ON MARTON'S ACHIEVABLE REGION: LOCAL TENSORIZATION FOR PRODUCT CHANNELS WITH A BINARY COMPONENT

CHANDRA NAIR

Abstract. We show that Marton's achievable rate region for product broadcast channels with one binary component satisfies a property called *local tensorization*. If a corresponding *global tensorization* property held for the same setting, then this would be equivalent to showing the optimality of Marton's achievable region for any two receiver broadcast channel with binary inputs.

1. Introduction

Network information theory studies the feasibility of transmitting multiple sources reliably over a network of users to their intended receivers. The point-to-point setting was considered by Shannon in his seminal paper [1]. In this setting, Shannon demonstrated that it was possible to do a source-channel separation, i.e. to decouple the compression of source to remove redundancy, and to expand the compressed source to combat channel errors. While such a source-channel separation usually does not hold in a multi-user setting, it is nevertheless interesting, both from a mathematical as well as an engineering perspective, to decouple the two problems. In the latter problem, called the channel-coding setting, determining a computable characterization of the capacity region for several fundamental scenarios remain open (see open problems in [2]).

For some of these settings, there are natural achievable regions whose optimality has not been determined. For certain other settings, for instance the capacity region of the interference channel and a three receiver broadcast channel with degraded message sets (see open problems 6.4 and 8.2 in [2]), the author and his collaborators have shown that a natural achievable rate region is strictly sub-optimal (see [3, 4]). The sub-optimality was demonstrated, in both instances, by taking a particular channel instantiation where the two-letter extension of the achievable region strictly outperformed the single-letter region.

However there are several problems for which the optimality of natural achievable regions are unknown. Here is a partial list, taken essentially from among the open problems listed in [2], whose answer is not yet determined:

- 5.1 *Is superposition coding region optimal for less-noisy broadcast channels with four or more receivers?*
- 6.1 *Is the Han–Kobayashi scheme with Gaussian signaling tight for the two user scalar Gaussian Interference channel with weak interference?*
- 8.3 *Does the Marton inner bound achieve the sum-capacity of the binary skew-symmetric broadcast channel?*
- 8.4 *Is the Marton inner bound tight in general for broadcast channels?*

To show the sub-optimality (or optimality), in each of the instance above as well as other similar instances, one can reduce the question to showing the additivity (or *tensorization*) of an associated functional. This article deals with such a question for the two-receiver broadcast channel.

1.1. **Two receiver broadcast channel.** A two receiver broadcast channel models communication from a single sender X to two receivers, say Y and Z, as shown in Figure 1. An (n, R_1, R_2)-code, C, for this setting consists of an encoder that maps $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \mapsto \mathcal{X}^n$, and two decoders that map received sequences $\mathcal{Y}^n \mapsto [1 : 2^{nR_1}]$ and $\mathcal{Z}^n \mapsto [1 : 2^{nR_2}]$, respectively. The probability of error for a given code, $P_{(e)}^{\mathsf{C}}$, is defined to be $\mathrm{P}((M_1, M_2) \neq (\hat{M}_1, \hat{M}_2))$ when (M_1, M_2) is uniformly distributed over $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. A non-negative rate pair (R_1, R_2) is said to be achievable if there exists a sequence (in blocklength n) of (n, R_1, R_2)-codes C_n such that $P_{(e)}^{\mathsf{C}_n} \to 0$ as $n \to \infty$. The closure of the set of achievable rate pairs is called the capacity region for the broadcast channel $(W_a(y|x), W_b(z|x))$.

Date: February 7, 2020.

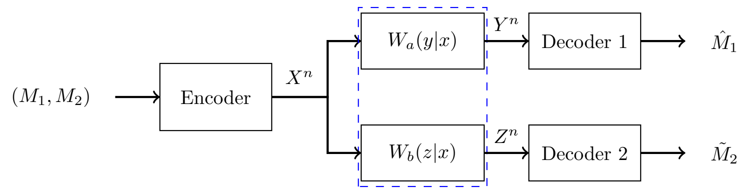

FIGURE 1. Discrete memoryless broadcast channel (no common message)

Theorem 1.1 (Marton's inner bound, [5])**.** *A non-negative rate pair (R_1, R_2) is achievable if it satisfies*

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

for some $p(q, u, v, x)$ such that $(Q, U, V) \to X \to (Y, Z)$ is Markov.

If we denote this region as $Mc(W_a, W_b)$, then it is easy to see that Marton's region is the capacity region for any broadcast channel *if and only if*

$$Mc(W_a \otimes W_a, W_b \otimes W_b) = Mc(W_a, W_b) \oplus Mc(W_a, W_b), \quad \forall W_a, W_b, \tag{1}$$

where $\oplus$ denotes the Minkowski sum of the two regions.

Remark 1.2. *This follows from the fact that n-letter extension of Marton's region tends to capacity from which the above condition for optimality is rather immediate.*

Let (W_a, W_b) and $(\hat{W}_a, \hat{W}_b)$ be two (not necessarily identical) broadcast channels from $X_1 \to (Y_1, Z_1)$ and $X_2 \to (Y_2, Z_2)$ respectively. For $\lambda \geq 1$ and $\alpha \in [0, 1]$, define the following:

$$\begin{aligned} F_1^{(\lambda,\alpha)}(\eta) &:= \max_{p_{UVX_1}} -(\lambda - \alpha)I(X_1; Y_1) - \alpha I(X_1; Z_1) + \lambda I(U; Y_1) + I(V; Z_1) - I(U; V) - \mathrm{E}(\eta(X_1)) \\ F_2^{(\lambda,\alpha)}(\zeta) &:= \max_{p_{UVX_2}} -(\lambda - \alpha)I(X_1; Y_1) - \alpha I(X_1; Z_1) + \lambda I(U; Y_1) + I(V; Z_1) - I(U; V) - \mathrm{E}(\zeta(X_2)) \\ F_{12}^{(\lambda,\alpha)}(\eta, \zeta) &:= \max_{p_{UVX_1X_2}} -(\lambda - \alpha)I(X_1X_2; Y_1Y_2) - \alpha I(X_1X_2; Z_1Z_2) + \lambda I(U; Y_1Y_2) + I(V; Z_1Z_2) \\ &\qquad - I(U; V) - \mathrm{E}(\eta(X_1)) - \mathrm{E}(\zeta(X_2)), \end{aligned}$$

where $\eta(X_1)$ and $\zeta(X_2)$ are any two functions. In [6] it was shown that one can restrict to the set of simplices satisfying $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}_1| + 1$ to compute the first maximum, $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}_2| + 1$ to compute the second maximum, and $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}_1||\mathcal{X}_2| + 1$ to compute the third maximum.

It follows from Lemma 2 in [6] that (1) would follow if we show the following statement:

$$F_{12}^{(\lambda,\alpha)}(\eta, \zeta) = F_1^{(\lambda,\alpha)}(\eta) + F_2^{(\lambda,\alpha)}(\zeta) \quad \forall \eta, \zeta. \tag{2}$$

Remark 1.3. *This implication follows from a dual representation of an upper concave envelope which follows via Fenchel duality. This statement is slightly more general than is needed to establish* (1)*. On the other hand such a statement with non-identical components will be useful as can be seen later in this article.*

Consider the following functionals defined by:

$$\begin{aligned} F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1}) &:= -(\lambda - \alpha)I(X_1; Y_1) - \alpha I(X_1; Z_1) + \lambda I(U; Y_1) + I(V; Z_1) - I(U; V) - \mathrm{E}(\eta(X_1)) \\ F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2}) &:= -(\lambda - \alpha)I(X_1; Y_1) - \alpha I(X_1; Z_1) + \lambda I(U; Y_1) + I(V; Z_1) - I(U; V) - \mathrm{E}(\zeta(X_2)) \\ F_{12}^{(\lambda,\alpha,\eta,\zeta)}(p_{UVX_1X_2}) &:= -(\lambda - \alpha)I(X_1X_2; Y_1Y_2) - \alpha I(X_1X_2; Z_1Z_2) + \lambda I(U; Y_1Y_2) + I(V; Z_1Z_2) \\ &\qquad - I(U; V) - \mathrm{E}(\eta(X_1)) - \mathrm{E}(\zeta(X_2)), \end{aligned}$$

Note that the equality in (2) is equivalent to requiring that if $p^*_{U_1V_1X_1}$ is a maximizer of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ and $p^\dagger_{U_2V_2X_2}$ is a maximizer of $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$ then $p^*_{U_1V_1X_1} \otimes p^\dagger_{U_2V_2X_2}$, with $U = (U_1, U_2)$ and $V = (V_1, V_2)$, is a maximizer of $F_{12}^{(\lambda,\alpha,\eta,\zeta)}(p_{UVX_1X_2})$. In other words, the product of the maximizing distributions in the marginal spaces is a maximizer of the joint distribution. This leads us to the following definition.

Definition 1.4. *We say that Márton's achievable region satisfies a global tensorization property if* (2) *holds for all functions* η, ζ.

As said earlier, since establishing the *global tensorization property* would establish the optimality of Marton's achievable region, one step towards verifying this property would be to verify if the distribution $p^*_{U_1V_1X_1} \otimes p^\dagger_{U_2V_2X_2}$ is a *local-maximizer* of $F_{12}^{(\lambda,\alpha,\eta,\zeta)}(p_{UVX_1X_2})$. An issue is that one has to identify the maximizers of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ and $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$. Alternately, one could hypothesize a more general statement which is called the *local tensorization property.*

Definition 1.5. *We say that Márton's achievable region satisfies a local tensorization property if: let* $p^*_{U_1V_1X_1}$ *be a local maximizer of* $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ *and* $p^\dagger_{U_2V_2X_2}$ *be a local maximizer of* $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$*, then we require that* $p^*_{U_1V_1X_1} \otimes p^\dagger_{U_2V_2X_2}$*, with* $U = (U_1, U_2)$ *and* $V = (V_1, V_2)$*, be a local maximizer of* $F_{12}^{(\lambda,\alpha,\eta,\zeta)}(p_{UVX_1X_2})$ *for all* η, ζ.

Remark 1.6. *The following points are worth noting:*

- *The local tensorization property can be tested using the local behaviour of a function, for instance using first and second order conditions, and there is no need for the identification of global maximizers of the non-convex functional in the marginal spaces.*
- *In all the network information theory settings that we have studied so far, it appears that for an associated functional, similar to the one considered here, either both the local tensorization property and the global tensorization property holds, or neither holds.*

2. The Local Tensorization Property

Below, we list some properties of local maximizers of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ that has been esyablished in previous works (see [6] and references there in).

Proposition 2.1. *Let* $p_{U_1V_1X_1}$ *be a local maximizer of* $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$*. Then we can assume that*

- $|\mathcal{U}_1| + |\mathcal{V}_1| \le |\mathcal{X}_1| + 1$
- X_1 *is a function of* (U_1, V_1)*, i.e.* $H(X_1|U_1V_1) = 0$.

First order conditions. Let $p_{U_1V_1X_1}$ be a local maximizer of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$, w.l.o.g. satisfying the properties in Proposition 2.1. We have the following first order conditions for local optimality that

$$\sum_{y_1,z_1} W_a(y_1|x_1)W_b(z_1|x_1)\log\left(\frac{p^\lambda(x_1)p^\lambda(u_1y_1)p(v_1z_1)e^{-\eta(x_1)}}{p^\alpha(y_1)p^{1-\alpha}(z_1)p^{\lambda-\alpha}(x_1y_1)p^\alpha(x_1z_1)p^{\lambda-1}(u_1)p(u_1v_1)}\right)\begin{cases} = C & p(u_1,v_1,x_1) > 0 \\ \le C & p(u_1,v_1,x_1) = 0 \end{cases}.$$

It is also immediate that $C = F_1^{(\lambda,\alpha)}(\eta)$. A similar first order conditions for local optimality also holds for the local maximizer of $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$. Therefore, it follows by summing these conditions that if $p^*_{U_1V_1X_1}$ is a local maximizer of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ and $p^\dagger_{U_2V_2X_2}$ is a local maximizer of $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$, then $p^*_{U_1V_1X_1} \otimes p^\dagger_{U_2V_2X_2}$, with $U = (U_1, U_2)$ and $V = (V_1, V_2)$, will also satisfy the first order conditions for local optimality.

Remark 2.2. *The above observation is generically true for all such functionals arising in the test of optimality of achievable regions. It is the second order conditions listed below that are seen to fail for instances where the achievable regions are sub-optimal.*

Second order conditions. Let $p_{U_1V_1X_1}$ be a local maximizer of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$. Let $L(u_1, v_1, x_1)$ be a function such that $E(L) = 0$. Then $p^\epsilon_{U_1,V_1,X_1} = p_{U_1V_1X_1}(1 + \epsilon L)$ defines a perturbation in the space of probability distributions, for ϵ small enough. Let us define

$$f_L(\epsilon) := F_1^{(\lambda,\alpha,\eta)}(p^\epsilon_{U_1V_1X_1}).$$

From Proposition 2.1 we know that it suffices to consider X_1 to be a function of U_1, V_1. Hence we can w.l.o.g. assume that the perturbation L is a function of (U_1, V_1) only for testing second order conditions.

Lemma 2.3. *Let* $p^*_{U_1,V_1,X_1}$ *be a maximizer of* $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$*. Then for any* $L(U_1, V_1)$ *we have*

$$\mathrm{E}(L^2) + (\lambda - 1)\,\mathrm{E}(\mathrm{E}(L|U_1)^2) - \lambda\,\mathrm{E}(\mathrm{E}(L|U_1, Y_1)^2) \ge 0.$$

Proof. Define a new distribution with $\tilde{U}_1 = (U_1, Q)$, where Q is binary and

$$p^{(\epsilon)}((u_1, q), v_1, x_1) = \begin{cases} \frac{1}{2} p^*(u_1, v_1, x_1)(1 + \epsilon L(u_1, v_1)) & q = 0 \\ \frac{1}{2} p^*(u_1, v_1, x_1)(1 - \epsilon L(u_1, v_1)) & q = 1 \end{cases} .$$

Hence $p^{(\epsilon)}(v_1, x_1) = p^*(v_1, x_1)$. Therefore

$$\begin{aligned} F_1^{(\lambda,\alpha,\eta)}(p^{\epsilon}_{\tilde{U}_1 V_1 X_1}) - F_1^{(\lambda,\alpha,\eta)}(p_{U_1 V_1 X_1}) &= (H_{p^{(\epsilon)}}(\tilde{U}_1, V_1) - H_{p^*}(U_1, V_1)) + (\lambda - 1)(H_{p^{(\epsilon)}}(\tilde{U}_1) - H_{p^*}(U_1)) \\ &\quad - \lambda(H_{p^{(\epsilon)}}(\tilde{U}_1, Y_1) - H_{p^*}(U_1, V_1)) \\ &= \frac{\epsilon^2}{2} \left(\lambda \operatorname{E}(\operatorname{E}(L|U_1, Y_1)^2) - (\lambda - 1) \operatorname{E}(\operatorname{E}(L|U_1)^2) - \operatorname{E}(L^2) \right) + O(\epsilon^3). \end{aligned}$$

The lemma then follows from the maximality of $p^*_{U_1, V_1, X_1}$. □

The second order conditions for $p_{U_1 V_1 X_1}$ to be a local maximizer is that $f_L''(0) \leq 0$ for all $L(U_1, V_1)$ such that $\operatorname{E}(L) = 0$. A bit of elementary calculations shows that this is equivalent to requiring that

$$\begin{aligned} &\lambda \operatorname{E}(\operatorname{E}(L|X_1)^2) + \lambda \operatorname{E}(\operatorname{E}(L|U_1 Y_1)^2) + \operatorname{E}(\operatorname{E}(L|V_1 Z_1)^2) - \alpha \operatorname{E}(\operatorname{E}(L|Y_1)^2) - (1 - \alpha) \operatorname{E}(\operatorname{E}(L|Z_1)^2) \\ &\quad - (\lambda - \alpha) \operatorname{E}(\operatorname{E}(L|X_1 Y_1)^2) - \alpha \operatorname{E}(\operatorname{E}(L|X_1 Z_1)^2) - (\lambda - 1) \operatorname{E}(\operatorname{E}(L|U_1)^2) - \operatorname{E}(\operatorname{E}(L|U_1 V_1)^2) \leq 0, \end{aligned} \tag{3}$$

for all $L(U_1, V_1)$ such that $\operatorname{E}(L) = 0$. Since $(U_1, V_1) \to X_1 \to (Y_1, Z_1)$ is Markov, we have $\operatorname{E}(L|X_1) = \operatorname{E}(L|X_1, Y_1) = \operatorname{E}(L|X_1, Z_1)$. Further any function $L(U_1, V_1)$ can be expressed as $L_1(U_1, V_1) + a$, with $\operatorname{E}(L_1) = 0$. Since the coefficients in (3) add to zero, the constant term a vanishes for a generic $L(U_1, V_1)$. Combining these three observations together, the second order conditions can be equivalently expressed as requiring

$$\begin{aligned} &\lambda \operatorname{E}(\operatorname{E}(L|U_1 Y_1)^2) + \operatorname{E}(\operatorname{E}(L|V_1 Z_1)^2) - \alpha \operatorname{E}(\operatorname{E}(L|Y_1)^2) - (1 - \alpha) \operatorname{E}(\operatorname{E}(L|Z_1)^2) \\ &\quad - (\lambda - 1) \operatorname{E}(\operatorname{E}(L|U_1)^2) - \operatorname{E}(\operatorname{E}(L|U_1 V_1)^2) \leq 0, \end{aligned} \tag{4}$$

for all $L(U_1, V_1)$. We now have a quadratic form which can be represented using the matrices of size $|\mathcal{U}_1||\mathcal{V}_1| \times |\mathcal{U}_1||\mathcal{V}_1|$.

For any $L(U_1, V_1)$ we can represent:

$$\begin{aligned} &\operatorname{E}(\operatorname{E}(L|U_1 V_1)^2) = L^T A_{U_1 V_1} L, \quad \operatorname{E}(\operatorname{E}(L|U_1 Y_1)^2) = L^T A_{U_1 Y_1} L, \quad \operatorname{E}(\operatorname{E}(L|V_1 Z_1)^2) = L^T A_{V_1 Z_1} L, \\ &\operatorname{E}(\operatorname{E}(L|Y_1)^2) = L^T A_{Y_1} L, \quad \operatorname{E}(\operatorname{E}(L|Z_1)^2) = L^T A_{Z_1} L, \operatorname{E}(\operatorname{E}(L|U_1)^2) = L^T A_{U_1} L \end{aligned}$$

where the matrices (of size $|\mathcal{U}_1||\mathcal{V}_1| \times |\mathcal{U}_1||\mathcal{V}_1|$) are defined by

- *(i)* $A_{U_1 V_1}$ is a diagonal matrix with entries $p(u_1, v_1)$
- *(ii)* $A_{U_1 Y_1}$ is a matrix with entries $\sum_{y_1} p(u_1, v_1, y_1) p(\hat{v}_1|u_1, y_1)$ at location $((u_1, v_1), (u_1, \hat{v}_1))$, and zeroes elsewhere
- *(iii)* $A_{V_1 Z_1}$ is a matrix with entries $\sum_{z_1} p(u_1, v_1, z_1) p(\hat{u}_1|v_1, z_1)$ at location $((u_1, v_1), (\hat{u}_1, v_1))$, and zeroes elsewhere
- *(iv)* A_{Y_1} is a matrix with entries $\sum_{y_1} p(u_1, v_1, y_1) p(\hat{u}_1, \hat{v}_1|y_1)$ at location $((u_1, v_1), (\hat{u}_1, \hat{v}_1))$
- *(v)* A_{Z_1} is a matrix with entries $\sum_{z_1} p(u_1, v_1, z_1) p(\hat{u}_1, \hat{v}_1|z_1)$ at location $((u_1, v_1), (\hat{u}_1, \hat{v}_1))$
- *(vi)* A_{U_1} is a matrix with entries $p(u_1, v_1) p(\hat{v}_1|u_1)$ at location $((u_1, v_1), (u_1, \hat{v}_1))$ and zeroes elsewhere.

The first two orderings are immediate from the definition in terms of their quadratic forms, and the third one, holds for a maximizer, follows from Lemma 2.3:

$$A_{U_1 V_1} \succeq A_{U_1 Y_1} \succeq A_{U_1}, A_{Y_1} \succeq 0 \tag{5a}$$

$$A_{U_1 V_1} \succeq A_{V_1 Z_1} \succeq A_{Z_1} \succeq 0 \tag{5b}$$

$$A_{U_1 V_1} + (\lambda - 1) A_{U_1} \succeq \lambda A_{U_1 Y_1} \tag{5c}$$

Remark 2.4. *Note apart from being positive semi-definite the matrices above also satisfy additional properties: all the entries are non-negative, and they are a convex combination of completely positive matrices.*

In terms of these matrices one can rewrite the second order conditions in (4) as

$$A_{U_1 V_1} + (\lambda - 1) A_{U_1} + \alpha A_{Y_1} + (1 - \alpha) A_{Z_1} \succeq \lambda A_{U_1, Y_1} + A_{V_1 Z_1}. \tag{6}$$

Suppose $p^*_{U_1V_1X_1}$ is a local maximizer of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ and $p^\dagger_{U_2V_2X_2}$ is a local maximizer of $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$, then we know that

$$A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} + \alpha A^*_{Y_1} + (1-\alpha)A^*_{Z_1} \succeq \lambda A^*_{U_1,Y_1} + A^*_{V_1Z_1} \tag{7}$$

$$A^\dagger_{U_2V_2} + (\lambda-1)A^\dagger_{U_2} + \alpha A^\dagger_{Y_2} + (1-\alpha)A^\dagger_{Z_2} \succeq \lambda A^\dagger_{U_2,Y_2} + A^\dagger_{V_2Z_2}. \tag{8}$$

from the previous section we know that $p^*_{U_1V_1X_1} \otimes p^\dagger_{U_2V_2X_2}$ satisfies the first order conditions for local optimality of $F_{12}^{(\lambda,\alpha,\eta,\zeta)}(p_{UVX_1X_2})$. To check if it satisfies the second order conditions, a bit of algebra shows that we have to verify if the following holds:

$$A^*_{U_1V_1}\otimes A^\dagger_{U_2V_2} + (\lambda-1)A^*_{U_1}\otimes A^\dagger_{U_2} + \alpha A^*_{Y_1}\otimes A^\dagger_{Y_2} + (1-\alpha)A^*_{Z_1}\otimes A^\dagger_{Z_2} \succeq \lambda A^*_{U_1,Y_1}\otimes A^\dagger_{U_2,Y_2} + A^*_{V_1Z_1}\otimes A^\dagger_{V_2Z_2}. \tag{9}$$

Remark 2.5. *There are positive semi-definite matrices A^* and $A^\dagger$, of size 2×2, that satisfy (7) and satisfy the ordering in (5), but fail to satisfy (9). Such an example, albeit with some negative entries in matrices, was discovered in collaboration with Venkat Anantharam, Amin Gohari, and Ali Yekhekhany. Therefore it seems essential that we need to use further properties of the matrices to establish the second order conditions.*

2.1. **Restriction to binary input broadcast channels.** Suppose we wish to seek the optimality of Marton's achievable region for broadcast channels where $|\mathcal{X}| = 2$, then it suffices to prove the *global-tensorization property* for product channels (as in (2)) where one of the component, say $\mathcal{X}_2$, is binary. This is because we can split an n-letter of a binary channel extension into the product of an $(n-1)$-letter extension and a binary component. Hence if the *global-tensorization property* holds when one of the components is binary, then we would demonstrate the optimality of Marton's achievable region for broadcast channels where $|\mathcal{X}| = 2$.

Therefore it is natural to ask if we can prove the *local-tensorization property* for product channels where one of the component, say $\mathcal{X}_2$, is binary. The following proposition is the main result of this article.

Proposition 2.6. *Consider a product broadcast channel where $\mathcal{X}_2$ is binary. Then for product of maximizers of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ and $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$, the second order conditionals for local optimality as stated in (9) are satisfied.*

Remark 2.7. *Since the first order conditions are more widely satisfied, it follows that it we have a product broadcast channel where $\mathcal{X}_2$ is binary, then for product of maximizers of $F_1^{(\lambda,\alpha,\eta)}(p_{UVX_1})$ and $F_2^{(\lambda,\alpha,\zeta)}(p_{UVX_2})$ is a local maximizer of $F_{12}^{(\lambda,\alpha,\eta,\zeta)}(p_{UVX_1X_2})$.*

Proof. From proposition 2.1 we know that at any local maximizer p_{U_2,V_2,X_2} we can assume that $|\mathcal{U}_2|+|\mathcal{V}_2| \le |\mathcal{X}_2|+1 = 3$. Hence it follows that either $(i) U_2 = X_2$ and V_2 is a constant; or $(i) V_2 = X_2$ and U_2 is a constant.

We treat the two cases separately below.

Case 1: At $p^\dagger_{U_2V_2X_2}$, we have $U_2 = X_2$ and V_2 is a constant. This, in particular, implies that

$$A^\dagger_{U_2,V_2} = A^\dagger_{U_2,Y_2} = A^\dagger_{U_2} = A^\dagger_{X_2},$$

where $A^\dagger_{X_2}$ is a 2×2 diagonal matrix with entries p_{X_2}. Further since V_2 is a constant we have $A^\dagger_{V_2,Z_2} = A^\dagger_{Z_2}$. Substituting the above, and using (8) obtain

$$A^\dagger_{X_2} \succeq A^\dagger_{Y_2} \succeq A^\dagger_{Z_2} \succeq 0. \tag{10}$$

Putting these together observe that

$$\begin{aligned}
&A^*_{U_1V_1}\otimes A^\dagger_{U_2V_2} + (\lambda-1)A^*_{U_1}\otimes A^\dagger_{U_2} + \alpha A^*_{Y_1}\otimes A^\dagger_{Y_2} + (1-\alpha)A^*_{Z_1}\otimes A^\dagger_{Z_2} - \lambda A^*_{U_1Y_1}\otimes A^\dagger_{U_2Y_2} - A^*_{V_1Z_1}\otimes A^\dagger_{V_2Z_2}\\
&\quad = A^*_{U_1V_1}\otimes A^\dagger_{X_2} + (\lambda-1)A^*_{U_1}\otimes A^\dagger_{X_2} + \alpha A^*_{Y_1}\otimes A^\dagger_{Y_2} + (1-\alpha)A^*_{Z_1}\otimes A^\dagger_{Z_2} - \lambda A^*_{U_1Y_1}\otimes A^\dagger_{X_2} - A^*_{V_1Z_1}\otimes A^\dagger_{Z_2}\\
&\quad = (A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} - \lambda A^*_{U_1Y_1})\otimes A^\dagger_{X_2} + \alpha A^*_{Y_1}\otimes A^\dagger_{Y_2} + (1-\alpha)A^*_{Z_1}\otimes A^\dagger_{Z_2} - A^*_{V_1Z_1}\otimes A^\dagger_{Z_2}\\
&\quad = (A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} - \lambda A^*_{U_1Y_1})\otimes (A^\dagger_{X_2} - A^\dagger_{Y_2}) + (A^*_{V_1Z_1} - (1-\alpha)A^*_{Z_1})\otimes (A^\dagger_{Y_2} - A^\dagger_{Z_2})\\
&\qquad + (A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} + \alpha A^*_{Y_1} + (1-\alpha)A^*_{Z_1} - \lambda A^*_{U_1,Y_1} - A^*_{V_1Z_1})\otimes A^\dagger_{Y_2}\\
&\quad \succeq 0.
\end{aligned}$$

The last inequality is a consequence of the fact that if $B, C \succeq 0$, then $B\otimes C \succeq 0$, and conditions (5), (6), and (10).

Case 2: At $p^{\dagger}_{U_2V_2X_2}$, we have $V_2 = X_2$ and U_2 is a constant. This, in particular, implies that

$$A^{\dagger}_{U_2,V_2} = A^{\dagger}_{V_2,Z_2} = A^{\dagger}_{V_2} = A^{\dagger}_{X_2} \succeq A^{\dagger}_{Z_2},$$

where $A^{\dagger}_{X_2}$ is a 2×2 diagonal matrix with entries p_{X_2}. Further since U_2 is a constant we have $A^{\dagger}_{U_2,Y_2} = A^{\dagger}_{Y_2} \succeq A^{\dagger}_{U_2}$.

Substituting the above, and using (8) obtain

$$(1-\alpha)A^{\dagger}_{Z_2} \succeq (\lambda-1)(A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) + (1-\alpha)A^{\dagger}_{Y_2} \succeq 0. \tag{11}$$

or equivalently $(1-\alpha)(A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) \succeq (\lambda-\alpha)(A^{\dagger}_{Y_2} - A^{\dagger}_{U_2})$. As $\lambda \geq 1$, the above observations also implies

$$A^{\dagger}_{X_2} \succeq A^{\dagger}_{Z_2} \succeq A^{\dagger}_{Y_2} \succeq A^{\dagger}_{U_2} \succeq 0.$$

Now, using the above, we have

$$\begin{aligned}
&A^*_{U_1V_1} \otimes A^{\dagger}_{U_2V_2} + (\lambda-1)A^*_{U_1} \otimes A^{\dagger}_{U_2} + \alpha A^*_{Y_1} \otimes A^{\dagger}_{Y_2} + (1-\alpha)A^*_{Z_1} \otimes A^{\dagger}_{Z_2} - \lambda A^*_{U_1Y_1} \otimes A^{\dagger}_{U_2Y_2} - A^*_{V_1Z_1} \otimes A^{\dagger}_{V_2Z_2} \\
&\quad = (A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes A^{\dagger}_{X_2} + (\alpha A^*_{Y_1} - \lambda A^*_{U_1Y_1}) \otimes A^{\dagger}_{Y_2} + (\lambda-1)A^*_{U_1} \otimes A^{\dagger}_{U_2} + (1-\alpha)A^*_{Z_1} \otimes A^{\dagger}_{Z_2} \\
&\quad = (A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} + \alpha A^*_{Y_1} + (1-\alpha)A^*_{Z_1} - \lambda A^*_{U_1,Y_1} - A^*_{V_1Z_1}) \otimes A^{\dagger}_{U_2} \\
&\qquad + (A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes (A^{\dagger}_{X_2} - A^{\dagger}_{U_2}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) + (1-\alpha)A^*_{Z_1} \otimes (A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) \\
&\quad = (A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} + \alpha A^*_{Y_1} + (1-\alpha)A^*_{Z_1} - \lambda A^*_{U_1,Y_1} - A^*_{V_1Z_1}) \otimes A^{\dagger}_{U_2} \\
&\qquad + A^*_{Z_1} \otimes ((1-\alpha)(A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) - (\lambda-\alpha)(A^{\dagger}_{Y_2} - A^{\dagger}_{U_2})) \\
&\qquad + (A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes (A^{\dagger}_{X_2} - A^{\dagger}_{U_2}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) + (\lambda-\alpha)A^*_{Z_1} \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) \\
&\quad = (A^*_{U_1V_1} + (\lambda-1)A^*_{U_1} + \alpha A^*_{Y_1} + (1-\alpha)A^*_{Z_1} - \lambda A^*_{U_1,Y_1} - A^*_{V_1Z_1}) \otimes A^{\dagger}_{U_2} \\
&\qquad + A^*_{Z_1} \otimes ((1-\alpha)(A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) - (\lambda-\alpha)(A^{\dagger}_{Y_2} - A^{\dagger}_{U_2})) + (A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes (A^{\dagger}_{X_2} - A^{\dagger}_{Z_2}) \\
&\qquad + (A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes (A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) + (\lambda-\alpha)A^*_{Z_1} \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) \\
&\quad \succeq (A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes (A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) + (\lambda-\alpha)A^*_{Z_1} \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}).
\end{aligned}$$

The last inequality is a consequence of the fact that if $B, C \succeq 0$, then $B \otimes C \succeq 0$, and conditions (5), (6), and (11).

Using $(1-\alpha)(A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) \succeq (\lambda-\alpha)(A^{\dagger}_{Y_2} - A^{\dagger}_{U_2})$ which follows from (11), we have

$$\begin{aligned}
&(A^*_{U_1V_1} - A^*_{V_1Z_1}) \otimes (A^{\dagger}_{Z_2} - A^{\dagger}_{U_2}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) + (\lambda-\alpha)A^*_{Z_1} \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}) \\
&\quad \succeq \left(\frac{\lambda-\alpha}{1-\alpha}(A^*_{U_1V_1} - A^*_{V_1Z_1}) + (\lambda-\alpha)A^*_{Z_1} - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1})\right) \otimes (A^{\dagger}_{Y_2} - A^{\dagger}_{U_2}).
\end{aligned}$$

Therefore to show the second order conditions, it suffices to show that

$$\frac{\lambda-\alpha}{1-\alpha}(A^*_{U_1V_1} - A^*_{V_1Z_1} + (1-\alpha)A^*_{Z_1}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \succeq 0.$$

Using (6) we have

$$\begin{aligned}
&\frac{\lambda-\alpha}{1-\alpha}(A^*_{U_1V_1} - A^*_{V_1Z_1} + (1-\alpha)A^*_{Z_1}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \\
&\quad \succeq \frac{\lambda-\alpha}{1-\alpha}(\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1} - (\lambda-1)A^*_{U_1}) - (\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1}) \\
&\quad = \frac{\lambda-1}{1-\alpha}(\lambda A^*_{U_1Y_1} - \alpha A^*_{Y_1} - (\lambda-\alpha)A^*_{U_1}) \succeq 0
\end{aligned}$$

since $\lambda \geq 1, \alpha \in (0,1)$ and $A^*_{U_1Y_1} \succeq A^*_{Y_1}, A^*_{U_1}$. This completes the proof of the Proposition 2.6. $\square$

Summary

In this article we show that a particular functional associated with testing the optimality of Marton's inner bound for the capacity region of a two-receiver broadcast channel satisfies a local-tensorization property, if one of the components is binary. In other words, we show that a product distribution obtained by taking the maximizers in each component channel is a local maximizer in the space of joint distributions. If the same distribution were a global maximizer, then this would imply the optimality of Marton's inner bound for binary input broadcast channels. Further it is not known whether there are other non-trivial local maximizers in the space of joint distributions.

In previous works in network information theory, it had been observed that the local maximizer condition and the global maximizer condition seemed to go hand in hand for similar functionals. Indeed it was by observing examples where the local condition failed that the author and his co-authors were able to find counter examples to the optimality of achievable regions in other fundamental network information theory settings, as reported in [3], [4].

Acknowledgements. The author owes immense gratitude for various discussions on this very same problem to Amin Gohari, Venkat Anantharam, Ali Yekkhekhany, Yannan Wang, and Liu Hang. The initial efforts along with Amin Gohari, Venkat Anantharam, and Ali Yekkhekhany, though unsuccessful, made us realize that we needed to use more than positive-semidefiniteness of the various matrices to obtain the result. The first positive result, for pairs of binary channels, were obtained by Liu Hang and coroborrated by Yannan Wang via observing a diagonal dominance property for certain matrices. This proof side-steps this argument and is more general but still owes some intuition to all the previous attempts.

References

[1] C E Shannon. A mathermatical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October, 1948.

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

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

[4] Chandra Nair and Mehdi Yazdanpanah. Sub-optimality of superposition coding region for three receiver broadcast channel with two degraded message sets. In *2017 IEEE International Symposium on Information Theory (ISIT) (ISIT'2017)*, pages 1038–1042, Aachen, Germany, June 2017.

[5] K. Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Transactions on Information Theory*, 25(3):306–311, May 1979.

[6] V. Anantharam, A. Gohari, and C. Nair. On the evaluation of marton's inner bound for two-receiver broadcast channels. *IEEE Transactions on Information Theory*, 65(3):1361–1371, March 2019.