Your Name	October 1, 2024
6.5240 Problem Set 2	
Collaborators: *Firstname Lastname*	Due: **10pm**, October 2, 2024

1. **Sublinear testing of connectedness.** Given a graph G of max degree d, and a parameter ϵ, give an algorithm which has the following behavior: if G is connected, then the algorithm should pass with probability 1, and if G is ϵ-far from connected (at least ϵdn edges must be added to connect G), then the algorithm should fail with probability at least 3/4. Your algorithm should look at a number of edges that is independent of n, and polynomial in d, $1/\epsilon$. For extra credit, try to make your algorithm as efficient as possible in terms of $n, d, 1/\epsilon$.

   For this homework set, when proving the correctness of your algorithm, it is acceptable to show that if the input graph G is likely to be passed, then it is ϵ-close to a graph G_0 which is connected, without requiring that G_0 has degree at most d.

2. **MST with non-integral weights.** In class we gave an MST approximation algorithm for graphs in which the weights on each edge were integers in the set $\{1, \ldots, w\}$. Show that one can get an approximation algorithm when the weights can be any value in the range $[1, w]$ (it is ok to get a slightly worse running time in terms of w, $1/\varepsilon$).

3. **Sublinear test of diameter.** The diameter of an unweighted graph is the maximum distance between any pair of nodes. Give an algorithm for graphs with degree at most d (where d is a constant and the graph is represented in the adjacency list model) that have low diameter.

   The algorithm should have the following specific behavior:

   (a) Graphs with diameter at most D are always accepted.

   (b) Graphs which are ϵ-far (that is, at least ϵdn edges must be added) from having diameter $4D + 2$ are failed with probability at least 2/3.

   (c) The query complexity of the algorithm should be $O(1/\epsilon^c)$ for some constant $c \geq 1$.

   For this homework set, when proving the correctness of your algorithm, it is acceptable to show that if the input graph G is likely to be passed, then it is ϵ-close to a graph G_0 which has diameter $4D + 2$, without requiring that G_0 has degree at most d.

4. **A simple algorithm for $\Delta + 1$ coloring.** Consider the following algorithm:

Randomly ordered coloring

```
  Require: Adjacency matrix access to unknown graph G, upper bound Δ on the degree
1 Let C_1, ..., C_{Δ+1} be sets initialized to ∅
2 Choose a uniformly random ordering of the vertices
3 for each vertex v in this order
4     Sample a color c ∈ [Δ + 1] uniformly at random
5     For each vertex in u ∈ C_c, i.e. colored c so far, check if u is a neighbor of v
6     If none of C_c is neighbor of v, add v to C_c; otherwise, go back to line 4
```

In this problem we will prove that this algorithm always outputs a proper $\Delta + 1$ coloring of G and runs in expected time $O(n^2 \log n/\Delta)$, which is sublinear in the graph size ($\leq n\Delta$) when $\Delta = \omega(\sqrt{n \log n})$.

a) Fix a vertex v and an ordering of the vertices, and define $d^<(v)$ as the number of neighbors of v that appear before v in the ordering. Define the random variable X_v to be the number of queries made by the algorithm while attempting to color v. Show that

$$\mathrm{E}\left[X_v\right] \leq \frac{n}{\Delta + 1 - d^<(v)}.$$

b) We will now randomize over the ordering, which we will denote by $\pi : V \to [n]$. This makes $d^<(v)$ a random variable that depends on π, which we will denote $d^<(v)_\pi$. Let v be a vertex with degree exactly Δ. Show that

$$\mathrm{E}_\pi\left[\frac{n}{\Delta + 1 - d^<_\pi(v)}\right] \leq \frac{n}{\Delta + 1} \cdot (\ln(\Delta + 1) + 1).$$

c) *Optional:* Show that (b) also holds for a vertex v with degree $< \Delta$.

d) Assuming (b) and (c), bound the expected query complexity of the algorithm by $O(n^2 \log n/\Delta)$.