On Three-Receiver More Capable Channels

Chandra Nair, Lingxiao Xia
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, NT, Hong Kong S.A.R.
Email:{chandra,lingxiao}@ie.cuhk.edu.hk

Abstract—In this paper we show that superposition coding is not optimal for three-receiver more capable channels. The optimality of superposition coding region has been open for k-receiver ($k \geq 3$) more capable broadcast channel. The main contribution is in identifying a counterexample to demonstrate that superposition coding is sub-optimal. We also compute the capacity region for the counter example. On the other hand, we show that the sum-capacity for a k-receiver more capable broadcast channel is obtained by transmitting all the information to the most capable receiver.

I. INTRODUCTION

Broadcast channel models a basic communication scenario where a single sender X wishes to communicate possibly different messages to multiple receivers $Y_1, ..., Y_k$ over a noisy medium. For details of previous results on this problem the readers are encouraged to refer to Chapters 5, 8, and 9 of [1]. This paper addresses an open question in the book (open problem 5.2), and we show that superposition coding technique (the main material presented in Chapter 5) is not optimal for the three-receiver more capable channel.

Determining the capacity region of a two-receiver discrete memoryless broadcast channel continues to remain open as one of the most fundamental unsolved problems in this field; yet, there has been remarkable success when there is an ordering between the decoding capabilities of the receivers. The first such ordering is the *degraded* broadcast channel[3] where $X \rightarrow Y_1 \rightarrow \cdots Y_k$ forms a Markov chain, and the optimality of superposition coding was established by Bergmans[2] and Gallager[6]. Two progressively less stringent orderings[8] called *less noisy* and *more capable* were introduced by Korner and Marton. The optimality of the superposition coding scheme for these classes for the two-receiver case were established in [8], [5], respectively. The optimality of superposition coding for three-receiver less noisy broadcast channels was established in [10]. Our results here show that superposition coding is not optimal for three-receiver more capable channels. This represents the first instance where an ordering on the decodability of the receivers is established, yet superposition coding is not optimal.

It has been shown in [7] that the more-capable ordering is a much weaker ordering than less noisy ordering. In particular it was shown that if one substitutes a receiver (in a two-receiver broadcast channel) by a more-capable receiver then the capacity region could strictly decrease (!). Further it was also shown that such a phenomenon would

not occur for less noisy ordering. Hence our result should not come as very surprising. However, based on the work in [7], a natural instinct for beating the superposition coding's achievable region would be to show that the maximum sum-rate achieved by superposition coding region is strictly smaller than the sum-rate capacity. This fails because, as we show later, the sum-rate capacity can be achieved by transmitting only to the best receiver. Furthermore, since the capacity region would indeed reduce to the time-division region when all the channels have the same point-to-point capacity, which is the class studied in [7], we cannot conclude in a straightforward manner from the results in [7] that superposition coding is sub-optimal. Therefore, even though this work is based on one of the authors' insights in [7], the counterexample is nevertheless interesting.

Definition 1. A receiver Y_1 is said to be *more capable* than receiver Y_2 if the following holds: for every ϵ-error channel codebook[1] of size 2^{nR} from sender X to Y_2, there exists an ϵ'-error channel codebook of size $2^{n(R-\delta)}$ from sender X to Y_1 where $\delta, \epsilon' \rightarrow 0$ as $\epsilon \rightarrow 0$.

Remark: This is essentially equivalent to saying that Y_1 could decode any codebook that Y_2 could decode.

Korner and Marton[8] showed that the above definition is *equivalent* to the following:

Definition 2. A receiver Y_1 is said to be *more capable* than receiver Y_2 if the following holds: $I(X;Y_1) \geq I(X;Y_2), \ \forall p(x)$.

El Gamal[5] showed that superposition coding is optimal, i.e. achieves the capacity region, for any two-receiver more capable broadcast channel. Note that the definition of more capable induces a partial ordering among the receivers (or equivalently probability transition matrices), hence we are assuming here that the three receivers satisfy an induced more capable ordering.

Definition 3. A three-receiver more capable channel consists of a single sender X and three receivers Y_1, Y_2 and Y_3 such that the mutual information $I(X;Y_1)$, $I(X;Y_2)$ and $I(X;Y_3)$ satisfies $I(X;Y_1) \geq I(X;Y_2) \geq I(X;Y_3)$ for any fixed distribution $p(x)$ on X.

First we state(without the standard proof) the superposi-

---

[1] An ϵ-error codebook of size 2^{nR} consists of a set of codewords $x^n(m)$, $m \in [1 : 2^{nR}]$ and disjoint decoding regions $\mathcal{B}(m) \in \mathcal{Y}_2^n$ such that $\mathrm{P}(y_2^n \notin \mathcal{B}(m)|x^n(m)$ is transmitted$) < \epsilon, \ \forall m$.

tion coding region for the three receiver broadcast channel. We select the natural order induced by the more capable ordering for superposition coding.

Theorem 1. *The following set, $\mathcal{S}$, obtained by taking the union of all non-negative rate triples (R_1, R_2, R_3) satisfying*

$$R_3 \leq I(U_3; Y_3)$$
$$R_2 + R_3 \leq I(U_2; Y_2|U_3) + I(U_3; Y_3)$$
$$R_2 + R_3 \leq I(U_2, U_3; Y_2)$$
$$R_1 + R_2 + R_3 \leq I(X; Y_1|U_2, U_3) + I(U_2; Y_2|U_3) + I(U_3; Y_3)$$
$$R_1 + R_2 + R_3 \leq I(X; Y_1|U_2, U_3) + I(U_2, U_3; Y_2)$$
$$R_1 + R_2 + R_3 \leq I(X; Y_1|U_3) + I(U_3; Y_3)$$
$$R_1 + R_2 + R_3 \leq I(X; Y_1)$$

over all pairs of random variables (U_2, U_3) such that $(U_2, U_3) \rightarrow X \rightarrow (Y_1, Y_2, Y_3)$ forms a Markov chain is achievable for the private-messages-only case. We call this region the superposition coding region.

Superposition coding yields the optimal sum-rate.

Theorem 2. *Any set of achievable rates $R_1, \ldots, R_k$ for a k-receiver more capable channel must satisfy*

$$R_1 + R_2 + \cdots + R_k \leq \max_{p(x)} I(X; Y_1). \qquad (1)$$

Proof: We will prove the theorem for three-receiver more capable channels, the proof for more receivers shall follow with similar steps. Note that

$n(R_1 + R_2 + R_3) - n\epsilon_n$
$\leq I(M_1; Y_{11}^n) + I(M_2; Y_{21}^n|M_1) + I(M_3; Y_{31}^n|M_2, M_1)$
$\leq I(M_1; Y_{11}^n) + I(M_2; Y_{21}^n|M_1) + \sum_{i=1}^{n} I(X_i; Y_{3i}|M_2, M_1, Y_{31}^{i-1})$
$= I(M_1; Y_{11}^n) + \sum_{i=1}^{n} I(X_i; Y_{3i}|M_2, M_1, Y_{2i+1}^n, Y_{31}^{i-1})$
$+ \sum_{i=1}^{n} I(Y_{2i+1}^n; Y_{3i}|M_2, M_1, Y_{31}^{i-1}) + \sum_{i=1}^{n} I(M_2; Y_{2i}|M_1, Y_{2i+1}^n)$
$\overset{(a)}{=} I(M_1; Y_{11}^n) + \sum_{i=1}^{n} I(X_i; Y_{3i}|M_2, M_1, Y_{2i+1}^n, Y_{31}^{i-1})$
$+ \sum_{i=1}^{n} I(Y_{31}^{i-1}; Y_{2i}|M_2, M_1, Y_{2i+1}^n) + \sum_{i=1}^{n} I(M_2; Y_{2i}|M_1, Y_{2i+1}^n)$
$= I(M_1; Y_1^n) + \sum_{i=1}^{n} I(X_i; Y_{3i}|M_2, M_1, Y_{2i+1}^n, Y_{31}^{i-1})$
$+ \sum_{i=1}^{n} I(M_2, Y_{31}^{i-1}; Y_{2i}|M_1, Y_{2i+1}^n)$
$\overset{(b)}{\leq} I(M_1; Y_1^n) + \sum_{i=1}^{n} I(X_i; Y_{2i}|M_2, M_1, Y_{2i+1}^n, Y_{31}^{i-1})$
$+ \sum_{i=1}^{n} I(M_2, Y_{31}^{i-1}; Y_{2i}|M_1, Y_{2i+1}^n)$
$= I(M_1; Y_{11}^n) + \sum_{i=1}^{n} I(X_i; Y_{2i}|M_1, Y_{2i+1}^n) \qquad (2)$
$\overset{(c)}{=} \sum_{i=1}^{n} \left(I(X_i; Y_{2i}|M_1, Y_{2i+1}^n, Y_{11}^{i-1}) + I(M_1, Y_{2i+1}^n; Y_{1i}|Y_{11}^{i-1}) \right) \qquad (3)$

$\overset{(d)}{\leq} \sum_{i=1}^{n} \left(I(X_i; Y_{1i}|M_1, Y_{2i+1}^n, Y_{11}^{i-1}) + I(M_1, Y_{2i+1}^n; Y_{1i}|Y_{11}^{i-1}) \right)$
$= \sum_{i=1}^{n} I(X_i; Y_{1i}|Y_{11}^{i-1}) \leq n \max_{p(x)} I(X; Y_1)$

As $\epsilon_n \rightarrow 0$ when $n \rightarrow \infty$, we have

$$R_1 + R_2 + R_3 \leq \max_{p(x)} I(X; Y_1),$$

which, as we mentioned before, is achieved by transmitting only to the best receiver with superposition coding. In the above chain of inequalities we have used Fano's inequality, chain-rule for mutual information, data-processing inequality, Csiszar-sum lemma (equalities (a),(c)) (reproduced below) and the more capable ordering (inequalities (b),(d)). The data-processing inequalities used above come from the following Markov chain

$$(Y_{11}^{i-1}, Y_{2i+1}^n, Y_{31}^{i-1}, M_1, M_2, M_3) \rightarrow X_i \rightarrow (Y_{1i}, Y_{2i}, Y_{3i}).$$

Furthermore, noting the similarity of the second inequality and (2), a k-receiver proof could be generated by eliminating one receiver at a time. ∎

Lemma 1. *(Csiszar-sum Lemma) For any $p(U, y_{11}^n, y_{21}^n)$ we have*

$$\sum_{i=1}^{n} I(Y_{11}^{i-1}; Y_{2i}|U, Y_{2i+1}^n) = \sum_{i=1}^{n} I(Y_{2i+1}^n; Y_{1i}|U, Y_{11}^{i-1}).$$

Originally presented in [4], this is one of the most commonly used identities to derive outer bounds and converses for discrete memoryless broadcast channels.

II. THE SUB-OPTIMALITY OF SUPERPOSITION CODING

Theorem 2 implies that it is not possible to beat the sum-rate. What we could try is to beat the achievable region along some other directions, which is what we will do in the counter example.

We consider a DM-BC with $X \in \{0, 1\}$, $Y_1 \in \{0, 1, e\}$, $Y_2 \in \{0, 1, e\}$, and $Y_3 \in \{0, 1\}$, where the channel from X to Y_1, Y_2 and Y_3 are BEC(ϵ_1), BEC(ϵ_2) and BSC(p), respectively (see Figure 1). Let $p \in [0, \frac{1}{2}]$, $\epsilon_1 = 4p(1-p)$ and $\epsilon_2 = H(p)$, then from [9] we know that this is a three-receiver more capable channel.

In fact, for this particular case, Y_1 is also less noisy than Y_3 and Y_2 is a degraded version of Y_1. However Y_2 is only more capable than Y_3. Let $\mathcal{C}$ denote the true (as yet unknown) capacity region and $\mathcal{S}$ denote the superposition coding region.

Suppose the private message rates are R_1, R_2 and R_3 for receivers Y_1, Y_2 and Y_3, respectively. We try to maximize the following equation

$$T = \max_{(R_1, R_2, R_3) \in \mathcal{C}} \frac{R_1}{1 - \epsilon_1} + \frac{R_2 + R_3}{1 - \epsilon_2}.$$

Lemma 2. *For all $(R_1, R_2, R_3) \in \mathcal{S}$ we have*

$$\frac{R_1}{1 - \epsilon_1} + \frac{R_2 + R_3}{1 - \epsilon_2} \leq 1.$$

Proof: Note that if $(R_1, R_2, R_3) \in \mathcal{S}$ is in the achievable region, then so is $(R_1, R_2', 0)$, where $R_2' = R_2 + R_3$.
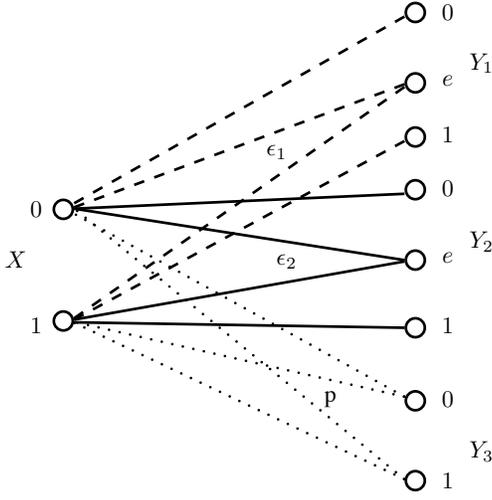
Fig. 1. 3-receiver more capable channel: $\epsilon_1 = 4p(1-p), \epsilon_2 = H(p)$

One can see this by plugging the choice into the region in Theorem 1. However the channel $X \to Y_1 \to Y_2$ is a degraded broadcast channel consisting of two BEC's and its capacity region consists of all non-negative pairs (R_1, R_2) satisfying

$$\frac{R_1}{1-\epsilon_1} + \frac{R_2}{1-\epsilon_2} \leq 1.$$

This well-known fact holds since for any $U \to X \to (Y_1, Y_2)$ we have $I(U; Y_2) = (1-\epsilon_2)I(U; X)$ and $I(X; Y_1|U) = (1-\epsilon_1)H(X|U)$ which can be obtained by routine manipulations. Hence

$$\frac{I(X; Y_1|U)}{1-\epsilon_1} + \frac{I(U; Y_2)}{1-\epsilon_2} = H(X|U) + I(U; X) = H(X) \leq 1. \quad \blacksquare$$

Lemma 2 implies that $T \leq 1$ if superposition coding were optimal.

Beating superposition coding region: Next, we will show that one can actually achieve $T > 1$. Instead of treating Y_2 as the second best receiver, we ignore Y_2 completely; i.e. it does not need to decode any message. This way the channel is transformed into a two-receiver less noisy broadcast channel with receivers Y_1, Y_3. Using superposition coding on this two-receiver channel, we can achieve $R_1 = I(X; Y_1|U), R_3 = I(U; Y_3)$ for any $U \to X \to (Y_1, Y_3)$. Hence

$$T \geq \max_{U \to X_1 \to (Y_1, Y_2, Y_3)} \frac{I(X; Y_1|U)}{(1-\epsilon_1)} + \frac{I(U; Y_3)}{1-\epsilon_2}$$
$$= \max_{U \to X_1 \to (Y_1, Y_2, Y_3)} \frac{I(X; Y_1|U)}{(1-\epsilon_1)} + \frac{I(U; Y_3)}{1-H(p)}.$$

Let $U \to X$ be a BSC with crossover probability $s, 0 < s < \frac{1}{2}$. Further, let $P(U = 0) = \frac{1}{2}$. We have,

$$T \geq \frac{I(X; Y_1|U)}{1-\epsilon_1} + \frac{I(U; Y_3)}{1-H(p)}$$
$$= \frac{(1-\epsilon_1)H(s)}{1-\epsilon_1} + \frac{1-H(s*p)}{1-H(p)}$$
$$= H(s) + \frac{1-H(s*p)}{1-H(p)}. \quad (4)$$

By setting p and s to $\frac{1}{10}$, we see that

$$T \geq H(s) + \frac{1-H(s*p)}{1-H(p)}$$
$$= H(0.1) + \frac{1-H(0.18)}{1-H(0.1)}$$
$$\geq 1.07.$$

Therefore, superposition coding region is not optimal for the three-receiver more capable channel.

A. An achievable rate region for the three-receiver more capable broadcast channel

Since the sum-capacity is bounded by what one can transmit to the receiver Y_1, a natural guess would be to allow the Y_1 to decode all the messages. Now, one can employ Marton's binning scheme to transmit non-nested messages to receivers Y_2 and Y_3.

Theorem 3. *Consider a three receiver more capable broadcast channel with Y_1 being the most capable receiver and Y_3 the least. Then any rate triple (R_1, R_2, R_3) satisfying*

$$R_2 \leq I(U_2, W; Y_2)$$
$$R_3 \leq I(U_3, W; Y_3)$$
$$R_2 + R_3 \leq \min\{I(W; Y_2), I(W; Y_3)\} + I(U_2; Y_2|W)$$
$$\quad + I(U_3; Y_3|W) - I(U_2; U_3|W)$$
$$R_1 + R_2 + R_3 \leq I(X; Y_1)$$
$$R_1 + R_2 + R_3 \leq I(U_2, W; Y_2) + I(X; Y_1|U_2, W)$$
$$R_1 + R_2 + R_3 \leq I(U_3, W; Y_3) + I(X; Y_1|U_3, W)$$
$$R_1 + R_2 + R_3 \leq \min\{I(W; Y_2), I(W; Y_3)\} + I(U_2; Y_2|W)$$
$$\quad + I(U_3; Y_3|W) + I(X; Y_1|U_2, U_3, W) - I(U_2; U_3|W)$$
$$R_1 + 2R_2 + 2R_3 \leq I(U_2, W; Y_2) + I(U_3, W; Y_3)$$
$$\quad + I(X; Y_1|W) - I(U_2; U_3|W)$$
$$2R_1 + 2R_2 + 2R_3 \leq I(U_2, W; Y_2) + I(U_3, W; Y_3)$$
$$\quad + I(X; Y_1|U_2, U_3, W) + I(X; Y_1|W) - I(U_2; U_3|W)$$

is achievable for any $(W, U_2, U_3) \to X \to (Y_1, Y_2, Y_3)$.

Proof: The proof follows standard techniques of random binning, superposition coding, and jointly typical decoding. Receiver Y_2 decodes U_2^n, W^n, receiver Y_3 decodes U_3^n, W^n, and receiver Y_1 decodes W^n, U_2^n, U_3^n, X^n. The analysis is routine and straightforward (but messy) and hence is omitted. $\quad \blacksquare$

Remark 1. A constraint

$$0 \leq I(U_2; Y_2|W) + I(U_3; Y_3|W) - I(U_2; U_3|W)$$

is obtained during the process of Fourier-Motzkin Elimination, but it is easy to see that this condition is redundant to the computation of the region.

In the event where Y_1 is less noisy than both Y_2 and Y_3 pairwise(as in the counter example), the achievable region in Theorem 3 reduces to

$$R_1 + R_2 + R_3 \leq \min\{I(W; Y_2), I(W; Y_3)\} + I(U_2; Y_2|W)$$
$$\quad + I(U_3; Y_3|W) + I(X; Y_1|U_2, U_3, W) - I(U_2; U_3|W)$$
$$R_2 + R_3 \leq \min\{I(W; Y_2), I(W; Y_3)\} + I(U_2; Y_2|W)$$
$$\quad + I(U_3; Y_3|W) - I(U_2; U_3|W) \quad (5)$$
$$R_2 \leq I(U_2, W; Y_2)$$
$$R_3 \leq I(U_3, W; Y_3)$$

Also, since Y_3 is *essentially less noisy*[9] than Y_2 in the counter example, by symmetrization argument[2] we can further assume $P(X = 0) = \frac{1}{2}$, and hence $I(U_2, W; Y_3) \geq I(U_2, W; Y_2)$. Therefore we can set $\tilde{W} = (U_2, W), \tilde{U}_2 = \emptyset, \tilde{U}_3 = U_3$ to obtain the following achievable region:

Theorem 4. *For the channel depicted in 1, the union of rate triples* (R_1, R_2, R_3) *satisfying*

$$R_1 + R_2 + R_3 \leq I(\tilde{W}; Y_2) + I(\tilde{U}_3; Y_3|\tilde{W}) + I(X; Y_1|\tilde{U}_3, \tilde{W})$$
$$R_2 + R_3 \leq I(\tilde{W}; Y_2) + I(\tilde{U}_3; Y_3|\tilde{W})$$
$$R_2 \leq I(\tilde{U}_3, W; Y_2)$$

over all $(\tilde{W}, \tilde{U}_3) \to X \to (Y_1, Y_2, Y_3)$ *is achievable.*

This is just superposition coding by treating Y_3 as the second best receiver. We will prove in the next section that this is indeed the capacity region for the counterexample.

B. The capacity region of the counterexample

In this section we show that the region presented in Theorem 4 is indeed the capacity region for the counterexample.

Note that it suffices to just show a converse to Theorem 4 to establish the capacity region. The arguments are reasonably routine once the identifications of the auxiliaries have been made:

Identify $\tilde{U}_{3i} = (M_3, Y_{11}^{i-1})$ and $\tilde{W}_i = (M_2, Y_{3i+1}^n, Y_{21}^{i-1})$. Observe that

$n(R_1 + R_2 + R_3) - n\epsilon_n$
$\leq I(M_2; Y_{21}^n) + I(M_3; Y_{31}^n|M_2) + I(M_1; Y_{11}^n|M_2, M_3)$
$= \sum_{i=1}^n I(M_1; Y_{1i}|M_2, M_3, Y_{11}^{i-1}) + I(M_3; Y_{31}^n|M_2) + I(M_2; Y_{21}^n)$
$\leq I(M_2; Y_{21}^n) + \sum_{i=1}^n \Big(I(M_1; Y_{1i}|M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1})$
$\quad + I(Y_{3i+1}^n; Y_{1i}|M_2, M_3, Y_{11}^{i-1}) + I(M_3; Y_{3i}|Y_{3i+1}^n, M_2) \Big)$
$\overset{(a)}{=} I(M_2; Y_{21}^n) + \sum_{i=1}^n \Big(I(M_1; Y_{1i}|M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1})$
$\quad + I(M_3, Y_{11}^{i-1}; Y_{3i}|Y_{3i+1}^n, M_2) \Big)$
$\leq \sum_{i=1}^n I(M_2; Y_{2i}|Y_{21}^{i-1}) + I(M_3, Y_{11}^{i-1}; Y_{3i}|Y_{3i+1}^n, M_2, Y_{21}^{i-1})$
$\quad + I(Y_{21}^{i-1}; Y_{3i}|Y_{3i+1}^n, M_2) + I(M_1; Y_{1i}|M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1})$
$\overset{(b)}{=} \sum_{i=1}^n I(Y_{3i+1}^n, M_2; Y_{2i}|Y_{21}^{i-1}) + I(M_1; Y_{1i}|M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1})$
$\quad + I(M_3, Y_{11}^{i-1}; Y_{3i}|Y_{3i+1}^n, M_2, Y_{21}^{i-1})$
$\leq \sum_{i=1}^n I(Y_{3i+1}^n, Y_{21}^{i-1}, M_2; Y_{2i}) + I(X_i; Y_{1i}|M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1})$
$\quad + I(M_2, Y_{11}^{i-1}; Y_{3i}|Y_{3i+1}^n, M_2, Y_{21}^{i-1})$

$\overset{(c)}{=} \sum_{i=1}^n I(Y_{3i+1}^n, Y_{21}^{i-1}, M_2; Y_{2i})$
$\quad + I(M_2, Y_{11}^{i-1}; Y_{3i}|Y_{3i+1}^n, M_2, Y_{21}^{i-1})$
$\quad + I(X_i; Y_{1i}|M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1}, Y_{21}^{i-1})$
$= \sum_{i=1}^n I(X_i; Y_{1i}|\tilde{U}_{3i}, \tilde{W}_i) + I(\tilde{U}_{3i}; Y_{3i}|\tilde{W}_i) + I(\tilde{W}_i; Y_{2i}).$

In the above the usual toolset: Fano's inequality, Csiszar sum-lemma (steps $(a), (b)$), data-processing inequality, and chain rule of mutual information. All the data processing inequalities come from the following Markov chain:

$$(M_1, M_2, M_3, Y_{3i+1}^n, Y_{11}^{i-1}, Y_{21}^{i-1}) \to X_i \to (Y_{1i}, Y_{2i}, Y_{3i}).$$

The equality (c) comes from the fact that Y_2 is a degraded version[3] of Y_1 and hence

$$Y_{21}^{i-1} \to Y_{11}^{i-1} \to (M_1, M_2, M_3, Y_{3i+1}^n, X_i, Y_{1i}, Y_{2i}, Y_{3i})$$

is Markov. Finally in the usual manner, let Q be an independent random variable distributed uniformly in $[1 : n]$ and set $\tilde{W} = (\tilde{W}_Q, Q), \tilde{U}_3 = \tilde{U}_{3Q}, X = X_Q$.

The other inequalities follow a similar line (but is simpler) of reasoning. Observe that

$n(R_2 + R_3) - n\epsilon_n$
$\leq I(M_2; Y_{21}^n) + I(M_3; Y_{31}^n|M_2)$
$= \sum_{i=1}^n I(M_2; Y_{2i}|Y_{21}^{i-1}) + I(M_3; Y_{3i}|M_2, Y_{3i+1}^n)$
$\leq \sum_{i=1}^n I(M_2; Y_{2i}|Y_{21}^{i-1}) + I(Y_{21}^{i-1}; Y_{3i}|M_2, Y_{3i+1}^n)$
$\quad + I(M_3; Y_{3i}|M_2, Y_{3i+1}^n, Y_{21}^{i-1})$
$= \sum_{i=1}^n I(M_2, Y_{3i+1}^n; Y_{2i}|Y_{21}^{i-1}) + I(M_3; Y_{3i}|M_2, Y_{3i+1}^n, Y_{21}^{i-1})$
$\leq \sum_{i=1}^n I(M_2, Y_{3i+1}^n, Y_{21}^{i-1}; Y_{2i}) + I(M_3, Y_{11}^{i-1}; Y_{3i}|M_2, Y_{3i+1}^n, Y_{21}^{i-1})$
$= \sum_{i=1}^n I(\tilde{W}_i; Y_{2i}) + I(\tilde{U}_{3i}; Y_{3i}|\tilde{W}_i).$

The last inequality (on R_2) is very straightforward with this identification and is omitted. This completes the proof for the capacity region of the channel in Figure 1.

Remark 2. It may appear a bit strange to see that even though superposition coding in the natural more-capable ordering (i.e. Y_1 better than Y_2 better than Y_3) is suboptimal, a re-ordering of the receivers, i.e. (i.e. Y_1 better than Y_3 better than Y_2) could make superposition coding optimal again. But of course, this is a carefully chosen counterexample and hence the peculiar situation. It is natural to ask whether there exists a three-receiver more capable broadcast channel where superposition coding is not optimal with either ordering. We will show such an example (a minor perturbation of the example in Figure 1) in the next section.

---

[2]Symmetrization argument can be found in [9], [11], [7] or in Chapter 5 of [1]. The main purpose of this argument is to show that points on the boundary for a binary input symmetric output channels can be computed using distributions that satisfy $P(X = 0) = \frac{1}{2}$.

[3]Since capacity region depends only on the marginals $p(y_1|x), p(y_2|x)$ and $p(y_3|x)$, we can assume without loss of generality that Y_2 is a physically degraded version of Y_1 in the example in Figure 1.

C. A modified counterexample

Consider the same channel as in Figure 1. Set $\epsilon_1 = 4 * (0.1) * 0.9 = 0.36$, $\epsilon_2 = H(0.1)$. Slightly change the value of p from 0.1 to 0.11. Clearly since the new receiver Y_3 is a degraded version of the old receiver Y_3 (which was $BSC(0.1)$), this setting is still a three-receiver more capable channel. As before, we try to maximize

$$T = \max_{(R_1, R_2, R_3) \in \mathcal{C}} \frac{R_1}{1 - \epsilon_1} + \frac{R_2 + R_3}{1 - \epsilon_2}.$$

If superposition coding in the more capable ordering were optimal, then again the same arguments would imply that $T \leq 1$. However, if we ignore Y_2 again and use superposition coding between receivers Y_1 and Y_3, we can obtain, taking $U \to X$ to be $BSC(0.1)$ with uniform distribution,

$$\begin{aligned}
T &\geq \frac{I(X; Y_1 | U)}{1 - \epsilon_1} + \frac{I(U; Y_3)}{1 - \epsilon_2} \\
&= H(0.1) + \frac{1 - H(0.11 * 0.1)}{1 - H(0.1)} \\
&\geq 1.039.
\end{aligned}$$

Hence, superposition coding in the more capable ordering is not optimal.

To show that superposition coding in the Y_1, Y_3, Y_2 ordering is not optimal, we can maximize

$$T = \max_{(R_1, R_2, R_3) \in \mathcal{C}} R_2 + R_3.$$

If superposition coding in Y_1, Y_3, Y_2 ordering were optimal, this would be the same as maximizing R_3, whose maximum is $1 - H(0.11) \approx 0.501$. On the other hand, by just transmitting to receiver Y_2 we can obtain $R_2 = 1 - \epsilon_2 = 1 - H(0.1) \geq 0.531$. Thus superposition coding in the Y_1, Y_3, Y_2 order is also not optimal for this modified counter example.

Remark 3. The converse in the last section continues to hold for this modified setting. However, since Y_3 is no longer an essentially less noisy receiver than Y_2, the achievability of the region depicted by Theorem 4 fails to hold.

A natural guess for the capacity region in this modified counterexample would be given by the constraints in Equations (5).

III. Conclusion

In this paper, we showed that superposition coding does not achieve the capacity region for a three-receiver broadcast channel. In fact, we presented a counterexample where capacity could be achieved by treating the least capable receiver as the intermediate receiver and the intermediately-capable receiver the worst receiver. The main purpose of this counterexample is to show that more capable is a very weak ordering that does not preserve the nested decoding properties a less noisy ordering would, at least in the three receiver case. Then we produced a modified counterexample to show that superposition coding (in whatever order one wishes) cannot yield the capacity region for general three-receiver more capable broadcast channels.

On the other hand, we showed that one can achieve the sum-rate capacity for any k-receiver more capable channel by just transmitting to the best receiver. Motivated by this result we presented certain achievable regions for the three-receiver more capable broadcast channels.

References

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

[2] P F Bergmans, *Coding theorem for broadcast channels with degraded components*, IEEE Trans. Info. Theory **IT-15** (March, 1973), 197–207.

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

[4] I Csizár and J Körner, *Broadcast channels with confidential messages*, IEEE Trans. Info. Theory **IT-24** (May, 1978), 339–348.

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

[6] R G Gallager, *Capacity and coding for degraded broadcast channels*, Probl. Peredac. Inform. **10(3)** (1974), 3–14.

[7] 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).

[8] J Körner and K Marton, *Comparison of two noisy channels*, Topics in Inform. Theory(ed. by I. Csiszar and P.Elias), Keszthely, Hungary (August, 1975), 411–423.

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

[10] C. Nair and Z.V. Wang, *The capacity region of the three receiver less noisy broadcast channel*, Information Theory, IEEE Transactions on **57** (2011), no. 7, 4058 –4062.

[11] Chandra Nair, Abbas El Gamal, and Yeow-Khiang Chia, *An achievability scheme for the compound channel with state noncausally available at the encoder*, CoRR **abs/1004.3427** (2010).