Ph219/CS219: Quantum Computation
Winter 2006
Solutions to Problem Set 4

Problem 4.1

a) Using the definition of quantum mutual information

$$\begin{gathered} I(X;R) + I(XR;B) - I(R;B) = \\ H(X) + H(R) - H(XR) + \\ H(XR) + H(B) - H(XRB) - \\ [\,H(R) + H(B) - H(RB)\,] = \\ H(X) + H(RB) - H(XRB) = \\ I(X;RB)\ . \end{gathered} \tag{1}$$

b) Before Bob's decoding POVM that produces the random variable Y, Alice and Bob share the quantum state σ^{XRB}. Bob's POVM can be viewed as a two step process: (a) a unitary rotation in his system RB and an ancillary system Y, followed by (b) a measurement on system Y. Step (a) does not change the mutual information, so $I(X;RB)_\sigma = I(X;RBY)$. Step (b) can only decrease the mutual information, so $I(X;RBY) \geq I(X;Y)$. In fact, the last inequality is strong subadditivity and we have just rederived the Holevo bound, $I(X;Y) \leq I(X;RB)_\sigma$.

¿From Eq. (1), $I(X;RB)_\sigma \leq I(XR;B)_\sigma$, since $I(R;B)_\sigma \geq 0$ (due to the subadditivity of von Neumann entropy) and $I(X;R)_\sigma = 0$ (because Bob's reference system R is prepared in a product state with the message X and Alice's operation $\mathcal{E}_x$ corresponds to a unitary change of basis in system A that does not alter X). Hence, $I(X;Y) \leq I(XR;B)_\sigma$.

c) Before the noisy channel $\mathcal{N}^{A\to B}$ acts, the state on XRA is

$$\rho^{XRA} = \sum_x p(x)\ (|x\rangle\langle x|)^X \otimes \phi_x^{RA}\ . \tag{2}$$

We can construct a purification $\phi^{\tilde{R}A}$ of ρ^{XRA}

$$\phi^{\tilde{R}A} = \sum_x \sqrt{p(x)}\ (|x\rangle)^X \otimes (|\phi_x\rangle)^{RAE}\ , \tag{3}$$

where we have realized Alice's operation $\mathcal{E}_x$ as an isometry in the extended space RAE, $U_x : (|x\rangle)^X \otimes (|\phi\rangle)^{RA} \otimes (|0\rangle)^E \mapsto (|x\rangle)^X \otimes (|\phi_x\rangle)^{RAE}$. $\tilde{R}$ is then a shorthand for XRE.

The output when $\mathcal{N}^{A\to B}$ acts on $\phi^{\tilde{R}A}$ is

$$\tilde{\sigma}^{\tilde{R}B} = \sum_x \sqrt{p(x)}\,(|x\rangle)^X \otimes (I \otimes \mathcal{N}^{A\to B})(|\phi_x\rangle)^{RAE} \ . \tag{4}$$

Furthermore, the map $\mathcal{E}^{\tilde{R}\to XR} : \tilde{\sigma}^{\tilde{R}B} \mapsto \sigma^{XRB}$ can be realized by decoherence acting on system X (e.g., measuring the x register in the computation basis and discarding the measurement outcome) and by tracing out the ancillary system E. By the monotonicity of quantum mutual information, $I(XR;B)_\sigma \leq I(\tilde{R};B)_{\tilde{\sigma}}$ since $\mathcal{E}^{\tilde{R}\to XR}$ cannot increase the mutual information between the system it acts on and the disjoint system B.

Note: Because the hint in part (d) in the first version of this problem set was misleading, this problem has been revised with an new part added. Please see the new version of the problem set before reading the rest of the solution.

d) From strong subadditivity

$$\begin{aligned} H(ABC) + H(B) &\leq H(AB) + H(BC) \Rightarrow \\ H(ABC) - H(BC) &\leq H(AB) - H(B) \Rightarrow \\ H(A|BC) &\leq H(A|B) \ . \end{aligned} \tag{5}$$

Interchanging B and C we find

$$\begin{aligned} H(ABC) + H(C) &\leq H(AC) + H(BC) \Rightarrow \\ H(ABC) - H(C) &\leq H(AC) - H(C) + H(BC) - H(C) \Rightarrow \\ H(AB|C) &\leq H(A|C) + H(B|C) \ . \end{aligned} \tag{6}$$

Finally, inequality (6) with the correspondence $A \mapsto A$, $B \mapsto B$, $C \mapsto CD$ yields

$$\begin{aligned} H(AB|CD) &\leq H(A|CD) + H(B|CD) \\ &\leq H(A|C) + H(B|D) \ , \end{aligned} \tag{7}$$

where in the last step we have used inequality (5) twice.

e) We can realize the noisy channels $\mathcal{N}^{A_1\to B_1}$ and $\mathcal{N}^{A_2\to B_2}$ as isometries on a larger space, $U_1^{A_1\to B_1E_1}$ and $U_2^{A_2\to B_2E_2}$ respectively. When acting together, the two channels take the input pure state $\phi^{RA_1A_2}$ on RA_1A_2 to the output pure state $(\phi')^{RB_1E_1B_2E_2}$ on $RB_1E_1B_2E_2$. We first evaluate $I(R;B_1B_2)$.

$$\begin{aligned} I(R;B_1B_2) &= H(R) + H(B_1B_2) - H(RB_1B_2) \\ &= H(B_1E_1B_2E_2) + H(B_1B_2) - H(E_1E_2) \ , \end{aligned} \tag{8}$$

where $H(R) = H(B_1E_1B_2E_2)$ and $H(RB_1B_2) = H(E_1E_2)$ follow from the purity of $(\phi')^{RB_1E_1B_2E_2}$ [1]. Similarly,

$$\begin{array}{c} I(RA_2;B_1) + I(RA_1;B_2) = \\ H(RA_2) + H(B_1) - H(RA_2B_1) + [\ H(RA_1) + H(B_2) - H(RA_1B_2)\ ] = \\ H(B_1E_1) + H(B_1) - H(E_1) + [\ H(B_2E_2) + H(B_2) - H(E_2)\ ]\ , \end{array} \tag{9}$$

where $H(RA_2) = H(B_1E_1)$, $H(RA_2B_1) = H(E_1)$ follow from the purity of the state after $U_1^{A_1\to B_1E_1}$ acts on the initial pure state $\phi^{RA_1A_2}$, and $H(RA_1) = H(B_2E_2)$, $H(RA_1B_2) = H(E_2)$ follow from the purity of the state after $U_2^{A_2\to B_2E_2}$ acts on $\phi^{RA_1A_2}$.

Using the subadditivity of von Neumann entropy,

$$H(B_1B_2) \quad \leq \quad H(B_1) + H(B_2)\ . \tag{10}$$

Also, inequality (7) from part (d) with the correspondence $A \mapsto B_1$, $B \mapsto B_2$, $C \mapsto E_1$ and $D \mapsto E_2$ yields

$$H(B_1E_1B_2E_2) - H(E_1E_2) \leq H(B_1E_1) - H(E_1) + H(B_2E_2) - H(E_2)\ . \tag{11}$$

Adding the inequalities (10) and (11), it follows that $I(R;B_1B_2)$ given by Eq. (8) is no larger than $I(RA_2;B_1) + I(RA_1;B_2)$ given by Eq. (9), as wanted.

f) We can purify the uniform density distribution over the 2^{nM} codewords (encoding Alice's messages) using a reference system $R^{(n)}$. Then using parts (b) and (c) we can write

$$I(X:Y) \leq I(R^{(n)}:B^{(n)})_{\tilde{\sigma}}\ , \tag{12}$$

where $B^{(n)} \equiv B_1B_2\ldots B_n$ is the output from n independent uses of the channel $\mathcal{N}$ on input $A^{(n)} \equiv A_1A_2\ldots A_n$.

¿From part (e) we can upper bound $I(R^{(n)}:B^{(n)})_{\tilde{\sigma}}$ as

$$\begin{aligned} I(R^{(n)}:B^{(n)})_{\tilde{\sigma}} &\leq \sum_{i=1}^{n} I(R_i;B_i) \\ &\leq nC_E\ , \end{aligned} \tag{13}$$

where $R_i \equiv R^{(n)}A_1A_2\ldots A_{i-1}A_{i+1}\ldots A_n$, and we have used the definition of C_E as the supremum over the achievable rates for a *single* use of the channel.

If the rate M is achievable when using the channel n times independently, then $I(X:Y) \to nM$ as $n \to \infty$. So, we can use inequalities (12) and (13) to upper bound

[1] Here we used the fact that for a bipartite *pure* state $|\psi_{AB}\rangle$, $H(A) = H(B)$. Indeed, $\rho_A \equiv \mathrm{Tr}_{\mathrm{B}}(|\psi\rangle\langle\psi|)$ and $\rho_B \equiv \mathrm{Tr}_{\mathrm{A}}(|\psi\rangle\langle\psi|)$ have identical non-trivial eigenvalues as can be seen from the Schmidt decomposition of $|\psi_{AB}\rangle$.

the achievable rate for the transmission through $\mathcal{N}^{\otimes n}$ of messages encoded using (possibly entangled) n-qubit codewords by the capacity for transmitting "single-letters" (i.e. n unentangled qubits), or

$$M \leq C_E(\mathcal{N}) \; . \tag{14}$$

Problem 4.2

a) For the GHZ state we compute

$$\begin{aligned} \rho^A = \rho^B = \rho^E &= \tfrac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) \; , \\ \rho^{AE} = \rho^{AB} = \rho^{BE} &= \tfrac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right) \; . \end{aligned} \tag{15}$$

Therefore, $H(A) = H(B) = H(E) = 1$ and $H(AE) = H(AB) = 1$. And so $I(A;E) = I(A;B) = 1$.

b) The GHZ state can be put in the form

$$|GHZ\rangle = \frac{1}{\sqrt{2}}|+\rangle^A \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)^{BE} + \frac{1}{\sqrt{2}}|-\rangle^A \otimes \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)^{BE} \; . \tag{16}$$

Therefore, if Alice gets outcome $+1$, then the state on ABE is $|+\rangle^A \otimes \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)^{BE}$. If her outcome is -1 instead, then the post-measurement state is $|-\rangle^A \otimes \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right)^{BE}$.

c) First note that when Alice gets outcome $+1$ then BE already share $|\phi^+\rangle$. If she gets outcome -1 then BE share $|\phi^-\rangle$ which Bob can rotate to the state $|\phi^+\rangle$ by applying a Pauli σ_z.

But Alice can send Bob a cobit (coherent bit), which is more powerful than a classical bit. So instead of measuring her qubit, she first applies a H rotation to change the state to

$$(H \otimes I \otimes I)|GHZ\rangle = \frac{1}{\sqrt{2}}|0\rangle^A \otimes \frac{1}{\sqrt{2}}(|\phi^+\rangle)^{BE} + \frac{1}{\sqrt{2}}|1\rangle^A \otimes \frac{1}{\sqrt{2}}(|\phi^-\rangle)^{BE} \; , \tag{17}$$

and then sends the value of her qubit coherently to Bob. That is, she performs the operation $(|x\rangle)^A \mapsto (|x\rangle)^A \otimes (|x\rangle)^B$ $(x = 0, 1)$ between her qubit and a *new* qubit that she sends to Bob.

Now the state is

$$\frac{1}{\sqrt{2}}|0\rangle^A \otimes |0\rangle^B \frac{1}{\sqrt{2}}(|\phi^+\rangle)^{BE} + \frac{1}{\sqrt{2}}|1\rangle^A \otimes |1\rangle^B \frac{1}{\sqrt{2}}(|\phi^-\rangle)^{BE} \; , \tag{18}$$

and all that Bob needs to do to complete the protocol is to apply σ_z on his second qubit (i.e. the one which is entangled with E) conditioned on his first qubit (i.e.

the one Alice sent him coherently). This will make both terms in the superposition have $|\phi^+\rangle$ shared between BE creating the state $(|\phi^+\rangle)^{AB} \otimes (|\phi^+\rangle)^{BE}$.

d) In the last homework we saw that one ebit and one qubit of communication can simulate two cobits, or

$$[qq] + [q \to q] \geq 2[q \to qq] \ . \tag{19}$$

So, if Alice and Bob borrow one ebit, they can use their quantum communication to simulate two cobits. They can then take the two copies of $|GHZ\rangle$ and consume these two cobits to run the protocol of part (c) twice. This will create two ebits between Alice and Bob, one of which must be repaid back. Eventually they have gained one ebit! Overall, using one ebit as a catalyst, they have managed to achieve the conversion

$$2\langle GHZ^{ABE}\rangle + [q \to q]^{AB} \geq [qq]^{AB} + 2\langle(\phi^+)^{BE}\rangle \ . \tag{20}$$

Problem 4.3

a) By the data processing inequality

$$\begin{aligned} H(R) - I_c(R\rangle B) &\leq H(R) - I_c(R\rangle C) \\ &= H(R) + H(RC) - H(C) \ . \end{aligned} \tag{21}$$

The decoder $\mathcal{D}^{B\to C}$ can be viewed as an isometry on a larger space that includes an "environment" system E'. Then, the action of the channel followed by the decoder is an isometry taking the input pure state (that lives on RA) to an output pure state (that lives on $RCEE'$). For this output pure state we can write

$$\begin{aligned} H(R) = H(CEE') &\leq H(C) + H(EE') \\ &= H(C) + H(RC) \ . \end{aligned} \tag{22}$$

Combining the two inequalities, $H(R) - I_c(R\rangle B) \leq 2H(RC)$.

b) In the first problem set we showed that the vector $\lambda(\rho)$ of eigenvalues of a density matrix ρ majorizes the probability vector $p(x)$ of any ensemble $\{|\psi_x\rangle, p(x)\}$ realizing the same density matrix. This implies that $H(\rho) \leq H(X)$, where X is the random variable that describes the outcomes of any orthogonal measurement performed on the state ρ.

Now, we can imagine an orthogonal measurement where $|\psi\rangle\langle\psi|$ is one of the projectors, so that this outcome will occur with probability $1-\epsilon$. The $d-1$ one-dimensional projectors in the orthogonal complement of $|\psi\rangle$ could be chosen in various ways, but we can upper bound the entropy of the measurement outcome variable X by assigning an equal probability $\frac{\epsilon}{d-1}$ to each of them. Thus,

$$\begin{aligned} H(\rho) \leq H(X) &\leq -(1-\epsilon)log(1-\epsilon) - (d-1) \times \tfrac{\epsilon}{d-1} log\left(\tfrac{\epsilon}{d-1}\right) \\ &= H_2(\epsilon) + \epsilon\ log(d-1)\ . \end{aligned} \tag{23}$$

c) Applying part (b) by making the correspondence $|\psi\rangle \mapsto |\phi\rangle^{RA}$, $\rho \mapsto \sigma^{RC}$ and noting that $dim(RC) = d^2$, we obtain

$$H(RC) \equiv H(\sigma^{RC}) \quad \leq \quad H_2(\epsilon) + \epsilon\ log(d^2-1)\ . \tag{24}$$

From part (a) it then follows that

$$H(R) - I_c(R\rangle B) \leq 2H_2(\epsilon) + 2\epsilon\ log(d^2-1)\ . \tag{25}$$

d) We consider an input density operator to the channel $\mathcal{N}^{\otimes n}$ that is uniform on a subspace of dimension $d = 2^{nM}$, where nM is the number of qubits that can be transmitted reliably as $n \to \infty$. Therefore each codeword will appear with probability 2^{-nM} and so $H(R^{(n)}) \to -2^{nM} \times 2^{-nM} log(2^{-nM}) = nM$ as $n \to \infty$.

Furthermore, the encoding space $B^{\otimes n}$ and the reference system $R^{(n)}$ each have dimension d^n, so $dim(R^{(n)}B^{\otimes n}) = d^{2n}$. Therefore, using part (c),

$$H(R) - I_c(R^{(n)}\rangle B^{\otimes n}) \quad \leq \quad 2H_2(\epsilon) + 2\epsilon\ log(d^{2n}-1)\ . \tag{26}$$

Then, in the limit $n \to \infty$ we can take $\epsilon \to 0$ and $H(R^{(n)}) \to nM$ to find

$$M \leq \lim_{n\to\infty} \left(\tfrac{1}{n} I_c(R^{(n)}\rangle B^{\otimes n})\right)\ . \tag{27}$$

Maximizing over all achievable rates, $C(\mathcal{N}) \leq \lim_{n\to\infty} \max_{\rho^{A^{\otimes n}}} \left(\frac{1}{n} I_c(R^{(n)}\rangle B^{\otimes n})\right)$.