Equivalence of Two Inner Bounds on the Capacity Region of the Two-Receiver Broadcast Channel

Yingbin Liang, Gerhard Kramer and H. Vincent Poor

***Abstract*— A recent inner bound on the capacity region of the two-receiver discrete memoryless broadcast channel is shown to be equivalent to the Marton-Gelfand-Pinsker region. The proof method is based on a result of Gelfand and Pinsker concerning channel input distributions.**

I. Introduction

The broadcast channel is introduced in [1], in which one transmitter sends common and individual information to multiple receivers. The performance measure of interest is the capacity region that characterizes the simultaneously and reliably achievable communication rates. Although the capacity region is obtained for many special cases, e.g., [2]–[11], it is still unknown for the general discrete memoryless model even for the simplest two-receiver case. Inner bounds on the capacity region have been obtained in, e.g., [1], [12]–[14], and outer bounds have been obtained in, e.g., [7], [13], [15]–[17].

In this paper, we focus on capacity inner bounds for two receivers. Marton's region [13, Th. 2] is the largest inner bound without a common message. Marton's region was extended to include a common message and appears in [18, P. 391, Prob. 10(c)] and [7, Th. 1]. We call this region the Marton-Gelfand-Pinsker (MGP) region. Another inner bound was derived recently in [14] that includes (but may not be strictly larger than) the MGP region. In this paper, we first review the two inner bounds and then show that the two bounds are equivalent. The technique we exploit is based on a property developed in [7, Prop. 1].

II. Channel Model

The two-receiver discrete memoryless broadcast channel depicted in Fig. 1 includes a transmitter and two receivers (receivers 1 and 2). The transmitter has a common message W_0 for both receivers, and private messages W_1 and W_2 for receivers 1 and 2, respectively. The messages W_0, W_1 and W_2 are independent of each other and are uniformly distributed over the message sets $\mathcal{W}_0$, $\mathcal{W}_1$, and $\mathcal{W}_2$, respectively. Let $\mathcal{X}$ be the channel input alphabet, $\mathcal{Y}_1$ and $\mathcal{Y}_2$ be the channel output alphabets of receivers 1 and 2, respectively, and $\mathcal{X}^n$ be the n-fold Cartesian product of $\mathcal{X}$. An encoder at the transmitter, i.e., f: $\mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathcal{X}^n$, maps each message triple $(w_0, w_1, w_2) \in \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2$ to a codeword $x^n \in \mathcal{X}^n$. The symbols x^n are transmitted over a broadcast channel with the transition probability $P_{Y_1Y_2|X}(\cdot)$ so there are two output sequences y_1^n and y_2^n at receivers 1 and 2, respectively. A decoder at receiver 1, i.e., g_1: $\mathcal{Y}_1^n \rightarrow \mathcal{W}_0 \times \mathcal{W}_1$, maps the received sequence $y_1^n \in \mathcal{Y}_1^n$ to a message pair $(\hat{w}_0^{(1)}, \hat{w}_1) \in \mathcal{W}_0 \times \mathcal{W}_1$, and a decoder at receiver 2, i.e., g_2: $\mathcal{Y}_2^n \rightarrow \mathcal{W}_0 \times \mathcal{W}_2$, maps the received sequence $y_2^n \in \mathcal{Y}_2^n$ to a message pair $(\hat{w}_0^{(2)}, \hat{w}_2) \in \mathcal{W}_0 \times \mathcal{W}_2$.

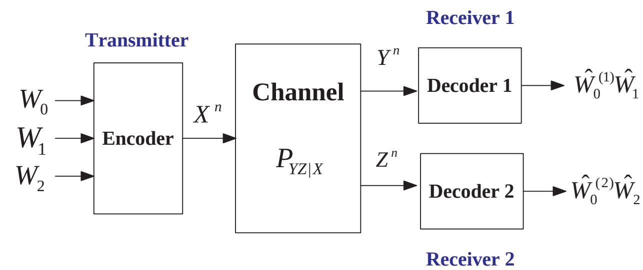

Fig. 1. The two-receiver broadcast channel

The average block probability of error for a length n code is defined as

$$P_e^{(n)} = \Pr\left\{(\hat{W}_0^{(1)}, \hat{W}_0^{(2)}, \hat{W}_1, \hat{W}_2) \neq (W_0, W_0, W_1, W_2)\right\}.$$

The rate triple (R_0, R_1, R_2) is *achievable* if there exists a sequence of message sets $(\mathcal{W}_{0n}, \mathcal{W}_{1n}, \mathcal{W}_{2n})$ with $|\mathcal{W}_{kn}| = 2^{nR_k}$ for $k = 0, 1, 2$, and encoder-decoder triples (f_n, g_{1n}, g_{2n}) such that the average error probability $P_e^{(n)} \rightarrow 0$ as n goes to infinity. The capacity region is the closure of the set of achievable rate triples.

III. Preliminary and Main Results

The first inner bound we consider is the MGP region in [18, P. 391, Prob. 10(c)] and [7, Th. 1], which is an extension of Marton's region in [13, Th. 2] to include $R_0 > 0$. The MGP region is given by

$$\mathcal{R}_{MGP} = \bigcup_{P_{TU_1U_2X}} \mathcal{R}_{MGP}(P_{TU_1U_2X}) \tag{1}$$

The material in this paper has been presented in part at the 46th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, Sept. 2008.

Yingbin Liang is with Department of Electrical Engineering, University of Hawaii, Honolulu, HI 96822, USA, yingbinl@hawaii.edu

Gerhard Kramer is with Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA, gkramer@usc.edu

H. Vincent Poor is with Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA, poor@princeton.edu

The work of Y. Liang was supported by a National Science Foundation CAREER Award under Grant CCF-08-46028. The work of H. V. Poor was supported in part by the National Science Foundation under Grants ANI-03-38807 and CNS-06-25637.

where

$$\mathcal{R}_{MGP}(P_{TU_1U_2X}) = \left\{ \begin{array}{l} (R_0,R_1,R_2): R_0 \geq 0, R_1 \geq 0, R_2 \geq 0, \\ R_0 \leq \min\{I(T;Y_1), I(T;Y_2)\} \\ R_0 + R_1 \leq I(T,U_1;Y_1), \\ R_0 + R_2 \leq I(T,U_2;Y_2), \\ R_0 + R_1 + R_2 \leq I(T,U_1;Y_1) + I(U_2;Y_2|T) \\ \qquad\qquad -I(U_1;U_2|T), \\ R_0 + R_1 + R_2 \leq I(U_1;Y_1|T) + I(T,U_2;Y_2) \\ \qquad\qquad -I(U_1;U_2|T) \end{array} \right\}. \tag{2}$$

Another inner bound was derived in [14, Sec. III-A] and is given by

$$\mathcal{R}_{LK} = \bigcup_{P_{TU_1U_2X}} \mathcal{R}_{LK}(P_{TU_1U_2X}) \tag{3}$$

where

$$\mathcal{R}_{LK}(P_{TU_1U_2X}) = \left\{ \begin{array}{l} (R_0,R_1,R_2): R_0 \geq 0, R_1 \geq 0, R_2 \geq 0, \\ R_0 + R_1 \leq I(T,U_1;Y_1), \\ R_0 + R_2 \leq I(T,U_2;Y_2), \\ R_0 + R_1 + R_2 \leq I(T,U_1;Y_1) + I(U_2;Y_2|T) \\ \qquad\qquad -I(U_1;U_2|T), \\ R_0 + R_1 + R_2 \leq I(U_1;Y_1|T) + I(T,U_2;Y_2) \\ \qquad\qquad -I(U_1;U_2|T), \\ 2R_0 + R_1 + R_2 \leq I(T,U_1;Y_1) + I(T,U_2;Y_2) \\ \qquad\qquad -I(U_1;U_2|T) \end{array} \right\}. \tag{4}$$

Comparing the regions $\mathcal{R}_{MGP}$ and $\mathcal{R}_{LK}$, it is clear that for a given distribution $P_{TU_1U_2X}$, the bounds defining $\mathcal{R}_{LK}(P_{TU_1U_2X})$ differ from the bounds defining $\mathcal{R}_{MGP}(P_{TU_1U_2X})$ in that the bound

$$2R_0 + R_1 + R_2 \leq I(T,U_1;Y_1) + I(T,U_2;Y_2) - I(U_1;U_2|T)$$

that appears in $\mathcal{R}_{LK}(P_{TU_1U_2X})$ replaces the bound

$$R_0 \leq \min\{I(T;Y_1), I(T;Y_2)\}$$

that appears in $\mathcal{R}_{MGP}(P_{TU_1U_2X})$. As commented in [14, Remark 6], the region $\mathcal{R}_{LK}$ includes $\mathcal{R}_{MGP}$. In particular, we note the following remark.

***Remark** 1:* The region $\mathcal{R}_{LK}(P_{TU_1U_2X})$ is strictly larger than $\mathcal{R}_{MGP}(P_{TU_1U_2X})$ for some distributions $P_{TU_1U_2X}$. For example, if T is a constant, then $R_0 = 0$ for all points in $\mathcal{R}_{MGP}(P_{TU_1U_2X})$ while $\mathcal{R}_{LK}(P_{TU_1U_2X})$ may include points with $R_0 > 0$.

Although $\mathcal{R}_{MGP} \subseteq \mathcal{R}_{LK}$, it is not easy to see whether $\mathcal{R}_{MGP}$ is a strict subset of $\mathcal{R}_{LK}$ or not, because the rate points that are in $\mathcal{R}_{LK}(P_{TU_1U_2X})$ but not in $\mathcal{R}_{MGP}(P_{TU_1U_2X})$ may be in $\mathcal{R}_{MGP}(P_{T'U_1'U_2'X'})$ for some $P_{T'U_1'U_2'X'} \neq P_{TU_1U_2X}$. The main result of this paper is stated in the following theorem, which establishes the equivalence of the two regions.

***Theorem** 1:* $\mathcal{R}_{MGP} = \mathcal{R}_{LK}$.

***Remark** 2:* Theorem 1 is also true if the channel has cost constraints, either an average cost constraint over a block of inputs and outputs, or a cost constraint over each channel use. This is because every step in the proof of Theorem 1 in the following section holds for such cost constraints. A detailed discussion about cost constraints can be found in [19, Chapter 3].

IV. Proof of Theorem 1

As we have argued in the previous section, $\mathcal{R}_{MGP} \subseteq \mathcal{R}_{LK}$. Hence we need to show that $\mathcal{R}_{LK} \subseteq \mathcal{R}_{MGP}$ to establish Theorem 1. We first state three lemmas that will be useful later on.

***Lemma** 1:* The region $\mathcal{R}_{MGP}$ is the capacity region for the broadcast channel with degraded message sets, i.e., the cases when R_1=0 or R_2=0.

Proof: Let $R_2 = 0$ and set $U_1 = X$ and $U_2 = T$ in $\mathcal{R}_{MGP}$. The region $\mathcal{R}_{MGP}$ reduces to the following region

$$\mathcal{C}_{d1} = \bigcup_{P_{TX}} \left\{ \begin{array}{l} (R_0,R_1): R_0 \geq 0, R_1 \geq 0 \\ R_0 \leq \min\{I(T;Y_1), I(T;Y_2)\} \\ R_0 + R_1 \leq I(X;Y_1), \\ R_0 + R_1 \leq I(X;Y_1|T) + I(T;Y_2) \end{array} \right\}. \tag{5}$$

To show that the region $\mathcal{C}_{d1}$ is the capacity region, we apply the outer bound given in [16, Lemma 2], for which we set $R_2 = 0$, and apply the bound on R_0, the first bound on $R_0 + R_1$, and the second bound on $R_0 + R_1 + R_2$ to obtain:

$$\begin{aligned} R_0 &\leq \min\left\{I(T;Y_1), I(T;Y_2)\right\} \\ &\leq \min\left\{I(T,U;Y_1), I(T,U;Y_2)\right\} \\ R_0 + R_1 &\leq I(X;Y_1) \\ R_0 + R_1 + R_2 &\leq I(X;Y_1|T,U) + I(T,U;Y_2). \end{aligned} \tag{6}$$

The above bounds coincide with the bounds in $\mathcal{C}_{d1}$ with (T,U) being replaced by T', which completes the proof. ■

We note that the capacity region for the broadcast channel with degraded message sets was established in [9] and [18, P. 360, Th. 4.1], and is given by

$$\mathcal{C}_{d} = \bigcup_{P_{TX}} \left\{ \begin{array}{l} (R_0,R_1): R_0 \geq 0, R_1 \geq 0 \\ R_0 \leq I(T;Y_2) \\ R_0 + R_1 \leq I(X;Y_1), \\ R_0 + R_1 \leq I(X;Y_1|T) + I(T;Y_2) \end{array} \right\}. \tag{7}$$

Thus, $\mathcal{C}_{d1}$ must be equivalent to $\mathcal{C}_d$. We further use $\mathcal{C}_{d2}$ to denote the capacity region of the broadcast channel when $R_1 = 0$.

We next state a lemma that will help to prove an important property of $\mathcal{R}_{LK}$ (see Lemma 3 below).

***Lemma** 2:* For a joint distribution $P_{TU_1Y_1Y_2}$, if $I(T;Y_1) < I(T;Y_2)$ and $I(T,U_1;Y_1) > I(T,U_1;Y_2)$, then there exists a function $f(U_1,Z)$ with Z being a random variable independent of T, U_1, X, Y_1 and Y_2, such that $I(T, f(U_1,Z);Y_1) = I(T, f(U_1,Z);Y_2)$.

Proof: See Appendix I. ■

***Remark** 3:* A statement similar to Lemma 2 has been made in [7], which claims the existence of a deterministic function $f(U_1)$ in contrast to a stochastic function $f(U_1,Z)$ in Lemma 2. However, such a deterministic function $f(U_1)$ does not always exist. A simple counter example is when U_1

is a binary random variable. Then a deterministic function $f(U_1)$ either has the same distribution as U_1 or has a constant value. Then $I(T, f(U_1); Y_1) = I(T, f(U_1); Y_2)$ cannot always be satisfied.

***Lemma** 3:* Let $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$ denote the sets of the following distributions:

$$\begin{aligned} \mathcal{P}_0 &= \{P_{TU_1U_2X} : I(T;Y_1) = I(T;Y_2)\} \\ \mathcal{P}_1 &= \{P_{TU_1U_2X} : I(T;Y_1) \le I(T;Y_2), U_1 = \phi\} \\ \mathcal{P}_2 &= \{P_{TU_1U_2X} : I(T;Y_1) \ge I(T;Y_2), U_2 = \phi\}. \end{aligned} \tag{8}$$

The region $\mathcal{R}_{LK}$ given in (3) can be obtained by taking the union over only $\mathcal{P}_0$, $\mathcal{P}_1$ and $\mathcal{P}_2$, i.e.,

$$\mathcal{R}_{LK} = \bigcup_{\mathcal{P}_0 \cup \mathcal{P}_1 \cup \mathcal{P}_2} \mathcal{R}_{LK}(P_{TU_1U_2X}). \tag{9}$$

Proof: See Appendix II, which follows the proof in [7, Appendix] where it is shown that the region $\mathcal{R}_{MGP}$ has the same type of property. ∎

We next consider $\mathcal{R}_{LK}(P_{TU_1U_2X})$, where $P_{TU_1U_2X}$ is in either $\mathcal{P}_0$, $\mathcal{P}_1$, or $\mathcal{P}_2$.

(1) If $P_{TU_1U_2X} \in \mathcal{P}_1$, then the reader can readily check that

$$\begin{aligned} &\mathcal{R}_{LK}(P_{TU_1U_2X}) = \mathcal{R}_{MGP}(P_{TU_1U_2X}) \\ &= \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_0 \ge 0, R_1 \ge 0, R_2 \ge 0, \\ R_0 + R_1 \le I(T;Y_1) \\ R_0 + R_1 + R_2 \le I(T;Y_1) + I(U_2;Y_2|T) \end{array} \right\}. \end{aligned} \tag{10}$$

(2) If $P_{TU_1U_2X} \in \mathcal{P}_2$, we also have $\mathcal{R}_{LK}(P_{TU_1U_2X}) = \mathcal{R}_{MGP}(P_{TU_1U_2X})$.

(3) If $P_{TU_1U_2X} \in \mathcal{P}_0$, i.e., $I(T;Y_1) = I(T;Y_2)$, we have

$$\begin{aligned} &\mathcal{R}_{LK}(P_{TU_1U_2X}) \\ &= \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_0 \ge 0, R_1 \ge 0, R_2 \ge 0 \\ R_0 + R_1 \le I(T, U_1; Y_1) \\ R_0 + R_2 \le I(T, U_2; Y_2) \\ R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T) \\ \qquad\qquad\qquad -I(U_1; U_2|T) \\ 2R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(T, U_2; Y_2) \\ \qquad\qquad\qquad -I(U_1; U_2|T) \end{array} \right\}. \end{aligned} \tag{11}$$

It is clear that the points that satisfy $R_0 \le I(T;Y_1)$ are in $\mathcal{R}_{MGP}$. We need to consider only the extreme points that satisfy $R_0 > I(T;Y_1)$. Under this condition, the last bound on $2R_0 + R_1 + R_2$ in (11) can be written as

$$\begin{aligned} &R_0 + R_1 + R_2 \\ &\quad \le I(T, U_1; Y_1) + I(T, U_2; Y_2) - I(U_1; U_2|T) - R_0. \end{aligned} \tag{12}$$

The right-hand side of the above equation is smaller than the bound on $R_0 + R_1 + R_2$ in (11), because $R_0 > I(T;Y_1)$. Hence the bound on $R_0 + R_1 + R_2$ in (11) is redundant when $R_0 > I(T;Y_1)$. The extreme points, which satisfy $R_0 > I(T;Y_1)$ and are on the plane determined by $R_1 = 0$ or $R_2 = 0$, are in $\mathcal{C}_{d1}$ or $\mathcal{C}_{d2}$, and are hence in $\mathcal{R}_{MGP}$. The

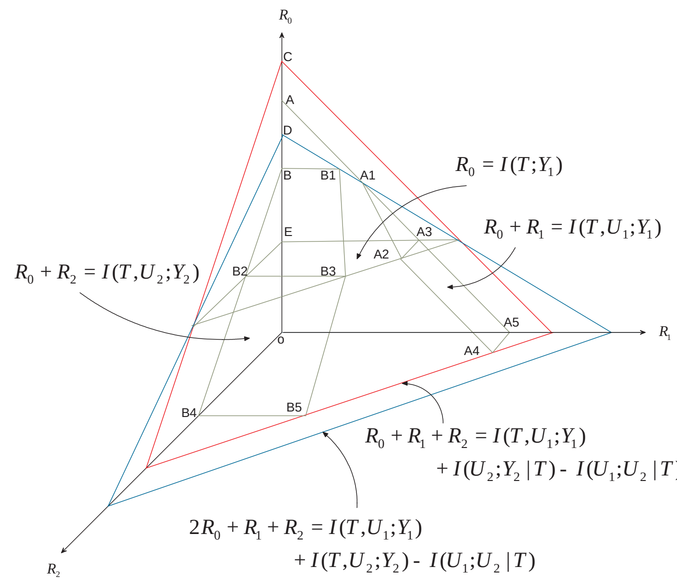

Fig. 2. An illustration of the regions $\mathcal{R}_{LK}(P_{TU_1U_2X})$ and $\mathcal{R}_{MGP}(P_{TU_1U_2X})$ when $I(T;Y_1) = I(T;Y_2)$

remaining extreme points are the intersections of the planes defined by the following bounds

$$\begin{aligned} R_0 + R_1 &= I(T, U_1; Y_1) \\ R_0 + R_2 &= I(T, U_2; Y_2) \\ 2R_0 + R_1 + R_2 &= I(T, U_1; Y_1) + I(T, U_2; Y_2) \\ &\quad - I(U_1; U_2|T). \end{aligned} \tag{13}$$

Now, if $I(U_1; U_2|T) = 0$, the sum of the first and second bounds is equal to the third one, and hence the three bounds become two and the intersection of the corresponding two planes is not an extreme point. If $I(U_1; U_2|T) > 0$, the above three bounds do not have common points because the sum of the first and second bounds is larger than the third one. Hence we have shown that all extreme points of $\mathcal{R}_{LK}(P_{TU_1U_2X})$ are in $\mathcal{R}_{MGP}$, which implies all points in $\mathcal{R}_{LK}(P_{TU_1U_2X})$ are in $\mathcal{R}_{MGP}$. This concludes the proof.

V. A Geometric Illustration

We now illustrate the region $\mathcal{R}_{LK}(P_{TU_1U_2X})$ in Fig. 2 for the case when $I(T;Y_1) = I(T;Y_2)$ (see (11)). The four bounds on $R_0 + R_1$, $R_0 + R_2$, $R_0 + R_1 + R_2$, and $2R_0 + R_1 + R_2$ in (11) determine four planes in the three-dimensional space. We use A, B, C, D to denote these respective planes. We also use A, B, C, D to denote points where these planes intersect the R_0-axis, and use R_{0i} to denote the R_0 values for $i = A, B, C, D$, respectively. Suppose that $R_{0B} < R_{0D} < R_{0A} < R_{0C}$, which is the case when the four planes have the greatest number of intersections with each other. In addition to these four planes, we also plot plane E in the figure, which is determined by $R_0 = I(T;Y_1)$, and intersects the R_0-axis at point E. We assume that $R_{0E} = I(T;Y_1) < R_{0B}$.

We first observe that the region $\mathcal{R}_{MGP}(P_{TU_1U_2X})$ is contained inside the planes EB_2B_4O (plane $R_1 = 0$),

$B_2B_4B_5B_3$ (plane B), $B_3B_5A_4A_2$ (plane C), $A_2A_4A_5A_3$ (plane A), EOA_5A_3 (plane $R_2 = 0$), $OB_4B_5A_4A_5$ (plane $R_0 = 0$), and $EB_2B_3A_2A_3$ (plane E). For the region $\mathcal{R}_{LK}(P_{TU_1U_2X})$, plane E is not a constraint, and plane D is an additional constraint. We also note that plane D intersects plane C at line B_3A_2, and hence plane C does not play a role (i.e., is not a part of the boundary for the region $\mathcal{R}_{LK}(P_{TU_1U_2X})$) above plane E. This demonstrates that the bound on the sum rate $R_0+R_1+R_2$ becomes redundant when $R_0 > I(T;Y)$. Above plane E, the region $\mathcal{R}_{LK}(P_{TU_1U_2X})$ has more rate points than the region $\mathcal{R}_{MGP}(P_{TU_1U_2X})$, and these rate points are contained in planes BB_2E (plane $R_1 = 0$), $BB_2B_3B_1$ (plane B), $B_1B_3A_2A_1$ (plane D), $A_1A_2A_3$ (plane A), $BEA_3A_1B_1$ (plane $R_2 = 0$), and $EB_2B_3A_2A_3$ (plane E). From the figure, it can be seen that all extreme points in this region are either on plane E or on plane $R_1 = 0$ or $R_2 = 0$. It is clear that the extreme points on plane E are contained in $\mathcal{R}_{MGP}(P_{TU_1U_2X})$. The extreme points on plane $R_1 = 0$ or $R_2 = 0$ are contained in $\mathcal{C}_{d1}$ or $\mathcal{C}_{d2}$, and hence must be in $\mathcal{R}_{MGP}$ (these points are not necessarily achieved by the distribution $P_{TU_1U_2X}$).

VI. Conclusions

We have showed that two seemingly different inner bounds on the capacity region of the two-receiver discrete memoryless broadcast channel are equivalent. Our proof is based on an important property, motivated by one shown in [7] for the MGP region, that $\mathcal{R}_{LK}$ can also be characterized by only a subset of joint input distributions. This property greatly facilitates the proof, which may be challenging otherwise. We also anticipate that this property is useful for studying rate regions for other multiuser channels.

Appendix I
Proof of Lemma 2

We define a binary random variable Z that is independent of T, U_1, X, Y_1 and Y_2, and satisfies

$$\Pr(Z = 0) = \alpha, \quad \text{and} \quad \Pr(Z = 1) = 1 - \alpha.$$

We define a function $c(U_1, Z)$ that satisfies

$$c(U_1, 0) = U_1, \quad \text{and} \quad c(U_1, 1) = 1.$$

We further define

$$f(U_1, Z) = (c(U_1, Z), Z).$$

It suffices to show that there exists a value $0 \le \alpha \le 1$ such that

$$I(T, f(U_1, Z); Y_1) = I(T, f(U_1, Z); Y_2). \tag{14}$$

We first compute the left-hand side of (14) to be

$$\begin{aligned} I(T, f(U_1, Z); Y_1) &= I(T, c(U_1, Z), Z; Y_1) \\ &= I(T, c(U_1, Z); Y_1|Z) \\ &= \alpha I(T, U_1; Y_1) + (1-\alpha) I(T; Y_1) \\ &:= g_1(\alpha) \end{aligned} \tag{15}$$

where the last step defines a function $g_1(\alpha)$. We next compute the right-hand side of (14) to be

$$\begin{aligned} I(T, f(U_1, Z); Y_2) &= I(T, c(U_1, Z), Z; Y_2) \\ &= I(T, c(U_1, Z); Y_2|Z) \\ &= \alpha I(T, U_1; Y_2) + (1-\alpha) I(T; Y_2) \\ &:= g_2(\alpha) \end{aligned} \tag{16}$$

where the last step defines a function $g_2(\alpha)$.

It is easy to see that the function $g_1(\alpha) - g_2(\alpha)$ is continuous, and $g_1(\alpha) - g_2(\alpha) > 0$ when $\alpha = 1$ and $g_1(\alpha) - g_2(\alpha) < 0$ when $\alpha = 0$. Hence there must exist a value $0 \le \alpha \le 1$ such that $g_1(\alpha) = g_2(\alpha)$, i.e., such that

$$I(T, f(U_1, Z); Y_1) = I(T, f(U_1, Z); Y_2).$$

Appendix II
Proof of Lemma 3

For a given distribution $P_{TU_1U_2X}$ if $I(T;Y_1) \neq I(T;Y_2)$, then assume $I(T;Y_1) < I(T;Y_2)$ without loss of generality. We wish to show that there exists a distribution $P_{T'U_1'U_2'X'}$ that is in $\mathcal{P}_0 \cup \mathcal{P}_1 \cup \mathcal{P}_2$ such that $\mathcal{R}_{LK}(P_{TU_1U_2X}) \subseteq \mathcal{R}_{LK}(P_{T'U_1'U_2'X'})$. We consider the following two cases.

Case 1: $I(T, U_1; Y_1) \le I(T, U_1; Y_2)$. Let $T' = (T, U_1)$, $U_1' = \phi$, $U_2' = U_2$, and $X' = X$. It is clear that $P_{T'U_1'U_2'X'} \in \mathcal{P}_1$, and we obtain

$$\begin{aligned} &\mathcal{R}_{LK}(P_{T'U_1'U_2'X'}) \\ &= \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_0 \ge 0, R_1 \ge 0, R_2 \ge 0, \\ R_0 + R_1 \le I(T, U_1; Y_1) \\ R_0 + R_2 \le I(T, U_1, U_2; Y_2) \\ R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T, U_1) \\ R_0 + R_1 + R_2 \le I(T, U_1, U_2; Y_2) \\ 2R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(T, U_1, U_2; Y_2) \end{array} \right\} \\ &= \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_0 \ge 0, R_1 \ge 0, R_2 \ge 0, \\ R_0 + R_1 \le I(T, U_1; Y_1) \\ R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T, U_1) \end{array} \right\}. \end{aligned} \tag{17}$$

In order to show $\mathcal{R}_{LK}(P_{TU_1U_2X}) \subseteq \mathcal{R}_{LK}(P_{T'U_1'U_2'X'})$, we consider a given point $(R_0, R_1, R_2) \in \mathcal{R}_{LK}(P_{TU_1U_2X})$. It is clear that (R_0, R_1, R_2) satisfies the bound on $R_0 + R_1$ in $\mathcal{R}_{LK}(P_{T'U_1'U_2'X'})$. We further compare the $R_0 + R_1 + R_2$ bound in (17) with the first $R_0 + R_1 + R_2$ bound in (4) and find that

$$\begin{aligned} &I(T, U_1; Y_1) + I(U_2; Y_2|T, U_1) \\ &- I(T, U_1; Y_1) - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad = I(U_2; Y_2|T, U_1) - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad = I(U_2; Y_2, U_1|T) - I(U_2; Y_2|T) \\ &\quad \ge 0. \end{aligned} \tag{18}$$

Thus, the rate triple (R_0, R_1, R_2) satisfies the bound on $R_0 + R_1 + R_2$ in $\mathcal{R}_{LK}(P_{T'U_1'U_2'X'})$.

Case 2: $I(T, U_1; Y_1) > I(T, U_1; Y_2)$. The conditions in Lemma 2 are satisfied, and hence there exists a function $f(U_1, Z)$ with Z being a random variable independent of everything else, such that

$$I(T, f(U_1, Z); Y_1) = I(T, f(U_1, Z); Y_2).$$

Let $T' = (T, f(U_1, Z))$, $U_1' = U_1$, $U_2' = U_2$, and $X' = X$. It is clear that $P_{T'U_1'U_2'X'} \in \mathcal{P}_0$, and we obtain (see (11))

$$\mathcal{R}_{LK}(P_{T'U_1'U_2'X'}) = \left\{ \begin{array}{l} (R_0, R_1, R_2) : R_0 \geq 0, R_1 \geq 0, R_2 \geq 0, \\ R_0 + R_1 \leq I(T, U_1; Y_1) \\ R_0 + R_2 \leq I(T, f(U_1, Z), U_2; Y_2) \\ R_0 + R_1 + R_2 \leq I(T, U_1; Y_1) \\ +I(U_2; Y_2|T, f(U_1, Z)) - I(U_1; U_2|T, f(U_1, Z)) \\ 2R_0 + R_1 + R_2 \leq I(T, U_1; Y_1) \\ +I(T, f(U_1, Z), U_2; Y_2) - I(U_1; U_2|T, f(U_1, Z)) \end{array} \right\} \tag{19}$$

where we have used the following equation

$$\begin{aligned} & I(T, f(U_1, Z), U_1; Y_1) \\ & = I(T, U_1; Y_1) + I(f(U_1, Z); Y_1|T, U_1) \\ & = I(T, U_1; Y_1). \end{aligned} \tag{20}$$

It is easy to see that a given point $(R_0, R_1, R_2) \in \mathcal{R}_{LK}(P_{TU_1U_2X})$ in (4) satisfies the first two bounds in the above region. Furthermore, we show that the third bound on $R_0 + R_1 + R_2$ in (19) is looser than the third bound on $R_0 + R_1 + R_2$ in (4) by considering

$$\begin{aligned} & I(T, U_1; Y_1) + I(U_2; Y_2|T, f(U_1, Z)) \\ & - I(U_1; U_2|T, f(U_1, Z)) \\ & - I(T, U_1; Y_1) - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ & \quad = I(U_2; Y_2|T, f(U_1, Z)) - I(U_1; U_2|T, f(U_1, Z)) \\ & \quad \quad - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ & \quad = I(U_2, f(U_1, Z); Y_2|T) - I(f(U_1, Z); Y_2|T) \\ & \quad \quad - I(U_1, f(U_1, Z); U_2|T) + I(f(U_1, Z); U_2|T) \\ & \quad \quad - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ & \quad \overset{(a)}{=} I(f(U_1, Z); Y_2|T, U_2) - I(f(U_1, Z); Y_2|T) \\ & \quad \quad + I(f(U_1, Z); U_2|T) \\ & \quad = I(f(U_1, Z); Y_2, U_2|T) - I(f(U_1, Z); Y_2|T) \\ & \quad \geq 0 \end{aligned} \tag{21}$$

where (a) follows because

$$\begin{aligned} & I(U_1, f(U_1, Z); U_2|T) \\ & = I(U_1; U_2|T) + I(f(U_1, Z); U_2|T, U_1) \\ & = I(U_1; U_2|T). \end{aligned} \tag{22}$$

Thus, the triples (R_0, R_1, R_2) that satisfy the third bound in (4) satisfy the third bound in (19). Based on the inequality (21), we can further show that the bound on $2R_0 + R_1 + R_2$ in (19) is larger than the bound on $2R_0 + R_1 + R_2$ in (4) by considering

$$\begin{aligned} & I(T, U_1; Y_1) + I(T, f(U_1, Z), U_2; Y_2) \\ & - I(U_1; U_2|T, f(U_1, Z)) \\ & - I(T, U_1; Y_1) - I(T, U_2; Y_2) + I(U_1; U_2|T) \\ & \quad \geq I(T, f(U_1, Z); Y_2) - I(T; Y_2) \\ & \quad \geq 0. \end{aligned} \tag{23}$$

Thus, the (R_0, R_1, R_2) that satisfy the fifth bound in (4) satisfy the fourth bound in (19). This concludes the proof.

REFERENCES

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

[2] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. 19, no. 2, pp. 197–207, Mar. 1973.

[3] ——, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 20, pp. 279–280, Mar. 1974.

[4] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Inform. Transm.*, vol. 10, pp. 185–193, Jul.-Sep. 1974.

[5] K. Marton, "The capacity region of deterministic broadcast channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Paris-Cachan, France, 1977.

[6] M. S. Pinsker, "The capacity region of noiseless broadcast channels," *Probl. Inform. Transm.*, vol. 14, pp. 97–102, 1978.

[7] S. I. Gelfand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 17–25, Jan.-Mar. 1980.

[8] A. A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inform. Theory*, vol. 25, no. 2, pp. 166–169, Mar. 1979.

[9] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. 23, pp. 60–64, Jan. 1977.

[10] A. A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 1–16, Jan.-Mar. 1980.

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

[12] B. E. Hajek and M. B. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 1, pp. 36–46, Jan. 1979.

[13] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[14] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Transactions on Information Theory, Special Issue on Models, Theory and Codes for Relaying and Cooperation in Communication Networks*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.

[15] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.

[16] Y. Liang, G. Kramer, and S. Shamai (Shitz), "Capacity outer bounds for broadcast channels," in *Proc. of IEEE Information Theory Workshop (ITW)*, Porto, Portugal, May 2008, pp. 2–4.

[17] C. Nair, "An outer bound for 2-receiver discrete memoryless broadcast channels," preprint, available at http://chandra.ie.cuhk.edu.hk/pub/papers/outerbound.pdf.

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

[19] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 4-5, pp. 1567–2190, 2008.