

Achievable Rates for the Relay Channel with Orthogonal Receiver Components

Abbas El Gamal

Dept. of Electrical Engineering
Stanford University
Email: abbas@ee.stanford.edu

Amin Gohari

Tehran Institute for Advanced Studies
Khatam University
Email: a.gohari@teias.institute

Chandra Nair

Dept. of Information Engineering
The Chinese University of Hong Kong
Email: chandra@ie.cuhk.edu.hk

Abstract—This paper studies lower bounds on the capacity of the relay channel with orthogonal receiver components (also referred to as primitive relay channel). We show that the lower bound in Theorem 7 of Cover and El Gamal, which uses mixed decode-forward and compress-forward strategies, is identical to the lower bound of Chong, Motani and Garg, and is always larger than or equal to the recent lower bound of Mondelli, Hassani and Urbanke. We provide a simplified expression for the lower bound in Theorem 7 of Cover and El Gamal and interpret one of its auxiliary variables as implementing the randomized time-sharing strategy. Next, we compare the lower bound for the Gaussian relay channel with orthogonal receiver components to existing upper bounds. Finally, we disprove a conjecture by Ahlswede and Han on the capacity of the subclass of relay channels with orthogonal receiver components and i.i.d. output.

A full version of this paper is accessible at: <http://chandra.ie.cuhk.edu.hk/pub/papers/NIT/Prim-Rel-LB.pdf>

I INTRODUCTION

The discrete memoryless relay channel consists of four alphabets \mathcal{X} , \mathcal{X}_r , \mathcal{Y}_r , \mathcal{Y} , and a collection of conditional pmfs $p(y_r, y|x, x_r)$ on $\mathcal{Y}_r \times \mathcal{Y}$. A $(2^{nR}, n)$ code for the discrete memoryless relay channel $p(y, y_r|x, x_r)$ consists of a message set $[1 : 2^{nR}]$, an encoder that assigns a codeword $x^n(m)$ to each message $m \in [1 : 2^{nR}]$, a relay encoder that assigns a symbol $x_{ri}(y_r^{i-1})$ to each past received sequence y_r^{i-1} for each time $i \in [1 : n]$, and a decoder that assigns an estimate \hat{M} or an error message ε to each received sequence y^n . We assume that the message M is uniformly distributed over $[1 : 2^{nR}]$. The definitions of the average probability of error, achievability and capacity follow those in [1].

Obtaining a computable characterization of the capacity for the relay channel $p(y, y_r|x, x_r)$ is one of the longstanding open problems in network information theory. Computable lower and upper bounds on the capacity which coincide in some special cases have been developed. In [2], a number of coding strategies were introduced (later termed decode-forward, partial decode-forward, and compress-forward). These strategies were combined to obtain a unified lower bound on the capacity in Theorem 7 of [2]. Subsequently, Chong, Motani and Garg [3], [4] put forth a lower bound and showed that it is greater than or equal to the bound in Theorem 7. Supported by some numerical simulations for Gaussian relay channels, Chong, Motani and Garg conjectured that their bound can be strictly larger than the bound in Theorem 7 [3, Remark 6], [4, Remark

6]. The cutset bound on the capacity of the relay channel was introduced in [2]. This bound was recently strengthened by the authors in [5] (see also [6]).

There has been renewed interest in the class of relay channels with orthogonal receiver components in which $Y = (Y_1, Y_2)$ and $p(y_1, y_2, y_r|x, x_r) = p(y_1, y_r|x)p(y_2|x_r)$. It is known that the capacity of this class of relay channels depends on $p(y_2|x_r)$ only through its capacity, hence it can be substituted with a noiseless link of the same capacity C_0 [7] as shown in Figure 1. The capacity of the relay channel with orthogonal receiver components, however, is also not known in general (see Section 16.7.3 in [1]). Recently, Mondelli, Hassani and Urbanke [8] proposed a lower bound for this class of relay channels. The cutset bound is known to be tight when Y_r is a function of (X, Y_1) [9] and the capacity for the Gaussian case coincides with the cutset bound for a certain range of channel parameters [1]. In [10], [11] examples of this class of relay channels are given for which the cutset bound is not tight. More recently, it was also shown that the cutset bound is not tight for the special class of relay channels with orthogonal receiver components introduced by Cover in [12], e.g., see [5], [13]–[16].

In this paper, we explore the unified lower bounds on the capacity of the relay channel with orthogonal receiver components in Theorem 7 of [17], [3], [4] and [8], and establish the following results:

- In Section II: We show that the lower bound in Theorem 7 of [17, Theorem 7] and the lower bound in [3], [4] are *identical* for relay channels with orthogonal receiver components. We also provide a simplified expression of this lower bound. We show that this simplified expression of the lower bound is larger than or equal to, and can be strictly larger than, the recent lower bound in [8].
- Section III: We compute our simplified expression of the lower bound for the Gaussian relay channels with orthogonal receiver components when the input random variables conditioned on Q are assumed to be Gaussian. Here Q plays the role of *randomized time-sharing* introduced for the binary skew-symmetric broadcast channel in [18] and is shown to strictly improve on pure time-sharing (power-control).
- Section IV: Finally, we restrict the discussion to the class of relay channels with orthogonal receiver components

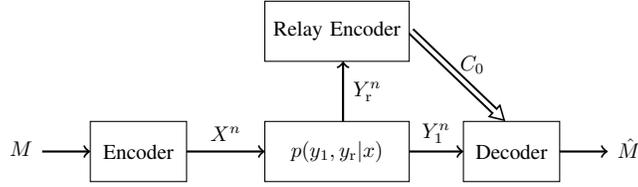


Fig. 1. Relay channel with orthogonal receiver components.

for which $p(y_r, y_1|x) = p(y_r)p(y_1|x, y_r)$ and show that the lower bound in Theorem 7 of [17, Theorem 7] reduces to the compress-forward bound with time-sharing. While it is known that time-sharing can increase the achievable rate in the compress-forward strategy, Ahlswede and Han conjectured in [19] that when Y_r is independent of the input signal X , time-sharing is not necessary and that the capacity of this class of channels is given by the pure compress-forward strategy. We disprove this conjecture by finding an example for which time-sharing strictly improves the compress-forward lower bound.

The proofs can be found in the full version of this paper in [20].

Notation: We adopt most of our notation from [1]. In particular, we use Y^i to denote the sequence (Y_1, Y_2, \dots, Y_i) , and Y_i^j to denote $(Y_i, Y_{i+1}, \dots, Y_j)$. All logarithms are to the base e . For a real number x , we use \bar{x} to denote $1 - x$. We use $h(\cdot)$ to denote differential entropy and $H_2(\cdot)$ to denote the binary entropy function.

II SIMPLIFYING THE LOWER BOUNDS

In [4, Lemma 2], the following equivalent characterization of the lower bound on the capacity of the general relay channel $p(y, y_r|x, x_r)$ in [17, Theorem 7] is given.

Theorem 1 ([17]). *A rate R is achievable if*

$$R \leq I(X, X_r; Y|Q) - I(Y_r; V|U, X, X_r, Y, Q), \quad (1)$$

$$R \leq I(X; V, Y|U, X_r, Q) + I(U; Y_r|W, X_r, Q), \quad (2)$$

$$R \leq I(X; V, Y|U, X_r, Q) + I(U; Y_r|W, X_r, Q) + I(X_r; Y|W, Q) - I(Y_r; V|U, X_r, Y, Q), \quad (3)$$

for some

$$p(q, w, u, v, x, x_r, y_r, y) = p(q)p(w|q)p(u|w, q)p(x|u, q) \times p(x_r|w, q)p(y_r, y|x, x_r)p(v|u, x_r, y_r, q). \quad (4)$$

Specific choices for the auxiliary random variables U, V, W, Q lead to the compress-forward, decode-forward and partial decode-forward lower bounds.

In [4], the following improved version of the above bound is given.

Theorem 2 ([4]). *A rate R is achievable if*

$$R \leq I(X, X_r; Y|Q) - I(Y_r; V|U, X, X_r, Y, Q), \quad (5)$$

$$R \leq I(X; V, Y|U, X_r, Q) + I(U; Y_r|W, X_r, Q), \quad (6)$$

$$R \leq I(X; V, Y|U, X_r, Q) + I(U; Y_r|W, X_r, Q) + I(X_r; Y|U, W, Q) - I(Y_r; V|U, X_r, Y, Q), \quad (7)$$

for some joint distribution as in (4).

Note that the only difference between the bounds in Theorems 1 and 2 is in one of the terms in equations (3) and (7). Since $U \rightarrow (W, Q) \rightarrow X_r$ forms a Markov chain, we have $I(X_r; Y|U, W, Q) = I(X_r; Y, U|W, Q) \geq I(X_r; Y|W, Q)$, which shows that the achievable rate in Theorem 2 is at least as large as that in Theorem 1 (see Remark 6 of [4]).

Remark 1. Letting $W' = (W, Q)$, $U' = (U, Q)$ and $Q' = \emptyset$ improves the region in Theorem 2. Thus, without loss of generality, we can set $Q = \emptyset$ in the statement of Theorem 2. Next, using the fact that $X_r \rightarrow W \rightarrow U \rightarrow X$ forms a Markov chain, one can verify that replacing U by U, W improves the rate constraints. Moreover, it relaxes the condition $p(v|u, x_r, y_r)$ to $p(v|u, w, x_r, y_r)$. This leads to the equivalent characterization of Theorem 2: A rate R is achievable if

$$R \leq I(X, X_r; Y) - I(Y_r; V|U, X, X_r, Y, W),$$

$$R \leq I(X; V, Y|U, X_r, W) + I(U; Y_r|W, X_r),$$

$$R \leq I(X; Y|U, X_r, W) + I(U; Y_r|W, X_r)$$

$$+ I(X_r; Y|U, W) - I(Y_r; V|U, X_r, X, Y, W),$$

for some $p(u, w)p(x|u, w)p(x_r|w)p(y_r, y|x, x_r)p(v|u, w, x_r, y_r)$.

We now state the main result of this section.

Theorem 3. *For any arbitrary relay channel $p(y_r, y_1|x)$ with orthogonal receiver components, the lower bounds in Theorem 1 and Theorem 2 are identical and have the equivalent form: A rate R is achievable if*

$$R \leq \min(I(U; Y_r), I(U; Y_1)) + I(X; Y_1|U) + C_0 - I(Y_r; V|U, X, Y_1), \quad (8)$$

$$R \leq I(X; V, Y_1|U) + I(U; Y_r), \quad (9)$$

for some $p(u, x, y_r, v, y_1) = p(u)p(x|u)p(y_r, y_1|x)p(v|u, y_r)$.

Further it suffices to consider $|\mathcal{V}| \leq |\mathcal{Y}_r| + 1$ and $|\mathcal{U}| \leq |\mathcal{X}| + 2$ when evaluating the bound.

Remark 2. Replacing U with (U, Q) , we can rewrite the bound in Theorem 3 in the equivalent form:

$$R \leq I(X; Y_1) + C_0 - I(Y_r; V|U, Q, X, Y_1), \quad (10)$$

$$R \leq I(Q; Y_r) + I(U; Y_r|Q) + I(X; Y_1|U, Q) + C_0 - I(Y_r; V|U, Q, X, Y_1), \quad (11)$$

$$R \leq I(Q; Y_r) + I(X; V, Y_1|U, Q) + I(U; Y_r|Q), \quad (12)$$

for some joint distribution satisfying $p(q, u, x, y_r, v, y_1) = p(q)p(x, u|q)p(y_r, y_1|x)p(v|u, q, y_r)$.

We can view Q as a time-sharing random variable and interpret the term $I(Q; Y_r)$ as implementing a randomized time-division strategy, that is, Q also carries part of the message as well. This decomposition into (U, Q) is seen to improve the region for the Gaussian setting over naive power control. The idea of exploiting Q , intuitively a time-sharing random variable, to convey additional message was first considered for the skew symmetric binary broadcast channel in [18]. In [21], it was shown that the randomized time-division region coincides with Marton's inner bound for the two receiver binary-input broadcast channel.

skip

II.1 Comparison of the simplified bound with the lower bound of Mondelli, Hassani and Urbanke

Consider the relay channels with orthogonal receiver components In [8], Mondelli, Hassani and Urbanke derived the following lower bound for this class of relay channels.

Theorem 4 (Theorem 1 in [8]). *Take some arbitrary joint distribution $p(x)p(y_1, y_r|x)p(v|y_r)$ satisfying $I(X; Y_r) < I(X; Y_1) + C_0$ and $C_0 \leq I(V; Y_r|Y_1)$. Then, the following rate is achievable:*

$$R = \frac{1}{I(V; Y_r|Y_1) - I_{max}} \left((C_0 - I_{max})I(X; V, Y_1) + \max\{I(X; Y_r), I(X; Y_1)\}(I(V; Y_r|Y_1) - C_0) \right)$$

where $I_{max} = \max\{0, I(X; Y_r) - I(X; Y_1)\}$.

The authors of [8] conjectured that this lower bound strictly improves over the lower bound in Theorem 1. The conjecture is justified in [8, Section IV] by an incomplete simulation. We show (in the full version) that this conjecture is false.

Lemma 1. *The achievable rate in Theorem 4 is subsumed by the achievable rate in Theorem 1.*

skip

III LOWER BOUND FOR GAUSSIAN RELAY CHANNEL

Consider the Gaussian relay channel with orthogonal receiver components described by

$$Y_1 = X + W_1, \quad \text{and} \quad Y_r = X + W_r,$$

where $W_1 \sim \mathcal{N}(0, N_1)$ and $W_r \sim \mathcal{N}(0, N_r)$ are independent of each other and of X , and a link of capacity C_0 from the relay to the destination. We assume average power constraint P on X and define $S_{12} = P/N_r$, $S_{13} = P/N_1$ and $S_{23} = 2^{2C_0} - 1$.

In the Gaussian setting, the channel $p(y_r|x)$ is less noisy than $p(y_1|x)$ if $N_1 \geq N_r$ and vice versa. Therefore, the term minimizing $\min(I(U; Y_r), I(U; Y_1))$ in (8) can be found depending on whether $N_1 \geq N_r$ holds or not. The following lemma gives an alternative characterization of Theorem 3 when there is a less noisy relation between the channels from the transmitter to the destination and to the relay.

Lemma 2. *If $p(y_r|x)$ is less noisy than $p(y_1|x)$, we can express the lower bound in Theorem 3 as*

$$R \leq \min_{\alpha \in [0,1]} \max \left[I(X; Y_1|U) + (1 - \alpha)I(U; Y_r) + \alpha I(U; Y_1) + \alpha C_0 - \alpha I(Y_r; V|U, X, Y_1) + (1 - \alpha)I(X; V|Y_1, U) \right] \quad (13)$$

where the inner maximum is over

$$p(u, x, y_r, v, y_1) = p(x, u)p(y_r, y_1|x)p(v|u, y_r). \quad (14)$$

If $p(y_1|x)$ is less noisy than $p(y_r|x)$, we can express the lower bound in Theorem 3 as

$$R \leq \min_{\alpha \in [0,1]} \max \left[I(X; Y_1|U) + I(U; Y_r) + \alpha C_0 - \alpha I(Y_r; V|U, X, Y_1) + (1 - \alpha)I(X; V|Y_1, U) \right],$$

where the inner maximum is over (14). Further it suffices to consider $|V| \leq |Y_r|$ and $|U| \leq |X|$ when evaluating the above expressions.

III.1 Evaluation for Gaussian inputs with randomized power control

The joint distribution that attains the inner maximum is not clear. In the rest of this section we replace U with (U, Q) in Lemma 2 and restrict the inner maximum to the case of Q being a discrete random variable. Moreover, for every $Q = q$ we assume that the conditional distribution of (U, V, X, Y_r, Y_1) given $\{Q = q\}$ is zero mean, jointly Gaussian random variables. Let $X|\{Q = q\} \sim \mathcal{N}(0, P_q)$. The auxiliary random variable Q allows us to perform power control by having P_q depend on the choice of q .¹ Since Q itself contains some part of the message, we call the class of such distributions as *Gaussian inputs with randomized power control*.

For the case of $N_1 \geq N_r$, we can write the bound in (13) (specialized to the above joint distribution on auxiliary variables) as

$$\min_{\alpha \in [0,1]} \max_{p(q,x)} \left[(1 - \alpha)I(Q; Y_r) + \alpha I(Q; Y_1) + \alpha C_0 + \sum_q G_\alpha^{(1)}(P_q) \right] \quad (15)$$

where Q is a discrete random variable, $(X|Q = q) \sim \mathcal{N}(0, P_q)$ and

$$G_\alpha^{(1)}(P) = \max_{p(u,v|x)} \left[I(X; Y_1|U) + (1 - \alpha)I(U; Y_r) + \alpha I(U; Y_1) - \alpha I(Y_r; V|U, X, Y_1) + (1 - \alpha)I(X; V|Y_1, U) \right].$$

In the definition of $G_\alpha^{(1)}(P)$ the maximum is over zero mean jointly Gaussian random variables (U, V, X, Y_r, Y_1)

¹Power control is known to be helpful to improve the compress-forward rate [1, Sec. 16.8].

satisfying $X \sim \mathcal{N}(0, P)$ and $p(u, x, y_r, v, y_1) = p(x, u)p(y_r, y_1|x)p(v|u, y_r)$. Similarly, for the case of $N_1 \leq N_r$, we can write the lower bound (specialized to the above joint distribution on auxiliary variables) as

$$\min_{\alpha \in [0,1]} \max_{p(q,x)} \left[I(Q; Y_r) + \alpha C_0 + \sum_q G_\alpha^{(2)}(P_q) \right] \quad (16)$$

where

$$G_\alpha^{(2)}(P) = \max_{p(u,v|x)} \left[I(X; Y_1|U) + I(U; Y_r) - \alpha I(Y_r; V|U, X, Y_1) + (1 - \alpha) I(X; V|Y_1, U) \right],$$

and maximum in $G_\alpha^{(2)}(P)$ is taken over the same set as $G_\alpha^{(1)}(P)$. Lemma 3 in the full version gives an explicit expression for $G_\alpha^{(1)}(P)$ and $G_\alpha^{(2)}(P)$.

We call the bound in equations (15) and (16) as the *Gaussian inputs with randomized power control* lower bounds. In Figure 2, we plot the these lower bound when Q is a constant random variable, Q is a binary random variable. As one can see, increasing the size of Q improves the lower bound. We also plot a theoretical upper bound on the Gaussian inputs with randomized power control lower bounds as the size of Q tends to infinity. This upper bound is computed as follows: consider, for instance, the bound in (15):

$$\begin{aligned} & \min_{\alpha \in [0,1]} \max_{p(q,x)} \left[(1 - \alpha)h(Y_r) - (1 - \alpha)h(Y_r|Q) + \alpha h(Y_1) \right. \\ & \quad \left. - \alpha h(Y_1|Q) + \alpha C_0 + \sum_q G_\alpha^{(1)}(P_q) \right] \\ & \leq \min_{\alpha \in [0,1]} \max_{p(q,x)} \left[\frac{1 - \alpha}{2} \log(2\pi e(P + N_r)) - (1 - \alpha)h(Y_r|Q) \right. \\ & \quad \left. + \frac{\alpha}{2} \log(2\pi e(P + N_1)) - \alpha h(Y_1|Q) + \alpha C_0 + \sum_q G_\alpha^{(1)}(P_q) \right] \end{aligned}$$

where we use the fact that Gaussian distribution has higher entropy than a mixture Gaussian distribution with the same variance. Note that it suffices to take binary Q to compute the above expression.

The upper bound in Figure 2 is as follows:

Theorem 5 ([5]). *Any achievable rate R for the relay channel with orthogonal receiver components $p(y_1, y_r|x)$ with a relay-to-receiver link of capacity C_0 must satisfy the following inequalities*

$$R \leq I(X; Y_1, Y_r) - I(V; Y_1|Y_r) - I(X; Y_r|V, Y_1), \quad (17)$$

$$R \leq I(X; Y_1) + C_0 - I(V; Y_r|X, Y_1), \quad (18)$$

for some $p(x)p(y_1, y_r|x)p(v|x, y_r)$. Further it suffices to consider $|\mathcal{V}| \leq |\mathcal{X}||\mathcal{Y}_r| + 1$ when evaluating the bound.

For the Gaussian relay channel, this upper bound simplifies to the expression given in (19) on the top of the next page [5] [22].

The figure also plots the decode-forward lower bound and the compress-forward lower bound for this relay channel

whose expressions are given in [1, Eq. 16.16] and [1, Eq. 16.17]. When $S_{12} \geq S_{13} + S_{23} + S_{13}S_{23}$, the upper bound in (19), the cut-set bound and the decode-forward lower bound all coincide. This corresponds to the top corner point in the figure.

skip

IV DISPROVING AHLWEDE-HAN CONJECTURE

A relay channel with orthogonal receiver components is said to be with i.i.d. output if its family of conditional probabilities have the form

$$p(y_1, y_2, y_r|x, x_r) = p(y_r)p(y_1|x, y_r)p(y_2|x_r). \quad (20)$$

For this class of channels, the bound in Theorem 3 reduces to:

$$R \leq I(X; V, Y_1|U) = I(X; Y_1|V, U)$$

$$R \leq I(X; Y_1|U) + C_0 - I(Y_r; V|U, X, Y_1)$$

for some $p(u, x, y_r, v, y_1) = p(u, x)p(y_r)p(y_1|x)p(v|u, y_r)$. The above rate is just the compress-forward rate with a time-sharing variable U . If we set $U = \emptyset$, we obtain the compress-forward rate without a time-sharing variable [1, Eq. 16.14].

Communication over a relay channel of the form (20) is equivalent to communication over a channel with rate-limited state information available at the receiver [10]. As such, the Ahlswede-Han conjecture in [19] for the channels with rate-limited state information is equivalent to the capacity for this relay channel with orthogonal receiver components being equal to $\max I(X; Y_1|V)$, where the maximum is over $p(x)p(y_r)p(y_1|x, y_r)p(v|y_r)$ such that $I(V; Y_r|Y_1) \leq C_0$. In other words, the Ahlswede-Han conjecture is equivalent to the optimality of the compress-forward strategy without time-sharing.

While it is known that the compress-forward rate is not concave in C_0 for general relay channel with orthogonal receiver components and time-sharing improves the compress-forward rate (e.g., see [1, Sec. 16.8]), the Ahlswede-Han conjecture implies that time-sharing is unnecessary for the special case of relay channels of the form (20). Interestingly, the Ahlswede-Han conjecture is valid for all channels for which the capacity is known [9]–[11].

In this section, we construct an example for which the compress-forward rate with time-sharing is strictly better than that without time-sharing, which shows that the Ahlswede-Han conjecture (as stated) is false. A key feature of our example is its non-symmetric construction.

Lemma 3. *Take a discrete memoryless relay channel with orthogonal receiver components of the form $p(y_r)p(y_1|x, y_r)$. Let $F(p(x), \lambda) = \max_{p(v|y_r)} I(X; Y_1|V) + \lambda(C_0 - I(V; Y_r|Y_1))$. Then, the compress-forward rate without time-sharing equals*

$$\max_{p(x)} \min_{\lambda \in [0,1]} F(p(x), \lambda)$$

while the compress-forward rate with time-sharing equals

$$\min_{\lambda \in [0,1]} \max_{p(x)} F(p(x), \lambda).$$

$$C \leq \begin{cases} \frac{1}{2} \log \left(1 + S_{13} + \frac{S_{12}(S_{13}+1)S_{23}}{(S_{13}+1)(S_{23}+1)-1} \right), & \text{for } S_{12} \leq S_{13} + S_{23} + S_{13}S_{23}, \\ \frac{1}{2} \log((1 + S_{13})(1 + S_{23})), & \text{otherwise.} \end{cases} \quad (19)$$

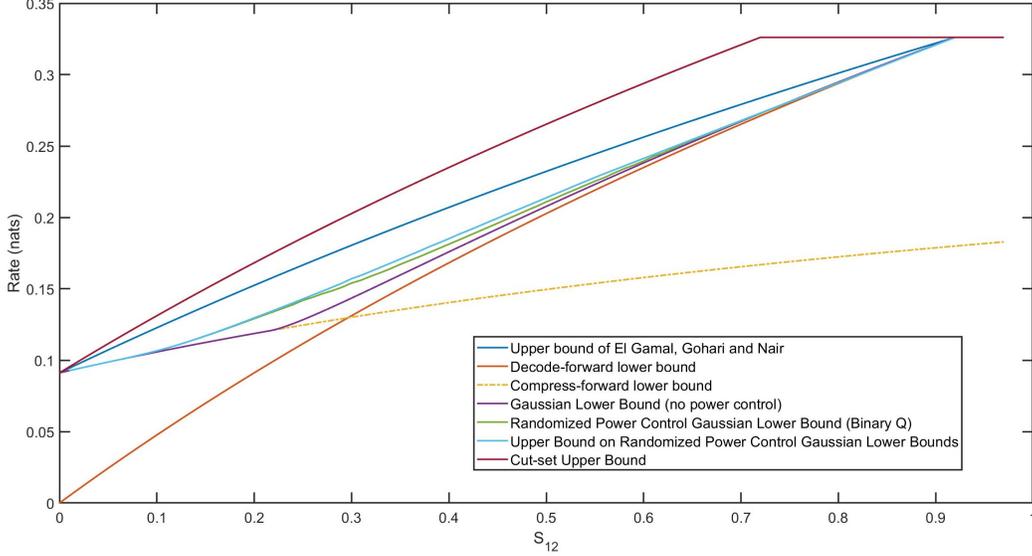


Fig. 2. Plots of the bounds for the Gaussian product-form relay channel with $S_{13} = 0.2$, $S_{23} = 0.6$.

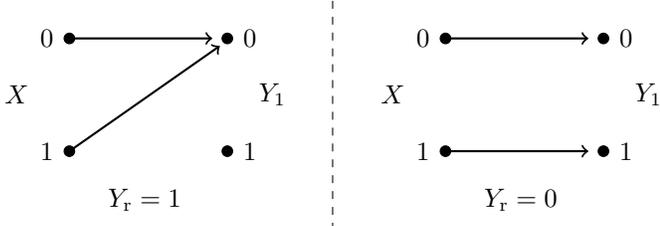


Fig. 3. Conditional probability $p(y_1|x, y_r)$. We assume $\mathbb{P}[Y_r = 0] = \beta$ and $\mathbb{P}[X = 0] = \theta$.

This lemma implies that the Ahlswede-Han conjecture does not hold if one can construct an example for which

$$\max_{p(x)} \min_{\lambda} F(p(x), \lambda) < \min_{\lambda} \max_{p(x)} F(p(x), \lambda).$$

We now construct such an example. Consider a channel with binary X , Y_1 and Y_r where $p(x, y_r, y_1) = p(x)p(y_r)p(y_1|x, y_r)$ is defined as follows: assume that $Y_1 = 0$ if $Y_r = 1$, and $Y_1 = X$ if $Y_r = 0$ (see Fig. 3). Let $\mathbb{P}[Y_r = 0] = \beta$.

Let $\mathbb{P}[Y_r = 0|v] = q$ and $\mathbb{P}[X = 0] = \theta$. Since $p(x)$ is specified by θ , we can denote $F(p(x), \lambda)$ by $F(\theta, \lambda)$ for simplicity. Observe that

$$\begin{aligned} I(X; Y_1|v) &= H_2(q\bar{\theta}) - \bar{\theta}H_2(q), \\ H(Y_r|v) - H(Y_1|v) &= H_2(q) - H_2(q\bar{\theta}), \end{aligned}$$

where $\bar{\theta} = 1 - \theta$. Thus,

$$\begin{aligned} F(\theta, \lambda) &= \lambda(C_0 - H_2(\beta) + H_2(\beta\bar{\theta})) \\ &\quad + \mathcal{C}[(1 - \lambda)H_2(q\bar{\theta}) + (\lambda - \bar{\theta})H_2(q)]_{q=\beta}, \end{aligned}$$

where $\mathcal{C}[(1 - \lambda)H_2(q\bar{\theta}) + (\lambda - \bar{\theta})H_2(q)]$ is the upper concave envelope of the function $q \mapsto (1 - \lambda)H_2(q\bar{\theta}) + (\lambda - \bar{\theta})H_2(q)$ defined for $q \in [0, 1]$.² An explicit expression for $\mathcal{C}[(1 - \lambda)H_2(q\bar{\theta}) + (\lambda - \bar{\theta})H_2(q)]$ is given in [20]. Let $C_0 = 0.06$ and $\beta = 0.3$. For this setting, we claim that $\max_{\theta} \min_{\lambda} F(\theta, \lambda) < \min_{\lambda} \max_{\theta} F(\theta, \lambda)$. Let

$$\mathcal{P} = \{t \in [0, 1] : t = 0.001 + 0.01k \text{ for some } k \in \mathbb{N} \cup \{0\}\}.$$

Observe that

$$\begin{aligned} \max_{\theta} \min_{\lambda} F(\theta, \lambda) &\leq \max_{\theta} \min_{\lambda \in \mathcal{P}} F(\theta, \lambda), \\ \min_{\lambda} \max_{\theta} F(\theta, \lambda) &\leq \min_{\lambda} \max_{\theta \in \mathcal{P}} F(\theta, \lambda). \end{aligned}$$

Maximum over θ in $\max_{\theta} \min_{\lambda \in \mathcal{P}} F(\theta, \lambda)$ yields a value in $(0.191232, 0.191234)$. On the other hand, minimum over λ in $\min_{\lambda} \max_{\theta \in \mathcal{P}} F(\theta, \lambda)$ yields a value in $(0.191898, 0.1919)$, showing that $\max_{\theta} \min_{\lambda \in \mathcal{P}} F(\theta, \lambda) < \min_{\lambda} \max_{\theta \in \mathcal{P}} F(\theta, \lambda)$. Thus we have $\max_{\theta} \min_{\lambda} F(\theta, \lambda) < \min_{\lambda} \max_{\theta} F(\theta, \lambda)$ and this completes the proof. Please see the full version for various plots and more details.

²The concave envelope of the function is the smallest concave function that dominates f from above.

REFERENCES

- [1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.
- [2] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on information theory*, vol. 25, no. 5, pp. 572–584, 1979.
- [3] H.-F. Chong, M. Motani, and H. K. Garg, "Generalized backward decoding strategies for the relay channel," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 394–401, 2007.
- [4] H.-F. Chong and M. Motani, "On achievable rates for the general relay channel," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1249–1266, 2011.
- [5] A. El Gamal, A. Gohari, and C. Nair, "Strengthened cutset upper bounds on the capacity of the relay channel and applications," *arXiv preprint arxiv:2101.11139*, 2021.
- [6] A. Gohari and C. Nair, "Outer bounds for multiuser settings: The auxiliary receiver approach," *submitted to the IEEE Transactions on Information Theory*, 2020.
- [7] Y.-H. Kim, "Coding techniques for primitive relay channels," in *Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput.*, 2007, p. 2007.
- [8] M. Mondelli, S. H. Hassani, and R. Urbanke, "A new coding paradigm for the primitive relay channel," *Algorithms*, vol. 12, no. 10, p. 218, 2019.
- [9] Y.-H. Kim, "Capacity of a class of deterministic relay channels," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 1328–1329, 2008.
- [10] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulosum relay channels," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 921–930, March 2009.
- [11] R. Tandon and S. Ulukus, "A new upper bound on the capacity of a class of primitive relay channels," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2008, pp. 1562–1569.
- [12] T. Cover, "The capacity of the relay channel," in *Open Problems in Communication and Computation*. Springer, 1987, pp. 72–73.
- [13] X. Wu, A. Ozgur, and L.-L. Xie, "Improving on the cut-set bound via geometric analysis of typical sets," *IEEE Transactions on Information Theory*, vol. 63, no. 4, pp. 2254–2277, 2017.
- [14] X. Wu, L. P. Barnes, and A. Ozgur, "'the capacity of the relay channel': Solution to cover's problem in the gaussian case," *IEEE Transactions on Information Theory*, vol. 65, no. 1, pp. 255–275, Jan 2019.
- [15] J. Liu and A. Ozgur, "New converses for the relay channel via reverse hypercontractivity," in *2019 IEEE International Symposium on Information Theory (ISIT)*, July 2019, pp. 2878–2882.
- [16] J. Liu, "Minorization via mixed volumes and cover's problem for general channels," *arXiv preprint arXiv:2012.14521*, 2020.
- [17] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [18] B. Hajek and M. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Trans. Info. Theory*, vol. IT-25, pp. 36–46, January, 1979.
- [19] R. Ahlswede and T. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 396–412, 1983.
- [20] A. El Gamal, A. Gohari, and C. Nair, "Achievable rates for the relay channel with orthogonal receiver components," <http://chandra.ie.cuhk.edu.hk/pub/papers/NIT/Prim-Rel-LB.pdf>, 2021.
- [21] Y. Geng, V. Jog, C. Nair, and Z. V. Wang, "An information inequality and evaluation of Marton's inner bound for binary input broadcast channels," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4095–4105, 2013.
- [22] X. Wu, L. P. Barnes, and A. Özgür, "The geometry of the relay channel," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 2233–2237.