March 12, 2019

REMINDERS

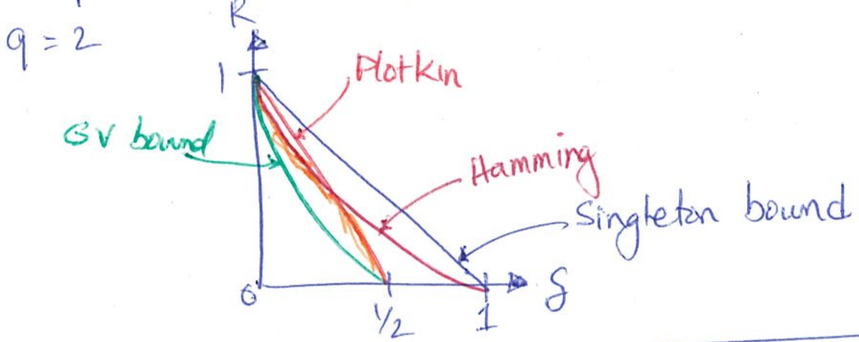
(1) HW 3 out this Th and due 2 weeks after that (Mar 28)
↳ Delay in HW 2 grading (during spring break)

(2) 2-page report due in 3 week (April 2)
↳ You can submit early & get comments sooner → email me after submission

(3) 2nd round of proof reading starts today
↳ Make sure you're aware when your 2nd proof reading is due

RECAP:

Big Q: Optimal tradeoff between R & δ. (recall: worst-case noise model due to Hamming)



TODAY: Proof reader → Amy

→ Revisit the worst-case noise model.

→ Noise model pioneered by Shannon ← things are pretty different compared to worst-case noise model

↳ Memoryless (ie noise model acts independently on each transmitted symbol)

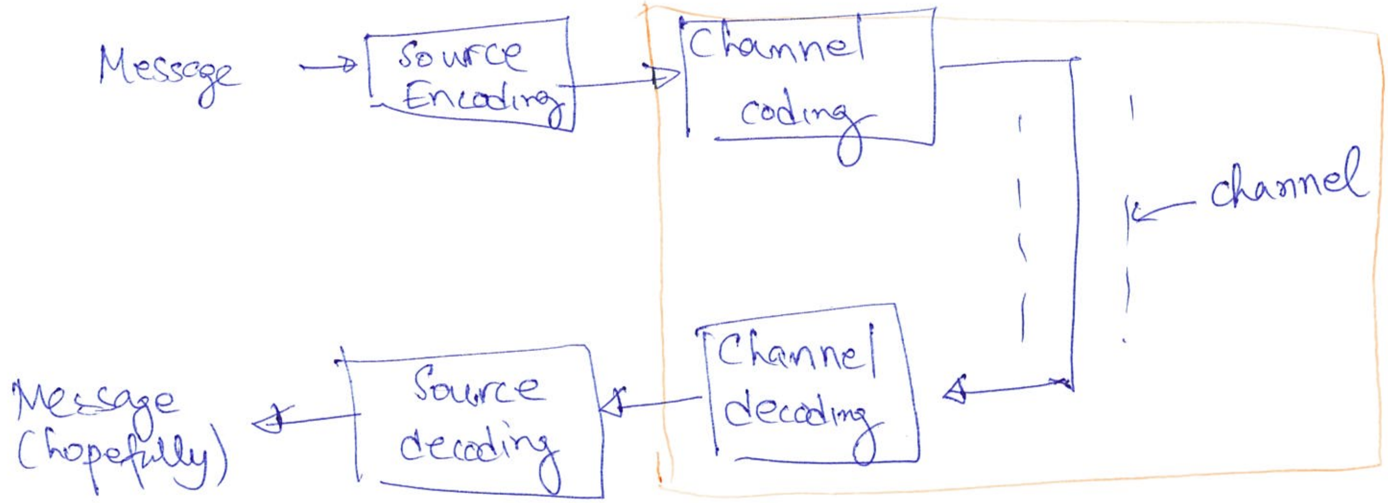
↳ (Fully specified) random noise.

Shannon's '48 paper:

(•) Noisy channel: (Channel coding) Our setup so far except noise is not worst-case (stochastic noise)

(•) Noiseless channel: (Source coding) No noise during transmission ⟹ do compression.

Generic setup:



→ In Shannon's setup can ~~decopt~~ decouple source and channel coding / decoding & optimize separately.
(More on channel coding next.)

Source coding: notion of entropy as the "right" notion of compressibility was introduced.

Shannon's noise model:

$$\mathcal{X} \ni x \xrightarrow[\text{channel}]{} y \in \mathcal{Y}$$

memoryless: same noise function acts on all the n symbols transmitted.

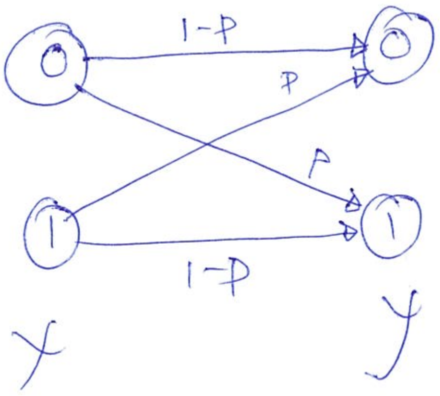
For us: $\mathcal{X}, \mathcal{Y}$: discrete.

Transition matrix $\mathcal{X} \ni x \rightarrow \begin{pmatrix} & \overset{y \in \mathcal{Y}}{\downarrow} & \\ \cdots & \cdot & \\ & & \end{pmatrix} = M$ $\quad M[x,y] = \Pr(y|x)$

① Binary Symmetric Channel (BSC) $0 \le p \le 1$ (BSC$_p$)

↑ crossover probability

$\mathcal{X} = \mathcal{Y} = \{0,1\}$



= each bit gets flipped w.p. p

Q: Assume $0 < p \le \frac{1}{2}$.

A: flip all bit received & use scheme for $0 \le p \le \frac{1}{2}$.

② q-ary symmetric channel (qSC_p) ($q=2 \rightarrow BSC_p$)

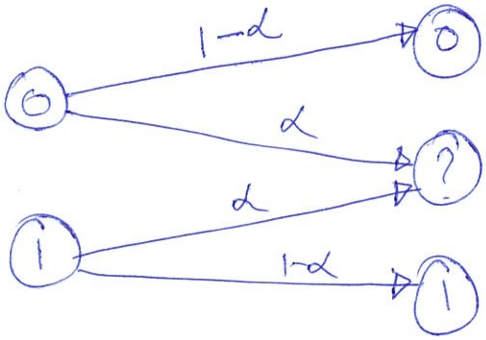
$0 \le p \le 1-\frac{1}{q}$ $\mathcal{X} = \mathcal{Y} = [q]$

$$M[x,y] = \begin{cases} 1-p & \text{if } y = x \\ \frac{p}{q-1} & y \neq x \end{cases}$$

So far $\mathcal{X} = \mathcal{Y}$

③ Binary erasure channel (BEC_α) $0 < \alpha \le 1$

$\mathcal{X} = \{0,1\}$; $\mathcal{Y} = \{0,1,?\}$



i.e each bit independant gets flipped with prob α.

Error correction: (BSC_p) → some non-zero prob that transmitted codeword gets changed to another codeword.

BEC_α → some non-zero prob that all symbols got erased.

WANT: Every message is decoded correctly w.p. $1-f(n)$

$\lim_{n\to\infty} f(n) = 0$ Ideally: $f(n) = 2^{-\Omega(n)}$

Shannon's general result

Big Q! R vs errors → error parameter

p: qSC_p

α: BEC_α

Shannon's thm: For every channel there is a specific $0 < C \le 1$ s.t $\forall$ all $R < C \Rightarrow$ reliable communication.

& for any $R > C \Rightarrow$ ——— is not possible

C: capacity of channel.

Next: Above for BSC_p

$\bar{e} \sim BSC_p$ $\quad \bar{e} \in \{0,1\}^n$

Shannon's Capacity thm for BSCp

$0 \le p < \frac{1}{2}$, $0 < \varepsilon \le \frac{1}{2} - p$

following are true for large enough n:

① $\exists\ \delta > 0$, an encoding function $E : \{0,1\}^k \longrightarrow \{0,1\}^n$
a decoding function $D : \{0,1\}^n \longrightarrow \{0,1\}^k$
for $k \le \lfloor (1 - H(p) - \varepsilon) n \rfloor$ s.t

$\forall \bar{m} \{0,1\}^k$ $\quad \Pr_{\bar{e} \sim BSC_p} \left[D\left(E(\bar{m}) + \bar{e} \right) \neq \bar{m} \right] \le 2^{-\delta n}$.

② If $k \ge \lceil (1 - H(p) + \varepsilon) n \rceil$ then $\forall E : \{0,1\}^k \longrightarrow \{0,1\}^n$
$D : \{0,1\}^n \longrightarrow \{0,1\}^k$
$\exists\ \bar{m} \in \{0,1\}^k$ s.t

$\Pr_{\bar{e} \sim BSC_p} \left[D(E(\bar{m}) + \bar{e}) \neq m \right] \ge \frac{1}{2}$ ← can make this $1 - 2^{-\beta n}$.

$\Rightarrow$ Capacity of BSCp : $1 - H(p)$
qSCp : $1 - H_q(p)$
BEC$_\alpha$: $1 - \alpha$

} same reason as GV bound: upper/lower bounds for $Vol_q(pn, n)$

Pf (sketch) of part ①

Via probabilistic method: pick E at random:
$\forall \bar{m} \in \{0,1\}^k$ $E(\bar{m})$ is uniformly random $\in \{0,1\}^n$
↳ independent choices for different $\bar{m}$.

D : MLD_E

2 steps:

Step 1: For any $\bar{m} \in \{0,1\}^k$, for random E prob of decoding error is $2^{-\Omega(n)}$
$\Rightarrow \exists$ a good E for a fixed $\bar{m}$

Step 2: show similar result $\forall \bar{m} \in \{0,1\}^n$.
Involves cropping $\overset{\exists E \text{ s.t}}{\text{half of}}$ codewords $(k \to k-1)$

2 sources of randomness:

① Choice of E ← for prob. method

② Randomness in BSC_p noise → contributes to decoding error prob.