**601.438 / 601.638 Algorithmic Foundations of Differential Privacy Spring 2025**
**Homework #3** **Due:** April 3, 2025, 9:00am

---

Reminder: you may work in groups of up to three people, but must write up solutions entirely on your own. Collaboration is limited to discussing the problems – you may not look at, compare, reuse, etc. any text from anyone else in the class. Please include your list of collaborators on the first page of your submission. Many of these problems have solutions which can be found on the internet – please don't look. You can of course use the internet (including the links provided on the course webpage) as a learning tool, but don't go looking for solutions.

Please include proofs with all of your answers, unless stated otherwise.

---

## 1 Stability of Densest Subgraph

Let $G = (V, E)$ be an undirected, unweighted graph. For $S \subseteq V$, let $E(S) = \{\{u, v\} \in E : u, v \in S\}$ be the edges in the subgraph induced by $S$, and let $\rho(S) = |E(S)|/|S|$ be the *density* of $S$. Slightly abusing notation, let $\rho(G) = \max_{S \subset V} \rho(S)$ be the density of the densest subgraph of $G$.

We will design an $\epsilon$-DP algorithm to compute $\rho(G)$ in the edge-DP model, i.e., where two graph $G$ and $G'$ are adjacent if they differ in one edge.

(a) What is the global sensitivity of $\rho(G)$? Using this, show how to use the Laplace mechanism to get an $\epsilon$-DP algorithm which computes a value $\hat{\rho}(G)$ with the property that $\mathbf{E}[|\rho(G) - \hat{\rho}(G)|] \leq O(1/\epsilon)$. Prove privacy and accuracy.

Now we'll want to do better by using stability, i.e., a modification of propose-test-release. The intuition is that if the densest subgraph is small then we can simply "round" the solution up (similar to stable histograms, where we rounded down), and if the densest subgraph is large then $\rho(G)$ is more stable (has lower sensitivity)

(b) For $x \in \mathbb{R}^+$, let $\rho_x(G) = \max(x, \rho(G))$. What is the global sensitivity of $\rho_x(G)$, as a function of $x$? Prove your bound.

(c) Use the previous part to design an $\epsilon$-DP algorithm which computes a value $\hat{\rho}(G)$ with the property that $\mathbf{E}[|\rho(G) - \hat{\rho}(G)|] \leq O(\sqrt{1/\epsilon})$. Prove privacy and accuracy.

## 2 Binary Tree as Factorization

We mentioned in class that the binary tree mechanism for interval queries can be viewed as a special case of the factorization mechanism. Let's show that here.

(a) Let's do a case with explicit small numbers first. Suppose that $X = [8]$, so all elements of the database are elements of $[8]$ and an interval query is of the form

$$f_{s,t}(D) = \# \text{ of elements in database in } [s, t]$$

where $1 \leq s \leq t \leq 8$. Recall that a threshold query is of the form $f_{1,t}$ for some value $t \in X$. For the set of threshold queries, write the matrix $F$ and its factorization into the product of matrices $RM$ corresponding to the binary tree mechanism.

(b) Do the same for the set of interval queries (again for $X = [8]$).

Now generalize to a $X = 2^\ell$ for some integer $\ell$. Specifically:

(c) Show the factorization $F = RM$ corresponding to the binary tree mechanism for threshold queries.

(d) Show the factorization $F = RM$ corresponding to the binary tree mechanism for interval queries.