The capacity region of classes of product broadcast channels

Yanlin Geng, Amin Gohari, Chandra Nair, Yuanming Yu

Abstract

We establish a new outer bound for the capacity region of product broadcast channels. This outer bound matches Marton's inner bound for a variety of classes of product broadcast channels whose capacity regions were previously unknown. These classes include product of reversely semi-deterministic and product of reversely more-capable channels. A significant consequence of this new outer bound is that it establishes, via an example, that the previously best known outer-bound is strictly suboptimal for the general broadcast channel. Our example is comprised of a product broadcast channel with two semi-deterministic components in reverse orientation.

1 Introduction

The broadcast channel refers to a communication scenario where a single sender wishes to communicate (possibly different messages) with multiple receivers. We consider a simple setting of the problem where the sender X, who has messages M_1, M_2, wishes to communicate message M_1 to receiver Y and M_2 to receiver Z over a noisy discrete memoryless broadcast channel $\mathfrak{q}(y,z|x)$. A set of rate pairs (R_1, R_2) is said to be achievable for this broadcast channel, $\mathfrak{q}$, if there is a sequence of codebooks, each consisting of:

- an encoder at the sender that maps the message pair (m_1, m_2) into a sequence X^n
- a decoder at receiver Y that maps the received sequence Y^n into an estimate $\hat{M}_1$ of its intended message M_1, and
- a decoder at receiver Z that maps the received sequence Z^n into an estimate $\hat{M}_2$ of its intended message M_2

such that $\mathrm{P}(\hat{M}_1 \neq M_1), \mathrm{P}(\hat{M}_2 \neq M_2) \to 0$ as $n \to \infty$, when the messages M_1, M_2 are uniformly distributed in $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. The capacity region is the closure of the set of all achievable rate pairs. An evaluable characterization of this capacity region is a well known open problem.

The best known inner and outer bounds for the capacity region of a general two-receiver discrete-memoryless broadcast channel with private messages are the following:

- *Inner bound*: (Marton [Mar79]) The union of rate pairs (R_1, R_2) satisfying the inequalities

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

  over all $(U,V,W,X) : (U,V,W) \to X \to (Y,Z)$ forms a Markov chain constitutes an inner bound to the capacity region. Further to compute this region it suffices [GA09] to consider $|U|, |V| \leq |X|, |W| \leq |X| + 4$.

- *Outer bound*: (UV outer bound [El 79, NE07]) The union of rate pairs (R_1, R_2) satisfying the inequalities

$$\begin{aligned} R_1 &\leq I(U;Y) \\ R_2 &\leq I(V;Z) \\ R_1 + R_2 &\leq I(U;Y) + I(X;Z|U) \\ R_1 + R_2 &\leq I(V;Z) + I(X;Y|V) \end{aligned}$$

  over all $(U, V, X) : (U, V) \to X \to (Y, Z)$ forms a Markov chain constitutes an outer bound to the capacity region. Further to compute this region it suffices to consider $|U|, |V| \leq |X| + 1$.

Remark 1. Although there have been outer bounds subsequent to this outer bound, it was shown recently [Nai10] that all these were equivalent to the UV-outer bound.

For all the classes of channels for which the capacity region was previously established the inner and outer bounds coincided. In a sequence of results that was established in the last few years it has been shown [NW08, GA09, JN10, GNSW10] that the inner and outer bounds can indeed be different for various channels. Hence there are three possibilities: 1) UV-outer bound is optimal, 2) Marton's inner bound is optimal, or 3) Neither is optimal in general.

In this paper we show that the UV-outer bound is *strictly* sub-optimal. One of the main contributions of this paper is a new outer bound for product broadcast channels. This outer bound matches Marton's inner bound (and hence is tight) in many instances including classes whose capacity regions were previously unknown, e.g. product of reversely semi-deterministic and reversely more-capable channels. We then construct a reversely semi-deterministic channel where the UV-outer bound is *strictly* weak.

Definition 1. A broadcast channel $\mathfrak{q}(y, z|x)$ is said to be a product broadcast channel if $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2), \mathcal{Y} = (\mathcal{Y}_1, \mathcal{Y}_2), \mathcal{Z} = (\mathcal{Z}_1, \mathcal{Z}_2)$ and $\mathfrak{q}(y_1, y_2, z_1, z_2|x_1, x_2) = \mathfrak{q}_1(y_1, z_1|x_1)\mathfrak{q}_2(y_2, z_2|x_2)$. Here we denote $\mathfrak{q} = \mathfrak{q}_1 \times \mathfrak{q}_2$.

Definition 2. A product broadcast channel $\mathfrak{q} = \mathfrak{q}_1 \times \mathfrak{q}_2$ is said to be reversely semi-deterministic if the channel to one of the receivers in the first component is deterministic, and the channel to the other receiver in the second component is deterministic. That is either both $\mathfrak{q}_1(y_1|x_1), \mathfrak{q}_2(z_2|x_2) \in \{0, 1\}$ or both $\mathfrak{q}_1(z_1|x_1), \mathfrak{q}_2(y_2|x_2) \in \{0, 1\}$.

Definition 3. A product broadcast channel $\mathfrak{q} = \mathfrak{q}_1 \times \mathfrak{q}_2$ is said to be reversely more capable if one of the following two holds:

- $I(X_1;Y_1) \geq I(X_1;Z_1),\ \forall p(x_1),\ I(X_2;Z_2) \geq I(X_2;Y_2),\ \forall p(x_2)$,
- $I(X_1;Z_1) \geq I(X_1;Y_1),\ \forall p(x_1),\ I(X_2;Y_2) \geq I(X_2;Z_2),\ \forall p(x_2)$.

The capacity results originated from the consideration of the 2-letter version of Marton's inner bound and particularly on the sum-rate. Consider a broadcast channel $\mathfrak{q}(y, z|x)$. Marton's sum-rate for this channel is given by

$$SR_{MIB}(\mathfrak{q}) = \max_{p(u,v,w,x)} \min\{I(W;Y), I(W;Z)\} + I(U;Y|W) + I(V;Z|W) - I(U;V|W). \tag{1}$$

Consider the product channel $\mathfrak{q} \times \mathfrak{q}$ obtained by choosing two identical copies of the original channel $\mathfrak{q}(y, z|x)$. If we could give an example where $SR_{MIB}(\mathfrak{q} \times \mathfrak{q}) > 2SR_{MIB}(\mathfrak{q})$, then we could indeed conclude that Marton's inner bound is *strictly* sub-optimal in general.

On the other hand, consider a class $\mathcal{C}$ of channels that is closed under the product operation; i.e. if $\mathfrak{q}_1, \mathfrak{q}_2 \in \mathcal{C}$ then $\mathfrak{q}_1 \times \mathfrak{q}_2 \in \mathcal{C}$. If one is able to show that $SR_{MIB}(\mathfrak{q} \times \mathfrak{q}) = 2SR_{MIB}(\mathfrak{q}), \forall \mathfrak{q} \in \mathcal{C}$; then one can conclude that Marton's sum rate is optimal for channels in $\mathcal{C}$. This is because on the one hand the normalized n-letter Marton's sum rate $\frac{1}{n} SR_{MIB}(\mathfrak{q} \otimes_n) \to SR^*$, the optimal sum-rate, as n converges to infinity. Here we have used the notation $\mathfrak{q} \otimes_n$ to denote $\prod_{i=1}^n \mathfrak{q}(y_i, z_i|x_i)$, the n-fold product channel. On the other hand, $SR_{MIB}(\mathfrak{q} \times \mathfrak{q}) = 2SR_{MIB}(\mathfrak{q}), \forall \mathfrak{q} \in \mathcal{C}$ shows that $SR_{MIB}(\mathfrak{q}) = \frac{1}{2^k} SR_{MIB}(\mathfrak{q} \otimes_{2^k}), \forall k \geq 0$.

See a recent related article [GGNY11] for elaborated discussion of this 2-letter approach. This approach is only superficially different from the usual technique of proving optimality, where one starts from Fano's inequality and directly identifies auxiliary random variables for a class to show the optimality. While these two approaches are identical, we believe that this alternate viewpoint has helped us get our new results. We also believe that the true utility of this approach however lies in the ability to test (at least via numerical simulations) whether a particular scheme is optimal or not.

The outline of this paper is as follows. Our first goal is to construct an example where the sum-rate given by the UV outer-bound is loose. Next, we provide a new outer bound for the capacity region product channels that is optimal for the product of reversely semi-deterministic and reversely more-capable channels. Lastly, we study some properties of the λ-sum rate that played an important role in the construction of the example where the outer-bound was loose.

2 The UV outer bound is not tight

From UV outer bound the sum-rate of a general broadcast channel can be bounded from above by

$$SR_{UV}(\mathfrak{q}) = \max_{p(u,v,x)} \min\{I(U;Y) + I(V;Z), I(U;Y) + I(X;Z|U), I(V;Z) + I(X;Y|V)\}. \quad (2)$$

In this example we will demonstrate a product of reversely semi-deterministic channel, $\mathfrak{q} = \mathfrak{q}_1 \times \mathfrak{q}_2$, such that the optimal sum-rate $SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2)$ satisfies

$$SR_{MIB}(\mathfrak{q}_1 \times \mathfrak{q}_2) = SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2) < SR_{UV}(\mathfrak{q}_1 \times \mathfrak{q}_2).$$

This unequivocally shows that the UV outer bound is *strictly* suboptimal for the general broadcast channel.

Remark 2. Even if one were to consider the best outer bound with a common message requirement, the UVW outer bound [Nai10], the fact that we are showing that the sum-rate is strictly weak for the UV outer bound immediately implies the *strict* sub optimality of the UVW outer bound as well. Reason: Its projection on the plane $R_0 = 0$ (which becomes the UV outer bound) is *strictly* suboptimal.

The flow of this section is as follows: in order to compute the sum-rate for product of reversely semi-deterministic channels, we first define a λ-parametrized family of functions, called the λ-sum rate, that are related to the sum rate given by Marton's inner bound. The fact that λ-sum rate factorizes, in the sense to be defined later, for product channels of this type is used to show that Marton's sum-rate is optimal for this class of product channels. The last step is to show that the UV outer bound is strictly suboptimal over this class of broadcast channels.

2.1 λ-sum rate

Define a λ-parametrized family of functions that is related to the sum rate given by Marton's inner bound as follows:

$$\lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)) = \lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(V;Z|W) - I(U;V|W).$$

We further define

$$\lambda\text{-}SR(\mathfrak{q}, p(x)) = \max_{p(u,v,w|x)} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)), \text{ and } \lambda\text{-}SR(\mathfrak{q}) = \max_{p(u,v,w,x)} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)).$$

Observe that $SR_{MIB} = \max_{p(u,v,w,x)} \min_{\lambda\in[0,1]} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x))$. The following claim allows us to shift the discussion from the sum-rate to λ-sum rate, and then return to the sum-rate at a later point to complete the argument.

Claim 1. *The following min-max theorem holds:*

$$\begin{aligned}\max_{p(u,v,w,x)} \min_{\lambda\in[0,1]} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)) &= \max_{p(x)} \min_{\lambda\in[0,1]} \max_{p(u,v,w|x)} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)) \\ &= \min_{\lambda\in[0,1]} \max_{p(u,v,w,x)} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)).\end{aligned}$$

This implies that the sum-rate of Marton's inner bound can be calculated using any of the three above expressions.

Proof. The fact that

$$\max_{p(u,v,w,x)} \min_{\lambda\in[0,1]} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x)) = \min_{\lambda\in[0,1]} \max_{p(u,v,w,x)} \lambda\text{-}SR(\mathfrak{q}, p(u,v,w,x))$$

was established in section 3.1.1 of [GEA10]. For completeness we present a (slightly different) proof in the Appendix which can be considered as an application of a min-max theorem of Terkelsen [Ter72]. The argument in the Appendix can also be used in other instances where a max-min occurs, such as compound channels. □

2.2 Factorization of λ-sum rate

Given two channels $\mathfrak{q}_1(y_1, z_1|x_1), \mathfrak{q}_2(y_2, z_2|x_2)$, we say that λ-SR factorizes if

$$\lambda\text{-}SR(\mathfrak{q}_1 \times \mathfrak{q}_2) = \lambda\text{-}SR(\mathfrak{q}_1) + \lambda\text{-}SR(\mathfrak{q}_2). \tag{3}$$

We have discussed the factorization of λ-sum rate in a recent related article [GGNY11]. We need a shortened version of lemma from that article:

Lemma 1. *The λ-$SR(\mathfrak{q}_1 \times \mathfrak{q}_2)$ factorizes if any one of the four channels $X_1 \to Y_1; X_1 \to Z_1; X_2 \to Y_2$ or $X_2 \to Z_2$ is deterministic.*

2.3 Marton's sum-rate is tight for product of reversely semi-deterministic channels

Claim 2. *Marton's sum rate is optimal for product of reversely semi-deterministic channels. Moreover the sum rate of such a product channel $\mathfrak{q}_1 \times \mathfrak{q}_2$ is given by*

$$\min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1) + \lambda\text{-}SR(\mathfrak{q}_2)\big).$$

Proof. Take two semi-deterministic channels $\mathfrak{q}_1(y_1, z_1|x_1)$ and $\mathfrak{q}_2(y_2, z_2|x_2)$ where Y_1 is a deterministic function of X_1 and Z_2 is a deterministic function of X_2. It is well-known (also see proof of Lemma 1 in [GGNY11]) that the optimal sum-rate $SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2)$ satisfies

$$SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2) \leq \liminf_{n\to\infty} \frac{1}{n} SR(\mathfrak{q}_1 \otimes_n \times \mathfrak{q}_2 \otimes_n).$$

Consider the n-letter λ-sum rate of the product channel $\mathfrak{q}_1 \times \mathfrak{q}_2$. Using Lemma 1 the n-letter product channel factorizes into λ- sum rate of two n-letter sub channels. Each term them again factorizes by repeated application of Lemma 1. More precisely,

$$\begin{aligned}\lambda\text{-}SR(\mathfrak{q}_1 \otimes_n \times \mathfrak{q}_2 \otimes_n) &= \lambda\text{-}SR(\mathfrak{q}_1 \otimes_n) + \lambda\text{-}SR(\mathfrak{q}_2 \otimes_n) \\ &= n \cdot \lambda\text{-}SR(\mathfrak{q}_1) + n \cdot \lambda\text{-}SR(\mathfrak{q}_2).\end{aligned}$$

Marton's inner bound sum rate for the n-letter of the product channel $\mathfrak{q}_1 \times \mathfrak{q}_2$ is equal to

$$\min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1 \otimes_n \times \mathfrak{q}_2 \otimes_n)\big).$$

We can write the above expression as

$$n \cdot \min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1) + \lambda\text{-}SR(\mathfrak{q}_2)\big).$$

Therefore, the actual sum rate satisfies

$$SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2) \leq \min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1) + \lambda\text{-}SR(\mathfrak{q}_2)\big).$$

On the other hand, this sum rate is achievable since it is equal to the single letter Marton's inner bound for $\mathfrak{q}_1 \times \mathfrak{q}_2$, i.e.

$$SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2) = \min_{\lambda\in[0,1]} \lambda\text{-}SR(\mathfrak{q}_1 \times \mathfrak{q}_2) = \min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1) + \lambda\text{-}SR(\mathfrak{q}_2)\big).$$

□

2.4 The UV outer bound is strictly suboptimal

Claim 3. *Consider the reversely semi-deterministic channel in Figure 1. Assume that the transition probabilities are uniform across the possible outputs, i.e the red edges have a probability $\frac{1}{3}$ in the first component and the blue edges have a probability $\frac{1}{3}$ in the second component. Then Marton's sum rate (the optimal sum rate) is given by $\frac{8}{3} = 3 - \frac{1}{3}$, while the UV sum-rate is at least by $3 - \frac{1}{15}$.*

Proof. We begin by showing that Marton's sum rate (the optimal sum rate) is given by $\frac{8}{3}$. Claim 2 shows that the sum rate of $\mathfrak{q}_1 \times \mathfrak{q}_2$ is

$$\min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1) + \lambda\text{-}SR(\mathfrak{q}_2)\big).$$

The result of Appendix B.1 implies that for any $\lambda \in [0, 1]$, $\lambda\text{-}SR(\mathfrak{q}_1)$ is equal to $\lambda\text{-}SR(\mathfrak{q}_1, u(x_1))$ where u is the uniform distribution on $\mathcal{X}_1$. A similar statement holds for $\lambda\text{-}SR(\mathfrak{q}_2)$. Therefore the sum rate of $\mathfrak{q}_1 \times \mathfrak{q}_2$ is equal to

$$\min_{\lambda\in[0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1, u(x_1)) + \lambda\text{-}SR(\mathfrak{q}_2, u(x_2))\big). \tag{4}$$

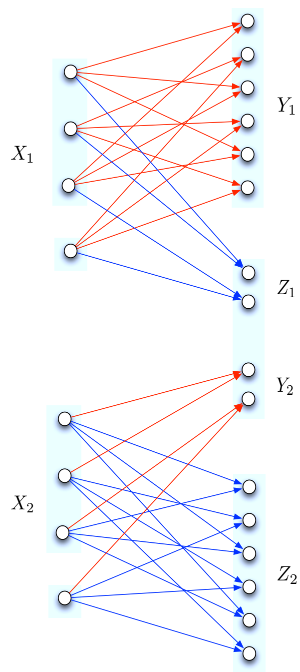

Figure 1: A reversely semi-deterministic channel

By symmetry, $\lambda\text{-}SR(\mathfrak{q}_2, u(x_2)) = (1-\lambda)\text{-}SR(\mathfrak{q}_1, u(x_1))$. Therefore we can express the sum rate as

$$\min_{\lambda \in [0,1]} \big(\lambda\text{-}SR(\mathfrak{q}_1, u(x_1)) + (1-\lambda)\text{-}SR(\mathfrak{q}_1, u(x_1))\big).$$

In Appendix B.2 we show that $\lambda\text{-}SR(\mathfrak{q}_1, u(x_1))$ is equal to

$$\lambda\text{-}SR(\mathfrak{q}_1, u(x_1)) = \begin{cases} \frac{5}{3} - \frac{2}{3}\lambda & \lambda \in [0, \frac{1}{2}] \\ \frac{4}{3} & \lambda \in [\frac{1}{2}, 1] \end{cases} .$$

Substituting this function into (4) we see that the minimum occurs at $\lambda = 0.5$ and the optimum sum-rate is equal to $\frac{8}{3}$.

To compute a lower bound on the UV sum-rate, let $p(x_1, x_2) = u(x_1)u(x_2)$, i.e. independent uniform distribution on $\mathcal{X}_1$ and $\mathcal{X}_2$. We define $U_1, V_1, X_1, U_2, V_2, X_2$ having a joint distribution of the form $p(u_1, v_1, x_1)p(u_2, v_2, x_2)$ as follows. Let $U_1 = Y_1$ and $p(u_2, x_2)$ to satisfy

$$\mathrm{P}(X_2 = 1|U_2 = 1) = \mathrm{P}(X_2 = 3|U_2 = 1) = \frac{1}{2},$$
$$\mathrm{P}(X_2 = 2|U_2 = 1) = \mathrm{P}(X_2 = 4|U_2 = 1) = \frac{1}{2},$$
$$\mathrm{P}(U_2 = 1) = \mathrm{P}(U_2 = 2) = \frac{1}{2}.$$

Similarly, let $V_2 = Z_2$ and $p(v_1, x_1)$ to satisfy

$$\mathrm{P}(X_1 = 1|V_1 = 1) = \mathrm{P}(X_1 = 3|V_1 = 1) = \frac{1}{2},$$
$$\mathrm{P}(X_1 = 2|V_1 = 1) = \mathrm{P}(X_1 = 4|V_1 = 1) = \frac{1}{2},$$
$$\mathrm{P}(V_1 = 1) = \mathrm{P}(V_1 = 2) = \frac{1}{2}.$$

Let Q_1 and Q_2 to binary random variable be mutually independent of each other, and of $U_1, V_1, X_1, U_2, V_2, X_2$. Furthermore assume that $\mathrm{P}(Q_1 = 0) = \mathrm{P}(Q_2 = 0) = \frac{4}{5}$. Define V_1' and U_2' as follows: When $Q_1 = 0$ set $V_1' = V_1$ and else set $V_1' = X_1$. When $Q_2 = 0$ set $U_2' = U_2$ and else set $U_2' = X_2$. Lastly set $\tilde{V}_1 = (V_1', Q_1)$ $\tilde{U}_2 = (U_2', Q_2)$.

We consider the UV region for the choice of $(U_1, \tilde{U}_2)$, $(\tilde{V}_1, V_2)$, (X_1, X_2). Note that

$$\begin{aligned} R_1 &\leq I(U_1, \tilde{U}_2; Y_1, Y_2) \\ &= I(U_1; Y_1) + I(\tilde{U}_2; Y_2) \\ &= H(Y_1) + \frac{4}{5} I(U_2; Y_2) + \frac{1}{5} I(X_2; Y_2) \\ &= 1 + \frac{4}{5} \cdot \frac{1}{3} + \frac{1}{5} \cdot 1 \\ &= \frac{22}{15}. \end{aligned}$$

Similarly, one can show that

$$\begin{aligned} R_2 &\leq I(\tilde{V}_1, V_2; Z_1, Z_2) \\ &= \frac{22}{15}. \end{aligned}$$

The sum rate constraint on $R_1 + R_2$ is as follows:

$$\begin{aligned} R_1 + R_2 &\leq I(U_1, \tilde{U}_2; Y_1, Y_2) + I(X_1, X_2; Z_1, Z_2 | U_1, \tilde{U}_2) \\ &= I(U_1; Y_1) + I(X_1; Z_1 | U_1) + I(\tilde{U}_2; Y_2) + I(X_2; Z_2 | \tilde{U}_2) \\ &= H(Y_1) + I(X_1; Z_1 | Y_1) + \frac{4}{5} I(U_2; Y_2) + \frac{1}{5} I(X_2; Y_2) + \frac{4}{5} H(Z_2 | U_2) \\ &= 1 + \frac{2}{3} + \frac{4}{5} \cdot \frac{1}{3} + \frac{1}{5} \cdot 1 + \frac{4}{5} \cdot 1 \\ &= \frac{44}{15}. \end{aligned}$$

Similarly, one can show that

$$\begin{aligned} R_1 + R_2 &\leq I(\tilde{V}_1, V_2; Z_1, Z_2) + I(X_1, X_2; Y_1, Y_2 | \tilde{V}_1, V_2) \\ &= \frac{44}{15}. \end{aligned}$$

Therefore the point $(R_1, R_2) = (\frac{22}{15}, \frac{22}{15})$ is in this region. Hence the UV sum rate is at least $\frac{44}{15} = 3 - \frac{1}{15}$. Thus for the product channel under consideration

$$\frac{8}{3} = SR_{MIB}(\mathfrak{q}_1 \times \mathfrak{q}_2) = SR^*(\mathfrak{q}_1 \times \mathfrak{q}_2) < \frac{44}{15} \leq SR_{UV}(\mathfrak{q}_1 \times \mathfrak{q}_2).$$

This shows that the UV outer bound is strictly suboptimal in general. □

3 Capacity regions for classes of product broadcast channels

In this section we establish the capacity region for classes of product broadcast channels. Indeed we consider a more general setting where in addition to the private messages M_1, M_2 that the receivers Y and Z wish to decode, the receivers also wish to decode a common message M_0. Hence we are interested in the achievable rate triples (R_0, R_1, R_2). The capacity region is defined in a similar fashion as in the case without common message.

3.1 An outer bound for product channels

We present a new outer bound for the product of two broadcast channels. The manipulations here are inspired by the manipulations that establish Lemma 1 in [GGNY11]. This outer bound matches the capacity region for a variety of product channels, including product of two reversely semideterministic and product of two reversely more-capable channels. Hence, from Claim 3, it follows that this is a *strictly* better bound for product broadcast channels as compared to the UVW outer bound.

Claim 4. *Given a product channel* $\mathfrak{q}(y_1y_2, z_1z_2|x_1x_2) = \mathfrak{q}_1(y_1, z_1|x_1) \times \mathfrak{q}_2(y_2, z_2|x_2)$*, the union over all* $p_1(w_1, v_1, u_1, x_1)p_2(w_2, v_2, u_2, x_2)$ *of triples* (R_0, R_1, R_2) *satisfying*

$$\begin{aligned}
R_0 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
R_0 + R_1 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(U_1; Y_1|W_1) + I(U_2; Y_2|W_2) \\
R_0 + R_2 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(V_1; Z_1|W_1) + I(V_2; Z_2|W_2) \\
R_0 + R_1 + R_2 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
&\quad + I(U_2; Y_2|W_2) + I(X_2; Z_2|U_2, W_2) \\
&\quad + \min\big\{I(U_1; Y_1|W_1) + I(X_1; Z_1|U_1, W_1), I(V_1; Z_1|W_1) + I(X_1; Y_1|V_1, W_1)\big\}, \\
R_0 + R_1 + R_2 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
&\quad + \min\big\{I(U_2; Y_2|W_2) + I(X_2; Z_2|U_2, W_2), I(V_2; Z_2|W_2) + I(X_2; Y_2|V_2, W_2)\big\} \\
&\quad + I(V_1; Z_1|W_1) + I(X_1; Y_1|V_1, W_1),
\end{aligned}$$

forms an outer bound to a product broadcast channel.

Remark 3. One can interchange the roles of Y_2 and Z_1 with Z_2 and Y_1 respectively to get another bound, and we can take the intersection of these two regions as an outer bound.

Proof. The proof of this outer bound is given in the Appendix C. □

Remark 4. The above outer bound is strictly sub-optimal. To see this first note that when one of the product channels is trivial, this outer bound does not give us anything beyond the UVW-outer bound [Nai10]. Now, consider a product of three channels, first one is trivial, the collection of two and three forms a reversely semi-deterministic pair. The new outer bound reduces to the UVW bound on the reversely semi-deterministic, and therefore it is strictly sub-optimal. An interesting *open* question would be to write a outer bound for a general broadcast channel that is at least as good as the new outer bound for the product broadcast channel.

3.1.1 An achievable region for a product broadcast channel

Given a product channel $\mathfrak{q}(y_1y_2, z_1z_2|x_1x_2) = \mathfrak{q}(y_1, z_1|x_1)\mathfrak{q}(y_2, z_2|x_2)$ the union of rate triples satisfying

$$\begin{aligned}
R_0 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
R_0 + R_1 &\leq I(W_1; Y_1) + I(W_2; Y_2) + I(U_1; Y_1|W_1) + I(U_2; Y_2|W_2) \\
R_0 + R_2 &\leq I(W_1; Z_1) + I(W_2; Z_2) + I(V_1; Z_1|W_1) + I(V_2; Z_2|W_2) \\
R_0 + R_1 + R_2 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
&\quad + I(U_1; Y_1|W_1) + I(U_2; Y_2|W_2) + I(V_1; Z_1|W_1) \\
&\quad + I(V_2; Z_2|W_2) - I(U_1; V_1|W_1) - I(U_2; V_2|W_2).
\end{aligned} \tag{5}$$

over all $p_1(w_1, v_1, u_1, x_1)p_2(w_2, v_2, u_2, x_2)$ constitutes an inner bound to the capacity region. The achievablility of these points are immediate from Marton's inner bound by letting $U = (U_1, U_2), V = (V_1, V_2), W = (W_1, W_2)$ and $p(u, v, w) \sim p_1(w_1, v_1, u_1, x_1)p_2(w_2, v_2, u_2, x_2)$.

3.2 Capacity regions for new classes of product broadcast channels

Theorem 1. *The capacity region for a product of reversely semi-deterministic (say, channels $X_1 \to Y_1, X_2 \to Z_2$ are deterministic) broadcast channel is given by the union of rate triples satisfying*

$$\begin{aligned} R_0 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\} \\ R_0+R_1 &\le I(W_1;Y_1)+I(W_2;Y_2)+H(Y_1|W_1)+I(U_2;Y_2|W_2) \\ R_0+R_2 &\le I(W_1;Z_1)+I(W_2;Z_2)+I(V_1;Z_1|W_1)+H(Z_2|W_2) \\ R_0+R_1+R_2 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\} \\ &\quad + I(V_1;Z_1|W_1)+H(Y_1|V_1,W_1)+I(U_2;Y_2|W_2)+H(Z_2|U_2,W_2) \end{aligned}$$

over all $p_1(w_1,v_1,x_1)p_2(w_2,u_2,x_2)$.

Proof. The achievability is immediate by setting $U_1 = Y_1$ and $V_2 = Z_2$ in (5). Note that these two choices of auxiliary random variables are possible since channels $X_1 \to Y_1, X_2 \to Z_2$ are deterministic.

The converse is also immediate from the outer bound in 4. Observe that for any $p_1(w_1,v_1,u_1,x_1)$, $p_2(w_2,v_2,u_2,x_2)$ we have

$$I(U_1;Y_1|W_1) \le H(Y_1|W_1), \quad I(V_2;Z_2|W_2) \le H(Z_2|W_2)$$

and each of the two sum-rate terms is bounded by

$$\begin{aligned} &\min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}+I(V_1;Z_1|W_1) \\ &\quad + H(Y_1|V_1,W_1)+I(U_2;Y_2|W_2)+H(Z_2|U_2,W_2). \end{aligned}$$

Thus the outer bound is contained in the inner bound (and hence they coincide). □

Theorem 2. *The capacity region for a product of reversely more-capable (say, receiver Z_1 is more capable than Y_1, and receiver Y_2 is more capable than Z_2) broadcast channel is given by the union of rate triples satisfying*

$$\begin{aligned} R_0 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\} \\ R_0+R_1 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}+I(U_1;Y_1|W_1)+I(X_2;Y_2|W_2) \\ R_0+R_2 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}+I(X_1;Z_1|W_1)+I(V_2;Z_2|W_2) \\ R_0+R_1+R_2 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}+I(X_2;Y_2|W_2) \\ &\quad + \min\big\{I(U_1;Y_1|W_1)+I(X_1;Z_1|U_1,W_1), I(X_1;Z_1|W_1)\big\}, \\ R_0+R_1+R_2 &\le \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\} \\ &\quad + \min\big\{I(X_2;Y_2|W_2), I(V_2;Z_2|W_2)+I(X_2;Y_2|V_2,W_2)\big\}+I(X_1;Z_1|W_1), \end{aligned}$$

over all $p_1(w_1,v_1,x_1)p_2(w_2,u_2,x_2)$.

Proof. The achievability is immediate by setting $W_1' = (U_1,W_1), U_1' = \emptyset, V_1' = X_1$ and $W_2' = (V_2,W_2), U_2' = X_2', V_2' = \emptyset$ in (5). Plugging these choices into (5) we obtain that one can achieve rate triples satisfying

$$\begin{aligned} R_0 &\le \min\{I(U_1,W_1;Y_1)+I(V_2,W_2;Y_2), I(U_1,W_1;Z_1)+I(V_2,W_2;Z_2)\} \\ R_0+R_1 &\le I(U_1,W_1;Y_1)+I(V_2,W_2;Y_2)+I(X_2;Y_2|V_2,W_2) \\ &= I(W_1;Y_1)+I(W_2;Y_2)+(U_1;Y_1|W_1)+I(X_2;Y_2|W_2) \\ R_0+R_2 &\le I(U_1,W_1;Z_1)+I(V_2,W_2;Z_2)+I(X_1;Z_1|U_1,W_1) \\ &= I(W_1;Z_1)+I(W_2;Z_2)+I(X_1;Z_1|W_1)+I(V_2;Z_2|W_2) \\ R_0+R_1+R_2 &\le \min\{I(U_1,W_1;Y_1)+I(V_2,W_2;Y_2), I(U_1,W_1;Z_1)+I(V_2,W_2;Z_2)\} \\ &\quad + I(X_2;Y_2|V_2,W_2)+I(X_1;Z_1|U_1,W_1). \end{aligned}$$

The last sum-rate term can be split into two terms as follows

$$\begin{aligned}
R_0 + R_1 + R_2 &\leq I(U_1, W_1; Y_1) + I(V_2, W_2; Y_2) + I(X_2; Y_2|V_2, W_2) + I(X_1; Z_1|U_1, W_1) \\
&= I(W_1; Y_1) + I(W_2; Y_2) + I(X_2; Y_2|W_2) + I(U_1; Y_1|W_1) + I(X_1; Z_1|U_1, W_1) \\
R_0 + R_1 + R_2 &\leq I(U_1, W_1; Z_1) + I(V_2, W_2; Z_2) + I(X_2; Y_2|V_2, W_2) + I(X_1; Z_1|U_1, W_1) \\
&= I(W_1; Z_1) + I(W_2; Z_2) + I(X_1; Z_1|W_1) + I(V_2; Z_2|W_2) + I(X_2; Y_2|V_2, W_2)
\end{aligned}$$

Thus we see, by comparing term by term, that this achievable region is at least as large as the region stated in Theorem 2, and hence the region in Theorem 2 is achievable.

The converse is also reasonably immediate from the outer bound in 4. Observe the following:

$$\begin{aligned}
&\min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(U_1; Y_1|W_1) + I(U_2; Y_2|W_2) \\
&\quad \leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(U_1; Y_1|W_1) + I(X_2; Y_2|W_2),
\end{aligned}$$

$$\begin{aligned}
&\min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(V_1; Z_1|W_1) + I(V_2; Z_2|W_2) \\
&\quad \{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(X_1; Z_1|W_1) + I(V_2; Z_2|W_2),
\end{aligned}$$

$$\begin{aligned}
&\min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(U_2; Y_2|W_2) + I(X_2; Z_2|U_2, W_2) \\
&\quad + \min\left\{I(U_1; Y_1|W_1) + I(X_1; Z_1|U_1, W_1), I(V_1; Z_1|W_1) + I(X_1; Y_1|V_1, W_1)\right\} \\
&\quad \leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(X_2; Y_2|W_2) \\
&\qquad + \min\left\{I(U_1; Y_1|W_1) + I(X_1; Z_1|U_1, W_1), I(X_1; Z_1|W_1)\right\},
\end{aligned}$$

$$\begin{aligned}
&\min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(V_1; Z_1|W_1) + I(X_1; Y_1|V_1, W_1) \\
&\quad + \min\left\{I(U_2; Y_2|W_2) + I(X_2; Z_2|U_2, W_2), I(V_2; Z_2|W_2) + I(X_2; Y_2|V_2, W_2)\right\} \\
&\quad \leq \{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(X_1; Z_1|W_1) \\
&\qquad + \min\left\{I(X_2; Y_2|W_2), I(V_2; Z_2|W_2) + I(X_2; Y_2|V_2, W_2)\right\}
\end{aligned}$$

Thus we see, by comparing term by term, that the region stated in Theorem 2 is at least as large as the outer bound in Claim 4. Hence the region in Theorem 2 is an outer bound, thus completing the converse. □

Remark 5. The achievable region in (5) also matches the outer bound in Claim 4 for a variety of other classes. For instance, say Z_1 is more capable than Y_1 and Y_2 is a deterministic function of X_2. In this case, one can show that the capacity region is given by the union of rate triples satisfying

$$\begin{aligned}
R_0 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
R_0 + R_1 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(U_1; Y_1|W_1) + H(Y_2|W_2) \\
R_0 + R_2 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} + I(X_1; Z_1|W_1) + I(V_2; Z_2|W_2) \\
R_0 + R_1 + R_2 &\leq \min\{I(W_1; Y_1) + I(W_2; Y_2), I(W_1; Z_1) + I(W_2; Z_2)\} \\
&\quad + I(V_2; Z_2|W_2) + H(Y_2|V_2, W_2) + I(X_1; Z_1|W_1).
\end{aligned}$$

The details are left to the reader.

4 Properties of λ-$SR(\mathfrak{q})$

In this section we discuss some of the properties of λ-$SR(\mathfrak{q})$ in addition to Claim 1:

Lemma 2. *λ-$SR(\mathfrak{q})$ is related to the optimal sum rate as follows:*

$$\min_{\lambda\in\{0,1\}} \lambda\text{-}SR(\mathfrak{q}) \geq SR^*(\mathfrak{q}),$$

i.e. the minimum value of λ-SR, for $\lambda = 0, l = 1$, yields an upper bound on the optimal sum rate, $SR^*(\mathfrak{q})$.

Proof. We prove the statement for $l = 0$; the proof for $l = 1$ is similar. We begin by showing that for any $p(x)$, $0\text{-}SR(\mathfrak{q}, p(x)) = \max_{p(w|x)} I(W;Z) + I(X;Y|W)$. This implies that $0\text{-}SR(\mathfrak{q}) = \max_{p(w,x)} I(W;Z) + I(X;Y|W)$ which is in turn an upper bound on the optimal sum rate by the UV outer bound (replace W by V). To see this first note that by setting $V = \emptyset, U = X$ we obtain

$$0\text{-}SR(\mathfrak{q}, p(x)) \geq \max_{p(w|x)} I(W;Z) + I(X;Y|W)$$

To obtain the other direction, observe that

$$\begin{aligned}
0\text{-}SR(\mathfrak{q}, p(x)) &= \max_{p(u,v,w|x)} \big\{I(W;Z) + I(U;Y|W) + I(V;Z|W) - I(U;V|W)\big\}\\
&= \max_{p(u,v,w|x)} \big\{I(VW;Z) + I(U;Y|W) - I(U;V|W)\big\}\\
&= \max_{p(u,v,w|x)} \big\{I(VW;Z) + I(U;Y|VW) - I(U;V|WY)\big\}\\
&\leq \max_{p(u,v,w|x)} \big\{I(VW;Z) + I(X;Y|VW)\big\}\\
&= \max_{p(w'|x)} I(W';Z) + I(X;Y|W').
\end{aligned}$$

Not that in the last step we replace (V, W) by W'.

Thus we have, as desired,

$$0\text{-}SR(\mathfrak{q}, p(x)) = \max_{p(w|x)} I(W;Z) + I(X;Y|W).$$

□

Corollary 1. *If the minimum value of λ-$SR(\mathfrak{q})$ is attained at $l = 0$ or $l = 1$ then $SR(\mathfrak{q}) = SR^*(\mathfrak{q})$, i.e. Marton's strategy achieves the optimal sum-rate.*

Proof. This follows from the relationships

$$\min_{\lambda\in[0,1]} \lambda\text{-}SR(\mathfrak{q}) = SR(\mathfrak{q}) \leq SR^*(\mathfrak{q}) \leq \min_{\lambda\in\{0,1\}} \lambda\text{-}SR(\mathfrak{q}).$$

□

Lemma 3. *For an given $\mathfrak{q}$ and $p(x)$, λ-$SR(\mathfrak{q}, p(x))$ and λ-$SR(\mathfrak{q})$ are convex in λ for $\lambda \in [0,1]$.*

Proof. To show that $\lambda \mapsto \lambda\text{-}SR(\mathfrak{q}, p(x))$ is convex, take arbitrary $\lambda_1, \lambda_2, \lambda_3$ satisfying $\lambda_2 = \frac{\lambda_1+\lambda_3}{2}$. Take some $p(w^*, u^*, v^*|x)$ maximizing $\lambda_2\text{-}SR(\mathfrak{q}, p(x))$. Note that

$$\lambda_2\text{-}SR(\mathfrak{q}, p(x)) =$$
$$\{\lambda_2 I(W^*; Y) + (1-\lambda_2) I(W^*; Z) + I(U^*; Y|W^*) + I(V^*; Z|W^*) - I(U^*; V^*|W^*)\} =$$
$$\frac{1}{2}\Big[\{\lambda_1 I(W^*; Y) + (1-\lambda_1) I(W^*; Z) + I(U^*; Y|W^*) + I(V^*; Z|W^*) - I(U^*; V^*|W^*)\}+$$
$$\{\lambda_3 I(W^*; Y) + (1-\lambda_3) I(W^*; Z) + I(U^*; Y|W^*) + I(V^*; Z|W^*) - I(U^*; V^*|W^*)\}\Big] \le$$
$$\frac{1}{2}\big[\lambda_1\text{-}SR(\mathfrak{q}, p(x)) + \lambda_3\text{-}SR(\mathfrak{q}, p(x))\big].$$

To show that $\lambda \mapsto \lambda\text{-}SR(\mathfrak{q})$ is convex, note that

$$\max_{p(x)} \lambda_2\text{-}SR(\mathfrak{q}) \le \max_{p(x)} \frac{1}{2}\big[\lambda_1\text{-}SR(\mathfrak{q}) + \lambda_3\text{-}SR(\mathfrak{q})\big] \le$$
$$\max_{p(x)} \frac{1}{2}\lambda_1\text{-}SR(\mathfrak{q}) + \max_{p(x)} \frac{1}{2}\lambda_3\text{-}SR(\mathfrak{q}) = \frac{1}{2}\big[\lambda_1\text{-}SR(\mathfrak{q}) + \lambda_3\text{-}SR(\mathfrak{q})\big].$$

□

Lemma 4. *To compute the $\lambda\text{-}SR(\mathfrak{q})$, it suffices to consider auxiliary random variables that satisfy $|\mathcal{U}| \le \min(|\mathcal{X}|, |\mathcal{Y}|)$, $|\mathcal{V}| \le \min(|\mathcal{X}|, |\mathcal{Z}|)$, $|\mathcal{W}| \le |\mathcal{X}|$.*

Proof. This is proved in Theorem 2 of [GEA10]. □

Lemma 5. *Take some arbitrary $p(x)$ and real λ^*. Then for any $p(w^*, u^*, v^*|x)$ maximizing $\lambda\text{-}SR(\mathfrak{q}, p(x))$, the line $\lambda \mapsto (\lambda - \lambda^*)(I(W^*; Y) - I(W^*; Z)) + l^*\text{-}SR(\mathfrak{q}, p(x))$ is a supporting hyperplane to the convex curve $\lambda \mapsto \lambda\text{-}SR(\mathfrak{q}, p(x))$.*

Proof. At $\lambda = \lambda^*$, the expression $(\lambda - \lambda^*)(I(W^*; Y) - I(W^*; Z)) + l^*\text{-}SR(\mathfrak{q}, p(x))$ is equal to $l^*\text{-}SR(\mathfrak{q}, p(x))$ which is a point on the curve $\lambda \mapsto \lambda\text{-}SR(\mathfrak{q}, p(x))$. We need to show that for any arbitrary λ,

$$\lambda\text{-}SR(\mathfrak{q}, p(x)) \ge (\lambda - \lambda^*)(I(W^*; Y) - I(W^*; Z)) + l^*\text{-}SR(\mathfrak{q}, p(x)).$$

The above inequality holds because it is equivalent with

$$\lambda\text{-}SR(\mathfrak{q}, p(x)) \ge \lambda I(W^*; Y) + (1-\lambda) I(W^*; Z) + I(U^*; Y|W^*) + I(V^*; Z|W^*) - I(U^*; V^*|W^*).$$

□

Lemma 6. *$\lambda\text{-}SR(\mathfrak{q}, p(x))$ is constant in λ for less noisy channels, deterministic channels, and linear in λ for more capable channels.*

Proof. Less Noisy: Assume that Y is less noisy than Z.

$$\begin{aligned}
\lambda\text{-}SR(\mathfrak{q}, p(x)) &= \max_{p(u,v,w|x)} \big[\lambda I(W; Y) + (1-\lambda) I(W; Z) + I(U; Y|W) + I(V; Z|W) - I(U; V|W)\big] \\
&\le \max_{p(u,v,w|x)} \big[\lambda I(W; Y) + (1-\lambda) I(W; Z) + I(U; Y|W) + I(V; Z|U, W)\big] \\
&\le \max_{p(u,v,w|x)} \big[\lambda I(W; Y) + (1-\lambda) I(W; Z) + I(U; Y|W) + I(V; Y|U, W)\big] \\
&\le \max_{p(w|x)} \big[\lambda I(W; Y) + (1-\lambda) I(W; Z) + I(X; Y|W)\big] \\
&\le \max_{p(w|x)} \big[I(W; Y) + I(X; Y|W)\big] = I(X; Y)
\end{aligned}$$

On the other hand setting $W = \emptyset$, $U = X$, $V = \emptyset$ shows that $\lambda\text{-}SR(\mathfrak{q}, p(x)) \geq I(X;Y)$. Hence $\lambda\text{-}SR(\mathfrak{q}, p(x))$ is a constant.

Remark 6. While the capacity region of the product of reversely degraded broadcast channels was known; however in this case UV-outer bound is tight since the $\lambda\text{-}SR$ is a straight line.

Deterministic:

$$\begin{aligned}
\lambda\text{-}SR(\mathfrak{q}, p(x)) &= \max_{p(u,v,w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(V;Z|W) - I(U;V|W)\big] \\
&\leq \max_{p(u,v,w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(V;Z|U,W)\big] \\
&\leq \max_{p(u,v,w|x)} \big[\lambda I(W;Y,Z) + (1-\lambda) I(W;Y,Z) + I(U;Y,Z|W) + I(V;Y,Z|U,W)\big] \\
&\leq I(X;Y,Z) = H(Y,Z)
\end{aligned}$$

One the other hand setting $W = \emptyset$, $U = Y$, $V = Z$ shows that $\lambda\text{-}SR(\mathfrak{q}, p(x)) \geq H)Y, Z)$. Hence $\lambda\text{-}SR(\mathfrak{q}, p(x))$ is a constant. (Note that these choices of auxiliaries, i.e. $U = Y$, $V = Z$, are permissible for deterministic channels since $(U, V) \to X \to (Y, Z)$ is a Markov chain.)

More capable: Assume that Y is more capable than Z.

$$\begin{aligned}
\lambda\text{-}SR(\mathfrak{q}, p(x)) &= \max_{p(u,v,w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(V;Z|W) - I(U;V|W)\big] \\
&\leq \max_{p(u,v,w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(V;Z|U,W)\big] \\
&\leq \max_{p(u,w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(X;Z|U,W)\big] \\
&\leq \max_{p(u,w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(X;Y|U,W)\big] \\
&\leq \max_{p(w|x)} \big[\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(X;Y|W)\big] \\
&= I(X;Y) + (1-\lambda) \max_{p(w|x)} (I(W;Z) - I(W;Y)).
\end{aligned}$$

On the other hand setting $U = X$, $V = \textit{constant}$ shows that $\lambda\text{-}SR(\mathfrak{q}, p(x)) \geq I(X;Y) + (1 - \lambda) \max_{p(w|x)}(I(W;Z) - I(W;Y))$. Hence $\lambda\text{-}SR(\mathfrak{q}, p(x))$ is linear in λ. $\square$

References

[El 79] A El Gamal. The capacity of a class of broadcast channels. *IEEE Trans. Info. Theory*, IT-25:166–169, March, 1979.

[GA09] A Gohari and V Anantharam. Evaluation of Marton's inner bound for the general broadcast channel. *International Symposium on Information Theory*, pages 2462–2466, 2009.

[GEA10] A Gohari, A El Gamal, and V Anantharam. On an outer bound and an inner bound for the general broadcast channel. *International Symposium on Information Theory*, 2010, arxiv.org/abs/1006.5166.

[GGNY11] Y Geng, A Gohari, C Nair, and Y Yu. On marton's inner bound for two receiver broadcast channels. *Proceedings of the ITA Workshop*, 2011.

[GNSW10] Y Geng, C Nair, S Shamai, and Z V Wang. On broadcast channels with binary inputs and symmetric outputs. *International Symposium on Information Theory*, 2010.

[JN10] V Jog and C Nair. An information inequality for the bssc channel. *Proceedings of the ITA Workshop*, 2010.

[Mar79] K Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Info. Theory*, IT-25:306–311, May, 1979.

[Nai10] C Nair. A note on outer bounds for broadcast channel. *Presented at International Zurich Seminar*, 2010, http://arXiv.org/abs/1101.0640.

[NE07] C Nair and A El Gamal. An outer bound to the capacity region of the broadcast channel. *IEEE Trans. Info. Theory*, IT-53:350–355, January, 2007.

[NW08] C Nair and Z V Wang. On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels. *Proceedings of the ITA Workshop*, 2008, cs.IT/0804.3825.

[Ter72] F Terkelsen. Some minimax theorems. *Mathematica Scandinavica*, 31:405–413, 1972.

A A Min-Max Theorem

Theorem 3 (Theorem 3 of [Ter72])**.** *Let X be a compact connected space, let Y be a set, and let $f : X \times Y \mapsto \mathbb{R}$ be a function satisfying:*

(i) For any $y_1, y_2 \in Y$ there exists $y_0 \in Y$ such that

$$f(x, y_0) \geq \frac{1}{2}\left(f(x, y_1) + f(x, y_2)\right), \forall x \in X.$$

(ii) Every finite intersection of sets of the form $\{x \in X : f(x, y) \leq \alpha)\}$ with $(y, \alpha) \in Y \times R$ is closed and connected.

Then

$$\sup_{y \in Y} \min_{x \in X} f(x, y) = \min_{x \in X} \sup_{y \in Y} f(x, y).$$

We now present a Corollary of the above theorem that can be potentially used in many information theory scenarios.

Corollary 2. *Let Λ_d be the d-dimensional simplex, i.e. $\lambda_i \geq 0$ and $\sum_{i=1}^d \lambda_i = 1$. Let $\mathcal{P}$ be a set of probability distributions $p(u)$. Let $T_i(p(u)), i = 1, .., d$ be a set of functions such that the set $\mathcal{A}$, defined by*

$$\mathcal{A} = \{(a_1, a_2, ..., a_d) \in \mathbb{R}^d : a_i \leq T_i(p(u)) \textit{ for some } \ p(u) \in \mathcal{P}\},$$

is a convex set.

Then

$$\sup_{p(u) \in \mathcal{P}} \min_{\lambda \in \Lambda_d} \sum_{i=1}^d \lambda_i T_i(p(u)) = \min_{\lambda \in \Lambda_d} \sup_{p(u) \in \mathcal{P}} \sum_{i=1}^d \lambda_i T_i(p(u)).$$

Proof. Let $f(l, p(u)) = \sum_{i=1}^{d} \lambda_i T_i(p(u))$. It suffices to verify that $f(l, p(u))$ satisfies the conditions of Theorem 3. Since the set $\mathcal{A}$ is convex, we know that for any $p_1(u), p_2(u) \in \mathcal{P}$ we have a distribution $p_c(u) \in \mathcal{P}$ such that

$$T_i(p_c(u)) \geq \frac{1}{2}\big(T_i(p_1(u)) + T_i(p_2(u))\big), i = 1, ..., d.$$

Hence (using linearity in λ and non-negativity of λ_i) we have

$$f(l, p_c(u)) \geq \frac{1}{2}\big(f(l, p_1(u)) + f(l, p_2(u))\big), \forall \lambda \in \Lambda_d.$$

Since $f(l, p(u))$ is a linear function of λ, it is immediate that the set

$$\mathcal{B}(p(u), \alpha) = \{\lambda \in \Lambda_d : f(l, p(u)) \leq \alpha\}$$

is closed for every pair $(p(u), \alpha) \in \mathcal{P} \times \mathbb{R}$. Further, due to the linearity in λ, if $\lambda_1, \lambda_2 \in \mathcal{B}(p(u), \alpha)$, then the line segment joining λ_1 and λ_2 belongs to $\mathcal{B}(p(u), \alpha)$. This implies that a finite intersection of sets, each containing λ_1 and λ_2 will also contain the line segment joining λ_1 and λ_2, showing that the finite intersection will be connected. Therefore finite intersections of the sets of the form $\mathcal{B}(p(u), \alpha)$ are closed and connected. Thus the Corollary 2 follows from Theorem 3. □

We will now show how one can use the Corollary 2 to establish Claim 1.

Proof. (Proof of Claim 1) It is clear that

$$\begin{aligned}\max_{p(u,v,w,x)} \min_{\lambda \in [0,1]} \lambda\text{-}SR(\mathfrak{q}, p(u, v, w, x)) &\leq \max_{p(x)} \min_{\lambda \in [0,1]} \max_{p(u,v,w|x)} \lambda\text{-}SR(\mathfrak{q}, p(u, v, w, x)) \\ &\leq \min_{\lambda \in [0,1]} \max_{p(u,v,w,x)} \lambda\text{-}SR(\mathfrak{q}, p(u, v, w, x)).\end{aligned}$$

Therefore suffices to show that

$$\max_{p(u,v,w,x)} \min_{\lambda \in [0,1]} \lambda\text{-}SR(\mathfrak{q}, p(u, v, w, x)) = \min_{\lambda \in [0,1]} \max_{p(u,v,w,x)} \lambda\text{-}SR(\mathfrak{q}, p(u, v, w, x)).$$

Here we take $d = 2$ and set

$$\begin{aligned}T_1(p(u, v, w, x)) &= I(W; Y) + I(U; Y|W) + I(V; Z|W) - I(U; V|W) \\ T_2(p(u, v, w, x)) &= I(W; Z) + I(U; Y|W) + I(V; Z|W) - I(U; V|W)\end{aligned}$$

It is clear that the set

$$\mathcal{A} = \{(a_1, a_2) : a_1 \leq T_1(p(u, v, w, x)), a_2 \leq T_2(p(u, v, w, x))\}$$

is a convex set. (In the standard manner, choose $\tilde{W} = (W, Q)$, and when $Q = 0$ choose $(U, V, W, X) \sim p_1(u, v, w, x)$ and $Q = 1$ choose $(U, V, W, X) \sim p_2(u, v, w, x)$). Hence from Corollary 2, we have the proof of Claim 1. □

Remark 7. The proof of this claim in section 3.1.1 of [GEA10] is very similar in flavor and uses the convexity of the set $\mathcal{A}$. However here we recover it from an application of some general theorems, and this technique and Corollary 2 may be helpful in other situations as well.

B The F function for the semi-deterministic channel of Figure 1

B.1 Maximum of F is obtained at the uniform input distribution

Consider the semi-deterministic channel of Figure 1. In this appendix we show that for any $\lambda \in [0,1]$, λ-$SR(\mathfrak{q}, p(x))$ is less than or equal to λ-$SR(\mathfrak{q}, u(x))$ where u is the uniform distribution on $\mathcal{X}$.

Note that λ-$SR(\mathfrak{q}, p(x))$ is concave in $p(x)$. To see this, take two marginal distributions $p_0(x)$ and $p_1(x)$, and assume that (U_0, V_0, W_0, X_0) and (U_1, V_1, W_1, X_1) are two set of random variables maximizing the expressions of λ-$SR(\mathfrak{q}, p_0(x))$ and λ-$SR(\mathfrak{q}, p_1(x))$ respectively. Take a uniform binary random variable Q, independent of all previously defined random variables. Let $U = U_Q$, $V = V_Q$, $W = (W_Q, Q)$, $X = X_Q$. Observe that X is distributed according to $\frac{p_0(x)}{2} + \frac{p_1(x)}{2}$, and furthermore if we compute the expression in λ-$SR(\mathfrak{q}, p(x))$ for (U, V, W, X), we get a value that is greater than or equal to the average of the corresponding values for (U_0, V_0, W_0, X_0) and (U_1, V_1, W_1, X_1). Therefore

$$\lambda\text{-}SR(\mathfrak{q}, p_0(x)) + \lambda\text{-}SR(\mathfrak{q}, p_1(x)) \leq \lambda\text{-}SR(\mathfrak{q}, \frac{p_0(x)}{2} + \frac{p_1(x)}{2}).$$

Thus, λ-$SR(\mathfrak{q}, p(x))$ is concave in $p(x)$.

Take an arbitrary $p(x)$ of the form $\mathrm{P}(X) = (a, b, c, d)$. Because of the symmetries of the channel of Figure 1 with respect to the two receivers, we have

$$\begin{aligned}
\lambda\text{-}SR(\mathfrak{q}, p(x) \sim (a,b,c,d)) &= \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (b,a,d,c)) \\
&= \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (c,d,a,b)) \\
&= \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (d,c,b,a)).
\end{aligned}$$

Here we have used the symmetry between inputs 1 and 2, and the symmetry between inputs 3 and 4, and the symmetry between the pair of inputs $(1,2)$ and $(3,4)$. Using the concavity of F, we have

$$\begin{aligned}
4\lambda\text{-}SR(\mathfrak{q}, p(x) \sim (a,b,c,d)) &= \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (a,b,c,d)) + \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (b,a,d,c)) + \\
&\quad \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (c,d,a,b)) + \lambda\text{-}SR(\mathfrak{q}, p(x) \sim (d,c,b,a)) \\
&\leq 4\lambda\text{-}SR(\mathfrak{q}, p(x) \sim (\frac{1}{4}(a+b+c+d), \frac{1}{4}(a+b+c+d), \\
&\qquad \frac{1}{4}(a+b+c+d), \frac{1}{4}(a+b+c+d))) \\
&= 4\lambda\text{-}SR(\mathfrak{q}, u(x)).
\end{aligned}$$

B.2 Computing the λ-sum rate at the uniform input distribution

In this appendix we compute λ-$SR(\mathfrak{q}, u(x))$ at the uniform input distribution for the semi-deterministic channel given in Figure 1.

Claim 5. *The $\lambda \mapsto \lambda$-$SR(\mathfrak{q}, u(x))$ curve for the channel under consideration consists of two lines,*

$$\lambda\text{-}SR(\mathfrak{q}, u(x)) = \begin{cases} \frac{5}{3} - \frac{2}{3}\lambda & \lambda \in [0, \frac{1}{2}] \\ \frac{4}{3} & \lambda \in [\frac{1}{2}, 1] \end{cases}.$$

Proof. Note that

$$\begin{aligned}
\lambda\text{-}SR(\mathfrak{q}, u(x)) &= \max_{p(u,v,w|x)} \big\{\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + I(V;Z|W) - I(U;V|W)\big\} \\
&= \max_{p(u,w|x)} \big\{\lambda I(W;Y) + (1-\lambda) I(W;Z) + I(U;Y|W) + H(Z|UW)\big\}.
\end{aligned}$$

In the last step we have used the inequality $I(V;Z|W) - I(U;V|W) \leq H(Z|UW)$ together with the fact that $I(Z;Z|W) - I(U;Z|W) = H(Z|UW)$. Therefore $\lambda\text{-}SR(\mathfrak{q}, u(x))$ can be written as

$$\max_{p(u,w|x)} \big\{\lambda H(Y) + (1-\lambda)H(Z) + (1-\lambda)\big(H(Y|W) - H(Z|W)\big) + H(Z|UW) - H(Y|UW)\big\},$$

which is equal to

$$\begin{aligned}
&\lambda H(Y) + (1-\lambda)H(Z) + \max_{p(w|x)} \big\{(1-\lambda)\big(H(Y|W) - H(Z|W)\big) + \\
&\max_{p(u|w,x)} \big(H(Z|UW) - H(Y|UW)\big)\big\}. \qquad (6)
\end{aligned}$$

Let $\mathrm{P}(X|W=i) = (a_i, b_i, c_i, d_i)$, and $f(a_i, b_i, c_i, d_i) = \max_{p(u|x)} H(Z|U) - H(Y|U)$ conditioned on $\mathrm{P}(X) = (a_i, b_i, c_i, d_i)$. Observe that f is concave. The argument is similar to the one given above in the first part of this appendix and we will not repeat it here. Further, observe that $f(a_i, b_i, c_i, d_i) = f(b_i, a_i, d_i, c_i)$ because the symmetry between inputs 1 and 2, and the symmetry between inputs 3 and 4.

Consider the transformation $(a_i, b_i, c_i, d_i) \to (b_i, a_i, d_i, c_i)$, for all i while leaving $\mathrm{P}(W=i)$ unchanged. This preserves expression in equation (6) because of the symmetry between inputs 1 and 2, and the symmetry between inputs 3 and 4. Thus the transformation $(a_i, b_i, c_i, d_i) \to (\frac{a_i+b_i}{2}, \frac{a_i+b_i}{2}, \frac{c_i+d_i}{2}, \frac{c_i+d_i}{2})$, for all i while leaving $\mathrm{P}(W=i)$ unchanged, does not decrease the λ-sum rate since $H(Y|W)$ and f are concave functions in (a_i, b_i, c_i, d_i), and $H(Z|W)$ that appears with a negative sign remains constant under this transformation. Therefore without loss of generality assume that $\mathrm{P}(X|W=i) = (\frac{x_i}{2}, \frac{x_i}{2}, \frac{1-x_i}{2}, \frac{1-x_i}{2})$ when optimizing the expression in equation (6). Let $\mathrm{P}(W=i) = w_i$. Then we require $\sum w_i x_i = \frac{1}{2}$.

Hence we can work out the expression as the maximum over w_i, x_i satisfying the above constraint of

$$\lambda \log 6 + (1-\lambda) + (1-\lambda)\sum_i w_i[\log 3 + \frac{2}{3} - \frac{2}{3}H(x_i, 1-x_i)] + \sum_i w_i f(\frac{x_i}{2}, \frac{x_i}{2}, \frac{1-x_i}{2}, \frac{1-x_i}{2}).$$

We now compute $f(\frac{x_i}{2}, \frac{x_i}{2}, \frac{1-x_i}{2}, \frac{1-x_i}{2})$. Observe that

$$\begin{aligned}
H(Z) - H(Y) &= H(a+b, c+d) - H(\frac{a+b}{3}, \frac{a+c}{3}, \frac{a+d}{3}, \frac{b+c}{3}, \frac{b+d}{3}, \frac{c+d}{3}) \\
&\overset{(a)}{\leq} H(a+b, c+d) - H(\frac{a+b}{3}, \frac{a+c+d}{3}, \frac{a}{3}, \frac{b+c+d}{3}, \frac{b}{3}, \frac{c+d}{3}) \\
&\overset{(b)}{\leq} H(a+b, c+d) - H(\frac{a+b}{3}, \frac{a+b+c+d}{3}, \frac{a+b}{3}, \frac{c+d}{3}, \frac{0}{3}, \frac{c+d}{3}) \\
&= \frac{1}{3}H(a+b, c+d) - \log 3.
\end{aligned}$$

The step (a) holds because the expression is convex in c and d once we fix $c+d$, therefore its maximum must occur at the boundaries. The step (b) holds because the expression is convex in a and b once we fix $a+b$, therefore its maximum must occur at the boundaries.

Therefore $H(Z)-H(Y) \leq \frac{1}{3}H(a+b,c+d)-\log 3$ for all permissible (a,b,c,d). Since the function $\frac{1}{3}H(a+b,c+d)-\log 3$ is concave, we conclude that $f(a_i,b_i,c_i,d_i) \leq \frac{1}{3}H(a+b,c+d)-\log 3$ for all permissible (a,b,c,d). Hence, at $(a,b,c,d) = (\frac{x_i}{2},\frac{x_i}{2},\frac{1-x_i}{2},\frac{1-x_i}{2})$, we have

$$f(\frac{x_i}{2},\frac{x_i}{2},\frac{1-x_i}{2},\frac{1-x_i}{2}) \leq \frac{1}{3}H(x_i,1-x_i)-\log 3.$$

The equality can be indeed achieved by taking with probability half $(0,x_i,0,1-x_i)$ and with probability half $(x_i,0,1-x_i,0)$. Thus, $f(\frac{x_i}{2},\frac{x_i}{2},\frac{1-x_i}{2},\frac{1-x_i}{2}) = \frac{1}{3}H(x_i,1-x_i)-\log 3$.

Substituting this in we get

$$1+(1-\lambda)\frac{2}{3}+(\frac{1}{3}-\frac{2}{3}(1-\lambda))\sum_i w_i H(x_i,1-x_i).$$

We need to maximize this subject to $\sum w_i x_i = \frac{1}{2}$. Clearly when $(1-\lambda) \leq \frac{1}{2}$ the optimal choice is to set $x_i = \frac{1}{2}$. in the other interval, it is optimal to set $x_i = 0$ w.p. $\frac{1}{2}$ and $x_i = 1$ w.p. $\frac{1}{2}$. In this case we get $1+(1-\lambda)\frac{2}{3}$. □

C Proof of the outer bound (Claim 4)

We wish to show that the union over all $p_1(w_1,v_1,u_1,x_1)p_2(w_2,v_2,u_2,x_2)$ of triples (R_0,R_1,R_2) satisfying

$$\begin{aligned}
R_0 &\leq \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}\\
R_0+R_1 &\leq \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}+I(U_1;Y_1|W_1)+I(U_2;Y_2|W_2)\\
R_0+R_2 &\leq \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}+I(V_1;Z_1|W_1)+I(V_2;Z_2|W_2)\\
R_0+R_1+R_2 &\leq \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}\\
&\quad + I(U_2;Y_2|W_2)+I(X_2;Z_2|U_2,W_2)\\
&\quad + \min\big\{I(U_1;Y_1|W_1)+I(X_1;Z_1|U_1,W_1), I(V_1;Z_1|W_1)+I(X_1;Y_1|V_1,W_1)\big\},\\
R_0+R_1+R_2 &\leq \min\{I(W_1;Y_1)+I(W_2;Y_2), I(W_1;Z_1)+I(W_2;Z_2)\}\\
&\quad + \min\big\{I(U_2;Y_2|W_2)+I(X_2;Z_2|U_2,W_2), I(V_2;Z_2|W_2)+I(X_2;Y_2|V_2,W_2)\big\}\\
&\quad + I(V_1;Z_1|W_1)+I(X_1;Y_1|V_1,W_1),
\end{aligned}$$

forms an outer bound to a product broadcast channel.

Proof. Take a code of length n. Let Q be a random variable independent of the code book such that Q is uniform in $[1:n]$. Identify

$$\begin{aligned}
&W_1 = (M_0, Z_{21}^n, Y_{11}^{Q-1}, Z_{1Q+1}^n, Q), W_2 = (M_0, Y_{11}^n, Y_{21}^{Q-1}, Z_{2Q+1}^n, Q), U_1 = U_2 = M_1,\\
&\quad V_1 = V_2 = M_2, X_1 = X_{1Q}, X_2 = X_{2Q}.
\end{aligned}$$

We need to verify that these choice of auxiliaries work. We begin with the sum rate. The manipulations on the sum-rate are the most unconventional, while the rest are quite standard.

Using the Fano inequality, for any $\lambda \in [0,1]$, we can write

$$
\begin{aligned}
& n(R_0+R_1+R_2) - n\epsilon_{1n} \\
& \leq \lambda I(M_0;Y_{11}^n,Y_{21}^n) + (1-\lambda)I(M_0;Z_{11}^n,Z_{21}^n) + I(M_1;Y_{11}^n,Y_{21}^n|M_0) + I(M_2;Z_{11}^n,Z_{21}^n|M_0) - I(M_1;M_2|M_0) \\
& = \lambda I(M_0;Y_{21}^n|Y_{11}^n) + (1-\lambda)I(M_0;Z_{21}^n) + I(M_1;Y_{21}^n|M_0,Y_{11}^n) + I(M_2;Z_{21}^n|M_1,M_0) \\
& \quad + \lambda I(M_0;Y_{11}^n) + (1-\lambda)I(M_0;Z_{11}^n|Z_{21}^n) + I(M_1;Y_{11}^n|M_0) + I(M_2;Z_{11}^n|Z_{21}^n,M_0) - I(M_1;M_2|M_0,Z_{21}^n) \\
& \leq \lambda I(M_0;Y_{21}^n|Y_{11}^n) + (1-\lambda)I(M_0,Y_{11}^n;Z_{21}^n) + I(M_1;Y_{21}^n|M_0,Y_{11}^n) + I(M_2;Z_{21}^n|M_1,M_0,Y_{11}^n) \\
& \quad + \lambda I(M_0,Z_{21}^n;Y_{11}^n) + (1-\lambda)I(M_0;Z_{11}^n|Z_{21}^n) + I(M_1;Y_{11}^n|M_0,Z_{21}^n) + I(M_2;Z_{11}^n|Z_{21}^n,M_0) - I(M_1;M_2|M_0,Z_{21}^n) \\
& \leq \lambda I(M_0,Y_{11}^n;Y_{21}^n) + (1-\lambda)I(M_0,Y_{11}^n;Z_{21}^n) + I(M_1;Y_{21}^n|M_0,Y_{11}^n) + I(X_{21}^n;Z_{21}^n|M_1,M_0,Y_{11}^n) \qquad (7) \\
& \quad + \lambda I(M_0,Z_{21}^n;Y_{11}^n) + (1-\lambda)I(M_0,Z_{21}^n;Z_{11}^n) + I(M_1;Y_{11}^n|M_0,Z_{21}^n) + I(M_2;Z_{11}^n|M_0,Z_{21}^n) - I(M_1;M_2|M_0,Z_{21}^n)
\end{aligned}
$$

where ϵ_{1n} is a function that converges to zero as $n \to \infty$.

Using the two inequalities stated below

$$
\begin{aligned}
& I(M_1;Y_{11}^n|M_0,Z_{21}^n) + I(M_2;Z_{11}^n|M_0,Z_{21}^n) - I(M_1;M_2|M_0,Z_{21}^n) \leq I(M_1;Y_{11}^n|M_0,M_2,Z_{21}^n) + I(M_2;Z_{11}^n|M_0,Z_{21}^n) \\
& I(M_1;Y_{11}^n|M_0,Z_{21}^n) + I(M_2;Z_{11}^n|M_0,Z_{21}^n) - I(M_1;M_2|M_0,Z_{21}^n) \leq I(M_1;Y_{11}^n|M_0,Z_{21}^n) + I(M_2;Z_{11}^n|M_0,M_1Z_{21}^n)
\end{aligned}
$$

using we obtain the following two constraints on the sum-rate.

$$
\begin{aligned}
& n(R_0+R_1+R_2) - n\epsilon_{1n} \\
& \leq \lambda I(M_0,Y_{11}^n;Y_{21}^n) + (1-\lambda)I(M_0,Y_{11}^n;Z_{21}^n) + I(M_1;Y_{21}^n|M_0,Y_{11}^n) + I(X_{21}^n;Z_{21}^n|M_1,M_0,Y_{11}^n) \qquad (8) \\
& \quad + \lambda I(M_0,Z_{21}^n;Y_{11}^n) + (1-\lambda)I(M_0,Z_{21}^n;Z_{11}^n) + I(X_{11}^n;Y_{11}^n|M_0,M_2,Z_{21}^n) + I(M_2;Z_{11}^n|M_0,Z_{21}^n)
\end{aligned}
$$

$$
\begin{aligned}
& n(R_0+R_1+R_2) - n\epsilon_{1n} \\
& \leq \lambda I(M_0,Y_{11}^n;Y_{21}^n) + (1-\lambda)I(M_0,Y_{11}^n;Z_{21}^n) + I(M_1;Y_{21}^n|M_0,Y_{11}^n) + I(X_{21}^n;Z_{21}^n|M_1,M_0,Y_{11}^n) \qquad (9) \\
& \quad + \lambda I(M_0,Z_{21}^n;Y_{11}^n) + (1-\lambda)I(M_0,Z_{21}^n;Z_{11}^n) + I(M_1;Y_{11}^n|M_0,Z_{21}^n) + I(X_{11}^n;Z_{11}^n|M_0,M_1,Z_{21}^n)
\end{aligned}
$$

Starting from (8), standard manipulations as in [El 79] or [NE07] will yield us

$$
\begin{aligned}
& n(R_0+R_1+R_2) - n\epsilon_{1n} \\
& \leq \sum_{i=1}^{n} \lambda I(M_0,Y_{11}^n,Y_{21}^{i-1},Z_{2i+1}^{\ n};Y_{2i}) + (1-\lambda)I(M_0,Y_{11}^n,Y_{21}^{i-1},Z_{2i+1}^{\ n};Z_{2i}) \\
& \quad + I(M_1;Y_{2i}|M_0,Y_{11}^n,Y_{21}^{i-1},Z_{2i+1}^{\ n}) + I(X_{2i};Z_{21}^n|M_1,M_0,Y_{11}^n,Y_{21}^{i-1},Z_{2i+1}^{\ n}) \\
& \quad + \lambda I(M_0,Z_{21}^n,Y_{11}^{i-1},Z_{1i+1}^{\ n};Y_{1i}) + (1-\lambda)I(M_0,Z_{21}^n,Y_{11}^{i-1},Z_{1i+1}^{\ n};Z_{1i}) \\
& \quad + I(X_{1i};Y_{1i}|M_0,M_2,Z_{21}^n,Y_{11}^{i-1},Z_{1i+1}^{\ n}) + I(M_2;Z_{1i}|M_0,Z_{21}^n,Y_{11}^{i-1},Z_{1i+1}^{\ n})
\end{aligned}
$$

This leads to the single letter bound, using our identification of the auxiliaries,

$$
\begin{aligned}
R_0+R_1+R_2 \leq\ & \lambda I(W_2;Y_2) + (1-\lambda)I(W_2;Z_2) + I(U_2;Y_2|W_2) + I(X_2;Z_2|U_2,W_2) \\
& + \lambda I(W_1;Y_1) + (1-\lambda)I(W_1;Z_1) + I(V_1;Z_1|W_1) + I(X_1;Y_1|V_1,W_1).
\end{aligned}
$$

Starting from (9), similar manipulations will yield us the single letter bound

$$
\begin{aligned}
R_0+R_1+R_2 \leq\ & \lambda I(W_2;Y_2) + (1-\lambda)I(W_2;Z_2) + I(U_2;Y_2|W_2) + I(X_2;Z_2|U_2,W_2) \\
& + \lambda I(W_1;Y_1) + (1-\lambda)I(W_1;Z_1) + I(U_1;Y_1|W_1) + I(X_1;Z_1|U_1,W_1).
\end{aligned}
$$

Thus we have the bound

$$
\begin{aligned}
R_0 + R_1 + R_2 \leq \lambda I(W_2;Y_2) + (1-\lambda) I(W_2;Z_2) + I(U_2;Y_2|W_2) + I(X_2;Z_2|U_2,W_2) \\
+ \lambda I(W_1;Y_1) + (1-\lambda) I(W_1;Z_1) + \min\big\{ I(U_1;Y_1|W_1) + I(X_1;Z_1|U_1,W_1), \\
I(V_1;Z_1|W_1) + I(X_1;Y_1|V_1,W_1) \big\}.
\end{aligned}
$$

Since the auxiliaries do not depend on λ we can evaluate the bounds at $\lambda = 0, \lambda = 1$ to yield

$$
\begin{aligned}
R_0 + R_1 + R_2 \leq \min\{I(W_1;Y_1) + I(W_2;Y_2), I(W_1;Z_1) + I(W_2;Z_2)\} + I(U_2;Y_2|W_2) + I(X_2;Z_2|U_2,W_2) \\
+ \min\big\{ I(U_1;Y_1|W_1) + I(X_1;Z_1|U_1,W_1), I(V_1;Z_1|W_1) + I(X_1;Y_1|V_1,W_1) \big\},
\end{aligned}
$$

Interchanging the roles $Y_{11}^n \leftrightarrow Z_{21}^n, Y_{21}^n \leftrightarrow Z_{11}^n, M_2 \leftrightarrow M_1, \lambda \leftrightarrow 1-\lambda$ in (7) and following a similar procedure we obtain

$$
\begin{aligned}
R_0 + R_1 + R_2 \leq \min\{I(W_1;Y_1) + I(W_2;Y_2), I(W_1;Z_1) + I(W_2;Z_2)\} \\
+ \min\big\{ I(U_2;Y_2|W_2) + I(X_2;Z_2|U_2,W_2), \\
I(V_2;Z_2|W_2) + I(X_2;Y_2|V_2,W_2) \big\} \\
+ I(V_1;Z_1|W_1) + I(X_1;Y_1|V_1,W_1).
\end{aligned}
$$

Remark 8. An observant reader may wonder whether we forgot another set of manipulations that would help us combine the two sum-rate bounds into the following term

$$
\begin{aligned}
& R_0 + R_1 + R_2 \\
& \leq \min\{I(W_1;Y_1) + I(W_2;Y_2), I(W_1;Z_1) + I(W_2;Z_2)\} \\
& \quad + \min\big\{ I(U_2;Y_2|W_2) + I(X_2;Z_2|U_2,W_2), I(V_2;Z_2|W_2) + I(X_2;Y_2|V_2,W_2) \big\} \\
& \quad + \min\big\{ I(V_1;Z_1|W_1) + I(X_1;Y_1|V_1,W_1), I(U_1;Y_1|W_1) + I(X_1;Z_1|U_1,W_1) \big\}.
\end{aligned}
$$

However this is not the case. Not only does our choice of auxiliaries not yield this, this term is not even an outer bound (for the product of reversely more capabe channels). We leave the details to the reader.

It remains to verify the following inequalities

$$
\begin{aligned}
R_0 &\leq I(W_1;Y_1) + I(W_2;Y_2), \\
R_0 &\leq I(W_1;Z_1) + I(W_2;Z_2), \\
R_0 + R_1 &\leq I(W_1;Y_1) + I(W_2;Y_2) + I(U_1;Y_1|W_1) + I(U_2;Y_2|W_2), \\
R_0 + R_1 &\leq I(W_1;Z_1) + I(W_2;Z_2) + I(U_1;Y_1|W_1) + I(U_2;Y_2|W_2), \\
R_0 + R_2 &\leq I(W_1;Y_1) + I(W_2;Y_2) + I(V_1;Z_1|W_1) + I(V_2;Z_2|W_2), \\
R_0 + R_2 &\leq I(W_1;Z_1) + I(W_2;Z_2) + I(V_1;Z_1|W_1) + I(V_2;Z_2|W_2).
\end{aligned}
$$

Of these the first, second, third, and sixth are really straightforward from standard manipulations for our choice of auxiliaries. Hence we show only the fourth term. The fifth term follows in a similar fashion as the fourth term.

To get the fourth inequality, from Fano's inequality we have

$$
\begin{aligned}
(R_0 + R_1) - \epsilon_{2n} &\le \frac{1}{n}\big(I(M_0; Z_{11}^n, Z_{21}^n) + I(M_1; Y_{11}^n, Y_{21}^n | M_0)\big) \\
&\le \frac{1}{n}\big(I(M_0, Z_{21}^n; Z_{11}^n) + I(M_0, Y_{11}^n; Z_{21}^n) + I(M_1; Y_{11}^n | M_0, Z_{21}^n) + I(M_1; Y_{21}^n | M_0, Y_{11}^n)\big) \\
&\overset{(a)}{\le} \frac{1}{n}\sum_{i=1}^{n} I(M_0, Z_{21}^n, Z_{1i+1}^n, Y_{11}^{i-1}; Z_{1i}) + I(M_0, Y_{11}^n, Z_{2i+1}^n, Y_{21}^{i-1}; Z_{2i}) \\
&\quad + I(M_1; Y_{1i} | M_0, Z_{21}^n, Z_{1i+1}^n, Y_{11}^{i-1}) + I(M_1; Y_{21}^n | M_0, Y_{11}^n, Z_{2i+1}^n, Y_{21}^{i-1}) \\
&\le I(W_1; Z_1) + I(W_2; Z_2) + I(U_1; Y_1 | W_1) + I(U_2; Y_2 | W_2),
\end{aligned}
$$

as desired. Here (a) again follows from standard manipulations. This completes the proof of the outer bound for product channels. □