

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 identical to a weaker form of the outer bound that was also presented in [2]. We are able to further express the best outer bound in a form that is computable, i.e. there are bounds on the cardinalities of the auxiliary random variables.

The inner and outer bounds coincide for all channels for which the capacity region is known and it is not known whether the regions described by these bounds are same or different. We present a channel, where assuming a certain conjecture backed by simulations and partial theoretical results, one can show 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 \mathcal{X} and output alphabets \mathcal{Y}_1 and \mathcal{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

$$x^n : 2^{nR_1} \times 2^{nR_2} \rightarrow \mathcal{X}^n,$$

and two decoders

$$\hat{\mathcal{W}}_1 : \mathcal{Y}_1^n \rightarrow 2^{nR_1}$$

$$\hat{\mathcal{W}}_2 : \mathcal{Y}_2^n \rightarrow 2^{nR_2}.$$

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.,

$$P_e^{(n)} = \mathbf{P} \left(\{\hat{\mathcal{W}}_1(Y_1^n) \neq \mathcal{W}_1\} \cup \{\hat{\mathcal{W}}_2(Y_2^n) \neq \mathcal{W}_2\} \right)$$

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 broadcast 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 literature. Here we state the best known inner and outer bounds for the region from the literature.

Bound 1: [Märton '79] The following rate pairs are achievable:

$$R_1 \leq I(U, W; Y_1)$$

$$R_2 \leq I(V, W; Y_2)$$

$$R_1 + R_2 \leq \min\{I(W; Y_1), I(W; Y_2)\} + I(U; Y_1|W) + I(V; Y_2|W) - I(U; V|W)$$

for any $p(u, v, w, x)$ such that $(U, V, W) \rightarrow X \rightarrow (Y_1, Y_2)$ form a Markov chain.

Bound 2: [Nair-El Gamal '07] The region \mathcal{R} defined by the union over the rate pairs satisfying

$$R_1 \leq I(U, W; Y_1)$$

$$R_2 \leq I(V, W; Y_2)$$

$$R_1 + R_2 \leq \min\{I(U, W; Y_1) + I(V; Y_2|U, W), I(V, W; Y_2) + I(U; Y_1|V, W)\}$$

over all $p(u)p(v)p(w, x|u, v)$ such that $(U, V, W) \rightarrow X \rightarrow (Y_1, Y_2)$ form a Markov chain forms 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:

Bound 3: [Nair-El Gamal '07] The region \mathcal{R}_1 defined by 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 $(U, V) \rightarrow X \rightarrow (Y_1, Y_2)$ 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 Bounds 2 and 3 are identical.*

The organization of the paper is as follows. In Section II we show that the regions described by Bound 2 and Bound 3

are the same. We also present a different representation of the bound which allows us to have bounds on the cardinalities of the auxiliary random variables. 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 OUTER BOUND

A. Identity of the bounds

Theorem 1: 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 proof¹ is as follows: Given a (U, V) we will produce a (U^*, V^*, W^*) with U^*, V^* being independent such that

$$\begin{aligned} I(U; Y_1) &= I(U^*, W^*; Y_1) \\ I(V; Y_2) &= I(V^*, W^*; Y_2) \\ I(U; Y_1|V) &= I(U^*; Y_1|V^*, W^*) \\ I(V; Y_2|U) &= I(V^*; Y_2|U^*, W^*). \end{aligned} \quad (1)$$

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

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

where $(\cdot)_m$ denotes the mod 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. ■

B. An alternate characterization

We reproduce some of the arguments in [2] to express the Bound 3 in an alternate manner to aid its evaluation.

Lemma 1: The region \mathcal{R}_1 is equivalent to the following region, \mathcal{R}_2 , defined by the union of rate pairs satisfying

$$\begin{aligned} R_1 &\leq I(U; Y_1) \\ R_2 &\leq I(V; Y_2) \\ R_1 + R_2 &\leq \min\{I(U; Y_1) + I(X; Y_2|U), \\ &\quad I(V; Y_2) + I(X; Y_1|V)\} \end{aligned}$$

over all $p(u, v, x)$ such that $(U, V) \rightarrow X \rightarrow (Y_1, Y_2)$

Proof: Since $(U, V) \rightarrow X \rightarrow (Y_1, Y_2)$ form a Markov chain, we have $I(U; Y_1|V) \leq I(X; Y_1|V)$ and $I(V; Y_2|U) \leq I(X; Y_2|U)$. Therefore it is clear that $\mathcal{R}_1 \subseteq \mathcal{R}_2$. Hence it suffices to show that $\mathcal{R}_2 \subseteq \mathcal{R}_1$.

Let l denote the size of \mathcal{X} . Given (U, V, X) it was shown in [2] that for the following triple (U^*, V^*, X) , having cardinalities $l\|\mathcal{U}\|, l\|\mathcal{V}\|, l$ respectively, defined according to

$$\begin{aligned} P(U^* = u_i, V^* = v_j) \\ = \frac{1}{l} P(U = u, V = v, X = (i-j)_l), \\ P(X^* = k|U^* = u_i, V^* = v_j) \\ = \begin{cases} 1 & \text{if } k = (i-j)_l \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (2)$$

one obtains

$$\begin{aligned} I(U; Y_1) &= I(U^*; Y_1) \\ I(V; Y_2) &= I(V^*; Y_2) \\ I(X; Y_1|V) &= I(X; Y_1|V^*) = I(U^*; Y_1|V^*) \\ I(X; Y_2|U) &= I(X; Y_2|U^*) = I(V^*; Y_2|U^*). \end{aligned} \quad (3)$$

Thus $\mathcal{R}_2 \subseteq \mathcal{R}_1$. ■

C. Cardinality bounds

Using the strengthened Carathéodory theorem by Fenchel and Eggleston [6] it can be readily shown that for any choice of the auxiliary random variable U , there exists a random variable U_1 with cardinality bounded by $\|\mathcal{X}\| + 1$ such that $I(U; Y_1) = I(U_1; Y_1)$, $I(X; Y_2|U) = I(X; Y_2|U_1)$ and preserves the distribution $p(X)$. Similarly one can find a V_1 with cardinality bounded by $\|\mathcal{X}\| + 1$ such that $I(V; Y_2) = I(V_1; Y_2)$, $I(X; Y_1|V) = I(X; Y_1|V_1)$ and preserves the distribution $p(X)$. Since both U_1 and V_1 share the same distribution $p(X)$ one can create a triple (U_1, V_1, X) (for e.g. by generating U_1 and V_1 conditionally independent of X). Thus one can assume without loss of generality that the cardinalities of U, V in Lemma 1 are bounded by $\|\mathcal{X}\| + 1$ each.

D. An outer bound formulation that can be evaluated

Putting all of these together we have the following characterization of the Bound 2.

Bound 4: The region \mathcal{R} consists of the union of rate pairs satisfying

$$\begin{aligned} R_1 &\leq I(U; Y_1) \\ R_2 &\leq I(V; Y_2) \\ R_1 + R_2 &\leq \min\{I(U; Y_1) + I(X; Y_2|U), \\ &\quad I(V; Y_2) + I(X; Y_1|V)\} \end{aligned}$$

over all $p(u, v, x)$ such that $(U, V) \rightarrow X \rightarrow (Y_1, Y_2)$ and constitutes the Bound 2. Further, one can assume that $\|\mathcal{U}\|, \|\mathcal{V}\| \leq \|\mathcal{X}\| + 1$.

Alternately one can also use construction (2) to restrict X to be a deterministic function of U, V while relaxing the cardinalities to $\|\mathcal{U}\|, \|\mathcal{V}\| \leq \|\mathcal{X}\| (\|\mathcal{X}\| + 1)$.

¹The idea of the construction is motivated in part by a similar construction (2) originally appearing in [2] and also from a conversation with Prof. Hajek about the tightness of Bound 2 for the deterministic broadcast channel.

III. THE BINARY SKEW-SYMMETRIC CHANNEL

A. On evaluating Märtton inner bound

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

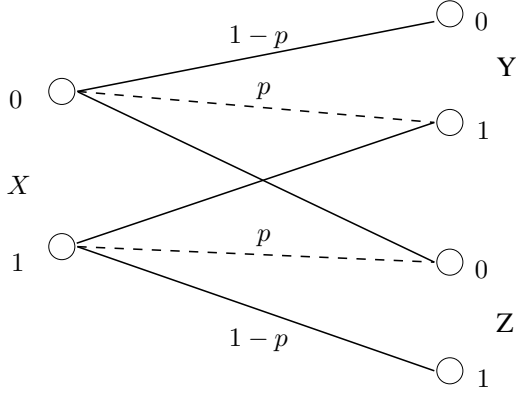


Fig. 1. Binary Skew Symmetric Channel

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 Märtton [1] was not tight and that the region prescribed by Theorem 2 was strictly contained inside the Korner-Märtton region.

Backed by numerical simulations we make the following conjecture about the BSSC with $p = \frac{1}{2}$.

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: It is easy to see that 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 $P(X = 0) = 1 - \eta$ and hence it suffices to consider η in the range $0 \leq \eta \leq \frac{1}{2}$.

Observe that

$$\begin{aligned} & I(U; Y_1) + I(V; Y_2) - I(U; V), \\ & \leq I(V; Y_2) + I(U; Y_1, V) - I(U; V), \\ & = I(V; Y_2) + I(U; Y_1|V), \\ & \leq I(V; Y_2) + I(X; Y_1|V), \\ & = I(X; Y_1) + I(V; Y_2) - I(V; Y_1). \end{aligned}$$

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

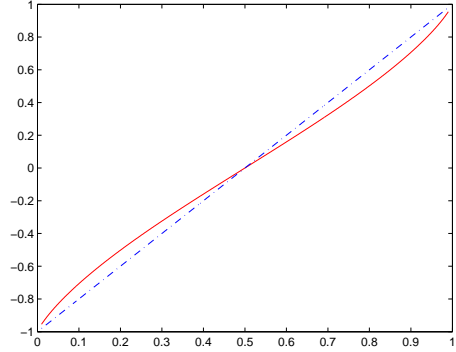


Fig. 2. The plot of the function $f(\eta) = H(\frac{\eta}{2}) - H(\frac{1-\eta}{2})$.

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 V that maximizes $I(V; Y_2) - I(V; Y_1)$ subject to $V \rightarrow X \rightarrow (Y_1, Y_2)$ being Markov and $P(X = 0) = \eta$. Then it is not difficult to see that the optimal choice would be to set $V = \Phi$ (the trivial random variable) for all $\eta \leq \eta_0 = \frac{1}{5}$ where η_0 is the unique point in $[0, \frac{1}{2}]$ at which the line joining $(\eta_0, f(\eta_0))$ to the point $(1, 1)$ is a tangent to the curve $f(\eta)$.

Lemma 2: Let $P(X = 0) = \eta \leq \eta_0$ where $\eta_0 = \frac{1}{5}$ 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 $f(\eta)$.

Then for all $V \rightarrow X \rightarrow (Y_1, Y_2)$ we have $I(V; Y_2) \leq I(V; Y_1)$.

Proof: Define $g(\eta)$ as follows:

$$g(\eta) = \begin{cases} f(\eta) & 0 \leq \eta \leq \eta_0 \\ \frac{1-\eta}{1-\eta_0} f(\eta_0) + \frac{\eta-\eta_0}{1-\eta_0} f(1) & \eta_0 \leq \eta \leq 1 \end{cases}.$$

Observe that $g(\eta)$ is concave and that $f(\eta) \leq g(\eta), 0 \leq \eta \leq 1$.

Let $P(V = i) = v_i$ and $P(X = 0|V = i) = \alpha_i$. We have $\sum_i v_i \alpha_i = \eta$. Observe that we have the following,

$$\begin{aligned} & I(V; Y_2) - I(V; Y_1) \\ & = H(Y_1|V) - H(Y_2|V) - (H(Y_1) - H(Y_2)) \\ & = \sum_i v_i f(\alpha_i) - f(\eta) \\ & \leq \sum_i v_i g(\alpha_i) - f(\eta) \\ & \stackrel{(a)}{\leq} g(\sum_i v_i \alpha_i) - f(\eta) \\ & = g(\eta) - f(\eta) = 0 \text{ as } 0 \leq \eta \leq \eta_0. \end{aligned}$$

Here (a) follows from the concavity of $g(\eta)$. This completes the proof of Lemma 2. \blacksquare

This implies that for $\eta \leq \eta_0 = \frac{1}{5}$, we have

$$\begin{aligned} & I(U; Y_1) + I(V; Y_2) - I(U; V), \\ & \leq I(X; Y_1) + I(V; Y_2) - I(V; Y_1), \\ & \leq I(X; Y_1) + 0, \\ & = I(X; Y_1). \end{aligned}$$

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 3: Conjecture 1 is true as long as $\max\{P(X = 0), P(X = 1)\} \leq \eta_0 = \frac{1}{5}$.

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

$$\begin{aligned} R_1 + R_2 & \leq \min\{I(W; Y_1), I(W; Y_2)\} \\ & \quad + I(U; Y_1|W) + I(V; Y_2|W) - I(U; V|W). \end{aligned}$$

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

$$\begin{aligned} R_1 + R_2 & \leq \min\{I(W, T; Y_1), I(W, T; Y_2)\} \\ & \quad + I(U; Y_1|W, T) + I(V; Y_2|W, T) \\ & \quad - I(U; V|W, T) \\ & \stackrel{(a)}{\leq} \min\{I(W, T; Y_1), I(W, T; Y_2)\} \\ & \quad + P(T = 0)I(X; Y_1|W, T = 0) \\ & \quad + P(T = 1)I(X; Y_2|W, T = 1) \\ & \stackrel{(b)}{\leq} \min\{I(T; Y_1), I(T; Y_2)\} \\ & \quad + P(T = 0)I(X; Y_1|T = 0) \\ & \quad + P(T = 1)I(X; Y_2|T = 1). \end{aligned}$$

Here (a) follows from Conjecture 1 and (b) follows from the fact that

$$\begin{aligned} P(T = 1)I(W; Y_1|T = 1) & \leq P(T = 1)I(W; Y_2|T = 1), \\ P(T = 0)I(W; Y_2|T = 0) & \leq P(T = 0)I(W; Y_1|T = 0). \end{aligned}$$

The above fact can be easily established by considering the convexity and concavity of $f(\eta)$ (see Figure 2) in the intervals $\eta \in [\frac{1}{2}, 1]$ and $\eta \in [0, \frac{1}{2}]$, respectively.

In [2] the bound on sum rate, $\min\{I(T; Y_1), I(T; Y_2)\} + P(T = 0)I(X; Y_1|T = 0) + P(T = 1)I(X; Y_2|T = 1)$ has been studied and the maximum was evaluated as ≈ 0.3616 . This could also be inferred from [5] and the evaluation of the Cover-van-der-Meulen region for this channel.

Thus assuming Conjecture 1 we have that the sum rate of the Marton inner bound is bounded by 0.3616... (correct to 4 decimal places).

B. Evaluating outer bound - BSSC

In [2] the sum rate of the pairs (R_1, R_2) described by Bound 4 was evaluated and it was shown that the maximum sum rate was bounded by 0.3711.. (correct to 4 decimal places). *Thus we have that the region described by Bound 2 is strictly larger than that described by Bound 1 (assuming Conjecture 1) and thus 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. 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

- [1] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Info. Theory*, vol. IT-25, pp. 306–311, May, 1979.
- [2] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Info. Theory*, vol. IT-53, pp. 350–355, January, 2007.
- [3] T. Cover, "Broadcast channels," *IEEE Trans. Info. Theory*, vol. IT-18, pp. 2–14, January, 1972.
- [4] —, "Comments on broadcast channels," *IEEE Trans. Info. Theory*, vol. IT-44, pp. 2524–2530, October, 1998.
- [5] B. Hajek and M. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Trans. Info. Theory*, vol. IT-25, pp. 36–46, January, 1979.
- [6] I. Csizár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Info. Theory*, vol. IT-24, pp. 339–348, May, 1978.