Lec 8, June 1, 2010

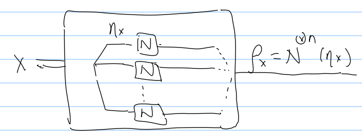
Note Title 31/05/2010

Last time:

$C(Q) = \max_{p_x} \chi(\{p_x, \rho_x\})$ where $\xrightarrow{x} \boxed{Q} \rho_x$

$C(N) = \max_{p_x, \eta_x} \frac{1}{n} \chi(\{p_x, N^{\otimes n}(\eta_x)\})$

interpreted as :



HSW Thm (after Holevo, Schumacher & Westmoreland)
Proved ≈ 1997.

Convenient to define

$$\chi(N) = \chi^{(1)}(N) = \max_{\{p_x, \rho_x\}} \chi(\underbrace{\{p_x, N(\rho_x)\}}_{\text{ensemble received by Bob after 1 use of channel}})$$

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

Thus $C(N) = \sup_n \chi^{(n)}(N) \geq \chi^{(n)}(N) \geq \chi^{(1)}(N)$

Consequences:

1. $C(N) = 0$ iff $\chi^{(1)}(N) = 0$
$\quad \| $
QMI of $\left[\sum_x p_x |x\rangle\langle x| \otimes N(\rho_x)\right]$

Recall $S(A:B) = 0$ iff A, B in a product state

$\therefore C(N) = 0$ iff $N(\rho_x)$ independent of x
for any choice of $\{p_x, \rho_x\}$

iff $N(\rho) = \rho_0$ const

N completely useless

(Bob can simulate N by creating ρ_0 on his side!)

2. We can prove continuity of $C(N)$ (as a function of N).
(Tricky for a regularized expression, but doable.) $\boxed{\text{later in the course}}$
So similar channels have similar capacities.

3. $\chi^{(1)}(N)$ is a useful lower bound to $C(N)$.
Will calculate this for several channels (this lecture).

4. For some special channels, $\chi^{(1)}(N) = \chi^{(n)}(N)$!
So we know $C(N)$ for these channels (this lecture).

5. Up until 2008, the hope was that $\chi^{(1)}(N) = \chi^{(n)}(N)$
for ALL channels, but Hastings gave a counterexample.

Will discuss these "additivity conjectures" (this lecture)

July 8th - Fukuda will give a guest lecture on the ideas behind Hasting's counterexample :)

Finding $\chi^{(1)}(N) = \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$

$$= \max_{\{p_x, \rho_x\}} \underbrace{S\left(\sum_x p_x N(\rho_x)\right)}_{\text{max this entropy}} - \underbrace{\sum_x p_x S(N(\rho_x))}_{\text{min these entropies}}$$

Reference: 9912122 Schumacher & Westmoreland

1. ρ_x can be chosen to be pure (elementary argument)
2. Suffices to have d^2 states in the ensemble where d = dim of input to N
(argument similar to those in Davies, IEEE TIT 24 (1978) 596)

eg1. N: d-dim erasure channel

$N(\rho) = (1-p)\rho + p|E\rangle\langle E|$ for $|E\rangle$ error symbol

Consider any input ensemble $\{p_x, \rho_x\}$ (ρ_x pure):

$\forall x,\ N(\rho_x) = (1-p)\rho_x + p|E\rangle\langle E|$ and $\boxed{S(N(\rho_x)) = H(p)}$ (1) no need to min

Average output state $= \sum_x p_x N(\rho_x) = N\left(\sum_x p_x \rho_x\right)$
$= (1-p)\left(\sum_x p_x \rho_x\right) + p|E\rangle\langle E|$

$$\boxed{S\left(\sum_x p_x N(\rho_x)\right) = (1-p)\, S\left(\sum_x p_x \rho_x\right) + p\, \cancel{S(|E\rangle\langle E|)} + H(p)} \quad (2)$$

max by $\sum_x p_x \rho_x = \frac{I}{d}$ eg. $p_x = \frac{1}{d}$, $\rho_x = |x\rangle\langle x|$

$\therefore \chi^{(1)}(N) = \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\}) = (1) - \sum_x p_x (2) = (1-p)\log d$. (cf classical case) eg2. N = d-dim depolarizing channel

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

For any pure state ensemble $\{p_x, \rho_x\}$

$$\boxed{\begin{array}{l} S(N(\rho_x)) = \text{Entropy of} \left\{ \left(1-p+\frac{p}{d}\right), \frac{p}{d}, \dots \frac{p}{d} \right\} \\ \quad \text{again independent of } \rho_x\text{, no need to minimize.} \end{array}} \quad ①$$

$\sum_x p_x N(\rho_x) = (1-p)\left(\sum_x p_x \rho_x\right) + p\frac{I}{d}$

Again $\underbrace{S\left(\sum_x p_x N(\rho_x)\right)}$ max by $\sum_x p_x \rho_x = \frac{I}{d}$

$p_x = \frac{1}{d}, \ \rho_x = |x\rangle\langle x|$

$\chi^{(1)}(N) = \log d - ①$

$= \log d - \left(1-p+\frac{p}{d}\right)\log\left(1-p+\frac{p}{d}\right) - \left(\frac{d-1}{d}\right)p\left[\log\frac{p}{d}\right] \dots$

When $d=2$, $\chi^{(1)}(N) = 1 - H\left(\frac{p}{2}\right)$ (cf Classical BSC)

eg3 $N(\rho) = 0.8\rho + 0.15 X\rho X + 0.05 Z\rho Z$

2 dim input & output, X, Z : Pauli's.

① $S(N(\rho))$ minimized by $\rho = |+\rangle\langle+|$ or $|-\rangle\langle-|$.

$S(N(|+\rangle\langle+|)) = H(0.05)$

② $S\left(\sum_x p_x N(\rho_x)\right)$ maximized by $\sum_x p_x N(\rho_x) = \frac{I}{2}$

Since $\sum_x p_x N(\rho_x) = N\left(\sum_x p_x \rho_x\right)$ and $N(I) = I$ (unital)

choose $\rho_0 = |+\rangle\langle+|$ & $\rho_1 = |-\rangle\langle-|$ with equal probs

$\therefore \chi^{(1)}(N) = 1 - H(0.05)$

In general, if N unital & there is a set of states $|\psi_i\rangle$

each $|\psi_i\rangle$ minimized $S(N(|\psi_i\rangle\langle\psi_i|))$ and

the convex hull of $|\psi_i\rangle\langle\psi_i|$ includes $\frac{I}{d}$

ie $\exists p_i$'s s.t. $\sum_i p_i |\psi_i\rangle\langle\psi_i| = \frac{I}{d}$

then $\{p_i, |\psi_i\rangle\}$ is an optimal ensemble for $\chi^{(1)}(N)$.

Qn: are orthogonal ρ_x always optimal?

They're after all the most distinguishable inputs.

Ans: No .... by Fuchs PRL 79 (1997) 1162.

There, a channel was cooked up for the demonstration

It turns out the case for the amplitude damping

channel also (9912122).

These Non orthogonal signal states are considered a

"Quantum" advantage...

eg4. Amplitude damping (AD) channel

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

$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}$	$A_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}$
$A_0(a|0\rangle + b|1\rangle)$ $= a|0\rangle + \sqrt{1-\gamma}\,|1\rangle$	$A_1(a|0\rangle + b|1\rangle)$ $= \sqrt{\gamma}\, b|0\rangle$.

$|0\rangle, |1\rangle$ are energy eigenstates and $|1\rangle$ is at a

higher energy level than $|0\rangle$.

AD represents the physical process of deexcitation.

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

- Restricting ρ_x to orthogonal states: $\chi(\{p_x, \rho_x\})$ max at $p_0 = p_1 = \frac{1}{2}$, $\rho_0 = \underset{|\psi_0\rangle}{|+\rangle}\langle+|$, $\rho_1 = \underset{|\psi_1\rangle}{|-\rangle}\langle-|$

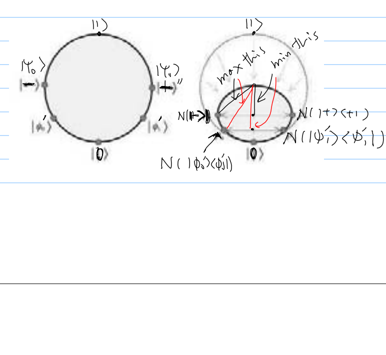
and $\chi = 0.4567$

- If instead, $|\psi_0'\rangle = a|0\rangle + b|1\rangle$, $|\psi_1'\rangle = a|0\rangle - b|1\rangle$

for a slightly larger than b ($\langle\psi_0'|\psi_1'\rangle \approx \cos 80°$)

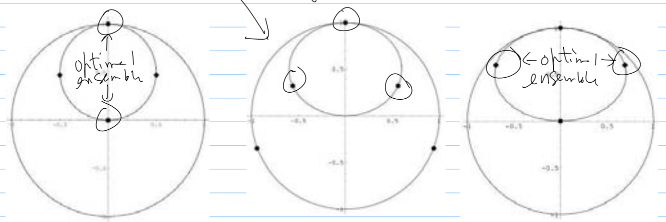
$\chi = 0.4717$

Intuition: "$|0\rangle$" state suffers less AD.



Qn: for $d=2$, will we need 3 or 4 states in the optimal ensemble?

0109079 King, Nathanson, & Ruskai: channel requiring 3 states



① Depolarizing + shift ② AD channel $\gamma = \frac{1}{2}$

4 states needed in 0403176

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

Ans: (Beigi & Shor 0709.2090)

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

NB this says little about hardness of estimating $\chi^{(n)}$ since tensor product channels $N^{\otimes n}$ can be easier to handle.

Back to $C(N)$:

We only know $C(N)$ if $\chi^{(1)}(N) = \chi^{(n)}(N)$.

Def: We say χ is <u>strongly</u> additive if $\forall N_1, N_2$, $\chi^{(1)}(N_1 \otimes N_2) = \chi^{(1)}(N_1) + \chi^{(2)}(N_2)$

Def: We say χ is <u>weakly</u> additive if $\forall N$, $\chi^{(n)}(N) = \chi^{(1)}(N)$

NB: strongly additive $\Rightarrow$ weakly additive

NB: "$\geq$" easily proved. The question is whether it can be $\gneq$. If we restrict to product states on the LHS, then $=$.

Known additivity results:

① If N_1 is entanglement breaking then $\forall N_2$, $\chi^{(1)}(N_1 \otimes N_2) = \chi^{(1)}(N_1) + \chi^{(1)}(N_2)$ (0302031?)

Cor: If N is entanglement breaking, then $C(N) = \chi^{(1)}(N)$.

Def: N is entanglement breaking if $\forall \rho_{AB}$, $I \otimes N(\rho_{AB})$ is separable.

$\boxed{N}$ ← not entangled with anything.

<u>Fact:</u> N ent breaking iff N = "QCQ" channel

ie $\rightarrow \boxed{N} \rightarrow \; = \; \rightarrow \boxed{M} \xrightarrow{x} \boxed{Q} \xrightarrow{\rho_x}$

Intuition: superadditivity of χ is due to entanglement between the inputs to the 2 channels. So this is expected....

Also, classical channels are entanglement breaking. So we recover Shannon's noisy coding theorem & the single letter capacity expression.

---

② If N_1 has 2-dim inputs & outputs <u>& N_1 unital</u> then $\forall N_2$, $\chi^{(1)}(N_1 \otimes N_2) = \chi^{(1)}(N_1) + \chi^{(2)}(N_2)$

Cor: If N_1 is a qubit unital channel, then $C(N) = \chi^{(1)}(N)$

(King 0103156)

NB: These include eg 2 & 3 (so our calculation of $\chi^{(1)}(N)$ gives $C(N)$ as well!).

Fact: N_1 is a qubit unital channel iff $\rightarrow \boxed{N_1} \rightarrow \; = \; \rightarrow \boxed{U} - \boxed{\Sigma} - \boxed{V} \rightarrow$

where U, V unitary & $\Sigma(\rho) = \sum_i p_i \sigma_i \rho \sigma_i$ ← Pauli's

③ If N_1 = d-dim depolarizing channel, then

$\forall N_2,\ \chi^{(1)}(N_1 \otimes N_2) = \chi^{(1)}(N_1) + \chi^{(1)}(N_2)$

⋮

Same consequences as before

④ for N = erasure channel,

I think $C(N) = \chi^{(1)}(N) = (1-p)\log d$.

⑤ $\chi^{(1)}(N_1 \otimes N_2)$ additive $\forall$ channels $N_1 N_2$

iff – – – – – $\forall N_2$ & $\forall N_1$ unital.

(Fukuda 0608010)

⑥ $\chi^{(1)}$ strongly additive ($\forall$ channels)

$\Updownarrow$

$\chi^{(1)}$ weakly additive ($\forall$ chanels).

⑦ $\exists\ N_1, N_2$ s.t.

$$\chi^{(1)}(N_1 \otimes N_2) \underset{\neq}{>} \chi^{(1)}(N_1) + \chi^{(2)}(N_2).$$

Hastings 0809.3972.

So, how bad is the nonadditivity? we don't know.

But we don't know $C(N)$ for N = AD channel

In fact, for γ small, we have no upper

bound to the capacity except for $C(N) \le 1$!!

(Eisert told me something last week....).