Continuity property (see TV properties, property 1 in page 59)

$$\mathbb{E}_e\left[\mathbb{E}_P\left[d(s^n,\hat{s}^n)\right]\right] \leq \mathbb{E}_e\left[\mathbb{E}_Q\, d(s^n,\hat{s}^n) + d_{max}\,\|P-Q\|_{TV}\right]$$

$$= \mathbb{E}_{\mathbb{E}_e Q}\, d(s^n,\hat{s}^n) + d_{max}\,\mathbb{E}_e\|P-Q\|_{TV}$$

$\mathbb{E}_e\|P-Q\|_{TV} \to 0$ by Soft Covering Lemma!

$$\mathbb{E}_e Q_{s^n\hat{s}^n}(s^n,\hat{s}^n) = \prod_{i=1}^{n} \bar{P}_{S\hat{S}}(s_i,\hat{s}_i)$$

where $Q_{\hat{S}^n|M}$ is a look up (deterministic)

Thus "ideal" $\to$ $\mathbb{E}_{\bar{P}}\, d(S,\hat{S})$

Choose a $\bar{P}_{\hat{S}|S}$ s.t.

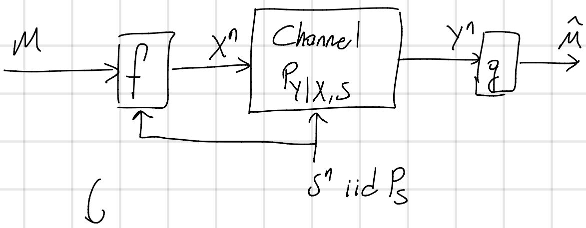
$$\mathbb{E}_{\bar{P}}\left[d(S,\hat{S})\right] < D$$

and

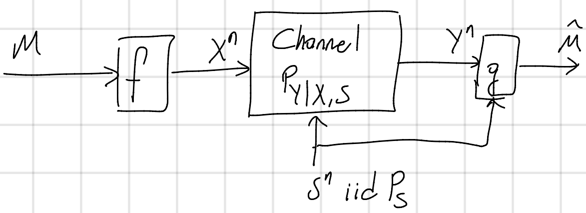
$$R < I_{\bar{P}}(S;\hat{S})$$

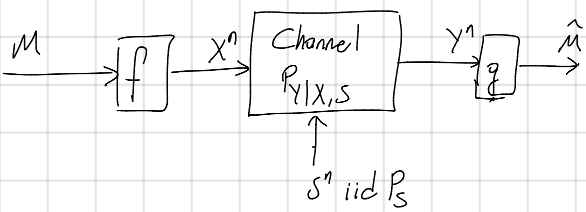
11/15/2016
Tuesday

Gelfand-Pinsker (1980)



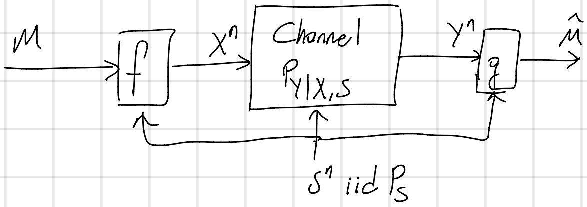
This is the most general b/c you may replace $P_{Y|X,S} \leftarrow P_{Y,S|X,S}$ this gives



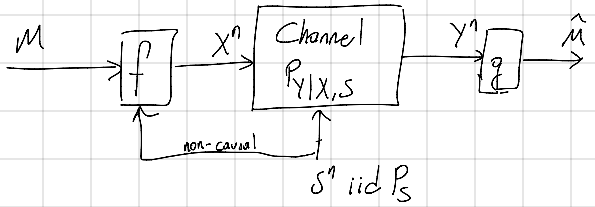


$\longrightarrow$ Capacity

$$C = \max_{P_X} I(X;Y)$$



$\longrightarrow \max_{P_{X|S}} I(X;Y|S)$



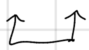
$\longrightarrow$ the answer is not the same as above!

Example 1: Write twice to optical disc.

in second time you use you can change the 0s to 1's but not o.w. around

in here encoder gets non-causal info (you can read the whole disk) but decoder has no info.

Example 2: Gaussian channel w/ known noise (e.g. stray interference)

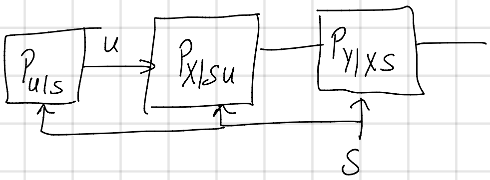
$\longrightarrow$ Ex 1 and 2 are equiv. in  and  versions.

Thm:

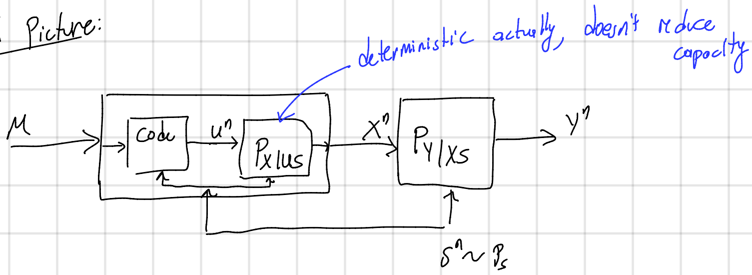
$$C = \max_{P_{U X|S}} I(U;Y) - I(U;S)$$

Again uses synthetic channel at encoder

↳ block diag. of Thm: (Not the problem)



Operational Picture:



Let's see Thm in action first

In Example 1: Channel: $Y = X \vee S$ $S \sim Ber(p)$

Let $U = X = \begin{cases} 1 & \text{if } S=1 \\ Z & \text{if } S=0 \end{cases}$ where $Z \sim Ber(1/2)$ $Z \perp\!\!\!\perp S$

i.e. $P_{U=1|S=1} = 1$ (as $X = U$ we can say $Y = X \vee S = U \vee S$)

$P_{U=1|S=0} = 1/2$

Therefore, $Y = U$ (see $Y = U \vee S$ when $S=0$ ✓ and $S=1$, $U=1$ and $Y=1$ ✓)

$I(Y;U) = H(U)$

$$I(Y;U) - I(U;S) = H(U) - I(U;S)$$

$$= H(U|S)$$

$$= \mathbb{P}[S=0] = \boxed{1-p}$$

→ This actually maximises $I(U;Y) - I(U;S)$, just consider that decoder also has S^n, then it is $1-p$ because you can ignore erased part

Example 2: "Writing on Dirty Paper" (Costa): (Channel: $Y = X + S + N$, $S \perp\!\!\!\perp N$) both Gaussian

Choose U, X, S to be Jointly Gaussian. and $X \perp\!\!\!\perp S$ and $H(U|X,S) = 0$

i.e. → Optimize U: WLOG $U = S + \lambda X$ + ~~Noise~~

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$$

N ↳ noise power without state ($N = \sigma^2$ here)

Gelfand – Pinsker

G-P Achievability proof

Start by choosing $\bar{P}_{U|X|S}$ s.t. $R < I_{\bar{P}}(U;Y) - I_{\bar{P}}(U;S)$

Let $\bar{P} = \bar{P}_{SXUY} = P_S \bar{P}_{UX|S} P_{Y|SX}$

$\bar{P}_{S^n X^n U^n Y^n} = \prod \bar{P}_{SXUY}$

<u>Encoder</u>: Transmit message and padding

m at rate R	m_p at rate R_p
Message	padding

↑ choose carefully

use padding to make transmission correlated with state

Choose $R_p \in (I(U,S), I(U;Y) - R)$

$\underbrace{R+R_p}_{\text{total rate}} < I(U;Y)$ (so that we can decode it by channel coding)

Random Codebook $\sim \bar{P}_U$: $\mathcal{C}\{u^n(m, m_p)\}$

Proof using Joint typicality encoder: We are suppressing X in the entire discussion below

Encoder: $P_{U^n|M,S^n}$ — Given m and S^n, find m_p s.t. $(U^n(m, m_p), S^n) \in \mathcal{J}_\varepsilon^{(n)}$ wrt $\bar{P}$

<u>Decoder</u>: Find unique m s.t. $\exists m_p$ s.t. $(u^n(m, m_p), Y^n) \in \mathcal{J}_{\varepsilon'}^{(n)}$ where $\varepsilon' > \varepsilon$.

⇝ <u>Error analysis</u>:

<u>Encoding error</u>: Small because $R_p > I(U;S)$

<u>Decoding error</u>:

~~For $(m, m_p) \neq$ truth $u^n(m, m_p) \perp\!\!\!\perp Y^n$~~ not true

For $m \neq$ truth $U^n(m, m_p) \perp\!\!\!\perp Y^n$

small error because $R + R_p < I(U;Y)$

↓

Proof with Likelihood encoder:

Encoder $P_{U^n|MS^n}$

Given m and S^n: $P_{M_p|MS^n} \propto \bar{P}_{S^n|U^n}(s^n|u^n(m,m_p))$

$U^n = u^n(m, m_p)$ ← look-up

Induced distribution $P_M P_{S^n} P_{M_p|S^nM} P_{U^n|MM_p} P_{X^n|U^nS^n} P_{Y^n|X^nS^n} P_{\hat{M}|Y^n}$

(enc. = $P_{M_p|S^nM} P_{U^n|MM_p} P_{X^n|U^nS^n}$)

- P_M: Unif
- P_{S^n}: $\prod P_S$
- $P_{M_p|S^nM}$: Likelihood encoder
- $P_{U^n|MM_p}$: lookup
- $P_{X^n|U^nS^n}$: Synthetic channel
- $P_{Y^n|X^nS^n}$: actual channel
- $P_{\hat{M}|Y^n}$: decoder

$P_{S^n} P_{M_p|S^nM}$ → $P_{S^nM_p|M}$ (we're just gonna change that part)

Ideal:

For each m Let $Q_{S^nM_p|M=m} = Q_{M_p|\cancel{M}} \, Q_{S^n|MM_p}$

- Q_{M_p}: unif. induced of M
- $Q_{S^n|MM_p}$: look up $u^n(m,m_p)$, $\bar{P}_{S^n|U^n}(s^n|u^n(m,m_p))$

$$\mathbb{E}_{e}\left[\| P_{S^nM_p|M=m} - Q_{S^nM_p|M=m} \|_{TV}\right] \to 0 \quad \forall m$$

because $R_p > I(U;S)$ (soft covering lemma)

now use property of TV: same channel, different input → output of channel is close!