Submitted by <**name**>

This problem set will cover concepts from quantum state tomography and shadow tomography.
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 (40 pts.) Tying up some loose threads about tomography (9/24, 9/29)

Motivation: In class we learned about a protocol for state tomography that achieves optimal sample complexity, up to logarithmic factors, using representation theory. In this exercise, we will go over a few isolated details that we did not have time to cover in class.

1.A. (5 pts.) **Power of entangled measurements.** The following was sketched briefly in class, and your goal is to provide a rigorous proof. We say that a learning protocol that is given a collection of N copies of an unknown state ρ uses *adaptive unentangled measurements* if it implements the following interaction:

- For $i = 1, \ldots, N$:
  - Based on the outcomes of the measurements performed so far on previous copies of ρ, choose a POVM $\{M_z\}$
  - Measure the i-th copy of ρ with POVM, and record the outcome, call it z_i, of the measurement. Discard the post-measurement state.

Prove that this interaction, which results in a sequence of outcomes $z_1, \ldots, z_N$, can be simulated with a single entangled measurement performed on $\rho^{\otimes N}$.

1.B. (5 pts.) **Block-diagonal structure.** *Prove or disprove:* Let $\rho \in \mathbb{C}^{d\times d}$ be an unknown density matrix which is known to be block-diagonal with blocks corresponding to known projectors $\Pi_1, \ldots, \Pi_m$. Consider measuring ρ with a POVM $\{M_z\}$. Then this is equivalent to instead performing the POVM $\{\Pi_1, \ldots, \Pi_m\}$ first, and then performing $\{M_z\}$ on the post-measurement state.

1.C. (12 pts.) **Specht modules are irreps of $\mathcal{S}_N$.** Prove the following facts stated in class, using Lemmas 74 and 75 from the lecture notes: (i) $V_\lambda = \mathbb{C}[\mathbb{S}_N]c_\lambda$ is an irrep of $\mathcal{S}_N$ for any $\lambda \vdash [N]$, (ii) For any distinct partitions $\lambda, \mu \vdash [N]$, V_λ and V_μ are not isomorphic.

Hint: For (i), you may find it helpful to use the fact that if $W \subseteq V_\lambda$ is a subspace, then $W \cdot W = 0$ implies $W = 0$.

1.D. (12 pts.) **Characterization of $\mathbb{S}_\lambda V_\lambda$.** Prove that for any $g \in \mathbb{C}[\mathcal{S}_N]$, the space $g \cdot \mathcal{H}$, where $\mathcal{H} = (\mathbb{C}^d)^{\otimes N}$, has the structure of a GL_d-representation.[1] Then use Lemma 74 to prove that the GL_d-representation $\mathbb{S}_\lambda V_\lambda$ is isomorphic to the representation $c_\lambda \cdot \mathcal{H}$. You may use without proof the fact that c_λ^2 is a nonzero scalar multiple of c_λ.

Hint: Note that any $f \in \mathbb{S}_\lambda V_\lambda$ is entirely determined by where it maps c_λ.

1.E. (6 pts.) **Saturating Fuchs-van de Graaf.** For any $0 \leqslant \epsilon \leqslant 1$, construct a pair of density matrices σ, σ' such that

$$\epsilon = \frac{1}{2}\|\sigma - \sigma'\|_{\mathsf{tr}} = \sqrt{1 - F(\sigma, \sigma')^2}\,,$$

where the fidelity $F(\sigma, \sigma')$ is defined by $F(\sigma, \sigma') \triangleq \mathrm{tr}(\sqrt{\sqrt{\sigma}\sigma'\sqrt{\sigma}})$.

Hint: You may find it helpful to focus on pure states.

Solution:

1.A.

1.B.

1.C.

1.D.

1.E.

[1] If $g \in \mathbb{C}[\mathcal{S}_N]$ is given by $g = \sum_{\pi \in \mathcal{S}_N} a_\pi \pi$, and $v \in \mathcal{H}$, then $g \cdot v$ denotes the element $\sum_\pi a_\pi \pi \cdot v \in \mathcal{H}$. The expression $c_\lambda \cdot \mathcal{H}$ then denotes the subspace of $\mathcal{H}$ consisting of all points of the form $c_\lambda \cdot v$ for $v \in \mathcal{H}$.

2 (40 PTS.) ESTIMATING MANY PAULI EXPECTATIONS IN PARALLEL (10/1)

Motivation: Suppose we are given m Pauli operators $C_1, \dots, C_m$ on n qubits, each of weight at most k. We want to estimate all expectations $\mu_i = \operatorname{tr}(C_i \rho)$ to additive accuracy ε, with failure probability at most δ, using as few copies of ρ as possible. An important idea (the basis of *quantum overlapping tomography*, a precursor to *classical shadow tomography*) is to measure, on each copy of ρ, a random full-weight Pauli string from $\{X, Y, Z\}^{\otimes n}$ and to reuse these outcomes to estimate every C_i supported on that string. You will prove that

$$N = O\left(\frac{3^k}{\varepsilon^2} \log \frac{m}{\delta}\right)$$

samples suffice.

Setup: On each of N copies of ρ, draw $P^{(t)} \in \{X, Y, Z\}^{\otimes n}$ uniformly at random and measure it, obtaining outcomes $o_j^{(t)} \in \{\pm 1\}$. For $C_i = \bigotimes_{j=1}^n \sigma_{i,j}$ with support $\operatorname{supp}(C_i) = \{j : \sigma_{i,j} \neq I\}$, set

$$\mathbb{1}_i^{(t)} = \mathbf{1}\left[\forall j \in \operatorname{supp}(C_i) : P_j^{(t)} = \sigma_{i,j}\right],$$
$$Y_i^{(t)} = \prod_{j \in \operatorname{supp}(C_i)} o_j^{(t)}.$$

If $\mathbb{1}_i^{(t)} = 1$, then $Y_i^{(t)}$ is a valid single-shot outcome for C_i. Let

$$S_i = \sum_{t=1}^N \mathbb{1}_i^{(t)}, \qquad \widehat{\mu}_i = \frac{1}{\max\{1, S_i\}} \sum_{t:\, \mathbb{1}_i^{(t)}=1} Y_i^{(t)}.$$

2.A. (4 PTS.) **Hit probability.** Show that for each i,

$$p_i = \mathbf{Pr}[\mathbb{1}_i^{(t)} = 1] = 3^{-\mathrm{wt}(C_i)}.$$

Deduce the uniform lower bound $p_i \geqslant 3^{-k}$ for all i.

2.B. (6 PTS.) **Enough hits.** Use the multiplicative Chernoff bound for $S_i \sim \operatorname{Bin}(N, p_i)$ and provide an N such that with probability at least $1 - \delta/3$

$$S_i \geqslant (1-\varepsilon) N p_{\min} \quad \text{for all } i$$

where $p_{\min} = \min_i p_i \geqslant 3^{-k}$.

2.C. (10 PTS.) **Accuracy given enough hits.** Condition on $S_i \geqslant T$. Show that Hoeffding's inequality implies

$$\mathbf{Pr}\left[|\widehat{\mu}_i - \mu_i| > \varepsilon \;\middle|\; S_i \geqslant T\right] \leqslant \frac{\delta}{3m}$$

provided $T \geqslant \frac{2}{\varepsilon^2} \log(3m/\delta)$.

2.D. (8 PTS.) **Putting it all together.** Combine the previous parts and a union bound over all i to conclude that if

$$N \geqslant \frac{2}{\varepsilon^2 (1-\varepsilon)\, p_{\min}} \log\left(\frac{3m}{\delta}\right),$$

then with probability at least $1 - \delta$, all estimates satisfy $|\widehat{\mu}_i - \mu_i| \leqslant \varepsilon$. Simplify using $p_{\min} \geqslant 3^{-k}$ to obtain the stated $O(3^k \varepsilon^{-2} \log(m/\delta))$ scaling.

2.E. (12 PTS.) **Extension: commutators.** Define $K_{jk} = \operatorname{tr}\big(i[P_j, P_k]\rho\big)$. Show that if each P_j has weight $\leqslant k$, then any nonzero commutator has weight at most $2k-1$. Deduce the number of samples needed so that all nonzero entries of K are estimated to accuracy ε with probability $\geqslant 1 - \delta$.

Solution:

2.A.

2.B.

2.C.

2.D.

2.E.

3 (40 PTS.) ADVANCED PAULI EXPECTATION ESTIMATION (10/1)

Motivation: In earlier problems, we analyzed methods for efficiently estimating expectation values of *low-weight* Pauli operators. We now extend our analysis to the more general and challenging case where the Pauli observables are not guaranteed to be low-weight.

Part I: Commuting and Partitionable Observables (26 points)

Consider a set of m distinct n-qubit Pauli observables $S = \{C_1, \ldots, C_m\}$ that pairwise commute, i.e. $[C_i, C_j] = 0$ for all $i \neq j$.

3.A. (8 PTS.) **Simultaneous Diagonalization.** Show that there exists a unitary U such that for every i, $U^\dagger C_i U$ is a tensor product of I and Z operators.

Hint: Try to construct U in a sequence of steps that gradually transform the C_i's into products of I and Z operators, as if you were performing Gaussian elimination.

3.B. (8 PTS.) **Sample Complexity for Commuting Sets.** Using the result from part (a), prove that $O(\log(m/\delta)/\varepsilon^2)$ measurements on individual copies of an unknown state ρ are sufficient to estimate all expectation values $\mathrm{tr}(C_i\rho)$ for $C_i \in S$ to additive error ε with total failure probability at most δ.

3.C. (10 PTS.) **Partitioned Sets.** Now consider an arbitrary set S of m Pauli observables that can be partitioned into M disjoint commuting subsets $S = S_1 \cup S_2 \cup \cdots \cup S_M$, where observables within each S_j mutually commute. Prove that $O(M \log(m/\delta)/\varepsilon^2)$ total measurements on individual copies of ρ are sufficient to estimate all expectation values in S to the desired accuracy.

Part II: The General Case and Fractional Coloring (14 points)

Consider an arbitrary set $S = \{C_1, \ldots, C_m\}$ of Pauli observables that may not commute with one another. To analyze this case, we introduce the concept of *fractional coloring* of a graph.

Definition: Let $G = (V, E)$ be a graph. A **fractional coloring** of G with parameter χ is a probability distribution q over the independent sets $I \subseteq V$ of the graph such that every vertex $v \in V$ has probability at least $1/\chi$ of being included in an independent set randomly sampled from the distribution q:

$$\forall v \in V : \quad \mathop{\mathbf{Pr}}_{I \sim q}[v \in I] \geqslant 1/\chi.$$

The smallest possible value of χ for which such a distribution exists is called the **fractional chromatic number** of G, denoted $\chi_f(G)$.

3.D. (14 PTS.) **Fractional Coloring and Sample Complexity.** Define the *anticommutation graph* $G(S)$ where each Pauli observable in S corresponds to a vertex, and an edge connects any two observables that **anticommute**. Prove that if $G(S)$ admits a fractional coloring with parameter χ, then a total of $O(\chi \log(m/\delta)/\varepsilon^2)$ measurements on individual copies of ρ suffice to estimate all expectation values in S to additive error ε with total failure probability at most δ.

Solution:

3.A.

3.B.

3.C.

3.D.