CMPUT 675 - Winter 2019
Assignment #5 - Due April 19 (Friday) by 3:00pm

All **Exercises** have the same weight, regardless of their difficulty. You are allowed to skip **one** exercise freely with no penalty. If you answer more than what is required, I will drop the one with the lowest marks when computing your mark for this assignment.

It is highly recommended that you typeset your solutions in LaTeX. I still want hard copies of your solution, so submit a printout if you do typeset it. As always, if any question is not clear then please feel free to ask me for clarification.

You can slide it under my door (ATH 3-06) if I'm not available when you want to hand it in (email me if you do this). In extreme cases, I will accept electronic copies if you are not able to hand in a hard copy for some reason.

Pages: 4

Exercise 1: Expanders of Any Size

We saw how there are constants $D > d \geq 3$ and $\lambda < 1$ such that for any k we can construct a (D^k, d, λ)-expander in time that is polynomial in D^k (i.e. the size of the graph). Now describe how to construct an (n, d', λ')-expander for any $n \geq 1$ where $d' \geq 3$ and $\lambda' < 1$ are constant.

Hint: The main idea to constructing the (D^k, d, λ)-expander for some D^k that is close to n and then "merge" some vertices to get exactly n nodes. You can either use an linear-algebraic argument via Rayleigh quotients to show it's second-largest eigenvalue is bounded for some appropriate λ', or you can use the following result from the textbook without proof (read it if you get the chance, it is neat!).

Theorem (essentially Theorem 21.9 from the book)
Let $G = (V; E)$ be a d'-regular graph with at least one loop at each node such that

$$\min_{\substack{S \subseteq V \\ |S| \leq |V|/2}} \frac{|\delta(S)|}{|S|} \geq \rho \cdot d'$$

for some $\rho > 0$. Then G is an $(n, d', 1 - \epsilon)$-expander where $\epsilon = \min\{2/d', \rho^2/2\}$.

Exercise 2: Missing PCP Pieces - Part 1

The following asks you to fill in the missing pieces behind the second preparation step we discussed in the proof of the PCP theorem.

Let $d \geq 3, \lambda < 1$ be constants such that a family $\{G_n\}_{n \geq 1}$ of (d, λ)-expanders exists where G_n can be constructed in poly(n) time.

- Prove that we can, in fact, get a family $\{G'_n\}_{n\geq 1}$ of (d', λ')-expanders for some constants $d' \geq 3, \lambda' < 1$ such that for each n and each $S \subseteq V(G_n)$ with $|S| \leq n/2$ we have $|\delta(S)| \geq 2 \cdot |S|$. The graphs G'_n must also be constructible in poly(n) time.
  Hint: We have a result from the lectures that talks about the ratio $|\delta(S)|/|S|$ for the G_n graphs. There is a very elementary way to fix the deficiency.
- For any n, consider the graph $G'_n = (V_n, E_n)$ from the first part. Consider some $\sigma : V_n \to \{0, \ldots, W-1\}$ where $W \geq 2$ is an integer.

  Prove there is some $\sigma' : V_n \to \{0, \ldots, W-1\}$ that is **constant** (i.e. there is some $a \in \{0, \ldots, W-1\}$ such that $\sigma'(v) = a$ for all $v \in V_n$) such that the number of v with $\sigma'(v) \neq \sigma(v)$ is at most the number of edges $uv \in E$ with $\sigma(u) \neq \sigma(v)$.
- Finally, show the following. In the 2nd preparation step in the PCP theorem, whether the resulting CSP instance is satisfiable or not we have there is some assignment σ' such that σ' is the same across all copies of a particular variable (the copies being the nodes from the expander that replaced the original node/variable).

  Refer to the posted lecture notes for notation used in this step, you should use the same notation.

Exercise 3: Some Linear Algebra

This fills in some of the small results we glossed over in the construction of expanders and analysis of random walks in expanders.

First Part
Recall the spectral norm for symmetric matrices A over $\mathbb{R}^{n\times n}$ is defined by:

$$||A|| = \max_{\substack{x\in\mathbb{R}^n \\ ||x||_2=1}} ||A \cdot x||_2.$$

We also use notation $[n] = \{1, 2, \ldots, n\}$.

Prove the following where $A, B \in \mathbb{R}^{n\times n}$ are symmetric matrices and x is any vector in $\mathbb{R}^n$ (not necessarily a unit vector).

- $||A + B|| \leq ||A|| + ||B||$
- $||A \cdot x||_2 \leq ||A|| \cdot ||x||_2$
- $||A \cdot B|| \leq ||A|| \cdot ||B||$
- $||A|| = \max_\lambda |\lambda|$ where the maximum is over all eigenvalues of A.

This was used in the analysis of random walks in expanders and also in the analysis of the replacement product.

Second Part
Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$ be symmetric matrices. Recall $A \otimes B$ is the $(nm) \times (nm)$ matrix whose rows and columns are indexed by pairs $(i, j) \in [n] \times [m]$ where entry $((i, j), (i', j'))$ of $A \otimes B$ is $A_{i,i'} \cdot B_{j,j'}$. We used this construction in the proof of the eigenvalue bound for the replacement product.

Let $\lambda_1, \ldots, \lambda_n$ and $\gamma_1, \ldots, \gamma_m$ be the eigenvalues for A and B, respectively, listed with multiplicity. Prove that $\{\lambda_i \cdot \gamma_j\}_{(i,j) \in [n] \times [m]}$ are the eigenvalues of $A \otimes B$ (enumerated with appropriate multiplicity).

Hint: Consider the eigenvectors coming from A and B.

Conclude that if A is the random walk matrix of an (n, d, λ)-expander and B is the random walk matrix of an (m, d', γ)-expander, then $A \otimes B$ is the random walk matrix of an $(nm, dd', \max\{\lambda, \gamma\})$-expander (you just have to show the eigenvalue bound).

Third Part
Conclude $||A \otimes B|| = ||A|| \cdot ||B||$ for symmetric $A, B \in \mathbb{R}^{n \times n}$. We also used this in the analysis of the replacement product.

Exercise 4: Missing PCP Pieces - Part 2

Let $G = (V; E)$ be an (n, d, λ)-expander with random walk matrix A. Assume G has **no loops**. Let $F \subseteq E$. Consider the following random walk in G.

- Select a random edge $e \in F$ and let v_0 be a random endpoint of e.
- Inductively, let v_{i+1} be a random neighbour of v_i.

Prove for any $t \geq 0$ that $\mathbf{Pr}[v_t v_{t+1} \in F] \leq \frac{|F|}{|E|} + \lambda^t$.

This is a "guided exercise": the steps are broken down carefully for you and you fill in the arguments. You may proceed with any part assuming the previous parts are proven, even if you did not prove them. I will mark the parts independently.

1. For any $v \in V$, let x_v be the probability that $v_0 = v$ (i.e. the walk starts at v). Also, for any $w \in V$ let y_w be the number of edges of F having w as an endpoint, divided by d.

   Prove for any $t \geq 0$ that $\mathbf{Pr}[v_t v_{t+1} \in F] = \langle A^t x, y \rangle$.
2. Show $\langle A^t x, y \rangle = \frac{2|F|}{d} \langle A^t x, x \rangle \leq \frac{2|F|}{d} \cdot \left(\frac{1}{n} + \lambda^t \cdot ||x||_2^2\right)$.
3. Then demonstrate $||x||_2^2 \leq \max_{v \in V} x_v$.
4. Finally, explain why $x_v \leq \frac{d}{2|F|}$ for each $v \in V$.
5. Put things together to finish the proof that $\mathbf{Pr}[v_t v_{t+1} \in F] \leq \frac{|F|}{|E|} + \lambda^t$.

Brownie Points
The expanders we used in the proof of the PCP theorem may have loops. The appropriate thing to prove in that case is the following. Suppose we perform a random walk except the first vertex v_0 is chosen as follows. Let $F' := \{(e, v) : e \in F, v \text{ an endpoint of } e\}$, in particular if $e \in F$ is a loop then there is only one (e, v) entry in F' for the single endpoint v of e.

Sample an entry (e, v) uniformly from F' and let $v_0 = v$. Show $\mathbf{Pr}[v_t v_{t+1} \in F] \leq O(1) \cdot \left(\frac{|F|}{|E|} + \lambda^t\right)$ by appropriately adapting the steps above. This would suffice for the proof of the PCP theorem when using expanders with loops. Clearly indicate if you are trying this so I don't get confused when marking!

Exercise 5: PCPs with 2 Bits

We discussed how $\mathbf{PCP}(O(\log n), 2) = \mathbf{P}$, essentially because we can decide if an instance of 2SAT is satisfiable in polynomial time.

On the other hand, it is hard to determine the maximum number of clauses in a 2SAT instance that can be satisfied in general. That is, the more general language

$$L = \{(\phi, k) : \phi \text{ is a 2-CNF formula such that some assignment satisfies at least } k \text{ clauses.}\}$$

is **NP**-complete. Let us extend this idea to showing we can introduce a gap in how many clauses can be satisfied in a 2CSP_2 instance between the "yes" case and the "no" case.

Show there are constants $\gamma < \rho$ such that for any language $L \in \mathbf{NP}$ there is a polynomial-time computable reduction f from L to instances of 2CSP_2 such that for $x \in \{0,1\}^*$,

- If $x \in L$, then $\text{sat}(f(x)) \geq \rho$,
- If $x \notin L$, then $\text{sat}(f(x)) \leq \gamma$.

Note, this would mean we cannot approximate the value of $\textsc{sat}(\phi)$ for arbitrary 2CSP_2 instances better than γ/ρ unless $\mathbf{P} = \mathbf{NP}$.

Hint
Feel free to look up classic **NP**-hardness reductions for MAX-2SAT, MAX-CUT, or any other 2CSP_2 problem, of course cite any resource you use. See if you can adapt them to introduce such a "γ vs. ρ" gap if you start with an instance of E3SAT that has a constant gap between yes and no cases (as proven last assignment).

If, for some reason, you want to perform the reduction from a different problem that we did not prove at any point has such a gap, ask me first and I'll let you know if that is ok.

Comment
Essentially this means there are constants $\gamma < \rho$ such that $\mathbf{NP} = \mathbf{PCP}_{\rho,\gamma}(O(\log n)2)$ where the subscripts mean in the "yes" case there is a proof π that is accepted with probability $\geq \rho$ and in the "no" case no proof is accepted with probability more than γ. So 2-bit **PCP**s are possible, as long as we do not demand perfect completeness.