Abstract—We consider a two-way relay channel (TRC) in which two terminals exchange their 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. We present with detailed proofs an outer bound on the capacity region of each of the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. We show that the outer bound for the discrete memoryless TRC with feedback is loose for some TRC.

I. INTRODUCTION

We consider a two-way relay channel (TRC) [1], in which two terminals exchange their 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 terminal nodes can use their previously received messages for encoding their messages, and the TRC described above is said to be with feedback.

Although several achievable rate regions for the TRC with feedback have been obtained in [2] and [3], the capacity region of the TRC is unknown. Knopp [4] stated an outer bound on the capacity region of each of the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. However, the proofs in [4] are incomplete, which motivates us to present detailed proofs of the outer bounds stated in [4]. The proofs are nontrivial.

This paper is organized as follows. Section II presents the notation of this paper. Section III presents the proof of the outer bound on the capacity region of the Gaussian TRC with feedback. Section IV shows that the outer bound for the discrete memoryless TRC with feedback is loose for some TRC. Section V presents the proof of the outer bound on the capacity region of the Gaussian TRC with feedback. Section VI concludes this paper.
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. Suppose node $t_1$ and node $t_2$ exchange information in $n$ time slots. Then, the variables involved in the communication through the TRC are:

- $W_i \in \{1, 2, \ldots, M_i\}$: messages of node $t_i$, $i = 1, 2$,
- $X_i^n = [X_{i,1}, X_{i,2}, \ldots, X_{i,n}]^T$: codewords transmitted by node $t_i$, $i = 1, 2$,
- $Y_i^n = [Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n}]^T$: channel outputs at the relay,
- $X_r^n = [X_{r,1}, X_{r,2}, \ldots, X_{r,n}]^T$: codewords transmitted by the relay,
- $Y_r^n = [Y_{r,1}, Y_{r,2}, \ldots, Y_{r,n}]^T$: channel outputs at node $t_r$, $i = 1, 2$,
- $\hat{W}_i \in \{1, 2, \ldots, M_i\}$: message estimates of message $W_i$, $i = 1, 2$.

Node $t_1$ and node $t_2$ choose messages $W_1$ and $W_2$ independently according to the uniform distribution. In the $k$th time slot, node $t_1$ constructs and transmits $X_{1,k}$, which is a function of $W_1$ and $Y_1^{k-1}$ for $i = 1, 2$, to node $r$ through the MAC. The received symbol at node $r$ is $Y_{r,k}$. In the same time slot, node $r$ constructs and transmits $X_{r,k}$, which depends on $Y_1^{k-1}$, to node $t_1$ and node $t_2$ through the BC. The received symbol at node $t_i$ is $Y_{r,k}$ for $i = 1, 2$. After $n$ time slots, node $t_1$ declares $\hat{W}_1$ to be the transmitted $W_2$ based on $Y_1^n$ and $W_1$, and node $t_2$ declares $\hat{W}_2$ 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 $X_1$, $X_2$, $X_r$, $Y_1$, $Y_2$ and $Y_r$, one probability mass function $p_1(y_1|x_1,x_2)$ 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 $x_1 \in X_1$, $x_2 \in X_2$, $x_r \in X_r$, $y_r \in Y_r$, $y_1 \in Y_1$ and $y_2 \in Y_2$, where

$$p(x_1,x_2,y_r) = p(x_1,x_2)p_1(y_1|x_1,x_2)$$

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 $(X_1,X_2,X_r,p_1(y_1|x_1,x_2),p_2(y_1,y_2|x_r),Y_1,Y_2,Y_r)$.

**Definition 2:** An $(n,M_1,M_2)$-code on the channel $(X_1,X_2,X_r,p_1(y_1|x_1,x_2),p_2(y_1,y_2|x_r),Y_1,Y_2,Y_r)$ 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. A set of $n$ encoding functions $f_1:k \rightarrow \mathcal{W}_1$ such that $f_1(x)$ is the encoding function in the $k$th time slot such that $X_{1,k} = f_1(W_1)$. In the $k$th time slot such that $X_{1,k} = f_1(W_1)$.
3. A set of $n$ encoding functions $f_2:k \rightarrow \mathcal{W}_2$ such that $f_2(x)$ is the encoding function in the $k$th time slot such that $X_{2,k} = f_2(W_2)$.
4. A set of $n$ encoding functions $f_3:k \rightarrow \mathcal{W}_3$ such that $f_3(x)$ is the encoding function in the $k$th time slot such that $X_{3,k} = f_3(W_3)$.
5. A decoding function $g_1: \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathcal{W}_1$ at node $t_1$ such that $g_1(W_1, W_2) = \hat{W}_1$.
6. A decoding function $g_2: \mathcal{W}_2 \times \mathcal{W}_3 \rightarrow \mathcal{W}_2$ at node $t_2$ such that $g_2(W_2, W_3) = \hat{W}_2$.

The transmissions of messages in the TRC are illustrated in Figure 1. The following proposition is cited from Proposition 2.5 of [5] to facilitate discussion.

**Proposition 1:** For the discrete random variables $X$, $Y$ and $Z$, $X \rightarrow Y \rightarrow Z$ forms a Markov Chain if and only if there exists two functions $\chi(x,y)$ and $\phi(y,z)$ such that $p(x,y,z) = \chi(x,y) \phi(y,z)$ for all $x$, $y$ and $z$ where $p(y) > 0$.

**Lemma 2:** For any $(n,M_1,M_2)$-code on the discrete memoryless TRC,

$$U \rightarrow (X_{1,k},X_{2,k}) \rightarrow Y_{r,k},$$

$$U \rightarrow X_{r,k} \rightarrow (Y_{1,k},Y_{2,k}),$$

and

$$(U,Y_{r,k}) \rightarrow (X_{1,k},X_{2,k}) \rightarrow Y_{r,k}$$

form four Markov Chains for $k = 1, 2, \ldots, n$, where $U = (W_1,W_2,X_{1,k}^{k-1},X_{2,k}^{k-1},Y_{1,k}^{k-1},Y_{2,k}^{k-1})$. In addition, $p(x_{1,k},x_{2,k},y_{r,k}) = p(x_{1,k},x_{2,k})p_1(y_{r,k}|x_{1,k},x_{2,k})$ and $p(x_{r,k},y_{1,k},y_{2,k}) = p(x_{r,k})p_2(y_{1,k},y_{2,k}|x_{r,k})$ for $k = 1, 2, \ldots, n$.

**Proof:** Fix an $(n,M_1,M_2)$-code on the discrete memoryless TRC and consider any $1 \leq k \leq n$. Figure 2 shows the dependency graph of the variables $W_1, W_2, X_{1,j}, X_{2,j}, X_{r,j}$,
Y_{i,j}, Y_{2,j} and Y_{r,j} for j = 1, 2, 3. We can then extend the dependency graph in such a way that the extended graph includes all the variables by the kth time slot and conclude from the extended graph that for any w_1 \in W_1, w_2 \in W_2, x_1^k \in X_1^k, x_2^k \in X_2^k, x_r^k \in X_r^k, y_1^k \in Y_1^k, y_2^k \in Y_2^k and y_r^k \in Y_r^k with \( p(u) > 0 \) where u = (w_1, w_2, x_1^k, x_2^k, x_r^k, x_r^{k-1}, y_1^{k-1}, y_r^{k-1}-y_r^{k-1}){
\begin{align*}
    p(u, x_1, x_2, x_r, y_1, y_2, y_r) &= p(u)p(x_1, x_2, | u)p_1(y_r | x_1, x_2) p(x_r | u) \\
    &= p(u) p_2(y_1, y_2 | x_r). 
\end{align*}
\tag{5}
\]

If \( p(u) > 0 \), then
\[ p(u, x_1, x_2, x_r, y_1, y_2, y_r) = \sum_{x_r, y_1, y_2, y_r} p(u, x_1, x_2, x_r, y_1, y_2, y_r) \tag{6} \]

where (a) follows from (5). If \( p(u) = 0 \), then
\[ p(u, x_1, x_2, x_r, y_1, y_2, y_r) = p(u, x_1, x_2, x_r, y_1, y_2, y_r) = 0. \tag{7} \]

It then follows from (6) and (7) that for all u \in U, x_1, x_2 \in X_r, and y_r \in Y_r,
\[ p(u, x_1, x_2, y_1, y_2, y_r) = p(u, x_1, x_2, x_r) p_1(y_r, x_1, x_2), \tag{8} \]

which implies from Proposition 1 that (1) is a Markov Chain. Similarly, it can be shown that
\[ p(u, x_r, y_1, y_2, y_r) = p(u, x_r) p_2(y_1, y_2, x_r), \tag{9} \]

which then implies from Proposition 1 that (2) is a Markov Chain.

In addition, if \( p(u) > 0 \), then
\[ p(u, x_1, x_2, y_1, y_2, y_r) = \sum_{x_r} p(u, x_1, x_2, x_r, y_1, y_2, y_r) \]
\[ = \sum_{x_r} p(u) p(x_1, x_2, x_r) p_1(y_r, x_1, x_2) p(x_r | u) \\
    = \sum_{x_r} p(u) p(x_1, x_2, x_r) p_1(y_r, x_1, x_2) p_2(y_1, y_2 | x_r). \tag{10} \]

where (a) follows from (5) and (b) follows from (9). If \( p(u) = 0 \),
\[ p(u, x_1, x_2, y_1, y_2, y_r) = 0. \tag{11} \]

Let \( \chi \) and \( \varphi \) be two functions such that
\[ \chi(u, x_1, x_2, y_1, y_2) = \begin{cases} p(u, y_1, y_2) p(x_1, x_2 | u) & \text{if } p(u) > 0, \\
    0 & \text{if } p(u) = 0 \end{cases} \]

and \( \varphi(x_1, x_2, y_r) = p_1(y_r, x_1, x_2). \) It then follows from (10) and (11) that for all u \in U, x_1, x_2 \in X_r, y_1 \in Y_1, y_2 \in Y_2 and y_r \in Y_r,
\[ p(u, x_1, x_2, y_1, y_2, y_r) = \chi(u, x_1, x_2, y_1, y_2) \varphi(x_1, x_2, y_r), \]

which then implies from Proposition 1 that (3) is a Markov Chain. Similarly, it can be shown from (5) and (8) that (4) is a Markov Chain.

The second statement of the lemma follows from summing u on both sides of (8) and (9).

**Definition 3**: For an \((n, M_1, M_2)\)-code on the discrete memoryless TRC, the average probabilities of decoding error of W_1 and W_2 are defined as \( P_{e,1}^n = P_{R} (g_2(W_2, Y_2) \neq W_1) \) and \( P_{e,2}^n = P_{R} (g_1(W_1, Y_1) \neq W_2) \) respectively.

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

**Definition 5**: The capacity region \( R^* \) of the discrete memoryless TRC with feedback is the closure of the set of all achievable rate pairs.

Let \( R^* \) denote
\[ \left\{ \left( R_1, R_2 \right) \in \mathbb{R}^2 : \begin{array}{ll}
    R_1 \geq 0, R_2 \geq 0, \\
    R_1 \leq \min \{ I(X_1; Y_r|X_2), I(X_r; Y_2) \}, \\
    R_2 \leq \min \{ I(X_2; Y_r|X_1), I(X_r; Y_1) \}
\end{array} \right\} \tag{12} \]

where \( p(x_1, x_2, x_r, y_1, y_2, y_r) = p(x_1, x_2) p(x_r) p_1(y_r, x_1, x_2) p_2(y_1, y_2 | x_r) \) for some input distribution \( p(x_1, x_2) \) for the MAC \( p_1(y_r, x_1, x_2) \) and some input distribution \( p(x_r) \) for the BC \( p_2(y_1, y_2 | x_r) \).
which is stated in [4] as an outer bound on $\mathcal{R}'$. We prove this outer bound in detail in the following theorem.

**Theorem 1**: $\mathcal{R}' \subset \mathcal{R}^*$. 

**Proof**: Suppose $(R_1, R_2)$ is in $\mathcal{R}'$. By Definition 4 and Definition 5, there exists a sequence of $(n, M_1, M_2)$-codes

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

and

$$\lim_{n \to \infty} \log_2 M_2 \geq R_2$$

such that

$$\lim_{n \to \infty} P^n_{e,1} = 0$$

and

$$\lim_{n \to \infty} P^n_{e,2} = 0.$$  \hspace{1cm} (13)

Fix $n$ and the corresponding $(n, M_1, M_2)$-code. It then follows from Lemma 2 that

$$U \rightarrow (X_{1,k}, X_{2,k}) \rightarrow Y_{r,k},$$  \hspace{1cm} (14)

$$U \rightarrow X_{r,k} \rightarrow Y_{2,k},$$  \hspace{1cm} (15)

and

$$(U, Y_{r,k}) \rightarrow X_{r,k} \rightarrow Y_{2,k}$$  \hspace{1cm} (16)

form three Markov Chains, where $U = (W_1, W_2, X_{k-1}, X_{2,k-1}, Y_{1,k-1}, Y_{2,k-1}, Y_{r,k-1})$. Since $W_1$ and $W_2$ are independent, we have

$$\log_2 M_1 = H(W_1 | W_2)$$

$$= I(W_1; Y_{r}^{n}, W_{2}^{n}|W_2) + H(W_1|Y_{r}^{n}, W_{2}^{n}, W_2)$$

$$\leq I(W_1; Y_{r}^{n}, W_{r}^{n}|W_2) + H(W_1|Y_{r}^{n}, W_{2}^{n})$$

$$\leq I(W_1; Y_{r}^{n}, W_{r}^{n}|W_2) + 1 + P^n_{e,1} \log_2 M_1,$$  \hspace{1cm} (17)

where the last inequality follows from the Fano’s inequality. Consider the following chain of inequalities:

$$I(W_1; Y_{r}^{n}, W_{r}^{n}|W_2)$$

$$= \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1})$$

$$= \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}) + I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}, Y_{2,k}^{k}, Y_{r,k}^{k})$$

$$= \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}) + H(Y_{2,k}|W_2, Y_{r,k}^{k-1}, Y_{2,k}^{k})$$

$$- H(Y_{2,k}|W_1, W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k})$$

$$\leq \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}) + H(Y_{2,k}|X_{r,k})$$

$$= \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}) + H(Y_{2,k}|X_{r,k})$$

$$= \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}) + H(Y_{2,k}|X_{r,k})$$

$$\leq \sum_{k=1}^{n} I(W_1; Y_{r,k}|W_2, Y_{2,k}^{k-1}, Y_{r,k}^{k-1}) + H(Y_{2,k}|X_{r,k})$$

for some functions $\chi$ and $\varphi$, where

\begin{align*}
(a) & \text{ follows from the construction of } Q \text{ that } Q \text{ is independent of } X_{1,k}, X_{2,k} \text{ and } Y_{r,k}, \\
(b) & \text{ follows from (22)}, \\
(c) & \text{ follows from the second statement of Lemma 2.}
\end{align*}
Following from the above,
\[
\frac{1}{n} \sum_{k=1}^{n} I(X_{1,k};Y_{r,k}|X_{2,k}) \\
\overset{(a)}{=} \frac{1}{n} \sum_{k=1}^{n} I(X_{1,k};Y_{r,k}|X_{2,k}, Q = k) \\
= I(X_{1}, Y_{r, Q}|X_{2}, Q) \\
= H(Y_{r, Q}|X_{2}, Q) - H(Y_{r, Q}|X_{1}, X_{2}, Q) \\
\leq H(Y_{r, Q}|X_{2}) - H(Y_{r, Q}|X_{1}, X_{2}, Q) \\
\overset{(b)}{=} H(Y_{r, Q}|X_{2}) - H(Y_{r, Q}|X_{1}, X_{2}, Q) \\
= I(X_{1}, Y_{r, Q}|X_{2}) \\
(25)
\]
where

(a) follows from (22),
(b) follows from the fact that \( Q \rightarrow (X_{1}, X_{2}, Q) \rightarrow Y_{r, Q} \)
forms a Markov Chain (cf. (24) and Proposition 1).

Using (20), (21), and (25), we obtain
\[
\log_{2} M_{1} \leq 1 + P_{e, 1}^{n} \log_{2} M_{1} + nI(X_{1};Y_{r};X_{2,Q}). \quad (26)
\]
By symmetry, we obtain
\[
\log_{2} M_{2} \leq 1 + P_{e, 2}^{n} \log_{2} M_{2} + nI(X_{2};Y_{r};X_{1,Q}). \quad (27)
\]
Since
\[
p_{X_{1,Q},X_{2,Q},Y_{r,Q}}(x_{1},x_{2},y_{r}) \\
= \sum_{k=1}^{n} pq_{X_{1,Q},X_{2,Q},Y_{r,Q}}(k,x_{1},x_{2},y_{r}) \\
\overset{(a)}{=} \sum_{k=1}^{n} pq_{k}p_{X_{1,k},X_{2,k}}(x_{1},x_{2})p_{r}(y_{r}|x_{1},x_{2}) \\
= \frac{1}{n} \sum_{k=1}^{n} p_{X_{1,k},X_{2,k}}(x_{1},x_{2})p_{r}(y_{r}|x_{1},x_{2}) \\
\]
where (a) follows from (23), it follows that
\( I(X_{1,Q};Y_{r,Q}|X_{2,Q}) \) and \( I(X_{2,Q};Y_{r,Q}|X_{1,Q}) \) are achieved by the distribution
\( p_{X_{1,k},X_{2,k}}(x_{1},x_{2}) = q_{n}(x_{1},x_{2}) p_{r}(y_{r}|x_{1},x_{2}) \)
where \( q_{n}(x_{1},x_{2}) = \sum_{k=1}^{n} \frac{1}{n} p_{X_{1,k},X_{2,k}}(x_{1},x_{2}) \)
is the input distribution for the MAC. Using (26), (27), (15), (16), (13) and (14), we obtain
\[
R_{1} \leq \lim inf_{n \to \infty} I(X_{1};Y_{r,Q}|X_{2,Q}) \quad (28)
\]
and
\[
R_{2} \leq \lim inf_{n \to \infty} I(X_{2};Y_{r,Q}|X_{1,Q}). \quad (29)
\]
Consider each distribution of \( (X_{1}, X_{2}) \) as a point in an \( |X_{1}| \times |X_{2}| \)-dimensional Euclidean space. Let \( \{q_{n}(x_{1},x_{2})\}_{k=1,2,...} \) be a convergent subsequence of \( \{q_{n}(x_{1},x_{2})\}_{n=1,2,...} \) with respect to the \( L_{2} \)-distance, where the \( L_{1} \)-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 joint distributions \( \{p_{X_{1},X_{2}|X_{1},x_{2}}(x_{1},x_{2})\} \) is closed with respect to the \( L_{1} \)-distance, there exists a joint distribution \( \bar{q}(x_{1},x_{2}) \) such that
\[
\lim_{k \to \infty} q_{n}(x_{1},x_{2}) = \bar{q}(x_{1},x_{2}). 
\]
Since \( I(X_{1};Y_{r,Q}|X_{2,Q}) \) is a continuous functional of \( p_{X_{1},X_{2}|X_{1},x_{2}}(x_{1},x_{2}) \), it then follows from (28) and (29) that
\[
R_{1} \leq \lim inf_{n \to \infty} I(X_{1}, Y_{r,Q}|X_{2,Q}) \leq I(X_{1};Y_{r,Q}|X_{2,Q}) \quad (30)
\]
and
\[
R_{2} \leq \lim inf_{n \to \infty} I(X_{2}, Y_{r,Q}|X_{1,Q}) \leq I(X_{2};Y_{r,Q}|X_{1,Q}) \quad (31)
\]
where \( p_{X_{1},X_{2},Y_{r}(x_{1},x_{2},y_{r})} = \bar{q}(x_{1},x_{2})p_{r}(y_{r}|x_{1},x_{2}) \).

On the other hand,
\[
\log_{2} M_{1} = H(W_{1}|W_{2}) \\
= I(W_{1};Y_{2,Q}^{n}|W_{2}) + H(W_{1}|Y_{2,Q}^{n}, W_{2}) \\
\leq I(W_{1};Y_{2,Q}^{n}|W_{2}) + 1 + P_{e, 1}^{n} \log_{2} M_{1} \quad (32)
\]
where the last inequality follows from the Fano’s inequality.

Consider the following chain of inequalities:
\[
I(W_{1};Y_{2,Q}^{n}|W_{2}) \\
= \sum_{k=1}^{n} I(W_{1};Y_{2,k}|W_{2}, Y_{2,k}^{k-1}) \\
= \sum_{k=1}^{n} H(Y_{2,k}|W_{2}, Y_{2,k}^{k-1}) - H(Y_{2,k}|W_{1}, Y_{2,k}^{k-1}, Y_{1}^{k-1}) \\
\leq \sum_{k=1}^{n} H(Y_{2,k}) - H(Y_{2,k}|W_{1}, Y_{2,k}^{k-1}, Y_{1}^{k-1}, Y_{r,k}^{k-1}) \\
\overset{(a)}{=} \sum_{k=1}^{n} H(Y_{2,k}) - H(Y_{2,k}|Y_{r,k}) \\
= \sum_{k=1}^{n} I(X_{r,k};Y_{2,k}) \\
(33)
\]
where (a) follows from the Markov Chain in (18). In addition,
\[
p_{X_{r,Q},Y_{1,Q},Y_{2,Q}|Q}(x_{r},y_{1},y_{2}|k) \\
= p_{X_{r,k},Y_{1,k},Y_{2,k}|Q}(x_{r},y_{1},y_{2}|k) \\
\overset{(a)}{=} p_{X_{r,k},Y_{1,k},Y_{2,k}}(x_{r},y_{1},y_{2}) \quad (34)
\]
and
\[
p_{Q,X_{r,Q},Y_{1,Q},Y_{2,Q}|Q}(k,x_{r},y_{1},y_{2}) \\
= p_{Q}(k)p_{X_{r,Q},Y_{1,Q},Y_{2,Q}|Q}(x_{r},y_{1},y_{2}|k) \\
\overset{(b)}{=} p_{Q}(k)p_{X_{r,k},Y_{1,k},Y_{2,k}}(x_{r},y_{1},y_{2}) \\
\overset{(c)}{=} p_{Q}(k)p_{X_{r,k}}(x_{r})p_{y_{1},y_{2}|x_{r}} \\
= \chi(k,x_{r})\varphi(x_{r},y_{1},y_{2}) \quad (35)
\]
for some functions \( \chi \) and \( \varphi \), where

(a) follows from the construction of \( Q \) that \( Q \) is independent of \( X_{r,k}, Y_{1,k} \) and \( Y_{2,k}, \)
(b) follows from (34),
(c) follows from the second statement of Lemma 2.
Following from the above,
\[
\frac{1}{n} \sum_{k=1}^{n} I(X_r; Y_2|k) = \sum_{k=1}^{n} \frac{1}{n} I(X_r, Q; Y_2|Q = k) = I(X_r, Q; Y_2|Q) = H(Y_2|Q) - H(Y_2|X_r, Q, Q) \leq H(Y_2|Q) - H(Y_2|X_r, Q)
\]
where

(a) follows from (34),
(b) follows from the fact that \( Q \rightarrow X_r, Q \rightarrow Y_2, Q \) forms a Markov Chain (cf. (36) and Proposition 1).

Using (32), (33) and (37), we obtain
\[
\log_2 M_1 \leq 1 + P_{c,1} \log_2 M_1 + nI(X_r; Q; Y_2).
\]
By symmetry, we obtain
\[
\log_2 M_2 \leq 1 + P_{n,2} \log_2 M_2 + nI(X_r; Q; Y_1).
\]
Since
\[
p_{X_r, Y_1, Q; Y_2}(x_r, y_1, y_2) = \sum_{k=1}^{n} p_{Q, X_r, Y_1, Q, Y_2}(k, x_r, y_1, y_2) = \sum_{k=1}^{n} p_{Q}(k)p_{X_r, k}(x_r)p_{2}(y_1, y_2|y_r)
\]
and
\[
p_{X_r, Q; Y_1, Q}(x_r, y_1) = \sum_{k=1}^{n} n^{-1} p_{X_r, k}(x_r)p_{2}(y_1, y_2|y_r)
\]
where (a) follows from (35), it follows that \( I(X_r; Q; Y_1, Q) \) and \( I(X_r; Q; Y_2, Q) \) are achieved by the distribution \( p_{X_r, Q; Y_1, Q, Y_2}(x_r, y_1, y_2) = s_n(x_r)p_{2}(y_1, y_2|x_r) \) where \( s_n(x_r) = \sum_{k=1}^{n} \frac{1}{n} p_{X_r, k}(x_r) \) is the input distribution for the BC. Using (38), (39), (15), (16), (13) and (14), we obtain
\[
R_1 \leq \liminf_{n \to \infty} I(X_r; Q; Y_2)
\]
and
\[
R_2 \leq \liminf_{n \to \infty} I(X_r; Q; Y_1).
\]
Consider each distribution of \( X_r \) as a point in an \( |X_r| \)-dimensional Euclidean space. Let \( \{ s_{n_k}(x_r) \}_{k=1,2,\ldots} \) be a convergent subsequence of \( \{ s_n(x_r) \}_{n=1,2,\ldots} \) with respect to the \( L_1 \)-distance and let \( \hat{s}(x_r) \) be the distribution such that \( \lim_{n \to \infty} s_{n_k}(x_r) = \hat{s}(x_r) \). Since \( I(X_r; Y_2) \) is a continuous functional of \( p_{X_r}(x_r) \), it then follows from (40) and (41) that
\[
R_1 \leq \liminf_{n \to \infty} I(X_r; Q; Y_2) \leq I(X_r; Y_2)
\]
and
\[
R_2 \leq \liminf_{n \to \infty} I(X_r; Q; Y_1) \leq I(X_r; Y_1)
\]
where \( p_{X_r, Y_1, Y_2}(x_r, y_1, y_2) = \hat{s}(x_r)p_2(y_1, y_2|x_r) \).

Let \( p(x_1, x_2, x_r, y_1, y_2, y_r) = \hat{q}(x_1, x_2)\hat{s}(x_r)p_1(y_1|x_1, x_2) \).

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{binary_multiplying_relay_channel.png}
\caption{Binary multiplying relay channel (BMRC).}
\end{figure}

The outer bound \( R^* \) is not tight in general. Consider the TRC with deterministic \( p_1(y_r|x_1, x_2) \) and \( p_2(y_1, y_2|x_r) \) such that \( X_1 = X_2 = Y_r = Y_1 = Y_2 = \{0, 1\} \) and \( \Pr\{X_r = X_1X_2 = 1 \} \) and \( \Pr\{Y_r = Y_1 = Y_2 = X_r = 1 \} \). We call the TRC described above the binary multiplying relay channel (BMRC), which is illustrated in Figure 3. Then,
\[
I(X_r; Y_1) = I(X_r; Y_2) = H(X_r).
\]
In addition,
\[
I(X_1; X_r|Y_2) \leq H(Y_r) \leq 1.
\]
Similarly,
\[
I(X_2; X_r|X_1) \leq 1.
\]
Then,
\[
R^* = \left\{ \begin{array}{ll}
R_1 \geq 0, & R_2 \geq 0, \\
R_1 \leq \min\{I(X_1; X_r|X_2), H(X_r)\}, & R_2 \leq \min\{I(X_2; X_r|X_1), H(X_r)\}
\end{array} \right.
\]
for the MAC \( p_1(y_1|x_1, x_2) \) and some input distribution \( p(x_r) \) for the BC \( p_2(y_1, y_2|x_r) \) is Bernoulli(1/2) for the BMRC.
Hekstra’s outer bound for the BMC as well as the BMRC, G Hekstra and Willems [7] obtained an outer bound for the BMC. Consequently, the capacity region of the BMRC is a fact that the terminal nodes in the BMC can construct \( X \) in any transmission scheme in the BMRC can be imitated in a single time slot. It then follows that any transmission scheme in the BMRC is assumed to be \( Y_{r,k} = X_{1,k} + X_{2,k} + Z_{r,k} \) where \( Z_{r,k} \) is a Gaussian random variable independent of \( (X_{1,k}, X_{2,k}) \). If \( X_{r,k} \) is transmitted at time \( t \) in the same time slot, the received signal at node \( t_1 \) in the time slot is assumed to be \( Y_{1,k} = X_{r,k} + Z_{1,k} \), where \( Z_{1,k} \) is a Gaussian random variable independent of \( X_{r,k} \), and the received signal at node \( t_2 \) in the time slot is assumed to be \( Y_{2,k} = X_{r,k} + Z_{2,k} \), where \( Z_{2,k} \) is a Gaussian random variable independent of \( X_{r,k} \). For any three codewords \( (x_{1,1}, x_{1,2}, \ldots, x_{1,n}) \), \( (x_{2,1}, x_{2,2}, \ldots, x_{2,n}) \) and \( (x_{r,1}, x_{r,2}, \ldots, x_{r,n}) \) that are transmitted over the channel, we require that \( \frac{1}{n} \sum_{k=1}^{n} x_{r,k} \leq P_1 \) and \( \frac{1}{n} \sum_{k=1}^{n} x_{r,k} \leq P_2 \) where \( P_1 \) and \( P_2 \) represent the maximum power constraints for node \( t_1 \), node \( t_2 \) and node \( r \) respectively.

**Definition 6:** The Gaussian TRC consists of the following:

1) Three Gaussian random variables \( Z_1 = \mathcal{N}(0,N_1), Z_2 = \mathcal{N}(0,N_2) \) and \( Z_r = \mathcal{N}(0,N_r) \).

2) A probability density function \( f_1(y_r|x_1, x_2) \) defined for all real numbers \( x_1, x_2 \) and \( y_r \) representing the MAC and a probability density function \( f_2(y_1, y_2|x_r) \) defined for all real numbers \( x_r, y_1 \) and \( y_2 \) representing the BC such that

\[
\begin{align*}
    f_1(y_r|x_1, x_2) &= f_Z(y_r - x_1 - x_2|x_1, x_2) \\
    &= f_Z(y_r - x_1 - x_2) \quad (49)
\end{align*}
\]

and

\[
\begin{align*}
    f_2(y_1, y_2|x_r) &= f_{Z_1, Z_2}(y_1 - x_r, y_2 - x_r|x_r) \\
    &= f_{Z_1, Z_2}(y_1 - x_r, y_2 - x_r). \quad (50)
\end{align*}
\]

In addition, we require that for any two finite discrete random variables \( X_1 \) and \( X_2 \) with an input distribution \( p(x_1, x_2) \) for the MAC and any finite discrete random variable \( X_r \) with an input distribution \( p(x_r) \) for the BC, 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

\[
\begin{align*}
    Pr\{X_1 = x_1, X_2 = x_2, X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2, Y_r \leq y_r\} \\
    &= Pr\{X_1 = x_1, X_2 = x_2, Y_r \leq y_r\} Pr\{X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2\}
\end{align*}
\]

for all \( x_1 \in X_1, x_2 \in X_2, x_r \in X_r, y_1 \in \mathbb{R}, y_2 \in \mathbb{R} \) and \( y_r \in \mathbb{R} \), where

\[
\begin{align*}
    Pr\{X_1 = x_1, X_2 = x_2, Y_r \leq y_r\} &= \int_{-\infty}^{y_r} f_1(v_r|x_1, x_2)p(x_1, x_2)dv_r
\end{align*}
\]
and
\[ Pr\{X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2 \} = \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} f_2(v_1, v_2|x_r)p(x_r)dv_1dv_2. \]

The Gaussian TRC is denoted by \((N_1, N_2, N_r)\).

**Definition 7:** An \((n, M_1, M_2, X_1, X_2, \ldots, X_r)\)-code on the channel \((N_1, N_2, N_r)\) with maximum power constraints \(P_1, P_2\) and \(P_r\) consists of the following:

1) A message set \(W_1 = \{1, 2, \ldots, M_1\}\) at node \(t_1\) and a message set \(W_2 = \{1, 2, \ldots, M_2\}\) at node \(t_2\).

2) A set of \(n\) encoding functions \(\alpha_{1,k} : W_1 \times Y_1^{k-1} \rightarrow X_1\) at node \(t_1\) for \(k = 1, 2, \ldots, n\), where \(\alpha_{1,k}\) is the encoding function in the \(k\)th time slot such that \(X_1 = \alpha_{1,k}(W_1, Y_1^{k-1})\). Without loss of generality, we assume that \(X_1\) is finite. In addition, every codeword \(x^n_1\) must satisfy the power constraint
\[ \sum_{k=1}^{n} x_{1,k}^2 \leq nP_1. \]

3) A set of \(n\) encoding functions \(\alpha_{2,k} : W_2 \times Y_2^{k-1} \rightarrow X_2\) at node \(t_2\) for \(k = 1, 2, \ldots, n\), where \(\alpha_{2,k}\) is the encoding function in the \(k\)th time slot such that \(X_2 = \alpha_{2,k}(W_2, Y_2^{k-1})\). Without loss of generality, we assume that \(X_2\) is finite. In addition, every codeword \(x^n_2\) must satisfy the power constraint
\[ \sum_{k=1}^{n} x_{2,k}^2 \leq nP_2. \]

4) A set of \(n\) encoding functions \(\alpha_{r,k} : Y_r^{k-1} \rightarrow X_r\) at \(r\) for \(k = 1, 2, \ldots, n\), where \(\alpha_{r,k}\) is the encoding function in the \(k\)th time slot such that \(X_r = \alpha_{r,k}(Y_r^{k-1})\). Without loss of generality, we assume that \(X_r\) is finite. In addition, every codeword \(x^n_r\) must satisfy the power constraint
\[ \sum_{k=1}^{n} x_{r,k}^2 \leq nP_r. \]

5) A decoding function \(\beta_1 : W_1 \times Y_1^n \rightarrow W_2\) at node \(t_1\) such that \(\beta_1(W_1, Y_1^n) = W_2\).

6) A decoding function \(\beta_2 : W_2 \times Y_2^n \rightarrow W_1\) at node \(t_2\) such that \(\beta_2(W_2, Y_2^n) = W_1\).

The transmissions of messages in the Gaussian TRC are illustrated in Figure 5.

**Definition 8:** For an \((n, M_1, M_2, X_1, X_2, \ldots, X_r)\)-code on the Gaussian TRC \((N_1, N_2, N_r)\) with maximum power constraints \(P_1, P_2\) and \(P_r\), the average probabilities of decoding error of \(W_1\) and \(W_2\) in \((N_1, N_2, N_r)\) are defined as \(P_{e,1}^n = Pr\{\beta_2(W_2, Y_2^n) \neq W_1\}\) and \(P_{e,2}^n = Pr\{\beta_1(W_1, Y_1^n) \neq W_2\}\) respectively.

**Definition 9:** A rate pair \((R_1, R_2)\) is Gaussian-achievable if there exists a sequence of \((n, M_1, M_2, X_1, X_2, \ldots, X_r)\)-codes for the Gaussian TRC \((N_1, N_2, N_r)\) with maximum power constraints \(P_1, P_2\) and \(P_r\) such that
\[ \lim_{n \to \infty} \frac{\log_2 M_1}{n} \geq R_1, \]
\[ \lim_{n \to \infty} \frac{\log_2 M_2}{n} \geq R_2, \]
\[ \lim_{n \to \infty} P^n_{e,1} = 0 \]
and
\[ \lim_{n \to \infty} P^n_{e,2} = 0. \]

**Theorem 2:** For the Gaussian TRC \((N_1, N_2, N_r)\) with maximum power constraints \(P_1, P_2\) and \(P_r\) \(R_C\) is a subset of
\[ \left\{ (R_1, R_2) \in \mathbb{R}^2 \mid R_1 \leq \min\left\{ \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N_1}\right), \frac{1}{2} \log_2 \left(1 + \frac{P_2}{N_2}\right) \right\}, \right. \]
\[ R_2 \leq \min\left\{ \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N_1}\right), \frac{1}{2} \log_2 \left(1 + \frac{P_2}{N_2}\right) \right\} \}\]

**Proof:** Suppose \((R_1, R_2)\) is in \(R_C\). By Definitions 9 and 10, there exists a sequence of \((n, M_1, M_2, X_1, X_2, \ldots, X_r)\)-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
\[ \lim_{n \to \infty} P^n_{e,1} = 0 \]
and
\[ \lim_{n \to \infty} P^n_{e,2} = 0. \]

Fix \(n\) and the corresponding \((n, M_1, M_2, X_1, X_2, \ldots, X_r)\)-code. Following similar procedures for proving (20), (21), (32) and (33), we can obtain
\[ \log_2 M_1 \leq 1 + \sum_{k=1}^{n} I(X_1,k; Y_r,k|X_2,k), \]
\[ \log_2 M_2 \leq 1 + \sum_{k=1}^{n} I(X_2,k; Y_r,k|X_1,k), \]
\[ \log_2 M_2 \leq 1 + \sum_{k=1}^{n} I(X_2,k; Y_r,k|X_1,k) \]
and
\[ \log_2 M_2 \leq 1 + P^n_{e,2} \log_2 M_2 + \sum_{k=1}^{n} I(X_{r,k}; Y_{1,k}). \]

It then follows from (56), (57), (54) and (55) that
\[ R_1 \leq \liminf_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} I(X_{1,k}; Y_{r,k} | X_{2,k}), \quad (58) \]
\[ R_1 \leq \liminf_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} I(X_{r,k}; Y_{2,k}), \quad (59) \]
\[ R_2 \leq \liminf_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} I(X_{2,k}; Y_{r,k} | X_{1,k}), \quad (60) \]
and
\[ R_2 \leq \liminf_{n \to \infty} \frac{1}{n} I(X_{r,k}; Y_{1,k}). \quad (61) \]

Consider the following chain of inequalities:
\[ I(X_{1,k}; Y_{r,k} | X_{2,k}) = h(Y_{r,k} | X_{2,k}) - h(Y_{r,k} | X_{1,k}, X_{2,k}) \]
\[ \overset{(a)}{=} h(X_{1,k} + X_{2,k} + Z_{r,k} | X_{2,k}) - h(Z_{r,k}) \]
\[ = h(X_{1,k} + Z_{r,k} | X_{2,k}) - h(Z_{r,k}) \]
\[ \leq h(X_{1,k} + Z_{r,k}) - h(Z_{r,k}) \]
\[ = h(X_{1,k} + Z_{r,k}) - \frac{1}{2} \log_2 2\pi e N_r \]
\[ \overset{(b)}{\leq} \frac{1}{2} \log_2 (1 + E[X_{1,k}^2]/N_r), \]
where
\[ (a) \text{ follows from } Y_{r,k} = X_{1,k} + X_{2,k} + Z_{r,k} \text{ and } (49), \]
\[ (b) \text{ follows from the fact that } X_{1,k} \text{ and } Z_{r,k} \text{ are independent, and the differential entropy of a random variable is upper bounded by the differential entropy of a Gaussian random variable with the same second moment.} \]

Therefore,
\[ \sum_{k=1}^{n} \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \leq \frac{1}{2} \sum_{k=1}^{n} \frac{1}{n} \log_2 (1 + E[X_{1,k}^2]/N_r) \]
\[ \overset{(a)}{\leq} \frac{1}{2} \log_2 \left( 1 + \sum_{k=1}^{n} \frac{E[X_{1,k}^2]}{nN_r} \right) \]
\[ = \frac{1}{2} \log_2 \left( 1 + \frac{E[\sum_{k=1}^{n} X_{1,k}^2]}{nN_r} \right) \]
\[ \overset{(b)}{\leq} \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_r} \right), \]
where
\[ (a) \text{ follows from applying Jenson’s inequality to the concave function } \log_2(1 + x), \]
\[ (b) \text{ follows from } (51) \text{ and the fact that } \log_2(1 + x) \text{ is increasing in } x. \]

Consequently,
\[ \liminf_{n \to \infty} \sum_{k=1}^{n} \frac{1}{n} I(X_{r,k}; Y_{2,k}) \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_r} \right), \]
which then implies from (58) that
\[ R_1 \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_r} \right). \quad (62) \]

By symmetry, we obtain from (60) that
\[ R_2 \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_r} \right). \quad (63) \]

Consider the following chain of inequalities:
\[ I(X_{r,k}; Y_{2,k}) = h(Y_{2,k}) - h(Y_{2,k} | X_{r,k}) \]
\[ \overset{(a)}{=} h(X_{r,k} + Z_{2,k}) - h(Z_{2,k}) \]
\[ = h(X_{r,k} + Z_{2,k}) - \frac{1}{2} \log_2 2\pi e N_2 \]
\[ \overset{(b)}{\leq} \frac{1}{2} \log_2 (1 + E[X_{2,k}^2]/N_2), \]
where
\[ (a) \text{ follows from } Y_{2,k} = X_{r,k} + Z_{2,k} \text{ and } (50), \]
\[ (b) \text{ follows from the fact that } X_{r,k} \text{ and } Z_{2,k} \text{ are independent, and the differential entropy of a random variable is upper bounded by the differential entropy of a Gaussian random variable with the same second moment.} \]

Therefore,
\[ \sum_{k=1}^{n} \frac{1}{n} I(X_{r,k}; Y_{2,k}) \leq \frac{1}{2} \sum_{k=1}^{n} \frac{1}{n} \log_2 (1 + E[X_{2,k}^2]/N_2) \]
\[ \overset{(a)}{\leq} \frac{1}{2} \log_2 \left( 1 + \sum_{k=1}^{n} \frac{E[X_{2,k}^2]}{nN_2} \right) \]
\[ = \frac{1}{2} \log_2 \left( 1 + \frac{E[\sum_{k=1}^{n} X_{2,k}^2]}{nN_2} \right) \]
\[ \overset{(b)}{\leq} \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_2} \right), \]
where
\[ (a) \text{ follows from applying Jenson’s inequality to the concave function } \log_2(1 + x), \]
\[ (b) \text{ follows from } (53) \text{ and the fact that } \log_2(1 + x) \text{ is increasing in } x. \]

Consequently,
\[ \liminf_{n \to \infty} \sum_{k=1}^{n} \frac{1}{n} I(X_{r,k}; Y_{2,k}) \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_r} \right), \]
which then implies from (59) that
\[ R_1 \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_r}{N_r} \right). \quad (64) \]
By symmetry, we obtain from (61) that

$$R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_r}{N_1}\right).$$

(65)

The theorem then follows from (62), (63), (64) and (65).

VI. CONCLUSION

We investigate the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. Knopp [4] stated an outer bound on the capacity region of each of the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. However, the proofs in [4] are incomplete. We therefore present detailed proofs of the outer bounds stated in [4]. We also show that the outer bound on the capacity region of the discrete memoryless TRC with feedback is loose for some TRC.

Further research includes obtaining achievable rate regions and capacity bounds for a general TRC, since the capacity region of a general TRC is still an open problem.

ACKNOWLEDGMENT

The authors would like to thank Profs. Abbas El Gamal and Young-Han Kim for the useful discussion. The work of the authors was partially supported by a grant from the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. AoE/E-02/08).

REFERENCES