An information inequality for the BSSC channel

Varun Jog   Chandra Nair

December 11, 2009

Abstract

We establish an information theoretic inequality concerning the binary skew-symmetric broadcast channel that was conjectured by one of the authors. This inequality helps to quantify the gap between the sum rate obtained by the inner bound and outer bound for the binary skew-symmetric broadcast channel.

1   Introduction

The broadcast channel is a fundamental network information theory setting modeling the communication of messages (private and common) from a single sender to multiple receivers. For formal definitions and early prior work the reader is referred to [1, 2]. There has been some recent progress for the discrete memoryless setting, and this work establishes a conjecture proposed in one of the recent papers [7].

We consider the broadcast channel where sender X wishes to communicate independent messages M_1, M_2 to two receivers Y_1, Y_2. The capacity region for the broadcast channel is an open problem and the best known achievable region is due to Marton[5] and is presented below.

Bound 1. *[5] The following region is achievable*

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

for any triple of random variables $p(u, v, w)$ such that $(U, V, W) \to X \to (Y_1, Y_2)$ form a Markov chain.

Capacity regions have been established for a number of special cases and in *every case* where capacity is known, the following outer bound and Marton's inner bound yields the same region.

Bound 2. *[6] The union of rate pairs*

$$R_1 \leq I(U; Y_1)$$
$$R_2 \leq I(V; Y_2)$$
$$R_1 + R_2 \leq I(U; Y_1) + I(V; Y_2|U)$$
$$R_1 + R_2 \leq I(V; Y_2) + I(U; Y_1|V)$$

over pairs of random variables $p(u, v)$ such that $(U, V) \to X \to (Y_1, Y_2)$ form a Markov chain constitutes an outer bound to the capacity region.
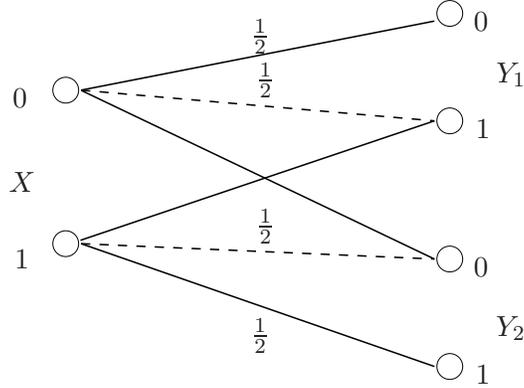
Figure 1: binary skew-symmetric broadcast channel

In [6] the authors studied Bound 2 for the binary skew symmetric channel and showed that the line segment[1] of $R_1 + R_2 = 0.3725..$ lies on the boundary of the outer bound. In [7] the authors studied Marton's inner bound for the binary skew-symmetric broadcast channel and showed that provided an information theoretic inequality (Conjecture 1) holds, a line segment of $R_1 + R_2 = 0.3616...$ lies on the boundary of the Marton's inner bound.

Conjecture 1. *[7] For the binary skew-symmetric channel shown in Figure 1,*

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

for all (U,V,X) such that $(U,V) - X - (Y_1,Y_2)$ forms a Markov chain.

It should be noted that this inequality was established in [4] when U,V were independent; and in [7] for dependent U,V and $P(X = 0) \in [0, \frac{1}{5}] \cup [\frac{4}{5}, 1]$.

The outline of the proof is as follows: (Parts 1 and 2 were established in [3] and is presented for completeness)

1. We show that to obtain the maximum sum-rate in Marton's region it is sufficient to consider X as a function of the auxiliary random variables U and V.

2. We further show that the cardinality of U and V can be restricted to $|\mathcal{X}|$, which in this case is 2.

3. As U and V are binary, we conclude that the conjecture is true iff it holds true for each of the 16 cases where $X = f(U, V)$. We further prove that it is sufficient to consider only the two cases $X = U \wedge V$ and $X = U \oplus V$ by showing that either the other cases are immediate or it reduces to the above two cases.

4. We prove the conjecture in $X = U \wedge V$ case using a multiplicative perturbation to derive properties about the distribution on U, V which achieves the maxima in LHS of 1.

5. Similarly, we prove the conjecture in $X = U \oplus V$ case using an additive perturbation to derive properties about the distribution on U, V which achieves the maxima in LHS of 1.

---

[1]There is a typo in the evaluation of this bound in the original paper though the main results are right. In the Appendix we will show the corrected derivation of this bound.

2   Main

Gohari and Anantharam [3] established bounds on the cardinalities of the auxiliary random variables needed to evaluate the Marton's achievable region. In this section, we present a modified version of their arguments for completeness.

Define the following three quantities:

- $M = \max(I(U;Y_1) + I(V;Y_2) - I(U;V))$ over all (U,V,X) such that $(U,V) - X - (Y_1,Y_2)$ forms a Markov chain.

- $M_d = \max(I(U;Y_1) + I(V;Y_2) - I(U;V))$ over all (U,V,X) such that $(U,V) - X - (Y_1,Y_2)$ forms a Markov chain and $X = f(U,V)$.

- $M_d^{(|\mathcal{X}|)} = \max(I(U;Y_1)+I(V;Y_2)-I(U;V))$ over all (U,V,X) such that $(U,V)-X-(Y_1,Y_2)$ forms a Markov chain; $|\mathcal{U}| \leq |\mathcal{X}|$, $|\mathcal{V}| \leq |\mathcal{X}|$, and $X = f(U,V)$.

We will show that $M = M_d = M_d^{(|\mathcal{X}|)}$ for any discrete-memoryless broadcast channel.

Fact 1. $M = M_d$

Proof. Using standard arguments, for e.g. [4], there exists a random variable W independent of U,V such that $X = f(U,V,W)$. Now set $V' = (V,W)$ and observe that $I(U;Y_1) + I(V;Y_2) - I(U;V) \leq I(U;Y_1) + I(V';Y_2) - I(U;V')$. □

Remark 1. One way to construct such a W is the following[2]: For every u,v consider the sequence $t_i(u,v) = \mathrm{P}(X \leq i|U = u, V = v), 1 \leq i \leq |\mathcal{X}|$. Mark the points $t_1(u,v),\ldots,t_{|\mathcal{X}|}(u,v)(= 1)$ for all choices of (u,v) along the unit interval $[0,1]$. The points define intervals (at most $|\mathcal{U}||\mathcal{V}||\mathcal{X}|$) and generate W as an independent random variable with probabilities defined by the length of the intervals. As the $\mathrm{P}(X = i|U = u, V = v)$ can be thought of as W falling in a certain consecutive set of appropriately chosen intervals, there is a natural mapping $(U,V,W) \mapsto X$.

Claim 1. $M_d = M_d^{(|\mathcal{X}|)}$

Proof. This is a simplified version of the arguments of Gohari and Anantharam [3], adapted to this setting. For a given $p(u,v,x)$ consider the multiplicative Lyapunov perturbation defined by $q(u,v,x) = p(u,v,x)(1 + \epsilon L(u))$. For $q(u,v,x)$ to be a valid probability distribution we require the following two conditions: $1 + \epsilon L(u) \geq 0, \forall u$ and $\sum_u p(u)L(u) = 0$. *Note:* If $p(u,v,x) = 0$ then $q(u,v,x) = 0$ and hence X continues to be a function of (U,V) under any such perturbation.

If distribution $p(u,v,x)$ maximizes $I(U;Y_1) + I(V;Y_2) - I(U;V)$ then we must have that for any valid perturbation

1. $\frac{\partial}{\partial \epsilon}I(U;Y_1) + I(V;Y_2) - I(U;V) = 0$,

2. $\frac{\partial^2}{\partial \epsilon^2}I(U;Y_1) + I(V;Y_2) - I(U;V) \leq 0$.

Consider a class of perturbations $L(u)$ such that

$$\mathrm{E}(L|X = x) = \sum_{u,v} p(u,v|x)L(u) = 0, \forall x \in \mathcal{X}. \tag{1}$$

Observe that these perturbations keep the distributions of $\mathcal{X}$ (hence Y_1, Y_2) unchanged.

---

[2]This construction was mentioned to one of the authors by Bruce Hajek.

Observation 1. There exists such a non-zero perturbation if $|\mathcal{U}| > |\mathcal{X}|$ since the null-space of the constraints have rank at most $|\mathcal{X}|$.

Observe that

$$
\begin{aligned}
&I_q(U;Y_1) + I_q(V;Y_2) - I_q(U;V) \\
&\quad = H_q(Y_1) + H_q(Y_2) + H_q(U,V) - H_q(U,Y_1) - H_q(V,Y_2) \\
&\quad = H_p(Y_1) + H_p(Y_2) + H_q(U,V) - H_q(U,Y_1) - H_q(V,Y_2) \\
&\quad = H_p(Y_1) + H_p(Y_2) + H_p(U,V) + \epsilon H_p^L(U,V) - H_p(U,Y_1) - \epsilon H_p^L(U,Y_1) - H_q(V,Y_2).
\end{aligned}
$$

Here $H_p^L(U,V) = -\sum_{u,v} p(u,v)L(u)\log p(u,v)$, $H_p^L(U,Y_1) = -\sum_{u,y_1} p(u,y_1)L(u)\log p(u,y_1)$.

Therefore, $\frac{\partial^2}{\partial \epsilon^2} I(U;Y_1) + I(V;Y_2) - I(U;V) \leq 0$ implies $\frac{\partial^2}{\partial \epsilon^2} H_q(V,Y_2) \geq 0$ and this implies

$$
E(E(L|V,Y_2)^2) \leq 0
$$

or in particular $E(L|V,Y_2) = 0$ whenever $p(v,y_2) \neq 0$. This, in turn, implies

$$
H_q(V,Y_2) = H_p(V,Y_2).
$$

Using this we obtain

$$
\begin{aligned}
&I_q(U;Y_1) + I_q(V;Y_2) - I_q(U;V) \\
&\quad = H_p(Y_1) + H_p(Y_2) + H_p(U,V) + \epsilon H_p^L(U,V) - H_p(U,Y_1) - \epsilon H_p^L(U,Y_1) - H_q(V,Y_2) \\
&\quad = H_p(Y_1) + H_p(Y_2) + H_p(U,V) + \epsilon H_p^L(U,V) - H_p(U,Y_1) - \epsilon H_p^L(U,Y_1) - H_p(V,Y_2).
\end{aligned}
$$

The first derivative being zero implies $H_p^L(U,V) - H_p^L(U,Y_1) = 0$ and this finally implies that if $p(u,v,x)$ attains the maximum of $I(U;Y_1) + I(V;Y_2) - I(U;V)$ then $I_q(U;Y_1) + I_q(V;Y_2) - I_q(U;V) = I_p(U;Y_1) + I_p(V;Y_2) - I_p(U;V)$ for any valid perturbation that satisfies (1).

Now we choose ϵ such that $\min_u 1 + \epsilon L(u) = 0$, and let $u = u^*$ achieve this minimum. Observe that $q(u^*) = 0$ and hence there exists an U with cardinality equal to $|\mathcal{U}| - 1$ (at most) such that $I(U;Y_1) + I(V;Y_2) - I(U;V)$ is constant. We can proceed by induction until $|\mathcal{U}| = |\mathcal{X}|$. Observe that when $|\mathcal{U}| = |\mathcal{X}|$, we are no longer guaranteed the existence of a non-trivial $L(u)$ satisfying (1).

The argument can then be repeated for V to make $|\mathcal{V}| \leq |\mathcal{X}|$ as well.

This completes the proof that $M_d = M_d^{(|\mathcal{X}|)}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 2. Use Fact 1 and Claim 1, to prove the conjecture 1 it suffices to consider binary U, V and $X = f(U,V)$. There are 16 possible boolean functions on binary (U,V) and we establish the conjecture for each such function.

We use the following notation: $U \wedge V$ (and), $U \vee V$ (or), $U \oplus V$ (xor), $\bar{U}$ (not).

Observation 2. Each of the following groups of functions are equivalent upto re-labeling (of either U or V or both)

- $X = U, X = \bar{U}$,

- $X = V, X = \bar{V}$,

- $X = U \wedge V, X = \bar{U} \wedge V, X = U \wedge \bar{V}, X = \bar{U} \wedge \bar{V}$,

- $X = U \vee V, X = \bar{U} \vee V, X = U \vee \bar{V}, X = \bar{U} \vee \bar{V}$,

- $X = U \oplus V$, $X = \bar{U} \oplus V$

Claim 2. *The conjecture is valid when $X = 0, X = 1, X = U, X = V$.*

Proof. In the first two cases, the conjecture reduces to $-I(U; V) \leq 0$ (true by non-negativity of mutual information). In the third case conjecture follows from data processing inequality as $I(V; Y_2) \leq I(V; X)$ and hence $I(X; Y_1) + I(V; Y_2) - I(V; X) \leq I(X; Y_1)$. The fourth case follows in a similar manner as the third. $\qquad\square$

Claim 3. *The conjecture is valid for all distributions $p(u, v)$ when $X = U \wedge V$, if and only if the conjecture is valid for all distributions $q(u, v)$ when $X = U \vee V$.*

Proof. This follows from the skew-symmetry of the channel and that $X = U \vee V$ is equivalent to $\bar{X} = \bar{U} \wedge \bar{V}$. Let $P(U = i, V = j) = p_{ij}$ for every $i, j \in \{0, 1\}$. For concreteness, when $X = U \wedge V$ the conjecture is equivalent to

$$h(\frac{p_{00} + p_{01} + p_{10}}{2}) - (p_{00} + p_{01})h(\frac{1}{2}) - (p_{10} + p_{11})h(\frac{p_{10}}{2(p_{10} + p_{11})}) + h(\frac{p_{11}}{2})$$

$$- (p_{01} + p_{11})h(\frac{p_{11}}{2(p_{01} + p_{11})}) - h(p_{00} + p_{01}) + (p_{00} + p_{10})h(\frac{p_{00}}{p_{00} + p_{10}}) + (p_{01} + p_{11})h(\frac{p_{01}}{p_{01} + p_{11}})$$

$$\leq \max\left\{h(\frac{p_{00} + p_{01} + p_{10}}{2}) - (p_{00} + p_{01} + p_{10})h(\frac{1}{2}), h(\frac{p_{11}}{2}) - p_{11}h(\frac{1}{2})\right\}, \tag{2}$$

where $h(x) = -x \log_2(x) - (1 - x) \log_2(1 - x)$ represents the binary entropy function.

For the $X = U \vee V$ case, let $P(U = i, V = j) = q_{ij}$ for every $i, j \in 0, 1$. The conjecture is now equivalent to

$$h(\frac{q_{00}}{2}) - (q_{01} + q_{00})h(\frac{q_{00}}{2(q_{01} + q_{00})}) + h(\frac{q_{11} + q_{01} + q_{10}}{2}) - (q_{11} + q_{01})h(\frac{1}{2})$$

$$- (q_{10} + q_{00})h(\frac{q_{10}}{2(q_{10} + q_{00})}) - h(q_{11} + q_{01}) + (q_{11} + q_{10})h(\frac{q_{11}}{q_{11} + q_{10}}) + (q_{01} + q_{00})h(\frac{q_{01}}{q_{01} + q_{00}})$$

$$\leq \max\left\{h(\frac{q_{11} + q_{01} + q_{10}}{2}) - (q_{11} + q_{01} + q_{10})h(\frac{1}{2}), h(\frac{q_{00}}{2}) - q_{00}h(\frac{1}{2})\right\}, \tag{3}$$

The bijection $p_{00} \leftrightarrow q_{11}, p_{01} \leftrightarrow q_{01}, p_{10} \leftrightarrow q_{10}, p_{11} \leftrightarrow q_{00}$ completes the proof of the equivalence of the conjectures under the constraints $X = U \wedge V$ and $X = U \vee V$. $\qquad\square$

Corollary 1. *From Remark 2, Observation 2, Claims 2, and 3 it follows that the conjecture is true provided it holds when $X = U \wedge V$ and $X = U \oplus V$.*

2.1   Case 1: X= $U \wedge V$

We prove the conjecture in this case by studying the local maxima. Clearly the conjecture is true when two of the three terms p_{00}, p_{01}, p_{10} are identically 0. When this happens, then the condition reduces to $X = U, V = 1$, $X = V, U = 1$, or $U = V = X$, each of which is solved by Claim 2. Clearly if $p_{11} = 0$ then $X = 0$; in which case the conjecture is valid. So we assume that $p_{11} > 0$. Therefore, we only establish the validity of the conjecture for the remaining cases.

Consider a perturbation $q(u, v, x) = p(u, v, x)(1 + \epsilon L(u, v))$ that maintains $\mathrm{P}(X = 0)$. This implies that the perturbation satisfies

$$L_{11} = 0, \quad p_{00}L_{00} + p_{01}L_{01} + p_{10}L_{10} = 0. \tag{4}$$

For any local maxima of $I(U; Y_1) + I(V; Y_2) - I(U; V)$, the derivative with respect to ϵ must be zero for all perturbations satisfying (4), i.e.

$$H_L(U, V) = H_{E(L|U,Y_1)}(U, Y_1) + H_{E(L|V,Y_2)}(V, Y_2). \tag{5}$$

The terms $H_L(U, V), H_{E(L|U,Y_1)}(U, Y_1), H_{E(L|V,Y_2)}(V, Y_2)$ correspond to

$$H_L(U, V) = -p_{00}L_{00} \log p_{00} - p_{10}L_{10} \log p_{10} - p_{01}L_{01} \log p_{01},$$

$$H_{E(L|U,Y_1)}(U, Y_1) = -(p_{00}L_{00} + p_{01}L_{01}) \log(\frac{p_{00} + p_{01}}{2}) - \frac{p_{10}L_{10}}{2} \log \frac{p_{10}}{2}$$

$$- \frac{p_{10}L_{10}}{2} \log(\frac{p_{10}}{2} + p_{11}),$$

$$H_{E(L|V,Y_2)}(V, Y_2) = -(p_{00}L_{00} + p_{10}L_{10}) \log(p_{00} + p_{10}) - p_{01}L_{01} \log(p_{01} + \frac{p_{11}}{2}).$$

2.1.1   Case 1.1 $p_{00}, p_{01}, p_{10}, p_{11} > 0$

In this case the conditions (4) and (5) imply that the following equalities hold:

$$\frac{p_{00}}{p_{01}} = \frac{p_{00} + p_{10}}{p_{01} + \frac{p_{11}}{2}},$$

$$\frac{p_{00}}{p_{10}} = \frac{p_{00} + p_{01}}{\sqrt{p_{10}(p_{10} + 2p_{11})}}.$$

These conditions are obtained by setting $L_{10} = 0$ and $L_{01} = 0$ respectively.

The above two conditions imply that

$$\frac{p_{01}}{p_{00}} = 2, \frac{p_{11}}{p_{10}} = 4.$$

These two equalities along with $p_{00} + p_{01} + p_{10} + p_{11} = 1$ implies that any non-trivial local maxima is of the form[3]

$$p_{00} = \frac{1 - t}{3}, p_{01} = \frac{2(1 - t)}{3}, p_{10} = \frac{t}{5}, p_{11} = \frac{4t}{5}.$$

We need to verify the conjecture at this point. It suffices to show that

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

This is equivalent to showing (for $0 \leq t \leq 1$)

$$0 \leq H(Y_1|U) - H(Y_1|X) - H(V|U) + H(V|Y_2)$$

$$= (1 - t) + th(\frac{1}{10}) - (1 - \frac{4t}{5}) - (1 - t)h(\frac{1}{3}) - th(\frac{1}{5}) + (1 - \frac{2t}{5})h(\frac{1}{3})$$

$$= t\left(\frac{4}{5} - 1 + h(\frac{1}{10}) + h(\frac{1}{3}) - h(\frac{1}{5}) - \frac{2}{5}h(\frac{1}{3})\right)$$

$$= \frac{3t}{5}\left(h(\frac{1}{3}) - \frac{3}{2}h(\frac{1}{9})\right),$$

and this is clearly true as $\frac{3}{2}h(\frac{1}{9}) \leq h(\frac{1}{3})$. This proves the validity of the conjecture when $p_{00}, p_{01}, p_{10} > 0$.

---

[3]This local maxima exists only when $P(X = 1) = p_{11} \leq \frac{4}{5}$, and hence there is no local maxima when $P(X = 1) > \frac{4}{5}$. When $P(X = 1) \geq \frac{4}{5}$, there was a simple argument in [7] that established the conjecture. It is curious that both the approaches lead to a simple proof in this regime.

2.1.2   Case 1.2 $p_{01} = 0; p_{00}, p_{10}, p_{11} > 0$

In this case the conditions (4) and (5) imply that the following equality holds:

$$\frac{p_{00}}{p_{10}} = \frac{p_{00}}{\sqrt{p_{10}(p_{10} + 2p_{11})}}.$$

However this cannot hold if $p_{00}, p_{10}, p_{11} > 0$.

2.1.3   Case 1.3 $p_{10} = 0; p_{00}, p_{01}, p_{11} > 0$

In this case the conditions (4) and (5) implies that

$$\frac{p_{00}}{p_{01}} = \frac{p_{00}}{p_{01} + \frac{p_{11}}{2}}.$$

Again this cannot hold if $p_{00}, p_{10}, p_{11} > 0$.

2.1.4   Case 1.4 $p_{00} = 0; p_{10}, p_{01}, p_{11} > 0$

In this case the conditions (4) and (5) implies

$$\sqrt{p_{10}(p_{10} + 2p_{11})} = p_{01} + \frac{p_{11}}{2}. \tag{6}$$

To eliminate this possibility, we show that any point that satisfies (6) cannot be a local maxima. Observe that for a local maxima one also requires $\frac{\partial^2}{\partial \epsilon^2} I(U; Y_1) + I(V; Y_2) - I(U; V) \leq 0$, i.e.

$$E[E[L_{UV}|UY_1]^2] + E[E[L_{UV}|VY_2]^2] - E[E[L_{UV}|UV]^2] \leq 0.$$

Equivalently for all perturbations satisfying $L_{11} = 0$ and $p_{01}L_{01} + p_{10}L_{10} = 0$, any local maxima must satisfy

$$p_{01}L_{01}^2 + p_{10}L_{10}^2 \geq p_{01}L_{01}^2 + \frac{1}{2}p_{10}L_{10}^2 + \frac{1}{2}\frac{p_{10}^2}{p_{10} + 2p_{11}}L_{10}^2$$

$$+ p_{10}L_{10}^2 + \frac{p_{01}^2}{p_{01} + \frac{p_{11}}{2}}L_{01}^2,$$

which is clearly not possible when $p_{10}, p_{01}, p_{11} > 0$. This completes the proof of Case 1.

2.2   Case 2: $X = U \oplus V$

We again prove the conjecture in this case by studying the local maxima. As before, the conjecture is true when two of the four terms $p_{00}, p_{01}, p_{10}, p_{11}$ are identically 0. When this happens, then the condition reduces to $X = \bar{U}, V = 1$, $X = \bar{V}, U = 1$, $X = 0$, $X = 1$, $X = U, V = 0$ or $X = V, U = 0$, each of which is solved by Claim 2. Therefore, we only establish the validity of the conjecture for the remaining cases.

Consider a perturbation[4] $q(u, v, x) = p(u, v, x) + \epsilon\lambda(u, v, x)$ for some $\epsilon > 0$. For this to be a valid perturbation we require

$$\lambda_{001}, \lambda_{010}, \lambda_{100}, \lambda_{111} \geq 0 \tag{7}$$

---

[4]Note that this perturbation is a more general perturbation that the one we have used so far, the multiplicative perturbation of the form $q(u, v, x) = p(u, v, x)(1 + \epsilon L(u, v, x))$. The multiplication perturbation ensures that if $p(u, v, x) = 0$ then $q(u, v, x) = 0$; however an additive one need not preserve this. Setting $\lambda(u, v, x) = p(u, v, x)L(u, v, x)$ shows that the multiplicative perturbation is a special case of the additive perturbation. It turns out that in the case $X = U \oplus V$, the analysis of the local maxima is greatly simplified if we consider an additive perturbation; as we are finding the local maxima over a possibly larger space.

as the corresponding $p(u, v, x)$ are zero. Further let us require that the perturbation maintains $P(X = 0)$. This implies that the perturbation satisfies

$$\lambda_{000} + \lambda_{010} + \lambda_{100} + \lambda_{110} = 0$$
$$\lambda_{001} + \lambda_{011} + \lambda_{101} + \lambda_{111} = 0 \tag{8}$$

For any perturbation that satisfies (15) and (8) at any local maximum it must be true that the first derivative cannot be positive. This implies

$$H_\lambda(U, V) - H_{E(\lambda|U,Y_1)}(U, Y_1) - H_{E(\lambda|V,Y_2)}(V, Y_2) \leq 0, \tag{9}$$

where

$$H_\lambda(U, V) = -(\lambda_{001} + \lambda_{000}) \log p_{00} - (\lambda_{010} + \lambda_{011}) \log p_{01} - (\lambda_{100} + \lambda_{101}) \log p_{10}$$
$$- (\lambda_{110} + \lambda_{111}) \log p_{11}$$

$$H_{E(\lambda|U,Y_1)}(U, Y_1) = -\frac{(\lambda_{000} + \lambda_{010})}{2} \log(\frac{p_{00}}{2}) - \left(\frac{(\lambda_{000} + \lambda_{010})}{2} + \lambda_{001} + \lambda_{011} \right) \log(\frac{p_{00}}{2} + p_{01})$$
$$- \frac{(\lambda_{110} + \lambda_{100})}{2} \log(\frac{p_{11}}{2}) - \left(\frac{(\lambda_{110} + \lambda_{100})}{2} + \lambda_{101} + \lambda_{111} \right) \log(\frac{p_{11}}{2} + p_{10})$$

$$H_{E(\lambda|V,Y_2)}(V, Y_2) = - \left(\frac{(\lambda_{001} + \lambda_{101})}{2} + \lambda_{100} + \lambda_{000} \right) \log(\frac{p_{10}}{2} + p_{00}) - \frac{(\lambda_{001} + \lambda_{101})}{2} \log(\frac{p_{10}}{2})$$
$$- \left(\frac{(\lambda_{011} + \lambda_{111})}{2} + \lambda_{010} + \lambda_{110} \right) \log(\frac{p_{01}}{2} + p_{11}) - \frac{(\lambda_{011} + \lambda_{111})}{2} \log(\frac{p_{01}}{2}).$$

2.2.1   Case 2.1 $p_{00}, p_{01}, p_{10}, p_{11} > 0$

Let $a, b, c, d \geq 0$ and let us choose $\lambda_{001} = a = -\lambda_{000}$, $\lambda_{100} = b = -\lambda_{101}$, $\lambda_{010} = c = -\lambda_{011}$, and $\lambda_{111} = d = -\lambda_{110}$. Observe that this choice satisfies (15) and (8). Therefore from the constraint (9), we must have for all choices of $a, b, c, d \geq 0$

$$0 \leq \frac{(a - c)}{2} \log \frac{p_{00}}{p_{00} + 2p_{01}} + \frac{(d - b)}{2} \log \frac{p_{11}}{p_{11} + 2p_{10}}$$
$$+ \frac{(b - a)}{2} \log \frac{p_{10}}{p_{10} + 2p_{00}} + \frac{(c - d)}{2} \log \frac{p_{01}}{p_{01} + 2p_{11}} \tag{10}$$

- Setting $a = c = k$, $b = d = l$ we require

$$\frac{(l - k)}{2} \log \frac{p_{10}(p_{01} + 2p_{11})}{p_{01}(p_{10} + 2p_{00})} \geq 0$$

  for all $l, k \geq 0$ which is true if and only if

$$p_{10}p_{11} = p_{01}p_{00}. \tag{11}$$

- Setting $a = b = l$, $c = d = k$ we require

$$\frac{(l - k)}{2} \log \frac{p_{00}(p_{11} + 2p_{10})}{p_{11}(p_{00} + 2p_{01})} \geq 0$$

  for all $l, k \geq 0$ which is true if and only if

$$p_{01}p_{11} = p_{10}p_{00}. \tag{12}$$

- Setting $a = d = l$, $b = c = k$ we require

$$\frac{(l-k)}{2} \log \frac{p_{00}}{p_{00} + 2p_{01}} \frac{p_{11}}{p_{11} + 2p_{10}} \frac{p_{10} + 2p_{00}}{p_{10}} \frac{p_{01} + 2p_{11}}{p_{01}} \geq 0 \tag{13}$$

for all $l, k \geq 0$. Observe that equations (11) and (12) imply that $p_{00} = p_{11}$ and $p_{01} = p_{10}$. Let $p = p_{00} = p_{11}, q = p_{01} = p_{10}$. Substituting this choice into (13) implies that

$$\frac{(l-k)}{2} \log \left(\frac{p}{p + 2q} \right)^2 \left(\frac{q + 2p}{q} \right)^2 \geq 0$$

for all $l, k \geq 0$ which is true if and only if $p = q$.

Therefore the only choice of $p_{00}, p_{01}, p_{10}, p_{10} > 0$ that satisfies the constraint (10) for all choices of $a, b, c, d \geq 0$ is when $p_{00} = p_{01} = p_{10} = p_{10} = \frac{1}{4}$. In this case observe that U, V and X are mutually independent, and the conjecture is trivially true as $I(U; Y_1) + I(V; Y_2) - I(U; V) = 0$.

2.2.2 Case 2.2 One among $p_{00}, p_{01}, p_{10}, p_{11}$ is zero

All these cases are similar to each other and reduces to a particular $X = U \wedge V$ case, and hence the validity of the conjecture follows.. For example, when $p_{00} = 0$, observe that $X = 0$ *if and only if $U = V = 1$*. Therefore this can also be viewed as a special case of $\bar{X} = U \wedge V$. (Note that we have already shown the equivalence between the $X = U \wedge V$ and $X = U \vee V$ cases.)

Condition	Equivalent $X = U \wedge V$ case
$p_{00} = 0$	$\bar{X} = U \wedge V$
$p_{01} = 0$	$X = U \wedge \bar{V}$
$p_{10} = 0$	$X = \bar{U} \wedge V$
$p_{11} = 0$	$\bar{X} = \bar{U} \wedge \bar{V}$

Since the conjecture was established when $X = U \wedge V$, this equivalence completes the proof when $X = U \oplus V$. Thus Conjecture 1 is established.

3 Sum-rate evaluations of inner and outer bounds for BSSC

We shall evaluate the inner and outer bounds for the BSSC from [6] and [7]. Apart from completeness, this section serves three purposes:

- We present a proof that to compute the maximum sum-rate of the Marton's achievable region it suffices to restrict ourselves to $|W| \leq |X|$

- We correct a minor typo in the evaluation of the maximum sum-rate of the outer bound presented in [6].

- We also compute the maximum sum-rate obtained via the Korner-Marton outer bound for the BSSC.

3.1   On sum-rate evaluation of Marton's inner bound

Though this evaluation was done in [7], assuming the conjectured inequality; we present a slightly different, albeit more general, argument that produces the same result. We first prove that for any broadcast channel it suffices to restrict ourselves to $|W| \leq |X|$ to compute the maximum sum-rate of the Marton's achievable region. In [7] we proved this fact using some properties of the BSSC channel and here we present a general argument.

Claim 4. *For a discrete memoryless broadcast channel, to compute the maximum of*

$$\lambda I(W; Y_1) + (1 - \lambda)I(W; Y_2) + I(U; Y_1|W) + I(V; Y_2|W) - I(U; V|W), 0 \leq \lambda \leq 1$$

over all choices of $(U, V, W) \to X \to (Y_1, Y_2)$ *it suffices to restrict to* $|\mathcal{W}| = |\mathcal{X}|$.

Proof. Let $p(u, v, w, x)$ achieve a maximum of the above expression. As before, we consider multiplicative Lyapunov perturbation defined by $q(u, v, w, x) = p(u, v, w, x)(1 + \varepsilon L(w))$. For $q(u, v, w, x)$ to be a valid probability distribution we require the conditions $1 + \varepsilon L(w) \geq 0$, $\forall w$ and $\sum_w p(w)L(w) = 0$. Further let us require that the perturbation maintains $\mathrm{P}(X = x)$, that is

$$E(L|X = x) = \sum_w p(w|x)L(w) = 0. \tag{14}$$

Remark: There exists nontrivial $L(w)$ if $|\mathcal{W}| > |\mathcal{X}|$.

Observe that

$$\lambda I_q(W; Y) + (1 - \lambda)I_q(W; Z) + I_q(U; Y|W) + I_q(V; Z|W) - I_q(U; V|W) \tag{15}$$
$$= +\lambda H_p(Y) + (1 - \lambda)H_p(Z) + \lambda\big(H_p(W, Z) + \varepsilon H_p^L(W, Z)\big) + (1 - \lambda)\big(H_p(W, Y) + \varepsilon H_p^L(W, Y)\big)$$
$$- H_p(U, W, Y) - \varepsilon H_p^L(U, W, Y) - H_p(V, W, Z) - \varepsilon H_p^L(V, W, Z) + H_p(U, V, W) + \varepsilon H_p^L(U, V, W)$$

where

$$H_p^L(W, Y) = -\sum_{w,y} p(w, y)L(w) \log p(w, y),$$

$$H_p^L(W, Z) = -\sum_{w,z} p(w, z)L(w) \log p(w, z),$$

$$H_p^L(U, V, W) = -\sum_{u,v,w} p(u, v, w)L(w) \log p(u, v, w),$$

$$H_p^L(U, W, Y) = -\sum_{u,w,y} p(u, w, y)L(w) \log p(u, w, y),$$

$$H_p^L(V, W, Z) = -\sum_{v,w,z} p(v, w, z)L(w) \log p(v, w, z).$$

The first derivative with respect to ε being zero implies

$$\lambda H_p^L(W, Z) + (1 - \lambda)H_p^L(W, Y) - H_p^L(U, W, Y) - H_p^L(V, W, Z) + H_p^L(U, V, W) = 0.$$

Substituting this into (15) we obtain

$$\lambda I_q(W; Y) + (1 - \lambda)I_q(W; Z) + I_q(U; Y|W) + I_q(V; Z|W) - I_q(U; V|W)$$
$$= \lambda I_p(W; Y) + (1 - \lambda)I_p(W; Z) + I_p(U; Y|W) + I_p(V; Z|W) - I_p(U; V|W)$$

for any valid perturbation that satisfies (14).

Now we choose ε such that $\min_w 1 + \varepsilon L(w) = 0$, and let $w = w^*$ achieve this minimum. Observe that $q(w^*) = 0$ and hence there exists an W with cardinality equal to $|\mathcal{W}| - 1$ such that $\lambda I(W; Y) + (1 - \lambda)I(W; Z) + I(U; Y|W) + I(V; Z|W) - I(U; V|W)$ is preserved. We can proceed by induction until $|\mathcal{W}| = |\mathcal{X}|$. This completes the proof of this claim. $\square$

3.1.1 Evaluation of the maximum sum-rate of Marton's region for BSSC

Clearly from the above claim we can assume that $|\mathcal{W}| = 2$. Observe that for the BSSC, $I(X; Y_1) \geq I(X; Y_2)$ if and only if $\mathrm{P}(X = 0) \leq \frac{1}{2}$. If $0 \leq \mathrm{P}(X = 0|W = 0), \mathrm{P}(X = 0|W = 1) \leq \frac{1}{2}$ it is easy to see, using the established inequality that

$$SR = \min(I(W; Y_1), I(W; Y_2)) + I(X; Y_1|W) \leq I(X; Y_1) \leq C,$$

where C is the single channel capacity given by $h(0.2) - 0.4 \approx 0.321928..$ Similarly if $\frac{1}{2} \leq \mathrm{P}(X = 0|W = 0), \mathrm{P}(X = 0|W = 1) \leq 1$ again the sum-rate will be bounded by C. Hence we can assume that $0 \leq \mathrm{P}(X = 0|W = 0) \leq \frac{1}{2} \leq \mathrm{P}(X = 0|W = 1) \leq 1$.

Let $d = \max_{p(x)} I(X; Y_1) - I(X; Y_2)$. Then we can solve for $d = 0.10072952..$ and the optimizing choice for $\mathrm{P}(X = 0) = 0.15843497...$ Now observe that

$$\begin{aligned} SR &\leq I(W; Y_1) + \mathrm{P}(W = 0)I(X; Y_1|W = 0) + \mathrm{P}(W = 1)I(X; Y_2|W = 1) \\ &= I(X; Y_1) + \mathrm{P}(W = 1)(I(X; Y_2|W = 1) - I(X; Y_1|W = 1)) \\ &\leq I(X; Y_1) + \mathrm{P}(W = 1)d. \end{aligned}$$

Similarly

$$\begin{aligned} SR &\leq I(W; Y_2) + \mathrm{P}(W = 0)I(X; Y_1|W = 0) + \mathrm{P}(W = 1)I(X; Y_2|W = 1) \\ &= I(X; Y_2) + \mathrm{P}(W = 0)(I(X; Y_1|W = 0) - I(X; Y_2|W = 0)) \\ &\leq I(X; Y_2) + \mathrm{P}(W = 0)d. \end{aligned}$$

From these two (by adding them) we can deduce that

$$2SR \leq I(X; Y_1) + I(X; Y_2) + d.$$

The maximum of $I(X; Y_1) + I(X; Y_2) = 0.6225562..$ occurs when $\mathrm{P}(X = 0) = \frac{1}{2}$ and hence substituting we obtain that $SR \leq 0.36164288...$

To show that it is indeed on the boundary of the achievable region consider the joint distribution on X and W as follows:
$p(W = 0) = p(W = 1) = \frac{1}{2}$
$p(X = 0|W = 0) = 0.15843497..$ and $p(X = 0|W = 1) = 0.84156502..$
For this distribution, all inequalities reduce to equalities and SR of 0.3616.. is achieved.

3.2 Sum-rate evaluations of the outer bounds for BSSC

3.2.1 Case 1: Bound 2

To evaluate maximum of the sum-rate of the outer bound (Bound 2) it was shown [6] that it suffices to consider $\mathrm{P}(X = 0) = \frac{1}{2}$. (It is immediate using the skew-symmetry of the channel and the inherent symmetry of the outer bound expressions.)

The sum-rate maximum is hence given by

$$\max_{p(u,x),\mathrm{P}(x=0)=\frac{1}{2}} I(U;Y_1) + I(X;Y_2|U)$$

or in other words maximize

$$\max_{p(u,x),\mathrm{P}(x=0)=\frac{1}{2}} I(X;Y_1) + I(X;Y_2|U) - I(X;Y_1|U)$$

Let $\mathrm{P}(x=0) = x$. In [7] it was shown that the curve $f(x) = I(X;Y_1) - I(X;Y_2) = h(\frac{x}{2}) - \frac{1-x}{2} + 1 - 2x$ is concave when $x \in [0, \frac{1}{2}]$ and convex when $x \in [0, \frac{1}{2}]$. Further it was also shown that the lower convex envelope[5] was given by

$$g(x) = \begin{cases} \frac{5x}{4} f(\frac{4}{5}) & 0 \le x \le \frac{4}{5} \\ f(x) & \frac{4}{5} \le x \le 1 \end{cases}.$$

From the definition of the lower convex envelope, we know that

$$I(X;Y_1|U) - I(X;Y_2|U) \ge g(\frac{1}{2})$$

and it easy to see that the equality is indeed achieved for a binary U.

Therefore

$$\max_{p(u,x),\mathrm{P}(x=0)=\frac{1}{2}} I(X;Y_1) + I(X;Y_2|U) - I(X;Y_1|U) = h(\frac{1}{4}) - 0.5 + g(0.5) \approx 0.3725562...$$

This is a correction to the implicit error I made in [6] while calculating the lower convex envelope and obtained a bound of 0.37111....

3.2.2   Case 2: Korner-Marton Bound

To show that this sum-rate is still strictly inside the Korner-Marton[5] outer bound observe that we need to evaluate the union over $p(u,x)$

$$R_1 \le I(U;Y_1)$$
$$R_2 \le I(X;Y_2)$$
$$R_1 + R_2 \le I(U;Y_1) + I(X;Y_2|U)$$

Further if a point $(R_1, R_2) = (a, a)$ belongs to this region, by the skew-symmetry of BSSC, it will also belong to the union over $p(v,x)$

$$R_1 \le I(X;Y_1)$$
$$R_2 \le I(V;Y_2)$$
$$R_1 + R_2 \le I(V;Y_1) + I(X;Y_2|V)$$

and hence to the intersection of the two regions. The key difference between the bounds is that while the former takes the intersection before the union, the latter takes the union prior to the intersection.

---

[5]more precisely, in [7] the upper concave envelope was characterized, and the characterization of the lower convex envelope follows by symmetry.

Suppose we wish to compute

$$\max_{p(u,x)} I(X;Y_1) + I(X;Y_2|U) - I(X;Y_1|U)$$

then from the earlier discussion, this will be the maximum over $x \in [0,1]$ of

$$h(\frac{x}{2}) - x - g(x)$$

It is easy to see that the global maximum will lie when $x \in [0, \frac{4}{5}]$ (otherwise maximum occurs when U is trivial and equals $I(X;Y_2)$). Taking derivatives we obtain that maximum occurs when

$$\frac{1}{2}\log_2 \frac{2-x}{x} - 1 - \frac{5}{4}f(\frac{4}{5}) = 0$$

or

$$x^* = \frac{2}{1 + 2^c} \approx 0.4571429...$$

where $c = 2(1 + \frac{5}{4}f(\frac{4}{5})) \approx 1.7548875...$

Thus the maximum sum rate given by

$$\max_{p(u,x)} I(X;Y_1) + I(X;Y_2|U) - I(X;Y_1|U) \approx 0.3743955...$$

The pair (U,X) that achieves the maximum can be characterized by

$$\mathrm{P}(U=0) = 1 - a, \mathrm{P}(X=0|U=0) = 0, \mathrm{P}(U=1) = a, \mathrm{P}(X=0|U=0) = \frac{4}{5}$$

where $0.8 * a = x^*$ or $a \approx 0.5714286..$

Observe that for this choice

$$I(U;Y_1) = h(\frac{x^*}{2}) - ah(0.4) \approx 0.2206837...$$
$$I(X;Y_2|U) \approx 0.1537118..$$
$$I(X;Y_2) \approx 0.3006499$$

Therefore the point $(R_1, R_2) = (0.1871978.., 0.1871978..)$ lies on the boundary of the Korner-Marton outer bound. In summary, the maximum sum rate given by Korner-Marton outer bound for the BSSC is 0.3743955....

Conclusion

We establish an inequality for the binary skew-symmeteric broadcast channel that was conjectured in [7]. Thus we have quantified the gap between the outer bounds and the inner bounds for this channel. It would be great to determine which of the bounds are weak, and if possible improve them at least for this interesting channel.

References

[1] T Cover, *Broadcast channels*, IEEE Trans. Info. Theory **IT-18** (January, 1972), 2–14.

[2] ⎯⎯⎯, *Comments on broadcast channels*, IEEE Trans. Info. Theory **IT-44** (October, 1998), 2524–2530.

[3] Amin Aminzadeh Gohari and Venkat Anantharam, *Evaluation of marton's inner bound for the general broadcast channel*, CoRR **abs/0904.4541** (2009).

[4] B Hajek and M Pursley, *Evaluation of an achievable rate region for the broadcast channel*, IEEE Trans. Info. Theory **IT-25** (January, 1979), 36–46.

[5] K Marton, *A coding theorem for the discrete memoryless broadcast channel*, IEEE Trans. Info. Theory **IT-25** (May, 1979), 306–311.

[6] C Nair and A El Gamal, *An outer bound to the capacity region of the broadcast channel*, IEEE Trans. Info. Theory **IT-53** (January, 2007), 350–355.

[7] C Nair and V W Zizhou, *On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels*, Proceedings of the ITA Workshop (2008).