COMPSCI 2233 / PHYSICS 272: Quantum Learning Theory, Fall 2025 (Sitan Chen, Jordan Cotler) Version: **1.0**

Submitted by <**name**>

This problem set will cover concepts from the first unit on quantum basics and the beginning of the second unit on learning general quantum states.

The questions have been labeled with the date of the lecture in which the relevant material is covered, to help you budget your time. The questions are meant to be challenging, so do not feel discouraged if you get stuck and are unable to solve some of them.

If you find that you are running low on time to finish all the problems, our recommendation is to try to aim for breadth rather than depth – e.g., it is better to complete a few parts of each of the three questions, than to completely solve one of the three questions and skip the others.

Below we provide hints for the various problems in this assignment. While these may help you solve the problems more easily, you are not required to follow the hints as long as the proofs you provide are correct.

1 (33 pts.) Fun with Post-Measurement States (9/10 and 9/15)

Motivation: The ways in which quantum states evolve as one performs measurements on them are incredibly subtle. In this exercise, we will explore some basic phenomena along these lines, with the ulterior motive of familiarizing you with some common linear algebraic manipulations that arise from playing around with the Born rule and inner products (fidelities) between states.

1.A. (6 pts.) Let $|\psi\rangle$ be an arbitrary n-qubit pure state, and let $\{M, \mathbb{1} - M\}$ denote a two-outcome projective measurement. Prove that the post-measurement state $|\psi'\rangle = (\mathbb{1} - M)|\psi\rangle \,/\, \langle\psi|(\mathbb{1} - M)|\psi\rangle^{1/2}$ upon observing the outcome corresponding to $\mathbb{1} - M$ satisfies

$$|\langle\psi'|\psi\rangle|^2 \geqslant 1 - \epsilon\,,$$

where ϵ is the probability of observing the outcome corresponding to M. Provide a short intuitive description of what this inequality is saying.

1.B. (10 pts.) Let Π_θ denote the single-qubit projector in the direction $\cos(\theta)|0\rangle + \sin(\theta)|1\rangle$.

Let $T \in \mathbb{N}$, and define $\epsilon = \frac{\pi}{2T}$, suppose that we start with the state $|0\rangle$ and apply the following sequence of measurements. First, we measure it with $\{\Pi_\epsilon, \mathbb{1} - \Pi_\epsilon\}$, then take the post-measurement state and measure it with $\{\Pi_{2\epsilon}, \mathbb{1} - \Pi_{2\epsilon}\}$, then take the post-measurement state and measure it with $\{\Pi_{3\epsilon}, \mathbb{1} - \Pi_{3\epsilon}\}$, etc., continuing until we measure with $\{\Pi_{T\epsilon}, \mathbb{1} - \Pi_{T\epsilon}\}$. Prove that the final post-measurement state is $|1\rangle$ with probability at least $1 - O(\epsilon)$. In a few sentences, briefly describe why this example is counterintuitive in light of Question **1.A.**.

1.C. (17 pts.) Motivated by the previous example, we now prove a version of Question **1.A.** where a *sequence* of two-outcome measurements is performed. Let $|\psi\rangle$ be an arbitrary n-qubit pure state as before, and let $\{M_1, \mathbb{1} - M_1\}, \ldots, \{M_s, \mathbb{1} - M_s\}$ denote a sequence of two-outcome projective measurements. If $|\psi'\rangle$ denotes the post-measurement state from performing these measurements in sequence and observing the outcomes corresponding to $\mathbb{1} - M_1, \ldots, \mathbb{1} - M_s$, then prove that

$$|\langle\psi'|\psi\rangle|^2 \geqslant 1 - \sum_{i=1}^{s} \langle\psi| M_i |\psi\rangle\,.$$

Hints: Proceed via induction on the number of measurements. You may find the following elementary inequality helpful: for any nonnegative scalars a, b, c, d, we have $\sqrt{ab} + \sqrt{cd} \leqslant \sqrt{a + c} \cdot \sqrt{b + d}$.

1.D. (0 pts.) **Optional**: Prove Question **1.C.** in the more general setting where the $\{M_i, \mathbb{1} - M_i\}$'s are arbitrary two-outcome POVMs (in this case, if a state $|\phi\rangle$ is measured with this POVM, the post-measurement state under observing $\mathbb{1} - M_i$ is, up to scaling, given by $\sqrt{\mathbb{1} - M_i}\,|\phi\rangle$ rather than $(\mathbb{1} - M_i)|\phi\rangle$).

Solution:

1.A.

1.B.

1.C.

1.D.

2 (34 PTS.) METRIC ENTROPY OF CLASSICAL AND QUANTUM STATE SPACES (9/10 AND 9/15)

Motivation: Covering numbers quantify how large a space is at resolution ε and power many counting arguments in quantum/classical information. In this problem you will develop bounds on epsilon-nets for the following spaces: (i) pure states on n qubits and (ii) classical probability distributions on n bits.

Setup and notation: Let $D = 2^n$. We write $\|\cdot\|_2$ for the Euclidean/Frobenius norm and $\|\cdot\|_1$ for the vector ℓ_1 norm or trace norm as appropriate. For pure states $\psi, \phi \in \mathbb{C}^D$ with $\|\psi\|_2 = \|\phi\|_2 = 1$ define

$$d_{\mathrm{proj}}(\psi,\phi) \overset{\mathrm{def}}{=} \min_{\theta\in[0,2\pi)} \left\|\psi - e^{i\theta}\phi\right\|_2, \qquad d_{\mathrm{tr}}(\psi,\phi) \overset{\mathrm{def}}{=} \tfrac{1}{2}\left\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \right\|_1.$$

For classical distributions $p, q \in \Delta_{D-1} = \{x \in \mathbb{R}^D_{\geqslant 0} : \sum_i x_i = 1\}$, define total variation distance $\mathrm{TV}(p,q) = \frac{1}{2}\|p-q\|_1$.

What is a covering number? Fix a metric space $(\mathcal{X}, d)$ and a tolerance $\varepsilon > 0$. An ε-net is any finite "catalog" $S \subseteq \mathcal{X}$ such that every point of $\mathcal{X}$ lies within distance ε of *some* catalog item. The covering number

$$N(\mathcal{X}, d, \varepsilon) \;=\; \min\{|S| : \; S \subseteq \mathcal{X} \text{ is an } \varepsilon\text{-net}\}$$

is the smallest possible size of such a catalog.

What is $\mathbb{CP}^{D-1}$ and why global phase doesn't matter? Two unit vectors $\psi, \phi \in \mathbb{C}^D$ that differ only by a global phase, $\phi = e^{i\theta}\psi$, represent the same physical pure state: for every POVM $\{M_k\}$ the probabilities $p_k = \langle\psi|M_k|\psi\rangle$ equal $\langle\phi|M_k|\phi\rangle$ because $|\phi\rangle\langle\phi| = |\psi\rangle\langle\psi|$. Thus, the physically distinct pure states are *rays* (one-dimensional complex subspaces) in $\mathbb{C}^D$, not individual vectors. The space of all rays is the *complex projective space* $\mathbb{CP}^{D-1}$; equivalently, take the unit sphere $S^{2D-1} \subset \mathbb{C}^D \cong \mathbb{R}^{2D}$ and identify points that differ by a phase $e^{i\theta}$. Choosing a phase convention (e.g. "make the first nonzero coordinate real and $\geqslant 0$") just picks one representative from each ray.

Throughout you may assume $0 < \varepsilon \leqslant 1/4$ and use universal constants $c, C > 0$ that may change from line to line.

2.A. (8 PTS.) **Warm-up: covering the Euclidean ball.** Let $B^m = \{x \in \mathbb{R}^m : \|x\|_2 \leqslant 1\}$. Prove that

$$(c/\varepsilon)^m \;\leqslant\; N\big(B^m, \|\cdot\|_2, \varepsilon\big) \;\leqslant\; (C/\varepsilon)^m.$$

Hints: For the lower bound, compare the volume of B^m to the volume of the union of ε-balls around points in an ε-net. For the upper bound, try constructing an ε-net in a greedy fashion and again reason about volume ratios.

2.B. (3 PTS.) **From ball to sphere.** Let $S^{m-1} = \{x \in \mathbb{R}^m : \|x\|_2 = 1\}$. Prove that

$$(c/\varepsilon)^{m-1} \;\leqslant\; N\big(S^{m-1}, \|\cdot\|_2, \varepsilon\big) \;\leqslant\; (C/\varepsilon)^{m-1}.$$

Hints: For the upper bound, how would you take an ϵ-net constructed for a unit ball and convert that into one for a unit sphere? For the lower bound, how would you go in the reverse direction?

2.C. (5 PTS.) **Metric equivalence for pure states.** Show that for any unit vectors $\psi, \phi \in \mathbb{C}^D$,

$$d_{\mathrm{tr}}(\psi,\phi) \;\leqslant\; d_{\mathrm{proj}}(\psi,\phi) \;\leqslant\; \sqrt{2}\, d_{\mathrm{tr}}(\psi,\phi).$$

Hint: Align the global phase to make $\langle\psi,\phi\rangle \geqslant 0$, note that $\frac{1}{2}\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_1 = \sqrt{1 - |\langle\psi,\phi\rangle|^2}$ and $\|\psi - \phi\|_2 = \sqrt{2 - 2\langle\psi,\phi\rangle}$.

2.D. (6 PTS.) **Covering number for n-qubit pure states.** Let $\mathbb{CP}^{D-1}$ denote the set of rays (global-phase equivalence classes) of unit vectors in $\mathbb{C}^D$. Using Parts **2.B.** and **2.C.**, prove

$$(c/\varepsilon)^{2D-2} \;\leqslant\; N\big(\mathbb{CP}^{D-1}, d_{\mathrm{tr}}, \varepsilon\big) \;\leqslant\; (C/\varepsilon)^{2D-2}.$$

Guidance: For the upper bound, start from an ε'-net of S^{2D-1} with $\varepsilon' = \Theta(\varepsilon)$ and fix a phase convention (e.g., first nonzero coordinate real and $\geqslant 0$) to pass to projective space, using Part **2.C.**. For the lower bound, argue via a packing subset of $\mathbb{CP}^{D-1}$ that behaves like an embedded $(2D-2)$-dimensional sphere up to constants.

2.E. (12 PTS.) **Classical distributions on n bits (TV distance).** Show that

$$(c/\varepsilon)^{D-1} \;\leqslant\; N\big(\Delta_{D-1}, \mathrm{TV}, \varepsilon\big) \;\leqslant\; (C/\varepsilon)^{D-1}.$$

Your bound for c can depend polynomially on D and still receive full credit. We will award 4 points of extra credit for solutions where c is independent of D.

Upper bound hint: Quantize each coordinate to a grid of step $\alpha = \Theta(\varepsilon/D)$ and adjust one coordinate to preserve the sum 1; count feasible integer compositions via stars-and-bars to get $\binom{O(D/\varepsilon)+D-1}{D-1} \leqslant (C/\varepsilon)^{D-1}$.
Lower bound hint: Let S be a collection of points in Δ_{D-1} such that the total variation distance between any pair of points is at least 2ε. Show that $N(\Delta_{D-1}, \mathrm{TV}, \varepsilon)$ must be at least $|S|$, and then construct a set S of size $(c/\varepsilon)^{D-1}$ using similar ideas as in the upper bound proof.

Solution:

2.A.

2.B.

2.C.

2.D.

2.E.

3 (33 PTS.) LEARNING A LOW-RANK STATE (9/17)

Motivation: In the first lecture on state tomography, we saw an algorithm for learning d-dimensional quantum states in trace distance using $O(d^3/\epsilon^2)$ random measurements. In this exercise, we will iron out some of the proof details about subexponential random variables that were deferred from lecture, and then we will refine this bound under the additional assumption that the unknown state has bounded rank.

Setup: Recall that the estimator that was given in class performs measures N copies of the unknown state ρ with the uniform POVM $\{d\,|v\rangle\langle v|\,\mathrm{d}v\}$, obtaining measurement outcomes $|v_1\rangle, \ldots, |v_N\rangle$, and computes

$$\hat{\rho} \triangleq \frac{1}{N}\sum_{i=1}^{N}((d+1)\,|v_i\rangle\langle v_i| - \mathbb{1}) \tag{1}$$

In the first part of the question, we will prove the concentration inequality that was claimed in class and which is reproduced below in Eq. (3) using some tools from probability theory.

A random variable Z with mean zero is said to be $O(1)$-subexponential if $\mathbb{E}[|Z|^k] \leqslant O(k)^k$ for all powers $k \geqslant 1$. You may use the *subexponential Bernstein inequality* which says that any collection of independent $O(1)$-subexponential random variables $X_1, \ldots, X_N$ with mean zero satisfies the tail bound

$$\mathbf{Pr}\Big[\Big|\frac{1}{N}\sum_i X_i\Big| > t\Big] \leqslant \exp(-N\Omega(\min(t, t^2)))\,.$$

In the last two parts of the question, we will show how to refine the sample complexity bound from class in the special case where the state has bounded rank.

3.A. (10 PTS.) Prove the inequality

$$\mathbf{E}\left[((d+1)|\langle u|v\rangle|^2)^k\right] \leqslant O(k)^k\,, \tag{2}$$

where $|u\rangle$ is an arbitrary fixed unit vector, and the expectation is taken with respect to the *distribution over outcomes $|v\rangle$ of measuring ρ with the uniform POVM*, i.e., the distribution with probability density function $\langle v|\,\rho\,|v\rangle$.

3.B. (8 PTS.) Prove that Eq. (2) implies that for any unit vector $|u\rangle$, each of the random variables $\langle u|\,((d+1)\,|v_i\rangle\langle v_i| - \mathbb{1} - \rho)\,|u\rangle$ for the measurement outcomes $|v_i\rangle$ in Eq. (1) is $O(1)$-subexponential (you may use without proof the fact, already shown in class, that these random variables have mean zero).

3.C. (2 PTS.) Conclude from Part **3.B.** and the subexponential Bernstein inequality that for any unit vector $|u\rangle$,

$$\mathbf{Pr}\Big[|\langle u|\,(\rho - \hat{\rho})\,|u\rangle| > t\Big] \leqslant \exp(-N\Omega(\min(t, t^2)))\,, \tag{3}$$

as was claimed in class.

Recall from class that we used Eq. (3) as a black box to conclude that for any $0 < \eta \leqslant 1$, with probability at least $1-\delta$,

$$\|\rho - \hat{\rho}\|_{\mathsf{op}} \leqslant \eta \qquad \text{if} \qquad N \geqslant \Omega(d/\eta^2)\,. \tag{4}$$

By taking $\eta = \epsilon/d$ and converting from operator norm to trace norm, we concluded that $O(d^3/\epsilon^2)$ samples suffice to learn ρ to trace distane ϵ. In the last two parts of this problem, we will refine this sample complexity bound in the special case where ρ has bounded rank.

Henceforth, assume that ρ has rank r for some $1 \leqslant r \leqslant d$. Let us define $\hat{\rho}_{\mathsf{LR}} \triangleq \mathrm{proj}(\hat{\rho})$, where proj is the projection to the space of rank-r density matrices given by removing all but the r largest eigenvalues. Concretely, if $\hat{\rho}$ has eigendecomposition $U\mathrm{diag}(\lambda_1, \ldots, \lambda_d)U^\dagger$ for $\lambda_1 \geqslant \cdots \geqslant \lambda_d$, then $\mathrm{proj}(\hat{\rho}) = U\mathrm{diag}(\lambda_1, \ldots \lambda_r, 0, \ldots, 0)U^\dagger$.

3.D. (6 PTS.) Show that $\|\rho - \hat{\rho}_{\mathsf{LR}}\|_{\mathsf{op}} \leqslant 2\|\rho - \hat{\rho}\|_{\mathsf{op}}$

3.E. (7 PTS.) If ρ has rank r, deduce from Eq. (4), for an appropriate choice of η, that $\|\rho - \hat{\rho}_{\mathsf{LR}}\|_{\mathsf{tr}} \leqslant \epsilon$ provided $N \geqslant \Omega(dr^2/\epsilon^2)$.

Solution:

3.A.

3.B.

3.C.

3.D.

3.E.