TODAY: DATA PROCESSING FOR QUANTUM RELATIVE ENTROPY.

Quantum channel $\mathcal{E}$

$$D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D(\rho\|\sigma).$$

Plan:
- Show that $(\rho,\sigma) \mapsto D(\rho\|\sigma)$ is convex
- Convexity $\Longrightarrow$ data processing.
- Implications.

<u>Th</u> (Joint convexity of quantum relative entropy)

The function $(\rho,\sigma) \longmapsto \overbrace{\mathrm{Tr}(\rho\log\rho - \rho\log\sigma)}^{D(\rho\|\sigma)}$

$\mathrm{Pos}(\mathcal{H})\times\mathrm{Pos}(\mathcal{H}) \longrightarrow \mathbb{R}\cup\{+\infty\}$

is convex, i.e., for any $p\in[0,1]$, $\rho_0,\rho_1,\sigma_0,\sigma_1\in\mathrm{Pos}(A)$:

$$D(p\rho_0+(1-p)\rho_1\|p\sigma_0+(1-p)\sigma_1) \leq pD(\rho_0\|\sigma_0)+(1-p)D(\rho_1\|\sigma_1).$$

Moreover:

$$(*)\quad D(\rho\|\sigma) = \sup_{\substack{\{Z_t\}_{t\geq 0}\\ Z_t\in L(\mathcal{H})}} \int_0^\infty \underbrace{\mathrm{Tr}\left[\rho\left(\frac{I}{t+1}-\frac{Z_tZ_t^*}{t}\right)-\sigma(I+Z^*)(I+Z)\right]}_{\text{linear function in } \rho,\,\sigma} dt$$

sup of linear functions

Rk: • Expression of D makes it manifestly convex as a sup of convex functions:

$$\sup_{f\in\mathcal{F}} f\big((1-p)\rho_0 + p\rho_1,\ (1-p)\sigma_0 + p\sigma_1\big) \overset{f \text{ linear}}{=} \sup_{f\in\mathcal{F}} (1-p) f(\rho_0,\sigma_0) + p f(\rho_1,\sigma_1)$$

$$\leq (1-p)\sup_{f\in\mathcal{F}} f(\rho_0,\sigma_0) + p \sup_{f\in\mathcal{F}} f(\rho_1,\sigma_1)$$

• $\log = \ln$ here.

• $(*)$ also allows one to prove data processing easily.

Proof: ρ and σ do not commute so

$$\log \rho - \log \sigma \neq \log \rho \sigma^{-1}$$

Trick: write $D(\rho\|\sigma)$ using one log.

$$\mathrm{Tr}(A \cdot B) = \langle \Phi | A \otimes B^T | \Phi \rangle \qquad |\Phi\rangle = \sum_i |i\rangle \otimes |i\rangle$$

$\{|i\rangle\}$ basis of $\mathcal{H}$.

$$D(\rho\|\sigma) = \langle \Phi | \rho \log \rho \otimes I - \rho \otimes \log(\sigma^T) | \Phi \rangle .$$

$$= \langle \Phi | (\rho \otimes I)(\log \rho \otimes I - I \otimes \log \sigma^T) | \Phi \rangle$$

$$= \langle \Phi | (\rho \otimes I) \log(\rho \otimes (\sigma^{-1})^T) | \Phi \rangle .$$

$\log(A \otimes B) = \log A \otimes I + I \otimes \log B$

$$\log(x) = \int_0^\infty \left(\frac{1}{t+1} - \frac{1}{t+x} \right) dt \quad \text{for } x > 0 .$$

$$A \otimes I \log(A \otimes B^{-1}) = (A \otimes I) \int_0^\infty \frac{1}{(t+1)(I \otimes I)} - \frac{1}{t I \otimes I + A \otimes B^{-1}} \, dt$$

$$= \int_0^\infty \left(\frac{A \otimes I}{t+1} - \frac{1}{t A^{-1} \otimes I + I \otimes B} \right) dt$$

Parallel sum: $(A : B) = \dfrac{1}{A^{-1} + B^{-1}}$ (operator harmonic mean)

Rk: Can define it for general $A, B \in \mathrm{Pos}(\mathcal{H})$. For simplity, here assume A, B invertible.

Proposition: (Variational expression for parallel sum)

For every $x \in \mathcal{H}$, $A, B \in \mathrm{Pos}(\mathcal{H})$. linear in A, B.

$$\langle x, (A:B)x\rangle = \inf\{\langle y, Ay\rangle + \langle z, Bz\rangle : y+z = x\}$$

Proof: $A:B = (A^{-1}+B^{-1})^{-1} = \left(B^{-1}(A+B)A^{-1}\right)^{-1} = A\,(A+B)^{-1}B = (A+B-B)(A+B)^{-1}B$

$$= B - B(A+B)^{-1}B.$$

$$\langle y, Ay\rangle + \langle x-y, B(x-y)\rangle - \langle x, (A:B)x\rangle$$
$$= \langle x, Bx\rangle + \langle y, (A+B)y\rangle - 2\,\mathrm{Re}\,\langle y, Bx\rangle - \langle x, (A:B)x\rangle$$
$$= \langle x, B(A+B)^{-1}Bx\rangle + \langle y, (A+B)y\rangle - 2\,\mathrm{Re}\,\langle y, Bx\rangle$$
$$= \|(A+B)^{-1/2}Bx\|^2 + \|(A+B)^{1/2}y\|^2 - 2\,\mathrm{Re}\,\langle (A+B)^{1/2}y, (A+B)^{-1/2}Bx\rangle$$
$$= \|(A+B)^{-1/2}Bx - (A+B)^{1/2}y\|^2$$
$$\geq 0$$

with equality if $y = (A+B)^{-1}Bx$ ∎

Back to our expression

$$D(\rho\|\sigma) = \langle\Phi|\int_0^\infty \left(\frac{\rho\otimes I}{t+1} - \left(\frac{\rho\otimes I}{t}\right):(I\otimes\sigma^T)\right)dt\,|\Phi\rangle$$

$$= \int_0^\infty \left(\langle\Phi|\frac{\rho\otimes I}{t+1}|\Phi\rangle - \langle\Phi|\left(\frac{\rho\otimes I}{t}\right):(I\otimes\sigma^T)|\Phi\rangle\right)dt. \quad (**)$$

- $\langle\Phi|\frac{\rho\otimes I}{t+1}|\Phi\rangle = \frac{1}{t+1}\mathrm{Tr}(\rho\cdot I^T) = \frac{1}{t+1}\mathrm{Tr}(\rho).$
- $\langle\Phi|\left(\frac{\rho\otimes I}{t}\right):(I\otimes\sigma^T)|\Phi\rangle = \inf\limits_{|z\rangle\in\mathcal{H}\otimes\mathcal{H}} \langle z, \frac{\rho\otimes I}{t} z\rangle + (\langle\Phi|-\langle z|)\, I\otimes\sigma^T(|\Phi\rangle-|z\rangle)$

Let us write $|z\rangle$ in the fixed basis $\{|i\rangle\}$. $|z\rangle = \sum_{i,j}' z_{ij} \cdot |i\rangle \otimes |j\rangle$

- $$\frac{1}{t}\langle z, \rho\otimes I z\rangle = \frac{1}{t}\sum_{\substack{i,j\\ i',j'}}' \langle i|\langle j|\, \overline{z_{ij}}\, (\rho\otimes I)\, z_{i'j'}\, |i'\rangle|j'\rangle$$

$$= \frac{1}{t}\sum_{\substack{i\,i'\\ j}}' \overline{z_{ij}}\, \langle i|\rho|i'\rangle\, z_{i'j}$$

$$= \frac{1}{t}\,\mathrm{Tr}(\rho Z Z^*) \quad \text{where} \quad Z = \sum_{ij}' z_{ij}\, |i\rangle\langle j|$$

$$Z^* = \sum_{ij}' \overline{z_{ij}}\, |j\rangle\langle i|$$

- $$\langle \Phi | I\otimes\sigma^T | \Phi\rangle = \mathrm{Tr}(\sigma)$$

- $$\langle z, I\otimes\sigma^T z\rangle = \sum_{\substack{ij\\ i'j'}}' \langle i|\langle j|\, \overline{z_{ij}}\, (I\otimes\sigma^T)\, z_{i'j'}\, |i'\rangle|j'\rangle$$

$$= \sum_{i\,j\,j'}' \overline{z_{ij}}\, z_{ij'}\, \langle j|\sigma^T|j'\rangle.$$

$$= \mathrm{Tr}(\sigma Z^* Z) \qquad Z^*Z = \sum_{j j' i}' \overline{z_{ij}}\, z_{ij'}\, |j\rangle\langle j'|$$

- $$\langle \Phi, I\otimes\sigma^T z\rangle = \sum_{ij}' z_{ij}\, \langle i|\sigma^T|j\rangle = \mathrm{Tr}(\sigma Z)$$

- $$\langle z, I\otimes\sigma^T \Phi\rangle = \sum_{ij}' \overline{z_{ij}}\, \langle j|\sigma^T|i\rangle = \mathrm{Tr}(\sigma Z^*)$$

Back to $(**)$

$$\langle\Phi|\frac{\rho\otimes I}{t+1}|\Phi\rangle - \langle\Phi|\left(\frac{\rho\otimes I}{t} : I\otimes\sigma^T\right)|\Phi\rangle$$

$$= \sup_{Z_t}\left\{ \frac{1}{t+1}\mathrm{Tr}(\rho) - \frac{1}{t}\mathrm{Tr}(\rho Z Z^*) - \mathrm{Tr}\big(\sigma\,(I + Z + Z^* + Z^*Z)\big)\right\} \blacksquare$$

From joint convexity to data processing

(Rather generic, works for many divergences).

Th (Data processing inequality for the quantum relative entropy).

$\rho, \sigma \in \mathrm{Pos}(A)$, $\mathcal{E}$ quantum channel $L(A) \to L(B)$

$$D(\mathcal{E}(\rho) \| \mathcal{E}(\sigma)) \leq D(\rho\|\sigma)$$

Proof: Consider a Stinespring dilation for $\mathcal{E}$

$$\mathcal{E}(S) = \mathrm{Tr}_E(VSV^*) \quad \text{with } V \in L(A, B\otimes E),\ V^*V = I_{\mathcal{H}}$$

$$\mathcal{E} = \mathrm{Tr}_E \circ \mathcal{V} \quad \text{where } \mathcal{V}(S) = VSV^*.$$

Claim: $D(\mathcal{V}(\rho) \| \mathcal{V}(\sigma)) = D(\rho\|\sigma)$.

$$\mathrm{Tr}(V\rho V^* \log V\rho V^* - V\rho V^* \log V\sigma V^*) = \mathrm{Tr}(V\rho \log \rho V^* - V\rho \log\sigma V^*)$$

$$= \mathrm{Tr}(\rho\log\rho - \rho\log\sigma)$$

So it suffices to analyze Tr_E map.

Let $\rho_{BE}, \sigma_{BE} \in \mathrm{Pos}(B\otimes E)$

Want to transform
$$\rho_{BE} \longrightarrow \rho_B \otimes \frac{I}{\dim E}$$
$$\sigma_{BE} \longrightarrow \sigma_B \otimes \frac{I}{\dim E}$$

$$D(\rho_{BE}\|\sigma_{BE}), \quad D\left(\rho_B\otimes\frac{I_E}{\dim E} \,\Big\|\, \sigma_B\otimes\frac{I_E}{\dim E}\right) = D(\rho_B\|\sigma_B).$$

Consider generalized Pauli operators $d = \dim E$, $[d] = \{0, \dots, \dim E - 1\}$

$$X_E|k\rangle = |k+1 \bmod d\rangle, \quad Z_E|k\rangle = e^{\frac{2\pi i k}{d}}|k\rangle$$

Consider the channel
$$\mathcal{D}(S) = \frac{1}{d^2}\sum_{\ell, m\in[d]} X_E^{\ell} Z_E^{m} S Z_E^{m*} X_E^{\ell*}$$

Then $\mathcal{D}(|k\rangle\langle k'|) = \frac{1}{d^2} \sum'_{\ell, m \in [d]} X_E^{\ell} e^{\frac{2i\pi(k-k')m}{d}} |k\rangle\langle k'| X_E^{\ell}$

$$= \begin{cases} 0 & \text{if } k \neq k' \\ \frac{1}{d} \sum'_{\ell} X_E^{\ell} |k\rangle\langle k| X_E^{\ell} = \frac{I}{d} & \text{if } k = k' \end{cases}$$

So

$$(I_B \otimes \mathcal{D})(\rho_{BE}) = \rho_B \otimes \frac{I_E}{d}$$

$$(I_B \otimes \mathcal{D})(\sigma_{BE}) = \sigma_B \otimes \frac{I_E}{d}.$$

Now use joint convexity:

$$D\Big(\overbrace{\frac{1}{d^2} \sum'_{\ell,m} X_E^{\ell} Z_E^{m} \rho_{BE} X_E^{\ell} Z_E^{m}}^{\rho_B \otimes \frac{I}{d}} \,\Big\|\, \overbrace{\frac{1}{d^2} \sum'_{\ell,m} X_E^{\ell} Z_E^{m} \sigma_{BE} X_E^{\ell} Z_E^{m}}^{\sigma_B \otimes \frac{I}{d}}\Big)$$

$$\leq \frac{1}{d^2} \sum'_{\ell,m} \underbrace{D\left(X_E^{\ell} Z_E^{m} \rho_{BE} X_E^{\ell} Z_E^{m} \,\|\, X_E^{\ell} Z_E^{m} \sigma_{BE} Z_E^{m} X_E^{\ell}\right)}_{D(\rho_{BE} \| \sigma_{BE})}$$

Conclude the proof. $\blacksquare$

Direct proof of data processing using ⊛

$$D(\rho\|\sigma) = \sup_{\substack{\{Z_t\}_{t\geq 0} \\ Z_t \in L(\mathcal{H})}} \int_0^\infty \mathrm{Tr}\left[\rho\left(\frac{I}{t+1} - \frac{Z_t Z_t^*}{t}\right) - \sigma (I+Z^*)(I+Z)\right] dt$$

$$D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) = \sup_{\substack{\{Z_t\}_{t\geq 0} \\ Z_t \in L(\mathcal{H})}} \int_0^\infty \mathrm{Tr}\left[\rho\left[\frac{\mathcal{E}^*(I)}{t+1} - \frac{\mathcal{E}^*(Z_t Z_t^*)}{t}\right] - \sigma\left[\mathcal{E}^*(I) + \mathcal{E}^*(Z^*) + \mathcal{E}^*(Z) + \mathcal{E}^*(Z^* Z)\right]\right] dt$$

<u>Claim:</u> For a unital completely positive map $\mathcal{F}$,

$$\mathcal{F}(ZZ^*) \geq \left[\mathcal{F}(Z)\right]^* \cdot \mathcal{F}(Z) \qquad \text{[Shwarz map]}$$

Indeed $T := |0\rangle\langle 0| \otimes I + |0\rangle\langle 1| \otimes Z + |1\rangle\langle 0| \otimes Z^* + |1\rangle\langle 1| \otimes Z^* Z \geq 0$

So $(I \otimes \mathcal{F})(T) \geq 0$ i.e., $\begin{bmatrix} \mathcal{F}(ZZ^*) & \mathcal{F}(Z^*) \\ \mathcal{F}(Z) & I \end{bmatrix} \geq 0$

$$\Updownarrow$$

$$\mathcal{F}(ZZ^*) \geq \mathcal{F}(Z)^* \mathcal{F}(Z)$$

So letting $Y_t = \mathcal{E}^*(Z_t)$

$$D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq \sup_{\substack{\{Y_t\}_{t\geq 0} \\ Y_t \in L(\mathcal{H})}} \int_0^\infty \mathrm{Tr}\left[\rho\left[\frac{I}{t+1} - \frac{Y_t Y_t^*}{t}\right] - \sigma\left[I + Y_t^* + Y_t + Y_t^* Y_t\right]\right] dt$$

$$= D(\rho\|\sigma).$$

Consequences

* Strong subadditivity:

$$H(A|C)_\rho + H(B|C)_\rho \leq H(AB|C)_\rho \qquad ①$$

$$\Updownarrow$$

$$H(A|BC)_\rho \leq H(A|C)_\rho \qquad ②$$

$$\Updownarrow$$

$$I(A:B|C)_\rho \geq 0 \qquad ③$$

Proof: ② $H(A|BC)_\rho = -D(\rho_{ABC} \| I_A \otimes \rho_{BC})$

$H(A|B)_\rho = -D(\rho_{AB} \| I_A \otimes \rho_B)$

Data processing: $D(\rho_{AB} \| I_A \otimes \rho_B) \leq D(\rho_{ABC} \| I_A \otimes \rho_{BC})$

② ⇒ ① Note that $H(A|BC) = H(ABC) - H(BC)$

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

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

② ⇒ ③ $I(A:B|C)_\rho = H(A|C) - H(A|BC)$.