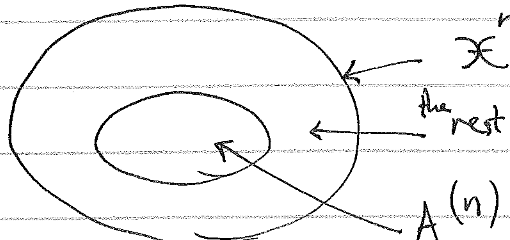
4

[illegible]

~~Channel Coding~~

Source Coding by typicality:

$X_1, \ldots, X_n$ iid $\sim p(x)$ over $\mathcal{X}$.



$A_\varepsilon^{(n)}$, $|A_\varepsilon^{(n)}| \le 2^{n(H(X)+\varepsilon)}$

* Encode each element in $A_\varepsilon^{(n)}$ by its index in $A_\varepsilon^{(n)}$ following a zero in $\lceil n(H(X)+\varepsilon)\rceil + 1$ $\le n[H(X)+\varepsilon] + 2$ bits.

* Anything else encoded as $(1, \text{trivial})$ in $\le n \log|\mathcal{X}| + 2$ bits.

* Easy to see: uniquely decodable.

* Take n large so that $\Pr(A_\varepsilon^{(n)} \ge 1-\varepsilon)$

Expected length: $\mathbb{E}(l(X_1^n)) = \sum_{x_1, \ldots, x_n} p(x_1, \ldots, x_n)\, l(x_1, \ldots, x_n)$

$=$?

$$= \sum_{x_1,\dots,x_n \in A_\varepsilon^{(n)}} p(x_1,\dots,x_n)\, l(x_1,\dots,x_n)$$

$$+ \sum_{x_1,\dots,x_n \notin A_\varepsilon^{(n)}} p(x_1,\dots,x_n)\, l(x_1,\dots,x_n)$$

$$\leq \sum_{x_1,\dots,x_n \in A_\varepsilon^{(n)}} p(x_1,\dots,x_n)\, \big(n(H+\varepsilon)+2\big)$$

$$+ \sum_{x_1,\dots,x_n \notin A_\varepsilon^{(n)}} p(x_1,\dots,x_n)\, \big(n \log|\mathcal{X}|+2\big)$$

$$= \Pr\big((X_1,\dots,X_n) \in A_\varepsilon^{(n)}\big)\big(n(H+\varepsilon)+2\big)$$

$$+ \Pr\big((X_1,\dots,X_n) \notin A_\varepsilon^{(n)}\big)\big(n \log|\mathcal{X}|+2\big)$$

$$\leq n(H+\varepsilon) + \varepsilon n\big(\log|\mathcal{X}|\big)+2$$

$$= n(H+\varepsilon') \longrightarrow nH(X)$$

$\varepsilon' = \varepsilon + \varepsilon \log|\mathcal{X}| + \frac{2}{n}$ arbitrarily small.

Postscript: * Extension to the bounded-entropy fixed-length universal source coding *

(Csiszár–Körner scheme)

Joint Typicality

Suppose $(X_1,Y_1), \dots, (X_n,Y_n)$ iid $\leftarrow (X,Y)$. (over $\mathcal{X}\times\mathcal{Y}$)

$A_\varepsilon^{(n)}$ for this joint distribution is defined as:

$$A_\varepsilon^{(n)} \stackrel{def}{:=} \Big\{ (x_1^n, y_1^n) \in \mathcal{X}^n \times \mathcal{Y}^n :$$

$$\left| -\frac{\log p(x_1^n)}{n} - H(X) \right| < \varepsilon,$$

$$\left| -\frac{\log p(y_1^n)}{n} - H(Y) \right| < \varepsilon,$$

$$\left| -\frac{\log p(x_1^n, y_1^n)}{n} - H(X,Y) \right| < \varepsilon \Big\}.$$

Joint AEP Theorem:

1. $P_r\left((X_1^n, Y_1^n) \in A_\varepsilon^{(n)} \right) \longrightarrow 1$ as $n \longrightarrow \infty$

2. $\left| A_\varepsilon^{(n)} \right| \leq 2^{n(H(X,Y)+\varepsilon)}$, (for large n) $\left| A_\varepsilon^{(n)} \right| \geq (1-\varepsilon) 2^{n(H(X,Y)-\varepsilon)}$.

3. If $(\tilde{X}_1^n, \tilde{Y}_1^n) \sim p(x_1^n)p(y_1^n)$, then

$$P_r\left((\tilde{X}_1^n, \tilde{Y}_1^n) \in A_\varepsilon^{(n)} \right) \leq 2^{-n(I(X;Y)-3\varepsilon)}.$$

&

$$\text{"} \geq (1-\varepsilon) 2^{-n(I(X;Y)+3\varepsilon)}.$$

preef of lower bound in (2).

for $\forall n$ large, ~~for~~ $Pr\left(A_\varepsilon^{(n)}\right) \geq 1-\varepsilon \Rightarrow$

$$1-\epsilon \leq \sum_{(x^n,y^n)\in A_\varepsilon^{(n)}} p(x^n,y^n)$$

$$\leq |A_\epsilon^{(n)}| \; 2^{-n(H(X,Y)-\varepsilon)} \Rightarrow \checkmark$$

12

upper bound

<u>Proof:</u> (2) ✓ follows from 1-variable AEP.
(lowerbound ←)

<u>Pf (1):</u> By weak law of large numbers:

$\forall \varepsilon > 0, \exists n_0$ s.t. $\forall n > n_0$,

$$\Pr\left(\left| -\frac{\log p(X_1^n)}{n} - H(X) \right| \geq \varepsilon \right) \leq \frac{\varepsilon}{3}$$

and the same for Y and (X,Y):

$$\Pr\left(\left| -\frac{\log p(Y_1^n)}{n} - H(Y) \right| \geq \varepsilon \right) \leq \frac{\varepsilon}{3},$$

$$\Pr\left(\left| -\frac{\log p(X_1^n, Y_1^n)}{n} - H(X,Y) \right| \geq \varepsilon \right) \leq \frac{\varepsilon}{3}.$$

Take a union bound and done! //

<u>Pf (3):</u>

$$\Pr\left((\tilde{X}_1^n, \tilde{Y}_1^n) \in A_\varepsilon^{(n)} \right) = \sum_{(x_1^n, y_1^n) \in A_\varepsilon^{(n)}} p(x_1^n)\, p(y_1^n)$$

$$\leq \underbrace{2^{n(H(X,Y)+\varepsilon)}}_{\text{by (2)}} \cdot 2^{-n(H(X)-\varepsilon)}\, 2^{-n(H(Y)-\varepsilon)}$$

$$= 2^{-n(I(X;Y) - 3\varepsilon)} \checkmark$$

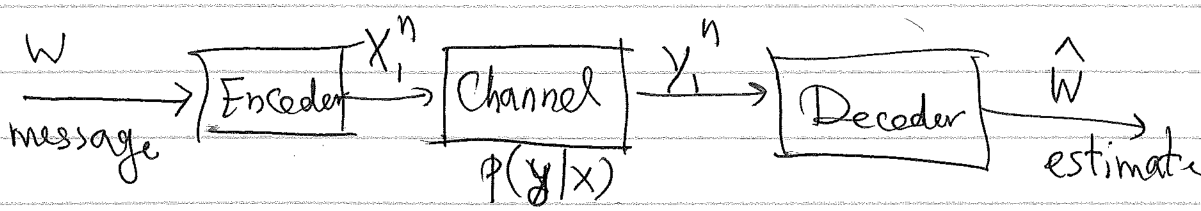
& (lower bound similar) □ ✓

~~Second part: For large n,~~ $\cancel{\Pr(A_\varepsilon^{(n)}) \geq 1-\varepsilon \Rightarrow}$

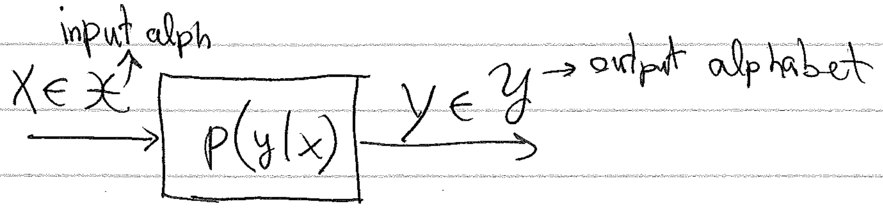
$\cancel{(1-\varepsilon) \leq \sum_{(x_1^n, y_1^n) \in A_\varepsilon^{(n)}} p(x_1^n, y_1^n) \leq |A_\varepsilon^{(n)}|\, 2^{-n(H(X,Y)-\varepsilon)}}$

Channel Coding

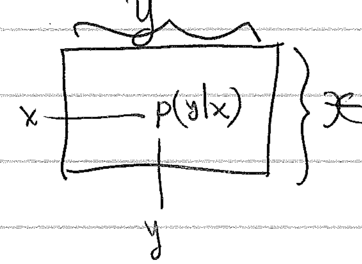
Basic "point-to-point", "no-feedback" Communication System:



(Discrete) Channel:



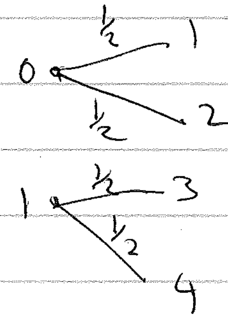
Channel can be described by a "probability transition matrix" $p(y|x)$:



Ex. Noiseless:

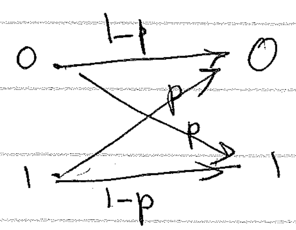
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

"practically noiseless":



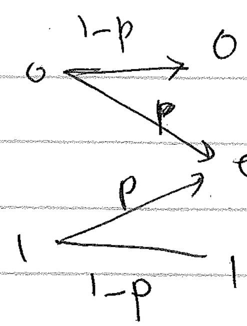
$$\begin{array}{c|cccc|c} & 1 & 2 & 3 & 4 & \\ \hline & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{array}$$

B.S.C(p):



$$p(y|x) = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}.$$

* BEC(p):



$$p(y|x) = \begin{array}{c} \\ \end{array}\begin{array}{cc} & \begin{array}{ccc} 0 & e & 1 \end{array} \\ & \begin{pmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{pmatrix} \begin{array}{c} 0 \\ 1 \end{array} \end{array}$$

* Symmetric Channel: Rows and col's are permutations of each other.

ex. BSC(p) and $p(y|x) = \begin{pmatrix} .3 & .2 & .5 \\ .5 & .3 & .2 \\ .2 & .5 & .3 \end{pmatrix}$

(doubly stochastic)

* Capacity

$$C \overset{\text{def}}{:=} \max_{p(x)} I(X;Y).$$

Ex: * Noise less : $C = 1$ $\left(p(x) = (\frac{1}{2}, \frac{1}{2}) \right)$.

* "practically noiseless": $C = 1$ $\left(p(x) = (\frac{1}{2}, \frac{1}{2}) \right)$

* BSC(p): $I(X;Y) = H(Y) - H(Y|X)$

$$= H(Y) - \sum p(x) \underbrace{H(Y|X=x)}_{h(p)}$$

$$= H(Y) - h(p) \leq 1 - h(p).$$

equality for $p(x) = (\frac{1}{2}, \frac{1}{2}) \Rightarrow C = 1 - h(p)$.

~~* BEC(p)~~

15

* BEC(p): same as before,

$$C = \max_{p(x)} H(Y) - h(p).$$

$$H(Y) = H(Y, \text{"erasure"}) = H(\text{erasure}) + H(Y|\text{erasure})$$

$$= h(p) + \underbrace{H(Y|\text{erasure})}_{(1-p)\,h(p_1)}, \quad p_1 := \Pr(X=1).$$

$$\Rightarrow C = \max_{p_1} (1-p)\,h(p_1) = 1-p.$$

$$(\text{for } p_1 = 1/2)$$

Ex: Symmetric Channel

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

$$= H(Y) - H(\text{row})$$

$$\leq \log|\mathcal{Y}| - H(\text{row}),$$

equality iff Y is uniform,

achieved by uniform X.