1 Lecture 2 exercise

Problem 1. Verify that parities have Fourier sparsity 1 and k-juntas have Fourier sparsity 2^k. Recall that the class of k-juntas is defined as

$$\mathcal{C} = \{c : \{0,1\}^n \to \{0,1\} | c(x) = c(x_S), S \subseteq [n] : |S| = k\},$$

i.e., there is an unknown set of k indices (call that S) such that $c(x)$ only depends on the values of x when restricted to S.

Hint: Write down the Fourier decomposition of parities and juntas and check when are they non-zero.

Problem 2. Let $f : \{0,1\}^n \to \{0,1\}$. Decribe a procedure that uses one copy of $\frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x, f(x)\rangle$, and with probability $1/2$, outputs an S drawn from the distribution $\{\widehat{f}(S)^2\}_S$, otherwise rejects

Hint: Apply Hadamard on all $n+1$ qubits, measure the last qubit and depending on the outcome bit, measure the remaining n qubits.

Problem 3. Show that $O(1)$ quantum example suffices to learn parities. Show that $O(n)$ classical examples suffice for learning parities.

Hint: Fourier sampling and Gaussian elimination.

Problem 4. In this exercise you will be showing that Learning parities with noise (LPN) on n bits is easy with quantum samples. Classically the best known algorithm given classical samples takes time $2^{O(n/\log n)}$ but we will see how it can be solved in quantum polynomial time.

In the LPN problem, a learner is given uniformly random $x \in \{0,1\}^n$ and $\langle a, x\rangle + b_x$ where b_x are iid random variables that equal 1 with probability $(1-\eta)$ and 0 otherwise. Show that, polynomially many copies of

$$|\psi_a\rangle = \frac{1}{\sqrt{2^n}} \sum_x |x, \langle a, x\rangle + b_x\rangle$$

and polynomial time suffices to learn a

Hint:

1. Apply Hadamards on all $n+1$-qubits and show that measuring the last qubit of $|\psi_a\rangle$ equals 0 with prob. 1/2.
2. With probability exponentially close to 1, show that measuring the first n bits equals a using Chernoff bound (recall that the Chernoff bound states the following: let $\mathbf{X}$ be a bounded random variable in $[-1,1]$ with $\mu = \mathbb{E}[\mathbf{X}]$, suppose we are given t independent samples $x_1, \ldots, x_t$, then we have that

$$\Pr[|\frac{1}{t}\sum_{i=1}^{t} x_i - \mu| \geq k] \leq \exp(-k^2 t)$$

.
3. For every $c \neq a$, with exponentially tiny probability, measuring the first n bits equals c

Problem 5. Here you will see how to approximate Fourier coefficients using just classical examples.

Recall that for a function $f : \{0,1\}^n \to \{0,1\}$, the Fourier coefficients are defined as

$$\widehat{f}(S) = \mathbb{E}_x[f(x)\chi_S(x)],$$

where $\chi_S(x) = (-1)^{S \cdot x}$. Show that there exists an algorithm that satisfies the following: the algorithm obtains $O(1/\varepsilon^2 \cdot \log(1/\delta))$ labelled examples $(x, f(x))$ where x is uniformly random and with probability $\geq 1-\delta$, outputs α such that $|\alpha - \widehat{f}(S)| \leq \varepsilon$.

Hint: Use Chernoff bound.

Problem 6. Consider the classical agnostic learning setup as follows: let $D : \{0,1\}^{n+1} \to [0,1]$ be an unknown distribution such that the marginal on the first n bits is uniform and the probability the last bit is 1 is $(1 + g(x))/2$, and it equals 0 with probability $(1 - g(x))/2$. The goal is to find S such that

$$\mathsf{err}_D(\chi_S) \leq \mathsf{OPT} + \varepsilon, \tag{1}$$

where $\mathsf{err}_D(\chi_S) = \Pr_{(x,b)\sim D}[\chi_S(x) \neq b]$, $\mathsf{OPT} = \min_T\{\mathsf{err}_D(\chi_T)\}$ and $\chi_S(x) = (-1)^{S\cdot x}$.

In quantum agnostic learning we are given

$$|\psi_g\rangle = \frac{1}{\sqrt{2^n}} \sum_x |x\rangle \otimes \Big(\sqrt{\frac{1+g(x)}{2}}|1\rangle + \sqrt{\frac{1-g(x)}{2}}|0\rangle\Big)$$

for an *arbitrary* $g : \{0,1\}^n \to [-1,1]$. Show that a quantum agnostic learner satisfying Eq. (1) satisfies

$$\widehat{g}(S) \in [\max_T \widehat{g}(T) - \varepsilon, \max_T \widehat{g}(T) + \varepsilon].$$