

Lecture 12 — Quantum Entropy

①

The Q notion of entropy corresponding to the Shannon entropy is the von Neumann entropy:

$$H(A) = -\text{Tr}\{\rho^A \log \rho^A\}$$
 ← often written $S(\rho)$ or $H(\rho)$

If we diagonalize ρ^A ,

$$\rho^A = \sum_x \lambda_x |\phi_x\rangle\langle\phi_x|$$

then $\log \rho^A = \sum_x \log(\lambda_x) |\phi_x\rangle\langle\phi_x|$

$$\begin{aligned} \Rightarrow H(A) &= \text{Tr}\left\{-\sum_x \lambda_x \log(\lambda_x) |\phi_x\rangle\langle\phi_x|\right\} \\ &= -\sum_x \lambda_x \log \lambda_x. \end{aligned}$$

So the von Neumann entropy is the Shannon entropy of the orthogonal ensemble. One can show that the minimum Shannon entropy over all POVMs with rank one POVM elements is precisely $H(A)$ (and achieved by the projective measurement $\{|\phi_x\rangle\langle\phi_x|\}$).

Properties

- i) Positivity $H(A) \geq 0$
- ii) Minimum value $H(A)=0$ iff ρ_A is a pure state $\rho_A = |\psi\rangle\langle\psi|$.
- iii) Maximum value $H(A) = \log D$ where $D=\dim H$ and ρ_A is the maximally mixed state π .
- iv) Concavity $\rho = \sum_x p_x(x) \rho_x$
 $\Rightarrow H(\rho) \geq \sum_x p_x(x) H(\rho_x)$
- v) Unitary invariance $H(\rho) = H(U\rho U^\dagger)$.
 Also true for isometries!

We can define joint entropy the same way:

$$H(A, B) = -\text{Tr} \{ \rho_{AB} \log \rho_{AB} \}$$

We can also define marginal entropies using the reduced density matrices:

$$\rho_A = \text{Tr}_B \{ \rho_{AB} \} \quad H(A) = -\text{Tr} \{ \rho_A \log \rho_A \}$$

But with quantum entropy, we can have

(3)

$H(A) > H(A, B)$. E.g.,

$$\rho_{AB} = |\Phi^+\rangle\langle\Phi^+| \Rightarrow H(AB) = 0$$

$$\rho_A = \text{Tr}_B \{\rho_{AB}\} = I/2 = \pi \Rightarrow H(A) = 1.$$

This is very different from Shannon entropy, and underlies many of the unintuitive properties of QIT.

On product states entropy is additive:

$$H(\rho \otimes \sigma) = H(\rho) + H(\sigma)$$

~~On separable states entropy is~~
~~additive~~

Conditional Entropy

The density matrix ρ_A plays a role analogous to the prob. density $p_{\mathcal{X}}(x)$ in classical entropy; but there is no real analogue to the conditional probability $p_{\mathcal{X}|\mathcal{Y}}(x|y)$. How, then, can we define quantum conditional entropy? One idea would be to condition on the outcome of a measurement $\{M_x\}$.

We could then choose the M_x to act on one subsystem only (say B): $M_x \equiv I_A \otimes N_x$ ④

$$\Rightarrow P_x = \text{Tr} \{ N_x^+ N_x P_B \}, \quad P_x = \frac{\text{Tr}_B \{ I_A \otimes N_x P_{AB} I_A \otimes N_x^+ \}}{P_x}$$

$$\Rightarrow H(A|B)_{\{M_x\}} = \sum_x P_x H(P_x)$$

The problem with this kind of definition is that it is strongly dependent on the choice of $\{M_x\}$, and doesn't give us the kind of measurement-independent properties we expect from a conditional entropy. (An idea like this has been used to define a quantity called the discord.)

Instead, we will take an expression for the classical conditional entropy ⑤

$$H(A|B) = H(AB) - H(B)$$

to be the definition of quantum conditional entropy.

This definition retains many of the properties of classical conditional entropy.

For example, $H(A) \geq H(A|B)$. For classical states $P_{AB} = \sum_{x,y} P_{x,y} |x\rangle\langle x| \otimes |y\rangle\langle y|$ if it agrees with the usual definition.

However, it differs in one crucial respect: ⑤
it can be negative.

Example: $P_{AB} = |\Psi_+\rangle\langle\Psi_+| \Rightarrow P_B = I/2$

$$H(A|B) = H(AB) - H(B) = 0 - 1 = -1.$$

Negative conditional entropy is a hallmark of quantum correlations (entanglement). In fact, this is used to define a measure of quantum correlations: the coherent information.

$$I(A>B) \equiv H(B) - H(AB) = -H(A|B)$$

The angle is to indicate that this is not symmetric between A & B.

If $I(A>B)$ is positive then quantum correlations are present.

Note that, positive or negative, there is a bound on the magnitude of $H(A|B)$:

$$|H(A|B)| \leq \log d_A.$$

We can also define conditional coherent info:

$$\begin{aligned} I(A>BC) &= I(A>B|C) \equiv H(B|C) - H(ABC|C) \\ &= (H(BC) - H(C)) - (H(ABC) - H(C)) \\ &= H(BC) - H(ABC). \end{aligned}$$

Quantum Mutual Information

We again make use of a classical formula for our definition:

$$I(A;B) \equiv H(A) + H(B) - H(AB)$$

This expression is manifestly symmetric: $I(A;B) = I(B;A)$. Unlike coherent information, it is also positive: $I(A;B) \geq 0$. We can relate it to coherent info:

$$\begin{aligned} I(A;B) &= H(A) + I(A>B) \\ &= H(B) + I(B>A). \end{aligned}$$

↑
This
follows
from ~~subadditivity~~
subadditivity.

$$\text{Example: } P_{AB} = |\Phi^+\rangle\langle\Phi^+|$$

$$\begin{aligned} I(A;B) &= H(A) + H(B) - H(AB) \\ &= 1 + 1 - 0 = 2. \end{aligned}$$

This is remarkable! Classically, the mutual info between 2 bits cannot be more than 1. This result of 2 is because an ebit can be used to transmit 2 classical bits by superdense coding.

Holevo information

⑦

Remember the accessible information?

$$\rho = \sum_x p_x(x) \rho_x, \quad p_y(y) = \text{Tr} \{ \Lambda_y \rho \},$$

$$p_{Y|X}(y|x) = \text{Tr} \{ \Lambda_y \rho_x \}$$

$$\Rightarrow I_{\text{acc}}(\underbrace{\{p_x(x), \rho_x\}}_{\text{ensemble } \mathcal{E}}) = \max_{\{\Lambda_y\}} I(X; Y)$$

POVMs

Because of the maximization this quantity is hard to compute. But there is an upper bound called the Holevo information:

$$X(\epsilon) \equiv H(\rho) - \sum_x p_x(x) H(\rho_x)$$

$$= H\left(\sum_x p_x(x) \rho_x\right) - \sum_x p_x(x) H(\rho_x)$$

This quantity is important in characterizing the ability to transmit classical information using quantum systems.

Conditional Q Mutual Info

$$I(A;B|C) \equiv H(A|C) + H(B|C) - H(AB|C)$$

$$= H(AC) - H(C) + H(BC) - H(C) \\ - H(ABC) + H(C)$$

$$= H(AC) + H(BC) - H(ABC) - H(C).$$

A very nonobvious property is positivity:

$$[I(A;B|C) \geq 0] \Leftrightarrow [H(A|B) \geq H(A|BC)].$$

This is equivalent to the strong subadditivity property of quantum entropy. We will sketch the proof of this next time.

Note the chain rule:

$$I(AB;C) = I(B;C|A) + I(A;C)$$

Quantum Relative Entropy

Again, we generalize the classical formula:

$$D(\rho||\sigma) \equiv \text{Tr}\{\rho \log \rho - \rho \log \sigma\}$$

Like classical relative entropy, this is a distance-like measure: it vanishes if $\rho = \sigma$ and $[D(\rho||\sigma) \geq 0]$.

Proof: IF $P = \sum_x p(x) |\phi_x \rangle \langle \phi_x|$, $Q = \sum_y q(y) |\psi_y \rangle \langle \psi_y|$,
 then with a little work we can show that

$$D(P||Q) = \sum_x p(x) (\log p(x) - \sum_y |\langle \phi_x | \psi_y \rangle|^2 \log q(y)).$$

We can treat $|\langle \phi_x | \psi_y \rangle|^2$ as a conditional probability

~~$p(\phi_x | \psi_y) = P(\phi_x | \psi_y)^2 = P(\phi_x | \psi_y)P(\psi_y) = P(\phi_x)$~~

$p(x|y) \equiv |\langle \phi_x | \psi_y \rangle|^2$ and define a new density

$$r(x) \equiv \sum_y p(x|y) q(y). \text{ By the convexity of } -\log$$

we then get

$$-\sum_x \sum_y p(x) |\langle \phi_x | \psi_y \rangle|^2 \log(q(y)) \geq -\sum_x p(x) \log(r(x))$$

$$\Rightarrow D(P||Q) \geq \sum_x p(x) (\log p(x) - \log r(x)) \\ = D(p(x)||r(x)) \geq 0.$$

classical relative entropy \square

This immediately implies subadditivity
 $[H(A) + H(B) \geq H(AB)]$ and $[I(A;B) \geq 0]$.

Next time we will sketch the proof of strong subadditivity, and look at the important Q information inequalities.