PSET 5

Due on **Tuesday November 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 touch upon agnostic tomography and learning tree lower bounds.
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 (72 pts.) Agnostic tomography (11/10 and Section 11/14)

Motivation: In class we saw an algorithm for agnostic tomography of stabilizer states in the regime where the fidelity between the unknown state and the closest stabilizer state exceeds some constant threshold. In this exercise, we will explore agnostic tomography of another class of quantum states, namely *discrete product states*. Whereas the algorithm in class heavily exploited the delicate algebraic structure of stabilizer states, the algorithm we will consider in this exercise is more general-purpose. The general outline of the algorithm was covered in the 11/14 section, and the objective of this problem is to work carefully through the details.

Setup: Let $\mathcal{K}$ be some known, finite set of pure 1-qubit states. We assume that the states in this state are somewhat far apart from each other; formally, there is a $\mu > 0$ such that for any $|\psi\rangle, |\psi'\rangle \in \mathcal{K}$,

$$|\langle\psi|\psi'\rangle|^2 \leqslant 1 - \mu\,.$$

Let $\mathcal{C} = \mathcal{K}^{\otimes n}$ denote the set of all product states of the form $|\psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$ where $|\psi_i\rangle \in \mathcal{C}$ for all $i \in [n]$. In this exercise, we will consider the question of agnostic tomography for this class of states. Throughout, let $|\phi\rangle$ denote an unknown pure state to which we have access to copies, and let

$$\tau := \max_{|\psi\rangle \in \mathcal{C}} |\langle\phi|\psi\rangle|^2\,.$$

Let $|\psi^*\rangle = |\psi_1^*\rangle \otimes \cdots \otimes |\psi_n^*\rangle \in \mathcal{C}$ denote the argmax for the above.

1.A. (4 pts.) **Warmup: non-agnostic tomography.** Suppose we know that $\tau = 1$, i.e., that $|\phi\rangle \in \mathcal{C}$. Then give an algorithm for learning $|\phi\rangle$ to zero error using $O(\log(n)/\mu^2)$ samples with probability at least $9/10$.

1.B. (8 pts.) **Geometry of $\mathcal{K}$.** Prove that for any 1-qubit pure state $|\tau\rangle$, there is at most one $|\psi'\rangle \in \mathcal{K}$ for which

$$|\langle\psi'|\tau\rangle|^2 > \frac{1+\sqrt{1-\mu}}{2}$$

For every $i \in [n]$, let $\rho_i \triangleq \mathrm{tr}_{[n]\backslash\{i\}}(|\psi\rangle\langle\psi|)$ denote the 1-qubit mixed state given by tracing out all but the i-th qubit of $|\psi\rangle\langle\psi|$.

1.C. (15 pts.) **Greedy local optimization.** Give an algorithm for finding a state $|\hat{\phi}\rangle \in \mathcal{C}$ using only $O(\log(n|\mathcal{K}|)/\mu^2)$ copies of $|\phi\rangle$ such that for every $i \in [n]$ and for every $|\psi_i\rangle \in \mathcal{K}\backslash\{|\hat{\phi}_i\rangle\}$,

$$\langle\psi_i|\rho_i|\psi_i\rangle \leqslant \frac{1+\sqrt{1-\mu}}{2} + \frac{\mu}{8} := g(\mu)\,. \tag{1}$$

Hint: Estimate the fidelities $\langle\psi_i|\rho_i|\psi_i\rangle$ for all $i \in [n]$ and $|\psi_i\rangle \in \mathcal{K}$, and for each i pick $|\hat{\phi}_i\rangle$ to be $|\psi_i\rangle$ which maximizes (up to estimation error) the fidelity w.r.t ρ_i. Argue that this satisfies Eq. (1) using Part 1.B.

1.D. (12 pts.) **Amplifying fidelity via measurement.** Suppose $|\hat{\phi}\rangle \neq |\psi^*\rangle$, and let $i \in [n]$ be an index for which $|\hat{\phi}_i\rangle \neq |\psi_i^*\rangle$. Let Π denote the projector which acts as $|\psi_i^*\rangle\langle\psi_i^*|$ on the i-th qubit and as $\mathbb{1}$ elsewhere. Consider measuring $|\phi\rangle$ using $\{\Pi, \mathbb{1} - \Pi\}$, and let $|\phi'\rangle$ denote the post-measurement state upon observing the outcome Π. Prove that outcome Π is observed with probability at least τ, and prove that

$$|\langle\phi'|\psi^*\rangle|^2 \geqslant g(\mu)^{-1}|\langle\phi|\psi^*\rangle|^2\,.$$

1.E. (5 pts.) **Finding the right measurement.** Under the assumption that $|\hat{\phi}\rangle \neq |\psi^*\rangle$, give an algorithm that with probability at least $1/(n|\mathcal{K}|)$ finds $i \in [n]$ such that $|\hat{\phi}_i\rangle \neq |\psi_i^*\rangle$, as well as $|\psi_i^*\rangle$.

Part 1.D. implies that the post-measurement state $|\phi'\rangle$ has higher fidelity with respect to $|\psi^*\rangle$ than $|\phi\rangle$ does, and Part 1.E. tells us that it is possible to find an appropriate measurement which induces this post-measurement state with non-negligible probability. The natural thing to do now is to iterate this: we replace $|\phi\rangle$ with $|\phi'\rangle$ and repeat the same reasoning, each time amplifying the fidelity by a factor $g(\mu)^{-1} > 1$ until the fidelity can't increase further, at which point we are guaranteed to have found $|\psi^*\rangle$. What remains is for you to make the above proof sketch formal.

1.F. (20 PTS.) **Putting everything together.** Complete the above reasoning to give an algorithm that uses $\log(n) \cdot \mathrm{poly}(1/\tau, 1/\mu)$ copies of $|\phi\rangle$, runs in time $\mathrm{poly}(n, 1/\tau, 1/\mu)$, and outputs $|\psi^*\rangle$ with probability at least

$$\left(\frac{1}{n|\mathcal{K}|}\right)^{O((1+\log(1/\tau))/\mu)}.$$

1.G. (8 PTS.) **Boosting the success probability.** The success probability above is quite small. Informally sketch how to use the algorithm above to derive an algorithm that outputs a state $|\psi\rangle \in \mathcal{C}$ for which $|\langle\phi|\psi\rangle|^2 \geqslant \tau - \epsilon$ with probability $9/10$. For full credit, you do not need to give a formal proof or runtime analysis, but you must give an algorithm whose runtime is $\mathrm{poly}(n, 1/\epsilon)$ when $\tau, \mu = \Theta(1)$, as well as high-level justification for why this is the case.

Solution:

1.A.

1.B.

1.C.

1.D.

1.E.

1.F.

1.G.

2 (28 pts.) Purity testing lower bound for computational basis measurements (11/17)

Motivation: In this problem you will explore a concrete version of an exponential copy lower bound for the purity testing problem if you are only allowed to perform computational basis measurements.

Setup: Let $d = 2^n$. You receive i.i.d. copies of an unknown state ρ with the promise

$$H_0: \ \rho = \tfrac{I}{d} \qquad \text{or} \qquad H_1: \ \rho = |\psi\rangle\langle\psi| \ \text{ with } |\psi\rangle \text{ Haar-random.}$$

You are restricted to *single-copy* measurements in the computational basis $\{|1\rangle, \ldots, |d\rangle\}$. On each copy you record the outcome $X_t \in [d]$, for $t = 1, \ldots, T$. Define the *collision count*

$$S := \sum_{1 \leqslant u < v \leqslant T} \mathbb{1}\{X_u = X_v\}.$$

which is the number of pairs of measurements that yield the same computational-basis outcome.

2.A. (10 pts.) **Expectations under H_0 and H_1.** Show that

$$\mathbb{E}_{H_0}[S] = \binom{T}{2}\frac{1}{d} \qquad \text{and} \qquad \mathbb{E}_{H_1}[S] = \binom{T}{2}\frac{2}{d+1}.$$

2.B. (10 pts.) **A variance bound.** Show that under H_0,

$$\mathrm{Var}_{H_0}(S) = \Theta\left(\frac{T^2}{d}\right).$$

2.C. (8 pts.) **A lower bound.** Let μ_0, μ_1 be the expectations of S under H_0 and H_1 from part (a), and let $\sigma_0^2 = \mathrm{Var}_{H_0}(S)$ from part (b). You may assume that under H_1,

$$\sigma_1^2 = \mathrm{Var}_{H_1}(S) = \Theta\left(\frac{T^2}{d} + \frac{T^3}{d^2} + \frac{T^4}{d^3}\right).$$

First note the scalings

$$\mu_1 - \mu_0 = \Theta\left(\frac{T^2}{d}\right) \qquad \text{and} \qquad \sigma_0 = \Theta\left(\frac{T}{\sqrt{d}}\right).$$

Observe that when $T \ll \sqrt{d}$, the gap in means $|\mu_1 - \mu_0|$ is much smaller than σ_0 (and also than σ_1).
It suffices to analyze *threshold tests* that decide H_1 if and only if $S \geqslant \tau$ for some $\tau \in \mathbb{R}_+$. Take $\tau \in [\mu_0, \mu_1]$ and apply the one-sided Chebyshev (Cantelli) inequality under each hypothesis:

$$\Pr_{H_0}[S \geqslant \tau] \leqslant \frac{\sigma_0^2}{\sigma_0^2 + (\tau - \mu_0)^2}, \qquad \Pr_{H_1}[S < \tau] \leqslant \frac{\sigma_1^2}{\sigma_1^2 + (\mu_1 - \tau)^2}.$$

Deduce that if both error probabilities are at most ε, then

$$\mu_1 - \mu_0 \ \geqslant \ \sqrt{\tfrac{1}{\varepsilon} - 1}\,(\sigma_0 + \sigma_1).$$

Using the scalings above (in particular, when $T \ll \sqrt{d}$ we have $\sigma_0, \sigma_1 = \Theta(T/\sqrt{d})$), conclude that for $\varepsilon = 1/3$ this inequality cannot hold unless $T = \Omega(\sqrt{d})$. Equivalently, since $d = 2^n$, one needs $T = \Omega(2^{n/2})$ to achieve success probability at least $2/3$.

Solution:

2.A.

2.B.

2.C.