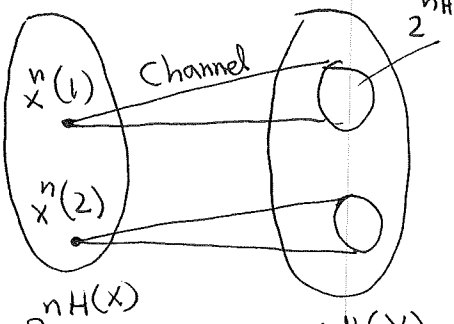
Lecture 8

* **Today:** (Weak) Converse
Source – Channel Coding Separation.
Constructing Codes for BEC/BSC.

Intuition for the Converse:

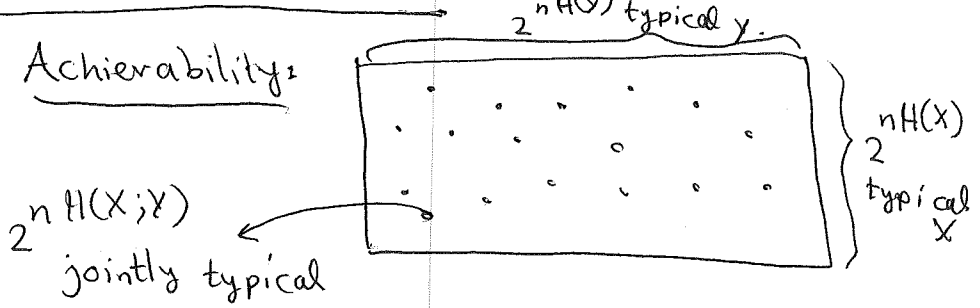


Packing:

$$\#\text{Codewords} \leq \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{nI(X;Y)}$$

$\Rightarrow$ Suggests: Capacity $\leq \max_{P(X)} I(X;Y)$.

Achievability:



$x^n(1)$ transmitted, y^n received.

if $x^n(2)$ & y^n are jointly typical $\Rightarrow Pr = \frac{2^{nH(X;Y)}}{\quad}$

Converse:

$\{1, \dots, 2^{Rn}\}$ → W → Enc → X^n → Chand → Y^n → $\hat{W}$ → Dec

If $P_e = Pr(\hat{W} \neq W) \to 0$ as $n \to \infty$,

Then $R \leq I(X;Y)$.

Pf: Warm up: Suppose $P_e = 0$.

$\Rightarrow W = \hat{W} = g(Y^n) \Leftrightarrow H(W|Y^n) = 0$.

(W uniform)

$$nR = H(W) = \underbrace{H(W|Y^n)}_{0} + I(W;Y^n)$$

$$= I(W;Y^n) \leq I(X^n;Y^n)$$

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

(Chain Rule) $$\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|Y_1 \dots Y_{i-1}, X^n)$$

(memoryless channel) $$= \sum_{i=1}^{n} H(Y_i) - \sum H(Y_i|X_i)$$

$$= \sum_{i=1}^{n} I(X_i;Y_i) \leq n \cdot C$$

Now, If P_e may be arbitrary,

$$P_e = P_r(\hat{W} \neq W)$$

Fano: $h(P_e) + P_e \log(2^{nR} - 1) \geq H(W|Y^n)$

$$\Rightarrow \quad 1 + nRP_e \geq H(W|Y^n).$$

Now just change the beginning with the above.

~~$nR \geq H(W) \geq H(W|Y^n) \geq \ldots$~~

~~$\Longrightarrow$~~

---

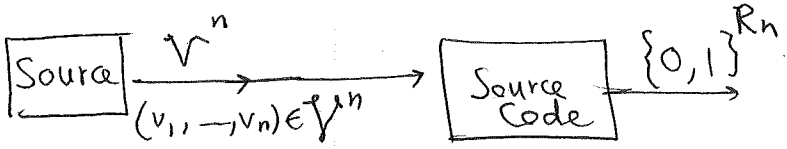
* Comment: $R < C \Rightarrow \exists$ code with

$$P_e \leq 2^{-\mathcal{C}(R,C) \cdot n}.$$

$R > C \Rightarrow \forall$ coding schemes, $P_e \to 1$.
as $n \to \infty$.

Joint Source-Channel Coding Thm:

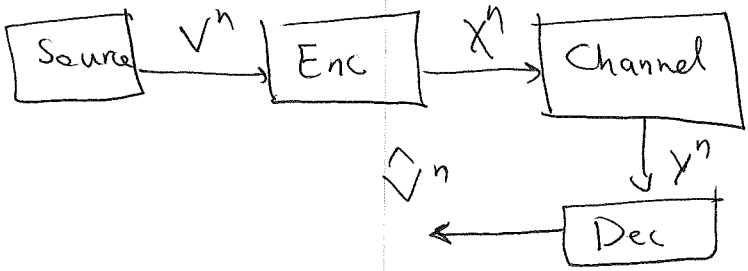
V source, $H := H(V)$



- Compression: $R > H$ is possible.
- Transmission: $R < C$ is possible.

Question: Is $C > H$ necessary & sufficient for Communication?

Using separate processes, $C > H$ is sufficient.



I.e., first use a source code to encode W. Then, pass through the Channel.

Is there a better "direct" code? No!

Namely,

Thm: If $V_1, \to V_n$ are iid from V,

Then $\exists$ a Source–Channel code

with $P_e = P_r(\hat{V}^n \neq V^n) \to 0$.

iff $C > H(V)$.

Pf: (in the book) via Fano's inequality.

(Converse)

$$n H(V) = H(V^n) = \underbrace{H(V^n \mid \hat{V}^n)}_{\to 0 \text{ (Fano)}} + I(V^n; \hat{V}^n)$$

$$\leq I(X^n; Y^n) \quad \text{(data proc.)}$$

$$(\leq nC)$$

□

* Explicit Codes.

We Consider BSC_p and BEC_α.
($C = 1 - h(p)$) ($C = 1 - \alpha$)

$$Enc: \{1, \ldots, 2^{Rn}\} \longrightarrow \{0,1\}^n$$

Linear maps: $k = Rn$ integer.

$$Enc: \{0,1\}^k \longrightarrow \{0,1\}^n \text{ linear.}$$

That is, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$.

$Enc(x) = G \cdot x \pmod 2$, G is $n \times k$ matrix.

Def: A code with a linear encoder is called a "linear code".
G is called the generator matrix.

Claim. For both BEC_α and BSC_p, we can pick a random G and this will also achieve capacity.

Joint typicality BSC_p: $(\underline{a}, \underline{b})$

$\underline{a}, \underline{b} \in \{0,1\}^n$.

$(a_1, \dots, a_n) \xrightarrow[BSC_p]{\text{Channel}} (b_1, \dots, b_n)$

$(\underline{a}, \underline{b})$ jointly typical if

$\Delta(\underline{a}, \underline{b}) \in (p \pm \varepsilon) n$.

Δ: Hamming Dist.

Decoding: Joint typical decoder:

Find $x \in \{0,1\}^k$ s.t.

$(p-\varepsilon)n \leq \Delta(G\underline{x}, \underline{y}) \leq (p+\varepsilon)n$.

Maximum Likelihood Decoding (for BSC$_p$)

Find $\underline{x}$ that minimizes $\Delta(G\underline{x}, \underline{y})$.

For BEC, decoding is easy!

$$\begin{pmatrix} & & \\ & G & \\ & & \end{pmatrix} \begin{pmatrix} x_1 \\ | \\ x_k \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ ? \\ 0 \\ 0 \\ 1 \\ ? \end{pmatrix} \leftarrow \approx qn \text{ ?'s}$$

linear system, solve it!

For a random G, solution is almost always unique! (Exercise).