Part III, Lent 2019
Quantum Information Theory

Example Sheet 3 Solutions

Exercise 1. Prove that the set of separable states form a closed, convex set.

Solution. Let

$$\rho_{AB} = \sum_{i=1}^{n} r_i \rho_A^{(i)} \otimes \rho_B^{(i)}$$

and

$$\sigma_{AB} = \sum_{i=1}^{m} s_i \sigma_A^{(i)} \otimes \sigma_B^{(i)}$$

be separable states, where $\sum_i r_i = \sum_i s_i = 1$ and each $\rho_A^{(i)}, \rho_B^{(i)}, \sigma_A^{(i)}$, and $\sigma_B^{(i)}$ are density matrices. Let $t \in (0,1)$. We need to show $t\rho_{AB} + (1-t)\sigma_{AB}$ is separable too (we know it is a density matrix by the previous example sheet). We have

$$\begin{aligned} t\rho_{AB} + (1-t)\sigma_{AB} &= \sum_{i=1}^{n} t r_i \rho_A^{(i)} \otimes \rho_B^{(i)} + sum_{i=1}^{m} (1-t) s_i \sigma_A^{(i)} \otimes \sigma_B^{(i)} \\ &= \sum_{i=1}^{n+m} p_i \omega_A^{(i)} \otimes \omega_B^{(i)} \end{aligned}$$

where

$$p_i = \begin{cases} t r_i & 1 \le i \le n \\ (1-t) s_{i-n} & n+1 \le i \le n+m \end{cases}$$

and similarly

$$\omega_A^{(i)} = \begin{cases} \rho_A^{(i)} & 1 \le i \le n \\ \sigma_A^{(i-n)} & n+1 \le i \le n+m \end{cases}$$

and the same for $A \to B$. Since each $\omega_A^{(i)}$, and $\omega_B^{(i)}$ is a density matrix and $\sum_{i=1}^{n+m} p_i = \sum_{i=1}^{n} t r_i + sum_{i=1}^{m} (1-t) s_i = t + 1 - t = 1$, this shows $t\rho_{AB} + (1-t)\sigma_{AB}$ is separable.

Next, to see the set $\mathrm{SEP}(A:B)$ of separable states on $\mathcal{H}_A \otimes \mathcal{H}_B$ is closed, we note

$$\mathrm{SEP}(A:B) = \bigcap_{\Lambda:\mathcal{B}(\mathcal{H}_A)\to\mathcal{B}(\mathcal{H}_B) \text{ positive}} \{\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) : (\Lambda \otimes \mathrm{id}_B)\rho_{AB} \geq 0\}.$$

Denoting $\mathcal{P}_{AB}$ as the set of positive semidefinite operators on $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we have

$$\{\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) : (\Lambda \otimes \mathrm{id}_B)\rho_{AB} \geq 0\} = (\Lambda\otimes\mathrm{id}_B)^{-1}(\mathcal{P}_{AB})\cap\mathcal{D}(\mathcal{H}_A\otimes\mathcal{H}_B)$$

is the preimage of $\mathcal{P}_{AB}$ under $\Lambda \otimes \mathrm{id}_B$ intersected with $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Since $\mathcal{P}_{AB}$ is closed and $\Lambda \otimes \mathrm{id}_B$ is continuous (as a linear map on a finite-dimensional space), the preimage is closed too. Since $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is closed and the intersection of closed is closed, we have $\{\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) : (\Lambda \otimes \mathrm{id}_B)\rho_{AB} \geq 0\}$ is a closed set too. Finally, using again that the (arbitrarily large) intersection of closed sets is closed, $\mathrm{SEP}(A:B)$ is closed.

You might wonder: why is $\mathcal{P}_{AB}$ closed? By definition,

$$\begin{aligned}\mathcal{P}_{AB} &= \{X_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) : \langle\psi, X_{AB}\psi\rangle \geq 0 \quad \forall|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B\} \\ &= \bigcap_{|\psi\rangle\in\mathcal{H}_A\otimes\mathcal{H}_B} \{X_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) : \langle\psi, X_{AB}\psi\rangle \geq 0\}\end{aligned}$$

which again is the intersection of closed sets (using $[0, +\infty)$ is closed in $\mathbb{R}$, and the set in question is the preimage of that interval under the linear (and hence continuous) map $f : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathbb{R}$ defined by $f(X_{AB}) = \langle\psi, X_{AB}\psi\rangle$).

Exercise 2. Consider the 2-qubit family of Werner states defined as follows:

$$\rho_{AB} = p|\Phi_-\rangle\langle\Phi_-| + \frac{1-p}{4}(I \otimes I).$$

Find the range of values of p for which it is entangled.

Solution. Since $\rho_{AB} \in \mathcal{D}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is a 2-qubit density matrix, it suffices to check the PPT criterion: namely, ρ_{AB} is entangled iff $(T \otimes \mathrm{id})\rho_{AB} \not\geq 0$, where T is the transpose.

We can write a matrix representation of $|\Phi_-\rangle\langle\Phi_-|$ as

$$|\Phi_-\rangle\langle\Phi_-| = \frac{1}{2}\left(\begin{array}{cc|cc} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right)$$

which, taking partial transpose yields

$$(T \otimes \mathrm{id})|\Phi_-\rangle\langle\Phi_-| = \frac{1}{2}\left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

We will use the fact that a matrix is positive-semidefinite iff it is Hermitian and has all non-negative eigenvalues[1]. To compute the eigenvalues of $(T \otimes \mathrm{id})\rho_{AB}$ (which is manifestly Hermitian as the linear combination of the above Hermitian matrix and a multiple of the identity), we first compute the eigenvalues of $(T \otimes \mathrm{id})|\Phi_-\rangle\langle\Phi_-|$. We can see it has a block diagonal structure (different than the block structure induced by the tensor product),

$$(T \otimes \mathrm{id})|\Phi_-\rangle\langle\Phi_-| = \frac{1}{2}\left(\begin{array}{c|cc|c} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

from which we can see the eigenvalues of this matrix are $\frac{1}{2}$ (with multiplicity 2) along with the eigenvalues of $\frac{1}{2}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$, which we can easily check are $\pm\frac{1}{2}$. Hence, the eigenvalues of $p|\Phi_-\rangle\langle\Phi_-|$ are $\frac{p}{2}$ with multiplicity 3, and $-\frac{p}{2}$ with multiplicity 1. Since adding a multiple of the identity only shifts the eigenvalues, we see that the eigenvalues of $(T \otimes \mathrm{id})\rho_{AB}$ are $\frac{p}{2} + \frac{1-p}{4} = \frac{1+p}{4}$ with multiplicity 3, and $-\frac{p}{2} + \frac{1-p}{4} = \frac{1-3p}{4}$ with multiplicity 1. We have $\frac{1-3p}{4} < 0$ whenever $p > \frac{1}{3}$. Hence, ρ_{AB} is entangled iff $p > \frac{1}{3}$.

Exercise 3. Prove that for any $d \times d$ complex matrix A, we have the identity

$$\mathrm{tr}[A]\frac{I}{d} = \frac{1}{d^2}\sum_{k,m=0}^{d-1} X^k Z^m A Z^{m\dagger} X^{k\dagger}$$

[1]The proof of this is fairly short, so it's worth verifying this fact for yourself if it's new to you!

where I is the identity matrix, and X and Z are the Heisenberg-Weyl operators satisfying $X|j\rangle = |j \oplus 1\rangle$, where $\oplus$ is addition mod d, and $Z|j\rangle = \exp(2\mathrm{i}j\pi/d)|j\rangle$.

Solution. We will use the identity

$$\sum_{j=0}^{d-1} \exp(2\mathrm{i}jm\pi/d) = d\delta_{m,0} = \begin{cases} 0 & m \neq 0 \pmod d \\ d & m = 0 \pmod d. \end{cases}$$

To see this, note if $m = 0 \pmod d$, we are simply summing 1 d times since $\exp(2\mathrm{i}\pi j) = 1$ for any integer j. On the other hand, if $m \neq 0$, we are summing the first d terms of the sequence $\{\exp(2\mathrm{i}m\pi/d)^j\}_{j=0}^{\infty}$ which is a geometric series and has the formula

$$\sum_{j=0}^{d-1} \exp(2\mathrm{i}m\pi/d)^j = \frac{1 - \exp(2\mathrm{i}m\pi)}{1 - \exp(2\mathrm{i}m\pi/d)} = \frac{0}{1 - \exp(2\mathrm{i}m\pi/d)} = 0.$$

Then

$$\sum_{m=0}^{d-1} Z^m|i\rangle\langle j|Z^{m\dagger} = \exp(2\mathrm{i}\pi(i-j)m/d)|i\rangle\langle j| = \delta_{i,j}d|i\rangle\langle j|.$$

Hence, writing $A = \sum_{i,j=0}^{d-1} a_{ij}|i\rangle\langle j|$, we have

$$\begin{aligned} \frac{1}{d^2}\sum_{k,m=0}^{d-1} X^k Z^m A Z^{m\dagger} X^{k\dagger} &= \frac{1}{d}\sum_{k,j=0}^{d-1} a_{jj} X^k|j\rangle\langle j|X^{k\dagger} \\ &= \frac{1}{d}\sum_{j=0}^{d-1} a_{jj} \sum_{k=0}^{d-1} |j+k\rangle\langle j+k| \\ &= \frac{1}{d}\sum_{j=0}^{d-1} a_{jj} I \\ &= \frac{1}{d} I \operatorname{tr}[A] \end{aligned}$$

as desired.

Exercise 4. Given two qubit states ρ and ω with Bloch vectors $\vec{r}$ and $\vec{s}$ respectively, show that

$$\|\rho - \omega\|_1 = \|\vec{r} - \vec{s}\|_2.$$

The two norm of a vector $\vec{v} = (v_x, v_y, v_z)$ is $\|\vec{v}\|_2 = \sqrt{v_x^2 + v_y^2 + v_z^2}$.

Solution. We have

$$\rho = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma}), \qquad \omega = \frac{1}{2}(I + \vec{s}\cdot\vec{\sigma}),$$

so

$$\|\rho - \sigma\|_1 = \frac{1}{2}\|(\vec{r}-\vec{s})\cdot\vec{\sigma}\|_1 = \frac{1}{2}\left\|\begin{pmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{pmatrix}\right\|_1$$

where $\vec{v} = \vec{r} - \vec{s}$, just as from exercise 1 from example sheet 2. We can use that for any self-adjoint matrix A, we have $\|A\|_1 = \mathrm{tr}[|A|] = \sum_j |\lambda_j|$ where λ_j are the eigenvalues of A. Thus, we simply need to calculate the eigenvalues of the above 2×2 matrix. We find the eigenvalues are

$$\lambda_\pm = \pm\sqrt{v_x^2 + v_y^2 + v_z^2} = \pm\|\vec{r} - \vec{s}\|_2.$$

Thus, $\|\rho - \sigma\|_1 = \frac{1}{2}(|\lambda_-| + |\lambda_+|) = \|\vec{r} - \vec{s}\|_2$.

Exercise 5. One Fuchs and van de Graaf inequality.

1. Show that for two pure *qubit* states ψ and ϕ, we have that

$$D(\psi, \phi) = \sqrt{1 - F(\psi, \phi)^2}.$$

   where D is the trace distance, and F the fidelity, defined in the lectures.
2. Using this, show that for any two mixed qubit states ρ and σ,

$$D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}.$$

Solution. 1. For two pure states, we have $F(\psi, \phi) = |\langle\psi|\phi\rangle|$. We may write

$$|\phi\rangle = \cos(\theta)|\psi\rangle + \sin(\theta)|\psi^\perp\rangle \tag{1}$$

for some vector $|\psi^\perp\rangle$ orthogonal to $|\psi\rangle$. In these terms, $F(\psi, \phi) = |\cos\theta|$. To calculate the trace distance, we want to simplify the form of $\psi - \phi = |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$. Using (1), we find

$$\phi = \cos^2(\theta)\,\psi + \sin(\theta)\cos(\theta)|\psi^\perp\rangle\langle\psi| + \cos(\theta)\sin(\theta)|\psi\rangle\langle\psi^\perp| + \sin^2(\theta)\,\psi^\perp.$$

In the basis $\{|\psi\rangle, |\psi^\perp\rangle\}$, we find

$$\psi - \phi = \begin{pmatrix} 1 - \cos^2(\theta) & -\sin(\theta)\cos(\theta) \\ -\sin(\theta)\cos(\theta) & -\sin^2(\theta) \end{pmatrix}$$

which has eigenvalues $\pm|\sin\theta|$. Thus,

$$\frac{1}{2}\|\psi-\phi\|_1 = |\sin\theta|.$$

Since $\sqrt{1-\cos^2(\theta)} = |\sin\theta|$, we therefore have

$$D(\psi,\phi) = \sqrt{1-\cos^2(\theta)} = \sqrt{1-F(\psi,\phi)^2}.$$

Remark. In fact, this proof holds for qudits.

2. Let $\rho = \rho_A$ have purification ϕ_{AR} and $\sigma = \sigma_A$ have purification ψ_{AR} such that
$$F(\rho_A,\sigma_A) = |\langle\phi_{AR}|\psi_{AR}\rangle|$$
using Uhlmann's theorem. Then
$$\frac{1}{2}\|\psi_{AR}-\phi_{AR}\|_1 = \sqrt{1-F(\psi_{AR},\phi_{AR})} = \sqrt{1-F(\rho_A,\sigma_A)}. \qquad (2)$$
Since the partial trace is a linear CPTP map, and therefore a quantum operation, we may use the monotonicity of the trace distance under quantum operations to find
$$\frac{1}{2}\|\psi_{AR}-\phi_{AR}\|_1 \geq \frac{1}{2}\|\rho_A-\sigma_A\|_1$$
which concludes the proof by (2).

Exercise 6. Show that for any two states ρ and σ and any unitary U,

$$F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho,\sigma)$$

by using the fact that for any positive semi-definite operator A and unitary U, we have $\sqrt{UAU^\dagger} = U\sqrt{A}U^\dagger$.

Solution. We have

$$\begin{aligned}
F(U\rho U^\dagger, U\sigma U^\dagger) &= \|\sqrt{U\rho U^\dagger}\sqrt{U\sigma U^\dagger}\|_1 \\
&= \|U\sqrt{\rho}U^\dagger U\sqrt{\sigma}U^\dagger\|_1 \\
&= \|U\sqrt{\rho}\sqrt{\sigma}U^\dagger\|_1 \\
&= \max_V |\operatorname{tr}[VU\sqrt{\rho}\sqrt{\sigma}U^\dagger]| \\
&= \max_V |\operatorname{tr}[U^\dagger VU\sqrt{\rho}\sqrt{\sigma}]| \\
&= \max_{UVU^\dagger} |\operatorname{tr}[U^\dagger UVU^\dagger U\sqrt{\rho}\sqrt{\sigma}]| \\
&= \max_{UVU^\dagger} |\operatorname{tr}[\sqrt{\rho}\sqrt{\sigma}]| \\
&= \|\sqrt{\rho}\sqrt{\sigma}\|_1 \\
&= F(\rho, \sigma).
\end{aligned}$$

We use the characterization of the trace distance as a maximum over unitary operators V, and along the way proved that the trace norm was invariant under unitary conjugation. We also used that maximizing over all V was the same as maximizing over all $UVU^\dagger$, since

$$\{UVU^\dagger : V \text{ unitary}\} = \{V : V \text{ unitary}\}.$$

Exercise 7. Show that the fidelity is jointly concave. That is, given any finite probability distribution $\{p_i\}_{i=1}^n$ and quantum states ρ_i and σ_i for $i = 1, \ldots, n$,

$$F\left(\sum_{i=1}^n p_i\rho_i, \sum_{i=1}^n p_i\sigma_i\right) \geq \sum_{i=1}^n p_i F(\rho_i, \sigma_i).$$

Hint: Choose purifications ϕ_i of ρ_i and ψ_i of σ_i such that $F(\rho_i, \sigma_i) = \langle\phi_i|\psi_i\rangle$, using Uhlmann's theorem.

Solution. As the hint suggests, we will take purifications ϕ_i of ρ_i and ψ_i of σ_i such that $F(\rho_i, \sigma_i) = \langle\phi_i|\psi_i\rangle$. Uhlmann's theorem gives that there exist purifications with $F(\rho_i, \sigma_i) = |\langle\phi_i|\psi_i\rangle|$, but we may always multiply one of them by the required phase to their inner product is real[2].

[2]I.e. if $\langle\phi_i|\psi_i\rangle = e^{i\theta}|\langle\phi_i|\psi_i\rangle|$ then we take $|\tilde{\psi}_i\rangle = e^{-i\theta}|\psi_i\rangle$ instead of $|\psi_i\rangle$, which is still a purification of σ_i

Then consider the states

$$|\Psi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle \otimes |i\rangle, \qquad |\Phi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle \otimes |i\rangle.$$

We notice

$$\langle\Phi|\Psi\rangle = \sum_i p_i \langle\phi_i|\psi_i\rangle = \sum_i p_i F(\rho_i, \sigma_i).$$

On the other hand,

$$\Phi = \sum_{i,j} \sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j| \otimes |i\rangle\langle j|,$$

and if we trace out the last system, we have

$$\sum_i p_i |\psi_i\rangle\langle\psi_i|.$$

Since each ψ_i is a purification of ρ_i, we can trace out again to find

$$\sum_i p_i \rho_i.$$

Therefore, Φ is a purification of $\sum_{i=1}^n p_i\rho_i$ (where the purifying reference is the tensor product of the purifying space for the ϕ_i's with the additional space spanned by $|i\rangle$ we added to make Φ). Similarly, Ψ is a purification of $\sum_{i=1}^n p_i\sigma_i$. Thus,

$$\sum_i p_i F(\rho_i, \sigma_i) = \langle\Phi, \Psi\rangle \leq |\langle\Phi, \Psi\rangle| \leq \max_{\phi,\psi} |\langle\phi, \psi\rangle| = F\left(\sum_{i=1}^n p_i\rho_i, \sum_{i=1}^n p_i\sigma_i\right),$$

where the maximum is over all purifications ϕ of $\sum_{i=1}^n p_i\rho_i$ and ψ of $\sum_{i=1}^n p_i\sigma_i$, and the final equality is by Uhlmann's theorem, again.

Exercise 8. Show that the fidelity is *multiplicative under tensor products*. That is,

$$F(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) = F(\rho_1, \sigma_1)F(\rho_2, \sigma_2).$$

Hint: first show that $\|A \otimes B\|_1 = \|A\|_1\|B\|_1$.

Solution. First, we have that

$$\begin{aligned}\|A \otimes B\|_1 &= \operatorname{tr}[\sqrt{(A^\dagger \otimes B^\dagger)(A \otimes B)}] \\ &= \operatorname{tr}[\sqrt{A^\dagger A \otimes B^\dagger B}] \\ &= \operatorname{tr}[\sqrt{A^\dagger A} \otimes \sqrt{B^\dagger B}] \\ &= \operatorname{tr}[\sqrt{A^\dagger A}] \operatorname{tr}[\sqrt{B^\dagger B}] \\ &= \|A\|_1 \|B\|_1.\end{aligned}$$

We used that the square root of the tensor product of two operators $X, Y \geq 0$ has $\sqrt{X \otimes Y} = \sqrt{X} \otimes \sqrt{Y}$. We can see this by using their spectral decompositions: let $X = \sum_i \lambda_i P_i$ where λ_i are the eigenvalues of X and P_i the associated eigenprojections, and similarly $Y = \sum_i \mu_i Q_i$. Then $X \otimes Y = \sum_{ij} \lambda_i \mu_j P_i \otimes Q_j$ is the spectral decomposition of $X \otimes Y$. Therefore

$$\sqrt{X \otimes Y} = \sum_{ij} \sqrt{\lambda_i \mu_j} P_i \otimes Q_j = \sum_i \sqrt{\lambda_i} P_i \otimes \sum_j \sqrt{\mu_j} Q_j = \sqrt{X} \otimes \sqrt{Y}.$$

Then,

$$\begin{aligned}F(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) &= \|\sqrt{\rho_1 \otimes \rho_2}\sqrt{\sigma_1 \otimes \sigma_2}\|_1 \\ &= \|(\sqrt{\rho_1} \otimes \sqrt{\rho_2})(\sqrt{\sigma_1} \otimes \sqrt{\sigma_2})\|_1 \\ &= \|(\sqrt{\rho_1}\sqrt{\sigma_1} \otimes \sqrt{\rho_2}\sqrt{\sigma_2})\|_1 \\ &= \|(\sqrt{\rho_1}\sqrt{\sigma_1}\|_1 \|\sqrt{\rho_2}\sqrt{\sigma_2})\|_1 \\ &= F(\rho_1, \sigma_1) F(\rho_2, \sigma_2).\end{aligned}$$

Exercise 9. Let $\rho = \sum_i p_i \rho_i$ where ρ_i are density matrices which have support on orthogonal subspaces. Then prove that

$$S\left(\sum_i p_i \rho_i\right) = H(p) + \sum_i p_i S(\rho_i)$$

where $p = \{p_i\}$ is a probability distribution, and $H(p)$ is the corresponding Shannon entropy.

Solution. Let us write the total number of states as n. Recall $S(\rho) := -\operatorname{tr}[\rho \log \rho]$. In fact, we don't need to restrict S to act on states; for any $X \geq 0$, we can define

$$S(X) = -\operatorname{tr}[X \log X].$$

We notice for $p > 0$ and a state ρ, we have

$$\begin{aligned} S(p\rho) &= -\operatorname{tr}[p\rho \log(p\rho)] = -p \operatorname{tr}[\rho(\log pI + \log \rho)] = -p \operatorname{tr}(\rho) \log p - p \operatorname{tr}[\rho \log \rho] \\ &= -p \log p + pS(\rho). \end{aligned} \tag{3}$$

Now, if the supports of the ρ_i are all orthogonal, then in particular their eigenvectors are too. We can thus make an orthonormal basis of the Hilbert space by the eigenvectors of all the ρ_i, so that each ρ_i is diagonal in this basis. We see that

$$\sum_i p_i \rho_i = \begin{pmatrix} p_1\rho_1 & & & \\ & p_2\rho_2 & & \\ & & \ddots & \\ & & & p_n\rho_n \end{pmatrix}$$

where here we write $\rho_1 \equiv \rho_1|_{\operatorname{supp} \rho_1}$ to mean the block matrix of ρ_1 restricted to its support. Then

$$-\Big(\sum_i p_i\rho_i\Big) \log \Big(\sum_i p_i\rho_i\Big) = \begin{pmatrix} -p_1\rho_1 \log(p_1\rho_1) & & & \\ & -p_2\rho_2 \log(p_2\rho_2) & & \\ & & \ddots & \\ & & & -p_n\rho_n \log(p_n\rho_n) \end{pmatrix},$$

and taking the trace, we see

$$S\left(\sum_i p_i\rho_i\right) = \sum_i S(p_i\rho_i) = -\sum_i p_i \log p_i + p_i S(\rho_i)$$

by (3).

Exercise 10. Prove that the von Neumann entropy is a concave function of its inputs, i.e., given probabilities $p_i \geq 0$, $\sum_{i=1}^r p_i = 1$, and corresponding density operators ρ_i:

$$S\left(\sum_{i=1}^r p_i\rho_i\right) \geq \sum_{i=1}^r p_i S\left(\rho_i\right). \tag{4}$$

Show that if each $p_i > 0$ and $\rho_i \neq \rho_j$ for $i \neq j$, then the inequality (4) is strict. Note: means the von Neumann entropy is *strictly concave*.

Hint: Consider the state $\sigma_{AB} := \sum_i p_i\rho_i \otimes |i\rangle\langle i|$, its reduced states σ_A, σ_B and use subadditivity.

Solution. As the hint says, we'll consider

$$\sigma_{AB} := \sum_i p_i \rho_i \otimes |i\rangle\langle i|.$$

We notice that the support of each $\rho_i \otimes |i\rangle\langle i|$ is orthogonal. If P_i is the projection onto the support of ρ_i, then $P_i \otimes |i\rangle\langle i|$ is the projection onto the support of $\rho_i \otimes |i\rangle\langle i|$. But

$$(P_i \otimes |i\rangle\langle i|)(P_j \otimes |j\rangle\langle j|) = P_i P_j \otimes |i\rangle\langle i||j\rangle\langle j| = 0$$

for $i \neq j$, so the supports are othogonal. Then

$$S(\sigma_{AB}) = \sum_i p_i S(\rho_i \otimes |i\rangle\langle i|) + H(p) = \sum_i p_i S(\rho_i) + H(p)$$

using that $S(\rho_i \otimes |i\rangle\langle i|) = S(\rho_i) + S(|i\rangle\langle i|) = S(\rho_i)$, since $|i\rangle\langle i|$ is a pure state. But by subadditivity

$$S(\sigma_{AB}) \leq S(\sigma_A) + S(\sigma_B). \tag{5}$$

Note $\sigma_A = \sum_i p_i \rho_i$ and $\sigma_B = \sum_i p_i |i\rangle\langle i|$. Then $S(\sigma_B) = H(p)$, and (5) gives

$$\sum_i p_i S(\rho_i) + H(p) = S(\sigma_{AB}) \leq S(\sigma_A) + S(\sigma_B) = S\Big(\sum_i p_i \rho_i\Big) + H(p)$$

which yields the result by subtracting $H(p)$ from each side.

Now, let us assume all the $p_i > 0$ and $\rho_i \neq \rho_j$ for $i \neq j$. The only step in our proof which had an inequality instead of equality was in subadditivity, (5). In fact, one has equality in (5) if and only if $\sigma_{AB} = \sigma_A \otimes \sigma_B$. One can see that by the fact that subadditivity is a rearrangement of terms of the inequality $D(\sigma_{AB} \| \sigma_A \otimes \sigma_B) \geq 0$, which has equality if and only if $\sigma_{AB} = \sigma_A \otimes \sigma_B$.

Therefore, we simply need to show that if $p_i > 0$ and $\rho_i \neq \rho_j$ for $i \neq j$, then $\sigma_{AB} \neq \sigma_A \otimes \sigma_B = \sum_{ij} p_i p_j \rho_i \otimes |j\rangle\langle j|$. Assume otherwise for the sake of contraction:

$$\sum_i p_i \rho_i \otimes |i\rangle\langle i|. = \sum_{ij} p_i p_j \rho_i \tag{6}$$

We can compare blocks (or, equivalently, act on the left of both sides by $I \otimes \langle k|$ and on the right by $I \otimes |k\rangle$) to find

$$p_k \rho_k = p_k \sum_i p_i \rho_i$$

which is equivalent to $\rho_k = \sum_i p_i \rho_i = \sigma_A$. But this must be true for every k (every block), so we must have $\sigma_A = \rho_1 = \rho_2 = \cdots = \rho_n$. This violates the assumption that the ρ_i are all different. Therefore, our assumption (6) must be wrong, and we have strict inequality in subadditivity, and therefore strict inequality in the final result.

Exercise 11. Consider the following *classical-quantum* or cq-states on the Hilbert space $\mathbb{C}^n \otimes \mathcal{H}$:

$$\rho = \sum_{i=1}^{n} p_i |i\rangle\langle i| \otimes \rho_i, \qquad \sigma = \sum_{i=1}^{n} p_i |i\rangle\langle i| \otimes \sigma_i$$

where $\rho_i, \sigma_i \in \mathcal{D}(\mathcal{H})$ and $\{p_i\}$ is a probability distribution. Evaluate the quantum relative entropy $D(\rho\|\sigma)$ and use the result to prove that

$$D\Big(\sum_i p_i \rho_i \| \sum_i p_i \sigma_i\Big) \leq \sum_i p_i D(\rho_i \| \sigma_i),$$

i.e. joint convexity of the quantum relative entropy.

Solution. First, let us prove joint convexity without using monotonicity of the relative entropy under partial trace. To do so, we'll use Lieb's concavity theorem: for any matrix X and $t \in (0,1)$, we have that the function

$$f(A,B) := \mathrm{tr}[X^\dagger A^t X B^{1-t}]$$

is jointly concave in positive matrices A and B.

That is, for any probability distribution p_i and positive matrices A_i and B_i, we have

$$f\Big(\sum_i p_i A_i, \sum_i p_i B_i\Big) \geq \sum_i p_i f(A_i, B_i).$$

We'll follow Nielsen and Chaung. We define

$$I_t(A,X) = \mathrm{tr}[X^\dagger A^t X A^{1-t}] - \mathrm{tr}[X^\dagger X A].$$

The first term is $f(A,A)$ which is thus concave in A, while the second term is linear in A. Therefore, $I_t(A,X)$ is concave in A. Now, $t \mapsto I_t(A,X)$ is a function from $[0,1] \to \mathbb{R}$. In fact, we can write it more simply as a function of t by writing A in its eigendecomposition, $A = \sum_i \lambda_i P_i$, where the λ_i are

the eigenvalues of A and the P_i the associated eigenprojection. Then $A^t = \sum_i \lambda_i^t P_i$. So,

$$I_t(A, X) = \operatorname{tr}[X^\dagger A^t X A^{1-t}] - \operatorname{tr}[X^\dagger X A] = \sum_{i,j} \lambda_i^t \lambda_j^{1-t} \operatorname{tr}[X^\dagger P_i X P_j] - \operatorname{tr}[X^\dagger X A]. \tag{7}$$

Notice that for each i and j, the quantity $\operatorname{tr}[X^\dagger P_i X P_j]$ is just a number. So I_t is a sum of eigenvalues of A to the powers t and $1-t$, weighted by some numbers, minus $\operatorname{tr}[X^\dagger X A]$, which has no t-dependence. Thus, I_t is differentiable at $t = 0$; we can take the derivative using the rule $\frac{\mathrm{d}}{\mathrm{d}t} x^t = \ln(x) x^t$ to find

$$\frac{\mathrm{d}}{\mathrm{d}t} I_t(A, X) = \sum_{i,j} [\ln(\lambda_i) \lambda_i^t \lambda_j^{1-t} - \lambda_i^t \ln(\lambda_j) \lambda_j^{1-t}] \operatorname{tr}[X^\dagger P_i X P_j]$$

and in particular, evaluating the derivative at $t = 0$,

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I_t(A, X) &= \sum_{i,j} [\ln(\lambda_i)\lambda_j - \ln(\lambda_j)\lambda_j] \operatorname{tr}[X^\dagger P_i X P_j] \\
&= \sum_{i,j} \operatorname{tr}[X^\dagger \ln(\lambda_i) P_i X \lambda_j P_j] - \sum_{i,j} \operatorname{tr}[X^\dagger P_i X \ln(\lambda_j) \lambda_j P_j] \\
&= \sum_i \operatorname{tr}[X^\dagger \ln(\lambda_i) P_i X A] - \sum_i \operatorname{tr}[X^\dagger P_i X \ln(A) A] \\
&= \operatorname{tr}[X^\dagger \ln(A) X A] - \operatorname{tr}[X^\dagger X \ln(A) A].
\end{aligned}$$

Let us define $I(A, X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I_t(A, X)$ as this derivative. In fact, $I(A, X)$ is concave in A as well. Using the definition of the derivative,

$$I(pA_1 + (1-p)A_2, X) = \lim_{t \to 0} \frac{I_t(pA_1 + (1-p)A_2, X) - I_0(pA_1 + (1-p)A_2, X)}{t}.$$

But $I_0(B, X) = 0$ for all $B \geq 0$, since the two terms in (7) cancel. And since $B \mapsto I_t(B, X)$ is concave, we have

$$\begin{aligned}
I(pA_1 + (1-p)A_2, X) &\geq \lim_{t \to 0} \frac{pI_t(A_1, X) + (1-p) I_t(A_2, X)}{t} \\
&= p \lim_{t \to 0} \frac{I_t(A_1, X)}{t} + (1-p) \lim_{t \to 0} \frac{I_t(A_2, X)}{t} \\
&= pI(A_1, X) + (1-p) I(A_2, X)
\end{aligned}$$

using the definition of the derivative again, and that
$I_0(A_1, X) = I_0(A_2, X) = 0$. Next, for any quantum states ρ and σ, we choose $A = \begin{pmatrix} \rho & 0 \\ 0 & \sigma \end{pmatrix}$, and $X = \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix}$ as block matrices. We calculate

$$\begin{aligned} I(A, X) &= \operatorname{tr}[X^\dagger \ln(A) X A] - \operatorname{tr}[X^\dagger X \ln(A) A] \\ &= \operatorname{tr}\Big[\begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \ln(\rho) & 0 \\ 0 & \ln(\sigma) \end{pmatrix} \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} \begin{pmatrix} \rho & 0 \\ 0 & \sigma \end{pmatrix} \Big] - \operatorname{tr}\Big[\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \rho \ln(\rho) & 0 \\ 0 & \sigma \ln(\sigma) \end{pmatrix} \Big] \\ &= \operatorname{tr}\Big[\begin{pmatrix} 0 & \ln(\sigma) \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \rho & 0 \end{pmatrix} \Big] + \ln(2) S(\rho) \\ &= \operatorname{tr}\Big[\begin{pmatrix} \ln(\sigma)\rho & 0 \\ 0 & 0 \end{pmatrix} \Big] + \ln(2) S(\rho) \\ &= \ln(2) \operatorname{tr}[\log(\sigma)\rho] + \ln(2) S(\rho) \\ &= -\ln(2) D(\rho \| \sigma). \end{aligned}$$

Therefore, defining $A_i = \begin{pmatrix} \rho_i & 0 \\ 0 & \sigma_i \end{pmatrix}$, we have

$$D\Big(\sum_i p_i \rho_i \| \sum_i p_i \sigma_i \Big) = -\frac{1}{\ln(2)} I\Big(\sum_i p_i A_i, X \Big) \leq -\frac{1}{\ln(2)} \sum_i p_i I(A_i, X) = \sum_i p_i D(\rho_i \| \sigma_i)$$

using the concavity of $A \mapsto I(A, X)$.

Proof of joint convexity using monotonicity of the relative entropy:

We notice that ρ has the block-diagonal form

$$\rho = \begin{pmatrix} p_1\rho_1 & & & \\ & p_2\rho_2 & & \\ & & \ddots & \\ & & & p_n\rho_n \end{pmatrix} \implies \log \rho = \begin{pmatrix} \log(p_1\rho_1) & & & \\ & \log(p_2\rho_2) & & \\ & & \ddots & \\ & & & \log(p_n\rho_n) \end{pmatrix}.$$

Since the same form holds for σ, we have

$$\log \rho - \log \sigma = \begin{pmatrix} \log(\rho_1) - \log(\sigma_1) & & & \\ & \log(\rho_2) - \log(\sigma_2) & & \\ & & \ddots & \\ & & & \log(\rho_n) - \log(\sigma_n) \end{pmatrix}$$

using properties of the logarithm to cancel the p_i's. Thus, we may write $\rho(\log\rho - \log\sigma)$ as

$$\begin{pmatrix} p_1\rho_1(\log(\rho_1) - \log(\sigma_1)) & & & \\ & p_2\rho_2(\log(\rho_2) - \log(\sigma_2)) & & \\ & & \ddots & \\ & & & p_n\rho_n(\log(\rho_n) - \log(\sigma_n)) \end{pmatrix}$$

Taking the trace, we find

$$D(\rho\|\sigma) = \operatorname{tr}[\rho(\log\rho - \log\sigma)] = \sum_i p_i D(\rho_i\|\sigma_i).$$

On the other hand, the data-processing inequality thus gives that

$$D\left(\sum_i p_i\rho_i\|\sum_i p_i\sigma_i\right) = D(\operatorname{tr}_1\rho\|\operatorname{tr}_1\sigma) \leq D(\rho\|\sigma) = \sum_i p_i D(\rho_i\|\sigma_i)$$

where we use tr_1 to denote the partial trace over the first system (which we haven't given an explicit label).

Exercise 12. Let $\rho_{AB} = \sum_i p_i\rho_{AB}^i$ be the state of a bipartite quantum system AB. Using the joint convexity of the relative entropy, prove that the quantum conditional entropy is concave in the state ρ_{AB}.

Solution. We want to show that for each $t \in (0,1)$ and pair of quantum states ρ_{AB} and σ_{AB} on some bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, show that

$$S(A|B)_\omega \geq tS(A|B)_\rho + (1-t)S(A|B)_\sigma. \tag{8}$$

where $\omega_{AB} = t\rho_{AB} + (1-t)\sigma_{AB}$.

We can use that $S(A|B)_\eta = -D(\eta_{AB}\|I_A \otimes \eta_B)$, for any state η. In this language, (8) becomes

$$D(\omega_{AB}\|I_A \otimes \omega_B) \leq tD(\rho_{AB}\|I_A \otimes \rho_B) + (1-t)D(\sigma_{AB}\|I_A \otimes \sigma_B)$$

upon multiplying by -1. But since $\omega_{AB} = t\rho_{AB} + (1-t)\sigma_{AB}$ and $I_A \otimes \omega_B = t(I_A \otimes \rho_A) + (1-t)(I_A \otimes \sigma_B)$, the previous exercise establishes this inequality immediately.

Exercise 13. Let $\mathcal{H}_A, \mathcal{H}_B$ and $\mathcal{H}_{B'}$ be Hilbert spaces and $\rho \in \mathcal{D}(\mathcal{H}_{AB})$ be a state. Further, let $\Lambda : \mathcal{H}_B \to \mathcal{H}_{B'}$ be a CPTP map and define

$$\sigma_{AB'} = (\mathrm{id}_A \otimes \Lambda)\rho_{AB}.$$

Prove that

$$S(A|B)_{\rho_{AB}} \leq S(A|B')_{\sigma_{AB'}}.$$

Hint: Use strong subadditivity.

Solution. To use subadditivity, we need three systems. We'll introduce a third: by Stinespring's dilation theorem, there is a Hilbert space $\mathcal{H}_C$, an isometry $U : \mathcal{H}_B \otimes \mathcal{H}_C \to \mathcal{H}_{B'} \otimes \mathcal{H}_C$ and a pure state $\varphi \in \mathcal{D}(\mathcal{H}_C)$ such that $\Lambda(\rho) = \mathrm{tr}_C(U(\rho \otimes \varphi)U^\dagger)$ for all $\rho \in \mathcal{D}(\mathcal{H}_B)$. Now let

$$\omega_{AB'C} = (I_A \otimes U)(\rho_{AB} \otimes \varphi)(I_A \otimes U^\dagger)$$

such that $\mathrm{tr}_C\, \omega_{AB'C} = \sigma_{AB'}$. Strong subadditivity of the von Neumann entropy implies

$$S(\omega_{AB'C}) + S(\omega_{B'}) \leq S(\sigma_{AB'}) + S(\omega_{B'C}).$$

We have $\omega_{B'} = \mathrm{tr}_A(\mathrm{tr}_C\, \omega_{AB'C}) = \mathrm{tr}_A\, \sigma_{AB'} = \sigma_{B'}$ and

$$S(\omega_{AB'C}) = S(\rho_{AB} \otimes \varphi) = S(\rho_{AB}) + S(\varphi) = S(\rho_{AB})$$

by invariance under unitaries of the von Neumann entropy and the fact that $S(\varphi) = 0$ for pure states φ. Furthermore,

$$\omega_{B'C} = U(\rho_B \otimes \varphi)U^\dagger$$

and hence $S(\omega_{B'C}) = S(\rho_B)$ by the same arguments.

In summary, we have

$$S(\rho_{AB}) - S(\rho_B) \leq S(\sigma_{AB'}) - S(\sigma_{B'})$$

which is equivalent to $S(A|B)_\rho \leq S(A|B')_\sigma$.

Exercise 14. Projective measurements do not decrease von Neumann entropy

Suppose a projective measurement described by a set of projection operators $\{P_i\}$ is performed on a quantum system, but we never learn the result of the

measurement. If the state of the system before the measurement was ρ then the state after the measurement is given by

$$\rho' = \sum_i P_i \rho P_i.$$

Prove that the entropy of this final state is at least as great as the original entropy:

$$S(\rho') \geq S(\rho),$$

with equality if and only if $\rho = \rho'$.

Solution. We consider

$$D(\rho \| \rho') = \mathrm{tr}[\rho(\log \rho - \log \rho')]. \tag{9}$$

We have that

$$\rho' P_j = \left(\sum_i P_i \rho P_i \right) P_j = P_j \rho P_j^2 = P_j^2 \rho P_j = P_j \left(\sum_i P_i \rho P_i \right) = P_j \rho'.$$

Thus, $[P_i, \rho'] = 0$, which implies $[P_i, \log \rho'] = 0$. Now, since $\sum_i P_i = I$,

$$\begin{aligned} \mathrm{tr}[\rho \log(\rho')] &= \sum_i \mathrm{tr}[P_i \rho \log(\rho')] = \sum_i \mathrm{tr}[P_i^2 \rho \log(\rho')] \\ &= \sum_i \mathrm{tr}[P_i \rho \log(\rho') P_i] \\ &= \sum_i \mathrm{tr}[P_i \rho P_i \log(\rho')] \\ &= \mathrm{tr}[\rho' \log(\rho')]. \end{aligned}$$

Thus, (9) becomes

$$D(\rho \| \rho') = -S(\rho) + S(\rho').$$

Klein's inequality gives $D(\rho \| \rho') \geq 0$. So we have

$$S(\rho) \leq S(\rho').$$