Ph 219c/CS 219c

Exercises
Due: Monday 1 June 2009

9.1 Positivity of quantum relative entropy

a) Show that $\ln x \leq x - 1$ for all positive real x, with equality iff $x = 1$.

b) The (classical) relative entropy of a probability distribution $\{p(x)\}$ relative to $\{q(x)\}$ is defined as

$$H(p \parallel q) \equiv \sum_x p(x) \left(\log p(x) - \log q(x)\right) \ . \tag{1}$$

Show that

$$H(p \parallel q) \geq 0 \ , \tag{2}$$

with equality iff the probability distributions are identical. **Hint**: Apply the inequality from (a) to $\ln \left(q(x)/p(x)\right)$.

c) The quantum relative entropy of the density operator ρ with respect to σ is defined as

$$H(\rho \parallel \sigma) = \mathrm{tr}\ \rho \left(\log \rho - \log \sigma\right) \ . \tag{3}$$

Let $\{p_i\}$ denote the eigenvalues of ρ and $\{q_a\}$ denote the eigenvalues of σ. Show that

$$H(\rho \parallel \sigma) = \sum_i p_i \left(\log p_i - \sum_a D_{ia} \log q_a \right) \ , \tag{4}$$

where D_{ia} is a doubly stochastic matrix. Express D_{ia} in terms of the eigenstates of ρ and σ. (A matrix is doubly stochastic if its entries are nonnegative real numbers, where each row and each column sums to one.)

d) Show that if D_{ia} is doubly stochastic, then (for each i)

$$\log \left(\sum_a D_{ia} q_a \right) \geq \sum_a D_{ia} \log q_a \ , \tag{5}$$

with equality only if $D_{ia} = 1$ for some a.

e) Show that

$$H(\rho \parallel \sigma) \geq H(p \parallel r) \ , \tag{6}$$

where $r_i = \sum_a D_{ia} q_a$.

f) Show that $H(\rho \parallel \sigma) \geq 0$, with equality iff $\rho = \sigma$.

9.2 Properties of Von Neumann entropy

a) Use nonnegativity of quantum relative entropy to prove the *subadditivity* of Von Neumann entropy

$$H(\rho_{AB}) \leq H(\rho_A) + H(\rho_B), \tag{7}$$

with equality iff $\rho_{AB} = \rho_A \otimes \rho_B$. **Hint**: Consider the relative entropy of ρ_{AB} and $\rho_A \otimes \rho_B$.

b) Use subadditivity to prove the concavity of the Von Neumann entropy:

$$H(\sum_x p_x \rho_x) \geq \sum_x p_x H(\rho_x) \ . \tag{8}$$

Hint: Consider

$$\rho_{AB} = \sum_x p_x \left(\rho_x\right)_A \otimes \left(|x\rangle\langle x|\right)_B \ , \tag{9}$$

where the states $\{|x\rangle_B\}$ are mutually orthogonal.

c) Use the condition

$$H(\rho_{AB}) = H(\rho_A) + H(\rho_B) \quad \text{iff} \quad \rho_{AB} = \rho_A \otimes \rho_B \tag{10}$$

to show that, if all p_x's are nonzero,

$$H\left(\sum_x p_x \rho_x\right) = \sum_x p_x H(\rho_x) \tag{11}$$

iff all the ρ_x's are identical.

d) Use subadditivity to prove the triangle inequality:

$$H(\rho_{AB}) \geq |H(\rho_A) - H(\rho_B)| \ . \tag{12}$$

Hint: Construct a "purification" of ρ_{AB} — introduce a third system C and consider $|\Phi\rangle_{ABC}$ such that

$$\text{tr}_C \left(|\Phi\rangle\langle\Phi|\right) = \rho_{AB} \ ; \tag{13}$$

then use the subadditivity relations $H(\rho_{BC}) \leq H(\rho_B) + H(\rho_C)$ and $H(\rho_{AC}) \leq H(\rho_A) + H(\rho_C)$.

9.3 Coherent information and entanglement fidelity

A criterion for the reversibility of the effect of a quantum channel $\mathcal{N}^{A\to B}$ on the input state ρ^A can be formulated in terms of the coherent information I_c. Let ϕ^{RA} be a purification of ρ^A, where R is a reference system; the channel maps ϕ^{RA} to the output σ^{RB}, which has purification ψ^{RBE}. Here system E can be regarded as the environment, where the channel $\mathcal{N}$ is realized as an isometry $U^{A\to BE}$. There is a decoding map $\mathcal{D}^{B\to C}$ that restores the purity of the state on RC if and only if

$$I_c(R\,\rangle B) \equiv H(B) - H(RB) = H(R)\ , \tag{14}$$

or equivalently

$$H(RE) = H(R) + H(E)\ , \tag{15}$$

where the entropy is evaluated in the output state ψ^{RBE}. That is, the effect of $\mathcal{N}^{A\to B}$ on input ρ^A is perfectly reversible if and only if the output state of the environment is uncorrelated with the reference system. That makes sense — there will be irreversible decoherence if and only if information about the input state leaks to the environment.

Naturally, we expect that if the effect of the channel is *nearly* reversible, then the criterion eq. (14) is *nearly* satisfied. The purpose of this problem is to make this observation more precise.

a) Show that

$$H(R) - I_c(R\,\rangle B) \le 2H(RC)\ . \tag{16}$$

Therefore, if the decoder's output (the state of RC) is almost pure, then the coherent information of the channel $\mathcal{N}$ comes close to matching the input entropy. **Hint**: Use the data processing inequality

$$I_c(R\,\rangle C) \le I_c(R\,\rangle B)\ , \tag{17}$$

and the subadditivity of von Neumann entropy. It is convenient to consider the pure joint state of the reference system, output, and environment. Do *not* assume (because it is not true) that the environment E used in the realization of the noisy channel $\mathcal{N}$ is uncorrelated with the environment E' used in the realization the decoder $\mathcal{D}$.

b) In a d-dimensional system, suppose that the state ρ has fidelity $F = 1 - \varepsilon$ with the pure state $|\psi\rangle$:

$$F = \langle\psi|\rho|\psi\rangle = 1 - \varepsilon \ . \tag{18}$$

Show that

$$H(\rho) \leq H_2(\varepsilon) + \varepsilon \log_2(d-1) \ , \tag{19}$$

where $H_2(\varepsilon) = -\varepsilon \log_2 \varepsilon - (1-\varepsilon)\log_2(1-\varepsilon)$ is the binary Shannon entropy. **Hint**: Recall that if the random variable X describes the outcome of a complete orthogonal measurement performed on the state ρ, then $H(\rho) \leq H(X)$, where $H(X)$ is the Shannon entropy of X.

c) The *entanglement fidelity* F_e provides a useful way to quantity how much the quantum channel $\mathcal{N}^{A\to B}$ deviates from the identity channel. Channel input ρ^A has purification ϕ^{RA}, which is mapped by $\mathcal{N}$ to σ^{RB}. The entanglement fidelity, defined as

$$F_e(\rho^A, \mathcal{N}) \equiv \mathrm{tr}\left(\phi^{RB}\sigma^{RB}\right) \ , \tag{20}$$

does not depend on how the purification is chosen. $F_e = 1$ if the channel preserves its input, and F_e is close to 1 if the output is close to the input. Suppose that

$$F_e(\rho^A, \mathcal{D} \circ \mathcal{N}) = 1 - \varepsilon \ , \tag{21}$$

where $\mathcal{N}^{A\to B}$ is a noisy channel and $\mathcal{D}^{B\to C}$ is the decoding map. Show that

$$H(R) - I_c(R\,\rangle B) \leq 2H_2(\varepsilon) + 2\varepsilon \log_2(d^2 - 1) \ , \tag{22}$$

where $d = \dim R = \dim C$.

d) To define the quantum capacity $Q(\mathcal{N})$ of a noisy channel $\mathcal{N}^{A\to B}$, we can use entanglement fidelity rather than fidelity as the criterion for asymptotically successful quantum communication. The rate M may be said to be achievable if, by using the channel n times, nM qubits can be transmitted from A to B such that, after decoding, the entanglement fidelity is arbitrarily close to 1 for n sufficiently large. The quantum capacity $Q(\mathcal{N})$ is the supremum of achievable rates. This definition of the capacity $Q(\mathcal{N})$ is equivalent to the definition where the fidelity rather than the entanglement fidelity is required to approach 1 as $n \to \infty$ (you are not asked to prove this equivalence). For n independent uses of the channel, let $R^{(n)}$ denote a reference system that purifies the input state on $\rho^{A^{\otimes n}}$. Show that

$$Q(\mathcal{N}) \leq \lim_{n\to\infty} \max_{\rho^{A^{\otimes n}}} \left(\frac{1}{n} I_c \left(R^{(n)} \rangle B^{\otimes n} \right) \right) \ . \tag{23}$$

Hint: Consider an input density operator to $\mathcal{N}^{\otimes n}$ that is uniform on the subspace of dimension 2^{nM} that can be transmitted reliably.

Remark: The resource inequality

$$\langle \mathcal{N}^{A\to B} : \rho^A \rangle \geq I_c(R\,\rangle B)[q \to q] \tag{24}$$

shows that the inequality eq. (23) is actually an equality. But unfortunately, because of the superadditivity of coherent information, this result does not yield a single-letter formula for the quantum capacity $Q(\mathcal{N})$.

9.4 Entanglement of typical bipartite pure states

Suppose that a pure state is chosen at random on the bipartite system AB, where $d_A/d_B \ll 1$. Then with high probability the density operator on A will be very nearly maximally mixed. The purpose of this problem is to derive this property.

To begin with, we will calculate the value of $\langle \text{tr}\ \rho_A^2 \rangle$, where $\langle \cdot \rangle$ denotes the average over all pure states $\{|\varphi\rangle\}$ of AB, and $\rho_A = \text{tr}_B\left(|\varphi\rangle\langle\varphi|\right)$.

a) It is convenient to evaluate tr ρ_A^2 using a trick. Imagine introducing a copy $A'B'$ of the system AB. Show that

$$\text{tr}_A\ \rho_A^2 = \text{tr}_{ABA'B'}\left[(S_{AA'} \otimes I_{BB'}) \left(|\varphi\rangle\langle\varphi|\right)_{AB} \otimes |\varphi\rangle\langle\varphi|_{A'B'}) \right] \ , \tag{25}$$

where $S_{AA'}$ denotes the swap operator

$$S_{AA'} : |\varphi\rangle_A \otimes |\psi\rangle_{A'} \mapsto |\psi\rangle_A \otimes |\varphi\rangle_{A'} \ . \tag{26}$$

b) We wish to average the expression found in (a) over all pure states $|\varphi\rangle$. Rather than go into the details of how such an average is defined, I will simply assert that

$$\langle |\varphi\rangle\langle\varphi|_A \otimes |\varphi\rangle\langle\varphi|_{A'} \rangle = C\ \Pi_{AA'} \ , \tag{27}$$

where C is a constant and $\Pi_{AA'}$ denotes the projector onto the subspace of AA' that is *symmetric* under interchange of A and A'. Eq. (27) can be proved using invariance properties of the average and some group representation theory, but I hope you will regard it as obvious. The state being averaged is symmetric, and the average should not distinguish any symmetric state from any other symmetric state. Express the constant C in terms of the dimension $d \equiv d_A = d_{A'}$.

c) Use the property $\Pi_{AA'} = \frac{1}{2}\left(I_{AA'} + S_{AA'}\right)$ to evaluate the expression found in (a). Show that

$$\left\langle \mathrm{tr}\ \rho_A^2 \right\rangle = \frac{d_A + d_B}{d_A d_B + 1} \ . \tag{28}$$

d) Now estimate the average L^2 distance of ρ_A from the maximally mixed density operator $\frac{1}{d_A} I_A$, where $\| \ M \ \|_2 = \sqrt{\mathrm{tr} M^\dagger M}$; show that

$$\left\langle \| \ \rho_A - \frac{1}{d_A} I_A \ \|_2 \right\rangle \le \frac{1}{\sqrt{d_B}} \ . \tag{29}$$

Hints: First estimate $\left\langle \|\rho_A - \frac{1}{d_A} I_A\|_2^2 \right\rangle$ using eq. (28) and the obvious property $\langle \rho_A \rangle = \frac{1}{d_A} I_A$. Then show that for any nonnegative function f, it follows from the Cauchy-Schwarz inequality that $\langle \sqrt{f} \rangle \le \sqrt{\langle f \rangle}$, and use this property to estimate $\left\langle \| \ \rho_A - \frac{1}{d_A} I_A \ \|_2 \right\rangle$.

e) Finally, estimate the average L^1 distance of ρ_A from the maximally mixed density operator, where $\| \ M \ \|_1 = \mathrm{tr} \ \sqrt{M^\dagger M}$. Use the Cauchy-Schwarz inequality to show that $\| \ M \ \|_1 \le \sqrt{d} \ \| \ M \ \|_2$, if M is a $d \times d$ matrix, and that therefore

$$\left\langle \| \ \rho_A - \frac{1}{d_A} I_A \ \|_1 \right\rangle \le \sqrt{\frac{d_A}{d_B}} \ . \tag{30}$$

It follows from (d) that the average entanglement entropy of A and B is close to maximal for $d_A/d_B \ll 1$: $\langle H(A) \rangle \ge \log_2 d_A - d_A/2 d_B \ln 2$, though you are not asked to prove this bound. Thus, if for example A is 50 qubits and B is 100 qubits, the typical entropy deviates from maximal by only about $2^{-50} \approx 10^{-15}$.