Sub-optimality of superposition coding region for three receiver broadcast channel with two degraded message sets

Chandra Nair
Dept. of Information Engineering
The Chinese University of Hong Kong,
Hong Kong

Mehdi Yazdanpanah
Dept. of Information Engineering
The Chinese University of Hong Kong,
Hong Kong

***Abstract*—In this article, we resolve open problem 8.2 in [1]. We show that superposition coding is sub-optimal for a three receiver broadcast channel with two message sets (M_0, M_1) where two of the three receivers need to decode messages (M_0, M_1) while the remaining one just needs to decode the message M_0.**

INTRODUCTION

Background

In a classical paper, by characterizing image-sizes of sets under two noisy channels, Korner and Marton [2] established that superposition coding achieves the capacity region for a two-receiver broadcast channel with degraded message sets, i.e. one of the receivers wants to decode one message and the other wants to decode both the messages. The characterization of the image sizes of sets under three noisy channels, remains a central unsolved problem (open problem 2 in [3]).

In [4] it was shown that superposition coding region was strictly sub-optimal for a three receiver broadcast channel with two message sets (M_0, M_1) where one of the three receivers, say Y, needs to decode messages (M_0, M_1) while the remaining two receivers, $Z, \hat{Z}$, only need to decode the message M_0. This was done through an indirect decoding idea which is presented in Section 8.2 of [1]. The basic idea of indirect decoding is to split the private message into (potentially) three parts; one each to help the receivers Z and $\hat{Z}$ and the rest as a private message. Receivers Z and $\hat{Z}$ utilize the parts of the private message to aid in recovering the common message M_0. There are also alternate ways to look at this scheme which do not explicitly need this indirect decoding but use a case based decoding analysis.

The setting we are considering here is different (like a dual): here, we have two receivers, say $Y, \hat{Y}$, that need to decode messages (M_0, M_1) while the remaining receiver, Z, needs to only decode the message M_0. In this case, the splitting idea of before does not improve on superposition coding region, and indeed the former intuition suggests that superposition coding may be optimal. In [5], the authors presented a channel where superposition coding is optimal and developed a tailor-made converse using novel arguments. Preliminary numerical simulations indicated a sub-additivity, which would then imply optimality of superposition coding region, which was conjectured by one of the authors in [6]. The result here, as a corollary, also disproves the conjecture.

Owing to the lack of evidence that the superposition coding region can be improved or of ideas on how to improve the superposition coding region, determining the optimality of superposition coding region is stated as open problem (8.2) in [1].

In this paper, we show that superposition coding region is strictly sub-optimal for this class by exhibiting a particular channel where the two-letter extension of superposition coding region improves on the single-letter region. It is hoped that by further examining this counter-example or similar counterexamples, one may get a better understanding behind the sub-optimality of superposition coding region, and of the interplay between the image sizes of sets under three noisy channels. Further, the computation of the superposition coding region as well as the two-letter region contains ideas, some developed recently in similar contexts by the authors, that may be of independent interest.

Problem Setting

A sender X, who has access to two independent messages (M_0, M_1), uniformly distributed over sets $[1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$, wishes to encode the message into a sequence X^n, which is then transmitted over a discrete memoryless broadcast channel, $W^{\otimes n}(y, \hat{y}, z|x)$. Three receivers which receive sequences $Y^n, \hat{Y}^n, Z^n$, respectively, wish to decode messages $(M_0, M_1), (M_0, M_1)$, and M_0. The setting is depicted in Figure 1.

The following region is achievable.

Bound 1 (Superposition coding achievable region)**.** *The union of the set of rate pairs (R_0, R_1) satisfying*

$$\begin{aligned} R_0 &\leq I(U; Z) \\ R_0 + R_1 &\leq I(U; Z) + \min\{I(X; Y|U), I(X; \hat{Y}|U)\} \\ R_0 + R_1 &\leq \min\{I(X; Y), I(X; \hat{Y})\} \end{aligned}$$

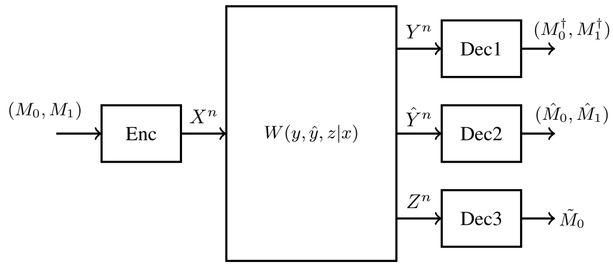

Fig. 1. Three receiver broadcast channel with two degraded message sets

where the union is taken over all pairs of random variables (U, X) *such that* $|\mathcal{U}| \leq |\mathcal{X}| + 2$ *and* $U \to X \to (Y, \hat{Y}, Z)$ *forms a Markov chain is achievable.*

The main result of this paper is that the above region can be strictly smaller than the capacity region.

I. Strict sub-optimality of the superposition coding inner bound

The example that shows the strict sub-optimality is a reversely degraded multi-level broadcast erasure channel, belonging to the class depicted in Figure 2.

Denote the erasure probability of each sub-channel by

$$X_a \to Y_a : BEC(e_a),\ X_b \to Y_b : BEC(e_b)$$
$$X_a \to \hat{Y}_a : BEC(\hat{e}_a),\ X_b \to \hat{Y}_b : BEC(\hat{e}_b)$$
$$X_a \to Z_a : BEC(f_a),\ X_b \to Z_b : BEC(f_b).$$

The order of channels in Figure 2 implies that $\hat{e}_a \geq f_a \geq e_a$ and $e_b \geq f_b \geq \hat{e}_b$.

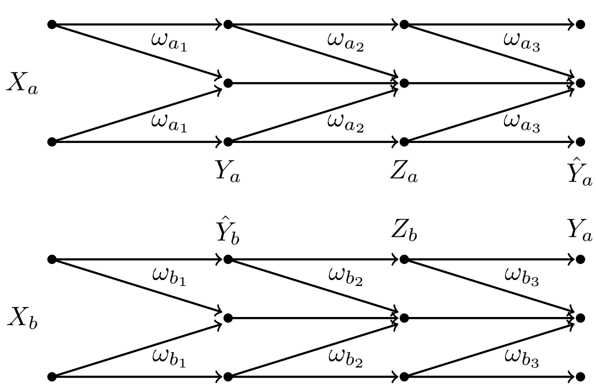

Fig. 2. Product broadcast erasure channel

One of the key contributions of this paper is the technique of computation of the superposition inner bound for this class of channels (product broadcast erasure channels).

Proposition 1. *For the broadcast channel in Figure 2, it suffices to consider uniform distribution on* $X = (X_a, X_b)$ *to obtain the superposition coding region (Bound 1).*

Proof. This is basically a symmetrization argument á la the one in [7]. The argument is also presented in a slightly more general fashion that is required for the proof of the proposition, in that the number of product components can be larger than 2.

Let π be either of the two permutations of $\{0, 1\}$. By an abuse of notation, let π also denote the induced permutation of $\{0, E, 1\}$ by mapping E to E. Then note for any generic symmetric erasure channel as in Figure 2, $W(Y = \pi(y)|X = \pi(x)) = W(Y = y|X = x)$. Now consider a product erasure channel structure where the inputs are $x_1, ..., x_k$ and let the corresponding outputs be $y_1, .., y_k$. Given any probability distribution $p(x_1, ..., x_k)$, let $p(y_1, .., y_k)$ denote the induced output distribution. Let $\pi_1, ..., \pi_k$ be any set of permutations of $\{0, 1\}$; and let $r(x_1, .., x_k) = p(\pi_1(x_1), ..., \pi_k(x_k))$ denote an induced input distribution. Then note that the induced output distribution is given by

$$\begin{aligned} r(y_1, ..., y_k) &= \sum_{x_1, .., x_k} r(x_1, .., x_k) \prod_{i=1}^{k} W_i(y_i|x_i) \\ &\overset{(a)}{=} \sum_{x_1, .., x_k} r(x_1, .., x_k) \prod_{i=1}^{k} W_i(\pi_i(y_i)|\pi_i(x_i)) \\ &= \sum_{x_1, .., x_k} p(\pi_1(x_1), .., \pi_k(x_k)) \prod_{i=1}^{k} W_i(\pi_i(y_i)|\pi_i(x_i)) \\ &= p(\pi_1(y_1), ..., \pi_1(y_k)). \end{aligned}$$

In the above, (a) follows from the symmetry of the component channels. Thus the output probability vector $r(\mathbf{y})$ is just a permutation of the original probability vector $p(\mathbf{y})$, and hence entropy of $Y_1, .., Y_k$ remains unchanged.

Given a joint distribution $p(u, x_1, .., x_k)$ (or on (U, X_a, X_b) as is this case), let Q denote a uniform random variable distributed over $[1 : 2^k]$. Identify with each Q a unique collection of permutations $\pi_1^q, .., \pi_k^q$. (for instance, using the binary representation). Define $\tilde{U} = (Q, U)$ and consider a joint distribution defined as follows:

$$r((q, u), x_1, .., x_k) = \frac{1}{2^k} p(u, \pi_1^q(x_1), ..., \pi_k^q(x_k)).$$

Note that the induced distributions on $(X_1, .., X_k)$ is uniform and that

$$\begin{aligned} r(x_1, .., x_k|(q, u)) &= p(\pi_1^q(x_1), ..., \pi_k^q(x_k)|u) \\ r(y_1, .., y_k|(q, u)) &= p(\pi_1^q(y_1), ..., \pi_k^q(y_k)|u) \end{aligned}$$

where the second equality follows the argument presented earlier.

Hence note the following inequalities for any collection of outputs of symmetric channels

$$\begin{aligned} H_p(Y_1, .., Y_k) &\overset{(a)}{\leq} H_r(Y_1, .., Y_k) \\ H_p(Y_1, .., Y_k|U) &\overset{(b)}{=} H_r(Y_1, .., Y_k|U, Q) \\ H_p(Y_1, .., Y_k|X_1, .., X_k, U) &\overset{(b)}{=} H_r(Y_1, .., Y_k|X_1, ..., X_k, U, Q), \end{aligned}$$

where (a) follows since uniform input distribution maximizes entropy for symmetric erasure channels, and equalities denoted by (b) is due to the fact that permutations of probability vectors do not change its entropy. Thus every term occurring in the superposition coding region is non-decreasing by virtue of this symmetrization using Q, which induces a uniform distribution on X. ☐

The following corollary is immediate.

Corollary 1. *Superposition coding region for the product broadcast erasure channel in Fig 2 is the intersection of* $\{(R_0, R_1)|R_0 + R_1 < \min(C_Y, C_{\hat{Y}})\}$ *and the region* $\mathcal{S}$ *defined as the union of the set of rate pairs* (R_0, R_1) *satisfying*

$$R_0 < I(U;Z)$$
$$R_0 + R_1 < I(U;Z) + \min\{I(X;Y|U), I(X;\hat{Y}|U)\}$$

where the union is taken over all pairs of random variables (U, X) *such that* $|\mathcal{U}| \leq |\mathcal{X}| + 2$, $U \to X \to (Y, \hat{Y}, Z)$ *forms a Markov chain, and* $X = (X_a, X_b)$ *is uniformly distributed.* $C_Y = 1 - e_a + 1 - e_b$ *and* $C_{\hat{Y}} = 1 - \hat{e}_a + 1 - \hat{e}_b$ *are the capacities for channels* $W(y|x)$ *and* $W(\hat{y}|x)$.

Thus the key difficulty in computation of the superposition coding region is reduced to computation of region $\mathcal{S}$.

Proposition 2. *For any* $\lambda > 1$, $\mathsf{SH}_\lambda^{\mathcal{S}} := \max_{\mathcal{S}}(\lambda R_0 + R_1)$ *is given by*

$$\lambda C_Z + \min_{\alpha \in [0,1]} \max_{p(x)} \{\alpha I(X;Y) + \bar{\alpha} I(X;\hat{Y}) - \lambda I(X;Z)\},$$

where C_Z *is the capacity for channel* $W(z|x)$.

Proof. We know that it suffices to consider X to be uniformly distributed. Thus $\max_{\mathcal{S}}(\lambda R_0 + R_1)$ is given by

$$\max_{p(u|x)} \left(\lambda I(U;Z) + \min\{I(X;Y|U), I(X;\hat{Y}|U)\}\right),$$

where X is uniform.

An immediate application of min-max result, Corollary 2 in [8], yields that

$$\begin{aligned} &\max_{p(u|x)} \left(\lambda I(U;Z) + \min\{I(X;Y|U), I(X;\hat{Y}|U)\}\right) \\ &= \min_{\alpha \in [0,1]} \max_{p(u|x)} \left(\lambda I(U;Z) + \alpha I(X;Y|U) + (1-\alpha) I(X;\hat{Y}|U)\right). \end{aligned}$$

Noting that $I(U;Z) = I(X;Z) - I(X;Z|U) = C_Z - I(X;Z|U)$ (since uniform X achieves C_Z), we re-write the above as

$$\begin{aligned} &\min_{\alpha \in [0,1]} \max_{p(u|x)} \left(\lambda I(U;Z) + \alpha I(X;Y|U) + (1-\alpha) I(X;\hat{Y}|U)\right) \\ &= \min_{\alpha \in [0,1]} \max_{p(u|x)} \Big(\lambda C_Z + \alpha I(X;Y|U) + (1-\alpha) I(X;\hat{Y}|U) \\ &\qquad - \lambda I(X;Z|U)\Big) \\ &= \lambda C_Z + \min_{\alpha \in [0,1]} \max_{p(x)} \{\alpha I(X;Y) + \bar{\alpha} I(X;\hat{Y}) - \lambda I(X;Z)\}. \end{aligned}$$

The last equality follows by applying the symmetrization argument to the constant U where $p(x|u)$ is the distribution that maximizes the quantity $\alpha I(X;Y) + \bar{\alpha} I(X;\hat{Y}) - \lambda I(X;Z)$. □

Remark 1. Note that the above two propositions regarding computation of superposition coding region for product broadcast channels apply to all symmetric (appropriately defined) channels and does not depend on the fact that the symmetric channel under consideration is an erasure channel. The next proposition on the other hand uses the erasure nature of the component channels.

The following lemma should be well-known but is provided here for completeness.

Lemma 1. *Consider a product erasure channel mapping* $X_1, .., X_k$ *to* $Y_1, .., Y_k$ *with erasure probabilities* $\epsilon_1, ..., \epsilon_k$. *Then*

$$I(X_1, ..., X_k; Y_1, ..., Y_k) = \sum_{S \subseteq [1:k]} \left(\prod_{i \in S}(1-\epsilon_i) \prod_{j \notin S} \epsilon_j\right) H(X_S),$$

where $X_S = (X_i : i \in S)$.

Proof. The proof is a simple exercise by induction on k, for instance, and is briefly outlined below. Observe that $k = 1$ is immediate. Note that $I(X_1, ..., X_k; Y_1, ..., Y_k) = I(X_1, ..., X_{k-1}; Y_1, ..., Y_{k-1}) + I(X_k; Y_k|Y_1, , , Y_{k-1})$. A simple calculation yields that

$$\begin{aligned} &I(X_k; Y_k|Y_1, ..., Y_{k-1}) \\ &= \sum_{S_1 \subseteq [1:k-1]} \left(\prod_{i \in S_1}(1-\epsilon_i) \prod_{j \notin S_1} \epsilon_j\right) (1-\epsilon_k) H(X_k|X_{S_1}). \end{aligned}$$

Combining this term with induction hypothesis completes the proof. □

Remark 2. Combining Proposition 2 with Lemma 1 shows that computation of the superposition coding region for a product erasure broadcast channel reduces to computation of the maximum of a linear combination of entropic-vectors, a subset of $\mathbb{R}^{2^k-1}$ generated by subsets of k binary random variables. When $k = 2$, for every $\alpha \in [0, 1]$, we wish to maximize a linear combination of the vector $[H(X_a), H(X_b), H(X_a, X_b)]$, where the co-efficients are determined using Proposition 2 and Lemma 1. Note X_a and X_b are binary random variables.

A specific example

There are many examples (though each one takes a fair bit of computer search) where two-letter superposition coding region beats the single-letter superposition coding region. However, below we produce a concrete example where using the machinery developed above we are able to explicitly demonstrate the gap between 1-letter and 2-letter regions.

Theorem 1. *For the three receiver broadcast channel with parameters*

$$\begin{aligned} e_a &= 1/2 & \hat{e}_a &= 1 & f_a &= 17/22 \\ e_b &= 1/2 & \hat{e}_b &= 0 & f_b &= 9/34 \end{aligned}$$

the non-trivial boundary (i.e. excluding axis) of the S.C region is determined by the two lines:

$$R_0 + R_1 = 1 \text{ and } \frac{11}{10} R_0 + R_1 = \frac{18}{17}.$$

The non-trivial boundary of the 2-letter S.C. region is determined by the two lines:

$$R_0 + R_1 = 1 \text{ and } \frac{484}{435} R_0 + R_1 = \frac{528}{493}.$$

Proof. From Corollary 1, and $C_Y = C_{\hat{Y}} = 1$, the line $R_0 + R_1 = 1$ is immediate. To compute the S.C region (1-letter or 2-letter) it remains to compute the region $\mathcal{S}$.

Computation of the 1-letter region.
For the 1-letter (usual) S.C. region, we first show that any $(R_0, R_1) \in \mathcal{S}$ satisfies

$$\frac{11}{10}R_0 + R_1 \leq \frac{18}{17}.$$

Since $C_Z = (1 - f_a) + (1 - f_b) = \frac{180}{187}$, from Proposition 2 (taking $\alpha = \frac{1}{2}$), the inequality above will follow if we show that

$$\frac{1}{2}I(X;Y) + \frac{1}{2}I(X;\hat{Y}) - \frac{11}{10}I(X;Z) \leq 0 \quad \forall p(x).$$

Here $X = (X_a, X_b), Y = (Y_a, Y_b), \hat{Y} = (\hat{Y}_a, \hat{Y}_b)$ and $Z = (Z_a, Z_b)$. Expanding the left hand side using Lemma 1 and substituting our choices of erasures yields

$$\frac{1}{2}I(X;Y) + \frac{1}{2}I(X;\hat{Y}) - \frac{11}{10}I(X;Z) = -\frac{1}{17}H(X_b|X_a),$$

implying the upper bound.

Next we show that the intersection of the two lines $R_0 + R_1 = 1$ and $\frac{11}{10}R_0 + R_1 = \frac{18}{17}$ belongs to the super-position coding region (completing the characterization).

Let U be a ternary random variable such that $\mathrm{P}(U = 0) = 13/34$, $\mathrm{P}(U = 1) = 7/34$, $\mathrm{P}(U = 2) = 14/34$. Conditionals are given by:

$$\begin{aligned}
(X_a, X_b)|(U = 0) &= (0, 0) \\
(X_a, X_b)|(U = 1) &= (M, 0) \\
(X_a, X_b)|(U = 2) &= (M, M),
\end{aligned}$$

where M is an unbiased binary random variable. Let Q be a random variable that symmetrizes the distribution of X (in the sense of the proof of Proposition 1) and let $\tilde{U} = (U, Q)$. Substituting $(\tilde{U}, X)$ into Bound 1 yields:

$$\begin{aligned}
R_0 &\leq I(\tilde{U}; Z_a, Z_b) = \frac{10}{17} \\
R_0 + R_1 &\leq \min\{I(X_a, X_b; Y|\tilde{U}), I(X_a, X_b; \hat{Y}_a, \hat{Y}_b|\tilde{U})\} \\
&\quad + I(\tilde{U}; Z_a, Z_b) = 1 \\
R_0 + R_1 &\leq \min\{I(X_a, X_b; Y), I(X_a, X_b; \hat{Y}_a, \hat{Y}_b)\} = 1.
\end{aligned}$$

Thus $(R_0, R_1) = (\frac{10}{17}, \frac{7}{17})$ lying at the intersection of the two lines $R_0 + R_1 = 1$ and $\frac{11}{10}R_0 + R_1 = \frac{18}{17}$ belongs to the superposition coding region. This establishes the superposition coding region.

Computation of the 2-letter region.
The proof mimics the 1-letter case. We first show that any $(R_0, R_1) \in \mathcal{S}_2$ satisfies

$$\frac{484}{435}R_0 + R_1 \leq \frac{528}{493}.$$

Since $C_Z = (1 - f_a) + (1 - f_b) = \frac{180}{187} = \frac{528 \times 435}{493 \times 484}$, from Proposition 2 (taking $\alpha = \frac{88}{174}$), the inequality above will follow if we show that

$$\frac{88}{174}I(X;Y) + \frac{86}{174}I(X;\hat{Y}) - \frac{484}{435}I(X;Z) \leq 0 \quad \forall p(x).$$

In the above, $X = (X_{a1}, X_{b1}, X_{a2}, X_{b2})$ and similarly for others. Expanding the left hand side using Lemma 1 and substituting our choices of erasures yields

$$\begin{aligned}
&-\frac{17}{174}\,I(X_{b1}; X_{b2}) - \frac{19}{2958}\big(I(X_{b1}; X_{b2}|X_{a1}) + I(X_{b1}; X_{b2}|X_{a2})\big) \\
&-\frac{2}{29}\,I(X_{a1}; X_{a2}|X_{b1}X_{b2}) - \frac{2543}{50286}\,I(X_{b1}; X_{b2}|X_{a1}X_{a2}) \\
&-\frac{35}{493}\big(H(X_{b1}|X_{a_1}X_{a2}X_{b2}) + H(X_{b2}|X_{a_1}X_{b1}X_{a2})\big) \\
&-\frac{1}{174}\big(I(X_{a1}; X_{b1}|X_{a2}) + I(X_{a1}; X_{b2}|X_{a2}) \\
&\qquad + I(X_{a2}; X_{b1}|X_{a1}) + I(X_{a2}; X_{b2}|X_{a1})\big) \\
&-\frac{1}{174}\big(I(X_{a1}; X_{b1}|X_{a2}X_{b2}) + I(X_{a1}; X_{b2}|X_{a2}X_{b1}) \\
&\qquad + I(X_{a2}; X_{b1}|X_{a1}X_{b2}) + I(X_{a2}; X_{b2}|X_{a1}X_{b1})\big),
\end{aligned}$$

which is term-by-term upper bounded by zero, implying the bound $\frac{484}{435}R_0 + R_1 \leq \frac{528}{493}$.

Let U be a ternary random variable such that $\mathrm{P}(U = 0) = 20/119$, $\mathrm{P}(U = 1) = 88/119$, $\mathrm{P}(U = 2) = 11/119$. Conditionals are given by:

$$\begin{aligned}
(X_{a1}, X_{b1}, X_{a2}, X_{b2})|(U = 0) &= (0, 0, 0, 0) \\
(X_{a1}, X_{b1}, X_{a2}, X_{b2})|(U = 1) &= (M_1, M_1, M_1, 0) \\
(X_{a1}, X_{b1}, X_{a2}, X_{b2})|(U = 2) &= (M_1, 0, M_2, 0),
\end{aligned}$$

where M_1 and M_2 are two independent unbiased binary random variables. Let Q be a random variable that symmetrizes the distribution of X (in the sense of the proof of Proposition 1) and let $\tilde{U} = (U, Q)$. Substituting $(\tilde{U}, X)$ into the normalized two-letter version of Bound 1 yields:

$$\begin{aligned}
R_0 &\leq \frac{1}{2}I(\tilde{U}; Z) = \frac{75}{119} \\
R_0 + R_1 &\leq \frac{1}{2}\left(I(\tilde{U}; Z) + \min\{I(X; Y|\tilde{U}), I(X; \hat{Y}|\tilde{U})\}\right) = 1 \\
R_0 + R_1 &\leq \frac{1}{2}\min\{I(X; Y), I(X; \hat{Y}\} = 1,
\end{aligned}$$

where $X = (X_{a1}, X_{b1}, X_{a2}, X_{b2})$ and similarly for others. Thus $(R_0, R_1) = (\frac{75}{119}, \frac{44}{119})$ lying at the intersection of the two lines $R_0 + R_1 = 1$ and $\frac{484}{435}R_0 + R_1 = \frac{528}{493}$ belongs to the two-letter superposition coding region. This establishes Theorem 1.

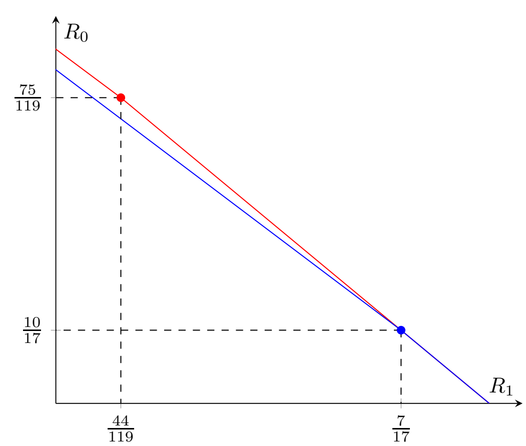

Fig. 3. Plots of the 1-letter and 2-letter superposition coding regions

Figure 3 shows the intersection points of the two lines that constitute the 1-letter and 2-letter superposition coding regions. The red line-segments indicates the 2-letter superposition coding region and the blue line-segments indicates the one-letter superposition coding region. □

DISCUSSION

The strict sub-optimality of the superposition coding region for this message setting is shown by demonstrating a channel for which the 2-letter superposition coding region strictly outperforms the 1-letter region. In this sense, the technique is similar to the one employed for showing the strict sub-optimality of the Han-Kobayashi achievable region for the interference channel [9]. However, there are some important differences that perhaps make the analysis in this setting more interesting.

Firstly, we are able to completely characterize the 2-letter superposition coding region. This boils down to finding choices of λ and α that make a certain linear combination of entropies of subsets of binary random variables negative for all probability distributions. This remains true even for higher letter computations. Hence there is a reasonable chance of being able to evaluate higher letter superposition coding regions and obtaining intuition as to the time-correlation in the optimizing distributions to see if it admits an alternate interpretation, yielding new achievability ideas.

Towards this end, notice that the optimizing distributions (that yield the non-trivial corner point along the $R_0+R_1 = 1$ line in the counterexample) have optimizers where the same X is transmitted across the parallel channels. Further in the 2-letter scheme, the same X is even transmitted across two consecutive-time slots, for some choices of U. This shows that superposition coding does not fully exploit the spatial and temporal diversity provided by the different channels to Y and $\hat{Y}$. Hidden in the optimizers should be some hint as to how best to exploit the missing diversity gains.

Secondly, using sub-modularity of entropy and the idea behind testing Shannon-type inequalities, one can get upper bounds on the critical λ, the slope of the capacity region around $(R_0, R_1) = (C_Z, 0)$. Even in the two-letter case (where the four variables $X_{a1}, X_{b1}, X_{a2}, X_{b2}$ involved are binary) restricting oneself to Shannon-type inequalities and maximizing the linear combination could have resulted in an non-entropic extreme point [10]. This would have led to an outer bound to the 2-letter superposition coding region.

Luckily for us, the extreme point in the region calculations using sub-modularity (Shannon-type) constraints turns out to be achievable; thus yielding a precise characterization of the 2-letter superposition coding region. One interesting question that is worth pursuing is whether this phenomenon continues to hold for higher letter computations as well. If so, would it be possible to make a well-informed guess of the structure of the optimizing distributions for higher letters. The simplicity of the expressions in this case and the rather vast literature about Shannon-type and non-Shannon type inequalities make the questions mentioned above an interesting and well motivated pursuit.

CONCLUSION

In this paper we show that superposition coding region is strictly sub-optimal for a three receiver broadcast channel with two degraded message sets; where two of the receivers need to decode both the messages, while the third receiver only needs to decode a common message. This solves open problem 8.2 in [1]. It also disproves a factorization conjecture, hypothesized by one of the authors in [6]. The key idea is to show that the 2-letter region strictly outperforms the 1-letter region and the main technical contribution is the bag of ideas (some of them developed by the authors previously) used in the evaluation of these regions for the channels. As a consequence, this opens interesting avenues for further exploration into new achievable schemes for this simple setting.

REFERENCES

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2012.

[2] J. Körner and K. Marton, "Images of a set via two channels and their role in mult-user communication," *IEEE Trans. Info. Theory*, vol. IT-23, pp. 751–761, Nov, 1977.

[3] I. Csiszár, *Some Basic Mathematical Problems of Multiuser Shannon Theory*, T. M. Cover and B. Gopinath, Eds. New York, NY: Springer New York, 1987.

[4] 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*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.

[5] C. Nair and Z. V. Wang, "On the inner and outer bounds of 3-receiver broadcast channels with 2-degraded message sets," *International Symposium on Information Theory*, pp. 1844–1848, 2009.

[6] C. Nair, "Upper concave envelopes and auxiliary random variables," *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 5, no. 1, pp. 12–20, 2013. [Online]. Available: http://dx.doi.org/10.1007/s12572-013-0081-7

[7] ——, "Capacity regions of two new classes of two-receiver broadcast channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4207–4214, sep. 2010.

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

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

[10] J. M. Walsh and S. Weber, "Relationships among bounds for the region of entropic vectors in four variables," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 1319–1326.