Quantum information theory

Lectures 29-31:

Quantum entropy

Classical information and Shannon entropy

Random variable: X, x_j, $p_j = p(x_j) = p_X(x_j)$

Shannon entropy: $H(\vec{p}) = H(X) = -\sum_j p(x_j) \log p(x_j)$

Ignorance and information $\qquad 0 \le H \le \log N$

Typical sequences and block coding

Yes-no questions and variable-length coding

Relative entropy: $\qquad H(\vec{p}\,\|\,\vec{q}) \to \infty$ if $p_j \neq 0 = q_j$ for some j

$$H(\vec{p}\,\|\,\vec{q}) = \sum_j p_j \log \frac{p_j}{q_j} = -H(\vec{p}) - \sum_j p_j \log q_j \ \underbrace{\geq}_{\text{iff } \vec{p} = \vec{q}} \ 0$$

Two random variables: X, Y

Conditional entropy: $H(X|Y) = \sum_y p(y) \left(-\sum_x p(x|y) \log p(x|y) \right)$

$$= \sum_{x,y} p(x,y) \log p(x|y)$$

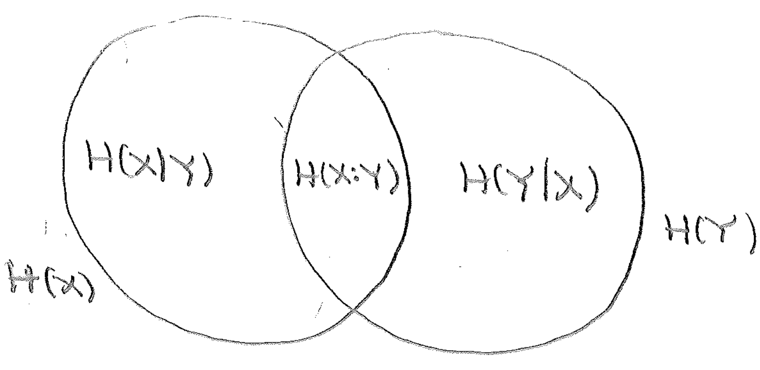
Mutual information:

$$H(X{:}Y) = H(X) - H(X|Y) = \sum_{x,y} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

$$= H(\vec{p}_{X,Y}\,\|\,\vec{p}_X \vec{p}_Y) \geq 0$$

$$= H(Y) - H(Y|X)$$

$$= H(X) + H(Y) - H(X,Y)$$



$H(X{:}Y) \le H(X), H(Y)$

Quantum information and von Neuman entropy

Quantum state: $\rho = \underbrace{\sum_j \lambda_j |e_j\rangle\langle e_j|}$ → eigendecomposition

Von Neuman entropy:

$$S(\rho) \equiv H(\vec{\lambda}) = -\sum_j \lambda_j \log \lambda_j = -\mathrm{tr}(\rho \ln \rho)$$

Why?
① Schumacher compression and block coding
② Accessible info
⋮

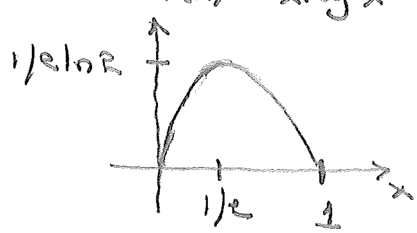
$$0 \leq S(\rho) \leq \log D$$

iff $\rho = |\psi\rangle\langle\psi|$ iff $\rho = I/D$

① ODOP inequality:

$$q_j = \langle f_j|\rho|f_j\rangle = \sum_k \underbrace{|\langle f_j|e_k\rangle|^2}_{P_{j|k}} \lambda_k$$

$f(x) = -x\log x$



$$\boxed{H(\vec{q}) \geq H(\vec{\lambda}) = S(\rho)}$$

iff $\vec{q} = \vec{\lambda}$

$f'(x) = -\log(ex)$

$f''(x) = -\dfrac{1}{x \ln 2} < 0$

doubly stochastic matrix

$$1 = \sum_j P_{j|k} = \sum_k P_{j|k}$$

Any concave function $f(x)$: $f(\vec{q}) \equiv \sum_j f(q_j) \geq f(\vec{\lambda}) = \mathrm{tr}(f(\rho))$

$$f(\vec{q}) = \sum_j f\Big(\underbrace{\sum_k |\langle f_j|e_k\rangle|^2 \lambda_k}\Big)$$

$$\geq \sum_k |\langle f_j|e_k\rangle|^2 f(\lambda_k)$$ ← concavity of f and double stochasticity of $P_{j|k}$

$$\geq \sum_k f(\lambda_k) \underbrace{\sum_j |\langle f_j|e_k\rangle|^2}_{1}$$

$$= f(\vec{\lambda}) = \mathrm{tr}(f(\rho))$$

⇓

If $q_j = \sum_k D_{jk} p_k$ (D: DS)

then $f(\vec{q}) \geq f(\vec{p})$

Direct proof for H:

$$H(\vec{q}) - H(\vec{\lambda}) = -\sum_k q_k \log q_k + \sum_j \lambda_j \log \lambda_j$$

$$= -\sum_{j,k} |\langle f_k | e_j \rangle|^2 \lambda_j \log \frac{q_k}{\lambda_j}$$

Use $\log x = \frac{\ln x}{\ln 2} \le \frac{x-1}{\ln 2}$

$$\ge \frac{1}{\ln 2} \sum_{j,k} |\langle f_k | e_j \rangle|^2 \lambda_j \left(1 - \frac{q_k}{\lambda_j}\right)$$

$$= \frac{1}{\ln 2} \sum_{j,k} |\langle f_k | e_j \rangle|^2 \lambda_j - \sum_{j,k} |\langle f_k | e_j \rangle|^2 q_k$$

$$= \frac{1}{\ln 2}\Big(\sum_k q_k - \underbrace{\sum_k q_k}_{\text{use DS}}\Big)$$

$$= 0$$

OR

$$= \sum_{j,k} |\langle f_k | e_j \rangle|^2 \lambda_j \times \log\left(\frac{|\langle f_k | e_j \rangle|^2 \lambda_j}{|\langle f_k | e_j \rangle|^2 q_k}\right)$$

This is a relative entropy

$$\ge 0$$

② Concavity:
$$\boxed{S(\underbrace{\mu \rho_1 + (1-\mu) \rho_2}_{\rho}) \ge \mu S(\rho_1) + (1-\mu) S(\rho_2)}$$

$0 < \mu < 1$: iff $\rho_1 = \rho_2$

$$\Rightarrow S(\underbrace{\sum_j p_j \rho_j}_{\rho}) \ge \sum_j p_j S(\rho_j)$$

OR

$$\chi = S(\rho) - \sum_j p_j S(\rho_j) \ge 0$$

Holevo quantity

Any concave function $f(x)$: $\operatorname{tr}(f(\rho)) \ge \mu \operatorname{tr}(f(\rho_1)) + (1-\mu) \operatorname{tr}(f(\rho_2))$

$$\operatorname{tr}(f(\rho)) = \sum_j f(\lambda_j) = \sum_j f(\langle e_j | \rho | e_j \rangle)$$

$$= \sum_j f\big(\mu \langle e_j | \rho_1 | e_j \rangle + (1-\mu) \langle e_j | \rho_2 | e_j \rangle\big)$$

concavity of f
$$\longrightarrow \ge \mu \sum_j f(\langle e_j | \rho_1 | e_j \rangle) + (1-\mu) \sum_j f(\langle e_j | \rho_2 | e_j \rangle)$$

ODOP inequality $\longrightarrow \geq \mu\, tr(f(\rho_1)) + (1-\mu)\, tr(f(\rho_2))$

③ Ensemble info inequality: $\rho = \sum_j p_j \rho_j$

$$\boxed{H(\vec{p}) \geq S(\rho) - \sum_j p_j S(\rho_j) \equiv \chi \geq 0}$$

iff ρ_j commute — iff ρ_j same

ⓐ $$\rho = \underbrace{\sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|}_{\text{Ensemble decomposition}} = \sum_j \lambda_j |e_j\rangle\langle e_j|$$

HJW theorem: $\sqrt{p_\alpha}|\psi_\alpha\rangle = \sum_j V_{\alpha j}\sqrt{\lambda_j}|e_j\rangle = \sum_\beta V_{\alpha\beta}\sqrt{\lambda_\beta}|e_\beta\rangle$

unitary matrix

$$\Longrightarrow \quad p_\alpha = \sum |V_{\alpha\beta}|^2 \lambda_\beta$$

doubly stochastic matrix

$$\therefore\ H(\vec{p}) \geq H(\vec{\lambda}) = S(\rho)$$

iff eigendecomposition

ⓑ $$\rho = \sum_j p_j \rho_j = \sum_{jk} p_j \lambda_{kj} |e_{kj}\rangle\langle e_{kj}|$$

$$ⓐ \Rightarrow H\begin{pmatrix}\text{joint distribution}\\ \lambda_{kj} p_j\end{pmatrix} \geq S(\rho)$$

$$\| $$

$$H(\vec{p}) + \sum_j p_j \underbrace{\left(-\sum_k \lambda_{kj}\log\lambda_{kj}\right)}_{S(\rho_j)} = H(\vec{p}) + \sum_j p_j S(\rho_j)$$

④ POVM inequality: POVM $\{E_\alpha\}$

$$q_\alpha = \mathrm{tr}(\rho E_\alpha)$$

$$\boxed{H(\vec{q}) + \sum_\alpha q_\alpha \log(\mathrm{tr}(E_\alpha)) = -\sum_\alpha q_\alpha \log\left(\frac{q_\alpha}{\mathrm{tr}(E_\alpha)}\right) \geq S(\rho)}$$

$$q_\alpha = \sum_j \lambda_j \underbrace{\langle e_j | E_\alpha | e_j \rangle}_{= P_{\alpha|j}} \qquad \sum_j P_{\alpha|j} = \mathrm{tr}(E_\alpha)$$

Let $f(x) = -x \log x$

$$H(\vec{q}) = \sum_\alpha f(q_\alpha) = \sum_\alpha f\left(\sum_j P_{\alpha|j} \lambda_j\right)$$

$$= \sum_\alpha f\left(\sum_j \frac{P_{\alpha|j}}{\mathrm{tr}(E_\alpha)} \lambda_j \,\mathrm{tr}(E_\alpha)\right)$$

concavity of f in average over j (distribution $P_{\alpha|j}/\mathrm{tr}(E_\alpha)$) →

$$\geq \sum_{\alpha,j} \frac{P_{\alpha|j}}{\mathrm{tr}(E_\alpha)} f(\lambda_j \,\mathrm{tr}(E_\alpha))$$

$$= -\sum_{\alpha,j} \frac{P_{\alpha|j}}{\mathrm{tr}(E_\alpha)} \lambda_j \,\mathrm{tr}(E_\alpha) \log(\lambda_j \,\mathrm{tr}(E_\alpha))$$

$$= \underbrace{-\sum_{\alpha,j} P_{\alpha|j} \lambda_j \log \lambda_j}_{= -\sum_j \lambda_j \log \lambda_j = H(\vec{\lambda})} - \underbrace{\sum_{\alpha,j} P_{\alpha|j} \lambda_j \log(\mathrm{tr}(E_\alpha))}_{= \sum_\alpha q_\alpha \log(\mathrm{tr}(E_\alpha))}$$

$$= S(\rho) - \sum_\alpha q_\alpha \log(\mathrm{tr}(E_\alpha))$$

5 Quantum relative entropy:

$$\boxed{S(\rho\|\sigma) = \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma) = -S(\rho) - \mathrm{tr}(\rho\log\sigma)}$$

Like a distance, but not symmetric

(a) $\rho = \sum_j p_j |e_j\rangle\langle e_j|, \quad \sigma = \sum_j q_j |f_j\rangle\langle f_j|$

$$S(\rho\|\sigma) = -S(\rho) - \mathrm{tr}(\rho\log\sigma)$$

$$= -H(\vec{p}) - \sum_j \underbrace{\langle f_j|\rho|f_j\rangle}_{\equiv r_j} \log q_j$$

$$= -H(\vec{p}) - \sum_j r_j \log r_j + \sum_j r_j \log(q_j/r_j)$$

$$\boxed{S(\rho\|\sigma) = \underbrace{\overbrace{-H(\vec{p})}^{-S(\rho)} + \overbrace{H(\vec{r}) + H(\vec{r}\|\vec{q})}^{-\mathrm{tr}(\rho\log\sigma)}}_{}\ \geq 0}$$

$-H(\vec{p}) + H(\vec{r}) \geq 0$ (ODOP), $\quad H(\vec{r}\|\vec{q}) \geq 0$, $\quad = 0$ iff $\rho = \sigma$

$S(\rho\|\sigma) \to \infty$ if $r_j \neq 0 = q_j$ for some j

(b) Unitary invariance

(c) $\chi = S(\rho) - \sum_j p_j S(\rho_j) = \sum_j p_j \underbrace{\left(\mathrm{tr}(\rho_j\log\rho_j) - \mathrm{tr}(\rho_j\log\rho)\right)}_{S(\rho_j\|\rho)}$

$$\boxed{\chi = \sum_j p_j S(\rho_j\|\rho)}$$

<u>Two systems:</u> ρ_{AB} ρ_A ρ_B

$S(\rho_{AB}) = S(A,B)$ $S(\rho_A) = S(A)$ $S(\rho_B) = S(B)$

① If ρ_{AB} is pure, $S(\rho_{AB}) = 0$ and $S(\rho_A) = S(\rho_B)$

Use Schmidt decomposition

② Subadditivity: $\boxed{S(A,B) \leq S(A) + S(B)}$ ($=$ iff $\rho_{AB} = \rho_A \otimes \rho_B$)

ⓐ

$$\left.\begin{aligned} \rho_A &= \sum_j p_j |e_j\rangle\langle e_j| \\ \rho_B &= \sum_k q_k |f_k\rangle\langle f_k| \end{aligned}\right] \to \text{Eigendecompositions}$$

$$\langle e_j, f_k | \rho_{AB} | e_j, f_k \rangle \equiv r_{jk} \qquad \sum_k r_{jk} = p_j \qquad \sum_j r_{jk} = q_k$$

$$S(A,B) \underset{\substack{\uparrow \\ \text{ODOP} \\ \text{ineq.}}}{\leq} H(\vec{r}) \underset{\substack{\uparrow \\ \text{classical} \\ \text{subadditivity}}}{\leq} H(\vec{p}) + H(\vec{q}) = S(\rho_A) + S(\rho_B)$$

<u>OR</u>

ⓑ

$$0 \leq S(\rho_{AB} \| \rho_A \otimes \rho_B) = -S(\rho_{AB}) \underbrace{- \operatorname{tr}\left(\rho_{AB} \log(\rho_A \otimes \rho_B)\right)}_{S(A) + S(B)}$$

Contrast classical
$H(X,Y) \geq H(X), H(Y)$

③ Araki-Lieb (triangle inequality)

$$S(A,B) \geq |S(A) - S(B)|$$

iff $\rho_{AB} = \sum_j p_j |\psi_j\rangle\langle\psi_j|$, with
$tr_A(|\psi_j\rangle\langle\psi_k|) = \rho_B \delta_{jk}$ $[S(A,B) = S(A) - S(B)]$
or $tr_B(|\psi_j\rangle\langle\psi_k|) = \rho_A \delta_{jk}$ $[S(A,B) = S(B) - S(A)]$,

ρ_{AB} is a mixture of orthogonal purifications of ρ_A or ρ_B

ⓐ Let $|\Phi_{RAB}\rangle$ be a purification of ρ_{AB}.

$$\underbrace{S(A) + S(R)}_{} \overset{\uparrow \text{ subadditivity}}{\geq} \underbrace{S(A,R)}_{}$$

$$S(A,B) \xleftarrow{\text{overall pure state}} \quad \longrightarrow S(B)$$

Here purification works because it changes joint entropies to single-system entropies & vice versa.

OR

ⓑ $\rho_{AB} = \sum_j \lambda_j |e_j\rangle\langle e_j|$ ⟵ eigendecomposition

$$\rho_A = \sum_j \lambda_j \rho_j^A, \quad \rho_B = \sum_j \lambda_j \rho_j^B$$

$$\boxed{S(\rho_j^A) = S(\rho_j^B)}$$

$\rho_j^A = tr_B(|e_j\rangle\langle e_j|)$ are $\perp$

$$\boxed{S(\rho_{AB}) = H(\vec{\lambda}) \geq S(\rho_A) - \sum_j \lambda_j S(\rho_j^A)} \geq 0$$

$$S(\rho_{AB}) = H(\vec{\lambda}) \geq \boxed{S(\rho_B) - \sum_j \lambda_j S(\rho_j^B)} \geq 0$$

$\rho_j^B = tr_A(|e_j\rangle\langle e_j|)$ are same

$$S(\rho_{AB}) \geq S(\rho_A) - S(\rho_B)$$

Quantum conditional entropy

$$S(A|B) = S(A,B) - S(B) \begin{cases} \geq -S(A), & S(B) \geq S(A) \\ \geq S(A) - 2S(B), & S(A) \geq S(B) \end{cases}$$

$$S(B|A) = S(A,B) - S(A)$$

Could we use negativity of these as entanglement criterion?

Quantum mutual information

$$S(A:B) = S(A) + S(B) - S(A,B) \geq 0$$

Subadditivity

$$S(A:B) \leq 2S(A),\ 2S(B)$$

$S(A) \geq S(B)$ $\quad$ $S(B) \geq S(A)$

Too correlated to be classical

Quantum conditional entropy and relative entropy

$$S\left(\rho_{AB} \| \frac{I}{d} \otimes \rho_B\right) = -S(\rho_{AB}) - \mathrm{tr}\left(\rho_{AB} \log\left(\frac{I}{d} \otimes \rho_B\right)\right)$$

$$\log\left(\sum_{j,k} \frac{1}{d} |e_j\rangle\langle e_j| \otimes \lambda_k |f_k\rangle\langle f_k|\right)$$

$$= \sum_{j,k} \log\left(\frac{\lambda_k}{d}\right) |e_j\rangle\langle e_j| \otimes |f_k\rangle\langle f_k|)$$

$$= I \otimes \left(\sum_k \log \lambda_k |f_k\rangle\langle f_k| - I \log d\right)$$

$$= I \otimes \log \rho_B - I \otimes I \log d$$

$$= -S(\rho_{AB}) - \mathrm{tr}(\rho_B \log \rho_B) + \log d$$

$$= -S(A,B) + S(B) + \log d$$

$$\boxed{S(A|B) = -S(\rho_{AB} \| (I/d) \otimes \rho_B) + \log d}$$

④ Another proof of concavity of S

$$\rho = \sum_j p_j \rho_j$$

Define $\rho_{AB} = \sum_j p_j \rho_j \otimes |f_j\rangle\langle f_j|$ ← extension (not a purification)

B states $|f_j\rangle$ <u>record</u> which state ρ_j applies to A

$$\rho_A = \rho \qquad \rho_B = \sum_j p_j |f_j\rangle\langle f_j|$$

$$S(A,B) = H(\vec{p}) + \sum_j p_j S(\rho_j)$$ ← Why?

$$S(A) = S(\rho)$$

$$S(B) = H(\vec{p})$$

Subadditivity $\Rightarrow \sum_j p_j S(\rho_j) \leq S(\rho)$

Extensions allow us to get the ensemble info & the average ensemble entropy into our equations.

⑤ Another proof of the ensemble information inequality

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$$

$|\Phi_{AB}\rangle = \sum_j \sqrt{p_j}\, |\psi_j\rangle \otimes |f_j\rangle$ is a purification of ρ

$$\rho_A = \rho, \qquad \rho_B = \sum_{j,k} \sqrt{p_j p_k}\, \langle\psi_j|\psi_k\rangle\, |f_j\rangle\langle f_k|$$

$$\Rightarrow \langle f_j|\rho_B|f_j\rangle = \sum_j p_j |f_j\rangle\langle f_j|$$

$$H(\vec{p}) \geq S(\rho_B) = S(\rho_A)$$

↑ QDOP ineq.

Three systems:

Strong subadditivity (SS):

$$\boxed{\begin{aligned} &S(A,B,C) + S(B) \leq S(A,B) + S(B,C) \\ &\underline{\text{OR}} \\ &S(A|B,C) \leq S(A|B) \longleftarrow \text{Trivial classically} \end{aligned}}$$

Proof: N&C App 6 and 11.4.1

Mother of all properties: use purifications and extensions

$\rho_{AB} \longrightarrow |\Phi_{ABC}\rangle$: $\quad 0 + S(B) \leq S(A,B) + S(A)$ Araki-Lieb

$\rho_{AC} \longrightarrow |\Phi_{ABC}\rangle$: $\quad 0 + S(A,C) \leq S(C) + S(A)$ Subadditivity

SS $\Longrightarrow$ concavity of conditional quantum entropy

$$\rho_{AB} = \sum_j p_j \rho_j^{AB}$$

Extend to $\rho_{ABC} = \sum_j p_j \rho_j^{AB} \otimes |f_j\rangle\langle f_j|$

Apply to $\underbrace{S(A,B,C)}_{= H(\vec{p}) + \sum_j p_j S(\rho_j^{AB})} + S(A) \leq S(A,B) + \underbrace{S(A,C)}_{= H(\vec{p}) + \sum_j p_j S(\rho_j^{A})}$

$$\rho_{AC} = \sum_j p_j \rho_j^{A} \otimes |f_j\rangle\langle f_j|$$

$$\Longrightarrow \underbrace{S(A,B) - S(A)}_{S(B|A)} \geq \sum_j p_j \Big(\underbrace{S(\rho_j^{AB}) - S(\rho_j^{A})}_{S_j(B|A)} \Big)$$

$$\boxed{S(B|A) \geq \sum_j p_j S_j(B|A)}$$

Concavity of quantum conditional entropy

Rewrite as $\underbrace{S(\rho_A) - \sum_j p_j S(\rho_j^A)}_{\chi_A} \leq \underbrace{S(\rho_{AB}) - \sum_j p_j S(\rho_j^{AB})}_{\chi_{AB}}$

Marginalization property of Holevo quantity

Note: $H(\vec{p}) \geq \chi_A$, $H(\vec{p}) \geq \chi_{AB}$ are ensemble info inequalities.
The marginalization property establishes that $H(\vec{p}) \geq \chi_{AB} \geq \chi_A$

Entanglement inequality: Specialize marginalization of χ to case where ρ_j^{AB} is a pure-state ensemble.

$$\boxed{\sum_j p_j S(\rho_j^A) \geq S(A) - S(A,B) = -S(B|A)}$$

If $S(A) > S(A,B)$, then ρ_{AB} cannot be written as a mixture of product states.

An equivalent form of strong subadditivity:

$$\left.\begin{matrix}\rho_{ABC} \\ \rho_{RBC}\end{matrix}\right] \longrightarrow |\Phi_{RABC}\rangle$$

$$0 \geq \underbrace{S(A,B,C)}_{S(R)} + S(B) - \underbrace{S(A,B)}_{S(R,C)} - S(B,C)$$

$$\boxed{\begin{aligned} & S(A) + S(B) \leq S(A,C) + S(B,C) \\ & \text{OR} \\ & S(C|A) + S(C|B) \geq 0 \iff S(A:C) + S(B:C) \leq 2S(C) \end{aligned}}$$

Monogamy of entanglement

$S(A|C) + S(B|C)$ can be negative: $|\psi_{ABC}\rangle = |\psi_A\rangle \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

$S(A,C) = 1$

$S(A) = 0, \quad S(C) = 1$

$S(A|C) = 0, \quad S(C|A) = 1$

$S(B,C) = 0$

$S(B) = S(C) = 1$

$S(B|C) = S(C|B) = -1$

Properties equivalent to SS (see pages A-C)

① Monotonicity of relative entropy: $S(\rho_A \| \sigma_A) \leq S(\rho_{AB} \| \sigma_{AB})$

② Double convexity of relative entropy:

$$S(\rho\|\sigma) \leq \sum_j p_j S(\rho_j \| \sigma_j), \quad \rho = \sum_j p_j \rho_j, \quad \sigma = \sum_j p_j \sigma_j$$

③ Concavity of conditional entropy.

① & ② are good distinguishability properties for relative entropy.

Related properties:

Contractivity of relative entropy:

$$\boxed{\begin{array}{l} S(\rho' \| \sigma') \le S(\rho \| \sigma) \\ \rho' = \mathcal{Q}(\rho),\ \sigma' = \mathcal{Q}(\sigma),\ \mathcal{Q} \text{ t-p qo} \end{array}}$$

$$S(\rho' \| \sigma') = S(\mathcal{Q}(\rho) \| \mathcal{Q}(\sigma))$$

$$= S\left(\mathrm{tr}_A(U \rho \otimes |0\rangle\langle 0| U^\dagger) \,\|\, \mathrm{tr}_A(U \sigma \otimes |0\rangle\langle 0| U^\dagger)\right)$$

monotonicity of relative entropy

$$\le S(U \rho \otimes |0\rangle\langle 0| U^\dagger \,\|\, U \sigma \otimes |0\rangle\langle 0| U^\dagger)$$

unitary invariance of relative entropy

$$= S(\rho \otimes |0\rangle\langle 0| \,\|\, \sigma \otimes |0\rangle\langle 0|)$$

$$= S(\rho \| \sigma)$$

(c) Since $\chi = \sum_j p_j S(\rho_j \| \rho)$, we get

$\chi_A \le \chi_{AB}$ from monotonicity of relative entropy

$\chi' \le \chi$ from contractivity of relative entropy

(d) Since $S(A|B) = -S(\rho_{AB} \| (I/d) \otimes \rho_B)$, we get

$S(A'|B') \ge S(A|B)$ from contractivity of relative entropy.

Holevo bound

classical info down a noiseless quantum channel

noiseless $\longrightarrow$ only "noise" is due to quantum measurement

Input: Alice sends ρ_x with probability p_x

Output: Bob measures POVM E_y

$$p_{y|x} = \mathrm{tr}(E_y \rho_x)$$

Holevo bound: $$H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x) = \chi$$

$=$ iff ρ_x commute

$H(X:Y)$: maximum value of this mutual information is called the accessible information

Proof: Let Q denote the quantum system. Introduce a reference system R and an ancilla A that record the input and output

Initial state: $$\rho_{RQA} = \underbrace{\left(\sum_x p_x |x\rangle\langle x| \otimes \rho_x \right)}_{\rho_{RQ}} \otimes |0\rangle\langle 0|$$

$$\rho_R = \sum_x p_x |x\rangle\langle x|$$

Final state: Quantum operation $\mathcal{E}$ on Q and A such that

$$\mathcal{A}(\sigma \otimes |0\rangle\langle 0|) = \sum_y \underbrace{\sqrt{E_y} \otimes U_y}_{\text{Kraus operators}} (\sigma \otimes |0\rangle\langle 0|) \sqrt{E_y} \otimes U_y^\dagger$$

$$= \sum_y \underbrace{\sqrt{E_y}\,\sigma\sqrt{E_y}}_{p_{y|\sigma}\sigma_y} \otimes |y\rangle\langle y|$$

makes an extension of ensemble $p_{y|\sigma}\sigma_y$

$$\rho'_{RQA} = \mathcal{D} \otimes \mathcal{A}(\rho_{RQA})$$

$$= \sum_{x,y} p_x |x\rangle\langle x| \otimes \sqrt{E_y}\,\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|$$

$$\rho'_{RA} = \sum_{x,y} p_{y|x}\, p_x |x\rangle\langle x| \otimes |y\rangle\langle y|$$

classical correlation of X and Y

$$\rho'_R = \sum_x p_x |x\rangle\langle x| = \rho_R$$

Our guess: $\underbrace{H(X:Y) = S(R':A')}_{\uparrow\ \text{obvious}} \leq S(R:Q) = \chi$

$$S(R:Q) = \underbrace{S(R)}_{H(\vec{p})} + \underbrace{S(Q)}_{S(\rho)} - \underbrace{S(R,Q)}_{H(\vec{p}) + \sum_x p_x S(\rho_x)} = \chi$$

$$S(R:Q) = S(R) + \overbrace{S(Q) - S(R,Q)}^{-S(R|Q)}$$

$$= \underbrace{S(R) + \underbrace{S(Q,A) - S(R,Q,A)}_{-S(R|Q,A)}}_{S(R:Q,A)}$$

because A is initially uncorrelated

conditional entropy increases under quantum operation

$$\geq S(R') - S(R'|Q',A')$$

$$SS \rightsquigarrow \geq S(R') - S(R'|A')$$

$$= S(R':A')$$

Properties equivalent to SS

① Monotonicity of relative entropy

② Double convexity of relative entropy

③ Concavity of conditional entropy

① $\Rightarrow$ ②: $\quad \rho = \sum_j p_j \rho_j, \quad \sigma = \sum_j p_j \sigma_j$

$$\rho_{AB} = \sum_j p_j \rho_j \otimes |f_j\rangle\langle f_j| \quad , \quad \rho_B = \sum_j p_j |f_j\rangle\langle f_j| = \sigma_B$$

$$\sigma_{AB} = \sum_j p_j \sigma_j \otimes |f_j\rangle\langle f_j| \quad ,$$

$$\log \rho_{AB} = I \otimes \log \rho_B + \sum_j p_j \log \rho_j \otimes |f_j\rangle\langle f_j|$$

$$\log \sigma_{AB} = I \otimes \log \rho_B + \sum_j p_j \log \sigma_j \otimes |f_j\rangle\langle f_j|$$

So $S(\rho_{AB} \| \sigma_{AB}) = -S(\rho_{AB}) - tr(\rho_{AB} \log \sigma_{AB})$

$$= -H(\vec{p}) - \sum_j p_j S(\rho_j)$$
$$+ H(\vec{p}) - \sum_j p_j tr(\rho_j \log \sigma_j)$$

$$= \sum_j p_j S(\rho_j \| \sigma_j)$$

① $\Rightarrow$ $S(\rho \| \sigma) \le S(\rho_{AB} \| \sigma_{AB}) = \sum_j p_j S(\rho_j \| \sigma_j)$

② $\Rightarrow$ ①: The completely depolarizing channel can be written as $\mathcal{Q} = \mathbb{I}/d = \frac{1}{d^2} \sum_\alpha U_\alpha \odot U_\alpha^\dagger$, where the U_α are a set of d^2 orthogonal unitary operators. Thus

$$\frac{1}{d^2} \sum_\alpha U_\alpha \rho U_\alpha^\dagger = \mathcal{Q}(\rho) = \frac{1}{d} \mathbb{I}(\rho) = \frac{I}{d} tr(\rho) = \frac{I}{d}.$$

So

$$\frac{1}{d^2}\sum_\alpha I\otimes U_\alpha \rho_{AB} I\otimes U_\alpha^\dagger = \mathcal{I}\otimes\mathcal{C}(\rho_{AB}) = tr_B(\rho_{AB})\otimes\frac{I}{d} = \rho_A\otimes\frac{I}{d}.$$

$$\Rightarrow S(\rho_A\|\sigma_A) = S(\rho_A\otimes(I/d)\|\sigma_A\otimes(I/d))$$

$$\leq \frac{1}{d^2}\sum_\alpha S(I\otimes U_\alpha\rho_{AB}I\otimes U_\alpha^\dagger\|I\otimes U_\alpha\sigma_{AB}I\otimes U_\alpha^\dagger)$$

convexity of relative entropy

$$= \frac{1}{d^2}\sum_\alpha S(\rho_{AB}\|\sigma_{AB})$$

$$= S(\rho_{AB}\|\sigma_{AB})$$

unitary invariance of relative entropy

Ⓔ $\Rightarrow$ Ⓑ: $S(A|B) = -S(\rho_{AB}\|(I/d)\otimes\rho_B) + \log d$

Ⓑ $\Rightarrow$ SS: Define $T(\rho_{ABC}) = S(A)+S(B)-S(A,C)-S(B,C)$

$$= -S(C|A) - S(C|B)$$

$$\rho_{ABC} = \sum_j \lambda_j |e_j\rangle\langle e_j|$$

$$T(\rho_{ABC}) = -S(C|A) - S(C|B)$$

$$\leq \sum_j \lambda_j\left(-S_j(C|A) - S_j(C|B)\right)$$

concavity of conditional entropy

$$= \sum_j \lambda_j\left(S_j(A)+S_j(B)-\underbrace{S_j(A,C)}_{S_j(B)}-\underbrace{S_j(B,C)}_{S_j(A)}\right)$$

$$= 0$$

SS $\Rightarrow$ (2): Already shown on pp. 11–12.

I haven't been able to find a proof that (3) or SS implies (1) or (2), so it remains possible that (1) and (2) are strictly stronger than (3) and SS.

Schumacher compression.

i.i.d. source: p_j, $\rho_j = |\psi_j\rangle\langle\psi_j|$

$$\rho = \sum_j p_j \rho_j = \begin{pmatrix}\text{unconditioned source} \\ \text{density operator}\end{pmatrix}$$

$$= \sum_x p(x)\, |x\rangle\langle x|$$

eigendecomposition

i.i.d. random variable X with probability $p(x)$

Key is to work with the eigenstates and their eigenvalues rather than the source states and their probabilities.

$$H(\vec{p}) = S(\rho) \qquad \rho^{\otimes N} = \sum_{x_1,\ldots,x_N} \overbrace{p(x_1)\ldots p(x_N)}^{p(x_1,\ldots,x_N) = p(x)} |x_1,\ldots,x_N\rangle\langle x_1,\ldots,x_N|$$

<u>ϵ-typical states:</u> A state $|x_1,\ldots,x_N\rangle$ is ϵ-typical if the sequence $x_1,\ldots,x_N$ is ϵ-typical, i.e., if

$$\left| \frac{1}{N} \log\left(\frac{1}{p(x_1,\ldots,x_N)}\right) - \underset{H(\vec{p})}{\underset{\uparrow}{S(\rho)}} \right| \leq \epsilon.$$

Let $T(N,\epsilon)$ be the <u>subspace</u> spanned by the ϵ-typical sequences, and $P(N,\epsilon)$ be the projector onto $T(N,\epsilon)$.

The AEP (typical sequences theorem) can be translated directly to a typical subspace theorem.

Typical subspace theorem.

(i) For any $\epsilon, \delta > 0$, there exists N_0 such that for all $N \geq N_0$, $tr\left(P(N,\epsilon)\rho^{\otimes N}\right) \geq 1-\delta$.

Proof: $$tr\left(P(N,\epsilon)\rho^{\otimes N}\right) = \sum_{x\ \epsilon\text{-typical}} p(x_1,\dots,x_N) \underset{\text{AEP}}{\geq} 1-\delta$$

(ii) $$(1-\delta)\, 2^{N(S(\rho)-\epsilon)} \leq |T(N,\epsilon)| = tr\left(P(N,\epsilon)\right) \leq 2^{N(S(\rho)+\epsilon)}$$

$$\left(\begin{matrix}\text{dimension of}\\ T(N,\epsilon)\end{matrix}\right) = \left(\begin{matrix}\#\text{ of }\epsilon\text{-typical}\\ \text{sequences}\end{matrix}\right)$$

(iii) Let S_N be any subspace of $\mathcal{H}^{\otimes N}$ of dimension 2^{NR}, where $R < S(\rho)$. Given any $\delta > 0$, there exists N_0 such that for all $N \geq N_0$,

$$tr\left(S_N \rho^{\otimes N}\right) \leq \delta.$$

Proof: $$tr\left(S_N\rho^{\otimes N}\right) \leq \sum_{\substack{2^{NR}\text{ largest}\\ \text{probabilities}\\ p(x)}} p(x) \underset{\text{AEP}}{\leq} \delta$$

uses lemma on next page

Lemma: Let $\vec{\lambda} = (\lambda_1, \dots, \lambda_D)$, with $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_D$.

$$\max\left(\vec{\lambda}\cdot\vec{g} = \sum_j \lambda_j g_j \;\middle|\; 0 \le g_j \le 1,\ j = 1,\dots,D \text{ and } \sum_{j=1}^{D} g_j = N\right) = \sum_{j=1}^{N} \lambda_j .$$

Proof: $\vec{\lambda}\cdot\vec{g}$ does not decrease whenever we decrement g_j by δ and increase g_k, $k < j$, by δ, consistent with $g_j \ge 0$ and $g_k \le 1$. So the maximum occurs when

$$g_j = \begin{cases} 1, & j = 1,\dots,N, \\ 0, & j = N+1,\dots,D. \end{cases}$$

Lemma: If $H = H^\dagger$ has eigendecomposition $H = \sum_{j=1}^{D} \lambda_j |e_j\rangle\langle e_j|$, where $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_D$, and Π is an N-dimensional projector, then

$$\begin{pmatrix}\text{Sum of } N \\ \text{smallest} \\ \text{eigenvalues}\end{pmatrix} = \sum_{j=D+1-N}^{D} \lambda_j \le \operatorname{tr}(H\Pi) \le \sum_{j=1}^{N} \lambda_j = \begin{pmatrix}\text{Sum of } N \\ \text{largest} \\ \text{eigenvalues}\end{pmatrix}$$

attained by $\Pi = \sum_{j=D+1-N}^{D} |e_j\rangle\langle e_j|$

attained by $\Pi = \sum_{j=1}^{N} |e_j\rangle\langle e_j|$

Proof:

$$\operatorname{tr}(H\Pi) = \sum_j \lambda_j \underbrace{\langle e_j|\Pi|e_j\rangle}_{g_j} \le \sum_{j=1}^{N} \lambda_j$$

$$0 \le g_j \le 1, \quad \sum_{j=1}^{N} g_j = \operatorname{tr}(\Pi) = N$$

$$\operatorname{tr}\big(H\underbrace{(1-\Pi)}_{(D-N)\text{-dimensional projector}}\big) \le \sum_{j=1}^{D-N} \lambda_j \implies \operatorname{tr}(H\Pi) \ge \operatorname{tr}(H) - \sum_{j=1}^{D-N} \lambda_j = \sum_{j=D-N+1}^{D} \lambda_j$$

$\rho = \sum_j p_j \rho_j$

i.i.d. Source: p_j, $\rho_j = |\psi_j\rangle\langle\psi_j|$

$$\rho^{\otimes N} = \sum_{\underbrace{j_1,\ldots,j_N}_{J}} \underbrace{p_{j_1}\cdots p_{j_N}}_{p_J} \underbrace{\rho_{j_1}\otimes\cdots\otimes\rho_{j_N}}$$

Compression scheme of "rate" R:

Coding $\mathcal{C}^N$: $\mathcal{H}^{\otimes N} \longrightarrow 2^{NR}$-dimensional subspace

Decoding $\mathcal{D}^N$

$$\rho_J = |\psi_{j_1}\rangle\otimes\cdots|\psi_{j_N}\rangle\langle\psi_{j_1}|\otimes\cdots\otimes\langle\psi_{j_N}| = |J\rangle\langle J|$$

Reliability:

$$\begin{pmatrix}\text{average}\\ \text{fidelity}\end{pmatrix} = \overline{F}^N = \sum_J p_J F(\rho_J, \mathcal{D}^N\circ\mathcal{C}^N(\rho_J)) \le F(\rho^{\otimes N}, \mathcal{D}^N\circ\mathcal{C}^N(\rho^{\otimes N}))$$

NBC use entanglement fidelity

double concavity of fidelity

Schumacher compression theorem:

(i) $R > S(\rho)$: For any $\delta > 0$, there exists N_0 such that for all $N \ge N_0$, $\overline{F}^N \ge 1-\delta$.

It is easy to push through the proof for entanglement fidelity, even though compressing half of entangled pairs is a harder task.

Proof: Let $R = S(\rho)+\epsilon$, $\epsilon > 0$. The ϵ-typical subspace $T(N,\epsilon)$ can be thought of as a subspace of the 2^{NR}-dimensional subspace.

Coding: $$\mathcal{C}^N = \underbrace{P(N,\epsilon)\odot P(N,\epsilon)}_{\text{projection into typical subspace}} + \sum_\alpha |0\rangle\langle\alpha|\odot|\alpha\rangle\langle 0|$$

orthonormal basis in atypical subspace

any state in typical subspace

Decoding: $\mathcal{D}^N$ is unit superoperator on $T(N,\epsilon)$.

$$\mathcal{C}^N(\rho_J) = \underbrace{P(N,\epsilon)\rho_J P(N,\epsilon)}_{P(N,\epsilon)|J\rangle\langle J|P(N,\epsilon)} + |0\rangle\langle 0| \underbrace{\sum_\alpha \langle\alpha|\rho_J|\alpha\rangle}$$

$$= tr(\rho_J \underbrace{\bar{P}(N,\epsilon)}_{1-P(N,\epsilon)})$$

$$= \langle J|\bar{P}(N,\epsilon)|J\rangle$$

$$\mathcal{C}^N(\rho_J) = P(N,\epsilon)|J\rangle\langle J|P(N,\epsilon) + |0\rangle\langle 0| \langle J|\bar{P}(N,\epsilon)|J\rangle$$

$$\mathcal{D}^N \circ \mathcal{C}^N(\rho_J) = \mathcal{C}^N(\rho_J)$$

$$\overline{F^N} = \sum_J p_J F\underbrace{(\rho_J, \mathcal{D}^N \circ \mathcal{C}^N(\rho_J))}$$

$$= \langle J|\mathcal{D}^N \circ \mathcal{C}^N(\rho_J)|J\rangle^{1/2} = \Big(\langle J|P(N,\epsilon)|J\rangle^2 + |\langle J|0\rangle|^2 \langle J|\bar{P}(N,\epsilon)|J\rangle\Big)^{1/2}$$

$$\geq \langle J|P(N,\epsilon)|J\rangle$$

$$\overline{F^N} \geq \sum_J p_J \langle J|P(N,\epsilon)|J\rangle = tr(\rho^{\otimes N} P(N,\epsilon)) \geq 1-\delta$$

If we use entanglement fidelity, we have

$$F_e(\rho^{\otimes N}, \mathcal{D}^N \circ \mathcal{C}^N) = (\rho^{\otimes N}| \mathcal{D}^N \circ \mathcal{C}^N | \rho^{\otimes N})$$

$$= (\rho^{\otimes N}| \mathcal{C}^N | \rho^{\otimes N})$$

$$= |tr(\rho^{\otimes N} P(N,\epsilon))|^2 + \sum_\alpha |tr(\rho^{\otimes N} |0\rangle\langle\alpha|)|^2$$

$$\geq |\mathrm{tr}(\rho^{\otimes N} P(N,\epsilon))|^2$$

$$\geq (1-\delta')^2$$

This square is the expression of how compressing half of entanglement pair is a harder task.

$$\geq 1 - 2\delta'$$

$$= 1 - \delta \qquad (\delta = 2\delta')$$

The entanglement fidelity result implies the average fidelity result because $\overline{F_N} \geq \sqrt{F_e} \geq 1 - \delta$.

(ii) $R < S(\rho)$: Given any $\delta > 0$, there exists N_0 such that $F_e(\rho^{\otimes N}, \mathcal{D}^N \circ \mathcal{C}^N) \leq \delta$ for all $N \geq N_0$.

<u>Proof:</u> Coding $\mathcal{C}^N = \sum_\alpha C_\alpha \odot C_\alpha^\dagger$

maps to 2^{NR}-dimensional subspace S_N, projector P_N

Decoding: $\mathcal{D}^N = \sum_\beta D_\beta \odot D_\beta^\dagger$

maps S_N to a subspace $S_{N\beta}$ of dimensionality $\leq \dim(S^N) = 2^{NR}$, projector $P_{N\beta}$.

$$F_e(\rho^{\otimes N}, \mathcal{D}^N \circ \mathcal{C}^N) = (\rho^{\otimes N} | \mathcal{D}^N \circ \mathcal{C}^N | \rho^{\otimes N})$$

$$= \sum_{\alpha,\beta} |\mathrm{tr}(\rho^{\otimes N} D_\beta C_\alpha)|^2$$

can put P_N here

can put $P_{N\beta}$ here

$$= \sum_{\alpha,\beta} \Big| \underbrace{\mathrm{tr}\left(\sqrt{\rho^{\otimes N}} P_{N\beta} D_\beta C_\alpha \sqrt{\rho^{\otimes N}}\right)}\Big|^2$$

$$\le \mathrm{tr}\left(\sqrt{\rho^{\otimes N'}}\, P_{N\beta} \sqrt{\rho^{\otimes N'}}\right)$$

$$\times\, \mathrm{tr}\left(\sqrt{\rho^{\otimes N'}}\, C_\alpha^\dagger D_\beta^\dagger D_\beta C_\alpha \sqrt{\rho^{\otimes N'}}\right)$$

Schwarz inequality

$$\le \sum_{\alpha,\beta} \underbrace{\mathrm{tr}\left(\rho^{\otimes N} P_{N\beta}\right)}_{\le \delta} \underbrace{\mathrm{tr}\left(\rho^{\otimes N} C_\alpha^\dagger D_\beta^\dagger D_\beta C_\alpha\right)}_{p_{\alpha\beta}}$$

$$\le \delta \underbrace{\sum_{\alpha,\beta} \mathrm{tr}\left(\rho^{\otimes N} C_\alpha^\dagger D_\beta^\dagger D_\beta C_\alpha\right)}_{1}$$

$$= \delta$$

Let's try to push through the same proof using average fidelity.

$$(\bar{F}^N)^2 = \left(\sum_J p_J F_J\right)^2$$

$$\le \sum_J p_J F_J^2$$

$$= \sum_J p_J \underbrace{F\left(\rho_J, \mathcal{D}^N \circ \mathcal{C}^N(\rho_J)\right)}_{\langle J| \mathcal{D}^N \circ \mathcal{C}^N(\rho_J) |J\rangle = \sum_{\alpha,\beta} |\langle J| \underset{\uparrow\, P_{N\beta}}{D_\beta} \underset{\uparrow\, P_N}{C_\alpha} |J\rangle|^2}$$

$$= \sum_J p_J \sum_{\alpha,\beta} \underbrace{|\langle J| P_{N\beta} D_\beta C_\alpha |J\rangle|^2}_{\le \langle J| P_{N\beta} |J\rangle \underbrace{\langle J| C_\alpha^\dagger D_\beta^\dagger D_\beta C_\alpha |J\rangle}_{p_{\alpha\beta|J}}}$$

$$\leq \sum_J p_J \langle J | P_{N\beta} | J \rangle \, p_{\beta|J}$$

We're stuck. If there is only one β (unitary decoding) or if all D_β map to same subspace S_N', then we have

$$(\overline{F^N})^2 \leq \sum_J p_J \langle J | P_N' | J \rangle \underbrace{\sum_\beta p_{\beta|J}}_{1}$$

$$= \operatorname{tr}(\rho^{\otimes N} P_N')$$

$$\leq \delta$$

It seems clear we should be able to do the general case, but I don't know how.