Probabilistic and extremal behavior in graphs and matrices

by

Gweneth McKinley

B.S., University of California, Davis, 2015
A.S., Pasadena City College, 2013

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2020

Author . . . .
Department of Mathematics
February 5, 2020

Certified by . . . .
Henry Cohn
Adjunct Professor of Mathematics
Thesis Supervisor

Accepted by . . . .
Jonathan Kelner
Professor of Mathematics
Chairman, Department Committee on Graduate Theses

Probabilistic and extremal behavior in graphs and matrices

by

Gweneth McKinley

Submitted to the Department of Mathematics
on February 5, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis deals with several related questions in probabilistic and extremal graph theory and discrete random matrix theory.

First, for any bipartite graph H containing a cycle, we prove an upper bound of $2^{O(\mathrm{ex}(n,H))}$ on the number of labeled H-free graphs on n vertices, given only a fairly natural assumption on the growth rate of $\mathrm{ex}(n, H)$. Bounds of the form $2^{O(\mathrm{ex}(n,H))}$ have been proven only for relatively few special graphs H, often with considerable difficulty, and our result unifies all previously known special cases.

Next, we give a variety of bounds on the clique numbers of random graphs arising from the theory of graphons. A *graphon* is a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$, and each graphon gives rise naturally to a random graph distribution, denoted $\mathbb{G}(n, W)$, that can be viewed as a generalization of the Erdős-Rényi random graph. Recently, Doležal, Hladký, and Máthé gave an asymptotic formula of order $\log n$ for the clique number of $\mathbb{G}(n, W)$ when W is bounded away from 0 and 1. We show that if W is allowed to approach 1 at a finite number of points, and displays a moderate rate of growth near these points, then the clique number of $\mathbb{G}(n, W)$ will be $\Theta(\sqrt{n})$ almost surely. We also give a family of examples with clique number $\Theta(n^\alpha)$ for any $\alpha \in (0, 1)$, and some conditions under which the clique number of $\mathbb{G}(n, W)$ will be $o(\sqrt{n})$, $\omega(\sqrt{n})$, or $\Omega(n^\alpha)$ for $\alpha \in (0, 1)$.

Finally, for an $n \times m$ matrix M of independent Rademacher (± 1) random variables, it is well known that if $n \le m$, then M is of full rank with high probability; we show that this property is resilient to adversarial changes to M. More precisely, if $m \ge n + n^{1-\varepsilon/6}$, then even after changing the sign of $(1 - \varepsilon)m/2$ entries, M is still of full rank with high probability. This is asymptotically best possible, as one can easily make any two rows proportional with at most $m/2$ changes. Moreover, this theorem gives an asymptotic solution to a slightly weakened version of a conjecture made by Van Vu in [Vu08].

Thesis Supervisor: Henry Cohn
Title: Adjunct Professor of Mathematics

Acknowledgments

When I started on this acknowledgments section, I was planning to write something short, punchy, and to-the-point. But after thinking about all the people who have had an impact on my life during grad school, or helped me get here in the first place, I've decided to spread out a little and take advantage of the closest opportunity to an Oscar speech I'm ever likely to have. Even so, of the people who are important to me, there are far more off this list than on it. With those disclaimers out of the way...

First, I would like to thank my advisor, Henry Cohn. Thank you for sharing your infectious joy and enthusiasm for mathematics with me. Thank you for teaching me *this much* about mathematics and how to be a mathematician. Thank you for your unfailing kindness and patience. Thank you for making me feel welcomed and at home in the MIT math department from the very first day. In brief, thank you for your guidance, encouragement, advice, and support.

Next, I'd like to thank Asaf Ferber. To paraphrase ABBA, "Thank you for the math." Thank you for taking me under your wing. Thank you for teaching me how to tell a story, and thank you for pushing me to do my best.

Thank you also to the many people in the MIT math department who have supported me personally and professionally. First, to my thesis committee members, Nike Sun and Laci Lovasz (and of course, my advisor Henry Cohn), thank you for reading this thing! Next, to many faculty and staff members for supporting me through every step of graduate school; in particular, to Jonathan Harmon, thank you for your wise guidance, for many interesting conversations, and for patiently listening to several years of large and small worries and complaints. And of course, thank you to my classmates, officemates, and friends in the department; you have made it a joy to come in to the office every day.

Next, thank you to the many teachers and professors who shaped my life before MIT, and without whom I would not have pursued this path. To René Williams, thank you for teaching me how to work, and for giving me confidence in my own abilities. Thank you to the faculty at Pasadena City College, and in particular, to Carrie Starbird and Kathleen Hogue, for first showing me the beauty of mathematics, and for encouraging me to pursue it – you have given me a passion and a profession. And to Les Reid at the Missouri State University REU program, and many faculty members at UC Davis, including but not certainly not limited to Anne Schilling, Monica Vazirani, and Lawrence Marx, thank you for all of the encouragement and guidance.

I would also like to thank the people outside of the MIT math department who have made my life in graduate school better and sweeter, among them Giselle and Andrew Lilley, Lorenza Bronkhorst, and Lauren Milling. You are dear friends, and I feel very lucky to have shared this time with you. Thank you also to the students on Burton 2, the MIT Gospel Choir, the students and staff at $\sqrt{}$mathroots, and Maya Stein, Marcos Kiwi, Daniel Quiroz, and Matías Pavez Signé at the Universidad de Chile. You have brought so many good things to my life.

And last of all, I would like to thank my family. Thank you to my Mom and Dad, to my sister Corrie, and to my brother-in-law Ben and nephews Reuel and Malcolm. Thank you for always supporting and encouraging me. Thank you for the many, many hours on the phone since I've come to MIT. Thank you for your love. I love you all so much.

Contents

Chapter 1

Introduction

This thesis deals with several distinct but closely related problems in discrete mathematics, primarily in extremal and probabilistic graph theory. Formally, a graph $G = (V, E)$ is a collection of *vertices* V, together with a collection E of 2-element subsets of V, called *edges*. For example, a social network such as Facebook can be viewed as a graph, where people are represented by vertices, and edges represent friendships between pairs of people.

This work addresses two slightly different types of problems in graph theory; the first type deals with extremal conditions, loosely speaking, the limiting cases in which a graph with certain properties can exist. For example, how densely can a graph be connected before it must contain a given substructure? Or, a closely related question: if we must avoid a certain substructure, how much are we constrained in building a graph? The second type of problem analyzes the structure of a *typical* graph, where the graph is chosen at random according to some natural probability distribution. For example, in a large network described by a given probabilistic model, how large a set of fully interconnected nodes will typically exist? Chapters 2 and 3 deal respectively with problems in these two areas, and Chapter 4 presents work in the closely related area of random matrix theory.

1.1 Counting H-free graphs

Chapter 2 is based on the paper [FMS], joint with Asaf Ferber and Wojciech Samotij.

Given a graph H, its *extremal number* $\mathrm{ex}(n, H)$ is the maximum number of edges that a graph G on n vertices may have without containing a copy of H as a subgraph. Such a graph G is called *H-free*. The behavior of $\mathrm{ex}(n, H)$ is well understood for graphs with $\chi(H) \geq 3$, where $\chi(H)$ is the *chromatic number* of H, the minimum number of colors necessary to color the vertices of H so that no two adjacent vertices are of the same color [Tur41], [ES46]. However, determining the asymptotic behavior of $\mathrm{ex}(n, H)$ in the case where H is bipartite (i.e., when $\chi(X) = 2$) remains a major open problem in extremal graph theory, and has been solved in very few cases. [FS13]

The work presented here gives a partial solution to the separate but closely related problem of counting H-free graphs. Explicitly, given a graph H, the problem is to determine $|\mathcal{F}_n(H)|$, the number of labeled graphs on n vertices that do not contain a copy of H as a subgraph. This is indeed very closely tied to the problem of finding the extremal number $\mathrm{ex}(n, H)$; it can be shown easily that for any graph H,

$$2^{\mathrm{ex}(n,H)} \leq |\mathcal{F}_n(H)| \leq 2^{O(\mathrm{ex}(n,H)\log(n))}.$$

It is conjectured that, except in the case where H is a forest, the lower bound is essentially correct: namely, $|\mathcal{F}_n(H)| = 2^{O(\mathrm{ex}(n,H))}$. This was proven in 1986 by Erdős, Frankl, and Rödl for non-bipartite graphs [EFR86]. For bipartite graphs, the work presented here unifies a number of earlier results that proved this conjecture in specific cases ([KW82], [KW96], [BS11b], and [MS16]), to show that it holds for any H whose extremal number $\mathrm{ex}(n, H)$ displays sufficiently regular growth. Namely, we prove a slightly stronger version of the following result.

Theorem (Ferber, M., Samotij [FMS])**.** *If H is any graph containing a cycle, and if $\mathrm{ex}(n, H) = \Theta(n^\alpha)$ for some constant $\alpha \in (1, 2)$, then $|\mathcal{F}_n(H)| = 2^{O(\mathrm{ex}(n,H))}$.*

We also show that for a bipartite graph H, even if this regularity condition is not satisfied, given a sufficiently strong lower bound on $\mathrm{ex}(n, H)$, the conjectured number of H-free graphs on n vertices is correct for an infinite sequence of values $n \in \mathbb{N}$. In addition, we prove generalizations of both results to hypergraphs.

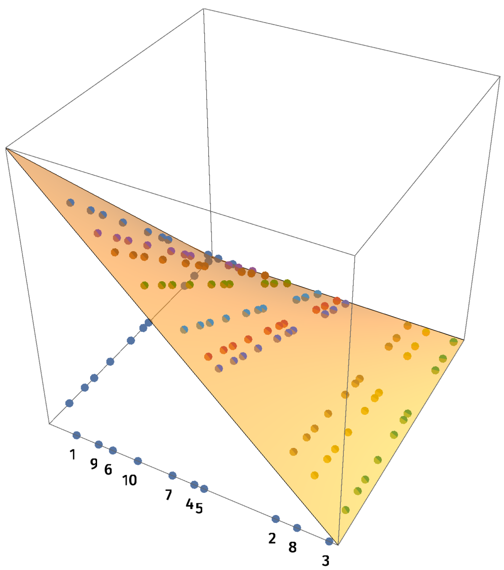
1.2 Random graphs generated from graphons

Chapter 3 is based on the paper [McK19].

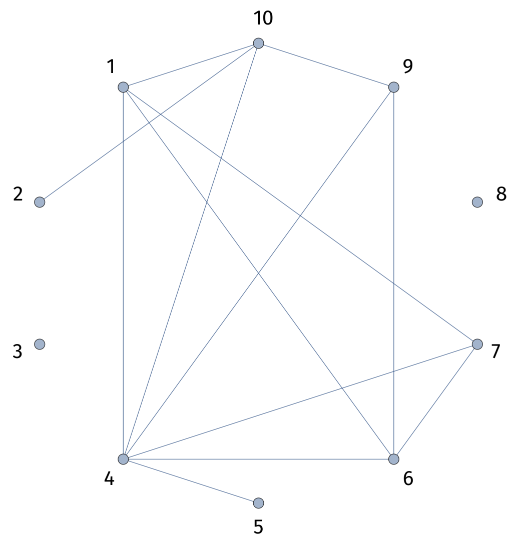
As just outlined, Chapter 2 will explore problems related to the extremal properties that a graph may possess. But one may also ask about the properties of a *typical* graph, averaging over graphs selected according to some natural probability distribution. Perhaps the simplest such distribution is given by the Erdős-Rényi random graph model. The Erdős-Rényi random graph $G_{n,p}$ is a graph on n vertices where, between each pair of vertices, an edge is placed independently with probability p. Since its introduction in 1959 by Gilbert [Gil59] and by Erdős and Rényi [ER59], it has become one of the fundamental objects of study in probabilistic combinatorics, and a wide variety of its properties are very well understood, including degree distribution, chromatic number, size of the largest connected component, and clique number. Of particular interest here, it was shown in [GM75] and [Mat76] that for a fixed $p \in (0, 1)$, the *clique number* of $G_{n,p}$ (the number of vertices in the largest complete subgraph of G) approaches $2 \log n \cdot \log(1/p)^{-1}$ asymptotically almost surely (abbreviated "a.a.s" hereafter).

The Erdős-Rényi random graph $G_{n,p}$ may be considered "homogeneous" in the following sense: between each pair of vertices, an edge is assigned with the *same* probability p. In recent years, interest has been developing in studying inhomogeneous random graphs, where edges are assigned between some pairs with higher or lower probabilities. Specifically, an inhomogeneous dense random graph model arises from the theory of graph limits. This theory studies objects called *graphons*; symmetric measurable functions $W : [0,1]^2 \to [0,1]$. To obtain a random graph $\mathbb{G}(n, W)$ from W, we sample points $x_1, \ldots, x_n$ uniformly from $[0,1]$, and connect vertices i and j by an edge with probability $W(x_i, x_j)$. This can be thought of as a generalization of $G_{n,p}$, which is equivalent to $\mathbb{G}(n, W)$ for the constant graphon $W = p$.

As an example, we illustrate a realization of $x_1, \ldots, x_n$ and $\mathbb{G}(n, W)$ for the graphon $W(x, y) = (1 - x)(1 - y)$, taking $n = 10$ (note: for legibility, only the subscripts are shown for the points $x_1, \ldots, x_{10}$).



$W(x, y) = (1 - x)(1 - y)$



$\mathbb{G}(n, W)$

With this theory in place, it is natural to ask whether we can determine the clique number of an inhomogeneous random graph $\mathbb{G}(n, W)$ for a given graphon W. In 2017, Doležal, Hladkỳ, and Máthé [DHM19] showed that for a graphon W that is bounded away from 1, if a certain technical condition is satisfied, then a.a.s.,

$$\omega(\mathbb{G}(n, W)) = (1 + o(1)) \cdot \kappa(W) \log n,$$

for a specific constant $\kappa(W)$, which they determined. Here, we consider the case of graphons W that are allowed to approach 1 at a finite number of points; we show that if such a graphon W displays a moderate rate of growth near these points, then the clique number of $\mathbb{G}(n, W)$ will be $\Theta(\sqrt{n})$ almost surely:

Theorem (M., [McK19])**.** *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at a collection of points $(a_1, a_1), \ldots, (a_k, a_k)$, and essentially bounded away from 1 in some neighborhood of (x, x) for all other $x \in [0,1]$. If all directional derivatives of W exist at the points $(a_1, a_1), \ldots, (a_k, a_k)$, and are uniformly bounded away from 0 and $-\infty$, then $\omega(\mathbb{G}(n, W)) = \Theta(\sqrt{n})$ a.a.s.*

We also give a family of examples with clique number $\Theta(n^\alpha)$ for any $\alpha \in (0,1)$, and some conditions (weaker than the existence of directional derivatives) under which the clique number of $\mathbb{G}(n, W)$ will be $o(\sqrt{n})$, $\omega(\sqrt{n})$, or $\Omega(n^\alpha)$ for $\alpha \in (0,1)$.

1.3 Discrete random matrices

Chapter 4 is based on the paper [FLM19], joint with Asaf Ferber and Kyle Luh.

Random discrete matrices, in particular 0/1 and ± 1 random matrices, have a distinguished history in random matrix theory. They have applications in computer science, physics, and random graph theory, among other areas, and numerous investigations have been tailored to this class of random matrices [BVW10, KKS95, Kom67, Ngu13, TV07, TV09, Tik]. Discrete random matrices are also of interest in their own right, as they pose combinatorial questions that are vacuous or trivial for other models such as the gaussian ensembles (e.g. singularity and simpleness of spectrum). For example, denoting by $M_{n,m}$ an $n \times m$ matrix with independent uniform ± 1 entries, it is already non-trivial to show that $M_{n,n}$ is non-singular with probability $1 - o(1)$ (this was first proved by Komlós in [Kom67]). It was a long standing conjecture that

$$p_n := \Pr(M_{n,n} \text{ is singular}) = \left(\frac{1}{2} + o(1)\right)^n,$$

which corresponds to the probability that any two rows or columns are identical. This problem has stimulated a great deal of activity [KKS95, TV07, BVW10], culminating in the recent resolution by Tikhomirov [Tik] of the above conjecture.

Here, we examine another aspect of the singularity problem for discrete random matrices. We look at the robustness of the non-singularity, meaning how many changes to the entries of the matrix need to be performed to make it singular. This has been called the "resilience" of a random matrix with respect to singularity [Vu08]. Note that an $n \times n$ matrix in singular if and only if its rank is less than n. Therefore, we may extend this notion to general matrices (not necessarily square) as follows:

Definition. *Given an $n \times m$ matrix M with entries in $\{\pm 1\}$, and with $m \geq n$, we denote by $\operatorname{Res}(M)$ the minimum number of sign flips necessary in order to make M of rank less than n.*

Note that for any two vectors $\boldsymbol{a}, \boldsymbol{b} \in \{\pm 1\}^m$ one can always achieve either $\boldsymbol{a} = \boldsymbol{b}$ or $\boldsymbol{a} = -\boldsymbol{b}$ by changing at most $m/2$ entries; so in particular, for an $n \times m$ matrix M, we have the deterministic upper bound $\operatorname{Res}(M) \leq m/2$. For the case $n = m$, it was conjectured by Vu [Vu08] that this is essentially tight; i.e., that

$$\operatorname{Res}(M_{n,n}) = \left(\frac{1}{2} + o(1)\right) n$$

a.a.s. Note that by a a simple union bound, using any exponential upper bound on p_n, one can easily show that a.a.s. we have $\operatorname{Res}(M_{n,n}) \geq cn/\log n$ for some appropriate choice of $c > 0$. Perhaps surprisingly, no better lower bound is known.

Here, we prove the following weakening of Vu's conjecture, resolving it in the case of random matrices that are a little wider than square.

Theorem (Ferber, Luh, M. [FLM19])**.** *For every $\varepsilon > 0$ and $m \geq n + n^{1-\varepsilon/6}$, a.a.s. we have*

$$Res(M_{n,m}) \geq (1 - \varepsilon)m/2.$$

The proof relies on some recent machinery introduced in [FJLS19], and uses the following basic strategy: first, notice that $M_{n,m}$ has rank less than n precisely if there is a nonzero vector $\boldsymbol{a}$ such that $\boldsymbol{a}^T M_{n,m} = \mathbf{0}$. We use the first n columns of $M_{n,m}$ to show that any such $\boldsymbol{a}$ will be "well-unstructured" or "pseudorandom" in some precise sense, making use of the machinery from [FJLS19] to consider some portions of this argument over a finite field $\mathbb{F}_p$. We then show that any "pseudorandom" vector $\boldsymbol{a}$ is unlikely to also be orthogonal to the remaining $m - n$ columns.

Chapter 2

Counting H-free graphs

2.1 Introduction

This chapter is based on the paper [FMS], joint with Asaf Ferber and Wojciech Samotij.

The *extremal number* of a graph H, denoted by $\mathrm{ex}(n, H)$, is the maximum possible number of edges in a graph G on n vertices which does not contain H as a (not necessarily induced) subgraph. Such a graph G is referred to as *H-free*. The study of the asymptotic behavior of $\mathrm{ex}(n, H)$ for various H is a central theme in extremal graph theory and goes back to the pioneering work of Turán [Tur41], who determined $\mathrm{ex}(n, H)$ exactly in the case when H is a complete graph. In fact, Turán's construction provides a lower bound on $\mathrm{ex}(n, H)$ that depends on the *chromatic number* of H, denoted by $\chi(H)$, which is the least integer k for which one can partition $V(H)$ into k *independent sets* (that is, sets which induce no edges). More precisely, Turán's construction gives

$$\mathrm{ex}(n, H) \geqslant \left(1 - \frac{1}{\chi(H) - 1}\right)\binom{n}{2}$$

for every nonempty graph H. A matching upper bound was proved several years later by Erdős and Stone [ES46], giving

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1}\right)\binom{n}{2} + o(n^2). \tag{2.1}$$

Note that (2.1) determines the asymptotics of $\mathrm{ex}(n, H)$ whenever $\chi(H) \geqslant 3$, but when $\chi(H) = 2$, that is, when H is *bipartite*, it only implies that $\mathrm{ex}(n, H) = o(n^2)$, whereas Turán's construction gives the trivial bound $\mathrm{ex}(n, H) \geqslant 0$.

Perhaps unsurprisingly, the bipartite case of Turán's problem is much more challenging and there are only a few bipartite graphs H for which even the order of magnitude of $\mathrm{ex}(n, H)$ has been determined. Among the known examples one can find trees, cycles of lengths four, six, and ten, and the complete bipartite graphs $K_{s,t}$ when $s \in \{2, 3\}$ or $t > (s-1)!$. For a generic bipartite H, there does not even seem to be a good guess for what $\mathrm{ex}(n, H)$ might be. The lower bounds in all the above examples are established by rather involved algebraic or geometric constructions. The strongest general upper bound on $\mathrm{ex}(n, H)$ is due to Füredi [Für91] who proved that $\mathrm{ex}(n, H) = O(n^{2-1/D})$ if all but one of the vertices in one of the color classes of some proper two-coloring of H have degree at most D. This generalizes the classical result of Kővári, Sós, and Turán [KST54], who showed that $\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/s})$ for all s and t. Treating a more general class of graphs than the one considered in [Für91], Alon, Krivelevich, and Sudakov [AKS03] proved that $\mathrm{ex}(n, H) = O(n^{2-1/4D})$ for every D-*degenerate* bipartite graph H (a graph is D-degenerate if every subgraph of it has minimum degree at most D). For a more detailed discussion and further references, we refer the reader to the excellent survey of Füredi and Simonovits [FS13].

Here we shall be concerned with the closely related problem of enumerating H-free graphs. That is, we are interested in the asymptotic size of the set $\mathcal{F}_n(H)$ consisting of all (labeled) H-free graphs with vertex set $[n] := \{1, \ldots, n\}$. Observing that every subgraph of an H-free graph is also H-free and that every n-vertex H-free graph has at most $\mathrm{ex}(n, H)$ edges, one obtains the trivial bounds

$$2^{\mathrm{ex}(n,H)} \leqslant |\mathcal{F}_n(H)| \leqslant \sum_{k=0}^{\mathrm{ex}(n,H)} \binom{\binom{n}{2}}{k} \leqslant \left(\frac{e\binom{n}{2}}{\mathrm{ex}(n,H)} \right)^{\mathrm{ex}(n,H)}. \tag{2.2}$$

This counting problem has been widely studied, and when H is not bipartite, bounds much tighter than (2.2) are known. It was proved by Erdős, Kleitman, and Rothschild [EKR76] (when H is a complete graph, but implicitly also for every non-bipartite H) and then by Erdős, Frankl, and Rchap1ödl [EFR86] that

$$|\mathcal{F}_n(H)| = 2^{\mathrm{ex}(n,H)+o(n^2)}. \tag{2.3}$$

In particular, if $\chi(H) \geqslant 3$, then (2.1), (2.3), and the lower bound in (2.2) imply that $|\mathcal{F}_n(H)| = 2^{(1+o(1))\mathrm{ex}(n,H)}$. On the other hand, if H is bipartite, then (2.3) is very weak and the trivial upper bound in (2.2) is still the state-of-the-art bound for a generic graph H (up to a constant multiplicative factor in the exponent), giving

$$2^{\mathrm{ex}(n,H)} \leqslant |\mathcal{F}_n(H)| \leqslant 2^{C\mathrm{ex}(n,H)\log n} \tag{2.4}$$

for some positive constant C that depends only on H. It is natural to ask whether the $\log n$ factor in the above upper bound can be removed. Indeed, this question was posed by Erdős some thirty five years ago (see [KW82]) for all bipartite H that contain a cycle.[1] Until very recently, it was even believed that the stronger bound $|\mathcal{F}_n(H)| \leqslant 2^{(1+o(1))\mathrm{ex}(n,H)}$ holds, as it does for non-bipartite H, but this was disproved by Morris and Saxton [MS16] in the case when H is the cycle of length six. In view of this, the following seems to be the right question to ask.

[1] The case when H has no cycles is very different as then $\mathrm{ex}(n, H) = O(n)$ while there could be n-vertex H-free graphs with as many as $n!$ different labelings. In particular, since there are $2^{\Omega(n \log n)}$ different labeled n-vertex graphs with maximum degree one, then $|\mathcal{F}_n(H)| \geqslant 2^{\Omega(\mathrm{ex}(n,H)\cdot\log n)}$ for every acyclic H with maximum degree at least two.

Question 2.1.1. *Suppose that H is a bipartite graph which contains a cycle. Is there a constant C such that*

$$|\mathcal{F}_n(H)| \leqslant 2^{C\mathrm{ex}(n,H)}$$

for all n?

Despite renewed interest in Question 2.1.1 in recent years, very little is known. To the best of our knowledge, it has been answered positively only in the cases when H is the cycle of length four [KW82], six [KW96], or ten [MS16], the complete bipartite graph $K_{s,t}$ with $s \in \{2,3\}$ or $t > (s-1)!$ (see [BS11a, BS11b]), or so-called theta-graphs [CT17]. Here we make a first attempt at addressing Question 2.1.1 for a generic bipartite graph H. Our methods also extend to the setting of uniform hypergraphs, which we shall discuss at the end of this section. The following is our first main result:

Theorem 2.1.2. *Let H be an arbitrary graph containing a cycle. Suppose that there are positive constants α and A such that $\mathrm{ex}(n,H) \leqslant An^{\alpha}$ for all n. Then there exists a constant C depending only on α, A, and H such that for all n,*

$$|\mathcal{F}_n(H)| \leqslant 2^{Cn^{\alpha}}.$$

Note that Theorem 2.1.2 answers Question 2.1.1 in the affirmative for every bipartite H such that $\mathrm{ex}(n,H) = \Theta(n^{\alpha})$ for some α. This is the case for each H for which Question 2.1.1 has been answered so far and therefore Theorem 2.1.2 reproves all the previously known results listed above. In fact, it is commonly believed that $\mathrm{ex}(n,H) = \Theta(n^{\alpha})$ for all bipartite H, as conjectured by Erdős and Simonovits (see for example [Erd81]):

Conjecture 2.1.3. *For every nonempty bipartite graph H, there exist a rational number $\alpha \in [1,2)$ and $c > 0$ such that*

$$\frac{\mathrm{ex}(n,H)}{n^{\alpha}} \to c.$$

Observe that if Conjecture 2.1.3 is true, then Theorem 2.1.2 resolves Question 2.1.1 for all H. Actually, the following weaker version of Conjecture 2.1.3 is sufficient. However, a solution to either of these conjectures is most likely unattainable in the near future.

Conjecture 2.1.4. *For every nonempty bipartite graph H, there exist $\alpha \in [1,2]$ and $c_2 > c_1 > 0$ such that*

$$c_1 \leqslant \frac{\mathrm{ex}(n,H)}{n^{\alpha}} \leqslant c_2$$

On a related note, we would like to mention a recent breakthrough of Bukh and Conlon [BC18], who used a random algebraic method, pioneered by Bukh [Buk15], to prove the following "inverse" version of Conjecture 2.1.4: for every rational $\alpha \in [1,2)$, there exists a finite family of graphs $\mathcal{L}$ for which $\mathrm{ex}(n,\mathcal{L}) = \Theta(n^{\alpha})$ (where $\mathrm{ex}(n,\mathcal{L})$ is the maximum possible number of edges in an n-vertex graph that does not contain any member of the family $\mathcal{L}$).

There are bipartite graphs H for which the best known upper bound on $\mathrm{ex}(n, H)$ is of the form $O(n^\alpha)$, for some explicit α, and is conjectured to be tight. For such graphs, it makes sense to establish the bound $|\mathcal{F}_n(H)| \leqslant 2^{O(n^\alpha)}$. Indeed, such results have been proved for even cycles [MS16], complete bipartite graphs [BS11a, BS11b], and theta graphs [CT17]. All these estimates follow as simple corollaries of Theorem 2.1.2 and the corresponding upper bounds on the extremal numbers [BS74, FS83, KST54].

Even if the asymptotic behavior of $\mathrm{ex}(n, H)$ is unknown, assuming a sufficiently strong lower bound on it, in Theorem 2.1.5, we are able to prove strong estimates for $|\mathcal{F}_n(H)|$ for an infinite sequence of n. A similar result for the number of k-arithmetic-progression-free subsets of $[n]$ was obtained by Balogh, Liu, and Sharifzadeh [BLS17]. This result served as an inspiration for our work. Before formally stating the theorem, we recall the notion of 2-*density* of a graph H:

$$m_2(H) := \max \left\{ \frac{e_F - 1}{v_F - 2} \colon F \subseteq H, v_F > 2 \right\}.$$

Theorem 2.1.5. *Let H be a graph and assume that $\mathrm{ex}(n, H) \geqslant \varepsilon n^{2-1/m_2(H)+\varepsilon}$ for some $\varepsilon > 0$ and all n. Then there exist a constant C depending only on ε and H and an infinite sequence of n for which*

$$|\mathcal{F}_n(H)| \leqslant 2^{C \cdot \mathrm{ex}(n,H)}.$$

The assumption on H stated in Theorem 2.1.5 is widely believed to hold for every H containing a cycle. In fact, it is known to hold for quite a few bipartite graphs. For example, it is known that for every ℓ,

$$\mathrm{ex}(n, C_{2\ell}) \geqslant \Omega \left(n^{1+\frac{2}{3\ell+3}} \right) = \Omega \left(n^{2-1/m_2(C_{2\ell})+\varepsilon_\ell} \right),$$

where $\varepsilon_\ell > 0$; see, for example, Terlep and Williford [TW12] and the references therein (in particular, the famous papers of Margulis [Mar88] and Lubotzky, Phillips, and Sarnak [LPS88]). To give another example, consider the case when H is the 3-dimensional hypercube graph Q_3. Theorem 2.1.5 applies to H because $2-1/m_2(Q_3) = 2 - 6/11 < 3/2$ and $\mathrm{ex}(n, Q_3) \geqslant \mathrm{ex}(n, C_4) = \Omega(n^{3/2})$. As a third example, note that $\mathrm{ex}(n, K_{4,4}) \geqslant \mathrm{ex}(n, K_{3,3}) = \Omega(n^{5/3})$ and $5/3 > 2 - 7/15 = 2 - 1/m_2(K_{4,4})$ and thus Theorem 2.1.5 also applies with $H = K_{4,4}$. Finally, it follows from the work of Ball and Pepe [BP12] that $K_{5,5}$ and $K_{6,6}$ also satisfy the assumptions of Theorem 2.1.5.

One may consider a natural extension of Question 2.1.1 to the setting of uniform hypergraphs, where $\mathrm{ex}(n, H)$ and $\mathcal{F}_n(H)$ are defined in the obvious way. However, the problem of enumerating hypergraphs without a forbidden subhypergraph has only been addressed fairly recently. Generalizing (2.3), Nagle, Rödl, and Schacht [NRS06] proved that for each r-uniform hypergraph H,

$$|\mathcal{F}_n(H)| = 2^{\mathrm{ex}(n,H)+o(n^r)}. \tag{2.5}$$

Analogously to the graph case, it is easy to see that an r-uniform hypergraph H that is not r-partite[2] satisfies $\mathrm{ex}(n, H) = \Omega(n^r)$. On the other hand, extending the result of Kővári, Sós, and Turán [KST54] to hypergraphs, Erdős [Erd64] proved that for every r-partite r-uniform H, there is an $\varepsilon > 0$ such that $\mathrm{ex}(n, H) = O(n^{r-\varepsilon})$. In particular, (2.5) implies that $|\mathcal{F}_n(H)| = 2^{(1+o(1))\mathrm{ex}(n,H)}$ for all non-r-partite r-uniform H, but it gives a very weak bound when H is r-partite. Therefore, the right generalization of Question 2.1.1 to the setting of hypergraphs with uniformity larger than two seems to be the following:

Question 2.1.6. *Suppose that $r \geqslant 3$ and suppose that H is an r-partite r-uniform hypergraph. Under what conditions can one expect the existence of a constant C such that*

$$|\mathcal{F}_n(H)| \leqslant 2^{C\mathrm{ex}(n,H)}$$

for all n?

As mentioned above, our proof method applies to hypergraphs and both Theorems 2.1.2 and 2.1.5 extend to this setting. Before stating them formally, we need the following definition, which generalizes the notion of 2-density to hypergraphs. The *r-density* of an r-uniform hypergraph H, denoted by $m_r(H)$, is defined by

$$m_r(H) = \max\left\{\frac{e_F - 1}{v_F - r} \colon F \subseteq H,\ v_F > r\right\}.$$

The hypergraph analog to Theorem 2.1.2 is the following:

Theorem 2.1.7. *Let H be an r-uniform hypergraph and let α and A be positive constants. Suppose that $\alpha > r - 1/m_r(H)$ and that $\mathrm{ex}(n, H) \leqslant An^\alpha$ for all n. Then there exists a constant C depending only on α, A, and H such that for all n,*

$$|\mathcal{F}_n(H)| \leqslant 2^{Cn^\alpha}.$$

The idea of investigating Question 2.1.6 was suggested in a recent work of Mubayi and Wang [MW19]. They conjectured that Question 2.1.6 has an affirmative answer in the case when H is $C_k^{(r)}$, the r-uniform expansion[3] of C_k, the (2-uniform) cycle of length k. Improving upon the result from [HK18, MW19], Balogh, Narayanan, and Skokan [BNS19] have recently solved the conjecture of Wang and Mubayi. As immediate corollaries from Theorem 2.1.7 we reprove this result along with two related estimates for expansions of paths and complete bipartite graphs. For further reading about Turán problems for graph expansions, we refer the reader to a recent survey of Mubayi and Verstraëte [MV16] and the references therein. Here is a summary of our results:

Corollary 2.1.8. *Suppose that H is any one of the following:*

[2] An r-uniform hypergraph H is r-partite if its vertex set admits a partition into r parts such that every edge of H contains one vertex from each of the parts.

[3] Given a graph G and an integer $r \geqslant 3$, we define the r-uniform *expansion* of G to be the hypergraph $G^{(r)}$ with edge set $\{e \cup S_e \colon e \in E(G)\}$, where $\{S_e\}_{e \in E(G)}$ are pairwise disjoint $(r-2)$-element sets disjoint from $V(G)$.

1. $P_k^{(r)}$ *for some* $k, r \geqslant 3$*, or*

2. $C_k^{(r)}$ *for some* $k, r \geqslant 3$*, or*

3. $K_{s,t}^{(3)}$ *for some* $s, t \geqslant 3$ *with* $t > (s-1)!$.

Then, there exists a constant C *depending only on* H *such that for all* n*,*

$$|\mathcal{F}_n(H)| \leqslant 2^{C \cdot \mathrm{ex}(n,H)}.$$

The straightforward verification of the fact that the three families of hypergraphs from the statement of Corollary 2.1.8 satisfy the assumptions of Theorem 2.1.7 is left to the reader. We conclude with the following analog of Theorem 2.1.5 in the hypergraph setting.

Theorem 2.1.9. *Let* H *be an* r*-uniform hypergraph and assume that* $\mathrm{ex}(n, H) \geqslant \varepsilon n^{r-1/m_r(H)+\varepsilon}$ *for some* $\varepsilon > 0$ *and all* n*. Then there exist a constant* C *depending only on* ε *and* H *and an infinite sequence of* n *for which*

$$|\mathcal{F}_n(H)| \leqslant 2^{C \cdot \mathrm{ex}(n,H)}.$$

All of our theorems are obtained as (more or less) simple corollaries of the more general but somewhat technical Theorem 2.4.1, which is stated (and proved) in Section 2.4.

The rest of the chapter is organized as follows: First, in Section 2.2 we give a short discussion of our proof method, including some comments about previous work. Then, in Section 2.3, we present the main tool to be used in our proofs, Lemma 2.3.2, which is a version of a similar lemma from [MS16] and is based on the method of hypergraph containers developed in [BMS15, ST15]. Next, in Section 2.4, we introduce our main technical theorem, Theorem 2.4.1, a "balanced supersaturation" result that complements Lemma 2.3.2. Finally, in Sections 2.5 and 2.6, we prove Theorems 2.1.2 and 2.1.7 and Theorems 2.1.5 and 2.1.9, respectively.

2.2 Discussion

As we have mentioned in the introduction, enumeration problems in the context of forbidden (hyper)graphs have been successfully addressed for non-bipartite graphs [BBS04, EFR86, EKR76] and non-r-partite r-uniform hypergraphs [NRS06]. A main difficulty in extending the results of [BBS04, EFR86, EKR76] to the bipartite case is that the proofs in [BBS04, EFR86] are based on Szemerédi's regularity lemma. Even though there are now sparse versions of the regularity lemma, it is unlikely that the regularity approach could be used for counting graphs without a bipartite subgraph. The proof method of [EKR76] is different, but it hinges on the fact that for non-bipartite H, the number of edges in most graphs in $\mathcal{F}_n(H)$ is $n^{2-o(1)}$; this is no longer true when H is bipartite. In the case of r-uniform hypergraphs ($r \geqslant 3$), the situation is even more complicated, as a hypergraph regularity lemma which is sufficiently strong to address the enumeration problem was proved only relatively recently and is quite involved.

A nowadays standard way of tackling enumeration problems of this type is by using the method of *hypergraph containers*. This method was introduced by Balogh, Morris, and Samotij [BMS15] and, independently, by Saxton and Thomason [ST15]. In particular, it can be used to reprove (2.5) for all r-uniform H in a simple way. The container method essentially reduces the problem of establishing upper bounds on $|\mathcal{F}_n(H)|$ to proving the following statement: If an n-vertex graph contains "slightly more" than $\mathrm{ex}(n, H)$ edges, then it has "many" copies of H (such property is known as *supersaturation*) that are moreover "well-distributed".

Keeping this in mind, it seems hopeless to provide a general solution to the counting problem, as it seems crucial to know the order of magnitude of $\mathrm{ex}(n, H)$ in order to establish a sufficiently strong supersaturation result. However, Balogh, Liu, and Sharifzadeh [BLS17] recently managed to settle a question that has a similar flavor without knowing the corresponding extremal function. Specifically, they showed that for infinitely many n, there are $2^{\Theta(\Gamma_k(n))}$ many subsets of $[n]$ that do not contain an arithmetic progression of length k; here $\Gamma_k(n)$ is the largest cardinality of a subset of $[n]$ without a k-term arithmetic progression. We have found this result very surprising, as the asymptotic behavior of $\Gamma_k(n)$ is unknown. It motivated us to investigate whether similar estimates can be obtained for the problem of counting H-free graphs. A fact that was crucially used in [BLS17] is that every pair of integers is contained in a constant number of k-term arithmetic progressions. This is not the case with copies of a fixed graph H in a large complete graph (and pairs of edges of this complete graph) and this was one of the main challenges that we had to overcome.

The main contribution of this work is a general supersaturation theorem for r-uniform r-partite hypergraphs, Theorem 2.4.1 below. Roughly speaking, it states the following. Suppose that $\mathrm{ex}(n, H) = O(n^\alpha)$ for some α such that the expected number of copies of (the densest subgraph of) the forbidden hypergraph H in the random hypergraph with n vertices and n^α edges is of larger order of magnitude than n^α. Then every n-vertex hypergraph with at least n^α edges contains "many" copies of H which are "well-distributed". Although the number of copies of H that we can guarantee is still very far from the value conjectured by Erdős and Simonovits [ES84], the lower bound we prove for this quantity is sufficiently strong to allow us to derive a strong upper bound on $|\mathcal{F}_n(H)|$ using the container method. This was in fact already observed by Morris and Saxton [MS16], who formulated the following conjecture and showed that it implies a positive answer to Question 2.1.1. For an r-uniform hypergraph $\mathcal{H}$ and $1 \leqslant \ell \leqslant r$, let $\Delta_\ell(\mathcal{H})$ be the maximum number of hyperedges of $\mathcal{H}$ that contain a given set of ℓ vertices.

Conjecture 2.2.1 ([MS16, Conjecture 1.6])**.** *Given a bipartite graph H, there exist constants $C > 0$, $\varepsilon > 0$, and $k_0 \in \mathbb{N}$ such that the following holds. Let $k \geqslant k_0$ and suppose that G is a graph on n vertices with $k \cdot \mathrm{ex}(n, H)$ edges. Then there exists a* *(non-empty) collection $\mathcal{H}$ of copies of H in G, satisfying*

$$\Delta_\ell(\mathcal{H}) \leqslant \frac{C \cdot e(\mathcal{H})}{k^{(1+\varepsilon)(\ell-1)}} \text{ for all } 1 \leqslant \ell \leqslant e_H.$$

Although we have not succeeded in resolving Conjecture 2.2.1, our Theorem 2.4.1 shows that the "balanced supersaturation" property asserted by it holds for every graph H for which Conjecture 2.1.4 (or the stronger Conjecture 2.1.3) is true.

2.3 A container lemma

Let H be an r-uniform hypergraph and let $\mathcal{H}$ denote the e_H-uniform hypergraph whose vertex set is the edge set of the complete r-uniform n-vertex hypergraph $K_n^{(r)}$ and whose hyperedges are (the edge sets of) all copies of H in $K_n^{(r)}$. Note that the edge set of every H-free hypergraph on n vertices corresponds to an *independent set* in $\mathcal{H}$ and vice versa. Therefore, any upper bound on the number of independent sets in $\mathcal{H}$ yields an upper bound on the number of H-free hypergraphs.

In order to obtain the desired bound on the number of independent sets, we will use a version of the container lemma due to Balogh, Morris, and Samotij [BMS15, Proposition 3.1]. Roughly speaking, the lemma states that if the edges of a uniform hypergraph $\mathcal{H}$ are "well-distributed", then the following holds: there is a "relatively small" collection $\mathcal{C}$ of subsets of $V(\mathcal{H})$ (referred to as *containers*), each of which induces "not too many" hyperedges, such that every independent set of $\mathcal{H}$ is a subset of at least one container. Here is the formal statement:

Proposition 2.3.1 (Container lemma [BMS15, Proposition 3.1])**.** *Let $\mathcal{H}$ be a k-uniform hypergraph and let K be a constant. There exists a constant δ depending only on k and K such that the following holds. Suppose that for some $p \in (0,1)$ and all $\ell \in \{1, \dots, k\}$,*

$$\Delta_\ell(\mathcal{H}) \leqslant K \cdot p^{\ell-1} \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})}. \tag{2.6}$$

Then, there exists a family $\mathcal{C} \subseteq \mathcal{P}(V(\mathcal{H}))$ of containers *with the following properties:*

(i) $|\mathcal{C}| \leqslant \binom{v(\mathcal{H})}{\leqslant kpv(\mathcal{H})} \leqslant \left(\frac{e}{kp}\right)^{kpv(\mathcal{H})}$,

(ii) $|G| \leqslant (1-\delta) \cdot v(\mathcal{H})$ *for each* $G \in \mathcal{C}$,

(iii) each independent set of $\mathcal{H}$ is contained in some $G \in \mathcal{C}$.

Clearly, the smaller the p we choose, the stronger the upper bound on the number of containers. On the other hand, as we decrease p, it becomes more difficult to satisfy the "density" condition (2.6).

To illustrate how the container lemma can be applied in our setting, let us assume that we have an upper bound of $O(M)$ on the largest size of a container and that (2.6) is fulfilled with p satisfying $p \log \frac{1}{p} = O(M/v(\mathcal{H}))$. Then, we immediately obtain

$$|\mathcal{F}_n(H)| \leqslant |\mathcal{C}| \cdot 2^{O(M)} = 2^{O(M)}.$$

Since one does not obtain strong bounds on the largest size of a container after one application of Proposition 2.3.1, it is natural to iterate it. Specifically, given a candidate G for a final container, we can either decide to keep it (if G is small enough for our purposes) or invoke Proposition 2.3.1 to the induced subhypergraph $\mathcal{H}[G]$ to break G down further. In order for this recursive process not to produce too many containers, we must prove that $\mathcal{H}[G]$ fulfills (2.6) with a "relatively small" p. Unfortunately, since we do not know anything about the structure of G, such a statement might be very hard, or even impossible to prove.

In order to overcome this difficulty, we employ the following simple, yet powerful strategy that was first used in this context by Morris and Saxton [MS16]. Given any subhypergraph $\mathcal{H}_G \subseteq \mathcal{H}[G]$, every independent set in $\mathcal{H}[G]$ is also independent in $\mathcal{H}_G$. Hence, any upper bound on the number of independent sets in $\mathcal{H}_G$ is also an upper bound on the number of independent sets in $\mathcal{H}[G]$. It thus follows that even if $\mathcal{H}[G]$ does not fulfill (2.6), we might hope to find a suitable subhypergraph $\mathcal{H}_G \subseteq \mathcal{H}[G]$ which does satisfy this condition, enabling us to continue the iteration.

With this strategy in mind, we are first going to show how the existence of such $\mathcal{H}_G$ for every G implies the desired upper bound on the number of independent sets. A similar statement appears in [MS16], but since we consider hypergraphs here as well (as opposed to [MS16]), for the convenience of the reader and in order to keep this work self-contained, we include a full proof.

Lemma 2.3.2. *Let H be a nonempty r-uniform hypergraph, let $n \in \mathbb{N}$, and let $\mathcal{H}$ be the e_H-uniform hypergraph comprising (the edge sets of) all copies of H in $K_n^{(r)}$. Let K be a constant and let $\gamma = \frac{1}{1-\delta}$, where $\delta := \delta(e_H, K)$ is defined in Proposition 2.3.1. Suppose that M and t_0 are such that the following holds: for all integers $t \geqslant t_0$ and all $G \subseteq V(\mathcal{H})$ satisfying*

$$\gamma^t M < |G| \leqslant \gamma^{t+1} M,$$

there exists a nonempty subhypergraph $\mathcal{H}_G \subseteq \mathcal{H}[G]$ for which

$$\Delta_\ell(\mathcal{H}_G) \leqslant K \cdot \left(\frac{b_t}{|G|}\right)^{\ell-1} \cdot \frac{e(\mathcal{H}_G)}{|G|} \tag{2.7}$$

where $b_t = \frac{M}{(t+1)^3}$, for all $\ell \in \{1, \ldots, e_H\}$. Then there is a constant C depending only on K, t_0, and e_H such that $|\mathcal{F}_n(H)| \leqslant 2^{C \cdot M}$.

Remark 2.3.3. *Since $\Delta_{e_H}(\mathcal{H}_G) = 1$ whenever G is nonempty, inequality (2.7) in particular implies a lower bound on $e(\mathcal{H}_G)$, which in turn is a lower bound on the number of copies of H in G.*

Proof. We are going to prove the claimed upper bound on $|\mathcal{F}_n(H)|$ by constructing a collection of $2^{O(M)}$ containers for independent sets in $\mathcal{H}$, each of size $O(M)$. We start with the trivial container $V(\mathcal{H})$ which we break down into smaller containers by repeatedly applying Proposition 2.3.1 to the subhypergraphs $\mathcal{H}_G$ from the assumption of the lemma. Formally, we shall construct a rooted tree $\mathcal{T}$ whose vertices are subsets of $V(\mathcal{H})$, that is, subgraphs of $K_n^{(r)}$, with the following properties:

(T1) The root of $\mathcal{T}$ is $V(\mathcal{H})$.

(T2) If G is a non-leaf vertex of $\mathcal{T}$, then every independent set of $\mathcal{H}[G]$ is an independent set of $\mathcal{H}[G']$ for some child G' of G in $\mathcal{T}$.

(T3) Every leaf of $\mathcal{T}$ is a subset of $V(\mathcal{H})$ with at most $\gamma^{t_0}M$ elements.

The existence of such a tree $\mathcal{T}$ clearly implies that

$$|\mathcal{F}_n(H)| \leqslant (\#\text{leaves of } \mathcal{T}) \cdot 2^{\gamma^{t_0}M}. \tag{2.8}$$

We construct $\mathcal{T}$ greedily by starting from a tree comprising just the root $V(\mathcal{H})$ and repeatedly ‘s‘splitting” every leaf vertex that corresponds to a subset of $V(\mathcal{H})$ with more than $\gamma^{t_0}M$ elements. Suppose that G is such a subset and let $t \geqslant t_0$ be the unique integer such that

$$\gamma^t M < |G| \leqslant \gamma^{t+1}M. \tag{2.9}$$

By our assumption, there is a subhypergraph $\mathcal{H}_G \subseteq \mathcal{H}[G]$ that satisfies condition (2.7). Observe that if we let $p = \frac{b_t}{|G|}$, then we obtain precisely (2.6). Therefore, we can apply Proposition 2.3.1 to $\mathcal{H}_G$ and obtain a family $\mathcal{C}_G$ of subsets of G such that

(i) $|\mathcal{C}_G| \leqslant \left(\frac{e\cdot|G|}{e_H b_t}\right)^{e_H b_t} \leqslant \left(\frac{e\gamma^{t+1}M}{e_H b_t}\right)^{e_H b_t}$,

(ii) $|G'| \leqslant (1-\delta)\cdot|G| \leqslant \gamma^t M$ for every $G' \in \mathcal{C}_G$,

and such that (T2) holds for G, as every independent set in $\mathcal{H}[G]$ is still independent in $\mathcal{H}_G$. Note that (ii) implies that as G ranges over the vertices of any path from the root to a leaf of $\mathcal{T}$, the sequence of t satisfying (2.9) is strictly decreasing. Moreover, $t \leqslant T$, where T is the smallest integer satisfying $\gamma^T M > v(\mathcal{H})$. It follows that

$$\begin{aligned}\#\text{leaves of } \mathcal{T} &\leqslant \prod_{t=t_0}^{T}\left(\frac{e\gamma^{t+1}M}{e_H b_t}\right)^{e_H b_t} \leqslant \prod_{t=t_0}^{T}\left(\frac{e\gamma^{t+1}(t+1)^3}{e_H}\right)^{\frac{e_H M}{(t+1)^3}} \leqslant \prod_{t=t_0}^{T}\left(A^{t+1}\right)^{\frac{e_H M}{(t+1)^3}} \\ &\leqslant \exp\left(e_H M \cdot \log A \cdot \sum_{t=1}^{\infty}\frac{1}{t^2}\right) \leqslant 2^{(C-\gamma^{t_0})M},\end{aligned} \tag{2.10}$$

where A and C are constants depending only on γ, e_H, and t_0. The assertion of the lemma now follows from (2.8) and (2.10). □

2.4 Supersaturation

In this section we establish our supersaturation statement for copies of a fixed hypergraph H. We shall be able to prove, for every n-vertex hypergraph G, the existence of an $\mathcal{H}_G$ as in the discussion before Lemma 2.3.2 using only a relatively mild and natural assumption on the growth rate of $\mathrm{ex}(s, H)$ for all s below some given n. As in the argument of [MS16], we build $\mathcal{H}_G$ by adding suitable copies of H in G one by one. The following technical statement is the main contribution of our work. The key idea in its proof, a double counting argument based on averaging over induced subhypergraphs of G, can be traced back to the seminal work of Erdős and Simonovits [ES83].

Theorem 2.4.1. *Let H be an r-uniform hypergraph, let $\gamma > 1$, and let $\alpha > r - 1/m_r(H)$. Suppose that M and n are such that for every $s \in \{1, \ldots, n\}$,*

$$\mathrm{ex}(s, H) \leqslant M \cdot \left(\frac{s}{n}\right)^{\alpha}.$$

Then there exists a constant t_0 depending only on α, γ, and H such that the following holds. If G is an n-vertex r-uniform hypergraph with

$$\gamma^t M < e(G) \leqslant \gamma^{t+1} M$$

for some integer $t \geqslant t_0$, then there is a collection $\mathcal{H}_G$ of copies of H in G for which, letting $b_t = \frac{M}{(t+1)^3}$,

$$\Delta_\ell(\mathcal{H}_G) \leqslant 2^{2e_H+3} \cdot \left(\frac{b_t}{e(G)}\right)^{\ell-1} \cdot \frac{e(\mathcal{H}_G)}{e(G)} \tag{2.11}$$

for every $\ell \in \{1, \ldots, e_H\}$. In particular, Lemma 2.3.2 implies the existence of a constant C depending only on α and H such that $|\mathcal{F}_n(H)| \leqslant 2^{C \cdot M}$.

Proof. Let $\mathcal{H}$ denote the hypergraph with vertex set $E(G)$ comprising all copies of H in G. We shall construct an $\mathcal{H}_G \subseteq \mathcal{H}$ from an initially empty hypergraph by adding to it copies of H one by one, in a sequence of N steps (N to be chosen shortly). We shall do it in such a way that after N steps, the obtained hypergraph $\mathcal{H}_G$ will have exactly N edges and will satisfy (2.11).

Let $m = e(G)$. Since we will add each copy of H to $\mathcal{H}_G$ only once, we will have $\Delta_{e_H}(\mathcal{H}_G) = 1$ and thus, isolating $e(\mathcal{H}_G)$ in (2.11) with $\ell = e_H$, the number of edges that we have to add to $\mathcal{H}_G$ satisfies

$$N \geqslant \left(\frac{m}{b_t}\right)^{e_H-1} \cdot 2^{-2e_H-3} \cdot m.$$

In particular, choosing

$$N := \left(\frac{\gamma^{t+1} M}{b_t}\right)^{e_H-1} \cdot m = \left((t+1)^3 \cdot \gamma^{t+1}\right)^{e_H-1} \cdot m,$$

we will guarantee that (2.11) holds for $\ell = e_H$.

We now make the above discussion precise. We shall construct a sequence $(\mathcal{H}_i)_{i=0}^N$ of subhypergraphs of $\mathcal{H}$ such that $\mathcal{H}_i \subseteq \mathcal{H}_{i+1}$ and $e(\mathcal{H}_i) = i$ for each i and let $\mathcal{H}_G = \mathcal{H}_N$. We let $\mathcal{H}_0$ be the empty hypergraph. Suppose that we have already defined $\mathcal{H}_i$ for some $i \in \{0, \ldots, N-1\}$. Our goal is not only to find some copy of H in $\mathcal{H} \setminus \mathcal{H}_i$ to be added to $\mathcal{H}_i$ in order to form $\mathcal{H}_{i+1}$, but also to choose this copy carefully so that at the end of the process, condition (2.11) is satisfied for every ℓ. To this end, for every nonempty $F \subsetneq H$, we let $\mathcal{B}_F(\mathcal{H}_i)$ denote the collection of 'bad' copies of F in G in the sense that they are already 'saturated' in $\mathcal{H}_i$. That is, the $\mathcal{H}_i$-degree of the set of e_F edges of G that form this copy of F is close to violating the bound (2.11), with $\ell = e_F$. More precisely, given $F' \subseteq V(\mathcal{H}_i)$, we define

$$\deg_{\mathcal{H}_i} F' = |\{E \in E(\mathcal{H}_i) \colon F' \subseteq E\}|,$$

and let

$$\mathcal{B}_F(\mathcal{H}_i) = \left\{F' \subseteq G \colon F' \simeq F \text{ and } \deg_{\mathcal{H}_i} F' \geqslant 2^{2e_H+2} \cdot \left((t+1)^3 \cdot \gamma^{t+1}\right)^{1-e_F} \cdot \frac{N}{m}\right\}.$$

Observe that

$$2^{e_H} \cdot N \geqslant \binom{e_H}{e_F} \cdot e(\mathcal{H}_i) \geqslant |\mathcal{B}_F(\mathcal{H}_i)| \cdot 2^{2e_H+2} \cdot \left((t+1)^3 \cdot \gamma^{t+1}\right)^{1-e_F} \cdot \frac{N}{m}$$

and therefore,

$$|\mathcal{B}_F(\mathcal{H}_i)| \leqslant 2^{-e_H-2} \left((t+1)^3 \cdot \gamma^{t+1}\right)^{e_F-1} \cdot m. \tag{2.12}$$

Suppose that there exists an $E \in \mathcal{H}$ such that $F' \notin \mathcal{B}_F(\mathcal{H}_i)$ for every nonempty $F \subsetneq H$ and every $F \simeq F' \subsetneq E$. Call each such E *good*, assuming that i is fixed. If there is a good E that is not already in $\mathcal{H}_i$, then letting $\mathcal{H}_{i+1} = \mathcal{H}_i \cup \{E\}$ guarantees that for every $\ell \in [e_H - 1]$,

$$\begin{aligned}
\Delta_\ell(\mathcal{H}_{i+1}) &\leqslant \max\left\{\Delta_\ell(\mathcal{H}_i),\ 2^{2e_H+2} \cdot \left((t+1)^3 \cdot \gamma^{t+1}\right)^{1-\ell} \cdot \frac{N}{m} + 1\right\} \\
&\leqslant \max\left\{\Delta_\ell(\mathcal{H}_i),\ 2^{2e_H+3} \cdot \left((t+1)^3 \cdot \gamma^{t+1}\right)^{1-\ell} \cdot \frac{N}{m}\right\} \\
&\leqslant \max\left\{\Delta_\ell(\mathcal{H}_i),\ 2^{2e_H+3} \cdot \left(\frac{b_t}{m}\right)^{\ell-1} \cdot \frac{N}{m}\right\},
\end{aligned}$$

where the second inequality holds because

$$\left((t+1)^3 \cdot \gamma^{t+1}\right)^{1-\ell} \cdot \frac{N}{m} = \left((t+1)^3 \cdot \gamma^{t+1}\right)^{e_H-\ell} \geqslant 1$$

and the last inequality uses the definition of b_t and the bound $m \leqslant \gamma^{t+1} M$. In particular, by the definition of N, if we succeed in finding such a good $E \in \mathcal{H} \setminus \mathcal{H}_i$ for every i, then the final hypergraph $\mathcal{H}_G = \mathcal{H}_N$ will satisfy (2.11) for every $\ell \in [e_H]$.

Fix some $p \in (0,1]$ such that pn is an integer and let R be a uniformly chosen random subset of pn vertices of G.[4] Denote by G' the subgraph of G induced by R. Let G'' be a graph obtained from G' by removing one edge from each copy of F in G' that belongs to $\mathcal{B}_F(\mathcal{H}_i)$, for every nonempty $F \subsetneq H$. Note that any copy of H in G'' is good by definition. Let X denote the (random) number of good copies of H in G'' and let Z be the total number of good copies of H in G. Even though we might have accidentally eliminated some good copies of H in G' while forming the subgraph G'', it is still true that

$$\mathbb{E}[X] \leqslant Z \cdot \binom{n-v_H}{pn-v_H} \Big/ \binom{n}{pn} = Z \cdot \binom{pn}{v_H} \Big/ \binom{n}{v_H} \leqslant Z \cdot p^{v_H}.$$

Since every copy of H in G'' is good and G'' has pn vertices, then

$$X \geqslant e(G'') - \mathrm{ex}(pn, H) \geqslant e(G'') - M \cdot p^{\alpha}.$$

Since clearly

$$e(G'') \geqslant e(G') - \sum_{F \subsetneq H} \sum_{F' \in \mathcal{B}_F(\mathcal{H}_i)} \mathbb{1}[F' \subseteq G'],$$

and for every $F' \subseteq G$ with $F' \simeq F$, we have $\Pr(F' \subseteq G') = \binom{n-v_F}{pn-v_F}/\binom{n}{pn} \leqslant p^{v_F}$, it follows that

$$Z \cdot p^{v_H} \geqslant \mathbb{E}[X] \geqslant \mathbb{E}[e(G'')] - M \cdot p^{\alpha} \geqslant \mathbb{E}[e(G')] - \sum_{F \subsetneq H} |\mathcal{B}_F(\mathcal{H}_i)| \cdot p^{v_F} - M \cdot p^{\alpha}. \quad (2.13)$$

Finally, if $pn \geqslant 2r^2$, then

$$\begin{aligned}\mathbb{E}[e(G')] &= m \cdot \binom{n-r}{pn-r} \Big/ \binom{n}{pn} = m \cdot \binom{pn}{r} \Big/ \binom{n}{r} \geqslant m \cdot \left(\frac{pn-r}{n}\right)^r \\ &= m \cdot p^r \cdot \left(1 - \frac{r}{pn}\right)^r \geqslant m \cdot p^r \cdot \left(1 - \frac{r^2}{pn}\right) \geqslant \frac{m \cdot p^r}{2},\end{aligned}$$

which substituted into (2.13) yields

$$Z \cdot p^{v_H} \geqslant \frac{m \cdot p^r}{2} - \sum_{F \subsetneq H} |\mathcal{B}_F(\mathcal{H}_i)| \cdot p^{v_F} - M \cdot p^{\alpha}. \quad (2.14)$$

We claim that there is a $p \in [2r^2/n, 1]$ such that pn is an integer and the right-hand side of (2.14) is at least $N \cdot p^{v_H}$, and thus $Z \geqslant N$. Since $e(\mathcal{H}_i) = i < N$, this inequality would imply that there is a good copy of H in G that does not belong to $\mathcal{H}_i$, completing the proof. Hence, it suffices to establish this claim. To this end, note

[4] The reason why we let R be a uniformly chosen random set of pn vertices and not a binomial random subset is that in the latter case, we did not see a clean way to bound $\mathbb{E}[\mathrm{ex}(|R|, H)]$ from above that would avoid estimating the r^{th} moment of $|R|$.

first that by (2.12), we have

$$\sum_{F\subsetneq H} |\mathcal{B}_F(\mathcal{H}_i)| \cdot p^{v_F} \leqslant \frac{m}{4} \cdot \max\left\{ \left((t+1)^3 \cdot \gamma^{t+1}\right)^{e_F-1} \cdot p^{v_F} : F \subsetneq H \right\}. \tag{2.15}$$

Thus it suffices to have the following three inequalities for every $F \subsetneq H$:

$$p^{r-v_F} \geqslant \left((t+1)^3 \cdot \gamma^{t+1}\right)^{e_F-1}, \tag{2.16}$$

$$p^{\alpha-r} \leqslant \frac{\gamma^t}{8} \leqslant \frac{m}{8M}, \tag{2.17}$$

$$p^{r-v_H} \geqslant 8\left((t+1)^3 \cdot \gamma^{t+1}\right)^{e_H-1} = \frac{8N}{m}. \tag{2.18}$$

Indeed, combining inequalities (2.14), (2.15), (2.16), and (2.17) yields

$$Z \cdot p^{v_H} \geqslant \frac{m \cdot p^r}{8},$$

which combined with (2.18) gives the desired lower bound on Z. Note also that, as $\gamma > 1$, $e_H > 1$, and $0 \le e_F - 1 \le e_H - 1$, both (2.16) and (2.18) would follow if the following was true for every $F \subseteq H$:

$$p^{r-v_F} \geqslant \left(8 \cdot (t+1)^3 \cdot \gamma^{t+1}\right)^{e_F-1}. \tag{2.19}$$

Observe that (2.17) holds trivially for all large enough t if $\alpha = r$. Moreover, (2.19) holds when $v_F = r$, as then $e_F = 1$. Hence, we may assume that $\alpha < r$ and verify (2.19) only for all $F \subseteq H$ with $e_F > 1$.

We claim that it suffices to show that for all $F \subseteq H$ with $e_F > 1$,

$$\left(8 \cdot (t+1)^3 \cdot \gamma^{t+1}\right)^{\frac{e_F-1}{v_F-r}} \leqslant \min\left\{ \frac{1}{2} \cdot \left(\frac{\gamma^t}{8}\right)^{\frac{1}{r-\alpha}}, \frac{n}{4r^2} \right\} \tag{2.20}$$

Indeed, assuming that (2.20) holds, we shall be able to show that every p in some interval $[p_0/2, p_0] \subseteq [2r^2/n, 1]$ satisfies both (2.17) and (2.19). Since every such interval must contain a p such that pn is an integer, we will be done. We let p_0 be the reciprocal of the left-hand side of (2.20), that is, $1/p_0 := (8 \cdot (t+1)^3 \cdot \gamma^{t+1})^{\frac{e_F-1}{v_F-r}}$, and note that $p_0 \leqslant 1$, as $\gamma > 1$ and $\frac{e_F-1}{v_F-r} > 0$, and that (2.20) implies that $p_0/2 \geqslant 2r^2/n$; in particular, $[p_0/2, p_0] \subseteq [2r^2/n, 1]$. Finally, inequalities (2.17) and (2.19) are equivalent to

$$\frac{1}{p} \leqslant \left(\frac{\gamma^t}{8}\right)^{\frac{1}{r-\alpha}} \qquad \text{and} \qquad \frac{1}{p} \geqslant \frac{1}{p_0},$$

respectively and thus (2.20) implies that every $p \in [p_0/2, p_0]$ satisfies both of them.

Finally, we show that (2.20) holds. The first of the two inequalities in (2.20) holds for all large t, as $\gamma > 1$ and by our hypothesis

$$\frac{e_F - 1}{v_F - r} \leqslant m_r(H) < \frac{1}{r - \alpha}. \tag{2.21}$$

To see that the second inequality in (2.20) holds as well, note first that

$$1 = \mathrm{ex}(r, H) \leqslant M \cdot \left(\frac{r}{n}\right)^{\alpha} \leqslant \frac{e(G)}{\gamma^t} \cdot \left(\frac{r}{n}\right)^{\alpha} \leqslant \frac{n^r}{\gamma^t} \cdot \left(\frac{r}{n}\right)^{\alpha},$$

and hence $t \leqslant \frac{r \log n}{\log \gamma}$. It follows that

$$\left(8 \cdot (t+1)^3 \cdot \gamma^{t+1}\right)^{\frac{e_F - 1}{v_F - r}} \leqslant \left(\frac{16 \cdot r^{\alpha+3} \cdot \gamma}{(\log \gamma)^3} \cdot n^{r-\alpha} \cdot (\log n)^3\right)^{\frac{e_F - 1}{v_F - r}} \leqslant \frac{n}{4r^2},$$

provided that t is sufficiently large (and thus n is sufficiently large), since $(r - \alpha) \cdot \frac{e_F - 1}{v_F - r} < 1$ by our hypothesis, see (2.21). This completes the proof. □

2.5 Proofs of Theorems 2.1.2 and 2.1.7

In this section we prove Theorems 2.1.2 and 2.1.7. Both will be obtained as (more or less) immediate corollaries of our technical Theorem 2.4.1.

Proof of Theorem 2.1.7. Let $\alpha > r - 1/m_r(H)$ and let A be such that

$$\mathrm{ex}(n, H) \leqslant An^{\alpha}$$

for all n. Define $M = An^{\alpha}$ and observe that for all $s \in [n]$,

$$\mathrm{ex}(s, H) \leqslant As^{\alpha} = \left(\frac{s}{n}\right)^{\alpha} \cdot An^{\alpha}.$$

Therefore, Theorem 2.4.1 implies the existence of some $C > 0$ such that

$$|\mathcal{F}_n(H)| \leqslant 2^{Cn^{\alpha}},$$

as claimed. □

Using a standard probabilistic argument, one can show that for every r-uniform hypergraph H with at least two edges, the bound $\mathrm{ex}(n, H) \geqslant c_H n^{r - 1/m_r(H)}$ holds for some positive constant c_H. In particular, if $\mathrm{ex}(n, H) \leqslant An^{\alpha}$ for all n, as in the statement of Theorem 2.1.7, then $\alpha \geqslant r - 1/m_r(H)$. It turns out that when H is a graph that contains a cycle, the stronger lower bound

$$\mathrm{ex}(n, H) \geqslant c_H n^{2 - 1/m_2(H)} (\log n)^{1/(e_H - 1)} \tag{2.22}$$

holds for all n. This was first proved by Bohman and Keevash [BK10] and later generalized to hypergraphs of higher uniformity by Bennett and Bohman [BB16].

Proof of Theorem 2.1.2. Suppose that H contains a cycle and α and A are such that $\mathrm{ex}(n, H) \leqslant An^{\alpha}$ for all n. It follows from (2.22) that $\alpha > 2 - 1/m_2(H)$. The assertion of the theorem now easily follows from Theorem 2.1.7. $\square$

2.6 Proofs of Theorems 2.1.5 and 2.1.9

In this section, we prove Theorems 2.1.5 and 2.1.9. That is, we show that if $\mathrm{ex}(n, H)$ exceeds the standard probabilistic lower bound of $c_H n^{r-1/m_r(H)}$ by a factor polynomial in n, then Theorem 2.4.1 implies that $|\mathcal{F}_n(H)| \leqslant 2^{C \cdot \mathrm{ex}(n,H)}$ for infinitely many n. Since Theorem 2.1.5 is simply the case $r = 2$ of Theorem 2.1.9, we only prove the latter.

Proof of Theorem 2.1.9. Let H be an r-uniform hypergraph and suppose that there is an $\varepsilon > 0$ such that

$$\mathrm{ex}(n, H) \geqslant \varepsilon n^{r-1/m_r(H)+\varepsilon}$$

for all n. We shall construct an infinite sequence of n satisfying the hypothesis of Theorem 2.4.1 with $M = \mathrm{ex}(n, H)$ and $\alpha = r - 1/m_2(H) + \varepsilon/2$. Then, for each n in the sequence, we obtain

$$|\mathcal{F}_n(H)| \leqslant 2^{C \cdot \mathrm{ex}(n,H)}$$

for some C that depends only on α, ε, and H. This will complete the proof.

Assume towards a contradiction that there are only finitely many n satisfying the hypothesis of Theorem 2.4.1. In particular, there exists an N such that for all $n_0 \geqslant N$,

$$\mathrm{ex}(n_1, H) > \mathrm{ex}(n_0, H) \cdot \left(\frac{n_1}{n_0}\right)^{\alpha}$$

for some $n_1 < n_0$. Choose a small $\delta > 0$, let $n_0 = \lceil N^{1/\delta} \rceil$, and suppose that we have defined $n_0, \ldots, n_{k-1}$ this way. If $n_{k-1} \geqslant n_0^{\delta} \geqslant N$, then there is some $n_k \leqslant n_{k-1} - 1 \leqslant n_0 - k$ such that

$$\mathrm{ex}(n_k, H) > \mathrm{ex}(n_{k-1}, H) \cdot \left(\frac{n_k}{n_{k-1}}\right)^{\alpha} > \mathrm{ex}(n_0, H) \cdot \left(\frac{n_k}{n_{k-1}}\right)^{\alpha} \cdot \left(\frac{n_{k-1}}{n_0}\right)^{\alpha}.$$

Note that if $n_k \leqslant n_0^{\delta}$, then the lower bound $\mathrm{ex}(n, H) \geqslant \varepsilon n^{r-1/m_r(H)+\varepsilon}$ implies

$$n_0^{r\delta} \geqslant n_k^r > \binom{n_k}{r} \geqslant \mathrm{ex}(n_k, H) > \mathrm{ex}(n_0, H) \cdot n_0^{(\delta-1)\alpha} \geqslant \varepsilon n_0^{\alpha+\varepsilon/2} \cdot n_0^{(\delta-1)\alpha} \geqslant \varepsilon n_0^{\alpha\delta+\varepsilon/2}.$$

This is clearly impossible, as $\varepsilon > 0$ is fixed and we may choose N as large as we want and δ as small as we want. Therefore, there must be some $n \geqslant N$ for which the hypothesis of Theorem 2.4.1 holds, a contradiction. $\square$

Chapter 3

Cliques in W-random graphs

3.1 Introduction

This chapter is based on the paper [McK19].

The Erdős-Rényi random graph $G_{n,p}$ is a graph on n vertices where an edge is placed independently with probability p between each pair of vertices. One of the most basic parameters of any graph G is the *clique number* $\omega(G)$, the number of vertices in the largest complete subgraph of G. It was shown independently by Grimmett and McDiarmid in 1975 [GM75] and Matula in 1976 [Mat76] that for a fixed $p \in (0,1)$, the clique number $\omega(G_{n,p})$ of $G_{n,p}$ satisfies

$$\omega(G_{n,p}) = (1 + o(1)) \cdot \frac{2 \log n}{\log(1/p)} \tag{3.1}$$

with probability $1 - o(1)$ as n approaches infinity. This can be proved roughly as follows: we obtain an upper bound on $\omega(G_{n,p})$ by finding k such that the expected number of k-cliques in $G_{n,p}$ is asymptotically zero (the first moment method). Then, to prove a matching lower bound, we show that for an appropriate, slightly smaller k, the number of k-cliques in $G_{n,p}$ approaches infinity in the limit and has low variance. This implies that the number of cliques of size k is highly concentrated around its expectation, and will be positive with high probability (the second moment method). Some variation on this method has been a standard technique for computing clique number in other random graph models as well. (See [GM75], [Mat76], [DGLU11], [DHM19], and [BCvdH20].)

Here, we turn our view from the "homogeneous" Erdős-Rényi random graph to an "inhomogeneous" setting, in which edges may be assigned between some pairs of vertices with higher or lower probabilities. This is both a better model of many real-world phenomena and an object of independent mathematical interest. However, with this greater flexibility comes greater difficulty in analysis. In what follows, we will characterize the clique numbers of a variety of inhomogeneous random graphs that arise from the theory of graphons.

A *graphon* W is defined as a symmetric, measurable function from Ω^2 to $[0, 1]$, where Ω is a probability space. To obtain a random graph from W, we sample n

points $x_1, \ldots, x_n$ independently according to the probability distribution on Ω, and connect vertices i and j by an edge with probability $W(x_i, x_j)$, independently for each pair (i, j). (For the sake of brevity, we will often identify the vertex i with the value x_i, and speak of "sampling vertices" from Ω). We denote this graph by $\mathbb{G}(n, W)$, and refer to it as a "W-random graph". Notice that in the case where W is equal to the constant function p, we simply have $\mathbb{G}(n, W) = G_{n,p}$. One of the main results in the theory of graphons, proved by Lovász and Szegedy in 2006 in [LS06], is that every infinite sequence of graphs contains a subsequence converging to some graphon W (in what is called the cut norm), and moreover, that every graphon can be achieved in this way, as the limit of some sequence of graphs. It is therefore reasonable to think of graphons as the correct limiting objects for sequences of graphs that are Cauchy sequences in an appropriate metric. See [Lov12] for a detailed survey of the theory of graphons.

It should be noted that we must take some care in defining a notion of clique number for graphons. We might hope that all sequences of graphs converging to a given graphon would have the same clique number asymptotically; however, as noted in [DHM19], this is not the case. Consider as an example the following two sequences of graphs.

Example 3.1.1.

- G_n consists of a clique on $\sqrt{n}$ vertices, and $n - \sqrt{n}$ isolated vertices.
- H_n consists of n isolated vertices.

Both sequences approach the zero graphon, as the density of edges approaches zero in both cases. However $\omega(G_n) = \sqrt{n}$, while $\omega(H_n) = 1$. Thus, instead of looking at all sequences of graphs converging to a given graphon W, we will consider only "typical" sequences, sampled according to the distribution $\mathbb{G}(n, W)$. (Note: an alternate notion of clique number for a graphon is presented in [HR17].)

This was the question considered by Doležal, Hladký, and Máthé in [DHM19], where they obtained a partial characterization of the clique number of $\mathbb{G}(n, W)$ for graphons W. They proved the following result. (Note: in the statement below, by "essentially bounded", we mean that the given bound holds everywhere except perhaps on some set of measure zero.)

Theorem 3.1.2 (Doležal, Hladký, and Máthé [DHM19, Cor. 2.8])**.** *For a graphon $W \colon \Omega^2 \to [0, 1]$ that is essentially bounded away from 0 and 1,*

$$\omega(\mathbb{G}(n, W)) = (1 + o(1))\kappa(W) \log n$$

a.a.s., where

$$\kappa(W) = \sup \left\{ \frac{2\|h\|_1^2}{\int_{(x,y)\in\Omega^2} h(x)h(y)\log(1/W(x,y))\, d(\nu^2)} \;:\; h \text{ is a nonnegative } L^1\text{-function on } \Omega \right\}.$$

Notice that, for a graphon W essentially bounded between $p_1 > 0$ and $p_2 < 1$, we can couple $\mathbb{G}(n, W)$ with the Erdős-Rényi random graphs G_{n,p_1} and G_{n,p_2} so that

$\omega(G_{n,p_1}) \le \omega(\mathbb{G}(n,W)) \le \omega(G_{n,p_2})$. Since the clique number of $G_{n,p}$ is $\Theta(\log n)$ for any value of p, this immediately tells us that the clique number of $\mathbb{G}(n,W)$ is also $\Theta(\log n)$ with probability approaching 1. Thus the key part of the result above is the characterization of the constant $\kappa(W)$ in $\Theta(\log n)$.

A similar question was considered by Bogerd, Castro, and van der Hofstad in [BCvdH20]; they studied clique number for rank-1 inhomogeneous random graphs, where a graph is formed by assigning a weight to each vertex according to some distribution, and then connecting each pair of vertices independently with a probability proportional to the product of their weights. They showed that, if all vertex weights are bounded away from 1 (analogous to the assumption in Theorem 3.1.2 that W is essentially bounded away from 1), then the clique number of such a graph is concentrated on at most two consecutive integers, for which they gave explicit expressions. This was proved in both the dense case and the sparse case, in which the edge density approaches zero as the number of vertices grows. It should be noted that a great deal of the work on inhomogeneous random graph models has centered on the sparse case, which gives a more accurate model for a variety of real-world networks, and it would be interesting to see more results in this direction. (See [BJR07] for one of the seminal sparse models, and [vdH16] and [vdH17] for a survey of other recent work.) Results have also been obtained for clique number in random graphs with a power-law distribution [JŁN10] and hyperbolic random graphs [BFK18].

Here, however, we will explore in a different (and in some sense, even opposite) direction. Namely, for graphons W that are not bounded away from 1, even the rough order of growth of $\omega(\mathbb{G}(n,W))$ is not apparent (we could think of this as producing a W-random graph with potentially very dense spots); for this reason, it is interesting to ask what may happen if W is allowed to approach 1. (Note, however, that if $W = 1$ on $S \times S$ for some set S of positive measure, then W will have linear clique number a.a.s., as the subset of vertices sampled from S will all be connected with probability 1.) Additionally, although the restriction to graphons essentially bounded away from 1 given in [DHM19] is a natural condition that precludes a variety of pathological examples, there is no reason to suppose that any particular graphon that might arise in an applied setting would necessarily satisfy it. It is still necessary, however, to impose some restrictions on the behavior of W in order to obtain a good characterization of $\omega(\mathbb{G}(n,W))$; the authors of [DHM19] also showed that for an arbitrary graphon W not bounded away from 1, $\omega(\mathbb{G}(n,W))$ may behave quite wildly as $n \to \infty$.

Example 3.1.3 (Doležal, Hladký, and Máthé [DHM19, Prop. 2.1])**.** There exists a graphon W and a sequence of integers $n_1 < n_2 < \cdots$ such that, a.a.s., $\omega(\mathbb{G}(n_i,W))$ alternates between at most $\log\log n_i$ and at least $\frac{n_i}{\log\log n_i}$ on elements of the sequence.

In fact, we may take any $\omega(1)$ function in place of $\log\log n$ in the example above. This behavior is shown in [DHM19] to be achieved by a highly discontinuous graphon $W\colon [0,1]^2 \to [0,1]$, which raises the question: even if W is not bounded away from 1, can we obtain a good characterization of $\omega(\mathbb{G}(n,W))$ as long as W is reasonably well-behaved? This is the central question of this chapter.

In order to characterize the "well-behavedness" of a graphon in a natural way, we must impose additional requirements on the underlying probability space Ω. For example, for continuity to make sense, we must assume Ω is a topological space, and smoothness requires even stronger hypotheses (for example, that Ω is a manifold). It is unclear what the most natural or general setting is, and for practical applications, the particular form of Ω should reflect the geometry of the underlying feature space for a network model. We will not take up such questions here. Rather, all of the examples and results here will deal with the uniform distribution on $\Omega = [0, 1]$, which is the simplest probability space capable of capturing a wide range of graphon behaviors and the most widely studied setting for graphons; indeed, this is the setting in which graphons were originally defined, in [LS06]. Some generalizations to, say, a compact topological space or Riemann manifold are fairly straightforward, but we leave open the question of how to formulate broadly applicable generalizations.

We also note that, for a graphon $W\colon [0,1]^2 \to [0,1]$, among points with $W(x, y) = 1$, we are primarily concerned with those points along the line $x = y$, as shown by the following lemma.

Lemma 3.1.4. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon whose essential supremum is strictly less than 1 in some neighborhood of each point (x, x) for $x \in [0, 1]$. Then $\omega(\mathbb{G}(n, W)) = O(\log n)$ a.a.s.*

This lemma is proved in Section 3.2. With this in mind, our goal is really to find the clique number associated to a "well-behaved" graphon that is equal to 1 at one or more points (x, x) with $x \in [0, 1]$. The main contribution of this chapter consists of several such results. Before presenting these results, however, one final observation: it is perhaps natural to ask whether graphons that are close in cut distance will produce W-random graphs whose clique numbers are close asymptotically. In general, however, this is not the case. This can be illustrated by a wide variety of examples, but perhaps the simplest is the following family of graphons on $[0, 1]^2$:

$$W_\varepsilon(x, y) = \begin{cases} 1 & \text{if } (x, y) \in [0, \varepsilon]^2 \\ 0 & \text{otherwise} \end{cases}$$

for each $\varepsilon > 0$. Under the cut norm, W_ε converges to the zero graphon as $\varepsilon \to 0$, but for any fixed ε, the clique number of $\mathbb{G}(n, W_\varepsilon)$ is $(1 + o(1))\varepsilon n = \Theta(n)$ a.a.s. (see Lemma 3.2.1) Indeed, the primary driver of clique number for a W-random graph is not global behavior (as measured by the cut norm), but local behavior near points (x, x) where W is maximized. Following are several results characterizing clique number in terms of this local behavior. First, and perhaps surprisingly, for a graphon equal to 1 at only a finite number of points (x, x), we will very often obtain a clique number of $\Theta(\sqrt{n})$.

Theorem 3.1.5. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some collection of points $(a_1, a_1), \ldots, (a_k, a_k)$, and essentially bounded away from 1 in some neighborhood of (x, x) for each other $x \in [0, 1]$. If all directional derivatives of W exist at the points $(a_1, a_1), \ldots, (a_k, a_k)$, and are uniformly bounded away from 0 and $-\infty$, then $\omega(\mathbb{G}(n, W)) = \Theta(\sqrt{n})$ a.a.s.*

We can expand this result to graphons W whose directional derivatives are not defined at the points where $W(a,a) = 1$. In Section 3.3, we give a more complete characterization in terms of the Dini derivatives of W (the limsup and liminf of the difference quotient that defines the ordinary derivative) at the points where $W(a,a) = 1$. In particular, this characterization will show that if W is "too steep" at the points where it is equal to 1, then the clique number of $\mathbb{G}(n,W)$ will be $o(\sqrt{n})$ (Lemma 3.3.6 (ii)), and if W is "too flat" at these points (derivatives equal to zero), then the clique number will be $\omega(\sqrt{n})$ (Lemma 3.3.5 (ii)). This expanded characterization will also yield the following.

Lemma 3.1.6. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some point (a,a). If W is locally Lipschitz continuous at (a,a), then $\omega(\mathbb{G}(n,W)) = \Omega(\sqrt{n})$ a.a.s.*

(We will recall the definition of local Lipschitz continuity at a point (a,a) in the proof of Lemma 3.1.6.) In addition, in Section 3.4, we present a family of graphons yielding clique numbers $\Theta(n^\alpha)$ for any constant $\alpha > 0$.

Theorem 3.1.7. *For any constant $r > 0$, define the graphon*

$$U_r(x,y) := (1-x^r)(1-y^r).$$

The random graph $\mathbb{G}(n,U_r)$ a.a.s. has clique number $\Theta(n^{\frac{r}{r+1}})$.

It will be shown in Section 3.4 that this implies the following more general result.

Theorem 3.1.8. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some point (a,a). If W is locally α-Hölder continuous at (a,a) for some constant α, then $\omega(\mathbb{G}(n,W)) = \Omega(n^{\frac{\alpha}{\alpha+1}})$ a.a.s.*

(We will recall the definition of local α-Hölder continuity at a point (a,a) immediately before the proof of Theorem 3.1.8.) We will also be able to use the characterization of $\omega(\mathbb{G}(n,U_r))$ given in Section 3.4 to show that if a graphon W has infinitely many derivatives equal to zero at a point (a,a) where $W(a,a) = 1$, then the clique number of $\mathbb{G}(n,W)$ will be $n^{1-o(1)}$ a.a.s. In other words, if W is "extremely flat" at the points where it is equal to 1, then the clique number of $\mathbb{G}(n,W)$ will be nearly linear. We will prove this for the following specific example, but the same reasoning can apply more generally.

Proposition 3.1.9. *For the graphon $W : [0,1]^2 \to [0,1]$ defined by*

$$W(x,y) = (1-f(x))(1-f(y)), \text{ where } f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases},$$

the clique number of $\mathbb{G}(n,W)$ is both $n^{1-o(1)}$ and $o(n)$ a.a.s.

It should be noted that, in contrast to Theorem 3.1.2 and the characterization of clique number for Erdős-Rényi random graphs, all the results above give the relevant clique number up to a constant. For any graphon W, however, the clique number of $\mathbb{G}(n, W)$ is highly concentrated for large values of n; the following was proved in [DHM19]. (This is slightly different from the original formulation, but follows directly the proof of Theorem 2.2 in [DHM19].)

Theorem 3.1.10 (Doležal, Hladký, and Máthé [DHM19, Thm. 2.2])**.** *For any graphon W, with probability $1 - o(1)$,*

$$\omega(\mathbb{G}(n, W)) = (1 + o(1)) \cdot \mathbb{E}[\omega(\mathbb{G}(n, W))].$$

From this, we know that a correct constant always exists. But although the clique number $\omega(\mathbb{G}(n, W))$ is almost always very close to its expectation, it may occasionally be very large. Indeed, for many of the examples we will consider, the number of cliques of an appropriate size will have quite high variance, making it impossible to directly apply the second moment method as with Erdős-Rényi random graphs. (This obstacle is detailed more explicitly in Sections 3.3 and 3.4, with proofs given in Appendix A.) Instead, in proving the results above, we use the first moment method to establish upper bounds (the standard technique), while for lower bounds, we directly attempt to predict which vertices are likely to form a large clique, and show that this does indeed happen with high probability. It seems likely that to improve these lower bounds, a different technique would be necessary.

It should be noted that the authors of [DHM19] did not use the second moment method directly to prove Theorem 3.1.2, but instead applied it to a carefully selected restriction of the graphon W, converting this back into a lower bound on $\omega(\mathbb{G}(n, W))$ via a somewhat complex argument. It is possible that a similar techinque could be used to improve some or all of the lower bounds given here. It is also possible that tighter bounds could obtained using techniques from the theory of large deviations as in [AP03]; in this case, instead of looking at the (random) number of cliques X of a given size in $\mathbb{G}(n, W)$ and attempting to give upper and lower bounds on X that hold with high probability, we would define a random variable X' that gives greater weight to those cliques arising from a "typical" configuration of vertices (e.g., not too many vertices sampled from a small interval), and that would thus have lower variance than X. If we could find upper and lower bounds on X', these could then be translated into upper and lower bounds on X.

Another potentially interesting extension of the results above could be to consider graphons with an infinite number (either countable or uncountable) of points (x, x) with $W(x, x) = 1$. For example, the following graphon is equal to 1 along the line $x = y$ and drops away from 1 off that line.

Proposition 3.1.11. *Let $W(x, y) = 1 - |x - y|$. The clique number of $\mathbb{G}(n, W)$ is $n^{1/2+o(1)}$ a.a.s.*

This will be proved in Section 3.5; the lower bound follows directly from applying Lemma 3.1.6 to any point on the line $x = y$, and the upper bound is a fairly straightforward calculation. Both arguments could be used on a wide variety of such examples, but both are likely not tight in general.

As a last note, we discuss briefly the related problems of finding a large clique or a planted clique in a random graph, and how they relate to the work here. It is a long-standing problem, proposed by Karp in 1976 [Kar76] to find a clique of size $(1+\varepsilon)\log_2(n)$ in $G_{n,1/2}$ in polynomial time. (A clique of size $\sim 2\log_2(n)$ will almost always exist.) There are several polynomial-time algorithms that find a clique of size $(1+o(1))\log_2(n)$ (e.g., [KS98]), but the original problem remains open. Of a similar flavor, but slightly different, the planted clique problem asks us to find a clique of size k that is "planted" in an Erdős-Rényi random graph $G_{n,p}$ by randomly selecting k vertices and adding all possible edges between them; we may ask for an algorithm that runs either in polynomial or unbounded time. In unbounded time, the planted clique can be recovered for k quite close to the expected clique number for $G_{n,p}$, but perhaps surprisingly, if we ask for a polynomial-time algorithm, the best known methods find the planted clique only for $k = c \cdot \sqrt{n}$, for some particular constant c (first proved in [AKS98], with a variety of simpler algorithms or algorithms improving the constant found later; see, for example, [FR10] and [DM15]). It could be interesting to explore these problems in the setting where the background graph is inhomogeneous (as opposed to $G_{n,p}$); it seems entirely possible that a large clique or hidden clique could be easier to recover in this setting. Indeed, this has been shown to be the case for several specific (mostly sparse) random graph models (see [FK12], [BFK18], and [JŁN10]). It is possible that a more general result along these lines could be established for some of graphs discussed here, or those in [DHM19] or [BCvdH20], especially given knowledge of the clique number in the background graph.

The remainder of this chapter is structured as follows. In Section 3.2, we prove Lemma 3.1.4 and a few other simple technical lemmas that will be used throughout the chapter. In Section 3.3, we present a family of graphons giving clique numbers $\Theta(\sqrt{n})$, and use this to prove Theorem 3.1.5, Lemma 3.1.6, and an extension to graphons satisfying a more general set of conditions. In Section 3.4, we prove Theorem 3.1.7, Theorem 3.1.8, and Proposition 3.1.9. In Section 3.5, we prove Proposition 3.1.11, and discuss possible extensions of this work. And finally, in Appendix A, we prove that for many of the W-random graphs discussed in this chapter, the number of cliques of an appropriate size has high variance, making a direct application of the second moment method to establish a lower bound on the clique number impossible in those cases.

3.2 Preliminaries

In this section, we establish some notation and technical lemmas that will be used throughout the rest of the chapter. We will often want to focus only on a small portion of a graphon W, typically a neighborhood around a point where W is equal to 1. In order to analyze how local behavior affects the clique number of a W-random graph, we first ascertain how many vertices will typically be sampled from a given neighborhood. Note: below, we write λ for the Lebesgue measure on $\mathbb{R}$.

Lemma 3.2.1. *Let $A(1), A(2), \ldots$ be measurable subsets of $[0,1]$ with $\lambda(A(n)) = \omega\left(\frac{1}{n}\right)$. Among n points uniformly distributed on the interval $[0,1]$, the number of points in $A(n)$ will a.a.s. be $(1+o(1))n\lambda(A(n))$.*

Proof. The number of points X in any given subset of $[0,1]$ of measure $\lambda(A(n))$ is a binomial random variable with parameters n and $p = \lambda(A(n))$. Therefore

$$\mathbb{E}[X] = np$$

and

$$\begin{aligned}\mathrm{Var}(X) &= np(1-p)\\ &\leq np.\end{aligned}$$

Thus, for any $\varepsilon > 0$, by Chebyshev's inequality

$$\Pr\left[|X - \mathbb{E}[X]| \geq \varepsilon\mathbb{E}[X]\right] \;\leq\; \frac{\mathrm{Var}(X)}{\varepsilon^2\mathbb{E}[X]^2} \;\leq\; \frac{np}{\varepsilon^2(np)^2} \;=\; \frac{1}{\varepsilon^2 np}.$$

By assumption, $np = n \cdot \lambda(A(n)) = \omega(1)$. So taking, for instance, $\varepsilon^2 = (np)^{-1/2}$, we have

$$\Pr\left[|X - \mathbb{E}[X]| \geq \tfrac{1}{(np)^{1/4}} \cdot \mathbb{E}[X]\right] \leq \tfrac{1}{(np)^{1/2}} = o(1).$$

Thus with probability $1-o(1)$, we have $X = (1+o(1))\mathbb{E}[X] = (1+o(1))n\lambda(A(n))$. □

In Section 3.5, we will need a slight strengthening of the result above; namely, if we sample n points uniformly from $[0,1]$, the lemma below guarantees that no relatively large subset of these points will occupy an interval much smaller than expected.

Lemma 3.2.2. *Let $\delta = \omega\left(\frac{1}{\sqrt{n}}\right)$. Among n points uniformly distributed on the interval $[0,1]$, with probability $1-o(1)$, every set of δn points will occupy an interval of length at least $\frac{\delta}{2}(1-o(1))$.*

Proof. We begin by dividing $[0,1]$ into consecutive intervals of length δ. By Lemma 3.2.1, with probability $1-o(1)$, there will be at most $(\delta + o(1))n$ vertices in any fixed one of these intervals, as $\delta = \omega\left(\frac{1}{n}\right)$. For $\delta = \omega\left(\frac{1}{\sqrt{n}}\right)$, there will in fact be at most $(\delta + o(1))n$ vertices in *each*; as shown in Lemma 3.2.1, if X is the number of vertices in a given interval of length δ, then for any $\varepsilon > 0$, we have

$$\Pr\left[X \geq (1+\varepsilon)\delta n\right] \leq \frac{1}{\varepsilon^2(\delta n)}.$$

Then, taking a union bound over the $\frac{1}{\delta}$ consecutive length-δ intervals, with probability $1 - \frac{1}{\delta} \cdot \frac{1}{\varepsilon^2(\delta n)} = 1 - \frac{1}{\varepsilon^2\delta^2 n}$, each of these intervals contains at most $(1+\varepsilon)\delta n$ vertices. So for any $\delta = \omega\left(\frac{1}{\sqrt{n}}\right)$, we can choose an appropriate $\varepsilon = o(1)$ to conclude that with probability $1-o(1)$, each of the $\frac{1}{\delta}$ consecutive intervals contain at most $(1+o(1))\delta n$ vertices.

Notice that any other interval of length δ in $[0,1]$ is contained entirely in at most two of these consecutive intervals. So with probability $1-o(1)$, any interval of length δ in $[0,1]$ will contain at most $(1+o(1))2\delta n$ vertices. Equivalently, and after a slight change of variables, with probability $1-o(1)$, every δn vertices will occupy an interval of length at least $\frac{\delta}{2}(1-o(1))$. $\square$

Now, we would like to be able to say something about the cliques in W-random graphs that we obtain by sampling points from smaller sets, for example, from neighborhoods around points at which $W=1$. As in [DHM19], we define a *subgraphon* of any graphon W to be the restriction obtained by "zooming in" on a subset of the sample space:

Definition 3.2.3. *Given a graphon $W\colon [0,1]^2 \to [0,1]$ and a subset $A \subseteq [0,1]$ of positive measure, define the subgraphon $W|_{A\times A}\colon A^2 \to [0,1]$ as the restriction of W to $A\times A$, where we sample uniformly from the set A to obtain a probability distribution on A.*

(Note that $W|_{A\times A}$ as defined above satisfies the definition of a graphon on a more general probability space Ω.) Intuitively, if we break a graphon W into subgraphons, its clique number will be at least the maximum clique number among the subgraphons and at most the sum of all their clique numbers. This intuition is formalized in the following lemma.

Lemma 3.2.4. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon, let $k \in \mathbb{N}$ be constant, and let $A_1,\ldots,A_k \subseteq [0,1]$ be measurable sets depending on n that partition $[0,1]$, where each $A_i = A_i(n)$ has measure $\lambda(A_i) = \omega\left(\frac{1}{n}\right)$. Then for each i, there exist $n_i^+, n_i^- = (1+o(1))\lambda(A_i)$, with $n_i^- \le n_i^+$, such that a.a.s.,*

(i) $\omega(\mathbb{G}(n,W)) \le (1+o(1))\left[\omega(\mathbb{G}(n_1^+, W|_{A_1\times A_1})) + \cdots + \omega(\mathbb{G}(n_k^+, W|_{A_k\times A_k}))\right]$, *and*

(ii) $\omega(\mathbb{G}(n,W)) \ge (1+o(1)) \cdot \omega(\mathbb{G}(n_i^-, W|_{A_i\times A_i}))$ *for each* $i \in \{1,\ldots,k\}$.

Proof. We will show that (i) and (ii) each hold for a specific coupling of $\mathbb{G}(n,W)$ with

$$\left(\mathbb{G}(n_1^\pm, W|_{A_1\times A_1}), \ldots, \mathbb{G}(n_k^\pm, W|_{A_k\times A_k})\right).$$

But before doing so, we briefly argue that this suffices to prove the lemma for any coupling. By Theorem 3.1.10, we know that the clique number for any graphon is highly concentrated: for any graphon U, with probability $1-o(1)$, we have $\omega(\mathbb{G}(n,U)) = (1+o(1))\mathbb{E}(\omega(\mathbb{G}(n,U)))$. So if (i) or (ii) holds for any specific choice of coupling, then it will hold for all, since each term in (i) and (ii) changes by at most a factor of $1+o(1)$ regardless of the choice of coupling.

With this in mind, we prove (ii). By Lemma 3.2.1, when sampling vertices of the W-random graph $\mathbb{G}(n,W)$, the number of vertices in the set A_i for each i will be at least $n_i^- = (1-o(1))\lambda(A_i)n$ a.a.s., for an appropriate $o(1)$ function. So there is a coupling of $\mathbb{G}(n,W)$ with $\left(\mathbb{G}(n_1^-, W|_{A_1\times A_1}), \ldots, \mathbb{G}(n_k^-, W|_{A_k\times A_k})\right)$ such that a.a.s. each $\mathbb{G}(n_1^-, W|_{A_1\times A_1})$ is contained as a subgraph in $\mathbb{G}(n,W)$. For this coupling, (ii) automatically holds.

Explicitly, the coupling is constructed as follows: for each n, we sample the n vertices of $\mathbb{G}(n,W)$. With probability $1-o(1)$, there will be at least n_i vertices sampled from each A_i (note: since k is constant, taking a union bound over all the A_i does not change this). Assume we are in this case (else, generate the other graphs independently). To generate each $\mathbb{G}\left(n_i^-, W|_{A_i \times A_i}\right)$, since at least n_i^- of the vertices of $\mathbb{G}(n, W)$ are in A_i, then uniformly sample exactly n_i^- of them. The subgraph induced on these vertices has distribution $\mathbb{G}\left(n_i^-, W|_{A_i \times A_i}\right)$. We place no additional edges. Note that we re-sample all the vertices for each n to generate this coupling, as opposed to adding on a vertex to $\mathbb{G}(n-1, W)$ to generate $\mathbb{G}(n, W)$. For this coupling, (ii) holds; thus as argued above, (ii) holds in general.

Now we show that (i) also holds for a similar coupling. By Lemma 3.2.1, in $\mathbb{G}(n, W)$, the number of vertices in the set A_i will be at most $n_i^+ = (1 + o(1))\lambda(A_i)n$ for an appropriate $o(1)$ function. We couple $\mathbb{G}(n, W)$ with

$$\left(\mathbb{G}(n_1^+, W|_{A_1 \times A_1}), \ldots, \mathbb{G}(n_k^+, W|_{A_k \times A_k})\right)$$

as follows: for each n, we sample n vertices for $\mathbb{G}(n, W)$. With probability $1 - o(1)$, there will be at most n_i^+ of them in each A_i. If this happens, then to generate each $\mathbb{G}\left(n_i^+, W|_{A_i \times A_i}\right)$, take these vertices, and add in enough extra vertices (uniformly sampled from A_i) to make exactly n_i^+ total vertices in A_i. On these n_i^+ vertices, place all edges belonging to the copy of $\mathbb{G}(n, W)$ that we have sampled, and add edges between a new vertex v and any other vertex w with probability $W(v, w)$. Now, the subgraph induced on these vertices has distribution $\mathbb{G}\left(n_i^+, W|_{A_i \times A_i}\right)$. And we see that the largest clique in $\mathbb{G}(n, W)$ is, at very most, the union of the largest cliques in each of the graphs $\mathbb{G}\left(n_i^+, W|_{A_i \times A_i}\right)$. So for this coupling, (i) holds, and as a consequence, holds for any coupling.

□

Note that in the previous lemma, the quantities n_i^- and n_i^+ are functions only of n and $\lambda(A_i)$, and not of the graphon W; we will use this fact in the proof of Lemma 3.3.6.

We finish this section with two lemmas showing that the clique number of $\mathbb{G}(n, W)$ is determined up to lower-order terms (in the case where this clique number is $\omega(\log n)$) by the local behavior of W near points (a, a) where $W(a, a) = 1$. In particular, if W is bounded above by U locally near points where $W(a, a) = 1$, the following lemma tells us that the clique numbers of W and U will satisfy the same inequality, up to lower-order terms. Note that in the proof of the lemma below, there is nothing special about using graphons W and U on $[0, 1]^2$; in particular, we can take graphons on A^2 for any interval $A \subseteq [0, 1]$, as long as both graphons have the *same* domain. We use this fact in the proof of Lemma 3.3.6.

Lemma 3.2.5. *Let $W, U : [0, 1]^2 \to [0, 1]$ be graphons equal to 1 at some point (a, a), and let W be essentially bounded away from 1 in some neighborhood of (x, x) for all other $x \in [0, 1]^2$. If there exists some neighborhood N of (a, a) on which $W(x, y) \leq U(x, y)$, then a.a.s.,*

$$\omega(\mathbb{G}(n, W)) \leq (1 + o(1)) \cdot \omega(\mathbb{G}(n, U)) + O(\log n).$$

Proof. As in Lemma 3.2.4, we will show that the result holds for a specific coupling of $\mathbb{G}(n, W)$ and $\mathbb{G}(n, U)$, and as argued in the proof of Lemma 3.2.4, this will in fact suffice to prove it for any choice of coupling. We define our coupling as follows: first, sample the *same* numbers $x_1, \ldots, x_n$ for the vertices of both $\mathbb{G}(n, W)$ and $\mathbb{G}(n, U)$, and couple their edges in such a way that every pair of vertices (i, j) with $(x_i, x_j) \in N$ is connected by an edge in $\mathbb{G}(n, W)$ only if (i, j) is also connected in $\mathbb{G}(n, U)$; this is possible because $W(x, y) \leq U(x, y)$ on the neighborhood N. The coupling of the remaining edges can be defined in any way (for concreteness, we may sample them independently for the two graphs).

Before proceeding further, we will restrict our view to a subset $I^2 \subseteq N$, for some open interval I containing a. Then, we define n_{in} to be the (random) number of vertices $x_1, \ldots x_n$ that are in the interval I, and consider the random graphs $\mathbb{G}(n_{\text{in}}, W|_{I^2})$ and $\mathbb{G}(n_{\text{in}}, U|_{I^2})$. We may generate them by taking the subgraphs of $\mathbb{G}(n, W)$ and $\mathbb{G}(n, U)$ respectively induced on the vertices in I (still using the coupling of $\mathbb{G}(n, W)$ and $\mathbb{G}(n, U)$ described above). With this coupling, $\mathbb{G}(n_{\text{in}}, W|_{I^2})$ is contained as a subgraph in $\mathbb{G}(n_{\text{in}}, U|_{I^2})$; thus we may write

$$\omega(\mathbb{G}(n_{\text{in}}, W|_{I^2})) \leq \omega(\mathbb{G}(n_{\text{in}}, U|_{I^2})). \tag{3.2}$$

And as $\mathbb{G}(n_{\text{in}}, U|_{I^2})$ is a subgraph of $\mathbb{G}(n, U)$ in the coupling we have just defined, we may also write $\mathbb{G}(n_{\text{in}}, U|_{I^2}) \leq \mathbb{G}(n, U)$, or combining this with (3.2),

$$\omega(\mathbb{G}(n_{\text{in}}, W|_{I^2})) \leq \omega(\mathbb{G}(n, U)). \tag{3.3}$$

This is the essence of our proof; however, we still need to deal with all the vertices in $\mathbb{G}(n, W)$ that do not fall into the interval I, and ensure that they will not change the clique number of $\mathbb{G}(n, W)$ by too much.

We deal with the remaining vertices as follows: let n_{out} be the number of vertices not in I (i.e., $n_{\text{out}} = n - n_{\text{in}}$). By the same reasoning just used, we may generate $\mathbb{G}(n_{\text{out}}, W|_{([0,1]\setminus I)^2})$ as the subgraph of $\mathbb{G}(n, W)$ induced on the vertices counted by n_{out}. And given any partition of the vertices of a graph G, the clique number of G is at most the sum of the clique numbers of the subgraphs induced on the parts of the partition. Here, given the partition of $[n]$ into I and $[0, 1] \setminus I$, this gives

$$\omega(\mathbb{G}(n, W)) \leq \omega(\mathbb{G}(n_{\text{in}}, W|_{I^2})) + \omega(\mathbb{G}(n_{\text{out}}, W|_{([0,1]\setminus I)^2})). \tag{3.4}$$

Combining (3.4) with (3.3), we see that

$$\omega(\mathbb{G}(n, W)) \leq \omega(\mathbb{G}(n, U)) + \omega(\mathbb{G}(n_{\text{out}}, W|_{([0,1]\setminus I)^2})). \tag{3.5}$$

Since n_{out} is deterministically bounded above by n, notice that $\omega(\mathbb{G}(n_{\text{out}}, W|_{([0,1]\setminus I)^2}))$ is always at most $\omega(\mathbb{G}(n, W|_{([0,1]\setminus I)^2}))$, provided we couple these two graphs so that $\mathbb{G}(n_{\text{out}}, W|_{([0,1]\setminus I)^2})$ is a subgraph of $\mathbb{G}(n, W|_{([0,1]\setminus I)^2})$. Now, since the subgraphon $W|_{([0,1]\setminus I)^2}$ is essentially bounded away from 1 in some neighborhood of each (x, x), we may apply Lemma 3.1.4 (proved below) to conclude that $\omega(\mathbb{G}(n, W|_{([0,1]\setminus I)^2})) = O(\log n)$ a.a.s. (Note that we are taking the slight liberty of applying Lemma 3.1.4 to a graphon defined on $([0,1]\setminus I)^2$ rather than $[0,1]^2$; as will be argued after the proof of Lemma 3.1.4, this can be justified formally since the set $[0,1]\setminus I$ is compact.) With this, (3.5) becomes

$$\omega(\mathbb{G}(n,W)) \le \omega(\mathbb{G}(n,U)) + O(\log n).$$

To finish, note that as shown in the proof of Lemma 3.2.4, for *any* choice of coupling of $\mathbb{G}(n,W)$ and $\mathbb{G}(n,U)$, the clique numbers of each of these graphs will change by a factor of at most $1+o(1)$ a.a.s. Therefore, regardless of the choice of coupling,

$$\omega(\mathbb{G}(n,W)) \le (1+o(1))\omega(\mathbb{G}(n,U)) + O(\log n)$$

with probability $1-o(1)$, as desired. □

We end this section with a proof of Lemma 3.1.4, restated here for the convenience of the reader.

Lemma 3.1.4. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon whose essential supremum is strictly less than 1 in some neighborhood of each point (x,x) for $x \in [0,1]$. Then $\omega(\mathbb{G}(n,W)) = O(\log n)$ a.a.s.*

Proof of Lemma 3.1.4. Associate to each point (x,x) an open neighborhood $N(x,x)$ in $[0,1]^2$ on which the essential supremum of W is $c(x) < 1$. These neighborhoods form an open cover of the (closed) diagonal line segment $D = \{(x,x) : x \in [0,1]\}$. Because this set is compact, we may find a finite subcover of D by neighborhoods $N(x,x)$. Taking the maximum essential supremum $c = c(x)$ of W on any of these neighborhoods, we see that for some $\varepsilon > 0$, the essential supremum of W is c on the region $\{(x,y) : |x-y| \le \varepsilon\}$.

Now consider any k vertices from $\mathbb{G}(n,W)$, and view them as points in $[0,1]$. By the pigeonhole principle, dividing $[0,1]$ into $1/\varepsilon$ disjoint intervals of length ε, of the k points, there must be at least εk points in some interval of length ε. The probability that this subset forms a clique is at most $c^{\binom{\varepsilon k}{2}} = c^{\Theta(k^2)}$, which also gives an upper bound on the probability that the original k vertices formed a clique. So, taking a union bound, the probability that there exists any k-clique is at most

$$\binom{n}{k} c^{\Theta(k^2)} \le \left(\frac{en}{(\frac{1}{c})^{\Theta(k)} \cdot k}\right)^k = \left(\frac{n}{(\frac{1}{c})^{\Theta(k)}}\right)^k.$$

The cutoff at which this approaches zero is $k = \Theta(\log n)$. So for any graphon W bounded away from 1 in some neighborhood of each point (x,x), the clique number of $\mathbb{G}(n,W)$ is a.a.s. $O(\log n)$. □

Notice that there was nothing special about the choice of $[0,1]^2$ in this result; the only property we used of the interval $[0,1]$ was its compactness. So in particular, Lemma 3.1.4 holds for a graphon defined on A^2 for any closed interval A, a fact that we will use several times throughout this chapter.

3.3 W-random graphs with clique number $\Theta(\sqrt{n})$

In this section, we prove Lemma 3.1.6 and Theorem 3.1.5, which characterize a variety of W-random graphs with clique number $\Theta(\sqrt{n})$ in terms of the local behavior of W at points where it is equal to 1. We begin by finding the clique number of a specific family of random graphs; this will in fact suffice to prove both Lemma 3.1.6 and a more general result, of which Theorem 3.1.5 is a special case.

3.3.1 A family of examples with clique number $\Theta(\sqrt{n})$

Lemma 3.3.1. *For any $r > 0$, define the graphon $W_r(x, y) = (1 - x)^r(1 - y)^r$. The clique number of $\mathbb{G}(n, W_r)$ is a.a.s. $\Theta(\sqrt{n})$.*

To prove Lemma 3.3.1, we begin by finding an upper bound on $\omega(\mathbb{G}(n, W_r))$; we will use the first moment method.

Lemma 3.3.2. *The clique number of $\mathbb{G}(n, W_r)$ is a.a.s. at most $(1+o(1)) \left(\frac{e}{r}\right)^{1/2} \cdot \sqrt{n}$.*

Proof. Write X_k for the number of cliques of size k in $\mathbb{G}(n, W_r)$. By Markov's inequality, $\omega(\mathbb{G}(n, W_r))$ is a.a.s. bounded above by any k for which $\mathbb{E}[X_k]$ is asymptotically 0. And for any k, writing $d\vec{x}$ in place of $dx_1 \cdots dx_k$, we have

$$\begin{aligned}
\mathbb{E}[X_k] &= \binom{n}{k} \int_{[0,1]^k} \prod_{\ell, m \in [k],\, \ell \neq m} W(x_\ell, x_m)\ d\vec{x} \\
&= \binom{n}{k} \int_{[0,1]^k} \prod_{\ell, m \in [k],\, \ell \neq m} (1 - x_\ell)^r \cdot (1 - x_m)^r\ d\vec{x} \\
&= \binom{n}{k} \left(\int_0^1 (1 - x)^{r(k-1)} dx \right)^k \\
&= \binom{n}{k} \left(\frac{1}{r(k-1)+1} \right)^k.
\end{aligned}$$

For any k that is $\omega(1)$, we have $\frac{1}{r(k-1)+1} = (1 + o(1))\frac{1}{rk}$. And for any k that is $\omega(1)$ but sublinear, it can be shown from Stirling's formula that $\binom{n}{k} = \left(\frac{en}{k}(1 - o(1))\right)^k$. Therefore the above expression becomes

$$\mathbb{E}[X_k] = \left(\frac{en}{k}(1 - o(1))\right)^k \left((1 + o(1))\frac{1}{rk}\right)^k = \left(\frac{en}{rk^2}(1 - o(1))\right)^k.$$

So the cutoff at which $\mathbb{E}[X_k]$ goes from asymptotically 0 to asymptotically infinity is when $k \sim \left(\frac{e}{r}\right)^{1/2} \cdot \sqrt{n}$, which, by Markov's inequality, gives an upper bound on the clique number of $\mathbb{G}(n, W_r)$ that will hold with probability $1 - o(1)$. $\square$

Ideally, we would like to prove a matching lower bound. However, such a bound may be difficult to establish, or even untrue, as the variance of the number of cliques in $\mathbb{G}(n, W_r)$ of any size of order $\Theta(\sqrt{n})$ is quite large (Corollary A.0.5 ((i)) in Appendix A). In particular, this means we cannot use the second moment method directly to prove a lower bound on the clique number $\omega(\mathbb{G}(n, W_r))$. (This argument is fleshed out more fully in Appendix A.) These difficulties notwithstanding, we can at least prove a lower bound that matches up to a constant.

Lemma 3.3.3. *The clique number of* $\mathbb{G}(n, W_r)$ *is a.a.s. at least* $\left(\frac{1}{12er}\right)^{1/2} \cdot \sqrt{n}$.

Proof. Our strategy is to directly compute a lower bound on the expected clique number for the graphon $W_r(x, y) = (1-x)^r(1-y)^r$ by guessing which vertices are most likely to form a large clique and showing that this does indeed happen with high probability. Suppose that for some constants s and t, there are $sn^{1/2}$ vertices less than $tn^{-1/2}$ in $\mathbb{G}(n, W_r)$. (Note: the expected number of vertices less than $tn^{-1/2}$ is $tn^{1/2}$.) By Lemma 3.2.1, this will happen a.a.s. for some $t = (1 + o(1))s$. We will show, for an appropriate choice of s, that if we do have such vertices, then a.a.s., the subgraph they induce will contain all but k possible edges (for some appropriate choice of k dependent on n). In total, this will show that the clique number is a.a.s. at least $s\sqrt{n} - k$, obtained by greedily deleting one vertex from each of the (up to) k missing edges.

Concretely, for any constants s and t, suppose that $\mathbb{G}(n, W_r)$ has $sn^{1/2}$ vertices less than $tn^{-1/2}$. The probability that any fixed set of k potential edges is missing from the subgraph of $\mathbb{G}(n, W_r)$ induced on those vertices is at most

$$\begin{aligned}\prod_{k \text{ edges}} \left[1 - (1 - tn^{-1/2})^r(1 - tn^{-1/2})^r\right] &= \left[1 - (1 - tn^{-1/2})^{2r}\right]^k \\ &= \left[1 - \left(1 - 2rtn^{-1/2} \cdot (1 + o(1))\right)\right]^k,\end{aligned}$$

where the last equality follows by taking a binomial series expansion. Simplifying this expression slightly, the probability that any fixed set of k edges is missing is at most

$$\left[2rtn^{-1/2} \cdot (1 + o(1))\right]^k.$$

Now to bound the probability that there are k or more edges missing, we take a union bound over all sets of k possible edges in the subgraph induced on the $s\sqrt{n}$ vertices under consideration. The number of such sets is

$$\binom{\binom{s\sqrt{n}}{2}}{k} \le \binom{s^2n/2}{k} \le \left(\frac{es^2n/2}{k}\right)^k.$$

So in total, the probability to have k or more missing edges is at most

$$\left(\frac{es^2n/2}{k}\right)^k \left(2rtn^{-1/2} \cdot (1 + o(1))\right)^k = \left(\frac{erts^2\sqrt{n}}{k} \cdot (1 + o(1))\right)^k.$$

As argued above, for any constant s, with probability $1 - o(1)$, there is a set of $s\sqrt{n}$ vertices less than $tn^{-1/2}$, for some $t = (1 + o(1))s$. Given such a set, as just shown, the probability that the induced subgraph on these vertices is missing k or more edges is at most $\left(\frac{erts^2\sqrt{n}}{k}\cdot(1+o(1))\right)^k$. Thus if this quantity is $o(1)$, then a.a.s., there is a clique of size at least $s\sqrt{n}-k$ in $\mathbb{G}(n, W_r)$, obtained by deleting one vertex from each missing edge. If we choose k to be, for example, $\frac{1}{2}s\sqrt{n}$, then

$$\left(\frac{erts^2\sqrt{n}}{k}\cdot(1+o(1))\right)^k = \left(2ers^2\cdot(1+o(1))\right)^{\frac{1}{2}s\sqrt{n}}.$$

This will be $o(1)$ as long as $2ers^2 = 1-\Omega(1)$, or equivalently, $s^2 = \frac{1-\Omega(1)}{2er}$. Taking any constant $s < \frac{1}{\sqrt{2er}}$ suffices, for instance $s = \frac{1}{\sqrt{3er}}$. Therefore, a.a.s., there will exist a clique of size at least $s\sqrt{n}-k = \frac{1}{2}\cdot\frac{1}{\sqrt{3er}}\sqrt{n} = \left(\frac{1}{12er}\right)^{1/2}\cdot\sqrt{n}$. □

Note that the bound in Lemma 3.4.2 can be tightened by optimizing the choice of k in the proof above, but not to the point of matching the upper bound given in Lemma 3.3.2. Together, the upper and lower bounds in Lemmas 3.3.2 and 3.3.3 imply the $\Theta(\sqrt{n})$ bound given in Lemma 3.3.1 for $\omega(\mathbb{G}(n, W_r))$.

3.3.2 More general $\mathbb{G}(n, W)$ with clique number $\Theta(\sqrt{n})$

We are now nearly ready to prove the main results of this section, Lemma 3.1.6 and Theorem 3.1.5. We will reframe both results in a slightly broader setting and prove a more general version of Theorem 3.1.5. This theorem characterizes the clique number of a W-random graph when W has moderate directional derivatives at the points where it is equal to 1. However, even if the directional derivatives of a graphon W do not exist at a given point, we can still have some notion of "bounded derivatives" by looking at the limsup and the liminf of the difference quotient that defines the derivative.

Definition 3.3.4. *For a function $W\colon \mathbb{R}^k \to \mathbb{R}$, a point $x \in \mathbb{R}^k$, and a unit direction vector $d \in \mathbb{R}^k$, the* upper Dini derivative *of W at x in direction d is defined as*

$$W'_+(x, d) = \limsup_{h\to 0^+} \frac{W(x+hd)-W(x)}{h}.$$

The lower Dini derivative *of W at x in direction d is*

$$W'_-(x, d) = \liminf_{h\to 0^+} \frac{W(x+hd)-W(x)}{h}.$$

Notice that if any directional derivative of a graphon W exists, then it is equal to both the corresponding upper and lower Dini derivatives. Also, we have defined Dini derivatives only in directions corresponding to unit vectors; this is not necessary, but it makes several of the results and their proofs below slightly neater. We will use these definitions throughout the remainder of this section. We now show that a bound on the lower Dini derivatives of a graphon W at a point (a, a) with $W(a, a) = 1$ provides a lower bound on the clique number of $\omega(\mathbb{G}(n, W))$.

Lemma 3.3.5. *Let* $W\colon [0,1]^2 \to [0,1]$ *be a graphon equal to 1 at some point* (a, a)*.*

(i) If all lower Dini derivatives of W *at* (a, a) *are bounded below by* $-c$ *for some constant* $c \geq 0$*, then* $\omega(\mathbb{G}(n, W)) = \Omega(\sqrt{n})$ *a.a.s.*

(ii) If all directional derivatives of W *at* (a, a) *exist and are equal to zero, then* $\omega(\mathbb{G}(n, W)) = \omega(\sqrt{n})$ *a.a.s.*

Before proving Lemma 3.3.5, we quickly show how it implies Lemma 3.1.6, which is restated here for the convenience of the reader.

Lemma 3.1.6. *Let* $W\colon [0,1]^2 \to [0,1]$ *be a graphon equal to 1 at some point* (a, a)*. If* W *is locally Lipschitz continuous at* (a, a)*, then* $\omega(\mathbb{G}(n, W)) = \Omega(\sqrt{n})$ *a.a.s.*

Proof of Lemma 3.1.6. We begin by recalling a definition: W is locally Lipschitz at the point (a, a) if there exists a neighborhood U of (a, a) and a constant $c \geq 0$ such that for all points (x, y) in U,

$$|W(x, y) - W(a, a)| \leq c \cdot \|(x, y) - (a, a)\|.$$

where in the line above, $\|\cdot\|$ represents the ℓ^2 norm. (Note that any other norm would produce an equivalent definition, as all norms on $\mathbb{R}^2$ are equivalent up to a constant.) If W satisfies this condition, and if $W(a, a) = 1$, then since $W(x, y) \leq 1$, the inequality above becomes

$$W(x, y) - W(a, a) \geq -c \cdot \|(x, y) - (a, a)\|.$$

Now for any $(x, y) \in U$, write $(x, y) - (a, a) = hd$ for a unit direction vector d; with this substitution, the inequality above is equivalent to

$$\frac{W((a, a) + hd) - W(a, a)}{h} \geq -c. \tag{3.6}$$

Indeed, for any unit direction vector d, and for h sufficiently small, the point $(a, a)+hd$ will be in the neighborhood U, and inequality (3.6) will hold. Thus, taking a liminf of inequality (3.6) for each d as $h \to 0^+$, we see that by definition, all the lower Dini derivatives of W at (a, a) are at least $-c$. Then Lemma 3.1.6 follows immediately from Lemma 3.3.5 (i). □

We now give the proof of Lemma 3.3.5.

Proof of Lemma 3.3.5. We begin with part (i). Roughly, our proof strategy will be to locally bound W from below by a graphon in the family $\{W_r\}_{r \in \mathbb{R}^+}$ defined in Lemma 3.3.1, thereby bounding the clique number of $\mathbb{G}(n, W)$ from below by $\omega(\mathbb{G}(n, W_r) = \Theta(\sqrt{n})$.

Take any constant $\varepsilon > 0$, and let $r = \sqrt{2}(c+\varepsilon)$. Notice that the graphon

$$W_r(x,y) = W_{\sqrt{2}(c+\varepsilon)}(x,y) = (1-x)^{\sqrt{2}(c+\varepsilon)}(1-y)^{\sqrt{2}(c+\varepsilon)}$$

has directional derivatives at most $-\frac{r}{\sqrt{2}} = -(c+\varepsilon)$ at $(0,0)$, achieved in the direction $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. Also, $W_r(0,0) = 1$. Then since $W(a,a) = 1$, and the lower Dini derivatives of W are at least $-c$, we see that up to translation of the domain, W is bounded below by W_r in some neighborhood of (a,a). Therefore, by Lemma 3.2.5, we have

$$\omega(\mathbb{G}(n,W)) \geq (1-o(1))\omega(\mathbb{G}(n,W_r)) - O(\log n) \tag{3.7}$$

a.a.s. Note that we have not assumed that W is bounded away from 1 away from the point (a,a); however, we may still apply Lemma 3.2.5, as we are only looking for a lower bound on $\omega(\mathbb{G}(n,W))$.

And by Lemma 3.3.3, the clique number of $\mathbb{G}(n,W_r)$ is at least $\left(\frac{1}{12er}\right)^{1/2} \cdot \sqrt{n}$; thus (3.7) becomes

$$\omega(\mathbb{G}(n,W)) \geq (1-o(1))\left(\tfrac{1}{12er}\right)^{1/2} \cdot \sqrt{n} - O(\log n) = \Theta(\sqrt{n}).$$

This proves part (i).

The proof of (ii) is similar; if all directional derivatives of W are equal to zero, then for any constant $r > 0$, consider the graphon $W_r(x,y) = (1-x)^r(1-y)^r$. Since the directional derivatives of W_r at $(0,0)$ are at most $-\frac{r}{\sqrt{2}}$, we have $W \geq W_r$ in some neighborhood of (a,a), up to translation of the domain. Thus, again by Lemma 3.2.5,

$$\omega(\mathbb{G}(n,W)) \geq (1-o(1))\omega(\mathbb{G}(n,W_r) - O(\log n)$$

And as above, this yields

$$\begin{aligned}\omega(\mathbb{G}(n,W)) &\geq (1-o(1)) \cdot \left(\tfrac{1}{12er}\right)^{1/2} - O(\log n) \\ &= (1-o(1))\left(\tfrac{1}{12er}\right)^{1/2} \cdot \sqrt{n}.\end{aligned}$$

Then since we can choose r arbitrarily small, we see that

$$\omega(\mathbb{G}(n,W)) = \omega(\sqrt{n}),$$

a.a.s., completing the proof of part (ii). $\square$

Just as a bound on the lower Dini derivatives of a graphon W gives us a lower bound on the clique number of a W-random graph, a bound on the upper Dini derivatives will give us an upper bound. Since we are proving an upper bound on the clique number, we will add the assumption that the graphon under consideration is only equal to 1 at a finite number of points (a,a). Together with Lemma 3.1.6, the following result will prove Theorem 3.1.5.

Lemma 3.3.6. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some collection of* *points $(a_1, a_1), \ldots, (a_k, a_k)$, and essentially bounded away from 1 in some neighborhood of each other point (x, x) for $x \in [0, 1]$. Then*

(i) if all upper Dini derivatives of W at $(a_1, a_1), \ldots, (a_k, a_k)$ are uniformly bounded away from zero then $\omega(\mathbb{G}(n, W)) = O(\sqrt{n})$ a.a.s., and

(ii) if all upper Dini derivatives of W at $(a_1, a_1), \ldots, (a_k, a_k)$ are equal to $-\infty$, then $\omega(\mathbb{G}(n, W)) = o(\sqrt{n})$ a.a.s.

Before giving the proof, we briefly show how Theorem 3.1.5 follows as a direct consequence of part (i) of this lemma, together with Lemma 3.1.6.

Proof of Theorem 3.1.5. If W is equal to 1 at the points $(a_1, a_1), \ldots, (a_k, a_k)$, and its directional derivatives exist and are uniformly bounded away from $-\infty$ at these points, then as argued in the proof of Lemma 3.1.6, W is locally Lipschitz at these points. Therefore, we can apply Lemma 3.1.6 and conclude that $\omega(\mathbb{G}(n, W)) = \Omega(\sqrt{n})$ a.a.s.

Similarly, if the directional derivatives of W are uniformly bounded away from 0 at the points $(a_1, a_1), \ldots, (a_k, a_k)$, and if W is essentially bounded away from 1 in some neighborhood of each other point (x, x) for $x \in [0, 1]$, then we may apply Lemma 3.3.6 (i) to obtain $\omega(\mathbb{G}(n, W)) = O(\sqrt{n})$ a.a.s. Therefore $\omega(\mathbb{G}(n, W)) = \Theta(\sqrt{n})$ a.a.s., as desired. □

We now prove the lemma above.

Proof of Lemma 3.3.6. First, divide $[0, 1]$ into subintervals so that each contains only one point of interest; specifically, divide $[0, 1]$ at each point a_i and at an arbitrary point between each pair a_i and a_{i+1}. This will produce a partition of $[0, 1]$ into a total of $2k$ subintervals $A_1, \ldots, A_{2k}$ so that for each A_i, either the left or right endpoint is one of the values a_j, and $W|_{A_i \times A_i}$ is essentially bounded away from 1 in some neighborhood of each other $(x, x) \neq (a_j, a_j)$. We will bound the clique number of $\mathbb{G}(n, W)$ in terms of the subgraphons $W|_{A_i \times A_i}$. By Lemma 3.2.4 (i), a.a.s.,

$$\omega(\mathbb{G}(n, W)) \leq (1 + o(1))\big[\omega(\mathbb{G}(n_1^+, W|_{A_1 \times A_1})) + \cdots + \omega(\mathbb{G}(n_{2k}^+, W|_{A_{2k} \times A_{2k}}))\big] \quad (3.8)$$

where each n_i^+ is of the form $n_i^+ = (1 + o(1))\lambda(A_i)n$, and is a function only of n and $\lambda(A_i)$, and not of W (this fact follows from the proof of Lemma 3.2.4).

Now, to prove part (i), take any A_i, and suppose the upper Dini derivatives of $W|_{A_i \times A_i}$ are at most $-c$ at (a, a), for the endpoint a of A_i at which $W(a, a) = 1$. For any $0 < \varepsilon < c$, take $r = c - \varepsilon$, and consider the graphon

$$W_r(x, y) = W_{c-\varepsilon}(x, y) = (1 - x)^{c-\varepsilon}(1 - y)^{c-\varepsilon},$$

as defined in Lemma 3.3.1. We have $W_r(0, 0) = 1$, and at the point $(0, 0)$, all the directional derivatives of $W_{c+\varepsilon}$ are at least $-r = -c + \varepsilon$ (achieved in the directions $(0, 1)$ and $(1, 0)$). Therefore, up to translation and/or reflection of its domain, $W|_{A_i \times A_i}$ is bounded above by W_r in some neighborhood of $(0, 0)$, and essentially bounded away from 1 near each $(x,x) \neq (0,0)$. The same statement also holds if we replace W_r with $W_r|_{[0,\ell]^2}$, for $\ell = \lambda(A_i)$, and in this case, up to translation and/or reflection of the domain, $W|_{A_i \times A_i}$ and $W_r|_{[0,\ell]^2}$ are graphons on the *same* interval. Therefore we may apply Lemma 3.2.5 to conclude that

$$\omega(\mathbb{G}(n_i^+, W|_{A_i \times A_i})) \leq (1+o(1)) \cdot \omega(\mathbb{G}(n_i^+, W_r|_{[0,\ell]^2})) + O(\log n_i^+) \tag{3.9}$$

a.a.s. And by Lemma 3.2.4 (ii),

$$(1+o(1)) \cdot \omega(\mathbb{G}(n_i^-, W_r|_{[0,\ell]^2})) \leq \omega(\mathbb{G}(n, W_r))$$

a.a.s. as well, where $n_i^- = (1-o(1))\lambda(A_i)n = (1-o(1))n_i^+$. Equivalently, rearranging slightly,

$$\begin{aligned} \omega(\mathbb{G}(n_i^+, W_r|_{[0,\ell]^2})) &\leq (1-o(1)) \cdot \omega(\mathbb{G}(n(1+o(1)), W_r)) \\ &\leq (1-o(1))\sqrt{\tfrac{e}{r}} \cdot \sqrt{n(1+o(1))} \qquad (3.10) \\ &= (1+o(1))\sqrt{\tfrac{e}{r}} \cdot \sqrt{n}, \qquad (3.11) \end{aligned}$$

where (3.10) is a direct application of Lemma 3.3.2. Together, (3.9) and (3.11) imply that

$$\omega(\mathbb{G}(n_i^+, W|_{A_i \times A_i})) \leq (1+o(1))\sqrt{\tfrac{e}{r}} \cdot \sqrt{n} = O(\sqrt{n})$$

a.a.s. Since this is true for each i, equation (3.8) becomes.

$$\omega(\mathbb{G}(n, W)) \leq 2k \cdot O(\sqrt{n}) = O(\sqrt{n}),$$

proving part (i).

To prove part (ii), recall that the directional derivatives of W_r at $(0,0)$ are at least $-r$, as mentioned in the proof of part (i). So if the upper Dini derivatives of W are $-\infty$ at each of the points (a_i, a_i), then for each A_i and any $r > 0$, we have (up to translation and/or reflection) $W|_{A_i \times A_i} \leq W_r$ locally on some neighborhood, and $W|_{A_i \times A_i}$ is essentially bounded away from 1 near each (x,x) outside that neighborhood. Then for any $r > 0$, as argued in the proof of part (i), equations (3.9) and (3.11) hold here as well, again implying that

$$\omega(\mathbb{G}(n_i^+, W|_{A_i \times A_i})) \leq (1-o(1))\sqrt{\tfrac{e}{r}} \cdot \sqrt{n}$$

a.a.s. Then since we can choose r arbitrarily large, we see that $\omega(\mathbb{G}(n_i^+, W|_{A_i \times A_i})) = o(\sqrt{n})$ a.a.s. Substituting into (3.8), this gives

$$\omega(\mathbb{G}(n, W)) \leq 2k \cdot o(\sqrt{n}) = o(\sqrt{n})$$

a.a.s., completing the proof of (ii). $\square$

3.4 A family of W-random graphs with clique number $\Theta(n^\alpha)$

As seen in the previous section, graphons W with moderate local growth near points where $W(x,x) = 1$ produce W-random graphs with clique numbers $\Theta(\sqrt{n})$. In this section, we prove Theorem 3.1.7, which introduces a family of graphons with clique numbers $\Theta(n^\alpha)$ for any $\alpha \in (0,1)$. The members W of this family corresponding to $\alpha \neq \frac{1}{2}$ have directional derivatives either 0 or $-\infty$ at points where $W(x,x) = 1$, consistent with the results of the previous section. In this section, we also prove Theorem 3.1.8, which characterizes a larger class of W-random graphs with clique numbers $\Omega(n^\alpha)$, and Proposition 3.1.9, which gives an example of a W-random graph with clique number $n^{1-o(1)}$. We begin by proving Theorem 3.1.7, restated here for the convenience of the reader.

Theorem 3.1.7. *For any constant $r > 0$, define the graphon*

$$U_r(x,y) := (1-x^r)(1-y^r).$$

The random graph $\mathbb{G}(n, U_r)$ a.a.s. has clique number $\Theta(n^{\frac{r}{r+1}})$.

We will prove Theorem 3.1.7 in very much in the same way as Lemma 3.3.1; first, we prove an upper bound on the clique number of $\mathbb{G}(n, U_r)$ by the first moment method.

Lemma 3.4.1. *For any $r > 0$, the clique number of the random graph $\mathbb{G}(n, U_r)$ is at most $(1+o(1)) \cdot \left(\Gamma\left(1+\frac{1}{r}\right)e\right)^{\frac{r}{r+1}} \cdot n^{\frac{r}{r+1}} = \Theta(n^{\frac{r}{r+1}})$ a.a.s.*

Proof. For any $r > 0$, write X_k for the number of cliques of size k in $\mathbb{G}(n, U_r)$. By Markov's inequality, the expected clique number of any random graph is a.a.s. bounded above by any value of k for which $\mathbb{E}[X_k] = o(1)$. And for any k,

$$\begin{aligned} \mathbb{E}[X_k] &= \binom{n}{k} \int_{[0,1]^k} \prod_{\ell, m \in [k],\, \ell \neq m} (1-x_\ell^r) \cdot (1-x_m^r)\, d\vec{x} \\ &= \binom{n}{k} \left(\int_0^1 (1-x^r)^{k-1}\, dx \right)^k. \end{aligned}$$

Using the change of variables $u = x^r$, this expression becomes

$$\binom{n}{k} \left(\frac{1}{r} \cdot \int_0^1 u^{\frac{1}{r}-1}(1-u)^{k-1}\, du \right)^k = \binom{n}{k} \left(\frac{1}{r} \cdot \frac{\Gamma(k)\Gamma(\frac{1}{r})}{\Gamma(k+\frac{1}{r})} \right)^k,$$

where the last equality follows from the definition of the beta function, and its relationship to the gamma function (see, for example, Definition 1.1.3 and Theorem 1.1.4

in [AAR99]). Simplifying slightly, we obtain

$$\mathbb{E}[X_k] = \binom{n}{k} \left(\frac{\Gamma(k)\Gamma(1+\frac{1}{r})}{\Gamma(k+\frac{1}{r})} \right)^k. \tag{3.12}$$

And for any k that is $\omega(1)$ but sublinear, we have $\binom{n}{k} = \left(\frac{en}{k}(1-o(1))\right)^k$; thus (3.12) becomes

$$\mathbb{E}[X_k] = \left(\frac{en}{k}(1-o(1))\right)^k \left(\frac{\Gamma(k)\Gamma(1+\frac{1}{r})}{\Gamma(k+\frac{1}{r})} \right)^k.$$

To obtain explicit asymptotics for this expression, we use Stirling's formula for the gamma function (for example, Theorem 1.4.1 in [AAR99]), which states that for $x \to \infty$,

$$\Gamma(x) = (1+o(1))\sqrt{2\pi}\, x^{x-1/2} e^{-x}.$$

From this, for any fixed $r > 0$ and $k \to \infty$, it follows that

$$\frac{\Gamma(k)}{\Gamma(k+\frac{1}{r})} = (1+o(1))k^{-\frac{1}{r}}.$$

Substituting this into (3.12), we see that

$$\begin{aligned}
\mathbb{E}[X_k] &= \left(\frac{en}{k}(1-o(1))\right)^k \left((1+o(1))k^{-\frac{1}{r}} \cdot \Gamma(1+\tfrac{1}{r})\right)^k \\
&= \left(\frac{en \cdot \Gamma(1+\frac{1}{r})}{k^{1+\frac{1}{r}}}(1+o(1)) \right)^k.
\end{aligned}$$

Therefore the cutoff at which $\mathbb{E}[X_k]$ goes from asymptotically 0 to asymptotically infinity is when $k^{1+\frac{1}{r}} \sim en \cdot \Gamma(1+\frac{1}{r})$, or equivalently, $k \sim \left(\Gamma\left(1+\frac{1}{r}\right)e\right)^{\frac{r}{r+1}} \cdot n^{\frac{r}{r+1}}$. Hence, with probability $1-o(1)$, the clique number of $\mathbb{G}(n, U_r)$ is at most

$$k = (1+o(1)) \cdot \left(\Gamma\left(1+\tfrac{1}{r}\right)e\right)^{\frac{r}{r+1}} \cdot n^{\frac{r}{r+1}} = \Theta(n^{\frac{r}{r+1}}).$$

□

Now we will prove a lower bound on the clique number of $\mathbb{G}(n, U_r)$ – as in the previous section, it will match the upper bound up to a constant.

Lemma 3.4.2. *The clique number of $\mathbb{G}(n, U_r)$ is a.a.s. at least $\frac{1}{2} \cdot e^{-\frac{2}{1+r}} \cdot n^{\frac{r}{r+1}}$.*

Proof. As in the proof of Lemma 3.3.3, we will directly compute a lower bound on the expected clique number for $\mathbb{G}(n, U_r)$ by guessing which vertices are most likely to form a large clique (after a few problematic vertices are greedily deleted), and showing that this does in fact happen with high probability. Suppose that there are $sn^{\frac{r}{r+1}}$ vertices less than $tn^{-\frac{1}{r+1}}$. (Note that the expected number of such vertices is $n \cdot tn^{-\frac{1}{r+1}} = tn^{\frac{r}{r+1}}$.) By Lemma 3.2.1, for any constant s, there is some $t = (1+o(1))s$, such that this will occur with probability $1 - o(1)$. Then, given a set of $sn^{\frac{r}{r+1}}$ such vertices, what is the probability that the subgraph they induce is missing at most k edges? The probability that any fixed set of k potential edges is missing is at most

$$\begin{aligned}
\prod_{k \text{ edges}} & \left[1 - (1 - (tn^{-\frac{1}{r+1}})^r)(1 - (tn^{-\frac{1}{r+1}})^r)\right] \\
&= \left[1 - (1 - (tn^{-\frac{1}{r+1}})^r)^2\right]^k \\
&= \left[t^r n^{-\frac{r}{r+1}} \left(2 - t^r n^{-\frac{r}{r+1}}\right)\right]^k \\
&\leq \left(t^r n^{-\frac{r}{r+1}} \cdot 2\right)^k .
\end{aligned}$$

Then, by a union bound, the probability that there exists any set of k edges missing from the induced subgraph on these $sn^{\frac{r}{r+1}}$ vertices is

$$\begin{aligned}
\binom{\binom{sn^{\frac{r}{r+1}}}{2}}{k} \cdot \left(t^r n^{-\frac{r}{r+1}} \cdot 2\right)^k &\leq \left(\frac{e(s^2 n^{\frac{2r}{r+1}}/2)}{k}\right)^k \cdot \left(t^r n^{-\frac{r}{r+1}} \cdot 2\right)^k \\
&= \left(\frac{et^r s^2 n^{\frac{r}{r+1}}}{k}\right)^k .
\end{aligned}$$

If we choose k to be, for example $\frac{1}{2}sn^{\frac{r}{r+1}}$, then this is equal to $(2et^r s)^{\frac{1}{2}sn^{\frac{r}{r+1}}}$. As long as $2et^r s = 1 - \Omega(1)$, or equivalently, $s^{1+r} = \frac{1-\Omega(1)}{2e}$, we will have $(2et^r s)^{\frac{1}{2}sn^{\frac{r}{r+1}}} = o(1)$. Taking any constant $s < (2e)^{-\frac{1}{1+r}}$ suffices, for example $s = e^{-\frac{2}{1+r}}$. Therefore, for such a constant s, the induced subgraph on the $sn^{\frac{r}{r+1}}$ vertices under consideration is missing at most $k = \frac{1}{2}sn^{\frac{r}{r+1}}$ edges with probability $1 - o(1)$. Deleting one vertex from each of these non-edges, we obtain a clique of size at least $sn^{\frac{r}{r+1}} - k = \frac{1}{2}sn^{\frac{r}{r+1}} = \frac{1}{2} \cdot e^{-\frac{2}{1+r}} \cdot n^{\frac{r}{r+1}}$ a.a.s. $\square$

Notice that, as in Lemma 3.3.3, this does not quite match the upper bound of $(\Gamma\left(1 + \frac{1}{r}\right) e)^{\frac{r}{r+1}} \cdot n^{\frac{r}{r+1}}$; we could narrow the gap somewhat by optimizing parameters in the proof just given, but not to the point of closing it entirely. And as in Section 3.3, the number of cliques of any size of the order $\Theta\left(n^{\frac{r}{r+1}}\right)$ in $\mathbb{G}(n, U_r)$ has quite high variance, which tells us that we cannot directly apply the second moment method to show that the lower bound we have given is tight (as indeed, it may not be). This argument is fleshed out more fully in Appendix A, with a variance bound given by Corollary A.0.5 ((ii)).

We now use Theorem 3.1.7 to prove Theorem 3.1.8, restated below. But first, we briefly discuss the continuity hypothesis in Theorem 3.1.8.

Definition 3.4.3. *A graphon W is* locally α-Hölder continuous *at (a, a) if there exists some neighborhood U of (a, a) and some constant $C > 0$ such that for all points*

$(a,a)+(x,y) \in U$,

$$\bigl|W\bigl((a,a)+(x,y)\bigr)-W(a,a)\bigr| < C\|(x,y)\|^{\alpha}, \tag{3.13}$$

where $\|\cdot\|$ may be taken to represent any fixed norm on $\mathbb{R}^2$.

Typically, local α-Hölder continuity is defined only for $\alpha \in [0,1]$; however, everything we do here will in fact hold and have meaning for larger α as well. On an interval, α-Hölder continuity with $\alpha > 1$ holds only for a constant function, but this is not the case for local α-Hölder continuity at a single point, which may be achieved by a non-constant function whose derivatives are equal to zero at the point in question.

Theorem 3.1.8. *Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some point (a,a). If W is locally α-Hölder continuous at (a,a) for some constant α, then $\omega(\mathbb{G}(n,W)) = \Omega(n^{\frac{\alpha}{\alpha+1}})$ a.a.s.*

Proof of Theorem 3.1.8. If W is α-Hölder continuous at (a,a), then there exist $C>0$ and a neighborhood U of (a,a) such that (3.13) is satisfied. For convenience, we will use the infinity norm. Then, since $W(a,a)=1$, (3.13) becomes

$$1-W\bigl((a,a)+(x,y)\bigr) < C\cdot\max(x,y)^{\alpha}. \tag{3.14}$$

With this in hand, we will prove a lower bound on the clique number of W by bounding W from below locally by a slightly modified member of the family $\{U_r\}$. Assume without loss of generality that the constant C is at least 1. Then we define

$$U_{\alpha,C}(x,y) = \begin{cases} (1-Cx^{\alpha})(1-Cy^{\alpha}) & \text{for } x,y \in [0,\frac{1}{C^{1/\alpha}}], \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Notice that for $x,y \in [0,\frac{1}{C^{1/\alpha}}]$,

$$\begin{aligned} 1-U_{\alpha,C}(x,y) &= 1-(1-Cx^{\alpha})(1-Cy^{\alpha}) \\ &= Cx^{\alpha}+Cy^{\alpha}(1-Cx^{\alpha}) \\ &\geq Cx^{\alpha} \end{aligned}$$

Similarly, we have $1-U_{\alpha,C}(x,y) \geq Cy^{\alpha}$; thus $1-U_{\alpha,C}(x,y) \geq C\cdot\max{(x,y)}^{\alpha}$. Therefore, by (3.14), we can write

$$1-W\bigl((a,a)+(x,y)\bigr) < C\cdot\max(x,y)^{\alpha} \leq 1-U_{\alpha,C}(x,y).$$

for (x,y) in some neighborhood of $(0,0)$. So up to translation, W is bounded below by $U_{\alpha,C}$ in some neighborhood of (a,a). (To be precise, we have only shown this in one quadrant, but this is sufficient for our purposes here.)

And indeed, the clique number of $\mathbb{G}(n,U_{\alpha,C})$ is $\Theta(n^{\frac{\alpha}{\alpha+1}})$, as with U_{α} (given in Theorem 3.1.7). To see this, first notice that in $\mathbb{G}(n,U_{\alpha,C})$, there will be some random number N of vertices selected from $[0,\frac{1}{C^{1/\alpha}}]$, and they will be uniform on this interval, while all other vertices of $\mathbb{G}(n, U_{\alpha,C})$ will be isolated. If x and y are uniform on $[0, \frac{1}{C^{1/\alpha}}]$, then $C^{1/\alpha}x, C^{1/\alpha}y$ are uniform on $[0, 1]$. And for $x, y \in [0, \frac{1}{C^{1/\alpha}}]$, we have $U_{\alpha,C}(x, y) = U_\alpha(C^{1/\alpha}x, C^{1/\alpha}y)$ by definition. So $\mathbb{G}(n, U_{\alpha,C})$ has the same distribution as $\mathbb{G}(N, U_\alpha)$ together with $n - N$ isolated vertices (which will not contribute to the size of the largest clique). And by Lemma 3.2.1, we will have $N = (1 + o(1))\frac{1}{C^{1/\alpha}}n$ a.a.s. Thus by Theorem 3.1.7, the clique number of $\mathbb{G}(N, U_\alpha)$ is $\Theta(N^{\frac{\alpha}{\alpha+1}}) = \Theta(n^{\frac{\alpha}{\alpha+1}})$ a.a.s. Therefore, we have $\omega(\mathbb{G}(n, U_{\alpha,C})) = \Theta(n^{\frac{\alpha}{\alpha+1}})$ a.a.s.

Now, given that W is locally bounded below by $U_{\alpha,C}$ at (a, a), and that $\mathbb{G}(n, U_{\alpha,C})$ has clique number $\Theta(n^{\frac{\alpha}{\alpha+1}})$, we may use the same argument as in Lemma 3.3.5; namely, we apply Lemma 3.2.5, which gives

$$\omega(\mathbb{G}(n, W)) \geq (1 - o(1)) \cdot \omega(\mathbb{G}(n, U_{\alpha,C})) - O(\log n) = \Theta(n^{\frac{\alpha}{\alpha+1}})$$

a.a.s. Therefore $\omega(\mathbb{G}(n, W) = \Omega(n^{\frac{\alpha}{\alpha+1}})$ a.a.s. □

We end this section with a proof of Proposition 3.1.9, restated here.

Proposition 3.1.9. *For the graphon $W : [0, 1]^2 \to [0, 1]$ defined by*

$$W(x, y) = (1 - f(x))(1 - f(y)), \text{ where } f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases},$$

the clique number of $\mathbb{G}(n, W)$ is both $n^{1-o(1)}$ and $o(n)$ a.a.s.

Proof. We begin by proving the $o(n)$ bound. For any $\varepsilon > 0$, define the graphon

$$W_\varepsilon(x, y) = \begin{cases} 1 & \text{if } (x, y) \in [0, \varepsilon]^2 \\ 0 & \text{otherwise} \end{cases}.$$

Clearly, W is bounded above by W_ε on the set $[0, \varepsilon]^2$, and observe also that W is bounded away from 1 in some neighborhood of each point $(x, x) \neq (0, 0)$. Therefore we may apply Lemma 3.2.5 to conclude that a.a.s.,

$$\omega(\mathbb{G}(n, W)) \leq (1 + o(1)) \cdot \omega(\mathbb{G}(n, W_\varepsilon)) + O(\log n). \tag{3.15}$$

Notice that the clique number of $\mathbb{G}(n, W_\varepsilon)$ is simply the number of vertices sampled from the interval $[0, \varepsilon]$, as all edges among these vertices are present deterministically. And with probability $1 - o(1)$, the number of such vertices will be $(1 + o(1))\varepsilon n$ (by Lemma 3.2.1). Thus (3.15) becomes

$$\begin{aligned} \omega(\mathbb{G}(n, W)) &\leq (1 + o(1)) \cdot \varepsilon n + O(\log n) \\ &= (1 + o(1)) \cdot \varepsilon n. \end{aligned}$$

Since this holds a.a.s. for any choice of $\varepsilon > 0$, we see that $\omega(\mathbb{G}(n, W)) = o(n)$ a.a.s.

We now prove the lower bound of $n^{1-o(1)}$. Our proof will consist of two parts: first, for each $r \in \mathbb{N}$, we will show that W is bounded below by U_r locally in some neighborhood of $(0,0)$. We will then use Lemma 3.2.5 and the bound on $\omega(\mathbb{G}(n,U_r))$ given by Lemma 3.4.2 to give a lower bound on $\omega(\mathbb{G}(n,W))$.

We begin by looking at the (two-variable) Taylor polynomial of $W(x,y)$ about (0,0) of order r, for $r \in \mathbb{N}$. It is well known that $f(x)$, as defined above, is smooth on $\mathbb{R}$; this implies that W is smooth on $\mathbb{R}^2$ as well. Thus Taylor's theorem tells us that

$$W(x,y) = \sum_{0 \le i+j \le r} \left(\frac{\partial^{i+j} W}{\partial x^i \partial y^j}(0,0) \cdot \frac{x^i y^j}{i! j!} \right) + R_r(x,y), \tag{3.16}$$

where the remainder term $R_r(x,y)$ is bounded in absolute value by

$$|R_r(x,y)| \le C \cdot \max(x,y)^{r+1} \tag{3.17}$$

for some constant $C = C(W,r)$. (Note: we may obtain a more precise bound on the remainder as a function of (x,y), but the bound above will be sufficient here.) It is also well known that $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$; thus for all $i, j \ge 1$,

$$\frac{\partial^{i+j} W}{\partial x^i \partial y^j}(0,0) = \left(-f^{(i)}(0)\right) \cdot \left(-f^{(j)}(0)\right) = 0.$$

In fact, if either $i \ge 1$ or $j \ge 1$, this will hold. So the only nonzero term of the sum in (3.16) is

$$\frac{\partial^0 W}{\partial x^0 \partial y^0}(0,0) = W(0,0) = 1.$$

Therefore, (3.16) becomes

$$W(x,y) = 1 + R_r(x,y).$$

Now recall that

$$U_r(x,y) = (1-x^r)(1-y^r) = 1 - (x^r + y^r - x^r y^r).$$

For any $(x,y) \in [0,1]^2$, we have $x^r + y^r - x^r y^r \ge 0$. So in order to show that W is bounded below by U_r in some neighborhood of $(0,0)$, it will be sufficient to show that $|R_r(x,y)| \le x^r + y^r - x^r y^r$ for (x,y) in the same neighborhood. And observe that

$$x^r + y^r - x^r y^r = x^r + y^r(1-x^r) \ge x^r.$$

Similarly, $x^r + y^r - x^r y^r \ge y^r$; thus

$$x^r + y^r - x^r y^r \ge \max(x,y)^r. \tag{3.18}$$

We may combine this with the bound on $|R_r(x,y)|$ given by (3.17) after making one last observation: for any constant $C = C(r,W)$, if (x,y) is sufficiently close to $(0,0)$, then $C \cdot \max(x,y) \le 1$. Therefore, for (x,y) sufficiently close to $(0,0)$, combining

(3.17) and (3.18), we obtain

$$\begin{aligned} x^r + y^r - x^r y^r &\geq \max(x,y)^r \\ &\geq C \max(x,y) \cdot \max(x,y)^r \\ &\geq |R_r(x,y)|. \end{aligned}$$

Thus, as argued above,

$$W(x,y) \geq U_r(x,y)$$

for (x,y) in some neighborhood of $(0,0)$. Therefore, we may apply Lemma 3.2.5, and conclude that

$$\begin{aligned} \omega(\mathbb{G}(n,W)) &\geq (1-o(1)) \cdot \omega(\mathbb{G}(n,U_r)) - O(\log n) \\ &\geq (1-o(1)) \cdot \tfrac{1}{2} \cdot e^{-\frac{2}{1+r}} \cdot n^{\frac{r}{r+1}} \end{aligned}$$

a.a.s., where the last line is the lower bound on $\omega(\mathbb{G}(n,U_r))$ from Lemma 3.4.2. Then, since r can be chosen to be arbitrarily large, we obtain

$$\omega(\mathbb{G}(n,W)) = n^{1-o(1)}$$

a.a.s., as desired. □

3.5 Graphons equal to 1 at infinitely many points

In this section, we prove Proposition 3.1.11, and discuss other directions in which this work could be extended. We have described the clique number of a wide variety of W-random graphs where $W(a,a) = 1$ for a finite number of $a \in [0,1]$. We could also ask for some characterization of clique numbers of W-random graphs when $W(a,a) = 1$ at an infinite number of points, either countable or uncountable. For example, what is the clique number of $\mathbb{G}(n,W)$ for the following graphon W?

Example 3.5.1. Let $W(x,y) = \left(1 - x\sin^2\left(\frac{1}{x}\right)\right) \cdot \left(1 - y\sin^2\left(\frac{1}{y}\right)\right)$.

In this case, we have $W(a,a) = 1$ at a countably infinite number of points, namely for all a with $\frac{1}{a} = k \cdot \pi$ for $k \in \mathbb{N}$. If we define $W(0,0) = 1$, we may also show that W is locally Lipschitz at $(0,0)$, giving $\omega(\mathbb{G}(n,W)) = \Omega(\sqrt{n})$. The upper Dini derivatives of W at $(0,0)$ are 0, however, so we cannot use Lemma 3.3.6 to give an upper bound. It could be interesting to find the correct order of growth of the clique number for this and other examples with a countably infinite number of points with $W(a,a) = 1$.

Proposition 3.1.11 (restated here) gives a rough estimate of the order of growth of $\omega(\mathbb{G}(n,W))$ for a graphon W with $W(a,a) = 1$ on an interval; the following graphon is equal to 1 along the line $x = y$ and drops off away from that line.

Proposition 3.1.11. *Let $W(x,y) = 1 - |x-y|$. The clique number of $\mathbb{G}(n,W)$ is $n^{1/2+o(1)}$ a.a.s.*

Before proving this proposition, let us note one difficulty in analyzing this and other graphons that are equal to 1 on a positive-measure portion of the line $x = y$. Namely, to obtain an upper bound on the clique number of such a graphon W, we will not easily be able to use the first moment method as with W_r and U_r in Sections 3.3 and 3.4. In order to do so, we would need to compute

$$\mathbb{E}[X_k] = \binom{n}{k} \int_{[0,1]^k} \prod_{\ell \neq m \in [k]} W(x_\ell, x_m)\, d\vec{x},$$

where X_k is the number of cliques in $\mathbb{G}(n, W)$ of size k. For W_r and U_r, we were able to simplify this integral by using the fact that $W_r(x, y)$ and $U_r(x, y)$ are of the form $f(x)f(y)$ for some function f. Graphons of this form are called "rank-1", and we can think of the W-random graphs that they produce as a more limited generalization of Erdős-Rényi random graphs than those produced by graphons generally; in a rank-1 graphon, edge probabilities are not fixed as in the Erdős-Rényi model, but the likelihood of each pair of vertices to be connected by an edge is determined only by how well-connected these vertices are overall, and not on any more complicated relationship between vertex weights.

The graphon W in the proposition above is not rank-1, so we cannot simplify the first moment calculation above by the same method we used for W_r and U_r. More generally, any rank-1 graphon that is equal to 1 on some positive-measure portion of the line $x = y$ is in some sense trivial; if we have a graphon W with $W(x_\ell, x_m) = f(x_\ell)f(x_m)$ and $W(a, a) = 1$ for all a in some positive-measure $A \subseteq [0, 1]$, then $f(a) = 1$ for $a \in A$. This would imply that W evaluates to 1 on the positive-measure set $A \times A$, and thus $\mathbb{G}(n, W)$ has a linear-size clique number.

Here, to obtain the rough order of growth of $\mathbb{G}(n, W)$ for W in Proposition 3.1.11, we will use a more direct approach; we expect that any set of vertices forming a large clique in $\mathbb{G}(n, W)$ would be sampled from a relatively small interval, as two vertices x_i and x_j are only likely to be connected in $\mathbb{G}(n, W)$ if $|x_i - x_j|$ is small. However, Lemma 3.2.2 tells us that a.a.s. there will be no very large set of vertices sampled from a very small interval. We then take a union bound over all sufficiently large sets of vertices (which must each be spread over a not-too-small interval) to show that a.a.s. we will not obtain a "large" clique. Following are the details of that argument.

Proof of Proposition 3.1.11. First, observe that W is locally Lipschitz at, for example, the point $(0, 0)$; all directional derivatives exist there and are bounded between -1 and 0. So by Lemma 3.1.6, $\mathbb{G}(n, W) = \Omega(\sqrt{n})$ a.a.s. Now we compute an upper bound on the clique number, using the method outlined in the previous paragraph.

Consider any set S of $k = 3\delta n$ vertices in $\mathbb{G}(n, W)$, with $\delta = \omega\left(\frac{1}{\sqrt{n}}\right)$ to be chosen later; we wish to show that no such set will form a clique. Partition S into S_1, S_2, and S_3, namely the first δn vertices, the middle, and the last, respectively, as they are ordered on the unit interval. By Lemma 3.2.2, with probability $1 - o(1)$, the vertices in each set, and in particular in S_2, occupy an interval of length at least $\frac{\delta}{2}(1 - o(1))$. Therefore each vertex in S_1 is at distance at least $\frac{\delta}{2}(1 - o(1))$ from each vertex in S_3, and hence by the definition of W, the probability that every such pair of vertices is connected is at most

$$\left(1-\frac{\delta}{2}(1-o(1))\right)^{(\delta n)^2},$$

which gives an upper bound on the probability that S is a clique. Taking a union bound over all sets of $3\delta n$ vertices in $\mathbb{G}(n,W)$, the probability that there exists a clique of size $k=3\delta n$ in $\mathbb{G}(n,W)$ is at most

$$\begin{aligned}\binom{n}{3\delta n}\left(1-\frac{\delta}{2}(1-o(1))\right)^{(\delta n)^2} &\le \left(\frac{en}{3\delta n}\right)^{3\delta n}\left(1-\frac{\delta}{2}(1-o(1))\right)^{(\delta n)^2}\\ &\le e^{3\delta n\cdot\log\frac{e}{3\delta}}\cdot e^{-(\delta n)^2\frac{\delta}{2}(1-o(1))}\\ &= e^{\delta n\left(3\log\frac{e}{3\delta}-\frac{\delta^2 n}{2}(1-o(1))\right)}.\end{aligned}$$

This will be $o(1)$ if $3\log\frac{e}{3\delta}\le\frac{\delta^2 n}{2}(1-\Omega(1))$, which is satisfied, for example, for $\delta=\frac{1}{\sqrt{n}}\log n=n^{-1/2+o(1)}$ (but not, say, for $\delta=\frac{1}{\sqrt{n}}(\log n)^{1/4}$). So with probability $1-o(1)$, the clique number of $\mathbb{G}(n,W)$ is at most $3\delta n=n^{-1/2+o(1)}\cdot n=n^{1/2+o(1)}$. $\square$

Chapter 4

Resilience of rank in random matrices

4.1 Introduction

This chapter is based on the paper [FLM19], joint with Asaf Ferber and Kyle Luh.

Let $M_{n,m}$ denote an $n \times m$ matrix with independent entries chosen uniformly from $\{\pm 1\}$. For any matrix with entries in $\{\pm 1\}$, we define its "resilience" as follows.

Definition 4.1.1. *Given an $n \times m$ matrix M with entries in $\{\pm 1\}$, and with $m \geq n$, we denote by* $\mathrm{Res}(M)$ *the minimum number of sign flips necessary in order to make M of rank less than n.*

Note that this quantity is always at most $m/2$, since for any two rows $\boldsymbol{a}, \boldsymbol{b} \in \{\pm 1\}^m$, one can achieve either $\boldsymbol{a} = \boldsymbol{b}$ or $\boldsymbol{a} = -\boldsymbol{b}$ by changing at most $m/2$ entries. Here we show that when m is slightly larger than n, this upper bound is essentially tight.

Theorem 4.1.2. *For every $\varepsilon > 0$ and $m \geq n + n^{1-\varepsilon/6}$, a.a.s. we have*

$$Res(M_{n,m}) \geq (1-\varepsilon)m/2.$$

Our proof strategy roughly goes as follows: Consider an outcome M of $M_{n,m}$. Note that if the rank of M is less than n, then in particular, writing $m' = m - n^{1-\varepsilon/6}$, there exists an $n \times m'$ submatrix M' of M with rank less than n. Moreover, as M' is not of full rank, there exists $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ which lies in the *left kernel* of M' (that is, with $\boldsymbol{a}^T M' = \boldsymbol{0}$). Our main goal is to show that for each such $\boldsymbol{a}$ (if it exists), and for a randomly chosen $\boldsymbol{x} \in \{\pm 1\}^n$, the probability

$$\rho(\boldsymbol{a}) := \Pr[\boldsymbol{a}^T \boldsymbol{x} = 0] \tag{4.1}$$

is typically very small.

Next, observe that a vector $\boldsymbol{a}$ will be in the left kernel of M if and only if it is in the left kernel of M' and is also orthogonal to the remaining $n^{1-\varepsilon/6}$ columns of M. With this in mind, using the bound on $\rho(\boldsymbol{a})$ and the extra $n^{1-\varepsilon/6}$ columns of M, we want to "boost" the probability and show that

$$\Pr[\exists \boldsymbol{a} \text{ such that } \boldsymbol{a}^T M = \boldsymbol{0}] = n^{-(1/2-o(1))m}. \tag{4.2}$$

Then, notice that there are at most $\sum_{s \le (1/2-o(1))m} \binom{nm}{s} \approx n^{(1/2+o(1))m}$ many matrices that can be obtained from M by changing $s \le (1/2 - o(1))m$ entries, and there are at most $2^m = n^{o(1)\cdot m}$ many choices for M'. Therefore, using the bound (4.2), we can complete the proof by a simple union bound (after, of course, showing that the $o(1)$ terms in (4.2) work in our favor).

The main challenge is to prove (4.2), as it involves a union bound over all possible kernel vectors $\boldsymbol{a} \in \mathbb{R}^n$. In order to overcome this difficulty, we use some recently developed machinery introduced in [FJLS19]. Roughly speaking, we embed the problem into a sufficiently large finite field $\mathbb{F}_p$. Then, as there are finitely many options for $\boldsymbol{a} \in \mathbb{F}_p$ in the left kernel of M, we can use a counting argument from [FJLS19] to bound the probability of encountering each possible kernel vector $\boldsymbol{a}$ according to the corresponding value of $\rho(\boldsymbol{a})$.

We mention that the approach of bounding $\rho(\boldsymbol{a})$ for possible null-vectors in the context of singularity is not new (see for example [KKS95, Ngu13, RV08, TV09, Ver14]). The novelty of our argument is that we utilize the methods from [FJLS19] to obtain the bound (4.2). Most of the previously used arguments yield exponential or polynomial probabilities which would only tolerate a sublinear number of modifications to the matrix. Although it is possible to modify these arguments to generate super-exponential bounds, the exact constant of 1/2 in (4.2) seems to be difficult to achieve via other arguments.

Lastly, we mention that the method in [FJLS19] has already been successfully applied to a variety of combinatorial problems in random matrix theory [CMMM19, FJ18, Jai19a, Jai19b, LMN19].

The remainder of this chapter is organized as follows. In Section 4.2, we provide the necessary background to state the counting lemma from [FJLS19]. In Section 4.2.3, we provide a convenient interface to apply this counting lemma. This is drawn from [FJLS19] as well. Finally, in Section 4.3, we provide the proof of Theorem 4.1.2.

4.2 Auxiliary results

Here we review some auxiliary results and introduce convenient notation to be used in the proof of our main result.

4.2.1 Halász inequality in $\mathbb{F}_p$

Let $\boldsymbol{a} := (a_1, \ldots, a_n) \in (\mathbb{Z} \setminus \{0\})^n$ and let $\epsilon_1, \ldots, \epsilon_n$ be independent and identically distributed (i.i.d.) Rademacher random variables; that is, each ϵ_i independently takes values ± 1 with probability $1/2$ each. We define the largest atom probability $\rho(\boldsymbol{a})$ by

$$\rho(\boldsymbol{a}) := \sup_{x\in\mathbb{Z}} \Pr\left(\epsilon_1 a_1 + \cdots + \epsilon_n a_n = x\right).$$

Similarly, if we are working over some finite field $\mathbb{F}_p$, let

$$\rho_{\mathbb{F}_p}(\boldsymbol{a}) := \sup_{x\in\mathbb{F}_p} \Pr\left(\epsilon_1 a_1 + \cdots + \epsilon_n a_n = x\right),$$

where, of course, the arithmetic is done over $\mathbb{F}_p$.

Now, let $R_k(\boldsymbol{a})$ denote the number of solutions to $\pm a_{i_1} \pm a_{i_2} \cdots \pm a_{i_{2k}} \equiv 0$, where repetitions are allowed in the choice of $i_1, \ldots, i_{2k} \in [n]$. A classical theorem of Halász [Hal77] gives an estimate on the atom probability based on $R_k(\boldsymbol{a})$. Here we need the following, slightly different version of this theorem, which can be applied to the finite field setting.

Theorem 4.2.1 (Halász's inequality over $\mathbb{F}_p$; Theorem 1.4 in [FJLS19])**.** *There exists an absolute constant C such that the following holds for every odd prime p, integer n, and vector $\boldsymbol{a} := (a_1, \ldots, a_n) \in \mathbb{F}_p^n \setminus \{\boldsymbol{0}\}$. Suppose that an integer $k \geqslant 0$ and positive real M satisfy $30M \leq |\operatorname{supp}(\boldsymbol{a})|$ and $80kM \leq n$. Then,*

$$\rho_{\mathbb{F}_p}(\boldsymbol{a}) \leq \frac{1}{p} + \frac{C R_k(\boldsymbol{a})}{2^{2k} n^{2k} \cdot M^{1/2}} + e^{-M}.$$

The proof of this theorem, which is essentially the same as the original one by Halász, can be found in [FJLS19].

4.2.2 Counting Lemma

In this section we state a counting lemma from [FJLS19] which plays a key role in our proof. First, we need the following definition:

Definition 4.2.2. *Suppose that $\boldsymbol{a} \in \mathbb{F}_p^n$ for an integer n and a prime p and let $k \in \mathbb{N}$. For every $\alpha \in [0,1]$, we define $R_k^\alpha(\boldsymbol{a})$ to be the number of solutions to*

$$\pm a_{i_1} \pm a_{i_2} \cdots \pm a_{i_{2k}} = 0 \mod p$$

that satisfy $|\{i_1, \ldots, i_{2k}\}| \geqslant (1+\alpha)k$.

It is easily seen that $R_k(\boldsymbol{a})$ cannot be much larger than $R_k^\alpha(\boldsymbol{a})$. This is formalized in the following simple lemma, which is proved in [FJLS19] (a proof is also given here, for the convenience of the reader).

Lemma 4.2.3. *For all $k, n \in \mathbb{N}$ with $k \leqslant n/2$, and any prime p, vector $\boldsymbol{a} \in \mathbb{F}_p^n$, and $\alpha \in [0,1]$,*

$$R_k(\boldsymbol{a}) \leq R_k^\alpha(\boldsymbol{a}) + \left(40 k^{1-\alpha} n^{1+\alpha}\right)^k.$$

Proof. By definition, $R_k(\boldsymbol{a})$ is equal to $R_k^\alpha(\boldsymbol{a})$ plus the number of solutions to $\pm a_{i_1} \pm a_{i_2} \cdots \pm a_{i_{2k}} = 0$ that satisfy $|\{i_1, \ldots, i_{2k}\}| < (1+\alpha)k$. The latter quantity is bounded from above by the number of sequences $(i_1, \ldots, i_{2k}) \in [n]^{2k}$ with at most $(1+\alpha)k$ distinct entries times 2^{2k}, the number of choices for the $\pm$ signs. Thus

$$R_k(\boldsymbol{a}) \leq R_k^\alpha(\boldsymbol{a}) + \binom{n}{(1+\alpha)k}\big((1+\alpha)k\big)^{2k} 2^{2k} \leq R_k^\alpha(\boldsymbol{a}) + \left(4e^{1+\alpha}(1+\alpha)^{1-\alpha} k^{1-\alpha} n^{1+\alpha}\right)^k,$$

where the final inequality follows from the well-known bound $\binom{a}{b} \leqslant (ea/b)^b$. Finally, noting that $4e^{1+\alpha}(1+\alpha)^{1-\alpha} \leq 40$ for all $\alpha \in [0,1]$ completes the proof. $\square$

Given a vector $\boldsymbol{a} \in \mathbb{F}_p^n$ and a subset of coordinates $I \subseteq [n]$, we define $\boldsymbol{a}_I$ to be its restriction to the coordinates in I; that is, $\boldsymbol{a}_I = (a_i)_{i \in I} \in \mathbb{F}_p^I$. We write $\boldsymbol{b} \subseteq \boldsymbol{a}$ if there exists an $I \subseteq [n]$ for which $\boldsymbol{b} = \boldsymbol{a}_I$. For $\boldsymbol{b} \subseteq \boldsymbol{a}$ we let $|\boldsymbol{b}|$ be the size of the subset I determining $\boldsymbol{b}$.

Now we are ready to state the counting lemma, which will allow us to give an upper bound on the number of "bad" vectors defined in the next secion.

Theorem 4.2.4 (Theorem 1.7 in [FJLS19])**.** *Let p be a prime, let $k, n \in \mathbb{N}$, $s \in [n]$, $t \in [p]$, and let $\alpha \in (0,1)$. Denoting*

$$\boldsymbol{B}_{k,s,\geq t}^\alpha(n) := \left\{\boldsymbol{a} \in \mathbb{F}_p^n : R_k^\alpha(\boldsymbol{b}) \geq t \cdot \frac{2^{2k} \cdot |\boldsymbol{b}|^{2k}}{p} \text{ for every } \boldsymbol{b} \subseteq \boldsymbol{a} \text{ with } |\boldsymbol{b}| \geq s\right\},$$

we have

$$|\,\boldsymbol{B}_{k,s,\geq t}^\alpha(n)| \leq \left(\frac{s}{n}\right)^{2k-1} (\alpha t)^{s-n} p^n.$$

4.2.3 "Good" and "bad" vectors

The purpose of this section is to formulate easy-to-use versions of Halász's inequality (Theorem 4.2.1) and our counting theorem (Theorem 4.2.4). This follows [FJLS19] closely, but requires a more delicate choice of parameters as we need to achieve the bound in (4.2) (and crucially, the constant $1/2$ in the exponent). We shall partition $\mathbb{F}_p^n$ into "good" and "bad" vectors. We shall then show that, on the one hand, every "good" vector $\boldsymbol{a}$ has a small $\rho(\boldsymbol{a})$ and that, on the other hand, there are relatively few "bad" vectors.[1] The formal statements now follow. In order to simplify the notation, we suppress the implicit dependence of the defined notions on n, k, p, and α.

Definition 4.2.5. *Let p be a prime, let $n, k \in \mathbb{N}$, and let $\alpha \in (0,1)$. For any $t > 0$, define the set $\boldsymbol{H}_t$ of t-*good *vectors by*

$$\boldsymbol{H}_t := \left\{\boldsymbol{a} \in \mathbb{F}_p^n : \exists \boldsymbol{b} \subseteq \boldsymbol{a} \text{ with } |\operatorname{supp}(\boldsymbol{b})| \geq n^{1-\varepsilon/2} \text{ and } R_k^\alpha(\boldsymbol{b}) \leq t \cdot \frac{2^{2k} \cdot |\boldsymbol{b}|^{2k}}{p}\right\}.$$

[1] In fact, we shall only show that there are relatively few "bad" vectors that have some number of nonzero coordinates. The number of remaining vectors (ones with very small support) is so small that even a very crude estimate will suffice for our needs.

The goodness *of a vector* $\boldsymbol{a} \in \mathbb{F}_p^n$*, denoted by* $h(\boldsymbol{a})$*, will be the smallest* t *such that* $\boldsymbol{a} \in \boldsymbol{H}_t$*. In other words*

$$h(\boldsymbol{a}) = \min \left\{ \frac{p \cdot R_k^\alpha(\boldsymbol{b})}{2^{2k} \cdot |\boldsymbol{b}|^{2k}} : \boldsymbol{b} \subseteq \boldsymbol{a} \text{ and } |\operatorname{supp}(\boldsymbol{b})| \geqslant n^{1-\varepsilon/2} \right\}.$$

Note that if a vector $\boldsymbol{a} \in \mathbb{F}_p^n$ has fewer than $n^{1-\varepsilon/2}$ nonzero coordinates, then it cannot be t-good for any t and thus $h(\boldsymbol{a}) = \infty$. On the other hand, trivially $R_k^\alpha(\boldsymbol{b}) \leqslant 2^{2k} \cdot |\boldsymbol{b}|^{2k}$ for every vector $\boldsymbol{b}$, as there are $2^{2k}|\boldsymbol{b}|^{2k}$ total possible choices of a sequence $\pm b_{i_1} \pm b_{i_2} \pm \cdots \pm b_{i_{2k}}$. Thus every $\boldsymbol{a} \in \mathbb{F}_p^n$ with at least $n^{1-\varepsilon/2}$ nonzero coordinates must be p-good, that is, $h(\boldsymbol{a}) \leqslant p$ for each such $\boldsymbol{a}$.

Having formalized the notion of a "good" vector, we are now ready to state and prove two corollaries of Theorems 4.2.1 and 4.2.4 that lie at the heart of our approach. (Note: the particular choice of parameters in Lemma 4.2.6 is made for convenience in a later application.)

Lemma 4.2.6. *Let* $\boldsymbol{a} \in \boldsymbol{H}_t$*, let* $\alpha \in (0,1)$*, and let* $\varepsilon < 1/100$*. Suppose that* $p = \Theta(2^{n^{\varepsilon/3}})$ *is a prime,* $t \geq n$*, and* $k = \Theta(n^{\varepsilon/3})$*. Then for sufficiently large* n *we have*

$$\rho_{\mathbb{F}_p}(\boldsymbol{a}) \leqslant \frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}},$$

where $C = C(\alpha, \varepsilon)$ *is a constant depending only on* α *and* ε*.*

Proof. As $\boldsymbol{a} \in \boldsymbol{H}_t$, we can find a subvector $\boldsymbol{b}$ of $\boldsymbol{a}$ such that $|\operatorname{supp}(\boldsymbol{b})| \geqslant n^{1-\varepsilon/2}$ and $R_k^\alpha(\boldsymbol{b}) \leqslant t \cdot 2^{2k} \cdot |\boldsymbol{b}|^{2k}/p$. Set $M = \lfloor n^{1-\varepsilon/2}/(80k) \rfloor = \Theta(n^{1-5\varepsilon/6})$ so that

$$\max\{30M, 80Mk\} = 80Mk \leqslant n^{1-\varepsilon/2} \leqslant |\operatorname{supp}(\boldsymbol{b})| \leqslant |\boldsymbol{b}|.$$

Thus we may apply Theorem 4.2.1 to obtain, for some absolute constant C_0,

$$\rho_{\mathbb{F}_p}(\boldsymbol{b}) \leqslant \frac{1}{p} + \frac{C_0 R_k(\boldsymbol{b})}{2^{2k} \cdot |\boldsymbol{b}|^{2k} \cdot M^{1/2}} + e^{-M}.$$

Now, using Lemma 4.2.3 we can upper bound the right hand side by

$$\begin{aligned}
\rho_{\mathbb{F}_p}(\boldsymbol{b}) &\leq \frac{1}{p} + \frac{C_0 R_k^\alpha(\boldsymbol{b}) + C_0 \left(40k^{1-\alpha}|\boldsymbol{b}|^{1+\alpha}\right)^k}{2^{2k} \cdot |\boldsymbol{b}|^{2k} \cdot M^{1/2}} + e^{-M} \\
&\leq \frac{1}{p} + \frac{C_0 t \cdot 2^{2k} \cdot |\boldsymbol{b}|^{2k}/p + C_0 \left(40k^{1-\alpha}|\boldsymbol{b}|^{1+\alpha}\right)^k}{2^{2k} \cdot |\boldsymbol{b}|^{2k} \cdot M^{1/2}} + e^{-M} \\
&= \frac{1}{p}\left(1 + \frac{C_0 t}{M^{1/2}} + C_0 \left(10(k/|\boldsymbol{b}|)^{1-\alpha}\right)^k \cdot \frac{p}{M^{1/2}}\right) + e^{-M}.
\end{aligned}$$

Now we wish to show that, with the parameter assignments above, the dominant term in this sum is $\frac{C_0 t}{pM^{1/2}}$. To this end, we bound each of the other terms as follows. First,

$$e^{-M} = e^{-\Theta(n^{1-5\varepsilon/6})} = o(2^{-n^{\varepsilon/3}}) = o\left(\frac{1}{p}\right).$$

(Here we use the upper bound assumption on ε.) Second,

$$\begin{aligned} C_0\left(10(k/|\boldsymbol{b}|)^{1-\alpha}\right)^k \cdot \frac{p}{M^{1/2}} &\leq C_0\left(10\left(n^{\varepsilon/3-(1-\varepsilon/2)}\right)^{1-\alpha}\right)^k \cdot p \\ &= \left(n^{-\Theta(1)}\right)^{\Theta(n^{\varepsilon/3})} \cdot p \\ &= 2^{-\Theta(n^{\varepsilon/3}\log n)} \cdot \Theta(2^{n^{\varepsilon/3}}) \\ &= o(1). \end{aligned}$$

And last, we observe that, as $t \geq n$,

$$\frac{C_0 t}{M^{1/2}} \geq \frac{n}{\Theta(n^{\frac{1}{2}(1-5\varepsilon/6)})} = \omega(1).$$

Therefore the dominant term in the sum above is indeed $\frac{C_0 t}{pM^{1/2}}$; then, choosing the constant $C = C(\alpha, \varepsilon) > C_0$ sufficiently large, we obtain

$$\rho_{\mathbb{F}_p}(\boldsymbol{b}) \leq \frac{Ct}{pM^{1/2}} \leq \frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}}$$

as desired. (Note: in the last step, we have incorporated the implicit constant in $M = \Theta(n^{1-5\varepsilon/6})$ into the constant C.) □

Lemma 4.2.7. *For every integer n and real $t \geqslant n$,*

$$\left|\left\{\boldsymbol{a} \in \mathbb{F}_p^n : |\operatorname{supp}(\boldsymbol{a})| \geqslant n^{1-\varepsilon/2} \text{ and } \boldsymbol{a} \notin \boldsymbol{H}_t\right\}\right| \leqslant 2^n \left(\frac{p}{\alpha t}\right)^n \cdot t^{n^{1-\varepsilon/2}}.$$

Proof. We may assume that $t \leqslant p$, as otherwise the left-hand side above is zero; see the comment below Definition 4.2.5. Let us now fix an $S \subseteq [n]$ with $|S| \geqslant n^{1-\varepsilon/2}$ and count only vectors $\boldsymbol{a}$ with $\operatorname{supp}(\boldsymbol{a}) = S$. Since $\boldsymbol{a} \notin \boldsymbol{H}_t$, the restriction $\boldsymbol{a}_S$ of $\boldsymbol{a}$ to the set S must be contained in the set $\boldsymbol{B}^{\alpha}_{k,n^{1-\varepsilon/2},\geq t}(|S|)$. Hence, Theorem 4.2.4 implies that the number of choices for $\boldsymbol{a}_S$ is at most

$$\left(\frac{n^{1-\varepsilon/2}}{|S|}\right)^{2k-1} (\alpha t)^{n^{1-\varepsilon/2}-|S|} p^{|S|} \leq \left(\frac{p}{\alpha t}\right)^n t^{n^{1-\varepsilon/2}},$$

where the second inequality follows as $n^{1-\varepsilon/2} \leq |S| \leq n$ and $\alpha t \leqslant t \leqslant p$. Since $\boldsymbol{a}_S$ completely determines $\boldsymbol{a}$, we obtain the desired conclusion by summing the above bound over all sets S. □

4.3 Proof of Theorem 4.1.2

In this section we gradually construct the entire proof of Theorem 4.1.2.

For convenience, we introduce some notation to indicate the distance of two Rademacher matrices.

Definition 4.3.1. *For two matrices $n \times m$ matrices M, M' we let $d(M, M')$ denote the number of entries where M and M' differ.*

With this definition in hand, Theorem 4.1.2 can be stated as follows:

Theorem 4.3.2. *For every $\varepsilon > 0$ and $m \geq n + n^{1-\varepsilon/6}$, a.a.s. we have $rank(M') = n$ for all $n \times m$, ± 1 matrices M' with $d(M_{n,m}, M') \leq (1-\varepsilon)m/2$.*

First, we will prove Theorem 4.1.2 under the assumption that $m = \omega(n)$.

4.3.1 Proof of Theorem 4.1.2 under the assumption $m = \omega(n)$

Let $\varepsilon > 0$ be any fixed constant, and let $m \geq C(\varepsilon)n$, where $C(\varepsilon)$ is a sufficiently large constant. We wish to show that a.a.s., $M = M_{n,m}$ is such that every $n \times m$ matrix M' with $d(M, M') \leq (1-\varepsilon)m/2$ has rank n.

In order to do so, let us take (say) $p = 3$ and work over $\mathbb{F}_3$. Observe that if the above statement holds over $\mathbb{F}_3$ then it trivially holds over $\mathbb{Z}$.

Let $\boldsymbol{a} \in \mathbb{F}_3^n \setminus \{\mathbf{0}\}$, and note that for a randomly chosen $\boldsymbol{x} \in \{\pm 1\}^n$ we have

$$\Pr[\boldsymbol{a}^T \boldsymbol{x} = 0\,] \leq \frac{1}{2}.$$

Therefore, as the columns of M are independent, it follows that the random variable $X_{\boldsymbol{a}} =$ "the number of zeroes in $\boldsymbol{a}^T M$" is stochastically dominated by $\mathrm{Bin}(m, \frac{1}{2})$. Hence, by Chernoff's bound, we obtain that

$$\Pr\left[X_{\boldsymbol{a}} \geq (1+\varepsilon)m/2\right] \leq e^{-C_1 m}$$

for some C_1 that depends on ε. By applying the union bound over all $\boldsymbol{a} \in \mathbb{F}_3^n \setminus \{\mathbf{0}\}$ we obtain that

$$\Pr\left[\exists \boldsymbol{a} \in \mathbb{F}_3^n \setminus \{\mathbf{0}\} \text{ with } X_{\boldsymbol{a}} \geq (1+\varepsilon)m/2\right] \leq 3^n e^{-C_1 m} = o(1),$$

where the last inequality follows from the fact that $m \geq C(\varepsilon)n$ and $C(\varepsilon)$ is sufficiently large.

Thus M is typically such that in *every* non-zero linear combination of its rows, there are less than $(1+\varepsilon)m/2$ many zeroes. In particular, since by changing at most $(1-\varepsilon)m/2$ many entries one can affect at most $(1-\varepsilon)m/2$ columns, it follows that for all M' with $d(M, M') \leq (1-\varepsilon)m/2$, no non-trivial combination of the rows of M' is the 0 vector. In particular, every such M' is of rank n. This completes the proof for this case. $\square$

4.3.2 Proof of Theorem 4.1.2 under the assumption $m = O(n)$

In what follows we always assume that $m = O(n)$. Therefore, whenever convenient, in appropriate asymptotic formulas we may switch between m and n without further explanation. This case is more involved than the case $m = \omega(n)$ and it will be further divided into a few subcases. From now on, we fix p to be some prime $p = \Theta(2^{n^{\varepsilon/3}})$, and concretely, we write $m \leq C(\varepsilon)n$ for some constant $C(\varepsilon)$.

Now, write $m' = m - n^{1-\varepsilon/6}$ (the width of the matrix under consideration minus the $n^{1-\varepsilon/6}$ "extra" columns). In the following two subsections, we will show that with high probability, for every $\boldsymbol{a} \in \mathbb{F}_p^n$, if $\boldsymbol{a}^T M' = \mathbf{0}$ for some $n \times m'$ matrix M' with $d(M', M_{n,m'}) \leq (1-\varepsilon)m/2$, then $\boldsymbol{a}$ has "many" nonzero entries, and is "pseudorandom" in some sense (Lemmas 4.3.3 and 4.3.4). From here, we can apply the Halász inequality (in the form of Lemma 4.2.6) almost directly, using the fact that there are $m - m' \geq n^{1-\varepsilon/6}$ extra columns, to conclude that for any such $\boldsymbol{a}$, the probability that $\boldsymbol{a}^T M_{n,m} = \mathbf{0}$ is small.

Eliminating Small Linear Dependencies

First, we wish to show that if $\boldsymbol{a}^T M' = \mathbf{0}$ (over $\mathbb{F}_p$) for some M' with $d(M_{n,m'}, M') \leq (1-\varepsilon)m/2$, then $\boldsymbol{a}$ has "many" non-zero entries (assuming $\boldsymbol{a} \neq \mathbf{0}$ of course).

Lemma 4.3.3. *Let $\varepsilon > 0$, let $p = \Theta(2^{n^{\varepsilon/3}})$ be a prime, and let $n + n^{1-\varepsilon/6} \leq m \leq C(\varepsilon)n$. Write $m' = m - n^{1-\varepsilon/6}$. Then, working in $\mathbb{F}_p$, the probability there exists a matrix M' with $d(M', M_{n,m'}) \leq (1-\varepsilon)m/2$ and a nonzero vector $\boldsymbol{a} \in \mathbb{F}_p^n$ with $|supp(\boldsymbol{a})| \leq n^{1-\varepsilon/2}$ and with $\boldsymbol{a}^T M' = \mathbf{0}$ is at most $2^{-\Theta(n)}$.*

Proof. Given a vector $\boldsymbol{a} \in \mathbb{F}_p^n$, we let $\ell := |\mathrm{supp}(\boldsymbol{a})|$. Note that for any $\boldsymbol{a} \neq \mathbf{0}$ and a uniformly chosen vector $\boldsymbol{x} \in \{\pm 1\}^n$ we trivially have

$$\Pr[\boldsymbol{a}^T \boldsymbol{x} = 0] \leq \frac{1}{2}.$$

Moreover, as we are only allowed to change at most $(1-\varepsilon)m/2$ coordinates of $M_{n,m'}$, it follows that at most $(1-\varepsilon)m/2$ entries of $\boldsymbol{a}^T M_{n,m'}$ can be altered. In particular, if there exists a vector $\boldsymbol{a}$ for which $\boldsymbol{a}^T M' = \mathbf{0}$, where $d(M_{n,m'}, M') \leq (1-\varepsilon)m/2$, then this implies that $\boldsymbol{a}^T M_{n,m'}$ already contained at least $m' - (1-\varepsilon)m/2 = (1+\varepsilon - o(1))m/2$ zero entries.

Now, since the random variable counting the number of 0 entries is stochastically dominated by $\mathrm{Bin}(n, \frac{1}{2})$, by Chernoff's bound we obtain that for a given $\boldsymbol{a} \neq \mathbf{0}$, the probability to have at least $(1+\varepsilon-o(1))m/2$ zeroes in $\boldsymbol{a}^T M_{n,m'}$ is at most $2^{-c(\varepsilon)m}$, where $c(\varepsilon)$ is some constant depending only on ε. Thus the probability that for a given nonzero vector $\boldsymbol{a}$, there exists some M' with $d(M', M_{n,m'}) \leq (1-\varepsilon)m/2$ and $\boldsymbol{a}^T M' = \mathbf{0}$ is at most $2^{-c(\varepsilon)m}$.

All in all, by applying the union bound over all $\boldsymbol{a} \neq \mathbf{0}$ with $\ell \leq n^{1-\varepsilon/2}$ nonzero entries, the probability that we are seeking to bound is at most

$$\sum_{\ell=1}^{n^{1-\varepsilon/2}} \binom{n}{\ell} p^\ell 2^{-c(\varepsilon)m} \leq \sum_{\ell=1}^{n^{1-\varepsilon/2}} 2^{\ell \log n + \ell n^{\varepsilon/3} - c(\varepsilon)m} = 2^{-\Theta(n)}$$

where the last equality holds due to the assumption $\ell \leq n^{1-\varepsilon/2}$. $\square$

Eliminating "bad" vectors

We now show that, almost surely, any vector $\boldsymbol{a}$ with many non-zero entries and with $\boldsymbol{a}^T M' = \mathbf{0}$ for some M' with $d(M', M_{n,m'}) \leq (1-\varepsilon)m/2$ will be "good" or "unstructured".

Lemma 4.3.4. *Let $1/100 > \varepsilon > 0$, let $p = \Theta(2^{n^{\varepsilon/3}})$ be a prime, and let $n + n^{1-\varepsilon/6} \leq m \leq C(\varepsilon)n$. Write $m' = m - n^{1-\varepsilon/6}$. Then, working in $\mathbb{F}_p$, the probability that there exists a matrix M' with $d(M', M_{n,m'}) \leq (1-\varepsilon)m/2$ and a vector $\boldsymbol{a} \in \mathbb{F}_p^n \setminus \boldsymbol{H}_n$ with at least $n^{1-\varepsilon/2}$ non-zero entries such that $\boldsymbol{a}^T M' = \mathbf{0}$ is at most $2^{-\Theta(n \log n)}$.*

Proof. Our first step is to take a union bound over choices of $\boldsymbol{a}$; we wish to bound the quantity

$$\sum_{\substack{\boldsymbol{a} \in \mathbb{F}_p^n \setminus \boldsymbol{H}_n \\ |\operatorname{supp}(\boldsymbol{a})| \geq n^{1-\varepsilon/2}}} \Pr[\exists M' \text{ with } d(M', M_{n,m'}) \leq (1-\varepsilon)m/2 \text{ and } \boldsymbol{a}^T M' = \mathbf{0}]. \tag{4.3}$$

Now we use the sets $\boldsymbol{H}_t$ to divide the vectors $\boldsymbol{a}$ into different classes. As observed after Definition 4.2.5, every $\boldsymbol{a} \in \mathbb{F}_p^n$ with at least $n^{1-\varepsilon/2}$ nonzero entries is in $\boldsymbol{H}_t$ for some $t \leq p$. Moreover, notice that $\boldsymbol{H}_t \subseteq \boldsymbol{H}_{t+1}$ for any $t > 0$. So we can write $\mathbb{F}_p^n \setminus \boldsymbol{H}_n$ as a union $\bigcup_{n+1 \leq t \leq p} \boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1}$. Therefore, taking a union bound over integers $t > n$, the probability Eq. (4.3) that we are trying to bound is at most

$$\sum_{t=n+1}^{p} \left(\sum_{\boldsymbol{a} \in \boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1}} \Pr[\exists M' \text{ with } d(M', M_{n,m'}) \leq (1-\varepsilon)m/2 \text{ and } \boldsymbol{a}^T M' = \mathbf{0}] \right).$$

Now, we take another union bound, this time over the possible edits to the matrix; by changing at most $(1-\varepsilon)m/2$ entries, an adversary can form

$$\sum_{i=0}^{(1-\varepsilon)m/2} \binom{nm'}{i} \leq \left((1+o(1)) \frac{2en}{1-\varepsilon} \right)^{(1-\varepsilon)m/2} = 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n}$$

many $n \times m'$ matrices. Thus Eq. (4.3) is at most

$$\sum_{t=n+1}^{p} \left(\sum_{\boldsymbol{a} \in \boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1}} 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot \Pr[\boldsymbol{a}^T M_{n,m'} = \mathbf{0}] \right).$$

(Note: this is possible because, by conditioning on the locations of the entries edited, each altered matrix M' is distributed identically to $M_{n,m'}$.)

We now wish to bound the probability that $\boldsymbol{a}^T M_{n,m'} = \mathbf{0}$ for any fixed $\boldsymbol{a} \in \boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1}$. By Lemma 4.2.6 (as $\boldsymbol{a} \in \boldsymbol{H}_t$), and by the independence of the columns in $M_{n,m'}$, this probability is at most $\left(\frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}}\right)^{m'}$. Therefore, Eq. (4.3) is at most

$$\sum_{t=n+1}^{p} \left(\sum_{\boldsymbol{a} \in \boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1}} 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot \left(\frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}}\right)^{m'} \right).$$

We now bound the number of vectors $\boldsymbol{a}$ in each $\boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1}$. By definition, $\boldsymbol{H}_t \setminus \boldsymbol{H}_{t-1} \subset \mathbb{F}_p^n \setminus \boldsymbol{H}_{t-1}$, and by Lemma 4.2.7, the size of $\mathbb{F}_p^n \setminus \boldsymbol{H}_{t-1}$ is bounded above by

$$\left(\frac{2p}{\alpha t}\right)^n \cdot t^{n^{1-\varepsilon/2}},$$

where $\alpha \in (0,1)$ is any fixed constant (note that the constant C above depends on α). Thus Eq. (4.3) is bounded by the following explicit expression:

$$\begin{aligned}
&\sum_{t=n+1}^{p} \left(\frac{2p}{\alpha t}\right)^n \cdot t^{n^{1-\varepsilon/2}} \cdot 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot \left(\frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}}\right)^{m'} \\
&= 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot n^{-(1-5\varepsilon/6)\frac{m'}{2}} \cdot \left(\frac{2}{\alpha}\right)^n \cdot C^{m'} \sum_{t=n+1}^{p} \left(\frac{t}{p}\right)^{m'-n} t^{n^{1-\varepsilon/2}}.
\end{aligned}$$

Now, bounding each piece separately, and recalling that $m' \geq n$,

$$\left(\frac{2}{\alpha}\right)^n C^{m'} = 2^{O(n)},$$

$$\sum_{t=n+1}^{p} \left(\frac{t}{p}\right)^{m'-n} t^{n^{1-\varepsilon/2}} \leq p \cdot 1 \cdot p^{n^{1-\varepsilon/2}} = 2^{n^{\varepsilon/3}} \cdot 2^{n^{\varepsilon/3} \cdot n^{1-\varepsilon/2}} = 2^{o(n)},$$

$$2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot n^{-(1-5\varepsilon/6)\frac{m'}{2}} = 2^{-(1-o(1))\frac{\varepsilon}{12}\cdot m\log n},$$

where in the last equality, we use the fact that $m' = m - n^{1-\varepsilon/6} = (1-o(1))m$. Thus in total, Eq. (4.3) is at most

$$2^{(-\varepsilon/12+o(1))m\log n} = 2^{-\Theta(n\log n)}.$$

This completes the proof of the lemma. $\square$

Completing the proof

Given the assumption $m \leq C(\varepsilon)n$, we will in fact prove something slightly stronger, namely that Theorem 4.1.2 holds over $\mathbb{F}_p$ for an appropriate choice of p. We wish to bound the probability that there exists some nonzero vector $\boldsymbol{a} \in \mathbb{F}_p^n$ with $\boldsymbol{a}^T M_{n,m} = \mathbf{0}$, even after at most $(1-\varepsilon)m/2$ edits. Let $p = \Theta(2^{n^{\varepsilon/3}})$ be prime. We begin by dividing into "structured" and "unstructured" vectors; for brevity, given a nonzero vector $\boldsymbol{a}$ and matrix M, we denote by $\mathcal{E}(\boldsymbol{a}, M)$ the event that there exists a matrix M' with $d(M', M) \leq (1-\varepsilon)m/2$ and $\boldsymbol{a}^T M' = \mathbf{0}$.

$$\begin{aligned}
\Pr[\exists\, \boldsymbol{a} \in \mathbb{F}_p^n \text{ with } \mathcal{E}(\boldsymbol{a}, M_{n,m})]& \\
&\leq \Pr[\exists\, \boldsymbol{a} \in \boldsymbol{H}_n \text{ with } \mathcal{E}(\boldsymbol{a}, M_{n,m})] \qquad (4.4)\\
&+ \Pr[\exists\, \boldsymbol{a} \in \mathbb{F}_p^n \setminus \boldsymbol{H}_n \text{ with } \mathcal{E}(\boldsymbol{a}, M_{n,m})], \qquad (4.5)
\end{aligned}$$

where $\boldsymbol{H}_n$ is the set of "good" or "unstructured" vectors defined in Section 4.2.3. The first summand (4.4) is bounded as follows: first, take a union bound over possible edits to $M_{n,m}$. There are

$$\sum_{i=0}^{(1-\varepsilon)m/2} \binom{nm}{i} = 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n}$$

possible choices for M'. Thus, for the first term (4.4), we obtain a bound of

$$2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot \Pr[\exists\, \boldsymbol{a} \in \boldsymbol{H}_n \text{ with } \boldsymbol{a}^T M_{n,m} = \mathbf{0}].$$

(As in the proof of Lemma 4.3.4, this is possible because, by conditioning on the locations of the entries edited, each M' is distributed identically to $M_{n,m}$.) And for $\boldsymbol{a} \in \boldsymbol{H}_n$, and $\boldsymbol{x} \in \{\pm 1\}^n$ chosen uniformly at random, Lemma 4.2.6 gives

$$\Pr\left[\boldsymbol{a}^T \boldsymbol{x} = 0\right] \leq \frac{Cn}{pn^{(1/2-5\varepsilon/12)}} < \frac{n}{p}.$$

So for $M_{n,m}$ with $m \geq n + n^{1-\varepsilon/6}$ columns, the probability of having $\boldsymbol{a}^T M_{n,m} = \mathbf{0}$ is at most $\left(\frac{n}{p}\right)^{n+n^{1-\varepsilon/6}}$. Therefore, as there are at most p^n vectors $\boldsymbol{a} \in \boldsymbol{H}_n$, and as $m \leq C(\varepsilon) \cdot n$, the first summand (4.4) is bounded by

$$\begin{aligned}
2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot p^n \left(\frac{n}{p}\right)^{n+n^{1-\varepsilon/6}} &= 2^{(1-\varepsilon+o(1))\frac{m}{2}\log n} \cdot p^{-n^{1-\varepsilon/6}} n^{n+n^{1-\varepsilon/6}} \\
&= 2^{O(n\log n)} \cdot 2^{-n^{\varepsilon/3} n^{1-\varepsilon/6}} \\
&= 2^{-\Theta(n^{1+\varepsilon/6})}.
\end{aligned}$$

Now we bound the second summand (4.5). We begin by restricting to the first $m' = m - n^{1-\varepsilon/6}$ columns of $M_{n,m}$. This gives a strictly larger probability, as it is more likely that there is a linear dependency among the rows of a matrix when we restrict to only a subset of its columns. So (4.5) is bounded above by

$$\begin{aligned}
&\Pr[\exists\, \boldsymbol{a} \in \mathbb{F}_p^n \setminus \boldsymbol{H}_n \text{ with } \mathcal{E}(\boldsymbol{a}, M_{n,m'})] \\
&\qquad \leq \Pr[\exists\, \boldsymbol{a} \in \mathbb{F}_p^n \setminus \boldsymbol{H}_n \text{ with } |\operatorname{supp}(a)| \geq n^{1-\varepsilon/2} \text{ and } \mathcal{E}(\boldsymbol{a}, M_{n,m'})] \\
&\qquad + \Pr[\exists\, \boldsymbol{a} \in \mathbb{F}_p^n \setminus \boldsymbol{H}_n \text{ with } |\operatorname{supp}(a)| < n^{1-\varepsilon/2} \text{ and } \mathcal{E}(\boldsymbol{a}, M_{n,m'})]
\end{aligned}$$

And these are respectively the precise probabilities bounded in Lemmas 4.3.4 and 4.3.3. Therefore this is at most

$$2^{-\Theta(n \log n)} + 2^{-\Theta(n)}.$$

Thus in total, the probability that there exists a nonzero vector $\boldsymbol{a} \in \mathbb{F}_p^n$ with $\boldsymbol{a}^T M_{n,m} = \mathbf{0}$, even after at most $(1-\varepsilon)m/2$ edits is at most

$$2^{-\Theta(n^{1+\varepsilon/6})} + 2^{-\Theta(n \log n)} + 2^{-\Theta(n)} = 2^{-\Theta(n)}. \qquad \square$$

Appendix A

Variance in number of cliques

In this section, we show that the numbers of k-cliques in $\mathbb{G}(n, W_r)$ and $\mathbb{G}(n, U_r)$ have high variance for k within a reasonable range (Corollary A.0.5). This makes it impossible to directly use the second moment method to find a useful lower bound on the clique number of these graphs.

In more detail, our setting is as follows: given any graphon W, we will write X_k for the number of k-cliques in $\mathbb{G}(n, W)$. Suppose that, for a given graphon W, we have found a cutoff value $k = k(n)$ at which $\mathbb{E}[X_k]$ goes from asymptotically infinite to asymptotically zero, giving an upper bound of $\omega(\mathbb{G}(n, W)) \leq (1 + o(1))k$ with probability $1 - o(1)$ by Markov's inequality. In order to prove a matching lower bound, we would like to show that the number of cliques of size $(1-o(1))k$ in $\mathbb{G}(n, W)$ is a.a.s. nonzero. Perhaps the simplest way to do this, and the technique used for Erdős-Rényi random graphs in [GM75] and [Mat76], is the second moment method; namely, Chebyshev's inequality gives the bound

$$\Pr[X_k = 0] \leq \frac{\mathrm{Var}(X_k)}{\mathbb{E}[X_k]^2} = \frac{\mathbb{E}[X_k^2]}{\mathbb{E}[X_k]^2} - 1 \tag{A.1}$$

for any k. If $\mathrm{Var}(X_k) = o(\mathbb{E}[X_k]^2)$, or equivalently $\mathbb{E}[X_k^2] = (1 + o(1))\mathbb{E}[X_k]^2$, then this shows that $X_k \geq 1$ with probability $1 - o(1)$, and thus $\omega(\mathbb{G}(n, W)) \geq k$ a.a.s.

The entire challenge of applying the second moment method lies in obtaining a good bound on the ratio $\mathbb{E}[X_k^2]/\mathbb{E}[X_k]^2$. The following lemma gives a slightly more explicit expression for this quantity; it is a standard result adapted slightly for this application (see Sections 4.3 and 4.5 of [AS08]).

Lemma A.0.1. *Let W be a graphon, and for $S \subseteq [n]$, let A_S be the event that the elements of S form a clique in $\mathbb{G}(n, W)$. Then*

$$\frac{\mathbb{E}[X_k^2]}{\mathbb{E}[X_k]^2} = \sum_{i=0}^{k} \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} \cdot \frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2},$$

where S_i is any subset of $[n]$ of size k that intersects $[k]$ in exactly i elements.

Proof. This lemma follows from a direct computation of the first and second moments;

first, write

$$X_k = \sum_{S \subseteq [n],\, |S|=k} I_S,$$

where I_S is the indicator variable for the vertices in S forming a clique. With this notation, we obtain

$$\mathbb{E}[X_k] = \sum_{S \subseteq [n],\, |S|=k} \mathbb{E}[I_S] = \binom{n}{k} \Pr[A_{[k]}]. \tag{A.2}$$

Similarly,

$$\mathbb{E}[X_k^2] = \sum_{\substack{S,T \in [n] \\ |S|,|T|=k}} \mathbb{E}\left[I_S I_T\right] = \sum_{\substack{S,T \in [n] \\ |S|,|T|=k}} \Pr[A_S \cap A_T]. \tag{A.3}$$

And notice that this last probability depends only on the size of the intersection of S and T; thus we can group the terms of the sum above by the size i of the intersection. The number of ways to choose two sets of k vertices that overlap in exactly i elements is $\binom{n}{k}\binom{k}{i}\binom{n-k}{k-i}$; so (A.3) becomes

$$\mathbb{E}[X_k^2] = \sum_{i=0}^{k} \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} \Pr[A_{[k]} \cap A_{S_i}]. \tag{A.4}$$

And combining (A.2) and (A.3), we see that

$$\frac{\mathbb{E}[X_k^2]}{\mathbb{E}[X_k]^2} = \sum_{i=0}^{k} \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} \cdot \frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2},$$

as desired. □

Now, in order to apply these results to any graphon W, we need to compute the sum given in the lemma above, and in particular, $\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2}$. For W of the right form, we can obtain a more explicit expression:

Lemma A.0.2. *For any graphon W of the form $W(x,y) = f(x)f(y)$, i.e., for any W that is rank-1,*

$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = \left(\int_0^1 f(x)^{k-1} dx\right)^{-2i} \cdot \left(\int_0^1 f(x)^{2k-i-1} dx\right)^{i},$$

where A_S is the event that the elements of S form a clique in $\mathbb{G}(n,W)$, and S_i is any subset of $[n]$ of size k that intersects $[k]$ in exactly i elements.

Proof. We begin by computing $\Pr[A_{S_i} \cap A_{[k]}]$, considering in three parts the edges of the graph consisting of a clique on $[k]$ and a clique on S_i. This is equal to

$$\begin{aligned}
\Pr[A_{S_i} \cap A_{[k]}] &= \int_{[0,1]^{2k-i}} \left(\prod_{\substack{\ell \neq m \in S \setminus (S\cap[k]) \\ \text{or } [k]\setminus(S\cap[k])}} f(x_\ell) f(x_m) \right) \cdot \left(\prod_{\ell \neq m \in S \cap [k]} f(x_\ell) f(x_m) \right) \\
&\qquad \cdot \left(\prod_{\substack{\ell \in S\cap[k], \\ m \in (S\cup[k])\setminus(S\cap[k])}} f(x_\ell) f(x_m) \right) d\vec{x} \\
&= \int_{[0,1]^{2k-i}} \left(\prod_{\ell \in (S\cup[k])\setminus(S\cap[k])} f(x_\ell)^{k-1} \right) \cdot \left(\prod_{\ell \in S\cap[k]} f(x_\ell)^{2k-i-1} \right) d\vec{x} \\
&= \left(\int_0^1 f(x)^{k-1}\, dx \right)^{2k-2i} \cdot \left(\int_0^1 f(x)^{2k-i-1}\, dx \right)^{i}
\end{aligned}$$

Without any further assumptions on $f(x)$, this is as far as $\Pr[A_{S_i} \cap A_{[k]}]$ can be evaluated. To finish off, we compute

$$\begin{aligned}
\Pr[A_{[k]}] &= \int_{[0,1]^k} \prod_{\ell \neq m \in [k]} f(x_\ell) f(x_m)\, d\vec{x} \\
&= \left(\int_0^1 f(x)^{k-1}\, dx \right)^{k}.
\end{aligned}$$

Therefore

$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = \left(\int_0^1 f(x)^{k-1} dx \right)^{-2i} \cdot \left(\int_0^1 f(x)^{2k-i-1} dx \right)^{i}.$$

□

For the graphons W_r and U_r, we can evaluate the integrals above and obtain more explicit expressions:

Lemma A.0.3. *Given any $k = \omega(1)$ and $1 \leq i \leq k-1$,*

1. *for the graphon W_r,*
$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = (\Theta(k))^i,$$

2. *and for the graphon U_r,*
$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = \left(\Theta(k^{1/r})\right)^i.$$

Proof. We begin with (1). For the graphon $W_r(x,y) = (1-x)^r(1-y)^r$, Lemma A.0.2

gives

$$\begin{aligned}
\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} &= \left(\int_0^1 (1-x)^{r(k-1)} dx\right)^{-2i} \cdot \left(\int_0^1 (1-x)^{r(2k-i-1)} dx\right)^{i} \\
&= \left(\frac{1}{r(k-1)+1}\right)^{-2i} \left(\frac{1}{r(2k-i-1)+1}\right)^{i} \\
&= (\Theta(k))^i.
\end{aligned}$$

Now we prove (2). Again by Lemma A.0.2, for $U_r(x,y) = (1-x^r)(1-y^r)$, we have

$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = \left(\int_0^1 (1-x^r)^{k-1} dx\right)^{-2i} \cdot \left(\int_0^1 (1-x^r)^{2k-i-1} dx\right)^{i}.$$

And as computed in the proof of Lemma 3.4.1,

$$\int_0^1 (1-x^r)^{k-1} dx = \frac{\Gamma(k) \cdot \Gamma(1+\frac{1}{r})}{\Gamma(k+\frac{1}{r})}.$$

Applying this to the expression above, we obtain

$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = \left(\frac{\Gamma(2k-i)\Gamma(1+\frac{1}{r})}{\Gamma(2k-i+\frac{1}{r})}\right)^{i} \left(\frac{\Gamma(k)\Gamma(1+\frac{1}{r})}{\Gamma(k+\frac{1}{r})}\right)^{-2i}. \tag{A.5}$$

Using the approximation $\frac{\Gamma(k)}{\Gamma(k+\frac{1}{r})} = k^{-\frac{1}{r}}(1+o(1))$ obtained from Stirling's formula, (A.5) becomes

$$\begin{aligned}
\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} &= \left((2k-i)^{-\frac{1}{r}} \cdot \Gamma(1+\tfrac{1}{r})(1+o(1))\right)^{i} \left(k^{-\frac{1}{r}} \cdot \Gamma(1+\tfrac{1}{r})(1+o(1))\right)^{-2i} \\
&= \left(\frac{k^{2/r}}{\Gamma(1+\frac{1}{r})(2k-i)^{1/r}}(1+o(1))\right)^{i} \\
&= \left(\Theta(k^{1/r})\right)^{i}.
\end{aligned}$$

□

We are now nearly ready to show that for the graphons W_r and U_r, and for any reasonably large k, the number of k-cliques in $\mathbb{G}(n, W)$ has large variance.

Theorem A.0.4. *For any $r > 0$ and any graphon W, if*

$$\frac{\Pr[A_{S_i} \cap A_{[k]}]}{\Pr[A_{[k]}]^2} = \left(\Omega(k^{1/r})\right)^{i},$$

then for any $k = \Theta\left(n^{\frac{r}{r+1}}\right)$, we have $\mathrm{Var}(X_k) = \omega(\mathbb{E}[X_k]^2)$.

Before proving the theorem, note that together with Lemma A.0.3, it directly implies the following corollary.

Corollary A.0.5. *Given $r > 0$,*

(i) for any $k = \Theta(\sqrt{n})$, if X_k is the number of k-cliques in $\mathbb{G}(n, W_r)$, then $\mathrm{Var}(X_k) = \omega(\mathbb{E}[X_k]^2)$*, and*

(ii) for any $k = \Theta\left(n^{\frac{r}{r+1}}\right)$, if X_k is the number of k-cliques in $\mathbb{G}(n, U_r)$, then $\mathrm{Var}(X_k) = \omega(\mathbb{E}[X_k]^2)$.

Now we prove the theorem.

Proof of Theorem A.0.4. We will apply Lemma A.0.1 to show that $\mathbb{E}[X_k^2]/\mathbb{E}[X_k]^2 = \omega(1)$, or equivalently, $\mathrm{Var}(X_k) = \omega(\mathbb{E}[X_k]^2)$. Recall that, by Lemma A.0.1 and by hypothesis,

$$\frac{\mathbb{E}[X_k^2]}{\mathbb{E}[X_k]^2} = \sum_{i=1}^{k-1} \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} \cdot \left(\Omega(k^{1/r})\right)^i.$$

We will show not only that this sum is $\omega(1)$, but in fact, that it always contains a term that is $\omega(1)$. This comes down almost entirely to appropriately estimating the three binomial coefficients appearing in the i^{th} term of the sum above. First, for any k that is $\omega(1)$ but sublinear,

$$\binom{n}{k} = \left(\frac{ne}{k}\right)^k e^{-o(k)}. \tag{A.6}$$

Next, observe that for all $0 \le i \le k$, since $k = o(n)$, we also have $k - i = o(n-k)$. If $i = \varepsilon k$ for some constant $0 < \varepsilon < 1$, then $(k - i) = \omega(1)$ as well, and we obtain

$$\begin{aligned}\binom{n-k}{k-i} &= \left(\frac{(n-k)e}{k-i}\right)^{k-i} e^{-o(k-i)} \\ &\ge \left(\frac{ne}{k}\right)^{(1-\varepsilon)k} e^{-o(k)}. \end{aligned} \tag{A.7}$$

We also have

$$\binom{k}{i} \ge \left(\frac{k}{i}\right)^i = e^{\varepsilon k \log \frac{1}{\varepsilon}}. \tag{A.8}$$

Together, (A.6), (A.7), and (A.8) imply that for $i = \varepsilon k = \Theta(k)$, the i^{th} term of the sum above is

$$\begin{aligned}
\frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} \cdot \left(\Omega(k^{1/r})\right)^i &\geq \frac{e^{\varepsilon k \log \frac{1}{\varepsilon}} \cdot \left(\frac{ne}{k}\right)^{(1-\varepsilon)k} e^{-o(k)}}{\left(\frac{ne}{k}\right)^k e^{-o(k)}} \cdot \left(\Theta(k^{1/r})\right)^i \\
&= e^{\varepsilon k \log \frac{1}{\varepsilon} - o(k)} \left(\frac{ne}{k}\right)^{-\varepsilon k} \cdot \left(\Theta\left(k^{1/r}\right)\right)^{\varepsilon k} \\
&= e^{\varepsilon k \log \frac{1}{\varepsilon} - o(k)} \left(\frac{ne}{k}\right)^{-\varepsilon k} \cdot \left(\Theta\left(\frac{n}{k}\right)\right)^{\varepsilon k}, \qquad \text{since } k = \Theta\left(n^{\frac{r}{r+1}}\right) \\
&= e^{\varepsilon k (\log \frac{1}{\varepsilon} - C) - o(k)}
\end{aligned}$$

for some constant C. Note that we can make C as large as we want by controlling the size of the implicit constant in $k = \Theta(n^{\frac{r}{r+1}})$. However, for any fixed choice of C, we can find some small but constant $\varepsilon = \varepsilon(C)$ such that $\log(1/\varepsilon) > \log(C)$. So for some ε, this expression will always be $\omega(1)$. Therefore $\mathbb{E}[X_k^2]/\mathbb{E}[X_k]^2 = \omega(1)$, or equivalently $\mathrm{Var}(X_k) = \omega(\mathbb{E}[X_k]^2)$, as desired. $\square$

Bibliography

[AAR99] G. Andrews, R. Askey, and R. Roy. *Special Functions*. Cambridge University Press, 1999.

[AKS98] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. *Random Structures Algorithms*, 13(3–4):457–466, 1998.

[AKS03] N. Alon, M. Krivelevich, and B. Sudakov. Turán numbers of bipartite graphs and related Ramsey-type questions. *Combin. Probab. Comput.*, 12:477–494, 2003. Special issue on Ramsey theory.

[AP03] D. Achlioptas and Y. Peres. The threshold for random k-SAT is $2^k \ln 2 - O(k)$. In *Proceedings of the thirty-fifth ACM symposium on theory of computing*, pages 223–231. ACM, 2003.

[AS08] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2008.

[BB16] P. Bennett and T. Bohman. A note on the random greedy independent set algorithm. *Random Structures Algorithms*, 49:479–502, 2016.

[BBS04] J. Balogh, B. Bollobás, and M. Simonovits. The number of graphs without forbidden subgraphs. *J. Combin. Theory Ser. B*, 91:1–24, 2004.

[BC18] B. Bukh and D. Conlon. Rational exponents in extremal graph theory. *J. Eur. Math. Soc.*, 20:1747–1757, 2018.

[BCvdH20] K. Bogerd, R. M. Castro, and R. van der Hofstad. Cliques in rank-1 random graphs: the role of inhomogeneity. *Bernoulli*, 26(1):253–285, 2020.

[BFK18] T. Bläsius, T. Friedrich, and A. Krohmer. Cliques in hyperbolic random graphs. *Algorithmica*, 80(8):2324–2344, 2018.

[BJR07] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 31(1):3–122, 2007.

[BK10] T. Bohman and P. Keevash. The early evolution of the H-free process. *Invent. Math.*, 181:291–336, 2010.

[BLS17] J. Balogh, H. Liu, and M. Sharifzadeh. The number of subsets of integers with no k-term arithmetic progression. *Int. Math. Res. Not. IMRN*, (20):6168–6186, 2017.

[BMS15] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28:669–709, 2015.

[BNS19] J. Balogh, B. Narayanan, and J. Skokan. The number of hypergraphs without linear cycles. *J. Combin. Theory Ser. B*, 134:309–321, 2019.

[BP12] S. Ball and V. Pepe. Asymptotic improvements to the lower bound of certain bipartite Turán numbers. *Combin. Probab. Comput.*, 21:323–329, 2012.

[BS74] J. A. Bondy and M. Simonovits. Cycles of even length in graphs. *J. Combin. Theory Ser. B*, 16:97–105, 1974.

[BS11a] J. Balogh and W. Samotij. The number of $K_{m,m}$-free graphs. *Combinatorica*, 31:131–150, 2011.

[BS11b] J. Balogh and W. Samotij. The number of $K_{s,t}$-free graphs. *J. Lond. Math. Soc.*, 83(2):368–388, 2011.

[Buk15] B. Bukh. Random algebraic construction of extremal graphs. *Bull. Lond. Math. Soc.*, 47:939–945, 2015.

[BVW10] J. Bourgain, V. H. Vu, and P. M. Wood. On the singularity probability of discrete random matrices. *J. Funct. Anal.*, 258(2):559–603, 2010.

[CMMM19] M. Campos, L. Mattos, R. Morris, and N. Morrison. On the singularity of random symmetric matrices. *arXiv:1904.11478*, 2019.

[CT17] J. Corsten and T. Tran. Balanced supersaturation for degenerate hypergraphs. *arXiv:1707.03788*, 2017.

[DGLU11] L. Devroye, A. György, G. Lugosi, and F. Udina. High-dimensional random geometric graphs and their clique number. *Electron. J. Probab.*, 16:2481–2508, 2011.

[DHM19] M. Doležal, J. Hladký, and A. Máthé. Cliques in dense inhomogeneous random graphs. *Random Structures Algorithms*, 51(2):275–314, 2019.

[DM15] Y. Deshpande and A. Montanari. Finding hidden cliques of size $\sqrt{N/e}$ in nearly linear time. *Found. Comput. Math.*, 15(4):1069–1128, 2015.

[EFR86] P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.*, 2:113–121, 1986.

[EKR76] P. Erdős, D. J. Kleitman, and B. L. Rothschild. Asymptotic enumeration of K_n-free graphs. In *Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II*, pages 19–27. Accad. Naz. Lincei, Rome, 1976.

[ER59] P. Erdős and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[Erd64] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. Math.*, 2:183–190, 1964.

[Erd81] P. Erdős. On the combinatorial problems which I would most like to see solved. *Combinatorica*, 1(1):25–42, 1981.

[ES46] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc*, 52:1087–1091, 1946.

[ES83] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3:181–192, 1983.

[ES84] P. Erdős and M. Simonovits. Cube-supersaturated graphs and related problems. In *Progress in graph theory (Waterloo, Ont., 1982)*, pages 203–218. Academic Press, Toronto, ON, 1984.

[FJ18] A. Ferber and V. Jain. Singularity of random symmetric matrices–a combinatorial approach to improved bounds. *arXiv:1809.04718*, 2018.

[FJLS19] A. Ferber, V. Jain, K. Luh, and W. Samotij. On the counting problem in inverse Littlewood–Offord theory. *arXiv:1904.10425*, 2019.

[FK12] T. Friedrich and A. Krohmer. Parameterized clique on scale-free networks. In *International Symposium on Algorithms and Computation*, pages 659–668. Springer, 2012.

[FLM19] A. Ferber, K. Luh, and G. McKinley. Resilience of the rank of random matrices. *arXiv:1910.03619*, 2019.

[FMS] A. Ferber, G. McKinley, and W. Samotij. Supersaturated sparse graphs and hypergraphs. *To appear in International Mathematics Research Notices.* doi:10.1093/imrn/rny030.

[FR10] U. Feige and D. Ron. Finding hidden cliques in linear time. In *21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*. Discrete Mathematics and Theoretical Computer Science, 2010.

[FS83] R. J. Faudree and M. Simonovits. On a class of degenerate extremal graph problems. *Combinatorica*, 3:83–93, 1983.

[FS13] Z. Füredi and M. Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdös centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 169–264. János Bolyai Math. Soc., Budapest, 2013.

[Für91] Z. Füredi. On a Turán type problem of Erdős. *Combinatorica*, 11:75–79, 1991.

[Gil59] E. Gilbert. Random graphs. *Ann. of Math. Stat.*, 30(4):1141–1144, 1959.

[GM75] G. Grimmett and C. McDiarmid. On colouring random graphs. In *Math. Proc. Cambridge Philos. Soc.*, volume 77, pages 313–324. Cambridge University Press, 1975.

[Hal77] G. Halász. Estimates for the concentration function of combinatorial number theory and probability. *Period. Math. Hungar.*, 8(3-4):197–211, 1977.

[HK18] J. Han and Y. Kohayakawa. On hypergraphs without loose cycles. *Discrete Math.*, 341(4):946–949, 2018.

[HR17] J. Hladký and I. Rocha. Independent sets, cliques, and colorings in graphons. *arXiv:1712.07367*, 2017.

[Jai19a] V. Jain. Approximate Spielman-Teng theorems for random matrices with heavy tailed entries: a combinatorial view. *arXiv:1904.11108*, 2019.

[Jai19b] V. Jain. Approximate Spielman-Teng theorems for the least singular value of random combinatorial matrices. *arXiv:1904.10592*, 2019.

[JŁN10] S. Janson, T. Łuczak, and I. Norros. Large cliques in a power-law random graph. *J. Appl. Probab.*, 47(4):1124–1135, 2010.

[Kar76] R. M. Karp. The probabilistic analysis of some combinatorial search algorithms. In *New directions and recent results in algorithms and complexity*, pages 1–19. Academic Press, New York, 1976.

[KKS95] J. Kahn, J. Komlós, and E. Szemerédi. On the probability that a random ±1-matrix is singular. *J. Amer. Math. Soc.*, 8(1):223–240, 1995.

[Kom67] J. Komlós. On the determinant of (0, 1) matrices. *Studia Sci. Math. Hungar*, 2:7–21, 1967.

[KS98] M. Krivelevich and B. Sudakov. Coloring random graphs. *Inform. Process. Lett.*, 67(2):71–74, 1998.

[KST54] T. Kővári, V. T. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloq. Math.*, 3:50–57, 1954.

[KW82] D. J. Kleitman and K. J. Winston. On the number of graphs without 4-cycles. *Discrete Math.*, 41:167–172, 1982.

[KW96] D. Kleitman and D. Wilson. On the number of graphs which lack small cycles. Manuscript, 1996.

[LMN19] K. Luh, S. Meehan, and H. H. Nguyen. Random matrices over finite fields: methods and results. *arXiv:1907.02575*, 2019.

[Lov12] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.

[LS06] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[Mar88] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24:51–60, 1988.

[Mat76] D. Matula. The largest clique size in a random graph. Technical report, Department of Computer Science, Southern Methodist University, 1976.

[McK19] G. McKinley. Superlogarithmic cliques in dense inhomogeneous random graphs. *SIAM J. Discrete Math.*, 33(3):1772–1900, 2019.

[MS16] R. Morris and D. Saxton. The number of $C_{2\ell}$-free graphs. *Adv. Math.*, 298:534–580, 2016.

[MV16] D. Mubayi and J. Verstraëte. A survey of Turán problems for expansions. In *Recent trends in combinatorics*, volume 159 of *IMA Vol. Math. Appl.*, pages 117–143. Springer, 2016.

[MW19] D. Mubayi and L. Wang. The number of triple systems without even cycles. *Combinatorica*, 39:679–704, 2019.

[Ngu13] H. H. Nguyen. On the singularity of random combinatorial matrices. *SIAM J. Discrete Math.*, 27(1):447–458, 2013.

[NRS06] B. Nagle, V. Rödl, and M. Schacht. Extremal hypergraph problems and the regularity method. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 247–278. Springer, Berlin, 2006.

[RV08] M. Rudelson and R. Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218(2):600–633, 2008.

[ST15] D. Saxton and A. Thomason. Hypergraph containers. *Invent. Math.*, 201:925–992, 2015.

[Tik] K. Tikhomirov. Singularity of random Bernoulli matrices. *To appear in Ann. of Math.* arXiv:1812.09016.

[Tur41] P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

[TV07] T. Tao and V. Vu. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.*, 20(3):603–628, 2007.

[TV09] T. Tao and V. H. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Ann. of Math.*, 169(2):595–632, 2009.

[TW12] T. A. Terlep and J. Williford. Graphs from generalized Kac-Moody algebras. *SIAM J. Discrete Math.*, 26:1112–1120, 2012.

[vdH16] R. van der Hofstad. *Random graphs and complex networks*, volume 1. Cambridge university press, 2016.

[vdH17] R. van der Hofstad. *Random graphs and complex networks*, volume 2. Preprint available on author's webpage., 2017+.

[Ver14] R. Vershynin. Invertibility of symmetric random matrices. *Random Structures Algorithms*, 44(2):135–182, 2014.

[Vu08] V. Vu. Random discrete matrices. In *Horizons of combinatorics*, volume 17 of *Bolyai Soc. Math. Stud.*, pages 257–280. Springer, Berlin, 2008.