On the Equivalence of Two Achievable Regions for the Broadcast Channel

Yingbin Liang, *Member, IEEE*, Gerhard Kramer, *Fellow, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***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.**

***Index Terms*—Broadcast channel, inner bound, rate region.**

I. Introduction

The broadcast channel, in which one transmitter sends common and individual information to multiple receivers, was introduced in [1]. The performance measure of interest for the broadcast channel is the capacity region, which characterizes the simultaneously and reliably achievable communication rates. Although the capacity region has been 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], [2], [7], [12]–[15], and outer bounds have been obtained in, e.g., [3]–[11], [13]–[18].

In this paper, we focus on capacity inner bounds, i.e., achievable regions, for the case of two receivers. Marton's region [13, Theorem 2] is the largest known inner bound without a common message. Marton's region has been extended to include a common message, a result that appears in [19, p. 391, Prob. 10(c)] and [7, Theorem 1]. We call this region the Marton-Gelfand-Pinsker (MGP) region. Another inner bound that was derived recently in [15] includes, but may not be strictly larger than, the MGP region. In this paper, we first review these two inner bounds and then show that the two bounds are equivalent. The technique we exploit is based on a property developed in [7, Proposition 1].

Manuscript received June 23, 2009; revised April 07, 2010; accepted July 04, 2010. Date of current version December 27, 2010. The material in this paper was presented (in part) at the 46th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 2008. The work of Y. Liang was supported by the National Science Foundation (NSF) under Grant CCF-09-15772. The work of G. Kramer was supported in part by the Board of Trustees of the University of Illinois Subaward 04-217 under NSF Grant CCR-0325673. The work of H. V. Poor was supported in part by the NSF under Grant CNS-09-05398.

Y. Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: yliang06@syr.edu).

G. Kramer is with the Department of Electrical Engineering and Information Technology, Technische Universität München, 80333 München, Germany (e-mail: gerhard.kramer@tum.de).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Communicated by E. Erkip, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2090236

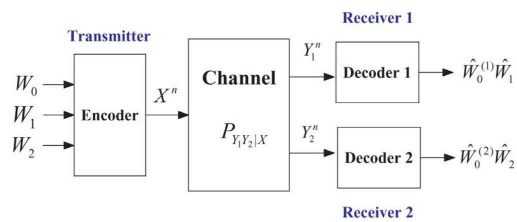

Fig. 1. Two-receiver broadcast channel.

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., $f\colon \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{X}^n$, 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 $P_{Y_1Y_2|X}(\cdot|\cdot)$ 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., $g_1 : \mathcal{Y}_1^n \to \mathcal{W}_0 \times \mathcal{W}_1$, maps the received sequence $y_1^n \in \mathcal{Y}_1^n$ to a message pair $\left(\hat{w}_0^{(1)}, \hat{w}_1\right) \in \mathcal{W}_0 \times \mathcal{W}_1$, and a decoder at receiver 2, i.e., $g_2\colon \mathcal{Y}_2^n \to \mathcal{W}_0 \times \mathcal{W}_2$, maps the received sequence $y_2^n \in \mathcal{Y}_2^n$ to a message pair $\left(\hat{w}_0^{(2)}, \hat{w}_2\right) \in \mathcal{W}_0 \times \mathcal{W}_2$.

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

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

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 $P_e^{(n)} \to 0$ as n goes to infinity. The capacity region is the closure of the set of achievable rate triples.

III. Preliminary and Main Results

Consider the MGP region in [19, P. 391, Prob. 10(c)] and [7, Theorem 1] which is an extension of Marton's region in [13, Theorem 2] to include $R_0 > 0$. The MGP region is given by

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

where $\mathcal{R}_{\rm MGP}(P_{TU_1U_2X})$ contains all nonnegative rate triples (R_0, R_1, R_2) that satisfy

$$R_0 \le \min\{I(T;T_1), I(T;Y_2)\} \tag{2}$$
$$R_0 + R_1 \le I(T, U_1; Y_1) \tag{3}$$
$$R_0 + R_2 \le I(T, U_2; Y_2) \tag{4}$$
$$R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T) - I(U_1; U_2|T) \tag{5}$$
$$R_0 + R_1 + R_2 \le I(U_1; Y_1|T) + I(T, U_2; Y_2) - I(U_1; U_2|T) \tag{6}$$

for the joint distribution $P_{TU_1U_2X}$.

Another inner bound was derived in [14] and [15, Section III-A] and is given by

$$\mathcal{R}_{LK} = \bigcup_{P_{TU_1U_2X}} \mathcal{R}_{LK}(P_{TU_1U_2X}) \tag{7}$$

where $\mathcal{R}_{LK}(P_{TU_1U_2X})$ contains all nonnegative rate triples (R_0, R_1, R_2) that satisfy

$$R_0 + R_1 \le I(T, U_1; Y_1), \tag{8}$$
$$R_0 + R_2 \le I(T, U_2; Y_2), \tag{9}$$
$$R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T) - I(U_1; U_2|T), \tag{10}$$
$$R_0 + R_1 + R_2 \le I(U_1; Y_1|T) + I(T, U_2; Y_2) - I(U_1; U_2|T), \tag{11}$$
$$2R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(T, U_2; Y_2) - I(U_1; U_2|T) \tag{12}$$

for the joint distribution $P_{TU_1U_2X}$.

Comparing the regions $\mathcal{R}_{\rm MGP}$ and $\mathcal{R}_{LK}$, it is clear that $\mathcal{R}_{LK}(P_{TU_1U_2X})$ differs from $\mathcal{R}_{\rm MGP}(P_{TU_1U_2X})$ in that the bound (12) replaces the bound (2). As commented in [15, Remark 6] the region $\mathcal{R}_{LK}$ includes $\mathcal{R}_{\rm MGP}$. In particular, we note the following.

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

Although $\mathcal{R}_{\rm MGP} \subseteq \mathcal{R}_{LK}$, it is not easy to see whether $\mathcal{R}_{\rm MGP}$ 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 $\mathcal{R}_{\rm MGP}(P_{TU_1U_2X})$ may be in $\mathcal{R}_{\rm MGP}(P_{T'U_1'U_2'X'})$ 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: $\mathcal{R}_{\rm MGP} = \mathcal{R}_{LK}$.

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 Section IV holds for such cost constraints. A detailed discussion about cost constraints can be found in [20, Chapter 3].

IV. Proof of Theorem 1

As we have argued in the previous section, $\mathcal{R}_{\rm MGP} \subseteq \mathcal{R}_{LK}$. Hence we need to show that $\mathcal{R}_{LK} \subseteq \mathcal{R}_{\rm MGP}$ to establish Theorem 1. We first state three lemmas that will be useful in the sequel. Lemmas 2 and 3 are new and of independent interest.

Lemma 1: The region $\mathcal{R}_{\rm MGP}$ 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 $\mathcal{R}_{\rm MGP}$. The region $\mathcal{R}_{\rm MGP}$ reduces to the region $\mathcal{C}_{d1}$ that contains all nonnegative rate pairs (R_0, R_1) satisfying

$$R_0 \le \min\{I(T;Y_1), I(T;Y_2)\} \tag{13}$$
$$R_0 + R_1 \le I(X;Y_1) \tag{14}$$
$$R_0 + R_1 \le I(X;Y_1|T) + I(T;Y_2) \tag{15}$$

for some joint distribution P_{TX}. To show that the region $\mathcal{C}_{d1}$ is the capacity region, we apply the outer bound given in [17, 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

$$R_0 \le \min\{I(T;Y_1), I(T;Y_2)\} \le \min\{I(T,U;Y_1), I(T,U;Y_2)\} \tag{16}$$
$$R_0 + R_1 \le I(X;Y_1) \tag{17}$$
$$R_0 + R_1 + R_2 \le I(X;Y_1|T,U) + I(T,U;Y_2). \tag{18}$$

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

Remark 3: The capacity region for the broadcast channel with degraded message sets was established in [9] and [19, p. 360, Theorem 4.1] and is given by the region $\mathcal{C}_d$ that contains all nonnegative rate pairs (R_0, R_1) satisfying

$$R_0 \le I(T;Y_2) \tag{19}$$
$$R_0 + R_1 \le I(X;Y_1) \tag{20}$$
$$R_0 + R_1 \le I(X;Y_1|T) + I(T;Y_2) \tag{21}$$

for some joint distribution P_{TX}. 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 A. ■

Remark 4: 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 arises 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. So $I(T, f(U_1); Y_1) = I(T, f(U_1); Y_2)$ cannot always be satisfied.

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

$$\mathcal{P}_0 = \{P_{TU_1U_2X} : I(T;Y_1) = I(T;Y_2)\} \quad (22)$$
$$\mathcal{P}_1 = \{P_{TU_1U_2X} : I(T;Y_1) \le I(T;Y_2), U_1 = \phi\} \quad (23)$$
$$\mathcal{P}_2 = \{P_{TU_1U_2X} : I(T;Y_1) \ge I(T;Y_2), U_2 = \phi\}. \quad (24)$$

The region $\mathcal{R}_{LK}$ given in (7) can be obtained by taking the union over only $\mathcal{P}_0$, $\mathcal{P}_1$ and $\mathcal{P}_2$, i.e., we have

$$\mathcal{R}_{LK} = \bigcup_{\mathcal{P}_0 \cup \mathcal{P}_1 \cup \mathcal{P}_2} \mathcal{R}_{LK}(P_{TU_1U_2X}). \quad (25)$$

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

Remark 5: The regions defined by the distributions in $\mathcal{P}_1$ and $\mathcal{P}_2$ in Lemma 3 can be achieved by superposition coding. The regions defined by distributions in $\mathcal{P}_0$ require both superposition coding and binning in general. Lemma 3 thus provides guidance on the required code structures.

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 $\mathcal{R}_{LK}(P_{TU_1U_2X}) = \mathcal{R}_{\mathrm{MGP}}(P_{TU_1U_2X})$, both containing all nonnegative rate triples (R_0, R_1, R_2) that satisfy

$$R_0 + R_1 \le I(T;Y_1), \quad (26)$$
$$R_0 + R_1 + R_2 \le I(T;Y_1) + I(U_2;Y_2|T). \quad (27)$$

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

(3) If $P_{TU_1U_2X} \in \mathcal{P}_0$, i.e., $I(T;Y_1) = I(T;Y_2)$, we obtain the region $\mathcal{R}_{LK}(P_{TU_1U_2X})$ that contains all nonnegative rate triples (R_0, R_1, R_2) satisfying

$$R_0 + R_1 \le I(T, U_1; Y_1) \quad (28)$$
$$R_0 + R_2 \le I(T, U_2; Y_2) \quad (29)$$
$$R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T) - I(U_1; U_2|T) \quad (30)$$
$$2R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(T, U_2; Y_2) - I(U_1; U_2|T). \quad (31)$$

It is clear that the points that satisfy $R_0 \le I(T;Y_1)$ are in $\mathcal{R}_{\mathrm{MGP}}$. We need to consider only the extreme points that satisfy $R_0 > I(T;Y_1)$. Under this condition, (31) can be written as

$$R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(T, U_2; Y_2) - I(U_1; U_2|T) - R_0 \quad (32)$$

where the right-hand side is smaller than that of (30), because $R_0 > I(T;Y_1)$. Hence, the bound (30) is redundant. 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 $\mathcal{R}_{\mathrm{MGP}}$. The remaining extreme points are the intersections of the planes defined by the following bounds:

$$R_0 + R_1 = I(T, U_1; Y_1) \quad (33)$$
$$R_0 + R_2 = I(T, U_2; Y_2) \quad (34)$$
$$2R_0 + R_1 + R_2 = I(T, U_1; Y_1) + I(T, U_2; Y_2) - I(U_1; U_2|T). \quad (35)$$

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 $\mathcal{R}_{\mathrm{MGP}}$, which implies all points in $\mathcal{R}_{LK}(P_{TU_1U_2X})$ are in $\mathcal{R}_{\mathrm{MGP}}$. 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 $I(T;Y_1) = I(T;Y_2)$. 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 (28)–(31) determine four planes in three-dimensional space. We use A, B, C, and D to denote these respective planes. We also use A, B, C, and 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 $R_{\mathrm{MGP}}(P_{TU_1U_2X})$ is contained within 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 $\mathcal{R}_{\mathrm{MGP}}(P_{TU_1U_2X})$, 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 $\mathcal{R}_{\mathrm{MGP}}(P_{TU_1U_2X})$. The extreme points on plane $R_1 = 0$ or $R_2 = 0$ are contained in $\mathcal{C}_{d1}$ or $\mathcal{C}_{d2}$ as defined in the proof for Lemma 1, and hence must be

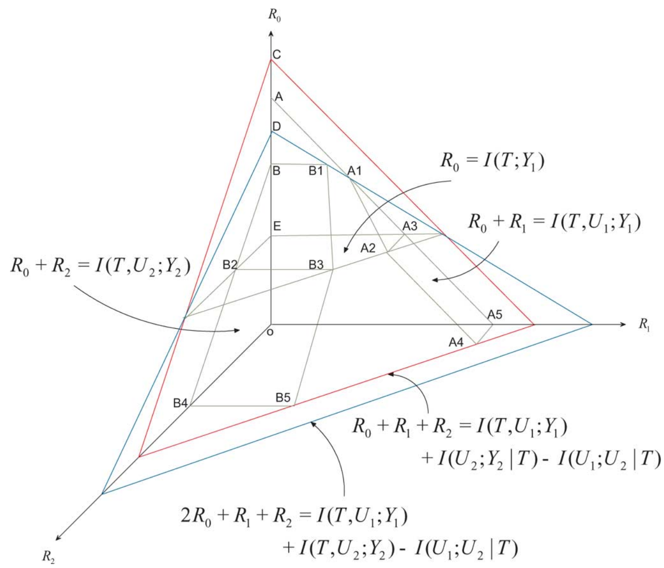

Fig. 2. Illustration of the regions $\mathcal{R}_{LK}(P_{TU_1U_2X})$ and $\mathcal{R}_{\mathrm{MGP}}(P_{TU_1U_2X})$ when $I(T;Y_1) = I(T;Y_2)$.

in $\mathcal{R}_{\mathrm{MGP}}$ (these points are not necessarily achieved by the distribution $P_{TU_1U_2X}$).

VI. Conclusions

We have shown 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 A
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

$$I(T, f(U_1, Z); Y_1) = I(T, f(U_1, Z); Y_2). \tag{36}$$

We first compute the left-hand side of (36) to be

$$\begin{aligned} I(T, f(U_1, Z); Y_1) &= I(T, c(U_1, Z), Z; Y_1) \\ &= I(T, c(U_1, Z); Y_1|Z) \\ &= \alpha I(T, U_1; Y_1) + (1-\alpha) I(T; Y_1) \\ &:= g_1(\alpha) \end{aligned} \tag{37}$$

where the last step defines a function $g_1(\alpha)$. We next compute the right-hand side of (36) to be

$$\begin{aligned} I(T, f(U_1, Z); Y_2) &= I(T, c(U_1, Z), Z; Y_2) \\ &= I(T, c(U_1, Z); Y_2|Z) \\ &= \alpha I(T, U_1; Y_2) + (1-\alpha) I(T; Y_2) \\ &:= g_2(\alpha) \end{aligned} \tag{38}$$

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 B
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 the region $\mathcal{R}_{LK}(P_{T'U_1'U_2'X'})$ that contains all nonnegative rate triples (R_0, R_1, R_2) satisfying

$$R_0 + R_1 \le I(T, U_1; Y_1), \tag{39}$$

$$R_0 + R_1 + R_2 \le I(T, U_1; Y_1) + I(U_2; Y_2|T, U_1). \tag{40}$$

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 (39). We further compare (40) with (10) and find that

$$\begin{aligned} &I(T, U_1; Y_1) + I(U_2; Y_2|T, U_1) \\ &\quad - I(T, U_1; Y_1) - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad = I(U_2; Y_2|T, U_1) - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad = I(U_2; Y_2, U_1|T) - I(U_2; Y_2|T) \ge 0. \end{aligned} \tag{41}$$

Thus, the rate triple (R_0, R_1, R_2) also satisfies (40) and is hence 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 all other random variables under consideration, 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 the region $\mathcal{R}_{LK}(P_{T'U_1'U_2'X'})$ that contains all nonnegative rate triples (R_0, R_1, R_2) satisfying

$$R_0 + R_1 \leq I(T, U_1; Y_1) \tag{42}$$

$$R_0 + R_2 \leq I(T, f(U_1, Z), U_2; Y_2) \tag{43}$$

$$R_0 + R_1 + R_2 \leq I(T, U_1; Y_1) + I(U_2; Y_2|T, f(U_1, Z)) - I(U_1; U_2|T, f(U_1, Z)) \tag{44}$$

$$2R_0 + R_1 + R_2 \leq I(T, U_1; Y_1) + I(T, f(U_1, Z), U_2; Y_2) - I(U_1; U_2|T, f(U_1, Z)) \tag{45}$$

where we have used the following equation:

$$\begin{aligned} I(T, f(U_1, Z), U_1; Y_1) &= I(T, U_1; Y_1) + I(f(U_1, Z); Y_1|T, U_1) \\ &= I(T, U_1; Y_1). \end{aligned} \tag{46}$$

It is easy to see that a given point $(R_0, R_1, R_2) \in \mathcal{R}_{LK}(P_{TU_1U_2X})$ characterized by (8)–(12) satisfies (42) and (43). Furthermore, we show that the bound (44) is looser than the bound (10) by considering

$$\begin{aligned} &I(T, U_1; Y_1) + I(U_2; Y_2|T, f(U_1, Z)) \\ &- I(U_1; U_2|T, f(U_1, Z)) \\ &- I(T, U_1; Y_1) - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad = I(U_2; Y_2|T, f(U_1, Z)) - I(U_1; U_2|T, f(U_1, Z)) \\ &\quad\quad - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad = I(U_2, f(U_1, Z); Y_2|T) - I(f(U_1, Z); Y_2|T) \\ &\quad\quad - I(U_1, f(U_1, Z); U_2|T) + I(f(U_1, Z); U_2|T) \\ &\quad\quad - I(U_2; Y_2|T) + I(U_1; U_2|T) \\ &\quad \overset{(a)}{=} I(f(U_1, Z); Y_2|T, U_2) - I(f(U_1, Z); Y_2|T) \\ &\quad\quad + I(f(U_1, Z); U_2|T) \\ &\quad = I(f(U_1, Z); Y_2, U_2|T) - I(f(U_1, Z); Y_2|T) \\ &\quad \geq 0 \end{aligned} \tag{47}$$

where (a) follows because

$$\begin{aligned} I(U_1, f(U_1, Z); U_2|T) &= I(U_1; U_2|T) + I(f(U_1, Z); U_2|T, U_1) \\ &= I(U_1; U_2|T). \end{aligned} \tag{48}$$

Thus, the triples (R_0, R_1, R_2) satisfying (10) satisfy (44). Based on the inequality (47), we can further show that the bound (45) is larger than the bound (12) by considering

$$\begin{aligned} &I(T, U_1; Y_1) + I(T, f(U_1, Z), U_2; Y_2) \\ &- I(U_1; U_2|T, f(U_1, Z)) - I(T, U_1; Y_1) \\ &- I(T, U_2; Y_2) + I(U_1; U_2|T) \\ &\geq I(T, f(U_1, Z); Y_2) - I(T; Y_2) \geq 0. \end{aligned} \tag{49}$$

This concludes the proof.

References

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

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

[3] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 279–280, Mar. 1974.

[4] R. R. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Inf. 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. Inf. 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. Inf. 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. Inf. Theory*, vol. IT-25, no. 2, pp. 166–169, Mar. 1979.

[9] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT–23, no. 1, pp. 60–64, Jan. 1977.

[10] A. A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Inf. 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. Inf. 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. Inf. Theory*, vol. IT-25, no. 1, pp. 36–46, Jan. 1979.

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

[14] Y. Liang, "Multiuser Communications With Relaying and User Cooperation," Ph.D. dissertation, Dept. Elect. Comput. Eng., University of Illinois, Urbana-Champaign, 2005.

[15] 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.

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

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

[18] C. Nair, An Outer Bound for 2-Receiver Discrete Memoryless Broadcast Channels [Online]. Available: http://chandra.ie.cuhk.edu.hk/pub/papers/outerbound.pdf

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

[20] G. Kramer, "Topics in multi-user information theory," *Found. Trends Commun. Inf. Theory*, vol. 4, no. 4–5, pp. 1567–2190, 2008.

Yingbin Liang (S'01–M'05) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 2005.

In 2005–2007, she was a Postdoctoral Research Associate at Princeton University, Princeton, NJ. In 2008–2009, she was an Assistant Professor with the Department of Electrical Engineering, University of Hawaii. Since December 2009, she has been an Assistant Professor at the Department of Electrical Engineering and Computer Science, Syracuse University. Her research interests include communications, wireless networks, and information theory.

Dr. Liang was a Vodafone Fellow at the University of Illinois at Urbana-Champaign during 2003–2005, and received the Vodafone-U.S. Foundation Fellows Initiative Research Merit Award in 2005. She also received the M.E. Van Valkenburg Graduate Research Award from the ECE Department, University of Illinois at Urbana-Champaign, in 2005. In 2009, she received the National Science Foundation CAREER Award and the State of Hawaii Governor Innovation Award.

Gerhard Kramer (S'91–M'94–SM'08–F'10) received the Dr. sc. techn. (Doktor der technischen Wissenschaften) degree from the Swiss Federal Institute of Technology (ETH), Zurich, in 1998.

From 1998 to 2000, he was with Endora Tech AG, Basel, Switzerland, as a Communications Engineering Consultant. From 2000 to 2008, he was with Bell Labs, Alcatel-Lucent, Murray Hill, NJ, as a Member of Technical Staff. He joined the University of Southern California (USC), Los Angeles, in 2009. Since 2010, he has been a Professor and Head of the Institute for Communications Engineering at the Technische Universität München (TUM), München, Germany.

Dr. Kramer was elected to the Board of Governors of the IEEE Information Theory Society for 2009–2011. He has served as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY, as Co-Chair of the Technical Program Committee of the 2008 IEEE International Symposium on Information Theory, and as Co-Chair of the 1st, 2nd, and 3rd Annual Schools of Information Theory in 2008–2010. He is serving as a member of the Emerging Technologies Committee of the IEEE Communications Society. He is a corecipient of the IEEE Communications Society 2005 Stephen O. Rice Prize paper award, a Bell Labs President's Gold Award in 2003, and a recipient of an ETH Medal in 1998. He was awarded the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research in 2010.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. His research interests are in the areas of stochastic analysis, statistical signal processing and information theory, and their applications in wireless networks and related fields. Among his publications in these areas are *Quickest Detection* (Cambridge, U.K.: Cambridge University Press, 2009), co-authored with O. Hadjiliadis, and *Information Theoretic Security* (Now Publishers, 2009), co-authored with Y. Liang and S. Shamai.

Dr. Poor is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U. K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, in 2004–2007 as the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY, and as General Co-Chair of the 2009 IEEE International Symposium on Information Theory, held in Seoul, South Korea. He is the recipient of the 2005 IEEE Education Medal. Recent recognition of his work includes the 2008 Aaron D. Wyner Distinguished Service Award of the IEEE Information Theory Society, the 2009 Edwin Howard Armstrong Achievement Award of the IEEE Communications Society, and the 2011 IEEE Eric E. Sumner Award.