Quantum Information Theory
Solutions 9

Exercise 9.1 Some properties of von Neumann entropy

The von Neumann entropy of a density operator $\rho \in \mathcal{S}(\mathcal{H})$ is defined as

$$H(\rho) = -Tr\big(\rho \log \rho\big) = -\sum_i \lambda_i \log \lambda_i, \tag{1}$$

where $\{\lambda_i\}_i$ are the eigenvalues of ρ.
Given a composite system $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ and $\rho_{AB} = Tr_C(\rho_{ABC})$ etc., we often write $H(AB)$ instead of $H(\rho_{AB})$ to denote the entropy of a subsystem.
The conditional *von Neumann entropy may be defined in a composed system $\mathcal{H}_A \otimes \mathcal{H}_B$ as*

$$H(A|B) = H(AB) - H(B). \tag{2}$$

The strong sub-additivity property of the von Neumann entropy proves very useful:

$$H(ABC) + H(B) \leq H(AB) + H(BC). \tag{3}$$

a) Prove the following general properties of the von Neumann entropy:

1. *If ρ_{AB} is pure, then $H(A) = H(B)$.*
   This becomes clear when you apply the Schmidt decomposition to the pure state ρ_{AB} — the reduced states of the two subsystems A and B have the same eigenvalues and therefore the same von Neumann entropy.

2. *If two subsystems are independent, $\rho_{AB} = \rho_A \otimes \rho_B$, then $H(AB) = H(A) + H(B)$.*
   We denote by $\{\lambda_i\}_i$ and $\{\gamma_j\}_j$ the eigenvalues of ρ_A and ρ_B respectively. Hence $\{\lambda_i \gamma_j\}_{i,j}$ are the eigenvalues of ρ_{AB} and we can write:

$$\begin{aligned} H(AB) &= -\sum_{i,j} \lambda_i \gamma_j \log(\lambda_i \gamma_j) \\ &= -\underbrace{\Big(\sum_i \lambda_i\Big)}_{=1} \cdot \Big(\sum_j \gamma_j \log \gamma_j\Big) - \underbrace{\Big(\sum_j \gamma_j\Big)}_{=1} \cdot \Big(\sum_i \lambda_i \log \lambda_i\Big) \\ &= H(A) + H(B). \end{aligned}$$

b) Consider a bipartite system that is classical on a subsystem Z, namely $\rho_{ZA} = \sum_z p_z |z\rangle\langle z|_Z \otimes \rho_A^z$ for some basis $\{|z\rangle Z\}_z$ of $\mathcal{H}_Z$. Show that:

1. *The entropy of the global state is given by*

$$H(AZ) = H(Z) + \sum_z p_z\, H(A|Z = z), \tag{4}$$

   where $H(A|Z = z) = H(\rho_A^z)$.

First, note that the eigenvalues of $\sum_z p_z |z\rangle\langle z| \otimes \rho_A^z$ are given by $\{p_z \lambda_k^z\}_{z,k}$, where $\{\lambda_k^z\}_k$ are the eigenvalues of $\rho_A^z \equiv \rho_{A|Z=z}$. We may now write:

$$\begin{aligned} H(AZ) &= -\sum_{z,k} p_z \lambda_k^z \log(p_z \lambda_k^z) \\ &= -\sum_z p_z \underbrace{\Big(\sum_k \lambda_k^z\Big)}_{=1} \log p_z - \sum_z p_z \Big(\sum_k \lambda_k^z \log \lambda_k^z\Big) \\ &= H(Z) + \sum_z p_z \, H(A|Z=z). \end{aligned}$$

2. *Systems A and Z do not share entanglement, i.e.,*

$$\sum_z p_z \, H(A|Z{=}z) \leq H(A). \tag{5}$$

First note that from strong sub-additivity follows sub-additivity, $H(AC) \leq H(A) + H(C)$, if $\mathcal{H}_B$ is empty. Applying this to a system classical in $\mathcal{H}_Z$, we get

$$H(AZ) = H(Z) + \sum_z p_z \;H(A|Z=z) \leq H(A) + H(Z) \tag{6}$$

from which the inequality follows immediately.

3. *Even if one has access to subsystem A the classical variable is not fully known,*

$$H(Z|A) \geq 0. \tag{7}$$

Let us introduce a copy of the classical subsystem Z, Y, as follows:

$$\rho_{AZY} = \sum_z p_z |z\rangle\langle z|_Z \otimes |z\rangle\langle z|_Y \otimes \rho_A^z.$$

Note that, for this state, $H(AZ) = H(AY) = H(AZY)$.
We may now appply the strong sub-additivity,

$$\begin{aligned} H(AZY) + H(A) &\leq H(AZ) + \underbrace{H(AY)}_{=H(AZY)} \\ \Leftrightarrow 0 &\leq H(AZ) - H(A) \\ \Leftrightarrow 0 &\leq H(Z|A) \end{aligned}$$

Remark: Eq (7) holds in general only for classical Z. Consider, e.g., the Bell-States as an immediate counterexample in the fully quantum case.

Exercise 9.2 Upper bound on von Neumann entropy

Given a state $\rho \in \mathcal{S}(\mathcal{H})$, show that

$$H(\rho) \leq \log |\mathcal{H}|. \tag{8}$$

Consider the state $\bar{\rho} = \int U \rho U^\dagger dU$, where the integral is over all unitaries $U \in \mathcal{U}(\mathcal{H})$ and dU is the Haar measure. Find $\bar{\rho}$ and use concavity (5) to show (8).
Hint: The Haar measure satisfies $d(UV) = d(VU) = dU$, where $V \in \mathcal{U}(\mathcal{H})$ is any unitary.

We use the properties of the Haar measure to verify that $\bar{\rho}$ commutes with all unitaries V on $\mathcal{H}$:

$$V\bar{\rho}V^\dagger = \int (VU)\rho(VU)^\dagger\ dU = \int \tilde{U}\rho\tilde{U}^\dagger\ d(V^\dagger\tilde{U}) = \int \tilde{U}\rho\tilde{U}^\dagger\ d\tilde{U} = \bar{\rho}.$$

The only density operator on $\mathcal{H}$ that has this property is the completely mixed state, so $\bar{\rho} = \mathbb{1}/|\mathcal{H}|$, . The concavity property of the von Neumann entropy (Eq. 5) naturally extends to integrals and we get

$$\log|\mathcal{H}| = H\left(\frac{\mathbb{1}}{|\mathcal{H}|}\right) = H(\bar{\rho}) \geq \int H(U\rho U^\dagger)\ dU = \int H(\rho)\ dU^{(*)} = H(\rho)\int dU = H(\rho),$$

where $^{(*)}$ stands because the entropy is independent of the basis.

Exercise 9.3 Data Processing Inequality

Random variables X, Y, Z form a Markov chain $X \to Y \to Z$ if the conditional distribution of Z depends only on Y: $p(z|x,y) = p(z|y)$. The goal in this exercise is to prove the data processing inequality, $I(X:Y) \geq I(X:Z)$ for $X \to Y \to Z$.

1. *First show the chain rule for mutual information: $I(X:YZ) = I(X:Z) + I(X:Y|Z)$, which holds for arbitrary X, Y, Z. The conditional mutual information is defined as*

$$I(X:Y|Z) = \sum_z p(z) I(X:Y|Z=z) = \sum_z p(z) \sum_{x,y} p(x,y|z) \log\frac{p(x,y|z)}{p(x|z)p(y|z)}.$$

   First observe that $\frac{p(x,y|z)}{p(y|z)} = \frac{p(x,y,z)}{p(y,z)} = p(x|y,z)$, which means $I(X{:}Y|Z) = H(X|Z) - H(X|YZ)$. Then

$$I(X{:}YZ) = H(X) - H(X|YZ) = H(X) + I(X{:}Y|Z) - H(X|Z) = I(X{:}Z) + I(X{:}Y|Z).$$

2. *Next show that in a Markov chain $X \to Y \to Z$, X and Z are conditionally independent given Y; that is, $p(x,z|y) = p(x|y)p(z|y)$.*

$$p(x,z|y) = \frac{p(x,y,z)}{p(y)} = \frac{p(x,y)p(z|x,y)}{p(y)} = \frac{p(x|y)p(y)p(z|y)}{p(y)} = p(x|y)p(z|y).$$

3. *By expanding the mutual information $I(X:YZ)$ in two different ways, prove the data processing in equality.*

   There are only two ways to expand this expression:

$$I(X{:}YZ) = I(X{:}Z) + I(X{:}Y|Z) = I(X{:}Y) + I(X{:}Z|Y).$$

   Since X and Z are conditionally independent given Y, $I(X{:}Z|Y) = 0$. Meanwhile, $I(X{:}Y|Z) \geq 0$, since it is a mixture (over Z) of positive quantities $I(X{:}Y|Z=z)$. Therefore $I(X{:}Y) \geq I(X{:}Z)$.