Note: Correction to page 308

X Lecture 13, Nov. 24, 2003

<u>Theorem</u>: The capacity region of a multiple access channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y})$ is the closure of the <u>convex hull</u> of the set of all rate pairs (R_1, R_2) satisfying:

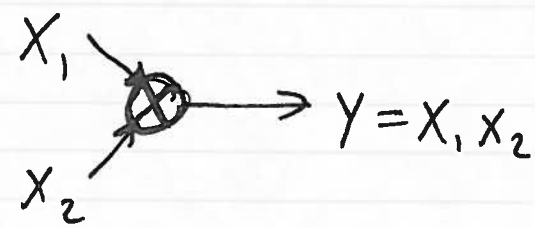
$$R_1 < I(X_1; Y | X_2),$$

$$R_2 < I(X_2; Y | X_1)$$

$$R_1 + R_2 < I(X_1, X_2; Y)$$

for some product distribution $p(x_1)p(x_2)$ on $\mathcal{X}_1 \times \mathcal{X}_2$.

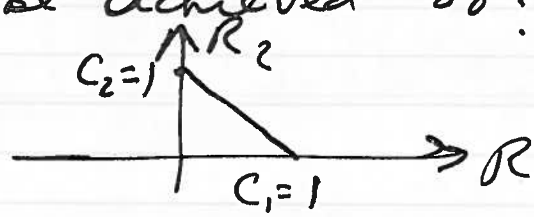
<u>Example</u>: Binary multiplier Channel



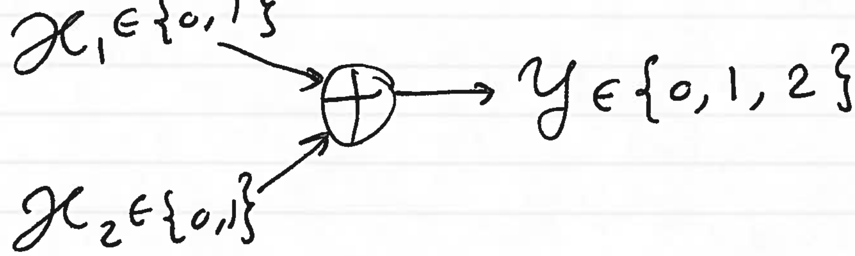
let $X_1 =$ ~~0 (or~~ 1) then $R_1 = 0$ and $R_2 = 1$.

Similarly ~~let~~ fix $X_2 =$ ~~0 (or~~ 1) then $R_2 = 0$, $R_1 = 1$

Any point on the line connecting the points $(0,1)$ and $(1,0)$ can be achieved so:



Example: Binary multiple access erasure channel:

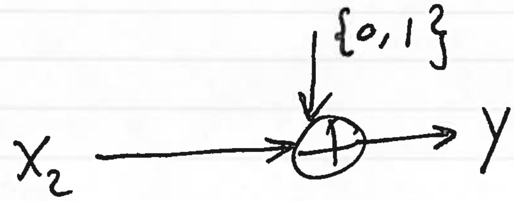


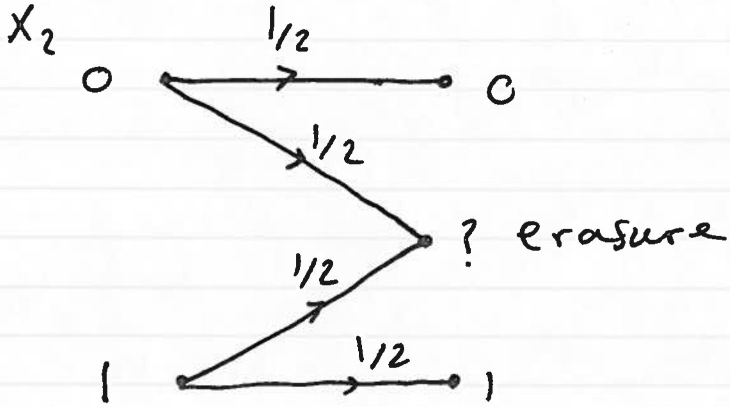
When $Y=0$ or $Y=2$, X_1 and X_2 can be found without ambiguity, but $Y=1$ can be either the result of $X_1=0$, $X_2=1$ or $(X_1=1, X_2=0)$.

Assume that we fix $X_2=0$, then rate ~~of~~ $R_2=0$ and $R_1=1$.

Similarly when we fix $X_1=0$, rate $R_1=0, R_2=1$

Now assume that X_1 is transmitted at the rate of 1 bits/transmission. Then the channel from X_2 is like

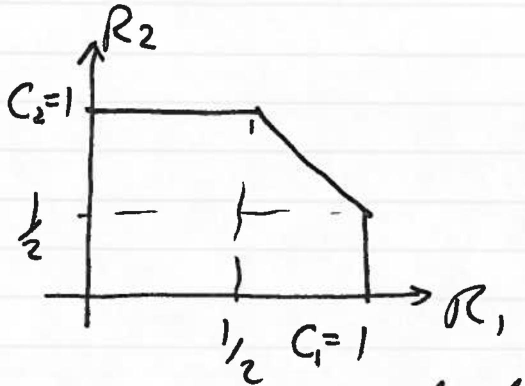




This is an erasure channel with erasure probabi 0.5, so its capacity is $1-\varepsilon = \frac{1}{2}$.

So, we can achieve the rate pair $(1, \frac{1}{2})$ and any rate $(1, R_2)$ where $R_2 < \frac{1}{2}$.

Similarly $(\frac{1}{2}, 1)$ rate can be achieved.



Main points in the proof of achievability of

$$R_1 \leq I(X_1; Y | X_2)$$

$$R_2 \leq I(X_2, Y | X_1)$$

$$R_1 + R_2 \leq I(X_1, X_2; Y)$$

- first fix $p(x_1, x_2) = P_1(x_1) P_2(x_2)$
- generate Codebook number 1 with

$$C_1(i) \quad i \in \{1, 2, \dots, 2^{nR_1}\}$$

Codebook 2

$$C_2(j) \quad j \in \{1, 2, \dots, 2^{nR_2}\}$$

or $X_1(i)$ {to be consistent with the text}

- Encoder 1 sends $C_1(i)$ to represent i.

  Encoder 2 " $C_2(j)$ " " j.

  or $X_2(j)$

- Decoder decides (i,j) if

$$(X_1(i), X_2(j), \underline{y}) \in A_\epsilon^{(n)}$$

where $\underline{y}$ is the noisy version of $\underline{x}_1(i) + \underline{x}_2(j)$.
(or any other way $\underline{x}_1(i)$ and $\underline{x}_2(j)$ are combined)

Analysis of the probability of error.

let $E_{ij} = \{ (X_1(i), X_2(j), Y) \in A_\epsilon^{(n)} \}$

$$P_e^{(n)} = P(E_{11}^c \cup_{(i,j)\neq(1,1)} E_{ij})$$

$$\leq P(E_{11}^c) + \sum_{\substack{i\neq 1\\ j=1}} P(E_{i1}) + \sum_{\substack{i=1\\ j\neq 1}} P(E_{1j}) + \sum_{\substack{i\neq 1\\ j\neq 1}} P(E_{ij})$$

assuming that $i=1$ and $j=1$ ~~were~~ the original messages (due to ~~[illegible]~~ symmetry, this does not entail any loss of generality).

it can be shown that

$$P(E_{i1}) \leq 2^{-n(I(X_1;Y|X_2)-3\varepsilon)}$$

similarly

$$P(E_{1j}) \leq 2^{-n(I(X_2;Y|X_1)-3\varepsilon)}$$

and

$$P(E_{ij}) \leq 2^{-n(I(X_1,X_2;Y)-4\varepsilon)}$$

So

$$P_e^{(n)} \leq P(E_{11}^c) + 2^{nR_1} \cdot 2^{-n(I(X_1;Y|X_2)-3\varepsilon)} + 2^{nR_2} \cdot 2^{-n(I(X_2;Y|X_1)-3\varepsilon)} + 2^{n(R_1+R_2)} \cdot 2^{-n(I(X_1,X_2;Y)-4\varepsilon)}$$

Since $\varepsilon > 0$ the conditions result in $P_e^{(n)} \to 0$ as $n \to \infty$

Discussion :

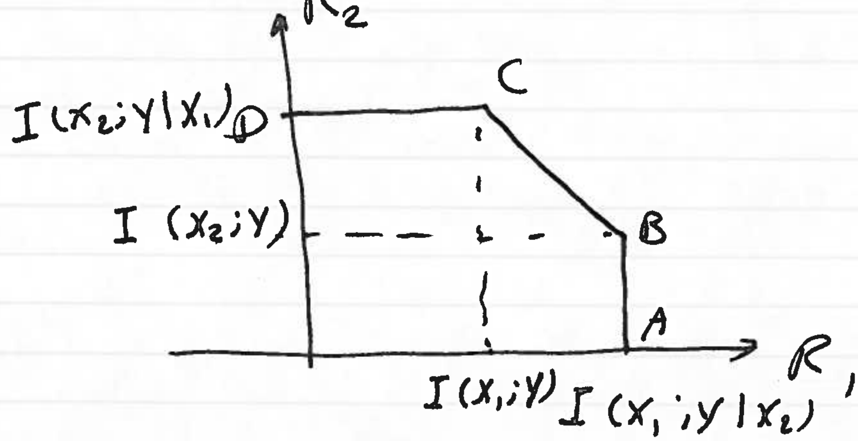
The maximum point on R_1 axis is (point A)

$$\max R_1 = \max_{P_1(x_1)P_2(x_2)} I(X_1;Y|X_2)$$

$$I(X_1;Y|X_2) = \sum_{x_2} P_2(x_2) I(X_1;Y|X_2=x_2)$$

$$\leq \max_{x_2} I(X_1;Y|X_2=x_2)$$

This point corresponds to choosing the value of X_2 (say x_2) that maximizes the value of



$I(X_1;Y|X_2=x_2)$, i.e., facilitates transmission of X_1.
Point B corresponds to the time where X_1 is being transmitted at the maximum possible rate and acts as noise for the transmission of X_2, in this case the maximum rate is $I(X_2;Y)$.

Points D and C are the counterparts of the points A and B, respectively. The points between B and C can be achieved by ~~time-sh~~ time-sharing.

receiver knows which X_2 was transmitted (since $R_2 < I(X_2;Y)$) can subtract from Y. We can look at the channel for X_1 as a X_2-indexed channel and

$$\sum_{x_2} p(x_2)\, I(X_1;Y|X_2=x_2) = I(X_1;Y|X_2)$$ is achievable for R_1

Theorem : (Convexity of the capacity region).

The capacity region is convex, i.e., if

$$(R_1, R_2) \in \mathcal{C} \text{ and } (R_1', R_2') \in \mathcal{C} \text{ then}$$

$$(\lambda R_1 + (1-\lambda) R_1', \lambda R_2 + (1-\lambda) R_2') \in \mathcal{C}$$

for any $0 \leq \lambda \leq 1$.

Proof : Assume we have two sequences of codes one for $\underline{R} = (R_1, R_2)$ and other for $\underline{R}' = (R_1', R_2')$.

Out of n bits to be transmitted, $n\lambda$ of them ~~are sent~~ we send ~~by~~ using the first codebook and $(1-n)\lambda$ bits by the second. The number of X_1 codewords is:

$$2^{n\lambda R_1} \cdot 2^{n(1-\lambda)R_1'} = 2^{n[\lambda R_1 + (1-\lambda)R_1']}$$

Similarly for X_2

$$2^{n\lambda R_2} \cdot 2^{n(1-\lambda)R_2'} = 2^{n[\lambda R_2 + (1-\lambda)R_2']}$$

The probability of error ~~is~~ does not exceed

$$P_e^{(n\lambda)} + P_e^{(n(1-\lambda))}$$ and tends to zero.

So, we have made an achievable rate pair $(\lambda R_1 + (1-\lambda) R_1', \lambda R_2 + (1-\lambda) R_2')$

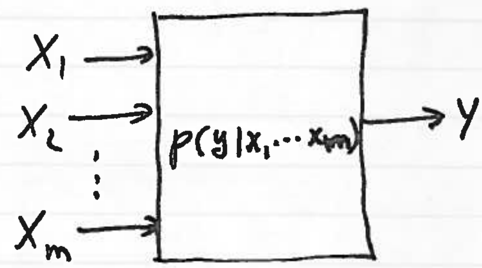
<u>Theorem</u>: The capacity region of the m-user multiple access channel is the closure of the convex hull of the rate vectors satisfying

$$R(S) \leq I(X(S); Y | X(S^c))$$

for all $S \subseteq \{1, 2, \ldots, m\}$

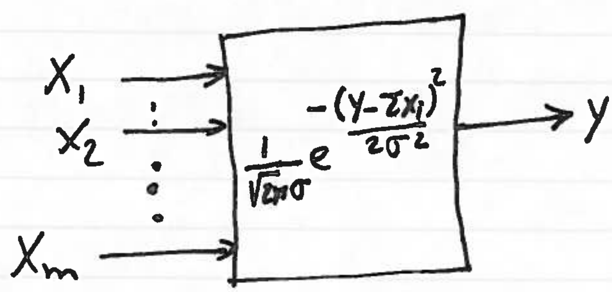
for some product of probabilities $p_1(x_1)p_2(x_2)\cdots p_m($

where $R(S) = \sum_{i \in S} R_i$

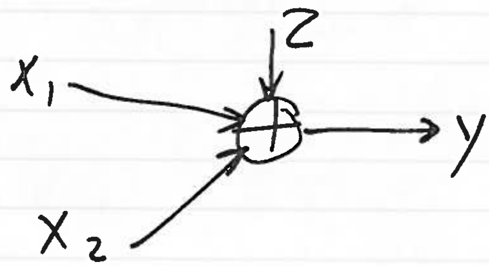


and $\quad X(S) = \{X_i : i \in S\}$

<u>Gaussian Multiple-Access Channel</u>



Take first $m = 2$



At time i,

$$Y_i = X_{1i} + \cancel{X_{2i}} + Z_i$$

where $\{Z_i\}$ is a sequence of i.i.d. normal, zero mean random processes with variance N.

Assume power constraint P_j,

$$\frac{1}{n}\sum_{i=1}^{n} x_{ji}^2(w_j) \le P_j, \quad w_j \in \{1, 2, \dots, 2^{nR_j}\}, j = 1, 2$$

The achievability conditions (rate-region) can be extended to the Gaussian case (note, so far, we have only talked about discrete case):

$$R_1 \le I(X_1; Y | X_2)$$

$$R_2 \le I(X_2; Y | X_1)$$

$$R_1 + R_2 \le I(X_1, X_2; Y)$$

We have

$$I(X_1; Y | X_2) = h(Y | X_2) - h(Y | X_1, X_2)$$

$$= h(X_1 + X_2 + Z | X_2) - h(X_1 + X_2 + Z | X_1, X_2)$$

$$= h(X_1 + Z | X_2) - h(Z)$$

$$= h(X_1 + Z) - h(Z) = h(X_1 + Z) - \frac{1}{2}\log(2\pi e)N$$

$$\le \frac{1}{2}\log(2\pi e)(P_1 + N) - \frac{1}{2}\log 2\pi e N = \frac{1}{2}\log\left(1 + \frac{P_1}{N}\right)$$

so,

$$R_1 \leq \frac{1}{2} \log\left(1+\frac{P_1}{N}\right)$$

~~call~~ Denote $\frac{1}{2}\log(1+x)$ by $C(x)$.

Then

$$R_1 \leq C\left(\frac{P_1}{N}\right)$$

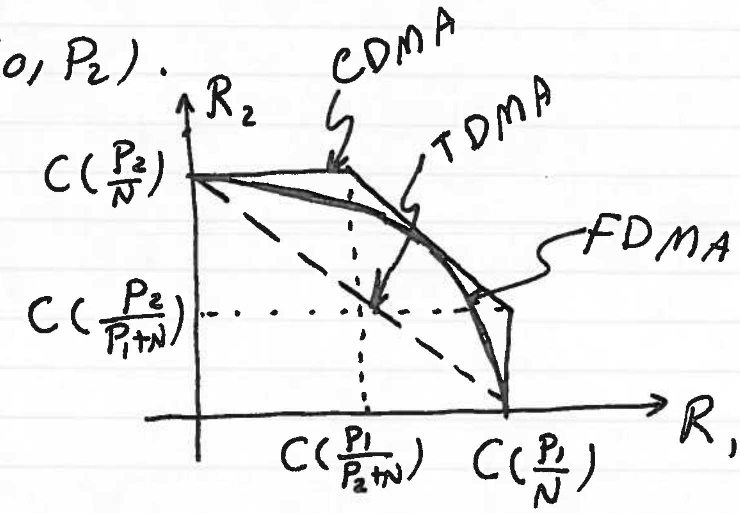
Similarly

$$R_2 \leq C\left(\frac{P_2}{N}\right)$$

and

$$R_1 + R_2 \leq C\left(\frac{P_1+P_2}{N}\right)$$

The upper bounds are achieved if $X_1 \sim \mathcal{N}(0, P_1)$ and $X_2 \sim \mathcal{N}(0, P_2)$.



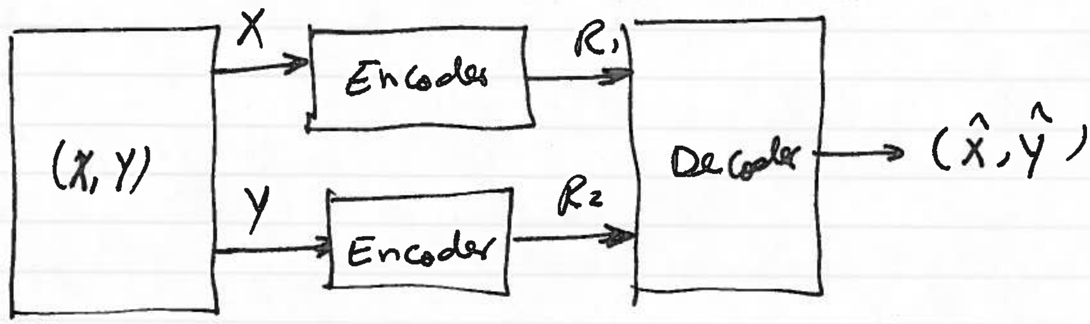
with FDMA

$$R_1 = \frac{W_1}{2} \log\left(1+\frac{P_1}{NW_1}\right)$$

$$W_1 + W_2 = W$$

$$R_2 = \frac{W_2}{2} \log\left(1+\frac{P_2}{NW_2}\right)$$

Encoding of correlated sources:



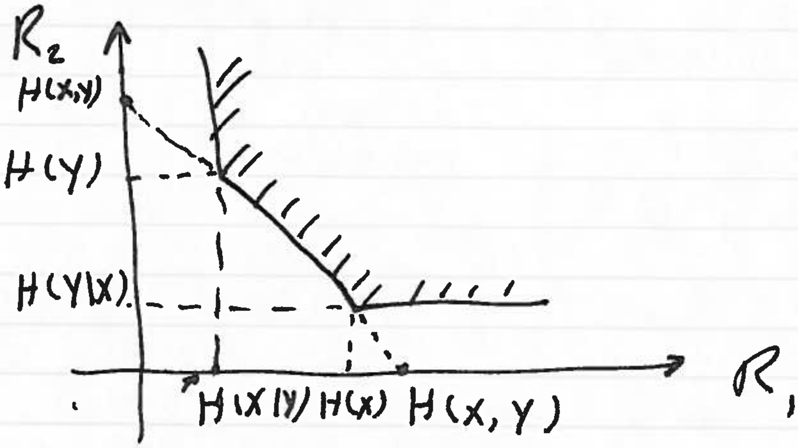
~~Step~~:

<u>Slepian-Wolf Theorem</u>: For the distributed source (X, Y) i.i.d. $\sim p(x,y)$, the <u>achievable rate region</u> is given by (for separate encoders)

$$R_1 \geq H(X|Y)$$

$$R_2 \geq H(Y|X)$$

$$R_1 + R_2 \geq H(X,Y)$$



This theorem corresponds to $(0,0,1;1)$ case in the original paper, i.e., decoder sees both I_X and I_Y but encoders only see their own input (X or Y).

by achievable rate-region, we mean the set of all (R_1, R_2) such that a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ code with $P_e^{(n)} \to 0$ exists.

General case: Slepian-Wolf algorithm for many sources:

Theorem: Let $(X_{1i}, X_{2i}, \ldots, X_{mi})$ be i.i.d. $\sim p(x_1, \ldots, x_m)$. Then the set of rate vectors achievable for distributed source coding with separate encoders is defined by:

$$R(S) > H(X(S) \mid X(S^c))$$

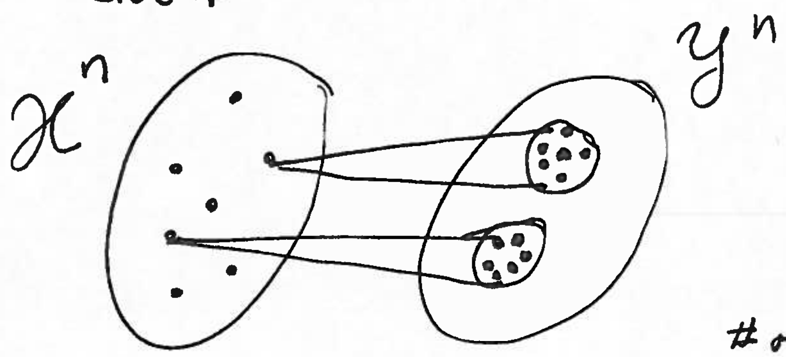
~~where~~ for all $S \subseteq \{1, 2, \ldots, m\}$ where

$$R(S) = \sum_{i \in S} R_i$$

and

$$X(S) = \{X_j : j \in S\}$$

Interpretation



Coloring scheme (relation to cosets)

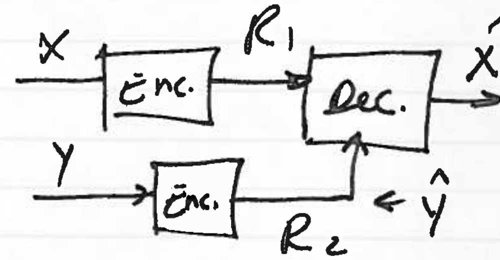
if $R_2 > H(Y|X)$
# of colors $> 2^{nH(Y|X)}$

Source Coding with side information

Theorem: Let $(X, Y) \sim p(x,y)$. If Y is encoded at rate R_2 and X is encoded at rate R_1, we can recover X with an arbitrarily small probability of error if and only if:

$$R_1 \geq H(X|U)$$

$$R_2 \geq I(Y;U)$$



for some joint probability ~~dens~~ mass function $p(x,y)\,p(u|y)$ where $|\mathcal{U}| \leq |\mathcal{Y}| + 2$.

Proof: refer to book.

Interpretation: if $R_2 > H(Y)$, then Y can be described perfectly and from Slepian-Wolf theorem $R_1 = H(X|Y)$ bits is sufficient to describe X. Also, if (on the other extreme), $R_2 = 0$, then $R_1 = H(X)$ is needed to describe X. In general, we can use $I(Y;\hat{Y})$ to represent Y approximately. Then, we need $H(X|\hat{Y})$ to represent X.

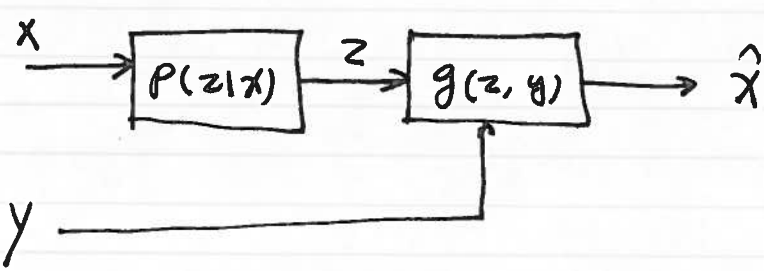
The rate-distortion function with side information:

Theorem: Let (X,Y) be drawn i.i.d. $\sim p(x,y)$ and let $d(x^n, \hat{x}^n) = \frac{1}{n}\sum_{i=1}^{n} d(x_i, \hat{x}_i)$ be given. The rate-distortion with side information is:

$$R(D) = \min_{M(D)} \left(I(X;Z) - I(Y;Z) \right)$$

where $M(D)$ is the set of all encoders $p(z|x)$, decoders $\hat{x} = g(z,y)$ such that the overall distortion is less than D, i.e.,

$$\sum_x \sum_z \sum_y p(x,y)\, p(z|x)\, d(x, g(z,y)) \leq D.$$



note that

$$I(X;Z) - I(Y;Z) = H(Z|Y) - H(Z|X)$$

$$= H(Z|Y) - H(Z|X,Y)$$

$$= I(X;Z|Y)$$

So, $R(D) = \min_{M(D)} I(X;Z|Y)$

For a Gaussian source X with side information $Y = X + U$ where $X \sim N(0, \sigma_x^2)$ and $U \sim N(0, \sigma_u^2)$

$$R^*(D) = R_{X|Y}(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma_u^2 \sigma_x^2}{(\sigma_u^2 + \sigma_x^2) D} & 0 < D \leq \frac{\sigma_u^2 \sigma_x^2}{\sigma_x^2 + \sigma_u^2} \\ 0 & D \geq \frac{\sigma_u^2 \sigma_x^2}{\sigma_x^2 + \sigma_u^2} \end{cases}$$

where $R^*(D)$ is for ~~separated~~ un-informed encoders and $R_{X|Y}(D)$ is for ~~w~~ the case where the encoder has knowledge of Y.