On the evaluation of Marton's inner bound for two-receiver broadcast channels

Venkat Anantharam, *Fellow, IEEE,* Amin Gohari, *Senior Member, IEEE,* Chandra Nair, *Fellow, IEEE,*

***Abstract*—Marton's inner bound is the best known achievable rate region for a general two-receiver discrete memoryless broadcast channel. In this paper, we establish improved bounds on the cardinalities of the auxiliary random variables appearing in this inner bound to the true rate region. We combine a perturbation technique, along with a representation using concave envelopes of information-theoretic functions that involve the use of auxiliary random variables, to achieve this improvement. The new cardinality bounds lead to a proof that a randomized time-division strategy achieves every rate triple in Marton's region for binary input broadcast channels. This extends the result by Hajek and Pursley which showed that the Cover-van der Muelen region was exhausted by the randomized time-division strategy.**

I. Introduction

A broadcast channel [1] models a communication scenario where a single sender wishes to communicate multiple messages to many receivers. A two-receiver discrete memoryless broadcast channel consists of a sender X and two receivers Y and Z. The sender maps a triple of messages M_0, M_1, M_2 to a transmit sequence $X^n(m_0, m_1, m_2)(\in \mathcal{X}^n)$ and the receivers each get a noisy version $Y^n(\in \mathcal{Y}^n), Z^n(\in \mathcal{Z}^n)$ respectively. Here $|\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}| < \infty$. Receiver Y wishes to decode the message pair (M_0, M_1) correctly and receiver Z wishes to decode the message pair (M_0, M_2) correctly, with probability approaching 1 as $n \to \infty$. We assume that the channel is memoryless and that there is no feedback from the receivers to the transmitter. This means that $p(y^n, z^n|x^n) = \prod_{i=1}^n p(y_i, z_i|x_i)$. We borrow most of our notation from Chapters 5 and 8 in [2], where the classical results on broadcast channels are reviewed. Let the rates of messages (M_0, M_1, M_2) be (R_0, R_1, R_2) respectively, i.e. $R_i := \frac{1}{n}\log M_i,\ i = 0, 1, 2$. A rate triple is said to be achievable if communication of messages with rates arbitrarily close to it is feasible. The best known achievable rate region for a broadcast channel is the following inner bound [12].

This paper was presented in part at the IEEE International Symposium on Information Theory, 2013.

V. Anantharam is with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: ananth@eecs.berkeley.edu). A. Gohari is with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (email: aminzadeh@sharif.edu). C. Nair is with the department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: chandra@ie.cuhk.edu.hk).

During the course of this work Venkat Anantharam was supported by the ARO MURI grant W911NF-08-1-0233: "Tools for the Analysis and Design of Complex Multi-Scale Networks", the NSF Science and Technology Center grant CCF-0939370: "Science of Information", and the NSF grants CNS-0910702, ECCS-1343398, CNS-1527846, and CIF-1618145, by Marvell Semiconductor Inc., by the U.C. Discovery program, and by the William and Flora Hewlett Foundation supported Center for Long Term Cybersecurity at Berkeley. The work of Amin Gohari was partially supported by the Sharif University of Technology under Grant QB950607. The work of Chandra Nair was partially supported by the following grants from the University Grants Committee of the Hong Kong Special Administrative Region, China: a) (Project No. AoE/E-02/08), b) GRF Projects 415715, 14231916, 1430714. He also acknowledges the support from the Institute of Theoretical Computer Science and Communications (ITCSC) at the Chinese University of Hong Kong, as well as a VC discretionary fund support.

Communicated by V. Prabhakaran, Associate Editor for Shannon Theory.

Theorem 1 (Marton '79)**.** *The union of nonnegative rate triples (R_0, R_1, R_2) satisfying the constraints*

$$
\begin{aligned}
R_0 &\leq \min\{I(W;Y), I(W;Z)\},\\
R_0 + R_1 &\leq I(U,W;Y),\\
R_0 + R_2 &\leq I(V,W;Z),\\
R_0 + R_1 + R_2 &\leq \min\{I(W;Y), I(W;Z)\} + I(U;Y|W)\\
&\qquad + I(V;Z|W) - I(U;V|W),
\end{aligned}
$$

for any triple of random variables (U, V, W) such that $(U, V, W) \to X \to (Y, Z)$ is achievable.

Denote this achievable region as $\mathcal{R}_M$. It was shown in [8], using a perturbation-based approach, that the extreme points of the region are obtained by $p(u, v, w, x)$ that satisfy $|\mathcal{W}| \leq |\mathcal{X}| + 4, |\mathcal{U}| \leq |\mathcal{X}|, |\mathcal{V}| \leq |\mathcal{X}|$. In addition, one can assume w.l.o.g. that X is a function of U, V and W.

A. Motivation

We summarize below our motivation for the work and its potential significance.

a) It is not known whether $\mathcal{R}_M$ is the true capacity region (i.e. the region of all achievable rate triples) for a two-receiver discrete memoryless broadcast channel. One way to establish the suboptimality of Marton's inner bound would be to consider two-letter extensions of $\mathcal{R}_M$ and see if this yields rates outside $\mathcal{R}_M$. The previous cardinality bounds established in [8] were not sufficiently strong to conduct an exhaustive numerical search to test this idea even for the two-letter extension of a binary channel, i.e. one where $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ each have cardinality 2. Theorem 4 and Proposition 1 (our main results) make a significant difference and now we can do rather convincing numerical optimizations over the space of auxiliaries for the two-letter extension of a binary alphabet broadcast channel. This numerical optimization approach has so far not yielded any improvements on the region by going to the two-letter extension.

b) In contrast to the above point, after this work had been done, a similar study was conducted for the interference channel and the Han-Kobayashi achievable region by one of the authors and his students. Due to the presence of more constraints, it is much harder to conduct numerical optimizations in a randomly chosen binary alphabet interference channel setting. However, by considering a subclass where numerical optimizations were possible, many examples showing the suboptimality of the Han-Kobayashi achievable region were obtained in [15].
c) In recent years, several definitive results [5], [13], [19] on channel capacity regions were established by evaluating achievable regions or outer bounds to the capacity region and showing that the inner and outer bounds matched. This raises the question of whether Marton's inner bound has an alternative representation that is better amenable to analysis or in establishing a converse to the coding theorem. We believe that central to answering this question is understanding properties of the joint distributions $p(u,v,w,x)$ corresponding to the extreme points of $\mathcal{R}_M$. We call such distributions of $p(u,v,w,x)$ *extremal distributions* of $\mathcal{R}_M$. In addition to the cardinality bounds on $\mathcal{U},\mathcal{V},\mathcal{W}$, results in [7], [10], [14] and in here further restrict the set of extremal distributions.

B. Evaluation using supporting hyperplanes

Since Marton's achievable region is a convex set in $\mathbb{R}^3_+$ the region can be characterized by determining the supporting hyperplanes, or equivalently by computing the function:

$$\Gamma_M(\gamma_0,\gamma_1,\gamma_2) := \max_{(R_0,R_1,R_2)\in\mathcal{R}_M} \gamma_0 R_0 + \gamma_1 R_1 + \gamma_2 R_2. \quad (1)$$

Observe that if $(R_0,R_1,R_2) \in \mathcal{R}_M$ then any point (R_0',R_1',R_2') with $0 \le R_0' \le R_0, 0 \le R_1' \le R_1, 0 \le R_2' \le R_2$ also belongs to $\mathcal{R}_M$. Hence it suffices to consider supporting hyperplanes of $\mathcal{R}_M$ that have $\gamma_0,\gamma_1,\gamma_2 \ge 0$ to characterize the nontrivial boundary of $\mathcal{R}_M$.

First, we recall the following preliminaries from convex analysis.

Definition 1. An *exposed* point of a closed convex set $\mathcal{C}$ is a point $x \in \mathcal{C}$ such that there is a supporting hyperplane that intersects the convex set only at x.

Definition 2. An *extreme* point of a closed convex set $\mathcal{C}$ is a point $x \in \mathcal{C}$ such that if $x = \lambda y + (1-\lambda) z$ for some $y,z \in \mathcal{C}$ and $\lambda \in (0,1)$, then either $x=y$ or $x=z$.

Theorem 2 (Strasziewicz, Theorem 18.6 in [17])**.** *For any closed convex set C, the set of exposed points of C is a dense subset of the set of extreme points of C. Thus every extreme point is the limit of some sequence of exposed points.*

Since any closed convex set is characterized by its extreme points, to characterize $\mathcal{R}_M$ it suffices to restrict to the set of nonnegative triples $(\gamma_0,\gamma_1,\gamma_2)$ that will lead to exposed points. Determining $\Gamma_M(\gamma_0,\gamma_1,\gamma_2)$ for such triples will, by continuity (using Theorem 2), determine all the extreme points of $\mathcal{R}_M$ and will thereby characterize the set.

From the rate constraints in Theorem 1 one sees that if $(R_0,R_1,R_2) \in \mathcal{R}_M$ then $(\alpha R_0, R_1 + (1-\alpha)\beta R_0, R_2 + (1-\alpha)(1-\beta)R_0) \in \mathcal{R}_M$ for any $0 \le \alpha,\beta \le 1$. An immediate consequence is that the supporting hyperplanes passing through an exposed point with $R_0 > 0$ must have $\gamma_0 \ge \max\{\gamma_1,\gamma_2\}$. Additionally any exposed point with $R_0 = 0$ has a supporting hyperplane satisfying $\gamma_0 = 0$.

From the above discussion, it suffices to compute only $\Gamma_M(\gamma_0,\gamma_1,\gamma_2)$ with $\gamma_0 \ge \max\{\gamma_1,\gamma_2\}$ or $\gamma_0 = 0$. We will now show how we can impose cardinality constraints on the auxiliary random variables showing up in these computations.

Let $\gamma_0 \ge \max\{\gamma_1,\gamma_2\}$. Assume $\gamma_1 \ge \gamma_2$ without loss of generality. Define nonnegative numbers δ_0,δ_1 according to $\gamma_1 = \gamma_2 + \delta_1$, $\gamma_0 = \gamma_1 + \delta_0$. We claim that the equation (2) on top of the next page holds where $t_\lambda(\cdot)$ is a function of the distribution of X defined by

$$\begin{aligned} t_\lambda(X) :=& -(\delta_1 + \lambda(\delta_0+\gamma_2)) I(X;Y) \\ & - (1-\lambda)(\delta_0+\gamma_2) I(X;Z) \\ & + \max_{\substack{p(u,v|x):\\ (U,V)-X-(Y,Z)}} \{(\delta_1+\gamma_2) I(U;Y) + \\ & \qquad + \gamma_2 I(V;Z) - \gamma_2 I(U;V)\}, \end{aligned} \quad (3)$$

and $\mathfrak{C}_X[\cdot]$ denotes the upper concave envelope of the function over the space of probability distributions on $\mathcal{X}$ evaluated at $p(x)$. Here step (a) can be justified by observing that for any $(U,V,W) \to X \to (Y,Z)$ the rate triple given by $R_0 = \min\{I(W;Y), I(W;Z)\}$, $R_1 = I(U,W;Y) - \min\{I(W;Y), I(W;Z)\}$, and

$$\begin{aligned} R_2 = \big[&\min\{I(W;Y), I(W;Z)\} + I(U;Y|W) + I(V;Z|W) \\ & - I(U;V|W) - I(U,W;Y) \big]_+ \end{aligned}$$

belongs to $\mathcal{R}_M$, where $[a]_+ := \max\{a,0\}$. Equality (b) follows by an application of Corollary 2 in [3] which, in turn, follows using a max-min theorem of Terkelsen [18] (our application of Corollary 2 of [3] follows similar lines as the proof of Lemma 1 of [3] using this corollary). Equality (c) follows from the interpretation of auxiliary variables in terms of upper concave envelopes as presented in [16].

Remark 1. It is useful to note that the random variable W plays the role of mixing between various distributions of $p(u,v,x)$, to achieve the concave envelope. In the main part of this section, we will establish cardinality bounds on the sizes of $\mathcal{U}$ and $\mathcal{V}$ that appear in these mixing distributions. This in turn, immediately translates to bounds on the sizes of $\mathcal{U}$ and $\mathcal{V}$ conditioned on the random variable W.

It follows immediately from the arguments in [9] that one can impose cardinality bounds $|\mathcal{U}| \le |\mathcal{X}|, |\mathcal{V}| \le |\mathcal{X}|$ while evaluating $t_\lambda(X)$ in (3). However, due to the non-convexity of the underlying expression, numerical exhaustive search is rather infeasible even when $|\mathcal{X}| = 4$, which would be required when considering a product of two binary-input broadcast channels. The main results of this paper (and the technique) are inspired by the idea of comparing the two-letter extension of Marton's region with the single-letter one to have an informed guess towards its optimality.

$$
\begin{aligned}
\Gamma_M(\gamma_0,\gamma_1,\gamma_2) &= \max_{(R_0,R_1,R_2)\in\mathcal{R}_M} \gamma_0 R_0 + \gamma_1 R_1 + \gamma_2 R_2 \\
&= \max_{(R_0,R_1,R_2)\in\mathcal{R}_M} \delta_0 R_0 + \delta_1 (R_0 + R_1) + \gamma_2 (R_0 + R_1 + R_2) \\
&\overset{(a)}{=} \max_{p(u,v,w,x)} \delta_0 \min\{I(W;Y), I(W;Z)\} + \delta_1 I(U,W;Y) \\
&\qquad + \gamma_2\big(\min\{I(W;Y), I(W;Z)\} + I(U;Y|W) + I(V;Z|W) - I(U;V|W)\big) \\
&= \max_{p(u,v,w,x)} \min_{\lambda\in[0,1]} (\delta_1 + \lambda(\delta_0+\gamma_2)) I(W;Y) + (1-\lambda)(\delta_0+\gamma_2) I(W;Z) \\
&\qquad + (\delta_1+\gamma_2) I(U;Y|W) + \gamma_2 I(V;Z|W) - \gamma_2 I(U;V|W) \\
&\overset{(b)}{=} \min_{\lambda\in[0,1]} \max_{p(u,v,w,x)} (\delta_1 + \lambda(\delta_0+\gamma_2)) I(W;Y) + (1-\lambda)(\delta_0+\gamma_2) I(W;Z) \\
&\qquad + (\delta_1+\gamma_2) I(U;Y|W) + \gamma_2 I(V;Z|W) - \gamma_2 I(U;V|W) \\
&\overset{(c)}{=} \min_{\lambda\in[0,1]} \max_{p(x)} \Big\{ (\delta_1 + \lambda(\delta_0+\gamma_2)) I(X;Y) + (1-\lambda)(\delta_0+\gamma_2) I(X;Z) + \mathfrak{C}_X[t_\lambda(X)] \Big\}, \qquad (2)
\end{aligned}
$$

Define the function

$$
\begin{aligned}
\hat{t}_\lambda(X) := &-(\delta_1 + \lambda(\delta_0+\gamma_2)) I(X;Y) \qquad (4) \\
&- (1-\lambda)(\delta_0+\gamma_2) I(X;Z) \\
&+ \max_{\substack{p(u,v|x):\\ (U,V)-X-(Y,Z)}} \big\{ (\delta_1+\gamma_2) I(U;Y) \\
&\qquad + \gamma_2 I(V;Z) - \gamma_2 I(U;V) \big\},
\end{aligned}
$$

where additionally the alphabets of the auxiliary variables U, V are assumed satisfy $|\mathcal{U}| + |\mathcal{V}| \le |\mathcal{X}| + 1$ and X is a function of (U,V). Clearly $\hat{t}_\lambda(X) \le t_\lambda(X)$.

Proposition 1. *Let $t_\lambda(X)$ and $\hat{t}_\lambda(X)$ be defined according to* (3) *and* (4) *respectively. Then the upper concave envelopes of $t_\lambda(X)$ and $\hat{t}_\lambda(X)$ match, i.e.*

$$\mathfrak{C}_X[t_\lambda(X)] = \mathfrak{C}_X[\hat{t}_\lambda(X)].$$

Remark 2. It is worth noting that the above bounds make it possible to estimate $\hat{t}_\lambda(X)$ when $|\mathcal{X}| = 4$. Further, it is also curious that the bounds on the cardinalities show an additive trade-off in their sizes. This is rather surprising both in the form and in that it captures a trade-off between the sizes of the transmission spaces of the two messages akin to the intrinsic trade-off of the broadcast channel.

In the preceding, we considered the case where $\gamma_0 \ge \max\{\gamma_1, \gamma_2\}$. The remaining interesting case happens when $\gamma_0 = 0$. As earlier, let $\gamma_1 \ge \gamma_2$, and let δ_1 be defined according to $\gamma_1 = \gamma_2 + \delta_1$. Then, mimicking the arguments above, one can see that (5) on top of the next page holds where $t^0_\lambda(\cdot)$ is a function of the distribution of X defined by

$$
\begin{aligned}
t^0_\lambda(X) := &-(\delta_1 + \lambda\gamma_2) I(X;Y) - (1-\lambda)\gamma_2 I(X;Z) \qquad (6) \\
&+ \max_{\substack{p(u,v|x):\\ (U,V)-X-(Y,Z)}} \big\{ (\delta_1+\gamma_2) I(U;Y) + \gamma_2 I(V;Z) \\
&\qquad - \gamma_2 I(U;V) \big\}.
\end{aligned}
$$

For the analog of step (a) in (2) observe that for any $(U,V,W) \to X \to (Y,Z)$ the rate triple given by $R_0 = 0$, $R_1 = I(U,W;Y)$, and

$$
\begin{aligned}
R_2 = \big[&\min\{I(W;Y), I(W;Z)\} + I(U;Y|W) + I(V;Z|W) \\
&- I(U;V|W) - I(U,W;Y)\big]_+
\end{aligned}
$$

belongs to $\mathcal{R}_M$.

As before, define the function

$$
\begin{aligned}
\hat{t}^0_\lambda(X) := &-(\delta_1 + \lambda\gamma_2) I(X;Y) - (1-\lambda)\gamma_2 I(X;Z) \qquad (7) \\
&+ \max_{\substack{p(u,v|x):\\ (U,V)-X-(Y,Z)}} \big\{ (\delta_1+\gamma_2) I(U;Y) + \gamma_2 I(V;Z) \\
&\qquad - \gamma_2 I(U;V) \big\},
\end{aligned}
$$

where additionally the alphabets of the auxiliary variables U, V are assumed to satisfy $|\mathcal{U}| + |\mathcal{V}| \le |\mathcal{X}| + 1$, and X is a function of (U,V). Clearly $\hat{t}^0_\lambda(X) \le t^0_\lambda(X)$.

Proposition 2. *Let $t^0_\lambda(X)$ and $\hat{t}^0_\lambda(X)$ be defined according to* (6) *and* (7) *respectively. Then the upper concave envelopes of $t^0_\lambda(X)$ and $\hat{t}^0_\lambda(X)$ match, i.e.*

$$\mathfrak{C}_X[t^0_\lambda(X)] = \mathfrak{C}_X[\hat{t}^0_\lambda(X)].$$

1) Idea of the proof: The proof builds on perturbation techniques introduced in [9]. A previously unused idea exploited in this paper is to use the dual form of the upper concave envelope of a function. To explain this idea, take an arbitrary function $f(p(x))$ defined on the probability distributions on a finite alphabet $\mathcal{X}$. The upper concave envelope of f, denoted by $\mathfrak{C}[f](p(x))$, is the smallest concave function on the probability simplex that dominates $f(p(x))$. It can be expressed via an optimization problem involving an auxiliary random variable (see [16]) as follows:

$$\mathfrak{C}[f](p(x)) := \sup_{p(w|x)} \sum_w p(w) f(p(x|w)). \qquad (8)$$

Note that one can restrict the maximization to $|\mathcal{W}| \le |\mathcal{X}| + 1$ using Carathéodory's theorem on convex hulls of sets in finite dimensions. Alternatively, upper concave envelopes can

$$\Gamma_M(0,\gamma_1,\gamma_2) = \min_{\lambda\in[0,1]} \max_{p(u,v,w,x)} (\delta_1+\lambda\gamma_2)I(W;Y) + (1-\lambda)\gamma_2 I(W;Z) \\ + (\delta_1+\gamma_2)I(U;Y|W) + \gamma_2 I(V;Z|W) - \gamma_2 I(U;V|W) \\ = \min_{\lambda\in[0,1]} \max_{p(x)} \Big\{(\delta_1+\lambda\gamma_2)I(X;Y) + (1-\lambda)\gamma_2 I(X;Z) + \mathfrak{C}_X[t^0_\lambda(X)]\Big\}, \tag{5}$$

be studied through duality. Given a vector $\mathbf{d} = (d_x, x \in \mathcal{X})$, the dual of the function $f(\cdot)$ is defined as

$$f^\dagger(\mathbf{d}) = \max_{p(x)} \left\{ f(p(x)) - \sum_x d_x p(x) \right\}. \tag{9}$$

Two properties that we exploit are that (i) the dual of a function f is the same as the dual of $\mathfrak{C}[f]$; (ii) the dual of f uniquely determines $\mathfrak{C}[f]$. In fact, the following equation holds:

$$\mathfrak{C}[f](p(x)) = \inf_{\mathbf{d}} \Big\{ f^\dagger(\mathbf{d}) + \sum_x d_x p(x) \Big\}. \tag{10}$$

Thus, if $f^\dagger(\mathbf{d}) = g^\dagger(\mathbf{d})$ for two functions f and g, then $\mathfrak{C}[f] = \mathfrak{C}[g]$. Also, for any functions f and g, one can use a proof by contradiction to show[1] that

$$\exists \mathbf{d}: \ f^\dagger(\mathbf{d}) < g^\dagger(\mathbf{d}) \iff \exists p(x) : \mathfrak{C}[f](p(x)) < \mathfrak{C}[g](p(x)). \tag{11}$$

□

The above argument shows that the following theorem implies Propositions 1 and 2. Proposition 2 follows by setting $\delta_0 = 0$ below.

Theorem 3. *Let $\mathbf{d} = (d_x, x \in \mathcal{X})$ be any vector. For $\delta_1, \delta_0, \gamma_2 \geq 0$ and $\lambda \in [0,1]$ consider the maximization problem defined by:*

$$\max_{p(x)} \Big\{ -(\delta_1 + \lambda(\delta_0+\gamma_2))I(X;Y) \\ -(1-\lambda)(\delta_0+\gamma_2)I(X;Z) \\ + \max_{\substack{p(u,v|x): \\ (U,V)-X-(Y,Z)}} \big\{(\delta_1+\gamma_2)I(U;Y) + \gamma_2 I(V;Z) \\ - \gamma_2 I(U;V)\big\} \\ - \sum_x d_x p(x) \Big\}.$$

Then there exist random variables (U,V) satisfying $|\mathcal{U}|+|\mathcal{V}| \leq |\mathcal{X}|+1$ that achieve the maximum. Further we can assume that X is a function of (U,V).

We will establish Theorem 3 in Section II.

Remark 3. We will henceforth assume that $\gamma_2 > 0$; else, an immediate inspection yields that setting $U = X$ and V to be a constant random variable is a maximizer, and this satisfies the cardinality constraints with equality.

[1] If $\mathfrak{C}[f](p(x)) < \mathfrak{C}[g](p(x))$ at some $p(x)$, then from (10) it cannot be that for all $\mathbf{d}$ we have $f^\dagger(\mathbf{d}) \geq g^\dagger(\mathbf{d})$. To see the reverse direction, since the dual of $\mathfrak{C}[f]$ is same as as that of f, we have $f^\dagger(\mathbf{d}) = \max_{p(x)} \{\mathfrak{C}[f](p(x)) - \sum_x d_x p(x)\}$ and hence if $f^\dagger(\mathbf{d}) < g^\dagger(\mathbf{d})$ then it cannot be that for all $p(x)$ we have $\mathfrak{C}[f](p(x)) \geq \mathfrak{C}[g](p(x))$.

Remark 4. The assumption that X is a function of U and V is an immediate consequence of the functional representation lemma [2]. Given $p(u,v,x)$, let Q be independent of (U,V) such that X is a function of (Q,U,V). Then setting $\tilde{U} = (U,Q)$ and $\tilde{V} = V$ does not decrease the expression in Theorem 3. Since the interest is in extremal distributions only; one can always begin the perturbations from a $p(u,v,x)$ where X is a function of U,V. Note that all the perturbations considered in this paper are of the multiplicative form, i.e. $p_\epsilon(u,v,x) = p(u,v,x)(1+\epsilon\, L(u,v,x))$, and hence preserve the functional relationship between X and U,V.

II. Proof of Theorem 3

Consider maximizing over $p(u,v,x)$ the expression:

$$S(p(u,v,x)) := -(\delta_1+\lambda(\delta_0+\gamma_2))I(X;Y) \\ -(1-\lambda)(\delta_0+\gamma_2)I(X;Z) \\ +(\delta_1+\gamma_2)I(U;Y) + \gamma_2 I(V;Z) \\ -\gamma_2 I(U;V) - \sum_x d_x p(x), \tag{12}$$

for $\delta_1, \delta_0 \geq 0$, $\gamma_2 > 0$, $\lambda \in [0,1]$ and arbitrary real numbers $d_x, x \in \mathcal{X}$. Here $(U,V) \to X \to (Y,Z)$

Proposition 3. *Let (U,V) be a cardinality minimal pair (in the sense of $|\mathcal{U}|+|\mathcal{V}|$) such that X is a function of U and V (see Remark 4) and $p(u,v,x)$ is a maximizer of $S(p(u,v,x))$. Then one cannot find $\omega_1(u)$ and $\omega_2(v)$ such that $\omega_1(u)$ and $\omega_2(v)$ are not simultaneously zero for all u and v and further*

$$\sum_u p(u)\omega_1(u) = 0, \quad \sum_v p(v)\omega_2(v) = 0,$$
$$\sum_{uv} p(u,v,x)\omega_1(u) - \sum_{uv} p(u,v,x)\omega_2(v) = 0 \quad \forall x.$$

Proof. For the main part of the proof, we will assume that $\lambda \in (0,1)$. The extreme cases, $\lambda \in \{0,1\}$, are rather easy and will be taken care of in Remark 5.

Suppose we are given that $p(u,v,x)$ is a (cardinality-minimal) global maximizer of $S(p(u,v,x))$. Let us first consider perturbations of the form $p^{(1)}_\epsilon(u,v,x) = p(u,v,x)(1 + \epsilon\, \omega_1(u))$ such that $\sum_u p(u)\omega_1(u) = 0$. In this case we are preserving $p(v,x|u)$ and perturbing the marginal distribution of U. Rewrite $S(p(u,v,x))$ as follows:

$$S(p(u,v,x)) = (1-\lambda)(\gamma_2+\delta_0)(H(Y)-H(Z)) - \delta_0 H(Y) \\ -(\delta_1+\gamma_2)H(Y|U) + \gamma_2 H(V|U) \\ -\gamma_2 H(V|Z) + (\delta_1+\lambda(\delta_0+\gamma_2))H(Y|X) \\ +(1-\lambda)(\delta_0+\gamma_2)H(Z|X) - \sum_x d_x p(x). \tag{13}$$

The terms $H(Y|X), H(Z|X), H(V|U), H(Y|U), \sum_x d_x p_x$ are linear in ϵ since we are preserving $p(v,x|u)$, and the terms $-H(V|Z), -H(Y)$ are convex in ϵ. The first derivative for this perturbation has to be zero, and the second derivative has to be nonpositive. Since $(1-\lambda)(\gamma_2+\delta_0) > 0$ (since $\gamma_2 > 0$ from Remark 3 and $\lambda \in (0,1)$) for the second derivative to be nonpositive we must have

$$\frac{d^2}{d\epsilon^2}\left[H(Y) - H(Z)\right]_{\epsilon=0} \le 0. \tag{14}$$

Note that the above quantity depends solely on $p_\epsilon^{(1)}(x) := \sum_{uv} p(u,v,x)(1+\epsilon\,\omega_1(u))$.

Next consider perturbations of the form $p_\epsilon^{(2)}(u,v,x) = p(u,v,x)(1+\epsilon\,\omega_2(v))$ such that $\sum_v p(v)\omega_2(v) = 0$. In this case we are preserving $p(u,x|v)$ and perturbing the marginal distribution of V. Rewrite $S(p(u,v,x))$ as follows:

$$\begin{aligned} S(p(u,v,x)) =& \lambda(\gamma_2+\delta_0)(H(Z) - H(Y)) - \delta_0 H(Z) \\ & - \delta_1 H(Y|U) - \gamma_2 H(U|Y) + \gamma_2 H(U|V) \\ & - \gamma_2 H(Z|V) + (\delta_1 + \lambda(\delta_0+\gamma_2))H(Y|X) \\ & + (1-\lambda)(\delta_0+\gamma_2)H(Z|X) - \sum_x d_x p(x). \end{aligned} \tag{15}$$

The terms $H(Y|X), H(Z|X), H(U|V), H(Z|V), \sum_x d_x p_x$ are linear in ϵ since we are preserving $p(u,x|v)$ and the terms $-H(Y|U), -H(U|Y), -H(Z)$ are convex in ϵ. The first derivative for this perturbation has to be zero, and the second derivative has to be nonpositive. That $\lambda\gamma_2 > 0$ implies

$$\frac{d^2}{d\epsilon^2}\left[H(Z) - H(Y)\right]_{\epsilon=0} \le 0, \tag{16}$$

and the above second derivative depends solely on $p_\epsilon^{(2)}(x) := \sum_{uv} p(u,v,x)(1+\epsilon\,\omega_2(v))$.

Now let us assume that one can find $\omega_1(u)$ and $\omega_2(v)$ such that $\omega_1(u)$ and $\omega_2(v)$ are not simultaneously zero for all u and v and further

$$\begin{aligned} &\sum_u p(u)\omega_1(u) = 0, \sum_v p(v)\omega_2(v) = 0, \\ &\sum_{uv} p(u,v,x)\omega_1(u) - \sum_{uv} p(u,v,x)\omega_2(v) = 0 \quad \forall x. \end{aligned} \tag{17}$$

We will arrive at a contradiction.

Consider the two perturbations induced by the nontrivial pair $\omega_1(u)$ and $\omega_2(v)$. Since we are at a global maximum, the second derivatives with respect to both of these perturbations have to be nonpositive. From (17) we know that they induce a common perturbation in $p(x)$, defined according to

$$\begin{aligned} p_\epsilon(x) &= \sum_{uv} p(u,v,x)(1+\epsilon\,\omega_2(v)) \\ &= \sum_{uv} p(u,v,x)(1+\epsilon\,\omega_1(u)). \end{aligned}$$

This implies that in (14), (16) we must have,

$$\frac{d^2}{d\epsilon^2}\left[H(Z) - H(Y)\right]_{\epsilon=0} = 0.$$

This forces that for the perturbation using $\omega_1(u)$, we must have (note $\gamma_2 > 0$)

$$\frac{d^2}{d\epsilon^2}\left[H(V|Z)\right]_{\epsilon=0} = \frac{d^2}{d\epsilon^2}\left[\delta_0 H(Y)\right]_{\epsilon=0} = 0,$$

and for the perturbation using $\omega_2(v)$, we must have (again note $\gamma_2 > 0$)

$$\begin{aligned} \frac{d^2}{d\epsilon^2}\left[H(U|Y)\right]_{\epsilon=0} &= \frac{d^2}{d\epsilon^2}\left[\delta_1 H(Y|U)\right]_{\epsilon=0} \\ &= \frac{d^2}{d\epsilon^2}\left[\delta_0 H(Z)\right]_{\epsilon=0} = 0. \end{aligned}$$

Lemma 3 from Appendix A implies that for the perturbation using $\omega_1(u)$ the terms $H_{p_\epsilon^{(1)}}(V|Z)$ and $H(Y)$ are linear in ϵ, and for the perturbation using $\omega_2(v)$, the terms $H_{p_\epsilon^{(2)}}(U|Y)$, $H_{p_\epsilon^{(2)}}(Y|U)$, and $H(Z)$ are linear in ϵ. We separate the argument into two cases.

Case 1: $\delta_0 > 0$. This implies that both $H(Y)$ and $H(Z)$ are linear in ϵ under both perturbations since both perturbations induce the same perturbation $p_\epsilon(x)$. Hence both terms $S(p_\epsilon^{(1)}(u,v,x))$ and $S(p_\epsilon^{(2)}(u,v,x))$ would be linear in ϵ and since the first derivative is zero, they must remain constant under both perturbations. Now, if we take a nontrivial perturbation, say $\omega_1(u)$, and take ϵ to its upper or lower limit[2], we get a distribution on U whose support is smaller than that for $\epsilon = 0$ while the value of $S(p(u,v,x))$ has been preserved. Thus we have been able to reduce $|\mathcal{U}| + |\mathcal{V}|$, which is a contradiction.

Case 2: $\delta_0 = 0$. Express $H(Y) - H(Z)$ as a function of ϵ induced by the common $p_\epsilon(x)$ as $H(Y) - H(Z) = g(\epsilon) + \hat{a}\epsilon + \hat{b}$, with $g(0) = g'(0) = 0$. Note that in both perturbations, since all other terms are linear in ϵ, we can write $S(p_\epsilon^{(1)}(u,v,x))$ as follows:

$$(1-\lambda)(\gamma_2+\delta_0)g(\epsilon) + a_1\epsilon + b_1,$$

and we can write $S(p_\epsilon^{(2)}(u,v,x))$ as follows:

$$-\lambda(\gamma_2+\delta_0)g(\epsilon) + a_2\epsilon + b_2.$$

The fact that at $\epsilon = 0$ the values under both perturbations match implies that $b_1 = b_2 = b$. Further since first derivatives with respect to both perturbations are 0 at $\epsilon = 0$, this forces $a_1 = a_2 = 0$. Thus we have

$$\begin{aligned} S(p_\epsilon^{(1)}(u,v,x)) &= (1-\lambda)(\gamma_2+\delta_0)g(\epsilon) + b, \\ S(p_\epsilon^{(2)}(u,v,x)) &= -\lambda(\gamma_2+\delta_0)g(\epsilon) + b. \end{aligned}$$

Thus, if $g(\epsilon)$ is nonzero for any valid ϵ, this will contradict the global maximality at $\epsilon = 0$.

Thus, as when $\delta_0 = 0$, we have that $S(p(u,v,x))$ remains constant under *both perturbations*. Now if we take a nontrivial perturbation, say $\omega_1(u)$ and take ϵ to its upper or lower limit, as said earlier, we get a distribution on U whose support is smaller than that for $\epsilon = 0$ while the value of $S(p(u,v,x))$ has been preserved. Thus we have been able to reduce $|\mathcal{U}| + |\mathcal{V}|$, which is a contradiction. $\square$

[2]The limits on ϵ are determined by the nonnegativity of the probability vector and at both of the positive(upper) or the negative(lower) limit for ϵ we have $\min_u p(u)(1+\epsilon\,\omega_1(u)) = 0$.

Remark 5. The case when $\lambda = 0$ or $\lambda = 1$ is much simpler. For $\lambda = 0$ observe that for any perturbation of the type $p(u,v,x)(1+\epsilon\omega_2(v))$, the resultant expression (see (15)) is convex in ϵ, and hence we can take the limiting ϵ (either the upper or lower limit) and reduce the cardinality of V. Indeed, it is not hard to see that an optimal choice is $V =$ constant. This suffices for the proof of Proposition 3, but also note that once V is a constant, an immediate inspection says that $U = X$ is a maximizer. A similar argument works for $\lambda = 1$.

We combine Proposition 3 with the Lemma 1 below to establish Theorem 3.

Lemma 1. *For any $p(u,v,x)$ where $|\mathcal{U}|+|\mathcal{V}| > |\mathcal{X}|+1$, one can find $\omega_1(u)$ and $\omega_2(v)$ such that $\omega_1(u)$ and $\omega_2(v)$ are not simultaneously zero for all u and v and further*

$$\sum_u p(u)\omega_1(u) = 0, \quad \sum_v p(v)\omega_2(v) = 0,$$
$$\sum_{uv} p(u,v,x)\omega_1(u) - \sum_{uv} p(u,v,x)\omega_2(v) = 0 \quad \forall x.$$

Proof of Lemma 1. These are $|\mathcal{X}|+2$ equations in total but one of the equations is redundant since the first pair of equations imply that

$$\begin{aligned}\sum_x \sum_{uv} p(u,v,x)\omega_1(u) &= \sum_u p(u)\omega_1(u) = 0 \\ &= \sum_v p(v)\omega_2(v) = \sum_x \sum_{uv} p(u,v,x)\omega_2(v).\end{aligned}$$

Thus the first pair of equations themselves imply that

$$\sum_x \left(\sum_{uv} p(u,v,x)\omega_1(u) - \sum_{uv} p(u,v,x)\omega_2(v) \right) = 0.$$

Thus, there are $|\mathcal{X}|+1$ independent equations at most. The choice $\omega_1(u) = \omega_2(v) = 0$ for all u, v solves the above system of linear equations. Therefore the system of linear equations is not inconsistent. Since the total number of free variables is $|\mathcal{U}|+|\mathcal{V}|$, which is strictly larger than the number of equations, i.e. $|\mathcal{X}|+1$, we must have a nontrivial solution for $\omega_1(u)$ and $\omega_2(v)$. ☐

A. Recasting the main results in a traditional form

Usually cardinality bounds on the auxiliary variables are stated for the region represented by a collection of inequalities, as in the bound (1); so we will recast the results in the previous section using traditional arguments à la the Fenchel-Bunt extension of Carathéodory's theorem.

Theorem 4. *The set of all triples (R_0, R_1, R_2) satisfying the constraints*

$$\begin{aligned} R_0 &\le \min\{I(W;Y), I(W;Z)\}, \\ R_0 + R_1 &\le I(UW;Y), \\ R_0 + R_2 &\le I(VW;Z), \\ R_0 + R_1 + R_2 &\le \min\{I(W;Y), I(W;Z)\} + I(U;Y|W) \\ &\quad + I(V;Z|W) - I(U;V|W), \end{aligned}$$

for any triple of random variables (U,V,W) such that $(U,V,W) \to X \to (Y,Z)$, can be computed under the restriction $|\mathcal{W}| \le |\mathcal{X}|+4$, and conditioned on any $W = w$ the sizes of $\mathcal{U}$ and $\mathcal{V}$ can be restricted to satisfy $|\mathcal{U}|+|\mathcal{V}| \le |\mathcal{X}|+1$, and further we can assume X to be function (possibly dependent on w) of U and V.

Proof. The proof comes from noting that each point in a convex set can be expressed as a convex combination of its extreme points. By our previous argument, every extreme point of Marton's region can be obtained by restricting sizes of to $\mathcal{U}$ and $\mathcal{V}$ to $|\mathcal{U}|+|\mathcal{V}| \le |\mathcal{X}|+1$ and X to be a function of U and V, conditioned on W (see Remark 1 and Theorem 2). Indeed, every boundary point is in the convex combination of extreme points that lie on a given supporting hyperplane of the form $\gamma_0 R_0 + \gamma_1 R_1 + \gamma_2 R_2$. As before, w.l.o.g., let us assume that $\gamma_1 \ge \gamma_2$.

Now consider the space of probability distributions $p(u,v,x)$ with $|\mathcal{U}|+|\mathcal{V}| \le |\mathcal{X}|+1$, and consider the continuous mapping from the space to $\mathbb{R}_+^{|\mathcal{X}|+4}$ defined by

$$\begin{aligned}&\big(p(X=1), \ldots, p(X=|\mathcal{X}|-1), H(Y), H(Z), \\ &\ I(U;Y), I(V;Z), I(U;V)\big).\end{aligned}$$

Let S denote the range of this mapping. By the Fenchel-Bunt extension of Carathéodory's theorem, every point in the convex hull of S can be expressed as a convex combination of at most $|\mathcal{X}|+4$ points in S. This yields a distribution $p(w,u,v,x)$ satisfying: $|\mathcal{W}| \le |\mathcal{X}|+4$; conditioned on any $W = w$ the sizes of $\mathcal{U}$ and $\mathcal{V}$ satisfy $|\mathcal{U}|+|\mathcal{V}| \le |\mathcal{X}|+1$ and X is a function of U and V; and the values of $I(W;Y), I(W;Z), I(U;Y|W), I(V;Z|W)$ and $I(U;V|W)$ are preserved. ☐

B. Application to binary input broadcast channels

If we restrict our attention to the case when X is binary, we obtain from Theorem 4 that it suffices to consider U and V such that conditioned on $W = w$, $|\mathcal{U}|+|\mathcal{V}| \le 3$, which implies one of U and V must be a constant random variable. It is immediate that the other variable must be X. This particular choice is referred to in the literature as the *randomized time-division* strategy.

Thus, Marton's region for binary input broadcast channels reduces to that given in the following result.

Corollary 1. *Marton's region for binary input broadcast channels can be written as the union of nonnegative rate* *triples* (R_0, R_1, R_2) *satisfying the constraints*

$$R_0 \le \min\{I(W;Y), I(W;Z)\},$$

$$R_0 + R_1 \le I(W;Y) + \sum_{i=1}^{k} p_W(i) I(X;Y|W=i),$$

$$R_0 + R_2 \le I(W;Z) + \sum_{i=k+1}^{\ell} p_W(i) I(X;Z|W=i),$$

$$\begin{aligned} R_0 + R_1 + R_2 \le \min\{I(W;Y), I(W;Z)\} \\ + \sum_{i=1}^{k} p_W(i) I(X;Y|W=i) \\ + \sum_{i=k+1}^{\ell} p_W(i) I(X;Z|W=i), \end{aligned}$$

over random variables W *with alphabet* $\mathcal{W} = \{1, 2, \ldots, \ell\}$ *such that* $W \to X \to (Y,Z)$ *is Markov, and* $k \le \ell \le 5$.

This region is known as the randomized time-division region [2] and this extends the result of [11] which showed that the Cover-van der Muelen region (which is at most as large as the Marton's region in general) reduced to the randomized-time-division region. The case for the sum rate was shown earlier in [6] using the inequality discussed in Section II-B1.

Remark 6. Note that the reduction to $|\mathcal{W}| = \ell \le 5$ is possible since $I(U;V|W) = 0$ whenever U or V is a constant. Thus we do not need to preserve this average when using Carathéodory's theorem. It may be possible to further reduce the cardinality of $\mathcal{W}$ in the binary setting. The main objective of this paper is to obtain sharper bounds on $\mathcal{U}$ and $\mathcal{V}$, so we have not pursued this direction of investigation. The reason for not being concerned about this is that, while evaluating Marton's region for binary input broadcast channels, the authors found that it is much easier to evaluate it via the method of supporting hyperplanes as outlined in Section I-B. In this case, for every fixed $\lambda \in [0,1]$ in (2) and (3), a binary-valued W suffices.

1) An observation: In [6] it was shown that it suffices to consider a randomized time-division strategy, as above, to obtain the maximum sum-rate of Marton's region for a binary input broadcast channel. The main step was to prove the following inequality, for any $(U,V) \to X \to (Y,Z)$ and binary X:

$$I(U;Y) + I(V;Z) - I(U;V) \le \max\{I(X;Y), I(X;Z)\}. \tag{18}$$

A direct and natural extension of this approach to the weighted case (i.e. all of Marton's region) to show optimality of randomized time-division would be to conjecture that the following weighted version of the above inequality held. Namely, one might hope that for every $\alpha > 1$ and X binary, $(U,V) \to X \to (Y,Z)$ would imply

$$\alpha I(U;Y) + I(V;Z) - I(U;V) \overset{?}{\le} \max\{\alpha I(X;Y), I(X;Z)\}. \tag{19}$$

However, this inequality is not correct, and a counterexample to the inequality is presented in Appendix A.

Remark 7. It is curious to observe that our new proof strategy has established that randomized time-division is optimal for all of Marton's region even though (19) is false in general. What our argument shows is that, for those $p(x)$ that are used to obtain the concave envelope as in (2) and (3), the above inequality indeed holds.

C. A study via numerical simulations

The main motivation for this research exercise was to be able to compare the two-letter Marton's achievable region and the single-letter achievable region and determine if there are channels for which the two-letter region strictly outperforms the single-letter region. Despite extensive simulations on binary input broadcast channels we were unable to find channels for which the two-letter region strictly improved on the single-letter region. In this section, we wish to provide some details on our simulations as well as our numerical observations so that it may help other researchers who may be interested in pursuing this line of attack.

We analyzed Marton's region using the expression in (2). In particular. Marton's inner bound is optimal if and only if for every channel $p(y,z|x)$ and for any $\delta_1, \delta_0, \gamma_2 \ge 0$, the following equality holds

$$\begin{aligned} &\min_{\lambda\in[0,1]} \max_{p(x)} \Big\{ (\delta_1 + \lambda(\delta_0+\gamma_2)) I(X;Y) \\ &\qquad + (1-\lambda)(\delta_0+\gamma_2) I(X;Z) + \mathfrak{C}_X[t_\lambda(X)] \Big\} \\ &= \frac{1}{2} \min_{\lambda\in[0,1]} \max_{p(x_1,x_2)} \Big\{ (\delta_1 + \lambda(\delta_0+\gamma_2)) I(X_1,X_2;Y_1,Y_2) \\ &\qquad + (1-\lambda)(\delta_0+\gamma_2) I(X_1,X_2;Z_1,Z_2) \\ &\qquad + \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1,X_2)] \Big\} \end{aligned} \tag{20}$$

where the second expression is evaluated for the two-letter channel defined according to $p(y_1,z_1|x_1) \otimes p(y_2,z_2|x_2)$. Since we were initially motivated to find examples that demonstrated the suboptimality of Marton's inner bound, we relaxed the above condition and searched for pairs of channels and distributions $p(x_1,x_2)$ where the following inequality held:

$$\begin{aligned} &(\delta_1 + \lambda(\delta_0+\gamma_2)) I(X_1;Y_1) + (1-\lambda)(\delta_0+\gamma_2) I(X_1;Z_1) \\ &+ \mathfrak{C}_{X_1}[t_\lambda(X_1)] + (\delta_1 + \lambda(\delta_0+\gamma_2)) I(X_2;Y_2) \\ &+ (1-\lambda)(\delta_0+\gamma_2) I(X_2;Z_2) + \mathfrak{C}_{X_2}[t_\lambda(X_2)] \\ &\quad < (\delta_1 + \lambda(\delta_0+\gamma_2)) I(X_1,X_2;Y_1,Y_2) \\ &\qquad + (1-\lambda)(\delta_0+\gamma_2) I(X_1,X_2;Z_1,Z_2) \\ &\qquad + \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1,X_2)] \end{aligned} \tag{21}$$

where the right-hand side is evaluated for the channel $p_1(y_1,z_1|x_1) \otimes p_2(y_2,z_2|x_2)$ at a joint input distribution $p(x_1,x_2)$ and the two terms on the left-hand side are evaluated for their corresponding channels at the corresponding marginal distributions.

Remark 8. Clearly, if there is no instance where (21) held, then it would imply that (20) held and Marton's region would be optimal. On the other hand, it is possible that there are instances that satisfy (21) while (20) continues to hold. Hence, our numerical searches were leaning towards gaining insight for instances where (21) (a weaker condition) is true.

Next, observe that for any $p_1(y_1, z_1|x_1) \otimes p_2(y_2, z_2|x_2)$ and any $p(x_1, x_2)$ we have

$$I(X_1, X_2; Y_1, Y_2) \leq I(X_1; Y_1) + I(X_2; Y_2),$$
$$I(X_1, X_2; Z_1, Z_2) \leq I(X_1; Z_1) + I(X_2; Z_2).$$

Therefore (21) holds only if there is an instance of $p(x_1, x_2)$ where

$$\mathfrak{C}_{X_1}[t_\lambda(X_1)] + \mathfrak{C}_{X_2}[t_\lambda(X_2)] < \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)]. \quad (22)$$

Note that the right-hand side is evaluated at a joint input distribution $p(x_1, x_2)$ and the two terms on the left-hand side are evaluated at the corresponding marginal distributions. This is an inequality that can be studied via the dual expression for the concave envelope.

Lemma 2. *There is an instance where the inequality in* (22) *holds if and only if there exist two functions $d_1(x_1)$ and $d_2(x_2)$ such that*

$$\max_{p_1(x_1)} \left\{ t_\lambda(X_1) - \sum_{x_1} p_1(x_1) d_1(x_1) \right\}$$
$$+ \max_{p_2(x_2)} \left\{ t_\lambda(X_2) - \sum_{x_2} p_2(x_2) d_2(x_2) \right\} <$$
$$\max_{p(x_1,x_2)} \left\{ t_\lambda(X_1, X_2) - \sum_{x_1,x_2} p(x_1, x_2)(d_1(x_1) + d_2(x_2)) \right\}.$$

Proof. The proof follows from the dual formulation of the concave envelope. The left-hand side of the inequality in (22) depends only on the marginal distributions $p_1(x_1), p_2(x_2)$, hence one can equivalently look for two instances of $p_1(x_1)$ and $p_2(x_2)$ such that

$$\begin{aligned}&\mathfrak{C}_{X_1}[t_\lambda(X_1)] + \mathfrak{C}_{X_2}[t_\lambda(X_2)] \\ &< \max_{p(x_1,x_2)\in\Pi(p_1,p_2)} \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)],\end{aligned} \quad (23)$$

where

$$\Pi(p_1, p_2) := \{p(x_1, x_2) : p(x_1) = p_1(x_1), p(x_2) = p_2(x_2), \quad \forall x_1, x_2\}$$

denotes the set of all joint distributions with the given marginals.

Next, we argue that for any instances of $p_1(x_1)$ and $p_2(x_2)$

$$\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)] = \mathfrak{C}_{X_1,X_2}\left[\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} t_\lambda(X_1, X_2)\right]. \quad (24)$$

Note that both sides of the equality are functions of pairs of distributions (p_1, p_2), and the concave envelope $\mathfrak{C}_{X_1,X_2}$ on the right-hand side should be interpreted as one over the pairs of distributions $(p_1(x_1), p_2(x_2))$.

Equality (24) follows from the following two observations: (i) if $f(x, y)$ is jointly concave in (x, y), then $g(x) := \sup_y f(x, y)$ is concave in x.[3] This implies that $\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)]$ is jointly concave in pairs of distributions (p_1, p_2). Since

$$\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)] \geq \max_{p(x_1,x_2)\in\Pi(p_1,p_2)} t_\lambda(X_1, X_2)$$

and the left-hand side is concave in pairs of distributions, we have from the definition of the upper concave envelope that

$$\max_{p(x_1,x_2)\in\Pi} \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)] \geq \mathfrak{C}_{X_1,X_2}\left[\max_{p(x_1,x_2)\in\Pi} t_\lambda(X_1, X_2)\right]. \quad (25)$$

(ii) On the other hand, for all $p(x_1, x_2) \in \Pi(p_1, p_2)$, we have $t_\lambda(X_1, X_2) \leq \max_{p(x_1,x_2)\in\Pi(p_1,p_2)} t_\lambda(X_1, X_2)$, implying immediately that

$$\mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)] \leq \mathfrak{C}_{X_1,X_2}\left[\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} t_\lambda(X_1, X_2)\right],$$

where the envelope on the right-hand side is evaluated with respect to pairs of distributions p_1, p_2, while the envelope on the left-hand side is evaluated with respect to joint distributions.

Since this holds for every $p(x_1, x_2) \in \Pi(p_1, p_2)$ we obtain the reverse direction that

$$\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} \mathfrak{C}_{X_1,X_2}[t_\lambda(X_1, X_2)] \leq \mathfrak{C}_{X_1,X_2}\left[\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} t_\lambda(X_1, X_2)\right]. \quad (26)$$

Equations (25) and (26) complete the proof of equation (24).

Equation (24) shows that (23) holds if and only if

$$\begin{aligned}&\mathfrak{C}_{X_1}[t_\lambda(X_1)] + \mathfrak{C}_{X_2}[t_\lambda(X_2)] \\ &< \mathfrak{C}_{X_1,X_2}\left[\max_{p(x_1,x_2)\in\Pi(p_1,p_2)} t_\lambda(X_1, X_2)\right],\end{aligned} \quad (27)$$

for some $p_1(x_1)$ and $p_2(x_2)$. Since both sides are functions defined on pairs of distributions (p_1, p_2), using equation (11) the statement of Lemma 2 is a reformulation of the inequality (27) in terms of their duals. This is because the right hand side of the expression in the lemma can be rewritten as

$$\max_{p_1(x_1),p_2(x_2)} \left\{ \left\{ \max_{p(x_1,x_2)\in\Pi} t_\lambda(X_1, X_2) \right\} - \sum_{x_1} p_1(x_1) d_1(x_1) - \sum_{x_2} p_2(x_2) d_2(x_2) \right\}.$$

□

Therefore, the initial target of our numerical simulations were to find product channels $p_1(y_1, z_1|x_1) \otimes p_2(y_2, z_2|x_2)$ and functions $d_1(x_1)$ and $d_2(x_2)$ such that the inequality in Lemma 2 held.

[3]To see this, take arbitrary x_0, x_1 and $\lambda \in [0, 1]$. For every $\epsilon > 0$, one can find y_0 and y_1 such that $g(x_i) \leq f(x_i, y_i) + \epsilon$ for $i = 0, 1$. Then, $\lambda g(x_1) + \bar{\lambda} g(x_0) \leq \lambda f(x_1, y_1) + \bar{\lambda} f(x_0, y_0) + \epsilon \leq f(x_\lambda, y_\lambda) + \epsilon$ by the joint concavity of f, where $x_\lambda = \lambda x_1 + \bar{\lambda} x_0$ and $y_\lambda = \lambda y_1 + \bar{\lambda} y_0$. From here, we conclude that $\lambda g(x_1) + \bar{\lambda} g(x_0) \leq g(x_\lambda) + \epsilon$. We obtain concavity of g by letting ϵ converge to zero.

- Proposition 1 implies that we can replace t_λ by (the cardinality-constrained) $\hat{t}_\lambda$ for testing the inequality in Lemma 2. Therefore when $p_1(y_1, z_1|x_1) \otimes p_2(y_2, z_2|x_2)$ has binary-input broadcast channels factors, the expression in Lemma 2, after expanding $\hat{t}_\lambda$ on the left-hand side, reduces to determining values of α, β such that

$$
\begin{aligned}
&\max_{p_1(x_1)} \Big\{ -(\delta_1 + \lambda(\delta_0+\gamma_2))I(X_1;Y_1) \\
&\qquad -(1-\lambda)(\delta_0+\gamma_2)I(X_1;Z_1) \\
&\qquad + \max\{(\delta_1+\gamma_2)I(X_1;Y_1), \gamma_2 I(X_1;Z_1)\} \\
&\qquad - \alpha p_1(X_1=0)\Big\} \\
&+ \max_{p_2(x_2)} \Big\{ -(\delta_1 + \lambda(\delta_0+\gamma_2))I(X_2;Y_2) \\
&\qquad -(1-\lambda)(\delta_0+\gamma_2)I(X_2;Z_2) \\
&\qquad + \max\{(\delta_1+\gamma_2)I(X_2;Y_2), \gamma_2 I(X_2;Z_2)\} \\
&\qquad - \beta p_2(X_2=0)\Big\} \\
&< \max_{p(x_1,x_2)} \big\{\hat{t}_\lambda(X_1,X_2) - \alpha p_1(X_1=0) \\
&\qquad\qquad -\beta p_2(X_2=0)\big\}. \qquad (28)
\end{aligned}
$$

Let us fix the broadcast channels $p_1(y_1, z_1|x_1), p_2(y_2, z_2|x_2)$ and the parameters $\gamma_2, \delta_0, \delta_1, \alpha, \beta$. Then, each of the two terms of the left-hand side of (28) is a continuous function of a single variable, and this is easy to optimize with high numerical accuracy (even with the previously known bounds on the alphabet of the auxiliary random variables, this part is only mildly more cumbersome and was clearly feasible). To compute the right-hand side, from Proposition 1 we need to compute

$$
\begin{aligned}
\max_{p(u,v,x_1,x_2)} &-(\delta_1 + \lambda(\delta_0+\gamma_2))I(X_1,X_2;Y_1,Y_2) \\
&-(1-\lambda)(\delta_0+\gamma_2)I(X_1,X_2;Z_1,Z_2) \\
&+(\delta_1+\gamma_2)I(U;Y_1,Y_2) + \gamma_2 I(V;Z_1,Z_2) \\
&-\gamma_2 I(U;V)
\end{aligned}
$$

over U and V such that $|\mathcal{U}| + |\mathcal{V}| \le |\mathcal{X}_1||\mathcal{X}_2| + 1$ and X_1, X_2 is a function of (U, V). For binary X_1, X_2 this implies that it suffices to consider $|\mathcal{U}| + |\mathcal{V}| \le 5$.
Remark 9. In contrast, the previous bound implied that it was sufficient to consider $|\mathcal{U}| \le 4$ and $|\mathcal{V}| \le 4$. This is a 15 dimensional non-convex optimization problem. Solving this optimization problem is infeasible by today's computing power considering the space of distributions of $p(u, v)$ we are left to deal with. This is the case even with the starting points of local optimization chosen from a very coarse grid of size, say 0.1 (here we even ignore the complexity of determining the maximizing function of X_1, X_2 over U, V).
The case of $|\mathcal{U}| = 4, |\mathcal{V}| = 1$ and $|\mathcal{U}| = 1, |\mathcal{V}| = 4$ are simple cases, as they correspond to $U = (X_1, X_2)$, $V =$ constant and $U =$ constant, $V = (X_1, X_2)$ respectively and it is immediate that there will be no instances satisfying Lemma 2. Therefore the main nontrivial cases are $|\mathcal{U}| = 2, |\mathcal{V}| = 3$ and $|\mathcal{U}| = 3, |\mathcal{V}| = 2$.
- Note that to show that there are no instances satisfying Lemma 2 we need to exhaustively compute the maximum over $p(u, v, x)$, where X is a function of (U, V) and, say, $|\mathcal{U}| = 3, |\mathcal{V}| = 2$. For each possible $p(u, v)$ naively there are 4^6 potential functional mappings.
However, from earlier results, we know that certain functions $x(u, v)$ cannot arise as maximizers, or, in some cases, if they do occur as maximizers then we know that they cannot satisfy Lemma 2. The study of such extremal functions started in [4], [14] and was continued in [10]. Here we recap the results from these papers, and we revert back to the notation where we have a single broadcast channel $p(y, z|x)$ rather than a product channel.
  - To compute the maximum, Theorem 1 in [10] (a generalization of the result regarding the infeasiblity of the XOR pattern as a maximizer [4]) imposes restrictions on the functional mappings that drastically reduce the number of mappings that need to be considered.
  - To compute the maximum, it suffices to consider mappings for which there do not exist symbols $u_0 \in \mathcal{U}$, $v_0 \in \mathcal{X}$ and $x_0 \in \mathcal{X}$ such that $x(u_0, v) = x_0 = x(u, v_0)$ for all u, v. For binary X this was called the AND mapping [4] and this result was generalized to arbitrarily sized $|\mathcal{X}|$ in [10]. Alternately, we can recover this result as an immediate consequence of Proposition 3. To see this, define $\omega_1(u) = 1 = \omega_2(v)$ as long as $u \ne u_0$ and $v \ne v_0$. Define $\omega_1(u_0) = \frac{p(u_0)-1}{p(u_0)}$ and $\omega_2(v_0) = \frac{p(v_0)-1}{p(v_0)}$. Thus clearly $\sum_u p(u)\omega_1(u) = 0 = \sum_v p(v)\omega_2(v)$. Further note that

$$
\begin{aligned}
&\sum_{u,v} p(u,v,x)(\omega_1(u) - \omega_2(v)) \\
&= \sum_{u=u_0 \text{ or } v=v_0} p(u,v,x)(\omega_1(u) - \omega_2(v)), \qquad \forall x.
\end{aligned}
$$

For any $x \ne x_0$, clearly $p(u, v, x) = 0$ when $u = u_0$ or $v = v_0$ (from the mapping). Finally, for $x = x_0$ we have

$$
\begin{aligned}
&\sum_{u=u_0 \text{ or } v=v_0} p(u,v,x_0)(\omega_1(u) - \omega_2(v)) \\
&= \sum_{u=u_0, v\ne v_0} p(u_0,v,x_0)(\omega_1(u_0) - 1) \\
&\quad + \sum_{u\ne u_0, v=v_0} p(u,v_0,x_0)(1 - \omega_2(v_0)) \\
&\quad + p(u_0,v_0,x_0)(\omega_1(u_0) - \omega_2(v_0)) \\
&= p(u_0)(\omega_1(u_0) - 1) + p(v_0)(1 - \omega_2(v_0)) = 0.
\end{aligned}
$$

Thus from Proposition 3, we see that it suffices to consider mappings such that there do not exist u_0, v_0 and x_0 such that $x(u_0, v) = x_0 = x(u, v_0)$ for all u, v.

These two results (especially the first) help to greatly reduce the number of functional mappings.

Remark 10. The following points are worth mentioning.

$$\begin{aligned} H_{p_\epsilon}(Y|Z) &= -\sum_{(y,z):p(y,z)>0} p(y,z)(1+\epsilon\ \mathrm{E}[L|Y=y,Z=z])\log\frac{p(y,z)(1+\epsilon\ \mathrm{E}[(L|Y=y,Z=z])}{p(z)(1+\epsilon\ \mathrm{E}[L|Z=z])} \\ &\overset{(a)}{=} -\sum_{(y,z):p(y,z)>0} p(y,z)(1+\epsilon\ \mathrm{E}[L|Y=y,Z=z])\log\frac{p(y,z)(1+\epsilon\ \mathrm{E}[L|Z=z])}{p(z)(1+\epsilon\ \mathrm{E}[L|Z=z])} \\ &= -\sum_{(y,z):p(y,z)>0} p(y,z)(1+\epsilon\ \mathrm{E}[L|Y=y,Z=z])\log p(y|z). \end{aligned} \tag{29}$$

- With all these results in the background, we performed extensive simulations trying to find an instance satisfying the inequality in Lemma 2. At least when all outputs of both broadcast channels are binary valued, we believe that we performed a rather exhaustive and intensive search, yet we were unable to find any such instance. We also generated some higher cardinality settings at random (especially when X_1 is ternary and X_2 is binary), and we were still unable to find any instance satisfying the inequality in Lemma 2.
- However, since this is a rather central problem in network information theory, we would also urge other interested researchers to independently confirm our numerical observations. We hope that the theoretical results mentioned about restricting the class of extremal functional mappings will provide useful to this end. Further note that Proposition 3 has implications beyond bounding cardinalities, as illustrated above.
- Finally, another interesting future work is suggested by our observation that when numerically optimizing the $|\mathcal{U}|=3$, $|\mathcal{V}|=2$ case in the statement of Lemma 2, only binary U and V showed up as the actual optimizers. Further the only distributions and functional mappings that arose even as local maximizers were the product distributions of the single-letter optimizers. All the results we have obtained regarding extremal distributions are for a generic broadcast channel and we have not been able to make use of the additional channel structure imposed by product channels to further restrict the extremal distributions or maximizing patterns. If further numerical simulations confirm similar behavior for product channels, then perhaps such a study is warranted.

Appendix

Lemma 3. *Consider three random variables X, Y, Z and a perturbation of their joint distribution defined according to*

$$p_\epsilon(x,y,z) = p(x,y,z)(1+\epsilon\ L(x,y,z)),$$

where $\sum_{x,y,z} p(x,y,z)L(x,y,z)=0$ and $|\epsilon|$ is small enough that $1+\epsilon L(x,y,z)\geq 0$ for all (x,y,z). Then

$$\frac{d^2}{d\epsilon^2}\left[H(Y|Z)\right]_{\epsilon=0} = 0$$

implies that $H_{p_\epsilon}(Y|Z)$ is linear in ϵ over the interval of ϵ where $p_\epsilon(x,y,z)$ is defined.

Proof of Lemma 3. Routine calculations yield that

$$\frac{d^2}{d\epsilon^2}\left[H(Y|Z)\right]_{\epsilon=0} = -\mathrm{E}[\mathrm{E}[L|Y,Z]^2] + \mathrm{E}[\mathrm{E}[L|Z]^2]$$

and hence the second derivative being zero at $\epsilon=0$ implies that

$$\mathrm{E}[\mathrm{E}[L|Y,Z]^2] - \mathrm{E}[\mathrm{E}[L|Z]^2] = 0.$$

Since $\mathrm{E}[L|Z] = \mathrm{E}[\mathrm{E}[L|Y,Z]|Z]$ the above equality can be written as $\mathrm{E}\left[(\mathrm{E}[L|Y,Z]-\mathrm{E}[L|Z])^2\right]=0$. Thus $\mathrm{E}[L|Y=y,Z=z]=\mathrm{E}[L|Z=z]$ whenever $p(y,z)>0$.

Observe that $H_{p_\epsilon}(Y|Z)$ is equal to the expression given in (29) on the top of this page where step (a) is justified because $\mathrm{E}[L|Y=y,Z=z]=\mathrm{E}[L|Z=z]$ whenever $p(y,z)>0$. Thus $H_{p_\epsilon}(Y|Z)$ is linear in ϵ. □

We will produce a counterexample to the following statement: for $|\mathcal{X}|=2$ the following inequality

$$\alpha I(U;Y) + I(V;Z) - I(U;V) \overset{?}{\leq} \max\{\alpha I(X;Y), I(X;Z)\}, \tag{30}$$

holds for any $\alpha>1$ and any Markov chain $(U,V)\to X\to(Y,Z)$.

Consider the following setting. The channels are:

$$p(Y|X) = \begin{bmatrix}0.5 & 0.5\\ 0 & 1\end{bmatrix}, \quad p(Z|X) = \begin{bmatrix}1 & 0\\ 0.1 & 0.9\end{bmatrix}.$$

The parameters are

$$p(X) = [0.8,\ 0.2], \quad \alpha = \frac{I(X;Z)}{I(X;Y)} = 3.429517.$$

The choice of α is actually the corner point for the right-hand side of (30). The result is

$$\begin{aligned} \alpha I(U;Y) + I(V;Z) - I(U;V) &= 0.593020 \\ &> 0.586278 \\ &= \max\{\alpha I(X;Y), I(X;Z)\}, \end{aligned}$$

where we use the probability mass function and mapping $X=f(U,V)$ given by

$$p(U,V) = \begin{bmatrix}0.05930 & 0.00005\\ 0.14065 & 0.80000\end{bmatrix}, \quad f(U,V) = \begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}.$$

Acknowledgment: The authors wish to thank Yanlin Geng for obtaining this counterexample.

REFERENCES

[1] T. Cover, *Broadcast channels*, IEEE Transactions on Information Theory **18** (1972), no. 1, 2–14.

[2] Abbas El Gamal and Young-Han Kim, *Network information theory*, Cambridge University Press, 2012.

[3] Y. Geng, A. Gohari, C. Nair, and Y. Yu, *On Marton's inner bound and its optimality for classes of product broadcast channels*, IEEE Transactions on Information Theory **60** (2014), no. 1, 22–41.

[4] Y Geng, V Jog, C Nair, and Z V Wang, *An information inequality and evaluation of marton's inner bound for binary input broadcast channel*, IEEE Trans. Info. Theory (accepted).

[5] Y. Geng and C. Nair, *The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages*, IEEE Transactions on Information Theory **60** (2014), no. 4, 2087–2104.

[6] Yanlin Geng, Varun Jog, Chandra Nair, and Zizhou Vincent Wang, *An information inequality and evaluation of Marton's inner bound for binary input broadcast channels*, IEEE Transactions on Information Theory **59** (2013), no. 7, 4095–4105.

[7] A. Gohari, C. Nair, and V. Anantharam, *On Marton's inner bound for broadcast channels*, 2012 IEEE International Symposium on Information Theory, July 2012, pp. 581–585.

[8] A. A. Gohari and V. Anantharam, *Evaluation of Marton's inner bound for the general broadcast channel*, 2009 IEEE International Symposium on Information Theory, June 2009, pp. 2462–2466.

[9] ——, *Evaluation of Marton's inner bound for the general broadcast channel*, IEEE Transactions on Information Theory **58** (2012), no. 2, 608–619.

[10] A. A. Gohari, A. El Gamal, and V. Anantharam, *On an outer bound and an inner bound for the general broadcast channel*, 2010 IEEE International Symposium on Information Theory, June 2010, pp. 540–544.

[11] 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.

[12] K. Marton, *A coding theorem for the discrete memoryless broadcast channel*, IEEE Transactions on Information Theory **25** (1979), no. 3, 306–311.

[13] C. Nair, *Capacity regions of two new classes of two-receiver broadcast channels*, IEEE Transactions on Information Theory **56** (2010), no. 9, 4207–4214.

[14] C. Nair, Z. V. Wang, and Y. Geng, *An information inequality and evaluation of marton's inner bound for binary input broadcast channels*, 2010 IEEE International Symposium on Information Theory, June 2010, pp. 550–554.

[15] C. Nair, L. Xia, and M. Yazdanpanah, *Sub-optimality of Han-Kobayashi achievable region for interference channels*, 2015 IEEE International Symposium on Information Theory (ISIT), June 2015, pp. 2416–2420.

[16] Chandra Nair, *Upper concave envelopes and auxiliary random variables*, International Journal of Advances in Engineering Sciences and Applied Mathematics **5** (2013), no. 1, 12–20 (English).

[17] R.T. Rockafellar, *Convex analysis*, Convex Analysis, Princeton University Press, 1997.

[18] F Terkelsen, *Some minimax theorems*, Mathematica Scandinavica **31** (1972), 405–413.

[19] H. Weingarten, Y. Steinberg, and S. S. Shamai, *The capacity region of the Gaussian multiple-input multiple-output broadcast channel*, IEEE Transactions on Information Theory **52** (2006), no. 9, 3936–3964.

Venkat Anantharam is on the faculty of the EECS department at the University of California at Berkeley, where he has been since 1994. From 1986 to 1994 he was on the faculty of the School of Electrical Engineering at Cornell University. He received the M.S. (1982) and PhD. (1986) degrees in Electrical Engineering and the M. A. (1983) and C. Phil. (1984) degrees in Mathematics from the University of California at Berkeley. He received the B. Tech. degree in Electrical Engineering (Electronics) from the Indian Institute of Technology Madras in 1980.

Amin Gohari is an Associate Professor at Sharif University of Technology, Tehran, Iran. Dr. Gohari received his M.Sc. and Ph.D. degrees in electrical engineering in 2010 from the University of California, Berkeley, and his B.Sc. degree in 2004 from Sharif University of Technology, Iran. He received the 2010 Eli Jury Award from UC Berkeley, Department of Electrical Engineering, for "outstanding achievement in the area of communication networks," and the 2009-2010 Bernard Friedman Memorial Prize in Applied Mathematics from UC Berkeley, Department of Mathematics, for "demonstrated ability to do research in applied mathematics." He also received the Gold Medal from the 41st International Mathematical Olympiad (IMO 2000) and the First Prize from the 9th International Mathematical Competition for University Students (IMC 2002). He is a co-author of a paper that received the 2013 Jack Keil Wolf ISIT Student Paper Award. He was selected as an exemplary reviewer for Transactions on Communications in 2016 and 2017. He is currently an associate editor for the IEEE Transactions on Information Theory.

Chandra Nair is an Associate Professor with the Information Engineering department at The Chinese University of Hong Kong, where he also serves as the Programme Director of the undergraduate program on Mathematics and Information Engineering and as the Director of the Institute of Theoretical Computer Science and Communications. He was ranked first in the Indian National Mathematics Olympiad (INMO 1994) and completed a four-year nurture programme in mathematics from the Institute of Mathematical Sciences in Madras, India. He obtained his degrees in Electrical Engineering: B.Tech degree from IIT Madras (India) in 1999, and his M.S. and Ph.D. degrees from Stanford University in 2002 and 2005 respectively. He was a Stanford graduate fellow (00-04) and a Microsoft graduate fellow (04-05).