18.226 PROBLEM SET (FALL 2020)

Helpful tips:

- Only turn in problems marked $\boxed{\textsf{ps1}}$ and $\boxed{\textsf{ps1}\star}$ for problem set 1, etc. You are recommended to try the other problems for practice, but do not submit them.
- In a multipart problem, if a later part is marked for submission, it may be helpful to think about the earlier unassigned parts first.
- **Bonus problems**, marked by $\star$, are more challenging. A grade of A- may be attained by only solving the non-starred problems. To attain a grade of A or A+, you should solve a substantial number of starred problems. No hints will be given for bonus problems, e.g., during office hours.
- **Start each solution on a new page**, and try to **fit your solution within one page** for each unstarred problem/part (without abusing font/margins). The spirit of this policy is to encourage you to think first before you write. Distill your ideas, structure your arguments, and eliminate unnecessary steps. If necessary, some details of routine calculations may be skipped provided that you give precise statements and convincing explanations.
- This file will be updated as the term progresses. Please check back regularly. There will be an announcement whenever each problem set is complete.
- You are encouraged to include figures whenever they are helpful. Here are some recommended ways to produce figures in decreasing order of learning curve difficulty:
  (1) TikZ
  (2) IPE (which supports LaTeX), Powerpoint, or other drawing app
  (3) drawing on a tablet (e.g., Notability on iPad)
  (4) photo/scan (I recommend the Dropbox app on your phone, which has a nice scanning feature that produces clear monochrome scans)

Problems begin on the next page　　　　Last updated: December 7, 2020

A. Introduction and linearity of expectations

A1. Verify the following asymptotic calculations used in Ramsey number lower bounds:

(a) For each k, the largest n satisfying $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ has $n = \left(\frac{1}{e\sqrt{2}} + o(1)\right) k2^{k/2}$.

(b) For each k, the maximum value of $n - \binom{n}{k}2^{1-\binom{k}{2}}$ as n ranges over positive integers is $\left(\frac{1}{e} + o(1)\right) k2^{k/2}$.

(c) For each k, the largest n satisfying $e\left(\binom{k}{2}\binom{n}{k-2} + 1\right) 2^{1-\binom{k}{2}} < 1$ satisfies $n = \left(\frac{\sqrt{2}}{e} + o(1)\right) k2^{k/2}$.

A2. Prove that, if there is a real $p \in [0,1]$ such that

$$\binom{n}{k}p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} < 1$$

then the Ramsey number $R(k,t)$ satisfies $R(k,t) > n$. Using this show that

$$R(4,t) \geq c\left(\frac{t}{\log t}\right)^{3/2}$$

for some constant $c > 0$.

ps1 A3. Let G be a graph with n vertices and m edges. Prove that K_n can be written as a union of $O(n^2(\log n)/m)$ copies of G (not necessarily edge-disjoint).

A4. *Generalization of Sperner's theorem.* Let $\mathcal{F}$ be a collection of subset of $[n]$ that does not contain $k+1$ elements forming a chain: $A_1 \subsetneq \cdots \subsetneq A_{k+1}$. Prove that $\mathcal{F}$ is no larger than taking the union of the k levels of the boolean lattice closest to the middle layer.

ps1 A5. Let $A_1, \ldots, A_m$ be r-element sets and $B_1, \ldots, B_m$ be s-element sets. Suppose $A_i \cap B_i = \emptyset$ for each i, and for each $i \neq j$, either $A_i \cap B_j \neq \emptyset$ or $A_j \cap B_i \neq \emptyset$. Prove that $m \leq (r+s)^{r+s}/(r^r s^s)$.

ps1 A6. Let G be a graph on $n \geq 10$ vertices. Suppose that adding any new edge to G would create a new clique on 10 vertices. Prove that G has at least $8n - 36$ edges.

Hint in white:

ps1⋆ A7. Prove that for every positive integer r, there exists an integer K such that the following holds. Let S be a set of rk points evenly spaced on a circle. If we partition $S = S_1 \cup \cdots \cup S_r$ so that $|S_i| = k$ for each i, then, provided $k \geq K$, there exist r congruent triangles where the vertices of the i-th triangle lie in S_i, for each $1 \leq i \leq r$.

ps1⋆ A8. Prove that $[n]^d$ cannot be partitioned into fewer than 2^d sets each of the form $A_1 \times \cdots \times A_d$ where $A_i \subsetneq [n]$.

A9. Let $k \geq 4$ and H a k-uniform hypergraph with at most $4^{k-1}/3^k$ edges. Prove that there is a coloring of the vertices of H by four colors so that in every edge all four colors are represented.

ps1 A10. Prove that there is an absolute constant $C > 0$ so that for every $n \times n$ matrix with distinct real entries, one can permute its rows so that no column in the permuted matrix contains an increasing subsequence of length at least $C\sqrt{n}$. (A subsequence does not need to be selected from consecutive terms. For example, $(1,2,3)$ is an increasing subsequence of $(1,5,2,4,3)$.)

ps1 A11. Given a set $\mathcal{F}$ of subsets of $[n]$ and $A \subseteq [n]$, write $\mathcal{F}|_A := \{S \cap A : S \in \mathcal{F}\}$ (its *projection* onto A). Prove that for every n and k, there exists a set $\mathcal{F}$ of subsets of $[n]$ with $|\mathcal{F}| = O(k2^k \log n)$ such that for every k-element subset A of $[n]$, $\mathcal{F}|_A$ contains all 2^k subsets of A.

ps1⋆ A12. Show that in every non-2-colorable n-uniform hypergraph, one can find at least $\frac{n}{2}\binom{2n-1}{n}$ unordered pairs of edges that intersect in exactly one vertex.

A13. Let A be a subset of the unit sphere in $\mathbb{R}^3$ (centered at the origin) containing no pair of orthogonal points.

ps1 (a) Prove that A occupies at most $1/3$ of the sphere in terms of surface area.

ps1⋆ (b) Prove an upper bound smaller than $1/3$ (give your best bound).

ps1⋆ A14. Prove that every set of 10 points in the plane can be covered by a union of disjoint unit disks.

A15. Let $\boldsymbol{r} = (r_1, \dots, r_k)$ be a vector of nonzero integers whose sum is nonzero. Prove that there exists a real $c > 0$ (depending on $\boldsymbol{r}$ only) such that the following holds: for every finite set A of nonzero *reals*, there exists a subset $B \subseteq A$ with $|B| \geq c|A|$ such that there do not exist $b_1, \dots, b_k \in B$ with $r_1 b_1 + \cdots + r_k b_k = 0$.

ps1 A16. Prove that every set A of n nonzero integers contains two disjoint subsets B_1 and B_2, such that both B_1 and B_2 are sum-free, and $|B_1| + |B_2| > 2n/3$. Can you do it if A is a set of nonzero reals?

A17. Let $M(n)$ denote the maximum number of edges in a 3-uniform hypergraph on n vertices without a clique on 4 vertices.

(a) Prove that $M(n+1)/\binom{n+1}{3} \leq M(n)/\binom{n}{3}$ for all n, and conclude that $M(n)/\binom{n}{3}$ approaches some limit α as $n \to \infty$.

(This limit is called the *Turán density* of the hypergraph $K_4^{(3)}$, and its exact value is currently unknown and is a major open problem.)

(b) Prove that for every $\delta > 0$, there exists $\epsilon > 0$ and n_0 so that every 3-uniform hypergraph with $n \geq n_0$ vertices and at least $(\alpha + \delta)\binom{n}{3}$ edges must contain at least $\epsilon\binom{n}{4}$ copies of the clique on 4 vertices.

A18. Prove that every graph with n vertices and $m \geq n^{3/2}$ edges contains a pair of vertex-disjoint and isomorphic subgraphs (not necessarily induced) each with at least $cm^{2/3}$ edges, where $c > 0$ is a constant.

B. Alteration method

B1. Using the alteration method, prove the Ramsey number bound

$$R(4,k) \geq c(k/\log k)^2$$

for some constant $c > 0$.

B2. Prove that every 3-uniform hypergraph with n vertices and $m \geq n$ edges contains an independent set (i.e., a set of vertices containing no edges) of size at least $cn^{3/2}/\sqrt{m}$, where $c > 0$ is a constant.

ps2 B3. Prove that every k-uniform hypergraph with n vertices and m edges has a transversal (i.e., a set of vertices intersecting every edge) of size at most $n(\log k)/k + m/k$.

ps2 B4. *Zarankiewicz problem.* Prove that for every positive integer $k \geq 2$, there exists a constant $c > 0$ such that for every n, there exists an $n \times n$ matrix with $\{0, 1\}$ entries, with at least $cn^{2-2/(k+1)}$ 1's, such that there is no $k \times k$ submatrix consisting of all 1's.

ps2 B5. Fix k. Prove that there exists a constant $c_k > 1$ so that for every sufficiently large $n > n_0(k)$, there exists a collection $\mathcal{F}$ of at least c_k^n subsets of $[n]$ such that for every k distinct $F_1, \dots, F_k \in \mathcal{F}$, all 2^k intersections $\bigcap_{i=1}^k G_i$ are nonempty, where each G_i is either F_i or $[n] \setminus F_i$.

B6. *Acute sets in* $\mathbb{R}^n$. Prove that, for some constant $c > 0$, for every n, there exists a family of at least $c(2/\sqrt{3})^n$ subsets of $[n]$ containing no three distinct members A, B, C satisfying $A \cap B \subseteq C \subseteq A \cup B$.

Deduce that there exists a set of at least $c(2/\sqrt{3})^n$ points in $\mathbb{R}^n$ so that all angles determined by three points from the set are acute.

Remark. The current best lower and upper bounds for the maximum size of an "acute set" in $\mathbb{R}^n$ (i.e., spanning only acute angles) are $2^{n-1} + 1$ and $2^n - 1$ respectively.

ps2⋆ B7. *Covering complements of sparse graphs by cliques*

(a) Prove that every graph with n vertices and minimum degree $n - d$ can be written as a union of $O(d^2 \log n)$ cliques.

(b) Prove that every bipartite graph with n vertices on each side of the vertex bipartition and minimum degree $n - d$ can be written as a union of $O(d \log n)$ complete bipartite graphs (assume $d \geq 1$).

ps2⋆ B8. Let $G = (V, E)$ be a graph with n vertices and minimum degree $\delta \geq 2$. Prove that there exists $A \subseteq V$ with $|A| = O(n(\log \delta)/\delta)$ so that every vertex in $V \setminus A$ contains at least one neighbor in A and at least one neighbor not in A.

ps2⋆ B9. Prove that every graph G without isolated vertices has an induced subgraph H on at least $\alpha(G)/2$ vertices such that all vertices of H have odd degree. Here $\alpha(G)$ is the size of the largest independent set in G.

C. Second moment method

C1. Let X be a nonnegative real-valued random variable. Suppose $\mathbb{P}(X = 0) < 1$. Prove that

$$\mathbb{P}(X = 0) \leq \frac{\operatorname{Var} X}{\mathbb{E}[X^2]}.$$

ps2 C2. Let X be a random variable with mean μ and variance σ^2. Prove that for all $\lambda > 0$,

$$\mathbb{P}(X \geq \mu + \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

ps2 C3. Let $x, y \in \mathbb{R}^n$ be two unit vectors each orthogonal to the all-1 vector. For a permutation σ of $[n]$, write $x^\sigma \in \mathbb{R}^n$ for the vector whose i-th coordinate is $x_{\sigma(i)}$. Write $\langle \cdot, \cdot \rangle$ for dot product.

(a) Compute $\operatorname{Var} \langle x^\sigma, y \rangle$ where σ is a uniformly chosen permutation of $[n]$.

(b) Prove that $\max_\sigma \langle x^\sigma, y \rangle - \min_\sigma \langle x^\sigma, y \rangle \geq 2/\sqrt{n-1}$, and that equality can be achieved for even n.

ps2 C4. Show that, for each fixed positive integer k, there is a sequence p_n such that

$$\mathbb{P}(G(n, p_n) \text{ has a connected component with exactly } k \text{ vertices}) \to 1 \quad \text{as } n \to \infty.$$

Hint in white:

C5. Let $p_n = (\log n + c_n)/n$. Show that, as $n \to \infty$,

$$\mathbb{P}(G(n,p_n) \text{ has no isolated vertices}) \to \begin{cases} 0 & \text{if } c_n \to -\infty, \\ 1 & \text{if } c_n \to \infty. \end{cases}$$

ps3 C6. Prove that, with probability approaching 1 as $n \to \infty$, $G(n, n^{-1/2})$ has at least $cn^{3/2}$ edge-disjoint triangles, where $c > 0$ is some constant.

ps3 C7. *Simple nibble.* Prove that, with probability approaching 1 as $n \to \infty$,
(a) $G(n, n^{-2/3})$ has at least $n/100$ vertex-disjoint triangles.
(b) $G(n, Cn^{-2/3})$ has at least $0.33n$ vertex-disjoint triangles, for some constant C.
Hint in white:

ps3$\star$ C8. Is the threshold for the bipartiteness of a random graph coarse or sharp?
(You are not allowed to use any theorems that we did not prove in class/notes.)

ps3$\star$ C9. Prove that there is an absolute constant $C > 0$ so that the following holds. For every prime p and every $A \subseteq \mathbb{Z}/p\mathbb{Z}$ with $|A| = k$, there exists an integer x so that $\{xa : a \in A\}$ intersects every interval of length at least $Cp/\sqrt{k}$ in $\mathbb{Z}/p\mathbb{Z}$.

ps3$\star$ C10. Prove that there is a constant $c > 0$ so that every hyperplane containing the origin in $\mathbb{R}^n$ intersects at least c-fraction of the 2^n closed unit balls centered at $\{-1,1\}^n$.

ps3 C11. Let $v_1 = (x_1, y_1), \ldots, v_n = (x_n, y_n) \in \mathbb{Z}^2$ with $|x_i|, |y_i| \le 2^{n/2}/(100\sqrt{n})$ for all $i \in [n]$. Show that there are two disjoint sets $I, J \subseteq [n]$, not both empty, such that $\sum_{i \in I} v_i = \sum_{j \in J} v_j$.

D. Chernoff bound

D1. Let $X \sim \text{Binomial}(n,p)$. Prove that for $0 < q \le p < 1$,

$$\mathbb{P}(X \le nq) \le e^{-nH(q\|p)} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X \le nq) = -H(\|p)$$

and for $0 < p \le q < 1$,

$$\mathbb{P}(X \ge nq) \le e^{-nH(q\|p)} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X \ge nq) = -H(\|p),$$

where

$$H(q\|p) := q \log \frac{q}{p} + (1-q) \log \frac{1-q}{1-p}.$$

is known as the *relative entropy* or *Kullback-Leibler divergence*, in this case, between two Bernoulli distributions.

ps3 D2. Prove that with probability $1 - o(1)$ as $n \to \infty$, every bipartite subgraph of $G(n, 1/2)$ has at most $n^2/8 + 10n^{3/2}$ edges.

D3. Show that for every k there is some $m = O(k \log k)$ such that for sufficiently large $n > n_0(k)$, there exists a collection $\mathcal{F}$ of m subsets of $[nk]$ such that for every partition $[nk] = A_1 \cup \cdots \cup A_k$ with $|A_i| = n$ for each $i \in [k]$, there exists some $F \in \mathcal{F}$ with $|A_i \cap F| \in [0.4n, 0.6n]$ for each $i \in [k]$.

D4. (a) Prove that there is some constant $c > 1$ so that there exists $S \subset \{0,1\}^n$ with $|S| \ge c^n$ so that every pair of points in S differ in at least $n/4$ coordinates.

(b) Prove that there is some constant $c > 1$ so that the the unit sphere in $\mathbb{R}^n$ contains at least c^n points, where each pair of points is at distance at least 1 apart.

ps3⋆ D5. Prove that there exists a constant $c > 1$ such that for every n, there are at least c^n points in $\mathbb{R}^n$ so that the angle spanned by every three distinct points is at most $61°$.

ps3 D6. *Planted clique.* Give a deterministic polynomial-time algorithm solving the following problem so that it succeeds over the random input with probability approaching 1 as $n \to \infty$:
Input: an n-vertex unlabeled graph G created as the union of $G(n, 1/2)$ and a clique on vertex subset of size $t = \lfloor 100\sqrt{n \log n} \rfloor$
Output: a clique in G of size t

D7. *Weighing coins*

ps3 (a) Prove that if $k \le 1.99n/\log_2 n$ and n is sufficiently large, then for every $S_1, \dots, S_k \subseteq [n]$, there are two distinct subsets $X, Y \subseteq [n]$ such that $|X \cap S_i| = |Y \cap S_i|$ for all $i \in [k]$.

ps3⋆ (b) Show that there is some constant C such that (a) is false if 1.99 is replaced by C.
(Bonus competition: additional points to the student who correctly proves the smallest C in the class; if m students acheieve the same lowest C, then each is awarded $\min\{10, \lfloor 24/m \rfloor\}$ points additional to the 10 points for a correct solution. You are not allowed to cheat by looking up the solution.)

E. Lovász local lemma

ps4 E1. Show that it is possible to color the edges of K_n with at most $3\sqrt{n}$ colors so that there are no monochromatic triangles.

E2. Prove that it is possible to color the vertices of every k-uniform k-regular hypergraph using at most $k/\log k$ colors so that every color appears at most $O(\log k)$ times on each edge.

ps4 E3. Prove that there is some constant $c > 0$ so that given a graph and a set of k acceptable colors for each vertex such that every color is acceptable for at most ck neighbors of each vertex, there is always a proper coloring where every vertex is assigned one of its acceptable colors.

ps4⋆ E4. Prove that there is a constant $C > 0$ so that for every sufficiently small $\epsilon > 0$, one can choose exactly one point inside each grid square $[n, n+1) \times [m, m+1) \subset \mathbb{R}^2$, $m, n \in \mathbb{Z}$, so that every rectangle of dimensions ϵ by $(C/\epsilon)\log(1/\epsilon)$ in the plane (not necessarily axis-aligned) contains at least one chosen point.

E5. Prove that there exists k_0 and a red/blue coloring of $\mathbb{Z}$ without any monochromatic k-term arithmetic progressions with $k \ge k_0$ and common difference less than 1.99^k.

ps4 E6. Prove that, for every $\epsilon > 0$, there exist ℓ_0 and some $(a_1, a_2, \dots) \in \{0,1\}^{\mathbb{N}}$ such that for every $\ell > \ell_0$ and every $i > 1$, the vectors $(a_i, a_{i+1}, \dots, a_{i+\ell-1})$ and $(a_{i+\ell}, a_{i+\ell+1}, \dots, a_{i+2\ell-1})$ differ in at least $(\frac{1}{2} - \epsilon)\ell$ coordinates.

E7. A *periodic path* in a graph G with respect to a vertex coloring $f \colon V(G) \to [k]$ is a path $v_1 v_2 \dots v_{2\ell}$ for some positive integer ℓ with $f(v_i) = f(v_{i+\ell})$ for each $i \in [\ell]$ (reminder: no repeated vertices allowed in a path).
Prove that for every Δ, there exists k so that every graph with maximum degree at most Δ has a vertex-coloring using k colors with no periodic paths.

ps4 E8. Prove that every graph with maximum degree Δ can be edge-colored using $O(\Delta)$ colors so that every cycle contains at least three colors and no two adjacent edges have the same color.

ps4⋆ E9. Prove that for every Δ, there exists g so that every bipartite graph with maximum degree Δ and girth at least g can be properly edge-colored using $\Delta + 1$ colors so that every cycle contains at least three colors.

ps4⋆ E10. Prove that for every positive integer r, there exists C_r so that every graph with maximum degree Δ has a *proper* vertex coloring using at most $C_r\Delta^{1+1/r}$ colors so that every vertex has at most r neighbors of each color.

E11. Let $H = (V, E)$ be a hypergraph satisfying, for some $\lambda > 1/2$,

$$\sum_{f \in E : v \in f} \lambda^{|f|} \le \frac{1}{2} - \frac{1}{4\lambda} \qquad \text{for every } v \in V$$

(here $|f|$ is then number of vertices in the edge f). Prove that H is 2-colorable.

E12. *Vertex-disjoint cycles in digraphs.* (Recall that a directed graph is *k-regular* if all vertices have in-degree and out-degree both equal to k. Also, cycles cannot repeat vertices.)

ps4 (a) Prove that every k-regular directed graph has at least $ck/\log k$ vertex-disjoint directed cycles, where $c > 0$ is some constant.

ps4⋆ (b) Prove that every k-regular directed graph has at least ck vertex-disjoint directed cycles, where $c > 0$ is some constant.

Hint in white:

E13. (a) *Generalization of Cayley's formula.* Using Prüfer codes, prove the identity

$$x_1 x_2 \cdots x_n (x_1 + \cdots + x_n)^{n-2} = \sum_T x_1^{d_T(1)} x_2^{d_T(2)} \cdots x_n^{d_T(n)}$$

where the sum is over all trees T on n vertices labeled by $[n]$ and $d_T(i)$ is the degree of vertex i in T.

(b) Let F be a forest with vertex set $[n]$, with components having $f_1, \dots, f_s$ vertices so that $f_1 + \cdots + f_s = n$. Prove that the number of trees on the vertex set $[n]$ that contain F is exactly $n^{n-2}\prod_{i=1}^{s}(f_i/n^{f_i - 1})$.

(c) *Independence property for uniform spanning tree of K_n.* Deduce from (b) that if H_1 and H_2 are vertex-disjoint subgraphs of K_n, then for a uniformly random spanning tree T of K_n, the events $H_1 \subseteq T$ and $H_2 \subseteq T$ are independent.

ps4⋆ (d) *Packing rainbow spanning trees.* Prove that there is a constant $c > 0$ so that for every edge-coloring of K_n where each color appears at most cn times, there exist at least cn edge-disjoint spanning trees, where each spanning tree has all its edges colored differently.

ps4 E14. Prove that there is a constant $c > 0$ so that if H is an n-vertex m-edge graph with maximum degree at most cn^2/m, then one can find two edge-disjoint copies of H in the complete graph K_n.

ps4⋆ E15. Prove that there is a constant $c > 0$ so that every $n \times n$ matrix where no entry appears more than cn times contains cn disjoint Latin transversals.

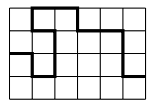
F. Correlation inequalities

F1. Let $G = (V, E)$ be a graph. Color every edge with red or blue independently and uniformly at random. Let E_0 be the set of red edges and E_1 the set of blue edges. Let $G_i = (V, E_i)$ for each $i = 0, 1$. Prove or disprove:

$$\mathbb{P}(G_0 \text{ and } G_1 \text{ are both connected}) \le \mathbb{P}(G_0 \text{ is connected})^2.$$

ps5 F2. A set family $\mathcal{F}$ is *intersecting* if $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{F}$. Let $\mathcal{F}_1, \dots, \mathcal{F}_k$ each be a collection of subsets of $[n]$ and suppose that each $\mathcal{F}_i$ is intersecting. Prove that $\left|\bigcup_{i=1}^k \mathcal{F}_i\right| \le 2^n - 2^{n-k}$.

F3. Let $G_{m,n}$ be the grid graph on vertex set $[m] \times [n]$ (m vertices wide and n vertices tall). A *horizontal crossing* is a path that connects some left-most vertex to some right-most vertex. See below for an example of a horizontal crossing in $G_{7,5}$.



Let $H_{m,n}$ denote the random subgraph of $G_{m,n}$ obtained by keeping every edge with probability $1/2$ independently.

Let $\mathsf{RSW}(k)$ denote the following statement: there exists a constant $c_k > 0$ such that for all positive integers n, $\mathbb{P}(H_{kn,n} \text{ has a horizontal crossing}) \ge c_k$.

ps5 (a) Prove $\mathsf{RSW}(1)$.

ps5 (b) Prove that $\mathsf{RSW}(2)$ implies $\mathsf{RSW}(100)$.

(c) (Challenging) Prove $\mathsf{RSW}(2)$.

F4. Let U_1 and U_2 be increasing events and D a decreasing event of independent boolean random variables. Suppose U_1 and U_2 are independent. Prove that $\mathbb{P}(U_1 | U_2 \cap D) \le \mathbb{P}(U_1 | U_2)$.

ps5⋆ F5. *Coupon collector.* Let $s_1, \dots, s_m$ be independent random elements in $[n]$ (not necessarily uniform or identically distributed; chosen with replacement) and $S = \{s_1, \dots, s_m\}$. Let I and J be disjoint subsets of $[n]$. Prove that $\mathbb{P}(I \cup J \subseteq S) \le \mathbb{P}(I \subseteq S)\mathbb{P}(J \subseteq S)$.

ps5⋆ F6. Prove that there exist $c < 1$ and $\epsilon > 0$ such that if $A_1, \dots, A_k$ are increasing events of independent boolean random variables with $\mathbb{P}(A_i) \le \epsilon$ for all i, then, letting X denote the number of events A_i that occur, one has $\mathbb{P}(X = 1) \le c$. (Give your smallest c. It is conjectured that any $c > 1/e$ works.)

G. Janson inequalities

G1. Prove that with probability $1 - o(1)$, the size of the largest subset of vertices of $G(n, 1/2)$ inducing a triangle-free subgraph is $\Theta(\log n)$.

ps5 G2. Show that for every constant $C > 16/5$, if $n^2 p^5 > C \log n$, then with probability $1 - o(1)$, every edge of $G(n, p)$ is contained in a K_4.

G3. *Vertex-disjoint triangles in $G(n,p)$ again.* Using Janson inequalities this time, give another solution to Problem C7 in the following generality.

ps5 (a) Prove that for every $\epsilon > 0$, there exists $C_\epsilon > 0$ such that such that with probability $1 - o(1)$, $G(n, C_\epsilon n^{-2/3})$ contains at least $(1/3 - \epsilon)n$ vertex-disjoint triangles.

(b) (Optional) Compare the the dependence of the optimal C_ϵ on ϵ you obtain using the method in Problem C7 versus this problem (don't worry about leading constant factors).

ps5 G4. *Lower tails of small subgraph counts.* Fix graph H and $\delta \in (0,1]$. Let X_H denote the number of copies of H in $G(n,p)$. Prove that for all n and $0 < p < 0.99$,

$$\mathbb{P}(X_H \le (1-\delta)\mathbb{E}X_H) = e^{-\Theta_{H,\delta}(\Phi_H)} \quad \text{where } \Phi_H := \min_{H' \subseteq H : e(H') > 0} n^{v(H')} p^{e(H')}.$$

Here the hidden constants in $\Theta_{H,\delta}$ may depend on H and δ (but not on n and p).

ps5⋆ G5. Show that the list chromatic number of $G(n,1/2)$ is $(1+o(1))\frac{n}{2\log_2 n}$ with probability $1-o(1)$.

H. Concentration of measure

H1. For each part, prove that there is some constant $c > 0$ so that, for all $\lambda > 0$,

$$\mathbb{P}(|X - \mathbb{E}X| \ge \lambda\sqrt{\operatorname{Var} X}) \le 2e^{-c\lambda^2}.$$

ps6 (a) X is the number of triangles in $G(n,1/2)$.

ps6 (b) X is the number of inversions of a uniform random permutation of $[n]$ (an *inversion* of $\sigma \in S_n$ is a pair (i,j) with $i < j$ and $\sigma(i) > \sigma(j)$).

H2. Prove that for every $\epsilon > 0$ there exists $\delta > 0$ and n_0 such that for all $n \ge n_0$ and $S_1, \dots, S_m \subset [2n]$ with $m \le 2^{\delta n}$ and $|S_i| = n$ for all $i \in [m]$, there exists a function $f\colon [2n] \to [n]$ so that $(1 - e^{-1} - \epsilon)n \le |f(S_i)| \le (1 - e^{-1} + \epsilon)n$ for all $i \in [m]$.

H3. *Simultaneous bisections.* Fix Δ. Let $G_1, \dots, G_m$ with $m = 2^{o(n)}$ be connected graphs of maximum degree at most Δ on the same vertex set V with $|V| = n$. Prove that there exists a partition $V = A \cup B$ so that every G_i has $(1+o(1))e(G_i)/2$ edges between A and B.

ps6⋆ H4. Prove that there exists a constant $c > 0$ so that the following holds. Let G be a d-regular graph and $v_0 \in V(G)$. Let $m \in \mathbb{N}$ and consider a simple random walk $v_0, v_1, \dots, v_m$ where each v_{i+1} is a uniform random neighbor of v_i. For each $v \in V(G)$, let X_v be the number times that v appears among $v_0, \dots, v_m$. For that for every $v \in V(G)$ and $\lambda > 0$

$$\mathbb{P}\left(\left|X_v - \frac{1}{d}\sum_{w \in N(v)} X_w\right| \ge \lambda + 1\right) \le 2e^{-c\lambda^2/m}$$

Here $N(v)$ is the neighborhood of v.

ps6⋆ H5. *Traveling salesman.* Let $x_1, \dots, x_n$ be random points in $[0,1]^2$ chosen independently and uniformly. Let Z denote the length (in Euclidean distance) of the shortest path that contains all n points. Prove that $\mathbb{P}(|Z - \mathbb{E}Z| \ge \lambda) \le 2\exp(-c\lambda^2/\log n)$ for all $\lambda \ge 0$, where $c > 0$ is some constant.

H6. Prove that for every k there exists a $2^{(1+o(1))k/2}$-vertex graph that contains every k-vertex graph as an induced subgraph.

ps6 H7. *Tighter concentration of chromatic number*

(a) Prove that with probability $1 - o(1)$, every vertex subset of $G(n,1/2)$ with at least $n^{1/3}$ vertices contains an independent set of size at least $c\log n$, where $c > 0$ is some constant.

(b) Prove that there exists some function $f(n)$ and constant C such that for all $n \ge 2$,

$$\mathbb{P}(f(n) \le \chi(G(n,1/2)) \le f(n) + C\sqrt{n}/\log n) \ge 0.99.$$

ps6 H8. Let $k \le n/2$ be positive integers and G an n-vertex graph with average degree at most n/k. Prove that a uniform random k-element subset of the vertices of G contains an independent set of size at least ck with probability at least $1 - e^{-ck}$, where $c > 0$ is a constant.

ps6⋆ H9. Let $G = (V, E)$ with chromatic number $\chi(G) = k$ and S a uniform random subset of V. Prove that for every $t \ge 0$,

$$\mathbb{P}(\chi(G[S]) \le k/2 - t) \le e^{-ct^2/k},$$

where $c > 0$ is a constant and $G[S]$ is the subgraph induced by S.

ps6⋆ H10. Prove that there is some constant $c > 0$ so that, with probability $1 - o(1)$, $G(n, 1/2)$ has a bipartite subgraph with at least $n^2/8 + cn^{3/2}$ edges.

ps6 H11. Show that for every $\epsilon > 0$ there exists $C > 0$ so that every $S \subset [4]^n$ with $|S| \ge \epsilon 4^n$ contains four elements whose pairwise Hamming distance at least $n - C\sqrt{n}$.

ps6⋆ H12. *Concentration of measure in the symmetric group.* Let $A \subset S_n$ be a set of at least $n!/2$ permutations of $[n]$. Let A_t denote the set of permutations that can be obtained starting from some element of A and then applying at most t transpositions. Prove that

$$|A_t| \ge (1 - e^{-ct^2/n})n!$$

for every $t \ge 0$, where $c > 0$ is some constant.

For the remaining exercises in this section, use Talagrand's inequality

H13. Let Q be a subset of the unit sphere in $\mathbb{R}^n$. Let $\boldsymbol{x} \in [-1,1]^n$ be a random vector with independent random coordinates. Let $X = \sup_{\boldsymbol{q} \in Q} \langle \boldsymbol{x}, \boldsymbol{q} \rangle$. Let $t > 0$. Prove that

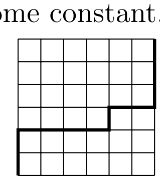
$$\mathbb{P}(|X - \mathbb{M}X| \ge t) \le 4e^{-ct^2}$$

where $c > 0$ is some constant.

ps6 H14. *First passage percolation.* Every edge in the grid graph on $\mathbb{Z}^2$ is independently assigned some random weight in $[0, 1]$ (not necessarily uniform or identically distributed). The weight of a path is defined to be the sum of the weights of its edges. Let X_n be the minimum weight of a path from $(0, 0)$ to (n, n) taking only rightward and upward steps. Show that

$$\mathbb{P}(|X_n - \mathbb{M}X_n| \ge t) \le 4e^{-ct^2/n},$$

for every n and $t \ge 0$, where $c > 0$ is some constant.



ps6⋆ H15. *Second largest eigenvalue of a random matrix.* Let A be an $n \times n$ random symmetric matrix whose entries on and above the diagonal are independent and in $[-1, 1]$. Show that the second largest eigenvalue $\lambda_2(A)$ satisfies

$$\mathbb{P}(|\lambda_2(A) - \mathbb{M}\lambda_2(A)| \ge t) \le 4e^{-ct^2},$$

for every $t \ge 0$, where $c > 0$ is some constant.

Hint in white:

H16. *Longest common subsequence.* Let $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_m)$ be two random sequences with independent entries (not necessarily identically distributed). Let X denote the length of the longest common subsequence, i.e., the largest k such that there exist $i_1 < \cdots < i_k$ and $j_1 < \cdots < j_k$ with $x_{i_1} = y_{j_1}, \ldots, x_{i_k} = y_{j_k}$. Show that, for all $t \geq 0$,

$$\mathbb{P}(X \geq \mathbb{M}X + t) \leq 2\exp\left(\frac{-ct^2}{\mathbb{M}X + t}\right) \quad \text{and} \quad \mathbb{P}(X \leq \mathbb{M}X - t) \leq 2\exp\left(\frac{-ct^2}{\mathbb{M}X}\right)$$

where $c > 0$ is some constant.

I. Entropy method

I1. *Submodularity.* Prove that $H(X,Y,Z) + H(X) \leq H(X,Y) + H(X,Z)$.

I2. Let $\mathcal{F}$ be a collection of subsets of $[n]$. Let p_i denote the fraction of $\mathcal{F}$ that contains i. Prove that

$$|\mathcal{F}| \leq \prod_{i=1}^{n} p_i^{-p_i}(1-p_i)^{-(1-p_i)}.$$

I3. *Uniquely decodable codes).* Let $[r]^*$ denote the set of all finite strings of elements in $[r]$. Let A be a finite subset of $[r]^*$ and suppose no two distinct concatenations of sequences in A can produce the same string. Prove that $\sum_{a \in A} r^{-|a|} \leq 1$ where $|a|$ is the length of $a \in A$.

ps6⋆ I4. *Triangles versus vees in a directed graph.* Let V be a finite set, $E \subseteq V \times V$, and

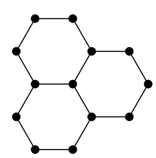
$$\triangle = \left|\left\{(x,y,z) \in V^3 : (x,y),(y,z),(z,x) \in E\right\}\right|$$

(i.e., cyclic triangles; note the direction of edges) and

$$\wedge = \left|\left\{(x,y,z) \in V^3 : (x,y),(x,z) \in E\right\}\right|.$$

Prove that $\triangle \leq \wedge$.

ps6⋆ I5. Prove Sidorenko's conjecture for the following graph.



ps6 I6. *Loomis–Whitney for sumsets.* Let A, B, C be finite subsets of some abelian group. Writing $A + B := \{a + b : a \in A, b \in B\}$, etc., prove that

$$|A+B+C|^2 \leq |A+B|\,|A+C|\,|B+C|.$$

I7. Let X, Y, Z be independent random elements of some abelian group. Prove that

$$2H(X+Y+Z) \leq H(X+Y) + H(X+Z) + H(Y+Z).$$

I8. Let $\mathcal{G}$ be a family of graphs on vertices labeled by $[2n]$ such that the intersection of every pair of graphs in $\mathcal{G}$ contains a perfect matching. Prove that $|\mathcal{G}| \leq 2^{\binom{2n}{2}-n}$.