Relay Channel

Min Xie

April 15, 2003

1 Introduction

In a relay channel, between the sender X and the the receiver Y lies at least one relay. Generally, the relay can both transmit its own information and help forwarding other sources' information. This summary considers the latter case, i.e., the relay intends solely to *help* the receiver. The relay and the transmitter cooperate to resolve the receiver's uncertainty.
Due to the presence of relay, the relay channel capacity is difficult to determine. The capacity is known only for some special cases, e.g., physically degraded relay channel [1, 2], Gaussian relay channel [3, 4] (asymptotic capacity). This summary starts from the simplest relay channel with only one relay. An outer bound for the capacity of the general relay channel is briefly described in Section 2, but the focus is on a degraded relay channel, whose capacity and coding construction leading to the achievable rate will be given in Section 3. Section 4 presents a special degraded relay channel, Gaussian degraded relay channel. The scenario with multiple relays is more complicated and we discuss a simple Gaussian channel in the last section.

2 General Relay Channel

Fig. 1 illustrates the simplest general relay channel which has only one relay. The channel consists of four finite sets $\mathcal{X}$, $\mathcal{X}_1$, $\mathcal{Y}$, and $\mathcal{Y}_1$, and a collection of probability mass functions $p(y, y_1|x, x_1)$. x is the input to the channel, y is the output of the channel; y_1 is the relay's observation and x_1 is the input chosen by the relay and depends only on the past observation $(y_{11}, y_{12}, \ldots, y_{1i-1})$. The capacity problem is to find the channel capacity between X and Y.
An (M, n) code for the relay channel consists of a set of integers $\mathcal{M} = \{1, 2, \ldots, M\}$, an encoding

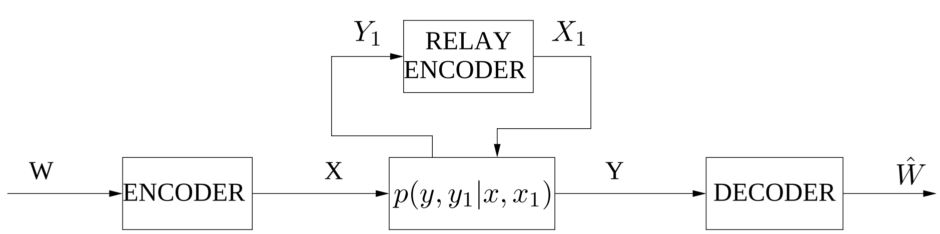


Figure 1: Relay Channel

function $X : \mathcal{M} \longrightarrow \mathcal{X}^n$, a set of relay functions $\{f_i\}_{i=1}^n$ such that

$$x_{1i} = f_i(Y_{11}, Y_{12}, \ldots, Y_{1i-1})$$

and a decoding function

$$g : \mathcal{Y}^n \longrightarrow \mathcal{M}.$$

The channel is memoryless in the sense that (Y_i, Y_{1i}) depends on the past only through the current transmitted symbols (X_i, X_{1i}). Thus, for any choice $p(w)$, $w \in \mathcal{M}$, code choice $x : \mathcal{M} \longrightarrow \mathcal{X}^n$ and relay functions $\{f_i\}_{i=1}^n$, the joint probability mass function on $\mathcal{M} \times \mathcal{X}^n \times \mathcal{X}_1^n \times \mathcal{Y}^n \times \mathcal{Y}_1^n$ is given by

$$p(w, \mathbf{x}, \mathbf{x_1}, \mathbf{y}, \mathbf{y_1}) = p(w) \prod_{i=1}^{n} p(x_i|w) p(x_{1i}|y_{11}, y_{12}, \ldots, y_{1i-1}) \cdot p(y_i, y_{1i}|x_i, x_{1i}). \quad (1)$$

The average probability of error is define as follows:

$$P_e^{(n)} = \frac{1}{2^{nR}} \sum_{w \in \mathcal{M}} \Pr\{g(\boldsymbol{Y}) \neq w | w \text{ sent}\} \triangleq \frac{1}{2^{nR}} \sum_{w \in \mathcal{M}} \lambda(w) \quad (2)$$

The relay channel combines a broadcast channel (X to Y and Y_1) and a multiple access channel (X_1 and X to Y). Directly applying the max-flow-min-cut theorem for general multiterminal networks to the relay channel, an upper bound of the capacity is obtained.

Theorem 1 *For any relay channel, the capacity is bounded above by*

$$C \leq \sup_{p(x,x_1)} \min\{I(X, X_1; Y), I(X; Y, Y_1|X_1)\} \quad (3)$$

The first term in (3) upper bounds the maximum rate of information transfer from senders X and X_1 to receiver Y (Multiple Access Channel); the second term bounds the rate from X to Y and Y_1 (Broadcast Channel, but the ultimate receiver Y should first decode the relay signal X_1 before decoding X, which contributes to the conditioning term X_1 in $I(X; Y, Y_1|X_1)$). The proof is referred to [1].

3 Degraded Relay Channel

The degraded relay channel, similar to the degraded broadcast channel, implies that one receiver is a degraded version of the other receiver. There are two degradednesses in the relay channel. What of interest is called *degraded relay channel*, in which the relay receiver y_1 is better than the ultimate receiver y and thus the relay can cooperate to send x. The other case, in which the relay y_1 is worse than y, is less interesting, because the relay can contribute no new information to the receiver.

3.1 Degraded Relay Channel

Definition 1 *The relay channel* $(\mathcal{X} \times \mathcal{X}_1, p(y, y_1|x, x_1), \mathcal{Y} \times \mathcal{Y}_1)$ *is said to be* degraded *if*

$$p(y, y_1|x, x_1) = p(y_1|x, x_1) p(y|y_1, x_1)$$

Equivalently, a relay channel is degraded if $p(y|y_1, x, x_1) = p(y|y_1, x_1)$, i.e., $X \to (X_1, Y_1) \to Y$ form a Markov chain. A degraded relay channel can be regarded as a family of *physically* degraded broadcast channels indexed by x_1.

Theorem 2 *The capacity C of the degraded relay channel is given by*

$$C = \sup_{p(x,x_1)} \min\{I(X, X_1; Y), I(X; Y_1|X_1)\} \tag{4}$$

where the supremum is over all joint distributions $p(x, x_1)$ on $\mathcal{X} \times \mathcal{X}_1$.

Here, due to the degradedness,

$$I(X; Y, Y_1|X_1) = I(X; Y_1|X_1). \tag{5}$$

Then the proof of the converse directly follows from Theorem 1. Next, we outline the proof of the achievability.

- **Random Coding:** First generate at random $M_0 = 2^{nR_0}$ i.i.d. n-sequences in $\mathcal{X}_1^n$, each drawn according to $p(\boldsymbol{x_1}) = \prod_{i=1}^n p(x_{1i})$, index them as $\boldsymbol{x}_1(s)$, $s \in [1, M_0]$. For each $\boldsymbol{x}_1(s)$, generate $M = 2^{nR}$ conditionally independent n-sequences $\boldsymbol{x}(w|s)$, $w \in [1, M]$ drawn according to $p(\boldsymbol{x}|\boldsymbol{x}_1(s)) = \prod_{i=1}^n p(x_i|x_{1i}(s))$. Thus we have a random code book $\mathcal{C} = \{\boldsymbol{x}(w|s), \boldsymbol{x}_1(s)\}$. Then randomly partition $\{1, 2, \ldots, 2^{nR}\}$ into 2^{nR_0} mutually exclusive subsets $\mathcal{S} = \{S_1, S_2, \ldots, S_{2^{nR_0}}\}$ such that each x message w has a corresponding x_1 index s to cell S_s. The relay, the receiver, and the transmitter agree on this partition.

- **Encoding:** At block i, let w_i be the new index to be sent and assume $w_{i-1} \in S_{s_i}$. The encoder then sends $\boldsymbol{x}(w_i|s_i)$. The relay estimates the previous w_{i-1} by $\hat{\hat{w}}_{i-1}$ and assumes that $\hat{\hat{w}}_{i-1} \in S_{\hat{\hat{s}}_i}$. Then the relay encoder sends $\boldsymbol{x}_1(\hat{\hat{s}}_i)$.

- **Decoding:** The decoding procedure at the end of block i is implemented according to the following steps.

  1. at the relay, upon estimating s_i by $\hat{\hat{s}}_i$ and receiving $\boldsymbol{y}_1(i)$, the relay claims that the message $\hat{\hat{w}}_i = w$ is sent iff there exists a *unique* w such that $(\boldsymbol{x}(w|\hat{\hat{s}}_i), \boldsymbol{x}_1(\hat{\hat{s}}_i), \boldsymbol{y}_1(i))$ is jointly ϵ-typical. $\hat{\hat{w}}_i = w$ with arbitrarily small probability of error if

  $$R < I(X; Y_1|X_1) \tag{6}$$

     and n is sufficiently large.

  2. at the receiver, upon receiving $\boldsymbol{y}(i)$, the receiver declares that $\hat{s}_i = s$ iff there exists a *unique* s such that $(\boldsymbol{x}_1(s), \boldsymbol{y}(i))$ is jointly ϵ-typical. $\hat{s}_i = s$ with arbitrarily small probability of error if

  $$R_0 < I(X_1; Y) \tag{7}$$

     and n is sufficiently large.

  3. at the receiver, the list code of the codeword w_{i-1} $\mathcal{L}(\boldsymbol{y}(i-1))$ is calculated to include all symbols w_{i-1} such that $(\boldsymbol{x}(w_{i-1}|\hat{s}_{i-1}), \boldsymbol{x}_1(\hat{s}_{i-1}), \boldsymbol{y}(i-1))$ are jointly ϵ-typical. Assuming that s_i is decoded successfully, $\hat{w}_{i-1} = w$ is declared iff there is a *unique* $w \in S_{s_i} \cap \mathcal{L}(\boldsymbol{y}(i-1))$. $\hat{w}_{i-1} = w$ with arbitrarily small probability of error if

  $$R < I(X; Y|X_1) + R_0 \tag{8}$$

Obviously, the receiver is always one block behind. In B blocks of transmission, a sequence of $B-1$ indices $\boldsymbol{W} = (W_1, W_2, \ldots, W_{B-1}, \emptyset)$ will be sent over the channel (the relay estimates $\hat{\hat{\boldsymbol{W}}} = (\hat{\hat{W}}_1, \hat{\hat{W}}_2, \ldots, \hat{\hat{W}}_{B-1}, \emptyset)$, and the receiver estimates $\hat{\boldsymbol{S}} = (\emptyset, \hat{S}_2, \hat{S}_3, \ldots, \hat{S}_B)$ and $\hat{\boldsymbol{W}} = (\emptyset, \hat{W}_1, \hat{W}_2, \ldots, \hat{W}_{B-1})$). The actual transmitting rate $R(B-1)/B$ is arbitrarily close to R as $B \to \infty$. Combining the encoding and decoding schemes and bounding the probability of error over the B blocks, we have

$$R < \min\{I(X; Y_1|X_1), I(X; Y|X_1) + I(X_1; Y)\} = \min\{I(X; Y_1|X_1), I(X, X_1; Y)\} \tag{9}$$

Note that all above contructions, including superposition coding, Slepian-Wolf partitioning, and coding for the cooperative MAC, can apply without change to arbitrary relay channels. However, the degradedness assumption is needed to establish that the achievable rate C is indeed the capacity.
The interpretation is as follows: given arbitrary $p(x, x_1)$, a rate $I(X, X_1; Y)$ can be achieved by complete cooperation of x and x_1. To set up this cooperation, x_1 must know x. Thus, the x rate of transmission should be less than $I(X; Y_1|X_1)$. Finally, both constraints lead to the minimum characterization.

3.2 Reversely Degraded Relay Channel

When the relay y_1 is worse than y, the channel is called *reversely degraded relay channel.*

Definition 2 *The relay channel is reversely degraded if $p(y, y_1|x, x_1)$ can be written in the form*

$$p(y, y_1|x, x_1) = p(y|x, x_1)p(y_1|y, x_1)$$

In this case, the relay cannot cooperate to send x, and thus just facilitates the transmission of x by sending the best x_1.

Theorem 3 *The capacity C_0 of the reversely degraded relay channel is given by*

$$C_0 = \max_{x_1 \in \mathcal{X}_1} \max_{p(x)} I(X; Y|x_1) \tag{10}$$

In other words, the relay y_1 sees a corrupted version of what y sees, then x_1 can contribute no new information to y. Thus x_1 is set constantly at the symbol that "opens" the channel for the transmission of x directly to y at rate $I(X; Y|X_1)$.

3.3 General Relay Channel with Feedback

In the general relay channel with feedback, although no degradedness relation between y and y_1 is assumed, the feedback nature can convert it to an ordinary degraded relay channel with the substitution of (Y, Y_1) for Y_1. Intuitively, the relay knows the y sequence through the feedback link, and the y sequence is an implication of what the transmitter x intends to send. Therefore, the relay has much better reliability of decoding x correctly than the ultimate receiver y. In this sense, y is a degraded version of the relay y_1 with feedback y. Like in the MAC with feedback, the relay unambiguously estimates x message index w, and then helps to resolve the uncertainty of the receiver in the subsequent block.

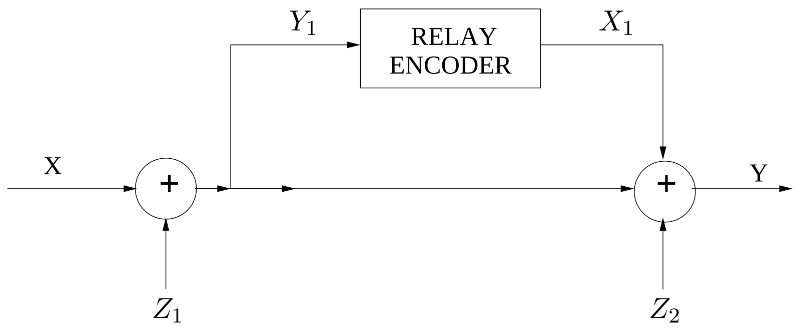


Figure 2: Gaussian Degraded Relay Channel

Theorem 4 *The capacity C_{FB} of an arbitrary relay channel with feedback is given by*

$$C_{FB} = \sup_{p(x,x_1)} \min\{I(X, X_1; Y), I(X; Y, Y_1|X_1)\} \tag{11}$$

Particularly, if the channel is degraded or reversely degraded, then feedback does not increase the capacity.

4 Gaussian Degraded Relay Channel

Consider a Gaussian degraded relay channel, shown in Fig.2, where $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are sequences of i.i.d. normal random variables with zero mean and variance N_1 and N_2, respectively. The ultimate receiver Y is a corrupted version of the relay Y_1, conditioning on X_1.

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y &= Y_1 + X_1 + Z_2 \end{aligned} \tag{12}$$

In addition, the transmitted power is constrained by

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} x_i^2(w) &\leq P \ , \ w \in \{1, 2, \ldots, M\} \\ \frac{1}{n}\sum_{i=1}^{n} x_{1i}^2(y_{11}, y_{12}, \ldots, y_{1i-1}) &\leq P_1 \ , \ (y_{11}, y_{12}, \ldots, y_{1n}) \in \mathbb{R}^n \end{aligned}$$

Theorem 5 *The capacity C^* of the Gaussian degraded relay channel is given by*

$$C^* = \max_{0\leq\alpha\leq 1} \min\left\{C\left(\frac{P + P_1 + 2\sqrt{\bar{\alpha}PP_1}}{N_1 + N_2}\right), C\left(\frac{\alpha P}{N_1}\right)\right\} \tag{13}$$

where $\bar{\alpha} = 1 - \alpha$ and $C(x) = \frac{1}{2}\log(1 + x)$.

We just sketch the random code that achieves C^*. For $0 \leq \alpha \leq 1$, let $X_1 \sim N(0, P_1)$, $\tilde{X} \sim N(0, \alpha P)$, with $\tilde{X}$, X_1 independent. The transmitter x generates the first codebook $\mathcal{M} = \{1, 2, \ldots, 2^{nR}\}$

according to $N_n(0, \alpha P)$, i.e., given a message index $w \in \mathcal{M}$, the transmitter sends the codeword $\tilde{\boldsymbol{x}}(w)$ with power αP. Given $R < C(\frac{\alpha P}{N_1})$, the relay knows w correctly, but the receiver (without the relay) has ambiguity because its capacity $C(\frac{\alpha P}{N_1+N_2}) < C(\frac{\alpha P}{N_1})$. It has a list of possible words of size $2^{n[R-C(\alpha P/(N_1+N_2))]}$.

In the next block, the transmitter and the relay cooperate to resolve the receiver's uncertainty. As mentioned in Section 3.1, they partition the first codebook $\mathcal{M}$ into 2^{nR_0} mutually exclusive subsets, which is exactly the second codebook with 2^{nR_0} codewords. They coherently transmit the partition index s through the codeword $\boldsymbol{x_1}(s)$. Specifically, the transmitter sends s with power $\bar{\alpha}P$, and the relay sends with power P_1. The power seen by the receiver y is therefore $(\sqrt{\bar{\alpha}P} + \sqrt{P_1})^2$, which leads to

$$R_0 < C\left(\frac{(\sqrt{\bar{\alpha}P} + \sqrt{P_1})^2}{\alpha P + N_1 + N_2}\right).$$

In addition, the transmitter also chooses a fresh codeword from the first codebook, adds it "on paper" to the cooperative codeword from the second codebook, and sends the sum over the channel. In summary,

$$\begin{aligned} \boldsymbol{x}(w|s) &= \tilde{\boldsymbol{x}}(w) + \sqrt{\frac{\bar{\alpha}P}{P_1}}\boldsymbol{x_1}(s) \ , \ w \in [1, 2^{nR}] \\ \boldsymbol{x_1}(s) & \quad\quad s \in [1, 2^{nR_0}] \end{aligned} \tag{14}$$

which results in a steady-state resolution of the past uncertainty and infusion of new information.

Remarks 1 *1. If $P_1/N_2 \geq P/N_1$, then $I(X, X_1; Y) \geq I(X; Y_1|X_1)$. The relay can forward the cooperative information s to the receiver without error. The capacity $C^* = C(P/N_1)$ is achieved when $\alpha = 1$, which implies that the transmitter does not need to allocate power to send the partition index s. The channel appears to be noise-free after the relay, and the capacity $C(P/N_1)$ from x to the relay can be achieved. Thus the rate without the relay $C(P/(N_1 + N_2))$ is increased by the relay to $C(P/N_1)$.*

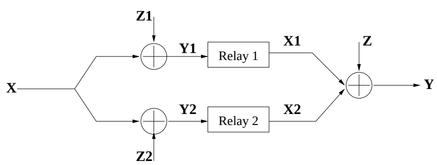
2. If $P_1/N_2 < P/N_1$, then $I(X, X_1; Y) < I(X; Y_1|X_1)$. The relay cannot guarantee perfect transmission of the cooperative information s, then the transmitter must cooperate to send s. Clearly the maximizing $\alpha = \alpha^*$ is strictly less than one, and is given by solving α in

$$\frac{1}{2}\ln\left(1 + \frac{P + P_1 + 2\sqrt{\bar{\alpha}PP_1}}{N_1 + N_2}\right) = \frac{1}{2}\ln\left(1 + \frac{\alpha P}{N_1}\right).$$

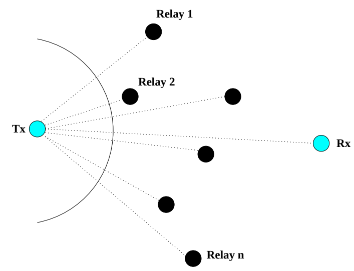
The capacity region of the degraded Gaussian relay channel with multiple relays can be obtained by building an inductive argument based on the single-relay capacity Theorem 5. The details are given in [5].

5 Gaussian Relay Network

In the Gaussian relay channel, when the channel is not degraded, or there exist more than one relays [3, 6], the capacity region is unknown yet, i.e., the upper bound and the lower bound are not coincident. [3] presented an asymptotic capacity in the general Gaussian relay network with



(a) General Gaussian Relay Network with Two Relays



(b) The Broadcast Cut in the general Gaussian Relay Network

Figure 3: General Guassian Relay Network

multiple relays. The "asymptotic" indicates that the upper bound and the lower bound meet asymptotically as the number of relays goes to infinity (see Fig.3).
The difference of Fig. 3(a) from previous schemes is that each relay acts as a simple transponder, amplifying both noise Z_i and signal X. The upper bound is derived from the max-flow-min-cut theorem (Fig.3(b)). The lower bound follows from a consideration of almost uncoded transmission of a particular source across the Gaussian relay channel. Two additional assumptions are required. First, there is a "dead zone" of nonzero radius around the source and the destination node. Second, the source node may only send half of the time.

Theorem 6 (upper bound) *For any particular realization of the random geometry of the network, the capacity of the considered relay network is upper bounded by*

$$C \leq C_{upper} = \frac{1}{4} \log_2 \big(1 + \frac{\|\boldsymbol{\beta}\|^2 P}{N}\big) \tag{15}$$

where $\boldsymbol{\beta}$ is a vector accounting for the fading process, and P is the power constraint for the source signal.

Theorem 7 (lower bound) *For any particular realization of the random geometry of the network, the capacity of the considered relay network is at least*

$$C \geq C_{lower} = \frac{1}{4} \log_2 \frac{P}{D_1} \tag{16}$$

where D_1 is the mean-square error of the decoded signal.

It is proved that the capacity C lies between C_{upper} and C_{lower}. Asymptotically, the capacity of the considered relay network is

$$C_\infty = \frac{1}{4} \log_2 \big(1 + \frac{\|\beta\|^2 P}{N}\big)$$

References

[1] T. M. Cover and A. A. E. Gamal, "Capacity Theorem for the Relay Channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 572–584, Sept. 1979.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, John and Sons, Incorporated, 1991. ISBN 0471062596.

[3] M. Gastpar and M. Vetterli, "On the Capacity of Wireless Networks: the Relay Case," in *IEEE INFOCOM*, vol. 3, pp. 1577–1586, 2002.

[4] B. Schein and R. Gallager, "The Gaussian Parallel Relay Network," in *ISIT 2000*, June 2000.

[5] A. Reznik, S. R. Kulkarni, and S. Verdú, "Capacity and Optimal Resource Allocation in the Degraded Gaussian Relay Channel with Multiple Relays," in *40th Allertion Conference on Communication, Control and Computing*, Oct. 2002.

[6] P. Gupta and P. R. Kumar, "Towards an Information Theory of Large Networks: An Achievable Rate Region," in *IEEE ISIT*, (Washington DC), p. 159, 2001.