Ph219a/CS219a

Solutions of Problem Set 7
March 18, 2009

Problem 1

(a) Clearly for $x = 1$, $\ln x = x - 1 = 0$. Since the function $f(x) = \ln x$ is strictly concave, $\ln x \leq x - 1$ for $x \neq 1$, if $x - 1$ is a tangent at $x = 1$. It is easy to see that this is indeed the case, since their slopes match at $x = 1$: $\frac{d(\ln x)}{dx}|_{x=1} = \frac{1}{x}|_{x=1} = 1$ and $\frac{d(x-1)}{dx} = 1$.

(b) Using the above inequality, we have,

$$
\begin{aligned}
\ln\left(\frac{p(x)}{q(x)}\right) &\geq 1 - \frac{q(x)}{p(x)} \\
\Rightarrow H(p \parallel q) &= \sum_x p(x) \log\left(\frac{p(x)}{q(x)}\right) \\
&\geq \sum_x p(x)\left(1 - \frac{q(x)}{p(x)}\right) = 0 \qquad (1)
\end{aligned}
$$

Equality holds iff $\frac{q(x)}{p(x)} = 1$ for all x, ie. iff the distributions $\{p(x)\}$ and $\{q(x)\}$ are identical. (Note that in the second step, we have omitted a factor of ln2 in going from the natural logarithm to the log function.)

(c) Expanding ρ and σ in their eigenbasis, as $\rho = \sum_i p_i |i\rangle \langle i|$ and $\sigma = \sum_a q_a |a\rangle \langle a|$, we have,

$$
\begin{aligned}
H(\rho \parallel \sigma) &= \mathrm{tr}\rho(\log\rho - \log\sigma) \\
&= \sum_i p_i \log p_i - \sum_{i,a} p_i \log q_a \langle i|a\rangle\langle a|i\rangle \\
&= \sum_i p_i \left(\log p_i - \sum_a D_{ia} \log q_a\right) \qquad (2)
\end{aligned}
$$

where we have defined the matrix with elements $D_{ia} = |\langle i|a\rangle|^2$. It is easy to see that D is a doubly stochastic matrix: $\sum_a D_{ia} = \sum_a \langle i|a\rangle\langle a|i\rangle = \langle i|i\rangle = 1 = \langle a|a\rangle = \sum_i D_{ia}$.

(d) Since the log function is strictly concave, and the set $\sum_a D_{ia} = 1$ for each i (implying $D_{ia} \leq 1 \; \forall \, i, a$), Jensen's inequality directly gives

$$\log \left(\sum_a D_{ia} q_a \right) \geq \sum_a D_{ia} \log q_a. \tag{3}$$

When all q's are different, then equality holds iff $D_{ia} = 1$ for some a. If all q's are the same, the equality trivially holds and we can't infer anything about D_{ij}.

(e) Putting Eqns.(2) and (3) together, we have,

$$\begin{aligned} H(\rho \parallel \sigma) &= \sum_i p_i \left(\log p_i - \sum_a D_{ia} \log q_a \right) \\ &\geq \sum_i p_i \left(\log p_i - \log \sum_a D_{ia} q_a \right) \\ &= \sum_i p_i \left(\log p_i - \log r_i \right) \\ &= H(p \parallel r) \end{aligned} \tag{4}$$

with $r_i = \sum_a D_{ia} q_a$, and equality iff $D_{ia} = 1$ for some a, for each i.

(f) Now that we have expressed the quantum relative entropy as a classical relative entropy between two classical distributions $p = \{p_i\}$ and $r = \{r_i\}$, we can use the result of part(b), so that

$$H(\rho \parallel \sigma) \geq H(p \parallel r) \geq 0 \tag{5}$$

The second inequality saturates iff $p_i = r_i = \sum_a D_{ia} q_a$, and the first inequality saturates iff $D_{ia} = 1$ for some a, for each i. Thus D has to be a permutation matrix, which implies that the set $\{p_i\}$ and $\{q_a\}$ are the same upto a permutation. Thus, equality holds iff $\rho = \sigma$.

Problem 2

(a) Consider the relative entropy of ρ_{AB} and the product state $\rho_A \otimes \rho_B$. Using the positivity of relative entropy, we have,

$$\begin{aligned} H(\rho_{AB} \parallel \rho_A \otimes \rho_B) &\geq 0 \\ \Rightarrow \mathrm{Tr}[\rho_{AB} \log \rho_{AB}] - \mathrm{Tr}[\rho_{AB} \log(\rho_A \otimes \rho_B)] &\geq 0 \end{aligned} \tag{6}$$

Using $\log(MN) = \log M + \log N$, we can simplify the second term as follows

$$\begin{aligned} \log(\rho_A \otimes \rho_B) &= \log(I_A \otimes \rho_B) + \log(\rho_A \otimes I_B) \\ &= I_A \otimes \log \rho_B + (\log \rho_A) \otimes I_B \end{aligned} \tag{7}$$

where the last step can be justified as follows: Suppose $\rho = \sum_a \lambda_a |a\rangle\langle a|$, then $\rho \otimes I = \sum_{a,b} \lambda_a |a,b\rangle\langle a,b|$, so that $\log(\rho \otimes I) = \sum_a \log(\lambda_a) |a,b\rangle\langle a,b| = (\log \rho) \otimes I$.
Thus we have,

$$\begin{aligned}\mathrm{Tr}[\rho_{AB} \log(\rho_A \otimes \rho_B)] &= \mathrm{Tr}(\rho_{AB}[I_A \otimes \log \rho_B + (\log \rho_A) \otimes I_B]) \\ &= \mathrm{Tr}(\rho_A \log \rho_A) + \mathrm{Tr}(\rho_B \log \rho_B) \\ &= -H(\rho_A) - H(\rho_B) \end{aligned} \tag{8}$$

Putting this back in Eqn.(6) and noting that the first term is simply $\mathrm{Tr}[\rho_{AB} \log \rho_{AB}] = -H(\rho_{AB})$, we get

$$H(\rho_{AB}) \leq H(\rho_A) + H(\rho_B) \tag{9}$$

where equality holds iff $\rho_{AB} = \rho_A \otimes \rho_B$.

(b) Suppose we construct the bipartite state $\rho_{AB} = \sum_x p_x (\rho_x)_A \otimes (|x\rangle\langle x|)_B$, which is obtained for example, by choosing one of a set of density operators ρ_x based on the outcome of a measurement in the orthogonal basis $\{|x\rangle_B\}$, where outcome x occurs with probability p_x. Then, we have,

$$\begin{aligned} H(\rho_{AB}) &= -\mathrm{Tr} \sum_x p_x (\rho_x)_A \otimes (|x\rangle\langle x|)_B \\ &= \sum_x p_x H(\rho_x) - \mathrm{Tr}(\sum_x \rho_x p_x \log \rho_x) \\ &= \sum_x p_x H(\rho_x) + H(X) \end{aligned} \tag{10}$$

where the second term is the Shannon entropy of the random variable X distributed according to the distribution $\{p(x)\}$.
The reduced density operators of the state ρ_{AB} are given by

$$\rho_A = \sum_x p_x \rho_x \ , \quad \rho_B = \sum_x p_x |x\rangle\langle x| \tag{11}$$

so that

$$H(\rho_A) = H(\sum_x p_x \rho_x) \ , \ H(\rho_B) = H(X) \tag{12}$$

Now, using subadditivity along with Eqns.(10) and (12), we get,

$$\begin{aligned} H(\rho_{AB}) &\leq H(\rho_A) + H(\rho_B) \\ \Rightarrow \sum_x p_x H(\rho_x) + H(X) &\leq H(\sum_x p_x \rho_x) + H(X) \\ \Rightarrow \sum_x p_x H(\rho_x) &\leq H(\sum_x p_x \rho_x) \end{aligned} \tag{13}$$

thus proving concavity of the Von Neumann entropy.

(c) From subadditivity, it follows that the equality

$$\sum_x p_x H(\rho_x) = H\left(\sum_x p_x \rho_x\right) \tag{14}$$

holds iff $\rho_{AB} = \rho_A \otimes \rho_B$, ie. iff

$$\begin{aligned} \sum_x p_x \rho_x \otimes (|x\rangle\langle x|) &= \sum_y p_y \rho_y \otimes \sum_x p_x |x\rangle\langle x| \\ &= \sum_x p_x \left(\sum_y p_y \rho_y\right) \otimes |x\rangle\langle x| \end{aligned} \tag{15}$$

For the cross terms on the RHS to vanish, given that $p_y \neq 0 \; \forall \; y$, we require that for each x, $\rho_x = \sum_y p_y \rho_y$.

(d) Given a bipartite state ρ_{AB}, we can construct a "purification" ($|\Phi\rangle_{ABC}$) by introducing a third system C, such that $\mathrm{tr}_C |\Phi\rangle\langle\Phi| = \rho_{AB}$. Since $|\Phi\rangle$ is pure, we know that $H(\rho_A) = H(\rho_{BC})$ and $H(\rho_{AB}) = H(\rho_C)$. Then, subadditivity gives

$$\begin{aligned} H(\rho_A) = H(\rho_{BC}) &\leq H(\rho_B) + H(\rho_C) \\ &= H(\rho_B) + H(\rho_{AB}) \\ \Rightarrow H(\rho_{AB}) &\geq H(\rho_A) - H(\rho_B) \end{aligned} \tag{16}$$

Similarly, using the subadditivity relation for systems A and C, we get,

$$\begin{aligned} H(\rho_{AC}) &\leq H(\rho_A) + H(\rho_C) \\ \Rightarrow H(\rho_B) &\leq H(\rho_A) + H(\rho_{AB}) \\ \Rightarrow H(\rho_{AB}) &\geq H(\rho_B) - H(\rho_A) \end{aligned} \tag{17}$$

Putting together Eqns.(16) and (17), we get the triangle inequality: $H(\rho_{AB}) \geq |H(\rho_A) - H(\rho_B)|$.

Problem 3

Following the hints provided in the problem, we can write

$$\sqrt{r_j}\,|e_j\rangle = \sum_b V_{jb}\,|\phi_b\rangle \sqrt{p_b}\,|\psi_b\rangle = \sum_{b,\nu} V_{jb}\,|\phi_b\rangle\, U_{b,\nu}\sqrt{s_\nu}\,|f_b\rangle\,. \tag{18}$$

We obtain

$$r_j = \langle e_j|\sqrt{r_j}\sqrt{r_j}\,|e_j\rangle = \sum_{a,b,\mu,\nu} V^*_{ja} V_{jb}\,\langle\phi_a|\,|\phi_b\rangle\, U^*_{a\mu} U_{b\nu} \sqrt{s_\mu s_\nu}\,\langle f_\mu|\,|f_\nu\rangle = \tag{19}$$

$$= \sum_\mu \left(\sum_{a,b} V^*_{ja} V_{jb} U^*_{a\mu} U_{b\mu}\,\langle\phi_a|\,|\phi_b\rangle\right) s_\mu \tag{20}$$

Now, we define

$$D_{j\mu} = \sum_{a,b} V_{ja}^* V_{jb} U_{a\mu}^* U_{b\mu} \langle\phi_a| |\phi_b\rangle \tag{21}$$

Observe, that

$$D_{j\mu} = \left(\sum_a V_{ja}^* U_{a\mu}^* \langle\phi_a|\right)\left(\sum_b V_{jb} U_{b\mu}^* |\phi_b\rangle\right) \geq 0 \tag{22}$$

since $\forall |\xi\rangle : \langle\xi| |\xi\rangle \geq 0$. Using the fact that V and U are unitary, i.e $\sum_j V_{ja}^* V_{jb} = \delta_{ab}$ and $\sum_j U_{ja}^* U_{jb} = \delta_{ab}$, we show

$$\sum_j D_{j\mu} = \sum_{a,b}(\sum_j V_{ja}^* V_{jb}) U_{a\mu}^* U_{b\mu} \langle\phi_a| |\phi_b\rangle = \sum_{a,b} \delta_{ab} U_{a\mu}^* U_{b\mu} \langle\phi_a| |\phi_b\rangle = \sum_a U_{a\mu}^* U_{a\mu} = \delta_{\mu\mu} = 1 \tag{23}$$

and similarly $\sum_\mu D_{j\mu} = 1$. Thus, D is a doubly stochastic matrix.

Problem 4

(a) Choosing orthonormal bases $\{|i\rangle_A\}$ and $\{|j\rangle_{A'}\}$ for systems A and A', the swap operator can be written as $S_{AA'} = \sum_{i,j}(|i\rangle\langle j|)_A \otimes (|j\rangle\langle i|)_{A'}$. Since $\rho_A = \text{tr}_B(|\psi\rangle\langle\phi|)_{AB} = \rho_{A'}$, we have,

$$\begin{aligned}
\text{tr}_{ABA'B'}\left[(S_{AA'} \otimes I_{BB'})\left(|\phi\rangle\langle\phi|_{AB} \otimes |\phi\rangle\langle\phi|_{A'B'}\right)\right] &= \text{tr}_{AA'}\left[S_{AA'}(\rho_A \otimes \rho_{A'})\right] \\
&= \sum_{k,l}\langle k|_A\langle l|_{A'}\left[\sum_{i,j}(|i\rangle\langle j|)_A \otimes (|j\rangle\langle i|)_{A'}\right](\rho_A \otimes \rho_{A'})|k\rangle_A|l\rangle_{A'} \\
&= \sum_{i,j,k,l} \delta_{ik}\delta_{jl}(\rho_A)_{jk}(\rho_A')_{il} \\
&= \sum_{k,l}(\rho_A)_{lk}(\rho_A)_{kl} = \sum_l (\rho_A^2)_{ll} = \text{tr}_A \rho_A^2
\end{aligned} \tag{24}$$

(b) The constant C multiplying $\Pi_{AA'}$ can be determined from the constraint that the average state must be normalized, ie. $\text{tr}(C\ \Pi_{AA'}) = 1$. But tr $\Pi_{AA'} = d_{\text{sym}}$ where $d_{\text{sym}} = d(d+1)/2$ is the dimension of the symmetric subspace of AA'. To see this, let's construct an explicit orthonormal basis for the symmetric subspace. The states $\{|i\rangle_A \otimes |i\rangle_{A'}\}$ with $i = 1, \ldots, d$ are clearly symmetric, and there are d such states. Furthermore, consider the states $(|i\rangle_A \otimes |j\rangle_{A'} + |j\rangle_A \otimes |i\rangle_{A'})/\sqrt{2}$ for $i \neq j$, which are also symmetric and linearly independent from the previous ones. There are $\binom{d}{2} = d(d-1)/2$ such states, giving a total of $d + d(d-1)/2 = d(d+1)/2$ linearly independent basis states spanning the symmetric subspace of AA'. Therefore,

$$C = \frac{1}{d_{\text{sym}}} = \left[\frac{d(d+1)}{2}\right]^{-1} \tag{25}$$

(c) Using part (b),

$$\left\langle |\phi\rangle\langle\phi|_{AB} \otimes |\phi\rangle\langle\phi|_{A'B'} \right\rangle = C\ \Pi_{AB;A'B'} \tag{26}$$

where $C = 2/d_A d_B(d_A d_B + 1)$, since $\dim(AB) = d_A d_B$. Taking the average on both sides of Eqn.(26),

$$\begin{aligned}\left\langle \mathrm{tr}_A \rho_A^2 \right\rangle &= \mathrm{tr}_{ABA'B'} \left[(S_{AA'} \otimes I_{BB'}) \left\langle |\phi\rangle\langle\phi|_{AB} \otimes |\phi\rangle\langle\phi|_{A'B'} \right\rangle \right] \\ &= \frac{C}{2}\, \mathrm{tr}_{ABA'B'} \left[(S_{AA'} \otimes I_{BB'})(I_{ABA'B'} + S_{AB;A'B'}) \right] \\ &= \frac{C}{2}\, \mathrm{tr}_{ABA'B'} \left[S_{AA'} \otimes I_{BB'} + I_{AA'} \otimes S_{BB'} \right] \end{aligned} \tag{27}$$

But,

$$\mathrm{tr}_{AA'}(S_{AA'}) = \mathrm{tr}(2\,\Pi_{AA'} - I_{AA'}) = 2d_A(d_A+1)2 - d_A^2 = d_A \tag{28}$$

And therefore, $\mathrm{tr}_{ABA'B'}\,(S_{AA'} \otimes I_{BB'}) = \mathrm{tr}_{AA'}(S_{AA'})\; d_B^2 = d_A d_B^2$. Similarly, we can calculate $\mathrm{tr}_{ABA'B'}\,(I_{AA'} \otimes S_{BB'}) = d_A^2 d_B$. Substituting these in Eqn.(29), we obtain

$$\left\langle \mathrm{tr}_A \rho_A^2 \right\rangle = \frac{d_A d_B^2 + d_A^2 d_B}{d_A d_B (d_A d_B + 1)} = \frac{d_A + d_B}{d_A d_B + 1} \tag{29}$$

(d) We first calculate

$$\begin{aligned}||\rho_a - \frac{1}{d_A} I_A||_2^2 &= \mathrm{tr}\left(\left(\rho_A - \frac{1}{d_A} I_A\right)^\dagger \left(\rho_A - \frac{1}{d_A} I_A\right) \right) \\ &= \mathrm{tr}\rho_A^2 + \frac{1}{d_A^2} \mathrm{tr} I_A - \frac{2}{d_A} \mathrm{tr}\rho_A \\ &= \mathrm{tr}\rho_A^2 + \frac{1}{d_A^2}\, d_A - \frac{2}{d_A} = \mathrm{tr}\rho_A^2 - \frac{1}{d_A} \end{aligned} \tag{30}$$

Now, using part (c),

$$\left\langle ||\rho_a - \frac{1}{d_A} I_A||_2^2 \right\rangle = \left\langle \mathrm{tr}\rho_A^2 \right\rangle - \frac{1}{d_A} = \frac{d_A + d_B}{d_A d_B + 1} - \frac{1}{d_A} \le \frac{1}{d_A} + \frac{1}{d_B} - \frac{1}{d_A} = \frac{1}{d_B} \tag{31}$$

Thus, using the Cauchy-Schwarz inequality,

$$\left\langle ||\rho_a - \frac{1}{d_A} I_A||_2 \right\rangle \le \sqrt{\frac{1}{d_B}} \tag{32}$$

(e) Let $M = \rho_A - \frac{1}{d_A} I_A$. Since M is Hermitian, we can write it in the basis $\{|\mu\rangle\}$ that diagonalizes it as $M = \sum_{\mu=1}^{d} \lambda_\mu |\mu\rangle\langle\mu|$. Then, $||M||_1 = \mathrm{tr}\sqrt{M^\dagger M} = \sum_{\mu=1}^{d} |\lambda_\mu|$. On the other hand, $||M||_2 = \sqrt{\mathrm{tr}(M^\dagger M)} = \sqrt{\sum_{\mu=1}^{d} |\lambda_\mu|^2}$. Now, using the Cauchy-Schwarz inequality again,

$$\left\langle |\lambda_\mu| \right\rangle \le \sqrt{\left\langle |\lambda_\mu|^2 \right\rangle} \;\Rightarrow \frac{1}{d} \sum_{\mu=1}^{d} |\lambda_\mu| \le \sqrt{\frac{1}{d} \sum_{\mu=1}^{d} |\lambda_\mu|^2} \;\Rightarrow ||M||_1 \le \sqrt{d}||M||_2 \tag{33}$$

Using the bound on $\left\langle ||M||_2 \right\rangle$ from part (d), we get

$$\left\langle ||M||_1 \right\rangle \le \sqrt{d_A} \left\langle ||M||_2 \right\rangle \le \sqrt{\frac{d_A}{d_B}} \tag{34}$$