1.1

a) Define $f(x) = x - \ln x - 1$

Then $f'(x) = 1 - \frac{1}{x} \begin{cases} < 0 & \text{when } 0 < x < 1 \\ > 0 & \text{when } x > 1 \end{cases}$

$\Rightarrow$ $f(x)$ has single minimum at $x = 0$.

$\Rightarrow$ $f(x) \geqslant f(1) = 0$

$\Rightarrow$ $\ln x \leqslant x - 1$. $\square$.

b) KL-divergence.

$$D(p||q) = \sum_x p(x) \ln\left(\frac{p(x)}{q(x)}\right) = -\sum_x p(x) \ln\left(\frac{q(x)}{p(x)}\right)$$

$$\because \ln\left(\frac{q(x)}{p(x)}\right) \leqslant \frac{q(x)}{p(x)} - 1$$

$$\begin{aligned} \therefore D(p||q) &\geqslant -\sum_x p(x)\left(\frac{q(x)}{p(x)} - 1\right) \\ &= -\sum_x (q(x) - p(x)) \\ &= 0 \end{aligned}$$

$\square$.

c) $D(\hat{\rho}\|\hat{\sigma}) = -\mathrm{tr}(\hat{\rho}\ln\hat{\sigma}) - H(\hat{\rho})$

$H(\rho) = -\mathrm{tr}(\hat{\rho}\ln\hat{\rho})$.

$\{|i\rangle\}$ and $\{|\tilde{i}\rangle\}$: eigenbasis of $\hat{\rho}$ and $\hat{\sigma}$.

$\mathrm{tr}(\hat{\rho}\ln\hat{\rho}) = \sum_i p_i \ln p_i$

$$\mathrm{tr}(\hat{\rho}\ln\hat{\sigma}) = \sum_i \langle i|\hat{\rho}\ln\hat{\sigma}|i\rangle$$

$$= \sum_i p_i \langle i|\ln\hat{\sigma}|i\rangle$$

$$= \sum_i \sum_j p_i \langle i|\tilde{j}\rangle\langle\tilde{j}|\ln\hat{\sigma}|i\rangle$$

$$= \sum_{ij} p_i \ln q_j \langle i|\tilde{j}\rangle\langle\tilde{j}|i\rangle$$

$$= \sum_i p_i \Big(\sum_a \underbrace{\langle i|\tilde{a}\rangle\langle\tilde{a}|i\rangle}_{D_{ia} = |\langle i|a\rangle|^2} \ln q_a\Big)$$

D_{ia} has nonnegative entries .

$\sum_i D_{ia} = \langle a|a\rangle = 1 = \sum_a D_{ia} = \langle i|i\rangle$.

$\Rightarrow$ D_{ia} is a doubly stochastic matrix . $\square$.

d) Jensen's inequality:

$$\ln(\mathbb{E}(X)) \geqslant \mathbb{E}[\ln(X)].$$

$$\Rightarrow \ln\Big(\sum_a D_{ia} q_a\Big) \geqslant \sum_a D_{ia} \ln q_a .$$

Equality holds iff $\mathbb{E}(X) = X$.

$$\Leftrightarrow \sum_a D_{ia} q_a = q_{a_0}$$

$$\Leftrightarrow \exists a_0, \ D_{ia_0} = 1 .$$

e) $$\ln\Big(\sum_a D_{ia} q_a\Big) \geqslant \sum_a D_{ia} \ln q_a .$$

Thus:

$$D(\hat{\rho} \| \hat{\sigma}) = \sum_i p_i \Big(\ln p_i - \sum_a D_{ia} \ln q_a\Big)$$

$$\geqslant \sum_i p_i \Big[\ln p_i - \ln\Big(\underbrace{\sum_a D_{ia} q_a}_{r_i}\Big)\Big]$$

$$= \sum_i p_i \ln\Big(\frac{p_i}{r_i}\Big)$$

$$= D(p \| r) \qquad \square.$$

Equality holds iff $\exists a, D_{ia} = 1$.

f) If r is a probability distribution, then $D(p\|r) \geq 0$, with equality iff $p=r$.

$\because \sum_i r_i = \sum_{ia} D_{ia} q_a = \sum_a q_a = 1$, $r_i \geq 0$.

$\therefore$ r is prob. distribution.

$\Rightarrow D(\hat{\rho}\|\hat{\sigma}) \geq D(p\|r) \geq 0$.

Equality: $r=p \Leftrightarrow r_i = p_i$.

$$
\begin{aligned}
r_i = \sum_a D_{ia} q_a &= \sum_a \langle i|a\rangle\langle a|i\rangle q_a \\
&= \sum_a \langle i|a\rangle\langle a|\hat{\sigma}|i\rangle \\
&= \langle i|\hat{\sigma}|i\rangle \\
&= p_i = \langle i|\hat{\rho}|i\rangle
\end{aligned}
$$

First equality holds iff $\exists a$, $D_{ia}=1$.

$\Leftrightarrow$ $r_i = q_a = \langle i|\hat{\sigma}|i\rangle = p_i \Leftrightarrow |\hat{a}\rangle = |i\rangle$, $q_a = p_i$.

$\Leftrightarrow$ In diagonal representation, $\hat{\rho}$ and $\hat{\sigma}$ have the same eigenvalues as well as the same eigenstates.

$\Leftrightarrow$ $\hat{\rho} = \hat{\sigma}$. $\square$.

1.2

a) $H(\rho_{AB}) \leq H(\rho_A) + H(\rho_B) \Leftrightarrow I(A;B) \geq 0$.

Notice for $\hat{\rho}_x = \hat{\rho}_A \otimes \hat{\rho}_B$, we have :

$$H(\hat{\rho}_x) = -tr[(\hat{\rho}_A \otimes \hat{\rho}_B)\ln(\hat{\rho}_A \otimes \hat{\rho}_B)]$$
$$= -\sum_{ij} p_{Ai}\, p_{Bj} \ln(p_{Ai}\, p_{Bj})$$
$$= -\sum_{ij} p_{Ai}\, p_{Bj} \ln p_{Ai} - \sum_{ij} p_{Ai}\, p_{Bj} \ln p_{Bj}$$
$$= H(\rho_A) + H(\rho_B)$$

Then the relative entropy :

$$D(\hat{\rho}_{AB} \| \hat{\rho}_x) = -tr(\hat{\rho}_{AB} \ln \hat{\rho}_x) + tr(\hat{\rho}_{AB} \ln \hat{\rho}_{AB})$$
$$= -\sum_{ij} \langle ij | \hat{\rho}_{AB} | ij \rangle \ln(p_{Ai}\, p_{Bj}) - H(\rho_{AB})$$
$$= -\sum_{i} \langle i | \hat{\rho}_A | i \rangle \ln p_{Ai} - \sum_{j} \langle j | \rho_B | j \rangle \ln p_{Bj} - H(\hat{\rho}_{AB})$$
$$= H(\hat{\rho}_A) + H(\hat{\rho}_B) - H(\hat{\rho}_{AB})$$
$$\geq 0 .$$

The equality holds when $\hat{\rho}_{AB} = \hat{\rho}_x = \hat{\rho}_A \otimes \hat{\rho}_B$.

b) $H(\sum_x p_x \hat{\rho}_x) \geq \sum_x p_x H(\hat{\rho}_x)$

Two ways to prove this:

①

$$H(\underbrace{\sum_x p_x \hat{\rho}_x}_{\hat{\rho}}) = -\sum_x p_x \operatorname{tr}(\rho_x \ln \hat{\rho})$$

$$= \sum_x p_x [\underbrace{D(\hat{\rho}_x \| \hat{\rho})}_{\geq 0} + H(\hat{\rho}_x)]$$

$$\geq \sum_x p_x H(\hat{\rho}_x)$$

Equality holds when $\hat{\rho}_x = \hat{\rho}$. □.

② Consider $\hat{\rho}_{AB} = \sum_x p_x (\hat{\rho}_x)_A \otimes (|x\rangle\langle x|)_B$.

$$\hat{\rho}_A = \sum_x p_x \hat{\rho}_x = \hat{\rho}$$

$$\hat{\rho}_B = \sum_x p_x |x\rangle\langle x| .$$

From subadditivity of Von Neumann entropy, we have:

$$H(\rho_{AB}) \leq H(\hat{\rho}_A) + H(\hat{\rho}_B) \quad \cdots \quad (*)$$

RHS: $H(\hat{\rho}_A) + H(\hat{\rho}_B) = H(\hat{\rho}) - \sum_x p_x \ln p_x$

Denote $\hat{\rho}_A^x = \hat{\rho}_x$, $\hat{\rho}_B^x = |x\rangle\langle x|$, with $\{|a\rangle\}_A$ as a complete basis in A, while $\{|b\rangle\}_B$ is just $\{|x\rangle\}_B$.

$\rho_B^x = |x\rangle\langle x|$ is diagonal in $\{|b\rangle\}_B$.

LHS: $H(\hat{\rho}_{AB}) = -\mathrm{tr}\left[\sum_x p_x \rho_A^x \otimes \rho_B^x \ln\left(\sum_y p_y \rho_A^y \otimes \rho_B^y\right)\right]$

$$= -\sum_x p_x \sum_{ab} \langle ab| \rho_A^x \otimes \rho_B^x \ln\left(\sum_y p_y \rho_A^y \otimes \rho_B^y\right)\Big]|ab\rangle$$

$$= -\sum_{ab} p_b \langle ab| \rho_A^b \otimes I_B \ln\left(\sum_y p_y \rho_A^y \otimes \rho_B^y\right)|ab\rangle .$$

Note that based on Taylor expansion and projective measurement effect, we have:

$$\langle x|_B \ln\left(\sum_y p_y \rho_A^y \otimes (|y\rangle\langle y|)_B\right)|x\rangle_B = \ln(p_y \rho_A^x)$$

Thus LHS of $(*)$ is:

$$H(\hat{\rho}_{AB}) = -\sum_{ab} p_b \langle a| \rho_A^b \ln(p_b \rho_A^b)|a\rangle$$

$$= -\sum_b p_b \,\mathrm{tr}\left[\rho_A^b \ln(p_b \rho_A^b)\right]$$

$$= \sum_x p_x \left[H(\rho_x) - \underbrace{\mathrm{tr}(\rho_A^x)}_{=1}\ln p_x\right]$$

$$= \sum_x p_x H(\rho_x) - \sum_x p_x \ln p_x .$$

Therefore, with RHS and LHS in (*), we have:

$$H\left(\sum_x p_x \hat{\rho}_x\right) \geq \sum_x p_x H(\hat{\rho}_x).$$

c) Equality holds iff:

$$\sum_x p_x (\hat{\rho}_x)_A \otimes (|x\rangle\langle x|)_B = \left[\sum_x p_x (\hat{\rho}_x)_A\right] \otimes \sum_y p_y (|y\rangle\langle y|)_B.$$

If all p_x are non-zero, we should have:

$$(\hat{\rho}_x)_A \otimes (|x\rangle\langle x|)_B = \sum_y p_y (\hat{\rho}_y)_A \otimes (|x\rangle\langle x|)_B$$

$$\Leftrightarrow \quad \hat{\rho}_x = \sum_y p_y \hat{\rho}_y, \qquad p_y > 0 \text{ for all } y.$$

$$\Leftrightarrow \quad \hat{\rho}_i = \hat{\rho}_j \text{ for } \forall\, i, j$$

$\Leftrightarrow$ $\hat{\rho}_x$ are all identical.

1.3 Monotonicity of Quantum relative entropy.

$$D(\hat{\rho}_A \| \hat{\sigma}_A) \leq D(\hat{\rho}_{AB} \| \hat{\sigma}_{AB}).$$

a) Consider tripartite system ABC.

Consider $\hat{\rho}_X = \hat{\rho}_A \otimes \hat{\rho}_{BC} = \hat{\sigma}_{ABC}$

Then $D(\hat{\rho}_{ABC} \| \hat{\sigma}_{ABC}) \geq D(\hat{\rho}_{AB} \| \hat{\sigma}_{AB})$.

Note $D(\hat{\rho}_{ABC} \| \hat{\sigma}_{ABC}) = -\mathrm{tr}(\hat{\rho}_{ABC} \ln \hat{\sigma}_{ABC}) - H(\rho_{ABC})$

$= H(A) + H(BC) - H(ABC)$.

$D(\hat{\rho}_{AB} \| \hat{\sigma}_{AB}) = -\mathrm{tr}(\hat{\rho}_{AB} \ln \hat{\sigma}_{AB}) - H(\hat{\rho}_{AB})$

$= -\sum_{ij} \langle ij | \rho_{AB} | ij \rangle \ln p_i^A p_j^B - H(AB)$.

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

Thus:

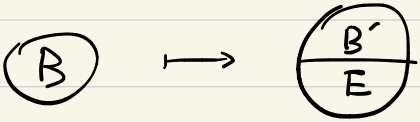
$H(A) + H(BC) - H(ABC) \geq H(A) + H(B) - H(AB)$

$\Rightarrow H(AB) + H(BC) - H(ABC) - H(B) \geq 0$.

$\Leftrightarrow I(A;BC) - I(A;B) = I(A;C|B) \geq 0$.

b). $D(\mathcal{N}(\hat{\rho}) \| \mathcal{N}(\hat{\sigma})) \leq D(\hat{\rho} \| \hat{\sigma})$

Proof: Dilation of $\mathcal{N}_{B \to B'}$: $U_{B \to B'E}$ (isometry).



Isometry keeps the eigenvalues of density operators unchanged.

$$\Rightarrow \quad E.V.(\hat{\rho}_B) = E.V.(\hat{\rho}_{B'E}).$$

$$\Rightarrow \quad H(\hat{\rho}_B) = H(\hat{\rho}_{B'E})$$

$$D[\mathcal{N}(\hat{\rho}) \| \mathcal{N}(\hat{\sigma})] = -\mathrm{tr}(\mathcal{N}(\hat{\rho}) \ln \mathcal{N}(\hat{\sigma})) - \underbrace{H(\mathcal{N}(\hat{\rho}))}_{= H(\hat{\rho}_{B'})}.$$

$$\begin{aligned} D(\hat{\rho} \| \hat{\sigma}) &= -\mathrm{tr}(\hat{\rho} \ln \hat{\sigma}) - H(\hat{\rho}) \\ &= -\mathrm{tr}(\hat{\rho} \ln \hat{\sigma}) - H(\hat{\rho}_{B'E}). \end{aligned}$$

Note that the Stinespring dilation is an isometry on B', we have:

$$\mathrm{tr}_B(\hat{\rho}_B \ln \hat{\sigma}_B) = \mathrm{tr}_{B'E}(\hat{\rho}_{B'E} \ln \hat{\sigma}_{B'E}) \quad \cdots \quad ①$$

① comes from the fact that $U_{B\to B'E}$ preserves the eigenvalues, thus we can use the same set of corresponding basis to perform trace in B and $B'E$, we have:

$$U \rho \cdot \ln\sigma U^{\dagger} = \rho_{B'E} \ln \sigma_{B'E}, \text{ then}$$

$$\mathrm{tr}(\rho_{B'E} \ln \sigma_{B'E}) = \mathrm{tr}(U\rho \cdot \ln\sigma U^{\dagger}) = \mathrm{tr}(\rho \ln \sigma).$$

$$\Rightarrow D(\hat{\rho}\|\hat{\sigma}) = -\mathrm{tr}(\hat{\rho}_{B'E} \ln \hat{\sigma}_{B'E}) - H(\hat{\rho}_{B'E})$$

$$= D(\hat{\rho}_{B'E} \| \hat{\sigma}_{B'E})$$

$$\geq D(\hat{\rho}_{B'} \| \hat{\sigma}_{B'}) \quad \cdots (\text{Monotonicity of relative entropy}).$$

$$\Rightarrow D(\mathcal{N}(\hat{\rho}) \| \mathcal{N}(\hat{\sigma})) \leq D(\hat{\rho}\|\hat{\sigma}).$$

$\square$.

1.4

Separable state has "parts that are less disordered than the whole".

$$\rho_{AB} = \sum_a p_a |\psi_a\rangle\langle\psi_a| \otimes |\varphi_a\rangle\langle\varphi_a|$$

$$= \sum_j r_j |e_j\rangle\langle e_j|$$

$$\rho_A = \sum_a p_a |\psi_a\rangle\langle\psi_a| = \sum_\mu s_\mu |f_\mu\rangle\langle f_\mu|$$

$\{|f_\mu\rangle\}$ is the othonormal eigenbasis of ρ_A.

From HJW theorem:

$$\sqrt{r_j}\,|e_j\rangle = \sum_a V_{ja}\sqrt{p_a}\,|\psi_a\rangle \otimes |\varphi_a\rangle$$

$$\sqrt{p_a}\,|\psi_a\rangle = \sum_\mu U_{a\mu}\sqrt{s_\mu}\,|f_\mu\rangle$$

$$\Rightarrow \sqrt{r_j}\,|e_j\rangle = \sum_a V_{ja}\left(\sum_\mu U_{a\mu}\sqrt{s_\mu}\,|f_\mu\rangle\right) \otimes |\varphi_a\rangle$$

$$\Rightarrow \sqrt{r_j r_{j'}}\, \delta_{jj'} = \sum_{\mu\alpha'} \sum_{\mu'\alpha'} V^{\dagger}_{\alpha' j'} \langle \varphi_{\alpha'} | \otimes U^{\dagger}_{\mu'\alpha'} \langle f_{\mu'} | \sqrt{s_{\mu'}} \cdot \sqrt{s_\mu}$$

$$\times V_{j\alpha} U_{\alpha\mu} | f_\mu \rangle \otimes | \varphi_\alpha \rangle$$

$$\Rightarrow$$

$$r_j = \sum_\mu \left[\sum_{\alpha\alpha'} V_{j\alpha} V^{\dagger}_{\alpha' j} U_{\alpha\mu} U^{\dagger}_{\mu\alpha'} \langle \varphi_{\alpha'} | \varphi_\alpha \rangle \right] s_\mu$$

$$= \sum_\mu D_{j\mu} s_\mu .$$

Here $\sum_j D_{j\mu} = \sum_{\alpha\alpha'} \underbrace{\left(\sum_j V^{\dagger}_{\alpha' j} V_{j\alpha} \right)}_{\delta_{\alpha\alpha'}} U_{\alpha\mu} U^{\dagger}_{\mu\alpha'} \langle \varphi_{\alpha'} | \varphi_\alpha \rangle$

$$= \sum_\alpha U^{\dagger}_{\mu\alpha} U_{\alpha\mu}$$

$$= 1$$

$$\sum_\mu D_{j\mu} = \sum_{\alpha\alpha'} V^{\dagger}_{\alpha' j} V_{j\alpha} \underbrace{\left(\sum_\mu U_{\alpha\mu} U^{\dagger}_{\mu\alpha'} \right)}_{\delta_{\alpha\alpha'}} \langle \varphi_{\alpha'} | \varphi_\alpha \rangle$$

$$= \sum_\alpha V_{j\alpha} V^{\dagger}_{\alpha j}$$

$$= 1$$

And note that $D^*_{j\mu}$ just switch the dummy variable $\alpha \longleftrightarrow \alpha'$, thus $D_{j\mu} = D^*_{j\mu}$ is real.

Notice:

$$D_{j\mu} = \left\| \sum_{\alpha} V_{j\alpha} U_{\alpha\mu} |\varphi_\alpha\rangle \right\| \geqslant 0 .$$

Thus $D_{j\mu}$ is a stochastic matrix, $\lambda(\rho_{AB}) \prec \lambda(\rho_A)$. □.

1.5

a) $\frac{d}{d\lambda} tr(A^n) = n\, tr(\frac{dA}{d\lambda} A^{n-1})$

Proof:
$$\frac{d}{d\lambda} tr(A^n) = \frac{d}{d\lambda} \sum_i \langle i| A^n |i\rangle$$
$$= \sum_i \langle i| \frac{d}{d\lambda} A^n |i\rangle$$
$$= \sum_i \langle i| \sum_k A^{k-1} \frac{dA}{d\lambda} A^{n-k} |i\rangle$$
$$= \sum_{k=1}^{n} tr(A^{k-1} \frac{dA}{d\lambda} A^{n-k})$$

Trace property: $tr(ABC) = tr(BCA)$

We have:
$$\frac{d}{d\lambda} tr(A^n) = \sum_{k=1}^{n} tr(\frac{dA}{d\lambda} A^{n-1}) = n\, tr(\frac{dA}{d\lambda} A^{n-1}) .$$

b) Define $\rho(\lambda) = \rho + \lambda\delta\rho$

Then $S(\rho(\lambda)) = S(\lambda) = -tr(\rho_\lambda \ln \rho_\lambda)$

$$\Rightarrow \quad \frac{dS}{d\lambda} = -\frac{d}{d\lambda} tr(\rho_\lambda \ln \rho_\lambda)$$

$$= -tr\left[(\ln\rho_\lambda + I)\cdot \frac{d\rho_\lambda}{d\lambda}\right]$$

$$= -tr\left[\frac{d\rho_\lambda}{d\lambda}\ln\rho_\lambda + \frac{d\rho_\lambda}{d\lambda}\right]$$

$$= \frac{S(\rho_{\lambda+d\lambda}) - S(\rho_\lambda)}{d\lambda}$$

Calculate $\frac{dS}{d\lambda}$ at $\lambda = 1$:

$$\delta S = S(\rho') - S(\rho) = -tr[\delta\rho\ln\rho + \delta\rho] \quad , \quad tr(\delta\rho) = 0 .$$

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

$$= -tr(\rho'\ln\rho) + tr(\rho\ln\rho)$$

$$= tr(\rho' K) - tr(\rho K)$$

$$= \delta\langle K\rangle$$

$$= \delta\langle H\rangle / T .$$

$$\Rightarrow \quad T\delta S = \delta\langle H\rangle = \delta E . \qquad \square .$$

We have used for $\rho = \frac{e^{-K}}{Z}$, $\ln\rho = -K - \ln Z$,
and $\mathrm{tr}(\rho' \ln Z) = \mathrm{tr}(\rho \ln Z) = \ln Z$.