Induced Subgraphs in Ramsey Graphs

University of Oxford

Mathematical Institute

Candidate Name: Ray Tsai
Supervisor: Professor Alex Scott

Degree Course: MSc in Mathematical Sciences
Submission: Trinity Term 2026

Submitted in partial fulfilment of the requirements for the MSc in Mathematical Sciences.

Abstract

We study the "random-like" behaviour of Ramsey graphs through the sizes of their induced subgraphs, with particular focus on a conjecture of Erdős and McKay.

Contents

1 Introduction

Classical Ramsey Theory asks for the minimum number of vertices required to guarantee the existence of a clique or an independent set in a graph. We refer to these numbers as *Ramsey numbers*.

Definition 1.1. The *Ramsey number* $R(s,t)$ is the minimum number N such that any graph on N vertices contains a clique of size s or an independent set of size t.

The existence of these numbers was proven by Ramsey [15] in 1930, but their quantitative behaviours remain largely unknown despite decades of intensive research. In particular, it is known that $R(3,3) = 6$ and $R(4,4) = 18$, but the exact values of $R(5,5)$ and beyond are yet to be determined. The cornerstone of Ramsey Theory due to Erdős and Szekeres in 1935 gives the following upper bound on $R(s,t)$:

Theorem 1.1 (Erdős–Szekeres [7])**.** *For $s,t \geq 2$,*

$$R(s,t) \leq \binom{s+t-2}{s-1}.$$

Since $\binom{2t-2}{t-1} < 2^{2t}$, we get $R(t,t) < 4^t$. This bound resisted improvement for decades until the recent breakthrough of Campos, Griffiths, Morris, and Sahasrabudhe [4], giving the first exponential improvement since Erdős–Szekeres. Lower bounds for Ramsey numbers have been even harder to determine. A rather basic result which already gives the best known exponents is the famous random graph construction due to Erdős in 1947, marking one of the first applications of the probabilistic method in combinatorics. Before stating the result, we define the *Erdős–Rényi random graph.*

Definition 1.2. For $n \geq 1$ and $p \in [0,1]$, the *random graph* $G_{n,p}$ is the graph on n vertices in which each of the $\binom{n}{2}$ edges is present independently with probability p.

Remark: For constant p, $G_{n,p}$ is dense in the sense that it has $\Theta(n^2)$ many edges almost surely.

Theorem 1.2 (Erdős [5])**.** *For $t \geq 3$,*

$$\sqrt{2}^t < R(t,t) < 4^t.$$

Proof. The upper bound follows from Theorem 1.1. For the lower bound, set $N := \lfloor \sqrt{2}^t \rfloor$ and consider the random graph $G_{N,1/2}$. The expected number of cliques of size t is

$$\binom{N}{t} \cdot 2^{-\binom{t}{2}} < \frac{N^t}{t!} 2^{-\binom{t}{2}} \leq \frac{\sqrt{2}^{t^2}}{t!} 2^{-\binom{t}{2}} = \frac{\sqrt{2}^t}{t!} = \frac{\sqrt{2}^{t-3}}{t(t-1)\cdots 4} \cdot \frac{\sqrt{2}^3}{3!} \leq \frac{\sqrt{2}^3}{3!} < \frac{1}{2}.$$

Similarly, the expected number of independent sets of size t is less than $1/2$. It now follows that there exists a graph on N vertices containing neither a clique nor an independent set of size t. □

We define a shorthand for cliques and independent sets.

Definition 1.3. A set $S \subseteq V(G)$ is *homogeneous* in a graph G if it induces a clique or an independent set in G. We denote the size of the largest homogeneous set in G by $\hom(G)$.

In terms of homogeneous sets, Theorem 1.2 says that any graph G on n vertices has $\hom(G) > \frac{1}{2}\log_2 n$, while some graph G on n vertices has $\hom(G) < 2\log_2 n$. Beyond improving these bounds, a major problem in Ramsey Theory is to give explicit constructions of graphs with only logarithmically small homogeneous sets. We call such graphs *Ramsey graphs.*

Definition 1.4. For constant $C > 0$, a graph G on n vertices is *C-Ramsey* if $\hom(G) < C\log n$.

Remark: For constant $p \in (0,1)$, $G_{n,p}$ is almost surely a Ramsey graph, as

$$\mathbb{P}\Big[\hom(G_{n,p}) \geq C\log n\Big] \leq \binom{n}{C\log n}\left(p^{\binom{C\log n}{2}} + (1-p)^{\binom{C\log n}{2}}\right) = o(1),$$

for some $C = C(p) > 0$.

Despite decades of intensive research, all known constructions of Ramsey graphs still rely on randomness, which raises the belief that Ramsey graphs are inherently "random-like." In particular, there are various evidence that Ramsey graphs share certain "richness" properties of dense random graphs. The first result that cemented this connection is due to Erdős and Szemerédi, showing that Ramsey graphs are also dense.

Theorem 1.3 (Erdős–Szemerédi [8])**.** *Let G be a graph on n vertices with edge density $d \in (0,1/2]$. Then G contains a homogeneous set of size greater than $K := \frac{c}{d\log(1/d)}\log n$, for some constant $c > 0$ and sufficiently large n.*

Proof. Fix $c < 1/32$ and suppose for contradiction that $\hom(G) \leq K$. Since G has $d\binom{n}{2}$ edges, by pigeonhole at least $m \geq n/2$ vertices have degree $\leq 2dn$. Let G' be the graph induced by these m vertices and let I be a largest independent set in G', so $k := |I| \leq K$. Set $s := 8dk$ and

$$S := 8dK = \frac{8c\log n}{\log(1/d)},$$

so that $s \leq S = O(\log n)$.

We first find a large subset $Z \subseteq V(G') \setminus I$ whose vertices share a common neighbourhood in I. Each $v \in I$ has $\deg(v) \leq 2dn$, so $\sum_{v\in I}\deg_{G'}(v) \leq 2dnk \leq 4dmk$. By pigeonhole, at least $m/2 - k$ vertices $v \in V(G') \setminus I$ satisfy $\deg_I(v) < s$. Call this set J. Partition J by the value of $N_I(v)$ to get at most $|\mathcal{F}|$ classes, where

$\mathcal{F} \coloneqq \{F \subseteq I : |F| < s\}$. The bound $\sum_{j<r} \binom{n}{j} \le (en/r)^r$ gives

$$|\mathcal{F}| \le \left(\frac{ek}{s}\right)^s = \left(\frac{e}{8d}\right)^s \le \left(\frac{1}{d}\right)^s \le \left(\frac{1}{d}\right)^S = n^{8c} < n^{1/3}.$$

By pigeonhole, some class $Z \subseteq J$ with common neighbourhood $N \in \mathcal{F}$ has size

$$|Z| \ge \frac{|J|}{|\mathcal{F}|} \ge \frac{m/2 - k}{n^{1/3}} \ge \sqrt{n}.$$

We now show that $G[Z]$ cannot contain large homogeneous sets. Since $|N| < s$, if $G[Z]$ contained an independent set I^* of size s, then $(I \setminus N) \cup I^*$ would be a strictly larger independent set in G', contradicting the maximality of I. By assumption $G[Z]$ also has no clique of size K, so the Erdős–Szekeres theorem (Theorem 1.1) gives

$$\sqrt{n} \le |Z| \le \binom{s+K-2}{s-1} \le \binom{s+K}{s} \le \binom{S+K}{S}.$$

By the bound $\binom{a+b}{a} \le (e(a+b)/a)^a$,

$$\binom{S+K}{S} \le \left(\frac{e(S+K)}{S}\right)^S \le \left(\frac{e(8d+1)}{8d}\right)^S \le \left(\frac{2}{d}\right)^S,$$

where the last inequality uses $e(1+8d)/8 \le 5e/8 \le 2$ for $d \le 1/2$. But then

$$\left(\frac{2}{d}\right)^S = n^{8c \cdot \frac{\log(2/d)}{\log(1/d)}} = n^{8c\left(1+\frac{\log 2}{\log(1/d)}\right)} \le n^{16c} < \sqrt{n},$$

for $d \le 1/2$ and $c < 1/32$, contradiction. $\square$

Recall that the Erdős–Szekeres gives $\hom(G) \ge \frac{1}{2}\log_2 n$, regardless of the edge density. Theorem 1.3 improves this when the density is small. For example, if $d = 1/n^{\varepsilon}$ for some $\varepsilon > 0$, then $\hom(G) = \Omega\left(\frac{n^{\varepsilon}}{\log n^{\varepsilon}}\log n\right)$, vastly larger than $\frac{1}{2}\log_2 n$. Intuitively, a sparse graph should contain a large independent set, and the theorem makes this notion precise.

If G is C-Ramsey, then applying Theorem 1.3 to whichever G or $\overline{G}$ has density $d \le 1/2$ yields $c/(d\log(1/d)) \le C$, forcing d to be bounded away from 0. The same argument applied to $\overline{G}$ forces d to be bounded away from 1. This gives the following corollary.

Corollary 1.4. *For any C-Ramsey graph G, there exists $\delta = \delta(C) \in (0, 1/2)$ such that the edge density of G lies in the interval $[\delta, 1-\delta]$.*

Due to the rich variety of induced subgraphs in dense random graphs, a popular angle to study the pseudorandom properties of Ramsey graphs is to consider their induced subgraphs. Some notable results are a theorem of Shelah [16] which asserts that every n-vertex Ramsey graph contains $2^{\Omega(n)}$ non-isomorphic induced subgraphs, and a theorem of Prömel and Rödl [14] which shows the existence of $\delta = \delta(C) > 0$ such that every C-Ramsey graph on n vertices contains an induced copy of all graphs on at most $\delta \log n$ vertices.

In this dissertation, we will focus on the edge statistics of induced subgraphs in Ramsey graphs. For a graph G, let

$$\Phi(G) \coloneqq \Big\{ e(H) : H \text{ is an induced subgraph of } G \Big\}.$$

Erdős and McKay [6] conjectured that for any Ramsey graph G on n vertices, there exists $\delta = \delta(C) > 0$ such that

$$\Big\{ 1, 2, \ldots, \delta n^2 \Big\} \subseteq \Phi(G).$$

Calkin, Frieze, and McKay [3] proved it for random graphs, but the conjecture for Ramsey graphs remained open for decades. Significant progress began with Narayanan, Sahasrabudhe, and Tomon [13], who established $|\Phi(G)| = n^{2-o(1)}$ for Ramsey graph G on n vertices. Their proof builds on ideas from Bukh and Sudakov [2], who showed that every Ramsey graph contains a linear-sized subgraph with many distinct degrees, a feature shared with dense random graphs. Kwan and Sudakov [10] later gave an optimal bound $|\Phi(G)| = \Omega(n^2)$ via a refined argument. Finally, Kwan, Sah, Sauermann, and Sawhney [11] closed the conjecture via Fourier-analytic arguments on graph polynomials.

This dissertation starts by studying the result of Bukh and Sudakov on large induced subgraphs in Ramsey graphs with many distinct degrees. This work not only tightens the connection between Ramsey graphs and random graphs, but also builds the foundation for the subsequent $|\Phi(G)| = \Omega(n^2)$ result by Kwan and Sudakov which we will present in the following chapter. Lastly, we outline the proof of the Erdős–McKay conjecture by Kwan, Sah, Sauermann, and Sawhney.

1.1 Anticoncentration

The core probabilistic mechanism in this dissertation is *anticoncentration*. In the following chapters, we will repeatedly sample a uniformly random subset $U \subseteq V(G)$ and study the graph statistics of $G[U]$, such as vertex degrees or edge counts. The guiding principle is that if a random variable depends on many independent choices, then it should not concentrate too heavily on any particular value.

For example, a question to ask in this setting is *how likely is it that two vertices $x, y \in V(G)$ have the same degree in $G[U]$?*

Observe that $\deg_U(x) - \deg_U(y)$ counts only vertices in $D \coloneqq N_G(x) \triangle N_G(y)$. Write $A \coloneqq N(x) \setminus N(y), B \coloneqq N(y) \setminus N(x)$. Then

$$\deg_U(x) - \deg_U(y) = |U \cap A| - |U \cap B|, \tag{1}$$

so x, y have the same degree in $G[U]$ precisely when $|U \cap A| = |U \cap B|$. Since each $v \in A$ contributes $+1$ to the difference $|U \cap A| - |U \cap B|$ with probability $1/2$ and each $v \in B$ contributes -1 with probability $1/2$, the difference is a sum of $|D| = |A| + |B|$ independent random variables, which has standard deviation $\Theta(\sqrt{n})$ whenever $|D| \geq \varepsilon n$ is large. Hence, by the Erdős–Littlewood–Offord Theorem,

$$\mathbb{P}[\deg_U(x) = \deg_U(y)] = O(1/\sqrt{n}).$$

From this, we may construct a *collision graph* and apply Turán's theorem to obtain many distinct degrees in U.

An intuition for this $O(1/\sqrt{n})$ bound is that $\deg_U(x) - \deg_U(y)$ is a *linear* polynomial of many independent indicators, which is approximately normally distributed with standard deviation $\Theta(\sqrt{n})$. Since the probability density function $\phi(x)$ of a normal distribution with standard deviation σ is at most $1/(\sigma\sqrt{2\pi})$ for all x, the probability that $\deg_U(x) - \deg_U(y)$ equals any particular value is $O(1/\sigma) = O(1/\sqrt{n})$, and this anticoncentration bound is the key to the results in Chapters 2 and 3.

In Chapter 4, we will study the induced edge count $e(U)$ itself, which is instead a *quadratic* polynomial of indicators. It turns out a central limit theorem still holds for $e(U)$ at the large scale with standard deviation $\sigma(e(U)) = \Theta(n^{3/2})$, and we prove its point probabilities follow the $O(1/\sigma) = O(n^{-3/2})$ bound that a Gaussian would give, even though $e(U)$ need not be Gaussian at the small scale. This is the same anticoncentration principle as in Chapters 2 and 3, sharpened to a statement about $e(U)$ rather than a degree difference.

1.2 Notations and tools

For $n \in \mathbb{N}$, write $[n] := \{1, 2, \ldots, n\}$. For set X, write $\binom{X}{2}$ as the set of unordered pairs of elements in X. Denote the union of sets X, Y as $X \sqcup Y$ if $X \cap Y = \emptyset$.

Given graph G on n vertices, we call $e(G) := |E(G)|$ the *size* of G, and the ratio

$$d(G) := e(G)\binom{n}{2}^{-1}$$

the *density* of G. For vertex $v \in V(G)$, denote $N(v) := \{w \in V(G) : \{v, w\} \in E(G)\}$ as the *neighbourhood* of v and $\deg(v) := |N(v)|$ as the *degree* of v. Write $\overline{G}$ as the *complement graph* of G and $\overline{N(v)} := V(G) \setminus N(v)$ as the *non-neighbourhood* of v. For $S \subseteq V(G)$, denote $G - S$ as the subgraph of G obtained by deleting S and all edges incident to S. For $U \subseteq V(G)$, we call $G[U] := G - (V(G) \setminus U)$ the *induced subgraph* of U in G. Write $e(U) := e(G[U])$ and $d(U) := d(G[U])$ as the size and density of U, respectively. Write $N_U(v) := N(v) \cap U$, $\deg_U(v) := |N_U(v)|$, and $\overline{N_U(v)} := U \setminus N_U(v)$ as the neighbourhood, degree, and non-neighbourhood of v in U, respectively. For $A, B \subseteq V(G)$, write $e(A, B) := e(A \cup B) - e(A) - e(B)$ and

$$d(A, B) := \frac{e(A, B)}{|A||B|}$$

as the edge count and density between A and B, respectively.

For functions $f, g : \mathbb{R} \to \mathbb{R}$, we use the standard asymptotic notations:

- $f = O(g)$ and $f \lesssim g$ mean there exists constant $C > 0$ such that $f \le C \cdot g$.
- $f = \Omega(g)$ and $f \gtrsim g$ mean there exists constant $c > 0$ such that $f \ge c \cdot g$.
- $f = \Theta(g)$ means $f = O(g)$ and $f = \Omega(g)$.
- $f = o(g)$ means $f/g \to 0$.
- $f = \omega(g)$ means $f/g \to \infty$.

The results in this dissertation mostly concern large graphs, so we omit floor and ceiling signs when they are not crucial.

We call subset $U \subseteq X$ *uniformly random* if each $x \in X$ is in U independently with probability $1/2$. For event A, write its indicator function as

$$\mathbb{1}_A \coloneqq \begin{cases} 1 & \text{if } A \text{ occurs,} \\ 0 & \text{otherwise.} \end{cases}$$

We will use the following probability bounds throughout the dissertation.

Proposition 1.5 (Markov's inequality)**.** *Let X be a non-negative random variable. Then for $t > 0$,*

$$\mathbb{P}\Big[X \ge t\Big] \le \frac{\mathbb{E}[X]}{t}.$$

□

Proposition 1.6 (Chebyshev's inequality)**.** *Let X be a random variable. Then for $t > 0$,*

$$\mathbb{P}\Big[|X - \mathbb{E}[X]| \ge t\Big] \le \frac{\mathrm{Var}(X)}{t^2}.$$

□

For $X \coloneqq \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_k}$, we often use the following inequality to bound the variance:

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \sum_i \sum_j \mathbb{P}[A_i \cap A_j] - \mathbb{P}[A_i]\mathbb{P}[A_j] \\ &= \sum_i \Big[\mathbb{P}[A_i] - \mathbb{P}[A_i]^2\Big] + \sum_i \sum_{j \sim i} \Big[\mathbb{P}[A_i \cap A_j] - \mathbb{P}[A_i]\mathbb{P}[A_j]\Big] \\ &\le \mathbb{E}[X] + \sum_i \sum_{j \sim i} \mathbb{P}[A_i \cap A_j] \end{aligned}$$

where we write $i \sim j$ if $i \neq j$ and A_i, A_j are dependent.

Proposition 1.7 (Hoeffding's inequality)**.** *Let $X_1, X_2, \ldots, X_n$ be independent random variables with $0 \le X_i \le 1$ for each i, and write $X \coloneqq \sum_{i=1}^n X_i$. Then for $t \ge 0$,*

$$\mathbb{P}\Big[|X - \mathbb{E}[X]| \ge t\Big] \le 2e^{-2t^2/n}.$$

$\square$

This inequality applies when X is binomial, in which case we call it the *Chernoff bound*.

Many random variables turn out to be linear polynomials of independent Bernoulli random variables, to which Erdős–Littlewood–Offord can be applied to show anti-concentration. We refer to these random variables as *Littlewood–Offord type*.

Definition 1.5. A random variable X is of (n, p)-*Littlewood–Offord type* if it is of the form $X = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n + C$, where $a_1, a_2, \ldots, a_n \in \mathbb{Z} \setminus \{0\}$ and $C \in \mathbb{Z}$ are fixed and $\xi_1, \xi_2, \ldots, \xi_n$ are independent Bernoulli random variables with parameter p.

Proposition 1.8 (Erdős–Littlewood–Offord theorem)**.** *Let $p \in (0, 1)$ be constant and X be a random variable of (n, p)-Littlewood–Offord type. Then for any $t \in \mathbb{Z}$,*

$$\mathbb{P}[X = t] = O\big(1/\sqrt{n}\big).$$

Proof. First consider the case where all coefficients a_i are positive. Without loss of generality, assume $X \coloneqq \sum_{i=1}^n a_i\xi_i$. For $t \in \mathbb{Z}$, let

$$\mathcal{A}_t \coloneqq \left\{A : \sum_{i \in A} a_i = t\right\} \subseteq [n].$$

Since $a_i \ge 1$ for all i, we have $1 + \sum_{i \in A} a_i \le \sum_{i \in B} a_i$ whenever $A \subset B$. Thus, no elements of $\mathcal{A}_t$ are contained in another. Consider the chain $C \coloneqq (S_0, S_1, \ldots, S_n)$ of n subsets of $[n]$ such that $\emptyset = S_0 \subset S_1 \subset \cdots \subset S_n = [n]$. There are $n!$ such chains C in $[n]$, and for $A \subseteq [n]$, there are $|A|!(n - |A|)!$ of them containing A. Since no chain can contain multiple elements of $\mathcal{A}_t$,

$$\sum_{A \in \mathcal{A}_t} |A|!(n - |A|)! \le n! \implies \sum_{A \in \mathcal{A}_t} \frac{1}{\binom{n}{|A|}} \le 1.$$

By Stirling's formula on binomial coefficients,

$$\begin{aligned}
\mathbb{P}\Big[X = t\Big] &= \sum_{A \in \mathcal{A}_t} p^{|A|}(1-p)^{n-|A|} \\
&\le \sum_{A \in \mathcal{A}_t} \frac{1}{\binom{n}{|A|}} \cdot \binom{n}{|A|} p^{|A|}(1-p)^{n-|A|} \\
&\le \left(\sum_{A \in \mathcal{A}_t} \frac{1}{\binom{n}{|A|}}\right) \cdot \binom{n}{np} p^{np}(1-p)^{n-np} \\
&\le \frac{1 + o(1)}{\sqrt{2\pi p(1-p)n}} = O\big(1/\sqrt{n}\big).
\end{aligned}$$

Now consider the general case. Let $I \coloneqq \{i : a_i \geq 1\}$ and $J \coloneqq \{i : a_i \leq -1\}$. We may assume $|I| \geq n/2$ by passing to $-X$ and swapping I with J if necessary. Conditioning on $(\xi_j)_{j\in J}$, X is of $(|I|, p)$-Littlewood–Offord type with positive coefficients, so

$$\mathbb{P}\Big[X = t \mid (\xi_j)_{j\in J}\Big] = O\Big(1/\sqrt{|I|}\Big) = O\Big(1/\sqrt{n}\Big).$$

The law of total probability gives

$$\mathbb{P}\Big[X = t\Big] = \mathbb{E}\left[\mathbb{P}\Big[X = t \mid (\xi_j)_{j\in J}\Big]\right] = O\Big(1/\sqrt{n}\Big).$$

□

We also use the following version of Turán's theorem.

Proposition 1.9 (Turán's theorem)**.** *Every graph G with n vertices and m edges contains an independent set of size at least $n^2/(2m+n)$.*

Proof. Consider a random ordering $<$ on $V(G)$. Define

$$I \coloneqq \Big\{v \in V : v < u,\ \forall u \in N(v)\Big\}.$$

For distinct $u, v \in V(G)$, either $u > v$ or $v > u$, so u, v cannot both be in I if they are adjacent. Thus I is an independent set. Since $v \in I$ if and only if v is the least element among v and its neighbours,

$$\mathbb{E}\Big[|I|\Big] = \sum_{v\in V} \mathbb{P}\Big[v \in I\Big] = \sum_{v\in V} \frac{1}{1+\deg(v)}.$$

The function $f(x) \coloneqq 1/(x+1)$ is convex on $[0, \infty)$, so by Jensen's inequality, G contains an independent set I of size

$$|I| \geq \sum_{v\in V} \frac{1}{1+\deg(v)} \geq \frac{n}{1+\frac{1}{n}\sum_{v\in V}\deg(v)} = \frac{n^2}{2m+n}.$$

□

2 Subgraphs with Many Degrees

In this chapter, we present the following result of Bukh and Sudakov.

Theorem 2.1 (Bukh–Sudakov [2])**.** *Every Ramsey graph on n vertices has an induced subgraph on $\Omega(n)$ vertices with $\Omega(\sqrt{n})$ distinct degrees.*

This was originally conjectured by Erdős, Faudree, and Sós [6] in light of the observation that the random graph $G_{n,1/2}$ exhibits such a property almost surely. In particular, Theorem 2.1 already has implications for the edge statistics of Ramsey graphs G: if $G[U]$ is the resulting induced subgraph, then $G[U] - \{v\}$ has $\Omega(\sqrt{n})$ different sizes for $v \in U$.

2.1 Diversity

To prove Theorem 2.1, Bukh and Sudakov introduced the notion of *diversity*, which ensures the neighbourhoods of most pairs of vertices are not too similar.

Definition 2.1. An n-vertex graph G is (ε, δ)-*diverse* if for each $x \in V(G)$, we have $|N(x) \triangle N(y)| \geq \varepsilon n$ for all but $\leq n^\delta$ vertices $y \in V(G)$.

Diversity matches our assumption $|N(x) \triangle N(y)| = D \geq \varepsilon n$ from Section 1.1. Hence, in a diverse graph we have $\mathbb{P}[\deg_U(x) = \deg_U(y)] = O(1/\sqrt{n})$ for most pairs of vertices $x, y \in V(G)$, and Turán's theorem then yields many distinct degrees. The following lemma formalizes this observation and reduces Theorem 2.1 to showing that every Ramsey graph contains a diverse subgraph of linear size.

Lemma 2.2. *Let $\delta \in (0, 1/2)$ and let G be an (ε, δ)-diverse graph on n vertices for large enough n. Then there is a subgraph on $\geq n/4$ vertices containing vertices of $\Omega(\sqrt{n})$ distinct degrees.*

Proof. Let U be a uniformly random subset of $V(G)$. It suffices to show that with positive probability, $|U| \geq n/4$ and $G[U]$ contains $\Omega(\sqrt{n})$ vertices of distinct degrees.

For distinct $x, y \in V(G)$, define $\mathbb{1}_{xy}$ to be the indicator for the event that $x, y \in U$ and $\deg_U(x) = \deg_U(y)$. Then $\mathbb{E}[\mathbb{1}_{xy}] \leq \mathbb{P}[\deg_U(x) = \deg_U(y)]$. Put $D \coloneqq N_G(x) \triangle N_G(y), A \coloneqq N_G(x) \setminus N_G(y)$, and $B \coloneqq N_G(y) \setminus N_G(x)$. Then the random variable

$$\deg_U(x) - \deg_U(y) = |A \cap U| - |B \cap U| = \sum_{v \in A} \mathbb{1}_{v \in U} - \sum_{v \in B} \mathbb{1}_{v \in U} \leq |D|$$

is of $(|D|, 1/2)$-Littlewood–Offord type. Thus, Erdős–Littlewood–Offord gives

$$\mathbb{P}\Big[\deg_U(x) = \deg_U(y)\Big] = O\Big(1/\sqrt{|D|}\Big).$$

In particular, $\mathbb{E}[\mathbb{1}_{xy}] = O(1/\sqrt{n})$, whenever $|D| \geq \varepsilon n$.

Now consider the degree collision graph H on U, where $x, y \in U$ are adjacent if and only if $\deg_U(x) = \deg_U(y)$. Since G is (ε, δ)-diverse, $\mathbb{E}[\mathbb{1}_{xy}] = O(1/\sqrt{n})$ for all but $\leq n^{1+\delta} < n^{3/2}$ pairs $x, y \in V(G)$. Hence, the expected number of edges in H is

$$\mathbb{E}[e(H)] = \sum_{\{x,y\} \in \binom{V(G)}{2}} \mathbb{E}[\mathbb{1}_{xy}] \leq n^{1+\delta} + n^2 \cdot O\Big(1/\sqrt{n}\Big) = O\Big(n^{3/2}\Big)$$

and Markov's inequality yields $\mathbb{P}[e(H) = O(n^{3/2})] \geq 1/2$. Since $|U| < n/4$ with probability $o(1)$ by the Chernoff bound, the union bound yields a positive probability that $|U| \geq n/4$ and $e(H) \leq O(n^{3/2})$ occur simultaneously, in which case Turán's theorem (Proposition 1.9) gives an independent set I in H of size at least

$$|I| \geq \frac{|U|^2}{2e(H) + |U|} = \frac{(n/4)^2}{O(n^{3/2}) + n/4} = \Omega\Big(\sqrt{n}\Big).$$

In particular, each vertex in I has a distinct degree in $G[U]$. □

2.2 Richness

It remains to show that every Ramsey graph contains a diverse subgraph of linear size. We prove the same statement on a strengthening of diversity called *richness*.

Definition 2.2. An n-vertex graph G is (ε, δ)-*rich* if, for $W \subseteq V(G)$ with $|W| \geq \delta n$, we have $|N(v) \cap W| \geq \varepsilon |W|$ and $|\overline{N(v)} \cap W| \geq \varepsilon |W|$ for all but $\leq n^{\delta}$ vertices $v \in V(G)$.

That is, almost every vertex $v \in V(G)$ satisfies

$$\varepsilon |W| \leq |N(v) \cap W| \leq (1 - \varepsilon)|W|,$$

whenever $|W|$ is large. Thus, richness ensures neighbourhoods are not unusually concentrated in nor absent from large sets, which mirrors the density control in random graphs like $G_{n,1/2}$, where $|N(v) \cap W| \approx |W|/2$ by the Chernoff bound. When W is the neighbourhood of another vertex, richness ensures one side of the symmetric difference $|N(x) \triangle N(y)| = |N(x) \cap \overline{N(y)}| + |\overline{N(x)} \cap N(y)|$ is large, implying diversity.

Lemma 2.3. *Let an n-vertex graph G be (ε, δ)-rich, with $\delta \leq 1/2$. Then G is $(\varepsilon/2, \delta)$-diverse.*

Proof. For each $x \in V(G)$, either $|N(x)| \geq n/2$ or $|\overline{N(x)}| \geq n/2$. If $|N(x)| \geq n/2$, then richness gives

$$|\overline{N(y)} \cap N(x)| \geq \varepsilon |N(x)| \geq \varepsilon n/2$$

for all but $\leq n^{\delta}$ vertices $y \in V(G)$. Similarly, if $|\overline{N(x)}| \geq n/2$, then

$$|N(y) \cap \overline{N(x)}| \geq \varepsilon |\overline{N(x)}| \geq \varepsilon n/2$$

for all but $\leq n^{\delta}$ vertices $y \in V(G)$. But then in either case,

$$|N(x) \triangle N(y)| = |\overline{N(y)} \cap N(x)| + |N(y) \cap \overline{N(x)}| \geq \varepsilon n/2.$$

□

Richness also gives the reversed notion of diversity.

Lemma 2.4. *Let an n-vertex graph G be (ε, δ)-rich. Then G has $\leq n^{1+\delta}$ pairs of vertices $x_1, x_2 \in V(G)$ with $|N(x_1) \triangle \overline{N(x_2)}| < (\varepsilon/2)n$.*

Proof. We show that for each $x_1 \in V(G)$, there are $\leq n^{\delta}$ vertices $x_2 \in V(G)$ with $|N(x_1) \triangle \overline{N(x_2)}| < (\varepsilon/2)n$. Since either $|N(x_1)| \geq n/2$ or $|\overline{N(x_1)}| \geq n/2$, we assume without loss of generality that $|N(x_1)| \geq n/2$. By (ε, δ)-richness, there are $\leq n^{\delta}$ vertices x_2 with $|N(x_2) \cap N(x_1)| < \varepsilon n/2$. For all other x_2,

$$|N(x_1) \triangle \overline{N(x_2)}| \geq |N(x_1) \cap N(x_2)| \geq (\varepsilon/2)n.$$

The result now follows from the fact that there are only n choices of x_1. $\square$

Remark: Since $|N(x_1) \triangle N(x_2)| + |N(x_1) \triangle \overline{N(x_2)}| = n$ for any $x_1, x_2 \in V(G)$, Lemma 2.4 gives the upper bound

$$|N(x_1) \triangle N(x_2)| \leq (1 - \varepsilon/2)n$$

for most pairs $x_1, x_2 \in V(G)$.

We now show that every Ramsey graph contains a rich subgraph of linear size. Specifically, we show that if this is not the case, we may inductively construct a sequence of small vertex sets $S_1, S_2, \ldots, S_K$ that are sparse/dense between one another, and the union S of some of these sets induces a sparse/dense subgraph, which contains a large homogeneous set by the Erdős–Szemerédi theorem.

Lemma 2.5. *For $\delta \in (0, 1)$, there exists $\varepsilon = \varepsilon(C) > 0$ such that every C-Ramsey graph G on n vertices contains an (ε, δ)-rich subgraph on $\Omega(n)$ vertices.*

Remark: $\varepsilon = \varepsilon(C)$ is independent of δ, so δ can be taken arbitrarily small for any given ε.

Proof. Suppose for contradiction that every set of $\geq cn$ vertices fails to induce a (ε, δ)-rich subgraph, for small $\varepsilon = \varepsilon(C)$ and $c = (\delta/4)^K$ with large $K = K(C, \delta)$ to be determined. We inductively construct a sequence of disjoint vertex sets $S_1, S_2, \ldots, S_K$ of size $s := (cn)^{\delta}/2$, such that for all S_i one of the following conditions holds:

(i) $d(S_i, S_j) < 4\varepsilon$ for $j > i$,

(ii) $d(S_i, S_j) > 1 - 4\varepsilon$ for $j > i$.

That is, each S_i is small, and the sets are extremely sparse or dense between one another. Alongside, we also construct a sequence of induced subgraphs $G = G[U_0] \supset G[U_1] \supset \cdots \supset G[U_K]$, such that $|U_i| \geq (\delta/4)|U_{i-1}|$ and $S_i \subseteq U_{i-1}$ for all i.

We first show how this construction gives the lemma. By working with $\overline{G}$ instead of G if necessary, we may assume that $r \geq K/2$ of the sets $S_1, \ldots, S_K$ satisfy condition (i), say $S_{i_1}, \dots, S_{i_r}$. Put $S \coloneqq \bigsqcup_{k=1}^r S_{i_k}$. Since each S_{i_k} has size s,

$$d(S) = \frac{\sum_{k=1}^r e(S_{i_k})}{\binom{|S|}{2}} + \frac{\sum_{k<l} e(S_{i_k}, S_{i_l})}{\binom{|S|}{2}} \le \frac{r\binom{s}{2}}{\binom{rs}{2}} + \max_{(k,l)} d(S_{i_k}, S_{i_l}) \le \frac{1}{r} + 4\varepsilon \le \frac{2}{K} + 4\varepsilon.$$

Set $d \coloneqq 2/K + 4\varepsilon$, which is a positive constant to be taken small. The Erdős–Szemerédi theorem (Theorem 1.3) now yields a homogeneous set $A \subseteq S$ of size

$$|A| \ge \frac{\gamma}{d\log(1/d)} \log|S| = \frac{\gamma\delta}{d\log(1/d)} \log cn = \frac{\gamma\delta}{d\log(1/d)} \log n + o(\log n),$$

for some absolute constant $\gamma > 0$ given by the theorem. Take d to be small enough so that $\frac{\gamma\delta}{d\log(1/d)} > C$. Then $|A| > C\log n$, contradiction.

We now construct the sets S_i and U_i inductively. Suppose $U_1, \dots, U_{i-1}$ and $S_1, \dots, S_{i-1}$ have been constructed. By induction, $|U_{i-1}| \ge (\delta/4)^{i-1}|U_0| \ge cn$. Since $G[U_{i-1}]$ is not (ε, δ)-rich by assumption, U_{i-1} contains subsets W_i, Y_i with $|W_i| \ge \delta|U_{i-1}|$ and $|Y_i| = (cn)^\delta$, such that for each $y \in Y_i$ we have $|N_{G_{i-1}}(y) \cap W_i| < \varepsilon|W_i|$ or $|\overline{N_{G_{i-1}}(y)} \cap W_i| < \varepsilon|W_i|$. We consider two cases.

Suppose $|N_{G_{i-1}}(y) \cap W_i| < \varepsilon|W_i|$ for half the vertices $y \in Y_i$, and let S_i be the corresponding subset of Y_i, so $|S_i| = (cn)^\delta/2$. Put $T_i \coloneqq W_i \setminus S_i$. Since $|S_i| \le |W_i|/2$ for large enough n, we have $|T_i| \ge |W_i|/2$. Thus for $y \in S_i$,

$$d(y, T_i) = \frac{|N_{G_{i-1}}(y) \cap T_i|}{|T_i|} \le \frac{|N_{G_{i-1}}(y) \cap W_i|}{|T_i|} < \frac{\varepsilon|W_i|}{|W_i|/2} = 2\varepsilon.$$

Partition T_i into $U_i \sqcup Z_i$ where $U_i \coloneqq \{v \in T_i : d(v, S_i) < 4\varepsilon\}$. We show that $|U_i| \ge (\delta/4)|U_{i-1}|$ by double counting $e(T_i, S_i)/|S_i|$. Since $d(v, S_i) \ge 4\varepsilon$ for $v \in Z_i$,

$$\frac{e(T_i, S_i)}{|S_i|} = \sum_{v \in T_i} d(v, S_i) \ge \sum_{v \in Z_i} d(v, S_i) \ge 4\varepsilon|Z_i|.$$

On the other hand, since $d(y, T_i) < 2\varepsilon$ for $y \in S_i$,

$$\frac{e(T_i, S_i)}{|S_i|} = \frac{1}{|S_i|} \sum_{y \in S_i} e(y, T_i) = \frac{|T_i|}{|S_i|} \sum_{y \in S_i} d(y, T_i) < 2\varepsilon|T_i|.$$

Combining both inequalities yields $|Z_i| < |T_i|/2$. Therefore

$$|U_i| = |T_i| - |Z_i| > \frac{|T_i|}{2} \ge \frac{|W_i|}{4} \ge \frac{\delta}{4}|U_{i-1}|.$$

For $j > i$, since $S_j \subseteq U_{j-1} \subseteq U_i$, we have $d(v, S_i) < 4\varepsilon$ for every $v \in S_j$, and averaging gives $d(S_i, S_j) < 4\varepsilon$. Thus S_i satisfies condition (i).

Otherwise, $|\overline{N_{G_{i-1}}(y)} \cap W_i| < \varepsilon |W_i|$ for half the vertices $y \in Y_i$. Let S_i be the corresponding subset of Y_i and put $T_i \coloneqq W_i \setminus S_i$, so again $|T_i| \geq |W_i|/2$ and $|S_i| = (cn)^\delta/2$. For $y \in S_i$,

$$1 - d(y, T_i) = \frac{|\overline{N_{G_{i-1}}(y)} \cap T_i|}{|T_i|} \leq \frac{|\overline{N_{G_{i-1}}(y)} \cap W_i|}{|T_i|} < \frac{\varepsilon |W_i|}{|W_i|/2} = 2\varepsilon,$$

so $d(y, T_i) > 1 - 2\varepsilon$ for every $y \in S_i$. Partition T_i into $U_i \sqcup Z_i$ where $U_i \coloneqq \{v \in T_i : d(v, S_i) > 1 - 4\varepsilon\}$. We double count $|T_i| - e(T_i, S_i)/|S_i| = \sum_{v \in T_i}(1 - d(v, S_i))$. Since $1 - d(v, S_i) \geq 4\varepsilon$ for $v \in Z_i$,

$$|T_i| - \frac{e(T_i, S_i)}{|S_i|} = \sum_{v \in T_i} \Big(1 - d(v, S_i)\Big) \geq \sum_{v \in Z_i} \Big(1 - d(v, S_i)\Big) \geq 4\varepsilon |Z_i|.$$

On the other hand, using $d(y, T_i) > 1 - 2\varepsilon$ for $y \in S_i$,

$$|T_i| - \frac{e(T_i, S_i)}{|S_i|} = |T_i| - \frac{|T_i|}{|S_i|} \sum_{y \in S_i} d(y, T_i) < |T_i| - (1 - 2\varepsilon)|T_i| = 2\varepsilon |T_i|.$$

Combining both inequalities gives $|Z_i| < |T_i|/2$, hence $|U_i| > |T_i|/2 \geq (\delta/4)|U_{i-1}|$ as in the other case. For $j > i$, each $v \in S_j \subseteq U_i$ satisfies $d(v, S_i) > 1 - 4\varepsilon$, so $d(S_i, S_j) > 1 - 4\varepsilon$. Thus S_i satisfies condition (ii). This completes the induction. $\square$

2.3 Proof of Theorem 2.1

Theorem 2.1 now follows immediately by applying Lemmas 2.5, 2.3, and 2.2 in sequence.

Let G be a C-Ramsey graph on n vertices, and fix any $\delta \in (0, 1/2)$. By Lemma 2.5, there exists an (ε, δ)-rich induced subgraph $G' \subseteq G$ on $m = \Omega(n)$ vertices. By Lemma 2.3, G' is $(\varepsilon/2, \delta)$-diverse. Applying Lemma 2.2 to G', we obtain an induced subgraph of G' on $\Omega(m) = \Omega(n)$ vertices containing $\Omega(\sqrt{m}) = \Omega(\sqrt{n})$ vertices of distinct degrees. $\square$

3 Subgraphs of Quadratically Many Sizes

Erdős and McKay conjectured that for any Ramsey graph G on n vertices, there exists $\delta = \delta(C) > 0$ such that

$$\Big\{1, 2, \ldots, \delta n^2\Big\} \subseteq \Phi(G) := \Big\{e(H) : H \text{ is an induced subgraph of } G\Big\}.$$

That is, for any $1 \leq t \leq \delta n^2$, there exists an induced subgraph with exactly t edges. The conjecture was widely open for decades. In 2003, Alon, Krivelevich and Sudakov [1] proved the conjecture with n^δ in place of δn^2.

Theorem 3.1 (Alon, Krivelevich and Sudakov [1])**.** *Fix $C > 0$. For any C-Ramsey graph G on n vertices,*

$$\Big\{1, 2, \ldots, n^\delta\Big\} \subseteq \Phi(G),$$

for some $\delta = \delta(C) > 0$.

Since the conjecture implies $|\Phi(G)| = \Omega(n^2)$ and trivially $|\Phi(G)| \leq \binom{n}{2}$, a natural relaxation is to ask whether $\Phi(G)$ indeed has quadratic size. In 2017, Narayanan, Sahasrabudhe, and Tomon [13] proved a near optimal bound of $|\Phi(G)| = n^{2-o(1)}$ by building on the notion of diversity from Chapter 2. The desired quadratic bound was achieved by Kwan and Sudakov [10] a few years later, and this chapter is dedicated to presenting their proof.

Theorem 3.2 (Kwan and Sudakov [10])**.** *For any $C > 0$, every C-Ramsey graph G on n vertices has*

$$|\Phi(G)| = \Theta(n^2).$$

Recall that Theorem 2.1 provides a bound of $|\Phi(G)| = \Omega(\sqrt{n})$ for any Ramsey graph G by removing a single vertex to construct each subgraph. Two ideas may be used to improve this construction. First, we apply the same idea to many different induced subgraphs rather than merely G itself. Second, instead of deleting only one vertex, we add and remove various combinations of vertices to produce many more distinct sizes.

Specifically, pick a target m and find two disjoint vertex sets U and W, where U is a linear-sized random set with $e(U) \approx m$ and W is an augmenting set of size $O(\sqrt{n})$ selected from a *candidate set* $W_0 \subseteq V(G) \setminus U$ whose degrees are of linear order and all lie in a narrow interval. Then the subsets $Z \subseteq W$ produce $\Omega(n^{3/2})$ distinct values of $e(U \sqcup Z)$ that all lie in an interval of length $O(n^{3/2})$ centred at $e(U)$. Applying this to $\Omega(\sqrt{n})$ well-separated targets m yields the full $\Omega(n^2)$ count.

Kwan and Sudakov construct W as three disjoint parts $\mathbf{S}$, $\mathbf{T}$, $\mathbf{X}$, each of size $\Theta(\sqrt{n})$, that act at complementary scales. Every element has $\Theta(n)$ edges into U, and the degrees from $\mathbf{S}$ into U are uniformly smaller than those from $\mathbf{T}$ by a gap of $\Omega(\sqrt{n})$. Three operations on $Z \subseteq W$ shift $e(U \sqcup Z)$ at three separate scales. Varying $k := |Z \cap (\mathbf{S} \sqcup \mathbf{T})|$ shifts $e(U \sqcup Z)$ at the coarse scale $\Theta(n)$. With k fixed, exchanging an element of $\mathbf{S}$ for an element of $\mathbf{T}$ shifts $e(U \sqcup Z)$ by $\Theta(\sqrt{n})$, so the achievable values at this level are spaced by $\Theta(\sqrt{n})$. The set $\mathbf{X}$ will contain $\Theta(\sqrt{n})$ distinct degrees into U, so varying the singleton $x \in \mathbf{X}$ fills these $\Theta(\sqrt{n})$ gaps. Hence, the different combinations of $\mathbf{S}, \mathbf{T}, \mathbf{X}$ yield $\Omega(n^{3/2})$ distinct values of $e(U \sqcup Z)$.

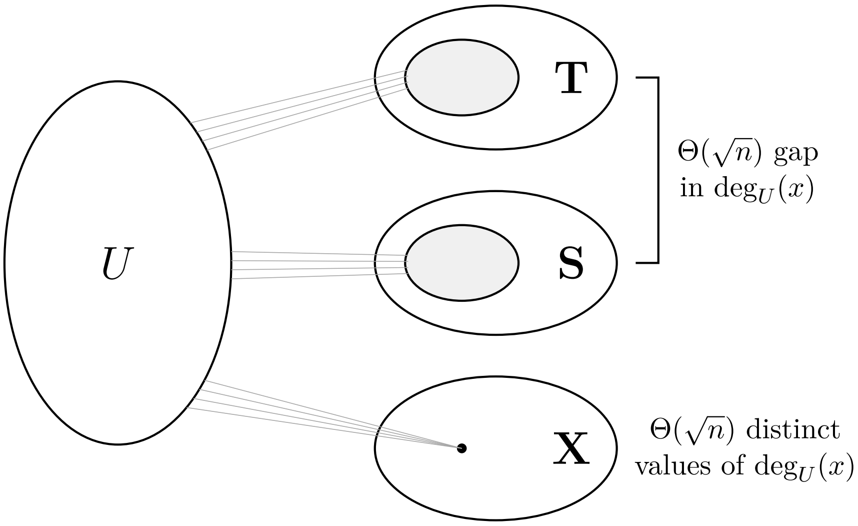

Figure 1: Construction of $U \sqcup Z$. The gray areas and singleton in $\mathbf{X}$ constitute Z, and $\deg_U(x) = \Theta(n)$ for all $x \in \mathbf{S} \sqcup \mathbf{T} \sqcup \mathbf{X}$.

Achieving the requirement for $\mathbf{X}$ is the main technical challenge, as vertex diversity defined in Chapter 2 does not produce enough distinct degrees in this case. In the proof of Lemma 2.2, we saw that $\mathbb{P}[\deg_U(x) = \deg_U(y)] = O(1/\sqrt{n})$ for all but $\le n^\delta$ pairs $x, y \in V(G)$ in an (ε, δ)-diverse graph, and the same construction yields the collision graph H on the candidate set W_0 with expected number of edges

$$\mathbb{E}[e(H)] = O\Big(|W_0|^2/\sqrt{n} + |W_0| \cdot n^\delta\Big),$$

producing an independent set $I \subseteq W_0$ of size

$$|I| = \frac{|W_0|^2}{2\mathbb{E}[e(H)] + |W_0|} = \Omega\left(\sqrt{n} \cdot \frac{|W_0|}{|W_0| + n^{1/2+\delta}}\right).$$

by Turán's theorem. We thus need $|W_0| = \Omega(n^{1/2+\delta})$ for $|I|$ to reach $\Omega(\sqrt{n})$. Unfortunately, W_0 is extracted by pigeonholing n vertices into $O(\sqrt{n})$ degree intervals of length $O(\sqrt{n})$, so at best $|W_0| = \Theta(\sqrt{n})$, yielding only $|I| = \Omega(n^{1/2-\delta})$.

To overcome this obstacle, Kwan and Sudakov's construction allows the candidate set W_0 to also contain disjoint *pairs* while treating them like singletons. Write $\mathbf{x}$ in bold to denote an object that is either a single vertex or a pair of vertices, and similarly for sets of such objects. For $\mathbf{x} = \{x_1, x_2\}$, define $N(\mathbf{x}) := N(x_1) \cup N(x_2)$ as a multiset union and $\deg(\mathbf{x}) := |N(\mathbf{x})| = \deg(x_1) + \deg(x_2)$ as the *pair-degree*. Note that there are $\binom{n}{2} = \Theta(n^2)$ pairs, but the pair-degrees take only $O(n)$ values.

Pigeonholing now yields $\Theta(n)$ disjoint pairs sharing a common pair-degree, vastly exceeding the $n^{1/2+\delta}$ threshold. The price is that we need the symmetric difference $|N(\mathbf{x})\triangle N(\mathbf{y})| = \Omega(n)$ not just for most pairs of vertices, but for most pairs of *pairs*. We call this *pair-diversity*.

Definition 3.1. An n-vertex graph G is $(\varepsilon, \delta, \alpha)_2$-*diverse* if for each pair of vertices $\mathbf{x} = \{x_1, x_2\}$ in G with $|N(x_1)\triangle\overline{N(x_2)}| \geq \alpha n$, there are $\leq n^\delta$ mutually disjoint pairs of vertices $\mathbf{y} = \{y_1, y_2\}$ with $\mathbf{x} \cap \mathbf{y} = \emptyset$, such that $|N(\mathbf{x})\triangle N(\mathbf{y})| < \varepsilon n$.

It turns out that richness also implies pair-diversity.

Lemma 3.3. *Let an n-vertex graph G be (ε, δ)-rich, with $\delta \leq 1/2$. Then G is $(\alpha\varepsilon/2, \delta, \alpha)_2$-diverse for any $\alpha \geq 2\delta$.*

Proof. Suppose not. Then there exists a pair $\mathbf{x} = \{x_1, x_2\}$ with $|N(x_1)\triangle\overline{N(x_2)}| \geq \alpha n$ and a set Y of $> n^\delta$ pairs with $|N(\mathbf{x})\triangle N(\mathbf{y})| < \alpha\varepsilon n/2$ for all $\mathbf{y} \in Y$. Since

$$|N(x_1)\triangle\overline{N(x_2)}| = |N(x_1) \cap N(x_2)| + |\overline{N(x_1)} \cap \overline{N(x_2)}| \geq \alpha n,$$

we have $|N(x_1) \cap N(x_2)| \geq (\alpha/2)n \geq \delta n$ or $|\overline{N(x_1)} \cap \overline{N(x_2)}| \geq (\alpha/2)n \geq \delta n$. Assume without loss of generality that the former case holds. Then for each vertex y in any $\mathbf{y} \in Y$,

$$\left|\overline{N(y)} \cap \Big(N(x_1) \cap N(x_2)\Big)\right| \leq |N(\mathbf{x})\triangle N(\mathbf{y})| < \alpha\varepsilon n/2 \leq \varepsilon|N(x_1) \cap N(x_2)|.$$

But then the set of all such y now contradicts the (ε, δ)-rich condition. $\square$

One last thing we need is to ensure $\mathbf{X}$ actually provides $\Theta(\sqrt{n})$ distinct degrees into $U \sqcup \mathbf{Z} := U \sqcup \bigsqcup_{\mathbf{z}\in\mathbf{Z}} \mathbf{z}$ for each choice of augmenting set $\mathbf{Z} \subseteq \mathbf{S} \sqcup \mathbf{T}$. Notice the current construction of $\mathbf{X}$ only ensures distinct degrees $\deg_U(\mathbf{x})$ into U instead of $\deg_{U\sqcup Z}(\mathbf{x})$. If we sampled U once and directly constructed $\mathbf{S}, \mathbf{T}, \mathbf{X}$ based on U, then the $\deg_U(\mathbf{x})$ would be fixed numbers given the outcome of U. However, there is no guarantee that $\deg_{U\sqcup Z}(\mathbf{x}) = \deg_U(\mathbf{x}) + \deg_Z(\mathbf{x})$ would remain distinct, as the addition of $\deg_Z(\mathbf{x})$ could cancel out the uniqueness of $\deg_U(\mathbf{x})$. Kwan and Sudakov overcome this with a *double-exposure* technique. They perform two rounds of sampling. First, they sample the random set U_0 and construct $\mathbf{S}, \mathbf{T}, \mathbf{X}$ relative to U_0. In the second round, they sample U from U_0 by independently keeping each vertex with probability 1/2. After the first sampling, the construction guarantees $|N_{U_0}(\mathbf{x})\triangle N_{U_0}(\mathbf{y})| = \Omega(n)$ for all $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, and the augmenting sets $\mathbf{Z} \subseteq \mathbf{S} \cup \mathbf{T}$ are fully determined by U_0. In the second round, the difference $\deg_U(\mathbf{x}) - \deg_U(\mathbf{y})$ becomes a sum of $\Omega(n)$ independent Bernoulli terms, and Erdős–Littlewood–Offord yields $\mathbb{P}[\deg_U(\mathbf{x}) - \deg_U(\mathbf{y}) = t] = O(1/\sqrt{n})$ for every integer t. In particular, since $t_{xy} := \deg_Z(\mathbf{y}) - \deg_Z(\mathbf{x})$ is deterministic for each $\mathbf{x}, \mathbf{y}$,

$$\mathbb{P}\Big[\deg_{U\sqcup Z}(\mathbf{x}) = \deg_{U\sqcup Z}(\mathbf{y})\Big] = \mathbb{P}\Big[\deg_U(\mathbf{x}) - \deg_U(\mathbf{y}) = t_{xy}\Big] = O(1/\sqrt{n}).$$

The collision graph argument now gives the desired set of U such that $\deg_{U\sqcup Z}(\mathbf{x})$ is distinct for $\mathbf{x} \in \mathbf{X}$.

3.1 The main lemma

As outlined above, we will find many induced subgraphs $G[U]$ whose edge counts are well-separated from each other, and show that each can be augmented in $\Omega(n^{3/2})$ ways. Theorem 3.2 will be an immediate consequence of the following lemma, which we dedicate the rest of this chapter to proving.

Lemma 3.4. *Let $C > 0$. Then there exists $c = c(C) > 0$, such that for any C-Ramsey graph G on n vertices and any m satisfying $cn^2 \leq m \leq 2cn^2$, there are disjoint subsets $U, W \subseteq V(G)$ with $|e(U) - m| = O(n^{3/2})$ and $|W| = O(\sqrt{n})$ such that*

$$\left|\left\{e(U \sqcup Z) : Z \subseteq W\right\}\right| = \Omega(n^{3/2}).$$

We first show how this implies Theorem 3.2.

Proof of Theorem 3.2. For any U, W given by Lemma 3.4,

$$|e(U \sqcup W) - e(U)| = O(n^{3/2}),$$

as $|W| = O(\sqrt{n})$ and each $w \in W$ has $\leq n$ neighbours in $U \sqcup W$. But then the $\Omega(n^{3/2})$ values of $e(U \sqcup Z)$ for $Z \subseteq W$ are contained in an interval of length $O(n^{3/2})$ centred at $e(U)$. Partition the interval between cn^2 and $2cn^2$ into $\frac{c}{k}\sqrt{n}$ subintervals of length $kn^{3/2}$, for some large constant k. Apply Lemma 3.4 to a suitable choice of m in each subinterval so that the obtained values of $e(U \sqcup Z)$ for $Z \subseteq W$ lie in the same subinterval as m, which is allowed for large enough k. Since we applied Lemma 3.4 to $\Omega(\sqrt{n})$ disjoint subintervals, we obtain $\Omega(n^2)$ distinct sizes of induced subgraphs. $\square$

3.2 The construction of augmenting sets

The following lemma constructs U_0 and the augmenting sets $\mathbf{S}, \mathbf{T}, \mathbf{X}$ with the properties needed for Lemma 3.4.

Lemma 3.5. *For any $C > 0$ there exists $c = c(C) > 0$ such that the following holds. For any C-Ramsey graph G on n vertices and any m satisfying $cn^2 \leq m \leq 2cn^2$, there is vertex set $U_0 \subseteq V(G)$ with $|e(U_0) - 4m| = O(n^{3/2})$, constant $c' = c'(C) \in (0, 1]$, and sets $\mathbf{S}$, $\mathbf{T}$, $\mathbf{X}$ of size $2c'\sqrt{n}$ such that*

1. *either $U_0, \mathbf{S}, \mathbf{T}, \mathbf{X} \subseteq V(G)$ are disjoint sets of vertices, or $\mathbf{S}, \mathbf{T}, \mathbf{X} \subseteq \binom{V(G)}{2}$ are sets of disjoint pairs of vertices with no vertex appearing in more than one of $U_0, \mathbf{S}, \mathbf{T}, \mathbf{X}$,*
2. *there is $d = \Theta(n)$ such that $\deg_{U_0}(\mathbf{x}) = d + O(\sqrt{n})$ for each $\mathbf{x} \in \mathbf{S} \sqcup \mathbf{T} \sqcup \mathbf{X}$,*
3. $\min_{\mathbf{x} \in \mathbf{T}} \deg_{U_0}(\mathbf{x}) - \max_{\mathbf{x} \in \mathbf{S}} \deg_{U_0}(\mathbf{x}) \geq 8c'\sqrt{n}$,
4. *for each $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{X}}{2}$, we have $|N_{U_0}(\mathbf{x}) \triangle N_{U_0}(\mathbf{y})| = \Omega(n)$.*

Property 2 ensures adding or removing a single element shifts the edge count by $\Theta(n)$. Property 3 guarantees swapping an element of $\mathbf{S}$ for an element of $\mathbf{T}$ shifts the edge count by $\Theta(\sqrt{n})$. Property 4 sets up the Littlewood–Offord argument in the double-exposure step.

The proof starts with reducing to a rich subgraph using Lemma 2.5 and using the pigeonhole principle to obtain a large candidate set $\mathbf{W}_0$ of vertices or pairs with similar degrees, satisfying Property 1 and 2. It then samples U_0 at random and verifies that all desired properties hold simultaneously with positive probability.

Proof of Lemma 3.5. Let $\varepsilon = \varepsilon(C)$ be given by Lemma 2.5 and set $\delta := \varepsilon/4$. By Lemma 2.5, G contains an (ε, δ)-rich induced subgraph $G[V']$ on $\Omega(n)$ vertices. Since $\log |V'| \geq \frac{1}{2} \log n$, the graph $G[V']$ is $2C$-Ramsey, so by adjusting constants it suffices to work inside $G[V']$. We may thus assume G itself is (ε, δ)-rich.

By richness, Lemmas 2.3, 2.4, and 3.3 with $\alpha := \varepsilon/2$ furnish the following properties:

(a) G is $(\varepsilon/2, \delta)$-diverse,

(b) G is $(\varepsilon^2/4, \delta, \varepsilon/2)_2$-diverse,

(c) there are $\leq n^{1+\delta}$ pairs $\{x_1, x_2\}$ with $|N(x_1) \triangle \overline{N(x_2)}| < (\varepsilon/2)n$.

Each of the $\binom{n}{2} = \Theta(n^2)$ vertex pairs $\{x_1, x_2\}$ has a pair-degree $\deg(\{x_1, x_2\}) = \deg(x_1) + \deg(x_2)$ taking values from 0 to $2n$, so by pigeonhole there exists d^* and a collection H of $\Omega(n^{3/2})$ pairs with $\deg(\mathbf{x}) = d^* + O(\sqrt{n})$ for all $\mathbf{x} \in H$.

Interpret H as a graph on $V(G)$ whose edges are the $\Omega(n^{3/2})$ pairs. By (c), we may delete $O(n^{1+\delta}) = o(n^{3/2})$ edges in H so that all remaining pairs have $|N(x_1) \triangle \overline{N(x_2)}| \geq (\varepsilon/2)n$. Suppose there exists vertex v in H with degree $\Omega(n^{3/4})$. Take $\mathbf{W}_0 := N_H(v)$ and set $d' := d^* - \deg(v)$. Then each $w \in \mathbf{W}_0$ is a single vertex with degree $\deg(w) = d' + O(\sqrt{n})$. Now suppose H has maximum degree $\Delta = o(n^{3/4})$. By iteratively picking an edge and removing $\leq 2\Delta$ edges adjacent to it, we obtain a matching of size $\Omega(n^{3/2})/o(n^{3/4}) = \Omega(n^{3/4})$. Take $\mathbf{W}_0$ to be the pairs in this matching and set $d' := d^*$. Thus, $\mathbf{W}_0$ is now the candidate set, such that $\deg(\mathbf{x}) = d' + O(\sqrt{n})$ for all $\mathbf{x} \in \mathbf{W}_0$ and all subsets of $\mathbf{W}_0$ satisfy Property 1.

Next, we extract a subset $\mathbf{A} \subseteq \mathbf{W}_0$ of size $\Theta(\sqrt{n})$ with pairwise large neighbourhood symmetric differences. Define graph F on $\mathbf{W}_0$ by joining $\mathbf{x}$ and $\mathbf{y}$ whenever $|N(\mathbf{x}) \triangle N(\mathbf{y})| < (\varepsilon^2/4)n$. By diversity, F has $|\mathbf{W}_0| = \Omega(n^{3/4})$ vertices and maximum degree at most $n^\delta < n^{1/4}$, so Turán's theorem gives an independent set $\mathbf{A} \subseteq \mathbf{W}_0$ in F of size $\Omega\Big(|\mathbf{W}_0|/n^{1/4}\Big) = \Omega(\sqrt{n})$. That is, for $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$, we have $|N(\mathbf{x}) \triangle N(\mathbf{y})| = \Omega(n)$, which implies $d' = \Omega(n)$. Trim $\mathbf{A}$ to size $\Theta(\sqrt{n})$.

By the Erdős–Szemerédi bound (Corollary 1.4), $e(G) \geq 800cn^2$ for some constant $c = c(C) > 0$. For $cn^2 \leq m \leq 2cn^2$, set $p := \sqrt{4m/e(G)}$, so that $p = \Omega(1)$ and $p \leq 0.1$. Let U_0 be a random subset of $V(G)$ by including each vertex with probability p.

Claim 3.6. *The following five events each hold with probability greater than 4/5, and therefore all hold simultaneously with positive probability by the union bound.*

(i) $|e(U_0) - 4m| = O(n^{3/2})$,

(ii) *There is a set* $\mathbf{Q} \subseteq \mathbf{A}$ *involving no vertices of* U_0, *with* $|\mathbf{Q}| \geq (2/3)|\mathbf{A}|$,

(iii) *There is a set* $\mathbf{R} \subseteq \mathbf{A}$ *with* $|\mathbf{R}| \geq (2/3)|\mathbf{A}|$ *and* $\deg_{U_0}(\mathbf{x}) = pd' + O(\sqrt{n})$ *for each* $\mathbf{x} \in \mathbf{R}$,

(iv) $|N_{U_0}(\mathbf{x}) \triangle N_{U_0}(\mathbf{y})| = \Omega(n)$ *for each* $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$,

(v) *The equality* $\deg_{U_0}(\mathbf{x}) = \deg_{U_0}(\mathbf{y})$ *holds for* $O(\sqrt{n})$ *pairs* $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$.

Proof of Claim. We prove each part in order.

(i) $\mathbb{E}[e(U_0)] = 4m$ and there are $O(n^3)$ pairs of edges in G whose presence are dependent in $G[U_0]$, as this occurs only if two edges share a vertex. Thus for $e \in E(G)$, the standard variance bound gives

$$\mathrm{Var}\Big(e(U_0)\Big) \leq \mathbb{E}[e(U_0)] + O(n^3) = O(n^3).$$

Chebyshev's inequality yields

$$\mathbb{P}\Big[|e(U_0) - 4m| = \omega(n^{3/2})\Big] \leq \frac{O(n^3)}{\omega(n^3)} \to 0,$$

as $n \to \infty$.

(ii) Let $\mathbf{Q}$ be the subset of $\mathbf{A}$ involving no vertices of U_0. Then $|\mathbf{Q}|$ follows a binomial distribution with $|\mathbf{A}|$ trials, which has mean at least $(1-p)^2|\mathbf{A}| \geq 0.8|\mathbf{A}|$ and variance $O(|\mathbf{A}|)$. By Chebyshev's,

$$\mathbb{P}\left[|\mathbf{Q}| < \frac{2}{3}|\mathbf{A}|\right] \leq \mathbb{P}\Big[\mathbb{E}[|\mathbf{Q}|] - |\mathbf{Q}| \geq 0.1|\mathbf{A}|\Big] \leq \frac{O\big(|\mathbf{A}|\big)}{0.01|\mathbf{A}|^2} = O\Big(1/|\mathbf{A}|\Big) \to 0,$$

as $n \to \infty$.

(iii) For $\mathbf{x} \in \mathbf{A}$, the degree $\deg_{U_0}(\mathbf{x})$ has mean $pd' + O(\sqrt{n})$ and variance $O(n)$. By Chebyshev's,

$$\begin{aligned}\mathbb{P}\Big[\deg_{U_0}(\mathbf{x}) = pd' + O(\sqrt{n})\Big] &\geq 1 - \mathbb{P}\Big[|\deg_{U_0}(\mathbf{x}) - \mathbb{E}[\deg_{U_0}(\mathbf{x})]| = \omega(\sqrt{n})\Big] \\ &\geq 1 - \frac{O(n)}{\omega(n)} \to 1,\end{aligned}$$

In particular, $\deg_{U_0}(\mathbf{x}) = pd' + O(\sqrt{n})$ with probability 0.99 for large enough n. Let $\mathbf{R} \subseteq \mathbf{A}$ be the set of $\mathbf{x}$ that satisfies this bound. Since $\mathbb{E}[|\mathbf{A} \setminus \mathbf{R}|] = 0.01|\mathbf{A}|$, Markov's inequality gives

$$\mathbb{P}\left[|\mathbf{R}| \geq \frac{2}{3}|\mathbf{A}|\right] = 1 - \mathbb{P}\left[|\mathbf{A} \setminus \mathbf{R}| \geq \frac{1}{3}|\mathbf{A}|\right] \geq 1 - \frac{\mathbb{E}\Big[|\mathbf{A} \setminus \mathbf{R}|\Big]}{\frac{1}{3}|\mathbf{A}|} \geq 0.8.$$

(iv) Fix $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$. Recall that $|N(\mathbf{x}) \triangle N(\mathbf{y})| = \Omega(n)$. Since $|N_{U_0}(\mathbf{x}) \triangle N_{U_0}(\mathbf{y})| = |(N(\mathbf{x}) \triangle N(\mathbf{y})) \cap U_0|$ has binomial distribution with parameters $|N(\mathbf{x}) \triangle N(\mathbf{y})|$ and p, Chernoff bound gives

$$\mathbb{P}\left[|N_{U_0}(\mathbf{x}) \triangle N_{U_0}(\mathbf{y})| < \frac{p}{2}|N(\mathbf{x}) \triangle N(\mathbf{y})|\right] = e^{-\Omega(n)} = o(n^{-1}).$$

The result follows from summing over $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$ and invoking the union bound.

(v) For $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$, the random variable $\deg_{U_0}(\mathbf{x}) - \deg_{U_0}(\mathbf{y})$ is of $(|N(\mathbf{x}) \triangle N(\mathbf{y})|, p)$-Littlewood–Offord type. Since $|N(\mathbf{x}) \triangle N(\mathbf{y})| = \Omega(n)$, the Littlewood–Offord theorem gives $\mathbb{P}[\deg_{U_0}(\mathbf{x}) = \deg_{U_0}(\mathbf{y})] = O(1/\sqrt{n})$, so the expected number of pairs $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{A}}{2}$ satisfying $\deg_{U_0}(\mathbf{x}) = \deg_{U_0}(\mathbf{y})$ is $O(\sqrt{n})$. The property now follows from Markov's inequality.

□

Fix an outcome of U_0 for which all five events hold. Since $|\mathbf{Q} \cap \mathbf{R}| \geq |\mathbf{A}|/3 = \Omega(\sqrt{n})$, the collision graph construction on $\mathbf{Q} \cap \mathbf{R}$ and Turán's theorem yield a set $\mathbf{B}$ of size $\Omega(\sqrt{n})$ that has distinct degrees into U_0. Partition $\mathbf{B}$ into three equal parts by the degree $\deg_{U_0}(\mathbf{x})$ into U_0, and let $\mathbf{T}$ be the top third, $\mathbf{S}$ be the bottom third, and $\mathbf{X}$ be the middle third. Since $\mathbf{B}$ has distinct degrees and $|\mathbf{X}| = \Omega(\sqrt{n})$, we may find constant $c' \in (0, 1]$ that satisfies Property 3. Since $\mathbf{X} \subseteq \mathbf{A}$, Property 4 holds. Set $d := pd'$. Recall that $d' = \Theta(n)$, so $\deg_{U_0}(\mathbf{x}) = d + O(\sqrt{n})$ for each $\mathbf{x} \in \mathbf{S} \sqcup \mathbf{T} \sqcup \mathbf{X}$, satisfying Property 2. Trimming the sizes of $\mathbf{S}, \mathbf{T}, \mathbf{X}$ to $2c'\sqrt{n}$ completes the proof. □

3.3 Double-exposure

Given $\mathbf{S}, \mathbf{T}, c'$ from Lemma 3.5, fix an ordering on $\mathbf{S}$ and $\mathbf{T}$ and define

$$\mathcal{P} := \left\{(k, i) : c'\sqrt{n} \leq k \leq 2c'\sqrt{n} \text{ and } 0 \leq i \leq c'\sqrt{n}\right\}.$$

For $(k, i) \in \mathcal{P}$, define $\mathbf{Z}_{k,i} \subseteq \mathbf{S} \sqcup \mathbf{T}$ by taking the first $k - i$ elements from $\mathbf{S}$ and the first i elements from $\mathbf{T}$. By Properties 2 and 3 of Lemma 3.5, the values of $e(U_0, \mathbf{Z}_{k,i}) := \sum_{\mathbf{z} \in \mathbf{Z}_{k,i}} \deg_{U_0}(\mathbf{z})$ are evenly spaced by $\Theta(\sqrt{n})$ as (k, i) vary lexicographically, which already forms $\Omega(n)$ distinct values of $e(U_0 \sqcup \mathbf{Z}_{k,i})$. It remains to fill the $\Theta(\sqrt{n})$ gaps between each $e(U_0 \sqcup \mathbf{Z}_{k,i})$ with $\mathbf{X}$ via the *double-exposure* technique while maintaining the gap structure of $\mathbf{Z}_{k,i}$. Specifically, we sample U from U_0 such that the values $e(U \sqcup \mathbf{Z}_{k,i} \sqcup \{\mathbf{x}\})$ are distinct as $\mathbf{Z}_{k,i} \subseteq \mathbf{S} \sqcup \mathbf{T}$ and $\mathbf{x} \in \mathbf{X}$ vary. Define $U_{k,i} := U \sqcup \mathbf{Z}_{k,i}$ and $e_{k,i} := e(U_{k,i}) - e(U) = e(\mathbf{Z}_{k,i}) + e(\mathbf{Z}_{k,i}, U)$. We need to ensure the following:

1. *Degree distinctness in* $\mathbf{X}$. For most (k, i), the elements of $\mathbf{X}$ should have $\Omega(\sqrt{n})$ pairwise distinct degrees into $U_{k,i}$, all lying in an interval of length $O(\sqrt{n})$. This is the payoff of double-exposure.

2. *Concentration of the edge count.* The value $e_{k,0}$ should be $O(n)$ around its expectation, so that the $\Theta(n)$ spacings between $e_{k,0}$ and $e_{k+1,0}$ are preserved.

3. *Regular medium-scale spacing.* For each k, the values $e_{k,0}, e_{k,1}, \ldots, e_{k,c'\sqrt{n}}$ should span a range of $\Omega(n)$ and increase roughly uniformly, so that $\Theta(\sqrt{n})$ of them are pairwise separated by $\Omega(\sqrt{n})$.

Write $\Delta_{k,i} \coloneqq e_{k,i} - e_{k,i-1}$, and we have the following lemma, whose first two properties correspond to the first two points above, and whose last two properties combined ensure the regular medium-scale spacing.

Lemma 3.7. *Let G be a C-Ramsey graph, let $\mathbf{S}$, $\mathbf{T}$, $\mathbf{X}$, U_0, c', d be obtained from Lemma 3.5. Let U be a uniformly random subset of U_0, and let $U_{k,i}, e_{k,i}, \Delta_{k,i}$ be defined as above. Then there exist positive constants $M, \beta, Q > 0$ depending on C, such that $Q \geq 3\beta$ and for each $c'\sqrt{n} \leq k \leq 2c'\sqrt{n}$ the following hold:*

1. *With probability ≥ 0.99, there is a set $\mathcal{I}_k$ of $(1 - \beta/(2M)) \cdot c'\sqrt{n}$ values with the following property: for each $i \in \mathcal{I}_k$, there exists $\mathbf{X}_{k,i} \subseteq \mathbf{X}$ of size $\Omega(\sqrt{n})$, such that the values $\deg_{U_{k,i}}(\mathbf{x})$ for $\mathbf{x} \in \mathbf{X}_{k,i}$ are distinct and all lie in the interval $[d/2 - Q\sqrt{n},\ d/2 + Q\sqrt{n}]$.*

2. *With probability ≥ 0.99,*
$$\Big| e_{k,0} - \mathbb{E}[e_{k,0}] \Big| \leq Qn.$$

3. *With probability $\geq 1/2$,*
$$e_{k,c'\sqrt{n}} - e_{k,0} \geq 3\beta n.$$

4. *With probability ≥ 0.99,*
$$\sum_{i\,:\,|\Delta_{k,i}| \geq M\sqrt{n}} |\Delta_{k,i}| \leq \beta n.$$

   That is, the unusually large increments contribute little to the total range $e_{k,c'\sqrt{n}} - e_{k,0}$.

Proof. We prove each property in the order of 3, 4, 1, 2, so that the constants β, M, Q are determined in their dependency order.

3. Note that $|e(\mathbf{Z}_{k,i}) - e(\mathbf{Z}_{k,i-1})| \leq 2c'\sqrt{n}$, as $|\mathbf{Z}_{k,i}| = k \leq 2c'\sqrt{n}$. This and the construction of U yields
$$\begin{aligned}\mathbb{E}\Big[e_{k,i} - e_{k,i-1}\Big] &= \mathbb{E}\Big[e(U, \mathbf{Z}_{k,i}) - e(U, \mathbf{Z}_{k,i-1})\Big] + \Big[e(\mathbf{Z}_{k,i}) - e(\mathbf{Z}_{k,i-1})\Big] \\ &\geq \frac{1}{2}\Big[e(U_0, \mathbf{Z}_{k,i}) - e(U_0, \mathbf{Z}_{k,i-1})\Big] - 2c'\sqrt{n}.\end{aligned}$$

   Property 3 of Lemma 3.5 gives $e(U_0, \mathbf{Z}_{k,i}) - e(U_0, \mathbf{Z}_{k,i-1}) \geq 8c'\sqrt{n}$, so
$$\mathbb{E}\Big[e_{k,i} - e_{k,i-1}\Big] \geq 2c'\sqrt{n},$$

for all i. Put $\beta := 2(c')^2/3$ so the $e_{k,c'\sqrt{n}} - e_{k,0}$ has expectation $\geq 3\beta n$. Since $e(U, \mathbf{Z}_{k,i}) = \sum_{u\in U_0} \mathbb{1}_{u\in U} \cdot \deg_{\mathbf{Z}_{k,i}}(u)$,

$$\begin{aligned} e_{k,c'\sqrt{n}} - e_{k,0} &= \Big[e(\mathbf{Z}_{k,c'\sqrt{n}}) - e(\mathbf{Z}_{k,0})\Big] + \Big[e(U, \mathbf{Z}_{k,c'\sqrt{n}}) - e(U, \mathbf{Z}_{k,0})\Big] \\ &= \Big[e(\mathbf{Z}_{k,c'\sqrt{n}}) - e(\mathbf{Z}_{k,0})\Big] + \sum_{u\in U_0} \mathbb{1}_{u\in U} \cdot \Big(\deg_{\mathbf{Z}_{k,c'\sqrt{n}}}(u) - \deg_{\mathbf{Z}_{k,0}}(u)\Big) \end{aligned}$$

is of Littlewood–Offord type with parameter $1/2$, which is symmetrically distributed around its expectation. The result now follows.

4. Since

$$\Delta_{k,i} = \Big[e(\mathbf{Z}_{k,i}) - e(\mathbf{Z}_{k,i-1})\Big] + \sum_{u\in U_0} \mathbb{1}_{u\in U} \cdot \Big(\deg_{\mathbf{Z}_{k,i}}(u) - \deg_{\mathbf{Z}_{k,i-1}}(u)\Big)$$

is of Littlewood–Offord type and measures the edge count difference of swapping one vertex from $\mathbf{S}$ to $\mathbf{T}$, Lemma 3.5 gives $\mathbb{E}\Big[\Delta_{k,i}\Big] = O(\sqrt{n})$ and each $u \in U_0$ contributes at most $|\deg_{\mathbf{Z}_{k,i}}(u) - \deg_{\mathbf{Z}_{k,i-1}}(u)| \leq 2$ to $\Delta_{k,i}$. Thus by Hoeffding's inequality,

$$\mathbb{P}[|\Delta_{k,i}| \geq t] \leq 2\exp\left(-\frac{2\left(t - \mathbb{E}\Big[|\Delta_{k,i}|\Big]\right)^2}{4|U_0|}\right) = \exp\Big(-\Omega(t^2/n)\Big).$$

By the identity $\mathbb{E}[X] = \sum_{t\geq 1} \mathbb{P}(X \geq t)$ for nonnegative integer random variable X,

$$\begin{aligned} \mathbb{E}\Big[|\Delta_{k,i}| \cdot \mathbb{1}_{|\Delta_{k,i}|\geq M\sqrt{n}}\Big] &= \sum_{t=1}^{\infty} \mathbb{P}\Big[|\Delta_{k,i}| \cdot \mathbb{1}_{|\Delta_{k,i}|\geq M\sqrt{n}} \geq t\Big] \\ &= M\sqrt{n} \cdot \mathbb{P}\Big[|\Delta_{k,i}| \geq M\sqrt{n}\Big] + \sum_{t=M\sqrt{n}}^{\infty} \mathbb{P}\Big[|\Delta_{k,i}| \geq t\Big] \\ &= M\sqrt{n} \cdot \exp\Big(-\Omega(M^2)\Big) + \sum_{t=M\sqrt{n}}^{\infty} \exp\Big(-\Omega(t^2/n)\Big) \\ &= \sqrt{n} \cdot \exp\Big(-\Omega(M^2)\Big), \end{aligned}$$

uniformly over M. For large enough M, Markov's inequality now yields

$$\mathbb{P}\left[\sum_{i\,:\,|\Delta_{k,i}|\geq M\sqrt{n}} |\Delta_{k,i}| > \beta n\right] \leq \frac{\sum_{i=1}^{c'\sqrt{n}} \mathbb{E}\Big[|\Delta_{k,i}| \cdot \mathbb{1}_{|\Delta_{k,i}|\geq M\sqrt{n}}\Big]}{\beta n} \leq \frac{e^{-\Omega(M^2)}}{\beta} < 0.01.$$

1. For $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{X}}{2}$,

$$\deg_U(\mathbf{x}) - \deg_U(\mathbf{y}) = \sum_{u\in N_{U_0}(\mathbf{x})} \mathbb{1}_{u\in U} - \sum_{u\in N_{U_0}(\mathbf{y})} \mathbb{1}_{u\in U} = \sum_{u\in N_{U_0}(\mathbf{x})\triangle N_{U_0}(\mathbf{y})} a_u \mathbb{1}_{u\in U}$$

with $a_u \in \mathbb{Z}\backslash\{0\}$. Thus for each k, i, the random variable $\deg_{U_{k,i}}(\mathbf{x}) - \deg_{U_{k,i}}(\mathbf{y})$ is of $\big(|N_{U_0}(\mathbf{x}) \triangle N_{U_0}(\mathbf{y})|, 1/2\big)$-Littlewood–Offord type, as it is a translation of $\deg_U(\mathbf{x}) - \deg_U(\mathbf{y})$. By Property 4 of Lemma 3.5, $|N_{U_0}(\mathbf{x}) \triangle N_{U_0}(\mathbf{y})| = \Omega(n)$, so by Erdős–Littlewood–Offord,

$$\mathbb{P}\Big[\deg_{U_{k,i}}(\mathbf{x}) = \deg_{U_{k,i}}(\mathbf{y})\Big] = O(1/\sqrt{n}).$$

Let $H_{k,i}$ be the graph on $\mathbf{X}$ with edges $\{\mathbf{x}, \mathbf{y}\} \in \binom{\mathbf{X}}{2}$ such that $\deg_{U_{k,i}}(\mathbf{x}) = \deg_{U_{k,i}}(\mathbf{y})$, so $\mathbb{E}[e(H_{k,i})] = |\mathbf{X}|^2 \cdot O(1/\sqrt{n}) = O(\sqrt{n})$. By Markov's inequality, for large enough n,

$$\mathbb{P}\Big[e(H_{k,i}) = O(\sqrt{n})\Big] \geq 1 - \frac{\mathbb{E}[e(H_{k,i})]}{\omega(\sqrt{n})} \geq 1 - \frac{\beta}{400M},$$

in which case Turán's theorem yields an independent set $\mathbf{Y}_{k,i}$ of size $n/(2O(\sqrt{n}) + \sqrt{n}) = 2\gamma\sqrt{n}$, for some constant $\gamma = \gamma(\beta, M) > 0$. In particular, $\deg_{U_{k,i}}(\mathbf{x})$ is distinct for $\mathbf{x} \in \mathbf{Y}_{k,i}$. The expected proportion of values of i for which this fails to occur is $\beta/400M$, and by Markov's inequality again, it fails only for $\beta/2M$ proportion of values of i with probability at least $1 - (\beta/400M)/(\beta/2M) = 0.995$, in which case put $\mathcal{I}_k$ as the set of i that succeeds.

By Property 2 of Lemma 3.5, $\mathbb{E}[\deg_U(\mathbf{x})] = d/2 + O(\sqrt{n})$ for each $\mathbf{x} \in \mathbf{X}$. Since $\deg_U(\mathbf{x})$ follows a binomial distribution with parameters $O(n)$ and $1/2$, we have $\mathrm{Var}(\deg_U(\mathbf{x})) = O(n)$. By triangle inequality,

$$|\deg_U(\mathbf{x}) - d/2| \leq \Big|\deg_U(\mathbf{x}) - \mathbb{E}[\deg_U(\mathbf{x})]\Big| + O(\sqrt{n}),$$

so for sufficiently large Q, Chebyshev's gives

$$\mathbb{P}\left[|\deg_U(\mathbf{x}) - d/2| > Q\sqrt{n}\right] \leq \frac{\mathrm{Var}\big(\deg_U(\mathbf{x})\big)}{Q^2 n} = \frac{O(1)}{Q^2} < \frac{\gamma}{200}.$$

Let $\mathbf{Y}$ be the set of all $\mathbf{x} \in \mathbf{X}$ such that $|\deg_U(\mathbf{x}) - d/2| \leq Q\sqrt{n}$. Then each $\mathbf{x} \in \mathbf{X}$ is not present in $\mathbf{Y}$ with probability $< \gamma/200$. By Markov's,

$$\mathbb{P}\Big[|\mathbf{Y}| \geq (1-\gamma)|\mathbf{X}|\Big] = 1 - \mathbb{P}\Big[|\mathbf{X} \setminus \mathbf{Y}| \geq \gamma|\mathbf{X}|\Big] \geq 1 - \frac{\gamma|\mathbf{X}|}{200\gamma|\mathbf{X}|} = 0.995.$$

Hence, with probability ≥ 0.995, for sufficiently large Q we obtain $\mathbf{Y} \subseteq \mathbf{X}$ of size $\geq (1-\gamma)|\mathbf{X}|$ whose elements lie in the interval $[d/2 - Q\sqrt{n},\ d/2 + Q\sqrt{n}]$.

By the union bound, both of the above events occur with probability ≥ 0.99, in which case take $\mathbf{X}_{k,i} := \mathbf{Y}_{k,i} \cap \mathbf{Y}$ for $i \in \mathcal{I}_k$. Since $|\mathbf{Y}_{k,i}| \geq 2\gamma\sqrt{n}$ and $|\mathbf{Y}| \geq (1-\gamma)|\mathbf{X}|$, we have $|\mathbf{X}_{k,i}| \geq \gamma\sqrt{n}$. This proves Property 1.

2. Since

$$e_{k,i} = e(\mathbf{Z}_{k,i}) + e(U, \mathbf{Z}_{k,i}) = e(\mathbf{Z}_{k,i}) + \sum_{u \in U_0} \mathbb{1}_{u \in U} \cdot \deg_{\mathbf{Z}_{k,i}}(u),$$

$e_{k,i}$ is of $(O(n), 1/2)$-Littlewood–Offord type with coefficients $\deg_{\mathbf{Z}_{k,i}}(u) = O(\sqrt{n})$. Thus $\mathrm{Var}(e_{k,0}) = O(n^2)$ and by Chebyshev's,

$$\mathbb{P}\Big[|e_{k,0} - \mathbb{E}[e_{k,0}]| \leq Qn\Big] \geq 1 - \frac{\mathrm{Var}(e_{k,0})}{Q^2n^2} = 0.99,$$

for large enough Q. We end this proof by noting that enlarging Q does not fail Property 1.

$\square$

3.4 Proof of the main lemma

We assemble Lemma 3.5 and Lemma 3.7 to prove the main lemma.

Apply Lemma 3.5 to obtain $U_0, \mathbf{S}, \mathbf{T}, \mathbf{X}, c', d$, and define $\mathcal{P}, U, \mathbf{Z}_{k,i}, U_{k,i}$ as in Section 3.3. Note that U is the only source of randomness. We prove the main lemma with W being the set of all vertices in $\mathbf{S}, \mathbf{T}, \mathbf{X}$.

Recall that k ranges over the interval between $c'\sqrt{n}$ and $2c'\sqrt{n}$. For each k, the union bound shows that all four properties of Lemma 3.7 hold with probability ≥ 0.4. Let $\mathcal{K}$ be the set of all such k. Then $\mathbb{E}\Big[c'\sqrt{n} - |\mathcal{K}|\Big] \leq 0.6\, c'\sqrt{n}$, and Markov's inequality yields

$$\mathbb{P}\left[|\mathcal{K}| \geq 0.2\, c'\sqrt{n}\right] = 1 - \mathbb{P}\left[c'\sqrt{n} - |\mathcal{K}| \geq 0.8\, c'\sqrt{n}\right] \geq 1 - \frac{\mathbb{E}\Big[c'\sqrt{n} - |\mathcal{K}|\Big]}{0.8\, c'\sqrt{n}} \geq 0.25.$$

We now show that $e(U)$ is close to m with high probability. Note that $|m - \mathbb{E}[e(U)]| = O(n^{3/2})$, as $\mathbb{E}[e(U)] = e(U_0)/4$ and $|e(U_0) - 4m| = O(n^{3/2})$ from Lemma 3.5. Recall that $\mathrm{Var}(e(U)) = O(n^3)$. Since $|e(U) - m| \leq |e(U) - \mathbb{E}[e(U)]| + |\mathbb{E}[e(U)] - m|$, Chebyshev's inequality yields

$$\mathbb{P}\left[|e(U) - m| = O(n^{3/2})\right] \geq \mathbb{P}\left[|e(U) - \mathbb{E}[e(U)]| = O(n^{3/2})\right] \geq 1 - \frac{\mathrm{Var}(e(U))}{\omega(n^3)} \geq 0.9,$$

for large enough n. Fix an outcome U satisfying both events above, which hold together with probability $\geq 0.25 + 0.9 - 1 = 0.15$. Let $Q, \beta, M, \mathcal{I}_k$ be given by Lemma 3.7.

For each k,

$$\mathbb{E}[e_{k,0}] - \mathbb{E}[e_{k-1,0}] = \frac{1}{2}\Big[e(\mathbf{Z}_{k,0}, U_0) - e(\mathbf{Z}_{k-1,0}, U_0)\Big] + O(\sqrt{n}) = \Theta(n),$$

as $\mathbb{E}[\deg_U(\mathbf{x})] = \deg_{U_0}(\mathbf{x})/2 = \Theta(n)$ for $\mathbf{x} \in \mathbf{S}$. We find a subset of $\mathcal{K}$ whose realized values $e_{k,0}$ are well-separated by $\Theta(n)$. Specifically, let $\mathcal{K}' \subseteq \mathcal{K}$ contain every qth element of $\mathcal{K}$, for sufficiently large q such that the values of $\mathbb{E}[e_{k,0}]$ are separated by $\geq 4Qn$ for $k \in \mathcal{K}'$. Then, $|\mathcal{K}'| = \Theta(\sqrt{n})$ and Property 2 of Lemma 3.7 ensures the values of $e_{k,0}$ are separated by $\geq 2Qn$ for $k \in \mathcal{K}'$.

We now partition the linear gaps between the values of $e_{k,0} : k \in \mathcal{K}'$ into square-root sized gaps. Let $k \in \mathcal{K}'$. By Property 3 of Lemma 3.7, $e_{k,c'\sqrt{n}} - e_{k,0} \geq 3\beta n$. Consider the range of integers between $e_{k,0}$ and $e_{k,0}+3\beta n$. Within this range, we take $2\beta\sqrt{n}/M$ intervals of length $M\sqrt{n}$, each separated by a distance of $M\sqrt{n}/2 = \Omega(\sqrt{n})$. By Property 4 of Lemma 3.7, $\leq \beta\sqrt{n}/M$ of these intervals contain no value of $e_{k,i}$. Consider the $\beta\sqrt{n}/M$ intervals that do contain a value of $e_{k,i}$ and pick a representative $e_{k,i}$ from each. Let $\mathcal{I}'_k$ be the set of indices i corresponding to these representatives, and let $\mathcal{I}''_k \coloneqq \mathcal{I}_k \cap \mathcal{I}'_k$. Then

$$|\mathcal{I}''_k| = |\mathcal{I}_k| + |\mathcal{I}'_k| - |\mathcal{I}_k \cup \mathcal{I}'_k| \geq \beta\sqrt{n}/M - \beta c'\sqrt{n}/2M = \Omega(\sqrt{n})$$

and each $e_{k,i}$ is separated by $\Omega(\sqrt{n})$ for $i \in \mathcal{I}''_k$. Recall that $Q \geq 3\beta$, so by construction $|e_{k,i} - e_{k,0}| \leq 3\beta n \leq Qn$ for $i \in \mathcal{I}''_k$. In particular, amongst choices of $k \in \mathcal{K}'$ and $\mathcal{I}''_k$, we already have $\Omega(n)$ distinct values of $e(U_{k,i})$ that are separated by $\Omega(\sqrt{n})$. It remains to fill the square-root sized gaps between these values.

In increasing order of $e(U_{k,i})$, let $\mathcal{P}' \subseteq \left\{(k,i) : k \in \mathcal{K}', i \in \mathcal{I}''_k\right\}$ be the corresponding index set of every qth of these values. We enlarge q so that $e(U_{k,i})$ from $\mathcal{P}'$ are separated by $\geq 2Q\sqrt{n}$. For each $(k,i) \in \mathcal{P}'$ we have $i \in \mathcal{I}_k$, so Property 1 of Lemma 3.7 furnishes a set $\mathbf{X}_{k,i}$ such that each $\deg_{U_{k,i}}(\mathbf{x})$ is unique for $\mathbf{x} \in \mathbf{X}_{k,i}$ while lying in a fixed interval of length $2Q\sqrt{n}$ centred at $d/2$, where d is independent of k and i. Thus for each $(k,i) \in \mathcal{P}'$, the set

$$\left\{e(U_{k,i} \sqcup \mathbf{x}_{k,i}) : \mathbf{x}_{k,i} \in \mathbf{X}_{k,i}\right\}$$

has size $|\mathbf{X}_{k,i}| = \Omega(\sqrt{n})$ and lies in an interval of length $2Q\sqrt{n}$ centred at $e(U_{k,i})+d/2$, so these sets do not overlap among different $(k,i) \in \mathcal{P}'$. Therefore,

$$\left|\left\{e(U \sqcup Z) : Z \subseteq W\right\}\right| \geq \sum_{(k,i)\in\mathcal{P}'} \left|\left\{e(U_{k,i} \sqcup \mathbf{x}_{k,i}) : \mathbf{x}_{k,i} \in \mathbf{X}_{k,i}\right\}\right| = \Omega(n^{3/2}).$$

$\square$

4 The Erdős–McKay Conjecture

Both proofs in Chapters 2 and 3 rely on the same anticoncentration principle: when the symmetric difference $|N(x)\triangle N(y)|$ is large, the degree difference $\deg_U(x) - \deg_U(y)$ is a sum of many independent bounded random variables, and its point probabilities are $O(1/\sqrt{n})$. The Erdős–McKay conjecture asks for a qualitatively stronger result that asserts $\Phi(G)$ contains every integer in a quadratic-length interval, not merely that it is large. The conjecture was resolved by Kwan, Sah, Sauermann, and Sawhney [11], and this chapter is dedicated to outlining their proof.

Theorem 4.1 (Erdős–McKay conjecture [11])**.** *Fix $C > 0$. Let G be a C-Ramsey graph on n vertices for sufficiently large n. Then there exists $\delta = \delta(C) > 0$ such that*

$$\Phi(G) \supseteq \Big\{1, 2, \ldots, \delta e(G)\Big\}.$$

The approach to this theorem fundamentally shifts from bounding degree-collision probabilities for pairs of vertices to directly controlling the point probabilities of $e(U)$ itself.

4.1 Point probabilities of $e(U)$

Let G be a Ramsey graph on $[n]$, and let U be a uniformly random vertex subset of G. By the Erdős–Szemerédi bound (Corollary 1.4), the random variable $e(U)$ has mean $\mu := \mathbb{E}[e(U)] = e(G)/4 = \Theta(n^2)$. For variance, the standard variance identity gives

$$\operatorname{Var}\Big(e(U)\Big) = \sum_{(e,e')\in E(G)^2} \operatorname{Cov}\Big(\mathbb{1}_{e\subseteq U}, \mathbb{1}_{e'\subseteq U}\Big).$$

The covariance is 0 if e and e' are disjoint and $\Theta(1)$ otherwise, so $\operatorname{Var}(e(U)) = \Theta(n^3)$. For a fixed vertex v, the number of pairs of edges sharing v is $\binom{\deg(v)}{2}$, and the Erdős–Szemerédi bound with Jensen's inequality gives

$$\operatorname{Var}\Big(e(U)\Big) \gtrsim \sum_{v\in V(G)} \binom{\deg(v)}{2} \geq n\binom{2e(G)/n}{2} = \Omega(n^3).$$

Thus the standard deviation is $\sigma := \sigma(e(U)) = \Theta(n^{3/2})$.

In Chapters 2 and 3, the anticoncentration analysis targeted the degree difference $\deg_U(x) - \deg_U(y)$, which is approximately normally distributed with standard deviation $\Theta(\sqrt{n})$, and the control of $e(U)$ was achieved indirectly through the construction of $\mathbf{S}, \mathbf{T}, \mathbf{X}$. It turns out that, via standard techniques, a CLT directly holds for $e(U)$, giving us

$$\mathbb{P}\Big[e(U) \leq x\Big] = \Phi\left(\frac{x-\mu}{\sigma}\right) + o(1/\sigma)$$

for all $x \in \mathbb{R}$ with Φ being the standard Gaussian cumulative distribution function. In [11], it is strengthened to an approximation of the local CLT, showing that the point probabilities of $e(U)$ are all of order $\Theta(1/\sigma) = \Theta(n^{-3/2})$ within the bulk of the distribution.

Theorem 4.2 (Kwan–Sah–Sauermann–Sawhney [11])**.** *Fix constant $C > 0$. Let G be a C-Ramsey graph on n vertices and let U be a uniformly random subset of $V(G)$. Then*

$$\sup_{x\in\mathbb{Z}} \mathbb{P}[e(U) = x] = O(n^{-3/2}),$$

and for every fixed $A > 0$,

$$\inf_{\substack{x\in\mathbb{Z} \\ |x-e(G)/4|\le An^{3/2}}} \mathbb{P}[e(U) = x] = \Omega(n^{-3/2}).$$

The Erdős–McKay conjecture follows directly from the lower bound of Theorem 4.2 combined with Theorem 3.1 on small values of x.

Proof of Theorem 4.1. By Theorem 3.1, there exists $\alpha = \alpha(C) > 0$ such that $\Phi(G) \supseteq \{0, 1, \dots, n^\alpha\}$. Fix integer x between n^α and $e(G)/4$. Let m be the smallest integer such that $e(\{1, \dots, m\}) \ge 4x$. Put $G' := G[\{1, \dots, m\}]$ and note that

$$e(G') \ge 4x \ge e\Big(\{1, \dots, m-1\}\Big) \ge e(G') - m,$$

so $|x - e(G')/4| \le m/4 \le m^{3/2}$. Since $m^2 \ge e(G') \ge 4x \ge n^\alpha$, we have $m \ge n^{\alpha/2}$, so G' is $(2C/\alpha)$-Ramsey. Applying Theorem 4.2 to G' with $A = 1$ gives

$$\mathbb{P}\Big[e(G'[U]) = x\Big] = \Omega(m^{-3/2}) = \Omega(n^{-3\alpha/4}) > 0.$$

Therefore, there exists $U \subseteq V(G') \subseteq V(G)$ such that $e(G[U]) = e(G'[U]) = x$, for sufficiently large n. This completes the proof. $\square$

Remark: Kwan, Sah, Sauermann, and Sawhney [11] prove the stronger statement $\Phi(G) \supseteq \{0, 1, \dots, (1-\eta)e(G)\}$ for any $\eta > 0$, at the cost of a slight complication in the argument. We settle for the weaker version that suffices for the conjecture.

We will outline the proof of Theorem 4.2 in the remainder of this chapter.

4.2 Proof overview

The standard way to approach local limit theorems is via Fourier transforms. Define the *graph polynomial* as

$$f(\xi_1, \dots, \xi_n) := \sum_{ij\in E(G)} \xi_i \xi_j.$$

For any $U \subseteq V(G)$, setting $\xi_v := \mathbb{1}_{v\in U}$ gives $f\Big(\vec{\xi}^{(U)}\Big) = e(U)$. The Erdős–McKay conjecture is thus equivalent to asking whether the range of this *quadratic* polynomial on $\{0,1\}^n$ covers every integer in $\{0, \dots, \delta n^2\}$.

For a random variable X, the *characteristic function* of X is defined as

$$\varphi_X(\tau) := \mathbb{E}\left[e^{i\tau X}\right] = \int_{-\infty}^{\infty} e^{i\tau X} \cdot \mathbb{P}[X = x]\, dx, \qquad \tau \in \mathbb{R},$$

and note that $|\varphi_X(\tau)| \leq 1$ for all τ. Applying the Fourier inversion formula yields

$$\mathbb{P}[X = x] = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\tau x}\varphi_X(\tau)\, d\tau \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |\varphi_X(\tau)|\, d\tau, \tag{2}$$

for integer-valued X. In particular, any upper bound on the integrated $|\varphi_X(\tau)|$ produces an upper bound on the point probability. For the matching lower bound, we compare φ_X with the characteristic function $\varphi_Z(\tau) := e^{i\tau\mu - \sigma^2\tau^2/2}$ of a Gaussian $Z \sim \mathcal{N}(\mu, \sigma^2)$, where $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$. Since the Gaussian density at x is of order $\Theta(1/\sigma)$ for any x within $O(\sigma)$ of the mean, $\mathbb{P}[X = x]$ inherits the same order of magnitude $\Theta(1/\sigma)$ through (2) whenever $|\varphi_X(\tau) - \varphi_Z(\tau)|$ is small. If a local CLT held for $X = e(U)$, the proof would amount to verifying suitable estimates on $|\varphi_X(\tau)|$ on $[-\pi, \pi]$ and we are done.

Unfortunately, an actual local CLT does not hold for $e(U)$ in general. The distribution of $e(U)$ can fail to be smooth at the unit scale despite the global shape being approximately Gaussian, which results in spikes in the characteristic function and prevents the integrand in (2) from being uniformly small. In particular, the authors in [11] observed a “two-scale” behaviour of the distribution of $e(U)$, depending on the additive structure of the degree sequence of G (see Figure 2).

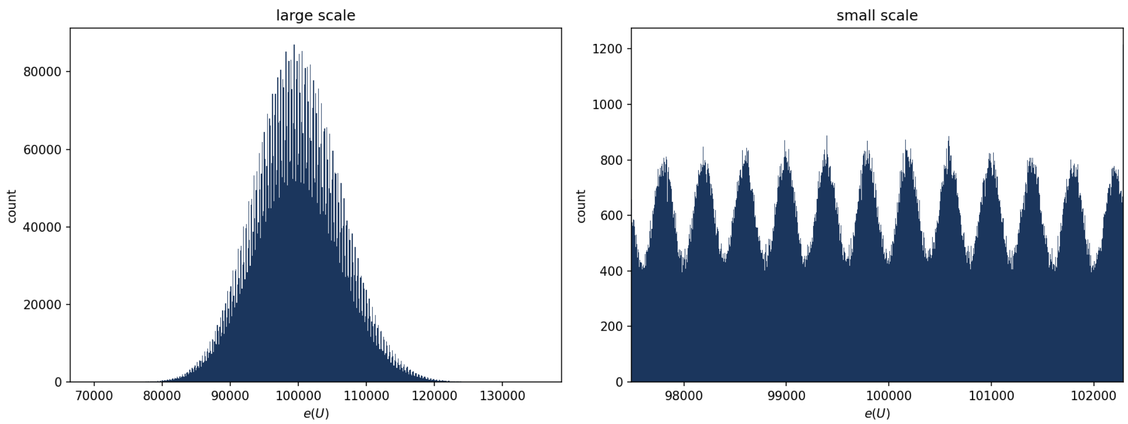

Figure 2: Histograms of $e(U)$ obtained from ten million independent samples of U from an outcome of $G_{1000,0.8}$. We see that on the large scale, the distribution of $e(U)$ is approximately Gaussian, whereas on the small scale it shows multiple smaller Gaussian-like curves. Experiment reproduced from [11]; code in Appendix A.

The proof of Theorem 4.2 overcomes this obstacle in two stages. The first stage splits into two cases according to whether the degree sequence of G is additively structured or not. The *unstructured case* can be handled by a standard Fourier-analytic argument without further intervention. The *structured case* is where $|\varphi_X(\tau)|$

spikes, which we will dedicate most of the effort to handling. These two cases yield the interval estimate

$$\mathbb{P}\Big[|X-x| \leq B\Big] = \Theta(n^{-3/2}), \tag{3}$$

for some $B = B(C)$. The second stage upgrades (3) to a pointwise estimate by showing that the probability mass cannot be too unevenly distributed within the interval via a *switching argument*.

4.3 Degree sequence dichotomy

We first decompose $e(U)$ into a sum of a linear and quadratic part. Set $\xi_v := \mathbb{1}_{v\in U}$ and write $x_v := 2\xi_v - 1 \in \{-1,+1\}$, so that $\xi_v = (1+x_v)/2$. Substituting into $e(U) = f\big(\vec{\xi}\big) = \sum_{ij\in E(G)} \xi_i\xi_j$,

$$X = \sum_{ij\in E(G)} \frac{(1+x_i)(1+x_j)}{4} = \underbrace{\frac{e(G)}{4}}_{\mathbb{E}[X]} + \underbrace{\frac{1}{4}\sum_{v\in V(G)} \deg(v)x_v}_{L} + \underbrace{\frac{1}{4}\sum_{uv\in E(G)} x_u x_v}_{Q}, \tag{4}$$

and denote the linear term as L and the quadratic term as Q. Since $\mathbb{E}[x_v] = 0$ and $\mathbb{E}[x_u x_v] = 0$ for $u \neq v$, we have $\mathbb{E}[L] = \mathbb{E}[Q] = 0$, the covariance of L and Q is zero, and the constant term $e(G)/4 = \mathbb{E}[X]$. It can be shown via the Erdős–Szemerédi bound and Jensen's inequality that $\mathrm{Var}(L) = \Theta(n^3), \mathrm{Var}(Q) = \Theta(n^2)$. The linear part L therefore dominates the total variance, so we expect L to drive the global behaviour of X and Q to act as a corrective role. Write $d_v := \deg(v)/4$ so that $L = \sum_v d_v x_v$. Since the variables x_v are independent,

$$\varphi_L(\tau) = \mathbb{E}\left[\exp\left(i\tau\sum_v d_v x_v\right)\right] = \prod_{v\in V(G)} \mathbb{E}\Big[\exp(i\tau d_v x_v)\Big] = \prod_{v\in V(G)} \cos\Big(\tau d_v\Big). \tag{5}$$

The decomposition (4) thus allows us to bound X by bounding the characteristic function of L which factors well.

For $y \in \mathbb{R}$, denote $\|y\| := \mathrm{dist}(y, \mathbb{Z})$ as the distance from y to the nearest integer, and note that $0 \leq \|y\| \leq 1/2$. By Jordan's inequality and $1-t \leq \exp(-t)$ for $t \geq 0$,

$$|\cos(\pi y)| = \cos(\pi\|y\|) = \sqrt{1-\sin^2(\pi\|y\|)} \leq \sqrt{1-4\|y\|^2} \leq 1-2\|y\|^2 \leq \exp(-2\|y\|^2),$$

for $y \in \mathbb{R}$, so applying to (5) gives

$$|\varphi_L(\tau)| \leq \exp\left(-2\sum_{v\in V(G)} \left\|\frac{\tau d_v}{\pi}\right\|^2\right). \tag{6}$$

The right-hand side decays exponentially in n unless many of the quantities $\tau d_v/\pi$ are simultaneously close to integers. Throughout what follows, let $\nu = \nu(C) > 0$ be a small constant. The low range $|\tau| \leq n^{-0.99}$ is controlled by comparison with the characteristic function of a Gaussian $\varphi_Z(\tau) := \exp\Big(i\mu\tau - \sigma^2\tau^2/2\Big)$ via a version of the Esseen's inequality [11, Lemma 6.1]. The high range $|\tau| \ge \nu$ is controlled by a *quadratic-cancellation* estimate on Q [11, Section 8]. What concerns us is the medium window $n^{-0.99} \le |\tau| \le \nu$, in which the bound (6) splits into two cases.

The unstructured case. In the case that

$$\sum_{v\in V(G)} \left\| \frac{\tau d_v}{\pi} \right\|^2 = \omega(\log n)$$

for all $n^{-0.99} < |\tau| \le \nu$, the bound (6) gives a superpolynomial decay $|\varphi_L(\tau)| \le n^{-\omega(1)}$ on the window. In this case, Esseen's inequality applied to the difference $\varphi_X - \varphi_Z$ delivers the interval estimate (3). This case admits a local CLT.

The structured case. Here we consider the opposite case that

$$\sum_{v\in V(G)} \left\| \frac{\tau d_v}{\pi} \right\|^2 = O(\log n),$$

for some $n^{-0.99} < |\tau| \le \nu$. This says that for this one frequency, many of the numbers $\tau d_v/\pi$ are simultaneously close to integers. Lemma 4.12 of [11] uses this condition to show that for small enough $\gamma > 0$, the vertex set admits a partition

$$V(G) = R \sqcup (B_1 \sqcup \cdots \sqcup B_m), \tag{7}$$

with $|R| \le n^{1-\gamma}$, $|B_j| = \lceil n^{1-2\gamma} \rceil$ for $j \in [m]$, and real numbers $\kappa_1, \ldots, \kappa_m \ge 0$ such that

$$|d_v - \kappa_j| \le n^{1/2+4\gamma},$$

for $j \in [m]$ and $v \in B_j$. That is, the vertex set splits into $m = \Theta(n^{2\gamma})$ many buckets $B_1, \ldots, B_m$, each with degrees clustered around κ_j, and a sublinear-sized exceptional set R. This allows us to write L as

$$L = \sum_{j=1}^{m} \sum_{v\in B_j} d_v x_v + \sum_{v\in R} d_v x_v = \underbrace{\sum_{j=1}^{m} \kappa_j \sum_{v\in B_j} x_v}_{L_{\mathrm{main}}} + \underbrace{\sum_{j=1}^{m} \sum_{v\in B_j} (d_v - \kappa_j) x_v}_{L_{\mathrm{err}}} + \underbrace{\sum_{v\in R} d_v x_v}_{L_R}.$$

Then conditional on

$$\vec{\Delta} := \Big(|U \cap B_1|, \ldots, |U \cap B_m|, (x_v)_{v\in R}\Big),$$

namely the sizes of each bucket's intersection with U and $(x_v)_{v\in R}$, makes L_{main} and L_R deterministic, reducing the conditional variance of L to $\mathrm{Var}(L_{\mathrm{err}} \mid \vec{\Delta}) \le n^{2+8\gamma}$. The effect of conditioning on $\vec{\Delta}$ can be thought of as conditioning on the approximate value of L which governs the large-scale behaviour of X, allowing us to "zoom in" to see the smaller-scale behaviour of the distribution which corresponds to one of the peaks in Figure 2. In particular, $U \setminus R$ is conditionally uniform over the product of Boolean slices

$$\binom{B_1}{\Delta_1} \times \cdots \times \binom{B_m}{\Delta_m}.$$

This reduces the interval estimate (3) to evaluating the small-ball probability of $L_{\mathrm{err}} + Q$ on this product of Boolean slices.

It is tempting at this point to hope that, after conditional on $\vec{\Delta}$, $L_{\mathrm{err}} + Q$ satisfies a local CLT of its own. This hope is too optimistic. Consider G to be the disjoint union of two independent copies of $G(n/2, 1/2)$. With high probability G is a Ramsey graph. Every vertex has degree approximately $n/4$, so all vertices fall into a single bucket and $\vec{\Delta} = |U|$. However, conditional on $|U| = n/2$ yields a normalized limiting distribution of the form $\alpha Z_1^2 + \beta Z_2$ for independent Gaussians Z_1, Z_2. Therefore, the asymptotic conditional distribution of X need not be Gaussian after conditional on the bucket, though we shall see that it can always be evaluated at independent Gaussians.

4.4 The structured case

The target now is to show

$$\mathbb{P}\left[|X - x| \leq B \mid \vec{\Delta}\right] = O(n^{-1}), \tag{8}$$

which gives the interval estimate (3) after averaging over $\vec{\Delta}$. Specifically, since $\sigma\Big(\mathbb{E}[X \mid \vec{\Delta}]\Big) = \Theta(n^{3/2})$, only $O(n/n^{3/2}) = O(n^{-1/2})$ fraction of $\vec{\Delta}$ values place $\mathbb{E}[X \mid \vec{\Delta}]$ within $O(n)$ of x, so this fraction contributes $O(n^{-1/2}) \cdot O(n^{-1}) = O(n^{-3/2})$ when averaged over $\vec{\Delta}$. The remaining tail values of $\vec{\Delta}$ are negligible by *hypercontractivity* [11, Theorem 4.15].

As shown in the previous section, the conditional distribution of X need not be approximately Gaussian, so the standard Fourier-analytic approach does not close on its own. The authors of [11] overcome this by combining the following three ingredients.

Ingredient 1. The Gaussian invariance principle We first replace the product of Boolean slices by a Gaussian space. The *invariance principle* of Mossel, O'Donnell and Oleszkiewicz [12] says that a low-degree polynomial of independent bounded random variables has a distribution close to that of the same polynomial evaluated on independent standard Gaussians. Applying this to $L_{\mathrm{err}} + Q$ produces a quadratic polynomial

$$f(\vec{Z}) = \vec{Z}^\top F \vec{Z} + \vec{f} \cdot \vec{Z} + f_0, \qquad \vec{Z} = (Z_1, \ldots, Z_n) \sim \mathcal{N}(0, I_n), \tag{9}$$

with a symmetric matrix $F \in \mathbb{R}^{n \times n}$ and $\vec{f} \in \mathbb{R}^n$ derived from the adjacency matrix A of G and the bucket partition, such that the characteristic functions of $L_{\mathrm{err}} + Q$

(given $\vec{\Delta}$) and of $f(\vec{Z})$ are close – see [11, Lemma 11.1]. The role of this step is to convert a discrete, dependent domain into a continuous, independent one, allowing us to apply the *Carbery–Wright theorem.*

Ingredient 2. The sharpened Carbery–Wright theorem Given the $f(\vec{Z})$ from ingredient 1, we now bound $\sup_x \mathbb{P}[|f(\vec{Z}) - x| \le 1]$. Recall that the *Frobenius norm* for matrix $M \in \mathbb{R}^{n\times n}$ is $\|M\|_{\mathrm{F}} := \left(\sum_{i,j} M_{ij}^2\right)^{1/2}$. In our application $\sigma(f(\vec{Z})) = \Omega(n)$, as $\sigma(f(\vec{Z}))^2 = 2\|F\|_{\mathrm{F}}^2 + \|\vec{f}\|_2^2$ by [11, Lemma 11.1(2)] and Ingredient 3 below forces $\|F\|_{\mathrm{F}}^2 = \Omega(n^2)$. The target $O(n^{-1})$ in (8) is therefore asking for a bound of the form $1/\sigma(f(\vec{Z}))$.

The classical result in this direction is the Carbery–Wright theorem, which states that for quadratic polynomials f in independent standard Gaussians and any $\varepsilon > 0$,

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left[|f(\vec{Z}) - x| \le \varepsilon\right] = O\left(\sqrt{\varepsilon/\sigma(f(\vec{Z}))}\right),$$

which falls short of our target $O\Big(1/\sigma(f(\vec{Z}))\Big)$. The square-root exponent cannot be improved in general, but it turns out that forcing the associated matrix F to have *effective rank* at least three yields the desired bound.

Theorem 4.3 (Sharpened Carbery–Wright, [11, Theorem 1.6])**.** *Let $f(\vec{Z}) = \vec{Z}^\top F \vec{Z} + \vec{f}\cdot\vec{Z} + f_0$ be a real quadratic polynomial in independent standard Gaussians, with F a nonzero symmetric matrix. Suppose that for some $\eta > 0$,*

$$\min_{\substack{\widetilde{F}\in\mathbb{R}^{n\times n}\\ \operatorname{rank}(\widetilde{F})\le 2}} \|F - \widetilde{F}\|_{\mathrm{F}}^2 \ge \eta\|F\|_{\mathrm{F}}^2. \tag{10}$$

Then for $\varepsilon > 0$,

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left[|f(\vec{Z}) - x| \le \varepsilon\right] = O\left(\varepsilon/\sigma(f(\vec{Z}))\right).$$

By the Eckart–Young theorem, the left-hand side of (10) equals $\sum_{i\ge 3}\sigma_i(F)^2$, where $\sigma_1(F) \ge \sigma_2(F) \ge \cdots \ge \sigma_n(F) \ge 0$ are the singular values of F. Condition (10) therefore asserts that the rank-(≥ 3) tail of F carries at least η-fraction of $\|F\|_{\mathrm{F}}^2$.

It remains to show (10) for the matrix F arising from the bucket structure, and this is where the Ramsey hypothesis enters.

Ingredient 3. The robust rank lemma

Lemma 4.4 (Robust rank, [11, Lemma 10.1])**.** *Fix $C > 0$, $0 < \delta < 1$ and $r \in \mathbb{N}$. Let G be a C-Ramsey graph on n vertices with adjacency matrix A, and let $V(G) = B_1 \sqcup \cdots \sqcup B_m$ be a partition into equal parts with $n^{\delta/2} \le m \le 2n^\delta$. Then for every matrix $M \in \mathbb{R}^{n\times n}$ with* $\operatorname{rank}(M[B_j \times B_k]) \le r$ *for all $j, k \in [m]$,*

$$\|A - M\|_{\mathrm{F}}^2 = \Omega(n^2). \tag{11}$$

The lemma says that the adjacency matrix of a Ramsey graph is at Frobenius distance $\Omega(n)$ from every matrix whose restriction to each bucket block has low rank. When M is binary, $\|A - M\|_{\mathrm{F}}^2$ counts the number of entries on which A, M disagree, so the lemma says no binary matrix with low-rank blocks can agree with A on almost all entries.

In the application of Theorem 4.3, the parts B_j are the buckets from (7) of size $\lceil n^{1-2\gamma} \rceil$. Write $\vec{x} := (x_v)_{v \in V}$. The matrix F from Ingredient 1 has the property that, conditional on $\vec{\Delta}$, the quadratic form $\vec{x}^\top F \vec{x}$ only concerns the deviation of $\vec{x}$ from its bucket averages. Specifically, the adjacency matrix A of $G[V \setminus R]$ may be written as

$$A = \alpha F + B$$

for some fixed α and correction matrix B satisfying $\mathrm{rank}(B[B_j \times B_k]) \leq b$ for any block $B_j \times B_k$ and some constant b. Suppose for contradiction that some $\tilde{F}$ with rank ≤ 2 violates (10). Set $\tilde{A} := \alpha \tilde{F} + B$. Then for $\eta > 0$, since $\|F\|_{\mathrm{F}}^2 = O(n^2)$,

$$\|A - \tilde{A}\|_{\mathrm{F}}^2 = |\alpha|^2 \cdot \|F - \tilde{F}\|_{\mathrm{F}}^2 < \eta \cdot |\alpha|^2 \cdot \|F\|_{\mathrm{F}}^2 = O(\eta n^2).$$

But then on every $B_j \times B_k$ we have $\mathrm{rank}(\tilde{A}[B_j \times B_k]) \leq b + 2$, contradicting (11) for small enough η. Hence (10) holds for some $\eta = \eta(C, \gamma)$. See [11, Section 12] for the explicit construction of F.

To sum it up, applying Ingredients 1, 3, and 2 in sequence yields (8).

4.5 The switching argument

The first stage delivers the interval estimate

$$\mathbb{P}\Big[|X - x| \leq B\Big] = \Theta(n^{-3/2}),$$

but leaves open the worst case where all of this mass sits on some single integer $x' \neq x$ inside the window $[x - B, x + B]$. We rule this out via a *switching argument* between different possibilities of the set U, which upgrades the interval estimate to a pointwise bound. The interval estimate (3) already implies an upper bound on $\mathbb{P}[X = x]$, so it remains to show a lower bound.

For $y \in U$ and $z \notin U$, the *switch* (y, z) replaces U by $U' = (U \setminus \{y\}) \cup \{z\}$. This operation shifts X by

$$s(y, z; U) := e(U') - e(U) = \deg_{U \setminus \{y\}}(z) - \deg_U(y).$$

Fix a set $T \subseteq V(G)^2$ of candidate ordered pairs, to be specified later, and let

$$Y_\ell := \#\Big\{(y, z) \in T : y \in U,\ z \notin U,\ s(y, z; U) = \ell\Big\}$$

count candidate switches at U that shift X by exactly $\ell \in \{-B, \dots, B\}$. Each switch $(y, z) \in T$ with $y \in U, z \notin U$, and $s(y, z; U) = \ell$ pairs U with the set

$U' = (U \setminus \{y\}) \cup \{z\}$. The reverse switch $(z, y) \in T$ from U' recovers U and shifts X by $-\ell$. In particular, pairing $(U, (y, z))$ with $(U', (z, y))$ in this way is a bijection. Since the uniform measure on $\{0, 1\}^n$ assigns the same probability 2^{-n} to every U, the bijection gives the symmetry identity

$$\mathbb{E}\Big[Y_\ell \cdot \mathbb{1}_{X=x}\Big] = \mathbb{E}\Big[Y_{-\ell} \cdot \mathbb{1}_{X=x+\ell}\Big], \tag{12}$$

for any $x \in \mathbb{Z}$ and $\ell \in \{-B, \dots, B\}$. By Cauchy–Schwarz,

$$\Big(\mathbb{E}\Big[Y_\ell \cdot \mathbb{1}_{X=x}\Big]\Big)^2 \leq \mathbb{E}\Big[Y_\ell^2 \cdot \mathbb{1}_{X=x}\Big] \cdot \mathbb{E}\Big[\mathbb{1}_{X=x}\Big],$$

and applying (12) yields

$$\mathbb{P}[X = x] = \mathbb{E}\Big[\mathbb{1}_{X=x}\Big] \geq \frac{\Big(\mathbb{E}\Big[Y_{-\ell} \cdot \mathbb{1}_{X=x+\ell}\Big]\Big)^2}{\mathbb{E}\Big[Y_\ell^2 \cdot \mathbb{1}_{X=x}\Big]}, \tag{13}$$

reducing the problem to bounding the numerator and denominator using the interval estimate (3). For a suitable T, Lemma 13.4 of [11] converts the interval estimate (3) to moment estimates for Y_ℓ,

$$\mathbb{E}\left[\prod_{\ell=-B}^{B} Y_\ell^{a_\ell} \cdot \mathbb{1}_{|X-x| \leq B}\right] = \Theta\left(\left(|T|/\sqrt{n}\right)^{\sum_\ell a_\ell} \cdot n^{-3/2}\right), \tag{14}$$

for $a_\ell \in \{0, 1, 2\}$. Below we apply (14) to bound both numerator and denominator in (13).

Since $\mathbb{1}_{X=x} \leq \mathbb{1}_{|X-x| \leq B}$, (14) bounds the denominator of (13) by

$$\mathbb{E}\Big[Y_\ell^2 \cdot \mathbb{1}_{X=x}\Big] \leq \mathbb{E}\Big[Y_\ell^2 \cdot \mathbb{1}_{|X-x| \leq B}\Big] \lesssim \left(\frac{|T|}{\sqrt{n}}\right)^2 \cdot n^{-3/2}. \tag{15}$$

As for the numerator, Cauchy–Schwarz yields

$$\mathbb{E}\Big[Y_{-\ell} \cdot \mathbb{1}_{X=x+\ell}\Big] \geq \frac{\left(\mathbb{E}\left[\prod_{\ell=-B}^{B} Y_\ell \cdot \mathbb{1}_{X=x+\ell}\right]\right)^2}{\mathbb{E}\Big[\prod_{i \neq -\ell} Y_i^2 \cdot Y_{-\ell} \cdot \mathbb{1}_{X=x+\ell}\Big]}.$$

By writing $\mathbb{1}_{|X-x| \leq B} = \sum_{\ell'=-B}^{B} \mathbb{1}_{X=x+\ell'}$ and pigeonholing on the largest term, there exists $\ell^* \in \{-B, \dots, B\}$ such that

$$\mathbb{E}\left[\prod_{\ell=-B}^{B} Y_\ell \cdot \mathbb{1}_{X=x+\ell^*}\right] \geq \frac{1}{2B+1} \mathbb{E}\left[\prod_{\ell=-B}^{B} Y_\ell \cdot \mathbb{1}_{|X-x| \leq B}\right] \gtrsim \left(\frac{|T|}{\sqrt{n}}\right)^{2B+1} \cdot n^{-3/2}. \tag{16}$$

On the other hand, for all $\ell \in \{-B, \dots, B\}$, (14) gives

$$\mathbb{E}\left[\prod_{i \neq -\ell} Y_i^2 \cdot Y_{-\ell} \cdot \mathbb{1}_{X=x+\ell}\right] \leq \mathbb{E}\left[\prod_{i \neq -\ell} Y_i^2 \cdot Y_{-\ell} \cdot \mathbb{1}_{|X-x| \leq B}\right] \lesssim \left(\frac{|T|}{\sqrt{n}}\right)^{4B+1} \cdot n^{-3/2}. \tag{17}$$

Combining (16) and (17) gives

$$\mathbb{E}\Big[Y_{-\ell^*} \cdot \mathbb{1}_{X=x+\ell^*}\Big] \gtrsim \Big(|T|/\sqrt{n}\Big) \cdot n^{-3/2}, \tag{18}$$

for some $\ell^* \in \{-B, \dots, B\}$. Thus, (18) and (15) give the desired pointwise bound

$$\mathbb{P}[X = x] \gtrsim \frac{\Big((|T|/\sqrt{n}) \cdot n^{-3/2}\Big)^2}{\Big(|T|/\sqrt{n}\Big)^2 \cdot n^{-3/2}} = n^{-3/2}.$$

It remains to find the suitable set T that enables (14). The two factors in (14) have concrete meanings: $n^{-3/2}$ is the interval estimate (3), whereas $\mathbb{E}[Y_\ell] = \Theta(|T|/\sqrt{n})$, as follows.

For $(y, z) \in T$, we have $\mathbb{P}[y \in U, z \notin U] = 1/4$. Conditional on this event, set $R_w \coloneqq 2\mathbb{1}_{w \in U} - 1$ and

$$a_w \coloneqq \mathbb{1}_{w \in N(z) \setminus N(y)} - \mathbb{1}_{w \in N(y) \setminus N(z)}.$$

Then $s(y, z; U)$ may be written as

$$\begin{aligned} s(y, z; U) &= \deg_{U \setminus \{y\}}(z) - \deg_U(y) \\ &= \sum_{w \in N(z),\, w \neq y} \mathbb{1}_{w \in U} - \sum_{w \in N(y)} \mathbb{1}_{w \in U} \\ &= \frac{\deg(z) - \deg(y)}{2} + \frac{1}{2} \sum_{w \in N(y) \triangle N(z)} a_w R_w, \end{aligned}$$

which is of $(|N(y) \triangle N(z)|, 1/2)$-Littlewood–Offord type. Now suppose we can find T such that for $(y, z) \in T$

1. $|N(y) \triangle N(z)| = \Theta(n)$,
2. $|\deg(z) - \deg(y)| = O(\sqrt{n})$

Condition 1 controls the variance of $s(y, z; U)$ at $\Theta(n)$, and condition 2 places the conditional mean $(\deg(z) - \deg(y))/2$ within $O(\sqrt{n})$ of zero, so $\ell \in \{-B, \dots, B\}$ falls within $O(\sqrt{n})$ around the mean. Erdős–Littlewood–Offord then gives the upper bound

$$\mathbb{P}\Big[s(y, z; U) = \ell \mid y \in U, z \notin U\Big] = O(1/\sqrt{n})$$

uniformly in ℓ, and Stirling's formula in the form $\binom{m}{m/2+k} 2^{-m} = \Theta(m^{-1/2})$ for $|k| = O(\sqrt{m})$ gives the matching lower bound

$$\mathbb{P}\Big[s(y, z; U) = \ell \mid y \in U, z \notin U\Big] = \Omega(1/\sqrt{n})$$

for ℓ around the mean. Hence, $\mathbb{P}[s(y, z; U) = \ell \mid y \in U, z \notin U] = \Theta(1/\sqrt{n})$ for $\ell \in \{-B, \dots, B\}$, and so

$$\mathbb{E}[Y_\ell] = \frac{1}{4} \sum_{(y,z) \in T} \mathbb{P}\Big[s(y, z; U) = \ell \mid y \in U, z \notin U\Big] = \Theta\Big(|T|/\sqrt{n}\Big).$$

It remains to construct such a set T and to prove (14). Both are carried out inside a linear-sized (ε, δ)-rich induced subgraph $S_0 \subseteq V(G)$. Condition 1 follows from two applications of richness inside S_0, condition 2 follows from pigeonholing. The proof of (14) further relies on a *dependent random choice* technique which supplies S_0 with a strong common-neighborhood property needed to control correlations between switches of multiple pairs in T – see [11, Lemmas 13.1, 13.3, 13.4] for details.

5 Concluding Remarks

This dissertation built progressively sharper statements about induced subgraphs of Ramsey graphs, confirming their strong resemblance to random graphs and culminating in the resolution of the Erdős–McKay conjecture. The core results all follow a similar *anticoncentration* principle that bounds the point probabilities of a random variable X associated with a Ramsey graph in roughly the same way as a local central limit theorem would:

$$\sup_{x\in\mathbb{Z}} \mathbb{P}[X = x] = O\Big(1/\sigma(X)\Big).$$

The underlying characteristic of Ramsey graphs that powers this anticoncentration principle is the *richness* property introduced in Chapter 2, which captures the pseudorandomness of Ramsey graphs and ensures linear polynomials of Bernoulli indicators carry enough non-degenerate terms for a Littlewood–Offord argument to apply.

The belief that graphs with small homogeneous sets have various random-like properties goes beyond the Ramsey graph constraint $\hom(G) = O(\log n)$. For instance, Erdős and Hajnal [9] famously conjectured that for any graph H, there exists $\varepsilon = \varepsilon(H) > 0$ such that any graph G on n vertices with $\hom(G) < n^{\varepsilon}$ contains an induced copy of H. This conjecture is widely open despite intensive research.

Alon, Krivelevich, and Sudakov [1] also conjectured a strengthening of Theorem 4.1 that $\hom(G) \le n/4$ is enough to force $\Phi(G) \supseteq \{0, 1, \ldots, \delta e(G)\}$ for some $\delta > 0$. They noted the constant $1/4$ cannot be relaxed below $1/3$ via a counterexample of G consisting of two vertex-disjoint $(n/3)$-cliques and $n/3$ isolated vertices. Kwan and Sudakov [10] also proposed the natural weakening that $\hom(G) \le n/4$ should already force $|\Phi(G)| = \Omega(e(G))$, which is open as well.

References

[1] Noga Alon, Michael Krivelevich, and Benny Sudakov. Induced subgraphs of prescribed size. *J. Graph Theory*, 43(4):239–251, August 2003. ISSN 0364-9024.

[2] Boris Bukh and Benny Sudakov. Induced subgraphs of Ramsey graphs with many distinct degrees. *Journal of Combinatorial Theory, Series B*, 97(4):612–619, 2007. ISSN 0095-8956. doi: https://doi.org/10.1016/j.jctb.2006.09.006. URL `https://www.sciencedirect.com/science/article/pii/S0095895606001080`.

[3] Neil Calkin, Alan Frieze, and Brendan McKay. On subgraph sizes in random graphs. *Combinatorics, Probability & Computing*, 1:123–134, 06 1992. doi: 10.1017/S0963548300000146.

[4] Marcelo Campos, Simon Griffiths, Robert Morris, and Julian Sahasrabudhe. An exponential improvement for diagonal ramsey, 2025. URL `https://arxiv.org/abs/2303.09521`.

[5] Paul Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53(4):292–294, 1947.

[6] Paul Erdős. Some of my favourite problems in various branches of combinatorics. *Le Matematiche; Vol 47, No 2 (1992); 231-240*, 51, 01 1993. doi: 10.1016/S0167-5060(08)70608-3.

[7] Paul Erdős and George Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.

[8] Paul Erdős and Endre Szemerédi. On a Ramsey type theorem. *Periodica Mathematica Hungarica*, 2:295–299, 1972. doi: 10.1007/BF02018669.

[9] P. Erdös and A. Hajnal. Ramsey-type theorems. *Discrete Applied Mathematics*, 25(1):37–52, 1989. ISSN 0166-218X. doi: https://doi.org/10.1016/0166-218X(89)90045-0. URL `https://www.sciencedirect.com/science/article/pii/0166218X89900450`.

[10] Matthew Kwan and Benny Sudakov. Ramsey graphs induce subgraphs of quadratically many sizes, 2021. URL `https://arxiv.org/abs/1711.02937`.

[11] Matthew Kwan, Ashwin Sah, Lisa Sauermann, and Mehtaab Sawhney. Anticoncentration in Ramsey graphs and a proof of the Erdős–McKay conjecture. *Forum of Mathematics, Pi*, 11, 2023. ISSN 2050-5086. doi: 10.1017/fmp.2023.17. URL `http://dx.doi.org/10.1017/fmp.2023.17`.

[12] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality, 2005. URL `https://arxiv.org/abs/math/0503503`.

[13] Bhargav Narayanan, Julian Sahasrabudhe, and István Tomon. Ramsey graphs induce subgraphs of many different sizes, 2017. URL https://arxiv.org/abs/1609.01705.

[14] Hans Jürgen Prömel and Vojtěch Rödl. Non-Ramsey graphs are $c \log n$-universal. *Journal of Combinatorial Theory, Series A*, 88(2):379–384, 1999. ISSN 0097-3165. doi: https://doi.org/10.1006/jcta.1999.2972. URL https://www.sciencedirect.com/science/article/pii/S0097316599929722.

[15] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286, 01 1930. ISSN 0024-6115. doi: 10.1112/plms/s2-30.1.264. URL https://doi.org/10.1112/plms/s2-30.1.264.

[16] Saharon Shelah. Erdős and Rényi conjecture, 1997. URL https://arxiv.org/abs/math/9707226.

A Code for Figure 2

```
import numpy as np
import matplotlib.pyplot as plt

n = 1000
p_edge = 0.8
p_sample = 0.5
num_samples = 10000000

A = np.zeros((n, n))
for i in range(n):
    for j in range(i + 1, n):
        if np.random.rand() < p_edge:
            A[i, j] = 1
            A[j, i] = 1

counts = np.zeros(num_samples)
for t in range(num_samples):
    u = (np.random.rand(n) < p_sample)
    counts[t] = int(u @ A @ u) // 2

fig, axes = plt.subplots(1, 2, figsize=(13, 5))

axes[0].hist(counts, bins=600, color='#1b365d')
axes[0].set_title("large scale")
axes[0].set_xlabel("e(U)")
axes[0].set_ylabel("count")

lo = int(counts.mean()) - 2500
hi = int(counts.mean()) + 2500
window = (counts >= lo) & (counts <= hi)
axes[1].hist(counts[window], bins=np.arange(lo, hi + 1), color='#1
    b365d')
axes[1].set_title("small scale")
axes[1].set_xlabel("e(U)")
axes[1].set_ylabel("count")
axes[1].margins(x=0)

plt.tight_layout()
plt.show()
```