General instructions. The exam has 2 parts (plus cheat-sheet). You have 2.5 hours. No books, notes, cell phones, or other external materials are allowed.

Part I (52 points)

Answer 4 of the following 5 questions. Give short answers, sketching the proof or giving a convincing justification in 2-5 sentences (even for true/false questions). You may use without proof theorems stated in class, provided you state the appropriate theorem that you are using. As usual, assume $n$ (or $|V|$) is large enough.

A. Markov’s inequality states that for every non-negative random variable $X$,

$$\forall t > 0, \quad \Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$ 

Does it hold even if $X$ is not restricted to be non-negative?

B. Let $G$ be a graph drawn from the distribution $G_{n,p}$ for $p = 8/n$.

Is it true that $\Pr[G$ is connected$] \geq 1/2$?

C. Fix an $n \times n$ matrix $A$ with 0-1 entries that has full rank, let $x$ be chosen uniformly at random from $\{0, 1\}^n$, and set $y = Ax$, where all the operations (the rank computation and the product $Ax$) are over GF[2].

Is it true that $y_1, \ldots, y_n$ (the $n$ coordinates of $y$) are fully independent bits?

D. Let $q$ and $x_1, \ldots, x_{n^2}$ be all random vectors in $\{0, 1\}^n$.

Is it true that with probability 90% or more, $q$ has a unique 1.1-approximate nearest neighbor among $x_1, \ldots, x_{n^2}$ (under Hamming distance)?

E. Let $G = (V, E, w)$ be an undirected graph with edge weights $w : E \rightarrow \mathbb{R}_+$, and let $G' = (V, E', w')$ be a $(1 + \varepsilon)$-cut-sparsifier of $G$ for $\varepsilon \in (0, 1)$.

Is it true that for every partition $V = V_1 \cup \cdots \cup V_k$, the total weight of edges connecting different $V_i$’s is the same in $G'$ as in $G$ up to factor $1 \pm \varepsilon$, formally, $\sum_{i<j} w'(V_i, V_j) \in (1 \pm \varepsilon) \sum_{i<j} w(V_i, V_j)$?
Part II (48 points)

Answer 2 of the following 3 questions.

1. Suppose Alice’s input is $x \in \{0, 1\}^n$ and Bob’s input is $y \in \{0, 1\}^n$, and the goal is to determine whether $x = y$. Design a non-trivial protocol where each of them sends a short message to a Referee which outputs an answer, assuming each party has private randomness, but no shared randomness.

   Hint: You may use the fact that there are “good” error correcting codes $C : \{0, 1\}^n \to \{0, 1\}^m$, which means that $m = O(n)$ and for all $x_1 \neq x_2 \in \{0, 1\}^n$ the Hamming distance between $C(x_1)$ and $C(x_2)$ is $\Omega(n)$.

2. Suppose the inputs of Alice and Bob are sets $E_A$ and $E_B$, respectively, of undirected edges on the same vertex set $V = [n]$. It is guaranteed that both $|E_A \setminus E_B|$ and $|E_B \setminus E_A|$ are at most $k := n^{1/3}$.

   Design a randomized protocol where each of them sends a short message to a Referee, whose goal is to output the precise symmetric difference $E_A \Delta E_B$ (e.g., not just most of the edges in this set). Assume the parties have access to shared randomness.

   Analyze the message-size and success probability of your protocol.

3. Let $G = (V, E)$ be an undirected graph on $n$ vertices, and denote the maximum hitting time in $G$ by $H := \max\{h_{u,v} : u, v \in V\}$. Prove that with probability at least $3/4$, a random walk of length $O(H \log n)$ (starting at some fixed $s \in V$) visits all the vertices of the graph.

   Hint: “Break” the walk into phases of length $2H$.

Good Luck.

Cheat Sheet

Chebychev’s inequality. Let $X$ be a random variable with finite variance $\sigma^2 > 0$. Then

$$\forall t > 0, \quad \Pr[|X - \mathbb{E}[X]| \geq t \sigma] \leq \frac{1}{t^2}.$$ 

Chernoff-Hoeffding bound. Let $X = \sum_{i \in [n]} X_i$, where $X_i \in [0, 1]$ for $i \in [n]$ are independently distributed random variables. Then

$$\forall t > 0, \quad \Pr[|X - \mathbb{E}[X]| \geq t] \leq 2e^{-2t^2/n}.$$  

$$\forall 0 < \varepsilon \leq 1, \quad \Pr[X \leq (1 - \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2\mathbb{E}[X]/2}. $$  

$$\forall 0 < \varepsilon \leq 1, \quad \Pr[X \geq (1 + \varepsilon)\mathbb{E}[X]] \leq e^{-\varepsilon^2\mathbb{E}[X]/3}. $$  

$$\forall t \geq 2e\mathbb{E}[X], \quad \Pr[X \geq t] \leq 2^{-t}.$$ 

Azuma’s inequality. Let $X_0, X_1, \ldots, X_m$ be a Martingale such that $|X_{i+1} - X_i| \leq 1$ for all $0 \leq i < m$. Then

$$\forall t > 0, \quad \Pr[|X_m - X_0| \geq t\sqrt{m}] \leq 2e^{-t^2/2}. $$ 

THE END.