An achievable rate region for the 2-receiver broadcast channel obtained by viewing it as an interference channel

Chandra Nair

Abstract

We derive an achievable region for the 2-receiver broadcast channel with private messages only by incorporating ideas of message splitting as in the Han-Kobayashi scheme for the interference channel. The achievable region we obtain is equivalent to the best-known achievable region under this scenario, obtained by Marton.

The scheme presented here is motivated by the following question: Why does a common random variable W help improve Marton's achievable rate region when only private messages are required? Motivated by this question, we produce a region without an explicit commonly generated random variable W; instead we think of each private message being split naturally into two parts: one part that is also decoded by the other receiver, and another that is only decoded by its intended receiver.

1 Introduction

In [1], Cover introduced the notion of a broadcast channel through which one sender transmits information to two or more receivers.

Definition: A *broadcast channel* (BC) consists of an input alphabet $\mathcal{X}$ and output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ and a probability transition function $p(y_1, y_2|x)$. A $((2^{nR_1}, 2^{nR_2}), n)$ code for a broadcast channel consists of an encoder

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

and two decoders

$$\hat{\mathcal{W}}_1 : \mathcal{Y}_1^n \to 2^{nR_1}$$
$$\hat{\mathcal{W}}_2 : \mathcal{Y}_2^n \to 2^{nR_2}.$$

The probability of error $P_e^{(n)}$ is defined to be the probability that the decoded message is not equal to the transmitted message, i.e.,

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

where the message is assumed to be uniformly distributed over $2^{nR_1} \times 2^{nR_2}$.

A rate pair (R_1, R_2) is said to be *achievable* for the broadcast channel if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$. The *capacity region* of the broadcast channel is the closure of the set of achievable rates.

The capacity region of the two user discrete memoryless channel is unknown. The capacity region is known for lots of special cases such as degraded, less noisy, more capable, deterministic, semi-deterministic, etc. The best known achievable region for the private messages is the one presented by Märton [5] and is stated below:

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

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

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

1.1 Motivation

The paper is motivated by the following observation: W is a random variable whose information content will be decoded by both receivers. Since there is only private messages requirement it seems quite natural to assume that it is optimal to set $W = \emptyset$ - the trivial random variable. This yields the following achievable region:

Bound 2. *[Märton '79] The following rate pairs are achievable:*

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

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

However it is known that the region represent by Bound 2 is strictly contained in the region denoted by Bound 1, even for private messages.

In the next section we will derive an achievable region from an intuitive stand point. It will be shown that the region thus obtained is equivakent to Marton's inner bound. It is hoped that the derivation will make the role played by W in the original region clear.

2 An achievable region

In the first part we walk the reader through the intuition behind the derivation of the region. A reader may skip to Section 2.3 if desired.

2.1 Intuition

For the derivation of this achievable region we start with the reduced Marton's region, i.e. the set of rate-pairs (R_1, R_2) satisfying

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

In this expression U^n sequences represent the private message for receiver Y_1, and V^n sequences represent the private message for receiver Y_2. Each sequence U^n, V^n is decoded separately at both receivers Y_1, Y_2.

Let us ask a hypothetical question: *What if the receivers can decode some part of the other receiver's message for free?* Then the receivers could use this additional information to decode the message intended for them.

To answer this need to make precise the following two notions: (i) what do we mean by part of the other receiver's message, and (ii) what do we mean by free?

The first part is clear; any random variable $U_1 = f(U)$ and $V_1 = g(V)$ can be regarded as parts of the random variables U and V as the information content in U and V will be the same as that in (U, U_1), (V, V_1). The second part is a bit more tricky. We say that receiver Y_1 can decode V_1 and Y_2 can decode U_1, *for free*, if the following two conditions are satisfied:

$$\begin{aligned} I(U_1; Y_2) &\geq I(U_1; Y_1, V_1) \\ I(V_1; Y_1) &\geq I(V_1; Y_2, U_1). \end{aligned} \tag{1}$$

Intuitively the first inequality means that Y_2 can decode U_1 better than Y_1 can decode U_1 even if it knew V_1. This condition may remind the reader about *very strong interference* and *strong interference* conditions, and is exactly in the same spirit.

2.2 An interference channel like reasoning

The transmission of the message M_1 can be viewed as causing interference to the transmission between X to Y_2. Similarly the transmission of the message M_2 can be viewed as causing interference to the transmission between X to Y_1. Unlike traditional interference channels, all of the interference is present at the same physical transmitter.

Just as in the Han-Kobayashi scheme for the interference channel it is feasible to think of $U = (U_1, U_2)$ where U_1 represents the part that can easily be decoded by Y_2 while the remaining part of U is the private part. A similar decomposition holds for $V = (V_1, V_2)$. Conditions (1) just ensures that necessitating Y_2 to decode U_1 and Y_1 to decode V_1 do not really impose constraints on the number of U_1^n and V_1^n sequences that may otherwise be generated.

Motivated by this reasoning we arrive at the achievable region[1] in the next section.

2.3 Main Result

Theorem 1. *Let (U, V) be any two (auxiliary) random variables such that*

$$(U, V) \to X \to (Y_1, Y_2)$$

forms a Markov chain. Further let $U_1 = f(U), V_1 = g(V)$ be two random variables that are deterministic functions of U and V respectively satisfying:

$$\begin{aligned} I(U_1; Y_2) &\geq I(U_1; Y_1, V_1) \\ I(V_1; Y_1) &\geq I(V_1; Y_2, U_1). \end{aligned}$$

Then any pair (R_1, R_2) satisfying the following constraints:

$$\begin{aligned} R_1 &\leq I(U; Y_1, V_1) \\ R_1 &\leq I(U; Y_1|V_1) + I(V_1; Y_2, U_1) \\ R_2 &\leq I(V; Y_2, U_1) \\ R_2 &\leq I(V; Y_2|U_1) + I(U_1; Y_1, V_1) \\ R_1 + R_2 &\leq I(U; Y_1, V_1) + I(V; Y_2, U_1) - I(U; V) \end{aligned} \tag{2}$$

is achievable.

Proof. The proof of this theorem closely mirrors that of Marton's theorem for obvious reasons. For completeness we present the proof in the Appendix. □

Remark 1. We present some remarks on the achievable region presented in Theorem 1.

[1] The author is thankful to Amin Gohari for pointing out an error in an earlier version

i) There is no explicit random variable W that is commonly generated; instead there is just U and V that are independently generated. The receivers may choose to decode parts of the interfering message if it helps them decode their own message.

ii) Setting $U_1 = V_1 = \emptyset$, the trivial random variable, (and is a valid choice satisfying (1) for every U, V) reduces the region to the one in Bound 2.

iii) For the degraded broadcast channel $X \to Y_1 \to Y_2$, a natural choice would to set $V_1 = V$ and $U_1 = \emptyset$ since receiver Y_1 can decode all that receiver Y_2 decodes. It is easy to see that this is a valid choice, i.e. satisfies (1), and leads us to the following region

$$\begin{aligned} R_1 &\leq I(U; Y_1, V) \\ R_1 &\leq I(U; Y_1|V) + I(V; Y_2) \\ R_2 &\leq I(V; Y_2) \\ R_1 + R_2 &\leq I(U; Y_1, V) + I(V; Y_2) - I(U; V) \\ &= I(U; Y_1|V) + I(V; Y_2) \end{aligned} \tag{3}$$

By setting $U = X$ it follows that the above region indeed coincides with the capacity region of the degraded broadcast channel.

iv) The following region is equivalent to the region presented in Theorem 1.

Bound 3. *Let* (U, U_1, V, V_1) *be any four (auxiliary) random variables such that*

$$(U, U_1, V, V_1) \to X \to (Y_1, Y_2)$$

forms a Markov chain. Further let U_1, V_1 *satisfy:*

$$\begin{aligned} I(U_1; Y_2) &\geq I(U_1; Y_1, V_1) \\ I(V_1; Y_1) &\geq I(V_1; Y_2, U_1). \end{aligned}$$

Then any pair (R_1, R_2) *satisfying the following constraints:*

$$\begin{aligned} R_1 &\leq I(U, U_1; Y_1, V_1) \\ R_1 &\leq I(U, U_1; Y_1|V_1) + I(V_1; Y_2, U_1) \\ R_2 &\leq I(V, V_1; Y_2, U_1) \\ R_2 &\leq I(V, V_1; Y_2|U_1) + I(U_1; Y_1, V_1) \\ R_1 + R_2 &\leq I(U, U_1; Y_1, V_1) + I(V, V_2; Y_2, U_1) \\ &\quad - I(U, U_1; V, V_1) \end{aligned} \tag{4}$$

is achievable.

Clearly this is at least as good as the region presented in Theorem 1 as this reduces to the region in Theorem 1 when $U_1 = f(U), V_1 = f(V)$. To show the other direction set $\tilde{U} = (U, U_1), \tilde{V} = (V, V_1)$ and observe that $U_1 = f(\tilde{U}), V_1 = g(\tilde{V})$.

We now present the two Lemmas that shows the equivalence between the regions in Theorem 1 and Bound 1.

Lemma 1. *The region presented in Theorem 1 contains the region in Bound 1.*

Proof. Consider a triple a triple $(\tilde{U}, \tilde{V}, \tilde{W})$ for Marton's scheme.

Case 1: $I(\tilde{V}, \tilde{W}; Y_1) \geq I(\tilde{V}, \tilde{W}; Y_2)$ (or $I(\tilde{U}, \tilde{W}; Y_2) \geq I(\tilde{U}, \tilde{W}; Y_1)$) If $I(\tilde{V}, \tilde{W}; Y_1) \geq I(\tilde{V}, \tilde{W}; Y_2)$, then by setting $U_1 = \emptyset, V_1 = V = (\tilde{V}, \tilde{W})$, $U = (\tilde{U}, \tilde{V}, \tilde{W})$, we obtain that we can achieve

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

which contains the region prescribed by Bound 1. The other case is dealt similarly.

Case 2: $I(\tilde{V}, \tilde{W}; Y_1) < I(\tilde{V}, \tilde{W}; Y_2)$ and $I(\tilde{U}, \tilde{W}; Y_2) < I(\tilde{U}, \tilde{W}; Y_1)$); In this case it is easy to see that for all boundary points of Martons region it suffices to consider $I(\tilde{W}; Y_1) = I(\tilde{W}; Y_2)$ (see also [4]). For our scheme set, $U = (\tilde{U}, \tilde{W}), V = (\tilde{V}, \tilde{W}), V_1 = \tilde{W}, U_1 = \emptyset$. Clearly (1) holds. Therefore we can achieve

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

which is same as the region prescribed by Bound 1. □

Lemma 2. *The region presented in Bound 1 contains the region in Theorem 1.*

Proof. We have $R_1 \leq I(U, U_1; Y_1|V_1) + I(V_1; Y_2, U_1) \leq I(U, U_1, V_1; Y_1)$ (using (1)). Similarly $R_2 \leq I(V, U_1, V_1; Y_2)$.

Observe that we have

$$
\begin{aligned}
&R_1 + R_2 \\
&\quad \leq I(V;Y_2,U_1) + I(U;Y_1,V_1) - I(U;V) \\
&\quad = I(V,V_1;Y_2,U_1) + I(U,U_1;Y_1,V_1) - I(U,U_1;V,V_1) \\
&\quad = I(V,V_1;Y_2|U_1) + I(U,U_1;Y_1,V_1) - I(U;V,V_1|U_1) \\
&\quad = I(V_1;Y_2|U_1) + I(V;Y_2|U_1,V_1) + I(U_1;Y_1,V_1) \\
&\qquad + I(U;Y_1,V_1|U_1) - I(U;V,V_1|U_1) \\
&\quad = I(V_1;Y_2,U_1) + I(U_1;Y_1,V_1) - I(U_1;V_1) \\
&\qquad + I(V;Y_2|U_1,V_1) + I(U;Y_1|V_1,U_1) \\
&\qquad - I(U;V|V_1,U_1)
\end{aligned}
$$

However (1) implies that

$$
\begin{aligned}
&I(V_1;Y_2,U_1) + +I(U_1;Y_1,V_1) - I(U_1;V_1) \\
&\quad \leq I(V_1;Y_1) + I(U_1;Y_1|V_1) = I(U_1,V_1;Y_1), \\
&I(V_1;Y_2,U_1) + I(U_1;Y_1,V_1) - I(U_1;V_1) \\
&\quad \leq I(U_1;Y_2) + I(V_1;Y_2|U_1) = I(U_1,V_1;Y_2).
\end{aligned}
$$

Therefore it follows that

$$
\begin{aligned}
R_1 + R_2 \leq \min\{&I(U_1,V_1;Y_2), I(U_1,V_1;Y_1)\} \\
&+ I(U;Y_1|V_1,U_1) + I(V;Y_2|U_1,V_1) \\
&- I(U;V|V_1,U_1).
\end{aligned}
$$

These rate pairs are achievable using Marton's scheme with $(\tilde{U},\tilde{V},\tilde{W})$ by setting $\tilde{W} = (U_1,V_1)$, $\tilde{U} = U$ and $\tilde{V} = V$. □

3 Conclusion

This paper presents an achievable region for the 2-receiver broadcast channel by borrowing on ideas of message splitting. The use of message splitting in broadcast channel is not new to this work, indeed it was introduced in [6] to establish the capacity region of a class of three receiver broadcast channels with degraded message sets. In this paper we use this idea to get a better understanding of the role played by W for the case when only private messages are required. It thus incorporates ideas from the interference channel model on to the broadcast channel setting.

Acknowledgements

The author is grateful to Prof. Abbas El Gamal, Zizhou Vincent Wang, Varun Jog for very fruitful discussions on various elements of this work. The author is also thankful to Prof. Young-Han Kim for an early observation of the conceptual similarity of the scheme in [6] to the Han-Kobayashi scheme of the interference channel. This observation was also later independently made by Prof. David Tse and his research group.

References

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

[2] T Cover and J Thomas, *Elements of information theory*, Wiley Interscience, 1991.

[3] Abbas A. El Gamal and Edward C. van der Meulen, *A proof of marton's coding theorem for the discrete memoryless broadcast channel*, IEEE Transactions on Information Theory **27** (1981), no. 1, 120–121.

[4] Y Liang, G Kramer, and H V Poor, *Equivalence of two inner bounds on the capacity region of the broadcast channel*, Proceedings of the 46th Annual Allerton Conference on Communication, Control and Computing (2008).

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

[6] C. Nair and A. El Gamal, *The capacity region of a class of three-receiver broadcast channels with degraded message sets*, Information Theory, IEEE Transactions on **55** (2009), no. 10, 4479–4493.

Proof of Theorem 1

The arguments are standard and are just presented for completeness.

Codebook generation: Generate $2^{n(I(U_1;Y_1,V_1)-\epsilon)}$ sequences U_1^n independently and identically distributed according to $\prod_i p(u_{1i})$. For each such sequences U_1^n generate $2^{n(I(U;Y_1,V_1|U_1)-\epsilon)}$ sequences U^n independently and identically distributed according to $\prod_i p(u_i|u_{1i})$. Consider the total collection of $2^{n(I(U;Y_1,V_1)-2\epsilon)}$ sequences thus generated. Randomly and uniformly bin them into 2^{nR_1} bins. It is easy to see that a randomly chosen bin will be non-empty with high probability as long as

$$R_1 < I(U;Y_1,V_1) - \epsilon \qquad (5)$$

Independently generate $2^{n(I(V_1;Y_2,U_1)-\epsilon)}$ sequences V_1^n independently and identically distributed according to $\prod_i p(v_{1i})$. For each such sequences V_1^n generate $2^{n(I(V;Y_2,U_1|V_1)-\epsilon)}$ sequences V^n independently and identically distributed according to $\prod_i p(v_i|v_{1i})$. Consider the total collection of $2^{n(I(V;Y_1,V_1)-2\epsilon)}$ sequences thus generated. Randomly and uniformly bin them into 2^{nR_2} bins. Similarly, a randomly chosen bin will be non-empty with high probability as long as

$$R_2 < I(V; Y_2, U_1) - \epsilon \tag{6}$$

For all $(m_1, m_2) \in 2^{nR_1} \times 2^{nR_2}$ define a product bin $\mathcal{B}_{m_1,m_2}$ comprising of pairs of sequences U^n that belong to bin numbered m_1 and V^n that belong to bin numbered m_2. We say that a product bin m_1, m_2 is *good* if there exists at least one jointly typical pair U^n, V^n. For all good product bins, pick one of the jointly typical pairs U^n, V^n in the product bin and generate a jointly typical sequence $X^n(m_1, m_2)$ according to $p(x|u, v)$.

Encoding and decoding strategy: To transmit a message pair $(m_1, m_2) \in 2^{nR_1} \times 2^{nR_2}$ the transmitter look at the product bin numbered by (m_1, m_2). The transmitter declares an error (error event $\mathcal{E}_1$) if the product bin is not good; otherwise we transmit the sequence $X^n(m_1, m_2)$.

Assuming the transmitter does not declare an error, the receiver Y_1 upon receiving y_1^n computes a list of sequences v_1^n such that it is jointly typical with the receiver y_1^n. The decoder at receiver Y_1 declares an error if

(a) Event $\mathcal{E}_{2,1}$: there is no sequence v_1^n that is jointly typical with the received y_1^n

(b) Event $\mathcal{E}_{3,1}$: there is more that one sequence v_1^n that is jointly typical with the received y_1^n.

If there is a unique v_1^n that is typical with the received y_1^n, then the receiver computes a list of sequences u^n that is jointly typical with the pair (y_1^n, v_1^n). The decoder at receiver Y_1 declares an error if

(c) Event $\mathcal{E}_{4,1}$: there is no sequence u^n that is jointly typical with the pair y_1^n, v_1^n.

(d) Event $\mathcal{E}_{5,1}$: there is more that one sequence u^n that is jointly typical with the pair y_1^n, v_1^n..

If there is a unique u^n the the receiver estimates $\hat{M}_1$ to be the bin number corresponding to the sequence U^n.

A similar decoding strategy occurs at receiver Y_2 as well. Let us call the corresponding error events $\mathcal{E}_{2,2}, \mathcal{E}_{3,2}, \mathcal{E}_{4,2}, \mathcal{E}_{5,2}$ respectively.

Analysis of error probability: Following the arguments in [3] (follows from an application of the second moment method) that the probability of the error event $\mathcal{E}_1$ goes to zero as $n \to \infty$ as long as

$$\begin{aligned} I(U; Y_1, V_1) - R_1 + I(V_1; Y_2, U_1) \\ > I(U; V_1) + 3\epsilon, \\ I(V; Y_2, U_1) - R_2 + I(U_1; Y_1, V_1) \\ > I(U_1; V) + 3\epsilon, \\ I(U; Y_1, V_1) - R_1 + I(V; Y_2, U_1) - R_2 \\ > I(U; V) + 3\epsilon. \end{aligned} \tag{7}$$

By the strong Markov property[2] the transmitted sequence v_1^n will be jointly typical with the received y_1^n with high probability. Therefore the probability of the error event $\mathcal{E}_{2,1}$ goes to zero as $n \to \infty$. The probability that a randomly generated sequence v_1^n is jointly typical with the received y_1^n is at most $2^{-n(I(V_1;Y_1)-\frac{\epsilon}{2})}$. Therefore by the union bound the probability that at least one randomly generated v_1^n is jointly typical with the received y_1^n goes to zero since

$$-I(V_1; Y_1) + \frac{\epsilon}{2} + I(V_1; Y_2, U_1) - \epsilon < 0,$$

which follows from (1). So only the transmitted v_1^n is jointly typical with the received y_1^n with high probability. Therefore the probability the error event $\mathcal{E}_{3,1}$ goes to zero as $n \to \infty$.

Again using the strong Markov property, we can show that the probability of the error event $\mathcal{E}_{4,1}$ goes to zero as $n \to \infty$. We split the error event $\mathcal{E}_{5,1}$ into two parts: (i) there is an u^n with an underlying sequence u_1^n different from the transmitted u_1^n that is jointly typical with the pair y_1^n, v_1^n. By union bound this probability goes to zero as the total number of such sequences u^n is bounded by $2^{n(I(U;Y_1,V_1)-2\epsilon)}$ and the probability that each randomly generated sequence is jointly typical with the pair y_1^n, v_1^n is at most $2^{n(I(U;Y_1,V_1)-\frac{\epsilon}{2})}$. (ii) there is a sequence u^n with an underlying sequence u_1^n same as the transmitted one that is jointly typical with the pair y_1^n, v_1^n. By union bound this probability goes to zero as the total number of such sequences u^n is bounded by $2^{n(I(U;Y_1,V_1|U_1)-\epsilon)}$ and the probability that each randomly generated sequence is jointly typical with the pair y_1^n, v_1^n is at most $2^{n(I(U;Y_1,V_1|U_1)-\frac{\epsilon}{2})}$. Combining these two, we see that the probability of the error event $\mathcal{E}_{5,1}$ goes to zero as $n \to \infty$.

This implies that receiver Y_1 decodes the correct U^n with high probability, and hence the right M_1. A similar analysis can be carried out for receiver Y_2 as well, showing that she also decodes the correct M_2 with high probability. Thus the estimates $\hat{M}_1, \hat{M}_2$ are correct with high probability, and this completes the proof of the achievability.

Setting $\epsilon \to 0$ and taking the closure of the rates completes the proof of Theorem 1.