2.2 Channel Capacity

recall BSC = $(\{0,1\}, \cdot, \{0,1\})$

→ note: universal model

(a) Context

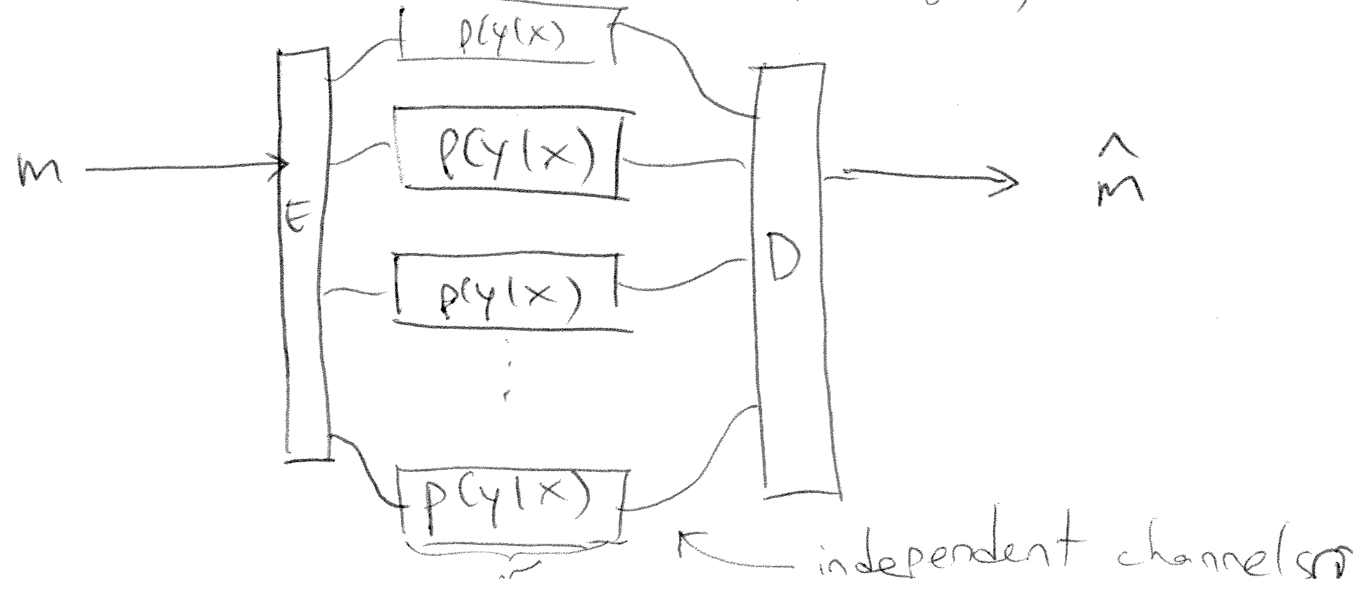
We are given access to some communication channel

$$(\mathcal{X}, p(y|x), \mathcal{Y}),$$

input alphabet · transition prob · output alphabet

and want to use it to send some message $m \in [M] = \{1, 2, \dots, M\}$, from sender to receiver ~~with~~ s.t.:

1. Vanishingly small probability of error, $P_e \leq \varepsilon$, for any $\varepsilon > 0$.
2. efficient (i.e. repetition code not good)



① Definitions

① Coding scheme $(\mathcal{C})$ consists of an encoder E

$$E : \{1, 2, \ldots, M\} \longrightarrow \mathcal{X}^n,$$

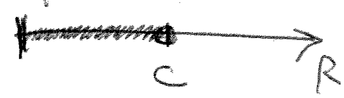
a decoder

$$D : \mathcal{Y}^n \longrightarrow \{1, 2, \ldots, M\}.$$

③ The rate of this scheme is

$$R = \frac{\#\text{bits of msg}}{\#\text{channel uses}} = \frac{\log_2 M}{n}$$

④ The capacity of a channel is the largest possible rate s.t. $P_e^* \to 0$.



② The probability of error

$$P_e(m, \mathcal{C}) = \Pr\{\overbrace{D(N(E(m)))}^{\hat{m}} \neq m\}$$
$$= \Pr\{\hat{m} \neq m\}$$

average prob error

$$\overline{P_e}(\mathcal{C}) = \frac{1}{M} \sum_{i=1}^{M} P_e(m, \mathcal{C})$$

max prob. error

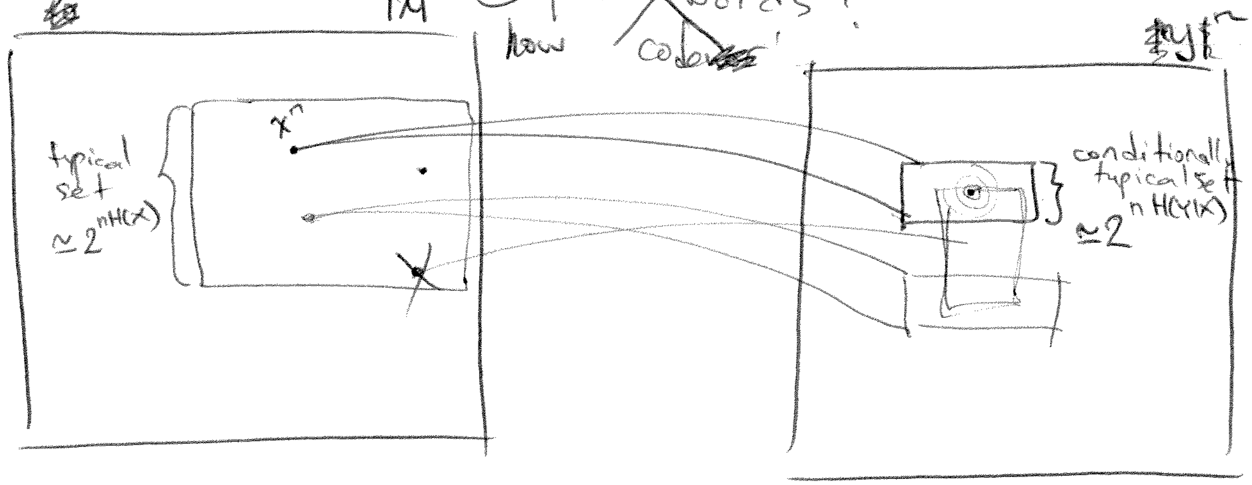
$$P_e^* = \max_m P_e(m, \mathcal{C})$$

(11) ~~Intuition~~ Clouds

(1) msg $i \to x^n(i)$, will get smeared out in output space $\mathcal{Y}$

(2) want no overlap

(3) pick words? how codewords



Defs

→ Conditional entropy

$$H(Y|X) \triangleq H(XY) - H(X) = -\sum_{x,y} p(x,y) \log p(y|x)$$

→ uncertainty in Y given X

while we are at it

→ mutual Information

$$I(X;Y) = H(X) + H(Y) - H(XY)$$

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

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

$$= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)\,p(y)} = D_{KL}\big(p(x,y) \,\|\, p(x)p(y)\big)$$

how ~~different~~ far from independent is $p(x,y)$

① ~~Shannon~~ random coding

Shannon's idea: Choose codewords at random.

① → Fix a prob dis $p_X(x)$

② $$\mathcal{C} = \begin{bmatrix} E(1) \\ E(2) \\ \vdots \\ \\ E(M) \end{bmatrix} = \left.\begin{bmatrix} \overbrace{X_1(1)\; X_2(1)\; X_3(1) \cdots \quad X_n(1)}^{\text{codeword}} \\ \\ \\ \\ \underbrace{X_1(M)\; X_2(M) \quad \cdots \quad X_n(M)}_{n \text{ symbols}} \end{bmatrix}\right\} M \text{ codewords}$$

→ note: every codeword is independent of message it encodes!

~~Probability for a particular code $\mathcal{C}_0$~~

$$\cancel{P_{\mathcal{C}}(\mathcal{C}_0) = \prod_{m=1}^{M} \prod_{i=1}^{n} p_X(x_i(m))}$$

(I) Exp. prob of error

$$\mathbb{E}_C\{\bar{P}_e(C)\} = \mathbb{E}_C\left\{ \frac{1}{M}\sum_{m=1}^{M} P_e(m,C) \right\}$$

$$= \frac{1}{M}\sum_{m=1}^{M} \underbrace{\mathbb{E}_C\{P_e(m,C)\}}_{\text{codeword choice indep of msg}}$$

$$= \frac{1}{M}\sum_{m=1}^{M} \mathbb{E}_C\{P_e(1,C)\}$$

$$= \mathbb{E}_C\{P_e(1,C)\} \leftarrow \text{simple to analyze!}$$

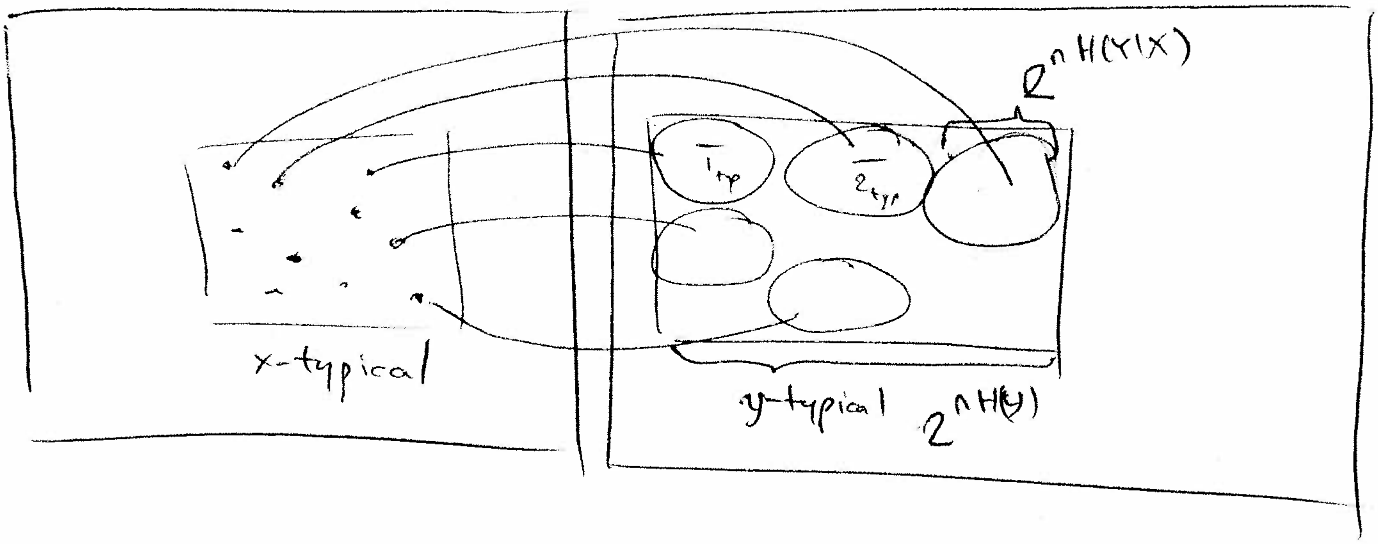
(II) Derandomization

if $\mathbb{E}_C\{\bar{P}_e(C)\} < \varepsilon$,

then $\exists\ C_0$ s.t $\bar{P}_e(C_0) < \varepsilon$.

→ Shannon proof (randomized arg in general) = existence proof not constructive.

Packing argument



~~Decode~~

Protocol

~~Generate random code~~

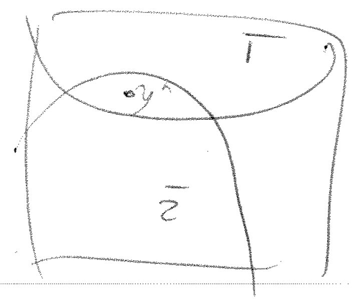
1. Alice & Bob exchange code "C" information.
2. ~~Alice sends msg m~~ To send message m, alice inputs $x^n(m)$ = "m^{th} row in C"
3. Channel acts
   $$x^n \to p(y|x)^{\otimes n} \to y^n$$
4. Bob look in which conditionally typical set y^n lies, to decode $\hat{m}$.
   if $\hat{m} = m$ success

(1) Errors

~~if x^n not typical~~

if y^n not typical $\rightarrow$ err_1 ← not likely

if y^n is ~~jointly~~ in multiple cond. typ sets $\rightarrow$ err_2



→ want clouds ~~cond. typical~~ ~~sets~~ not to overlap.

(1) Rate ~~how?~~ how many!

~~→ Rate = how many messages cond~~

M = # messages

$\leq$ # of ~~d~~cond. typ spaces we can pack into

$$M \simeq \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$$

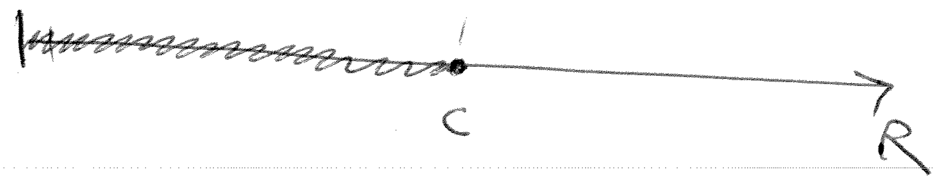
$$\text{R}\cancel{\text{ate}} \triangleq \frac{\log M}{n} = \frac{\log_2\left(2^{nI(X;Y)}\right)}{n} = I(X;Y)$$

→ note $p_X(x)$ was a free parameter.

~~Thm:~~ Channel coding Thm (Shannon '48)

The capacity of ~~Ros~~ a discrete memoryless channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ ~~the capacity~~ is

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



coding $\longrightarrow$ & "hand wavy exp"

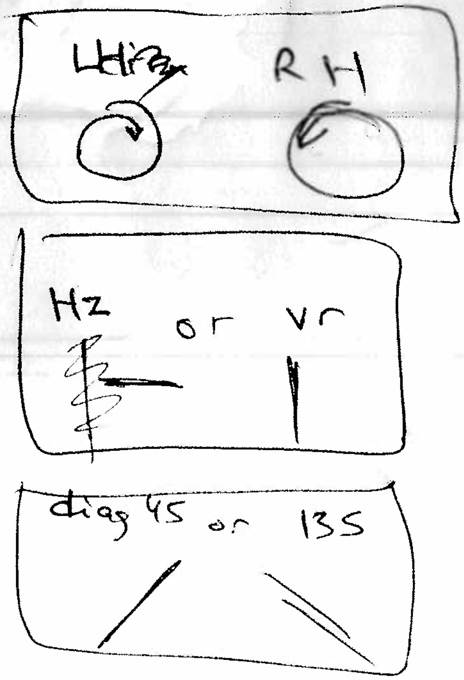
$\longleftarrow$ converse

(I) Demo

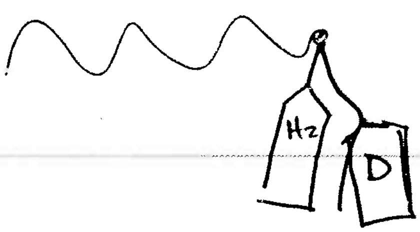
→ light 

→ explain polarizers

= property of light



H_0: Each photon has definite (Hz,V) polarization, and definite (D,AD)-polarization, and properties are independent



H_1: The $\overbrace{\text{polarization}}^{\text{joint}}$ is a QM phenomenon i.e. described by a $\vec{\text{vector}}$

$\vec{Hz} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\vec{V} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\vec{D} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\vec{AD} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, thus properties are not independent