On the inner and outer bounds for 2-receiver
discrete memoryless broadcast channels

Chandra Nair, CUHK and Vincent Wang Zizhou, CUHK

Abstract—We study the best known general inner bound[1] and
outer bound[2] for the capacity region of the two user discrete
memory less channel. We prove that a seemingly stronger outer
bound is equivalent to a weaker form of the outer bound that
was also presented in [2].

The bounds matched for all channels for which the capacity
region was known and it was not known whether the regions
described by these bounds are same or different. In this paper,
under the assumption of a conjecture on a particular channel,
we prove that the bounds are different.

I. INTRODUCTION

In [3], Cover introduced the notion of a broadcast channel through which one sender transmits information to two or
more receivers. For the purpose of this paper we focus our
attention on broadcast channels with precisely two receivers.

Definition: A broadcast channel (BC) consists of an input
alphabet $X$ and output alphabets $Y_1$ and $Y_2$ and a probability
transition function $p(y_1, y_2|x)$. A $((2^{nR_1}, 2^{nR_2}), n)$ code
for a broadcast channel consists of an encoder
\[ x^n : 2^{nR_1} \times 2^{nR_2} \rightarrow X^n, \]
and two decoders
\[ \hat{Y}_1 : Y^n_1 \rightarrow 2^{nR_1}, \]
\[ \hat{Y}_2 : Y^n_2 \rightarrow 2^{nR_2}. \]

The probability of error $P_e^{(n)}$ is defined to be the probability
that the decoded message is not equal to the transmitted
message, i.e.,
\[ P_e^{(n)} = P \left( \{ \hat{Y}_1(Y_1^n) \neq W_1 \} \cup \{ \hat{Y}_2(Y_2^n) \neq W_2 \} \right) \]
where the message is assumed to be uniformly distributed
over $2^{nR_1} \times 2^{nR_2}$.

A rate pair $(R_1, R_2)$ is said to be achievable for the broad-
cast channel if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$
codes with $P_e^{(n)} \rightarrow 0$. The capacity region of the broadcast
channel with is the closure of the set of achievable rates. The capacity region of the two user discrete memoryless channel
is unknown.

The capacity region is known for lots of special cases such
as degraded, less noisy, more capable, deterministic, semi-
deterministic, etc. - see [4] and the references therein.

General inner and outer bounds for the two-user discrete
memoryless broadcast channel have also been known in liter-

\[ R_1 \leq I(U; W_1) \]
\[ R_2 \leq I(V; W_2) \]
\[ R_1 + R_2 \leq \min \{ I(W; Y_1), I(W; Y_2) \} + I(U; Y_1|W) \]
\[ + I(V; Y_2|W) - I(U; V|W) \]
for any $p(u, v, w, x)$ such that $(U, V, W) \rightarrow X \rightarrow (Y_1, Y_2)$
form a Markov chain.

Theorem 1: [Mártón ’79] The following rate pairs are
achievable:
\[ R_1 \leq I(U, W_1; Y_1) \]
\[ R_2 \leq I(V, W_2; Y_2) \]
\[ R_1 + R_2 \leq \min \{ I(W; Y_1), I(W; Y_2) \} + I(U; Y_1|W) \]
\[ + I(V; Y_2|W) - I(U; V|W) \]
\[ R_1 + R_2 \leq I(U, W_1; Y_1) + I(V; Y_2; U, W) \]
\[ I(V; Y_2) + I(U; Y_1|V) \]
over all $p(u, v, x)$ such that $(U, V, W) \rightarrow X \rightarrow (Y_1, Y_2)$
form a Markov chain.

Theorem 2: [Nair-El Gamal ’07] The union over the rate
pairs satisfying
\[ R_1 \leq I(U, W_1; Y_1) \]
\[ R_2 \leq I(V, W_2; Y_2) \]
\[ R_1 + R_2 \leq \min \{ I(U, W_1; Y_1) + I(V; Y_2; U, W), \]
\[ I(V; Y_2) + I(U; Y_1|V) \} \]
form a Markov chain constitutes an outer bound to the
capacity region.

Remark 1: Both the bounds are tight for all the special
classes of two-user broadcast channels for which the capacity
region is known. However, since the bounds are difficult to
evaluate in general it is not known whether the tightness
of these bounds is specific to the scenarios or whether they
coincide yielding the capacity region.

A possibly weaker form of the outer bound was also
presented in [2] by removing the independence between $U$ and $V$. Under this relaxation we have the following:

Theorem 3: [Nair-El Gamal ’07] The union over the rate
pairs satisfying
\[ R_1 \leq I(U; Y_1) \]
\[ R_2 \leq I(V; Y_2) \]
\[ R_1 + R_2 \leq \min \{ I(U; Y_1) + I(V; Y_2|U), \]
\[ I(V; Y_2) + I(U; Y_1|V) \} \]
over all $p(u, v, x)$ such that $(U, V, W) \rightarrow X \rightarrow (Y_1, Y_2)$
form a Markov chain constitutes an outer bound to the
capacity region.

One of the main results of the paper is the following: The
regions described by Theorems 2 and 3 are equivalent.

This simplifies the evaluation of the outer bound as it was
shown in [2] that we can make either one of the following
two assumptions (without loss of generality) for the purpose
of evaluating the region described by Theorem 3.

- One can assume $|U|, |V| \leq |X| + 2.$ or
• $X$ is a deterministic function of $U, V$ and $\|U\|, \|V\| \leq \|X\|(\|X\| + 2)$.

The organization of the paper is as follows. In Section II we show that the regions described by Theorem 2 and Theorem 2 are the same. In Section III we study the binary skew-symmetric channel [5] and conjecture that the inner and outer bounds are different for this channel.

II. ON EVALUATION OF THE BOUNDS

A. Evaluating the outer bound

Let $\mathcal{R}$ denote the region described by the outer bound in Theorem 2. Let $\mathcal{R}_1$ be the region described by the outer bound in Theorem 3.

**Theorem 4:** The regions $\mathcal{R}$ and $\mathcal{R}_1$ coincide, i.e. $\mathcal{R} = \mathcal{R}_1$.

**Proof:** Clearly, by setting $U' = (U, W)$ and $V' = (V, W)$, we have that $\mathcal{R} \subseteq \mathcal{R}_1$. Therefore it suffices to show that $\mathcal{R}_1 \subseteq \mathcal{R}$.

The idea of the proof is as follows: Given a $(U, V)$ we will produce a $(U^*, V^*, W^*)$ with $U^*, V^*$ being independent such that

$$
\begin{align*}
I(U; Y_1) &= I(U^*, W^*; Y_1) \\
I(V; Y_2) &= I(V^*, W^*; Y_2) \\
I(U; Y_1|V) &= I(U^*; Y_1|V^*, W^*) \\
I(V; Y_2|U) &= I(V^*; Y_2|U^*, W^*)
\end{align*}
$$

(1)

Let $(U, V, X)$ be a triple such that $(U, V) \rightarrow X \rightarrow (Y_1, Y_2)$ form a Markov chain. Let $\mathcal{V} = \{0, 1, \ldots, m - 1\}$. Define new random variables $U^*, V^*, W^*$ and a distribution $p(u^*, v^*, w^*, x)$ according to

$$
P(U^* = u, V^* = i, W^* = j, X = x) = \frac{1}{m}P(U = u, V = (i + j)_m, X = x),$$

where $(\cdot)_m$ denotes the $\bmod$ operation.

It is straightforward to check the following:

$$
P(U^* = u, V^* = i) = \frac{1}{m}P(U = u)$$

and hence independent,

$$
P(U^* = u, W^* = i, X = x) = \frac{1}{m}P(U = u, X = x),$$

$$
P(V^* = i, W^* = j, X = x) = \frac{1}{m}P(V = (i + j)_m, X = x).$$

From the above it follows in a straightforward manner that (1) holds and thus completes the proof. \hfill \blacksquare

B. On evaluating the inner bound

In a conversation between Bruce Hajek and the authors they concluded that choosing $X$ to be a deterministic function of $U, V, W$ suffices for Theorem 1. To show this first consider an independent random variable $\Theta$ uniformly distributed on $[0, 1]$. Use $\Theta$ that to make $X$ a deterministic function of $U, V, W$. Let $\Theta_k$ be a discretization of $\Theta$ into intervals of length $\frac{1}{k}$.

Set $\hat{W} = (\Theta_k, W)$ and choose $X$ to be a deterministic function of $(U, V, \hat{W}, \Theta_k)$ (ignoring the discrepancies caused due to discretization) as prescribed by the tuple $(U, V, \hat{W}, \Theta).$ Observe that for sufficiently large $k$ the region determined by $(U, V, \hat{W}, X)$ will approximately cover the region determined by $(U, V, W, X)$ (errors due to discretization can be made sufficiently small). This will imply that $\mathcal{R} \subseteq \mathcal{R}_d$ for Theorem 1.

III. THE BINARY SKEW-SYMMETRIC CHANNEL

A. On evaluating Märtön inner bound

We consider the following channel [5] called the Binary skew-symmetric channel, BSSC. For ease we restrict ourselves to the case $p = \frac{1}{2}$.

![Fig. 1. Binary Skew Symmetric Channel](image)

**Remark 2:** The channel, BSSC, has already appeared in a couple of instances to produce the following surprising results:

- In [5] BSSC was used to show that using the auxiliary random variable $W$ in the Cover-van der Meulen achievable region, even in the absence of rate $R_0$ (common information), enhanced the achievable region.
- In [2] BSSC was used to show that an outer bound to 2-user broadcast channel by Korner and Märtön [1] was not tight and could be tightened by using a weaker version of Theorem 2.

Backed by numerical simulations we make the following conjecture about the BSSC.

**Conjecture 1:** Let $(U, V)$ be auxiliary random variables such that $(U, V) \rightarrow X \rightarrow (Y_1, Y_2)$ form a Markov chain. Then the following holds:

$$I(U; Y_1) + I(V; Y_2) - I(U; V) \leq \max\{I(X; Y_1), I(X; Y_2)\}.$$ 

**Remark 3:** This conjecture implies that Marton’s bound without the random variable $W$ reduces to the time-division region.

When $U$ and $V$ are independent, this conjecture has been established in the appendix of [5]. In this paper, we shall establish the validity of the conjecture for some ranges of $P(X = 0).$

By symmetry of BSSC the maximum of the term $I(U; Y_1) + I(V; Y_2) - I(U; V)$ is same for $P(X = 0) = \eta$ and $P(X = 0) = \eta.$
0) = 1 − η and hence it suffices to consider η in the range 0 ≤ η ≤ \frac{1}{2}.

Observe that
\[ I(U; Y_1) + I(V; Y_2) - I(U; V), \]
\[ = I(U; Y_1) + I(V; Y_2, U) - I(U; V), \]
\[ = I(U; Y_1) + I(V; Y_2 | U), \]
\[ = I(X; Y_2) + I(U; Y_1) - I(U; Y_2). \]

Figure 2 plots \( H(Y_1) - H(Y_2) \) and the line \( 2\eta - 1 \) as a function of \( P(X = 0) = \eta \). Let \( f(\eta) = H(\frac{\eta}{2}) - H(\frac{1-\eta}{2}) \).

Fig. 2. The plot of the function \( f(\eta) = H(\frac{\eta}{2}) - H(\frac{1-\eta}{2}) \).

where \( H(\cdot) \) denotes the binary entropy function. Then it is easy to see that \( f(\eta) \) is concave in \( 0 \leq \eta \leq \frac{1}{2} \) and convex in the remaining region, \( \frac{1}{2} \leq \eta \leq 1. \)

Suppose that \( P(X = 0) = \eta \) and we seek the \( U \) that maximizes \( I(U; Y_1) - I(U; Y_2) \) subject to \( U \rightarrow X \rightarrow (Y_1, Y_2) \) being Markov and \( P(X = 0) = \eta \). Then it is not difficult to see that the optimal choice would be to set \( U = X \) for all \( \eta \leq \eta_0 \approx 0.269607 \). where \( \eta_0 \) is the unique solution of the equation
\[ f'(\eta) = \frac{1-f(\eta)}{1-\eta}, \]

or in other words the point at which the line joining \((\eta_0, f(\eta_0))\) to the point \((1,1)\) is a tangent to the curve \( f(\eta) \).

This implies that for \( \eta \leq \eta_0 \approx 0.269607 \), we have
\[ I(U; Y_1) + I(V; Y_2) - I(U; V), \]
\[ = I(X; Y_2) + I(U; Y_1) - I(U; Y_2), \]
\[ = I(X; Y_2) + I(X; Y_1) - I(X; Y_2), \]
\[ = I(X; Y_1). \]

Further using the symmetry of BSSC and the fact that the maximum of \( I(U; Y_1) + I(V; Y_2) - I(U; V) \) is same for \( P(X = 0) = \eta \) or \( 1-\eta \), we have the following result.

Lemma 1: Conjecture 1 is true as long as
\[ \max\{P(X = 0), P(X = 1)\} \leq \eta_0 \approx 0.269607. \]

Assuming Conjecture 1 is true we can now analyze the sum rate of the Marton inner bound with the random variable \( W \).

Theorem 1 implies
\[ R_1 + R_2 \leq \min\{I(W; Y_1), I(W; Y_2)\} \]
\[ + I(U; Y_1 | W) + I(V; Y_2 | W) - I(U; V | W). \]

Let \( \mathcal{W}_0 = \{w : P(X = 0 | W = w) \leq 0.5\} \) and \( \mathcal{W}_1 = \{w : P(X = 0 | W = w) > 0.5\} \). Let \( T \) be a function of \( W \) defined by
\[ T = \begin{cases} 0 & \text{if } w \in \mathcal{W}_0, \\ 1 & \text{if } w \in \mathcal{W}_1. \end{cases} \]

We have the following bound on the sum rate
\[ R_1 + R_2 \leq \min\{I(W, T; Y_1), I(W, T; Y_2)\} \]
\[ + I(U; Y_1 | W, T) + I(V; Y_2 | W, T) \]
\[ - I(U; V | W, T) \]
\[ \leq \min\{I(W, T; Y_1), I(W, T; Y_2)\} \]
\[ + P(T = 0)I(X; Y_1 | W, T = 0) \]
\[ + P(T = 1)I(X; Y_2 | W, T = 1) \]
\[ \leq \min\{I(T; Y_1), I(T; Y_2)\} \]
\[ + P(T = 0)I(X; Y_1 | T = 0) \]
\[ + P(T = 1)I(X; Y_2 | T = 1). \]

Here \((a)\) follows from Conjecture 1 and \((b)\) follows from the fact that
\[ P(T = 1)I(W; Y_1 | T = 1) \leq P(T = 1)I(W; Y_1 | T = 0), \]
\[ P(T = 0)I(W; Y_2 | T = 0) \leq P(T = 0)I(W; Y_1 | T = 0). \]

In [2] the sum rate of the pairs \((R_1, R_2)\) described by Theorem 3 was evaluated and it was shown that the maximum sum rate was bounded by 0.3711. (correct to 4 decimal places).

B. Evaluating outer bound - BSSC

In [2] the sum rate of the pairs \((R_1, R_2)\) described by Theorem 3 was evaluated and it was shown that the maximum sum rate was bounded by 0.3711. (correct to 4 decimal places). By Theorem 4 we know that the regions described by Theorem 2 and Theorem 3 coincide. Thus we have that the outer bound described in Theorem 2 is strictly larger than the Marton inner bound described by Theorem 1 (assuming Conjecture 1) and the inner and outer bounds differ for BSSC.

IV. Conclusion

In this paper, we study the inner and outer bounds for the 2-user discrete memoryless broadcast channel. We prove that for the purpose of evaluating the outer bound the region described by a weaker version (which is easier to evaluate) indeed coincides with a stronger version.

The bounds matched for all the special classes of channels for which the capacity was known. It is not known if the
bounds were inherently different or not. We then studied the bounds for the particular case of the binary skew symmetric channel (BSSC). We present a conjecture that, if proved, would establish that the inner and the outer bounds are indeed not tight for BSSC. Numerical simulations also indicate that the bounds differ for BSSC.

This definitely indicates that one of the bounds or possibly both are weak. We have demonstrated that resolving the capacity region for the BSSC would definitely give a strong hint on the capacity region of the broadcast channel for two users.

ACKNOWLEDGMENTS

The authors would like to acknowledge Prof. Bruce Hajek for very stimulating discussions during his visit to CUHK during which we realized that deterministic $X$ also suffices for the evaluation of the Márton inner bound. The authors would also like to acknowledge some valuable suggestions and stimulating exchanges on the broadcast channel and on BSSC by Prof. Abbas El Gamal.

REFERENCES