Homework 2 – Due Thursday, February 16 *before* 10am on Angel

Instructions

- Solutions written in LaTeX are strongly preferred, but you can upload any pdf files, including scanned hand-written solutions. Template latex files are on the course webpage.
- Collaboration is allowed and encouraged. However, each of you should think about a problem before discussing it with others and write up your solution independently. You may consult books and on-line sources to get information about well-known theorems, such as the Chernoff bound. **But you are not allowed to look up solutions directly in papers or any other sources.** And you *must* list all collaborators and sources!
- Correctness, clarity, and succinctness of the solution will determine your score.

Problems

1. (**Yao's Principle**) To show that every probabilistic algorithm needs $q = q(n, \epsilon)$ queries to test property $\mathcal{P}$ on inputs of length n with probability of error $\leq \frac{1}{3}$ (over the algorithm's coin tosses), it is enough to show that there is a distribution $\mathcal{D}$ on the inputs of length n such that every deterministic algorithm needs q queries to test $\mathcal{P}$ over inputs from $\mathcal{D}$ with probability of error $\leq \frac{1}{3}$ (over the choice of the input).

   Prove this statement for 2-sided error adaptive algorithms. Notice (but do not hand in) that the same reasoning applies to 1-sided error and/or non-adaptive algorithms.

2. (**Lipschitz functions with a small range**) Recall that a function $f : \{0,1\}^d \to R$ is called *Lipschitz* if for d-bit strings x and y that differ in one bit, $|f(x) - f(y)| \leq 1$.
   (a) Give an ϵ-tester for the Lipschitz property of functions of the form $f : \{0,1\}^d \to \{0,1,2\}$. Your tester should run in time $O\left(\frac{d}{\epsilon}\right)$.
   (b) Prove by giving a reduction from a communication complexity problem that every $\frac{1}{4}$-tester for the Lipschitz property of functions of the form $f : \{0,1\}^d \to \{0,1,2\}$ must make $\Omega(d)$ queries.

3. This is an exercise in using Chernoff Bounds.
   (a) (**Amplification of the success probability**) Given an algorithm A that produces a good approximation (additive or multiplicative) with probability at least $\frac{2}{3}$ and runs in time $T(n)$ on inputs of size n, convert it to an algorithm that produces a good approximation (in the same sense as A) with error probability $\leq \delta$ and runs in time $O(T(n) \log \frac{1}{\delta})$. (Hint: run the algorithm $\Theta(\log \frac{1}{\delta})$ times and output the median answer.) Note that for algorithms with 0/1 output (e.g., property testers), taking the median corresponds to taking the majority.

(b) (**Witness Lemma extension**) Suppose that an α fraction of vertices in a graph are "witnesses". Prove that a sample of $s = \max\left\{\frac{2m}{\alpha}, \frac{8\ln(1/\delta)}{\alpha}\right\}$ vertices (selected uniformly and independently) will have at least m "witnesses" with probability at least $1-\delta$.

4. This is a collection of questions with short answers on testers in the adjacency matrix model.

   (a) Based on our analysis in class, give a new ϵ-test for bipartiteness with query and time complexity $O(\frac{\log^2 1/\epsilon}{\epsilon^3})$ instead of $O(\frac{\log^2 1/\epsilon}{\epsilon^4})$. Provide a one-line explanation.

   (b) In class we saw an ϵ-additive approximation algorithm for the edge density of the max cut with running time exponential in $\frac{1}{\epsilon}$. Explain why it is unlikely that there is such an algorithm with $poly(\frac{1}{\epsilon})$ running time that works for all ϵ. (Hint: it would imply $\mathrm{NP} \subseteq \mathrm{BPP}$.)

   (c) Improve the running time of the probabilistic algorithm (given in class) that finds a cut with density at least $\mu(G) - \epsilon$ from $2^{poly\frac{1}{\epsilon}} n^2$ to $2^{poly\frac{1}{\epsilon}} + O(n \cdot poly\frac{1}{\epsilon})$. Explain the new algorithm and give a one-line justification. Hint: we proved all the necessary lemmas in class.