# Cooperative Relay Broadcast Channels

Yingbin Liang, *Member, IEEE*, and Venugopal V. Veeravalli, *Fellow, IEEE*

***Abstract*—The capacity regions are investigated for two relay broadcast channels (RBCs), where relay links are incorporated into two-user broadcast channels to support user cooperation. In the first channel, the partially cooperative RBC, only one user in the system acts as a relay. An achievable rate region is derived based on the relay using the decode-and-forward scheme. An outer bound on the capacity region is derived and is shown to be tighter than the cut-set bound. For the special case where the partially cooperative RBC is degraded, the achievable rate region is shown to be the capacity region. Two Gaussian cases of the partially cooperative RBC are studied. For the system where the additive white Gaussian noise (AWGN) term at one receiver is a degraded version of the other, which we refer to as the D-AWGN partially cooperative RBC, the capacity region is established. For the system where the AWGN term at one receiver is independent of the other, which we refer to as the AWGN partially cooperative RBC, inner and outer bounds on the capacity region are derived and are shown to be close. Furthermore, it is shown that feedback does not increase the capacity region for the degraded partially cooperative RBC, but that it improves the capacity region for the nondegraded version. In particular, feedback improves the capacity region for the AWGN partially cooperative RBC. In the second channel model being studied in the paper, the fully cooperative RBC, both users can act as relay nodes. All the results for the partially cooperative RBC are correspondingly generalized to the fully cooperative RBC. In particular, capacity regions are established for the degraded and D-AWGN fully cooperative RBCs. The capacity region is also established for the fully cooperative RBC with feedback. It is further shown that the AWGN fully cooperative RBC has a larger achievable rate region than its partially cooperative counterpart. The results illustrate that relaying and user cooperation are powerful techniques for improving the capacity of broadcast channels.**

***Index Terms*—Cut-set bound, degraded channel, feedback capacity, Gaussian channel, rate region, user cooperation, wireless downlink.**

## I. Introduction

COOPERATIVE relaying of information between users is emerging as a powerful technique for improving the reliability and throughput of wireless networks. The building block of such relay networks, the three-terminal relay channel, was first introduced by van der Meulen [1], and was comprehensively studied by Cover and El Gamal [2]. Recently, this channel has been further studied in a variety of contexts including Gaussian relay channels (e.g., [3], [4]), fading relay channels (e.g., [5]–[12]), relay channels with complexity constraints [13], relay channels with multiple antennas (e.g.,[7], [8]), and relay channels with orthogonal components (e.g., [14], [9], [15]–[19]). More complicated relay networks have also been studied including relay networks with multiple relay nodes simultaneously relaying information to the destination (e.g.,[7], [20]–[23]), relay networks with multiple levels of relay nodes forwarding information from one level to next (e.g.,[24]–[26], [7], [27]), and relay networks with multiple cooperative sources or destinations (e.g., [24], [28]–[30]). Furthermore, these information-theoretic studies of relay networks have motivated practical relaying protocols and coding design to achieve user-cooperative diversity (e.g., [31], [17], [32]–[34]).

For centralized networks, to date much of the work on this topic has focused on the uplink (from the users to a base station or access point). Cooperative diversity schemes, where one user may share another user's resources to improve its transmission rate, have been explored in a number of recent works (see, e.g., [5], [6], [14]). The use of a relay node to assist all the users in a multiple-access channel has been studied in [35], [7], [36], [37], and bounds on the corresponding capacity regions have been derived.

In cellular and WiFi data networks, mobile users have been demanding increasingly higher data rates on the downlink. This application motivates us to study the downlink or broadcast channel that exploits the techniques of relaying and user cooperation to achieve higher throughput. We introduce and study two such systems of relay broadcast channels (RBCs), where relay links are incorporated into the standard broadcast channel [38], [39] to assist broadcast transmission. These RBC models represent the most fundamental user-cooperative downlink systems and capture the essential roles of user cooperation in downlink communications. We focus on the two-user case of these RBCs, and study the gains in capacity region offered by relaying and user cooperation.

We first study the partially cooperative relay broadcast channel, which is based on the standard two-user broadcast channel with one source attempting to transmit both common and private information to two users. In addition, user 1 acts as a standard relay node [1], [2] and transmits cooperative information to user 2 through a relay link (see Fig. 1). A possible motivation for studying this channel is that in a two-user broadcast system usually one user (denoted by user 1) has a "better" channel from the source than the other user (denoted by user 2), and hence user 1 may decode the information

Manuscript received July 19, 2005; revised December 3, 2006. This work was supported by the National Science Foundation CAREER/PECASE award CCF 0049089 and by a Vodafone Foundation Graduate Fellowship. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Chicago, IL, June/July 2004 and in part at the 2005 International Conference on Wireless Networks, Communications and Mobile ComputingMaui, HI, June 2005.

Y. Liang is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: yingbinl@princeton.edu).

V. V. Veeravalli is with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: vvv@uiuc.edu).

Communicated by K. Kobayashi, Associate Editor for Shannon Theory.

Color versions of Figures 1, 2, 4, 5, 9, 11, and 12 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2006.890726

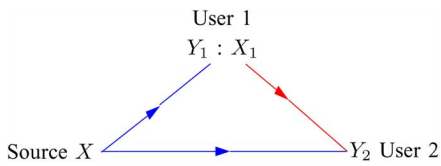


Fig. 1. Partially cooperative RBC.

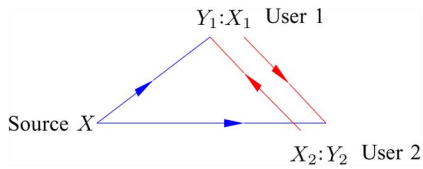


Fig. 2. Fully cooperative RBC.

intended for user 2 in addition to its own information. In this case, user 2 should benefit from having a relay link from user 1. Such user cooperation is particularly useful when a mobile user experiences a deep fading state, where it may be possible to maintain reliable communication only with the help of another user (the relay node).

We further study the partially cooperative RBC with feedback, where, as in the relay channel with feedback [2, Sec. V], the outputs at user 2 are provided to user 1 and the outputs at both users 1 and 2 are provided to the source, all through perfect feedback links. Our motivation for studying this feedback channel is not only because feedback is a natural topic in the study of broadcast channels [40]–[42] and relay channels [2], but also because the study of feedback channels provides useful insights into the type of information that is useful for user cooperation. Furthermore, the proof of the converse for the capacity region of a feedback channel suggests a way to obtain a tighter outer bound than the cut-set bound on the capacity region for the corresponding relay broadcast channel without feedback.

We then move on to study a more general model, the fully cooperative relay broadcast channel, where both users can transmit cooperative information to each other through relay links (see Fig. 2). In this channel, both users can potentially gain in capacity due to this cooperative relaying, which we illustrate via a Gaussian example. We also study the fully cooperative RBC with feedback and derive results that are parallel to those for the partially cooperative RBC.

For both the partially and fully cooperative RBCs, we specialize our results for two Gaussian cases. The first is the degraded case, where the outputs at the two receivers are corrupted by (physically) degraded AWGN (D-AWGN) Gaussian noise terms, i.e., if the random noise terms are denoted by $Z_1$ and $Z_2$, then $Z_2 = Z_1 + Z'$, and $Z_1$ and $Z'$ are independent Gaussian random variables. Such degraded channel models are of information-theoretic interest, because we can usually establish the capacity regions for these models. Furthermore, the proofs of the converse for the capacity regions of these degraded models suggest techniques for obtaining outer bounds on the capacity regions for the corresponding nondegraded channels. The second Gaussian case is the more practically relevant one, where the outputs at the two receivers are corrupted by independent AWGN terms.

We note that a model of the broadcast channel with cooperating receivers was simultaneously studied in [43] (further studied in [44]) with the conference version of this paper [45]. The model studied in [43], [44] is a special case of the fully cooperative RBC being studied in this paper in that the relay links for user cooperation are orthogonal to the original broadcast channel. We also note that a related relay broadcast channel model, where an additional relay node is introduced to broadcast systems to assist all users, has been introduced and studied in [7], [45], [46].

We now summarize the main results of this paper. For the discrete memoryless partially and fully cooperative RBCs, we derive inner and outer bounds on the capacity regions, and we show that the outer bounds that we derive are tighter than the cut-set bounds. We then establish the capacity regions for the degraded partially and fully cooperative RBCs, where the previous inner and outer bounds match. We further establish the capacity regions for the partially and fully cooperative RBCs with feedback. We show that feedback does not increase the capacity regions for these RBCs, but that it improves the capacity regions for their nondegraded counterparts.

For the Gaussian partially and fully cooperative RBCs, we show that the D-AWGN partially and fully cooperative RBCs (with/without feedback) have the same capacity region. However, we show that the achievable rate region of the AWGN fully cooperative RBC is larger than the achievable rate region of the AWGN partially cooperative RBC. In particular, we show that the outer bounds on the capacity regions of the AWGN partially and fully cooperative RBCs are contained in the capacity regions of the corresponding channels with feedback. This indicates that feedback improves the capacity regions for the AWGN partially and fully cooperative RBCs.

We also summarize our main results in Tables I and II for easy reference. The notation in the two tables (also throughout the paper) is described as follows. We use $\mathcal{C}$ to denote the capacity region. The subscripts "$P$" and "$F$" denote partially and fully cooperative RBCs, respectively. The subscripts "$PF$" and "$FF$" denote partially and fully cooperative RBCs with feedback, respectively. The subscript "$D$" denotes degraded cooperative RBCs, and "$G$" denotes AWGN cooperative RBCs. Hence, the subscript "$DG$" denotes D-AWGN cooperative RBCs. We use $\mathcal{C}^{(i)}$ to denote the inner bound and use $\mathcal{C}^{(o)}$ to denote the outer bound.

The notation in this paper mainly follows the following rules. Upper case letters indicate random variables, and lower case letters indicate deterministic variables or realizations of the corresponding random variables. There are some exceptions, but these will be clarified where they appear in the paper. We use $\underline{x}$ or $x^n$ to indicate the vector $(x_1, \ldots, x_n)$, and use $x_i^n$ to indicate the vector $(x_i, \ldots, x_n)$. We define two functions: $\bar{x} := 1 - x$ and $\mathcal{C}(x) := \frac{1}{2}\log(1+x)$. We adopt the notation in [47] and use $H(X)$ to denote the entropy of $X$, $h(X)$ to denote the differential entropy of $X$, and $I(X;Y)$ to denote the mutual information between $X$ and $Y$. Throughout the paper, the logarithmic function is to the base 2.

In Sections II–VII, we first present the results for the partially cooperative RBC and the results for this channel with feedback. We then present the results for the fully cooperative RBC and the results for this channel with feedback. We finally discuss and compare the achievable regions for the AWGN case of these

TABLE I
RESULTS ON DISCRETE MEMORYLESS PARTIALLY AND FULLY COOPERATIVE RBCS

| | Partially Cooperative | | Fully Cooperative | |
|---|---|---|---|---|
| | No Feedback | Feedback | No Feedback | Feedback |
| (General) RBC | $\mathcal{C}_P^{(i)}$ (Th. 1)<br>$\mathcal{C}_P^{(o)}$ (Th. 2) | $\mathcal{C}_{PF}$ (Th. 6) | $\mathcal{C}_F^{(i1)}$ (Th. 9)<br>$\subset \mathcal{C}_F^{(i2)}$ (Th. 10)<br>$\mathcal{C}_F^{(o)}$ (Th. 11) | $\mathcal{C}_{FF}$ (Th. 16) |
| (Degraded) RBC | $\mathcal{C}_{D,P}$ (Th. 3) | $\mathcal{C}_{D,PF}$ (Cor. 2)<br>$(= \mathcal{C}_{D,P})$ | $\mathcal{C}_{D,F}$ (Th. 12) | $\mathcal{C}_{D,FF}$ (Cor. 5)<br>$(= \mathcal{C}_{D,F})$ |

TABLE II
RESULTS ON GAUSSIAN PARTIALLY AND FULLY COOPERATIVE RBCS

| | Partially Cooperative | | Fully Cooperative | |
|---|---|---|---|---|
| | No Feedback | Feedback | No Feedback | Feedback |
| (General) AWGN RBC | $\mathcal{C}_{G,P}^{(i)}$ (Cor. 1)<br>$(= \mathcal{C}_{DG,P})$<br>$\mathcal{C}_{G,P}^{(o)}$ (Th. 5) | $\left(\mathcal{C}_{G,P}^{(o)} \subset\right)$<br>$\mathcal{C}_{G,PF}$ (Th. 8) | $\left(\mathcal{C}_{G,P}^{(i)} \subset\right)$<br>$\mathcal{C}_{G,F}^{(i)}$ (Th. 14)<br>$\mathcal{C}_{G,F}^{(o)}$ (Th. 15) | $\mathcal{C}_{G,FF}$ (Th. 18)<br>$(= \mathcal{C}_{G,PF})$ |
| (Degraded) AWGN RBC | $\mathcal{C}_{DG,P}$ (Th. 4) | $\mathcal{C}_{DG,PF}$ (Th. 7)<br>$(= \mathcal{C}_{DG,P})$ | $\mathcal{C}_{DG,F}$ (Th. 13)<br>$(= \mathcal{C}_{DG,P})$ | $\mathcal{C}_{DG,FF}$ (Th. 17)<br>$(= \mathcal{C}_{DG,F})$ |

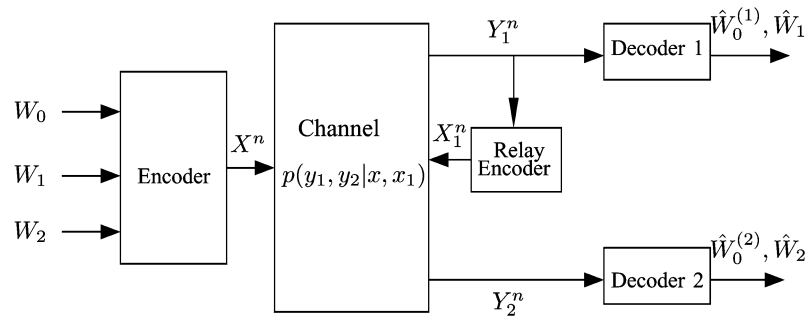


Fig. 3. Partially cooperative RBC.

RBCs with all nodes being subject to a total power constraint. We end the paper with some concluding remarks.

## II. PARTIALLY COOPERATIVE RELAY BROADCAST CHANNELS

In this section, we first introduce the channel model for the partially cooperative RBC, and then present the main results. We further illustrate the results via two Gaussian channel examples.

### A. System Model

*Definition 1:* A partially cooperative RBC (see Fig. 3) consists of three messages $(W_0, W_1, W_2)$, a source input $X \in \mathcal{X}$ with $\mathcal{X}$ being a finite alphabet set, a relay input $X_1 \in \mathcal{X}_1$ with $\mathcal{X}_1$ being a finite alphabet set, two channel outputs $Y_1 \in \mathcal{Y}_1$ and $Y_2 \in \mathcal{Y}_2$ with $\mathcal{Y}_1$ and $\mathcal{Y}_2$ being two finite alphabet sets, and two message estimate pairs $(\hat{W}_0^{(1)}, \hat{W}_1)$ and $(\hat{W}_0^{(2)}, \hat{W}_2)$. The channel transition probability distribution is $p(y_1, y_2|x, x_1)$.

We note that in the above definition, $W_0$ indicates the common message that needs to be decoded at both users, and $W_1$ and $W_2$ are private messages that need to be decoded at users 1 and 2, respectively. We also note that the channel input–output relationship is similar to that of the relay channel [2], but now the relay node (user 1) also has its own message to decode. We assume throughout the paper that the channel is memoryless.

*Definition 2:* A $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n\right)$ code for a partially cooperative RBC consists of the following.

- Three message sets: $\mathcal{W}_0 = \{1, 2, \ldots, 2^{nR_0}\}$, $\mathcal{W}_1 = \{1, 2, \ldots, 2^{nR_1}\}$, and $\mathcal{W}_2 = \{1, 2, \ldots, 2^{nR_2}\}$.
- An encoder: $\mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{X}^n$, which maps each message tuple $(W_0, W_1, W_2) \in \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2$ to a codeword $x^n \in \mathcal{X}^n$.
- A set of relay functions $\{f_i\}_{i=1}^n$ such that
$$x_{1,i} = f_i(y_{1,1}, \ldots, y_{1,i-1}), \qquad 1 \le i \le n.$$
- Two decoders: one at user 1, $\mathcal{Y}_1^n \to \mathcal{W}_0 \times \mathcal{W}_1$, which maps a received sequence $y_1^n$ to a message pair $(\hat{W}_0^{(1)}, \hat{W}_1) \in \mathcal{W}_0 \times \mathcal{W}_1$; and the other at user 2, $\mathcal{Y}_2^n \to \mathcal{W}_0 \times \mathcal{W}_2$, which maps $y_2^n$ to a message pair $(\hat{W}_0^{(2)}, \hat{W}_2) \in \mathcal{W}_0 \times \mathcal{W}_2$.

The probability of error when the message tuple $(W_0, W_1, W_2)$ is sent is defined as

$$\begin{aligned} P_e^{(n)}&(W_0, W_1, W_2) \\ &= Pr\left(\left(\hat{W}_0^{(1)}, \hat{W}_1\right) \neq (W_0, W_1)\right. \\ &\qquad \left. \text{or } \left(\hat{W}_0^{(2)}, \hat{W}_2\right) \neq (W_0, W_2)\right) \end{aligned} \tag{1}$$

and the average probability of error is defined by assuming that the message $(W_0, W_1, W_2)$ is uniformly distributed over $\mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2$ and is given by

$$P_e^{(n)} = \frac{1}{2^{nR_0}2^{nR_1}2^{nR_2}} \sum_{W_0=1}^{2^{nR_0}} \sum_{W_1=1}^{2^{nR_1}} \sum_{W_2=1}^{2^{nR_2}} P_e^{(n)}(W_0, W_1, W_2). \tag{2}$$

The rate tuple $(R_0, R_1, R_2)$ is said to be achievable for the partially cooperative RBC if there exists a sequence of $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n\right)$ codes with average error probability $P_e^{(n)} \to 0$ as $n$ goes to infinity.

*Definition 3:* A partially cooperative RBC is *degraded* if the channel satisfies

$$p(y_1, y_2|x, x_1) = p(y_1|x, x_1)p(y_2|y_1, x_1). \tag{3}$$

### B. Discrete Memoryless Partially Cooperative RBCs

A motivation for the study of the partially cooperative RBC is that in many cases one user in the broadcast channel has the capability to decode at higher rate than the other user. Hence, this user may also decode the message for the other user in addition to the message for itself, and then forward this information to the other user. The following achievable rate region for the partially cooperative RBC is based on this idea, where the relay (user 1) employs the decode-and-forward relaying scheme [2, Sec. II] to help user 2.

*Theorem 1:* An inner bound on the capacity region of the partially cooperative RBC is given by

$$\mathcal{C}_P^{(i)} = \bigcup_{p(u,x_1,x)p(y_1,y_2|x,x_1)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 < \min\{I(U, X_1; Y_2), I(U; Y_1 \mid X_1)\} \\ R_1 < I(X; Y_1 \mid U, X_1) \end{array} \right\} \tag{4}$$

where the auxiliary random variable $U$ is bounded in cardinality by $|\mathcal{U}| \leq |\mathcal{X}| \cdot |\mathcal{X}_1| + 2$.

*Proof:* See Appendix I for an outline of the proof. □

The achievable region based on the decode-and-forward scheme in Theorem 1 serves as an example to show that relaying from user 1 to user 2 indeed helps to enlarge the capacity region of the original broadcast channel. This will be further demonstrated by Gaussian channels later. More importantly, we will show that this achievable region gives the capacity region of the degraded RBC, and we hence derive the capacity regions of the degraded Gaussian channel and feedback channel. Even for the nondegraded Gaussian channel, we will show that this achievable rate region is close to the outer bound on the capacity region.

Other achievable regions can also be derived based on the relay node (user 1) using other relaying schemes to assist user 2, for example, the estimate-and-forward scheme [2, Sec. IV], the amplify-and-forward scheme (e.g., see [21], [20]), or combinations of these schemes. The derivations of these achievable rate regions follow steps that are similar to those used in deriving the achievable rates based on these relaying schemes for the three-terminal relay channel as in [2], [7], [21], [20]. Which scheme results in the largest achievable rate region depends on the particular channel of interest. In general, none of these schemes provides the capacity region.

For the general discrete memoryless partially cooperative RBC, we provide the following outer bound on the capacity region.

*Theorem 2:* An outer bound on the capacity region of the partially cooperative RBC is given by

$$\mathcal{C}_P^{(o)} = \bigcup_{p(u,u',x_1,x)p(y_1,y_2|x,x_1)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 \leq \min\{I(U, X_1; Y_2), I(U; Y_1, Y_2 \mid X_1)\} \\ R_1 \leq I(X; Y_1, Y_2 \mid U, X_1) \\ R_0 + R_1 \leq I\left(U'; Y_1|X_1\right) \\ R_2 \leq I\left(X; Y_1, Y_2|U', X_1\right) \end{array} \right\} \tag{5}$$

where the random variables satisfy two Markov chain conditions: $X_1 \to U \to X$ and $X_1 \to U' \to X$. The auxiliary random variables $U$ and $U'$ are bounded in cardinality by $|\mathcal{U}| \leq |\mathcal{X}| \cdot |\mathcal{X}_1| + 2$ and $|\mathcal{U}'| \leq |\mathcal{X}| \cdot |\mathcal{X}_1| + 2$, respectively.

*Proof:* See Appendix II. □

*Remark 1:* The outer bound given in Theorem 2 is tighter than the outer bound based on the general max-flow min-cut theorem [47, Theorem 14.10.1] (the cut-set bound).

To compare the two outer bounds, we first write the cut-set bound as follows:

$$R_0 + R_1 \leq I(X; Y_1|X_1) \tag{6a}$$
$$R_0 + R_2 \leq I(X, X_1; Y_2) \tag{6b}$$
$$R_0 + R_1 + R_2 \leq I(X; Y_1, Y_2|X_1) \tag{6c}$$

for some joint distribution $p(x, x_1)p(y_1, y_2|x, x_1)$.

We now show that the outer bound given in (5) in Theorem 2 is contained in (tighter than) the cut-set bound given in (6).

The third bound in (5) implies (6a) because

$$\begin{aligned} R_0 + R_1 &\leq I\left(U'; Y_1|X_1\right) \leq I\left(U', X; Y_1|X_1\right) \\ &= I(X; Y_1|X_1). \end{aligned} \tag{7}$$

The first bound in (5) implies (6b) because

$$\begin{aligned} R_0 + R_2 &\leq I(U, X_1; Y_2) \leq I(U, X, X_1; Y_2) \\ &= I(X, X_1; Y_2). \end{aligned} \tag{8}$$

Finally, the first and second bounds in (5) imply (6c) because

$$\begin{aligned} R_0 + R_1 + R_2 &\leq I(U; Y_1, Y_2|X_1) + I(X; Y_1, Y_2|U, X_1) \\ &= I(U, X; Y_1, Y_2|X_1) = I(X; Y_1, Y_2|X_1) \end{aligned} \tag{9}$$

where we have used the Markov property: $U \to (X_1, X) \to (Y_1, Y_2)$ in the last equality. Therefore, the outer bound given

in Theorem 2 is at least as tight as the cut-set bound. In fact, one can find an example partially cooperative RBC (e.g., the D-AWGN partially cooperative RBC we study in Section II-C) for which the outer bound given in Theorem 2 is strictly contained in the cut-set bound, and is hence tighter than the cut-set bound.

The inner and outer bounds given in Theorems 1 and 2 may not be tight for the general partially cooperative RBC. However, for the degraded partially cooperative RBC, which satisfies the condition given in Definition 3, we immediately see that the outer bound (bounds on $R_0 + R_2$ and $R_1$) given in Theorem 2 reduces to a form that matches the inner bound given in Theorem 1. Thus, we have the following capacity region.

*Theorem 3:* The capacity region of the degraded partially cooperative RBC that satisfies the condition given in Definition 3 is given by

$$\mathcal{C}_{D,P} = \bigcup_{p(u,x_1)p(x|u)p(y_1,y_2|x,x_1)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 \leq \min\{I(U, X_1; Y_2), I(U; Y_1 \mid X_1)\} \\ R_1 \leq I(X; Y_1 \mid U, X_1) \end{array} \right\} \tag{10}$$

where $U$ is bounded in cardinality by $|\mathcal{U}| \leq |\mathcal{X}| \cdot |\mathcal{X}_1| + 2$.

*Proof:* The achievability is given by Theorem 1. To show the converse, we apply Theorem 2.

The first bound in (5) implies the first bound in (10) because

$$\begin{aligned} R_0 + R_2 &\leq I(U; Y_1, Y_2|X_1) \\ &= I(U; Y_1|X_1) + I(U; Y_2|X_1, Y_1) \\ &= I(U; Y_1|X_1) \end{aligned} \tag{11}$$

where $I(U; Y_2|X_1, Y_1) = 0$ follows from the degradedness condition given in Definition 3.

The second bound in (5) implies the second bound in (10) because

$$\begin{aligned} R_1 &\leq I(X; Y_1, Y_2|U, X_1) \\ &\leq I(X; Y_1|U, X_1) + I(X; Y_2|U, X_1, Y_1) \\ &= I(X; Y_1|U, X_1) \end{aligned} \tag{12}$$

where $I(X; Y_2|U, X_1, Y_1) = 0$ also follows from the degradedness condition given in Definition 3. This concludes the proof of the converse. □

*Remark 2:* The bounds on $R_0 + R_1$ and on $R_2$ in (5) are not necessary for the degraded channel. They are still useful for the nondegraded channels as we will demonstrate in Section II-C via the AWGN partially cooperative RBC.

### C. Gaussian Partially Cooperative RBCs

We study two Gaussian partially cooperative RBCs, where the outputs at the two users are corrupted by AWGNs.

We first define what it means for one Gaussian noise variable to be degraded with respect to another, and then define the two Gaussian partially cooperative RBCs with degraded noise terms and independent noise terms, respectively.

*Definition 4:* The Gaussian random variable $Z_2$ is (physically) degraded with respect to the Gaussian random variable $Z_1$ if $Z_2$ can be expressed as $Z_2 = Z_1 + Z'$, where $Z'$ is a Gaussian random variable that is independent of $Z_1$.

*Definition 5:* The D-AWGN partially cooperative RBC is a partially cooperative RBC with the channel outputs being corrupted by degraded Gaussian noise terms, i.e., the channel outputs at the two users are given by

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y_2 &= X + X_1 + Z_1 + Z' \end{aligned} \tag{13}$$

where $Z_1$ and $Z'$ are independent real Gaussian random variables with variances $N_1$ and $N_2 - N_1$, respectively, where $N_1 < N_2$. The channel input sequences $\{X_n\}$ and $\{X_{1,n}\}$ are subject to the average power constraints $P$ and $P_1$, respectively, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[X_i^2\right] \leq P \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[X_{1,i}^2\right] \leq P_1. \tag{14}$$

*Definition 6:* The AWGN partially cooperative RBC is the partially cooperative RBC with the channel outputs being corrupted by independent Gaussian noise terms, i.e., the channel outputs at the two users are given by

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y_2 &= X + X_1 + Z_2 \end{aligned} \tag{15}$$

where $Z_1$ and $Z_2$ are independent real Gaussian random variables with variances $N_1$ and $N_2$, respectively. The channel input sequences $\{X_n\}$ and $\{X_{1,n}\}$ are subject to the power constraints given in (14).

Note that the D-AWGN partially cooperative RBC is degraded (satisfies the condition given in Definition 3) due to the degraded Gaussian noise terms at the two outputs. For the D-AWGN partially cooperative RBC, we have the following theorem for the capacity region.

*Theorem 4:* The capacity region of the D-AWGN partially cooperative RBC is given by (16) at the bottom of the page, where $\bar{\alpha} := 1 - \alpha$, $\bar{\beta} := 1 - \beta$, and $\mathcal{C}(x) := \frac{1}{2}\log(1 + x)$, as defined at the end of Section I.

$$\mathcal{C}_{DG,P} = \bigcup_{0 \leq \alpha \leq 1} \left\{ \begin{array}{l} (R_0, R_1, R_2): R_0 + R_2 \leq \max\limits_{0 \leq \beta \leq 1} \min\left\{ \mathcal{C}\left(\frac{P_1 + \bar{\alpha}P + 2\sqrt{\beta\bar{\alpha}PP_1}}{\alpha P + N_2}\right), \mathcal{C}\left(\frac{\beta\bar{\alpha}P}{\alpha P + N_1}\right) \right\} \\ R_1 \leq \mathcal{C}\left(\frac{\alpha P}{N_1}\right) \end{array} \right\} \tag{16}$$

In (16), the parameter $\alpha$ indicates the fraction of source power that is used to transmit information intended for user 1, and $\beta$ is the correlation coefficient between the source and relay signals.

*Proof:* The proof of the achievability follows by evaluating the mutual information terms in Theorem 1 using the following input distributions: $X_1 \sim N(0, P_1)$, $U' \sim N(0, \beta\bar{\alpha}P)$, $X' \sim N(0, \alpha P)$, where $X_1, U', X'$ are independent. Furthermore, we let $U = \sqrt{\frac{\bar{\beta}\bar{\alpha}P}{P_1}}X_1 + U'$ and $X = U + X'$.

The proof of the converse follows directly from the proof of the converse for the D-AWGN partially cooperative RBC with feedback (proof given in Appendix IV), because the feedback channel provides an outer bound on the capacity region for the original channel without feedback. □

We now study the property of the boundary of the capacity region of the D-AWGN partially cooperative RBC. In the following discussion, we let $R_0 = 0$ for convenience. In (16), the optimization over $\beta$ can be evaluated by considering the following two cases.

**Case 1:** If $\frac{P_1}{N_2-N_1} \geq \frac{P}{N_1}$, then $\beta = 1$ achieves the maximum in (16) for any $\alpha \in [0, 1]$, and the capacity region is defined by the rate pairs $(R_1, R_2)$ that satisfy

$$R_1 \leq \mathcal{C}\left(\frac{\alpha P}{N_1}\right), \quad R_2 \leq \mathcal{C}\left(\frac{\bar{\alpha}P}{\alpha P + N_1}\right). \tag{17}$$

Note that in this case $R_1 + R_2 \leq \mathcal{C}\left(\frac{P}{N_1}\right)$, and hence, the boundary of the capacity region is a straight line. The capacity region for the D-AWGN partially cooperative RBC hence, coincides with the capacity region of a broadcast channel where the two users have symmetric channels (both have noise levels $N_1$). This means that if the relay power is large enough, user 2 effectively sees the same level of noise as user 1 due to relaying. Also note that the value $\frac{P(N_2-N_1)}{N_1}$ is a threshold on $P_1$ beyond which the capacity region will not be further enhanced by relaying.

**Case 2:** If $\frac{P_1}{N_2-N_1} < \frac{P}{N_1}$, then define

$$\alpha_0 := \frac{\frac{P}{N_1} - \frac{P_1}{N_2-N_1}}{\frac{P}{N_1}\frac{P_1}{N_2-N_1} + \frac{P}{N_1}}.$$

The optimizing $\beta$ will depend on the value of $\alpha$ compared to $\alpha_0$.

(i) If $\alpha \geq \alpha_0$, then $\beta = 1$ achieves the maximum in (16) and again the rate pair given in (17) defines one part of the boundary of the capacity region corresponding to $R_1 \geq \mathcal{C}(\frac{\alpha_0 P}{N_1})$. It is clear that this part of the boundary is straight line.

(ii) If $0 \leq \alpha < \alpha_0$, then $\beta^*$ that achieves the maximum satisfies the following equation:

$$\mathcal{C}\left(\frac{P_1 + \bar{\alpha}P + 2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P + N_2}\right) = \mathcal{C}\left(\frac{\beta\bar{\alpha}P}{\alpha P + N_1}\right). \tag{18}$$

Hence, the other part of the boundary of the capacity region corresponding to $0 \leq R_1 \leq (\frac{\alpha_0 P}{N_1})$ is defined by the rate pairs $(R_1, R_2)$ that satisfy

$$R_1 \leq \mathcal{C}\left(\frac{\alpha P}{N_1}\right), \quad R_2 \leq \mathcal{C}\left(\frac{\beta^*\bar{\alpha}P}{\alpha P + N_1}\right). \tag{19}$$

We summarize the properties of the boundary of the capacity region in the following proposition.

*Proposition 1:* If $\frac{P_1}{N_2-N_1} \geq \frac{P}{N_1}$, the boundary of the capacity region of the D-AWGN partially cooperative RBC is a straight line defined by (17). If $\frac{P_1}{N_2-N_1} < \frac{P}{N_1}$, the boundary of the capacity region of the D-AWGN partially cooperative RBC consists of one straight-line segment defined by (17), where $R_1 \geq \mathcal{C}(\frac{\alpha_0 P}{N_1})$; and one curved segment defined by (19), where $0 \leq R_1 \leq (\frac{\alpha_0 P}{N_1})$.

We now compare the capacity region for the D-AWGN partially cooperative RBC with the capacity region for the Gaussian broadcast channel without user cooperation. The capacity region of the latter channel is given by [47, Ch. 14.6]

$$R_1 \leq \mathcal{C}\left(\frac{\alpha P}{N_1}\right), \quad R_2 \leq \mathcal{C}\left(\frac{\bar{\alpha}P}{\alpha P + N_2}\right) \tag{20}$$

for some $\alpha \in [0, 1]$. In Fig. 4, we plot this region with the dashed curve as its boundary. We also plot the capacity regions (boundaries with solid lines) for the D-AWGN partially cooperative RBC under different relay signal-to-noise ratios (SNRs) $\frac{P_1}{N_2}$. Note that for simplicity, we only plot the region for the case where $R_0 = 0$. It is clear from the figure that the D-AWGN partially cooperative RBC has a larger capacity region, and the improvement becomes more significant as $\frac{P_1}{N_2}$ increases. However, as suggested by discussion under Case 1, no further improvement is possible for values of $\frac{P_1}{N_2}$ greater than 14.54 dB. Thus, the outer most solid line with $\frac{P_1}{N_2} = 15$ dB defines the best capacity region.

We now consider the AWGN partially cooperative RBC defined in Definition 6, where the noise terms at the two outputs are independent. This channel does not satisfy the degradedness condition given in Definition 3. We consider the case where $N_1 < N_2$ in the following analysis, and only comment on the case where $N_1 \geq N_2$ at the end of this section. An achievable rate region for this channel can be derived, which is exactly the same as the capacity region for the D-AWGN partially cooperative RBC.

*Corollary 1:* An inner bound $\mathcal{C}_{G,P}^{(i)}$ on the capacity region of the AWGN partially cooperative RBC is given by the capacity region $\mathcal{C}_{DG,P}$ of the D-AWGN partially cooperative RBC given in Theorem 4, i.e.,

$$\mathcal{C}_{G,P}^{(i)} = \mathcal{C}_{DG,P}. \tag{21}$$

Hence, the boundary of the region $\mathcal{C}_{G,P}^{(i)}$ has the properties described in Proposition 1.

The proof follows the steps that are the same as those in the achievability proof for Theorem 4.

The achievable region given in Corollary 1 may not be a tight inner bound on the capacity region for the AWGN partially cooperative RBC. In the following, we further provide an outer bound on the capacity region for this channel.

*Theorem 5:* An outer bound on the capacity region of the AWGN partially cooperative RBC is given by (22) at the bottom

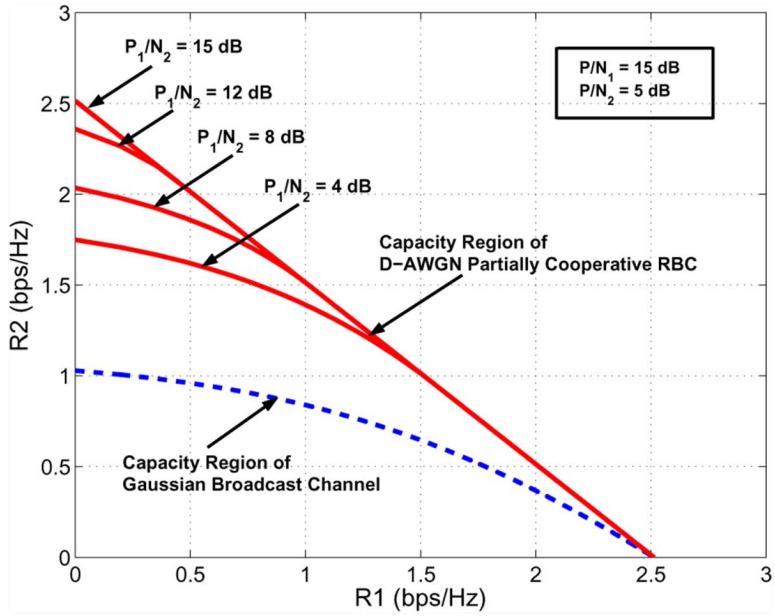


Fig. 4. Comparison of the capacity regions for the Gaussian broadcast channel and the D-AWGN partially cooperative RBCs.

of the page, where $\bar{\alpha} := 1-\alpha$, $\bar{\beta} := 1-\beta$, and $\mathcal{C}(x) := \frac{1}{2}\log(1+x)$ as defined at the end of Section I.

*Proof:* The proof is relegated to Appendix V since part of the proof is similar to the proof of Theorem 7 given in Appendix IV. □

In Fig. 5, we plot the inner bound on the capacity region (achievable region) of the AWGN partially cooperative RBC with dot-dashed line as its boundary and the outer bound on the capacity region with the solid line as its boundary. We plot the region for the case where $R_0 = 0$ for simplicity. It is clear from the figure that the outer and inner bounds are very close. The gap between the two bounds varies with the particular SNRs chosen for the transmission links in the system. In Fig. 5, we also plot the capacity region of the original broadcast channel with dashed line as its boundary. It is clear that the AWGN partially cooperative RBC has a significantly larger capacity region than the broadcast channel.

*Remark 3:* For the AWGN partially cooperative RBC, we have restricted our attention to the channel where $P/N_1 > P/N_2$, i.e., the channel from the source to the relay (user 1) is better than the channel from the source to user 2. This is a case for which it is reasonable to introduce a relay transmission from user 1 to user 2. Nevertheless, even if the relay (user 1) has a weaker channel from the source than user 2, i.e., $P/N_1 < P/N_2$, it can still assist user 2. However, under this condition, the relay needs to use schemes other than the decode-and-forward scheme. For example, the relay can employ the estimate-and-forward scheme [2, Sec. VI] to assist user 2, and an achievable rate region based on this scheme is given as follows:

$$\bigcup_{0\le\alpha,\eta\le 1}\left\{\begin{array}{l}(R_0,R_1,R_2):\\ R_0+R_1<\min\left\{\mathcal{C}\left(\frac{\bar{\alpha}P}{\alpha P+N_1}\right),\mathcal{C}\left(\frac{\bar{\alpha}P}{\alpha P+\eta P_1+N_2}\right)\right\}\\ R_2<\mathcal{C}\left(\frac{\alpha P}{N_2}+\frac{\alpha\eta PP_1}{\eta P_1N_1+\alpha P(N_1+N_2)+N_1N_2}\right)\end{array}\right\}. \tag{23}$$

It is clear that the above achievable region is larger than the capacity region of the original broadcast channel. This achievable rate region can be viewed as a special case of the region given in Theorem 14, with the roles of user 1 and user 2 being switched, and the corresponding notations for rates and noise variances also being switched.

## III. Partially Cooperative RBCs With Feedback

In this section, we study the partially cooperative RBC with feedback, where the outputs at user 2 are provided to user 1 and the outputs at both users 1 and 2 are provided to the source all through perfect feedback links (see Fig. 6). Note that this definition for feedback channel follows the definition for the relay channel with feedback [2, Sec. V]. We will show that feedback

$$\mathcal{C}_{G,P}^{(o)} = \bigcup_{0\le\alpha,\beta\le 1}\left\{\begin{array}{l}(R_0,R_1,R_2): R_0+R_2\le\min\left\{\mathcal{C}\left(\frac{P_1+\bar{\alpha}P+2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P+N_2}\right),\mathcal{C}\left(\frac{\beta\bar{\alpha}P}{\alpha P+\frac{N_1N_2}{N_1+N_2}}\right)\right\}\\ R_1\le\mathcal{C}\left(\frac{\alpha P}{\frac{N_1N_2}{N_1+N_2}}\right)\\ R_0+R_1\le\mathcal{C}\left(\frac{\alpha P+\beta\bar{\alpha}P}{N_1}\right)\end{array}\right\} \tag{22}$$

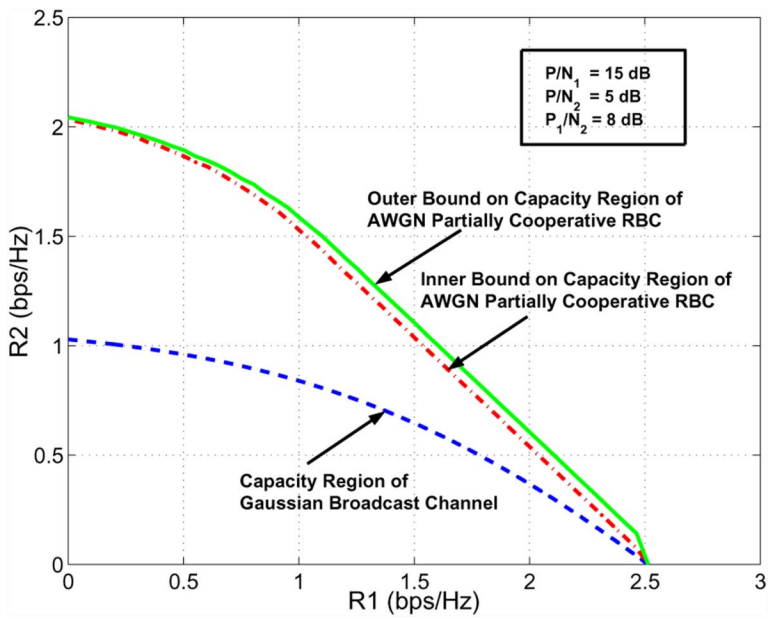


Fig. 5. Inner and outer bounds on the capacity region of the AWGN partially cooperative RBC.

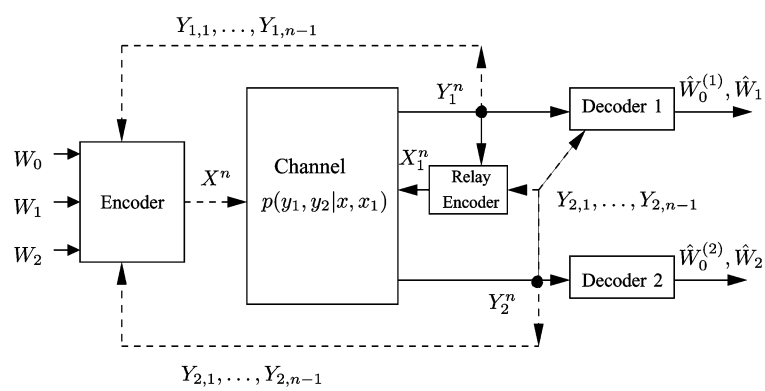


Fig. 6. Partially cooperative RBC with feedback.

in general may improve the capacity region for the partially cooperative RBCs, but does not affect the capacity region for the degraded partially cooperative RBCs.

For a distribution on the message set $p(w_0, w_1, w_2)$, the following joint distribution is induced for a partially cooperative RBC with feedback:

$$\begin{aligned} &p\left(w_0, w_1, w_2, x^n, x_1^n, y_1^n, y_2^n\right) \\ &\quad = p(w_0, w_1, w_2) \\ &\qquad \cdot \prod_{i=1}^{n} \Big[ p\left(x_i | w_0, w_1, w_2, y_1^{i-1}, y_2^{i-1}\right) \\ &\qquad\qquad \cdot p\left(x_{1,i} | y_1^{i-1}, y_2^{i-1}\right) p(y_{1,i}, y_{2,i} | x_i, x_{1,i}) \Big]. \end{aligned} \tag{24}$$

### A. Discrete Memoryless Partially Cooperative RBCs With Feedback

From the definition for the degraded partially cooperative RBC given in Definition 3, it is clear that the partially cooperative RBC with feedback is degraded. For this channel, we have the following capacity theorem.

*Theorem 6:* The capacity region of the partially cooperative RBC with feedback is given by

$$\mathcal{C}_{PF} = \bigcup_{p(x_1)p(u|x_1)p(x|u)p(y_1,y_2|x,x_1)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 \le \min\{I(U, X_1; Y_2), I(U; Y_1, Y_2 \mid X_1)\} \\ R_1 \le I(X; Y_1, Y_2 \mid U, X_1) \end{array} \right\} \tag{25}$$

where $U$ is bounded in cardinality by $|\mathcal{U}| \le |\mathcal{X}| \cdot |\mathcal{X}_1| + 2$.

*Proof:* Theorem 1 provides the achievability with $(Y_1, Y_2)$ replacing $Y_1$ as the output at user 1. The proof of the converse follows steps that are similar to those in the proof for Theorem 2, and is hence only outlined in Appendix III. □

*Remark 4:* From the above achievability proof, it is clear that in the capacity region achieving scheme, the source need not exploit the feedback information from the two users. However,

the relay (user 1) makes use of the feedback information from user 2 to improve on its decoding and relaying.

It is clear that if the original partially cooperative RBC is degraded, then the capacity region given in Theorem 6 is the same as the capacity region for the original channel without feedback.

*Corollary 2:* Feedback does not increase the capacity region for the degraded partially cooperative RBC, i.e.,

$$\mathcal{C}_{D,PF} = \mathcal{C}_{D,P}$$

where $\mathcal{C}_{D,PF}$ denotes the capacity region of the degraded partially cooperative RBC with feedback, and $\mathcal{C}_{D,P}$ is given in Theorem 3.

This result is intuitive. Since the original channel is degraded, the output $y_2$ at user 2 does not provide user 1 with more information other than the information already contained in the output $y_1$ at user 1. Hence, feedback of $y_2$ to user 1 does not help. The reason that feedback of $y_1$ and $y_2$ to the source does not help follows from the result in [40] that feedback does not increase the capacity for the physically degraded broadcast channel.

### B. Gaussian Partially Cooperative RBCs With Feedback

In this subsection, we consider two Gaussian feedback channels: the D-AWGN and AWGN partially cooperative RBCs with feedback. We study how feedback affects the capacity regions of these Gaussian channels.

For the D-AWGN partially cooperative RBC with feedback, we have the following theorem, which is consistent with the result given in Corollary 2 for the discrete memoryless channel.

*Theorem 7:* Feedback does not increase the capacity region for the D-AWGN partially cooperative RBC, i.e., $\mathcal{C}_{DG,PF} = \mathcal{C}_{DG,P}$, where $\mathcal{C}_{DG,PF}$ denotes the capacity region of the D-AWGN partially cooperative RBC with feedback, and $\mathcal{C}_{DG,P}$ is given in Theorem 4.

*Proof:* The achievability proof is the same as that for the D-AWGN partially cooperative RBC. The proof for the converse is provided in Appendix IV. □

We note that in the definition for the D-AWGN partially cooperative RBC, the output $y_1$ at user 1 is not affected by the input signal $x_1$ transmitted by user 1 to user 2. In practice, $x_1$ may cause interference to $y_1$. We hence define the following self-interfered D-AWGN partially cooperative RBC model:

$$\begin{aligned} Y_1 &= X + aX_1 + Z_1 \\ Y_2 &= X + X_1 + Z_1 + Z' \end{aligned} \tag{26}$$

where $a$ is a real constant number indicating how strong the interference is, and $Z_1$ and $Z'$ are independent real Gaussian random variables with variances $N_1$ and $N_2 - N_1$, respectively.

The self-interfered D-AWGN partially cooperative RBC is degraded. It is also easy to check that the self-interference does not affect the capacity region of the D-AWGN partially cooperative RBC with/without feedback. This result is summarized in the following corollary, and it will be useful in proving the capacity region for the AWGN partially cooperative RBC with feedback given in Theorem 8.

*Corollary 3:* The capacity region for the self-interfered D-AWGN partially cooperative RBC with/without feedback is the same as the capacity region of the D-AWGN partially cooperative RBC with/without feedback, and is given in Theorem 4.

We now consider the second Gaussian feedback channel: the AWGN partially cooperative RBC with feedback. For this channel, we have the following capacity theorem.

*Theorem 8:* The capacity region of the AWGN partially cooperative RBC with feedback is given by (27) at the bottom of the page.

*Proof:* The idea of the proof is to follow the argument in [42, Ch. 3.2.2] to change the AWGN partially cooperative RBC with feedback to an equivalent D-AWGN partially cooperative RBC with feedback, with $N_1$ being replaced by $\frac{N_1 N_2}{N_1+N_2}$.

We first define

$$S := \frac{N_1 Y_2 + N_2 Y_1}{N_1 + N_2} \tag{28}$$

and note that the mapping from $(Y_1, Y_2)$ to $(S, Y_2)$ is one-to-one. Hence, the channel with the outputs being $(Y_1, Y_2)$ and $Y_2$ is equivalent to the channel with the outputs being $(S, Y_2)$ and $Y_2$.

We now want to show that given $(S, X_1)$, $Y_2$ is independent of $X$, i.e., $Y_2$ is a degraded version of $S$. We express $S$ in the following form:

$$S = X + \frac{N_1}{N_1 + N_2} X_1 + \hat{Z}_1 \tag{29}$$

where $\hat{Z}_1 := \frac{N_1 Z_2 + N_2 Z_1}{N_1+N_2}$. We can express $Y_2$ as

$$Y_2 = S + \frac{N_2}{N_1 + N_2} X_1 + \hat{Z} \tag{30}$$

where $\hat{Z} := \frac{N_2}{N_1+N_2}(Z_2 - Z_1)$.

It is clear that $\hat{Z}$ is independent of $X$ and $X_1$, and it is easy to check that $\hat{Z}$ is independent of $\hat{Z}_1$. Hence, $\hat{Z}$ is independent of $S$. Therefore, given $(S, X_1)$, $Y_2$ is independent of $X$.

$$\mathcal{C}_{G,PF} = \bigcup_{0 \le \alpha \le 1} \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_0 + R_2 \le \max\limits_{0 \le \beta \le 1} \min\left\{ \mathcal{C}\left( \frac{P_1 + \bar{\alpha} P + 2\sqrt{\bar{\beta}\bar{\alpha} P P_1}}{\alpha P + N_2} \right), \mathcal{C}\left( \frac{\beta \bar{\alpha} P}{\alpha P + \frac{N_1 N_2}{N_1+N_2}} \right) \right\} \\ R_1 \le \mathcal{C}\left( \frac{\alpha P}{\frac{N_1 N_2}{N_1+N_2}} \right) \end{array} \right\}. \tag{27}$$

Now, for the equivalent channel with outputs being $(S, Y_2)$ and $Y_2$, we have

$$\begin{aligned} S &= X + \frac{N_1}{N_1+N_2} X_1 + \hat{Z}_1 \\ Y_2 &= X + X_1 + \hat{Z}_1 + \hat{Z} \end{aligned} \tag{31}$$

where $\hat{Z}_1$ and $\hat{Z}$ are independent real Gaussian random variables with variances $\frac{N_1 N_2}{N_1+N_2}$ and $\frac{N_2^2}{N_2+N_1}$, respectively.

This is a self-interfered D-AWGN partially cooperative RBC with feedback, and hence Corollary 3 can be applied to obtain the capacity region. □

*Corollary 4:* Feedback enlarges the capacity region for the AWGN partially cooperative RBC, i.e.,

$$\mathcal{C}_{G,P} \subset \mathcal{C}_{G,PF}.$$

This corollary can be shown by comparing the capacity region $\mathcal{C}_{G,PF}$ of the AWGN partially cooperative RBC with feedback given in Theorem 8 with the outer bound $\mathcal{C}_{G,P}^{(o)}$ on the capacity region of the AWGN partially cooperative RBC given in Theorem 5. This result is reasonable because for the AWGN partially cooperative RBC, the output $y_2$ at user 2 is not a degraded version of the output $y_1$ at user 1, and hence, feedback of $y_2$ to user 1 provides further information and results in an enlargement of the capacity region.

### C. Comparison With Results on Broadcast Channels With Feedback

At this point, it is instructive to compare the results on the capacity regions for broadcast channels with feedback (see [39] for a review) with our results on the capacity regions of the partially cooperative RBCs with feedback.

We have obtained the capacity region of the general discrete memoryless partially cooperative RBC with feedback in Theorem 6. However, the capacity region is still not known for the general discrete memoryless broadcast channel with feedback. We have shown that feedback does not enlarge the capacity region for the degraded partially cooperative RBC and its Gaussian example in Corollary 2 and Theorem 7. This result is consistent with the result obtained in [40], [41] that feedback does not enlarge the capacity region for the physically degraded broadcast channel.

We have obtained the capacity region of the AWGN partially cooperative RBC with feedback in Theorem 8. However, the capacity region of the AWGN broadcast channel with feedback is still not known. We have shown that feedback enlarges the capacity region for the AWGN partially cooperative RBC. This is consistent with the result in [48]–[50] that feedback enlarges the capacity region for the AWGN broadcast channel. However, the reasons for the enlargement are different for the two channels. For the AWGN partially cooperative RBC with feedback, the source does not make use of the feedback information. It is user 1 that utilizes the feedback information from user 2 to improve on its decoding and relaying. Such a strategy achieves the capacity region. Whereas for the broadcast channel with feedback, the source needs to make use of the feedback information to improve the encoding. In [48]–[50], the authors provided example encoding schemes for the source to exploit feedback information to improve the capacity region. However, the optimal encoding scheme that achieves the capacity region for the AWGN broadcast channel with feedback is still not known.

We finally note that the structure of feedback in partially cooperative RBCs is different from that in broadcast channels. In partially cooperative RBCs with feedback, the output at user 2 is also fed back to user 1, but this feedback is not available in broadcast channels with feedback. Hence, obtaining the capacity region for the partially cooperative RBCs with feedback does not necessarily imply obtaining the capacity region for the broadcast channel with feedback.

## IV. Fully Cooperative Relay Broadcast Channels

In the previous sections, we studied the partially cooperative RBC where one user (usually the user with "better channel" from the source) in the broadcast system helps the other user by sending relay signals. In this case, we have seen that the partially cooperative RBC has a larger capacity region than the original broadcast channel due to user cooperation. It is then natural to explore whether the capacity region can be further enlarged if we allow both users to help each other by sending cooperative signals through relay links.

In this section, we study the fully cooperative RBC, where not only user 1 serves as a relay to help user 2, but user 2 serves as a relay to assist user 1 as well. We will first describe the channel model, and then present our main results. We further illustrate the results on achievable rate/capacity regions via Gaussian examples, and compare these regions with those of the partially cooperative RBC.

### A. System Model

*Definition 7:* A fully cooperative RBC (see Fig. 7) consists of three messages $(W_0, W_1, W_2)$, a source input $X \in \mathcal{X}$ with $\mathcal{X}$ being a finite-alphabet set, two relay inputs $X_1 \in \mathcal{X}_1$ and $X_2 \in \mathcal{X}_2$ with $\mathcal{X}_1$ and $\mathcal{X}_2$ being two finite-alphabet sets, two channel outputs $Y_1 \in \mathcal{Y}_1$ and $Y_2 \in \mathcal{Y}_2$ with $\mathcal{Y}_1$and $\mathcal{Y}_2$ being two finite-alphabet sets, and message estimate pairs $(\hat{W}_0^{(1)}, \hat{W}_1)$ and $(\hat{W}_0^{(2)}, \hat{W}_2)$. The channel transition probability distribution is $p(y_1, y_2|x, x_1, x_2)$.

*Definition 8:* A fully cooperative RBC is degraded if it either satisfies the condition

$$p(y_1, y_2|x, x_1, x_2) = p(y_1|x, x_1, x_2)p(y_2|y_1, x_1, x_2) \tag{32}$$

i.e., $y_2$ is independent of $x$, conditioned on $y_1$, $x_1$ and $x_2$; or satisfies the condition

$$p(y_1, y_2|x, x_1, x_2) = p(y_2|x, x_1, x_2)p(y_1|y_2, x_1, x_2) \tag{33}$$

i.e., $y_1$ is independent of $x$, conditioned on $y_2$, $x_1$ and $x_2$.

Without loss of generality, in this paper we consider only the degraded channel that satisfies the first condition.

The definition for a $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n\right)$ code for the fully cooperative RBC is similar to that for the partially cooperative RBC, except that it includes another set of relay functions $\{g_i\}_{i=1}^n$ such that

$$x_{2,i} = g_i(y_{2,1}, \ldots, y_{2,i-1}), \qquad 1 \le i \le n. \tag{34}$$

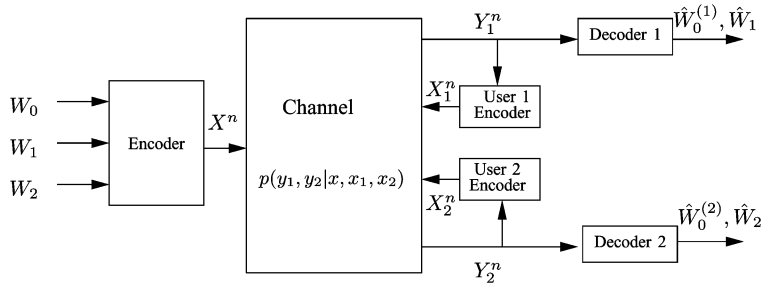


Fig. 7. Fully cooperative RBC.

### B. Discrete Memoryless Fully Cooperative RBCs

To derive an achievable rate region for the fully cooperative RBC, we first need to choose relaying schemes for user 1 and user 2 to assist each other. A simple choice is one where one of the users employs the decode-and-forward scheme, and the other user always sends a single codeword (that may vary according to the target rate tuple) which results in the best achievable rate region. We obtain the following achievable rate region which is the union of two achievable rate regions derived by switching these two relaying schemes for the two users.

*Theorem 9:* An inner bound $\mathcal{C}_F^{(i1)}$ on the capacity region of the fully cooperative RBC is given by the convex hull of the union of the following two rate regions $\mathcal{R}_1$ and $\mathcal{R}_2$:

$$\mathcal{R}_1 = \bigcup_{p(u,x_1,x_2,x)p(y_1,y_2|x,x_1,x_2)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 < \min\{I(U, X_1; Y_2|X_2), I(U; Y_1|X_1, X_2)\} \\ R_1 < I(X; Y_1 \mid U, X_1, X_2) \end{array} \right\} \tag{35}$$

$$\mathcal{R}_2 = \bigcup_{p(u',x_1,x_2,x)p(y_1,y_2|x,x_1,x_2)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_1 < \min\left\{I\left(U', X_2; Y_1|X_1\right), I\left(U'; Y_2|X_1, X_2\right)\right\} \\ R_2 < I\left(X; Y_2 \mid U', X_1, X_2\right) \end{array} \right\} \tag{36}$$

where $U$ and $U'$ are bounded in cardinality by

$$|\mathcal{U}| \leq |\mathcal{X}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2$$

and

$$|\mathcal{U}'| \leq |\mathcal{X}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2$$

respectively.

The proof is similar to the proof for Theorem 1 and is omitted.

The achievable rate region given in Theorem 9 will be shown to give the capacity region of the degraded fully cooperative RBC in Theorem 12. However, this achievable rate region may not be tight for a general fully cooperative RBC. For example, for the AWGN fully cooperative RBC which we will consider later, this achievable rate region is not tight, and the relay node that sends a single codeword does not help at all. This relay needs to employ a better relaying scheme, for example, the estimate-and-forward scheme [2, Sec. VI], to be of help. Hence, motivated by the AWGN fully cooperative RBC, we provide the following achievable rate region which is based on the scheme where one user in the system employs the decode-and-forward scheme and the other user in the system employs the estimate-and-forward scheme to relay information.

*Theorem 10:* An inner bound $\mathcal{C}_F^{(i2)}$ on the capacity region of the fully cooperative RBC is given by the convex hull of the union of the following two rate regions $\mathcal{R}_1$ and $\mathcal{R}_2$:

$$\mathcal{R}_1 = \bigcup_{p(x_2)p(u,x_1,x)p(y_1,y_2|x,x_1,x_2)p(\hat{y}_2|y_2,x_1,x_2,u)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 < \min\{I(U, X_1; Y_2|X_2), I(U; Y_1|X_1)\}, \\ R_1 < I(X; \hat{Y}_2, Y_1|U, X_1, X_2) \\ \texttt{subject to:} \\ I(X_2; Y_1|U, X_1) \geq I(\hat{Y}_2; Y_2|Y_1, U, X_1, X_2) \end{array} \right\} \tag{37}$$

$$\mathcal{R}_2 = \bigcup_{p(x_1)p(u',x_2,x)p(y_1,y_2|x,x_1,x_2)p(\hat{y}_1|y_1,x_1,x_2,u')} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_1 < \min\left\{I\left(U', X_2; Y_1|X_1\right), I\left(U'; Y_2|X_2\right)\right\}, \\ R_2 < I\left(X; \hat{Y}_1, Y_2|U', X_1, X_2\right) \\ \texttt{subject to:} \\ I\left(X_1; Y_2|U', X_2\right) \geq I\left(\hat{Y}_1; Y_1|Y_2, U', X_1, X_2\right) \end{array} \right\}. \tag{38}$$

*Proof:* See Appendix VI for an outline of the proof. □

The achievable rate region given in Theorem 10 serves as an example to demonstrate that the fully cooperative RBC can achieve larger rate region than the partially cooperative RBC due to an additional relay link. This will be clear when we apply Theorem 10 to the AWGN fully cooperative RBC in the next subsection.

There are other relaying schemes that the system can choose for the two users to assist each other, and each of these relaying schemes results in an achievable rate region. In general, these achievable regions are not tight.

In the following theorem, we provide an outer bound on the capacity region for the fully cooperative RBC. Note that this outer bound is tighter than the cut-set bound.

*Theorem 11:* An outer bound on the capacity region of the fully cooperative RBC is given by (39) at the top of the following page, where the random variables satisfy the Markov chain conditions: $(X_1, X_2) \to U \to X$ and $(X_1, X_2) \to U' \to$

$$\mathcal{C}_F^{(o)} = \bigcup_{p(u,u',x_1,x_2,x)p(y_1,y_2|x_1,x_2,x)} \left\{ \begin{array}{l} (R_0,R_1,R_2): R_0+R_2 \le \min\{I(U,X_1;Y_2|X_2), I(U;Y_1,Y_2|X_1,X_2)\} \\ R_1 \le I(X;Y_1,Y_2|U,X_1,X_2) \\ R_0+R_1 \le \min\{I(U',X_2;Y_1|X_1), I(U';Y_1,Y_2|X_1,X_2)\} \\ R_2 \le I(X;Y_1,Y_2|U',X_1,X_2) \end{array} \right\} \tag{39}$$

$X$. Furthermore, the auxiliary random variables $U$ and $U'$ are bounded in cardinality by $|\mathcal{U}| \le |\mathcal{X}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2$ and $|\mathcal{U}'| \le |\mathcal{X}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2$, respectively.

The proof for Theorem 11 is similar to the proof for Theorem 2, and is hence omitted.

For the special case of the degraded fully cooperative RBC, we have the following capacity theorem.

*Theorem 12:* The capacity region of the degraded fully cooperative RBC is given by

$$\mathcal{C}_{D,F} = \bigcup_{p(x_1,x_2,u)p(x|u)p(y_1,y_2|x,x_1,x_2)} \left\{ \begin{array}{l} (R_0,R_1,R_2): \\ R_0+R_2 \le \min\{I(U,X_1;Y_2|X_2), I(U;Y_1|X_1,X_2)\} \\ R_1 \le I(X;Y_1 \mid U,X_1,X_2) \end{array} \right\} \tag{40}$$

where $U$ is bounded in cardinality by $|\mathcal{U}| \le |\mathcal{X}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2$.

*Proof:* The achievability is given by $\mathcal{R}_1$ in Theorem 9. The converse proof follows by applying the degraded condition (32) to the first two bounds in (39) given in Theorem 11. □

*Remark 5:* Theorem 12 includes Theorem 1 in [43], [44] as a special case, because the model studied in [43], [44] is a special case of the fully cooperative RBC. In fact, if we let $Y_1 = (Y_{1a}, Y_{1b})$ and $Y_2 = (Y_{2a}, Y_{2b})$, define the channel probability distribution to be $p(y_{1a}, y_{2a}|x)p(y_{1b}|x_2)p(y_{2b}|x_1)$, assume that $p(y_{1b}|x_2)$ and $p(y_{2b}|x_1)$ are deterministic probability distributions, i.e., having values either 0 or 1, and further assume $H(Y_{1b}) = C_{21}$ and $H(Y_{2b}) = C_{12}$, then the fully cooperative RBC reduces to the channel studied in [43], [44]. For this channel, $\mathcal{C}_{D,F}$ given in Theorem 12 reduces to the following region

$$\bigcup_{p(u,x)p(y_{1a},y_{2a}|x)} \left\{ \begin{array}{l} (R_0,R_1,R_2): \\ R_0+R_2 \le \min\{I(U;Y_{2a})+C_{12}, I(U;Y_{1a})\} \\ R_1 \le I(X;Y_{1a} \mid U) \end{array} \right\} \tag{41}$$

which is the capacity region given in Theorem 1 in [43], [44].

Note that for the degraded fully cooperative RBC, the output at user 2 is a degraded version of the output at user 1, and it does not receive any more information than user 1. Hence, user 2 does not need to relay any information for user 1, and all it does is to send a single codeword (may vary according to the target rate tuple) which results in the best achievable rate region. The converse for Theorem 12 shows that this scheme is optimal.

### C. Gaussian Fully Cooperative RBCs

As for the partially cooperative RBC, we study two Gaussian channels for the fully cooperative RBC: the D-AWGN and AWGN fully cooperative RBCs.

For the D-AWGN fully cooperative RBC, the channel outputs at the two users are given by

$$\begin{aligned} Y_1 &= X + X_2 + Z_1 \\ Y_2 &= X + X_1 + Z_1 + Z' \end{aligned} \tag{42}$$

where $Z_1$ and $Z'$ are independent real Gaussian random variables with variances $N_1$ and $N_2 - N_1$, respectively, where $N_1 < N_2$. The channel input sequences $\{X_n\}$, $\{X_{1,n}\}$ and $\{X_{2,n}\}$ are subject to the average power constraints $P$, $P_1$ and $P_2$, respectively, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[X_i^2\right] \le P \qquad \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[X_{1,i}^2\right] \le P_1$$

and

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[X_{2,i}^2\right] \le P_2. \tag{43}$$

Note that the D-AWGN fully cooperative RBC satisfies the condition (32), and is hence degraded. The following theorem provides the capacity region for the D-AWGN fully cooperative RBC.

*Theorem 13:* The capacity region $\mathcal{C}_{DG,F}$ of the D-AWGN fully cooperative RBC is the same as the capacity region $\mathcal{C}_{DG,P}$ of the D-AWGN partially cooperative RBC given in Theorem 4, i.e.,

$$\mathcal{C}_{DG,F} = \mathcal{C}_{DG,P}.$$

*Proof:* The proof of the achievability is straightforward based on the proof of the achievability for the partially cooperative RBC with user 2 being silent. The proof of the converse follows similarly as the proof of the converse for Theorem 7. □

*Remark 6:* The relay link from user 2 to user 1 does not help to enlarge the capacity region for the D-AWGN fully cooperative RBC.

The intuition behind Theorem 13 is as follows. For the D-AWGN fully cooperative RBC, user 2 receives a degraded version of the output at user 1, and hence, cannot provide further information for user 1 other than the information that user 1 already has. This result is also consistent with Theorem 7 that feedback does not increase the capacity region for the D-AWGN partially cooperative RBC, i.e., perfectly providing the output at user 2 to user 1 cannot enlarge the capacity region for the D-AWGN partially cooperative RBC. It is then reasonable that sending information based on the output at user 2 to user 1 through a noisy channel cannot enlarge the capacity

region for the D-AWGN partially cooperative RBC. This is exactly what Theorem 13 concludes.

However, Theorem 13 is not necessarily true for the discrete memoryless degraded fully cooperative RBC. First of all, it is clear that the fully cooperative RBC achieves at least the capacity region for the corresponding partially cooperative RBC with user 2 "being silent" (sending a fixed alphabet symbol). We now explore whether the degraded fully cooperative RBC can achieve a better rate region. Although for the discrete memoryless degraded fully cooperative RBC, user 2 still cannot provide further information to user 1 through the relay link, it can affect the channel by sending a predetermined codeword (this is not the case for the D-AWGN fully cooperative RBC, because the channel noise term is independent of the input of user 2). Hence, user 2 can send a single codeword through the relay link to result in the best achievable region. Therefore, the relay input from user 2 can assist in enlarging the capacity region for the discrete memoryless partially cooperative RBC.

This fact can also be seen from the following simple example depicted in Fig. 8. For the partially cooperative RBC, we assume that the source node transmits to user 1 only and user 1 transmits to user 2 over an orthogonal link. We further assume that both links are deterministic with link capacities 1 and 0.5, respectively. It is clear this channel is degraded. The capacity region follows from Theorem 3 and is given by

$$\mathcal{C}_{D,P} = \bigcup_{0\le\alpha\le 1} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 \le \min\{1-\alpha, 0.5\}, \\ R_1 \le \alpha \end{array} \right\}. \quad (44)$$

For the fully cooperative RBC, we assume that the channel input $X_2$ from user 2 is binary taking values 0 or 1. If $X_2 = 0$, the channel is the same as the above partially cooperative RBC, and if $X_2 = 1$, the channel differs from the above partially cooperative RBC only in that the link from user 1 to user 2 has capacity 1. It is clear that for the fully cooperative RBC, it is optimal to always set the input $X_2 = 1$ which results in the following capacity region

$$\mathcal{C}_{D,F} = \{(R_0, R_1, R_2) : R_0 + R_1 + R_2 \le 1\}. \quad (45)$$

It is clear that $\mathcal{C}_{D,F}$ given in (45) is larger than $\mathcal{C}_{D,P}$ given in (44).

We now consider the AWGN fully cooperative RBC where the Gaussian noise terms at the two receivers are independent. For this channel, the outputs at the two users are given by

$$\begin{aligned} Y_1 &= X + X_2 + Z_1 \\ Y_2 &= X + X_1 + Z_2 \end{aligned} \quad (46)$$

where $Z_1$ and $Z_2$ are independent real Gaussian random variables with variances $N_1$ and $N_2$, respectively, where $N_1 < N_2$. The channel input sequences $\{X_n\}$, $\{X_{1,n}\}$, and $\{X_{2,n}\}$ are subject to the power constraints given in (43).

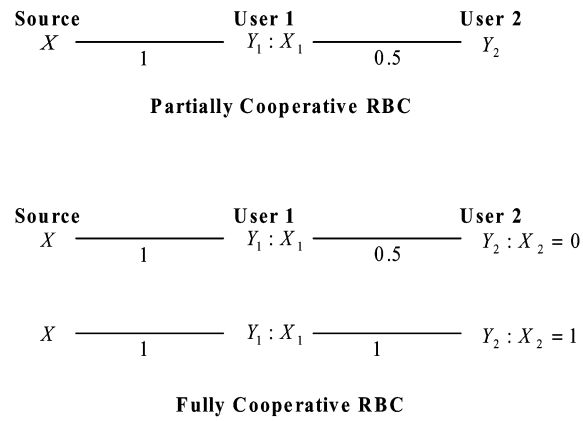


Fig. 8. Example partially and fully cooperative RBCs.

It is clear that the AWGN fully cooperative RBC can achieve the rate region of the AWGN partially cooperative RBC given in Corollary 1 by using the same coding scheme and keeping user 2 silent. Can it achieve a better rate region? Since the outputs at user 1 and user 2 are corrupted by stochastically independent noise terms, user 2 indeed receives some additional information other than the information received at user 1. Hence, potentially user 2 can assist in enlarging the rate for user 1 by sending this additional information to user 1 through the relay link. Although user 2 cannot decode this additional information (receiver noise level at user 2 is higher than that at user 1), it can first compress this information and then forward it to user 1, i.e., user 2 can employ the estimate-and-forward relaying scheme.

We provide an achievable rate region for the AWGN fully cooperative RBC based on user 1 employing the decode-and-forward relaying scheme and user 2 employing the estimate-and-forward relaying scheme in the following theorem.

*Theorem 14:* An inner bound on the capacity region of the AWGN fully cooperative RBC is given by (47) at the bottom of the page, where "Convex" indicates the convex hull of the rate region.

*Proof:* See Appendix VII. □

Note that the parameter $\eta$ in (47) is the fraction of the relay power at user 2 being used for relaying transmission. This parameter can be used as tradeoff between the rates $R_2$ and $R_1$. Enlarging $\eta$ sacrifices $R_2$ to improve $R_1$. This is because larger $\eta$ causes more interference to user 1 in the decoding of the information for user 2, and hence, user 1 becomes less helpful for user 2. On the other hand, user 1 gets more benefit from user 2 due to the larger relaying power.

*Remark 7:* Theorem 14 shows that the AWGN fully cooperative RBC indeed achieves a larger rate region than the AWGN partially cooperative RBC. This is because the relay link from user 2 to user 1 assists in enlarging the rate region. Theorem

$$C_{G,F}^{(i)} = \text{Convex} \bigcup_{0\le\alpha,\eta\le 1} \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_1 < \mathcal{C}\left(\frac{\alpha P}{N_1} + \frac{\alpha\eta P P_2}{\eta P_2 N_2 + \alpha P(N_1+N_2) + N_1 N_2}\right) \\ R_0 + R_2 < \max\limits_{0\le\beta\le 1} \min\left\{\mathcal{C}\left(\frac{\bar{\alpha}P + P_1 + 2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P + N_2}\right), \mathcal{C}\left(\frac{\beta\bar{\alpha}P}{\alpha P + \eta P_2 + N_1}\right)\right\} \end{array} \right\} \quad (47)$$

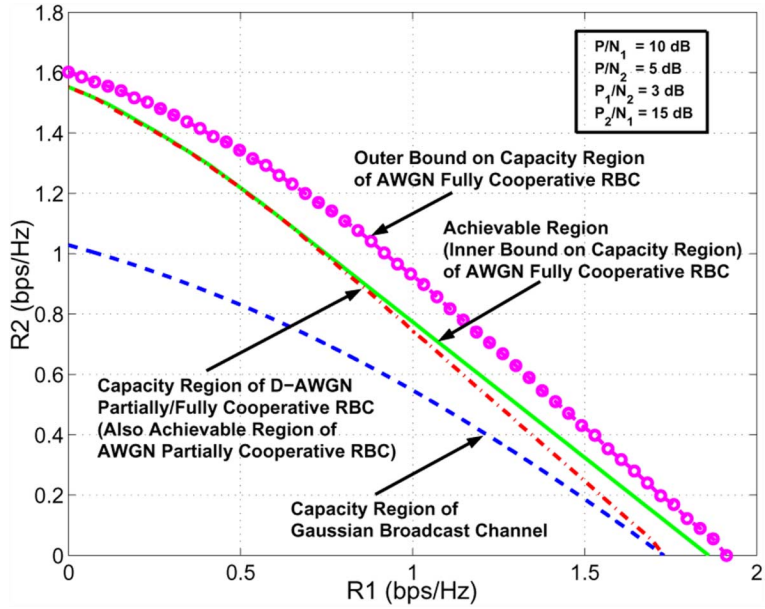


Fig. 9. Comparison of rate regions for Gaussian RBCs.

14 also shows that the AWGN fully cooperative RBC achieves a larger rate region than the D-AWGN fully cooperative RBC. These two facts will be demonstrated by the numerical results at the end of this section.

We further provide an outer bound on the capacity region of the AWGN fully cooperative RBC.

*Theorem 15:* An outer bound on the capacity region of the AWGN fully cooperative RBC is given by (48), at the bottom of the page.

*Proof:* See Appendix VIII. □

We now compare the achievable rate region of the AWGN fully cooperative RBC with the capacity region of the D-AWGN partially/fully cooperative RBC numerically. In Fig. 9, we plot the inner bound (boundary with solid line) and outer bound (boundary with circled line) on the capacity region of the AWGN fully cooperative RBC, the capacity region (boundary with dot-dashed line) for the D-AWGN partially/fully cooperative RBC, and the capacity region (boundary with dashed line) of the original Gaussian broadcast channel without relay links. Since the capacity region of the AWGN fully cooperative RBC lies between its inner and outer bounds, it is clear from the figure that this capacity region is larger than the capacity region of the D-AWGN partially/fully cooperative RBC.

In Fig. 9, we also plot the achievable region for the AWGN partially cooperative RBC (boundary also with dot-dashed line). It is clear from the figure that the achievable region of the AWGN fully cooperative RBC is larger than the achievable region of the AWGN partially cooperative RBC. Furthermore, for the AWGN fully cooperative RBC, the maximum rates of both users 1 and 2 are improved relative to the original Gaussian broadcast channel. However, for the AWGN partially cooperative RBC, the maximum rate of only user 2 is improved. This is because user 2 in the AWGN fully cooperative RBC helps user 1 through a relay link, which is not allowed for the AWGN partially cooperative RBC.

## V. Fully Cooperative RBCs With Feedback

In this section, we study the fully cooperative RBC with feedback, where the outputs at user 2 are provided to user 1 and the outputs at both users 1 and 2 are provided to the source all through perfect feedback links (see Fig. 10). We study how feedback affects the fully cooperative RBC.

$$C_{G,F}^{(o)} = \bigcup_{0\le\alpha,\beta,\gamma\le 1} \left\{ \begin{array}{l} (R_0,R_1,R_2): R_0+R_2 \le \min\left\{\mathcal{C}\left(\frac{P_1+\bar{\alpha}P+2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P+N_2}\right), \mathcal{C}\left(\frac{\beta\bar{\alpha}P}{\alpha P+\frac{N_1N_2}{N_1+N_2}}\right)\right\} \\ R_1 \le \mathcal{C}\left(\frac{\alpha P}{\frac{N_1N_2}{N_1+N_2}}\right) \\ R_0+R_1 \le \min\left\{\mathcal{C}\left(\frac{P+P_2+2\sqrt{\bar{\beta}\bar{\alpha}PP_2}}{N_1}\right), \mathcal{C}\left(\frac{\bar{\gamma}(\alpha P+\beta\bar{\alpha}P)}{\gamma(\alpha P+\beta\bar{\alpha}P)+\frac{N_1N_2}{N_1+N_2}}\right)\right\} \\ R_2 \le \mathcal{C}\left(\frac{\gamma(\alpha P+\beta\bar{\alpha}P)}{\frac{N_1N_2}{N_1+N_2}}\right) \end{array} \right\}. \quad (48)$$

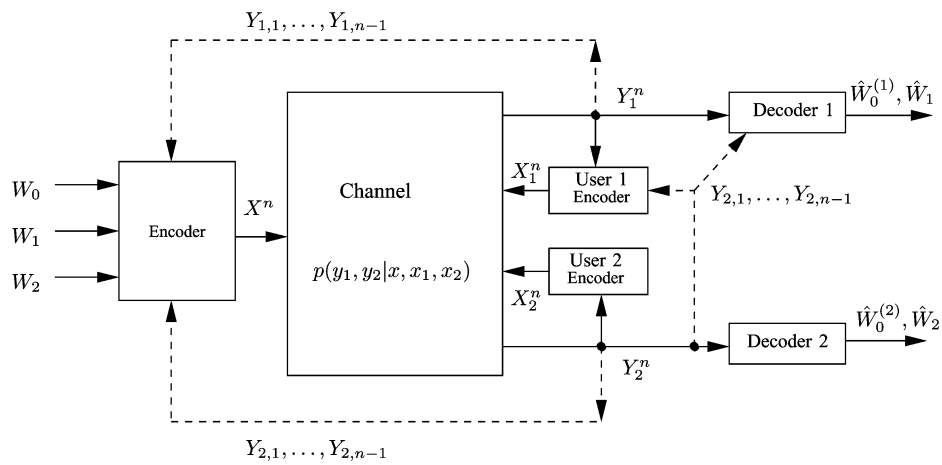


Fig. 10. Fully cooperative RBC with feedback.

For a distribution on the message set $p(w_0, w_1, w_2)$, the following joint distribution is induced for the fully cooperative RBC with feedback:

$$\begin{aligned} &p\left(w_0, w_1, w_2, x^n, x_1^n, x_2^n, y_1^n, y_2^n\right) \\ &\quad = p(w_0, w_1, w_2) \\ &\qquad \cdot \prod_{i=1}^{n} \Bigg[ p\left(x_i | w_0, w_1, w_2, y_1^{i-1}, y_2^{i-1}\right) \\ &\qquad\qquad \cdot p\left(x_{1,i} | y_1^{i-1}, y_2^{i-1}\right) p\left(x_{2,i} | y_2^{i-1}\right) \\ &\qquad\qquad \cdot p(y_{1,i}, y_{2,i} | x_i, x_{1,i}, x_{2,i}) \Bigg]. \end{aligned} \tag{49}$$

### A. Discrete Memoryless Fully Cooperative RBCs With Feedback

Since the fully cooperative RBC with feedback is degraded, we can obtain the capacity region of this feedback channel.

*Theorem 16:* The capacity region of the fully cooperative RBC with feedback is given by

$$C_{FF} = \bigcup_{p(x_1,x_2,u)p(x|u)p(y_1,y_2|x,x_1,x_2)} \left\{ \begin{array}{l} (R_0, R_1, R_2): \\ R_0 + R_2 \le \min\{I(U, X_1; Y_2|X_2), I(U; Y_1, Y_2|X_1, X_2)\} \\ R_1 \le I(X; Y_1, Y_2 \mid U, X_1, X_2) \end{array} \right\} \tag{50}$$

where the auxiliary random variable $U$ is bounded in cardinality by $|\mathcal{U}| \le |\mathcal{X}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2$.

*Proof:* Theorem 9 provides the achievability with $(Y_1, Y_2)$ replacing $Y_1$ as the output at user 1. The proof of the converse is similar to those for Theorems 2 and 6, and is hence omitted. □

*Remark 8:* From the proof of the achievability, the source does not exploit feedback information to achieve the capacity region of the channel. Hence even if the channel only has feedback from user 2 to user 1, the capacity region is still the same as the channel with additional feedback from both users to the source. This is similar to what we have remarked for the partially cooperative RBC with feedback.

*Remark 9:* For the fully cooperative RBC, since the output at user 2 is fed back to user 1, it may seem that the relay link from user 2 to user 1 would not serve a useful purpose. This is indeed true for Gaussian channels as we will study in the next subsection. However, this may not be the case for the discrete memoryless channel. The reason is that although user 2 does not need to forward any information to user 1, the relay input sent by user 2 may still affect the channel. Hence, user 2 can send a single codeword through the relay link to result in the best achievable region.

Note that if the original fully cooperative RBC is degraded, the capacity region given in Theorem 16 is the same as the capacity region of the same channel without feedback given in Theorem 12.

*Corollary 5:* Feedback does not increase the capacity region of the degraded fully cooperative RBC, i.e.,

$$\mathcal{C}_{D,FF} = \mathcal{C}_{D,F}$$

where $\mathcal{C}_{D,FF}$ indicates the capacity region of the degraded fully cooperative RBC with feedback, and $\mathcal{C}_{D,F}$ is given in Theorem 12.

### B. Gaussian Fully Cooperative RBCs With Feedback

In this subsection, we study two Gaussian fully cooperative RBCs with feedback: the D-AWGN and the AWGN cases.

We first have the following capacity theorem for the D-AWGN fully cooperative RBC with feedback.

*Theorem 17:* Feedback does not increase the capacity region for the D-AWGN fully cooperative RBC, i.e.,

$$\mathcal{C}_{DG,FF} = \mathcal{C}_{DG,F}$$

where $\mathcal{C}_{DG,FF}$ indicates the capacity region of the D-AWGN fully cooperative RBC with feedback, and $\mathcal{C}_{DG,F}$ is given in Theorem 13.

*Proof:* The proof of the achievability is given by that for the D-AWGN fully cooperative RBC without feedback. The proof

of the converse is similar to that for the D-AWGN partially cooperative RBC with feedback, which is given in Appendix IV. □

Note that for the D-AWGN partially cooperative RBC, we have studied a self-interfered channel, which is a reasonable model from a practical point of view. Similarly, it is of interest to study the following self-interfered D-AWGN fully cooperative RBC model, where the outputs at users 1 and 2 are given by

$$\begin{aligned} Y_1 &= X + aX_1 + bX_2 + Z_1 \\ Y_2 &= X + X_1 + dX_2 + Z_1 + Z' \end{aligned} \tag{51}$$

where $Z_1$ and $Z'$ are independent real Gaussian random variables with variances $N_1$ and $N_2 - N_1$, respectively, and parameters $a$, $b$, and $d$ are real numbers.

Note that the self-interfered D-AWGN fully cooperative RBC is also degraded. Theorem 17 still holds for this channel with feedback.

*Corollary 6:* The capacity region of the self-interfered D-AWGN fully cooperative RBC with/without feedback is the same as the capacity region of the D-AWGN fully cooperative RBC with/without feedback, and is given in Theorem 17.

We now consider the AWGN fully cooperative RBC with feedback. The following theorem provides the capacity region for this channel.

*Theorem 18:* The capacity region of the AWGN fully cooperative RBC with feedback is the same as the capacity region for the AWGN partially cooperative RBC with feedback given in Theorem 8, i.e.,

$$\mathcal{C}_{G,FF} = \mathcal{C}_{G,PF}.$$

*Proof:* The proof follows similarly as that for Theorem 8, and is briefly summarized as follows.

We define

$$S := \frac{N_1 Y_2 + N_2 Y_1}{N_1 + N_2}, \tag{52}$$

and the mapping from $(Y_1, Y_2)$ to $(S, Y_2)$ is one-to-one. Hence, the channel with the outputs being $(Y_1, Y_2)$ and $Y_2$ is equivalent to the channel with the outputs being $(S, Y_2)$ and $Y_2$.

We express $S$ and $Y_2$ in the following form:

$$S = X + \frac{N_1}{N_1 + N_2} X_1 + \frac{N_2}{N_1 + N_2} X_2 + \hat{Z}_1 \tag{53}$$

$$Y_2 = S + \frac{N_2}{N_1 + N_2} X_1 - \frac{N_2}{N_1 + N_2} X_2 + \hat{Z} \tag{54}$$

where $\hat{Z}_1 := \frac{N_1 Z_2 + N_2 Z_1}{N_1 + N_2}$ and $\hat{Z} := \frac{N_2}{N_1 + N_2}(Z_2 - Z_1)$.

It is clear that $\hat{Z}$ is independent of $X$, $X_1$ and $X_2$, and is also independent of $\hat{Z}_1$. Hence, $\hat{Z}$ is independent of $S$. Therefore, $Y_2$ is independent of $X$, given $(S, X_1, X_2)$, i.e., $Y_2$ is a degraded version of $S$.

Now, for the equivalent channel with outputs being $(S, Y_2)$ and $Y_2$, we have

$$\begin{aligned} S &= X + \frac{N_1}{N_1 + N_2} X_1 + \frac{N_2}{N_1 + N_2} X_2 + \hat{Z}_1 \\ Y_2 &= X + X_1 + \hat{Z}_1 + \hat{Z} \end{aligned} \tag{55}$$

where $\hat{Z}_1$ and $\hat{Z}$ are independent real Gaussian random variables with variances $\frac{N_1 N_2}{N_1 + N_2}$ and $\frac{N_2^2}{N_2 + N_1}$, respectively. This is a self-interfered D-AWGN fully cooperative RBC defined in (51) with feedback, and hence Corollary 6 can be applied to yield the capacity region. □

We have the following remarks for the capacity region of the AWGN fully cooperative RBC with feedback.

*Remark 10:* Theorem 18 implies that feedback effectively changes the AWGN fully cooperative RBC with feedback to a D-AWGN fully cooperative RBC with feedback but with noise variance $N_1$ being replaced by $\frac{N_1 N_2}{N_1 + N_2}$.

*Remark 11:* For both D-AWGN and AWGN fully cooperative RBCs with feedback, their capacity regions are the same as the corresponding partially cooperative RBCs with feedback. Hence, for these two Gaussian channels with feedback, the relay link from user 2 to user 1 does not help, because all the useful information at user 2 has been conveyed to user 1 through feedback.

We now compare the capacity region of the AWGN fully cooperative RBC with feedback given in Theorem 18 with the outer bound on the capacity region of the AWGN fully cooperative RBC given in Theorem 15. It is clear that the capacity region of the AWGN fully cooperative RBC with feedback contains the outer bound on the capacity region of the AWGN fully cooperative RBC without feedback. In particular, for small values of the relay power $P_2$, the capacity region of the AWGN fully cooperative RBC with feedback strictly contains the outer bound on the capacity region of the same channel without feedback, i.e., feedback enlarges the capacity region for the AWGN fully cooperative RBC for these cases.

## VI. Comments on Power Constraints

In Section II-C, we showed that the achievable region of the AWGN partially cooperative RBC (given in Corollary 1) is larger than the capacity region of the original Gaussian broadcast channel. This is also demonstrated by a numerical example in Fig. 5. However, this comparison is based on the assumption that the power constraint at the source for the AWGN partially cooperative RBC is the same as the power constraint at the source for the Gaussian broadcast channel, and that there is an additional power $P_1$ for the relay node (user 1) to transmit relaying information. Hence, it is conceivable that the improvement in the capacity region for the AWGN partially cooperative RBC is due to this additional power at the relay node. We now consider a case for the AWGN partially cooperative RBC where the total power available for the source and relay is the same as the power available for the source in the broadcast channel, i.e., the source and the relay node need to share the amount of power $P$. We explore whether the AWGN partially cooperative RBC still has a larger capacity region than the Gaussian broadcast channel.

In Fig. 11, we plot the achievable rate region (boundary with solid line) for the AWGN partially cooperative RBC, where $N_1 < N_2$, and where the source and the relay node share the amount of power $P$. We compare this achievable rate region with the capacity region (boundary with dashed line) of the

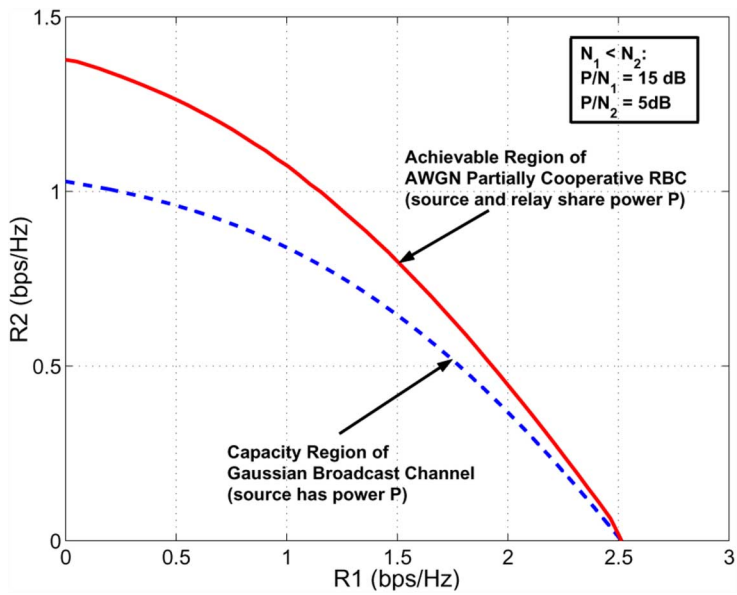


Fig. 11. Comparison of achievable rate region of AWGN partially cooperative RBC under total power constraint with capacity region of Gaussian broadcast channel: case where $N_1 < N_2$.

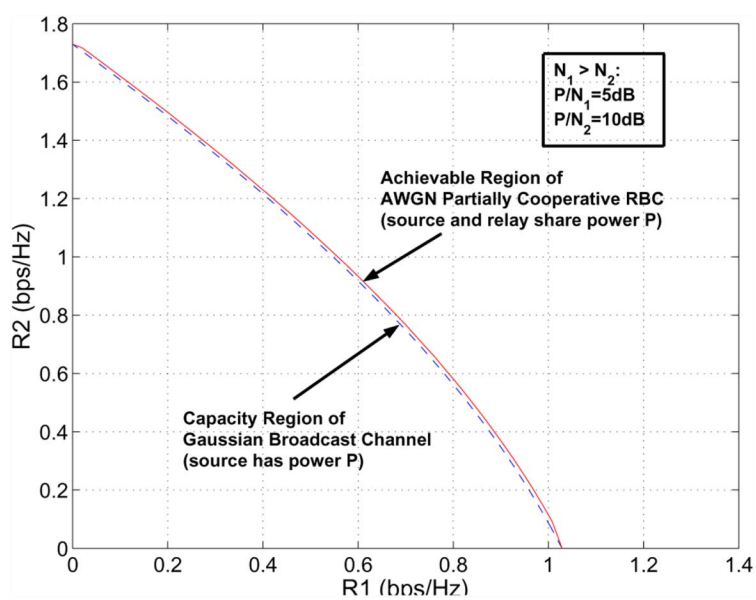


Fig. 12. Comparison of achievable rate region of AWGN partially cooperative RBC under total power constraint with capacity region of Gaussian broadcast channel: case where $N_1 > N_2$.

Gaussian broadcast channel with the power constraint $P$ for the source. It is clear from the graph that the achievable region of the AWGN partially cooperative RBC is larger than the capacity region of the Gaussian broadcast channel. Clearly, this enlargement is not due to the additional power at the relay node any more. The gain comes from the fact that the source and relay can coherently transmit information to user 2. Of course this gain due to coherent combining is limited by how much information the source can forward to the relay node before the two nodes can actually cooperate. Hence the stronger the link from the source to relay is, the larger the coherent combining gain that can be achieved. Since we have assumed that $N_1 < N_2$, which means that the source can forward more information to the relay than to user 2 through the direct link from the source to user 2, the coherent combining gain always exists even if it may be small.

We next consider the AWGN partially cooperative RBC where $N_1 > N_2$, i.e., where the relay has a worse channel from the source than user 2. We note that when the relay has an additional power $P_1$, the achievable region given in (23) is larger than the capacity region of the original Gaussian broadcast channel. Now even when the source and relay are subject to a total power constraint $P$, the achievable rate region based on the estimate-and-forward scheme for the AWGN partially cooperative RBC is larger than the capacity region of the original broadcast channel (see Fig. 12). However, this improvement is not due to coherent combining between the source and the relay as in the case where $N_1 < N_2$. The reason

for the improvement is that in this case the source needs to split some power for the relay and hence the source signals cause less interference at the relay. Thus, the relay is able to decode at a higher rate. Fig. 12 also suggests that this improvement is small, and that the maximum rate of user 2 is not improved. This suggests that in the case where the relay has a worse channel from the source than user 2, the relay transmission may not help much in enlarging the capacity region unless an additional amount of power is available at the relay. On the other hand, in this case, letting user 2 be the relay node makes the relaying more helpful, because now the relay (user 2) has a better channel from the source than user 1, and coherent combining helps to enlarge the capacity region.

Similarly, the achievable rate region of the AWGN fully cooperative RBC, where the source and two users share the amount of power $P$, is larger than the capacity region of the original Gaussian broadcast channel with the source subject to power constraint $P$. In the AWGN fully cooperative RBC, the relay node with a better channel from the source helps more toward enlarging the capacity region due to coherent combining. The relay node with a worse channel from the source helps only a little due to less interference in decoding as in the preceding discussion.

## VII. Concluding Remarks

We performed a comprehensive information-theoretic study of two relay broadcast channels, the partially cooperative RBC, and the fully cooperative RBC. We derived bounds on the capacity region for these channels, and established the capacity region for the special cases of degraded channels. We demonstrated via Gaussian examples that these RBCs have significant gains in capacity region compared to standard broadcast channels. Our results suggest that cooperative relaying is a powerful technique in achieving high-speed communication for wireless downlink systems and other networks that include broadcast transmissions.

In analyzing the capacity regions of RBCs, we provided an alternative to the cut-set bound approach to obtain outer bounds. We showed that our outer bounds are tighter than those based on the cut-set bound, and that they are close to the corresponding inner bounds in Gaussian examples. We believe that our technique for deriving these outer bounds is applicable more generally, and may be useful in deriving tighter outer bounds than the cut-set bound in other network information theory problems.

For the RBCs studied in the paper, the relay is allowed to transmit and receive at the same time in the same frequency band. In practice, models where the relays transmit and receive in orthogonal channels may be of interest. These RBCs have been studied in recent papers from an information-theoretic viewpoint [44], [43], and in the fading channel context [51].

In this paper, we have focused purely on the information-theoretic aspect of the RBCs. Further studies on this topic from coding and networking viewpoints will allow for the implementation of relaying and user cooperation in future wireless networks.

## Appendix I
## Outline of Proof for Theorem 1

We assume that the source uses the superposition coding which is optimal for the degraded broadcast channel [47, Ch. 14.6 ]. We also assume that the relay (user 1) uses the decode-and-forward relaying scheme [2, Sec. II ]. We adopt the regular encoding/sliding-window decoding strategy [52] for the decode-and-forward scheme which is different from the irregular encoding/successive decoding strategy used in [2, Sec. II ]. A review of three decode-and-forward strategies can be found in [7, Sec. I ].

We first prove that without common message $W_0$, the following rate pair is achievable:

$$\begin{aligned} R_2 &< \min\{I(U, X_1; Y_2), I(U; Y_1 \mid X_1)\} \\ R_1 &< I(X; Y_1 \mid U, X_1). \end{aligned} \tag{56}$$

Then, from the following proof, it is easily seen that user 1 decodes the messages for both users 1 and 2. We can hence view part of the rate $R_2$ to be the common rate $R_0$, and the rate region given in Theorem 1 is achievable.

We consider a transmission over $B$ blocks, each with length $n$. At each of the first $B-1$ blocks, a message pair $(W_{1,i}, W_{2,i}) \in [1, 2^{nR_1}] \times [1, 2^{nR_2}]$ is encoded and sent from the source, where $i$ denotes the index of the block, and $i = 1, 2, \ldots, B-1$. The rate pair $\left(R_1 \frac{B-1}{B}, R_2 \frac{B-1}{B}\right)$ approaches $(R_1, R_2)$ as $B \to \infty$.

We use random codes for the proof. Fix a joint probability distribution of $(X_1, U, X, Y_1, Y_2)$

$$p(x_1)p(u|x_1)p(x|u, x_1)p(y_1, y_2|x, x_1) \tag{57}$$

where $U$ is an auxiliary random variable that stands for the information being carried by the source input that is intended for user 2. In the following, we use $T_\epsilon^n(P_{UX_1XY_1Y_2})$ to denote the strongly jointly $\epsilon$-typical set (see [53, Sec. 1.2 ] for definition) based on the joint distribution (57).

Random Codebook Generation: We generate two statistically independent random codebooks 1 and 2 by the following same steps.

1. Generate $2^{nR_2}$ independent and identically distributed (i.i.d.) $\underline{x}_1$ each with distribution $\prod_{i=1}^n p(x_{1,i})$. Index $\underline{x}_1(w_2')$, $w_2' \in [1, 2^{nR_2}]$.

2. For each $\underline{x}_1(w_2')$, generate $2^{nR_2}$ i.i.d. $\underline{u}$ each with distribution $\prod_{i=1}^n p(u_i|x_{1,i}(w_2'))$. Index $\underline{u}(w_2', w_2)$, $w_2 \in [1, 2^{nR_2}]$.

3. For each $\underline{x}_1(w_2')$ and $\underline{u}(w_2', w_2)$, generate $2^{nR_1}$ i.i.d. $\underline{x}$ each with distribution $\prod_{i=1}^n p(x_i|u_i(w_2', w_2), x_{1,i}(w_2'))$. Index $\underline{x}(w_2', w_2, w_1)$, $w_1 \in [1, 2^{nR_1}]$.

Encoding: We encode messages using codebooks 1 and 2, respectively, for blocks with odd and even indices. This is because some of the following decoding steps are performed jointly over two adjacent blocks, and having independent codebooks makes the error events corresponding to these blocks independent, thus making the probabilities of these error events easy to calculate.

At the beginning of block $i$, let $(w_{1,i}, w_{2,i})$ be the new message pair to be sent from the source in block $i$, and $(w_{1,i-1}, w_{2,i-1})$ be the message pair being sent from the source in previous block $i-1$. The source encoder then sends $\underline{x}(w_{2,i-1}, w_{2,i}, w_{1,i})$.

TABLE III
CODEWORDS BEING SENT IN THE FIRST THREE BLOCKS TO ACHIEVE THE RATE REGION $\mathcal{C}_P^{(i)}$ WITH $R_0 = 0$

| block 1 | block 2 | block 3 |
|---|---|---|
| $\underline{x}_1(1)$ | $\underline{x}_1(w_{2,1})$ | $\underline{x}_1(w_{2,2})$ |
| $\underline{u}\ (1, w_{2,1})$ | $\underline{u}\ (w_{2,1}, w_{2,2})$ | $\underline{u}\ (w_{2,2}, w_{2,3})$ |
| $\underline{x}\ (1, w_{2,1}, w_{1,1})$ | $\underline{x}\ (w_{2,1}, w_{2,2}, w_{1,2})$ | $\underline{x}\ (w_{2,2}, w_{2,3}, w_{1,3})$ |

At the beginning of block $i$, user 1 (relay node) has decoded the message $w_{2,i-1}$ transmitted from the source in previous block $i-1$. It then sends the codeword $\underline{x}_1(w_{2,i-1})$.

For convenience, we list the codewords that are sent in the first three blocks in Table III.

Decoding: The decoding procedures at the end of block $i$ are as follows.

1. User 1, having known $w_{2,i-1}$, declares the message $\hat{w}_{2,i}$ is sent if there is a unique $\hat{w}_{2,i}$ such that

$$\left(\underline{x}_1(w_{2,i-1}), \underline{u}(w_{2,i-1}, \hat{w}_{2,i}), \underline{y}_1(i)\right) \in T_\epsilon^n(P_{UX_1Y_1}).$$

It can be shown that the decoding error in this step is small for sufficiently large $n$ if

$$R_2 < I(U; Y_1|X_1). \tag{58}$$

2. User 1, having known $w_{2,i-1}$ and $w_{2,i}$, declares the message $\hat{w}_{1,i}$ is sent if there is a unique $\hat{w}_{1,i}$ such that

$$\left(\underline{x}_1(w_{2,i-1}), \underline{u}(w_{2,i-1}, w_{2,i}), \underline{x}(w_{2,i-1}, w_{2,i}, \hat{w}_{1,i}), \underline{y}_1(i)\right) \in T_\epsilon^n(P_{UX_1XY_1}).$$

It can be shown that the decoding error in this step is small for sufficiently large $n$ if

$$R_1 < I(X; Y_1 \mid U, X_1). \tag{59}$$

3. User 2, having known $w_{2,i-2}$, decodes $w_{2,i-1}$ based on the information received in blocks $i-1$ and $i$. It declares that the message $\hat{\hat{w}}_{2,i-1}$ is sent if there is a unique $\hat{\hat{w}}_{2,i-1}$ such that

$$(\underline{x}_1(w_{2,i-2}), \underline{u}(w_{2,i-2}, \hat{\hat{w}}_{2,i-1}), \underline{y}_2(i-1)) \in T_\epsilon^n(P_{UX_1Y_2})$$

and

$$(\underline{x}_1(\hat{\hat{w}}_{2,i-1}), \underline{y}_2(i)) \in T_\epsilon^n(P_{X_1Y_2}).$$

It can be shown that the decoding error in this step is small for sufficiently large $n$ if

$$R_2 < I(U; Y_2|X_1) + I(X_1; Y_2) = I(X_1, U; Y_2). \tag{60}$$

Combining (58)–(60), we conclude that the rate region given in (56) is achievable.

Finally, the cardinality of $U$ can be bounded by applying standard techniques (e.g., see [53, Lemma 3.4 ]).

## APPENDIX II
## PROOF OF THEOREM 2

The proof uses techniques that are used in proving the converse of the capacity region of the degraded broadcast channel [47, Ch. 14, Problem 11], and in proving the upper bound on the capacity region for the relay channel [2, Sec. III ].

We consider a sequence of $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n\right)$ codes for a partially cooperative RBC with $P_e^{(n)} \to 0$. Then the probability distribution on the joint ensemble space $W_0 \times W_1 \times W_2 \times X^n \times X_1^n \times Y_1^n \times Y_2^n$ is given by

$$\begin{aligned} &p\left(w_0, w_1, w_2, x^n, x_1^n, y_1^n, y_2^n\right) \\ &\quad = p(w_0)p(w_1)p(w_2)p\left(x^n|w_0, w_1, w_2\right) \\ &\qquad \cdot \prod_{i=1}^n \left[p\left(x_{1,i}|y_1^{i-1}\right) p\left(y_{1,i}, y_{2,i}|x_i, x_{1,i}\right)\right]. \end{aligned} \tag{61}$$

By Fano's inequality, we have

$$\begin{aligned} H\left(W_1|Y_1^n\right) &\le H\left(W_0, W_1|Y_1^n\right) \\ &\le n(R_0 + R_1)P_e^{(n)} + 1 := n\delta_{1,n} \end{aligned} \tag{62}$$

$$\begin{aligned} H\left(W_2|Y_2^n\right) &\le H\left(W_0, W_2|Y_2^n\right) \\ &\le n(R_0 + R_2)P_e^{(n)} + 1 := n\delta_{2,n}. \end{aligned} \tag{63}$$

Note that $\delta_{1,n}, \delta_{2,n} \to 0$ as $n \to \infty$ if $P_e^{(n)} \to 0$ as $n \to \infty$.

We first consider

$$\begin{aligned} &nR_0 + nR_2 \\ &\quad = H(W_0, W_2) = I\left(W_0, W_2; Y_2^n\right) + H\left(W_0, W_2|Y_2^n\right) \\ &\quad \le I\left(W_0, W_2; Y_2^n\right) + n\delta_{2,n} \\ &\quad = \sum_{i=1}^n I\left(W_0, W_2; Y_{2,i}|Y_2^{i-1}\right) + n\delta_{2,n} \\ &\quad = \sum_{i=1}^n \left[H\left(Y_{2,i}|Y_2^{i-1}\right) - H\left(Y_{2,i}|Y_2^{i-1}, W_0, W_2\right)\right] + n\delta_{2,n} \\ &\quad \le \sum_{i=1}^n \left[H(Y_{2,i}) - H\left(Y_{2,i}|Y_2^{i-1}, W_0, W_2, Y_1^{i-1}, X_{1,i}\right)\right] \\ &\qquad + n\delta_{2,n} \\ &\quad = \sum_{i=1}^n \left[H(Y_{2,i}) - H(Y_{2,i}|U_i, X_{1,i})\right] + n\delta_{2,n} \\ &\quad = \sum_{i=1}^n I(U_i, X_{1,i}; Y_{2,i}) + n\delta_{2,n} \end{aligned} \tag{64}$$

where we define $U_i := \left(W_0, W_2, Y_1^{i-1}, Y_2^{i-1}\right)$. Note that $X_{1,i} \to U_i \to X_i$ form a Markov chain, and $U_i \to (X_i, X_{1,i}) \to (Y_{1,i}, Y_{2,i})$ also form a Markov chain.

We then consider

$$\begin{aligned} &nR_0 + nR_2 \\ &\quad \le I\left(W_0, W_2; Y_2^n\right) + n\delta_{2,n} \le I\left(W_0, W_2; Y_2^n, Y_1^n\right) + n\delta_{2,n} \\ &\quad \overset{(p)}{=} \sum_{i=1}^n I\left(W_0, W_2; Y_{2,i}, Y_{1,i}|Y_2^{i-1}, Y_1^{i-1}\right) + n\delta_{2,n} \\ &\quad = \sum_{i=1}^n \left[H\left(W_0, W_2|Y_2^{i-1}, Y_1^{i-1}\right) - H\left(W_0, W_2|Y_2^i, Y_1^i\right)\right] \\ &\qquad + n\delta_{2,n} \\ &\quad \overset{(a)}{\le} \sum_{i=1}^n \Bigg[H\left(W_0, W_2|Y_2^{i-1}, Y_1^{i-1}, X_{1,i}\right) \\ &\qquad\qquad - H\left(W_0, W_2|Y_2^i, Y_1^i, X_{1,i}\right)\Bigg] + n\delta_{2,n} \end{aligned}$$

$$\overset{(q)}{=}\sum_{i=1}^{n} I\left(W_0,W_2;Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},X_{1,i}\right)+n\delta_{2,n} \tag{65}$$

where (a) follows from the fact that conditioned on $\left(Y_2^{i-1},Y_1^{i-1}\right)$, $X_{1,i}$ is independent of $W_0,W_2$. We proceed to bound $nR_0+nR_2$ and obtain

$$\begin{aligned}
&nR_0+nR_2\\
&\le\sum_{i=1}^{n}\Bigg[H\left(Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},X_{1,i}\right)\\
&\qquad -H\left(Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},X_{1,i},W_0,W_2\right)\Bigg]+n\delta_{2,n}\\
&\le\sum_{i=1}^{n}\left[H(Y_{1,i},Y_{2,i}|X_{1,i})-H(Y_{1,i},Y_{2,i}|U_i,X_{1,i})\right]+n\delta_{2,n}\\
&=\sum_{i=1}^{n}I(U_i;Y_{1,i},Y_{2,i}|X_{1,i})+n\delta_{2,n}.
\end{aligned} \tag{66}$$

We next consider (67) at the bottom of the page, where (b) follows from the same reasoning as in the steps from (p) to (q) in (65).

We now consider

$$\begin{aligned}
&nR_0+nR_1\\
&\quad=H(W_0,W_1)=I\left(W_0,W_1;Y_1^n\right)+H\left(W_0,W_1|Y_1^n\right)\\
&\quad\le I\left(W_0,W_1;Y_1^n\right)+n\delta_{1,n}\\
&\quad=\sum_{i=1}^{n}I\left(W_0,W_1;Y_{1,i}|Y_1^{i-1}\right)+n\delta_{1,n}.
\end{aligned} \tag{68}$$

We proceed to bound $nR_0+nR_1$ and obtain

$$\begin{aligned}
&nR_0+nR_1\\
&\overset{(c)}{\le}\sum_{i=1}^{n}I\left(W_0,W_1;Y_{1,i}|Y_1^{i-1},X_{1,i}\right)+n\delta_{1,n}\\
&\le\sum_{i=1}^{n}\Bigg[H\left(Y_{1,i}|Y_1^{i-1},X_{1,i}\right)\\
&\qquad -H\left(Y_{1,i}|W_0,W_1,Y_1^{i-1},Y_2^{i-1},X_{1,i}\right)\Bigg]+n\delta_{1,n}\\
&\le\sum_{i=1}^{n}\left[H(Y_{1,i}|X_{1,i})-H\left(Y_{1,i}|U_i',X_{1,i}\right)\right]+n\delta_{1,n}\\
&\le\sum_{i=1}^{n}I\left(U_i';Y_{1,i}|X_{1,i}\right)+n\delta_{1,n}
\end{aligned} \tag{69}$$

where (c) follows from the same reasoning as in the steps from (p) to (q) in (65). We define $U_i':=\left(W_0,W_1,Y_1^{i-1},Y_2^{i-1}\right)$. Note that $X_{1,i}\to U_i'\to X_i$ form a Markov chain, and $U_i'\to(X_i,X_{1,i})\to(Y_{1,i},Y_{2,i})$ also form a Markov chain.

We finally consider (70) at the top of the following page, where (d) follows from the same reasoning as in the steps from (p) to (q) in (65).

Now in order to change the upper bounds that we have derived in to single-letter characterizations, we introduce a random variable $Q$ which is independent of $W_0$, $W_1$, $W_2$, $X^n$, $X_1^n$, $Y_1^n$, $Y_2^n$, and is uniformly distributed over $\{1,2,\ldots,n\}$. Define $U=(Q,U_Q)$, $U'=\left(Q,U_Q'\right)$, $X=X_Q$, $X_1=X_{1,Q}$, $Y_1=Y_{1,Q}$, and $Y_2=Y_{2,Q}$. Clearly, we have Markov chains: $X_1\to U\to X$, $X_1\to U'\to X$, and $(U,U')\to(X,X_1)\to(Y_1,Y_2)$. By using the above definitions, (64), (66), (67), (69), and (70) become

$$\begin{aligned}
R_0+R_2&\le\frac{1}{n}\sum_{i=1}^{n}I(U_i,X_{1,i};Y_{2,i})+\delta_{2,n}\\
&=I(U_Q,X_{1,Q};Y_{2,Q}|Q)+\delta_{2,n}\\
&\le I(Q,U_Q,X_{1,Q};Y_{2,Q})+\delta_{2,n}\\
&=I(U,X_1;Y_2)+\delta_{2,n}
\end{aligned} \tag{71}$$

$$\begin{aligned}
nR_1&=H(W_1)=I\left(W_1;Y_1^n\right)+H\left(W_1|Y_1^n\right)\\
&\le I\left(W_1;Y_1^n,Y_2^n,W_0,W_2\right)+n\delta_{1,n}\\
&=I\left(W_1;Y_1^n,Y_2^n|W_0,W_2\right)+n\delta_{1,n}\\
&=\sum_{i=1}^{n}I\left(W_1;Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},W_0,W_2\right)+n\delta_{1,n}\\
&\overset{(b)}{\le}\sum_{i=1}^{n}I\left(W_1;Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},W_0,W_2,X_{1,i}\right)+n\delta_{1,n}\\
&\le\sum_{i=1}^{n}\left[H(Y_{1,i},Y_{2,i}|U_i,X_{1,i})-H(Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},W_0,W_2,W_1,X_{1,i},X_i)\right]+n\delta_{1,n}\\
&=\sum_{i=1}^{n}\left[H(Y_{1,i},Y_{2,i}|U_i,X_{1,i})-H\left(Y_{1,i},Y_{2,i}|Y_1^{i-1},Y_2^{i-1},W_0,W_2,X_{1,i},X_i\right)\right]+n\delta_{1,n}\\
&=\sum_{i=1}^{n}\left[H(Y_{1,i},Y_{2,i}|U_i,X_{1,i})-H(Y_{1,i},Y_{2,i}|U_i,X_{1,i},X_i)\right]+n\delta_{1,n}\\
&=\sum_{i=1}^{n}I(X_i;Y_{1,i},Y_{2,i}|U_i,X_{1,i})+n\delta_{1,n}
\end{aligned} \tag{67}$$

$$\begin{aligned}
nR_2 &= H(W_2) = I\left(W_2; Y_2^n\right) + H\left(W_2|Y_2^n\right) \\
&\le I\left(W_2; Y_2^n\right) + n\delta_{2,n} \\
&\le I\left(W_2; Y_1^n, Y_2^n, W_0, W_1\right) + n\delta_{2,n} \\
&= I\left(W_2; Y_1^n, Y_2^n | W_0, W_1\right) + n\delta_{2,n} \\
&= \sum_{i=1}^{n} I\left(W_2; Y_{1,i}, Y_{2,i} | Y_1^{i-1}, Y_2^{i-1}, W_0, W_1\right) + n\delta_{2,n} \\
&\overset{(d)}{\le} \sum_{i=1}^{n} I\left(W_2; Y_{1,i}, Y_{2,i} | Y_1^{i-1}, Y_2^{i-1}, W_0, W_1, X_{1,i}\right) + n\delta_{2,n} \\
&\le \sum_{i=1}^{n} \left[ H\left(Y_{1,i}, Y_{2,i} | U_i', X_{1,i}\right) - H\left(Y_{1,i}, Y_{2,i} | Y_1^{i-1}, Y_2^{i-1}, W_0, W_1, W_2, X_{1,i}, X_i\right) \right] + n\delta_{2,n} \\
&= \sum_{i=1}^{n} \left[ H\left(Y_{1,i}, Y_{2,i} | U_i', X_{1,i}\right) - H\left(Y_{1,i}, Y_{2,i} | U_i', X_{1,i}, X_i\right) \right] + n\delta_{2,n} \\
&= \sum_{i=1}^{n} I\left(X_i; Y_{1,i}, Y_{2,i} | U_i', X_{1,i}\right) + n\delta_{2,n}
\end{aligned} \tag{70}$$

$$\begin{aligned}
R_0 + R_2 &\le \frac{1}{n} \sum_{i=1}^{n} I(U_i; Y_{1,i}, Y_{2,i} | X_{1,i}) + \delta_{2,n} \\
&= I(U_Q; Y_{1,Q}, Y_{2,Q} | X_{1,Q}, Q) + \delta_{2,n} \\
&\le I(Q, U_Q; Y_{1,Q}, Y_{2,Q} | X_{1,Q}) + \delta_{2,n} \\
&= I(U; Y_1, Y_2 | X_1) + \delta_{2,n}
\end{aligned} \tag{72}$$

$$\begin{aligned}
R_1 &\le \frac{1}{n} \sum_{i=1}^{n} I(X_i; Y_{1,i}, Y_{2,i} | U_i, X_{1,i}) + \delta_{1,n} \\
&= I(X_Q; Y_{1,Q}, Y_{2,Q} | U_Q, X_{1,Q}, Q) + \delta_{1,n} \\
&= I(X; Y_1, Y_2 | U, X_1) + \delta_{1,n}
\end{aligned} \tag{73}$$

$$\begin{aligned}
R_0 + R_1 &\le \frac{1}{n} \sum_{i=1}^{n} I\left(U_i'; Y_{1,i} | X_{1,i}\right) + \delta_{1,n} \\
&= I\left(U_Q'; Y_{1,Q} | X_{1,Q}, Q\right) + \delta_{1,n} \\
&\le I\left(Q, U_Q'; Y_{1,Q} | X_{1,Q}\right) + \delta_{1,n} \\
&= I\left(U'; Y_1 | X_1\right) + \delta_{1,n}
\end{aligned} \tag{74}$$

$$\begin{aligned}
R_2 &\le \frac{1}{n} \sum_{i=1}^{n} I\left(X_i; Y_{1,i}, Y_{2,i} | U_i', X_{1,i}\right) + \delta_{2,n} \\
&= I\left(X_Q; Y_{1,Q}, Y_{2,Q} | U_Q', X_{1,Q}, Q\right) + \delta_{2,n} \\
&= I\left(X; Y_1, Y_2 | U', X_1\right) + \delta_{2,n}.
\end{aligned} \tag{75}$$

This concludes the proof.

## Appendix III
## Proof of the Converse For Theorem 6

The proof is similar to that for Theorem 2 given in Appendix II. We hence provide only an outline, which will be useful in the following proof for the Gaussian case in Appendix IV.

We consider a sequence of $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n\right)$ code with $P_e^{(n)} \to 0$. Then the probability distribution on the joint ensemble space $W_0 \times W_1 \times W_2 \times X^n \times X_1^n \times Y_1^n \times Y_2^n$ is given by

$$\begin{aligned}
&p\left(w_0, w_1, w_2, x^n, x_1^n, y_1^n, y_2^n\right) \\
&\quad = p(w_0)p(w_1)p(w_2) \\
&\qquad \cdot \prod_{i=1}^{n} \Big[ p\left(x_i | w_0, w_1, w_2, y_1^{i-1}, y_2^{i-1}\right) \\
&\qquad\qquad \cdot p\left(x_{1,i} | y_1^{i-1}, y_2^{i-1}\right) p(y_{1,i}, y_{2,i} | x_i, x_{1,i}) \Big].
\end{aligned} \tag{76}$$

By Fano's inequality, we have

$$\begin{aligned}
H\left(W_1 | Y_1^n, Y_2^n\right) &\le H\left(W_0, W_1 | Y_1^n, Y_2^n\right) \\
&\le n(R_0 + R_1)P_e^{(n)} + 1 := n\delta_{1,n} \\
H\left(W_0, W_2 | Y_2^n\right) &\le n(R_0 + R_2)P_e^{(n)} + 1 := n\delta_{2,n}
\end{aligned} \tag{77}$$

where $\delta_{1,n}, \delta_{2,n} \to 0$ if $P_e^{(n)} \to 0$. The proof for the following two bounds follows steps that are identical to those in (64) and (66) in Appendix II.

$$nR_0 + nR_2 \le \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{2,i}) + n\delta_{2,n} \tag{78}$$

$$nR_0 + nR_2 \le \sum_{i=1}^{n} I(U_i; Y_{1,i}, Y_{2,i} | X_{1,i}) + n\delta_{2,n}. \tag{79}$$

We now consider

$$\begin{aligned}
nR_1 &= H(W_1) \\
&= I\left(W_1; Y_1^n, Y_2^n\right) + H\left(W_1 | Y_1^n, Y_2^n\right) \\
&\le I\left(W_1; Y_1^n, Y_2^n, W_0, W_2\right) + n\delta_{1,n} \\
&\le \sum_{i=1}^{n} I(X_i; Y_{1,i}, Y_{2,i} | U_i, X_{1,i}) + n\delta_{1,n}
\end{aligned} \tag{80}$$

where we omitted the steps that are identical to those in (67) in Appendix II.

## APPENDIX IV
## PROOF OF THE CONVERSE FOR THEOREM 7

The techniques in the proof of the converse for the capacity of the physically degraded Gaussian relay channel [2, Sec. IV] are useful here, but are not sufficient. In particular, the parameters $\alpha$ and $\beta$ need to be carefully chosen. Moreover, this proof applies the entropy power inequality to the components of the two random vectors, which is different from applying the entropy power inequality to two independent random vectors as in the proof of the converse for the capacity region of the degraded Gaussian broadcast channel. A similar idea has been used in establishing the capacity region of the physically degraded Gaussian broadcast channel with feedback [41].

For the D-AWGN partially cooperative RBC, the power constraints at the source and relay imply that the codewords satisfy

$$\sum_{i=1}^{n} \mathsf{E}\left[X_i^2\right] \le nP, \quad \sum_{i=1}^{n} \mathsf{E}\left[X_{1,i}^2\right] \le nP_1. \tag{81}$$

We apply the degradedness condition in Definition 3 to the bounds (78)–(80) in Appendix III and obtain the following bounds:

$$nR_0 + nR_2 \le \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{2,i}) + n\delta_{2,n} \tag{82}$$

$$nR_0 + nR_2 \le \sum_{i=1}^{n} I(U_i; Y_{1,i}|X_{1,i}) + n\delta_{2,n} \tag{83}$$

$$nR_1 \le \sum_{i=1}^{n} I(X_i; Y_{1,i}|U_i, X_{1,i}) + n\delta_{1,n}. \tag{84}$$

We now apply the bounds (82)–(84) for the D-AWGN partially cooperative RBC. We start with (83), and obtain

$$\begin{aligned} &nR_0 + nR_2 \\ &\quad \le \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}) - \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}, U_i) + n\delta_{2,n}. \end{aligned} \tag{85}$$

For the second term in (85), we have

$$\begin{aligned} \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}, U_i) &\le \sum_{i=1}^{n} h(Y_{1,i}) = \sum_{i=1}^{n} h(X_i + Z_{1,i}) \\ &\le \sum_{i=1}^{n} \frac{1}{2} \log 2\pi e \left(\mathsf{E}X_i^2 + N_1\right) \\ &\le \frac{n}{2} \log 2\pi e \left(\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}X_i^2 + N_1\right) \\ &\le \frac{n}{2} \log 2\pi e(P + N_1). \end{aligned} \tag{86}$$

On the other hand

$$\begin{aligned} \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}, U_i) &\ge \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}, U_i, X_i) \\ &= \sum_{i=1}^{n} h(Y_{1,i}|X_i) = \frac{n}{2} \log 2\pi e N_1 \end{aligned} \tag{87}$$

where we used that given $X_i$, $Y_{1,i}$ is independent of $X_{1,i}, U_i$.

Combining (86) and (87), we establish that there exists some $\alpha \in [0, 1]$ such that

$$\sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}, U_i) = \frac{n}{2} \log 2\pi e(\alpha P + N_1). \tag{88}$$

For the first term in (85)

$$\begin{aligned} &\sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}) \\ &\quad = \sum_{i=1}^{n} \mathsf{E}_{X_{1,i}} h(Y_{1,i}|X_{1,i}) \\ &\quad \le \sum_{i=1}^{n} \mathsf{E}_{X_{1,i}} \frac{1}{2} \log 2\pi e \mathsf{Var}(Y_{1,i}|X_{1,i}) \\ &\quad \le \frac{1}{2} \sum_{i=1}^{n} \log 2\pi e \mathsf{E}_{X_{1,i}}[\mathsf{Var}(X_i + Z_{1,i}|X_{1,i})] \\ &\quad = \frac{1}{2} \sum_{i=1}^{n} \log 2\pi e \mathsf{E}_{X_{1,i}}[\mathsf{Var}(X_i|X_{1,i}) + N_1] \\ &\quad \le \frac{n}{2} \log 2\pi e \left(\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}_{X_{1,i}}[\mathsf{Var}(X_i|X_{1,i})] + N_1\right) \\ &\quad = \frac{n}{2} \log 2\pi e \left(\frac{1}{n}\sum_{i=1}^{n} \left[\mathsf{E}\left(X_i^2\right) - \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right]\right] + N_1\right) \\ &\quad \le \frac{n}{2} \log 2\pi e \left(P - \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right] + N_1\right). \end{aligned} \tag{89}$$

On the other hand, we know that

$$\begin{aligned} \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}) &\ge \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}, U_i) \\ &= \frac{n}{2} \log 2\pi e(\alpha P + N_1) \end{aligned} \tag{90}$$

where we have used (88).

Combining (89) and (90), we obtain

$$\alpha P + N_1 \le P - \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right] + N_1 \tag{91}$$

and hence

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right] \le \bar{\alpha} P \tag{92}$$

where $\bar{\alpha} = 1 - \alpha$. Therefore, there exists some $\beta \in [0, 1]$ such that

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right] = \bar{\beta}\bar{\alpha} P \tag{93}$$

where $\bar{\beta} = 1 - \beta$.

We plug the preceding equation in (89), and obtain

$$\begin{aligned} \sum_{i=1}^{n} h(Y_{1,i}|X_{1,i}) &\le \frac{n}{2} \log 2\pi e \left(P - \bar{\beta}\bar{\alpha} P + N_1\right) \\ &= \frac{n}{2} \log 2\pi e \left(\alpha P + \beta\bar{\alpha} P + N_1\right). \end{aligned} \tag{94}$$

We plug (88) and (94) in (85), and obtain

$$\begin{aligned} nR_0 + nR_2 &\le \frac{n}{2}\log 2\pi e\left(\alpha P + \beta\bar{\alpha}P + N_1\right) \\ &\quad - \frac{n}{2}\log 2\pi e\left(\alpha P + N_1\right) + n\delta_{2,n} \\ &= \frac{n}{2}\log\left(1 + \frac{\beta\bar{\alpha}P}{\alpha P + N_1}\right) + n\delta_{2,n}. \end{aligned} \tag{95}$$

We next consider the bound (82), and obtain

$$nR_0 + nR_2 \le \sum_{i=1}^{n} h(Y_{2,i}) - \sum_{i=1}^{n} h(Y_{2,i}|U_i, X_{1,i}) + n\delta_{2,n}. \tag{96}$$

The first term in (96) can be bounded

$$\begin{aligned} \sum_{i=1}^{n} h(Y_{2,i}) &= \sum_{i=1}^{n} h\left(X_i + X_{1,i} + Z_{1,i} + Z_i'\right) \\ &\le \sum_{i=1}^{n} \frac{1}{2}\log 2\pi e\left(\mathsf{E}(X_i + X_{1,i})^2 + N_2\right) \\ &\le \frac{n}{2}\log 2\pi e\left(\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}(X_i + X_{1,i})^2 + N_2\right). \end{aligned} \tag{97}$$

For the sum in the preceding equation, we have

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}(X_i + X_{1,i})^2 \\ &\quad = \frac{1}{n}\sum_{i=1}^{n}\mathsf{E}X_i^2 + \frac{1}{n}\sum_{i=1}^{n}\mathsf{E}X_{1,i}^2 + \frac{2}{n}\sum_{i=1}^{n}\mathsf{E}X_iX_{1,i} \\ &\quad \le P + P_1 + \frac{2}{n}\sum_{i=1}^{n}\mathsf{E}\left[X_{1,i}\mathsf{E}(X_i|X_{1,i})\right] \\ &\quad \le P + P_1 + \frac{2}{n}\sum_{i=1}^{n}\sqrt{\mathsf{E}X_{1,i}^2 \cdot \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right]} \\ &\quad \le P + P_1 \\ &\qquad + 2\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}X_{1,i}^2\right)\cdot\left(\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right]\right)} \\ &\quad \le P + P_1 + 2\sqrt{P_1\bar{\beta}\bar{\alpha}P}. \end{aligned} \tag{98}$$

Hence, we obtain

$$\sum_{i=1}^{n} h(Y_{2,i}) \le \frac{n}{2}\log 2\pi e\left(P + P_1 + 2\sqrt{P_1\bar{\beta}\bar{\alpha}P} + N_2\right). \tag{99}$$

The second term in (96) can be expressed as

$$\begin{aligned} &\sum_{i=1}^{n} h(Y_{2,i}|U_i, X_{1,i}) \\ &\quad = \sum_{i=1}^{n} h\left(Y_{1,i} + X_{1,i} + Z_i'|U_i, X_{1,i}\right) \\ &\quad = \sum_{i=1}^{n} h(Y_{1,i} + Z_i'|U_i, X_{1,i}) \\ &\quad = \sum_{i=1}^{n} \mathsf{E}h\left(Y_{1,i} + Z_i'|U_i = u_i, X_{1,i} = x_{1,i}\right). \end{aligned} \tag{100}$$

We now use the entropy power inequality, and obtain

$$\begin{aligned} &2^{2h\left(Y_{1,i} + Z_i'|U_i = u_i, X_{1,i} = x_{1,i}\right)} \\ &\quad \ge 2^{2h\left(Y_{1,i}|U_i = u_i, X_{1,i} = x_{1,i}\right)} + 2^{2h\left(Z_i'|U_i = u_i, X_{1,i} = x_{1,i}\right)} \\ &\quad = 2^{2h\left(Y_{1,i}|U_i = u_i, X_{1,i} = x_{1,i}\right)} + 2\pi e(N_2 - N_1). \end{aligned}$$

Then

$$\begin{aligned} &h\left(Y_{1,i} + Z_i'|U_i = u_i, X_{1,i} = x_{1,i}\right) \\ &\quad \ge \frac{1}{2}\log\left(2^{2h\left(Y_{1,i}|U_i = u_i, X_{1,i} = x_{1,i}\right)} + 2\pi e(N_2 - N_1)\right). \end{aligned}$$

Hence

$$\begin{aligned} &\mathsf{E}h\left(Y_{1,i} + Z_i'|U_i = u_i, X_{1,i} = x_{1,i}\right) \\ &\quad \ge \frac{1}{2}\mathsf{E}\log\left(2^{2h\left(Y_{1,i}|U_i = u_i, X_{1,i} = x_{1,i}\right)} + 2\pi e(N_2 - N_1)\right) \\ &\quad \overset{(a)}{\ge} \frac{1}{2}\log\left(2^{2\mathsf{E}h\left(Y_{1,i}|U_i = u_i, X_{1,i} = x_{1,i}\right)} + 2\pi e(N_2 - N_1)\right) \\ &\quad \ge \frac{1}{2}\log\left(2^{2h\left(Y_{1,i}|U_i, X_{1,i}\right)} + 2\pi e(N_2 - N_1)\right) \end{aligned}$$

where (a) follows because $\log(2^x + c)$ is a convex function of $x$.

We plug the preceding equation in (100), and obtain

$$\begin{aligned} &\sum_{i=1}^{n} h(Y_{2,i}|U_i, X_{1,i}) \\ &\quad \ge \frac{1}{2}\sum_{i=1}^{n}\log\left(2^{2h\left(Y_{1,i}|U_i, X_{1,i}\right)} + 2\pi e(N_2 - N_1)\right) \\ &\quad \overset{(a)}{\ge} \frac{n}{2}\log\left(2^{2\frac{1}{n}\sum_{i=1}^{n} h\left(Y_{1,i}|U_i, X_{1,i}\right)} + 2\pi e(N_2 - N_1)\right) \\ &\quad \overset{(b)}{=} \frac{n}{2}\log\left(2\pi e(\alpha P + N_1) + 2\pi e(N_2 - N_1)\right) \\ &\quad = \frac{n}{2}\log\left(2\pi e(\alpha P + N_2)\right) \end{aligned} \tag{101}$$

where (a) also follows because $\log(2^x + c)$ is a convex function of $x$, and (b) follows from (88).

We plug (99) and (101) in (96), and obtain

$$\begin{aligned} &nR_0 + nR_2 \\ &\quad \le \frac{n}{2}\log\left(1 + \frac{\bar{\alpha}P + P_1 + 2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P + N_2}\right) + n\delta_{2,n}. \end{aligned} \tag{102}$$

We now consider (84), and obtain

$$\begin{aligned} nR_1 &\le \sum_{i=1}^{n}\left[h(Y_{1,i}|U_i, X_{1,i}) - h(Y_{1,i}|U_i, X_{1,i}, X_i)\right] + n\delta_{1,n} \\ &= \sum_{i=1}^{n}\left[h(Y_{1,i}|U_i, X_{1,i}) - h(Y_{1,i}|X_i)\right] + n\delta_{1,n} \\ &= \frac{n}{2}\log 2\pi e(\alpha P + N_1) - \frac{n}{2}\log(2\pi e N_1) + n\delta_{1,n} \\ &= \frac{n}{2}\log\left(1 + \frac{\alpha P}{N_1}\right) + n\delta_{1,n}. \end{aligned} \tag{103}$$

Therefore, (95), (102), and (103) provide the converse for Theorem 7.

## APPENDIX V
## OUTLINE OF PROOF FOR THEOREM 5

The proof uses the techniques in proving Theorem 8 and techniques in Appendix IV.

In the proof for Theorem 8, we have shown that the mapping from $(Y_1, Y_2)$ to $(S, Y_2)$ is one-to-one, where

$$\begin{aligned} S &= X + \frac{N_1}{N_1+N_2} X_1 + \hat{Z}_1 \\ Y_2 &= X + X_1 + \hat{Z}_1 + \hat{Z} \end{aligned} \tag{104}$$

where $\hat{Z}_1$ and $\hat{Z}$ are independent real Gaussian random variables with variances $\frac{N_1 N_2}{N_1+N_2}$ and $\frac{N_2^2}{N_2+N_1}$, respectively. We have also shown that given $(S, X_1)$, $Y_2$ is independent of $X$.

From (64)–(69) in Appendix II, we have the following upper bounds:

$$nR_0 + nR_2 = \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{2,i}) + n\delta_{2,n} \tag{105}$$

$$\begin{aligned} nR_0 + nR_2 &\le \sum_{i=1}^{n} I(U_i; Y_{1,i}, Y_{2,i}|X_{1,i}) + n\delta_{2,n} \\ &\overset{(a)}{=} \sum_{i=1}^{n} I(U_i; S_i, Y_{2,i}|X_{1,i}) + n\delta_{2,n} \\ &\overset{(b)}{=} \sum_{i=1}^{n} I(U_i; S_i|X_{1,i}) + n\delta_{2,n} \end{aligned} \tag{106}$$

$$\begin{aligned} nR_1 &\le \sum_{i=1}^{n} I(X_i; Y_{1,i}, Y_{2,i}|U_i, X_{1,i}) + n\delta_{1,n} \\ &\overset{(c)}{=} \sum_{i=1}^{n} I(X_i; S_i, Y_{2,i}|U_i, X_{1,i}) + n\delta_{1,n} \\ &\overset{(d)}{=} \sum_{i=1}^{n} I(X_i; S_i|U_i, X_{1,i}) + n\delta_{1,n} \end{aligned} \tag{107}$$

$$nR_0 + nR_1 \le \sum_{i=1}^{n} I\left(U_i'; Y_{1,i}|X_{1,i}\right) + n\delta_{1,n} \tag{108}$$

where (a) and (c) follow because the mapping from $(Y_{1,i}, Y_{2,i})$ to $(S_i, Y_{2,i})$ is one-to-one, and (b) and (d) follow because given $S_i$ and $X_{1,i}$, $Y_{2,i}$ is independent of $U_i$ and $X_i$.

We further bound (105)–(107) by following similar steps in bounding (82)–(84) in Appendix IV with $Y_1$ being replaced by $S$, and $N_1$ being replaced by $\frac{N_1 N_2}{N_1+N_2}$. We then obtain the following bounds:

$$nR_0 + nR_2 \le \frac{n}{2} \log\left(1 + \frac{P_1 + \bar{\alpha}P + 2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P + N_2}\right) + n\delta_{2,n} \tag{109}$$

$$nR_0 + nR_2 \le \frac{n}{2} \log\left(1 + \frac{\beta\bar{\alpha}P}{\alpha P + \frac{N_1 N_2}{N_1+N_2}}\right) + n\delta_{2,n}, \tag{110}$$

$$nR_1 \le \frac{n}{2} \log\left(1 + \frac{\alpha P}{\frac{N_1 N_2}{N_1+N_2}}\right) + n\delta_{1,n}, \tag{111}$$

where parameters $\alpha$ and $\beta$ are defined by

$$\sum_{i=1}^{n} h(S_i|X_{1,i}, U_i) = \frac{n}{2} \log 2\pi e \left(\alpha P + \frac{N_1 N_2}{N_1+N_2}\right)$$

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left(\left[\mathsf{E}^2(X_i|X_{1,i})\right]\right) = \bar{\beta}\bar{\alpha}P. \tag{112}$$

We now further bound (108), and obtain

$$\begin{aligned} &nR_0 + nR_1 \\ &\le \sum_{i=1}^{n} I\left(U_i'; Y_{1,i}|X_{1,i}\right) + n\delta_{1,n} \\ &= \sum_{i=1}^{n} \left[h(Y_{1,i}|X_{1,i}) - h\left(Y_{1,i}|U_i', X_{1,i}\right)\right] + n\delta_{1,n} \\ &\le \sum_{i=1}^{n} \mathsf{E}\left[h(X_i + Z_{1,i}|X_{1,i}) - h\left(Y_{1,i}|U_i', X_{1,i}, X_i\right)\right] + n\delta_{1,n} \\ &\le \sum_{i=1}^{n} \mathsf{E}_{X_{1,i}} \frac{1}{2} \log 2\pi e \mathsf{Var}(X_i + Z_{1,i}|X_{1,i}) \\ &\quad - \frac{n}{2} \log 2\pi e N_1 + n\delta_{1,n} \\ &\le \frac{1}{2}\sum_{i=1}^{n} \log 2\pi e(\mathsf{E}_{X_{1,i}}\left[\mathsf{Var}(X_i|X_{1,i})\right] + N_1) \\ &\quad - \frac{n}{2} \log 2\pi e N_1 + n\delta_{1,n} \\ &\le \frac{n}{2} \log 2\pi e \left(\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}_{X_{1,i}}\left[\mathsf{Var}(X_i|X_{1,i})\right] + N_1\right) \\ &\quad - \frac{n}{2} \log 2\pi e N_1 + n\delta_{1,n}. \end{aligned}$$

We proceed to bound $nR_0 + nR_1$ and obtain

$$\begin{aligned} &nR_0 + nR_1 \\ &\le \frac{n}{2} \log 2\pi e \left(\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left(X_i^2\right) - \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i})\right] + N_1\right) \\ &\quad - \frac{n}{2} \log 2\pi e N_1 + n\delta_{1,n} \\ &= \frac{n}{2} \log 2\pi e \left(P - \bar{\beta}\bar{\alpha}P + N_1\right) - \frac{n}{2} \log 2\pi e N_1 + n\delta_{1,n} \\ &= \frac{n}{2} \log\left(1 + \frac{\alpha P + \beta\bar{\alpha}P}{N_1}\right) + n\delta_{1,n}. \end{aligned} \tag{113}$$

Therefore, the bounds given in (109)–(111) and (113) provide the outer bound given in Theorem 5.

## APPENDIX VI
## OUTLINE OF PROOF FOR THEOREM 10

We assume that the source uses the superposition coding [47, Ch. 14.6 ]. We also assume that user 1 employs the decode-and-forward relaying scheme [2, Sec. II ], and user 2 employs the estimate-and-forward relaying scheme [2, Sec. VI ]. As in Appendix I, we adopt the regular encoding/sliding-window decoding strategy [52] for the decode-and-forward scheme. We also use the regular encoding/sliding-window decoding strategy for the estimate-and-forward scheme, which is different from the irregular encoding/successive decoding strategy originally used in [2, Sec. VI ].

We first prove that the rate region $\mathcal{R}_1$ is achievable for the case where $R_0 = 0$ (without common message $W_0$). It is then easily seen from the following proof that user 1 decodes the messages for both users 1 and 2. Hence, we can always view part of the rate

TABLE IV
CODEWORDS BEING SENT IN THE FIRST FOUR BLOCKS TO ACHIEVE $\mathcal{R}_1$ IN THEOREM 10 WITH $R_0 = 0$

| block 1 | block 2 | block 3 | block 4 |
|---|---|---|---|
| $\underline{x}_1(1)$ | $\underline{x}_1(w_{2,1})$ | $\underline{x}_1(w_{2,2})$ | $\underline{x}_1(w_{2,3})$ |
| $\underline{u}\ (1, w_{2,1})$ | $\underline{u}\ (w_{2,1}, w_{2,2})$ | $\underline{u}\ (w_{2,2}, w_{2,3})$ | $\underline{u}\ (w_{2,3}, w_{2,4})$ |
| $\underline{x}\ (1, w_{2,1}, w_{1,1})$ | $\underline{x}\ (w_{2,1}, w_{2,2}, w_{1,2})$ | $\underline{x}\ (w_{2,2}, w_{2,3}, w_{1,3})$ | $\underline{x}\ (w_{2,3}, w_{2,4}, w_{1,4})$ |
| $\underline{x}_2(\quad 1)$ | $\underline{x}_2(\quad 1)$ | $\underline{x}_2(\quad z_1)$ | $\underline{x}_2(\quad z_2)$ |
| $\hat{\underline{y}}_2\,(1, w_{2,1}, 1, z_1)$ | $\hat{\underline{y}}_2\,(w_{2,1}, w_{2,2}, 1, z_2)$ | $\hat{\underline{y}}_2\,(w_{2,2}, w_{2,3}, z_1, z_3)$ | $\hat{\underline{y}}_2\,(w_{2,3}, w_{2,4}, z_2, z_4)$ |

$R_2$ to be the common rate $R_0$. Therefore, we will have proven that the rate region $\mathcal{R}_1$ given in Theorem 10 is achievable. The achievability of $\mathcal{R}_2$ can be shown by the following steps similar to those used in proving the achievability of $\mathcal{R}_1$ with the roles of user 1 and user 2 being switched.

We consider a transmission over $B$ blocks, each with length $n$. At each of the first $B-2$ blocks, a message pair $(W_{1,i}, W_{2,i}) \in [1, 2^{nR_1}] \times [1, 2^{nR_2}]$ is encoded and sent from the source, where $i$ denotes the index of the block, and $i = 1, 2, \ldots, B-2$. The rate pair $\left(R_1 \frac{B-2}{B}, R_2 \frac{B-2}{B}\right)$ approaches $(R_1, R_2)$ as $B \to \infty$.

We use a random coding argument. Fix a probability distribution

$$p(x_1)p(u|x_1)p(x|u,x_1)p(x_2) \cdot p(y_1, y_2|x, x_1, x_2)p(\hat{y}_2|y_2, u, x_1, x_2) \quad (114)$$

and we use $T_\epsilon^n(P_{UX_1X_2XY_1Y_2\hat{Y}_2})$ to denote the jointly $\epsilon$-typical set based on the joint distribution (114).

Random Codebook Generation: We generate three statistically independent random codebooks by following the same steps.

1. Generate $2^{nR_2}$ i.i.d. $\underline{x}_1$ each with distribution $\prod_{i=1}^n p(x_{1,i})$. Index $\underline{x}_1\,(w_2')$, $w_2' \in [1, 2^{nR_2}]$.
2. For each $\underline{x}_1\,(w_2')$, generate $2^{nR_2}$ i.i.d. $\underline{u}$ each with distribution $\prod_{i=1}^n p(u_i|x_{1,i}\,(w_2'))$. Index $\underline{u}\,(w_2', w_2)$, $w_2 \in [1, 2^{nR_2}]$.
3. For each $\underline{x}_1\,(w_2')$ and $\underline{u}\,(w_2', w_2)$, generate $2^{nR_1}$ i.i.d. $\underline{x}$ each with distribution $\prod_{i=1}^n p(x_i|u_i\,(w_2', w_2), x_{1,i}\,(w_2'))$. Index $\underline{x}\,(w_2', w_2, w_1)$, $w_1 \in [1, 2^{nR_1}]$.
4. Generate $2^{n\hat{R}_1}$ i.i.d. $\underline{x}_2$ each with distribution $\prod_{i=1}^n p(x_{2,i})$. Index $\underline{x}_2(z')$, $z' \in [1, 2^{n\hat{R}_1}]$.
5. For each $\underline{x}_1\,(w_2')\,, \underline{u}\,(w_2', w_2)\,, \underline{x}_2(z')$, generate $2^{n\hat{R}_1}$ i.i.d. $\hat{\underline{y}}_2$ each with distribution

$$\prod_{i=1}^n p(\hat{y}_{2,i}|x_{1,i}\,(w_2')\,, u_i\,(w_2', w_2)\,, x_{2,i}(z'))$$

where the distribution $p(\hat{y}_2|u, x_1, x_2)$ is induced by the joint distribution given by (114). Index $\hat{\underline{y}}_2\,(w_2', w_2, z', z)$, $z \in [1, 2^{n\hat{R}_1}]$.

Encoding: We encode messages using codebooks 1, 2, and 3, respectively, for adjacent three blocks. This is because some of the following decoding steps are performed jointly over two or three adjacent blocks, and having independent codebooks makes the error events corresponding to these blocks independent, thus making the probabilities of these error events easy to calculate.

At the source, let $(w_{1,i}, w_{2,i})$ be the new message pair to be sent in current block $i$, and let $(w_{1,i-1}, w_{2,i-1})$ be the message pair being sent in previous block $i-1$. The source then sends $\underline{x}(w_{2,i-1}, w_{2,i}, w_{1,i})$.

At the beginning of block $i$, user 1 should have an estimation $\hat{w}_{2,i-1}$ of the message $w_{2,i-1}$ sent in the previous block $i-1$. It then sends $\underline{x}_1(\hat{w}_{2,i-1})$.

At the beginning of block $i$, user 2 should have an estimation $\hat{z}_{i-2}$ of the index $z_{i-2}$ of the compressed signal $\hat{\underline{y}}_2$. It then sends $\underline{x}_2(z_{i-2})$.

For convenience, we list the codewords sent in the first four blocks in Table IV.

Decoding: The decoding procedures at the end of block $i$ are as follows.

1. User 1, having known $w_{2,i-1}$, declares message $\hat{w}_{2,i}$ is sent if there is a unique $\hat{w}_{2,i}$ such that

$$\left(\underline{x}_1(w_{2,i-1}), \underline{u}(w_{2,i-1}, \hat{w}_{2,i}), \underline{y}_1(i)\right) \in T_\epsilon^n(P_{UX_1Y_1}).$$

The decoding error in this step is small for sufficiently large $n$ if

$$R_2 < I(U; Y_1|X_1). \quad (115)$$

2. User 1, having known $w_{2,i-3}, \ldots, w_{2,i}$ and $z_{i-4}$, determines that $\hat{\underline{y}}_2$ indexed by $\hat{z}_{i-2}$ is picked to compress $\underline{y}_2(i-2)$ by user 2 based on the information received in blocks $i-2$ and $i$. User 1 declares the index to be $\hat{z}_{i-2}$ if there is a unique $\hat{z}_{i-2}$ such that

$$\left(\underline{x}_1(w_{2,i-3}), \underline{u}(w_{2,i-3}, w_{2,i-2}), \underline{x}_2(z_{i-4}), \underline{y}_1(i-2), \hat{\underline{y}}_2(w_{2,i-3}, w_{2,i-2}, z_{i-4}, \hat{z}_{i-2})\right) \in T_\epsilon^n(P_{UX_1X_2Y_1\hat{Y}_2}),$$

and $\left(\underline{x}_1(w_{2,i-1}), \underline{u}(w_{2,i-1}, w_{2,i}), \underline{x}_2(\hat{z}_{i-2}), \underline{y}_1(i)\right) \in T_\epsilon^n(P_{UX_1X_2Y_1}).$

The decoding error in this step is small for sufficiently large $n$ if

$$\hat{R}_1 < I(\hat{Y}_2; Y_1|U, X_1, X_2) + I(X_2; Y_1|X_1, U). \quad (116)$$

3. User 1, having known $w_{2,i-3}, w_{2,i-2}$ and $z_{i-4}, z_{i-2}$, determines that the message $\hat{w}_{1,i-2}$ is sent based on the information received in block $i-2$. It declares the index to be $\hat{w}_{1,i-2}$ if there is a unique $\hat{w}_{1,i-2}$ such that

$$\left(\underline{x}_1(w_{2,i-3}), \underline{u}(w_{2,i-3}, w_{2,i-2}), \underline{x}(w_{2,i-3}, w_{2,i-2}, \hat{w}_{1,i-2}), \underline{x}_2(z_{i-4}), \hat{\underline{y}}_2(w_{2,i-3}, w_{2,i-2}, z_{i-4}, z_{i-2}), \underline{y}_1(i-2)\right) \in T_\epsilon^n(P_{UX_1X_2XY_1\hat{Y}_2}).$$

The decoding error in this step is small for sufficiently large $n$ if

$$R_1 < I(X; \hat{Y}_2, Y_1|U, X_1, X_2). \quad (117)$$

4. User 2, having known $w_{2,i-2}$, $z_{i-3}$, and $z_{i-2}$, determines that the message $\hat{\hat{w}}_{2,i-1}$ is sent based on the information received in blocks $i-1$ and $i$. It declares the index to be $\hat{\hat{w}}_{2,i-1}$ if there is a unique $\hat{\hat{w}}_{2,i-1}$ such that

$$\left(\underline{x}_1(w_{2,i-2}), \underline{u}(w_{2,i-2}, \hat{\hat{w}}_{2,i-1}), \underline{x}_2(z_{i-3}), \underline{y}_2(i-1)\right) \in T_\epsilon^n(P_{UX_1X_2Y_2})$$
$$\text{and } \left(\underline{x}_1(\hat{\hat{w}}_{2,i-1}), \underline{x}_2(z_{i-2}), \underline{y}_2(i)\right) \in T_\epsilon^n(P_{X_1X_2Y_2}).$$

The decoding error in this step is small for sufficiently large $n$ if

$$R_2 < I(U;Y_2|X_1,X_2) + I(X_1;Y_2|X_2) = I(X_1,U;Y_2|X_2). \tag{118}$$

5. User 2, having known $w_{2,i-2}$, $w_{2,i-1}$, and $z_{i-3}$, declares that the estimate signal $\hat{\underline{y}}_2$ for $\underline{y}_2(i-1)$ is indexed by $\hat{\hat{z}}_{i-1}$ if there is a unique $\hat{\hat{z}}_{i-1}$ such that

$$\left(\underline{x}_1(w_{2,i-2}), \underline{u}(w_{2,i-2}, w_{2,i-1}), \underline{x}_2(z_{i-3}), \hat{\underline{y}}_2(w_{2,i-2}, w_{2,i-1}, z_{i-3}, \hat{\hat{z}}_{i-1}), \underline{y}_2(i-1)\right) \in T_\epsilon^n(P_{UX_1X_2Y_2\hat{Y}_2}).$$

There exists such a $z_{i-1}$ with high probability for sufficiently large $n$ if

$$\hat{R}_1 > I(\hat{Y}_2;Y_2|U,X_1,X_2). \tag{119}$$

Combining (116) and (119), we obtain

$$\begin{aligned} &I(X_2;Y_1|X_1,U) \\ &\quad > I(\hat{Y}_2;Y_2|U,X_1,X_2) - I(\hat{Y}_2;Y_1|U,X_1,X_2) \\ &\quad = I(\hat{Y}_2;Y_2|Y_1,U,X_1,X_2), \end{aligned} \tag{120}$$

which is exactly the constraint given in $\mathcal{R}_1$ in Theorem 10.

Combining (115), (117), (118), and (120), we obtain the rate region $\mathcal{R}_1$.

## Appendix VII
## Outline of Proof for Theorem 14

Let $\hat{Y}_2 = Y_2 + \check{Z}$ where the variance of $\check{Z}$ is denoted by $\check{N}$ that will be determined later in the proof.

We compute the achievable rate region $\mathcal{R}_1$ given in Theorem 10 based on the following distributions and relationships for those random variables in the expression of $\mathcal{R}_1$:

$$\begin{aligned} X_1 &\sim \mathcal{N}(0,P_1) \\ U' &\sim \mathcal{N}(0,\beta\bar{\alpha}P) \\ U &= \sqrt{\frac{\bar{\beta}\bar{\alpha}P}{P_1}}X_1 + U' \\ X' &\sim \mathcal{N}(0,\alpha P) \\ X &= U + X' \\ X_2 &\sim \mathcal{N}(0,\eta P_2) \end{aligned} \tag{121}$$

where the random variables $X_1$, $U'$, $X'$, $X_2$ are independent.

It is straightforward to compute the following two mutual information terms that provide the expression for $R_0 + R_2$:

$$\begin{aligned} I(U,X_1;Y_2|X_2) &= \mathcal{C}\left(\frac{\bar{\alpha}P + P_1 + 2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P + N_2}\right) \\ I(U;Y_1|X_1) &= \mathcal{C}\left(\frac{\beta\bar{\alpha}P}{\alpha P + \eta P_2 + N_1}\right). \end{aligned} \tag{122}$$

To derive $R_1$, we first have

$$R_1 < I(X;\hat{Y}_2,Y_1|X_1,U,X_2) = \mathcal{C}\left(\frac{\alpha P}{N_1} + \frac{\alpha P}{N_2 + \check{N}}\right). \tag{123}$$

To determine $\check{N}$ in the preceding equation, we use the following constraint which is given in the expression of $\mathcal{R}_1$:

$$\begin{aligned} I(X_2;Y_1|U,X_1) &\geq I(\hat{Y}_2;Y_2|Y_1,U,X_1,X_2) \\ &= I(\hat{Y}_2;Y_2|U,X_1,X_2) \\ &\quad - I(\hat{Y}_2;Y_1|U,X_1,X_2). \end{aligned} \tag{124}$$

We evaluate the mutual information terms in (124), and have

$$\begin{aligned} I(X_2;Y_1|U,X_1) &= \frac{1}{2}\log\left(1 + \frac{\eta P_2}{\alpha P + N_1}\right), \\ I(\hat{Y}_2;Y_2|U,X_1,X_2) &= \frac{1}{2}\log\left(1 + \frac{\alpha P + N_2}{\check{N}}\right), \\ I(\hat{Y}_2;Y_1|U,X_1,X_2) &= \frac{1}{2}\log\left(\frac{\alpha P + N_2 + \check{N}}{\alpha P + N_2 + \check{N} - \frac{(\alpha P)^2}{\alpha P + N_1}}\right). \end{aligned} \tag{125}$$

We plug the three mutual information terms given in (125) into (124), and derive the following constraint on $\check{N}$:

$$\check{N} \geq \frac{\alpha P(N_1 + N_2) + N_1N_2}{\eta P_2}. \tag{126}$$

We now plug the preceding bound on $\check{N}$ in the expression (123), and obtain

$$R_1 < \mathcal{C}\left(\frac{\alpha P}{N_1} + \frac{\alpha\eta PP_2}{\eta P_2N_2 + \alpha P(N_1 + N_2) + N_1N_2}\right) \tag{127}$$

which concludes the proof.

## Appendix VIII
## Outline of Proof for Theorem 15

In the proof for Theorem 18, we have shown that the mapping from $(Y_1, Y_2)$ to $(S, Y_2)$ is one-to-one, where

$$\begin{aligned} S &= X + \frac{N_1}{N_1 + N_2}X_1 + \frac{N_2}{N_1 + N_2}X_2 + \hat{Z}_1 \\ Y_2 &= X + X_1 + \hat{Z}_1 + \hat{Z} \end{aligned} \tag{128}$$

where $\hat{Z}_1$ and $\hat{Z}$ are independent real Gaussian random variables with variances $\frac{N_1N_2}{N_1+N_2}$ and $\frac{N_2^2}{N_2+N_1}$, respectively.

We define the following two auxiliary random variables:

$$\begin{aligned} U_i &:= \left(W_0, W_2, Y_1^{i-1}, Y_2^{i-2}\right) \\ U_i' &:= \left(W_0, W_1, Y_1^{i-1}, Y_2^{i-2}\right). \end{aligned} \tag{129}$$

We follow the steps similar to those in Appendix II, and derive the following upper bounds:

$$nR_0 + nR_2 \leq \sum_{i=1}^{n} I(U_i, X_{1,i}; Y_{2,i}|X_{2,i}) + n\delta_{2,n} \tag{130}$$

$$\begin{aligned} nR_0 + nR_2 &\leq \sum_{i=1}^{n} I(U_i; Y_{1,i}, Y_{2,i}|X_{1,i}, X_{2,i}) + n\delta_{2,n} \\ &= \sum_{i=1}^{n} I(U_i; S_i, Y_{2,i}|X_{1,i}, X_{2,i}) + n\delta_{2,n} \\ &= \sum_{i=1}^{n} I(U_i; S_i|X_{1,i}, X_{2,i}) + n\delta_{2,n} \end{aligned} \tag{131}$$

$$\begin{aligned} nR_1 &\leq \sum_{i=1}^{n} I(X_i; Y_{1,i}, Y_{2,i}|U_i, X_{1,i}, X_{2,i}) + n\delta_{1,n} \\ &= \sum_{i=1}^{n} I(X_i; S_i, Y_{2,i}|U_i, X_{1,i}, X_{2,i}) + n\delta_{1,n} \\ &= \sum_{i=1}^{n} I(X_i; S_i|U_i, X_{1,i}, X_{2,i}) + n\delta_{1,n} \end{aligned} \tag{132}$$

$$nR_0 + nR_1 \leq \sum_{i=1}^{n} I\left(U_i', X_{2,i}; Y_{1,i}|X_{1,i}\right) + n\delta_{1,n} \tag{133}$$

$$\begin{aligned} nR_0 + nR_1 &\leq \sum_{i=1}^{n} I\left(U_i'; Y_{1,i}, Y_{2,i}|X_{1,i}, X_{2,i}\right) + n\delta_{1,n} \\ &= \sum_{i=1}^{n} I\left(U_i'; S_i, Y_{2,i}|X_{1,i}, X_{2,i}\right) + n\delta_{1,n} \\ &= \sum_{i=1}^{n} I\left(U_i'; S_i|X_{1,i}, X_{2,i}\right) + n\delta_{1,n} \end{aligned} \tag{134}$$

$$\begin{aligned} nR_2 &\leq \sum_{i=1}^{n} I\left(X_i; Y_{1,i}, Y_{2,i}|U_i', X_{1,i}, X_{2,i}\right) + n\delta_{2,n} \\ &= \sum_{i=1}^{n} I\left(X_i; S_i, Y_{2,i}|U_i', X_{1,i}, X_{2,i}\right) + n\delta_{2,n} \\ &= \sum_{i=1}^{n} I\left(X_i; S_i|U_i', X_{1,i}, X_{2,i}\right) + n\delta_{2,n} \end{aligned} \tag{135}$$

where for (131), (132), (134), and (135), we have used the fact that the mapping from $(Y_1, Y_2)$ to $(S, Y_2)$ is one-to-one and the fact that given $(S_i, X_{1,i}, X_{2,i})$, $Y_{2,i}$ is independent of $U_i$, $U_i'$ and $X_i$.

We further bound (130)–(132) by following similar steps in bounding (82)–(84) in Appendix IV with $Y_1$ being replaced by $S$ and $N_1$ being replaced by $\frac{N_1 N_2}{N_1+N_2}$. We then obtain the following bounds:

$$\begin{aligned} nR_0 + nR_2 &\leq \frac{n}{2}\log\left(1 + \frac{P_1 + \bar{\alpha}P + 2\sqrt{\bar{\beta}\bar{\alpha}PP_1}}{\alpha P + N_2}\right) + n\delta_{2,n} \\ nR_0 + nR_2 &\leq \frac{n}{2}\log\left(1 + \frac{\beta\bar{\alpha}P}{\alpha P + \frac{N_1 N_2}{N_1+N_2}}\right) + n\delta_{2,n} \\ nR_1 &\leq \frac{n}{2}\log\left(1 + \frac{\alpha P}{\frac{N_1 N_2}{N_1+N_2}}\right) + n\delta_{1,n} \end{aligned} \tag{136}$$

where parameters $\alpha$ and $\beta$ are defined by

$$\begin{aligned} \sum_{i=1}^{n} h(S_i|X_{1,i}, X_{2,i}, U_i) &= \frac{n}{2}\log 2\pi e\left(\alpha P + \frac{N_1 N_2}{N_1 + N_2}\right) \\ \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i}, X_{2,i})\right] &= \bar{\beta}\bar{\alpha}P. \end{aligned} \tag{137}$$

We also obtain the following intermediate bound which will be useful for the rest of the proof:

$$\sum_{i=1}^{n} h(S_i|X_{1,i}, X_{2,i}) \leq \frac{n}{2}\log 2\pi e\left(\alpha P + \beta\bar{\alpha}P + \frac{N_1 N_2}{N_1 + N_2}\right). \tag{138}$$

We now further bound (133)

$$\begin{aligned} &nR_0 + nR_1 \\ &\quad\leq \sum_{i=1}^{n} I\left(U_i', X_{2,i}; Y_{1,i}|X_{1,i}\right) + n\delta_{1,n} \\ &\quad= \sum_{i=1}^{n} \left[h(Y_{1,i}|X_{1,i}) - h\left(Y_{1,i}|U_i', X_{1,i}, X_{2,i}\right)\right] + n\delta_{1,n} \\ &\quad\leq \sum_{i=1}^{n} \left[h(X_i + X_{2,i} + Z_{1,i}) - h\left(Y_{1,i}|U_i', X_{1,i}, X_{2,i}, X_i\right)\right] \\ &\qquad + n\delta_{1,n} \\ &\quad\leq \sum_{i=1}^{n} \frac{1}{2}\log 2\pi e(\mathsf{E}(X_i + X_{2,i})^2 + N_1) \\ &\qquad - \frac{n}{2}\log 2\pi e N_1 + n\delta_{1,n} \\ &\quad\leq \frac{n}{2}\log 2\pi e\left(\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}(X_i + X_{2,i})^2 + N_1\right) \\ &\qquad - \frac{n}{2}\log 2\pi e N_1 + n\delta_{1,n}. \end{aligned} \tag{139}$$

For the sum in the preceding equation, we have (140) at the top of the following page. Hence, we have

$$nR_0 + nR_1 \leq \frac{n}{2}\log\left(1 + \frac{P + P_2 + 2\sqrt{\bar{\beta}\bar{\alpha}PP_2}}{N_1}\right) + n\delta_{1,n}. \tag{141}$$

We next bound (134), and obtain

$$\begin{aligned} nR_0 + nR_1 &\leq \sum_{i=1}^{n} I\left(U_i'; S_i|X_{1,i}, X_{2,i}\right) + n\delta_{1,n} \\ &= \sum_{i=1}^{n} h(S_i|X_{1,i}, X_{2,i}) \\ &\qquad - \sum_{i=1}^{n} h\left(S_i|U_i', X_{1,i}, X_{2,i}\right) + n\delta_{1,n}. \end{aligned} \tag{142}$$

For the second term in (142), we have

$$\begin{aligned} &\sum_{i=1}^{n} h(S_i|U_i', X_{1,i}, X_{2,i}) \\ &\quad\leq \sum_{i=1}^{n} h(S_i|X_{1,i}, X_{2,i}) \\ &\quad\leq \frac{n}{2}\log 2\pi e\left(\alpha P + \beta\bar{\alpha}P + \frac{N_1 N_2}{N_1 + N_2}\right) \end{aligned} \tag{143}$$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}(X_i+X_{2,i})^2 &= \frac{1}{n}\sum_{i=1}^{n}\mathsf{E}X_i^2+\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}X_{2,i}^2+\frac{2}{n}\sum_{i=1}^{n}\mathsf{E}X_iX_{2,i}\\ &\le P+P_2+\frac{2}{n}\sum_{i=1}^{n}\mathsf{E}\left(X_{2,i}\mathsf{E}(X_i|X_{1,i},X_{2,i})\right)\\ &\le P+P_2+\frac{2}{n}\sum_{i=1}^{n}\sqrt{\mathsf{E}X_{2,i}^2\cdot\mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i},X_{2,i})\right]}\\ &\le P+P_2+2\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}X_{2,i}^2\right)\cdot\left(\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\left[\mathsf{E}^2(X_i|X_{1,i},X_{2,i})\right]\right)}\\ &\le P+P_2+2\sqrt{P_2\bar{\beta}\bar{\alpha}P}. \end{aligned}\tag{140}$$

and

$$\begin{aligned}\sum_{i=1}^{n}h\left(S_i|U_i',X_{1,i},X_{2,i}\right) &\ge \sum_{i=1}^{n}h\left(S_i|U_i',X_{1,i},X_{2,i},X_i\right)\\ &= \frac{n}{2}\log 2\pi e\frac{N_1N_2}{N_1+N_2}.\end{aligned}\tag{144}$$

Hence, there exists a $\gamma\in[0,1]$ such that

$$\begin{aligned}&\sum_{i=1}^{n}h\left(S_i|U_i',X_{1,i},X_{2,i}\right)\\ &\quad=\frac{n}{2}\log 2\pi e\left(\gamma(\alpha P+\beta\bar{\alpha}P)+\frac{N_1N_2}{N_1+N_2}\right).\end{aligned}\tag{145}$$

We apply (138) and (145) to (142) and obtain

$$\begin{aligned}&nR_0+nR_1\\ &\quad\le\frac{n}{2}\log 2\pi e\left(\alpha P+\beta\bar{\alpha}P+\frac{N_1N_2}{N_1+N_2}\right)\\ &\qquad-\frac{n}{2}\log 2\pi e\left(\gamma(\alpha P+\beta\bar{\alpha}P)+\frac{N_1N_2}{N_1+N_2}\right)+n\delta_{1,n}\\ &\quad\le\frac{n}{2}\log\left(1+\frac{\bar{\gamma}(\alpha P+\beta\bar{\alpha}P)}{\gamma(\alpha P+\beta\bar{\alpha}P)+\frac{N_1N_2}{N_1+N_2}}\right)+n\delta_{1,n}.\end{aligned}\tag{146}$$

We finally bound (135), and obtain

$$\begin{aligned}nR_2 &\le \sum_{i=1}^{n}I(X_i;S_i|U_i',X_{1,i},X_{2,i})+n\delta_{2,n}\\ &=\sum_{i=1}^{n}\left[h\left(S_i|U_i',X_{1,i},X_{2,i}\right)-h\left(S_i|U_i',X_{1,i},X_{2,i},X_i\right)\right]\\ &\quad+n\delta_{2,n}\\ &=\frac{n}{2}\log 2\pi e\left(\gamma(\alpha P+\beta\bar{\alpha}P)+\frac{N_1N_2}{N_1+N_2}\right)\\ &\quad-\frac{n}{2}\log 2\pi e\frac{N_1N_2}{N_1+N_2}+n\delta_{2,n}\\ &=\frac{n}{2}\log\left(1+\frac{\gamma(\alpha P+\beta\bar{\alpha}P)}{\frac{N_1N_2}{N_1+N_2}}\right)+n\delta_{2,n}.\end{aligned}\tag{147}$$


## ACKNOWLEDGMENT

The authors would like to thank Dr. Gerhard Kramer at Bell Labs, Alcatel-Lucent for useful discussions. They would also like to thank the reviewers for their comments and suggestions, which have led to improvements in both the content and presentation of the paper. In particular, the authors are grateful to reviewer B for suggesting the inclusion of Tables I and II in Section I to summarize our results.


## REFERENCES


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

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

[3] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.

[4] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Lower bounds on the capacity of Gaussian relay channel," in *Proc. Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2004, pp. 597–602.

[5] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[6] ——, "User cooperation diversity-part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

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

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

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

[10] D. Gunduz and E. Erkip, "Outage minimization by opportunistic cooperation," in *Proc. 2005 Int. Conf. Wireless Networks, Communications and Mobile Computing*, Maui, HI, Jun. 2005, vol. 2, pp. 1436–1442.

[11] Y. Yao, X. Cai, and G. B. Giannakis, "On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2917–2927, Nov. 2005.

[12] Y. Liang and V. V. Veeravalli, "Resource allocation for wireless relay channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2004, vol. 2, pp. 1902–1906.

[13] U. Mitra and A. Sabharwal, "On achievable rates of complexity constrained relay channels," in *Proc. 41st Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep./Oct. 2003, pp. 551–560.

[14] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[15] A. A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

[16] M. A. Khojastepour and B. Aazhang, "'Cheap' relay channel: A unifying approach to time and frequency division relaying," in *Proc. 42nd Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep./Oct. 2004, pp. 1792–1801.

[17] R. U. Nabar, H. Bölckskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[18] G. Kramer, "Models and theory for relay channels with receiver constraints," in *Proc. 42nd Annu.Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep./Oct. 2004, pp. 1312–1321.

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

[20] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.

[21] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000, p. 22.

[22] I. Maric and R. Yates, "Forwarding strategies for Gaussian parallel-relay networks," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 269.

[23] ——, "Bandwidth and power allocation for cooperative strategies in Gaussian relay networks," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2004, vol. 2, pp. 1907–1911.

[24] P. Gupta and P. R. Kumar, "Toward an information theory of large networks: An achievable rate region," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1877–1894, Aug. 2003.

[25] L.-L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[26] ——, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[27] A. Reznik, S. Kulkarni, and S. Verdú, "Degraded Gaussian multirelay channel: Capacity and optimal power allocation," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3037–3046, Dec. 2004.

[28] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[29] N. Jindal, U. Mitra, and A. Goldsmith, "Capacity of ad-hoc networks with node cooperation," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 271.

[30] M. Katz and S. Shamai (Shitz), "Transmitting to colocated users in wireless *ad hoc* and sensor networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3540–3563, Oct. 2005.

[31] J. N. Laneman and G. W. Wornell, "Distributed space-time coded protocols for exploring cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[32] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1470–1476, Sep. 2004.

[33] N. Pradad and M. K. Varanasi, "Diversity and multiplexing tradeoff bounds for cooperative diversity protocols," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 268.

[34] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoffs in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, submitted for publication.

[35] G. Kramer and A. J. van Wijngaarden, "On the white Gaussian multiple-access relay channel," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000, p. 40.

[36] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Capacity theorems for the multiple-access relay channel," in *Proc. 42nd Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep./Oct. 2004, pp. 1782–1791.

[37] ——, "Hierarchical wireless networks: Capacity bounds using the constrained multiple-access relay channel model," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2004, vol. 2, pp. 1912–1916.

[38] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

[39] ——, "Comments on broadcast channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.

[40] A. A. El Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 379–381, May 1978.

[41] ——, "The capacity of the physically degraded Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 4, pp. 508–511, Jul. 1981.

[42] L. Ozarow, "Coding and Capacity for Additive White Gaussian Noise Multi-User Channels With Feedback," Ph.D. dissertation, Dept. Elec. Eng. Comp. Sci., MIT, Cambridge, MA, 1983.

[43] R. Dabora and S. Servetto, "Broadcast channels with cooperating receivers: A downlink for the sensor reachback problem," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 176.

[44] ——, "Broadcast channels with cooperating decoders," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5438–5454, Dec. 2006.

[45] Y. Liang and V. V. Veeravalli, "The impact of relaying on the capacity of broadcast channels," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 403.

[46] Y. Liang, "Multiuser Communications With Relaying and User Cooperation," Ph.D. dissertation, Dept. Elec. Comp. Eng., Univ. Illinois at Urbana-Champaign, Urbana, IL, 2005.

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

[48] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 667–671, Jul. 1984.

[49] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, Jun. 2002.

[50] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Trans. Automat. Contr*, vol. 49, no. 9, pp. 1477–1488, Sep. 2004.

[51] Y. Liang and V. V. Veeravalli, "Distributed optimal resource allocation for fading relay broadcast channels," in *Proc. IEEE Int. Workshop on Signal Processing Advances for Wireless Communication (SPAWC)*, New York, Jun. 2005, pp. 550–554.

[52] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 841–850, Nov. 1982.

[53] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Akadémiai Kiadó, 1981.