Relay Channels

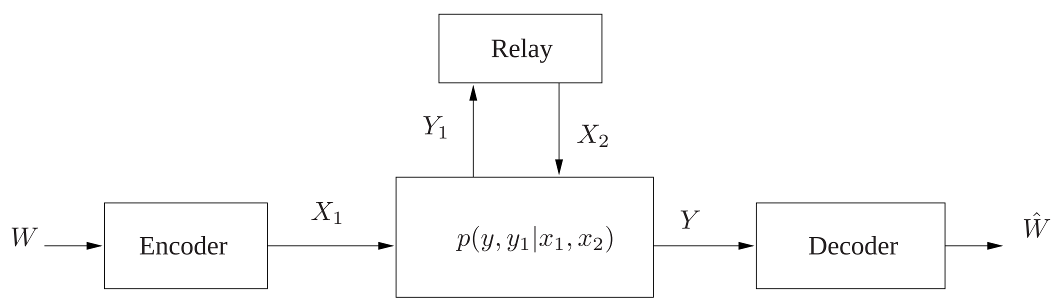
Radha Krishna Ganti

1 Introduction

Relay channel was firs introduced by van der Meulen [1, 2, 3]. Time sharing approaches were used to find inner bounds for the capacity. Most of the known results stem from the publication [4] by Dr. Thomas M. Cover and Dr. Abbas A. El Gammal. In this paper, capacity of degraded relay channels, relay channels with feedback is computed. Also lower bounds using Block Markov coding (decode and forward) and side-information coding (compress-and-forward) were calculated. A new lower bound was proposed in [5] using partial-decode and forward. More history about the relay channel is given in the introduction section of the doctoral thesis of Sina Zahedi [6].

In this report, the bounds on the capacity results [4] for the relay channel are studied. These results are in most cases the best bounds and the only known capacities of some classes of channels. Most of the material in this report is taken from [4, 7, 6]. The thesis of Dr. Sina Zahedi and the presentation slides [8] of Dr. Abbas El Gammal in MSRI 2006 are a good reference.

2 Channel Model



A discrete memoryless relay channel denoted by $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$, consists of the four finite sets, $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$ and a collection of probability distributions $P(., .|x_1, x_2)$ on $\mathcal{Y}_1 \times \mathcal{Y}_2$, one for each (x_1, x_2). Each x_1 is the input from the encoder, and x_2 is the input from the relay. y_1 is the output of the channel to the relay, and y is the output to the decoder. An (M, n) code for the relay channel consists of a set of integers

$$\mathcal{M} = \{1, 2, \ldots, M\} \quad (1)$$

an encoding function

$$x_1 : \mathcal{M} \to \mathcal{X}_1^n \tag{2}$$

a set of relay functions $\{f_i\}_{i=1}^n$ such that

$$x_{2i} = f_i(Y_{11}, Y_{12}, \ldots, Y_{1i-1}), \;\; 1 \le i \le n \tag{3}$$

and a decoding function

$$d : \mathcal{Y}^n \to \mathcal{M} \tag{4}$$

Observe that the relay function is casual, *i.e.*, depends only on the previous outputs, up to including $i-1$. The rate R of the (M, n) code is defined by

$$R = \frac{1}{n} \log M \quad \text{bits/transmission} \tag{5}$$

There exists a variant of the relay, called as Relay Without Delay (RWD) which is non causal in the sense that, the relay can instantaneously use the information to transmit, *i.e.*

$$x_{2i} = f_i(Y_{11}, Y_{12}, \ldots, Y_{1i-1}, y_{1,i}), \;\; 1 \le i \le n$$

The relay channel we are considering is also called the *classical relay*. A more general form of the relay can be introduced by the concept of Sleep-Listen-or-Talk **SloT** introduced in [9].

Definition The relay channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is said to be **degraded** (or rather Y is a degraded form of Y_1) if $p(y, y_1|x_1, x_2)$ can be written in the form

$$p(y, y_1|x_1, x_2) = p(y_1|x_1, x_2)p(y|y_1, x_2) \tag{6}$$

i.e., $X_1 \to (X_2, Y_1) \to Y$ form a Markov chain.

Definition The relay channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is said to be **reversely degraded** (or rather Y_1 is a degraded form of Y) if $p(y, y_1|x_1, x_2)$ can be written in the form

$$p(y, y_1|x_1, x_2) = p(y|x_1, x_2)p(y_1|y, x_2) \tag{7}$$

i.e., $X_1 \to (X_2, Y) \to Y_1$ form a Markov chain.

The following lemma [4] will be used to bound the error events in the subsequent proofs. Let S_1', S_2' be conditionally independent given S_3, with the marginals

$$p(s_1|s_3) = \sum_{s_2} p(s_1, s_2, s_3)/p(s_3)$$
$$p(s_2|s_3) = \sum_{s_1} p(s_1, s_2, s_3)/p(s_3)$$

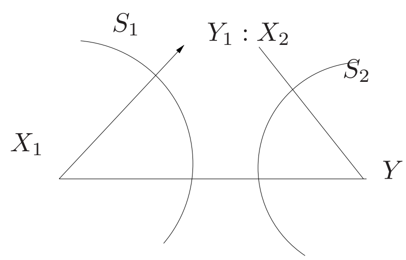
Lemma 2.1. *Let* $(S_1, S_2, S_3) \sim \prod_{i=1}^{n} p(s_1, s_2, s_3)$ *and* $(S_1', S_2', S_3) \sim \prod_{i=1}^{n} p(s_{1i}|s_{3i})p(s_{2i}|s_{3i})p(s_{3i})$. *Then, for* n *such that* $P(A_\epsilon(S_1, S_2, S_3)) \geq 1 - \epsilon$,

$$(1-\epsilon)2^{-n(I(S_1;S_2|S_3)+7\epsilon)} \leq P\{(S_1', S_2', S_3) \in A_\epsilon(S_1, S_2, S_3)\} \leq 2^{-n(I(S_1;S_2|S_3)-7\epsilon)}$$

Theorem 2.2. *For a general relay channel*

$$C \leq \max_{p(x_1,x_2)} \min\{I(X_1, X_2; Y), I(X_1; Y, Y_1|X_2)\} \tag{8}$$

Proof. This bound can be obtained by applying the cutset bound [10] [7]. Consider the following cuts S_1 and S_2



For the cut S_1 we have,

$$R_{X_1Y} + R_{X_1Y_1} \leq I(X_1; Y_1, Y|X_2) \tag{9}$$

For the cut S_2 we have

$$R_{X_1Y} + R_{X_2Y} \leq I(X_1, X_2; Y) \tag{10}$$

Observing that $R_{X_2Y}, R_{X_1Y_1} \geq 0$ and combining both the above inequalities, we have the required result. For an alternate proof using Fano's inequality (was developed before the cutset bound) see [4]. This upper bound is tight in all the cases in which capacity is known. □

Theorem 2.3. *For any relay channel [4, 6]*

$$C \geq \sup_{p(x_1,x_2)} \min\{I(X_1, X_2; Y), I(X_1; Y_1|X_2)\} \tag{11}$$

Proof. The basic idea of proving achievability is as follows

- Encoding consists of "Random coding, Block Markov encoding and Binning"
- Relay **decodes the complete message** and sends the bin number of previous block message
- Decoder decodes the present message and removes the ambiguity about the previous message by using the bin number send by the relay.

Codebook Generation: We generate a sequence of codebooks $\mathcal{C}_n$

$$\mathcal{C}_n = \{(x_1^n(w|s), x_2^n(s)) : w \in \{\phi, 1, 2, \ldots, 2^{\lceil nR \rceil}\}, s \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}\} \tag{12}$$

such that $P_e^n \to 0$ as $n \to \infty$ if equation 11 is satisfied. Let the random codebook be generated as follows:

- For $s \in \{\phi, 1, 2, \ldots, 2^{nR_0}\}$, generate an i.i.d sequence $x_2^n(s)$, with $x_2^i(s) \sim p(x_2), \forall s, i \in [1, n]$.
- for each s and $w \in \{\phi, 1, 2, \ldots, 2^{nR}\}$, generate $x_1^n(w|s)$conditionally independent from $x_2^n(s)$, with distribution $p(x_{1j}|x_{2j}(s))$

Random Binning: For each message $\{1, 2, \ldots, 2^{\lceil nR \rceil}\}$, assign an index at random from $\{1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}$. Denote this assignment (now no longer random once fixed) by a *mapping* $g : \{1, 2, \ldots, 2^{\lceil nR \rceil}\} \to \{1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}$. Also let $g^{-1}(y) = \{x : g(x) = y\}$ denote the inverse mapping . Also $\hat{a}$ denotes the estimate of a and belongs to the same space to which a belongs.[1]

[1]The sequence of messages transmitted are as follows $(W_1, \ldots, W_{B-1}, \phi)$. The relay estimates are $(\hat{W}_1, \ldots, \hat{W}_{B-1}, \phi)$. The receiver estimates $(\phi, \hat{W}_1, \ldots, \hat{W}_{B-1})$. The first message send by the encoder is $x_1(w_1|\phi)$.

	Block i, Message w_i	Typicality Decoding	Rate required for correct decoding
Encoder :			
Transmits	$x_1^n(w_i \| g(w_{i-1}))$		
Relay :			
Already Knows	$(\hat{w}_1, \ldots, \hat{w}_{i-1})$		
Receives	$y_1^n(i)$		
Transmits	$x_2^n(g(\hat{w}_{i-1}))$		
Decodes	$\hat{w}_i$	Declares $\hat{w}_i$ is sent if $(x_1^n(w_i \| g(\hat{w}_{i-1}), x_2^n(g(\hat{w}_{i-1})), y_1^n(i)) \in A_\epsilon^n(X_1, X_2, Y_1)$	$R \leq I(X_1; Y_1 \| X_2)$
Decoder :			
Already Knows	$(\hat{w}_1, \ldots, \hat{w}_{i-2})$, $L(y^n(i-1))$		
Receives	$y^n(i)$		
Decodes	$\hat{g}(w_{i-1})$	Declares $\hat{g}(w_{i-1})$ as the bin number in $i-1$ block if $(x_2^n(g(w_{i-1})), y^n(i)) \in A_\epsilon^n$	$R_0 \leq I(X_2; Y)$
	$\hat{w}_{i-1}$	Declares $\hat{w}_{i-1}$ is sent in block $i-1$ if it is the unique index in $\{g^{-1}(\hat{g}(w_{i-1}))\} \cap L(y^n(i-1))$	$R \leq I(X_1; Y \| X_2) + R_0$
	The ambiguity set $L(y^n(i))$	Set of all w_i such that $(x_1^n(w_i \| g(\hat{w}_{i-1})), x_2^n(g(\hat{w}_{i-1})), y(i)) \in A_\epsilon^n$	

$\square$

Theorem 2.4. *If Y is a degraded form of Y_1, then*

$$C = \max_{p(x_1,x_2)} \min\{I(X_1, X_2; Y), I(X_1; Y_1|X_2)\} \tag{13}$$

Proof. From the definition of degradedness, we have $I(X_1; Y, Y_1|X_2) = I(X_1; Y_1|X_2)$. Hence the cutset bound simplifies to

$$C \leq \max_{p(x_1,x_2)} \min\{I(X_1, X_2; Y), I(X_1; Y_1|X_2)\}$$

Also from general block Markov encoding with complete decoding, i.e theorem 2.3, we have the reverse inequality. $\square$

Theorem 2.5. *If $p(y, y_1|x_1, x_2)$ is an arbitrary relay channel with feedback from (y, y_1) to both x_1 and x_2, then*

$$C = \max_{p(x_1,x_2)} \min\{I(X_1, X_2; Y), I(X_1; Y, Y_1|X_2)\} \tag{14}$$

Proof. Feedback changes an arbitrary relay channel to a degraded channel. Clearly Y is a degraded form of (Y, Y_1) □

Theorem 2.6. *Let $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ be any discrete memoryless relay channel. Then the capacity C [5, 4, 6] is lower bounded by*

$$C \geq \sup_{p(u,x_1,x_2)} \{\min\{I(X_1, X_2; Y), I(U; Y_1|X_1) + I(X; Y|X_1, U)\}\} \tag{15}$$

where the supremum is taken over all joint distributions $p(u, x_1, x_2)$ on $\mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2$ such that $U \to (X_1, X_2) \to (Y, Y_1)$ form a Markov chain.

Proof. This method is also called as *Partial decode and forward*. The basic idea of achievability is as follows.

- Encoding consists of "Random coding, Block Markov encoding and Binning". Message is divided into two independent parts *i.e.*$w = (w', w'')$. Do the encoding and decoding of w' by using the same procedure as in previous theorem with the help of the relay. w'' is decoded straight by the decoder.
- Only the **first part of the message w' is decoded by the relay**.
- Decoder takes the help of the relay to decode the first part w' (remove the ambiguity, using the bin number send by the relay) and decodes the second part of the message w'' directly.

Codebook Generation: Let $R = R_1 + R_2$. We generate a sequence of codebooks $\mathcal{C}_n$

$$\mathcal{C}_n = \{(u^n(w'|s), x_1^n(w''|w', s), x_2^n(s)) : w' \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$$
$$, w'' \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_2 \rceil}\}, s \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}\}$$

such that $P_e^n \to 0$ as $n \to \infty$. The random codebook $\mathcal{C}_n$ is generated as follows

- For $s \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}$, generate an i.i.d sequence $x_2^n(s)$ according to $p(x_2^n)$..
- For each s and $w' \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$, generate an i.i.d sequences $u^n(w'|s)$ according to $p(u^n|x_2^n(s))$. So u is the cooperative information.
- For all pairs (w', s) and for each $w'' \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_2 \rceil}\}$ randomly generate an i.i.d. sequence $x_1^n(w''|w', s)$ according to probability mass function $p(x_1^n|u^n(w'|s), x_2^n(s))$.

Random Binning: For each message $\{1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$, assign an index at random from $\{1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}$. Denote this assignment (now no longer random once fixed) by a mapping $g : \{1, 2, \ldots, 2^{\lceil nR_1 \rceil}\} \to \{1, 2, \ldots, 2^{\lceil nR_0 \rceil}\}$. Also let $g^{-1}(y) = \{x : g(x) = y\}$ denote the inverse mapping.

	Block i, Message $w_i = (w_i', w_i'')$	Typicality Decoding	Rate required for correct decoding
Encoder :			
Transmits	$x_1^n(w_i''\|g(w_{i-1}'), w_i')$		
Relay :			
Already Knows	$(\hat{w}_1', \ldots, \hat{w}_{i-1}')$		
Receives	$y_1^n(i)$		
Transmits	$x_2^n(g(\hat{w}_{i-1}'))$		
Decodes	$\hat{w}_i'$	Declares $\hat{w}_i' = w_i'$ is sent if it is the unique index such that $(u^n(w_i'\|g(\hat{w}_{i-1}'), x_2^n(g(\hat{w}_{i-1}')), y_1^n(i)) \in A_\epsilon^n(X_1, X_2, Y_1)$	$R \leq I(U; Y_1\|X_2)$
Decoder :			
Already Knows	$(\hat{w}_1, \ldots, \hat{w}_{i-2})$, $L(y^n(i-1))$		
Receives	$y^n(i)$		
Decodes	$\hat{g}(w_{i-1}')$	Declares $\hat{g}(w_{i-1}')$ as the bin number in $i-1$ block if $(x_2^n(g(w_{i-1}')), y^n(i)) \in A_\epsilon^n$	$R_0 \leq I(X_2; Y)$
	$\hat{w}_{i-1}'$	Declares $\hat{w}_{i-1}'$ is sent in block $i-1$ if it is the unique index in $\{g^{-1}(\hat{g}(w_{i-1}'))\} \cap L(y^n(i-1))$	$R_1 \leq I(X_1; Y\|X_2) + R_0$
	$\hat{w}_{i-1}''$	Declares $\hat{w}_{i-1}'' = w_{i-1}''$ is sent in block $i-1$ if it is the unique index such that $(*)$ holds (Require Markovian property here)	$R_2 \leq I(X_1; Y\|U, X_2)$
	The ambiguity set $L(y^n(i))$	Set of all w_i' such that $(u^n(w_i'\|g(\hat{w}_{i-1}')), x_2^n(g(\hat{w}_{i-1}')), y(i)) \in A_\epsilon^n$	

$(*) \equiv (x_1^n(w''_{i-1}|\hat{w}'_{i-1}, \hat{g}(w'_{i-1})), u^n(\hat{w}'_{i-1}|\hat{g}(w'_{i-1})), x_2^n(\hat{g}(w'_{i-1})), y^n(i-1)) \in A_\epsilon^n$. Colored regions of the table denote the encoding and decoding operations different to the previous theorem. Combining the required rates and using the Markov property we get the required inequality. □

Theorem 2.7. *Let $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ be any discrete memoryless relay channel. Then the rate R_1^* is achievable [6, 4], where*

$$R_1^* = \sup_{p(x_1)p(x_2)p(\hat{y}_1|y_1,x_2)} I(X_1; Y, \hat{Y}_1|X_2) \tag{16}$$

subject to the constraint

$$I(X_2; Y) \geq I(Y_1; \hat{Y}_1|X_2, Y) \ \ (**) \tag{17}$$

Proof. This method is also called *compress and Forward* method.

- Use block Markov coding.
- Relay "quantizes"' its received sequence in previous block and sends it to the receiver. The receiver considers the information received directly form the encoder as side information and decodes the quantized information received from the relay (Wyner Ziv). It then uses this decoded information to decode the directly received information from the encoder.
- Relay and sender do not cooperate. Independent codebooks are used by the sender and relay.

Codebook Generation: We generate a sequence of codebooks $\mathcal{C}_n$

$$\mathcal{C}_n = \{(x_1^n(w|s), x_2^n(s), y_1^n(z|s)) : w \in \{\phi, 1, 2, \ldots, 2^{\lceil nR \rceil}\}, \tag{18}$$

$$s \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_1 \rceil}, z \in \{\phi, 1, 2, \ldots, 2^{\lceil n\hat{R} \rceil}\}\} \tag{19}$$

such that $P_e^n \to 0$ as $n \to \infty$ if equations 16 and 17 are satisfied. For each fixed joint probability mass function of the form $p(x_1)p(x_2)p(\hat{y}_1|y_1, x_2)$, the random codebook is generated as follows

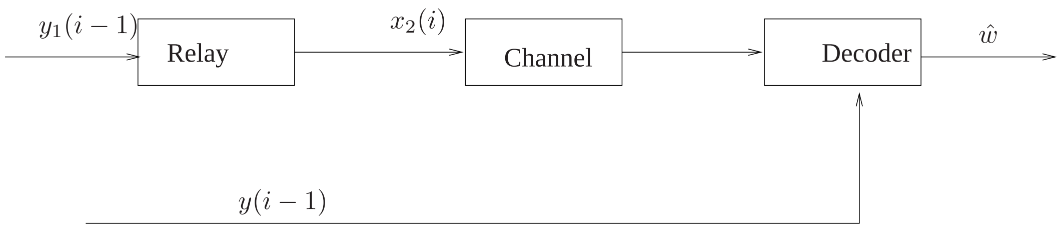
- For $w \in \{\phi, 1, 2, \ldots, 2^{\lceil nR \rceil}\}$, generate an i.i.d sequence $x_1^n(s)$ according to $p(x_1^n)$..
- For $s \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$, generate an i.i.d sequence $x_2^n(s)$ according to $p(x_2^n)$..
- For each $s' \in \{\phi, 1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$ and for each index $z \in \{\phi, 1, 2, \ldots, 2^{\lceil n\hat{R} \rceil}\}$ generate an i.i.d sequences $\hat{y}_1^n(z|s)$ according to $p(\hat{y}_1^n|x_2^n(s))$, where $p(\hat{y}_1|x_2(s)) = \sum_{y_1 \in \mathcal{Y}_1} p(y_1|x_2)p(\hat{y}_1|x_2(s), y_1)$.

Random Binning: For each message $\{1, 2, \ldots, 2^{\lceil n\hat{R} \rceil}\}$, assign an index at random from $\{1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$. Denote this assignment (now no longer random once fixed) by a mapping $g : \{1, 2, \ldots, 2^{\lceil n\hat{R} \rceil}\} \to \{1, 2, \ldots, 2^{\lceil nR_1 \rceil}\}$. Also let $g^{-1}(y) = \{x : g(x) = y\}$ denote the inverse mapping.

	Block i, Message w_i	Typicality Decoding	Rate required for correct decoding
Encoder :			
Transmits	$x_1^n(w_i)$		
Relay :			
Already Knows	$(z_1, z_2, \ldots, z_{i-1})$		
Receives	$y_1^n(i)$		
Transmits	$x_2^n(g(z_{i-1})) \doteq (*)$		
Decodes	z_i	$z_i = z$ is estimated $[\hat{y}_1^n(z\|g(z_{i-1})), y_1^n(i), x_2^n(g(z_{i-1}))] \in A_\epsilon^n$	$\hat{R} \geq I(Y_1; \hat{Y}_1\|X_2)$
Decoder :			
Already Knows	$(\hat{w}_1, \ldots, \hat{w}_{i-2})$ $(z_1, \ldots, z_{i-2})$		
Receives	$y^n(i)$		
Decodes	$\hat{g}(z_{i-1})$ (Removing the channel effects from the quantized relay transmission)	Declares $\hat{g}(z_{i-1})$ as the quantized estimate send by the relay in i block if $(x_2^n(g(\hat{z}_{i-1})), y^n(i)) \in A_\epsilon^n$	$R_1 \leq I(X_2; Y)$
	The ambiguity set $L(y^n(i-1))$	Set of all z such that $[\hat{y}_1^n(z\|\hat{g}(z_{i-1})), x_2^n(\hat{g}(z_{i-1})), y^n(i-1)] \in A_\epsilon^n$	
	$\hat{z}_{i-1}$ (unquantizing the quantized value from the relay)	Receiver declares that it is $\hat{z}_{i-1}$, that was quantized and send by the relay if it is the unique index such that $\hat{z}_{i-1} \in \{g^{-1}(\hat{g}(z_{i-1}))\} \cap L(y^n(i-1))$	$\hat{R} \leq I(\hat{Y}_1; Y\|X_1) + R_1$
	$\hat{w}_{i-1}$	Receiver declares that $\hat{w}_{i-1}$ is the message send in block $i-1$, if it is the unique message such that $[x_1^n(\hat{w}_{i-1}), y_1^n(\hat{z}_{i-1}\|\hat{g}(z_{i-1})), y^n(i-1), x_2^n(\hat{g}(z_{i-1}))] \in A_\epsilon^n$	$R \leq I(X_1; Y, \hat{Y}_1\|X_2)$

$(*)$: Quantization and encoding at the relay. The colored steps are the only **non** - Wyner-Ziv steps. Combining the rate constraints on $\hat{R}$ and R_1 we get the required constraint. □

Proof using block Markov and Wyner-Ziv coding: For further information regarding Wyner-Ziv see the report by Shiva Kotagiri [11] and the references within.



The relay at block i sends information about message $i-1$. Hence the information that the relay wants to send at block i is correlated to $y(i-1)$. Now the problem translates to the following. Suppose we are able to correctly decode the information send by the receiver at the decoder (i.e mitigate the channel or give the copy of x_2 to the decoder or in probabilistic sense condition on the event of having x_2) , what is the minimum rate with which the relay can encode the information. So we have the following Wyner-Ziv problem. The decoder has side information $y(i-1)$. The relay wants to encode $y_1(i-1)$ which is correlated with $y(i-1)$ (all of these under the condition given x_2). What is the minimum rate of quantization? Also we don't bother about the distortion. (Having some extra information at the receiver is better than having nothing). So from the result of Wyner-Ziv we have that the quantization rate R_1 should be

$$R_1 \geq I(\hat{Y}_1; Y_1|X_2) - I(\hat{Y}_1; Y|X_2)$$

where $Y \to Y_1 \to \hat{Y}$. We want to transmit this information across the channel $X_2 : Y$ and decode it correctly. This would require

$$\begin{aligned}
I(X_2; Y) &\geq R_1 \\
&\geq I(\hat{Y}_1; Y_1|X_2) - I(\hat{Y}_1; Y|X_2). \\
&= H(\hat{Y}_1|X_2) - H(\hat{Y}_1|X_2, Y_1) - H(\hat{Y}_1|X_2) + H(\hat{Y}_1|X_2, Y) \\
&= H(\hat{Y}_1|X_2, Y) - H(\hat{Y}_1|X_2, Y_1) \\
&= H(\hat{Y}_1|X_2, Y) - H(\hat{Y}_1|X_2, Y_1, Y) \\
&= I(\hat{Y}_1; Y_1|X_2, Y)
\end{aligned}$$

After having the information about $\hat{Y}_1(i-1)$ the decoder uses joint information $(Y, \hat{Y}_1)$ to decode the message. Hence the rate $R \leq I(X_1; Y, \hat{Y}_1|X_2)$.

References

[1] E. Van der Meulen, “Three-terminal communication channels,” *Advances in Applied Probability*, vol. 3, pp. 120–154, 1971.

[2] ——, “Transmission of Information in a T-terminal Discrete Memoryless Channel.” Ph.D. dissertation, University of California at Berkeley., 1969.

[3] ——, “A survey of multi-way channels in information theory: 1961-1976,” *IEEE Transactions on Information Theory*, vol. 23, no. 1, pp. 1–37, 1977.

[4] T. Cover and A. Gamal, “Capacity theorems for the relay channel,” *Information Theory, IEEE Transactions on*, vol. 25, no. 5, pp. 572–584, 1979.

[5] M. Aref, “The capacity of semideterministic relay channels,” *IEEE Transactions on Information Theory*, vol. 28, may 1982.

[6] S. Zahedi, “On reliable communication over the relay channel,” Ph.D. dissertation, Stanford University, 2005.

[7] J. Thomas and T. Cover, *Elements of Information Theory*. Wiley-Interscience, 1991.

[8] A. E. Gammal, “Capacity theorems for relay channels,” MSRI, Tech. Rep., 2006.

[9] G. Kramer, “Models and theory for relay channels with receive constraints,” in *Allerton*, Sept. 2004.

[10] L. Xiao, “Cut-set bounds, dependence-balance bounds,” University of Notre Dame, http://www.nd.edu/ jnl/ee87005/materials/summaries/lei.pdf, Tech. Rep., 2006.

[11] S. P. Kotagiri, “Source coding with decoder side information,” University of Notre Dame, http://www.nd.edu/ jnl/ee87005/materials/summaries/shiva.pdf, Tech. Rep., 2006.