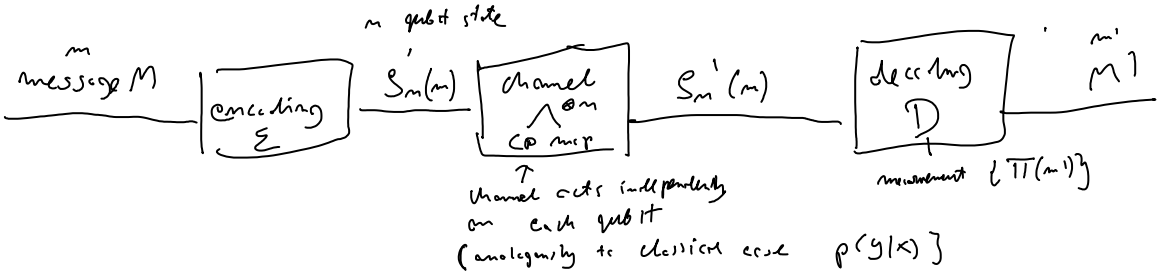
12 - Holevo bound, HSW theorem

5 grudnia 2010
15:20

We know the quantum analogue of Shannon coding theorem. What about Shannon channel theorem? What is the capacity of a quantum channel? How much information we can send using quantum states?



How many messages we can faithfully transmit. Rate $R = \frac{\log|M|}{n}$

12.1 Some definitions.

Def $S(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$ – q. mutual information

Def $S(A|B) = S(\rho_{AB}) - S(\rho_B)$ – q. conditional entropy

Notice: unlike in classical case can be negative !

Def: $S(\rho\|\sigma) = \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log\sigma)$ – q. relative entropy

Properties of S:

(i) $S(\rho\|\sigma) \geq 0$

(ii) subadditivity $S(AB) \leq S(A) + S(B)$

(iii) strong subadditivity $S(A,B,C) + S(B) \leq S(A,B) + S(B,C)$

$S(A:B) \leq S(A:BC)$ – equivalently

(iv) Let $\mathcal{E}$ be a q. channel on subsystem $B \to B'$, $\mathcal{E}: \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{B'})$

$S(A:B') \leq S(A:B)$

(v)

$$S\left(\sum_x p_x |x\rangle\langle x| \otimes \rho_x\right) = H(X) + \sum_x p_x S(\rho_x)$$ $\{|x\rangle\}$ – basis

Proof:

(i) $\rho = \sum_i p_i |i\rangle\langle i|$

$S(\rho\|\sigma) = \sum_i p_i \log p_i - \sum_i p_i \langle i|\log\sigma|i\rangle$ $\quad \sigma = \sum_k q_k |e_k\rangle\langle e_k|$

$= \sum_i p_i \left(\log p_i - \sum_k |\langle i|e_k\rangle|^2 \log q_k\right) \geq$

$\geq \sum_i p_i \left(\log p_i - \log \underbrace{\sum_k q_k |\langle i|e_k\rangle|^2}_{q_i}\right) \geq 0$ by classical relative entropy. ∎

(ii) $\rho = \rho_{AB}$ $\quad \sigma = \rho_A \otimes \rho_B$

$S(\rho\|\sigma) = \mathrm{Tr}\,\rho_{AB}\log\rho_{AB} - \mathrm{Tr}(\rho_{AB}\log\rho_A\otimes\rho_B) =$

$= -S(A,B) - \mathrm{Tr}(\rho_{AB} \log \rho_A \otimes \mathbb{1} \cdot \mathbb{1} \otimes \rho_B) = -S(A,B) + S(A) + S(B) \geq 0$ ∎

(iii) proof is a bit difficult unlike in classical case but intuitively it is clear

(iv)
$$S(A:B') = S(A) + S(B') - S(A,B')$$
$$\overset{(iii)}{\leq} S(A) + S(B',C') - S(A,B,C')$$
$S(B',C') = S(B)$ (because unitary operation cannot change entropy), $S(A,B,C') = S(A,B)$
$$\leq S(A) + S(B) - S(A,B) = S(A:B)$$

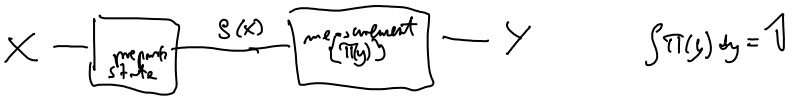
$\rho'_{AB} = \sum_i K_i \rho_{AB} K_i^\dagger$

$\rho'_{ABC} = U \rho_{AB} \otimes |0\rangle\langle 0| U^\dagger$ ← purification of the CP map

(v)
$$S\left(\sum_x p_x |x\rangle\langle x| \otimes \rho_x\right) = \sum_{x,r} p_x \lambda_{x,r} \log \lambda_{x,r} \cdot p_x = H(X) + \sum_x p_x S(\rho_x)$$

12.2 Holevo bound

Let X, $p(x)$ be classical random variable, according to which we prepare states $\rho(x)$. How much information we can learn about X by measuring $\rho(x)$



$$p(x,y) = p(x) \cdot \mathrm{Tr}(\Pi(y)\rho(x))$$

Theorem: $I(X:Y) \leq \underbrace{S(\rho) - \sum_x p_x S(\rho_x)}_{\chi\text{-Holevo quantity}}$ where $\rho = \sum_x p_x \rho_x$

(bound on accessible information.

Proof

Consider a state $\rho_{ABC} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|$ where $|x\rangle$ orthonormal basis

Consider the following CP map

$$\rho'_{ABC} = \mathcal{E}(\rho_{ABC}) = \sum_x p_x |x\rangle\langle x| \otimes \sum_y \sqrt{\Pi_y}\, \rho_x \sqrt{\Pi_y} \otimes \underbrace{U_y |0\rangle\langle 0| U_y^\dagger}_{|y\rangle\langle y|}$$

by (iv)
$$S(A:BC) \geq S(A:B'C') \geq S(A:C') = I(X:Y)$$

$S(A:BC) = H(X) + S(\rho) - \sum_x p_x S(\rho_x) - H(X)$

$\rho_{AC'} = \sum_{x,y} p_x \mathrm{Tr}(\Pi_y \rho_x) |x\rangle\langle x| \otimes |y\rangle\langle y|$

$$S(\rho) - \sum_x p_x S(\rho_x) \geq I(X:Y)$$ ∎

Fact: N qubit channel cannot transmit more than N bits of classical information

clearly $I(X:Y) \leq S(\rho) \leq N$

12.3 Capacity of a noisy quantum channel

(Holevo - Schumacher - Westmoreland theorem)

Max probability of error for a given code $[|M|, S_m]$

$$p_e = \max_{m \in M} \left[1 - \mathrm{Tr}\left[\Pi(m) S'^{\otimes n}(m) \right] \right]$$

Def: (classical) capacity $C(\Lambda)$ of a quantum channel Λ is a maximum rate R for which there exist sequence of codes $[2^{nR}, S_n]$ for which $\lim_{n\to\infty} p_e = 0$.

Theorem (HSW, 1995)

$$C^{(1)}(\Lambda) = \chi(\Lambda) = \max_{\{p_x, S_x\}} \left[S\Big(\Lambda\Big(\underbrace{\textstyle\sum_x p_x S_x}_{S'}\Big)\Big) - \underbrace{\textstyle\sum_x p_x S(\Lambda(S_x))}_{\overline{S}} \right]$$

(S': q. equivalent of $H(Y)$; $\overline{S}$: q. equivalent of $H(Y|X)$)

$C^{(1)}(\Lambda)$ — capacity if we restrict to encoding into product states (this would mean that entangled states are not needed for communication over independent channels)

{ It was long believed that $C(\Lambda) = C^{(1)}(\Lambda)$, however recently an example was shown where this is not the case.

Proof:

($\Rightarrow$) We present a construction that achieves $K = \chi(\Lambda)$

Of course we will use random coding: to each of messages $m \in \{1, \dots, 2^{nR}\}$ we assign a state:

$S_m^n = S_{x_1} \otimes \dots \otimes S_{x_n}$ where S_x are taken at random with p_x

The output state $S_m'^n = \Lambda^{\otimes n}(S_m^n)$

Let $S' = \sum_x p_x \Lambda(S_x)$ Let P – projection on ε typical subspace of $S'^{\otimes n}$

Let $S_x' = \Lambda(S_x) = \sum_k \lambda_k^x |e_k^x\rangle\langle e_k^x|$

$S_m'^n = \sum_K \lambda_K^n |e_K^n\rangle\langle e_K^n|$

$\lambda_K^n = \lambda_{k_1}^{x_1} \cdot \dots \cdot \lambda_{k_n}^{x_n}$ $\quad |e_K^n\rangle = |e_{k_1}^{x_1}\rangle \otimes \dots |e_{k_n}^{x_n}\rangle$

Let P^n be projection on subspace of $|e_K^n\rangle$ for which

$$\left| -\frac{1}{n} \log \lambda_K^n - \overline{S} \right| \le \varepsilon \qquad \overline{S} = \sum_x p_x S(\Lambda(S_x))$$

{ $\overline{S}$ is q. equivalent of classical $H(Y|X)$

T^n – set of all K for which the above is satisfied

Let us define a measurement

$$\Pi_m = Q^{-\frac{1}{2}} P P^m P Q^{-\frac{1}{2}}$$

{ notice $\sum \Pi_m \le 1$

inverse on the support $Q = \left(\sum_m p p^m p\right)$ — since we add $\Pi_0 = 1 - \sum_m \Pi_m$.

We calculate $\langle p_e \rangle$ — max pr. of errors averaged over all codebooks (will not depend on m so we can take any m)

$$\langle p_e \rangle = 1 - \left\langle \mathrm{Tr}\left(\Pi_m S_m^{\prime m}\right)\right\rangle = 1 - \frac{1}{2^{mR}}\left\langle \sum_m \mathrm{Tr}(\Pi_m S_m^{\prime m})\right\rangle$$

Let $P^m = \sum_{k\in T^m} |e_k^m\rangle\langle e_k|$ Let $|\tilde{e}_k^m\rangle = p|e_k^m\rangle$

$$\mathrm{Tr}\left(\Pi_m S_m^{\prime m}\right) = \mathrm{Tr}\Big[Q^{-\frac12} \sum_{k\in T^m} |\tilde{e}_k^m\rangle\langle \tilde{e}_k^m| \; Q^{-\frac12} \cdot \sum_k \lambda_k^m |e_k^m\rangle\langle e_k^m| \Big] =$$

$$= \sum_k \lambda_k^m \langle e_k^m| Q^{-\frac12} \sum_{k'\in T^m} |\tilde{e}_{k'}^m\rangle\langle \tilde{e}_{k'}^m| Q^{-\frac12} |e_k^m\rangle =$$

$$\geq \sum_{k,k'\in T^m} \lambda_k^m |\alpha_{km}^{k'm}|^2 \quad \Big\{ \alpha_{km}^{k'm'} = \langle \tilde{e}_{k'}^{m'}| Q^{-\frac12} |\tilde{e}_k^m\rangle$$

$$\geq \sum_{k\in T^m} \lambda_k^m |\alpha_{km}^{km}|^2 \quad \Big\{ x^2 \geq 2x - 1$$

$$\geq 2\sum_{k\in T^m} \lambda_k^m \alpha_{km}^{km} - 1$$

$\alpha_{km}^{k'm'} = \langle \tilde{e}_k^m | \Big(\sum_{m''} \sum_{k''} |\tilde{e}_{k''}^{m''}\rangle\langle \tilde{e}_{k''}^{m''}| \Big)^{-\frac12} | \tilde{e}_k^m \rangle$

We treat $\alpha_{km}^{k'm'}$ as a matrix $\hat{\alpha}$, $\sum_{k,m} \alpha_{km}^{km} = \mathrm{Tr}\hat{\alpha}$ Notice $\hat{\alpha} \geq 0$

$$\left(\hat{\alpha}^2\right)_l^{l'} = \sum_{l''} \langle \tilde{e}_{l'}| \Big(\sum_{\tilde{l}} |\tilde{e}_{\tilde{l}}\rangle\langle \tilde{e}_{\tilde{l}}| \Big)^{-\frac12} |\tilde{e}_{l''}\rangle \langle \tilde{e}_{l''}| \Big(\sum_{\tilde{l}} |\tilde{e}_{\tilde{l}}\rangle\langle \tilde{e}_{\tilde{l}}| \Big)^{-\frac12} |\tilde{e}_l\rangle$$

and we can commute $(\cdot\cdot)^{-\frac12}$

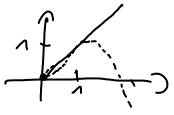
$$= \langle \tilde{e}_{l'} | \tilde{e}_l \rangle$$

$(1-\alpha)^2 = 1 - 2\alpha + \alpha^2$

$\frac{(1-\alpha)^2}{(1+\alpha)^2} = 1 - 2\alpha + \alpha^2$ $\qquad \frac{(1-\alpha^2)^2}{(1+\alpha)^2} = 1 - 2\alpha + \alpha^2$

$1 - 2\alpha + \alpha^2 \leq (1-\alpha^2)^2$ $\qquad \alpha \geq \frac12(1 + \alpha^2 - 1 + 2\alpha^2 - \alpha^4)$

$\alpha \geq \frac12(3\alpha^2 - \alpha^4)$



The same holds for matrices

So $\sum_{k\in T^m, m} \lambda_k^m \alpha_{km}^{km} = \mathrm{Tr}(\hat{\lambda}\cdot\hat{\alpha}) - \sum_{k\notin T^m, m} \lambda_k^m \alpha_{km}^{km}$ where $\hat{\lambda}$ – diagonal matrix with λ_k^m

$$\geq \frac12 \mathrm{Tr}\left(3\hat{\alpha}^2\hat{\lambda} - \hat{\lambda}\hat{\alpha}^4\right) - \sum_{k\notin T^m, m} \lambda_k^m = \quad \Big\{ \text{since } \alpha_{km}^{km} \leq 1$$

$$= \frac12 \Big(3 \cdot \sum_{k,m} \lambda_k^m \cdot \langle \tilde{e}_{k,m} | \tilde{e}_{km}\rangle - \sum_{\substack{k,m \\ k'm'}} \lambda_k^m \langle \tilde{e}_{k,m} | \tilde{e}_{k'm'}\rangle\langle \tilde{e}_{k'm'} | \tilde{e}_{km}\rangle \Big)$$

$$- \sum_{k \notin T^m} \lambda_k^m =$$

$$= \frac{1}{2}\left[3 \underbrace{\sum_m \mathrm{Tr}(S_m^{\prime m} P)}_{\geq (1-\delta)2^{mR}} - \sum_{m,m'} \mathrm{Tr}(P S_m^{\prime m} P P_{m'}) \right] - \sum_{k \notin T^m} \lambda_k^m$$

$$\geq \frac{1}{2}\left[3\sum_m \mathrm{Tr}(S_m^{\prime m} P) - \sum_m \mathrm{Tr}(P S_m^{\prime m} P P_m) - \sum_{m \neq m'} \mathrm{Tr}(P S_m^{\prime m} P P_{m'}) \right] - \underbrace{\sum_m \sum_{k \neq T^m} \lambda_k^m}_{\leq \delta \cdot 2^{-n}}$$

$$\geq \frac{1}{2}\left[2 \underbrace{\sum_m \mathrm{Tr}(S_m^{\prime m} P)}_{\geq 2^{mR}(1-\delta)} - \delta 2^{mR} - \sum_{m \neq m'} \mathrm{Tr}(P S_m^{\prime m} P P_{m'}) \right]$$

$$\langle p_e \rangle \leq 2 - 2(1-\delta) + 2\delta + \left(2^{mR} - 1\right) \mathrm{Tr}\left(P \langle S_m^{\prime m} \rangle P \langle P_{m'} \rangle\right)$$

since by construction of codebooks for different m we have independent distributions

$$\leq 4\delta + 2^{mR} \, 2^{-m(S(\delta)-\varepsilon)} \cdot 2^{m(\bar{S}+\varepsilon)} =$$

$2^{-m(S(\delta)-\varepsilon)}$: max eigenvalues of $P \langle S_m^{\prime m} \rangle P$ (the $S^{\prime\prime \otimes m}$)

$2^{m(\bar{S}+\varepsilon)}$: from bound on $\mathrm{Tr} \langle P_m \rangle$

$$= 4\delta + 2^{m(R - S(\delta) + \bar{S} + 2\varepsilon)}$$

So provided $R < S(\delta) - \bar{S}$ we have reliable communication.

Since $\langle p_e \rangle$ is arbitrary small i.e $< \epsilon$ it means that at least half of codes have $p_{av} < 2\epsilon$ so we can take one of them.

($\Leftarrow$)

Let us consider a code $[2^{mR}, m]$ for which (using product states)

$p_e \xrightarrow{m \to \infty} 0$

this implies that

$$\begin{cases} X = M \\ Y = M' \end{cases}$$

$I(M : M') \to mR$

on the other hand from Holevo bound we

on the other hand from Holevo bound we know that

$$I(M:M') \leq S(\mathcal{E}(\rho)) - \sum_m p(m)\, S(\rho_1^m \otimes \dots \otimes \rho_n^m)$$

$$\leq \left(S(\rho_1') + \dots + S(\rho_n')\right) - \sum_m p(m)\left(S(\rho_1^m) + \dots + S(\rho_n^m)\right)$$

$$nR \leq n\chi(\Lambda) \qquad R \leq \chi(\Lambda)$$

more formally we should use Fano inequality

Note HSW theorem tells us how much information we can communicate using product state, but most general collective measurement at the output → this will allow us to analyze collective attacks.

Examples:

- proof it is enough to encode into pure states
- capacity of depolarizing channel
- capacity of dephasing channel

Note: Notice that we can also think that A has a state $\rho_{AA'} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x$ half of which she sends to B

→ $\rho_{AB} = \sum_x p_x |x\rangle\langle x| \otimes \Lambda(\rho_x)$,

and measures her subsystem at some later stage even after B measurements

And the $\chi(\Lambda) = \max_{p_x, \rho_x} S(A:B)$