PSET 1 Due on **Thursday September 25, 2025 at 11:59 pm ET**

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

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

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

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

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

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

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

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

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.*

3.B. (8 PTS.) Prove that Eq. (2) implies that for any unit vector $|u\rangle$, each of the random variables $\langle u|\,((d+1)\,|v_i\rangle\langle v_i| - \mathbb{1})\,|u\rangle$ 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 $\|\rho - \hat{\rho}_{\mathsf{LR}}\|_{\mathsf{op}} \leqslant \|\rho - \hat{\rho}\|_{\mathsf{op}}$

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.

4 (50 PTS.) QUANTUM COMPUTING ENHANCED SENSING FOR OSCILLATING FIELDS (9/3, 9/8, AND 9/10)

Motivation: Sensing oscillating signals is a fundamental task in many areas of science and technology, from searching for dark matter to medical imaging. A key challenge is when the signal's frequency is unknown, requiring a search over a wide range of possibilities. In this problem, we will explore how quantum computers can provide a significant speedup for this task. We will compare the conventional sensing approach, where a spin sensor (a quantum sensor) is controlled by classical means, to a quantum computing enhanced approach, where the spin sensor is coherently controlled by a quantum computer. This problem is inspired by the recent paper *Quantum Computing Enhanced Sensing* by Allen et al. (arXiv:2501.07625).

Setup: We model the signal as a time-dependent magnetic field that couples to a single-spin sensor (model as a single qubit) via the following time-dependent single-qubit Hamiltonian,

$$H(t) = B(t)Z,$$

where Z is the Pauli-Z operator and $B(t)$ is a real-valued function. We want to distinguish between two cases:

- **Null Hypothesis:** No signal is present, so $B(t) = 0$.
- **Alternative Hypothesis:** An oscillating signal is present, $B(t) = B\cos(\omega t + \phi)$, for a known field strength $B > 0$, but an unknown frequency ω and phase $\phi \in [0, 2\pi)$. (in practice, B will also be unknown and there are known techniques for handling them; for this problem we will focus on B being known)

We will analyze two different models for performing this sensing task:

- **Classically-Controlled Sensor:** We can prepare an arbitrary 1-qubit state $|\psi\rangle$, let it evolve under the Hamiltonian $f(t)H(t)$ for a chosen duration, and then perform a measurement. The classical filter function $f(t) \in [0, 1]$ can be modulated in time to control the sensor's interaction with the field ($f(t) = 0$ means the sensor is removed from the field while $f(t) = 1$ means the sensor is fully immersed in the field). This process can be repeated, and later experiments can be chosen adaptively based on previous measurement outcomes. The total sensing time is the sum of all evolution periods.
- **Quantum Computing (QC) Enhanced Sensor:** The sensor qubit can be controlled by an n-qubit universal quantum computer. We can apply arbitrary quantum gates to the sensor qubit and the n computing qubits, interleaved with periods of quantum evolution under the Hamiltonian $H(t)$ on the sensor qubit. The output state can be formally written as

$$U_J(e^{-i\int_{t_{J-1}}^{t_J} dt B(t) Z} \otimes I)U_{J-1} \dots U_3(e^{-i\int_{t_2}^{t_3} dt B(t) Z} \otimes I)U_2(e^{-i\int_{t_1}^{t_2} dt B(t) Z} \otimes I)U_1 |0^{1+n}\rangle\,.$$

4.A. (10 PTS.) As a warm-up, suppose both the frequency ω and phase ϕ are known. Design a protocol that achieves $\tau = O(1/B)$.

4.B. (5 PTS.) What is the minimum sensing time τ required to distinguish the null and alternative hypotheses with a constant success probability? In quantum hypothesis testing between quantum states ρ and σ, the maximal probability of distinguishing between them is bounded by $\frac{1}{2} + \frac{1}{2} \cdot \|\rho - \sigma\|_1$, where $\|\rho - \sigma\|_1$ is the trace norm between ρ and σ. Show that $\tau = \Omega(1/B)$ is necessary for *both* sensor models and briefly explain why there is no quantum computational advantage in this simple case.

For the next few parts, suppose the signal is periodic under a known period of 1, i.e., $B(t) = B(t+1), \forall t$. This is equivalent to the frequency ω being an integer multiple of 2π. As we do not expect the oscillation to be infinitely fast, we bound the largest frequency to be $2\pi W$. Hence, ω is an unknown frequency in the discrete set $\{2\pi, 4\pi, \dots, 2\pi W\}$. The phase $\phi \in [0, 2\pi)$ remains fully unknown.

4.C. (10 PTS.) For the **classically-controlled sensor**, design a protocol to determine if a signal is present. Show that your protocol requires a total sensing time of $\mathcal{O}(W/B)$. Briefly explain how your protocol handles the unknown phase ϕ.

4.D. (10 PTS.) We now consider implementing a **QC-enhanced sensor** using Grover's unstructured search algorithm.

Here is a brief recap of Grover's unstructured search algorithm. Grover's algorithm addresses the unstructured search problem: given an oracle function $f : \{0,1\}^n \to \{0,1\}$ with a unique marked element $x^\star$ such that only $f(x^\star) = 1$, the task is to identify $x^\star$. The algorithm begins with the uniform superposition $|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle$, and iteratively implement Grover operator $G = \big(2|\psi_0\rangle\langle\psi_0| - I\big)\big(I - 2|x^\star\rangle\langle x^\star|\big)$. Each application of G rotates the state vector within the two-dimensional subspace spanned by $|x^\star\rangle$ and $|\psi_0\rangle$. A measurement in the computational basis yields $x^\star$ with probability close to one after $\mathcal{O}(\sqrt{2^n})$ rotations.

Design a protocol that constructs a multi-qubit Grover's oracle over W elements and uses Grover's unstructured search algorithm to achieve a total sensing time of $\mathcal{O}(\sqrt{W}/B)$.

4.E. (10 PTS.) Here, we are to use Rudin–Shapiro sequence to design an improved **classically-controlled sensor**. The Rudin-Shapiro sequence is a deterministic binary sequence $\{a_j\}_{j=0}^{M-1}$ with $a_j \in \{\pm 1\}$. Its discrete Fourier transform $S(\ell) = \sum_{j=0}^{M-1} a_j e^{i2\pi\ell j/M}, \ell = 0, 1, \dots, M-1$ satisfies two contrasting bounds. At $\ell = 0$, we have $|S(0)| = M$, while for all $\ell \neq 0$, we have $|S(\ell)| \leqslant C\sqrt{M}$ for a universal constant C. Design an improved protocol over the one in **4.C.** to achieve a total sensing time of $\mathcal{O}(\sqrt{W}/B)$.

4.F. (5 PTS.) Prove that any QC-enhanced sensing protocol for this task requires $\Omega(\sqrt{W}/B)$ sensing time, which shows that your protocol from Part **4.D.** is asymptotically optimal. Describe why the answers to the questions above show that quantum computers offer no asymptotic advantage when the signal is periodic and the frequency is discretized.

(Optional) Now, we consider the more realistic scenario where the frequency ω is an unknown real value in the continuous range $[1, W]$, and the phase $\phi \in [0, 2\pi)$ is also unknown.

4.G. (10,OPTIONAL PTS.) For the **classically-controlled sensor**, design a protocol to achieve a sensing time of $\mathcal{O}(W/B^2)$.

4.H. (2.5,OPTIONAL PTS.) What is the optimal sensing time for a **classically-controlled sensor** and a **QC-enhanced sensor** for this more realistic scenario when ω is continuous? [Hint: Read Allen et al.]

4.I. (2.5,OPTIONAL PTS.) Suppose W and B depend on a problem parameter $N \to \infty$: $W(N) = N$ and $B(N) = N^{\mu}$. Find the condition on μ such that the classical sensing time T_C is the *fourth power* of the QC-enhanced sensing time T_Q: $T_C = \Theta(T_Q^4)$ (a quartic advantage), and the condition on μ that only gives a tiny sub-quadratic advantage: $T_C = \Theta(T_Q^{1.01})$.

Solution:

4.A.

4.B.

4.C.

4.D.

4.E.

4.F.

4.G.

4.H.

4.I.