



PUMaC 2019 Power Round Solutions

November 9, 2019

1 1.4.1

[5 points]

1. Note that $A \cup A^c = \Omega$, so $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1$.
2. Note that $B = A \cup (B \setminus A)$, so $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$.

1.1 Grading

2 points for the first part, 3 points for the second.

2 1.4.2

[5 points]

1. Since $f(t) \geq 0$ and $(X - \mathbb{E}(X))^2 \geq 0$, we know $\text{Var}(X) \geq 0$.
2. If $\text{Var}(X) = 0$, it follows for every term in the variance sum, $f(t) = 0$ or $\mathbb{P}(X = t) = 0$. This means $t = \mathbb{E}(X)$ or $\mathbb{P}(X = t) = 0$. If X is nonrandom, then $t = \mathbb{E}(X)$ for all t , so $\text{Var}(X) = 0$.
3. We find

$$\begin{aligned}
 \text{Var}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) \\
 &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\
 &= \mathbb{E}(X^2) - \mathbb{E}(2X\mathbb{E}(X)) + \mathbb{E}(X)^2 \\
 &= \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 \\
 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2.
 \end{aligned}$$

4. We know $\mathcal{E}(XY) = \mathcal{E}(X)\mathcal{E}(Y)$ if and only if X and Y are independent random variables. Then

$$\begin{aligned}
 \text{Var}(X + Y) &= \mathcal{E}((X + Y - \mathcal{E}(X + Y))^2) \\
 &= \mathcal{E}((X + Y)^2) - \mathcal{E}(X + Y)^2 && \text{(Part 3)} \\
 &= \mathcal{E}(X^2) + \mathcal{E}(XY) + \mathcal{E}(YX) + \mathcal{E}(Y^2) - (\mathcal{E}(X) + \mathcal{E}(Y))^2 \\
 &= 0 && \text{(If and only if } \mathcal{E}(XY) = \mathcal{E}(X)\mathcal{E}(Y)\text{)}
 \end{aligned}$$

5. Using part 3 and linearity of expectation, note that $\text{Var}(aX) = \mathcal{E}((aX)^2) - \mathcal{E}(aX)^2 = a^2\mathcal{E}(X^2) - a^2\mathcal{E}(X)^2 = a^2(\mathcal{E}(X^2) - \mathcal{E}(X)^2) = a^2 \text{Var}(X)$.

2.1 Grading

1 points for each part.

3 1.4.3

[5 points] We induct on the number of sets. Note that

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 \cup (A_2 \setminus A_1)) = \mathbb{P}(A_1) + \mathbb{P}(A_2 \setminus A_1) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2).$$

By induction we can find that

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) \leq \mathbb{P}(A_1 \cup \cdots \cup A_{n-1}) + \mathbb{P}(A_n),$$

hence the desired result.

3.1 Grading

2 for having the idea to use induction, 5 for complete proof.

4 1.4.4

[10 points] We find that

$$\begin{aligned} \mathcal{E}(X) &= \sum_x x\mathbb{P}(X = x) \\ &\geq \sum_{x \geq t} x\mathbb{P}(X = x) && (t \text{ is nonnegative}) \\ &\geq t \sum_{x \geq t} \mathbb{P}(X = x) \\ &= t\mathbb{P}(X \geq t). \end{aligned}$$

This proves the desired statement.

4.1 Grading

4 points for writing out the formula for expected value, 10 for finishing the proof.

5 1.5.1

[5 points] If X is the value of the roll of a die, then $\mathcal{E}(X) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2}$. The variance is $\text{Var}(X) = \mathcal{E}(X^2) - \mathcal{E}(X)^2 = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \frac{7^2}{2^2} = \frac{35}{12}$. For one hundred rolls, let the values be X_1, \dots, X_{100} . Since the rolls are independent $\mathcal{E}(X_1 + \cdots + X_{100}) = 100\mathcal{E}(X_1) = 350$. For the variance, we find $\text{Var}(X_1 + \cdots + X_{100}) = 100 \text{Var}(X_1) = \frac{875}{3}$.

5.1 Grading

2 points for expected value, 3 for variance.

6 1.5.2

[5 points] We have three cases: if we roll RR we are done; if we roll RB we have to start over with two extra rolls; if we roll B we start over with one extra roll. Therefore, $\mathcal{E}(R) = \frac{1}{4}(2) + \frac{1}{4}(2 + \mathcal{E}(R)) + \frac{1}{2}(1 + \mathcal{E}(R))$, so $\mathcal{E}(R) = 6$.

Similarly, $\mathcal{E}(K) = \frac{1}{26^2}(2) + \frac{1}{26} \cdot \frac{25}{26}(\mathcal{E}(K) + 2) + \frac{25}{26}(\mathcal{E}(K) + 1)$. Solving for $\mathcal{E}(K)$ gives $\mathcal{E}(K) = 702$.

For the random variable M , we claim that this is the random variable representing the minimal number of draws needed to draw one red card plus 1. To observe this, note that $\mathbb{P}(M = t)$ implies that the last two cards drawn are either RR or RB , and that for every pair of adjacent cards before, they are either BB or BR . But this requires that the first $t - 2$ cards are B and the $t - 1$ st card is red. Hence, it suffices to compute the minimal number of draws it takes to draw a red card, and then add 1. To do this, we repeat the same argument as before: $\mathcal{E} = \frac{1}{2} + \frac{1}{2}(\mathcal{E} + 1)$, which means that $\frac{\mathcal{E}}{2} = 1$ and so $\mathcal{E} = 2$, which means that $\mathcal{E}(M) = 3$.

6.1 Grading

1 point for each of R, B, K and 2 points for each M . Subtract at most 1 point if there is a calculation error but the reasoning is correct.

7 1.5.3

[10 points] First, $\mathcal{E}(X_{10}X_{20}) = \frac{11}{2} \cdot \frac{21}{2} = \frac{231}{4}$.

For the variance,

$$\begin{aligned}\text{Var}(X_{10}X_{20}) &= \mathcal{E}((X_{10}X_{20})^2) \\ &= \mathcal{E}(X_{10}^2 X_{20}^2) + \mathcal{E}(-2X_{10}X_{20})^2 + \mathcal{E}(X_{10}X_{20})^2 \\ &= \left(\frac{1^2 + \dots + 10^2}{10}\right) \left(\frac{1^2 + \dots + 20^2}{20}\right) - \mathcal{E}(X_{10}X_{20})^2 \\ &= \frac{11 \cdot 21}{6} \frac{21 \cdot 41}{6} - \left(\frac{11 \cdot 21}{4}\right)^2 \\ &= \frac{35035}{16}.\end{aligned}$$

By manual computation, we observe $\mathcal{E}(X_{10}^2) = \frac{21 \cdot 11}{6}$, $\mathcal{E}(X_{10}^3) = \frac{11^2 \cdot 10}{4}$, and $\mathcal{E}(X_{10}^4) = \frac{25333}{10}$. This means

$$\begin{aligned}\text{Var}(X_{10}(X_{20} - X_{10})) &= \mathcal{E}(X_{10}^2 X_{20}^2) - 2\mathcal{E}(X_{10}^3 X_{20}) + \mathcal{E}(X_{10}^4) - \mathcal{E}(X_{10}X_{20})^2 \\ &\quad + 2\mathcal{E}(X_{10}X_{20})\mathcal{E}(X_{10}^2) - \mathcal{E}(X_{10}^2)^2 = \frac{106799}{80}\end{aligned}$$

7.1 Grading

3 for expectation of X_1 , 3 for variance of X_1 , 3 for variance of $X_2 - X_1$. Subtract up to 2 points for an error in calculation with correct reasoning.

8 1.5.4

[5 points]

A is the event that the second coin lands heads. B is the event that the first coin lands head. We want to compute

$$\mathbb{P}[A|B] = \mathbb{P}[A \cap B]/\mathbb{P}[B].$$

We have that

$$\begin{aligned}\mathbb{P}[A] &= \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2} \\ \mathbb{P}[A \cap B] &= \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{3}\end{aligned}$$

So, the answer is $(1/3)/(1/2) = \boxed{1/3}$. If you return the first coin, the second coin is independent of the first coin. From symmetry, answer is $1/2$.

8.1 Grading

3 points for first part and 2 points for the second part.

9 1.5.5

[20 points] Let $p = 3k + 1$ be a prime sufficiently large (so that S have distinct modulus in p). Consider $T = \{k + 1, \dots, 2k + 1\}$. Modulo p , T is sum-free. We will pick a random $a \in \mathbb{Z}_p^*$ and compute the number of elements in S which land in T (modulo p). This subset will be sum-free. The probability that $s \in S$ will land in T is $(k + 1)/(p - 1) \geq 1/3$ there is a unique a which makes s equal to $k+1, k+2, \dots, 2k+1$. From Linearity of Expectation, the expected number of elements that land in T is $(1/3)|S|$, which suffices for the proof.

9.1 Grading

5 points for trying to use expectation, another 5 for looking at the probability that $s \in S$ land in T .

10 1.5.6

[10 points]

There are $\binom{n}{3}$ triangles. Each triangle has a probability of p^3 of being in the graph. So, there are $\binom{n}{3}p^3$ expected triangles.

10.1 Grading

5 points for looking at each triangle separately. 10 for complete solution. Subtract up to 2 points for error in computation with correct reasoning.

11 1.5.7

[30 points] Take $32n$ edges, randomly, uniformly, independently with repetitions from A to B , then fix $X \subseteq A, Y \subseteq B, |X| = |Y| = \frac{n}{4}$. Now we have that $P(\text{no edge is from } X \text{ to } Y) = (1 - \frac{\frac{n^2}{16}}{n^2})^{32n} = (1 - \frac{1}{16})^{32n} < e^{-2n}$. There exists only $(\frac{n}{4})^2$ pairs of X, Y which is $< 2^{2n}$. Thus since $2^{2n}e^{-2n} \ll 1$ there is such graph.

11.1 Grading

5 points for looking at $32n$ edges randomly uniformly independently and 10 looking at the probability at $P(\text{no edge is from } X \text{ to } Y)$. The rest for finishing the proof.

12 2.1.1

[20 points] Second part implies the first part. The second part is equivalent to proving that as m goes to infinity the probability of P being not ε -regular goes to 0, and this follows from the union bound, namely $P(U, W \text{ not } \varepsilon\text{-regular}) \leq \sum_{\text{any pair } X, Y, |X| \geq \varepsilon|U, |Y| \geq \varepsilon|W|} P(X, Y \text{ is not } \varepsilon \text{ regular}) \leq 1$.

12.1 Grading

5 points for proving that second part implies the first part. Another 5 for doing manipulation with probabilities. The rest for finishing the proof.

13 2.1.2

[10 points] Since $d^2(U, W) \leq 1$ for all U, W and $\sum_{U, W, \text{unordered}} |U||W| \leq \binom{n}{2}$, where n is the number of vertices of the graph it follows that $q(P) \geq 0, q(P) \leq \frac{1}{2}$.

13.1 Grading

4 for looking at $d(U, W) \leq 1$ and 4 for looking at unordered pairs. The rest for finishing the proof.

14 2.1.3

[10 points] Follows immediately from lemma 2.1.1.

14.1 Grading

At most 5 points if a solution mentions lemma 2.1.1 but is not completed.

15 2.1.4

[30 points] Since (U, W) is not ε -regular there exists $U_1, W_1 \subseteq U, W$, $|U_1| \geq \varepsilon|U|$, $|W_1| \geq \varepsilon|W|$, $|d(U_1, W_1) - d(U, W)| \geq \varepsilon$. Let Z be as before. We have that $\text{var}(z) = \frac{n^2}{|U||W|}(q(\mathcal{U}, \mathcal{W}) - q(U, W))$. On the other hand since Z differs from $E(z) = d(U, W)$ by $\geq \varepsilon$ with probability $\geq \frac{|U_1||W_1|}{|U||W|}$ thus $\text{var}(z) \geq \frac{|U_1||W_1|}{|U||W|}\varepsilon^2$. The result follows from this.

15.1 Grading

10 points for showing that $\text{var}(z) = \frac{n^2}{|U||W|}(q(\mathcal{U}, \mathcal{W}) - q(U, W))$. Another 10 for looking at $E(z)$ and difference with $d(U, W)$. At most 20 for looking at these but not finishing the proof.

16 2.1.5

[50 points] Put $\frac{n}{k} \geq c = |V_1| = \dots = |V_k| \geq \frac{(1-\varepsilon)n}{k}$. For all $1 \leq i < j \leq k$ define a partition V_{ij} of V_i into ≤ 2 parts and a partition of V_j into ≤ 2 parts as follows: if (V_i, V_j) is ε -regular V_{ij}, V_{ji} are trivial (only one part) otherwise we partition as in previous proof. Take a partition \mathcal{V}_i of each V_i defined by the Venn diagram of all $k-1$ partitions of $V_{i,j}$. \mathcal{V}_i has at most 2^{k-1} parts. To make their sizes equal let $b = \lfloor \frac{c}{4^k} \rfloor = \frac{|V_i|}{4^k}$ and split each part of \mathcal{V}_i into disjoint sets of sizes b each, and if there is something remaining add it to the exceptional sets V'_0 . Now $|V'_0| \leq |V_0| + bk2^{k-1} \leq |V_0| + \frac{c}{4^k}k2^{k-1} \leq |V_0| + \frac{n}{2^k}$ (n is the number of vertices of the graph), since $ck \leq n$. Also note that $l \leq k\frac{c}{b} \leq k4^k$. Since P is not ε -regular by problem 2.1.4 and by lemma 2.1.1 we have that $q(P') \geq q(P) + \varepsilon k^2 \frac{c^2}{n^2} \varepsilon^4 \geq q(P) + \frac{1}{2}\varepsilon^5$, where the last inequality follows from $kc \geq (1-\varepsilon)n \geq \frac{3}{4}n$.

16.1 Grading

35 for showing that there exists a new partition of V_i such that it's equitable and $|V'_0| \leq |V_0| + \frac{n}{2^k}$. 10 for mentioning problem 2.1.4 and lemma 2.1.1

17 2.1.6

[15 points] Start with any partition into t equal parts such that $|V_0| \leq t(\leq \frac{\varepsilon n}{2}$ for large n). As long as it is not ε -regular apply 2.1.5 and note that index increases by at least $\frac{1}{2}\varepsilon^5$ and is never bigger than $\frac{1}{2}$ so it stops in at most s steps and the size of the exceptional set increases by at most $s\frac{n}{2^t} \leq (1 + \frac{1}{\varepsilon^5})\frac{n\varepsilon^6}{4} = \frac{n\varepsilon^6}{4} + \frac{n\varepsilon}{4} \leq \frac{n\varepsilon}{2}$. (because $2^{t-2} \geq \frac{1}{\varepsilon^6}$).

17.1 Grading

10 points for applying 2.1.5 by starting from a partition.

18 2.2.1

[50 points] Take, say, $\frac{\varepsilon}{10}$ regular partition of G with $t = \frac{10}{\varepsilon}$, which exists by Collegiate Splitting theorem. $V = v_0 \cup V_1 \dots V_k$, $t \leq k \leq T = T(t, \varepsilon)$. Remove now the following edges:

1. All edges incident with $V_0 (\geq \frac{\varepsilon n^2}{10}$ edges)
2. all edges inside sets V_i , $\leq \frac{\varepsilon n^2}{20}$ edges, as graph of edges removed here is of max degree $|V_1| \leq \frac{n}{t} = \frac{\varepsilon n}{10}$
3. all edges between irregular pairs which are not ε -regular, $\leq \frac{\varepsilon}{10} k^2 (\frac{n}{k})^2 = \frac{\varepsilon n^2}{10}$ edges
4. all edges between pairs of density $\leq \frac{3\varepsilon}{10}$, at most $\binom{k}{2} \frac{3\varepsilon}{10} (\frac{n}{k})^2 \leq \frac{3}{20} \varepsilon n^2$.

Altogether number of edges removed is $\leq \varepsilon n^2$. Thus at least one triangle is left. This triangle has vertices in 3 distinct parts V_p, V_q, V_r so that all pairs of parts are $\frac{\varepsilon}{10}$ -regular and of density $\geq \frac{3\varepsilon}{10}$. Wlog let pairs V_1, V_2, V_3 . Note that all vertices $v \in V_1$ but $\leq \frac{\varepsilon |V_1|}{10}$ have $\geq (d(V_1, V_2) - \frac{\varepsilon}{10}) |V_2|$ neighbors in V_2 , since otherwise those that do not, and V_2 violate $\varepsilon/10$ -regularity of (V_1, V_2) . Similarly all $v \in V_1$ have $\geq (d(V_1, V_3) - \frac{\varepsilon}{10}) |V_3| \geq \frac{2\varepsilon}{10} |V_3|$ neighbors in V_3 . Fix $v \in V_1$ with $|N(v) \cap V_2| \geq \frac{2\varepsilon}{10} |V_2|$ $|N(v) \cap V_3| \geq \frac{2\varepsilon}{10} |V_3|$ ($N(v)$ is neighbors of v). Then by $\frac{\varepsilon}{10}$ regularity of V_2, V_3 there exists $\geq (\frac{2\varepsilon}{10} |V_2| |V_3| \frac{2\varepsilon}{10})$ edges between $N(v) \cap V_2$ and $N(v) \cap V_3$, so v belongs to $\geq \frac{8\varepsilon^3}{10^3} |V_1| \geq \frac{8\varepsilon^3}{10^3} (\frac{n}{2T})^2$. Number of such v is $\geq (1 - \frac{2\varepsilon}{10}) |V_1| > \frac{n}{2T}$ which implies that the number of triangles is $\geq \frac{8\varepsilon^3}{10^3} (\frac{n}{2T})^3$ proving the theorem with $\delta = \frac{8\varepsilon^3}{1000} \frac{1}{8T^3}$.

18.1 Grading

5 points for looking at $c\varepsilon$ regular partition. 10 removing the edges as in the proof such that there's at most εn^2 of them. 35 for finishing the proof.

19 2.2.2

[60 points] Bob wins. We will prove that for any $\varepsilon > 0$ there exists n_0 such that for any $|A| \subseteq \varepsilon n$ A contains a three term arithmetic progression. Let G be a three-partite graph with classes of vertices of sizes h_1, h_2, h_3 , H_1, H_2, H_3 each beign a copy of $[n]$. For any $k \in [n], a \in A$ add to the graph the three edges of a triangle on vts $h \in H_1, h + a \in H_2, h + 2a \in H_3$ note that all the make $\geq \varepsilon n^2$ triangles which are pairwise edge disjoint. Since knowing two vertices among $h, h + a, h + 2a$ we can reconstruct h, a uniquely, we have to delete $\geq \varepsilon n^2 = \frac{\varepsilon}{9} (3n)^2$ edges to destroy all triangles. By Coupon theorem there is $\geq \delta(\varepsilon) n^3$ triangles, and in particular there is at least one triangle besides the initial ones, this gives a, b, c distinct and $h \in [n]$ with $h + a + b = h + 2c$ which implies $a + b = 2c$.

19.1 Grading

5 for reducing to arithmetic progression. Another 10 for wiritng that with an equivalent graph. 45 for finishing the proof.

20 2.2.3

[60 points] Let $|A| \subset [n]$ Fix d which we will choose later, $|A| = \{\sum_{i=0}^s x_i d^i \leq n, 0 \leq x_i < \frac{d}{2}, \sum x_i^2 = K\}$, where K is chosen to maximize $|A|$. Note now that by pigeonhole $|A| \geq \frac{(\frac{d}{2})^{s+1}}{(s+1)(\frac{d}{2})^2}$, choose $d = 2^{\sqrt{\log n}}$ and then $|A| \geq \frac{n}{2^{O(\sqrt{\log n})} > n^{1-\varepsilon}}$ for any ε , $|A|$ has no three term arithmetic progression so $|A|$ does not have two numbers such that one is the mean of the other two.

20.1 Grading

40 for choosing the right A , 20 for finishing the proof.

21 2.2.4 and 2.2.5

[30+30 points] For any $a_1, a_2 \in A$ there exists a unique $(x, y) \in \{2, 3, \dots, 2n\} \times \{2, 3, \dots, 2n\}$ such that $a_1 = (x, y) - a_2$. This means that for any $a_1 \in A$ $P(a_1 \in A_{x,y}) = \frac{|A|}{(2n-1)^2} \geq \frac{\varepsilon n^2}{4n^2} \geq \frac{\varepsilon}{4}$. By linearity of expectation expected size of $|A \cap A_{x,y}|$ is $\geq |A| \frac{\varepsilon}{4} \geq \frac{\varepsilon^2}{4} = \delta n^2$. Put $B = A \cap A'$. Suffices to show that contains the triple for either $d > 0$ or $d < 0$. Define a three partite graph on classes of vertices $V_1 =$ vertical lines, $V_2 =$ horizontal lines, $V_3 =$ line with slope -1 (through lattice points). Two vertices are adjacent if and only if the corresponding intersection of the corresponding lines lies in B . For any b in B we get a triangle corresponding to the three lines in V_1, V_2, V_3 that pass through b . The triangles are edge disjoint which implies that we have to remove $|B| \geq \delta n^2 \geq \frac{\delta}{16} (|V_1|^2 + |V_2|^2 + |V_3|^2)$ to remove all triangles which implies that by Coupon theorem there exists another triangle. Similarly, other two cases hold as well.

21.1 Grading

2.2.4. 10 points for looking at the probability that $a_1 \in A$. 20 for the rest of the solution.

2.2.5 15 for looking at the graph. 15 for finishing the proof.

22 2.2.6

[40 points]

We work in $\mathbb{Z}/(Mn+1)\mathbb{Z}$ for $M = \text{lcm}(2019, 2020, 2039)$. Take the first n elements of this set and take a subset of this that has $|X| \geq \varepsilon n$ elements. Take sets H_1, H_2, H_3 subsets of $\{1, \dots, 4040n\}$ in the integers mod $(Mn+1)$. Join vertices in the following way: if $a_1 \in H_1$ and $a_2 \in H_2$ join them iff $a_2 - a_1 \in 2019X$. Similarly join two elements in H_2, H_3 iff their difference $a_3 - a_2$ is in $2020X$ and elements in H_1 and H_3 iff difference in $4039X$. If we remove all the triangles of the form $h, h + 2019x, h + 4039x$ for $x \in X$, we have removed at least $n|X| \geq \varepsilon n^2$ triangles (pairwise disjoint). By triangle removal lemma, there are at least $\delta(\varepsilon)n^3$ of them, so for n big enough there is at least another triangle besides the ones we have removed. \exists distinct a, b, c , s.t. $h + 2019a + 2020b = h + 4039c$, which is what we want.

22.1 Grading

20 for looking at the right graph and using Triangle removal lemma. 20 for finishing the proof.

23 2.3.1

[50 points] We prove the statement that it holds if we need to remove at least εn^2 triangles. This is stronger since H contains a triangle. Refer to the proof of triangle removal lemma and choose an $(\frac{\varepsilon}{10})^h$ regular partition and remove parts which are not $\frac{\varepsilon}{10}$ regular and of density at least 3ε . Then count the number of copies of K_{h_1, h_2, h_3} in three pairs which are regular and of big enough size, so that there is enough copies of K_{h_1, h_2, h_3} , hence of H .

23.1 Grading

24 2.3.2

[50 points] The theorem is called Erdős-Stone-Simonovits. Solutions are available online. Google just in case. Proof: The lower bound is witnessed by the bipartite graph with almost equal number of vertices ≤ 1 . Suppose G has $\geq (\frac{1}{4} + \varepsilon)n^2$. By Turan's theorem this is ε far from triangle-free. By the previous problem, G contains $K(h, h, h)$ and hence H , where $|H| = h$.

24.1 Grading

25 2.3.3

[75 points] Take an $\frac{\varepsilon}{10}$ -regular partition. Overcount the number of such G , $N(G)$ by counting the number of ways to construct G from G_1 , $M(G : G_1)$. $N(G) \leq N(G_1)M(G : G_1)$. Now we specifically select the graph G_1 to be graph that we get from G when we do the $\frac{\varepsilon}{10}$ partition and then take out all the edges incident to V_0 , between pairs which are not regular and between pairs with density less than $\frac{\varepsilon}{10}$. We want to find $A(G_1)$. Choose $\frac{k^2}{4}$ pairs out of $\binom{k}{2}$ in $\binom{\binom{k}{2}}{\frac{k^2}{4}}$ ways. Furthermore we look at how to reconstruct G from G_1 to get that $B(G : G_1) \leq 2^k 2^{\varepsilon n^2}$, and choosing the sets V_0, V_1, \dots, V_k adds in total $A(G) \leq 2^k n! 2^{(1+4\varepsilon)\frac{n^2}{4}}$. $\log A(G) = 2^{\frac{n^2}{4}(1+o(1))}$.

25.1 Grading

26 2.3.4

[60 points] We use the bound T of the regularity lemma and let $\varepsilon = \frac{c^3}{100}$ and $\frac{1}{T(\varepsilon, \frac{1}{\varepsilon})} \frac{c^3}{100}$.

We do the triangle removal proof and use the following fact: Note that all but $\leq \frac{\varepsilon|V_1|}{10}$ vertices in V_1 have degree $\geq (d(V_1, V_2) - \frac{\varepsilon}{10})|V_2|$ (otherwise the $\varepsilon/10$ regularity would be violated).

Now we find H as a subgraph of the induced subgraph on V_1 and V_2 . We algorithmically find vertices u_1, \dots, u_h , where $h = |H|$, such that in step i we choose u_i , and that they alternate between V_1 and V_2 . In particular we use ε regularity to prove that we can always choose such a vertex. We maintain sets $C_{i,j}$ vertices that u_j can be chosen from after step i , at the start they are V_1 or V_2 . In each step we choose u_i from $C_{i-1,i}$ and update the rest. If $k = \lfloor \frac{1}{\varepsilon} \rfloor$ we prove that $|C_{i-1,i}| \geq 4\varepsilon \frac{n}{k}$ by induction. Updating the ones from the same set V_1 is just deducting the currently chosen vertex. Updating the others uses the fact that the degree is at most 3, and that so far at most δn have been removed to prove that $|C_{i,j}| \geq \dots \geq 4\varepsilon \frac{n}{k}$.

26.1 Grading

27 2.3.5

[75 points]

We are basically looking for a k -regular graph. We apply Colleague partition theorem with the argument for equitable partition (secretly Szemerédi's regularity lemma). There \exists a $\varepsilon/10$ -regular partition. It is enough to prove the claim for a small ε . Start imitating the proof of the triangle removal lemma and take a $\varepsilon/10$ regular pair V_1, V_2 . **Important:** Note that all but $\leq \frac{\varepsilon|V_1|}{10}$ vertices in V_1 have degree $\geq (d(V_1, V_2) - \frac{\varepsilon}{10})|V_2|$ (otherwise the $\varepsilon/10$ regularity would be violated). Let the set of those that have a high degree in V_1 be U_1 . Similarly define U_2 . For $v \in U_1$, $N(v) \cap U_2 \geq (d(V_1, V_2) - \frac{2\varepsilon}{10})|V_2|$. A symmetric statement holds for 2, 1. WLOG $|U_1| = |U_2|$. Then when restricted to U_1, U_2 , the pair is $\frac{\varepsilon}{9}$ regular by a short calculation.

Then the idea is to periodically find a perfect matching between vertices in U_1, U_2 , and remove those edges and show that the perfect matching conditions still hold, performing this in total $\varepsilon^2|U_1|$ times. The removed edges will make a $\varepsilon^2|U_1|$ regular subgraph.

The minimal degree after removing these edges in U_1 will be $d'_{min} = d_{min} - \varepsilon^2|U_1| \geq \varepsilon|U_1|$.

Using regularity prove **Hall's** condition. (There are three cases that depend on the size of the subset of U_1 that we pick. The first one is $|X| \leq \varepsilon|U_1|$, the second one is $|X|$ between this and $(1 - \frac{\varepsilon}{9})|U_2|$, and the third one is bigger than this.) Refer to Igor for clarification.

27.1 Grading

28 3.1.1

[10 points] When we remove or add an edge in graph G , the new graph is isomorphic to the graph H with the corresponding edge removed so this part holds. The second part is not true.

28.1 Grading

29 3.1.2

[15 points]

This part requires some sort of explanation, i.e. description of the tester and reasoning why is it true.

- This is testable.
- This is also testable.
- Trivially testable, since for large n no graphs are ε -far from having k -clique.
- Trivially testable, no graph is ε -far from this property since you always only add a linear number of vertices, so it's less than n^2 for large enough n .
- This is testable. They are only supposed to fill in the details here in the theorem after this problem, i.e. the following text and the theorem that follows it:

Take the ε -tester to be the following: choose randomly uniformly $\frac{100}{\varepsilon^2}$ vertices and accept if and only if the induced subgraph on them is 3-colorable. Here, the tester is also one-sided (if G is 3-colorable it will always accept). In order to prove correctness we have to show the following.

If $G = (V, E)$ with $|V| = n$ is ε -far from 3-colorable then the induced subgraph on a random set of $\frac{100}{\varepsilon^2}$ vertices is with high probability *not* 3-colorable.

29.1 Grading

30 3.1.3

[30 points] By triangle removal if G is ε -far from triangle free and has n vertices then has $\geq \delta n^3$ ($\delta = \delta(\varepsilon)$) triangles. If we take randomly 3 vertices, probability that they form a triangle is $\geq \frac{\delta n^3}{\binom{n}{3}} \geq 6\delta$. Doing it $\frac{1}{\delta}$ times we have that $P(\text{no triangle found} | G \text{ is } \varepsilon\text{-far from triangle free}) \leq (1 - 6\delta)^{\frac{1}{\delta}} \leq e^{-6}$. Thus deciding according to such $\frac{1}{\delta}$ samplings is a good 1-sided ε -tester. Taking randomly $\frac{3}{\delta}$ vertices and deciding according to the induced subgraph on them is a 1-sided ε -tester.

30.1 Grading

31 3.1.4

[40 points] Let $K_{1,t}$ be the star with t edges. Let $d_1 \geq d_2 \geq \dots \geq d_n$ be the degrees of vertices in G let $\bar{d} = \frac{\sum d_i}{n} \geq 2\varepsilon n$ be the average degree. Number of homomorphisms from $K_{1,t}$ to G is $\sum_{i=1}^n d_i^t \geq n\bar{d}^t \geq n(2\varepsilon)^t$ by Jensen. Now we can classify these homomorphisms according to ordered set of t vertices to which the leaves are mapped. There exists $N = n^t$ such classes. Let D_1, D_2, \dots, D_N be numbers of homomorphisms in the classes, let \bar{D} be their average, $\bar{D} \geq \frac{n(2\varepsilon)^t}{N} = n(2\varepsilon)^t$. The number of homomorphisms from $K_{s,t}$ to G is $\sum_{i=1}^N D_i^s \geq N\bar{D}^s \geq N(n(2\varepsilon)^t)^s = (2\varepsilon)^{st} n^{s+t}$. Number of homomorphisms which are not 1-1 is $\leq o(n^{s+t})$ and thus the number of ordered copies is at least $(1 - o(1))(2\varepsilon)^{st} n^{s+t} \geq c(s, t)\varepsilon^{h^2/4} n^h$.

31.1 Grading

32 3.1.5

[15 points] Let m, X be as in the previous lemma. Let G be a 5 partite graph on vertices classes V_1, \dots, V_5 , and $|V_i| = im$. Denote $V_i = \{1, \dots, im\}$, V_i are pairwise disjoint. For any $j \in [m]$ and $x \in X$ take a copy of C_5 on V_i vertices $j \in V_1, j+x \in V_2, \dots, j+4x \in V_5$. There exists $m|X| \geq \frac{m^2}{2^{10(\sqrt{\log m})}}$ copies of C_5 which are pairwise disjoint. Every edge is in exactly one copy of C_5 in G or equivalently G contains exactly $m|X|$ copies of C_5 . Let $r = \frac{n}{15m}$. Let G' be the r -blow up of G . Let $\varepsilon = \frac{1}{2^{10\sqrt{\log m}}}$ which implies $m = 2^{\frac{1}{100}(\log(\frac{1}{\varepsilon}))^2} = (\frac{1}{\varepsilon})^{\frac{1}{100} \log \frac{1}{\varepsilon}}$. Note that now G' contains $m|X|r^5$ copies of $C_5(r^5$ for each C_5 of G). Thus to destroy all copies of C_5 in G we have to delete at least $m|X|r^5/r^3 \geq m|X|r^2 \geq \Omega(\varepsilon n^2)$ edges so G is $\Omega(\varepsilon)C_5$ -free. The total number of copies in G is $m|X|r^5 \leq m^2(n/15m)^5 \leq n^5/m^3 = \varepsilon^{\Theta(\log \frac{1}{\varepsilon})}n^5$ which proves the claim.

32.1 Grading

33 3.1.6

[50 points] Note that if we may assume that any tester pick(perhaps adaptevely) some set of y vertices and queries the $\binom{y}{2}$ pairs on them. Also adaptiveness does not help since the input can be given after randomly permuting the vertices. Thus, we can WLOG assume that any tester simply looks at an induced subgraph and decides based on this since. By 3.1.4 we have proved the direction H is bipartite $\implies P_H$ is easily testable. To prove the other implication we use the theorem 3.1.IV. If we choose y random vertices G is ε -far from P_H and no of H in G is $\leq \varepsilon^{\theta \log(\frac{1}{\varepsilon})}n^h$, probability of seeing a copy of H is $\leq \varepsilon^{\theta(\log \frac{1}{\varepsilon})}n^h \frac{\binom{y}{h}}{\binom{n}{h}} = o(1)$ for $y = \frac{1}{\varepsilon}^{o(\log(\frac{1}{\varepsilon}))}$, which implies that we need at least $(\frac{1}{\varepsilon})^{\Omega(\log \frac{1}{\varepsilon})}$ queries, i.e. it's not easily testable if H is not bipartite.

33.1 Grading

34 3.2.1

[25 points] The first way to solve this is just to prove theorem 3.1.II, by showing that there exists n such that with less than $c\sqrt{n}$ queries we can't distinguish between two random graphs on $\{1, 2, \dots, \frac{n}{2}\}$ and $\{\frac{n}{2} + 1, \dots, n\}$ and two copies of the same random graph. Alternatively, take a similar construction as in theorem 3.1.II, but with 2 parts instead of 2. Then the proof is similar, i.e. prove that we cannot distinguish between two random graphs on n vertices with three parts with less than $o(\sqrt{n})$ queries(or any other function of n).

34.1 Grading

35 3.2.2

[15 points]