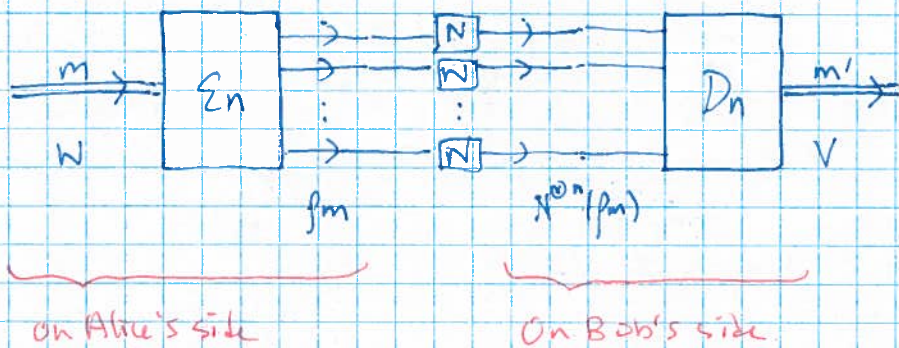
CO781 / QIC890 Lec 14 Nov 01, 2016

Classical capacity of a quantum channel.

Most general protocol:



NB : the above is a Q-box $\xrightarrow{m}$ $\square$ — $N^{\otimes n}(\rho_m)$

① Achievable rate $\geq \max_{\rho_m} \max_{p_m} \chi(\{p_m, N^{\otimes n}(\rho_m)\})$

② $nR \lesssim I(W, V) \leq \max_{\rho_m} \max_{p_m} I_{acc}(\{p_m, N^{\otimes n}(\rho_m)\})$

↑ Fano's ineq

$\leq \max_{\rho_m} \max_{p_m} \chi(\{p_m, N^{\otimes n}(\rho_m)\})$

Thm (Holevo – Schumacher – Westmoreland)* (HSW Thm)

$$C(N) = \sup_r \chi^{(r)}(N)$$

where $\chi^{(r)}(N) = \frac{1}{r}\chi(N^{\otimes r})$ $\longleftarrow \frac{1}{r}\max_{\{p_x, \rho_x\}} \chi(\{p_x, N^{\otimes r}(\rho_x)\})$

$$\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$$

NB Optimize over r, ρ_x (over r input spaces), p_x.

The capacity expression is called "regularized", with an optimization over r and optimizing p_x, ρ_x over r uses of N.

This is in contrast to the capacity expression for classical channels or Q-boxes.

NB When proving converse for $C(Q)$

$$S(B_1 B_2 \ldots B_n | X_1 X_2 \ldots X_n)$$

$$= \sum_{x_1 \cdots x_n} p(x_1 \ldots x_n)\, S(\rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n})$$

$$= \sum_{x_1 \cdots x_n} p(x_1 \ldots x_n) \sum_i S(\rho_{x_i})$$

not the case for the HSW thm

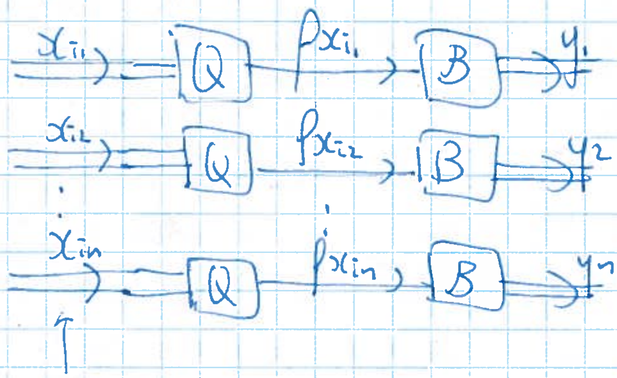
* Holevo: IEEE TIT 44 p269 (1998)

Schumacher & Westmoreland: PRA 56 p131 (1997)

Comparisons:

To send a message i:

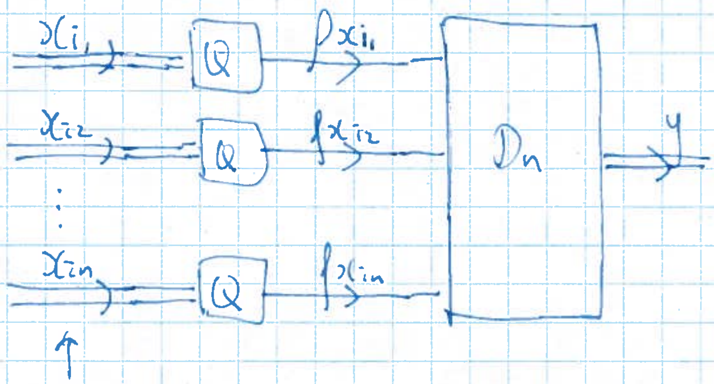
(1)



$C_i = x_{i_1} x_{i_2} \cdots x_{i_n}$, $x_{ij} \sim p(x)$

Capacity $= \underset{p_x}{\max}\, I_{acc}(\{p_x, \rho_x\})$ given

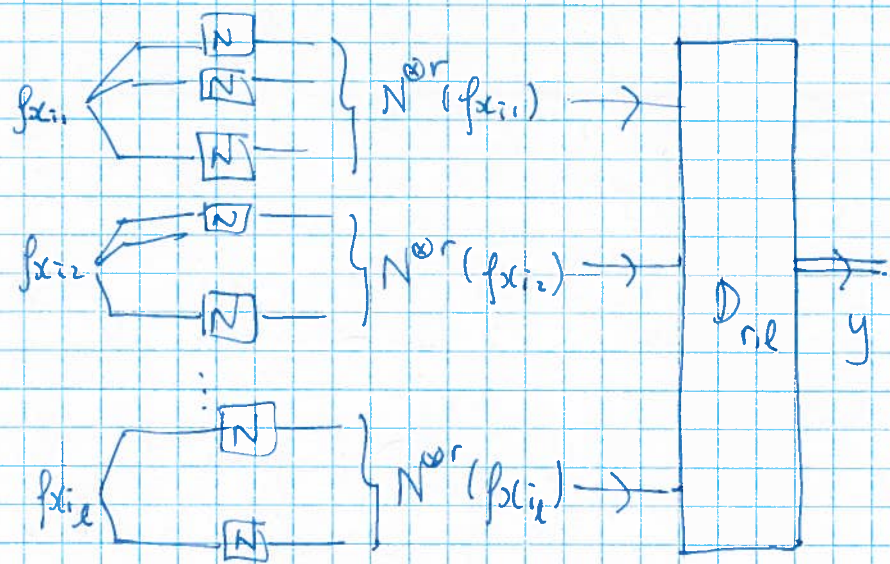
(2)



$C_i = x_{i_1} x_{i_2} \cdots x_{i_n}$, $x_{ij} \sim p(x)$, restrict to strongly typical sequences

Capacity $= C(Q) = \underset{p_x}{\max}\, \chi(\{p_x, \rho_x\})$

3



$C_i = x_{i_1} \cdots x_{i_\ell}$ $\quad x_{i_j} \sim p_{i_j}$, restrict to strongly typical sequences

Capacity $C(N) = \sup_r \frac{1}{r} \max_{\{p_x, \rho_x\}} \chi(p_x, N^{\otimes r}(\rho_x))$

Consequences of the HSW theorem:

① $C(N) = \sup_r \chi^{(r)}(N) \geq \chi^{(r)}(N) \geq \chi^{(1)}(N)$ ⟵ useful lower bound

② $C(N) = 0 \iff \chi^{(1)}(N) = 0$

[$\Rightarrow$] obvious

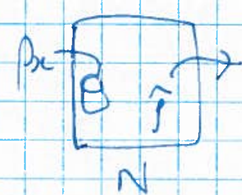
[$\Leftarrow$] $\chi^{(1)}(N) = \max_{p_x, \rho_x} S(X:B)$

$\Lambda = \sum_x p_x |x\rangle\langle x| \otimes N(\rho_x)$

$= 0 \iff \forall\, p_x, \rho_x,\ \Lambda$ product state

$\iff \forall\, \rho_x,\ N(\rho_x)$ indep of x

$\iff N$ trivial, Bob can create output $N(\rho_x)$ without x, N, or Alice



③ $C(N)$ continuous in N. (Hope we have time to prove it, or project?)

④ Is $\chi^{(1)}(N) = \chi^{(r)}(N)$? Depends on the channel. Will discuss more.

⑤ How to find $\chi^{(1)}(N)$?

Def: In the expression $\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$

$\{p_x, \rho_x\}$ is called the "optimal ensemble" for N if it achieves the max.

Properties of the optimal ensemble:

Uhlmann 9701014, Schumacher & Westmoreland 9912122

1. ρ_x's can be chosen pure
2. d^2 pure states are sufficient, where d = output dim of N

Pf: A3 (similar to Iacc)

eg 1 $\mathcal{N}_p$ = d-dim depolarizing channel

$$\mathcal{N}_p(\rho) = (1-p)\rho + p\frac{I}{d}$$

$$\chi^{(1)}(\mathcal{N}_p) = \max_{p_x, |\psi_x\rangle} S\Big(\sum_x p_x \mathcal{N}(|\psi_x\rangle\langle\psi_x|)\Big) - \sum_x p_x S\big(\mathcal{N}(|\psi_x\rangle\langle\psi_x|)\big)$$

For the 2nd term:

$$\mathcal{N}(|\psi_x\rangle\langle\psi_x|) = (1-p)|\psi_x\rangle\langle\psi_x| + p\frac{I}{d}.$$

$$\text{Spectrum} = \Big(1-p+\frac{p}{d}, \underbrace{\frac{p}{d}, \dots \frac{p}{d}}_{d-1 \text{ times}}\Big) \quad \text{independent of } |\psi\rangle.$$

$$S(\mathcal{N}(|\psi_x\rangle\langle\psi_x|)) = -\Big(1-p+\frac{p}{d}\Big)\log\Big(1-p+\frac{p}{d}\Big) - \frac{d-1}{d}p\log\frac{p}{d} =: s$$

$$\therefore \chi^{(1)}(\mathcal{N}_p) = \underbrace{\max_{p_x, |\psi_x\rangle} S\Big(\sum_x p_x \mathcal{N}(|\psi_x\rangle\langle\psi_x|)\Big)}_{\text{attained when } \sum_x p_x \mathcal{N}(|\psi_x\rangle\langle\psi_x|) = \frac{I}{d}} - s$$

when $p_x = \frac{1}{d}$, $|\psi_x\rangle = |x\rangle$.

$$= \log d - s.$$

Cor $d=2$, $\chi^{(1)}(\mathcal{N}_p) = \underbrace{1 - h\big(\tfrac{p}{2}\big)}_{\shortparallel}$ (will see this is $C(\mathcal{N}_p)$)

capacity of classical BSC with flipping prob $\frac{p}{2}$.

What is the code for achieving this? (d-dim case)

Draw $\approx 2^{n\chi}$ code words from all possible n-tuples of symbols in $\{1, 2, \dots, d\}$ and restrict to strongly-typical sequences.

Due to overlaps of the out puts, expect T entangling & so is the decoding measurements.

In general, if ① N unital $\quad (N(I)=I)$

② each $|\psi_i\rangle$ minimizes $S(N(|\psi_i\rangle\langle\psi_i|)) =: S$

③ $\frac{I}{d} \in \mathrm{conv}(|\psi_i\rangle)$

then $\{p_i, |\psi_i\rangle\}$ optimal ensemble and $\chi^{(1)}(N) = \log d - S$

where $\sum_i p_i |\psi_i\rangle\langle\psi_i| = \frac{I}{d}$.

Ex: find $\chi^{(1)}(N)$ for $N(\rho) = 0.8\rho + 0.15\, X\rho X + 0.05\, Z\rho Z$

Ex: find $\chi^{(1)}(E_p)$ for $E_p(\rho) = (1-p)\rho + p\, \underbrace{|E\rangle\langle E|}_{\text{error state ortho to input space}}$

d-dim erasure channel

erasure prob

Compare $\chi(E_p)$ to the classical analogue.

Note: E_p not unital.

Properties (ctd):

(3) The most distinguishable inputs need not be optimal.

ie non ortho ρ_x's may be needed.

- Fuchs PRL 79 (1997) 1162 found the first channel that requires non orthogonal optimal ensemble
- Amplitude damping channel shares the same property (9912122).

<u>Amplitude damping channel:</u>

$$N(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$$

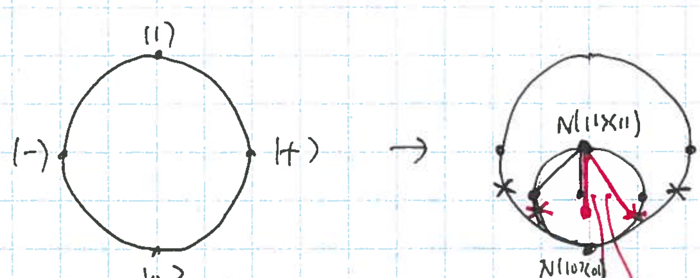
$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}$$

$$a|0\rangle + b|1\rangle \xrightarrow{A_0} a|0\rangle + \sqrt{1-\gamma}\, b|1\rangle$$

$$a|0\rangle + b|1\rangle \xrightarrow{A_1} \sqrt{\gamma}\, b|0\rangle$$

If $|0\rangle, |1\rangle$ are energy eigenstates,
(ground, excited)

then AD represents de-excitation.



distance from original should be small for the average, large for individual.

$\gamma = \frac{1}{2}$

- Restricting to $|\psi_0\rangle$ & $|\psi_1\rangle$, max @ $p_0 = p_1 = \frac{1}{2}$, $\chi = 0.4567$.
- Instead, $\chi = 0.4717$ for some $|\psi_0'\rangle$, $|\psi_1'\rangle$ with $\langle\psi_0'|\psi_1'\rangle \approx \cos 80^\circ$.

How hard is it to calculate or estimate $\chi^{(1)}(N)$?

Beigi & Shor 0707.2090:

Let $C \in \mathbb{R}$. To decide whether $\chi^{(1)}(N) > C$ or $< C-\varepsilon$ for $\varepsilon = \frac{1}{\mathrm{poly}(d)}$ (d = input dim) is NP complete.

NB The above does not imply hardness of estimating $\chi(N^{\otimes n})$.

more structure