

Monotonicity of Relative Entropy: $D(\rho \parallel \sigma) \geq D(\mathcal{N}(\rho) \parallel \mathcal{N}(\sigma))$

Consequences:

1.) $I(A; B | C) \geq 0$ (Equivalently, $H(A, B) + H(B, C) \geq H(A, B, C) + H(B)$)
↖ Last lecture

2.) Joint Convexity:

$$D(\rho \parallel \sigma) \leq \sum_x p(x) D(\rho_x \parallel \sigma_x)$$

where $\rho = \sum_x p(x) \rho_x$ and $\sigma = \sum_x p(x) \sigma_x$

← Similar property for trace distance and fidelity

Proof: Consider classical-quantum state.

$$\rho^{xA} = \sum_x p(x) |x\rangle\langle x|^X \otimes \rho_x^A$$

$$\sigma^{xA} = \sum_x p(x) |x\rangle\langle x|^X \otimes \sigma_x^A$$

$$\text{RHS} = D(\rho^{xA} \parallel \sigma^{xA}) \geq D(\rho^A \parallel \sigma^A) = D(\rho \parallel \sigma)$$

3.) Dephasing Increases Entropy:

$$\text{Let } \sigma = \sum_y |y\rangle\langle y| \rho |y\rangle\langle y|$$

$$H(\sigma) \geq H(\rho)$$

This is complete dephasing.
Equivalent to measuring in the $\{|y\rangle\}$ basis and ignoring measurement value.

Proof: Let $\Delta_y(\rho)$ be the dephasing channel.

$$D(\rho \parallel \pi) \geq D(\Delta_y(\rho) \parallel \pi) = D(\sigma \parallel \pi)$$

$$\text{Also, } D(\rho \parallel \pi) = \log d - H(\rho)$$

$$D(\sigma \parallel \pi) = \log d - H(\sigma) \quad \square$$

4.) Pinsker's Inequality:

$$\|\rho - \sigma\|_1 \leq \sqrt{2 D(\rho \parallel \sigma)} \log 2$$

← Change of units if not done in bits.

Proof: Prove first for classical binary distribution (calculus).

For general states:

Recall that $\rho - \sigma = (\rho - \sigma)_+ - (\rho - \sigma)_-$ by splitting positive and negative eigenvalues.
 such that $\|\rho - \sigma\|_1 = \|(\rho - \sigma)_+\|_1 + \|(\rho - \sigma)_-\|_1$

Let Π_+ and Π_- be the complete orthogonal projections such that

$$(\rho - \sigma)_+ = \Pi_+ (\rho - \sigma)$$

$$(\rho - \sigma)_- = \Pi_- (\rho - \sigma)$$

Measure according to $\{\Pi_+, \Pi_-\}$ and prepare states $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ respectively.

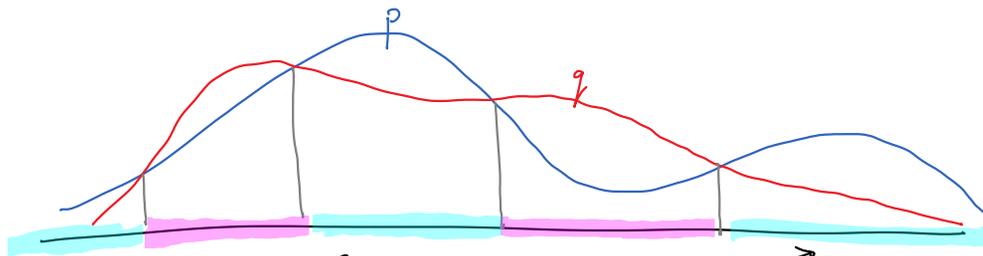
$$\text{That is, } M(\rho) = \text{Tr}(\Pi_+ \rho) |0\rangle\langle 0| + \text{Tr}(\Pi_- \rho) |1\rangle\langle 1|$$

$$M(\sigma) = \text{Tr}(\Pi_+ \sigma) |0\rangle\langle 0| + \text{Tr}(\Pi_- \sigma) |1\rangle\langle 1|$$

$$\begin{aligned} \text{Notice: } \|M(\rho) - M(\sigma)\|_1 &= |\text{Tr}(\Pi_+ \rho) - \text{Tr}(\Pi_+ \sigma)| + |\text{Tr}(\Pi_- \rho) - \text{Tr}(\Pi_- \sigma)| \\ &= |\text{Tr}(\rho - \sigma)_+| + |\text{Tr}(\rho - \sigma)_-| \\ &= \|\rho - \sigma\|_1 \end{aligned}$$

$$\begin{aligned} \text{Final step: } \|\rho - \sigma\|_1 &= \|M(\rho) - M(\sigma)\|_1 \\ &\leq \sqrt{2 D(M(\rho) \| M(\sigma))} \\ &\leq \sqrt{2 D(\rho \| \sigma)} \quad \text{by monotonicity} \end{aligned}$$

Intuition from classical proof:



$$A = \{x : p(x) \geq q(x)\}$$

$$\|\rho - q\|_1 = P(A) - Q(A) + Q(A^c) - P(A^c)$$

Indicator functions take the role of Π_+ and Π_-

Let $f(x) = \begin{cases} 0, & x \in A \end{cases}$ and p' be the dist. of $f(x)$ when $X \sim p$

Let $f(x) = \begin{cases} 0 & x \in A \\ 1 & x \notin A \end{cases}$ and p' be the dist. of $f(x)$ when $X \sim p$
 q' " " " " $f(x)$ when $X \sim q$

Then $\|p' - q'\|_1 = \|p - q\|_1$

5.) Quantum Data Processing Ineq.

Given $(A, B_1) \sim \rho^{AB_1}$ and channel $\mathcal{N}^{B_1 \rightarrow B_2}$

Let $\sigma^{AB_2} = (\mathbb{I} \otimes \mathcal{N}^{B_1 \rightarrow B_2})(\rho^{AB_1})$ and denote $(A, B_2) \sim \sigma^{AB_2}$

That is, a channel acts on B_1 to produce B_2 .

Then, $\mathcal{I}(A; B_1)_\rho \geq \mathcal{I}(A; B_2)_\sigma$ (Equivalently, $\mathcal{I}(A > B_1)_\rho \geq \mathcal{I}(A > B_2)_\sigma$)

Proof: Let E_1 be the reference system for the purification of ρ^{AB_1} ,
 and E_2 be the environment for the isometric extension of $\mathcal{N}^{B_1 \rightarrow B_2}$.

Then $\rho^{AB_1 E_1}$ is a pure state consistent with ρ^{AB_1} .

and $\sigma^{AB_2 E_1 E_2}$ is a pure state consistent with σ^{AB_2} .

Notice that $\sigma^{A E_1} = \rho^{A E_1}$.

$$\begin{aligned} \Rightarrow \mathcal{I}(A; B_1) &= H(A) + H(B_1) - H(A, B_1) \\ &= H(A) + H(A, E_1) - H(E_1) \\ &= H(A) + H(B_2, E_2) - H(A, B_2, E_2) \\ &= \mathcal{I}(A; B_2, E_2) \\ &= \mathcal{I}(A; B_2) + \mathcal{I}(A; E_2 | B_2) \\ &\geq \mathcal{I}(A; B_2) \quad \square \end{aligned}$$

Classical:
 $X - Y - Z$

$$\begin{aligned} \mathcal{I}(X; Y) &= \mathcal{I}(X; Y) + \cancel{\mathcal{I}(X; Z | Y)}^0 \\ &= \mathcal{I}(X; YZ) \\ &= \mathcal{I}(X; Z) + \mathcal{I}(X; Y | Z) \\ &\geq \mathcal{I}(X; Z). \end{aligned}$$

Accessible Information and Holevo Information

Given an ensemble $\mathcal{E} = \{(\rho(x), p_x)\}_x$

Accessible Information: $\mathcal{I}_{\text{acc}}(\mathcal{E}) = \max_{\Lambda_Y} \mathcal{I}(X; Y)$
(from Ch. 10)

Holevo Information: $\mathcal{X}(\mathcal{E}) = H(\rho) - \sum_x p(x) H(\rho_x)$

Notice that $\mathcal{X}(\mathcal{E}) = \mathcal{I}(X; A)$

where $\rho^{xA} = \sum_x p(x) |x\rangle\langle x| \otimes \rho_x^A$ ← The classical-quantum state corresponding to the ensemble.

Holevo Bound: $\mathcal{I}_{\text{acc}}(\mathcal{E}) \leq \mathcal{X}(\mathcal{E})$ ← Prove in problem set.

Continuity of Entropy:
(Fano-like Inequality)

Fannes' Inequality: $|H(\rho) - H(\sigma)| \leq 2 \|\rho - \sigma\|_1 \log d + 2 h(\|\rho - \sigma\|_1)$
binary entropy function.

Slight Improvement: $|H(\rho) - H(\sigma)| \leq \frac{1}{2} \|\rho - \sigma\|_1 \log(d-1) + h(\frac{1}{2} \|\rho - \sigma\|_1)$

Generalization: $|H(A|B)_\rho - H(A|B)_\sigma| \leq 4 \|\rho^{AB} - \sigma^{AB}\|_1 \log d_A + 2 h(\|\rho^{AB} - \sigma^{AB}\|_1)$

Uncertainty Principle:

Consider ρ^{AB} and two possible measurements on A: $\{\Lambda_x\}$ or $\{\Gamma_z\}$.

Then $H(X|B) + H(Z|B) \geq \log \frac{1}{c} + H(A|B)$
measurement state.

One of these two exist, not both.

$$\text{where } c = \max_{x, z} \left\| \sqrt{\Lambda_x} \sqrt{\Lambda_z} \right\|_{\infty}^2$$

Ch. 12: Information of a channel.

Classical setting: Let \mathcal{N} be a classical channel (a conditional pmf)

$$\mathcal{I}(\mathcal{N}) \cong \max_{p(x)} \mathcal{I}(X; Y)$$

where $X \xrightarrow{\mathcal{N}} Y$

in other words $p(x, y) = p(x) p(y|x)$

↑
channel

↑
Mutual information of a classical channel.
("capacity")

Quantum Information Quantities:

$$\text{Holevo Information: } \chi(\mathcal{N}) = \max_{\rho^{xA} \text{ separable}} \mathcal{I}(X; B)_{\rho}$$

$$\text{where } \rho^{xB} = (\mathcal{I} \otimes \mathcal{N})(\rho^{xA})$$

Notice, cannot maximize over state of A only as in $\max_{\rho^A} \mathcal{I}(A; B)$
because ρ^{AB} is undefined (don't both exist together).

$$\text{Mutual Information: } \mathcal{I}(\mathcal{N}) = \max_{\rho^{xA}} \mathcal{I}(X; B)$$

$$\text{where } \rho^{xB} = (\mathcal{I} \otimes \mathcal{N})(\rho^{xA})$$

$$\mathcal{I}(\mathcal{N}) \geq \chi(\mathcal{N})$$

$$\text{Coherent Information: } Q(\mathcal{N}) = \max_{\text{pure state}} \mathcal{I}(X > B)$$

Private Information: $P(N) = \max_{\rho^{XA}} \mathbb{I}(X;B) - \mathbb{I}(X;E)$
seperable ↑
isometric extension.

$Q(N) \leq P(N)$ with equality if degradable channel

$$\mathbb{I}(N_1 \otimes N_2) \geq \mathbb{I}(N_1) + \mathbb{I}(N_2)$$

$$\mathbb{I}(N_1 \otimes N_2) \leq \mathbb{I}(N_1) + \mathbb{I}(N_2)$$

$$\chi(N_1 \otimes N_2) \neq \chi(N_1) + \chi(N_2)$$