Pseudorandomness

Problem Set 2

Due date: September 22^{nd}, 2018

Instructions

1. The problem set has **4 questions** with a total score of **100 points**.
2. You are welcome to collaborate with other classmates. But if you do, please mention who all you collaborated with.

   I'd suggest you discuss only after you have spent enough time thinking about the problems independently.
3. Solutions are expected as a LaTeX document. You may use this very file by obtaining the source files from course webpage.
4. The deadline is **22nd September 2018, 2359 hrs IST**. For each day of delay, you lose **7 points** of your total score in this assignment. So if you plan to delay, be smart about it.

Questions

Question 1. (Spectrum of a tree) Let T_d be the undirected infinite d-ary tree. This, in some sense, is the *best* d-regular vertex expander. In this problem we will try to compute its spectrum. Recall that for a finite graph d-regular undirected graph G, we defined $\lambda(G)$ as

The factor of two is because each edge is counted twice in $x^T Ax$.

$$\lambda(G) = \max_{x \perp u} \left| \frac{x^T Ax}{x^T x} \right| = \left(\frac{1}{d}\right) \max_{x \perp u} \left| \frac{2 \sum_{(i,j) \in E} x_i x_j}{\sum x_i^2} \right|$$

where A is the normalised adjacency matrix for G. For an infinite graph such as T_d, we shall define

$$\lambda(T_d) = \left(\frac{1}{d}\right) \sup_{\substack{x \in \mathbb{R}^\infty \\ \sum_i x_i^2 < \infty}} \left| \frac{2 \sum_{(i,j) \in E} x_i x_j}{\sum x_i^2} \right|.$$

(We do not need to worry about the vector **1** here since it does not have finite norm.)

1. **[10 points]** Prove that $\lambda(T_d) \leq \frac{2\sqrt{d-1}}{d}$.

   (Hint: You might want to use the fact that, for any $\alpha > 0$, we have $2\left|x_i x_j\right| \leq \left(\alpha x_i^2 + \frac{x_j^2}{\alpha}\right)$.)

2. **[10 points]** Prove that $\lambda(T_d) = \frac{2\sqrt{d-1}}{d}$ by showing that, for every $\epsilon > 0$, there is a vector vector $x \in \mathbb{R}^\infty$ of bounded norm such that

$$2\left|\sum_{(i,j)\in E} x_i x_j\right| \geq (2\sqrt{d-1} - \epsilon)\left(\sum_i x_i^2\right).$$

Question 2. (Problem 4.8 in Vadhan's manuscript) 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 $y = ax - b$.

1. **[10 points]** Show that G is $|\mathbb{F}|$-regular and $\lambda(G) \leq \frac{1}{\sqrt{|\mathbb{F}|}}$.

   (Hint: Might be easier to understand G^2.)

2. **[5 points]** 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.

Question 3. In this problem you would be deriving the Chernoff bounds for expanders.

1. **[5 points]** Let $G = (V, E)$ be an undirected graph d-regular and let A be its (unnormalised) adjacency matrix. For any function $f : V \to \mathbb{R}$ on the vertices, define the diagonal matrix D_f whose (v,v)-th entry is $f(v)$. When we say ρ is an s-t path of length k, we shall treat $\rho = (v_0, \ldots, v_k)$ such that $v_0 = s$, $v_k = t$ and $(v_i, v_{i+1}) \in E$ for each i. Suppose $W_{s,t}$ is defined as

$$W_{s,t} = \sum_{\substack{\rho \text{ is an } s \text{ to } t \\ \text{paths of length } k}} \prod_{i=0}^{k} f(\rho(i)).$$

   Show that $W_{s,t} = e_s^T (D_f A)^k D_f e_t$, where e_s and e_t are the standard basis vectors with a 1 at position s and t respectively (and zero everywhere else).

2. **[5 points]** For a function $f : V \to [0,1]$ and a positive integer k, consider the random variable X that takes the value $\sum_{i=0}^{k} f(v_k)$ where $(v_0, \ldots, v_k)$ is a uniformly random path of length k in G.

   Let $0 \le r \le 1$ be some parameter. Express $\mathbb{E}[e^{rX}]$ linear algebraically.

3. **[10 points]** If G is an (n, d, λ)-expander and $\mu = \mathbb{E}[f]$, show that

$$\mathbb{E}[e^{rX}] \le e^{(\mu+\lambda)rk + O(r^2 k)}.$$

   (Hint: You might want to use the matrix decomposition for the normalized adjacency matrix for G, and also the fact that $e^x \le 1 + x + O(x^2)$ when for $|x| \le 1$.)

4. **[5 points]** Complete the proof of the Chernoff's bound on a λ-expander.

$$\Pr\left[X - \mu k \ge (\lambda + \epsilon) \cdot k\right] \le e^{-\Omega(\epsilon^2 k)}.$$

Question 4. (Modified form of Problem 4.5 in Vadhan's manuscript) 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

$$\Pr[|\hat{\mu} - \mu| > \epsilon] \le \delta.$$

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 *strongly-explicit* $(2^m, d, 0.5)$-spectral expander for some constant d.

1. **[10 points]** Using expanders, show how one can obtain an (ϵ, δ)-sampler that makes at most $O\left(\frac{1}{\epsilon^2} \log \frac{1}{\delta}\right)$ queries and uses at most

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

2. **[10 points]** Suppose we have an $(\epsilon, (1/8))$-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 $t = O\left(\log \frac{1}{\delta}\right)$.

3. **[20 points]** Construct an (ϵ, δ)-sampler that makes at most $O\left(\frac{1}{\epsilon^2} \log \frac{1}{\delta}\right)$ queries and uses at most

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

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