The capacity region of a class of broadcast channels with a sequence of less noisy receivers

Zizhou "Vincent" Wang   Chandra Nair

Abstract

The capacity region of a broadcast channel consisting of k-receivers that lie in a less noisy sequence is an open problem, when $k \geq 3$. We prove that superposition coding is indeed optimal for a class of broadcast channels with a sequence of less noisy receivers. This class contains the $k = 3$ case of the open problem mentioned above, thus resolving its capacity region.

1 Introduction

Consider a discrete memoryless broadcast channel with k-receivers $Y_1, .., Y_k$. For formal definitions and discussion of previous results on broadcast channels please refer to [1, 2]. A receiver Y_s is said to be less noisy[4] than receiver Y_t if $I(U; Y_s) \geq I(U; Y_t)$ for all $U \to X \to (Y_s, Y_t)$. We denote this relationship(partial-order) by $Y_s \succeq Y_t$.

The new idea is the use of *virtual receivers* in the identification of auxiliary random variables in the converse.

A k-receiver broadcast channel is said to belong to class $\mathcal{C}$ if there exists $k-1$ virtual receivers $V_1, ..., V_{k-1}$ satisfying:

- $X \to V_1 \to ... \to V_{k-1}$ forms a Markov chain and
- The following"interleaved" less noisy condition holds:

$$Y_1 \succeq V_1 \succeq Y_2 \succeq \cdots Y_{k-1} \succeq V_{k-1} \succeq Y_k. \qquad (1)$$

This class contains some interesting sequences of less noisy receivers as mentioned below. The following broadcast channels are some examples belonging to class $\mathcal{C}$:

1. A sequence of degraded receivers, i.e. $X \to Y_1 \to ... \to Y_k$; set $V_i = Y_{i+1}$,
2. A sequence of "nested" less noisy receivers, i.e. $Y_s \succeq (Y_{s+1}, ..., Y_k)$; set $V_i = (Y_{i+1}, ..., Y_k)$,
3. A 3-receiver less noisy sequence, i.e. $Y_1 \succeq Y_2 \succeq Y_3$; set $V_1 = V_2 = Y_2$.

We present a couple of results before we prove the capacity region for the independent message requirement for class $\mathcal{C}$.

Fact 1. *From the definition of less noisy receiver, by conditioning on U_2, it follows that whenever $(U_1, U_2) \to X \to (Y_s, Y_t)$ forms a Markov chain*

$$I(U_1; Y_t|U_2) \leq I(U_1; Y_s|U_2). \qquad (2)$$

Lemma 1. *If a receiver $Y_s \succeq Y_t$ then*[1]

$$I(Y_{t,1}^{i-1}; Y_{s,i}|U) \leq I(Y_{t,1}^{i-2}, Y_{s,i-1}; Y_{s,i}|U) \leq \cdots \leq I(Y_{t,1}, Y_{s,2}^{i-1}; Y_{s,i}|U) \leq I(Y_{s,1}^{i-1}; Y_{s,i}|U)$$

whenever $(U, Y_{t,1}^{p-1}, Y_{s,p+1}^{i-1}, Y_{s,i}) \to X_p \to (Y_{t,p}, Y_{s,p})$ forms a Markov chain for $1 \leq p \leq i-1$.

Proof. For all p such that $1 \leq p \leq i-1$, observe that

$$\begin{aligned} &I(Y_{t,1}^{p}, Y_{s,p+1}^{i-1}; Y_{s,i}|U) \\ &= I(Y_{t,1}^{p-1}, Y_{s,p+1}^{i-1}; Y_{s,i}|U) \\ &\quad + I(Y_{t,p}; Y_{s,i}|U, Y_{t,1}^{p-1}, Y_{s,p+1}^{i-1}) \\ &\leq I(Y_{t,1}^{p-1}, Y_{s,p+1}^{i-1}; Y_{s,i}|U) \\ &\quad + I(Y_{s,p}; Y_{s,i}|U, Y_{t,1}^{p-1}, Y_{s,p+1}^{i-1}) \\ &= I(Y_{t,1}^{p-1}, Y_{s,p}^{i-1}; Y_{s,i}|U) \end{aligned}$$

where the inequality follows from (2) as $(U, Y_{t,1}^{p-1}, Y_{s,p+1}^{i-1}, Y_{s,i}) \to X_p \to (Y_{t,p}, Y_{s,p})$ forms a Markov chain for $1 \leq p \leq i-1$. □

1.1 Main result

Theorem 1. *For any broadcast channel belonging to class $\mathcal{C}$ with independent message requirements, the capacity region is the set of rate tuples $R_1, ..., R_k$ such that $R_s \leq I(U_s; Y_s|U_{s+1})$ where $U_1 = X, U_{k+1} = \emptyset$ and the sequence $U_k \to U_{k-1} \cdots \to U_2 \to X \to (Y_1, ..., Y_k)$ forms a Markov chain.*

Proof. `Achievability`: The achievability is straightforward using superposition coding and jointly typical decoding. We shall refer the reader to [3] for

[1] The notation $Y_{t,p}^{i}$ denotes $(Y_{t,p}, Y_{t,p+1}, \ldots, Y_{t,i})$.

details. Since $Y_s \succeq Y_j, s \leq j \leq k$, the receiver Y_s successively decodes messages M_j (equivalently the sequences U_j^n) from $j = k$ to $j = s$. Each step is correct with high probability since

$$\begin{aligned} R_j &= I(U_j; Y_j | U_{j+1}) - \epsilon \\ &\leq I(U_j; Y_s | U_{j+1}) - \epsilon, \end{aligned}$$

when $s \leq j \leq k$. Therefore the rate tuples given in Theorem 1 are indeed achievable.

`Converse`: Let $M_{s+1}^k = (M_{s+1}, ..., M_k)$. Using Fano's inequality, observe that for $2 \leq s \leq k$

$$\begin{aligned} nR_s &\leq I(M_s; Y_{s,1}^n | M_{s+1}^k) + n\epsilon_n \\ &= \sum_{i=1}^n I(M_s; Y_{s,i} | M_{s+1}^k, Y_{s,1}^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n I(M_s, V_{s-1,1}^{i-1}; Y_{s,i} | M_{s+1}^k, Y_{s,1}^{i-1}) \\ &\quad - I(V_{s-1,1}^{i-1}; Y_{s,i} | M_s^k, Y_{s,1}^{i-1}) + n\epsilon_n \\ &\overset{(a)}{=} \sum_{i=1}^n I(M_s, V_{s-1,1}^{i-1}, Y_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k, V_{s,1}^{i-1}) \\ &\quad + I(V_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k) - I(Y_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k) \\ &\quad - I(V_{s-1,1}^{i-1}; Y_{s,i} | M_s^k, Y_{s,1}^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n I(M_s, V_{s-1,1}^{i-1}; Y_{s,i} | M_{s+1}^k, V_{s,1}^{i-1}) \\ &\quad + I(V_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k) - I(Y_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k) \\ &\quad + I(Y_{s,1}^{i-1}; Y_{s,i} | M_s^k, V_{s-1,1}^{i-1}) \\ &\quad - I(V_{s-1,1}^{i-1}; Y_{s,i} | M_s^k, Y_{s,1}^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n I(M_s, V_{s-1,1}^{i-1}; Y_{s,i} | M_{s+1}^k, V_{s,1}^{i-1}) \\ &\quad + I(V_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k) - I(Y_{s,1}^{i-1}; Y_{s,i} | M_{s+1}^k) \\ &\quad + I(Y_{s,1}^{i-1}; Y_{s,i} | M_s^k) - I(V_{s-1,1}^{i-1}; Y_{s,i} | M_s^k) \\ &\quad + n\epsilon_n \\ &\overset{(b)}{\leq} \sum_{i=1}^n I(M_s, V_{s-1,1}^{i-1}; Y_{s,i} | M_{s+1}^k, V_{s,1}^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n I(U_{s,i}; Y_{s,i} | U_{s+1,i}) + n\epsilon_n, \end{aligned}$$

where $U_{s,i} = (M_s^k, V_{s-1,1}^{i-1})$. Here the equality (a) follows from the fact that $X \to V_1 \to \cdots \to V_{k-1}$ is a Markov chain. The inequality (b) follows from (1) and Lemma 1.

For $s = 1$, similarly using Fano's inequality, observe that

$$\begin{aligned} nR_1 &\leq I(M_1; Y_{1,1}^n | M_2^k) + n\epsilon_n \\ &= \sum_{i=1}^n I(M_1; Y_{1,i} | M_2^k, Y_{1,1}^{i-1}) + n\epsilon_n \\ &\leq \sum_{i=1}^n I(X_i; Y_{1,i} | M_2^k, Y_{1,1}^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n I(X_i; Y_{1,i} | M_2^k, V_{1,1}^{i-1}) + I(V_{1,1}^{i-1}; Y_{1,i} | M_2^k) \\ &\quad - I(Y_{1,1}^{i-1}; Y_{1,i} | M_2^k) + n\epsilon_n \\ &\leq \sum_{i=1}^n I(X_i; Y_{1,i} | M_2^k, V_{1,1}^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n I(X_i; Y_{1,i} | U_{2i}) + n\epsilon_n, \end{aligned}$$

where the final inequality again follows from (1) and Lemma 1.

Define Q to be a uniform random variable taking values in $\{1, .., n\}$ and independent of all other random variables. As usual, we set $U_s = (U_{s,Q}, Q)$ and $X = X_Q$. Since $X \to V_1 \to \cdots \to V_{k-1}$ is a Markov chain it follows that $U_k \to U_{k-1} \to \cdots \to U_2 \to X$ forms a Markov chain as well. This completes the proof of the converse. □

2 Conclusion

We establish the capacity region for the 3-receiver less noisy broadcast channel. We also compute the capacity region for a class of k-receiver less noisy sequences that contain the previously mentioned scenario. As mentioned earlier, a new idea of virtual receivers is used to fashion a converse for the capacity region.

References

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

[2] ———, *Comments on broadcast channels*, IEEE Trans. Info. Theory **IT-44** (October, 1998), 2524–2530.

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

[4] 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.