Problem Set 2

- Due Date: **18 Oct, 2021**
- The points for each problem is indicated on the side. This problem set has **100** points but you may solve any **70** points worth of questions among these for a full score (the remaining **30** points are bonus). Any additional points obtained will still count towards your final aggregate.
- Turn in your problem sets electronically (PDF; either LaTeXed or scanned etc.) on Acadly.
- Collaboration is encouraged, but all writeups must be done individually and must include names of all collaborators.
- Referring to sources other than the text book and class notes is strongly discouraged. But if you do use an external source (eg., other text books, lecture notes, or any material available online), ACKNOWLEDGE all your sources in your writeup, with a brief remark on *why* you sought that source. This **will not** affect your grades. However, not acknowledging will be treated as a serious case of academic dishonesty.
- Be clear in your writing.

1. **[The Affine Line graph]** (7 + 3)

   Let $\mathbb{F}$ be a finite field. Consider the following graph G whose vertex set is $\mathbb{F}^2$ and edges set E defined as

   $$E = \{((a,b),(c,d)) \ : \ ac = b + d\}\,.$$

   One way to interpret this is the point (a,b) is connected to all points (c,d) on the line $y = ax - b$.

   (a) Show that G is $|\mathbb{F}|$-regular and $\lambda(G) \leq \frac{1}{\sqrt{|\mathbb{F}|}}$.

   [Hint: It might be easier to understand G^2.]

   (b) Starting with this, and using the graph operations seen in class, show that you can construct a $(D^8, D, 1/8)$-spectral expander for some suitably large constant D.

2. **[Chernoff's bound for expander walks]** (3 + 3 + 3 + 6)

   In this problem, you will see a generic way to go from a *hitting-set tail* to a *Chernoff tail* due to Kabanets and Impagliazzo which we can instantiate for expander walks.

   Lemma 1. *Let $X_1, \ldots, X_t \in \{0,1\}$ be random variables (possibly correlated) such that for any subset $S \subseteq [t]$, we have*

   $$\Pr\left[\bigwedge_{i \in S} X_i = 1\right] \leq \mu^{|S|}.$$

   Then, for any $\varepsilon > 0$ with $\mu + \varepsilon < 1$, we also have

   $$\Pr\left[\frac{\sum X_i}{t} > \mu + \varepsilon\right] \leq e^{-\mathrm{KL}(\mu+\varepsilon \| \mu)\cdot t} \leq \exp(-\Omega(\varepsilon^2 t)).$$

Base version: (2021-10-08 19:10:36 +0530) , f704c93

The quantity $\mathrm{KL}(p\|q)$, called the Kullback-Leibler divergence, or relative entropy, equals $p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q}$. Standard Taylor arguments show that $\mathrm{KL}(\mu+\varepsilon\|\mu) = O(\varepsilon^2)$.

In this problem, you will prove the above lemma, and then instantiate it with expander random walks to get the upper-tail bound.

Let $p \in [0,1]$ be a parameter to be chosen shortly. Consider the following event where a set $S \subseteq [t]$ is chosen at random by adding each element $i \in [t]$ to S independently with probability p. Let M be the following expression:

$$M = \Pr_{S, X_1, \ldots, X_t} \left[\bigwedge_{i \in S} X_i = 1 \right]$$

(a) Show that $M \leq (p\mu + (1-p))^t$.

(b) Show that

$$M \geq \Pr\left[\sum X_i > (\mu+\varepsilon)\cdot t\right] \cdot (1-p)^{(1-\mu-\varepsilon)t}.$$

[Hint: What happens if you condition on the event that $\sum X_i > (\mu+\varepsilon)t$? Can you lower bound M under that condition?]

(c) By setting $p = \frac{\varepsilon}{(\mu+\varepsilon)(1-\mu)}$, argue that

$$\begin{aligned}\Pr\left[\sum X_i > (\mu+\varepsilon)\cdot t\right] &\leq \left(\left(\frac{\mu}{\mu+\varepsilon}\right)^{\mu+\varepsilon} \cdot \left(\frac{1-\mu}{1-\mu-\varepsilon}\right)^{1-\mu-\varepsilon}\right)^t \\ &= e^{-\mathrm{KL}(\mu+\varepsilon\|\mu)\cdot t}\end{aligned}$$

(d) Instantiate this for expander random walks to prove the following result:

> Let G be an (N, D, λ)-expander. Suppose $B \subseteq [N]$ with $\mu = \frac{|B|}{N}$. Let $\mu' = \mu(1-\lambda)+\lambda$, and $\varepsilon > 0$ so that $\mu' + \varepsilon < 1$. If $v_1, \ldots, v_t$ is a random walk in G (pick v_1 uniformly at random, and keep choosing a uniformly random neighbour), then
>
> $$\Pr\left[\frac{|\{v_1, \ldots, v_t\} \cap B|}{|B|} > \mu' + \varepsilon\right] \leq \exp(-\Omega(\varepsilon^2 t)).$$

Although this problem worked with only indicator random variables (where $X_i = \mathbf{1}[v_i \in B]$), one can work with more general "weight" functions $f : V \to [-1,1]$ and expander random walks give a pretty good estimate for $\mathbb{E}_v[f(v)]$. For more details see Thm 4.22 in Vadhan's manuscript.

3. **[An optimal non-averaging sampler]** (4 + 4 + 12)

Suppose $f : \{0,1\}^m \to \{0,1\}$ is some function and $\mu = \mathbb{E}_x[f(x)]$. An (ε, δ)-*sampler* is a randomized algorithm that queries f at various points and outputs some estimate $\hat{\mu}$ with the property that

$$\Pr[|\hat{\mu} - \mu| > \varepsilon] \leq \delta.$$

We are primarily interested in two parameters of such samplers — how many queries did it make, and how many random bits did it use. For this entire problem, assume that we have a *super-explicit* $(2^m, d, 0.5)$-spectral expander for some constant d.

(a) Using expanders, show how one can obtain an (ε, δ)-sampler that makes at most $O\left(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}\right)$ queries and uses at most

$$m + O\left(\frac{\log(1/\varepsilon)}{\varepsilon^2} \cdot \log \frac{1}{\delta}\right) \text{ random bits.}$$

(You may assume Theorem 4.22 from Vadhan's manuscript, which is a stronger form of the previous question, for this problem. You may also assume that there are strongly explicit constant-degree expanders on 2^m vertices.)

(b) Suppose we have an $(\varepsilon, (1/8))$-sampler $\mathcal{S}$ that makes Q queries and uses R random bits. Consider the following "median of averages sampler" built from $\mathcal{S}$:

> Run the sampler $\mathcal{S}$ for t independent trials to obtain $\hat{\mu_1}, \ldots, \hat{\mu_t}$. Output the *median* of these estimates.

Prove that this new sampler will be an (ε, δ)-sampler if $t = O\left(\log \frac{1}{\delta}\right)$.

(c) Construct an (ε, δ)-sampler that makes at most $O\left(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}\right)$ queries and uses at most

$$O\left(m + \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}\right) \text{ random bits.}$$

[Hint: Recall the sampler using pairwise independence. Can we work with median of *correlated* estimates in (2)?]

You have now seen a sampler that has *optimal* number of queries and random bits used (up to constants) but is **NOT** an averaging sampler! Obtaining an averaging sampler with the same performance is an open problem.

4. **[Spectral gap of general regular graphs]** $(4 + 1 + 2 + 4 + 4 + 3 + 2)$

In this problem we will show that $\lambda(G) \leq 1 - \frac{1}{\text{poly}(n,d)}$ for any d-regular n-vertex non-bipartite graph. In the process also learn about a very useful object called the Laplacian of a graph.

For an n-vertex d-regular undirected graph, define the *Laplacian* of the graph G (denoted by L_G) as

$$L_G = d \cdot I - A_G$$

where A_G is the adjacency matrix of G.

(More commonly, the version studied is the *normalised Laplacian* given by $I - M$ where M is the random walk matrix, since this also makes sense in the non-regular case. However, for this question, it would be more convenient to work with the above form.)

For a symmetric matrix M, we shall write $M \succeq 0$ to mean that $x^T M x \geq 0$ for all $x \in \mathbb{R}^n$. This is equivalent to stating that all eigenvalues of M are non-negative, and such matrices are also called positive semi-definite matrices (PSD) matrices.

We will extend this to a partial order between matrices to say that $A \succeq B$ if and only if $A - B \succeq 0$.

(a) Show that for any $x \in \mathbb{R}^n$, we have

$$x^T L_G x = \sum_{(i,j) \in G} (x_i - x_j)^2.$$

Hence, in some sense, the quadratic form corresponding to L_G measures the total "variation" of x across edges. Sometimes this is also called the *energy* of x.

As a corollary, observe that $L_G \succeq 0$.

(b) If H is a subgraph of G, show that $L_G \succeq L_H$.

(c) If K_n is the complete graph on n-vertices (without self-loops), what are the eigenvalues of L_{K_n}?

(d) If P_n is a path graph consisting of edges $\{(1,2),(2,3),\ldots,(n-1,n)\}$ and $E_{1,n}$ is the graph with a single edge $(1,n)$, show that

$$L_{P_n} \succeq \frac{1}{n-1} \cdot L_{E_{1,n}}.$$

(e) For a connected graph G on n-vertices, show that

$$L_G \succeq \frac{1}{(n-1)\cdot\binom{n}{2}} L_{K_n}$$

[Hint: For each pair of vertices (i,j), take the path $P_{i,j}$ in the graph G and sum up the previous subdivision over all such pairs $(i,j) \in \binom{[n]}{2}$.]

(f) Let G be an n-vertex d-regular, all of whose eigenvalues are non-negative. Show that $\lambda(G) \leq 1 - \Omega\left(\frac{1}{dn^2}\right)$.

(g) Let G be an n-vertex d-regular non-bipartite graph. Show that

$$\lambda(G) \leq 1 - \Omega\left(\frac{1}{d^2n^2}\right)$$

[Hint: Consider $G' = G^2$.]

One of the exercises in Vadhan's notes is a far shorter way to get the above spectral gap bound, but we felt that perhaps the above route shines more light on some of the steps.

5. **[Hitting Set Lemma via Expander Mixing Lemma]** (4 + 4 + 7)

In this problem, we will try to give an alternate proof of the *hitting set lemma* for random walks on expanders via the expander mixing lemma. This alternate proof is due to Silas Richelson and Sourya Roy

Recall the Expander Mixing Lemma. Let $G = (V,E)$ be an undirected N-vertex D-regular graph[1] with spectral gap $\gamma = 1 - \lambda$ and stationary distribution π. Let $f, g \colon V \to \mathbb{R}$ be any two functions. Then

$$\left| \mathop{\mathbb{E}}_{\{u,v\}\sim E} [f(u)\cdot g(v)] - \mu(f)\cdot\mu(g) \right| \leq \lambda \cdot \sigma(f) \cdot \sigma(g), \tag{EML}$$

where $\mu(f)$ and $\sigma(f)$ are the mean and standard-deviation respectively of the function f with respect to the distribution π (i.e., $\mu(f) := \mathbb{E}_{v\sim\pi}[f(v)]$ and $\sigma^2(f) := \mathbb{E}_{v\sim\pi}[f^2(v)] - \mu^2(f)$).

Let $B \subseteq V$ such that $\pi(B) = \mu$. For any positive integer t, define the function $g_t \colon V \to \mathbb{R}$ as follows:

$$g_t(v) := \mathop{\mathbb{E}}_{v=V_0,V_1,V_2,\cdots,V_t} [\forall i \ , \ V_i \in B],$$

where $V_0, V_1, V_2, \cdots, V_t$ is a t-step random walk starting at the fixed vertex $V_0 = v$. In other words, $g_t(v)$ is the probability that a t-step random walk starting at v lies completely in the set B.

Let $\mu_t := \mu(g_t)$ and $\sigma_t := \sigma(g_t)$.

(a) Prove that $\sigma_t^2 + \mu_t^2 = \mathbb{E}_{v\sim\pi}[g_t^2(v)] = \mu_{2t}$.

[1] The result holds for the more general setting of *reversible Markov chains*, but let's only deal with undirected regular graphs here.

[Hint: Does the distribution of choosing two length t random walks starting from a vertex $v \sim \pi$ look familiar?]

(b) Let r, s, t be non-negative integers such that $r + s = t$. Prove that

$$\mu_{t+1} = \mathop{\mathbb{E}}_{\{u,w\}\sim E} [g_r(u) \cdot g_s(w)] .$$

[Hint: Does the distribution of choosing a uniformly random edge $\{u, v\}$ and then choosing two walks of length r and s starting at u and v respectively look familiar?]

(c) Conclude that

$$\mu_{t+1} \leq \sqrt{(1-\lambda) \cdot \mu_r^2 + \lambda \cdot \mu_{2r}} \cdot \sqrt{(1-\lambda) \cdot \mu_s^2 + \lambda \cdot \mu_{2s}}.$$

[Hint: The expander mixing lemma and Cauchy-Schwarz inequality might come useful here.]

(With a bit more work, you can prove the Hitting Set Lemma from the above statement by inducting on t. The inductive step for the case of odd t is easy but the case of even t requires a bit more work.)

6. **[Cayley graphs and epsilon-biased sets]** (3 + 4 + 4 + 4 + 5)

In this problem, we will show how to generate graphs from a finite group and understand their spectrum.

Let G be a finite group, not necessarily abelian. Given such a group G and a set S closed under inverses (i.e., $s \in S \iff s^{-1} \in S$), the *Cayley graph* of G and S, denoted by $C(G, S)$, is the (undirected) graph (V, E) defined as follows:

$$V := G,$$
$$E := \{\{g, gs\} \mid g \in G, s \in S\} .$$

A subset $S \subseteq G$ is said to be a generating set of the group G (which is not necessarily abelian), if every $g \in G$ can be generated using elements from S. In other words, $g = s_1 \cdot s_2 \cdots s_k$ for some non-negative integer k and (not necessarily distinct) $s_1, s_2, \cdots, s_k \in S$.

(a) Show that if S is a generating set of G (closed under inverses), the corresponding Cayley graph $C(G, S)$ is connected.

A function $\chi \colon G \to \mathbb{C}$ is said to be a character of G if χ preserves all group operations (also called a *homomorphism)*, i.e. for all $g, h \in G$, we have $\chi(gh) = \chi(g) \cdot \chi(h)$, and $\chi(g^{-1}) = (\chi(g))^{-1}$. The character that maps every group element to 1 is called the trivial (or empty) character.

(b) i. Let G be the group $\{0,1\}^n$ with the binary operation addition. For each $\alpha \in \{0,1\}^n$, define the parity function $\chi_\alpha \colon G \to \mathbb{R}$ as follows:

$$\chi_\alpha(x) := (-1)^{\sum_i \alpha_i \cdot x_i}.$$

Show that the χ_α's are characters of G and χ_{0^n} is the trivial character.

ii. Let G be the group $\mathbb{Z}/m\mathbb{Z} = \{0, 1, 2, , \ldots, m-1\}$ with the binary operation addition (modulo m). Let $\omega_m := e^{2\pi i/m}$ (i.e., a primitive m-th root of unity). For each $h \in G$, define $\chi_h \colon G \to \mathbb{R}$ as follows:

$$\chi_h(g) := \omega_m^{hg}.$$

Show that the χ_h's are characters of G and χ_0 is the trivial character.

(c) i. Show that for any non-trivial character χ of a group, we have $\mathbb{E}_{g\in G}[\chi(g)] = 0$.

ii. Show that if χ_1 and χ_2 are characters so is $\chi_1 \cdot \overline{\chi_2}$. (Here $(\chi_1 \cdot \overline{\chi_2})(g) := \chi_1(g) \cdot \overline{\chi_2(g)}$ and $\overline{a}$ denotes the complex conjugate of a).

iii. Show that $\mathbb{E}_{g\in G}[\chi_1(g) \cdot \overline{\chi_2(g)}] = \mathbb{1}[\chi_1 = \chi_2]$.

Hence, the set of characters are pairwise orthogonal under the inner product $\langle f, h\rangle := \mathbb{E}_{g\in G}[f(g) \cdot \overline{h(g)}]$.

It is known that for any finite *abelian* group G, the set $\hat{G}$ of all characters are in 1-1 correspondence with G. Combining this with the above part, we get that the set $\hat{G}$ of characters of an abelian group form an orthonormal basis for the $|G|$-dimensional $\mathbb{C}$-vector space consisting of all functions $f\colon G \to \mathbb{C}$.

We will now study the random-walk matrix $M_{G,S}$ corresponding to the Cayley graph $C(G, S)$.

(d) Show that the characters of G are the (right) eigenvectors of the random-walk matrix $M_{G,S}$ for every S. In other words, show that $M_{G,S} \cdot \chi = \lambda_{S,\chi} \cdot \chi$ for some $\lambda_{S,\chi} \in \mathbb{R}$. What are the corresponding eigenvalues $\lambda_{S,\chi}$?

We now extend the definition of ε-biased sets to arbitrary groups. Let $\varepsilon \in (0, 1)$ A set $S \subset G$ is said to be an ε-biased set of G if for every non-trivial character χ of G, we have

$$\frac{1}{|S|} \cdot \left|\sum_{s\in S} \chi(s)\right| \leq \varepsilon.$$

(Note: There might be a discrepancy of a factor of 2, like there was in problem set 1, depending on what you think of as the definition of an ε-biased set.)

(e) Let G be an abelian group. Show that S is an ε-biased set of G if and only if the corresponding Cayley graph $C(G, S)$ is an $|S|$-regular $(1 - \varepsilon)$-spectral expander (i.e. the spectral gap is $1 - \varepsilon$).

This suggests a natural way to construct an expander graph with spectral gap $1 - \varepsilon$. Choose an ε-biased set of G and construct the Cayley graph $C(G, S)$. In particular, if we instantiate this scheme with the AGHP construction of ε-biased sets for the group $\{0, 1\}^n$, we get an n^2/ε^2-regular $(1 - \varepsilon)$-spectral expander on 2^n vertices. Unfortunately, this is not a constant-degree graph; the degree is logarithmic. The lower bound from Problem 7 in Problem Set 1 says that this is unavoidable. On the contrary, if one works with *non-abelian* groups, we can in fact construct constant degree Cayley graphs that are also "good" expanders.
