Entropies & entropic inequalities

Recap:

- VN entropy: $H(\rho) = -\mathrm{Tr}[\rho \log(\rho)]$

  $\rightsquigarrow$ operational interpretation: Compression of q-states

- Relative entropy: $D(\rho\|\sigma) = \begin{cases} \mathrm{Tr}[\rho(\log(\rho) - \log(\sigma))] & \text{if } \ker(\sigma) \subseteq \ker(\rho) \\ +\infty & \text{else.} \end{cases}$

$$\boxed{\text{Lemma: } D(\rho\|\sigma) = \frac{1}{\ln(2)} \lim_{\varepsilon \searrow 0} \frac{1 - \mathrm{Tr}[\rho^{1-\varepsilon}\sigma^{\varepsilon}]}{\varepsilon}}$$

Central inequality:

$$\boxed{\begin{array}{l} \text{Data-processing inequality} \\ D(T(\rho)\|T(\sigma)) \leq D(\rho\|\sigma) \\ \forall \text{ q-channels } T. \end{array}}$$

Immediate consequences of data-processing:

(1) Joint convexity of relative entropy

$\rho_1, \rho_2, \sigma_1, \sigma_2 \in D(H)$, $\lambda \in [0,1]$

Define: $\rho = (1-\lambda)\rho_1 \otimes |1\rangle\langle 1| + \lambda \rho_2 \otimes |2\rangle\langle 2| \in D(H \otimes \mathbb{C}^2)$

$\sigma = (1-\lambda)\sigma_1 \otimes |1\rangle\langle 1| + \lambda \sigma_2 \otimes |2\rangle\langle 2| \in D(H \otimes \mathbb{C}^2)$

Then: $\mathrm{Tr}_2[\rho] = (1-\lambda)\rho_1 + \lambda \rho_2$

$\mathrm{Tr}_2[\sigma] = (1-\lambda)\sigma_1 + \lambda \sigma_2$

$$D((1-\lambda)\rho_1 + \lambda\rho_2 \,\|\, (1-\lambda)\sigma_1 + \lambda\sigma_2) = D(\mathrm{Tr}_2[\rho] \,\|\, \mathrm{Tr}_2[\sigma])$$

$$\overset{\text{Data-processing}}{\leq} D(\rho \| \sigma) = (1-\lambda) D(\rho_1 \| \sigma_1) + \lambda D(\rho_2 \| \sigma_2)$$

Uses that

$$\log((1-\lambda)\rho_1 \otimes |1\rangle\langle 1| + \lambda \rho_2 \otimes |2\rangle\langle 2|)$$

$$= \log(1-\lambda)\, \mathbb{1}_H \otimes |1\rangle\langle 1| + \log(\lambda)\, \mathbb{1}_H \otimes |2\rangle\langle 2|$$

$$+ \log(\rho_1) \otimes |1\rangle\langle 1| + \log(\rho_2) \otimes |2\rangle\langle 2|$$

(2) Klein's inequality

$$D(\rho\|\sigma) \underset{\substack{\uparrow \\ \text{Data-processing}}}{\geq} D(\mathrm{Tr}[\rho]\,\|\,\mathrm{Tr}[\sigma]) = 0.$$

$\rightsquigarrow$ Tomorrow: Strengthening of this inequality.

(3) Concavity of vN entropy

Note that $H(\rho) = \log(d) - D(\rho \| \mathbb{1}/d)$

Therefore: $H((1-\lambda)\rho + \lambda\sigma) = \log(d) - D((1-\lambda)\rho + \lambda\sigma \| (1-\lambda)\mathbb{1}/d + \lambda \mathbb{1}/d)$

$$\geq \log(d) - (1-\lambda) D(\rho \| \mathbb{1}/d) - \lambda D(\sigma \| \mathbb{1}/d)$$

$$= (1-\lambda) H(\rho) + \lambda H(\sigma).$$

$\square$

(4) Subadditivity of vN entropy

$$\varrho_{AB} \in \mathcal{D}(H_A \otimes H_B) \implies \ker(\varrho_A \otimes \varrho_B) \overset{\substack{\text{spectral}\\ \text{theorem}}}{=} \operatorname{span}\{ |v\rangle \otimes |w\rangle \in H_A \otimes H_B : |v\rangle \in \ker(\varrho_A) \text{ or } |w\rangle \in \ker(\varrho_B)$$

$$\subseteq \ker(\varrho_{AB})$$

If $|v\rangle \in \ker(\varrho_A)$, then $\langle v|\varrho_A|v\rangle = \sum_{i=1}^{\dim(H_B)} \underbrace{(\langle v| \otimes \langle i|)\varrho_{AB}(|v\rangle \otimes |i\rangle)}_{\geq 0} = 0$

$\implies \varrho_{AB}|v\rangle \otimes |i\rangle = 0 \quad \forall i \in \{1, \dots, \dim(H_B)\}$

$\implies \varrho_{AB}|v\rangle \otimes |w\rangle = 0 \quad \forall |w\rangle \in H_B$

$\implies |v\rangle \otimes |w\rangle \in \ker(\varrho_{AB}) \quad \forall |w\rangle \in H_B.$

Same argument works if $|w\rangle \in \ker(\varrho_B)$.

$$D(\varrho_{AB} \| \varrho_A \otimes \varrho_B) = \operatorname{Tr}\Big[\varrho_{AB}\big(\log(\varrho_{AB}) - \underbrace{\log(\varrho_A \otimes \varrho_B)}_{= \log(\varrho_A) \otimes \Pi_{\operatorname{Im}(\varrho_B)} + \Pi_{\operatorname{Im}(\varrho_A)} \otimes \log(\varrho_B)}\big)\Big]$$

$$= -H(\varrho_{AB}) + H(\varrho_A) + H(\varrho_B)$$

≥ 0 by Klein's inequality.

$\square$

(5) Strong subadditivity of vN entropy

As before, compute

$$D\left(\rho_{ABC} \,\Big\|\, \frac{\mathbb{1}_{H_A}}{d_A} \otimes \rho_{BC}\right) = \mathrm{Tr}\left[\rho_{ABC}\left(\log(\rho_{ABC}) - \underbrace{\log\left(\frac{\mathbb{1}_{H_A}}{d_A} \otimes \rho_{BC}\right)}_{\substack{= -\log(d_A)\mathbb{1}_{H_A} \otimes \Pi_{\mathrm{Im}(\rho_{BC})} \\ + \mathbb{1}_{H_A} \otimes \log(\rho_{BC})}}\right)\right]$$

$$= -H(\rho_{ABC}) + \log(d_A) + H(\rho_{BC})$$

$$D\left(\rho_{AB} \,\Big\|\, \frac{\mathbb{1}_{H_A}}{d_A} \otimes \rho_B\right) = -H(\rho_{AB}) + \log(d_A) + H(\rho_B)$$

Data-processing inequality

$$\Longrightarrow \quad D\left(\rho_{AB} \,\Big\|\, \frac{\mathbb{1}_{H_A}}{d_A} \otimes \rho_B\right) \leq D\left(\rho_{ABC} \,\Big\|\, \frac{\mathbb{1}_{H_A}}{d_A} \otimes \rho_{BC}\right)$$

$$\Longrightarrow \quad \boxed{H(\rho_{ABC}) + H(\rho_B) \leq H(\rho_{AB}) + H(\rho_{BC})}$$

$\square$

Exercise 3: (conditional entropy)

For $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ define

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

(1) "Classically, discarding systems does not increase the entropy"

Recall product rule: $p_{AB}(x,y) = p_B(y)\, p_{A|B}(x|y)$

$$H(p_{AB}) = -\sum_{x,y} p_B(y)\, p_{A|B}(x|y) \log\left(p_B(y)\, p_{A|B}(x|y)\right)$$

$$= -\sum_{x,y} p_B(y)\, p_{A|B}(x|y) \left(\log(p_B(y)) + \log(p_{A|B}(x|y))\right)$$

$$= H(p_B) + \underbrace{\sum_{y} p_B(y)\, H(p_{A|B}(\cdot|y))}_{\geq 0} \geq H(p_B)$$

$$\Longrightarrow H(A|B)_{\rho_{AB}} \geq 0. \qquad \square$$

(2) "Quantumly, discarding systems does increase the entropy in general"

$\varrho_{AB} = \omega_d \in D(\mathbb{C}^d \otimes \mathbb{C}^d)$ normalized maximally ent. state.

ω_d is pure $\Longrightarrow$ $H(\omega_d) = 0$

$\varrho_B = \mathrm{Tr}_2[\omega_d] = \frac{\mathbb{1}_d}{d}$ $\Longrightarrow$ $H(\varrho_B) = \log(d)$

$$H(A|B)_{\varrho_{AB}} = -\log(d) < 0 .$$

$\square$

(3) "Conditioning does not increase the vN entropy"

$$H(A|B,C)_{\rho_{ABC}} = H(\rho_{ABC}) - H(\rho_{BC})$$

$$H(A|B)_{\rho_{ABC}} = H(\rho_{AB}) - H(\rho_B)$$

SSA $\Longrightarrow$ $H(\rho_{ABC}) + H(\rho_B) \leq H(\rho_{AB}) + H(\rho_{BC})$

$$\longrightarrow \boxed{H(A|B,C)_{\rho_{ABC}} \leq H(A|B)_{\rho_{ABC}}}$$

(4) Cool application of conditional entropy

Thm: (Quantum Shearer's inequality)

For $t \in \mathbb{N}$ consider a family $\mathcal{F} = \{F_1, \dots, F_K\}$ of subsets of $\{1, \dots, n\}$ s.th.

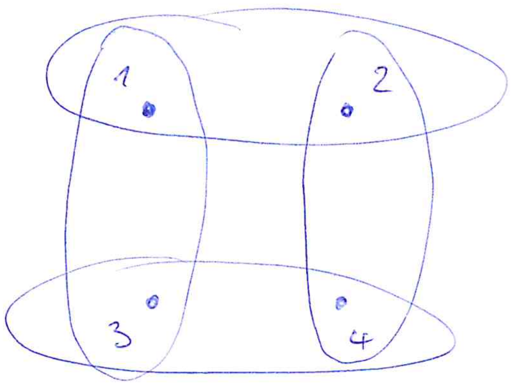
$$\left| \left\{ k \in \{1, \dots, K\} : i \in F_k \right\} \right| = t \qquad \forall i \in \{1, \dots, n\}.$$

Then:

$$H(\rho_{1,2,\dots,n}) \leq \frac{1}{t} \sum_{F \in \mathcal{F}} H(\rho_F)$$

for any $\rho_{1,2,\dots,n} \in \mathcal{D}(H_1 \otimes \dots \otimes H_n)$.

Example:



$$H(\rho_{1,2,3,4}) \leq \frac{1}{2}\Big(H(\rho_{1,2}) + H(\rho_{2,4}) + H(\rho_{3,4}) + H(\rho_{1,3}) \Big)$$

Proof: Consider $F \subseteq \{1,2,\dots,n\}$ and denote its elements by $(i_1,\dots,i_\ell)$ s.th. $i_1 < i_2 < \dots < i_\ell$.

$$\sum_{j=1}^{|F|} H(i_j \mid i_1,\dots,i_{j-1})_{\rho_{1,2,\dots,n}} = H(\rho_{i_1}) + H(i_2 \mid i_1)_{\rho_{1,\dots,n}} + H(i_3 \mid i_1, i_2)_{\rho_{1,\dots,n}} + \dots$$

$$= \cancel{H(\rho_{i_1})} + \cancel{H(\rho_{i_1 i_2})} - \cancel{H(\rho_{i_1})} + H(\rho_{i_1 i_2 i_3}) - \cancel{H(\rho_{i_1 i_2})} + \dots$$

$$= H(\rho_{i_1 \dots i_{|F|}}) = H(\rho_F).$$

$$\frac{1}{t} \sum_{F \in \mathcal{F}} H(\rho_F) = \frac{1}{t} \sum_{F \in \mathcal{F}} \sum_{j=1}^{|F|} H(i_j^F \mid i_1^F, \dots, i_{j-1}^F)_{\rho_{1,\dots,n}}$$

Conditioning decreases entropy

$$\geq \frac{1}{t} \sum_{F \in \mathcal{F}} \sum_{j=1}^{|F|} H(i_j^F \mid 1, \dots, i_j^F - 1)_{\rho_{1,\dots,n}}$$

Every i_j^F occurs in t F's

$$= \frac{1}{\cancel{t}} \cdot \cancel{t} \sum_{i=1}^{n} H(i \mid 1, \dots, i-1)_{\rho_{1,\dots,n}} = H(\rho_{1,\dots,n})$$

$\square$

Another example: $\mathcal{F} = \mathcal{F}_k = \{ F \subseteq \{1,\dots,n\} : |F| = k \}$

$\rightsquigarrow$ all k-element subsets $\Longrightarrow$ $t = \binom{n-1}{k-1} = \frac{k}{n}\binom{n}{k}$

Quantum Shearer's inequality $\longrightarrow$ $\frac{k}{n} H(\rho_{1,\dots,n}) \leq \frac{1}{\binom{n}{k}} \sum_{F\in\mathcal{F}_k} H(\rho_F)$

Application: Entropy production

Depolarizing channel : $T_p = (1-p)\, id + p \operatorname{Tr}[\cdot] \mathbb{1}/d$

$$T_p^{\otimes n}(\rho) = \sum_{k=0}^{n} \sum_{F\in\mathcal{F}_k} (1-p)^k p^{n-k} \left(\bigotimes_{\ell\in F^c} (\mathbb{1}/d)_\ell \otimes \rho_F \right)$$

How much entropy does $T_p^{\otimes n}$ produce?

$$\underline{\underline{H(T_p^{\otimes n}(\rho))}} \overset{\text{Concavity}}{\geq} \sum_{k=0}^{n} \sum_{F\in\mathcal{F}_k} (1-p)^k p^{n-k} \underbrace{H\left(\bigotimes_{\ell\in F^c} (\mathbb{1}/d)_\ell \otimes \rho_F\right)}_{=(n-k)\log(d) + H(\rho_F)}$$

$$\geq \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} (n-k)\log(d) + \sum_{k=0}^{n} (1-p)^k p^{n-k} \underbrace{\sum_{F\in\mathcal{F}_k} H(\rho_F)}_{\underset{\text{Q-Shearer}}{\geq} \binom{n}{k}\frac{k}{n} H(\rho) \,=\, \binom{n-1}{k-1} H(\rho)}$$

$$\underline{\underline{\geq pn\log(d) + (1-p) H(\rho)}}.$$

Use
$\sum_{k=0}^{n} \binom{n}{k}(n-k)(1-p)^k p^{n-k}$
$= \sum_{k=0}^{n} \binom{n-1}{k} n (1-p)^k p^{n-k}$
$= n(1-p)^n + \left(\sum_{k=0}^{n-1} \binom{n-1}{k} (1-p)^k p^{n-k-1}\right) np$
$\geq np$

Exercise 4: (Fannes' inequality)

Want to quantify the continuity of $\rho \longmapsto H(\rho)$.

Thm: For $\rho, \sigma \in D(H)$ s.th. $\|\rho-\sigma\|_1 \leq \frac{1}{2}$ we have

$$|H(\rho) - H(\sigma)| \leq \|\rho-\sigma\|_1 \log(\dim(H)) + \eta(\|\rho-\sigma\|_1) \qquad (*)$$

where $\eta(x) = -x\log(x)$.

Two remarks:

(1) $\eta(x) \longrightarrow 0$ as $x \to 0$ $\Longrightarrow$ $(*)$ quantifies the continuity of vN entropy

(2) RHS only depends on logarithm of $\dim(H)$.

$\Longrightarrow$ Even if $d \sim$ exponentially large, then $(*)$ is still useful.

Proof: Set $d = \dim(H)$

$$\varrho = \sum_{i=1}^{d} \lambda_i |v_i\rangle\langle v_i| \quad \& \quad \sigma = \sum_{j=1}^{d} \mu_j |w_j\rangle\langle w_j| \qquad \text{spectral decomposition}$$

s.th. $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ and $\mu_1 \geq \mu_2 \geq \dots \geq \mu_d$.

Step 1: Show that $\|\varrho - \sigma\|_1 \geq \sum_i |\lambda_i - \mu_i|$.

Write $\varrho - \sigma = P - Q$ for $P, Q \in \mathcal{B}(H)^+$ s.th. $PQ = 0$ (Jordan-Hahn decomposition)

$$\Longrightarrow \|\varrho - \sigma\|_1 = \mathrm{Tr}\left[(\varrho - \sigma)^+\right] = \mathrm{Tr}[P] + \mathrm{Tr}[Q].$$

Define $X := P + \varrho = Q + \sigma \Longrightarrow P + Q = 2X - \varrho - \sigma$

$\rightsquigarrow$ Spectral decomposition: $X = \sum_{k=1}^{d} \tau_k |u_k\rangle\langle u_k|$

s.th. $\tau_1 \geq \dots \geq \tau_d$.

Show: $\boxed{\tau_i \geq \max(\lambda_i, \mu_i) \quad \forall i}$

$\Downarrow$

$2\tau_i \geq \lambda_i + \mu_i + |\lambda_i - \mu_i|$ $\forall i$

$$\Longrightarrow \|\varrho - \sigma\|_1 = \mathrm{Tr}[P + Q]$$
$$= \mathrm{Tr}[2X] - \mathrm{Tr}[\varrho] - \mathrm{Tr}[\sigma]$$
$$= \sum_i 2\tau_i - \lambda_i - \mu_i \geq \sum_i |\lambda_i - \mu_i|.$$

Left to show: $\tau_i \geq \max(\lambda_i, \mu_i) \quad \forall i$

Let's show $\tau_i \geq \lambda_i \rightsquigarrow \tau_i \geq \mu_i$ follows in the same way.

Setup: $X = \sum_k \tau_k |u_k\rangle\langle u_k| \qquad \tau_1 \geq \tau_2 \geq \dots \geq \tau_d$

$$X = P + \varrho \geq \varrho. \quad (*)$$

$\varrho = \sum_i \lambda_i |v_i\rangle\langle v_i| \qquad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$

Consider two subspaces of H

$\mathcal{S}_\ell = \mathrm{span}\{|v_1\rangle, \dots, |v_\ell\rangle\} \rightsquigarrow \dim(\mathcal{S}_\ell) = \ell$

$\mathcal{M}_\ell = \mathrm{span}\{|u_\ell\rangle, \dots, |u_d\rangle\} \rightsquigarrow \dim(\mathcal{M}_\ell) = d - \ell + 1.$

$\dim(\mathcal{M}_\ell) + \dim(\mathcal{S}_\ell) = d - \ell + 1 + \ell = d + 1 > d$

$\Longrightarrow \exists |y\rangle \in \mathcal{S}_\ell \cap \mathcal{M}_\ell, \quad \langle y|y\rangle = 1$

Two expansions: $|y\rangle = \sum_{j=\ell}^{d} \alpha_j |u_j\rangle = \sum_{i=1}^{\ell} \beta_i |v_i\rangle \qquad \sum_j |\alpha_j|^2 = \sum_i |\beta_i|^2 = 1.$

$$\Longrightarrow \quad \tau_\ell \overset{\text{ordering}}{\geq} \sum_{j=\ell}^{d} |\alpha_j|^2 \tau_j = \langle y|X|y\rangle \overset{(*)}{\geq} \langle y|\varrho|y\rangle = \sum_{i=1}^{\ell} |\beta_i|^2 \lambda_i \geq \lambda_\ell.$$

$\square$

Step 2: Show that $|H(\varrho) - H(\sigma)| \leq \sum_i \eta(|\lambda_i - \mu_i|)$

Calculus $\Longrightarrow$ $|\eta(\lambda_i) - \eta(\mu_i)| \leq \eta(|\lambda_i - \mu_i|)$ $(*)$

with $(*)$ we have

$$|H(\varrho) - H(\sigma)| = \left|\sum_i \eta(\lambda_i) - \eta(\mu_i)\right| \leq \sum_i |\eta(\lambda_i) - \eta(\mu_i)| \leq \sum_i \eta(|\lambda_i - \mu_i|)$$

Step 3: Show that $|H(\varrho)-H(\sigma)| \leq \|\varrho-\sigma\|_1 \log(d) + \eta(\|\varrho-\sigma\|_1)$

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

$$S := \sum_{i=1}^{d} |\lambda_i - \mu_i| \overset{\text{Step 1}}{\leq} \|\varrho-\sigma\|_1 .$$

$$\boxed{\eta(|\lambda_i-\mu_i|) = S\,\eta(|\lambda_i-\mu_i|/S) - |\lambda_i-\mu_i| \log(S)}$$

$$\rightsquigarrow S\,\eta\left(\frac{|\lambda_i-\mu_i|}{S}\right) = -\cancel{S}\cdot\frac{|\lambda_i-\mu_i|}{\cancel{S}}\Big(\log(|\lambda_i-\mu_i|) - \log(S)\Big) = \eta(|\lambda_i-\mu_i|) + |\lambda_i-\mu_i|\cdot\log$$

$$|H(\varrho)-H(\sigma)| \overset{\text{Step 2}}{\leq} \sum_i \eta(|\lambda_i-\mu_i|) = \underbrace{S\sum_{i=1}^{d}\eta\left(\frac{|\lambda_i-\mu_i|}{S}\right)}_{=H(p)\leq\log(d)} - \underbrace{\sum_{i=1}^{d}|\lambda_i-\mu_i|\log(S)}_{-S\log(S)=\eta(S)}$$

$$p_i = \frac{|\lambda_i-\mu_i|}{S}$$

$$\leq S\log(d) + \eta(S)$$

$$\overset{\|\varrho-\sigma\|_1\leq\frac{1}{2}}{\leq} \|\varrho-\sigma\|_1\log(d) + \eta(\|\varrho-\sigma\|_1) .$$

η monotone increasing

$\square$