Quantum Computing and Information - Problem Set 4 Solutions

Exercise 1. **Fidelity and trace distance**

a) Give an exact relation between $F(\alpha, \beta)$ and $T := \frac{1}{2}\|\alpha - \beta\|_1$ for pure states $\alpha = |\alpha\rangle\langle\alpha|$ and $\beta = |\beta\rangle\langle\beta|$.

b) Use this to prove that $F(\rho, \sigma)^2 \leq 1 - \frac{1}{4}\|\rho - \sigma\|_1^2$ for general density matrices ρ, σ.

a) Define $p, \left|\alpha^\perp\right\rangle$ by $\left\langle\alpha\middle|\alpha^\perp\right\rangle$ and $|\beta\rangle = \sqrt{p}\,|\alpha\rangle + \sqrt{1-p}\,\left|\alpha^\perp\right\rangle$. Thus $F = \sqrt{p}$. In the $|\alpha\rangle, \left|\alpha^\perp\right\rangle$ basis, we have

$$\begin{aligned} T &= \frac{1}{2}\left\| \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} p & \sqrt{p(1-p)} \\ \sqrt{p(1-p)} & 1-p \end{pmatrix} \right\|_1 \\ &= \frac{1}{2}\left\| \begin{pmatrix} 1-p & -\sqrt{p(1-p)} \\ -\sqrt{p(1-p)} & p-1 \end{pmatrix} \right\|_1 \\ &= \frac{1}{2}\left\| (1-p)\sigma_z - \sqrt{p(1-p)}\sigma_x \right\|_1 \end{aligned}$$

Since $(1-p)\sigma_z - \sqrt{p(1-p)}\sigma_x$ has eigenvalues $\pm\sqrt{(1-p)^2 + 2p(1-p)} = \pm\sqrt{1-p^2}$, we calculate $T = \sqrt{1-p^2} = \sqrt{1-F^2}$. Equivalently, $F^2 = 1 - T^2$.

b) For general density matrices ρ, σ, Uhlmann's theorem implies that there exist purifications $|\alpha\rangle, |\beta\rangle$ satisfying $|\langle\alpha|\beta\rangle| = F(\rho, \sigma)$. By part (a), $\frac{1}{2}\|\alpha - \beta\|_1 = \sqrt{1 - F(\alpha, \beta)^2} = \sqrt{1 - F(\rho, \sigma)^2}$. Next, tracing out subsystems can only decrease trace distance, so

$$\frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)^2}.$$

Rearranging, we obtain the desired inequality.

Exercise 2. **Optimality of super-dense coding and teleportation**

a) Suppose that Alice would like to transmit an n-bit message x to Bob, but has access only to m uses of a noiseless bit channel, for $m \leq n$. Assume that x is drawn uniformly at random. Prove that for any encoding/decoding strategy, Bob's probability of guessing x is $\leq 2^{m-n}$.

b) Show that this bound still holds if Alice and Bob share an arbitrary entangled state $|\psi\rangle \in \mathbb{C}^{d \times d}$.

c) Can the communication cost of teleportation be improved, possibly at the cost of using more entanglement? Specifically, is it possible to exactly teleport n qubits using some large amount of entanglement, but using $< 2n$ bits of communication?

d) Similarly, can the communication cost of super-dense coding be improved, again possibly at the cost of using more entanglement? Specifically, is it possible to transmit $2n$ cbits using some large amount of entanglement, but $< n$ qubits of communication?

e) Optional: *Prove that n qubits cannot be teleported using fewer than n copies of $|\Phi_2\rangle$ and an unlimited amount of classical communication.* Hint: show that local operations and classical communication has zero probability of increasing the number of nonzero Schmidt coefficients of an entangled state.

a) We prove the claim first when $m = 0$. In this case, Bob simply must output a guess without *any* input from Alice. His guess may be random (based on randomness that is independent from x) . However, this will not increase his success probability. To see this, let Bob guess x' with probability $p_{x'}$. Let $C(x')$ be the probability that x' is the correct guess. (This is in fact 2^{-n} for any x' that is a

possible input of Alice, but the same argument can apply in more general settings where this may not be the case.) Then Bob's success probability is $\sum_{x'} p_{x'} C(x') \leq \max_{x'} C(x')$. Thus, Bob can simply deterministically guess $\arg\max_{x'} C(x')$ and he will not decrease his success probability.

However, a deterministic guess by Bob has probability 2^{-n} of being correct (or 0, if his guess isn't one of Alice's possible inputs). This proves the claim for the case of $m = 0$.

To extend this to general m, we consider an arbitrary protocol $\mathcal{P}$ that uses m bits of communication, and achieves success probability p. We would like to prove that $p \leq 2^{m-n}$. Consider a modified protocol $\mathcal{P}'$ in which the communication of $\mathcal{P}$ is replaced with Bob simply guessing Alice's m-bit message. If he guesses randomly, his guess will be correct with probability 2^{-m}. Thus, he will guess x correctly with probability $\geq p2^{-m}$. Since he achieved this without any communication, by our previous result, we must have $p2^{-m} \leq 2^{-n}$, and thus $p \leq 2^{m-n}$ as desired.

b) In the $m = 0$ case, the shared entanglement simply contributes a random variable that is uncorrelated with x. Thus, it is covered by the previous analysis. For the general case, the same guessing protocol works.

If this is unsatisfying, here is a direct argument. Alice's encoding strategy can be described as a measurement E_y^x satisfying $E_y^x \geq 0$ for each x, y and $\sum_{y \in \{0,1\}^m} E_y^x = I_d$ for each $x \in \{0,1\}^n$. Bob's decoding strategy is to, upon receiving message y, perform the measurement $D_{x'}^y$, where for each y, $\{D_{x'}^y\}_{x' \in \{0,1\}^n}$ is a valid measurement.

We can directly calculate the success probability p to be

$$\begin{aligned}
p &= \sum_{x,x' \in \{0,1\}^n} \sum_{y \in \{0,1\}^m} 2^{-n} \delta_{x,x'} \operatorname{tr}(E_y^x \otimes D_{x'}^y)\psi \\
&= 2^{-n} \sum_{y \in \{0,1\}^m} \sum_{x \in \{0,1\}^n} \operatorname{tr}(E_y^x \otimes D_x^y)\psi \\
&\leq 2^{-n} \sum_{y \in \{0,1\}^m} \sum_{x \in \{0,1\}^n} \operatorname{tr}(I \otimes D_x^y)\psi \\
&= 2^{-n} \sum_{y \in \{0,1\}^m} = 2^{m-n}
\end{aligned}$$

c) This isn't possible. If it were, then we could perform super-dense coding using the qubits teleported in this improved protocol, and transmit $2n$ classical bits using $< 2n$ cbits and some amount of entanglement. This would contradict (b).

d) This is also impossible, for a similar reason. If this super-duper-dense-coding protocol existed, then we could use $< 2n$ cbits plus entanglement to teleport the $< n$ qubits used in the protocol. In this way, we would communicate $2n$ cbits using $< 2n$ cbits plus entanglement.

e) Our strategy is as follows. First, we show that starting with m copies of $|\Phi_2\rangle$ and using LOCC, we can only create mixtures of states with Schmidt rank $\leq 2^m$. Second, we show that any such state (and thus any such mixture) has fidelity $\leq \sqrt{2^{m-n}}$ with $|\Phi_2\rangle^{\otimes n}$. Since a low-entanglement teleportation protocol could be used to turn a small amount of entanglement into a larger amount (using LOCC), this bound on the fidelity rules out such protocols.

For the first claim, use part (c) of exercise 3 (below) to show that any LOCC protocol maps $|\Phi_2\rangle^{\otimes m}$ to a density matrix that is a mixture of states proportional to $(X_j \otimes Y_j) |\Phi_2\rangle^{\otimes m}$ for X_j, Y_j arbitrary operators. Write $|\Phi_2\rangle^{\otimes m} := \frac{1}{\sqrt{2^m}} \sum_{i \in \{0,1\}^m} |i\rangle \otimes |i\rangle$. Then

$$(X_j \otimes Y_j) |\Phi_2\rangle^{\otimes m} = \frac{1}{\sqrt{2^m}} \sum_{i \in \{0,1\}^m} X_j |i\rangle \otimes Y_j |i\rangle .$$

This state is a superposition of 2^m product states, and thus has Schmidt rank $\leq 2^m$.

For the second claim, consider first the $m = 0$ case. In this case, we consider the maximum overlap of a product state $|\alpha\rangle \otimes |\beta\rangle$ with $|\Phi_2\rangle^{\otimes n}$. A direct calculation shows that this is $\sqrt{2^{-n}}\,\langle\bar{\alpha}|\beta\rangle$ which has absolute value $\leq \sqrt{2^{-n}}$. For general m, if $|\psi\rangle$ has Schmidt rank 2^m, then it can be written as $\sum_{i=1}^{2^m} \sqrt{p_i}\,|\alpha_i\rangle\,|\beta_i\rangle$. Thus

$$|\left\langle\psi\middle|\Phi_2^{\otimes n}\right\rangle| \leq \sum_{i=1}^{2^m} \sqrt{p_i 2^{-n}} \leq \sqrt{2^{m-n}},$$

where $\sum_{i=1}^{2^m} \sqrt{p_i} \leq \sqrt{2^m}$ due to Cauchy-Schwartz.

Finally, if $\rho = \sum_j q_j \psi_j$ where each $|\psi_j\rangle$ has Schmidt rank $\leq 2^m$, then we can bound

$$\begin{aligned} F(\rho, \Phi_2^{\otimes n}) &= \sqrt{\left\langle\Phi_2^{\otimes n}\middle|\,\rho\,\middle|\Phi_2^{\otimes n}\right\rangle} \\ &= \sqrt{\sum_j q_j |\left\langle\Phi_2^{\otimes n}\middle|\psi_j\right\rangle|^2} \\ &\leq \sqrt{\sum_j q_j 2^{m-n}} \\ &= \sqrt{2^{m-n}} \end{aligned}$$

Exercise 3. **Partial Transpose and Data Hiding**

a) *Define the transpose map $T : M_d \to M_d$ by $T(X) = X^T$. Show that T is positive but not completely positive.*

b) *Show that $(\mathrm{id} \otimes T)(\rho^{AB}) \geq 0$ for any $\rho \in \mathrm{SEP}(d_A, d_B)$, where* SEP *is the set of separable states defined in problem set 3. The operator $\mathrm{id} \otimes T$ is called the* partial transpose.

c) *Define the class of LOCC (Local Operations + Classical Communication) operations on $\mathcal{D}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ to consist of all finite-length sequences of measurements by Alice on her system (followed by sending the measurement result to Bob) and measurement by Bob of his system (followed by sending the measurement result to Alice). Note that measurements can be chosen based on the previous communication record. Prove that every quantum operation $\mathcal{N}$ in LOCC has the form*

$$\mathcal{N}(\rho) = \sum_j (X_j^A \otimes Y_j^B)\rho^{AB}(X_j^A \otimes Y_j^B)^\dagger.$$

d) *Suppose that $\{M, I - M\}$ is a 2-outcome measurement that is implemented by LOCC. Prove that*

$$0 \leq (\mathrm{id} \otimes T)(M) \leq I \tag{1}$$

e) *Let $F = \sum_{i,j=1}^d |i,j\rangle\,\langle j,i|$ denote the unitary operator that swaps the states of two quantum system. Compute $(\mathrm{id} \otimes T)(F)$ and write down its eigenvalues.*

f) *Since $F^2 = I$, it follows that the eigenvalues of F are ± 1. Define the projectors $\Pi_\pm = (I \pm F)/2$ and the* data-hiding *states $\rho_\pm = \Pi_\pm / \operatorname{tr}\Pi_\pm$. Consider a measurement*

$$M = m_+\Pi_+ + m_-\Pi_-. \tag{2}$$

Define the bias *of M to be $\operatorname{tr} M(\rho_+ - \rho_-)$. Calculate the maximum bias for (i) any valid measurement M, and (ii) any M satisfying Eq. (1). What can you say about the distinguishability of $\rho_\pm$ when Alice and Bob are restricted to LOCC measurements? Is the "data-hiding" name appropriate?*

g) Optional: *Prove that the optimal bias (either with or without the requirement that Eq. (1) be satisfied) is achieved by M of the form in Eq. (2).*

a) Beware that transpose is basis-dependent. So we fix a basis $|1\rangle, \ldots, |d\rangle$, called the "standard basis" and define transpose in this basis.

To show positivity, assume $X \geq 0$. Then $X = \sum_i \lambda_i |v_i\rangle\langle v_i|$ with $\lambda_i \geq 0$ and $X^T = \bar{X} = \sum_i \lambda_i |\bar{v}_i\rangle\langle \bar{v}_i| \geq 0$.

To show the lack of complete positivity, we note that $(\mathrm{id} \otimes T)$ applied to the maximally entangled state is proportional to F (see (e), below), which is not positive.

b) If $\rho = \sum_i p_i \alpha_i \otimes \beta_i$ for density matrices α_i, β_i, then $(\mathrm{id} \otimes T)(\rho) = \sum_i p_i \alpha_i \otimes \beta_i^T$. Since the β_i^T are all density matrices, then $(\mathrm{id} \otimes T)(\rho)$ is as well.

c) Use induction on the number of rounds, and note that LOCC is defined to include only protocols with a finite number of rounds. Consider a protocol with at most m rounds. We can model this as alternating quantum operations by Alice and Bob. First Alice applies an operation with Kraus operators $\{X_{a_1}^{()}\}$ (i.e. sending ρ to $\sum_{a_1} (X_{a_1}^{()} \otimes I)\rho(X_{a_1}^{()} \otimes I)^\dagger$) and then she sends the measurement outcome a_1 to Bob. Sending the complete outcome to Bob is WLOG since the most general thing Alice could do would be to (a) add some random bits to the message, and (b) apply some deterministic maps to the message. However, (a) can be simulated by adding more measurement outcomes, and (b) can be simulated by Bob ignoring part of the message.

Then Bob does a measurement conditioned on a_1, which we call $\{Y_{b_1}^{(a_1)}\}$, and he sends the outcome b_1 to Alice. She performs a measurement $\{X_{a_2}^{(a_1,b_1)}\}$ conditioned on a_1, b_1, sends the outcome a_2 to Bob and so on. After m rounds, we have mapped ρ to

$$\sum_{a_1,\ldots,a_m,b_1,\ldots,b_m} (X_{a_m}^{(a_1,b_1,\ldots,b_{m-1})} \cdots X_{a_2}^{(a_1,b_1)} X_{a_1}^{()} \otimes Y_{b_m}^{(a_1,b_1,\ldots,a_m)} \cdots Y_{b_2}^{(a_1,b_1,a_2)} Y_{b_1}^{(a_1)})\rho$$
$$(X_{a_m}^{(a_1,b_1,\ldots,b_{m-1})} \cdots X_{a_2}^{(a_1,b_1)} X_{a_1}^{()} \otimes Y_{b_m}^{(a_1,b_1,\ldots,a_m)} \cdots Y_{b_2}^{(a_1,b_1,a_2)} Y_{b_1}^{(a_1)})^\dagger. \quad (3)$$

Let $j = (a_1, \ldots, a_m, b_1, \ldots, b_m)$ and define

$$X_j = X_{a_m}^{(a_1,b_1,\ldots,b_{m-1})} \cdots X_{a_2}^{(a_1,b_1)} X_{a_1}^{()}$$
$$Y_j = Y_{b_m}^{(a_1,b_1,\ldots,a_m)} \cdots Y_{b_2}^{(a_1,b_1,a_2)} Y_{b_1}^{(a_1)}$$

d) We can achieve any LOCC measurement $\{M, I - M\}$ by performing an LOCC operation and grouping together measurement outcomes. So both M and $I - M$ can be written in the form $\sum_j X_j \otimes Y_j$ where now X_j, Y_j are psd operators. Thus $(\mathrm{id} \otimes T)(M) = \sum_j X_j \otimes Y_j^T \geq 0$ and similarly $(\mathrm{id} \otimes T)(I - M) \geq 0$. Rearranging we find

$$0 \leq (\mathrm{id} \otimes T)(M) \leq I.$$

e) $(\mathrm{id} \otimes T)(F) = \sum_{i,j} |i\rangle\langle j| \otimes |i\rangle\langle j| = d\Phi_d$, where $\Phi_d = |\Phi\rangle\langle\Phi|_d$ and $|\Phi_d\rangle = \sqrt{1}d \sum_{i=1}^d |i\rangle \otimes |i\rangle$. Thus $(\mathrm{id} \otimes T)(F)$ has a single non-zero eigenvalue, equal to d.

f) The bias is $m_+ - m_-$. For (i), we have the constraints $0 \leq m_+ \leq 1$ and $0 \leq m_- \leq 1$. Thus we can take $m_+ = 1$ and $m_- = 0$ and achieve bias of 1. For (ii) we have the additional constraint that

$$0 \leq (\mathrm{id} \otimes T)(M) = \frac{m_+ + m_-}{2} I + \frac{m_+ - m_-}{2} d\Phi_d = \frac{m_+ + m_-}{2}(I - \Phi_d) + \frac{(d+1)m_+ - (d-1)m_-}{2}\Phi_d \leq I.$$

This implies the constraints $0 \leq m_+ + m_- \leq 2$ (which is redundant) and $0 \leq (d+1)m_+ - (d-1)m_- \leq 2$. Rearranging the upper bound, we obtain $m_+ \leq \frac{2}{d+1} + \frac{d-1}{d+1} m_-$ and thus $m_+ - m_- \leq \frac{2}{d+1} + \frac{d-1}{d+1} m_- - m_- = \frac{2}{d+1}(1 - m_-) \leq \frac{2}{d+1}$. On the other hand, this bias is achieved by taking $m_+ = 2/(d+1)$ and $m_- = 0$. (This argument is an example of LP duality, lest it seem mysterious. Of course in 2-D, everything is easy.)

We conclude that the states are indeed "hiding", at least against LOCC, since $\rho_\pm$ are nearly indistinguishable via LOCC, even though unrestricted measurements can distinguish them with certainty.

g) Without Eq. (1) the problem is trivial: measurements of the form Eq. (2) already achieve bias 1, so relaxing the contraint cannot improve things. Now, consider the case when we require Eq. (1). First, observe that $\operatorname{tr}\Pi_\pm = d(d\pm 1)/2$. Thus, $\Delta := \rho_+ - \rho_- = \frac{I+F}{d(d+1)} - \frac{I-F}{d(d-1)} = 2\frac{dF-I}{d(d^2-1)}$. The bias is thus $\operatorname{tr} M\Delta = \frac{2}{d(d^2-1)}(d\operatorname{tr} MF - \operatorname{tr} M)$. Guided by our use of LP duality from before, we will write $d\operatorname{tr} MF - \operatorname{tr} M = (d-1)\operatorname{tr} MF - \operatorname{tr} M(I-F) \leq (d-1)\operatorname{tr} MF$, where the last inequality is because $M \geq 0$ and $I - F \geq 0$. Thus, the bias is $\leq \frac{2}{d(d+1)} \operatorname{tr} MF$.

Next, let $A^\Gamma := (\mathrm{id}\otimes T)(A)$, and note that $\operatorname{tr} AB = \operatorname{tr} A^\Gamma B^\Gamma$. In this language, Eq. (1) means that $0 \leq M^\Gamma \leq I$. Now express the bias as

$$\operatorname{tr} M\Delta \leq \frac{2}{d(d+1)} \operatorname{tr} MF \leq \frac{2}{d+1} \operatorname{tr} M^\Gamma \Phi_d \leq \frac{2}{d+1}, \tag{4}$$

where in the last step we have used the fact that $M^\Gamma \leq I$.

An alternate proof is to use symmetry to show that any measurement can WLOG be taken of the form in Eq. (2). This approach, and indeed the entire problem, is taken from quant-ph/0203004.