Achievable rates for the relay channel with orthogonal receiver components

Abbas El Gamal, Amin Gohari and Chandra Nair

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.

1. Introduction

The capacity of the relay channel $p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}})$ is one of the longstanding open problems in network information theory. Lower and upper bounds on the capacity which coincide in some special cases have been developed. In [CEG79], 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 [CEG79]. Subsequently, Chong, Motani and Garg [CMG07], [CM11] 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 [CMG07, Remark 6], [CM11, Remark 6]. The cutset bound on the capacity of the relay channel was introduced in [CEG79]. This bound was recently strengthened by the authors in [EGGN21] (see also [GN20]).

There has been renewed interest in the class of relay channels with orthogonal receiver components in which $Y = (Y_1, Y_2)$ and $p(y_1, y_2, y_{\mathrm{r}}|x, x_{\mathrm{r}}) = p(y_1, y_{\mathrm{r}}|x)p(y_2|x_{\mathrm{r}})$. It is known that the capacity of this class of relay channels depends on $p(y_2|x_{\mathrm{r}})$ only through its capacity, hence it can be substituted with a noiseless link of the same capacity C_0 [Kim07] 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 [EK11]). Recently, Mondelli, Hassani and Urbanke [MHU19] 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) [Kim08] and the capacity for the Gaussian case coincides with the cutset bound for a certain range of channel parameters [EK11]. In [ARY09], [TU08] 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 [Cov87], e.g., see [WOX17], [WBO19], [LO19], [EGGN21], [Liu20].

In this paper, we explore the unified lower bounds on the capacity of the relay channel with orthogonal receiver components in Theorem 7 of [CG79], [CMG07], [CM11] and [MHU19], and establish the following results:

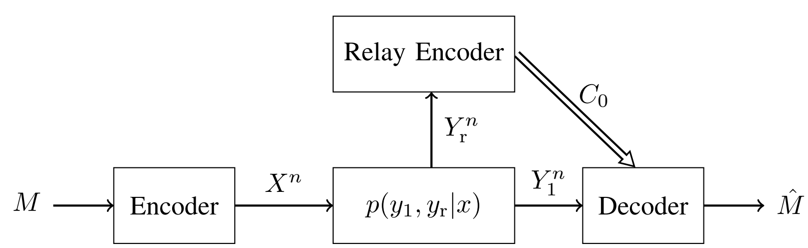

Fig. 1. Relay channel with orthogonal receiver components.

- In Section 2: We show that the lower bound in Theorem 7 of [CG79, Theorem 7] and the lower bound in [CMG07], [CM11] 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 [MHU19].
- Section 3: 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 [HP79] and is shown to strictly improve on pure time-sharing (power-control).
- Section 4: Finally, we restrict the discussion to the class of relay channels with orthogonal receiver components for which $p(y_{\mathrm{r}}, y_1|x) = p(y_{\mathrm{r}})p(y_1|x, y_{\mathrm{r}})$ and show that the lower bound in Theorem 7 of [CG79, 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 [AH83] 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.

Notation: We adopt most of our notation from [EK11]. In particular, we use Y^i to denote the sequence $(Y_1, Y_2, \cdots, Y_i)$, and Y_i^j to denote $(Y_i, Y_{i+1}, \cdots, Y_j)$. 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.

2. Simplifying the Lower Bounds

In [CM11, Lemma 2], the following equivalent characterization of the lower bound on the capacity of the general relay channel $p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}})$ in [CG79, Theorem 7] is given.

Theorem 1 ([CG79]). *A rate R is achievable if*

$$R \leq I(X, X_{\mathrm{r}}; Y|Q) - I(Y_{\mathrm{r}}; V|U, X, X_{\mathrm{r}}, Y, Q), \tag{1}$$

$$R \leq I(X; V, Y|U, X_{\mathrm{r}}, Q) + I(U; Y_{\mathrm{r}}|W, X_{\mathrm{r}}, Q), \tag{2}$$

$$R \leq I(X; V, Y|U, X_{\mathrm{r}}, Q) + I(U; Y_{\mathrm{r}}|W, X_{\mathrm{r}}, Q) + I(X_{\mathrm{r}}; Y|W, Q) - I(Y_{\mathrm{r}}; V|U, X_{\mathrm{r}}, Y, Q), \tag{3}$$

for some $p(q, w, u, v, x, x_{\mathrm{r}}, y_{\mathrm{r}}, y) = p(q)p(w|q)p(u|w, q)p(x|u, q)p(x_{\mathrm{r}}|w, q)p(y_{\mathrm{r}}, y|x, x_{\mathrm{r}})p(v|u, x_{\mathrm{r}}, y_{\mathrm{r}}, q)$.

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 [CM11], the following improved version of the above bound is given.

Theorem 2 ([CM11]). *A rate R is achievable if*

$$R \leq I(X, X_{\mathrm{r}}; Y|Q) - I(Y_{\mathrm{r}}; V|U, X, X_{\mathrm{r}}, Y, Q), \tag{4}$$

$$R \leq I(X; V, Y|U, X_{\mathrm{r}}, Q) + I(U; Y_{\mathrm{r}}|W, X_{\mathrm{r}}, Q), \tag{5}$$

$$R \leq I(X; V, Y|U, X_{\mathrm{r}}, Q) + I(U; Y_{\mathrm{r}}|W, X_{\mathrm{r}}, Q) + I(X_{\mathrm{r}}; Y|U, W, Q) - I(Y_{\mathrm{r}}; V|U, X_{\mathrm{r}}, Y, Q), \tag{6}$$

for some $p(q, w, u, v, x, x_{\mathrm{r}}, y_{\mathrm{r}}, y) = p(q)p(w|q)p(u|w, q)p(x|u, q)p(x_{\mathrm{r}}|w, q)p(y_{\mathrm{r}}, y|x, x_{\mathrm{r}})p(v|u, x_{\mathrm{r}}, y_{\mathrm{r}}, q)$.

Note that the only difference between the bounds in Theorems 1 and 2 is in one of the terms in equations (3) and (6). Since $U \to (W, Q) \to X_{\mathrm{r}}$ forms a Markov chain, we have $I(X_{\mathrm{r}}; Y|U, W, Q) = I(X_{\mathrm{r}}; Y, U|W, Q) \geq I(X_{\mathrm{r}}; Y|W, Q)$, which shows that the achievable rate in Theorem 2 is at least as large as that in Theorem 1 (see Remark 6 of [CM11]).

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 $X_{\mathrm{r}} \to W \to U \to X$ forms a Markov chain, one can verify that replacing U by U, W improves the rate constraints. Moreover, it relaxes the condition $p(v|u, x_{\mathrm{r}}, y_{\mathrm{r}})$ to $p(v|u, w, x_{\mathrm{r}}, y_{\mathrm{r}})$. This leads to the equivalent characterization of Theorem 2: A rate R is achievable if

$$\begin{aligned}
&R \leq I(X, X_{\mathrm{r}}; Y) - I(Y_{\mathrm{r}}; V|U, X, X_{\mathrm{r}}, Y, W), \\
&R \leq I(X; V, Y|U, X_{\mathrm{r}}, W) + I(U; Y_{\mathrm{r}}|W, X_{\mathrm{r}}), \\
&R \leq I(X; Y|U, X_{\mathrm{r}}, W) + I(U; Y_{\mathrm{r}}|W, X_{\mathrm{r}}) + I(X_{\mathrm{r}}; Y|U, W) - I(Y_{\mathrm{r}}; V|U, X_{\mathrm{r}}, X, Y, W),
\end{aligned} \tag{7}$$

for some $p(w,u,v,x,x_{\mathrm{r}},y_{\mathrm{r}},y) = p(u,w)p(x|u,w)p(x_{\mathrm{r}}|w)p(y_{\mathrm{r}},y|x,x_{\mathrm{r}})p(v|u,w,x_{\mathrm{r}},y_{\mathrm{r}})$.

We now state the main result of this section.

Theorem 3. *For any arbitrary relay channel $p(y_{\mathrm{r}},y_1|x)$ 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*

$$R \leq \min(I(U;Y_{\mathrm{r}}), I(U;Y_1)) + I(X;Y_1|U) + C_0 - I(Y_{\mathrm{r}};V|U,X,Y_1), \tag{8}$$
$$R \leq I(X;V,Y_1|U) + I(U;Y_{\mathrm{r}}), \tag{9}$$

for some $p(u,x,y_{\mathrm{r}},v,y_1) = p(u)p(x|u)p(y_{\mathrm{r}},y_1|x)p(v|u,y_{\mathrm{r}})$.
Further it suffices to consider $|\mathcal{V}| \leq |\mathcal{Y}_{\mathrm{r}}|+1$ and $|\mathcal{U}| \leq |\mathcal{X}|+2$ when evaluating the bound.

The proof of this theorem can be found in Section 5.1.

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

$$R \leq I(X;Y_1) + C_0 - I(Y_{\mathrm{r}};V|U,Q,X,Y_1), \tag{10}$$
$$R \leq I(Q;Y_{\mathrm{r}}) + I(U;Y_{\mathrm{r}}|Q) + I(X;Y_1|U,Q) + C_0 - I(Y_{\mathrm{r}};V|U,Q,X,Y_1), \tag{11}$$
$$R \leq I(Q;Y_{\mathrm{r}}) + I(X;V,Y_1|U,Q) + I(U;Y_{\mathrm{r}}|Q), \tag{12}$$

for some joint distribution satisfying $p(q,u,x,y_{\mathrm{r}},v,y_1) = p(q)p(x,u|q)p(y_{\mathrm{r}},y_1|x)p(v|u,q,y_{\mathrm{r}})$. We can view Q as a time-sharing random variable and interpret the term $I(Q;Y_{\mathrm{r}})$ 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 [HP79]. In [GJNW13], it was shown that the randomized time-division region coincides with Marton's inner bound for the two receiver binary-input broadcast channel.

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

Consider the relay channels with orthogonal receiver components In [MHU19], Mondelli, Hassani and Urbanke derived the following lower bound for this class of relay channels.

Theorem 4 (Theorem 1 in [MHU19])**.** *Take some arbitrary joint distribution $p(x)p(y_1,y_{\mathrm{r}}|x)p(v|y_{\mathrm{r}})$ satisfying $I(X;Y_{\mathrm{r}}) < I(X;Y_1) + C_0$ and $C_0 \leq I(V;Y_{\mathrm{r}}|Y_1)$. Then, the following rate is achievable:*

$$R = \frac{(C_0 - I_{max})I(X;V,Y_1) + \max\{I(X;Y_{\mathrm{r}}), I(X;Y_1)\}(I(V;Y_{\mathrm{r}}|Y_1) - C_0)}{I(V;Y_{\mathrm{r}}|Y_1) - I_{max}}$$

where

$$I_{max} = \max\{0, I(X;Y_{\mathrm{r}}) - I(X;Y_1)\}.$$

The authors of [MHU19] conjectured that this lower bound strictly improves over the lower bound in Theorem 1. The conjecture is justified in [MHU19, Section IV] by an incomplete simulation. We show that this conjecture is false.

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

The proof of this lemma can be found in Section 5.2.

3. Lower bound for Gaussian relay channel

Consider the Gaussian relay channel with orthogonal receiver components described by

$$Y_1 = X + W_1,$$
$$Y_{\mathrm{r}} = X + W_{\mathrm{r}},$$

where $W_1 \sim \mathcal{N}(0,N_1)$ and $W_{\mathrm{r}} \sim \mathcal{N}(0,N_{\mathrm{r}})$ 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 $S_{12} = P/N_{\mathrm{r}}$, $S_{13} = P/N_1$ and $S_{23} = 2^{2C_0} - 1$.

In the Gaussian setting, the channel $p(y_\mathrm{r}|x)$ is less noisy than $p(y_1|x)$ if $N_1 \geq N_\mathrm{r}$ and vice versa. Therefore, the term minimizing $\min(I(U;Y_\mathrm{r}), I(U;Y_1))$ in (8) can be found depending on whether $N_1 \geq N_\mathrm{r}$ 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 $p(y_\mathrm{r}|x)$ is less noisy than $p(y_1|x)$, we can express the lower bound in Theorem 3 as*

$$R \leq \min_{\alpha\in[0,1]} \max \Big[I(X;Y_1|U) + (1-\alpha)I(U;Y_\mathrm{r}) + \alpha I(U;Y_1) + \alpha C_0 - \alpha I(Y_\mathrm{r};V|U,X,Y_1) + (1-\alpha)I(X;V|Y_1,U) \Big] \tag{13}$$

where the inner maximum is over

$$p(u,x,y_\mathrm{r},v,y_1) = p(x,u)p(y_\mathrm{r},y_1|x)p(v|u,y_\mathrm{r}). \tag{14}$$

If $p(y_1|x)$ is less noisy than $p(y_\mathrm{r}|x)$, we can express the lower bound in Theorem 3 as

$$R \leq \min_{\alpha\in[0,1]} \max \Big[I(X;Y_1|U) + I(U;Y_\mathrm{r}) + \alpha C_0 - \alpha I(Y_\mathrm{r};V|U,X,Y_1) + (1-\alpha)I(X;V|Y_1,U) \Big],$$

where the inner maximum is over (14). Further it suffices to consider $|\mathcal{V}| \leq |\mathcal{Y}_\mathrm{r}|$ and $|\mathcal{U}| \leq |\mathcal{X}|$ when evaluating the above expressions.

The proof of this lemma can be found in Section 5.3.

3.1. Evaluation for Gaussian inputs with randomized power control

The joint distribution attaining 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 $N_1 \geq N_\mathrm{r}$, we can write the bound in (13) (specialized to the above joint distribution on auxiliary variables) as

$$\min_{\alpha\in[0,1]} \max_{p(q,x)} \Big[(1-\alpha)I(Q;Y_\mathrm{r}) + \alpha I(Q;Y_1) + \alpha C_0 + \sum_q G_\alpha^{(1)}(P_q) \Big] \tag{15}$$

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

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

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

$$\min_{\alpha\in[0,1]} \max_{p(q,x)} \Big[I(Q;Y_\mathrm{r}) + \alpha C_0 + \sum_q G_\alpha^{(2)}(P_q) \Big] \tag{16}$$

where

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

and maximum in $G_\alpha^{(2)}(P)$ is taken over the same set as $G_\alpha^{(1)}(P)$.

[1] Power control is known to be helpful to improve the compress-forward rate [EK11, Sec. 16.8].

The following lemma gives an explicit expression for $G_\alpha^{(1)}(P)$ and $G_\alpha^{(2)}(P)$:

Lemma 3. *For $N_{\rm r} \leq N_1$ we have*

$$G_\alpha^{(1)}(P) = \begin{cases} \frac{\alpha}{2}\log(1+S_{13}) + \max\Big(\frac{1-\alpha}{2}\log(1+S_{12}), & \alpha \in \left[0, \frac{S_{12}}{1+S_{12}+S_{13}+1}\right] \\ \quad \frac{1}{2}\log(S_{12}+S_{13}+1) - \frac{\alpha}{2}\log(S_{12}) - \frac{1}{2}h(\alpha)\Big) & \\ & \\ \frac{\alpha}{2}\log(1+S_{13}) + \frac{1-\alpha}{2}\log(1+S_{12}) & \alpha \in \left[\frac{S_{12}}{1+S_{12}+S_{13}+1}, 1\right] \end{cases},$$

where $h(\alpha)$ is the binary entropy function and $S_{12} = P/N_{\rm r}$, $S_{13} = P/N_1$.
For $N_{\rm r} \geq N_1$ we have

$$G_\alpha^{(2)}(P) = \begin{cases} \frac{\alpha}{2}\log(1+S_{13}) + \frac{\alpha}{2}\log\frac{1+S_{13}}{S_{12}} + \frac{1}{2}\log(1+S_{12}+S_{13}) - \frac{1}{2}h(\alpha), & \alpha \in \left[0, \frac{S_{12}}{1+S_{12}+S_{13}+1}\right] \\ \frac{1}{2}\log(1+S_{13}) & \alpha \in \left[\frac{S_{12}}{1+S_{12}+S_{13}+1}, 1\right] \end{cases},$$

The proof of this lemma can be found in Section 5.4.

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):

$$\begin{aligned} &\min_{\alpha\in[0,1]}\max_{p(q,x)}\left[(1-\alpha)h(Y_{\rm r}) - (1-\alpha)h(Y_{\rm r}|Q) + \alpha h(Y_1) - \alpha h(Y_1|Q) + \alpha C_0 + \sum_q G_\alpha^{(1)}(P_q)\right] \\ &\leq \min_{\alpha\in[0,1]}\max_{p(q,x)}\left[\frac{1-\alpha}{2}\log(2\pi e(P+N_{\rm r})) - (1-\alpha)h(Y_{\rm r}|Q) + \frac{\alpha}{2}\log(2\pi e(P+N_1)) - \alpha h(Y_1|Q)\right. \\ &\qquad\qquad \left. + \alpha C_0 + \sum_q G_\alpha^{(1)}(P_q)\right], \end{aligned} \tag{17}$$

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 ([EGGN21]). *Any achievable rate R for the relay channel with orthogonal receiver components $p(y_1, y_{\rm r}|x)$ with a relay-to-receiver link of capacity C_0 must satisfy the following inequalities*

$$R \leq I(X;Y_1,Y_{\rm r}) - I(V;Y_1|Y_{\rm r}) - I(X;Y_{\rm r}|V,Y_1), \tag{18}$$
$$R \leq I(X;Y_1) + C_0 - I(V;Y_{\rm r}|X,Y_1), \tag{19}$$

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

For the Gaussian relay channel, this upper bound simplifies to the following expression [EGGN21, Proposition 4] [WBÖ17]:

$$C \leq \begin{cases} \frac{1}{2}\log\left(1 + S_{13} + \frac{S_{12}(S_{13}+1)S_{23}}{(S_{13}+1)(S_{23}+1)-1}\right), & \text{for } S_{12} \leq S_{13} + S_{23} + S_{13}S_{23}, \\ \frac{1}{2}\log((1+S_{13})(1+S_{23})), & \text{otherwise.} \end{cases} \tag{20}$$

The figure also plots the decode-forward lower bound and the compress-forward lower bound for this relay channel whose expressions are given in [EK11, Eq. 16.16] and [EK11, Eq. 16.17]. When $S_{12} \geq S_{13} + S_{23} + S_{13}S_{23}$, the upper bound in (20), the cut-set bound and the decode-forward lower bound all coincide. This corresponds to the top corner point in the figure.

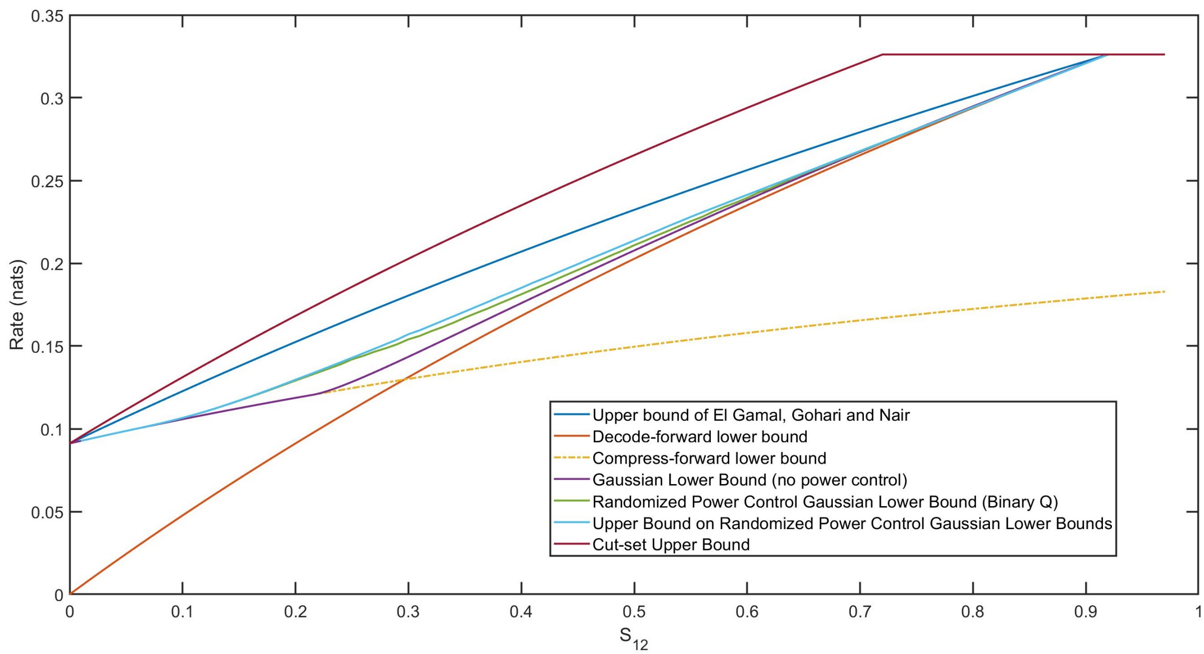

Fig. 2. Plots of the bounds for the Gaussian product-form relay channel with $S_{13} = 0.2$, $S_{23} = 0.6$.

4. Disproving Ahlswede-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

$$p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}}) = p(y_{\mathrm{r}})p(y_1|x, y_{\mathrm{r}})p(y_2|x_{\mathrm{r}}). \tag{21}$$

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

$$\begin{aligned} R &\leq I(X; V, Y_1|U) = I(X; Y_1|V, U) \\ R &\leq I(X; Y_1|U) + C_0 - I(Y_{\mathrm{r}}; V|U, X, Y_1) \end{aligned}$$

for some

$$p(u, x, y_{\mathrm{r}}, v, y_1) = p(u, x)p(y_{\mathrm{r}})p(y_1|x)p(v|u, y_{\mathrm{r}}).$$

The above rate is just the compress-forward rate with a time-sharing variable U. If we set $U = \emptyset$, we obtain he compress-forward rate without a time-sharing variable [EK11, Eq. 16.14].

Communication over a relay channel of the form (21) is equivalent to communication over a channel with rate-limited state information available at the receiver [ARY09]. As such, the Ahlswede-Han conjecture in [AH83] 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 $p(x)p(y_{\mathrm{r}})p(y_1|x, y_{\mathrm{r}})p(v|y_{\mathrm{r}})$ such that $I(V; Y_{\mathrm{r}}|Y_1) \leq C_0$. In other words, the Ahlswede-Han conjecture is equivalent to the optimally 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 [EK11, Sec. 16.8]), the Ahlswede-Han conjecture implies that time-sharing is unnecessary for the special case of relay channels of the form (21). Interestingly, the Ahlswede-Han conjecture is valid for all channels for which the capacity is known [Kim08], [ARY09], [TU08].

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.

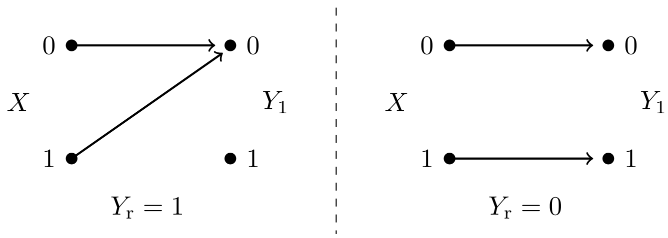

Fig. 3. Conditional probability $p(y_1|x, y_{\mathrm{r}})$. We assume $\mathbb{P}[Y_{\mathrm{r}}=0]=\beta$ and $\mathbb{P}[X=0]=\theta$.

Lemma 4. *Take a discrete memoryless relay channel with orthogonal receiver components of the form* $p(y_{\mathrm{r}})p(y_1|x, y_{\mathrm{r}})$. *Let*

$$F(p(x),\lambda)=\max_{p(v|y_{\mathrm{r}})} I(X;Y_1|V)+\lambda(C_0-I(V;Y_{\mathrm{r}}|Y_1)). \tag{22}$$

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

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

The proof of this lemma can be found in Section 5.5. 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 $p(x,y_{\mathrm{r}},y_1)=p(x)p(y_{\mathrm{r}})p(y_1|x,y_{\mathrm{r}})$ is defined as follows: assume that $Y_1=0$ if $Y_{\mathrm{r}}=1$, and $Y_1=X$ if $Y_{\mathrm{r}}=0$ (see Fig. 3). Let $\mathbb{P}[Y_{\mathrm{r}}=0]=\beta$.

Let $\mathbb{P}[Y_{\mathrm{r}}=0|v]=q$ 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

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

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

$$F(\theta,\lambda)=\lambda(C_0-H_2(\beta)+H_2(\beta\bar{\theta}))+\mathscr{C}[(1-\lambda)H_2(q\bar{\theta})+(\lambda-\bar{\theta})H_2(q)]_{q=\beta}, \tag{23}$$

where $\mathscr{C}[(1-\lambda)H_2(q\bar{\theta})+(\lambda-\bar{\theta})H_2(q)]$ is the upper concave envelope[2] of the function $q\mapsto(1-\lambda)H_2(q\bar{\theta})+(\lambda-\bar{\theta})H_2(q)$ defined for $q\in[0,1]$. An explicit expression for $\mathscr{C}[(1-\lambda)H_2(q\bar{\theta})+(\lambda-\bar{\theta})H_2(q)]$ is given in Appendix A.

Let $C_0=0.06$, $\beta=0.3$. We claim that

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

Let

$$\mathcal{P}=\{t: t=0.001+0.01k \text{ for } k\in[0:99]\}.$$

Define

$$\begin{aligned} G_{\mathcal{P}}(\theta)&:=\min_{\lambda\in\mathcal{P}} F(\theta,\lambda)\\ \hat{G}_{\mathcal{P}}(\lambda)&:=\max_{\theta\in\mathcal{P}} F(\theta,\lambda).\end{aligned}$$

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

Observe that

$$\min_{\lambda\in[0,1]} F(\theta,\lambda) \le G_{\mathcal{P}}(\theta)$$
$$\max_{\theta\in[0,1]} F(\theta,\lambda) \ge \hat{G}_{\mathcal{P}}(\lambda).$$

Therefore $\max_\theta \min_\lambda F(\theta,\lambda) < \min_\lambda \max_\theta F(\theta,\lambda)$ holds if

$$\max_{\theta\in[0,1]} G_{\mathcal{P}}(\theta) < \min_{\lambda\in[0,1]} \hat{G}_{\mathcal{P}}(\lambda).$$

The plots of $G_{\mathcal{P}}(\theta)$ and $\hat{G}_{\mathcal{P}}(\lambda)$ are given in Figures 5 and 7 respectively (in Appendix B). The plots are zoomed in near their global maximum and minimum in Figures 6 and 8.

In particular, the maximum over θ of $G_{\mathcal{P}}(\theta)$ occurs around 0.5772631 and yields a value in $(0.191232, 0.191235)$. On the other hand, minimum over λ in $\hat{G}_{\mathcal{P}}(\lambda)$ occurs around 0.24868889 and yields a value in $(0.191898, 0.1919)$. This demonstrates that time-sharing does improve the rate and provides a counterexample to the Ahlswede-Han conjecture.

5. Proofs of the results

In the following sections we present the proofs of the results stated in the previous section in their order of appearance.

5.1. *Proof of Theorem 3*

Let $\mathcal{R}_0$ be the bound of Theorem 1, $\mathcal{R}_1$ be the region given in Remark 1 and $\mathcal{R}_2$ be the region defined in the statement of Theorem 3. We know that $\mathcal{R}_0 \subseteq \mathcal{R}_1$. It is clear that $\mathcal{R}_2 \subseteq \mathcal{R}_0$ by taking $W = Q = \emptyset$ and X_{r} independent of (U, V, X) and $p(x_{\mathrm{r}})$ a capacity achieving distribution for the channel $p(y_2|x_{\mathrm{r}})$. It remains to show that $\mathcal{R}_1 \subseteq \mathcal{R}_2$.

Take some arbitrary U, V, W for $\mathcal{R}_1$. Observe that

$$I(X, X_{\mathrm{r}}; Y) \le I(X; Y_1) + C_0, \tag{24}$$
$$I(X_{\mathrm{r}}; Y|U, W) \le C_0. \tag{25}$$

Let $\tilde{W} = (W, X_{\mathrm{r}})$. Then, the constraints on R in $\mathcal{R}_1$ along with (24) and (25) imply that the rate R satisfies

$$\begin{aligned}
R &\le I(X; Y_1) + C_0 - I(Y_{\mathrm{r}}; V|U, X, Y, \tilde{W})\\
R &\le I(X; V, Y|U, X_{\mathrm{r}}, \tilde{W}) + I(U, \tilde{W}; Y_{\mathrm{r}})\\
&= I(X; V, Y_1|U, X_{\mathrm{r}}, \tilde{W}) + I(U, \tilde{W}; Y_{\mathrm{r}})\\
R &\le I(X; Y|U, \tilde{W}) + I(U, \tilde{W}; Y_{\mathrm{r}}) + C_0 - I(Y_{\mathrm{r}}; V|U, X, Y, \tilde{W})\\
&= I(X; Y_1|U, \tilde{W}) + I(U, \tilde{W}; Y_{\mathrm{r}}) + C_0 - I(Y_{\mathrm{r}}; V|U, X, Y_1, \tilde{W})
\end{aligned}$$

The constraint

$$p(w, u, x, x_{\mathrm{r}}, y_{\mathrm{r}}, v, y) = p(u, w)p(x|u, w)p(x_{\mathrm{r}}|w)p(y_{\mathrm{r}}, y|x, x_{\mathrm{r}})p(v|u, x_{\mathrm{r}}, y_{\mathrm{r}}, w)$$

implies

$$p(u, \tilde{w}, x, y_{\mathrm{r}}, v, y_1) = p(u, \tilde{w})p(x|u, \tilde{w})p(y_{\mathrm{r}}|x)p(y_1|x)p(v|u, \tilde{w}, y_{\mathrm{r}}).$$

Setting $\tilde{U} = (U, \tilde{W})$ shows that $\mathcal{R}_1 \subseteq \mathcal{R}_2$. This completes the proof for $\mathcal{R}_0 = \mathcal{R}_1 = \mathcal{R}_2$. The cardinality bounds on the auxiliary random variables in $\mathcal{R}_2$ come from the standard Caratheodory-Bunt [Bun34] arguments and is omitted.

5.2. *Proof of Lemma 1*

It turns out that the rate in Theorem 4 corresponds to a particular choice of variables Q, U, V in Remark 2.

Take some arbitrary $p(x)$ and $p(v'|y_{\mathrm{r}})$ satisfying the constraints of Theorem 4:

$$I(X; Y_{\mathrm{r}}) < I(X; Y_1) + C_0, \tag{26}$$
$$I(V'; Y_{\mathrm{r}}|Y_1) \ge C_0. \tag{27}$$

Consider the following choice for auxiliary variables U, V and Q in Remark 2:
Let $Q \in \{0,1\}$ be a Bernouli random variable independent of (X, Y_1, Y_r) satisfying

$$p(Q=1) = \frac{C_0 - I_{\max}}{I(V'; Y_\mathrm{r}|Y_1) - I_{\max}},$$

where $I_{\max} = \max\{0, I(X;Y_\mathrm{r}) - I(X;Y_1)\}$.
(i) If $I(X, Y_\mathrm{r}) \le I(X; Y_1)$, define U and V as follows: when $Q = 0$ set $U = V = \emptyset$ and when $Q = 1$ set $U = \emptyset$ and $V = V'$.
(ii) If $I(X, Y_\mathrm{r}) > I(X; Y_1)$, define U and V as follows: when $Q = 0$ set $U = X$, $V = \emptyset$, and when $Q = 1$ set $U = \emptyset$ and $V = V'$.
One can verify that for this choice for auxiliary variables U, Q and V, the lower bound in Remark 2 reduces to the one given in [MHU19].

5.3. Proof of Lemma 2

If $p(y_\mathrm{r}|x)$ is less noisy than $p(y_1|x)$ we can write the bound in Theorem 3 as

$$R \le I(U; Y_1) + I(X; Y_1|U) + C_0 - I(Y_\mathrm{r}; V|U, X, Y_1), \tag{28}$$

$$R \le I(X; V, Y_1|U) + I(U; Y_\mathrm{r}), \tag{29}$$

for some joint distribution distribution $p(u, v, x, y_1, y_\mathrm{r})$ satisfying

$$p(u, x, y_\mathrm{r}, v, y_1) = p(x, u)p(y_\mathrm{r}, y_1|x)p(v|u, y_\mathrm{r}). \tag{30}$$

Equivalently, we can write this bound as

$$\max \min_{\alpha\in[0,1]} \Big[I(X; Y_1|U) + (1-\alpha)I(U; Y_\mathrm{r}) + \alpha I(U; Y_1) + \alpha C_0 - \alpha I(Y_\mathrm{r}; V|U, X, Y_1) + (1-\alpha)I(X; V|Y_1, U) \Big]$$

where the outer maximum is over distributions of the form given in (30). It remains to prove that the maximum and minimum can be exchanged. The proof follows from [GGNY13, Corollary 2] and mimics the proof of the exchange of max and min in the proof of Lemma 1 in [GGNY13] by noting that the auxiliary variable U convexifies the region.

The other case of $p(y_1|x)$ being less noisy than $p(y_\mathrm{r}|x)$ similar. The cardinality bounds on the auxiliary random variables come from the standard Caratheodory-Bunt [Bun34] arguments and is omitted.

5.4. Proof of Lemma 3

Let $\mathsf{Var}[X|U] = aP$ for some $a \in [0,1]$, and $\mathsf{Var}[Y_\mathrm{r}|U, V] = b(aP + N_\mathrm{r})$ for some $b \in [0,1]$. Then,

$$K_{X,Y_\mathrm{r}|U} = \begin{bmatrix} aP & aP \\ aP & aP + N_\mathrm{r} \end{bmatrix}$$

and

$$K_{X,Y_\mathrm{r}|U,V} = \begin{bmatrix} \frac{aP(abP+N_\mathrm{r})}{aP+N_\mathrm{r}} & abP \\ abP & b(aP+N_\mathrm{r}) \end{bmatrix}$$

Case of $N_1 \ge N_\mathrm{r}$: Then, we obtain

$$G_\alpha^{(1)} = \max_{a,b\in[0,1]} \alpha \left(\frac{1}{2}\log(\frac{P+N_1}{N_1}) + \frac{1}{2}\log\frac{b(aP+N_\mathrm{r})}{abP+N_\mathrm{r}} \right) + (1-\alpha)\left(\frac{1}{2}\log\frac{P+N_\mathrm{r}}{N_1} + \frac{1}{2}\log\left(\frac{aP}{aP+N_\mathrm{r}} + \frac{N_1}{abP+N_\mathrm{r}} \right) \right).$$

To obtain the optimal a, b, we would like to maximize the expression

$$\alpha \log\frac{b(aP+N_\mathrm{r})}{abP+N_\mathrm{r}} + \bar{\alpha}\log\left(\frac{aP}{aP+N_\mathrm{r}} + \frac{N_1}{abP+N_\mathrm{r}} \right).$$

Let us reparameterize according to

$$x := \frac{aP}{aP+N_{\mathrm{r}}}, \quad y := \frac{b(aP+N_{\mathrm{r}})}{abP+N_{\mathrm{r}}}$$

which imposes the constraints $0 \le x \le \frac{P}{P+N_{\mathrm{r}}}$ and $0 \le y \le 1$. Now we would like to maximize the expression

$$\alpha \log y + \bar{\alpha} \log\left(x + \frac{N_1}{N_{\mathrm{r}}}(1-yx)\right),$$

and it is clear that the maximum is attained at $x=0$ or $x = \frac{P}{P+N_{\mathrm{r}}}$. If $x=0$, then $y=1$ is the maximizer; while if $x = \frac{P}{P+N_{\mathrm{r}}}$ (the expression is concave in y), we have $y^* = \min\left\{1, \frac{\alpha(P(N_1+N_{\mathrm{r}})+N_1N_{\mathrm{r}})}{PN_1}\right\}$ as the maximizer.

Therefore putting things together, substituting, and simplifying, we obtain the following:
(i) When $\alpha \in [\frac{S_{12}}{1+S_{12}+S_{13}}, 1]$ then

$$G_\alpha^{(1)} = \frac{\alpha}{2}\log(1+S_{13}) + \frac{1-\alpha}{2}\log(1+S_{12}),$$

(ii) When $\alpha \in [0, \frac{S_{12}}{1+S_{12}+S_{13}})$ then

$$G_\alpha^{(1)} = \frac{\alpha}{2}\log(1+S_{13}) + \max\left(\frac{1-\alpha}{2}\log(1+S_{12}), \frac{1}{2}\log(S_{12}+S_{13}+1) - \frac{\alpha}{2}\log(S_{12}) - \frac{1}{2}h(\alpha)\right).$$

Case of $N_1 \le N_{\mathrm{r}}$: In this case we can write

$$\begin{aligned} G_\alpha^{(2)}(P) = \max_{a,b\in[0,1]} \alpha &\left(\frac{1}{2}\log\frac{P+N_{\mathrm{r}}}{N_1} + \frac{1}{2}\log\frac{b(aP+N_1)}{abP+N_{\mathrm{r}}}\right) \\ &+ (1-\alpha)\left(\frac{1}{2}\log\frac{P+N_{\mathrm{r}}}{N_1} + \frac{1}{2}\log\left(\frac{aP}{aP+N_{\mathrm{r}}} + \frac{N_1}{abP+N_{\mathrm{r}}}\right)\right). \end{aligned}$$

To obtain the optimal a, b, we would like to maximize the expression

$$\alpha \log\frac{b(aP+N_1)}{abP+N_{\mathrm{r}}} + (1-\alpha)\log\left(\frac{aP}{aP+N_{\mathrm{r}}} + \frac{N_1}{abP+N_{\mathrm{r}}}\right). \tag{31}$$

Again, let us reparameterize according to

$$x := \frac{aP}{aP+N_{\mathrm{r}}}, \quad t := \frac{b(aP+N_1)}{abP+N_{\mathrm{r}}}$$

which imposes the constraints $0 \le x \le \frac{P}{P+N_{\mathrm{r}}}$ and $0 \le t \le \frac{xN_{\mathrm{r}}+(1-x)N_1}{N_{\mathrm{r}}}$. Now we would like to maximize the expression

$$\alpha \log t + \bar{\alpha}\log\left(\frac{xN_{\mathrm{r}}(xN_{\mathrm{r}}+(1-x)N_1) + N_1x(1-t)N_{\mathrm{r}} + N_1^2(1-x)}{N_{\mathrm{r}}(xN_{\mathrm{r}}+(1-x)N_1)}\right),$$

This expression is increasing with respect to t at $t=0$ and hence is maximized at

$$t_* = \min\left\{\frac{xN_{\mathrm{r}}+(1-x)N_1}{N_{\mathrm{r}}}, \frac{\alpha(xN_{\mathrm{r}}+N_1)(xN_{\mathrm{r}}+(1-x)N_1)}{xN_1N_{\mathrm{r}}}\right\}$$

where the latter point is the unique stationary point in t. This breaks into two cases:
(i): If $\alpha \in \left[\frac{S_{12}}{1+S_{12}+S_{13}}, 1\right]$ then $t_* = \frac{xN_{\mathrm{r}}+(1-x)N_1}{N_{\mathrm{r}}}$.

Now we have to maximize (after simplification) the expression

$$\log\left(\frac{xN_{\mathrm{r}}+(1-x)N_1}{N_{\mathrm{r}}}\right).$$

The above expression is increasing in x and maximized at $x = \frac{P}{P+N_{\mathrm{r}}}$ leading to

$$G_\alpha^{(2)}(P) = \frac{1}{2}\log\frac{P+N_{\mathrm{r}}}{N_1} + \frac{1}{2}\log\frac{P+N_1}{P+N_{\mathrm{r}}} = \frac{1}{2}\log(1+S_{13}).$$

(ii): If $\alpha \in \left[0, \frac{S_{12}}{1+S_{12}+S_{13}}\right]$ then we decompose into two sub-cases:

(a): If $x \in \left[0, \frac{\alpha N_1}{N_1 - \alpha N_{\mathrm{r}}}\right]$ then again $t^* = \frac{xN_{\mathrm{r}}+(1-x)N_1}{N_{\mathrm{r}}}$. As in the earlier case the resultant expression is increasing in x, and hence achieves it maximum at the boundary. By continuity of the expression in x, this implies that the global maximum in x will lie in $x \in \left[\frac{\alpha N_1}{N_1-\alpha N_{\mathrm{r}}}, \frac{P}{P+N_{\mathrm{r}}}\right]$.

(b): If $x \in \left[\frac{\alpha N_1}{N_1-\alpha N_{\mathrm{r}}}, \frac{P}{P+N_{\mathrm{r}}}\right]$ then $t^* = \frac{\alpha(xN_{\mathrm{r}}+N_1)(xN_{\mathrm{r}}+(1-x)N_1)}{xN_1N_{\mathrm{r}}}$. Now we have to maximize (after simplification) the expression

$$\alpha \log\left(\frac{\alpha(xN_{\mathrm{r}}+N_1)(xN_{\mathrm{r}}+(1-x)N_1)}{xN_1N_{\mathrm{r}}}\right) + (1-\alpha)\log\left(\frac{(1-\alpha)(xN_{\mathrm{r}}+N_1)}{N_{\mathrm{r}}}\right).$$

The derivative of the above expression in x is non-negative when

$$\alpha \le \frac{xN_{\mathrm{r}}}{xN_{\mathrm{r}}+N_1} + \frac{\alpha x(N_{\mathrm{r}}-N_1)}{x(N_{\mathrm{r}}-N_1)+N_1}.$$

It is immediate to verify that this holds when $x \in \left[\frac{\alpha N_1}{N_1-\alpha N_{\mathrm{r}}}, \frac{P}{P+N_{\mathrm{r}}}\right]$ as it suffices to verify that it holds when $x = \frac{\alpha N_1}{N_1-\alpha N_{\mathrm{r}}}$. Thus the overall maximizer is when $x = \frac{P}{P+N_{\mathrm{r}}}$. Plugging in and simplifying, when $\alpha \in \left[0, \frac{PN_1}{PN_{\mathrm{r}}+N_1(P+N_{\mathrm{r}})}\right]$, we obtain that

$$G_\alpha^{(2)}(P) = -\frac{1}{2}h(\alpha) + \frac{\alpha}{2}\log\frac{1+S_{13}}{S_{12}} + \frac{1}{2}\log(1+S_{12}+S_{13}).$$

5.5. Proof of Lemma 4

The compress-forward rate with time-sharing has the following equivalent characterization (see Proposition 3 of [Kim07]): it equals $\max I(X;Y_1,V|U)$ over all $p(u|x)p(v|u,y_{\mathrm{r}})$ subject to $I(V;Y_{\mathrm{r}}|U,Y_1) \le C_0$. This expression is equivalent to

$$\begin{aligned}
&\max_{p(u,x)p(v|u,y_{\mathrm{r}})} \min_{\lambda\ge 0} I(X;Y_1,V|U) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|U,Y_1)) \\
&= \min_{\lambda\ge 0} \max_{p(u,x)p(v|u,y_{\mathrm{r}})} I(X;Y_1,V|U) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|U,Y_1)) \qquad (32)\\
&= \min_{\lambda\ge 0} \max_{p(x)} \max_{p(v|y_{\mathrm{r}})} I(X;Y_1|V) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|Y_1)) \\
&= \min_{\lambda\in[0,1]} \max_{p(x)} \max_{p(v|y_{\mathrm{r}})} I(X;Y_1|V) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|Y_1)) \qquad (33)
\end{aligned}$$

where (33) holds because if $\lambda \ge 1$ the optimal choice of V is a constant. The exchange of maximum and minimum is justified as in the proof of Lemma 2 in Section 5.3. To apply the min-max interchange in (32), one can use [Sim95, Theorem 12] which allows for the exchange of max and min without requiring λ to be in a compact set. The set of distributions $p(u,x)p(v|u,y_{\mathrm{r}})$ can be restricted to a compact set as cardinality bounds can be imposed on the alphabet sets of U and V. This theorem is applied to the function $-[I(X;Y_1,V|U) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|U,Y_1))]$ which satisfies the conditions of [Sim95, Theorem 12] by a similar argument as in the proof of Lemma 1 in [GGNY13].

The compress-forward rate without time-sharing has the following equivalent characterization (see Proposition 3 of [Kim07]): it equals $\max I(X;Y_1,V)$ over $p(v|y_{\mathrm{r}})$ subject to $I(V;Y_{\mathrm{r}}|Y_1) \le C_0$. This expression is equal to

$$\begin{aligned}
&\max_{p(x)} \max_{p(v|y_{\mathrm{r}})} \min_{\lambda\ge 0} I(X;Y_1|V) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|Y_1)) \\
&= \max_{p(x)} \min_{\lambda\ge 0} \max_{p(v|y_{\mathrm{r}})} I(X;Y_1|V) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|Y_1)) \\
&= \max_{p(x)} \min_{\lambda\in[0,1]} \max_{p(v|y_{\mathrm{r}})} I(X;Y_1|V) + \lambda(C_0 - I(V;Y_{\mathrm{r}}|Y_1)).
\end{aligned}$$

Justification for restriction to $\lambda \in [0,1]$ and the exchange of maximum and minimum is similar to the earlier case.

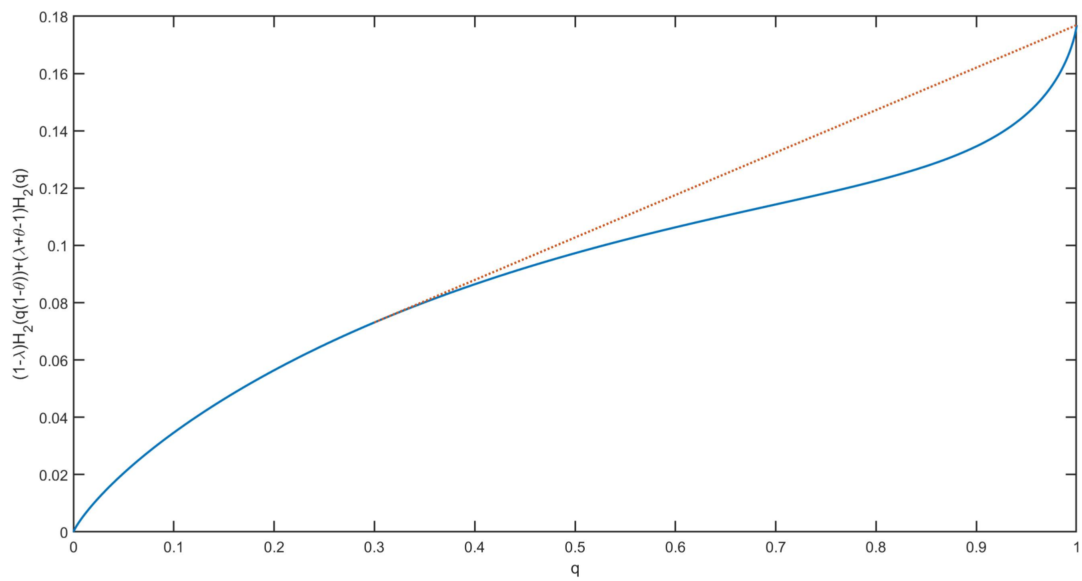

Fig. 4. Plot of $q \mapsto (1-\lambda)H_2(q(1-\theta)) + (\lambda+\theta-1)H_2(q)$ for $\theta = 0.1$ and $q' = 0.3$ (λ is found from (34)).

Appendix A
Upper concave envelope calculation

In this appendix we compute the upper concave envelope of the function $q \mapsto (1-\lambda)H_2(q(1-\theta)) + (\lambda+\theta-1)H_2(q)$ at $q = \beta \in [0, 1]$. If $\lambda \geq 1-\theta$, this function is concave in q. Thus,

$$\mathscr{C}[(1-\lambda)H_2(q\bar{\theta}) + (\lambda-\bar{\theta})H_2(q)]_{q=\beta} = (1-\lambda)H_2(\beta\bar{\theta}) + (\lambda-\bar{\theta})H_2(\beta).$$

If $\lambda < \bar{\theta}$, the second derivative of the function $q \mapsto (1-\lambda)H_2(q\bar{\theta}) + (\lambda-\bar{\theta})H_2(q)$ will have a unique root in $(0, 1)$. The function is initially concave and then becomes convex; see Figure 4 for an illustration where the function and its upper concave envelope are plotted. Since $\lambda < \bar{\theta}$, one can parameterize λ as

$$\lambda = \frac{\log \frac{1-q'\bar{\theta}}{\theta}}{\log \frac{1-q'\bar{\theta}}{\theta} + \log \frac{1}{q'}} \tag{34}$$

for some $q' \in [0, 1)$. In this case, the upper concave envelope of the function $q \mapsto (1-\lambda)H_2(q\bar{\theta}) + (\lambda-\bar{\theta})H_2(q)$ can be computed as follows:

$$\mathscr{C}\left[(1-\lambda)H_2(q\bar{\theta}) + (\lambda-\bar{\theta})H_2(q)\right] = \begin{cases} (1-\lambda)H_2(q\bar{\theta}) + (\lambda-\bar{\theta})H_2(q) & q \in [0, q'), \\ \omega\left[(1-\lambda)H_2(q'\bar{\theta}) + (\lambda-\bar{\theta})H_2(q')\right] + \bar{\omega}(1-\lambda)H_2(\bar{\theta}) & q \in [q', 1], \end{cases}$$

where

$$\omega = \frac{1-q}{1-q'}.$$

In other words, the tangent points of the line in the upper concave envelope are at $q = q'$ and $q = 1$.

Appendix B
Plots of the figures

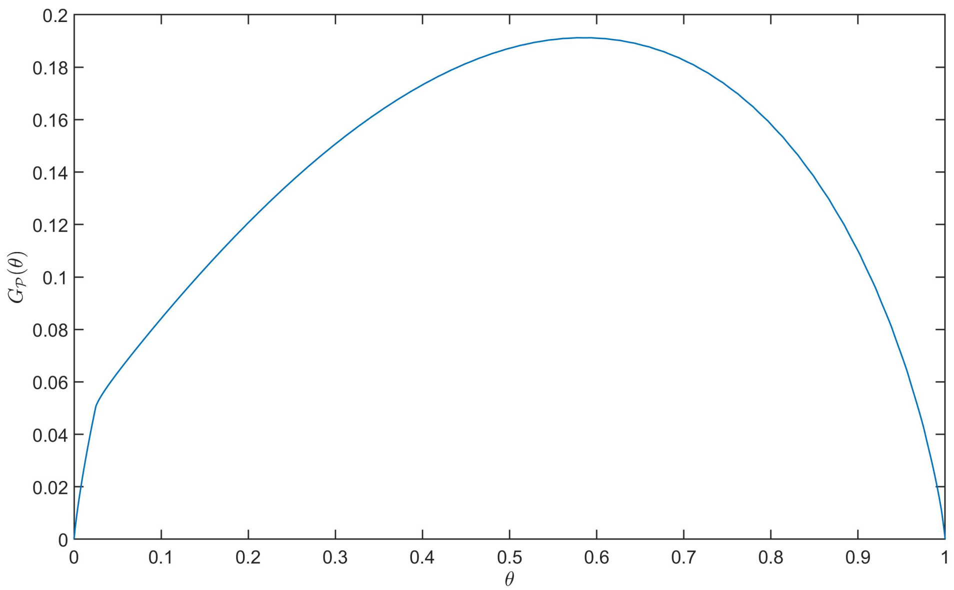

Fig. 5. Plot of the function $G_{\mathcal{P}}(\theta) = \min_{\lambda \in \mathcal{P}} F(\theta, \lambda)$.

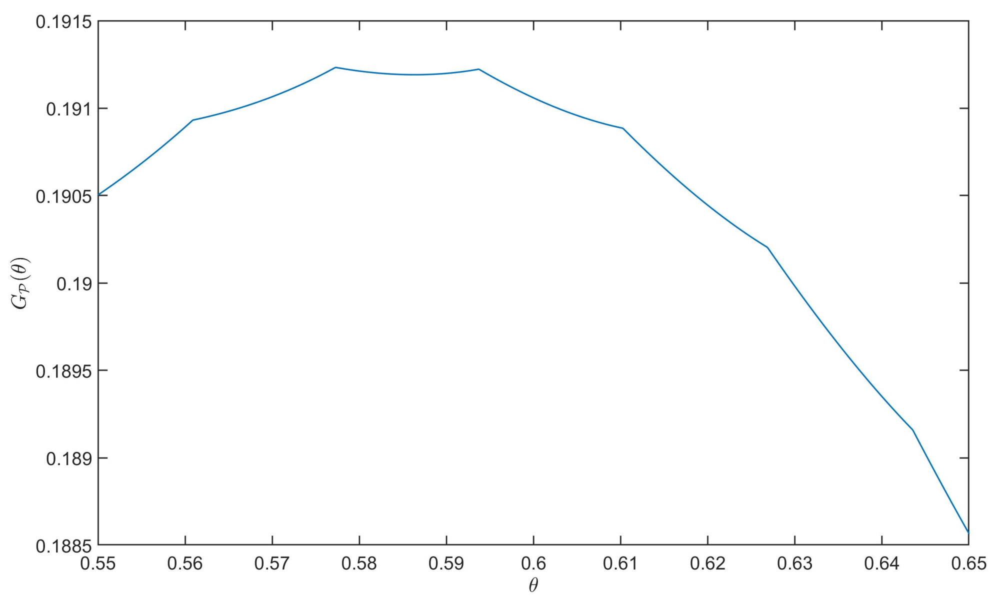

Fig. 6. Plot of the function $G_{\mathcal{P}}(\theta)$ zoomed in near the maximum. Observe that the curve is below 0.1915

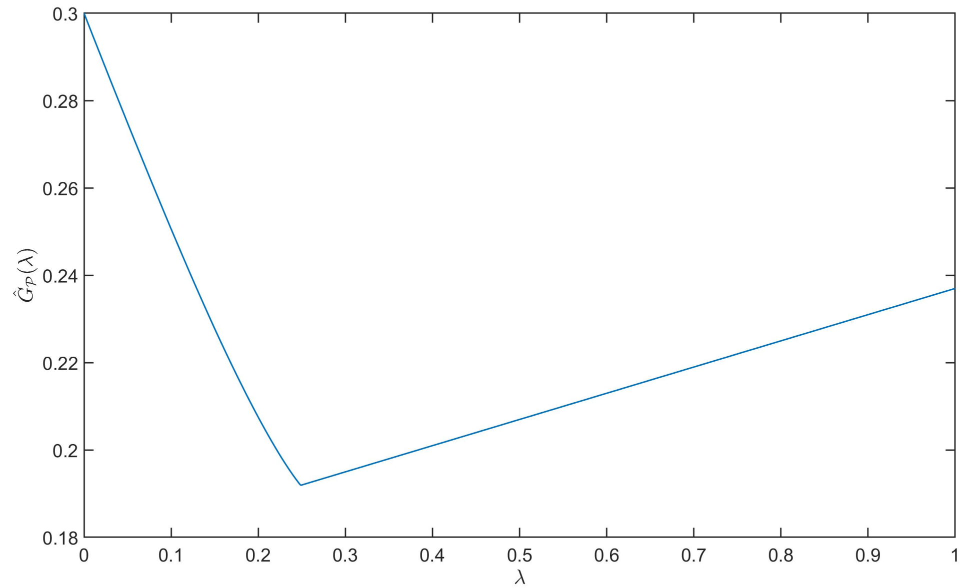

Fig. 7. Plot of the function $\hat{G}_{\mathcal{P}}(\lambda) = \max_{\theta \in \mathcal{P}} F(\theta, \lambda)$.

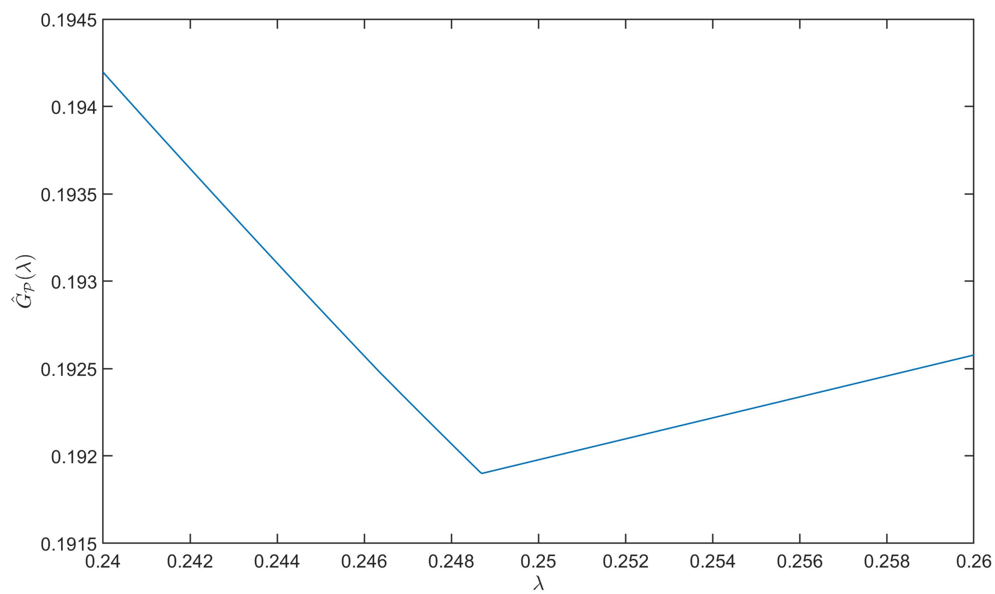

Fig. 8. Plot of the function $\hat{G}_{\mathcal{P}}(\lambda)$ zoomed in near the minimum. Observe that the curve is above 0.1915

References

[AH83] Rudolf Ahlswede and Te 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 **29** (1983), no. 3, 396–412.

[ARY09] Marko Aleksic, Peyman Razaghi, and Wei Yu, *Capacity of a class of modulo-sum relay channels*, IEEE Transactions on Information Theory **55** (2009), no. 3, 921–930.

[Bun34] Lucas Nicolaas Hendrik Bunt, *Bijdrage tot de theorie der convexe puntverzamelingen*, Ph.D. thesis, Univ. Groningne, Amsterdam, 1934.

[CEG79] Thomas Cover and Abbas El Gamal, *Capacity theorems for the relay channel*, IEEE Transactions on information theory **25** (1979), no. 5, 572–584.

[CG79] Tom Cover and Abbas El Gamal, *Capacity theorems for the relay channel*, IEEE Transactions on Information Theory **25** (1979), no. 5, 572–584.

[CM11] Hon-Fah Chong and Mehul Motani, *On achievable rates for the general relay channel*, IEEE Transactions on Information Theory **57** (2011), no. 3, 1249–1266.

[CMG07] Hon-Fah Chong, Mehul Motani, and Hari Krishna Garg, *Generalized backward decoding strategies for the relay channel*, IEEE Transactions on Information Theory **53** (2007), no. 1, 394–401.

[Cov87] Thomas Cover, *The capacity of the relay channel*, Open Problems in Communication and Computation, Springer, 1987, pp. 72–73.

[EGGN21] Abbas El Gamal, Amin Gohari, and Chandra Nair, *Strengthened cutset upper bounds on the capacity of the relay channel and applications*, arXiv preprint arxiv:2101.11139 (2021).

[EK11] Abbas El Gamal and Young-Han Kim, *Network information theory*, Cambridge University Press, 2011.

[GGNY13] Yanlin Geng, Amin Gohari, Chandra Nair, and Yuanming Yu, *On marton's inner bound and its optimality for classes of product broadcast channels*, IEEE transactions on information theory **60** (2013), no. 1, 22–41.

[GJNW13] Yanlin Geng, Varun Jog, Chandra Nair, and Zizhou Vincent Wang, *An information inequality and evaluation of Marton's inner bound for binary input broadcast channels*, IEEE Transactions on Information Theory **59** (2013), no. 7, 4095–4105.

[GN20] Amin Gohari and Chandra Nair, *Outer bounds for multiuser settings: The auxiliary receiver approach*, submitted to the IEEE Transactions on Information Theory (2020).

[HP79] B Hajek and M Pursley, *Evaluation of an achievable rate region for the broadcast channel*, IEEE Trans. Info. Theory **IT-25** (January, 1979), 36–46.

[Kim07] Young-Han Kim, *Coding techniques for primitive relay channels*, Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput, 2007, p. 2007.

[Kim08] ______, *Capacity of a class of deterministic relay channels*, IEEE Transactions on Information Theory **54** (2008), no. 3, 1328–1329.

[Liu20] Jingbo Liu, *Minorization via mixed volumes and cover's problem for general channels*, arXiv preprint arXiv:2012.14521 (2020).

[LO19] Jingbo Liu and Ayfer Ozgur, *New converses for the relay channel via reverse hypercontractivity*, 2019 IEEE International Symposium on Information Theory (ISIT), July 2019, pp. 2878–2882.

[MHU19] Marco Mondelli, S Hamed Hassani, and Rüdiger Urbanke, *A new coding paradigm for the primitive relay channel*, Algorithms **12** (2019), no. 10, 218.

[Sim95] Stephen Simons, *Minimax theorems and their proofs*, Minimax and applications, Springer, 1995, pp. 1–23.

[TU08] Ravi Tandon and Sennur Ulukus, *A new upper bound on the capacity of a class of primitive relay channels*, 2008 46th Annual Allerton Conference on Communication, Control, and Computing, IEEE, 2008, pp. 1562–1569.

[WBÖ17] Xiugang Wu, Leighton Pate Barnes, and Ayfer Özgür, *The geometry of the relay channel*, 2017 IEEE International Symposium on Information Theory (ISIT), IEEE, 2017, pp. 2233–2237.

[WBO19] Xiugang Wu, Leighton Pate Barnes, and Ayfer Ozgur, *"the capacity of the relay channel": Solution to cover's problem in the gaussian case*, IEEE Transactions on Information Theory **65** (2019), no. 1, 255–275.

[WOX17] Xiugang Wu, Ayfer Ozgur, and Liang-Liang Xie, *Improving on the cut-set bound via geometric analysis of typical sets*, IEEE Transactions on Information Theory **63** (2017), no. 4, 2254–2277.