline{m} \neq \underline{1}$, (where $m_s \neq 1, \forall s \in S$ and $m_s = 1, \forall s \notin S$), such that $\underline{m} \in L(y_1^N)$ and $B(\underline{m})$ contains $\hat{y}_2^N(1)$.

Probability of $E_S^{(1)}$: Consider the probability of the error event $E_S^{(1)}$: it can be upper bounded by (27) at the bottom of the page, as where (a) is due to the independent uniform binning.

For notational convenience we use $\underline{x}^N(\underline{m})$ to denote the vector of codewords corresponding to message $\underline{m}$, that is, $\left(x_{1c}^N(m_{1c}), x_1^N(m_{1c}, m_{1p}), x_{2c}^N(m_{2c})\right)$.

Note that for $k \neq 1$, $\hat{y}_2^N(k)$ is independent of $\left(\underline{x}^N(\underline{m}), y_1^N\right)$. We then make use of [30, Th. 15.2.2], which upper bounds the volume of conditional joint ϵ-typical set $A_\epsilon^{(N)}\left(\hat{y}_2 \middle| \underline{x}^N, y_1^N\right)$ given that $\left(\underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}$, to further upper bound $\sum_{k\neq 1} \Pr\left\{\left(\hat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\}$ by (28) at the bottom of the page, where (b) is due to in [30, Th. 15.2.2]. Besides, according to the results in [31],

$$\Pr\left\{\underline{m} \in L(y_1^N)\right\} \le \begin{cases} 2^{-N\left(I(x_1;y_1|x_{1c},x_{2c})-\epsilon'\right)} & S=\{1p\} \\ 2^{-N\left(I(x_1;y_1|x_{2c})-\epsilon'\right)} & S=\{1c\} \\ 2^{-N\left(I(x_{2c};y_1|x_1)-\epsilon'\right)} & S=\{2c\} \\ 2^{-N\left(I(x_{2c},x_1;y_1|x_{1c})-\epsilon'\right)} & S=\{1p,2c\} \\ 2^{-N\left(I(x_1;y_1|x_{2c})-\epsilon'\right)} & S=\{1p,1c\} \\ 2^{-N\left(I(x_1,x_{2c};y_1)-\epsilon'\right)} & S=\{2c,1c\} \\ 2^{-N\left(I(x_1,x_{2c};y_1)-\epsilon'\right)} & S=\{1p,2c,1c\} \end{cases}$$

where $\epsilon' = 4\epsilon$. Note that unlike in the interference channel without cooperation as in [31], here we require receiver 1 to

$$\begin{aligned} \Pr\left\{E_S^{(1)}\right\} &\le \sum_{\substack{\underline{m}:m_s\neq 1,\\ \forall s\in S}} \sum_{k\neq 1} \Pr\left\{\underline{m} \in L(y_1^N), \hat{y}_2^N(k) \in B(\underline{m}), b_2(\hat{y}_2^N(k)) = l_{21}\right\} \\ &= \sum_{\substack{\underline{m}:m_s\neq 1,\\ \forall s\in S}} \sum_{k\neq 1} \Pr\left\{\left(\hat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}, b_2(\hat{y}_2^N(k)) = l_{21}\right\} \\ &\overset{(a)}{=} 2^{-N\mathsf{C}_{21}^{\mathsf{B}}} \sum_{\substack{\underline{m}:m_s\neq 1,\\ \forall s\in S}} \sum_{k\neq 1} \Pr\left\{\left(\hat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\} \\ &\le 2^{N\left(\sum_{s\in S} R_s\right)} 2^{-N\mathsf{C}_{21}^{\mathsf{B}}} \sum_{k\neq 1} \Pr\left\{\left(\hat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\}; \end{aligned} \tag{27}$$

$$\begin{aligned} &\sum_{k\neq 1} \Pr\left\{\left(\hat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\} \\ &\le 2^{N\hat{R}_2} \int_{\left(\hat{y}_2^N, \underline{x}^N, y_1^N\right)\in A_\epsilon^{(N)}} p\left(\hat{y}_2^N\right) p\left(\underline{x}^N, y_1^N\right) d\hat{y}_2^N d\underline{x}^N dy_1^N \\ &\le 2^{N\hat{R}_2} \int_{\left(\underline{x}^N, y_1^N\right)\in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \int_{\hat{y}_2^N \in A_\epsilon^{(N)}\left(\hat{y}_2 \middle| \underline{x}^N, y_1^N\right)} p\left(\hat{y}_2^N\right) dy_2^N \\ &\le 2^{N\hat{R}_2} \int_{\left(\underline{x}^N, y_1^N\right)\in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \int_{\hat{y}_2^N \in A_\epsilon^{(N)}\left(\hat{y}_2 \middle| \underline{x}^N, y_1^N\right)} 2^{-N(h(\hat{y}_2)-\epsilon)} dy_2^N \\ &\overset{(b)}{\le} 2^{N(h(\hat{y}_2|x_{1c},x_1,x_{2c},y_1)+2\epsilon)} \cdot 2^{-N(h(\hat{y}_2)-\epsilon)} \cdot 2^{N\hat{R}_2} \int_{\left(\underline{x}^N, y_1^N\right)\in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \\ &= 2^{N\hat{R}_2} 2^{-N(I(\hat{y}_2;x_{1c},x_1,x_{2c},y_1)-3\epsilon)} \int_{\left(\underline{x}^N, y_1^N\right)\in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \\ &= \Pr\left\{\underline{m} \in L(y_1^N)\right\} \cdot 2^{N\hat{R}_2} 2^{-N(I(\hat{y}_2;x_{1c},x_1,x_{2c},y_1)-3\epsilon)} \end{aligned} \tag{28}$$

decode m_{2c} correctly. Hence, the event when $S=\{2c\}$ does cause an error. Therefore, the probability of the first kind of error event vanishes as $N\longrightarrow\infty$ if

$$\begin{aligned}R_{1p} &\le I(x_1;y_1|x_{1c},x_{2c})+\varphi\\ R_{2c} &\le I(x_{2c};y_1|x_1)+\varphi\\ R_{2c}+R_{1p} &\le I(x_{2c},x_1;y_1|x_{1c})+\varphi\\ R_{1c}+R_{1p} &\le I(x_1;y_1|x_{2c})+\varphi\\ R_{1c}+R_{2c}+R_{1p} &\le I(x_1,x_{2c};y_1)+\varphi,\end{aligned}$$

where $\varphi=\mathsf{C}_{21}^{\mathsf{B}}-\hat{R}_2+I(\hat{y}_2;x_{1c},x_1,x_{2c},y_1)$.

On the other hand, since we can alternatively upper bound $\Pr\left\{E_S^{(1)}\right\}$ by (29) at the bottom of the page, the probability of the first kind of error event vanishes as $N\longrightarrow\infty$ if

$$\begin{aligned}R_{1p} &\le I(x_1;y_1|x_{1c},x_{2c})+\varphi^+\\ R_{2c} &\le I(x_{2c};y_1|x_1)+\varphi^+\\ R_{2c}+R_{1p} &\le I(x_{2c},x_1;y_1|x_{1c})+\varphi^+\\ R_{1c}+R_{1p} &\le I(x_1;y_1|x_{2c})+\varphi^+\\ R_{1c}+R_{2c}+R_{1p} &\le I(x_1,x_{2c};y_1)+\varphi^+.\end{aligned}$$

Finally, plug in $\hat{R}_2=I(\hat{y}_2;y_2)$ and by Markov relation: $(x_{1c},x_1,x_{2c},y_1)-y_2-\hat{y}_2$, we get the rate loss term

$$\begin{aligned}\xi_1 &:= \hat{R}_2-I(\hat{y}_2;x_{1c},x_1,x_{2c},y_1)\\ &= I(\hat{y}_2;y_2)-I(\hat{y}_2;x_{1c},x_1,x_{2c},y_1)\\ &= I(\hat{y}_2;y_2|x_{1c},x_1,x_{2c},y_1).\end{aligned}$$

Probability of $E_S^{(2)}$*:* Consider the probability of the error event $E_S^{(2)}$:

$$\begin{aligned}&\Pr\left\{E_S^{(2)}\right\}\\ &\quad\le \sum_{\underline{m}:m_s\neq 1,\forall s\in S}\Pr\left\{\hat{y}_2^N(1)\in B(\underline{m}),\underline{m}\in L(y_1^N)\right\}\\ &\quad= \sum_{\underline{m}:m_s\neq 1,\forall s\in S}\Pr\left\{\left(\hat{y}_2^N(1),\underline{x}^N(\underline{m}),y_1^N\right)\in A_\epsilon^{(N)}\right\}\\ &\quad\le \begin{cases} 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_1;y_1,\hat{y}_2|x_{1c},x_{2c})-\epsilon'\right)}, & S=\{1p\}\\ 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_1;y_1,\hat{y}_2|x_{2c})-\epsilon'\right)}, & S=\{1c\}\\ 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_{2c};y_1,\hat{y}_2|x_1)-\epsilon'\right)}, & S=\{2c\}\\ 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_{2c},x_1;y_1,\hat{y}_2|x_{1c})-\epsilon'\right)}, & S=\{1p,2c\}\\ 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_1;y_1,\hat{y}_2|x_{2c})-\epsilon'\right)}, & S=\{1p,1c\}\\ 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_1,x_{2c};y_1,\hat{y}_2)-\epsilon'\right)}, & S=\{2c,1c\}\\ 2^{N\left(\sum_{s\in S}R_s\right)}\cdot 2^{-N\left(I(x_1,x_{2c};y_1,\hat{y}_2)-\epsilon'\right)}, & S=\{1p,2c,1c\}\end{cases}\end{aligned}$$

where $\epsilon'=4\epsilon$. Note that the event when $S=\{2c\}$ does cause an error. Hence, the probability of the second kind of error event vanishes as $N\longrightarrow\infty$ if

$$\begin{aligned}R_{1p} &\le I(x_1;y_1,\hat{y}_2|x_{1c},x_{2c})\\ R_{2c} &\le I(x_{2c};y_1,\hat{y}_2|x_1)\\ R_{2c}+R_{1p} &\le I(x_{2c},x_1;y_1,\hat{y}_2|x_{1c})\\ R_{1c}+R_{1p} &\le I(x_1;y_1,\hat{y}_2|x_{2c})\\ R_{1c}+R_{2c}+R_{1p} &\le I(x_1,x_{2c};y_1,\hat{y}_2).\end{aligned}$$

Error Probability Analysis at Receiver 2: After receiving the two bin indices, receiver 2 can decode (m_{1c},m_{2c},m_{2p}), with effectively smaller candidate message sets, (namely, the bins,) for m_{1c} and m_{2c}. Following the same line as [31], it can be shown that (we omit the detailed analysis here), for all $0\le\lambda_1^{(ic)}\le 1$ and $\lambda_1^{(1c)}+\lambda_1^{(2c)}=1$, the following region is achievable:

$$\begin{aligned}R_{2p} &\le I(x_2;y_2|x_{2c},x_{1c})\\ R_{1c}+R_{2p} &\le I(x_{1c},x_2;y_2|x_{2c})+\lambda_1^{(1c)}\mathsf{C}_{12}^{\mathsf{B}}\\ R_{2c}+R_{2p} &\le I(x_2;y_2|x_{1c})+\lambda_1^{(2c)}\mathsf{C}_{12}^{\mathsf{B}}\\ R_{2c}+R_{1c}+R_{2p} &\le I(x_2,x_{1c};y_2)+\mathsf{C}_{12}^{\mathsf{B}}.\end{aligned}$$

Note that the performance of decoding the private message m_{2p} does not gain from cooperation, since receiver 1 does not decode the private message m_{2p}.

Taking convex hull over all possible $\lambda_1^{(1c)}\in[0,1]$. Note that the bounds for R_{2p} and $R_{2c}+R_{1c}+R_{2p}$ remain unchanged. Project the three-dimensional rate region to a two-dimensional

$$\begin{aligned}\Pr\left\{E_S^{(1)}\right\} &\le \sum_{\substack{m:m_s\neq 1\\ \forall s\in S}}\Pr\left\{\underline{m}\in L(y_1^N)\right\}\cdot\Pr\left\{\exists\, k\neq 1,\hat{y}_2^N(k)\in B(\underline{m}),b_2(\hat{y}_2^N(k))=l_{21}\,\Big|\,\underline{m}\in L(y_1^N)\right\}\\ &\le 2^{N\left(\sum_{s\in S}R_s\right)}\Pr\left\{\underline{m}\in L(y_1^N)\right\},\end{aligned}\tag{29}$$

space for any fixed $R_{2p} = r_{2p}$, we see that the convexifying procedure results in the following region:

$$\begin{aligned} R_{1c} + r_{2p} &\leq I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \\ R_{2c} + r_{2p} &\leq I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \\ R_{2c} + R_{1c} + r_{2p} &\leq I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}^{\mathsf{B}}_{12}. \end{aligned}$$

Hence the following rate region is achievable for receiver 2 to decode successfully:

$$\begin{aligned} R_{2p} &\leq I\left(x_2; y_2 | x_{2c}, x_{1c}\right) \\ R_{1c} + R_{2p} &\leq I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \\ R_{2c} + R_{2p} &\leq I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \\ R_{2c} + R_{1c} + R_{2p} &\leq I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}^{\mathsf{B}}_{12}. \end{aligned}$$

APPENDIX B
PROOF OF LEMMA 5.1

Bounds (4) on R_1 and (5) on R_2:

Proof: One can directly use cut-set bounds. As an alternative, we give the following proof in which the decomposition of mutual informations is made clear.

We have the following bounds by Fano's inequality, data-processing inequality, and chain rule: if R_1 is achievable,

$$\begin{aligned} N(R_1 - \epsilon_N) &\overset{\text{(a)}}{\leq} I(x_1^N; y_1^N, u_{21}^N) \\ &\overset{\text{(b)}}{\leq} I(x_1^N; y_1^N, u_{21}^N, x_2^N) \\ &\overset{\text{(c)}}{=} I(x_1^N; y_1^N, u_{21}^N | x_2^N) \\ &\overset{\text{(d)}}{=} I(x_1^N; y_1^N | x_2^N) + I(x_1^N; u_{21}^N | y_1^N, x_2^N) \\ &= h(h_{11} x_1^N + z_1^N) - h(z_1^N) \\ &\quad + I(x_1^N; u_{21}^N | y_1^N, x_2^N) \\ &\overset{\text{(e)}}{\leq} N \log(1 + \mathsf{SNR}_1) + I(x_1^N; u_{21}^N | y_1^N, x_2^N), \end{aligned}$$

where $\epsilon_N \longrightarrow 0$ as $N \longrightarrow \infty$. (a) is due to Fano's inequality and data processing inequality. (b) is due to the genie giving side information x_1^N to receiver 1, ie., *conditioning reduces entropy*. (c) is due to the fact that x_1^N and x_2^N are independent. (d) is due to chain rule. (e) is due to the fact that i.i.d. Gaussian distribution maximizes differential entropy under covariance constraints.

To upper bound $I(x_1^N; u_{21}^N | y_1^N, x_2^N)$, which corresponds to the enhancement from cooperation, we make use of the fact that u_{21}^N is a function of (y_1^N, y_2^N)

$$\begin{aligned} &I(x_1^N; u_{21}^N | y_1^N, x_2^N) \\ &\quad = h(x_1^N | y_1^N, x_2^N) - h(x_1^N | u_{21}^N, y_1^N, x_2^N) \\ &\quad \overset{\text{(a)}}{\leq} h(x_1^N | y_1^N, x_2^N) - h(x_1^N | u_{21}^N, y_1^N, x_2^N, y_2^N) \\ &\quad \overset{\text{(b)}}{=} h(x_1^N | y_1^N, x_2^N) - h(x_1^N | y_1^N, x_2^N, y_2^N) \\ &\quad = I(x_1^N; y_2^N | y_1^N, x_2^N) \\ &\quad = h(y_2^N | y_1^N, x_2^N) - h(y_2^N | y_1^N, x_2^N, x_1^N) \\ &\quad = h(h_{21} x_1^N + z_2^N | h_{11} x_1^N + z_1^N) - h(z_2^N) \\ &\quad \leq N \log\left(1 + \frac{\mathsf{INR}_2}{1 + \mathsf{SNR}_1}\right). \end{aligned}$$

(a) is due to the fact that conditioning reduces entropy. (b) is due to the fact that u_{21}^N is a function of (y_1^N, y_2^N).

Besides, it is trivial to see that $I(x_1^N; u_{21}^N | y_1^N, x_2^N) \leq H(u_{21}^N) \leq N\mathsf{C}^{\mathsf{B}}_{21}$. Hence, (and similarly for R_2), we have shown bounds (4) and (5). ■

Bound (6) on $R_1 + R_2$:

Proof: Define

$$\begin{aligned} s_1 &:= h_{21} x_1 + z_2,\ s_2 := h_{12} x_2 + z_1, \\ \tilde{s}_1 &:= h_{21} x_1 + \tilde{z}_2,\ \tilde{s}_2 := h_{12} x_2 + \tilde{z}_1, \end{aligned}$$

where $\tilde{z}_1, \tilde{z}_2$ are i.i.d. $\mathcal{CN}(0,1)$'s, independent of everything else. Note that s_i and $\tilde{s}_i$ have the same marginal distribution, for $i = 1, 2$.

A genie gives side information $\tilde{s}_i^N$ to receiver i (refer to Fig. 12.) Making use of Fano's inequality, data processing inequality, and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$\begin{aligned} &N(R_1 + R_2 - \epsilon_N) \\ &\quad \overset{\text{(a)}}{\leq} I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) \\ &\quad \overset{\text{(b)}}{=} I(x_1^N; y_1^N) + I(x_2^N; y_2^N) + I(x_1^N; u_{21}^N | y_1^N) \\ &\qquad + I(x_2^N; u_{12}^N | y_2^N) \\ &\quad \overset{\text{(c)}}{\leq} I(x_1^N; y_1^N, \tilde{s}_1^N) + I(x_2^N; y_2^N, \tilde{s}_2^N) + H(u_{21}^N) + H(u_{12}^N) \\ &\quad \overset{\text{(d)}}{\leq} h(y_1^N, \tilde{s}_1^N) - h(s_2^N, \tilde{z}_2^N) + h(y_2^N, \tilde{s}_2^N) - h(s_1^N, \tilde{z}_1^N) \\ &\qquad + N\mathsf{C}^{\mathsf{B}}_{21} + N\mathsf{C}^{\mathsf{B}}_{12} \\ &\quad \overset{\text{(e)}}{=} h(y_1^N | \tilde{s}_1^N) + h(\tilde{s}_1^N) - h(s_2^N) - h(\tilde{z}_2^N) + h(y_2^N | \tilde{s}_2^N) \\ &\qquad + h(\tilde{s}_1^N) - h(s_1^N) - h(\tilde{z}_1^N) + N\mathsf{C}^{\mathsf{B}}_{21} + N\mathsf{C}^{\mathsf{B}}_{12} \\ &\quad = h(y_1^N | \tilde{s}_1^N) - h(\tilde{z}_2^N) + h(y_2^N | \tilde{s}_2^N) - h(\tilde{z}_1^N) \\ &\qquad + N(\mathsf{C}^{\mathsf{B}}_{21} + \mathsf{C}^{\mathsf{B}}_{12}) \\ &\quad \overset{\text{(f)}}{\leq} N\{\text{RHS of (6)}\}, \end{aligned}$$

where $\epsilon_N \longrightarrow 0$ as $N \longrightarrow \infty$. (a) follows from Fano's inequality and data processing inequality. (b) is due to chain rule. (c) is due to the genie giving side information $\tilde{s}_i^N$ to receiver i, $i = 1, 2$ and $I(x_i^N; u_{ji}^N | y_i^N) \leq H(u_{ji}^N)$. (d) is due to the fact that $H(u_{ji}^N) \leq N\mathsf{C}^{\mathsf{B}}_{ji}$. (e) is due to chain rule. (f) is due to the fact that i.i.d. Gaussian distribution maximizes conditional entropy subject to conditional variance constraints. Note that alternatively the genie can give side informations s_i^N to receiver i, as in [3].

Hence, we have shown bound (6). ■

Bounds (7) and (8) on $R_1 + R_2$:

Proof: A genie gives side information x_2^N and y_2^N to receiver 1 (refer to Fig. 13.) Making use of Fano's inequality, data processing inequality, the fact that u_{21}^N is a function of (y_1^N, y_2^N) and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$\begin{aligned} &N(R_1 + R_2 - \epsilon_N) \\ &\quad = I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) \end{aligned}$$

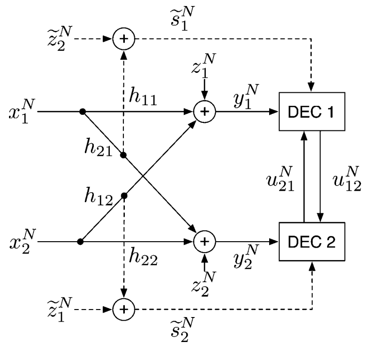


Fig. 12. Side information structure for bound (6).

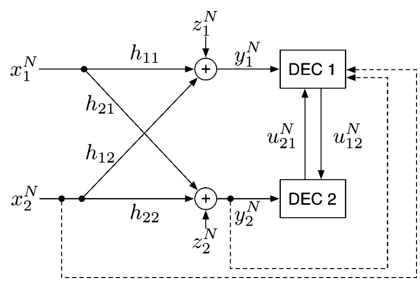


Fig. 13. Side information structure for bound (7).

$$
\begin{aligned}
&\overset{(a)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N, x_2^N) + I(x_2^N; y_2^N) + I(x_2^N; u_{12}^N | y_2^N) \\
&\overset{(b)}{\leq} I\left(x_1^N; y_1^N, u_{21}^N, y_2^N | x_2^N\right) + h\left(y_2^N\right) - h\left(s_1^N\right) + H\left(u_{12}^N\right) \\
&\overset{(c)}{=} I\left(x_1^N; y_1^N, y_2^N | x_2^N\right) + h\left(y_2^N\right) - h\left(s_1^N\right) + H\left(u_{12}^N\right) \\
&= h\left(h_{11}x_1^N + z_1^N, s_1^N\right) - h\left(z_1^N, z_2^N\right) + h\left(y_2^N\right) - h\left(s_1^N\right) \\
&\quad + H\left(u_{12}^N\right) \\
&= h\left(h_{11}x_1^N + z_1^N | s_1^N\right) - h\left(z_1^N, z_2^N\right) + h\left(y_2^N\right) + H\left(u_{12}^N\right) \\
&\leq N\{\text{RHS of (7)}\},
\end{aligned}
$$

where $\epsilon_N \longrightarrow 0$ as $N \longrightarrow \infty$. (a) is due to chain rule and the genie giving side information x_2^N and y_2^N to receiver 1. (b) is due to the fact that x_1^N and x_2^N are independent and $I(x_2^N; u_{12}^N | y_2^N) \leq H(u_{12}^N)$. (c) is due to the fact that u_{21}^N is a function of (y_1^N, y_2^N).

Hence, (and similarly if we gives side information x_1^N to receiver 2), we have shown bounds (7) and (8). ■

Bound (9) on $R_1 + R_2$*:*

Proof: This is straightforward cut-set upper bound: if (R_1, R_2) is achievable,

$$
\begin{aligned}
N(R_1 + R_2 - \epsilon_N) &\leq I\left(x_1^N, x_2^N; y_1^N, y_2^N\right) \\
&= h\left(y_1^N, y_2^N\right) - h\left(z_1^N, z_2^N\right) \\
&\leq N\{\text{RHS of (9)}\},
\end{aligned}
$$

where $\epsilon_N \longrightarrow 0$ as $N \longrightarrow \infty$.

Hence, we have shown bound (9). ■

Bounds (10) on $2R_1 + R_2$ *and (11) on* $R_1 + 2R_2$*:*

Proof: A genie gives side information x_2^N and y_2^N to one of the two receiver 1's and side information s_2^N to receiver 2 (refer to Fig. 14). Making use of Fano's inequality, data processing inequality, the fact that u_{21}^N is a function of (y_1^N, y_2^N), and the fact that Gaussian distribution maximizes conditional entropy

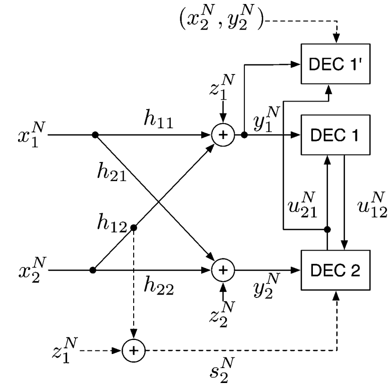


Fig. 14. Side information structure for bound (10).

subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$
\begin{aligned}
&N(2R_1 + R_2 - \epsilon_N) \\
&\overset{(a)}{\leq} I(x_1^N; y_1^N, u_{21}^N) + I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) \\
&\overset{(b)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N, x_2^N) + I(x_1^N; y_1^N) \\
&\quad + I(x_2^N; y_2^N, s_2^N) + I(x_1^N; u_{21}^N | y_1^N) + I(x_2^N; u_{12}^N | y_2^N) \\
&\overset{(c)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N | x_2^N) + I(x_1^N; y_1^N) + I(x_2^N; y_2^N, s_2^N) \\
&\quad + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&\overset{(d)}{=} I(x_1^N; y_1^N, y_2^N | x_2^N) + I(x_1^N; y_1^N) + I(x_2^N; y_2^N, s_2^N) \\
&\quad + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&\quad = h\left(h_{11}x_1^N + z_1^N, s_1^N\right) - h\left(z_1^N, z_2^N\right) + h(y_1^N) - h(s_2^N) \\
&\qquad + h(y_2^N, s_2^N) - h(s_1^N, z_1^N) + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&\quad = h\left(h_{11}x_1^N + z_1^N | s_1^N\right) - h\left(z_1^N, z_2^N\right) + h(y_1^N) \\
&\qquad + h(y_2^N | s_2^N) - h(z_1^N) + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&\quad \leq N\{\text{RHS of (10)}\},
\end{aligned}
$$

where $\epsilon_N \longrightarrow 0$ as $N \longrightarrow \infty$. (a) follows from Fano's inequality and data processing inequality. (b) is due to chain rule and the genie giving side information x_2^N and y_2^N to one of the receiver 1's and side information s_2^N to receiver 2. (c) is due to the fact that x_1^N, x_2^N are independent and $I(x_i^N; u_{ji}^N | y_i^N) \leq H(u_{ji}^N)$. (d) is due to the fact that u_{21}^N is a function of (y_1^N, y_2^N). Hence, (and similarly for $R_1 + 2R_2$), we have shown bounds (10) and (11). ■

Bounds (12) on $2R_1 + R_2$ *and (13) on* $R_1 + 2R_2$*:*

Proof: A genie gives side information $\tilde{s}_1^N, y_2^N$ to receiver 1 and side information y_1^N to one of the receiver 2's (refer to Fig. 15). Making use of Fano's inequality, data processing inequality, the fact that u_{12}^N, u_{21}^N are functions of (y_1^N, y_2^N), and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$
\begin{aligned}
&N(R_1 + 2R_2 - \epsilon_N) \\
&\quad \leq I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) + I(x_2^N; y_2^N, u_{12}^N) \\
&\overset{(a)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N, \tilde{s}_1^N) + I(x_2^N; y_2^N, u_{12}^N, y_1^N) \\
&\quad + I(x_2^N; y_2^N) + I(x_2^N; u_{12}^N | y_2^N)
\end{aligned}
$$

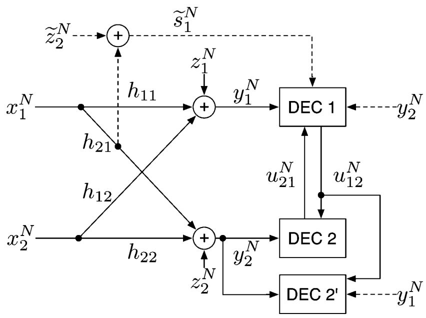


Fig. 15. Side information structure for bound (13).

$$
\begin{aligned}
&\overset{(b)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N | \tilde{s}_1^N) + I(x_1^n; \tilde{s}_1^N) \\
&+ I(x_2^N; y_2^N, u_{12}^N, y_1^N) + I(x_2^N; y_2^N) + H(u_{12}^N) \\
&\overset{(c)}{\leq} I(x_1^N; y_1^N, y_2^N | \tilde{s}_1^N) + I(x_2^N; y_1^N, y_2^N) + h(\tilde{s}_1^N) - h(z_2^N) \\
&+ h(y_2^N) - h(s_1^N) + N\mathsf{C}_{12}^{\mathsf{B}} \\
&\overset{(d)}{\leq} I(x_1^N; y_1^N, y_2^N | \tilde{s}_1^N) + I(x_2^N; y_1^N, y_2^N | x_1^N, \tilde{s}_1^N) \\
&+ h(y_2^N) - h(z_2^N) + N\mathsf{C}_{12}^{\mathsf{B}} \\
&= I(x_1^N, x_2^N; y_1^N, y_2^N | \tilde{s}_1^N) + h(y_2^N) - h(z_2^N) + N\mathsf{C}_{12}^{\mathsf{B}} \\
&= h(y_1^N, y_2^N | \tilde{s}_1^N) + h(y_2^N) - h(z_1^N, z_2^N) - h(z_2^N) + N\mathsf{C}_{12}^{\mathsf{B}} \\
&\overset{(e)}{\leq} N\{\text{RHS of (13)}\},
\end{aligned}
$$

where $\epsilon_N \longrightarrow 0$ as $N \longrightarrow \infty$. (a) is due to the genie giving side information $\tilde{s}_1^N$, y_2^N to receiver 1 and side information y_1^N to one of the receiver 2's. (b) is due to chain rule and the fact that $I(x_2^N; u_{12}^N | y_2^N) \leq H(u_{12}^N)$. (c) is due to the fact that u_{21}^N and u_{12}^N are both functions of (y_1^N, y_2^N) and that $H(u_{12}^N) \leq N\mathsf{C}_{12}^{\mathsf{B}}$. (d) is due to the fact that conditioning reduces entropy and that x_2^N and $(x_1^N, \tilde{s}_1^N)$ are independent. (e) is due to the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints.

Hence, (and similarly for $2R_1 + R_2$), we have shown bounds (13) and (12). ∎

Appendix C
Proof of Claim 5.6, Claim 5.7, Claim 5.9, and Claim 5.12

A) Proof of Claim 5.6:

Proof: To show (a), since we have four possible $R_1 + 2R_2$ bounds, we distinguish into 4 cases:

1) If the bound

$$
\begin{aligned}
R_1 + 2R_2 \leq\ & I(x_1, x_{2c}; y_1 | x_{1c}) + I(x_{1c}, x_2; y_2) \\
& + I(x_2; y_2 | x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+
\end{aligned}
$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (21) intersect, satisfies

$$
\begin{aligned}
3R_1^* + 3R_2^* &= \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1 | x_{1c}, x_{2c}) \\ + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&+ \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1 | x_{1c}) + I(x_{1c}, x_2; y_2) \\ + I(x_2; y_2 | x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \\
&= \{ I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \} \\
&+ \left\{ I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}_{12}^{\mathsf{B}} \right\} \\
&+ \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1 | x_{1c}) + I(x_{1c}, x_2; y_2 | x_{2c}) \\ + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \\
&= (16) + (19) + (17),
\end{aligned}
$$

which is greater than three times the active sum rate bound.

2) If the bound

$$
\begin{aligned}
R_1 + 2R_2 \leq\ & I(x_1, x_{2c}; y_1 | x_{1c}) + I(x_{2c}; y_1 | x_1) \\
& + I(x_{1c}, x_2; y_2 | x_{2c}) + I(x_2; y_2 | x_{1c}, x_{2c}) \\
& + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+
\end{aligned}
$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (21) intersect, satisfies

$$
\begin{aligned}
3R_1^* + 3R_2^* &= \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1 | x_{1c}, x_{2c}) \\ + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&+ \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1 | x_{1c}) + I(x_{2c}; y_1 | x_1) \\ + I(x_{1c}, x_2; y_2 | x_{2c}) + I(x_2; y_2 | x_{1c}, x_{2c}) \\ + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \\
&= \{ I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \} \\
&+ \left\{ \begin{array}{l} I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{2c}; y_1 | x_1) \\ + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&+ \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1 | x_{1c}) + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \\
&= (16) + (20) + (17),
\end{aligned}
$$

which is greater than three times the active sum rate bound.

3) If the bound

$$
\begin{aligned}
R_1 + 2R_2 \leq\ & I(x_1, x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_{1c}, x_2; y_2) \\
& + I(x_2; y_2 | x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}}
\end{aligned}
$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (21) intersect, satisfies

$$
\begin{aligned}
3R_1^* + 3R_2^* &= \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1 | x_{1c}, x_{2c}) \\ + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&+ \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_{1c}, x_2; y_2) \\ + I(x_2; y_2 | x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&= \{ I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \} \\
&+ \left\{ I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}_{12}^{\mathsf{B}} \right\} \\
&+ \left\{ I(x_1, x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \right\} \\
&= (16) + (19) + (18),
\end{aligned}
$$

which is greater than three times the active sum rate bound.

4) If the bound

$$
\begin{aligned}
R_1 + 2R_2 \leq\ & I(x_1, x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_{2c}; y_1 | x_1) \\
& + I(x_{1c}, x_2; y_2 | x_{2c}) + I(x_2; y_2 | x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}}
\end{aligned}
$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (21) intersect, satisfies

$$\begin{aligned}
&3R_1^* + 3R_2^* \\
&= \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1 | x_{1c}, x_{2c}) \\ +I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&\quad + \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_{2c}; y_1 | x_1) \\ +I(x_{1c}, x_2; y_2 | x_{2c}) + I(x_2; y_2 | x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&= \{ I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \} \\
&\quad + \left\{ \begin{array}{l} I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{2c}; y_1 | x_1) \\ +I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&\quad + \left\{ I(x_1, x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \right\} \\
&= (16) + (20) + (18),
\end{aligned}$$

which is greater than three times the active sum rate bound.

Hence, we conclude that in case (a), the corner point where $R_1 + R_2$ bound and $R_1 + 2R_2$ bound intersect can be achieved.

To show (b), since we have two possible R_2 bounds, we distinguish into 2 cases:

1) If the bound

$$R_2 \le I(x_2; y_2 | x_{1c}) + \mathsf{C}_{12}^{\mathsf{B}}$$

is active, note that the point (R_1^*, R_2^*) where the R_2 bound and the $2R_1 + R_2$ bound (21) intersect, satisfies

$$\begin{aligned}
&2R_1^* + 2R_2^* \\
&= \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1 | x_{1c}, x_{2c}) \\ +I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&\quad + \left\{ I(x_2; y_2 | x_{1c}) + \mathsf{C}_{12}^{\mathsf{B}} \right\} \\
&= I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \\
&\quad + I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}_{12}^{\mathsf{B}} \\
&\quad + I(x_2; y_2 | x_{1c}) + I(x_{1c}; y_2 | x_{2c}) - I(x_{1c}, x_2; y_2) + xs\mathsf{C}_{12}^{\mathsf{B}} \\
&= (16) + (19) + \left[I(x_{1c}; y_2 | x_{2c}) - I(x_{1c}; y_2) + \mathsf{C}_{12}^{\mathsf{B}} \right] \\
&\overset{(**)}{\ge} (16) + (19),
\end{aligned}$$

which is greater than two times the active sum rate bound. $(**)$ is due to

$$\begin{aligned}
I(x_{1c}; y_2 | x_{2c}) &= I(x_{1c}; y_2, x_{2c}) - I(x_{1c}; x_{2c}) \\
&= I(x_{1c}; y_2, x_{2c}) \ge I(x_{1c}; y_2),
\end{aligned}$$

since x_{1c} and x_{2c} are independent.

2) If the bound

$$R_2 \le I(x_{2c}; y_1 | x_1) + I(x_2; y_2 | x_{1c}, x_{2c})$$

is active, note that the point (R_1^*, R_2^*) where the R_2 bound and the $2R_1 + R_2$ bound (21) intersect, satisfies

$$\begin{aligned}
2R_1^* + 2R_2^* &= \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_1; y_1 | x_{1c}, x_{2c}) \\ +I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&\quad + \{ I(x_{2c}; y_1 | x_1) + I(x_2; y_2 | x_{1c}, x_{2c}) \} \\
&= \{ I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \} \\
&\quad + \left\{ \begin{array}{l} I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{2c}; y_1 | x_1) \\ +I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&= (16) + (20),
\end{aligned}$$

which is greater than two times the active sum rate bound.

Hence, we conclude that in case (b), the corner point where $R_1 + R_2$ bound and R_2 bound intersect can be achieved. ■

B) Proof of Claim 5.7:

Proof: (Keep in mind $\Delta_2 = 1 + \mathsf{SNR}_{2p}$ and $\mathsf{INR}_{ip} \le 1$, $i = 1, 2$)

1) R_1 bound: We have two bounds. First, $I(x_1; y_1 | x_{2c}) = \log\left(1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_{1p}}\right)$, which is within 2 bits to the upper bound $\log(1 + \mathsf{SNR}_1 + \mathsf{INR}_2)$. Second,

$$\begin{aligned}
&I(x_1; y_1 | x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2 | x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\
&= \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1 + \mathsf{INR}_{1p}}\right) \\
&\quad + \log\left(\frac{1 + \mathsf{SNR}_{2p} + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&\ge \log\left(\frac{1 + \mathsf{SNR}_1 + \mathsf{INR}_2}{1 + \mathsf{INR}_{1p}}\right) - 1.
\end{aligned}$$

Hence, if the second bound is active, it is within 2 bits to the upper bound $\log(1 + \mathsf{SNR}_1 + \mathsf{INR}_2)$.

2) R_2 bound: We have two bounds. First, $I(x_2; y_2 | x_{1c}) + \mathsf{C}_{12}^{\mathsf{B}} = \log\left(1 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}}$. If the first bound is active, it is within 1 bit to the upper bound $\log(1 + \mathsf{SNR}_2) + \mathsf{C}_{12}^{\mathsf{B}}$. Second,

$$\begin{aligned}
&I(x_{2c}; y_1 | x_1) + I(x_2; y_2 | x_{1c}, x_{2c}) \\
&= \log\left(\frac{1 + \mathsf{INR}_1}{1 + \mathsf{INR}_{1p}}\right) + \log\left(\frac{1 + \mathsf{SNR}_{2p} + \mathsf{INR}_{2p}}{1 + \mathsf{INR}_{2p}}\right) \\
&\ge \log\left(\frac{1 + \mathsf{SNR}_2 + \mathsf{INR}_1}{1 + \mathsf{INR}_{2p}}\right) - 1.
\end{aligned}$$

Hence, the second bound is within 2 bits to the upper bound $\log(1 + \mathsf{SNR}_2 + \mathsf{INR}_1)$.

3) $R_1 + R_2$ bound: We have six bounds for $R_1 + R_2$, investigated as follows:

- First,

$$\begin{aligned}
&I(x_1, x_{2c}; y_1) + I(x_2; y_2 | x_{1c}, x_{2c}) + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \\
&\qquad = \left\{ \begin{array}{l} \log\left(\frac{1+\mathsf{SNR}_1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(1 + \frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) \\ + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\},
\end{aligned}$$

which is within $2 + 1 = 3$ bits to the upper bound (8).

- Second,

$$\begin{aligned}
&I(x_1, x_{2c}; y_1, \hat{y}_2) + I(x_2; y_2 | x_{1c}, x_{2c}) \\
&= \log\left(\frac{\begin{array}{c}(1+\Delta_2)(1+\mathsf{SNR}_1+\mathsf{INR}_1) \\ +\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\end{array}}{(1+\Delta_2)(1+\mathsf{INR}_{1p}) + \mathsf{SNR}_{2p}}\right) \\
&\quad + \log\left(1 + \frac{\mathsf{SNR}_{2p}}{1 + \mathsf{INR}_{2p}}\right)
\end{aligned}$$

$$\overset{(a)}{\geq} \log\left(\frac{\substack{1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2\\ +|h_{11}h_{22}-h_{12}h_{21}|^2}}{4\Delta_2}\right) + \log\left(1+\mathsf{SNR}_{2p}\right) - 1$$
$$= \log\left(\substack{1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2\\ +|h_{11}h_{22}-h_{12}h_{21}|^2}\right) + \log\left(\frac{1+\mathsf{SNR}_{2p}}{\Delta_2}\right) - 3$$
$$\overset{(b)}{\geq} \log\left(\begin{array}{l}1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2\\ +|h_{11}h_{22}-h_{12}h_{21}|^2\end{array}\right) - 3,$$

where (a) is due to $(1+\Delta_2)(1+\mathsf{INR}_{1p})+\mathsf{SNR}_{2p} \leq (1+\Delta_2)2\ \mathsf{SNR}_{2p} = 4+3\mathsf{SNR}_{2p} \leq 4\Delta_2$ since $\Delta_2 = 1+\mathsf{SNR}_{2p}$ and $\mathsf{INR}_{1p} \leq 1$. (b) is due to $\Delta_2 = 1+\mathsf{SNR}_{2p}$. This lower bound is within 3 bits to the upper bound (9).

- Third,

$$I(x_1,x_{2c};y_1|x_{1c}) + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+$$
$$= \log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \log\left(\frac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+,$$

which is within $2+1=3$ bits to the upper bound (6).

- Fourth,

$$\begin{aligned} &I(x_1,x_{2c};y_1,\hat{y}_2|x_{1c}) + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ &= I(x_{2c};y_1,\hat{y}_2|x_{1c}) + I(x_1;y_1,\hat{y}_2|x_{1c},x_{2c}) \\ &\quad + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ &\geq I(x_{2c};\hat{y}_2|x_{1c}) + I(x_1;y_1|x_{1c},x_{2c}) \\ &\quad + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ &\overset{(a)}{\geq} I(x_{2c};y_2|x_{1c}) - 1 + I(x_1;y_1|x_{1c},x_{2c}) \\ &\quad + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ &\overset{(b)}{\geq} I(x_1;y_1|x_{1c},x_{2c}) + I(x_{1c},x_2;y_2) + \mathsf{C}_{12}^{\mathsf{B}} - 1 \\ &= \left\{\begin{array}{l}\log\left(1+\frac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) + \log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) \\ +\mathsf{C}_{12}^{\mathsf{B}} - 1\end{array}\right\}, \end{aligned}$$

which is within 3 bits to the upper bound (7). Note that (a) is due to

$$\begin{aligned} &I(x_{2c};\hat{y}_2|x_{1c}) \\ &= \log\left(\frac{1+\Delta_2+\mathsf{INR}_{2p}+\mathsf{SNR}_2}{1+\Delta_2+\mathsf{INR}_{2p}+\mathsf{SNR}_{2p}}\right) \\ &\geq \log\left(\frac{1+\mathsf{INR}_{2p}+\mathsf{SNR}_2}{1+(1+\mathsf{SNR}_{2p})+\mathsf{INR}_{2p}+\mathsf{SNR}_{2p}}\right) \\ &\geq \log\left(\frac{1+\mathsf{INR}_{2p}+\mathsf{SNR}_2}{1+\mathsf{INR}_{2p}+\mathsf{SNR}_{2p}}\right) - 1 \\ &= I(x_{2c};y_2|x_{1c}) - 1. \end{aligned}$$

(b) is due to

$$\begin{aligned} &I(x_{2c};y_2|x_{1c}) + I(x_{1c},x_2;y_2|x_{2c}) \\ &\quad = I(x_{2c};y_2,x_{1c}) + I(x_{1c},x_2;y_2|x_{2c}) \\ &\quad \geq I(x_{2c};y_2) + I(x_{1c},x_2;y_2|x_{2c}) \\ &\quad = I(x_{1c},x_2,x_{2c};y_2) = I(x_{1c},x_2;y_2). \end{aligned}$$

- Fifth,

$$I(x_1;y_1|x_{1c},x_{2c}) + I(x_{1c},x_2;y_2) + \mathsf{C}_{12}^{\mathsf{B}} = \left\{\begin{array}{l}\log\left(1+\frac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}}\end{array}\right\},$$

which is within 2 bits to the upper bound (7).

- Sixth,

$$\begin{aligned} &I(x_1;y_1|x_{1c},x_{2c}) + I(x_{2c};y_1|x_1) \\ &\quad + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} \\ &= \left\{\begin{array}{l}\log\left(1+\frac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) + \log\left(\frac{1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}}\end{array}\right\} \\ &\geq \left\{\begin{array}{l}\log\left(1+\frac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) + \mathsf{C}_{12}^{\mathsf{B}}\end{array}\right\}, \end{aligned}$$

which is within 3 bits to the upper bound (7).

4) $2R_1+R_2$ bound: The bound

$$\begin{aligned} &I(x_1,x_{2c};y_1) + I(x_1;y_1|x_{1c},x_{2c}) \\ &\quad + I(x_{1c},x_2;y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+ \\ &= \left\{\begin{array}{l}\log\left(\frac{1+\mathsf{SNR}_1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(1+\frac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ +\log\left(\frac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\end{array}\right\}, \end{aligned}$$

which is within $3+1=4$ bits to the upper bound (10).

5) R_1+2R_2 bound: We have six bounds for R_1+2R_2, investigated as follows:

- First,

$$\begin{aligned} &I(x_1,x_{2c};y_1|x_{1c}) + I(x_{1c},x_2;y_2) \\ &\quad + I(x_2;y_2|x_{1c},x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+ \\ &= \left\{\begin{array}{l}\log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) \\ +\log\left(1+\frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\end{array}\right\}, \end{aligned}$$

which is within $3+1=4$ bits to the upper bound (11).

- Second,

$$\begin{aligned} &I(x_1,x_{2c};y_1|x_{1c}) + I(x_{2c};y_1|x_1) + I(x_{1c},x_2;y_2|x_{2c}) \\ &\quad + I(x_2;y_2|x_{1c},x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+ \\ &= \left\{\begin{array}{l}\log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\frac{1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \log\left(1+\frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) \\ +\mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\end{array}\right\} \\ &\geq \left\{\begin{array}{l}\log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) \\ +\log\left(1+\frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\end{array}\right\}, \end{aligned}$$

which is within $4+1=5$ bits to the upper bound (11).

- Third,

$$\begin{aligned} &I(x_1,x_{2c};y_1,\hat{y}_2|x_{1c})+I(x_{1c},x_2;y_2) \\ &\quad +I(x_2;y_2|x_{1c},x_{2c})+\mathsf{C}_{12}^{\mathsf{B}} \\ &= \left\{ \begin{array}{l} \log\left(\frac{(1+\Delta_2)(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{(1+\Delta_2)(1+\mathsf{INR}_{1p})+\mathsf{SNR}_{2p}}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\log\left(1+\frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) \\ +\mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\ &\geq \left\{ \begin{array}{l} \log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{4\Delta_2}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\log(1+\mathsf{SNR}_{2p}) \\ +\mathsf{C}_{12}^{\mathsf{B}}-1 \end{array} \right\} \\ &\geq \left\{ \begin{array}{l} \log\left(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}}-3 \end{array} \right\}, \end{aligned}$$

which is within $1+3=4$ bits to the upper bound (13).

- Fourth,

$$\begin{aligned} &I(x_1,x_{2c};y_1,\hat{y}_2|x_{1c})+I(x_{2c};y_1|x_1) \\ &\quad +I(x_{1c},x_2;y_2|x_{2c})+I(x_2;y_2|x_{1c},x_{2c})+\mathsf{C}_{12}^{\mathsf{B}} \\ &= \left\{ \begin{array}{l} \log\left(\frac{(1+\Delta_2)(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{(1+\Delta_2)(1+\mathsf{INR}_{1p})+\mathsf{SNR}_{2p}}\right) \\ +\log\left(\frac{1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right)+\log\left(\frac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) \\ +\log\left(1+\frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\ &\geq \left\{ \begin{array}{l} \log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{4\Delta_2}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) \\ +\log(1+\mathsf{SNR}_{2p})+\mathsf{C}_{12}^{\mathsf{B}}-1 \end{array} \right\} \\ &\geq \left\{ \begin{array}{l} \log\left(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right)+\mathsf{C}_{12}^{\mathsf{B}}-3 \end{array} \right\}, \end{aligned}$$

which is within $2+3=5$ bits to the upper bound (13).

Therefore, we see that the bounds in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ except (21) satisfies:

- R_1 bound is within 2 bits to outer bounds;
- R_2 bound is within 2 bits to outer bounds;
- R_1+R_2 bound is within 3 bits to outer bounds;
- $2R_1+R_2$ bound is within 4 bits to outer bounds;
- R_1+2R_2 bound is within 5 bits to outer bounds. ∎

C) Proof of Claim 5.9:

Proof: (Keep in mind $\Delta_2=1$ and $\mathsf{INR}_{2p}\leq 1$)

(1) R_1 bound: We have two bounds. First, $I(x_1;y_1|x_2)=\log(1+\mathsf{SNR}_1)$, which is within 1 bit to the upper bound $R_1\leq\log(1+\mathsf{SNR}_1+\mathsf{INR}_2)$. Second,

$$\begin{aligned} &I(x_1;y_1|x_{1c},x_2)+I(x_{1c};y_2|x_2)+\mathsf{C}_{12}^{\mathsf{B}} \\ &\quad =\log(1+\mathsf{SNR}_{1p})+\log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}} \\ &\quad \geq\log\left(\frac{1+\mathsf{SNR}_1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}}. \end{aligned}$$

Hence, if the second bound is active, it is within 1 bit to the upper bound $\log(1+\mathsf{SNR}_1+\mathsf{INR}_2)$.

(2) R_2 bound: We have two bounds. First, $I(x_1;y_1|x_1)=\log(1+\mathsf{INR}_1)$, which is within 1 bit to the upper bound $R_2\leq\log(1+\mathsf{SNR}_2+\mathsf{INR}_1)$. Second, $I(x_2;y_2|x_{1c})+\mathsf{C}_{12}^{\mathsf{B}}=\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_{2p}}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}}$, which is within 1 bit to the upper bound $R_2\leq\log(1+\mathsf{SNR}_2)+\mathsf{C}_{12}^{\mathsf{B}}$.

(3) R_1+R_2 bound: We have five bounds, investigated as follows:

- First,

$$\begin{aligned} &I(x_1,x_2;y_1)+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+ \\ &\qquad =\log(1+\mathsf{SNR}_1+\mathsf{INR}_1)+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+, \end{aligned}$$

which is within $1+\xi_1=2$ bits to the upper bound (8).

- Second,

$$I(x_1,x_2;y_1,\hat{y}_2)=\log\left(\frac{2(1+\mathsf{SNR}_1+\mathsf{INR}_1)+\mathsf{SNR}_2 +\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2}{2}\right)$$

which is within 1 bit to the upper bound (9).

- Third,

$$\begin{aligned} &I(x_1;y_1|x_{1c},x_2)+I(x_{1c},x_2;y_2)+\mathsf{C}_{12}^{\mathsf{B}} \\ &\qquad =\log(1+\mathsf{SNR}_{1p})+\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}} \end{aligned}$$

which is within 1 bit to the upper bound (7).

- Fourth,

$$\begin{aligned} &I(x_1,x_2;y_1|x_{1c})+I(x_{1c};y_2|x_2)+\mathsf{C}_{12}^{\mathsf{B}}+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+ \\ &\qquad =\left\{ \begin{array}{l} \log(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) \\ \qquad +\mathsf{C}_{12}^{\mathsf{B}}+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+ \end{array} \right\}, \end{aligned}$$

which is within $2+\xi_1=3$ bits to the upper bound (6).

- Fifth,

$$\begin{aligned} &I(x_1,x_2;y_1,\hat{y}_2|x_{1c})+I(x_{1c};y_2|x_2)+\mathsf{C}_{12}^{\mathsf{B}} \\ &\quad =I(x_2;y_1,\hat{y}_2|x_{1c})+I(x_1;y_1,\hat{y}_2|x_{1c},x_2) \\ &\qquad +I(x_{1c};y_2|x_2)+\mathsf{C}_{12}^{\mathsf{B}} \\ &\quad \geq I(x_2;y_1,\hat{y}_2|x_{1c})+I(x_1;y_1|x_{1c},x_2) \\ &\qquad +I(x_{1c};y_2|x_2)+\mathsf{C}_{12}^{\mathsf{B}} \\ &\quad =\left\{ \begin{array}{l} \log\left(\frac{2(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{2(1+\mathsf{SNR}_{1p})+\mathsf{INR}_{2p}}\right) \\ +\log(1+\mathsf{SNR}_{1p})+\log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \end{aligned}$$

$$\geq \left\{ \begin{array}{c} \log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2Q_{1p}}{3(1+\mathsf{SNR}_{1p})}\right) \\ +\log\left(1+\mathsf{SNR}_{1p}\right)+\log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\}$$

$$\geq \left\{ \begin{array}{c} \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2 +|h_{11}h_{22}-h_{12}h_{21}|^2\right) \\ +\log\left(\frac{1}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12}-\log 3 \end{array} \right\}.$$

Hence, if this bound is active, it is within $1+\log 3 = \log 6$ bits to the upper bound (9).

(4) $R_1 + 2R_2$ bound: We have two bounds. First,

$$I\left(x_1,x_2;y_1|x_{1c}\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}^{\mathsf{B}}_{12}+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ = \left\{ \begin{array}{c} \log\left(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1\right)+\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) \\ +\mathsf{C}^{\mathsf{B}}_{12}+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ \end{array} \right\},$$

which is within $2+\xi_1 = 3$ bits to the upper bound (11) Second,

$$I\left(x_1,x_2;y_1,\hat{y}_2|x_{1c}\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}^{\mathsf{B}}_{12} = \left\{ \begin{array}{c} \log\left(\frac{2(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p} +|h_{11}h_{22}-h_{12}h_{21}|^2Q_{1p}}{2}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\},$$

which is within 2 bits to the upper bound (13).

Therefore, we see that the bounds in $\mathscr{R}_{2\longrightarrow 1\longrightarrow 2}$ satisfies:
- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- $R_1 + R_2$ bound is within 3 bits to outer bounds;
- $R_1 + 2R_2$ bound is within 3 bits to outer bounds.

■

D) Proof of Claim 5.12:

Proof: (Keep in mind that $\Delta_1 = \Delta_2 = 1$)

(1) R_1 bound: We have four bounds. First,

$$I\left(x_1;y_1|x_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ = \log\left(1+\mathsf{SNR}_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+$$

which is within $\xi_1 = 1$ bit to the upper bound $\log\left(1+\mathsf{SNR}_1\right)+\mathsf{C}^{\mathsf{B}}_{21}$. Second,

$$I\left(x_1;y_2|x_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2\right)^+ = \log\left(1+\mathsf{INR}_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2\right)^+ \geq \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)-1.$$

Hence if this bound is active, it is within 1 bit to the upper bound $\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)$. Finally,

$$I\left(x_1;y_1,\hat{y}_2|x_2\right) = \log\left(\frac{2+2\mathsf{SNR}_1+\mathsf{INR}_2}{2}\right)$$
$$I\left(x_1;y_2,\hat{y}_1|x_2\right) = \log\left(\frac{2+\mathsf{SNR}_1+2\mathsf{INR}_2}{2}\right),$$

which are both within 1 bit to the upper bound $\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)$.

(2) R_2 bound: By symmetry we have the same gap result as (1).

(3) $R_1 + R_2$ bound: We have four bounds. First,

$$I\left(x_1,x_2;y_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ = \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+,$$

which is within $1+\xi_1 = 2$ bits to the upper bound (8). Second,

$$I\left(x_2,x_1;y_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2\right)^+ = \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2\right)^+,$$

which is within $1+\xi_2 = 2$ bits to the upper bound (7). Finally,

$$I\left(x_1,x_2;y_1,\hat{y}_2\right) = \log\left(\frac{2(1+\mathsf{SNR}_1+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_2 +|h_{11}h_{22}-h_{12}h_{21}|^2}{2}\right)$$
$$I\left(x_2,x_1;y_2,\hat{y}_1\right) = \log\left(\frac{2(1+\mathsf{SNR}_2+\mathsf{INR}_2)+\mathsf{SNR}_1+\mathsf{INR}_1 +|h_{11}h_{22}-h_{12}h_{21}|^2}{2}\right),$$

which are both within 1 bit to the upper bound (9).

Therefore, we see that the bounds in $\mathscr{R}_{\mathrm{OneRound}}$ satisfies:
- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- $R_1 + R_2$ bound is within 2 bits to outer bounds.

■

APPENDIX D
PROOF OF THEOREM 6.2

From Section V-E, we have shown that when $\mathsf{SNR} \leq \mathsf{INR}$,

$$R_{\mathrm{sym,OneRound}} \leq C_{\mathrm{sym}} \leq \bar{C}_{\mathrm{sym}} \leq R_{\mathrm{sym,OneRound}} + 1.$$

Hence we focus on the case $\mathsf{SNR} > \mathsf{INR}$ in the rest of the proof.

By symmetry and by Theorem 5.10, if $R_{\mathrm{sym,OneRound}} \geq 0$ satisfies the following, it is achievable:

$$R_{\mathrm{sym,OneRound}} \leq \min\left\{ \begin{array}{l} I\left(x_{2c},x_1;y_1|x_{1c}\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ \\ ,I\left(x_{2c},x_1;y_1,\hat{y}_2|x_{1c}\right) \end{array} \right\},$$

$$R_{\mathrm{sym,OneRound}} \leq \min\left\{I\left(x_1;y_1|x_{2c}\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_1;y_1,\hat{y}_2|x_{2c}\right)\right\},$$

$$2R_{\mathrm{sym,OneRound}} \leq \min\left\{I\left(x_1,x_{2c};y_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_1,x_{2c};y_1,\hat{y}_2\right)\right\} + \min\left\{ \begin{array}{l} I\left(x_1;y_1|x_{1c},x_{2c}\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ \\ ,I\left(x_1;y_1,\hat{y}_2|x_{1c},x_{2c}\right) \end{array} \right\}.$$

Note that since

$$I\left(x_1;y_1|x_{1c},x_{2c}\right) \leq I\left(x_1;y_1,\hat{y}_2|x_{1c},x_{2c}\right) \leq I\left(x_1;y_1|x_{1c},x_{2c}\right)+\text{constant},$$
$$I\left(x_1;y_1|x_{2c}\right) \leq I\left(x_1;y_1,\hat{y}_2|x_{2c}\right) \leq I\left(x_1;y_1|x_{2c}\right)+\text{constant},$$

a sufficient condition for achievable $R_{\text{sym,OneRound}}$ is

$$R_{\text{sym,OneRound}} \leq \min \left\{ \begin{array}{l} I(x_{2c}, x_1; y_1 | x_{1c}) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+ \\ , I(x_{2c}, x_1; y_1, \hat{y}_2 | x_{1c}), I(x_1; y_1 | x_{2c}) \end{array} \right\}$$

$$R_{\text{sym,OneRound}} \leq \frac{1}{2} \min \left\{ \begin{array}{l} I(x_1, x_{2c}; y_1) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+ \\ , I(x_1, x_{2c}; y_1, \hat{y}_2) \end{array} \right\} + \frac{1}{2} I(x_1; y_1 | x_{1c}, x_{2c})$$

(1) $I(x_{2c}, x_1; y_1 | x_{1c}) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+$:

$$I(x_{2c}, x_1; y_1 | x_{1c}) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+ = \log\left(\frac{1 + \mathsf{SNR}_p + \mathsf{INR}}{1 + \mathsf{INR}_p}\right) + (\mathsf{C}^{\mathsf{B}} - \xi)^+,$$

and its gap to the outer bound $\log\left(1 + \mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \mathsf{C}^{\mathsf{B}}$:

$$\text{gap} \leq \log(1 + \mathsf{INR}_p) + \xi \leq 1 + 1 = 2.$$

(2) $I(x_{2c}, x_1; y_1, \hat{y}_2 | x_{1c})$:

$$\begin{aligned} I(x_{2c}, x_1; y_1, \hat{y}_2 | x_{1c}) &= I(x_{2c}; y_1, \hat{y}_2 | x_{1c}) + I(x_1; y_1, \hat{y}_2 | x_{2c}, x_{1c}) \\ &\geq I(x_{2c}; \hat{y}_2 | x_{1c}) + I(x_1; y_1 | x_{2c}, x_{1c}) \\ &= \log\left(\frac{1 + \Delta + \mathsf{SNR} + \mathsf{INR}_p}{1 + \Delta + \mathsf{SNR}_p + \mathsf{INR}_p}\right) + \log\left(\frac{1 + \mathsf{SNR}_p + \mathsf{INR}_p}{1 + \mathsf{INR}_p}\right) \\ &\overset{(a)}{\geq} \log\left(\frac{1 + \mathsf{SNR}}{2 + 2\mathsf{SNR}_p + 2\mathsf{INR}_p}\right) + \log\left(\frac{1 + \mathsf{SNR}_p + \mathsf{INR}_p}{1 + \mathsf{INR}_p}\right) \\ &= \log\left(\frac{1 + \mathsf{SNR}}{1 + \mathsf{INR}_p}\right) - 1, \end{aligned}$$

where (a) is due to $\Delta = 1 + \mathsf{SNR}_p$.
Therefore, the gap to the outer bound $\log(1 + \mathsf{SNR} + \mathsf{INR})$:

$$\begin{aligned} \text{gap} &\leq 1 + \log\left(\frac{1 + \mathsf{SNR} + \mathsf{INR}}{1 + \mathsf{SNR}}\right) + \log(1 + \mathsf{INR}_p) \\ &\leq 1 + \log\left(\frac{2 + 2\mathsf{SNR}}{1 + \mathsf{SNR}}\right) + \log(1 + 1) = 3, \end{aligned}$$

since $\mathsf{SNR} > \mathsf{INR}$ and $\mathsf{INR}_p \leq 1$.

(3) $I(x_1; y_1 | x_{2c})$:

$$I(x_1; y_1 | x_{2c}) = \log(1 + \mathsf{SNR} + \mathsf{INR}_p) - \log(1 + \mathsf{INR}_p),$$

and its gap to the outer bound $\log(1 + \mathsf{SNR} + \mathsf{INR})$:

$$\begin{aligned} \text{gap} &\leq \log\left(\frac{1 + \mathsf{SNR} + \mathsf{INR}}{1 + \mathsf{SNR} + \mathsf{INR}_p}\right) + \log(1 + \mathsf{INR}_p) \\ &\leq \log\left(\frac{2 + 2\mathsf{SNR}}{1 + \mathsf{SNR}}\right) + \log(1 + 1) = 2. \end{aligned}$$

(4) $\frac{1}{2} I(x_1, x_{2c}; y_1) + \frac{1}{2}(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+ + \frac{1}{2} I(x_1; y_1 | x_{1c}, x_{2c})$:

$$\begin{aligned} &\frac{1}{2} I(x_1, x_{2c}; y_1) + \frac{1}{2}(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+ + \frac{1}{2} I(x_1; y_1 | x_{1c}, x_{2c}) \\ &= \frac{1}{2}\log(1 + \mathsf{SNR} + \mathsf{INR}) + \frac{1}{2}(\mathsf{C}^{\mathsf{B}} - \xi)^+ \\ &\quad + \frac{1}{2}\log(1 + \mathsf{SNR}_p + \mathsf{INR}_p) - \log(1 + \mathsf{INR}_p), \end{aligned}$$

and its gap to the outer bound $\frac{1}{2}\log(1 + \mathsf{SNR} + \mathsf{INR}) + \frac{1}{2}\log\left(1 + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \frac{1}{2}\mathsf{C}^{\mathsf{B}}$:

$$\text{gap} \leq \frac{1}{2}\xi + \log(1 + \mathsf{INR}_p) \leq 1.5.$$

(5) $\frac{1}{2} I(x_1, x_{2c}; y_1, \hat{y}_2) + \frac{1}{2} I(x_1; y_1 | x_{1c}, x_{2c})$:

$$\begin{aligned} &\frac{1}{2} I(x_1, x_{2c}; y_1, \hat{y}_2) + \frac{1}{2} I(x_1; y_1 | x_{1c}, x_{2c}) \\ &= \frac{1}{2}\log\left(\frac{\Delta(1+\mathsf{SNR}+\mathsf{INR})+1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2}{\Delta(1 + \mathsf{INR}_p) + 1 + \mathsf{SNR}_p + \mathsf{INR}_p}\right) \\ &\quad + \frac{1}{2}\log\left(\frac{1 + \mathsf{SNR}_p + \mathsf{INR}_p}{1 + \mathsf{INR}_p}\right) \\ &\geq \frac{1}{2}\log\left(\frac{1 + 2\mathsf{SNR} + 2\mathsf{INR} + |h_{11}h_{22} - h_{12}h_{21}|^2}{4\Delta}\right) \\ &\quad + \frac{1}{2}\log\left(\frac{\Delta}{1 + \mathsf{INR}_p}\right) \\ &= \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + |h_{11}h_{22} - h_{12}h_{21}|^2\right) \\ &\quad - \frac{1}{2}\log(1 + \mathsf{INR}_p) - 1. \end{aligned}$$

Therefore, the gap to the outer bound $\frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + |h_{11}h_{22} - h_{12}h_{21}|^2\right)$:

$$\text{gap} \leq \frac{1}{2}\log(1 + \mathsf{INR}_p) + 1 \leq 1.5.$$

From (1)–(5), we conclude that when $\mathsf{SNR} > \mathsf{INR}$,

$$R_{\text{sym,OneRound}} \leq C_{\text{sym}} \leq \bar{C}_{\text{sym}} \leq R_{\text{sym,OneRound}} + 3.$$

This completes the proof.

Appendix E
Proof of Lemma 7.2

Proof: From Corollary 7.1 we see that except the term

$$V := \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + |h_{11}h_{22} - h_{12}h_{21}|^2\right),$$

all terms scaled by $\log \mathsf{SNR}$ converges *everywhere* as $\mathsf{SNR} \longrightarrow \infty$ with α, κ fixed. Note that

$$\begin{aligned} &|h_{11}h_{22} - h_{12}h_{21}|^2 \\ &= |g_{11}e^{j\Theta_{11}}g_{22}e^{j\Theta_{22}} - g_{12}e^{j\Theta_{12}}g_{21}e^{j\Theta_{21}}|^2 \\ &= \Big[g_{11}g_{22}\cos(\Theta_{11}+\Theta_{22}) - g_{12}g_{21}\cos(\Theta_{12}+\Theta_{21})\Big]^2 \\ &\quad + \Big[g_{11}g_{22}\sin(\Theta_{11}+\Theta_{22}) - g_{12}g_{21}\sin(\Theta_{12}+\Theta_{21})\Big]^2 \\ &= g_{11}^2g_{22}^2 + g_{12}^2g_{21}^2 \\ &\quad - 2g_{11}g_{22}g_{12}g_{21}\cos(\Theta_{11}+\Theta_{22}-\Theta_{12}-\Theta_{21}) \\ &= \mathsf{SNR}^2 + \mathsf{INR}^2 - 2(\cos\Theta)\mathsf{SNR}\mathsf{INR}, \end{aligned}$$

where $\Theta = \Theta_{11} + \Theta_{22} - \Theta_{12} - \Theta_{21} \bmod 2\pi$. Obviously Θ is uniformly distributed over $[0, 2\pi]$. Now, consider the limit

$$L(\alpha, \kappa) := \lim_{\substack{\text{fix } \alpha,\kappa \\ \mathsf{SNR} \longrightarrow \infty}} \frac{V}{\log \mathsf{SNR}}.$$

We have the following upper and lower bounds for V due to the fact that $\big||h_{11}||h_{22}| - |h_{12}||h_{21}|\big| \le |h_{11}h_{22} - h_{12}h_{21}| \le |h_{11}||h_{22}| + |h_{12}||h_{21}|$:

$$\begin{aligned} V &\ge \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + (\mathsf{SNR} - \mathsf{INR})^2\right); \\ V &\le \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + (\mathsf{SNR} + \mathsf{INR})^2\right). \end{aligned}$$

Hence, when $\alpha < 1$, taking limits at both sides yields $1 \le L(\alpha,\kappa) \le 1$ and implies $L(\alpha,\kappa) = 1$. Similarly, when $\alpha > 1$, taking limits at both sides yields $\alpha \le L(\alpha,\kappa) \le \alpha$ and implies $L(\alpha,\kappa) = \alpha$. When $\alpha = 1$, note that

$$\begin{aligned} V &= \frac{1}{2}\log\begin{pmatrix} 1 + 2\mathsf{SNR} + 2\mathsf{INR} + \mathsf{SNR}^2 + \mathsf{INR}^2 \\ -2(\cos\Theta)\mathsf{SNR}\mathsf{INR} \end{pmatrix} \\ &= \frac{1}{2}\log\left((1 + \mathsf{SNR} + \mathsf{INR})^2 - 4\cos^2\frac{\Theta}{2}\mathsf{SNR}\mathsf{INR}\right), \end{aligned}$$

and therefore $L(\alpha,\kappa) = 1$ if $\Theta \ne 0, 2\pi$. Since the event $\{\Theta = 0, 2\pi\}$ is of zero measure, the limit $L(\alpha,\kappa)$ exists almost surely. ■

REFERENCES

[1] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, pp. 786–788, Nov. 1981.

[2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, pp. 49–60, Jan. 1981.

[3] R. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[4] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, pp. 620–643, Feb. 2009.

[5] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, pp. 689–699, Feb. 2009.

[6] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3032–3050, Jul. 2009.

[7] F. M. J. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. IT-29, pp. 441–445, May 1983.

[8] S. I. Bross, A. Lapidoth, and M. A. Wigger, "The Gaussian MAC with conferencing encoders," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008.

[9] I. Marić, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3536–3548, Oct. 2007.

[10] R. Dabora and S. D. Servetto, "Broadcast channels with cooperating decoders," *IEEE Trans. Inf. Theory*, vol. 52, pp. 5438–5454, Dec. 2006.

[11] O. Simeone, D. Gündüz, H. V. Poor, A. J. Goldsmith, and S. Shamai, "Compound multiple access channels with partial cooperation," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2425–2441, Jun. 2009.

[12] L. Zhou and W. Yu, "Gaussian z-interference channel with a relay link: Achievability region and asymptotic sum capacity," *IEEE Trans. Inf. Theory*, Jun. 2010 [Online]. Available: http://arxiv.org/abs/1006.5087, submitted for publication

[13] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1522–1544, Apr. 2006.

[14] V. Prabhakaran and P. Viswanath, "Interference channels with destination cooperation," *IEEE Trans. Inf. Theory*, vol. 57, pp. 187–209, Jan. 2011.

[15] I. Marić, R. Dabora, and A. J. Goldsmith, "On the capacity of the interference channel with a relay," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 554–558.

[16] O. Sahin, E. Erkip, and O. Simeone, "Interference channel with a relay: Models, relaying strategies, bounds," presented at the Inf. Theory Appl. Workshop, San Diego, CA, Feb. 2009.

[17] P. Razaghi and W. Yu, "Universal relaying for the interference channel," presented at the Inf. Theory Appl. Workshop, San Diego, CA, 2010.

[18] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. N. C. Tse, "Approximate capacity of a class of Gaussian interference-relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2837–2864, May 2011.

[19] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 572–584, Sep. 1979.

[20] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, Sep. 2005.

[21] T. M. Cover and Y.-H. Kim, "Capacity of a class of deterministic relay channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 591–595.

[22] Y.-H. Kim, "Coding techniques for primitive relay channels," presented at the Allerton Conf. Commun., Control, Comput., Monticello, IL, Sep. 2007.

[23] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3008–3023, Jul. 2008.

[24] C. Schnurr, S. Stanczak, and T. J. Oechtering, "Coding theorems for the restricted half-duplex two-way relay channel with joint decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 2688–2692.

[25] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1872–1905, Apr. 2011.

[26] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow," presented at the Allerton Conf. Commun., Control, Comput., Monticello, IL, Sep. 2007.

[27] G. Bresler and D. N. C. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecommun.*, vol. 19, pp. 333–354, Jun. 2008.

[28] S. C. Draper, B. J. Frey, and F. R. Kschischang, "Interactive decoding of a broadcast message," presented at the Allerton Conf. Commun., Control, Comput., Monticello, IL, Oct. 2003.

[29] S. H. Lim, Y.-H. Kim, A. A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, Feb. 2010 [Online]. Available: http://arxiv.org/abs/1002.3188, submitted for publication

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.

[31] H.-F. Chong, M. Motani, H. K. Garg, and H. E. Gamal, "On the han-kobayashi region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3188–3195, Jul. 2008.

[32] I-H. Wang and D. N. C. Tse, "Interference mitigation through limited transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2941–2965, May 2011.

I-Hsiang Wang (S'10) received the B.Sc. degree in electrical engineering from National Taiwan University, Taiwan, in 2006. Since 2006, he has been with University of California at Berkeley and is currently working towards the Ph.D. degree in the Department of Electrical Engineering and Computer Sciences.

His research interests include network information theory, wireless networks, coding theory and network coding.

Mr. Wang received a 2-year Vodafone Graduate Fellowship in 2006.

David N. C. Tse (M'96–SM'07–F'09) received the B.A.Sc. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1991 and 1994, respectively.

From 1994 to 1995, he was a Postdoctoral Member of Technical Staff at AT&T Bell Laboratories. Since 1995, he has been with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, where he is currently a Professor.

Dr. Tse received a 1967 NSERC 4-year graduate fellowship from the government of Canada in 1989, a NSF CAREER award in 1998, the Best Paper Awards at the Infocom 1998 and Infocom 2001 conferences, the Erlang Prize in 2000 from the INFORMS Applied Probability Society, the IEEE Communications and Information Theory Society Joint Paper Award in 2001, the Information Theory Society Paper Award in 2003, and the 2009 Frederick Emmons Terman Award from the American Society for Engineering Education. He has given plenary talks at international conferences such as ICASSP in 2006, MobiCom in 2007, CISS in 2008, and ISIT in 2009. He was the Technical Program Cochair of the International Symposium on Information Theory in 2004 and was an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY from 2001 to 2003. He is a coauthor, with P. Viswanath, of the text *Fundamentals of Wireless Communication*, which has been used in over 60 institutions around the world.