Feedback Enlarges Capacity Region of Two-Way Relay Channel

Silas L. Fong
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
Email: lhfong5@ie.cuhk.edu.hk

Raymond W. Yeung
Institute of Network Coding and
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
Email: whyeung@ie.cuhk.edu.hk

***Abstract*—We consider a two-way relay channel (TRC) in which two terminals exchange messages with the help of a relay between them. The two terminals transmit messages to the relay through the Multiple Access Channel (MAC) and the relay transmits messages to the two terminals through the Broadcast Channel (BC). We assume that the MAC and the BC do not interfere with each other, and each terminal receives signals only from the relay but not the other terminal. All the nodes are assumed to be full-duplex, which means that they can transmit and receive information at the same time. The TRC is said to be *without feedback* if each terminal node cannot use its previously received information for encoding its message. Otherwise, the TRC is said to be *with feedback*. We obtain an outer bound on the capacity region of the discrete memoryless TRC without feedback and prove that the outer bound is tighter than the cut-set outer bound. In addition, we show that using feedback can enlarge the capacity region of some discrete memoryless TRC.**

I. Introduction

We consider a two-way relay channel (TRC) [1], in which two terminals exchange messages with the help of a relay between them. The two terminals transmit messages to the relay through the Multiple Access Channel (MAC) and the relay transmits messages to the two terminals through the Broadcast Channel (BC). We assume that the MAC and the BC do not interfere with each other, and each terminal receives signals only from the relay but not the other terminal. All the nodes are assumed to be full-duplex, which means that they can transmit and receive information at the same time. The TRC is said to be *without feedback* if each terminal node cannot use its previously received information for encoding its message. Otherwise, the TRC is said to be *with feedback*.

Although several outer bounds on the capacity rate region have been proved in [2,3] and several achievable rate regions have been obtained in [4,5] for the TRC described above, the capacity region of the TRC is unknown. In this paper we investigate the capacity region of the discrete memoryless TRC and establish a new outer bound on the capacity region of the discrete memoryless TRC without feedback.

This paper is organized as follows. Section II presents the notation of this paper. Section III establishes an outer bound on the capacity region of the discrete memoryless TRC without feedback. Section IV proves that the outer bound is always contained in the cut-set outer bound [3]. Section V shows by applying our outer bound that using feedback can enlarge the capacity region of some discrete memoryless TRC. This indirectly shows that our outer bound in Section III is tighter than the cut-set bound. Section VI concludes this paper.

II. Notation

We use $Pr\{E\}$ to represent the probability of an event E. We use a capital letter X to denote a random variable with alphabet $\mathcal{X}$, and use the small letter x to denote the realization of X. We use $E[X]$ to represent the expectation of a random variable X. We use X^n to denote a random column vector $[X_1\ X_2\ \ldots\ X_n]^T$, where the components X_k have the same alphabet. We let $p_X(x)$ and $p_{X^n}(x^n)$ denote the probability mass functions of the discrete random variables X and X^n respectively. We let $p_{Y|X}(y|x)$ denote the conditional probability $Pr\{Y = y|X = x\}$ for any discrete random variables X and Y. For simplicity, we drop the subscript of a notation if there is no ambiguity. The closure of a set S is denoted by $\overline{S}$ and the convex hull of S is denoted by conv(S). To facilitate discussion, we let $\mathbb{R}_+^2$ denote the set of all pairs of non-negative real numbers.

III. Discrete Memoryless Two-Way Relay Channel without Feedback

The TRC consists of two terminal nodes t_1 and t_2 and a relay node r between them. Node t_1 and node t_2 do not communicate directly, but communicate through node r using two different channels. In each time slot, node t_1 and node t_2 transmit a symbol to node r through the Multiple Access Channel (MAC), and node r transmits a symbol to node t_1 and node t_2 through the Broadcast Channel (BC). The MAC and the BC do not interfere with each other. The exchange of one message between node t_1 and node t_2 is conducted in n time slots as follows. Node t_1 and node t_2 choose messages W_1 and W_2 independently according to the uniform distribution. Based on message W_i, node t_i constructs codeword X_i^n for $i = 1, 2$ and transmits $X_{i,k}$ through the MAC in the k^{th} time slot. The received symbol in the k^{th} time slot at r is $Y_{r,k}$ and in the same time slot, r constructs and transmits $X_{r,k}$ through the BC, which depends on $Y_{r,1}, Y_{r,2}, \ldots, Y_{r,k-1}$. The received symbol in the k^{th} time slot at node t_i is $Y_{i,k}$ for $i = 1, 2$. After n time slots, node t_1 declares $\hat{W}_2$ to be the transmitted W_2 based on Y_1^n and W_1, and node t_2 declares $\hat{W}_1$ to be the transmitted W_1 based on Y_2^n and W_2.

Definition 1: The discrete memoryless TRC consists of six finite sets $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{X}_r$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Y}_r$, one probability mass function $p_1(y_r|x_1,x_2)$ representing the MAC and one probability mass function $p_2(y_1,y_2|x_r)$ representing the BC. For any two inputs X_1 and X_2 to the MAC with a joint distribution $p(x_1,x_2)$ and any input X_r to the BC with a distribution $p(x_r)$, the relationship among X_1, X_2, X_r, the output Y_r of the MAC and the outputs Y_1 and Y_2 of the BC satisfies $p(x_1,x_2,x_r,y_1,y_2,y_r) = p(x_1,x_2,y_r)p(x_r,y_1,y_2)$ for all $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, $x_r \in \mathcal{X}_r$, $y_1 \in \mathcal{Y}_1$, $y_2 \in \mathcal{Y}_2$ and $y_r \in \mathcal{Y}_r$, where $p(x_1,x_2,y_r) = p(x_1,x_2)p_1(y_r|x_1,x_2)$ and $p(x_r,y_1,y_2) = p(x_r)p_2(y_1,y_2|x_r)$. The discrete memoryless TRC is denoted by $(\mathcal{X}_1,\mathcal{X}_2,\mathcal{X}_r,p_1(y_r|x_1,x_2),p_2(y_1,y_2|x_r),\mathcal{Y}_1,\mathcal{Y}_2,\mathcal{Y}_r)$.

When the discrete memoryless TRC is used without feedback, for any integer n, $x_1^n \in \mathcal{X}_1^n$, $x_2^n \in \mathcal{X}_2^n$, $x_r^n \in \mathcal{X}_r^n$, $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$ and $y_r^n \in \mathcal{Y}_r^n$, $p(y_1^n,y_2^n,y_r^n|x_1^n,x_2^n,x_r^n) = \prod_{k=1}^n p(y_{1,k},y_{2,k},y_{r,k}|x_{1,k},x_{2,k},x_{r,k})$. In particular,

$$p(y_r^n|x_1^n,x_2^n) = \prod_{k=1}^{n} p_1(y_{r,k}|x_{1,k},x_{2,k}) \tag{1}$$

and

$$p(y_1^n,y_2^n|x_r^n) = \prod_{k=1}^{n} p_2(y_{1,k},y_{2,k}|x_{r,k}). \tag{2}$$

Unless otherwise specified, the discrete memoryless TRC in this paper is assumed to be without feedback.

Definition 2: An (n,M_1,M_2)-code on the channel $(\mathcal{X}_1,\mathcal{X}_2,\mathcal{X}_r,p_1(y_r|x_1,x_2),p_2(y_1,y_2|x_r),\mathcal{Y}_1,\mathcal{Y}_2,\mathcal{Y}_r)$ consists of the following:

1) A message set $\mathcal{W}_1 = \{1,2,\ldots,M_1\}$ at node t_1 and a message set $\mathcal{W}_2 = \{1,2,\ldots,M_2\}$ at node t_2.
2) An encoding function $f_1 : \mathcal{W}_1 \to \mathcal{X}_1^n$ at node t_1, yielding codewords $x_1^n(1),x_1^n(2),\ldots,x_1^n(M_1)$.
3) An encoding function $f_2 : \mathcal{W}_2 \to \mathcal{X}_2^n$ at node t_2, yielding codewords $x_2^n(1),x_2^n(2),\ldots,x_2^n(M_2)$.
4) A set of n encoding functions $f_{r,k} : \mathcal{Y}_r^{k-1} \to \mathcal{X}_r$ at r, $k = 1,2,\ldots,n$, where $f_{r,k}$ is the encoding function in the k^{th} time slot such that $X_{r,k} = f_{r,k}(Y_r^{k-1})$.
5) A decoding function $g_1 : \mathcal{W}_1 \times \mathcal{Y}_1^n \to \mathcal{W}_2$ at node t_1 such that $g_1(W_1,Y_1^n) = \hat{W}_2$.
6) A decoding function $g_2 : \mathcal{W}_2 \times \mathcal{Y}_2^n \to \mathcal{W}_1$ at node t_2 such that $g_2(W_2,Y_2^n) = \hat{W}_1$.

From Definition 2, $X_1^n = f_1(W_1)$, $X_2^n = f_2(W_2)$, $X_r^n = f_r(Y_r^n)$, $\hat{W}_1 = g_2(W_2,Y_2^n)$ and $\hat{W}_2 = g_1(W_1,Y_1^n)$ for an (n,M_1,M_2)-code, where f_r is a function completely determined by $f_{r,k}$, $k = 1,2,\ldots,n$. The transmissions of messages in the TRC are illustrated in Figure 1.

Definition 3: The average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = Pr\{g_2(W_2,Y_2^n) \neq W_1\}$ and $P_{e,2}^n = Pr\{g_1(W_1,Y_1^n) \neq W_2\}$ respectively.

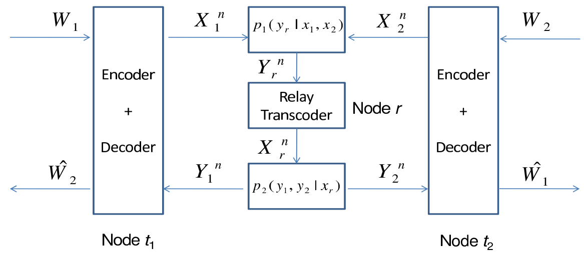

Fig. 1. A discrete memoryless TRC without feedback.

Definition 4: A rate pair (R_1,R_2) is achievable if there exists a sequence of (n,M_1,M_2)-codes with $\lim_{n\to\infty} \frac{\log_2 M_1}{n} \geq R_1$ and $\lim_{n\to\infty} \frac{\log_2 M_2}{n} \geq R_2$ such that $\lim_{n\to\infty} P_{e,1}^n = 0$ and $\lim_{n\to\infty} P_{e,2}^n = 0$.

Definition 5: The capacity region $\mathcal{R}$ of the discrete memoryless TRC without feedback is the closure of the set of all achievable rate pairs.

Let $\mathcal{R}^1$ and $\mathcal{R}^2$ denote the sets

$$\left\{ \begin{array}{c|l} (R_1,R_2) & R_1 \leq I(X_1;Y_r|X_2), R_2 \leq I(X_2;Y_r|X_1) \\ \in \mathbb{R}_+^2 & \text{where } p(x_1,x_2,y_r) = p(x_1)p(x_2) \\ & p_1(y_r|x_1,x_2) \text{ for some input distribution} \\ & p(x_1)p(x_2) \text{ for the MAC } p_1(y_r|x_1,x_2). \end{array} \right\} \tag{3}$$

and

$$\left\{ \begin{array}{c|l} (R_1,R_2) & R_1 \leq I(X_r;Y_2), R_2 \leq I(X_r;Y_1) \\ \in \mathbb{R}_+^2 & \text{where } p(x_r,y_1,y_2) = p(x_r)p_2(y_1,y_2|x_r) \\ & \text{for some input distribution } p(x_r) \text{ for the} \\ & \text{BC } p_2(y_1,y_2|x_r). \end{array} \right\} \tag{4}$$

respectively.

Lemma 1: $\mathcal{R} \subseteq \overline{\text{conv}(\mathcal{R}^1)} \cap \overline{\text{conv}(\mathcal{R}^2)}$.

Proof: Suppose (R_1,R_2) is achievable. By Definition 4, there exists a sequence of (n,M_1,M_2)-codes with

$$\lim_{n\to\infty} \frac{\log_2 M_1}{n} \geq R_1 \tag{5}$$

such that

$$\lim_{n\to\infty} P_{e,1}^n = 0. \tag{6}$$

Fix n and the corresponding (n,M_1,M_2)-code. Then,

$$(W_1,W_2) \to (X_1^n,X_2^n) \to Y_r^n \to X_r^n \to Y_2^n \tag{7}$$

forms a Markov Chain. If W_2 is fixed in the channel,

$$W_1 \to X_1^n \to Y_r^n \to X_r^n \to Y_2^n \to \hat{W}_1 \tag{8}$$

forms a Markov Chain. Since W_1 and W_2 are independent,

$$\begin{aligned} &\log_2 M_1 \\ &= H(W_1|W_2) \\ &= H(W_1|W_2,\hat{W}_1) + I(W_1;\hat{W}_1|W_2) \\ &\leq H(W_1|\hat{W}_1) + \sum_{w_2\in\mathcal{W}_2} p(w_2)I(W_1;\hat{W}_1|W_2=w_2) \\ &\leq 1 + P_{e,1}^n \log_2 M_1 + \sum_{w_2\in\mathcal{W}_2} p(w_2)I(W_1;\hat{W}_1|W_2=w_2), \end{aligned} \tag{9}$$

where the last inequality follows from Fano's inequality. Then,

$$\begin{aligned}I(X_1^n;Y_r^n|W_2) &= \sum_{w_2\in\mathcal{W}_2} p(w_2)I(X_1^n;Y_r^n|W_2=w_2)\\ &\geq \sum_{w_2\in\mathcal{W}_2} p(w_2)I(W_1;\hat{W}_1|W_2=w_2) \quad (10)\end{aligned}$$

and

$$\begin{aligned}I(X_r^n;Y_2^n|W_2) &= \sum_{w_2\in\mathcal{W}_2} p(w_2)I(X_r^n;Y_2^n|W_2=w_2)\\ &\geq \sum_{w_2\in\mathcal{W}_2} p(w_2)I(W_1;\hat{W}_1|W_2=w_2) \quad (11)\end{aligned}$$

where the two inequalities follow from applying the data processing inequality on the Markov Chain in (8). Consider the following chain of inequalities:

$$\begin{aligned}&I(X_1^n;Y_r^n|W_2)\\ &= H(Y_r^n|W_2) - H(Y_r^n|W_2,X_1^n)\\ &\overset{(a)}{=} H(Y_r^n|W_2,X_2^n) - H(Y_r^n|W_2,X_1^n)\\ &\leq H(Y_r^n|X_2^n) - H(Y_r^n|W_1,W_2,X_1^n,X_2^n)\\ &\overset{(b)}{=} H(Y_r^n|X_2^n) - H(Y_r^n|X_1^n,X_2^n)\\ &\overset{(c)}{=} H(Y_r^n|X_2^n) - \sum_{k=1}^n H(Y_{r,k}|X_{1,k},X_{2,k})\\ &\leq \sum_{k=1}^n H(Y_{r,k}|X_{2,k}) - H(Y_{r,k}|X_{1,k},X_{2,k})\\ &= \sum_{k=1}^n I(X_{1,k};Y_{r,k}|X_{2,k}) \quad (12)\end{aligned}$$

where

- (a) follows from the fact that X_2^n is a function of W_2,
- (b) follows from the Markov Chain in (7),
- (c) follows from (1).

Using (9), (10), (12), (6) and (5), we obtain

$$R_1 \leq \liminf_{n\to\infty}\sum_{k=1}^n \frac{1}{n} I(X_{1,k};Y_{r,k}|X_{2,k}), \quad (13)$$

where each $I(X_{1,k};Y_{r,k}|X_{2,k})$ is attained by some

$$\begin{aligned}&p_{X_{1,k},X_{2,k},Y_{r,k}}(x_{1,k},x_{2,k},y_{r,k})\\ &= \sum_{x_{1,i},x_{2,i},y_{r,i},i\neq k} p_{X_1^n,X_2^n}(x_1^n,x_2^n)p(y_r^n|x_1^n,x_2^n)\\ &\overset{(a)}{=} \sum_{x_{1,i},x_{2,i},y_{r,i},i\neq k} p_{X_1^n}(x_1^n)p_{X_2^n}(x_2^n)p(y_r^n|x_1^n,x_2^n)\\ &\overset{(b)}{=} \sum_{x_{1,i},x_{2,i},i\neq k} p_{X_1^n}(x_1^n)p_{X_2^n}(x_2^n)\sum_{y_{r,i},i\neq k}\prod_{j=1}^n p_1(y_{r,j}|x_{1,j},x_{2,j})\\ &= p_1(y_{r,k}|x_{1,k},x_{2,k})\sum_{x_{1,i},x_{2,i},i\neq k} p_{X_1^n}(x_1^n)p_{X_2^n}(x_2^n)\\ &= p_{X_{1,k}}(x_{1,k})p_{X_{2,k}}(x_{2,k})p_1(y_{r,k}|x_{1,k},x_{2,k}), \quad (14)\end{aligned}$$

where (a) follows from the fact that X_1^n and X_2^n are independent and (b) follows from (1).

On the other hand,

$$\begin{aligned}I(X_r^n;Y_2^n|W_2) &= H(Y_2^n|W_2) - H(Y_2^n|W_2,X_r^n)\\ &\leq H(Y_2^n) - H(Y_2^n|W_1,W_2,X_1^n,X_2^n,Y_r^n,X_r^n)\\ &\overset{(a)}{=} H(Y_2^n) - H(Y_2^n|X_r^n)\\ &\overset{(b)}{=} H(Y_2^n) - \sum_{k=1}^n H(Y_{2,k}|X_{r,k})\\ &\leq \sum_{k=1}^n H(Y_{2,k}) - H(Y_{2,k}|X_{r,k})\\ &= \sum_{k=1}^n I(X_{r,k};Y_{2,k}). \quad (15)\end{aligned}$$

where (a) follows from (7) and (b) follows from (2). Using (9), (11), (15), (6) and (5), we obtain

$$R_1 \leq \liminf_{n\to\infty}\sum_{k=1}^n \frac{1}{n} I(X_{r,k};Y_{2,k}), \quad (16)$$

where each $I(X_{r,k};Y_{2,k})$ is attained by some

$$p_{X_{r,k},Y_{1,k},Y_{2,k}}(x_{r,k},y_{1,k},y_{2,k}) = p_{X_{r,k}}(x_{r,k})p_2(y_{1,k},y_{2,k}|x_{r,k}),$$

which can be obtained by following similar procedures in (14). By symmetry, we also obtain

$$R_2 \leq \liminf_{n\to\infty}\sum_{k=1}^n \frac{1}{n} I(X_{2,k};Y_{r,k}|X_{1,k}) \quad (17)$$

and

$$R_2 \leq \liminf_{n\to\infty}\sum_{k=1}^n \frac{1}{n} I(X_{r,k};Y_{1,k}). \quad (18)$$

Using (13) and (17), we obtain $(R_1,R_2) \in \overline{\mathrm{conv}(\mathcal{R}^1)}$. Using (16) and (18), we obtain $(R_1,R_2) \in \overline{\mathrm{conv}(\mathcal{R}^2)}$. Consequently, $(R_1,R_2) \in \overline{\mathrm{conv}(\mathcal{R}^1)} \cap \overline{\mathrm{conv}(\mathcal{R}^2)}$. The theorem then follows from Definition 5 and the fact that $\overline{\mathrm{conv}(\mathcal{R}^1)} \cap \overline{\mathrm{conv}(\mathcal{R}^2)}$ is closed. ■

Proposition 2: $\mathcal{R}^2$ is convex.

Proof: It suffices to show that any convex combination of the points $(I(X_r;Y_2),I(X_r;Y_1))$ lies in $\mathcal{R}^2$. Suppose $(I(X_r';Y_2'),I(X_r';Y_1')) \in \mathcal{R}^2$ and $(I(X_r^*;Y_2^*),I(X_r^*;Y_1^*)) \in \mathcal{R}^2$ are attained by some distributions $p'(x_r)$ and $p^*(x_r)$ respectively. Fix $0 \leq \lambda \leq 1$. Let $(I_\lambda^1,I_\lambda^2)$ denote

$$\lambda(I(X_r';Y_2'),I(X_r';Y_1')) + (1-\lambda)(I(X_r^*;Y_2^*),I(X_r^*;Y_1^*)),$$

where $(I_\lambda^1,I_\lambda^2)$ is a convex combination of $(I(X_r';Y_2'),I(X_r';Y_1'))$ and $(I(X_r^*;Y_2^*),I(X_r^*;Y_1^*))$. Let $(I(\bar{X}_r;\bar{Y}_2),I(\bar{X}_r;\bar{Y}_1))$ be the point attained by the distribution $\bar{p} = \lambda p' + (1-\lambda)p^*$. Since $I(X_r;Y_2)$ and $I(X_r;Y_1)$ are concave with respect to the input distribution $p(x_r)$ for the BC channels $p_2(y_1|x_r)$ and $p_2(y_2|x_r)$ respectively (cf. [3, p.33]), it follows that $I_\lambda^1 \leq I(\bar{X}_r;\bar{Y}_2)$ and $I_\lambda^2 \leq I(\bar{X}_r;\bar{Y}_1)$, which implies $(I_\lambda^1,I_\lambda^2) \in \mathcal{R}^2$. ■

Proposition 3: $\mathcal{R}^2$ is closed.

Proof: It suffices to show that the limit of any Cauchy sequence of $(I(X_r;Y_2), I(X_r;Y_1))$ in $\mathcal{R}^2$ lies in $\mathcal{R}^2$. Let $\{(I(X_{r,n};Y_{2,n}), I(X_{r,n};Y_{1,n}))\}_{n=1,2,...}$ be a Cauchy sequence of points in $\mathcal{R}^2$ with respect to the Euclidean distance and let $\{s_n(x_r)\}_{n=1,2,...}$ be the corresponding sequence of distributions that attain the points. Regard each distribution of X_r as a point in an $|X_r|$-dimensional Euclidean space. Let $\{s_{n_k}(x_r)\}_{k=1,2,...}$ be a convergent subsequence of $\{s_n(x_r)\}_{n=1,2,...}$ with respect to the $\mathcal{L}_1$-distance, where the $\mathcal{L}_1$-distance between two distributions $u(x)$ and $v(x)$ on the same discrete alphabet $\mathcal{X}$ is defined as $\sum_{x\in\mathcal{X}}|u(x)-v(x)|$. Since the set of all distributions $\{p_{X_r}(x_r)\}$ is closed and bounded with respect to the $\mathcal{L}_1$-distance, there exists a distribution $\bar{s}(x_r)$ such that $\lim_{k\to\infty} s_{n_k}(x_r) = \bar{s}(x_r)$. Let $(I(\bar{X}_r;\bar{Y}_2), I(\bar{X}_r;\bar{Y}_1))$ denote the point attained by $\bar{s}(x_r)$. Since $(I(X_r;Y_2), I(X_r;Y_1))$ is a continuous functional of $p_{X_r}(x_r)$, it follows that

$$\lim_{n\to\infty}(I(X_{r,n};Y_{2,n}), I(X_{r,n};Y_{1,n})) = (I(\bar{X}_r;\bar{Y}_2), I(\bar{X}_r;\bar{Y}_1)),$$

which lies in $\mathcal{R}^2$. ∎

Theorem 1: $\mathcal{R} \subseteq \overline{\mathrm{conv}(\mathcal{R}^1)} \cap \mathcal{R}^2$.

Proof: It follows from Lemma 1, Proposition 2 and Proposition 3. ∎

Remark: The convexity proof for $\mathcal{R}^2$ in Proposition 2 is not applicable to $\mathcal{R}^1$ because the mixture of two independent input distributions for the MAC is in general not an independent input distribution.

IV. Comparison with Cut-Set Bound

The following proposition is reproduced from Proposition 2.5 in [6] to facilitate discussion.

Proposition 4: For discrete random variables X, Y and Z, $X \to Y \to Z$ forms a Markov Chain if and only if there exist two functions $\chi(x,y)$ and $\varphi(y,z)$ such that $p(x,y,z) = \chi(x,y)\varphi(y,z)$ for all x, y and z where $p(y) > 0$.

Let $\mathcal{C}$ denote

$$\left\{ (R_1,R_2) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{c} R_1 \le \min\{I(X_1;Y_r|X_2), I(X_r;Y_2)\}, \\ R_2 \le \min\{I(X_2;Y_r|X_1), I(X_r;Y_1)\}, \\ \text{where } p(x_1,x_2,x_r,y_1,y_2,y_r) = p(x_1,x_2) \\ p(x_r)p_1(y_r|x_1,x_2)p_2(y_1,y_2|x_r) \text{ for some input distribution } p(x_1,x_2) \text{ for the MAC and} \\ \text{some input distribution } p(x_r) \text{ for the BC.} \end{array} \right\},$$

which is proved in [2] as an outer bound for the discrete memoryless TRC. As pointed out by Young-Han Kim [7], the cut-set outer bound [3] for the TRC, namely

$$\left\{ (R_1,R_2) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{c} R_1 \le I(X_1;Y_r,Y_2|X_r,X_2), \\ R_1 \le I(X_1,X_r;Y_2|X_2), \\ R_2 \le I(X_2;Y_r,Y_1|X_r,X_1), \\ R_2 \le I(X_2,X_r;Y_1|X_1) \\ \text{where } p(x_1,x_2,x_r,y_1,y_2,y_r) = p(x_1,x_2) \\ p_1(y_r|x_1,x_2)p(x_r)p_2(y_1,y_2|x_r) \text{ for some input distribution } p(x_1,x_2,x_r) \text{ for the TRC.} \end{array} \right\},$$

can be simplified to $\mathcal{C}$ based on the fact that $(X_r,Y_1,Y_2) \to (X_1,X_2) \to Y_r$ and $(X_1,X_2,Y_r) \to X_r \to (Y_1,Y_2)$ form two Markov Chains for any distribution $p(x_1,x_2,x_r,y_1,y_2,y_r) = p(x_1,x_2)p(x_r)p_1(y_r|x_1,x_2)$ $p_2(y_1,y_2|x_r)$ (cf. Proposition 4). Let $\mathcal{C}^1$ denote the set

$$\left\{ (R_1,R_2) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{c} R_1 \le I(X_1;Y_r|X_2), R_2 \le I(X_2;Y_r|X_1) \\ \text{where } p(x_1,x_2,y_r) = p(x_1,x_2)p_1(y_r|x_1,x_2) \text{ for} \\ \text{some input distribution } p(x_1,x_2) \text{ for the MAC.} \end{array} \right\}. \tag{19}$$

Proposition 5: $\mathcal{C} = \mathcal{C}^1 \cap \mathcal{R}^2$.

Proof: Suppose $(R_1,R_2) \in \mathcal{C}$. Then, $R_1 \le \min\{I(X_1;Y_r|X_2), I(X_r;Y_2)\}$ and $R_2 \le \min\{I(X_2;Y_r|X_1), I(X_r;Y_1)\}$ for some distribution $p(x_1,x_2,x_r,y_1,y_2,y_r)$. Since $I(X_1;Y_r|X_2)$ and $I(X_2;Y_r|X_1)$ only depend on the marginal distribution $p(x_1,x_2,y_r)$, it follows that $(R_1,R_2) \in \mathcal{C}^1$. Similarly, $(R_1,R_2) \in \mathcal{R}^2$. Therefore, $\mathcal{C}^* \subseteq \mathcal{C}^1 \cap \mathcal{R}^2$. Conversely, suppose $(R_1,R_2) \in \mathcal{C}^1 \cap \mathcal{R}^2$. Then, $R_1 \le \min\{I(X_1;Y_r|X_2), I(X_r;Y_2)\}$ and $R_2 \le \min\{I(X_2;Y_r|X_1), I(X_r;Y_1)\}$ for the distribution defined as $p(x_1,x_2,y_r)p(x_r,y_1,y_2)$. Consequently, $\mathcal{C}^1 \cap \mathcal{R}^2 \subseteq \mathcal{C}^*$. ∎

Proposition 6: $H(Y|X)$ is concave with respect to $p(x,y)$.

Proof: Let $u(y) = 1/|\mathcal{Y}|$ for all $y \in \mathcal{Y}$. Then,

$$D(p(x,y)||p(x)u(y)) = -H(Y|X) + \log_2|\mathcal{Y}|, \tag{20}$$

where $D(p(x,y)||p(x)u(y))$ is the Kullback-Leibler divergence between the distributions $p(x,y)$ and $p(x)u(y)$. Since $D(p(x,y)||p(x)u(y))$ is convex with respect to $p(x,y)$ [3], it follows from (20) that $H(Y|X)$ is concave with respect to $p(x,y)$. ∎

Proposition 7: $I(X_1;Y_r|X_2)$ and $I(X_2;Y_r|X_1)$ are concave with respect to the input distribution $p(x_1,x_2)$ for the MAC $p_1(y_r|x_1,x_2)$.

Proof: Suppose $I(X_1';Y_r'|X_2')$ and $I(X_1^*;Y_r^*|X_2^*)$ are attained by some distributions $p'(x_1,x_2)$ and $p^*(x_1,x_2)$ respectively. Fix $0 \le \lambda \le 1$. Let $I(\bar{X}_1;\bar{Y}_r|\bar{X}_2)$ be the mutual information attained by the distribution $\bar{p} = \lambda p' + (1-\lambda)p^*$. Let h denote $H(Y_r|X_1 = x_1, X_2 = x_2)$, which is a function of $p_1(y_r|x_1,x_2)$. Then,

$$\begin{aligned}
&I(\bar{X}_1;\bar{Y}_r|\bar{X}_2) \\
&\quad= H(\bar{Y}_r|\bar{X}_2) - \sum_{\bar{x}_1,\bar{x}_2} \bar{p}(x_1,x_2)h \\
&\quad= H(\bar{Y}_r|\bar{X}_2) - \sum_{\bar{x}_1,\bar{x}_2} \lambda p'(x_1,x_2)h - \sum_{\bar{x}_1,\bar{x}_2}(1-\lambda)p^*(x_1,x_2)h \\
&\quad= H(\bar{Y}_r|\bar{X}_2) - \lambda H(Y_r'|X_1',X_2') - (1-\lambda)H(Y_r^*|X_1^*,X_2^*) \\
&\quad\overset{(a)}{\ge} \lambda H(Y_r'|X_2') + (1-\lambda)H(Y_r^*|X_2^*) - \lambda H(Y_r'|X_1',X_2') - \\
&\qquad (1-\lambda)H(Y_r^*|X_1^*,X_2^*) \\
&\quad= \lambda I(X_1';Y_r'|X_2') + (1-\lambda)I(X_1^*;Y_r^*|X_2^*)
\end{aligned}$$

where (a) follows from Proposition 6 that $H(Y_r|X_2)$ is concave with respect to $p(x_2,y_r)$. Consequently, $I(X_1;Y_r|X_2)$ is concave with respect to $p(x_1,x_2)$. By symmetry, $I(X_2;Y_r|X_1)$ is also concave with respect to $p(x_1,x_2)$. ∎

Proposition 8: $\mathcal{C}^1$ is convex.

Proof: Using the concavity of $I(X_1; Y_r|X_2)$ and $I(X_2; Y_r|X_1)$ proved in Proposition 7, we can prove the convexity of $\mathcal{C}^1$ as in Proposition 2. The details are omitted. ■

Proposition 9: $\mathcal{C}^1$ is closed.

Proof: Similar to the proof of Proposition 3. ■

Theorem 2: $\overline{\text{conv}(\mathcal{R}^1)} \cap \mathcal{R}^2 \subseteq \mathcal{C}$.

Proof: Since $\mathcal{R}^1 \subseteq \mathcal{C}^1$ (cf. (3) and (19)), it follows that $\text{conv}(\mathcal{R}^1) \subseteq \text{conv}(\mathcal{C}^1)$, which implies from Proposition 8 and Proposition 9 that $\overline{\text{conv}(\mathcal{R}^1)} \subseteq \mathcal{C}^1$. Consequently, $\overline{\text{conv}(\mathcal{R}^1)} \cap \mathcal{R}^2 \subseteq \mathcal{C}^1 \cap \mathcal{R}^2 = \mathcal{C}$. ■

In the next section, we will show that $\overline{\text{conv}(\mathcal{R}^1)} \cap \mathcal{R}^2$ is in fact a tighter outer bound than $\mathcal{C}$ on $\mathcal{R}$.

V. Benefit of Feedback

Consider the TRC with deterministic $p_1(y_r|x_1, x_2)$ and deterministic $p_2(y_1, y_2|x_r)$ such that all the random variables are binary,

$$Y_r = X_1 X_2 \text{ and } Y_1 = Y_2 = X_r. \tag{21}$$

We call the TRC described above the *binary multiplying relay channel* (BMRC). Consider any independent inputs X_1 and X_2 to the MAC with joint distribution $p(x_1, x_2) = p(x_1)p(x_2)$ where $Pr\{X_1 = 1\} = p$ and $Pr\{X_2 = 1\} = q$. For any $0 \leq \alpha \leq 1$, let $H(\alpha)$ denote the entropy function of a Bernoulli(α) random variable. Then,

$$\begin{aligned} I(X_1; Y_r|X_2) &\overset{(a)}{=} Pr\{X_2 = 1\} I(X_1; Y_r|X_2 = 1) \\ &\overset{(b)}{=} qI(X_1; X_1|X_2 = 1) \\ &\overset{(c)}{=} qH(X_1) \\ &= qH(p), \end{aligned}$$

where

(a) follows from the fact that $p_{Y_r|X_2}(0|0) = 1$ (cf. (21)),

(b) follows from (21),

(c) follows from the fact that X_1 and X_2 are independent.

Similarly, $I(X_2; Y_r|X_1) = pH(q)$. Then, $\mathcal{R}^1$ for the BMRC (cf. (3)) becomes

$$\left\{ \begin{matrix} (R_1, R_2) \\ \in \mathbb{R}_+^2 \end{matrix} \middle| \begin{matrix} R_1 \leq qH(p), R_2 \leq pH(q) \\ \text{for some } 0 \leq p \leq 1 \text{ and } 0 \leq q \leq 1. \end{matrix} \right\},$$

which is denoted by $\mathcal{K}$. It can be shown by standard optimization procedures that the largest value for $qH(p) + pH(q)$ is 1.234, which implies that the largest possible sum rate in $\text{conv}(\mathcal{K})$ is 1.234. Therefore, the maximum equal-rate pair in $\text{conv}(\mathcal{K})$ is less than or equal to $(1.234/2, 1.234/2) = (0.617, 0.617)$ when the BMRC is used without feedback.

In the general setting with feedback, the BMRC can be used as the binary multiplying channel (BMC) [8], shown in Figure 2, as follows. The BMC is almost a special case of the BMRC except that the product of the two inputs of the channel is broadcast without a unit delay. Suppose a codebook $\mathcal{B}$ designed for the BMC is given and the terminals of the BMRC want to communicate using the same codebook. Specifically, each terminal i transmits two messages $W_{i,1}$ and $W_{i,2}$ using $\mathcal{B}$ by interleaving the codewords: each terminal in the BMRC transmits $X_i^n(w_{i,1})$ during odd time slots and transmits $X_i^n(w_{i,2})$ during even time slots, while r forwards every bit Y_{k-1} to the two terminals in the k^{th} time slot.

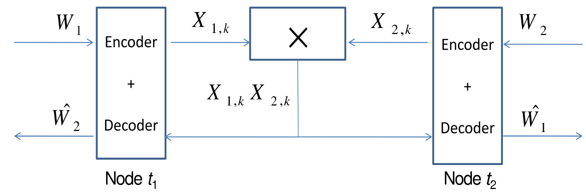

Fig. 2. Binary multiplying channel (BMC).

It is shown in [8] that the rate pair $(0.63056, 0.63056)$ is in the capacity region of the BMC and it is achievable by some coding scheme with feedback. Therefore, $(0.63056, 0.63056)$ is in the capacity region of the BMRC with feedback. However, the rate pair $(0.63056, 0.63056)$ is outside $\overline{\text{conv}(\mathcal{K})}$ and is therefore outside $\overline{\text{conv}(\mathcal{R}^1)} \cap \mathcal{R}^2$. It then follows that using feedback can enlarge the capacity region of the BMRC and $\overline{\text{conv}(\mathcal{R}^1)} \cap \mathcal{R}^2 \subsetneq \mathcal{C}$.

VI. Conclusion

We investigate the full-duplex discrete memoryless TRC without feedback and prove a new outer bound on the capacity region, which is tighter than the cut-set outer bound. We also investigate a particular TRC called BMRC and show by applying our outer bound that using feedback can enlarge the capacity region of the BMRC. The capacity region of a general TRC is still an open problem. Further research includes enhancing capacity bounds for a general TRC and determining the capacity region for some special TRCs.

Acknowledgment

The authors would like to thank Young-Han Kim for useful discussions on the cut-set outer bound for the discrete memoryless TRC. The work of Raymond Yeung was partially supported by a grant from the University Grants Committee of the HKSAR, China (Project No. AoE/E-02/08).

References

[1] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE ISIT'06*, Jul. 2006, pp. 1668–1672.

[2] S. L. Fong and R. W. Yeung, "Capacity bounds for full-duplex two-way relay channel with feedback," in *Proc. IEEE Information Theory and Applications Workshop (ITA), UCSD, San Diego, CA*, Feb. 2011.

[3] T. M. Cover, *Elements of Information Theory*, 2nd ed. Wiley, 2006.

[4] W. Nam, S.-Y. Chung and Y. H. Lee, "Capacity bounds for two-way relay channels," in *Proc. IEEE International Zurich Seminar on Communications*, Mar. 2008, pp. 144–147.

[5] A. S. Avestimehr, A. Sezgin and D. N. C. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," in *Proc. Allerton Conference on Communication, Control, and Computing*, Sep. 2008, pp. 1582–1589.

[6] R. W. Yeung, *Information Theory and Network Coding*. Springer, 2008.

[7] Private communication.

[8] J. P. M. Schalkwijk, "On an extension of an achievable rate region for the binary multiplying channel," *IEEE Trans. Inf. Theory*, vol. 29, pp. 445–448, May 1983.