Quantum Computing and Information - Problem Set 3 Solutions

Exercise 1. *Prove that* $\operatorname{tr}\rho^2 \leq 1$ *with equality iff* ρ *is pure (i.e. of the form* $|\psi\rangle\langle\psi|$.Let the eigenvalues of ρ be $\lambda_1, \ldots, \lambda_d$. Then $\operatorname{tr}\rho^2 = \sum_{i=1}^d \lambda_i^2 \leq \sum_{i=1}^d \lambda_i = 1$. The inequality $\lambda_i^2 \leq \lambda_i$ is tight iff $\lambda_i \in \{0, 1\}$. The case when all eigenvalues are 0 or 1 is equivalent to ρ being a pure state.

Exercise 2. *Prove that the extreme points of* $\mathcal{D}(\mathbb{C}^d)$ *are the pure states.* If ρ is not pure, then it can be written as $\sum_{i=1}^d \lambda_i |\psi_i\rangle\langle\psi_i|$ with at least two $\lambda_i > 0$. In particular, suppose that $0 < \lambda_1 < 1$. Then we can decompose ρ as a convex combination of two other density matrices as

$$\rho = \lambda_1 |\psi_1\rangle\langle\psi_1| + \sum_{i=2}^{d} \frac{\lambda_i}{1-\lambda_1} |\psi_i\rangle\langle\psi_i|.$$

Conversely, suppose a pure state $\psi = |\psi\rangle\langle\psi|$ can be written as $\psi = p\sigma + (1-p)\omega$ for $0 < p < 1$ and $\sigma, \omega \in \mathcal{D}(\mathbb{C}^d)$. Since $\omega \geq 0$, we have $\psi \geq p\sigma$. Thus for any $|\varphi\rangle$ orthogonal to $|\psi\rangle$, we have $\operatorname{tr}\sigma\varphi = 0$. This implies that $\sigma = \psi$. A similar argument shows that $\omega = \psi$. Thus, ψ is not an extreme point.

Exercise 3. *Alice and Bob share the state*

$$|\psi\rangle^{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} A_{i,j} \, |i\rangle^A \otimes |j\rangle^B .$$

Calculate Bob's reduced density matrix. Like the expression derived in class for Alice's reduced density matrix, your expression should not have any subscripts or summation signs in it.

$$\begin{aligned}
\operatorname{tr}_A |\psi\rangle\langle\psi| &= \operatorname{tr}_A \sum_{i,i',j,j'} A_{i,j} \bar{A}_{i',j'} \, |i\rangle \langle i'|^A \otimes |j\rangle \langle j'|^B \\
&= \sum_{i,j,j'} A_{i,j} \bar{A}_{i,j'} \, |j\rangle \langle j'| \\
&= \sum_{i,j,j'} A_{i,j} A^\dagger_{j',i} \, |j\rangle \langle j'| \\
&= (A^\dagger A)^T = A^T \bar{A}
\end{aligned}$$

Exercise 4. **Bit commitment** *Alice and Bob have been playing a grueling game of chess and by the end of the first day, it's Alice's move and they've only reached the midgame. Alice has only two choices of move (0 or 1), but if she tells Bob then he'll be able to spend all night planning his response. On the other hand, if Alice doesn't tell him her move until morning then* she *could get an unfair advantage by thinking about her move all night.*

Bob suggests that Alice could write her move on a piece of paper and give it to him in a sealed envelope. But Alice knows that Bob could easily steam the envelope open, read the paper and reseal the envelope. Instead she proposes to use quantum mechanics.

Her idea is to prepare one of two distinguishable states $|\psi_0\rangle^{AB}$ *or* $|\psi_1\rangle^{AB}$ *and give system B to Bob at night, keeping A for herself. Thus she* commits *to her bit* $a \in \{0, 1\}$. *Then she* reveals *a in the morning by sending system A to Bob and he performs a measurement to determine whether the state of AB is* $|\psi_0\rangle$ *or* $|\psi_1\rangle$.

Ideally the protocol would be concealing *if Bob could not learn any information about a after Alice commits her bit and before she reveals it (i.e. from system B alone). On the other hand, it should also be* binding, *meaning that after committing her bit, Alice is unable to change its value.*

Show that both properties cannot simultaneously hold: no commitment protocol can be both concealing and binding. If the protocol is concealing then $\mathrm{tr}_A |\psi_0\rangle\langle\psi_0| = \mathrm{tr}_A |\psi_1\rangle\langle\psi_1|$. (Otherwise Bob could learn something about a from $\mathrm{tr}_A |\psi_a\rangle\langle\psi_a|$.) Thus, both purifications $|\psi_0\rangle$ and $|\psi_1\rangle$ are related by a unitary transformations on Alice's side and she can cheat with no chance of being caught, e.g. by committing to $|\psi_0\rangle$ and then locally transforming the state to $|\psi_1\rangle$ before the reveal phase.

Exercise 5. **Separable states**

a) Let S be a set in $\mathbb{R}^d$. Prove that any $x \in \mathrm{conv}(S)$ can be written as a convex combination of $d+1$ points in S. That is, there exist $p_1, \ldots, p_{d+1} \geq 0$, $y_1, \ldots, y_{d+1} \in S$ such that $\sum_{i=1}^{d+1} p_i = 1$ and

$$x = \sum_{i=1}^{d+1} p_i y_i.$$

b) Let $\mathrm{SEP}(d_A, d_B) \subset \mathcal{D}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ denote the set of separable states*, defined to be the set of states ρ^{AB} that can be written in the form*

$$\rho^{AB} = \sum_i p_i \sigma_i^A \otimes \omega_i^B,$$

where $\sum_i p_i = 1$, each $p_i \geq 0$, $\sigma_i \in \mathcal{D}(\mathbb{C}^{d_A})$ and $\omega_i \in \mathcal{D}(\mathbb{C}^{d_B})$. We call states of the form $\sigma \otimes \omega$ product states *and can equivalently say that separable states are the convex hull of product states. States that are not separable are said to be* entangled*.*

Prove that any $\rho^{AB} \in \mathrm{SEP}(d_A, d_B)$ can be written as a convex combination of $d_A^2 d_B^2$ product pure states.

a) By the definition of convex hull and convex combination, x can be written as $x = \sum_{i=1}^m p_i y_i$ for some probability distribution $p_1, \ldots, p_m$ and some $y_1, \ldots, y_m \in S$. The only issue is that we may have $m > d+1$. If $m \leq d+1$ then we are done. Otherwise we will show that we can reduce m by one or more. Using induction (and the fact that m is, by definition, initially finite) this will prove our claim.

Now, suppose that $m > d+1$. Define the $d+1$-dimensional vectors $\tilde{y}_i = 1 \oplus y_i$. That is, they have a 1 in the first position, and y_i in the remaining d positions. Since there are m of them in a $d+1$-dimensional space, they must be linearly dependent. Thus, there exists $q \in \mathbb{R}^m$ s.t. $\sum_{i=1}^m q_i \tilde{y}_i = 0$. From the first position, we have $\sum_{i=1}^m q_i = 0$. From the remaining d positions, we have $\sum_{i=1}^m q_i y_i = 0$. Thus, for all real t, if we define $p' = p - tq$, then $\sum_{i=1}^m p'_i y_i = x$ and $\sum_{i=1}^m p'_i = 1$. Thus, p' is still a probability distribution as long as we have $p'_i \geq 0$ for each i.

Let t be the largest number satisfying $p_i \geq t q_i$ for each i (equivalently $p'_i \geq 0$). Since each $p_i > 0$, our choice of t is strictly positive. Since t cannot be increased, there must be at least one i such that $p_i = t q_i$ and thus $p'_i = 0$. Thus, p' is a probability distribution, supported on $< m$ elements, with $x = \sum_{i=1}^m p'_i y_i$. This satisfies the induction hypothesis and proves the claim.

b) Starting with the decomposition $\rho^{AB} = \sum_i p_i \sigma_i^A \otimes \omega_i^B$, we can further decompose each σ_i and ω_i into pure states. Thus, we can write ρ^{AB} as a convex combination of pure product states. This is contained in the space of Hermitian $d_A d_B \times d_A d_B$ matrices, which is a $d_A^2 d_B^2$-dimensional real vector space, so by part (a), we can decompose ρ^{AB} into a convex combination of $d_A^2 d_B^2 + 1$ pure states.

To remove the $+1$, we need to project ρ^{AB}, as well as all pure product states, onto the subspace of traceless matrices. Thus, we replace ρ^{AB} with $\rho^{AB} - I_{d_A d_B}/d_A d_B$ and write it as a convex combination of states of the form $|\alpha\rangle\langle\alpha| \otimes |\beta\rangle\langle\beta| - I_{d_A d_B}/d_A d_B$. These states live in a space of dimension $d_A^2 d_B^2 - 1$, so by part (a) of this problem, we can write

$$\rho^{AB} - I_{d_A d_B}/d_A d_B = \sum_{i=1}^{d_A^2 d_B^2} p_i (|\alpha_i\rangle\langle\alpha_i| \otimes |\beta_i\rangle\langle\beta_i| - I_{d_A d_B}/d_A d_B).$$

Adding $I_{d_A d_B}/d_A d_B$ to both sides we obtain the desired decomposition.

Exercise 6. **Trace distance** *Suppose that you are given one of two possible d-dimensional states σ_1 or σ_2, with probabilities p_1 and $p_2 = 1 - p_1$ respectively. Your task is to perform a two-outcome measurement and then try to guess which state you had been given, minimising the probability of error.*
If the measurement elements are nonnegative Hermitian matrices M_1 and $M_2 = I - M_1$ then the probability of guessing wrong is

$$P_{err} = p_1 \operatorname{tr}(\sigma_1 M_2) + p_2 \operatorname{tr}(\sigma_2 M_1).$$

a) Show that

$$P_{err} = p_1 - \sum_{i=1}^{d} \lambda_i \langle i| M_1 |i\rangle ,$$

where $|i\rangle$ denotes the orthonormal basis of eigenstates of the Hermitian operator $p_2\sigma_2 - p_1\sigma_1$ and the λ_i are the corresponding eigenvalues.

b) Find the nonnegative operator M_1 that minimizes P_{err}. Show that the resulting error probability is $P_{err,opt} = p_1 - \sum_{i:\lambda_i<0} |\lambda_i|$. Hint: Express M_1 in the $|i\rangle$ basis.

c) For a Hermitian matrix A, define $|A|$, the absolute value of A, as follows: write $A = UDU^\dagger$ for

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots 0 & \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ & & & & \\ 0 & 0 & \ddots & 0 & 0 \\ & & & & \\ 0 & \dots & & & \lambda_d \end{pmatrix}$$

and U unitary, and then

$$|A| = U \begin{pmatrix} |\lambda_1| & 0 & 0 & \cdots 0 & \\ 0 & |\lambda_2| & 0 & \cdots & 0 \\ & & & & \\ 0 & 0 & \ddots & 0 & 0 \\ & & & & \\ 0 & \dots & & & |\lambda_d| \end{pmatrix} U^\dagger.$$

Express the trace norm $\|p_2\sigma_2 - p_1\sigma_1\|_1 := \operatorname{tr}|p_2\sigma_2 - p_1\sigma_1|$ in terms of the eigenvalues λ_i. Use this, together with the fact that $\operatorname{tr}(p_2\sigma_2 - p_1\sigma_1) = \sum_i \lambda_i = p_2 - p_1$, to express $P_{err,opt}$ as a function of $\|p_2\sigma_2 - p_1\sigma_1\|_1$.

d) Evaluate $P_{err,opt}$ in the following cases:

i) $p_1 = 1$, $p_2 = 0$ and σ_1, σ_2 are arbitrary.

ii) $p_1 = p_2 = 1/2$, $\sigma_1 = |\psi_1\rangle\langle\psi_1|$, $\sigma_2 = |\psi_2\rangle\langle\psi_2|$, with $|\psi_1\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$ and $|\psi_2\rangle = \sin(\theta)|0\rangle + \cos(\theta)|1\rangle$. Check that your answer makes sense when θ is 0 or $\pi/4$.

a) $P_{err} = p_1 \operatorname{tr}(\sigma_1(I - E_1)) + p_2 \operatorname{tr}(\sigma_2 E_1) = p_1 \operatorname{tr}\sigma_1 + \operatorname{tr}(E_1(p_2\sigma_2 - p_1\sigma_1)) = p_1 + \operatorname{tr}(E_1(p_2\sigma_2 - p_1\sigma_1)) = p_1 + \operatorname{tr}(E_1 \sum_i \lambda_i |i\rangle\langle i|) = p_1 + \sum_i \lambda_i \langle i| E_1 |i\rangle$.

b) Write E_1 in the $|i\rangle$ basis as $E_1 = \sum_{i,j} e_{ij} |i\rangle\langle j|$. The second term in P_{err} is $\sum_i \lambda_i e_{ii}$. Since $0 \preceq E_1 \preceq I$, we have $0 \le e_{ii} \le 1$ for each i. Using this we bound $\sum_i \lambda_i e_{ii} = \sum_{i:\lambda_i>0} \lambda_i e_{ii} - \sum_{i:\lambda_i<0} |\lambda_i| e_{ii} \ge -\sum_{i:\lambda_i<0} |\lambda_i|$, yielding the desired result.

c)

$$\operatorname{tr}|p_2\sigma_2 - p_1\sigma_1| = \operatorname{tr}(\sum_i |\lambda_i| |i\rangle\langle i|) = \sum_i |\lambda_i| = \sum_{i:\lambda_i>0} \lambda_i - \sum_{i:\lambda_i<0} \lambda_i.$$

On the other hand,

$$p_2 - p_1 = \operatorname{tr}(p_2\sigma_2 - p_1\sigma_1) = \sum_{i:\lambda_i>0} \lambda_i + \sum_{i:\lambda_i<0} \lambda_i.$$

Taking the difference of these two equations yields that

$$p_2 - p_1 - \operatorname{tr}|p_2\sigma_2 - p_1\sigma_1| = 2\sum_{i:\lambda_i<0} \lambda_i = 2(P_{err,opt} - p_1),$$

Rearranging and using $p_1 + p_2 = 1$ we find that

$$P_{err,opt} = \frac{1 - \|p_2\sigma_2 - p_1\sigma_1\|_1}{2}.$$

d) i) $P_{err,opt} = 0$.

ii)

$$p_2\sigma_2 - p_1\sigma_1 = \frac{1}{2}\begin{pmatrix} \cos 2\theta & 0 \\ 0 & -\cos 2\theta \end{pmatrix}$$

so $P_{err,opt} = \frac{1}{2}(1 - \cos 2\theta) = \sin^2(\theta)$. If $\theta = 0$, the states are orthogonal and can be perfectly distinguished, while if $\theta = \pi/4$ then $|\psi_1\rangle = |\psi_2\rangle$ and the best guessing probability is $1/2$.

Exercise 7. **Purifications** *Let ρ^A be a density matrix and $|\psi\rangle^{AB}$ an arbitrary purification of ρ.*

a) Consider a decomposition $\rho = \sum_i p_i|\varphi_i\rangle\langle\varphi_i|$, where $|\varphi_i\rangle$ are not necessarily orthogonal to each other, and $\{p_i\}$ is a probability distribution. Find a measurement on B such that when applied to half of $|\psi\rangle$ outcome i occurs with probability p_i and Alice's residual state is $|\varphi_i\rangle$.

b) What if we decompose ρ into $\rho = \sum_i p_i\sigma_i$ for general density matrices σ_i? Is it still possible to find a measurement on B such that outcome i occurs with probability p_i and the residual state for Alice is σ_i?

c) Let $\mathcal{N}$ be a quantum operation. The entanglement fidelity *measures how well it approximates the identity on ensembles with density matrix ρ, and is defined*

$$F_e(\mathcal{N}, \rho) := \sqrt{\langle\psi|\,(\mathcal{N}\otimes \mathrm{id})(\psi)\,|\psi\rangle},$$

where $|\psi\rangle$ is an arbitrary purification of ρ and $\psi := |\psi\rangle\langle\psi|$. Prove that F_e does not depend on the purification chosen, and therefore that F_e is well defined.

d) Prove that $\sum_i p_i \langle\varphi_i|\mathcal{N}(\varphi_i)|\varphi_i\rangle \geq F_e$ for any ensemble satisfying $\rho = \sum_i p_i\varphi_i$.

a) Use the definition of measurements in which the outcomes are $M_1, \ldots, M_k$ and the matrices M_i are positive semidefinite and sum to the identity. For any psd matrix M, we define $\sqrt{M}$ to be the psd square root.

Let $d = \operatorname{rank}\rho$. Assume WLOG that $\rho \in \mathcal{D}(\mathbb{C}^d)$. Any purification of ρ can be written in the form $\sqrt{d}(A\otimes I)\,|\Phi_d\rangle$, where $|\Phi_d\rangle = \sum_{i=1}^d |i\rangle\otimes|i\rangle$ and $AA^\dagger = \rho$.

If we now perform the measurement $\{M_1, \ldots, M_k\}$ on system B and obtain outcome i, then the unnormalized residual state for Alice is

$$\begin{aligned} d\operatorname{tr}_B(A\otimes M_i)\Phi_d(A^\dagger\otimes I) &= d\operatorname{tr}_B(A\otimes\sqrt{M_i})\Phi_d(A^\dagger\otimes\sqrt{M_i}) \\ &= d\operatorname{tr}_B(A\sqrt{M_i}^T\otimes I)\Phi_d(\sqrt{M_i}^T A^\dagger\otimes I) \\ &= A\sqrt{M_i}^T\sqrt{M_i}^T A^\dagger \\ &= AM_i^T A^\dagger \end{aligned}$$

This can be thought of as $p_i\rho_i$ where $p_i \geq 0$ is the probability of outcome i and ρ_i is Alice's residual density matrix. We would like p_i to be the p_i given in the problem statement and would like ρ_i to be φ_i. Thus, we have $AM_i^T A^\dagger = p_i\varphi_i$. Since we have assumed that ρ is full rank, $A, A^\dagger$ are as well, and we have

$$M_i = ((A^\dagger)^{-1} p_i \varphi_i A^{-1})^T = (A^T)^{-1} p_i \varphi_i^T \bar{A}^{-1}.$$

To verify positivity, note that $p_i\varphi_i^T \geq 0$ and for any psd matrix B and any matrix A, $A^\dagger BA \geq 0$. To verify normalization, calculate $\sum_i M_i = (A^T)^{-1} \sum_i p_i \varphi_i^T \bar{A}^{-1} = (A^T)^{-1}\rho^T \bar{A}^{-1} = (A^T)^{-1}(A^\dagger A)^T \bar{A}^{-1} = I$.

An alternate, and arguably simpler, proof is obtained by performing a local change of basis so that $\rho = \sum_{i=1}^d \lambda_i |i\rangle\langle i|$ and the purification is $|\psi\rangle = \sum_{i=1}^d \sqrt{\lambda_i}\, |i\rangle \otimes |i\rangle$. In this case, we have the simpler situation that A is psd and thus $A = \sqrt{\rho}$, implying that $M_i^T = \rho^{-1/2} p_i \varphi_i \rho^{-1/2}$.

b) Essentially the same argument shows that $M_i = (A^T)^{-1} p_i \sigma_i^T \bar{A}^{-1}$ does the job. Again, if we choose the purification with $A = \sqrt{\rho}$ then we have $M_i^T = \rho^{-1/2} p_i \sigma_i \rho^{-1/2}$.

c)

$$\begin{aligned}
\langle\psi|\,(\mathcal{N}\otimes \mathrm{id})(\psi)\,|\psi\rangle &= \sum_i \langle\psi|\,(E_i \otimes I)\psi(E_i^\dagger \otimes I)\,|\psi\rangle \\
&= d^2 \sum_i \langle\Phi_d|\,(A^\dagger E_i A \otimes I)\Phi_d(A^\dagger E_i^\dagger A \otimes I)\,|\Phi_d\rangle \\
&= d^2 \sum_i |\,\langle\Phi_d|\,(A^\dagger E_i A \otimes I)\,|\Phi_d\rangle\,|^2 \\
&= d^2 \sum_i |\,\mathrm{tr}(A^\dagger E_i A \otimes I)\Phi_d|^2 \\
&= \sum_i |\,\mathrm{tr}\, A^\dagger E_i A|^2 \\
&= \sum_i |\,\mathrm{tr}\, A A^\dagger E_i|^2 \\
&= \sum_i |\,\mathrm{tr}\,\rho E_i|^2
\end{aligned}$$

which depends only on ρ and not on the choice of purification.

d) By part (a), there exists a measurement in which outcome i occurs with probability i and leaves the residual state φ_i. Consider this measurement to be a quantum operation called $\mathcal{M}$ which, upon outcome i, leaves the state $|i\rangle\langle i|$. This can be achieved by taking $E_{i,j} = |i\rangle\,\langle j|\,E_i$, for i running over all measurement outcomes and $j \in [d]$. Applying $\mathcal{M}$ to the B register of $|\psi\rangle$ leaves the state $\sum_i p_i\varphi_i \otimes |i\rangle\langle i|$. By the monotonocity of fidelity,

$$\begin{aligned}
F_e &= F(\psi, (\mathcal{N}\otimes \mathrm{id})\psi) \\
&\leq F((\mathrm{id}\otimes\mathcal{M})(\psi), (\mathcal{N}\otimes\mathcal{M})(\psi)) \\
&= F\left(\sum_i p_i\varphi_i \otimes |i\rangle\langle i|, \sum_i p_i \mathcal{N}(\varphi_i)\otimes |i\rangle\langle i|\right) \\
&= \mathrm{tr}\sqrt{\left(\sum_i \sqrt{p_i}\varphi_i \otimes |i\rangle\langle i|\right)\left(\sum_i p_i \mathcal{N}(\varphi_i)\otimes |i\rangle\langle i|\right)\left(\sum_i \sqrt{p_i}\varphi_i \otimes |i\rangle\langle i|\right)} \\
&= \mathrm{tr}\sqrt{\sum_i p_i^2 \varphi_i \mathcal{N}(\varphi_i)\varphi_i \otimes |i\rangle\langle i|} \\
&= \sum_i p_i\, \mathrm{tr}\sqrt{\varphi_i \mathcal{N}(\varphi_i)\varphi_i} \\
&= \sum_i p_i F(\varphi_i, \mathcal{N}(\varphi_i))
\end{aligned}$$