Def: The quantum entropy of a state $\rho \in D(H)$ is defined as

$$H(\rho) = -tr(\rho \ln \rho)$$

- $\rho \ln \rho = f(\rho)$ where $f(t) = t \ln t$ and $f(0) = \lim_{t \to 0} t \ln t = 0$
- Given $\rho = \sum p_i |\varphi_i\rangle\langle\varphi_i|$ orthogonal decomposition

  then $f(\rho) = \sum p_i \ln p_i \, |\varphi_i\rangle\langle\varphi_i|$

  $$H(\rho) = -tr(\rho \ln \rho) = -\sum_i p_i \ln p_i \, tr(|\varphi_i\rangle\langle\varphi_i|)$$
  $$= -\sum_i p_i \ln p_i = H(P)$$

  where $P = \{p_i\}_{i=1}^{d}$ prob. distribution of spectrum of ρ

  Why $\{p_i\}_{i=1}^{d}$ is a prob distribution?
- Also called von Neumann entropy (or simply entropy) as introduced by John von Neumann in 1930s.
- The unit is called "Qubit" Quantum bit.

Example: ① Pure state $\rho = |\varphi\rangle\langle\varphi|$ $\quad H(|\varphi\rangle\langle\varphi|) = 0 \quad \forall\, |\varphi\rangle \in H_{unit}$

② Maximally Mixed state: $\rho = \frac{1}{d} I = \sum \frac{1}{d} |i\rangle\langle i|$ where $d = \dim(H)$

$$H(\rho) = H(\{\tfrac{1}{d}, \tfrac{1}{d} \cdots \tfrac{1}{d}\}) = \log d$$

In particular, $d=2$. $\quad \rho = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad H(\frac{1}{2} I_2) = \log_2 2 = 1$ qubit

③ Embedding classical state into quantum system.

Given a classical information source P_X on $\mathcal{X}$, fix an O.N.B $\{|x\rangle\}$ in $\mathbb{C}^{\mathcal{X}}$

$$P_X \longrightarrow \rho_X = \sum P_X(x) |x\rangle\langle x|$$

Preparation process: $x \longmapsto |x\rangle\langle x|$.

Note that if $\rho = \sum \lambda_i |\varphi_i\rangle\langle\varphi_i|$ with $|\varphi_i\rangle$ not mutually orthogonal

then $H(\rho) \neq H(\{\lambda_1 \cdots \lambda_n\})$

Indeed $\rho = \frac{1}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| + \frac{1}{4}|+\rangle\langle +| + \frac{1}{4}|-\rangle\langle -|$

$= \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \qquad H(\rho) = \log 2 \neq \log 4$

(Properties of H) ① $H(\rho) \geq 0$ with equality iff ρ pure

② $H(\rho) \leq \log d_H$, - —— - iff $\rho = \frac{1}{d_H}$

③ $\sum p_i H(\rho_i) \leq H(\sum p_i \rho_i)$

Namely, entropy is a concave function

④ $H(\rho) = H(V\rho V^*)$ for any $V^*V = I$ isometry

In particular $H(\rho) = H(U\rho U^*)$ for $\forall$ U unitary

⑤ $H(\rho) = \min\limits_{\{E_y\}} H(P_Y)$

where the minimum is over all POVM $\{E_y\}$

and $P_Y(y) = tr(E_y \rho)$ is the measurement outcome.

Pf: ① & ② Note that $H(\rho) = H(P)$ where P is the distribution given by the eigenvalues of ρ. Then $0 \leq H(\rho) \leq \log d$ as $|spec(\rho)| \leq d$

$H(P) = 0 \Leftrightarrow P$ is point mass $\{1, 0 \cdots 0\}$

$\Leftrightarrow \rho = |\varphi\rangle\langle\varphi|$

$H(P) = \log d \Leftrightarrow P = \{\frac{1}{d} \cdots \frac{1}{d}\}$

$\Leftrightarrow \rho = \sum \frac{1}{d}|\varphi_i\rangle\langle\varphi_i|$ O.N.B $|\varphi_i\rangle$.

$= \frac{1}{d}$

③ (Nontrivial) Postpone

④ If $\rho = \sum p_i |\varphi_i\rangle\langle\varphi_i|$ $\quad V\rho V^* = \sum p_i V|\varphi_i\rangle\langle\varphi_i|V^*$

$$= \sum p_i |V\varphi_i\rangle\langle V\varphi_i|$$

$\langle\varphi_j|V^*V|\varphi_i\rangle = \langle\varphi_j|\varphi_i\rangle = \delta_{i,j} \Rightarrow |V\varphi_i\rangle$ O.N.B

where we used $V^*V = I$

⑤ The minimum if attained by choosing $E_i = |\varphi_i\rangle\langle\varphi_i|$ then $P_i = tr(\rho|\varphi_i\rangle\langle\varphi_i|)$

$$H(\rho) = H(P)$$

For the other direction, we need $H(\rho) \leq H(P)$ for any $P = \Lambda(\rho)$ measurement

Given $\{E_i\}$ POVM, Λ is the measurement map: $\Lambda : B(H) \to \mathbb{C}^{|\Lambda|}$

$$\rho \longrightarrow P_i = tr(\rho E_i)$$

Joint Entropy.

Def The joint entropy of a joint state $\rho_{AB} \in D(H_A \otimes H_B)$ is defined as

$$H(AB)_\rho = H(\rho_{AB})$$

Example: ① Product state $\rho_{AB} = \rho_A \otimes \rho_B$

$$H(AB)_\rho = H(\rho_A \otimes \rho_B) = H(\rho_A) + H(\rho_B) = H(A)_\rho + H(B)_\rho$$

where $H(A)_\rho = H(\rho_A)$, $\rho_A = tr_B(\rho_{AB})$ similarly for $H(B) = H(\rho_B)$

② Classical-quantum state

$$\rho_{XB} = \sum p_X(x)|x\rangle\langle x| \otimes \rho_x \qquad H(XB)_\rho = H(X) + \sum_x p_X(x) H(\rho_x)$$

Indeed, suppose $\rho_x = \sum p_{x,y} |\varphi_{x,y}\rangle\langle\varphi_{x,y}|$

$$\rho_{XB} = \begin{bmatrix} p_1\rho_1 & & & \\ & p_2\rho_2 & & \\ & & \ddots & \\ & & & p_n\rho_n \end{bmatrix} = \begin{bmatrix} \sum p_{1,y}|\varphi_{1,y}\rangle\langle\varphi_{1,y}| & & \\ & \ddots & \\ & & \ddots \end{bmatrix}$$

$$H(\rho_{XB}) = -\sum p_{x,y} \log p_{x,y}$$

$$= H(XY) = H(X) + H(Y|X) = H(X) + \sum p_X(x)\, H(P_{Y|X=x})$$

$$= H(X) + \sum p_X(x)\, H(\rho_x)$$

③ Pure state $\quad \rho_{AB} = |\varphi\rangle_{AB}\langle\varphi_{AB}|$

$$H(AB)_\rho = H(\rho_{AB}) = 0$$

Conditional Entropy

Def: Given $\rho_{AB} \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$, the conditional entropy of ρ conditional on B is

$$H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$$

Example: ① Classical case. P_{XY} joint distribution

$$\rightsquigarrow \quad \rho_{AB} = \sum p_{x,y}\, |x\rangle\langle x| \otimes |y\rangle\langle y| \quad \text{for any O.N.B } \{|x\rangle\} \subseteq \mathcal{H}_A,\ \{|y\rangle\} \subseteq \mathcal{H}_B$$

$$H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$$

$$= H(XY)_\rho - H(Y)_\rho = H(X|Y)_\rho$$

② Semi classical case: $\quad \rho_{XB} = \sum p_x |x\rangle\langle x| \otimes \rho_x^B$

$$H(B|X)_\rho = H(BX)_\rho - H(X) = \sum p_x H(\rho_x)$$

③ Entangled pure state: $\quad \rho_{AB} = |\varphi_{AB}\rangle\langle\varphi_{AB}|$ and $|\varphi_{AB}\rangle \neq |\varphi_A\rangle \otimes |\varphi_B\rangle$

$$H(A|B)_\rho = H(AB)_\rho - H(B)_\rho = 0 - H(B)_\rho \leq 0 \quad !$$

Negative entropy ~ Negative uncertainty
(How to intepret this? An final project)

Property: ① $-H(B) \leq H(A|B) \leq H(A)$

② $H(B|X) \geq 0$, $H(X|B) \geq 0$ for cq state $\rho_{XB} = \sum p_x |x\rangle\langle x| \otimes \rho_x^B$

② $H(A|BC) \leq H(A|B)$ conditioning does not increase uncertainty

Mutual Information:

$H(A) - H(A|B) = H(A) - (H(AB) + H(B)) \geq 0$. $\Rightarrow H(A) + H(B) - H(AB) \geq 0$

Def The mutual Information of ρ_{AB} between A and B is

$$I(A:B)_\rho = H(A)_\rho + H(B)_\rho - H(AB)_\rho$$

$$H(A|BC) \leq H(A|B)$$

$$H(ABC) - H(BC) \leq H(AB) - H(B) \Rightarrow \underbrace{H(AB) + H(BC) - H(B) - H(ABC)}_{\geq 0}$$

Def: The conditional mutual information of ρ_{ABC} between A and C conditional on B is

$$I(A:C|B)_\rho := H(AB) + H(BC) - H(B) - H(ABC) \geq 0$$

Highly nontrivial
final project on this.