1 Lecture 1 exercise

Problem 1. In this problem you will be showing that the pretty good measurement is very-closely related to the optimal measurement strategy for all quantum learning problems.

Consider an ensemble of size m, consisting of real d-dimensional states, $\mathscr{E} = \{(p_i, |\psi_i\rangle)\}_{i\in[m]}$ (where $\sum_{i\in[m]} p_i = 1$). Suppose we are given an unknown state $|\psi_?\rangle$ sampled according to the probabilities $\{p_i\}$ and we are interested in maximizing the average probability of success to identify the state that we are given. For a POVM specified by positive semidefinite matrices $\mathsf{M} = \{M_i\}_{i\in[m]}$, the probability of obtaining outcome j, given the state $|\psi_i\rangle$ is given by $\langle\psi_i|M_j|\psi_i\rangle$ and the average success probability is defined as

$$P_{\mathsf{M}}(\mathscr{E}) = \sum_{i=1}^{m} p_i \langle\psi_i|M_i|\psi_i\rangle.$$

Let $P^{opt}(\mathscr{E}) = \max_{\mathsf{M}} P_{\mathsf{M}}(\mathscr{E})$ denote the optimal average success probability of $\mathscr{E}$, where the maximization is over the set of valid POVMs.

The so-called *Pretty Good Measurement* (PGM) is a specific POVM, which we defined in the lecture as $\mathscr{M} = \{|\nu_i\rangle\langle\nu_i|\}_{i\in[m]}$ where $|\nu_i\rangle = \rho^{-1/2}|\psi_i'\rangle$, $|\psi_i'\rangle = \sqrt{p_i}|\psi_i\rangle$ and $\rho = \sum_{i\in[m]} |\psi_i'\rangle\langle\psi_i'|$. Suppose $P^{PGM}(\mathscr{E})$ is defined as the average success probability of distinguishing the states in $\mathscr{E}$ using the pretty good measurement, then the goal is to show the following

$$P^{opt}(\mathscr{E})^2 \leq P^{PGM}(\mathscr{E}) \leq P^{opt}(\mathscr{E}).$$

Hint:

1. First verify that $\mathscr{M}$ is a POVM
2. Write down the success probability of the PGM in terms of the entries of the Gram matrix $G(i,j) = \langle\psi_i|\psi_j\rangle$
3. Write down the success probability $P^{opt}(\mathscr{E})$ and then use Cauchy-Schrwartz inequalities to prove the inequality.

Problem 2. In this exercise we will formally prove the PAC learning lower bound. To that end, we will use the following theorem, which we will be the main technical theorem one needs.

Theorem 1. *Let $f : \{0,1\}^m \to \mathbb{R}$ be defined as $f(z) = (1 - \beta\frac{|z|}{m})^T$ for some $0 < \beta < 1$ and $1 \leq T \leq m/(e^3\beta)$. For $k \leq m$, let $M \in \mathbb{F}_2^{m\times k}$ be a matrix with rank k. Suppose $G \in \mathbb{R}^{2^k\times 2^k}$ is defined as $G(x,y) = (f\circ M)(x+y)$ for $x, y \in \{0,1\}^k$, then*

$$\sqrt{G}(x,x) \leq \exp(T^2\beta^2/m + \sqrt{Tm\beta} - T\beta - k/2) \qquad \text{for all } x \in \{0,1\}^k.$$

Let $\mathscr{C} \subseteq \{c : \{0,1\}^n \to \{0,1\}\}$ be an arbitrary concept class with VC dimension d. Prove that the quantum PAC sample complexity is $\Omega(d/\varepsilon)$.

Hint:

1. Fix a distribution D on the shattered set
2. Reduce the PAC learning task to exact state identification.
3. Write down the success probability of the PGM and use the theorem to show that if $T = o(d/\varepsilon)$, then the success probability of the PAC learner is small.

Problem 3. In this problem you will be showing quantum examples help the "coupon collector problem". Suppose a quantum learner is given copies of the state

$$|S\rangle = \frac{1}{\sqrt{S}} \sum_{i \in S} |i\rangle,$$

where $S \subseteq [n]$ is of size $|S| = n - 1$. The learner is given copies of $|S\rangle$ (for a uniformly random S) and needs to learn the unknown S.

1. Suppose a classical learner chooses to measure $|S\rangle$ on receiving every copy, how many copies are sufficient for learning the unknown S?

2. If a quantum learner is allowed to perform entangled measurements on copies of $|S\rangle$, how many copies suffice to learn $|S\rangle$?

   Hint: Analyze the problem using the pretty good measurement on the ensemble $\{|S\rangle^{\otimes T} : S \subseteq [n], |S| = n-1\}$.

   - Write down the POVM elements of the PGM (in order to take the square-root of the Gram matrix, observe that it is structured enough to diagonalize easily).
   - Write down the average success probability of the PGM and show that it is large if $T = O(n)$.