13.09.05.

* Randomly generated codes are good:

$X \geq 0, \quad E[X] < \infty$

$$P(X \geq \delta) = E[1(X \geq \delta)]$$

$$\leq E\left[\frac{X}{\delta} 1(X \geq \delta)\right]$$

$$\leq \frac{E[X]}{\delta}$$

($\frac{X}{\delta} \geq 1$ on those events)

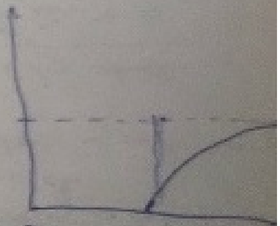
$$E[P_e^{(n)}(e)] \leqslant 10^{-20}$$

$$P(P_e^{(n)}(e) > 10^{-10}) = P(P_e^{(n)}(e) > 10^{-10})$$

$$\leqslant \frac{E[P_e^{(n)}(e)]}{10^{-10}} \leqslant 10^{-10}$$

* Strong Converse.

$$\lim_{n \to \infty} P_e^{(n)} \geqslant 1 - \frac{C}{R}, \quad \text{for } R > C.$$



$$X = (Y - E[Y])^2. \quad \text{[2nd order moment exists].}$$

$$P(|Y - E[Y]| > \epsilon) = P(X > \epsilon^2) \leqslant E[X]/\epsilon^2.$$

$$= \frac{\mathrm{Var}[Y]}{\epsilon^2}.$$

* Let $\{q(x)\} = \arg\max_{\{P(x)\}} I(X;Y)$

Let $q_Y(y) = \sum_x q(x)\, p(y|x)$ → transition probability of the Channel.

$$J(x;y) = \log\left[\frac{p(y|x)}{q_Y(y)}\right]$$

$$E_x[J(x;Y)] = \sum_y p(y|x) \log\left[\frac{p(y|x)}{q_Y(y)}\right] \leqslant C$$

input has all the mass at x.

$$q_{Y^n}(y^n) = \prod_{i=1}^{n} q_Y(y_i)$$

$$(\ldots) = \sum_{i}^{n} J(x_i; y_i).$$

Let $\{q(x)\} = \arg\max_{\{P(x)\}} I(X;Y)$

Let $q_Y(y) = \sum_x q(x)\, p(y|x)$ → transition probability of Channel.

$$J(x;y) = \log\left[\frac{p(y|x)}{q_Y(y)}\right]$$

$$E_x\left[J(x;Y)\right] = \sum_y p(y|x) \log\left[\frac{p(y|x)}{q_Y(y)}\right] \leq C$$

input has all the mass at x

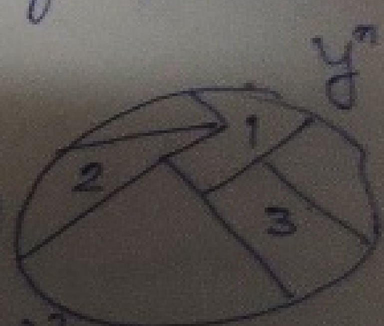
$$q_{Y^n}(y^n) = \prod_{i=1}^{n} q_Y(y_i)$$

$$J(x^n;y^n) = \sum_{i=1}^{n} J(x_i;y_i).$$

$M = 2^{nR}$

$\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_M$ are the decoding regions, i.e.,

$g(Y^n) = m$ if $Y^n \in \mathcal{Y}_m$.



$\epsilon > 0$ arbitrary.

$$B_m := \{y^n \in \mathcal{Y}^n : J(\underbrace{x^n(m)}_{m\text{-th Codeword}}, y^n) > n(C+\epsilon)\}$$

$$P_c = \frac{1}{M}\sum_{m=1}^{M} \sum_{y^n \in \mathcal{Y}_m} p(y^n|x^n(m))$$ (Probability of making a correct decision)

$$= \frac{1}{M}\sum_{m=1}^{M} \sum_{y^n \in \mathcal{Y}_m \cap B_m} p(y^n|x^n(m)) + \frac{1}{M}\sum_{m}\sum_{y \in \mathcal{Y}_m \cap B_m^c} p(y^n|$$

* Let $\{q(x)\} = \arg\max_{\{P(x)\}} I(X;Y)$

Let $q_Y(y) = \sum_x q(x)\, p(y|x)$ $\hookrightarrow$ transition probability of the Channel.

$$J(x;y) = \log\left[\frac{p(y|x)}{q_Y(y)}\right]$$

$$E_x\left[J(x;Y)\right] = \sum_y p(y|x) \log\left[\frac{p(y|x)}{q_Y(y)}\right] \leq C$$

input has all the mass at x

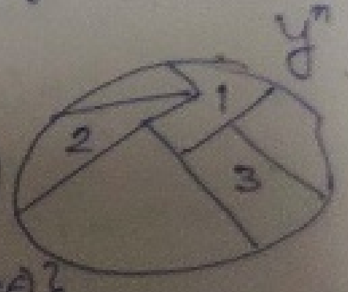
$$q_{Y^n}(y^n) = \prod_{i=1}^{n} q_Y(y_i)$$

$$J(x^n; y^n) = \sum_{i=1}^{n} J(x_i; y_i).$$

$M = 2^{nR}$

$\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_M$ are the decoding regions, i.e.,

$g(Y^n) = m$ if $Y^n \in \mathcal{Y}_m$.



$\epsilon > 0$ arbitrary.

$$B_m := \{y^n \in \mathcal{Y}^n : \underbrace{J(x^n(m), y^n)}_{m\text{-th codeword}} > n(C+\epsilon)\}$$

$$P_c = \frac{1}{M} \sum_{m=1}^{M} \sum_{y^n \in \mathcal{Y}_m} p(y^n | x^n(m))$$ (Probability of making a correct decision)

$$= \frac{1}{M} \sum_{m=1}^{M} \sum_{y^n \in \mathcal{Y}_m \cap B_m} p(y^n | x^n(m)) + \frac{1}{M} \sum_m \sum_{y \in \mathcal{Y}_m \cap B_m^c} p(y^n | x^n(m))$$

$=: P_{c,1} + P_{c,2}$

$R > C$

$$J(x^n(m); y^n) = \sum_{i=1}^{n} J(x_i(m), y_i)$$

$$= \sum_i \log\left[\frac{p(y_i | x_i(m))}{q_Y(y_i)}\right]$$

$$= \log\left[\frac{p(y^n | x^n(m))}{q_{Y^n}(y^n)}\right] \geq n(C+\epsilon) \text{ on } B_m$$

On B_m^c, $p(y^n | x^n(m)) \leq 2^{n(C+\epsilon)} \cdot q_{Y^n}(y^n)$.

$$P_{c,2} \leq \frac{2^{n(C+\epsilon)}}{M} \sum_{m=1}^{M} \sum_{y^n \in \mathcal{Y}_m \cap B_m^c} q_{Y^n}(y^n)$$

$$P_{c,2} \leq 2^{n(C-R+\epsilon)} \quad \ldots (*)$$

B_m are those pts w.r.t. which we are trying to do the average.

$$P_{c,1} = \frac{1}{M} \sum_{m=1}^{M} \sum_{y^n \in B_m \wedge \mathcal{Y}_m} p(y^n | x^n(m))$$

$$\leq \frac{1}{M} \sum_{m=1}^{M} \sum_{y^n \in B_m} p(y^n | x^n(m)) -$$

$$\sum_{y^n \in B_m} p(y^n | x^n(m)) = P(Y^n \in B_m | x^n(m))$$

$$= P(J(x^n(m); Y^n) > n(C+\epsilon) | x^n(m))$$

$$= P\left(\sum_{i=1}^{n} J(x_i(m), Y_i) > n(C+\epsilon) \,\Big|\, x^n(m)\right)$$

[we have m-random variables which are independent, but not i.i.d.'s]

$E_q[J(X,Y)] = C$ [when J is distributed at q].

$$\operatorname{var}\left[\sum_{i=1}^{n} J(x_i(m), Y_i)\right] = \sum_{i=1}^{n} \operatorname{var}[J(x_i(m); Y_i)]$$

$$\leq nA, \quad A = \max_i \operatorname{var}[J(x; Y)]$$

everything is only a finite number of times

$$P_{c,1} \leq \frac{nA}{n^2\epsilon^2} = \frac{A}{n\epsilon^2}$$

$$P_c \leq \frac{A}{n\epsilon^2} + 2^{n(C-R+\epsilon)}, \quad \epsilon > 0$$

$$\epsilon = \frac{R-C}{2}$$

$$P_c \leq \underbrace{\frac{4A}{(R-C)^2} \cdot \frac{1}{n}}_{\to 0} + \underbrace{2^{-n(R-C)/2}}_{\to 0}$$

for any code which is operating at a rate $> C$

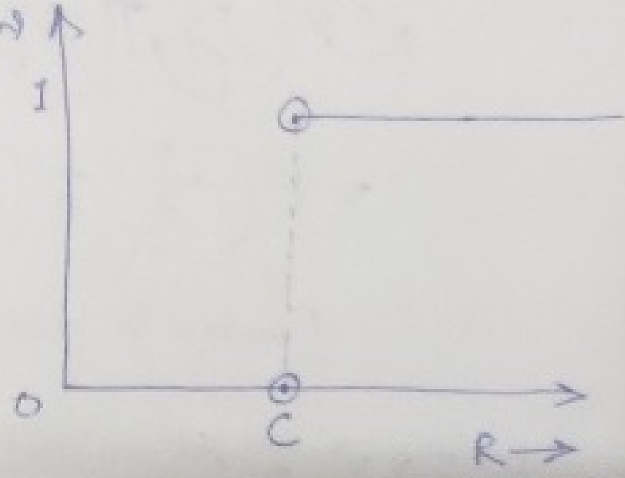
Probability of Correct Decision is going to 0

" " Error " " 1

Strong Converse

We don't know what happens at C.

Probability of error decays exponentially fast.



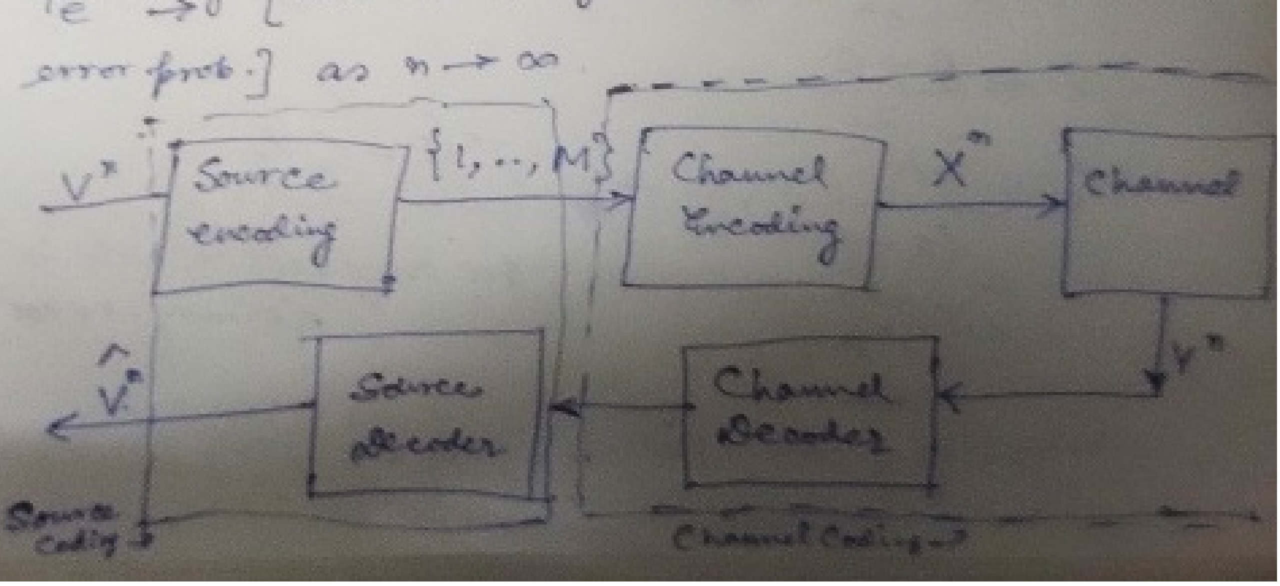
✶ Source - Channel Separation Thm:

Let $\{V_k\}$ be a stationary source that satisfies the AEP & to be transmitted over a DMC with capacity C.

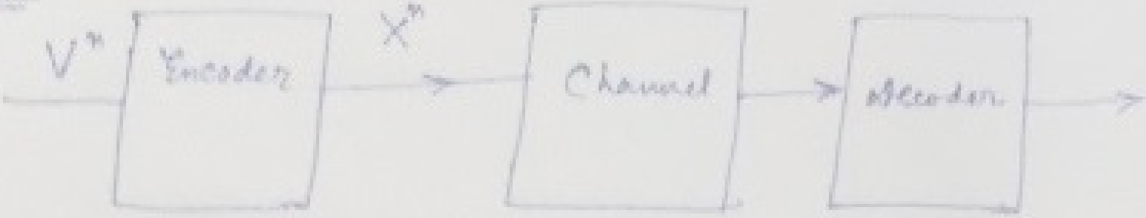
Achievability: If $H(\{V_k\}) < C$, then

$\exists$ a seperate source & channel coding scheme s.t.

$P_e^{(n)} \to 0$ [We are looking at block error prob, not symbol error prob.] as $n \to \infty$



Converse:



Minimize the data to transmitted
/ Maximize the rate of transfer

If $P_e^{(n)} \to 0$ for some sequence of codes, then the entropy rate $H(\{V_k\}) \leq C$. (This is a weak converse)

Proof: Achievability:

$$P\left(V^n \notin A_\epsilon^{(n)}\right) < \epsilon$$

Number of typical resources

(we are assuming that stationary sources are satisfying the AEP)

$$= \left|A_{\epsilon}^{(n)}\right| \leq 2^{\underbrace{n(H+\epsilon)}_{\substack{\shortparallel \\ R}}}$$

Source Code: $v^n \to \{1, \ldots, 2^{nR}\}$.

$v^n \in A_\epsilon^{(n)} \mapsto m(v^n)$

Anything within typical set we are performing no error.
error will be outside it

$P($ source code error$) < \epsilon$ for n large enough

suppose $R = H(\{V_k\}) + \epsilon < C$

Choose that ϵ s.t the above is going to be true

Then $\exists$ a channel code for n large enough s.t.

$P_e($Channel Code$) < \epsilon$

$\therefore P_e = P_e($source code error$) + P_e($Channel Code$) < 2\epsilon$

Converse: (* is going to be Fano's inequality)

$$H(\{V_k\}) \leq \frac{H(V^n)}{n}$$ [We have monotonically decreasing sequence]

$$= \frac{H(V^n | \hat{V}^n)}{n} + \frac{I(V^n; \hat{V}^n)}{n}$$

Fano's inequality $\leq \frac{(1 + P_e^{(n)} \log |\mathcal{V}|^n)}{n} + \frac{I(V^n; \hat{V}^n)}{n}$

Markov chain

$V^n \to X^n \to Y^n \to \hat{V}^n$

data-processing inequality

$$\leq \frac{1}{n} + P_e^{(n)} \log |\mathcal{V}| + I(X^n; Y^n)$$

$$\leq \frac{1}{n} + P_e^{(n)} \log |\mathcal{V}| + C$$

$n \to \infty$, $\frac{1}{n} \to 0$, $P_e^{(n)} \to 0$ $\therefore H(\{V_k\}) \leq C$