On the Evolution of Random Discrete Structures

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(dr. rer. nat.)
im Fach Informatik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Deryk Simeon Osthus, B.A.
geboren am 10.10.1974 in Cardiff

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Hans Meyer

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Bodo Krause

Gutachter:

1. Prof. Dr. Hans Jürgen Prömel
2. Prof. Joel H. Spencer, PhD
3. Prof. Dr. Angelika Steger

eingereicht am:	10.2.2000
Tag der mündlichen Prüfung:	26.4.2000

Preface

In this thesis, we study the evolution of random discrete structures, an area of research which was initiated by Erdős and Rényi in 1959. By now, it has an immense literature and is still very much alive.

One of the main motivations for their approach was the question of *typical behaviour*: suppose we are given a class of objects and are interested in whether they have some property or not. The traditional "worst case" approach would be to require that *all* objects in this class have this property (or that an algorithm performs well on *all* objects in this class). But in many circumstances it suffices to know whether at least a high proportion or "almost all" objects have this property. Thus, in probabilistic terms, the question is whether an object chosen at random from this class is very likely to have this property (or whether an algorithm performs well *on average*).

It turns out that there is a close connection between the above question and the study of the *evolution* of random discrete structures. Such evolution processes usually fit into the following general framework. Initially (say at time 0), we start with a very simple structure (e.g. a graph on n vertices with no edges) and a set of "building blocks" (e.g. the set of edges of the complete graph on n vertices). As time increases, we randomly add (possibly subject to certain rules) more and more elements from our set of building blocks. The process terminates when no more blocks can be added. The basic question which we shall investigate is the following:

> *For large instances (e.g. large values of n), what are the likely properties of the random structure produced by the process at any given time?*

One of the remarkable discoveries of Erdős and Rényi was the fact that such evolution processes exhibit so-called *threshold phenomena* or *phase transitions*: given some property, up to some critical time the random structure generated by that time is very unlikely to have this property, whereas shortly afterwards, it is extremely likely to have it. We shall see that in many cases, knowledge about these threshold functions and the likely behaviour of an evolution process also tells us about the typical properties of objects in a given class. This has consequences for instance for the average case analysis of algorithms: the expected running time or performance of an algorithm may be very sensitive to the input structure. For instance, a random graph with n vertices and $0.49n$ edges can be optimally coloured in expected polynomial time using a simple brute force approach, whereas the same brute force method is expected to require exponentially many steps if the random graph has $0.51n$ edges.

Although this thesis is concerned with the evolution of random structures, the results obtained can also be summarized according to the following keywords:

- *Random greedy algorithms:* we study the output of a random greedy algorithm which, for a given graph H, produces a random H-free graph.
- *Extremal results:* improving on previous bounds, we prove the existence of graphs with high girth and high chromatic number – they are thus both locally sparse and globally complex.
- *Asymptotic enumeration:* we prove sharp asymptotic bounds on the number of triangle-free graphs with n vertices and m edges for a large range of m – while it is easy to count the number of elements in a class whose elements can be generated by a series of independent decisions (like the total number of graphs on n vertices), counting elements of classes defined by structural side constraints (such as triangle-freeness) turns out to be more difficult.
- *Probabilistic versions of "classical" theorems:* we prove a probabilistic version of Sperner's theorem on finite sets – while classical results (like Sperner's theorem) apply to all objects in a given class, probabilistic theorems are concerned with the "typical" behaviour of objects in a given class.

The thesis is organized as follows. In Chapter 1, we give a very brief introduction to the theory of random graphs and introduce some of the definitions which we shall need later on. We also outline some of the applications and connections of the topic to other fields of research, such as the analysis of algorithms. In Chapter 2, we then summarize the results obtained in this thesis and describe their relationship to previous research. In many cases, we confine ourselves to stating only a special case of the result or just describing it informally, postponing the detailed statement to the appropriate chapter. In Chapter 3, we state some well-known large deviation inequalities which we shall make frequent use of later on. The remaining chapters are then devoted to the proofs of the results. In Chapter 4, we consider random greedy H-free graphs, Chapter 5 deals with the evolution of random triangle-free graphs, and Chapters 6 and 7 are concerned with random subsets of a finite set.

Chapter 4 of this thesis is joint work with Anusch Taraz [64], Chapter 5 is joint work with Hans Jürgen Prömel and Anusch Taraz [63], and Chapter 6 is joint work with Yoshi Kohayakawa and Bernd Kreuter [51]. The material in Chapter 7 is taken from [62].

I am very grateful to Hans Jürgen Prömel, my supervisor, and to Yoshi, Bernd, and Anusch, for their collaboration and for fruitful discussions. I would also like to thank my other collegues at the Humboldt-University for their help. I am indebted to the "Graduiertenkolleg Algorithmische Diskrete Mathematik" for enabling me to write this thesis by providing me with a scholarship. Finally, I would like to take this opportunity to thank Daniela, my parents and my sister for their love and support.

Contents

Chapter 1

Random graphs – an introduction

1.1 Random graph evolution

This section contains a brief introduction to the theory of random graphs. We will also give some definitions which will be used later on. For a more in-depth introduction to random graphs and related topics, we refer the reader to the books [6, 8, 41] and the survey [44]. For graph theoretical terminology not defined here, we refer the reader to the books [11, 20].

The study of the evolution of random graphs was initiated in a series of fundamental papers by Erdős and Rényi published in 1959–1961 (see e.g. the survey [45]). In these papers, they introduced the so-called "binomial model" for a random graph. Here a random graph $G_{n,p}$ on n vertices is obtained by including each edge with probability p (independently of all other edges). There is also a "dynamic" description of $G_{n,p}$ – to each edge of the complete graph on n vertices we assign a birthtime which is uniformly distributed in the interval $[0, 1]$ and which is independent of the birthtimes of the other edges. Then $G_{n,p}$ is the graph on n vertices whose edges are those whose birthtime is at most p. We say that $G_{n,p}$ has a given property $\mathcal{Q}$ *almost surely* if the probability that $G_{n,p}$ has $\mathcal{Q}$ tends to one as n tends to infinity.

The following paragraph – quoted from Diestel [20] – highlights a few of the stages in the evolution of a random graph. It demonstrates that, while one cannot say very much about the class of *all* graphs on n vertices which have some fixed density, one can say a surprising amount about their likely properties.

> *Let us see then what happens if p is allowed to vary with n. Almost immediately, a fascinating picture unfolds. For edge probabilities whose order of magnitude lies below n^{-2}, a random graph $G_{n,p}$ almost surely has no edges at all. As p grows, $G_{n,p}$ acquires more and more structure: from about $p = \sqrt{n}n^{-2}$ onwards, it almost surely has a component with more than two vertices, these components grow into trees, and around $p = n^{-1}$ the first cycles are born. Soon, some of these will have several crossing chords, making the graph non-planar. At the same time, one component outgrows the others, until it devours them around $p = (\log n)n^{-1}$, making*

> *the graph connected. Hardly later, at $p = (1+\varepsilon)(\log n)n^{-1}$, our graph almost surely has a Hamilton cycle!*

As the above description indicates, one of the most striking facts about random graphs is that their evolution exhibits phases of rapid change: for many graph properties one can find a function $t = t(n)$ so that if $p = p(n)$ is a little smaller than t, then almost surely $G_{n,p}$ does not have this property, while if p is only a little larger than t, then $G_{n,p}$ has this property almost surely. More formally, we say that t is a *threshold function* for a given property $\mathcal{Q}$ if the probability that $G_{n,p}$ has $\mathcal{Q}$ satisfies

$$\mathbb{P}[\, G_{n,p} \text{ has } \mathcal{Q}\,] \to \quad \begin{cases} 0 & \text{if } p/t \to 0 \\ 1 & \text{if } p/t \to \infty. \end{cases}$$

Bollobás and Thomason [16] proved that every monotone property (i.e. one that is preserved by adding edges) has a threshold function. However, there are properties for which one can prove an even "sharper" threshold behaviour. Here we say that t is a *sharp threshold* for a given property $\mathcal{Q}$ if for any fixed $\varepsilon > 0$, we have

$$\mathbb{P}[\, G_{n,p} \text{ has } \mathcal{Q}\,] \to \quad \begin{cases} 0 & \text{if } p \leq (1-\varepsilon)t \\ 1 & \text{if } p \geq (1+\varepsilon)t. \end{cases}$$

We say that a threshold is *coarse* if it is not sharp. For example, $1/n$ is a coarse threshold function for the property that $G_{n,p}$ contains a triangle, and $(\log n)/n$ is a sharp threshold for the property that $G_{n,p}$ is connected.

In Section 2.2 and Chapter 5 we will consider a slightly different but equally natural model of random graphs. Let $\mathcal{G}(n,m)$ denote the set of all graphs with n vertices and m edges and let $G_{n,m}$ be a graph chosen uniformly at random from $\mathcal{G}(n,m)$. We say that *almost all* graphs in $\mathcal{G}(n,m)$ have some property $\mathcal{Q}$ if the proportion of graphs in $\mathcal{G}(n,m)$ which have $\mathcal{Q}$ tends to one if $n \to \infty$. Unsurprisingly, it turns out that $G_{n,p}$ and $G_{n,m}$ are very likely to have similar properties if m is close to $p\binom{n}{2}$, the expected number of edges of $G_{n,p}$.

Of course, besides $G_{n,p}$ and $G_{n,m}$, there are many other models of random graphs – for instance random directed graphs, random trees, random regular graphs, and random subgraphs of the hypercube. We will consider two models for random triangle-free graphs below.

Similarly, there are random models for many objects other than graphs: for example random matrices, random partially ordered sets, and random hypergraphs. In Chapters 6 and 7, we will consider a model for a random hypergraph (which can be viewed also as a model for a random partially ordered set). The definitions of a threshold function and terms like "almost surely" etc. carry over to these models in the obvious way.

Finally, given two functions $f = f(n)$ and $g = g(n)$, we write $f = o(g)$ if $f(n)/g(n) \to 0$ as n tends to infinity, and we write $f = \mathcal{O}(g)$ if there exists a constant C so that $f(n)/g(n) \leq C$ for all n. We write $f = \Theta(g)$ if $f = \mathcal{O}(g)$ and $g = \mathcal{O}(f)$.

1.2 Applications and related fields

1.2.1 Extremal results

The application of random graphs to extremal problems actually predates the systematic study of the evolution of random graphs. For instance, a textbook example of an application of the so-called "probabilistic method" (see e.g. Alon and Spencer [6]), is the first exponential lower bound on the Ramsey number $R(t, t)$, which was proved by Erdős in 1947 [23]. He obtained this result by showing that, if $p = 1/2$, then a random graph $G_{n,p}$ is likely to contain neither a large complete subgraph nor a large independent set of vertices. ($R(s, t)$ is defined to be the smallest integer n so that every graph on n vertices contains either an independent set of size s or a complete subgraph of size t.)

Another famous result – proved by Erdős in 1959 [24] – is that there are graphs of arbitrarily large girth and arbitrarily large chromatic number. (The *girth* of a graph is the length of its shortest cycle and its *chromatic number* is the minimum number of colours needed to colour its vertices so that adjacent vertices receive different colours.) He proved this by considering the likely properties of $G_{n,p}$ for p a little smaller than $n^{-1/\ell}$. In Chapter 4, we give a proof of this result which yields improved bounds on the possible values of the chromatic number for graphs with a given number of vertices and given girth.

We now discuss some results obtained by the "nibble" method pioneered by Rödl. Answering an old question of Erdős and Hanani, in 1985 Rödl [77] proved the existence of asymptotically "near-perfect" partial Steiner (t, k, n)-systems (where k and $t < k$ are fixed). Here a *partial Steiner* (t, k, n)*-system* is a collection of k-sets (a k*-set* is a k-element subset of $\{1, \ldots, n\}$) so that every t-set is contained in at most one k-set. For example, a partial Steiner $(2, 3, n)$-system can be viewed as a collection of edge-disjoint triangles in the complete graph K_n on n vertices. A central problem in design theory is to find partial Steiner systems which are as large as possible. Rödl proved that there are partial Steiner (t, k, n)-systems so that the proportion of t-sets not contained in them (i.e. those which are not "packed" into some k-set) tends to zero. This bound was improved and generalized by several authors, including Pippenger and Spencer [65], Kahn [43], Grable [35], and Alon, Kim, and Spencer [5]. However, it remains an important open question how well such systems can be generated by random processes. A partial result in this direction was obtained by Grable [34], who (improving on earlier results by Rödl and Thoma [78] and Spencer [83]) showed that the following random greedy process almost surely produces a partial Steiner $(2, 3, n)$-system which contains all but at most $n^{7/4+o(1)}$ edges: starting with the empty graph, repeatedly pick a triangle uniformly at random from all triangles which are edge disjoint from all previously selected triangles, until there are no more candidate triangles.

Finally, one of the most well-known applications of the nibble method is due to Kim [47]. He applied it to solve a long outstanding problem in Ramsey theory – that of determining the order of magnitude of the Ramsey number $R(3, t)$. He achieved this by proving that with nonzero probability, a suitable random greedy process (see also Section 2.1) produces a triangle-free graph containing no "large" independent set. As a corollary, he also proved that there exist triangle-free graphs whose chromatic number is very large indeed: for sufficiently large n, there is a triangle-free graph whose chromatic number is at least $(n/\log n)^{1/2}/9$, which (by a result of Ajtai, Komlós, and Szemerédi [1]) is best possible up the value of the constant factor.

For further results and details on the material of this section, we refer the reader to the books [7, 6] and the surveys [10, 80].

1.2.2 Algorithms

While initially, the main applications of random graphs were to extremal graph theory, in the meantime random graphs have become important as models for the *average case analysis* of algorithms: one is interested in the likely performance of a given algorithm on a random graph. As Frieze and McDiarmid put it in their survey article [33]:

> *One attractive feature of average case analysis is that it banishes the pessimism of worst case analysis. NP-Completeness casts a much smaller shadow. Problems like finding a Hamilton cycle may become tractable. Of course, one can criticize the models as being unrealistic, but they are probably no more so than the pathological examples used in the proofs of NP-Completeness and the study of performance guarantees. Furthermore, the models can be close to those used in the empirical testing of algorithms.*

We illustrate this for the problem of graph colouring algorithms. Feige and Kilian [30] proved that unless NP $\subseteq$ ZPP, it is not possible to approximate the chromatic number of a graph on n vertices in polynomial time within a ratio $\mathcal{O}(n^{1-\varepsilon})$ for any fixed $\varepsilon > 0$. On the other hand, results of Grimmett and McDiarmid [37] and Bollobás [9] imply that if p is constant, then if one runs a simple greedy algorithm on $G_{n,p}$, almost surely the number of colours used will be about twice the chromatic number of $G_{n,p}$ (analogous results are also known if $p \to 0$). It is a major open question (see e.g. [33, 46]) to determine whether there is a polynomial time algorithm which needs significantly less colours. Although there are some results indicating that such an algorithm might not exist (see Vu [92]), a result of Prömel and Steger [67] shows that there are models of random graphs where a colouring algorithm exists that is almost surely optimal: extending results by previous authors [90, 21], they gave an algorithm whose (expected) running time is $\mathcal{O}(n^2)$, and which almost surely outputs an optimal colouring of a K_s-free graph chosen uniformly at random from the set of K_s-free graphs on n vertices. Their proof illustrates the use of structural results about the typical shape of objects in some class – it is based on the fact that all but a tiny proportion of K_s-free graphs have a simple structure.

1.2.3 Random models for complex systems

Physicists have become increasingly interested in random models which capture the dynamics of (typically large) complex systems. The strategy here is that an understanding of a simpler random model gives insight into the complex physical system under consideration. For example, the well-known Ising model (see e.g. the book by Grimmett [36]) is used to study the properties of ferromagnets. Related to this is the field of percolation theory (see [36]), which is perhaps best described as the study of permeability of random media. Note that Fill and Pemantle [31] proved the analogue of Theorem 2.10 in the context of oriented bond percolation on the hypercube (see the discussion at the beginning of Chapter 6).

Another example is the investigation of random models for large networks such as that formed by the nervous system (where the vertices represent nerve cells and the edges connecting them are the axons, see e.g. the book by Valiant [91]), the world wide web (where the vertices represent websites and the edges connecting them are links pointing from one page to another), or various social networks (where for instance the vertices represent people and edges represent acquaintances).

One motivation for studying these models is that one hopes to gain information about the spread of information or epidemics in such networks. In a paper that sparked off considerable interest (see e.g. [3, 2, 48]), Strogatz and Watts [88] observed that the networks mentioned above have several features in common, such as high local connectivity and short average distance between two vertices, but low overall density. They proposed a random model which captures some of these features and investigated this model using numerical simulations. In [61], we show how classical methods from the theory of random graphs can be applied to give some rigorous asymptotic results about this model.

Chapter 2

Summary of results

2.1 Random greedy H-free graphs

Chapter 4 is concerned with the study of a random greedy process, which, for any fixed graph H, produces a random H-free graph.

The process is defined as follows. To each edge of the complete graph K_n on n vertices assign a birthtime which is uniformly distributed in $[0, 1]$ and where the birthtimes of the edges are mutually independent. For $p = 0$ start with the empty graph on n vertices. Now increase p gradually. Each time a new edge is born, add it to the existing graph if this does not create a copy of H. Edges with equal birthtime (which occur with probability zero) are considered in arbitrary order. Denote the graph at time p by $M_{n,p}(H)$ and denote the final graph $M_{n,1}(H)$ by $M_n(H)$.

It is easily seen that $M_n(H)$ is a maximal H-free graph, i.e. an H-free graph to which no edge can be added without creating a copy of H. The question now arises what $M_n(H)$ looks like – in particular how many edges it is likely to have. Note that the above process can also be interpreted as a "constrained" evolution process. Due to the high dependence between the inclusion of different edges, it is harder to analyse than the "standard" $G_{n,m}$ model.

The above process was first studied by Ruciński and Wormald [79] for the case when the forbidden graph H is a star with $d + 1$ edges. They proved that in this case $M_n(H)$ almost surely contains at most one vertex of degree $d - 1$, and that all other vertices will have degree d. Thus if n is even, this provides an efficient way of producing a d-regular graph. Recently, Steger and Wormald [87] showed that a modification of the above process gives the same result even if d tends to infinity like some small power of n. The case when H is a triangle was studied by Erdős, Suen and Winkler [28] and Spencer [84].

Also, Kim [47] proved his celebrated lower bound on the Ramsey number $R(3, t)$ by studying a random greedy process (with forbidden triangles) similar to the one defined above. Furthermore, as mentioned in Section 1.2.1, random greedy processes have been proven to produce extremely efficient packings of small graphs in larger graphs.

Consider what might happen in the above process if for instance H is a complete graph K_s on s vertices. As for instance the complete $(s - 1)$-partite graph whose vertex classes have size as equal as possible (Fig. 2.1) shows, $M_n(K_s)$ might have as many as $\Theta(n^2)$ edges. On the other hand, there are families of maximal K_s-free graphs where the number of edges grows only linearly in the number of vertices. For example, consider the graphs obtained from a K_{s-2} by adding arbitrarily many new vertices which are adjacent to all vertices in the K_{s-2} but to none of the other new vertices (Fig. 2.1). However, it turns out that the number of edges of $M_n(K_s)$ is likely to be far away from either of these two extremes. For any graph G, we write $e(G)$ for the number of edges of G.

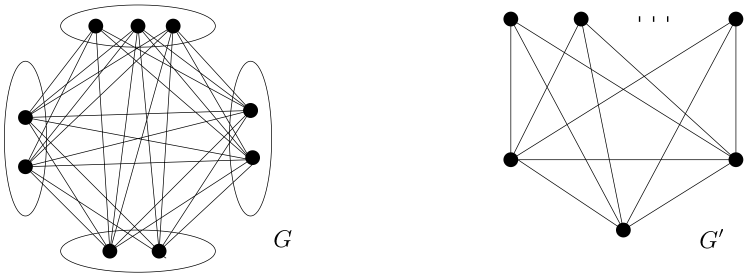

Figure 2.1: Two maximal K_5-free graphs: G shows that the final graph $M_n(K_5)$ produced by the random greedy process might have as many as $\Theta(n^2)$ edges, while G' shows that it might have only $\mathcal{O}(n)$ edges.

Theorem 2.1 *If H is a complete graph on s vertices, then almost surely*

$$cn^{2s/(s+1)} \leq e(M_n(H)) \leq Cn^{2s/(s+1)}(\log n)^{1/(s-2)},$$

for some constants c and C depending on s.

The case when $s = 3$ in Theorem 2.1 is already due (with slightly better bounds) to Erdős, Suen and Winkler [28] and Spencer [84]. The case when $s = 4$ (and the case when $\ell = 4$ in Theorem 2.2) were proven independently (with slightly weaker upper bounds) by Bollobás and Riordan [15]. Up to the logarithmic factor, the $s \geq 4$ cases solve a problem of Erdős, Suen, and Winkler (see [19]). For cycles, we obtain an analogous result.

Theorem 2.2 *If H is a cycle of length ℓ, then almost surely*

$$cn^{\ell/(\ell-1)} \leq e(M_n(H)) \leq Cn^{\ell/(\ell-1)} \log n,$$

for some constants c and C depending on ℓ.

In fact, we will prove an analogous result (see Theorems 4.1 and 4.2) for the more general case that H is strictly 2-balanced. This class of graphs will be defined in Chapter 4. It contains, for example, complete graphs, cycles, and hypercubes.

The common idea of the proofs is that as long as p is not too large, a random graph $G_{n,p}$ does not contain too many copies of H and thus $G_{n,p}$ and $M_{n,p}(H)$ will have similar properties. (Note that $G_{n,p}$ may be identified with the graph consisting of those edges whose birthtimes are at most p.) Beyond some critical point, the growth of $M_{n,p}(H)$ slows down very quickly. Quite naturally, the critical value of p turns out to be the point at which an edge in $G_{n,p}$ has a small but constant probability of being contained in a copy of H. We will prove our results by studying the random graph $G_{n,p}(H)$ obtained from $G_{n,p}$ by deleting *all* edges which lie in a copy of H, where p is suitably chosen (Fig. 2.2).

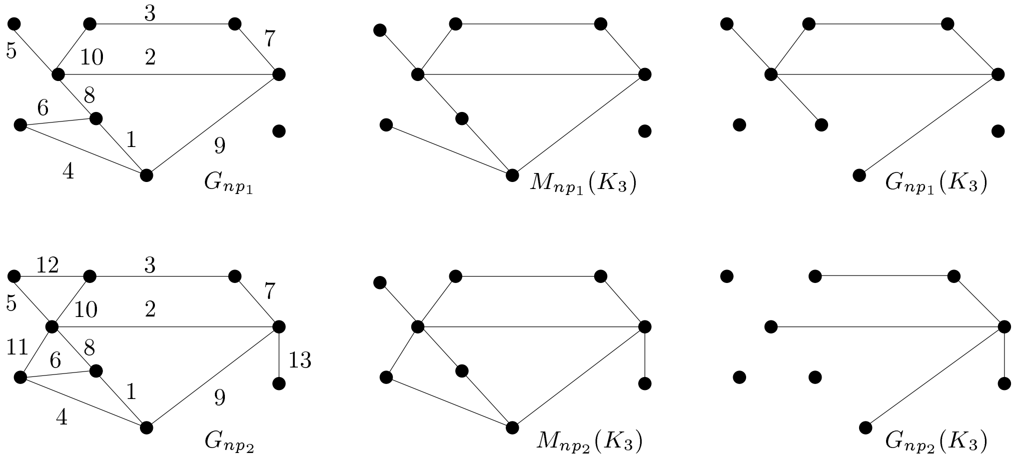

Figure 2.2: The edges of G_{n,p_i} are labelled according to the order of their birthtimes (where $p_1 < p_2$). Note that $G_{n,p_i} \supseteq M_{n,p_i}(K_3) \supseteq G_{n,p_i}(K_3)$.

The proofs build upon ideas introduced in Spencer [81] and Krivelevich [55]. As a simple application of the methods used in [55], we also obtain the following extremal result.

Theorem 2.3 *Given $\ell \in \mathbb{N}$ with $\ell \geq 3$, there is a graph G on n vertices with girth greater than ℓ and chromatic number*

$$\chi(G) \geq cn^{1/(\ell-1)}/\log n,$$

where c depends only on ℓ.

For $\ell = 3$, the logarithmic factor is improved by Kim [47], but for $\ell > 3$, the above theorem improves the best previously known exponent of $1/\ell$. The latter bound is contained in the original proof of the existence of graphs with high girth and high chromatic number due to Erdős [24]. Note that on the other hand, Kostochka (see e.g. Jensen and Toft [42]) proved that for fixed $\ell \geq 3$, the chromatic number of a graph on n vertices and girth greater than ℓ is $\mathcal{O}(n^{2/(\ell-1)})$.

2.2 The evolution of triangle-free graphs

As we have outlined in Section 1.1, the evolution of the random graphs $G_{n,m}$ and $G_{n,p}$ has been the subject of much study. Intimately connected to this problem, but much less understood, is the question of what happens if we restrict our attention to certain subclasses of graphs. In contrast to the random greedy approach outlined in the previous section, here and in Chapter 5 we consider the "uniform" model for the evolution of triangle-free graphs. We denote by $\mathcal{T}(n,m)$ the set of all triangle-free graphs with n vertices and m edges. Information about the evolution and typical structure of graphs in $\mathcal{T}(n,m)$ will also allow us to estimate $|\mathcal{T}(n,m)|$. Since by definition

$$\mathbb{P}[\, G_{n,m} \text{ is triangle-free}\,] = \frac{|\mathcal{T}(n,m)|}{|\mathcal{G}(n,m)|},$$

this in turn gives a more complete picture of the evolution of graphs in $\mathcal{G}(n,m)$.

To see what results one might expect, we first recall some fundamental results on triangle-free graphs. The first one was discovered by Mantel [60] as early as 1906. He showed that a triangle-free graph on n vertices can have at most $\lfloor n^2/4 \rfloor$ edges (and thus $\mathcal{T}(n,m)$ is empty for $m > \lfloor n^2/4 \rfloor$). His proof also shows that any triangle-free graph with $\lfloor n^2/4 \rfloor$ edges must be bipartite (i.e. its vertices can be partitioned into two sets, A and B, so that all edges have one endpoint in A and the other in B). An analogous result for forbidding complete graphs of arbitrary size was proven by Turán [89] in 1941. The above results mark the beginning of *extremal graph theory*, which by now has a vast literature (see also Section 1.2.1).

We now turn to random triangle-free graphs again. In 1976, Erdős, Kleitman and Rothschild [26] proved that *almost all triangle-free graphs are bipartite.*

Thus one might think that the same result holds also for $\mathcal{T}(n,m)$. Indeed, the following simple argument shows that this is certainly true if $m/n \to 0$. Erdős and Rényi proved that if $m/n \to 0$, almost surely $G_{n,m}$ is bipartite. Since every bipartite graph is triangle-free, the same holds for graphs in $\mathcal{T}(n,m)$ if $m/n \to 0$. However, the following result by Prömel and Steger [70] implies that this is not always the case.

Theorem 2.4 (Prömel and Steger [70]) *There exist constants c_1, c_2, and c_3 such that*

$$\mathbb{P}\,[\, G_{n,m} \textit{ is bipartite} \mid G_{n,m} \textit{ is } K_3\textit{-free}] \to \begin{cases} 1 & \textit{if} \quad m = o(n) \\ 0 & \begin{array}{l} \textit{if} \quad m \geq c_1\, n \\ \textit{and } m \leq c_2\, n^{3/2} \end{array} \\ 1 & \textit{if} \quad m \geq c_3\, n^{7/4} \log n. \end{cases}$$

Note that the above probability equals the proportion of triangle-free graphs which are bipartite. Prömel and Steger conjectured that the function involving the $n^{7/4}$ term could be replaced by a function growing like $n^{3/2+o(1)}$. Recently, Łuczak proved a related result, which implies that $n^{3/2}$ is the threshold for the property that almost all graphs in $\mathcal{T}(n,m)$ are "almost" bipartite.

Theorem 2.5 (Łuczak [59]) *Given $\delta > 0$, there exists a constant $C > 0$ so that almost all triangle-free graphs with n vertices and $m \geq Cn^{3/2}$ edges can be made bipartite by deleting at most δm edges.*

In Chapter 5, we give a complete answer to the above problem by proving the following sharp threshold for bipartiteness. Quite naturally, it turns out that this threshold is a little larger than that for “almost” bipartiteness. As in Theorem 2.4, the main result is the second 1-statement.

Theorem 2.6 *Let*

$$t_3 = t_3(n) = \frac{\sqrt{3}}{4}\, n^{3/2}\sqrt{\log n}.$$

Then for any $\varepsilon > 0$*,*

$$\mathbb{P}\,[\,G_{n,m} \text{ is bipartite} \mid G_{n,m} \text{ is } K_3\text{-free}\,] \to \begin{cases} 1 & \text{if} \quad m = o(n) \\ 0 & \begin{array}{l}\text{if} \quad m \geq n/2 \\ \text{and } m \leq (1-\varepsilon)\, t_3\end{array} \\ 1 & \text{if} \quad m \geq (1+\varepsilon)\, t_3. \end{cases}$$

Independently, using methods based on those in [70], Steger [86] proved that the second threshold for bipartiteness is $\Theta(t_3)$. The above results are summarized in Figure 2.3.

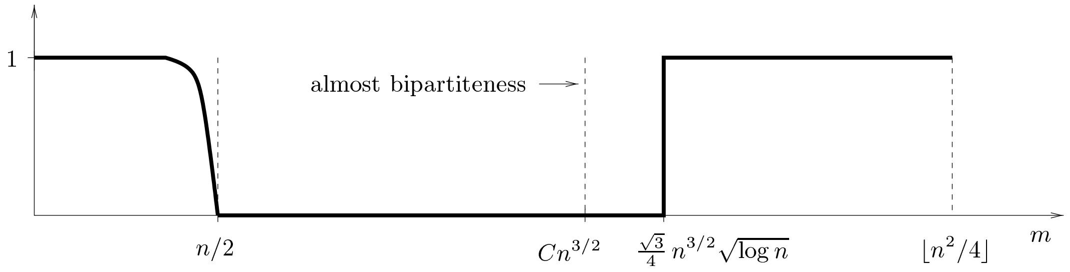

Figure 2.3: The proportion of triangle-free graphs with n vertices and m edges which are bipartite as $n \to \infty$

We emphasize that while the second threshold for bipartiteness is sharp, the first one is coarse – the condition that $m = o(n)$ cannot be weakened. The techniques used in the proof of the second 1-statement of Theorem 2.6 (which are based on large deviation inequalities) are quite different from those used in the proof of the second 1-statement of Theorem 2.4 (which is proved by an extension of the Kleitman-Rothschild method) and those used in the proof of Theorem 2.5 (which are based on a sparse version of the Regularity Lemma of Szemerédi). Our techniques can also be modified to obtain analogous results for odd cycles. The precise statement is deferred to Chapter 5.

As indicated earlier, information about the typical structure of triangle-free graphs allows us to determine asymptotically the number of graphs in $\mathcal{T}(n,m)$ for a large range of m. The following theorem (whose proof is based on Corollaries 2.9 and 2.10 of Prömel and Steger [68]) provides a precise asymptotic estimate for the number of bipartite graphs with sufficiently many edges. Together with Theorem 2.6, this gives us the asymptotic for the number of graphs in $\mathcal{T}(n,m)$ for $m \geq (1+\varepsilon)t_3$, which improves on the bounds implied by Theorem 2.5 for these m.

Theorem 2.7 *Almost all bipartite graphs with* n *vertices and* $m \geq 20n \log n$ *edges have a bipartition where the size of the vertex classes differs by at most* $n \log n/\sqrt{m}$. *Moreover, if* $m/n^2 \to 0$, *the number of such graphs is*

$$(1+o(1))\frac{\sqrt{\pi}}{4}\frac{n}{\sqrt{m}}\binom{n}{\lfloor n/2 \rfloor}\binom{\lfloor n^2/4 \rfloor}{m}.$$

If $m/n^2 \not\to 0$, *the number of such graphs is*

$$\Theta\left(\binom{n}{\lfloor n/2 \rfloor}\binom{\lfloor n^2/4 \rfloor}{m}\right).$$

For $m = o(n^{4/3})$, the best bounds on $|\mathcal{T}(n,m)|$ are due to Wormald [95]. If $m/n^{4/3} \not\to 0$ and if m is not large enough for Theorem 2.5 to apply, the best bounds are those which follow from the results in Janson, Łuczak and Ruciński [40].

2.3 Random subsets of a finite set

Finally, in Chapters 6 and 7 we turn from graphs to finite sets. While for a graph, basic parameters to consider are for instance its chromatic number and its independence number, the basic parameters of (partially ordered) finite sets are different – we will consider the width and the length, which are defined below. By $\mathcal{P}(n)$ we denote the set of all subsets of the integers $\{1, \dots, n\}$ and we say that $\mathcal{A} \subseteq \mathcal{P}(n)$ is an *antichain* if $\mathcal{A}$ contains no distinct elements x and y so that $x \subseteq y$. For any set $\mathcal{Q} \subseteq \mathcal{P}(n)$, we say that the *width* of $\mathcal{Q}$, denoted by $\operatorname{width} \mathcal{Q}$, is the size of the largest antichain contained in $\mathcal{Q}$. The following theorem, proved by Sperner in 1928, is probably the most well-known result in extremal set theory (see e.g. the book by Engel [22]).

Theorem 2.8 (Sperner [85])

$$\operatorname{width} \mathcal{P}(n) = \binom{n}{\lfloor n/2 \rfloor}.$$

Moreover, this bound is achieved by the antichain containing all those elements of $\mathcal{P}(n)$ *which have cardinality* $\lfloor n/2 \rfloor$.

Since for all $\mathcal{Q} \subseteq \mathcal{P}(n)$, the width of $\mathcal{Q}$ is at most that of $\mathcal{P}(n)$, Theorem 2.8 gives an upper bound on the width of any subset of $\mathcal{P}(n)$. But what happens if we consider "typical" (i.e. random) subsets of $\mathcal{P}(n)$? We obtain such a random subset $\mathcal{P}(n,p)$ by selecting each element of $\mathcal{P}(n)$ with probability p independently of all other elements. If we order the elements of $\mathcal{P}(n,p)$ by inclusion, we may regard $\mathcal{P}(n,p)$ as a random partially ordered set. This model was first considered by Rényi [75] in 1961, who, answering a question of Erdős, obtained the threshold for the property that $\mathcal{P}(n,p)$ is not an antichain itself. In Chapters 6 and 7, we will discuss further related results on $\mathcal{P}(n,p)$ by several authors, including Kreuter [54] and Kohayakawa and Kreuter [50]. In Chapter 6 we will also discuss other models of random partially ordered sets (see e.g. Brightwell [17] and Prömel, Steger, and Taraz [72]). In Chapter 7 we will prove that if p is sufficiently large, then the following probabilistic analogue of Sperner's theorem holds:

Theorem 2.9 *If $pn/\log n \to \infty$, then almost surely we have*

$$\operatorname{width} \mathcal{P}(n,p) = (1+o(1))p\binom{n}{\lfloor n/2 \rfloor}.$$

Moreover, almost surely this bound is achieved by the antichain containing all those elements of $\mathcal{P}(n,p)$ which have cardinality $\lfloor n/2 \rfloor$.

Up to the logarithmic factor, this answers a question of Kohayakawa and Kreuter [50] – in Section 7.1 we will show that the assertion of Theorem 2.9 does not hold if $pn \not\to \infty$. In fact (see Theorem 7.1), we give asymptotic bounds on the width of $\mathcal{P}(n,p)$ for the entire range of p. Surprisingly, it turns out that the key to proving these results is to investigate the structure and length of *chains* in $\mathcal{P}(n,p)$. Here a chain is a subset $\mathcal{C}$ of $\mathcal{P}(n)$ with the property that for all pairs of elements x and y of $\mathcal{C}$, we either have $x \subseteq y$ or $y \subseteq x$. The *length* of a chain is the number of its elements minus one and the *length* of a set $\mathcal{Q} \subseteq \mathcal{P}(n)$, denoted by length $\mathcal{Q}$, is equal to the length of the longest chain contained in $\mathcal{Q}$. Thus one immediately sees that the length of $\mathcal{P}(n)$ is equal to n. Rényi's result, mentioned above, then gives the threshold for the appearance of a chain of length one in $\mathcal{P}(n,p)$. Combined with the results of Kreuter [54], we obtain bounds on the length of $\mathcal{P}(n,p)$ for the entire range of p.

Let us give a very rough sketch of the evolution of $\mathcal{P}(n,p)$. For very small p, $\mathcal{P}(n,p)$ is an antichain. As p increases, so do the length and width of $\mathcal{P}(n,p)$. However, $\mathcal{P}(n,p)$ "looks" less and less like an antichain – the ratio of the width of $\mathcal{P}(n,p)$ to the number of elements in $\mathcal{P}(n,p)$ is steadily decreasing. Indeed, shortly after the appearance of a chain of length k in $\mathcal{P}(n,p)$, almost surely the width of $\mathcal{P}(n,p)$ will be not much more than the number of elements of $\mathcal{P}(n,p)$ whose cardinality is between $(1-1/k)n/2$ and $(1+1/k)n/2$, i.e. the n/k levels of $\mathcal{P}(n,p)$ which lie "in the middle". A (nearly) largest antichain can be obtained by choosing nearly all elements of $\mathcal{P}(n,p)$ contained in these n/k levels (Fig. 2.4).

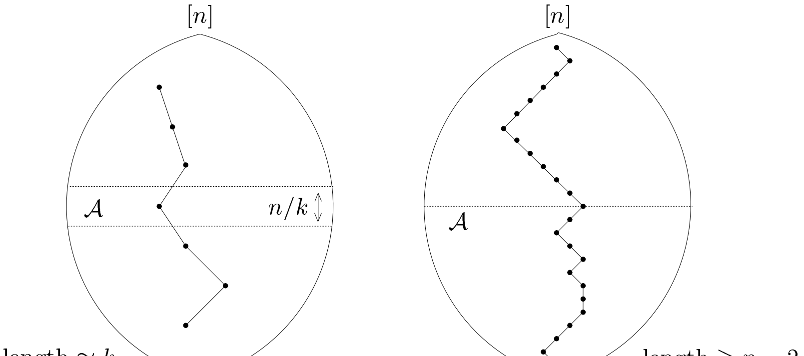

Figure 2.4: A nearly largest antichain $\mathcal{A}$ can almost surely be obtained by choosing nearly all elements of $\mathcal{P}(n,p)$ in the middle region.

The following theorem gives a detailed picture of the final stage of the evolution of the length of $\mathcal{P}(n,p)$:

Theorem 2.10 *Let* $c \geq \mathrm{e}$ *be a given constant. Define* $p(n) = c/n$*. Let* $\eta = \eta(c)$ *satisfy* $\eta = \mathrm{e}^{c(\eta-1)}$ *and* $0 < \eta < 1$*. Then*

$$\begin{aligned}\mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p)=n-2\,] &\to (1-\eta)^2,\\ \mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p)=n-3\,] &\to 2\eta(1-\eta),\\ \mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p)=n-4\,] &\to \eta^2.\end{aligned}$$

We remark that η is the extinction probability of a branching process whose family size has Poisson distribution with parameter c, and thus $\eta \to 0$ as $c \to \infty$. Also, note that since $\mathcal{P}(n)$ contains only a single element of cardinality zero and of cardinality n, almost surely we have that $\mathrm{length}\,\mathcal{P}(n,p) < n-1$ as long as $p \to 0$.

Using the idea of double-counting chains – first introduced in the elegant proof of Sperner's theorem due to Lubell [57] – we can then apply a technical lemma on the typical structure of chains in $\mathcal{P}(n,p)$, used in the proof of Theorem 2.10, as the main tool in the proof of Theorem 2.9. It seems rather surprising that even for the case when $p = 1/2$, no argument which would yield a direct proof of Theorem 2.9 is known.

We remark that while – as in $G_{n,p}$ – the inclusion of elements in $\mathcal{P}(n,p)$ is independent, what makes it more difficult to study $\mathcal{P}(n,p)$ than $G_{n,p}$ is its "nonhomogeneity": for instance, it turns out that the length of $\mathcal{P}(n,p)$ is influenced to a large extent by those elements in $\mathcal{P}(n,p)$ whose cardinalities are very small or very large, while to give bounds on the width of $\mathcal{P}(n,p)$, it suffices to consider those elements whose cardinalities are not too large and not too small.

Chapter 3

Inequalities

In this chapter, we introduce some well-known large deviation inequalities. We will make use of them several times in the subsequent chapters in order to show that the probability that the value of some random variable X is far from its expected value $\mathbb{E}[X]$ is small. In fact, a common feature of the proofs of the main results of Chapters 4, 5 and 7 is that the inequalities in Section 3.2 below (or some related inequalities) are essential ingredients. These related inequalities will be introduced at the appropriate places in the proofs (see Sections 4.2 and 7.2).

For later reference, we also state two simple estimates for binomial coefficients. Suppose that $j \leq j+r < i$. Then

$$\mathrm{e}^{-jr/(i-j-r)} \leq \left(1 - \frac{j}{i-r}\right)^r \leq \binom{i-r}{j} \Big/ \binom{i}{j} \leq \mathrm{e}^{-jr/i}. \tag{3.1}$$

The proof of (3.1) is elementary. The following crude estimate follows from Stirling's formula (see page 4 of [8]). For all $j \leq i$, we have

$$\binom{i}{j} \leq \left(\frac{\mathrm{e}i}{j}\right)^j. \tag{3.2}$$

3.1 Chernoff bounds – independent summands

Suppose that $X = \sum_{i=1}^N X_i$, where the X_i are independent 0-1-variables with $\mathbb{P}[X_i = 1] = p'$ and $\mathbb{P}[X_i = 0] = 1 - p'$ for all i. In other words, X is binomially distributed with parameters N and p', and its expected value is $\mathbb{E}[X] = Np'$. Then for any $\delta > 0$,

$$\mathbb{P}[X \leq (1-\delta)\mathbb{E}[X]] \leq \mathrm{e}^{-\delta^2 \mathbb{E}[X]/2}; \tag{3.3}$$

$$\mathbb{P}[X \geq (1+\delta)\mathbb{E}[X]] \leq \left(\mathrm{e}^{\delta}(1+\delta)^{-(1+\delta)}\right)^{\mathbb{E}[X]}. \tag{3.4}$$

Proofs can be found (for instance) in [6] and [41]. Here and elsewhere, we write e for the Euler number. All logarithms are base e.

3.2 Summands with limited dependence

The following large deviation inequalities for the case when X is the sum of dependent summands are due to Janson, Łuczak, and Ruciński [40] (see also [39] or [6]).

Let $\{I_i\}_{i\in\mathcal{J}}$ be independent 0-1-variables where $\mathcal{J}$ is an arbitrary index set. For every subset α of $\mathcal{J}$, let $I_\alpha = \prod_{i\in\alpha} I_i$. Let $\mathcal{S}$ be a collection of subsets of $\mathcal{J}$ and set

$$X = X_{\mathcal{S}} = \sum_{\alpha\in\mathcal{S}} I_\alpha.$$

Furthermore set $\mu = \mathbb{E}[X]$ and let

$$\Delta = \Delta(\mathcal{S}) = \sum_{\alpha\sim\beta} \mathbb{E}[I_\alpha I_\beta],$$

where the sum is over all ordered pairs (α, β) of elements of $\mathcal{S}$ so that $\alpha\cap\beta \neq \emptyset$ and $\alpha \neq \beta$. Then

$$\mathbb{P}[X = 0] \leq \exp\left\{-\mu + \frac{\Delta}{2}\right\}. \tag{3.5}$$

If $\Delta \geq \mu$, then we have

$$\mathbb{P}[X = 0] \leq \exp\left\{-\frac{\mu^2}{\mu+\Delta}\right\} \leq \exp\left\{-\frac{\mu^2}{2\Delta}\right\}. \tag{3.6}$$

Chapter 4

Random greedy H-free graphs

4.1 Results

The main result of this chapter is Theorem 4.2, which gives an upper bound on the number of edges in the graph $M_n(H)$ when H is strictly 2-balanced. Recall from Section 2.1 that $M_n(H)$ is the graph produced by the following random greedy process: To each edge of the complete graph K_n on n vertices assign a birthtime which is uniformly distributed in $[0, 1]$ and where the birthtimes of the edges are mutually independent. For $p = 0$ start with the empty graph on n vertices. Now increase p gradually. Each time a new edge is born, add it to the existing graph if this does not create a copy of H. Edges with equal birthtime (which occur with probability zero) are considered in arbitrary order. Denote the graph at time p by $M_{n,p}(H)$ and denote the final graph $M_{n,1}(H)$ by $M_n(H)$.

To be able to state our results, we first recall the following standard definitions. Given a graph G, $e(G)$ denotes the number of edges in G and $v(G)$ denotes the number of vertices. We say that G is *strictly* 2*-balanced* if $v(G) \geq 3$ and $e(G) \geq 3$ and if for all proper subgraphs G' of G with $v(G') \geq 3$ we have

$$\frac{e(G) - 1}{v(G) - 2} > \frac{e(G') - 1}{v(G') - 2}.$$

Furthermore, we say that a graph G is *balanced* if $e(G)/v(G) \geq e(G')/v(G')$ for all subgraphs G' of G. We remark that strictly 2-balanced graphs are balanced, and it is easily seen that cycles, complete graphs, and the complete r-partite graph $K_{t,\ldots,t}$ are strictly 2-balanced, whereas trees and disconnected graphs are not. Another example of a strictly 2-balanced graph is the d-dimensional cube (for $d \geq 2$). This fact may easily be derived from a lemma proven in Chung et al. [18], which states that every subgraph of the d-dimensional cube with v' vertices has at most $(v' \log_2 v')/2$ edges.

Throughout this chapter, H denotes the forbidden graph and if there is no danger of confusion, we will write $e = e(H)$ and $v = v(H)$. For any graph G, let $\Delta(G)$ denote the maximum degree of G.

The following lower bound on the number of edges in $M_n(H)$ is easily proved, while the main result is the upper bound.

Theorem 4.1 *Suppose that H is balanced. Then there exists a constant $c = c(H)$ so that almost surely $M_n(H)$ has average degree at least*

$$cn^{1-\frac{v-2}{e-1}}.$$

Theorem 4.2 *Suppose that H is strictly 2-balanced. Then there exists a constant $C = C(H)$ so that almost surely $M_n(H)$ has maximum degree at most*

$$Cn^{1-\frac{v-2}{e-1}} (\log n)^{1/(\Delta(H)-1)}.$$

The proofs of Theorems 4.1 and 4.2 easily generalize to forbidding a bounded number of strictly 2-balanced graphs. If the family of forbidden graphs is $\mathcal{H}$, we define the random maximal $\mathcal{H}$-free graph $M_n(\mathcal{H})$ analogously to $M_n(H)$. We state the following theorem separately, as we shall give only a sketch of its proof.

Theorem 4.3 *Given $s \in \mathbb{N}$ and a family of strictly 2-balanced graphs $\mathcal{H} = \{H_0, H_1, \ldots, H_{s-1}\}$ which is ordered so that*

$$\frac{e(H_0)-1}{v(H_0)-2} \leq \frac{e(H_1)-1}{v(H_1)-2} \leq \cdots \leq \frac{e(H_{s-1})-1}{v(H_{s-1})-2},$$

then almost surely $M_n(\mathcal{H})$ has average degree at least

$$cn^{1-\frac{v(H_0)-2}{e(H_0)-1}}$$

and maximum degree at most

$$Cn^{1-\frac{v(H_0)-2}{e(H_0)-1}} (\log n)^{1/(\Delta(H_0)-1)}$$

for some constants c and C depending on $\mathcal{H}$.

As a special case this answers (up to the factor of $\log n$ again) a problem of Erdős, Suen, and Winkler listed in [19], namely:

Corollary 4.4 *Let $\mathcal{H}$ be the set of all cycles of length up to ℓ. Then almost surely $M_n(\mathcal{H})$ has average degree at least $cn^{1/(\ell-1)}$ and maximum degree at most $Cn^{1/(\ell-1)} \log n$ for some constants c and C depending on ℓ.*

It is easily seen that the proofs yield error bounds which decay like $\mathcal{O}(n^{-c})$ for any fixed constant c. Also, since birthtimes are distinct with probability one, we also obtain corresponding results for the processes where instead of random birthtimes we are given a random permutation of the edge set of the K_n.

We have made no attempt to optimize the constants involved in our results, as we believe that none of them give the correct order of magnitude. Indeed, Kim's proof of the lower bound on $R(3,t)$ in [47] shows that the average degree of the final random maximal triangle-free graph in a similar process is almost surely of order $\sqrt{n \log n}$. The result of Kim is obtained through an extremely sophisticated application of the so-called "nibble method" introduced by Rödl. It is likely that these techniques could be applied to improve our results, but in view of the complexity of Kim's proof, closing the gap between the upper and lower bounds even for triangles seems to be a formidable task indeed. More generally, concerning cycles, we believe that in fact, the number of edges of $M_n(C_\ell)$ may be almost surely of order $n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)}$. Note that for odd ℓ, in the next chapter we will prove that this is exactly the (second) threshold for bipartiteness in C_ℓ-free graphs.

Note that if one interprets the birthtimes of the edges as edge weights (which are uniformly and independently distributed in $[0, 1]$), then the random greedy algorithm which forbids the set of *all* cycles produces a minimum spanning tree of the complete graph K_n. The (almost sure) weight of this tree was determined by Frieze [32], and other properties of this tree were considered by Aldous [4].

Finally, it is worth pointing out that Erdős, Suen and Winkler [28] proved that if one forbids the set of *all odd* cycles, then the graph obtained from the corresponding random greedy process almost surely has $\Theta(n^2)$ edges. On the other hand, Theorem 4.3 implies that if we forbid only all odd cycles up to length ℓ (where ℓ is fixed and odd) the resulting graph almost surely has far fewer edges – namely $n^{1+1/(\ell-1)+o(1)}$.

This chapter is organized as follows. In Section 4.2, we introduce some large deviation inequalities which we shall need in the proofs. In Section 4.3, we prove Theorem 2.3. In Section 4.4, we prove Theorem 4.1. In Section 4.5, we illustrate the ideas and techniques involved in the proof of Theorem 4.2 by proving the special case when the forbidden graph is a K_4. In Section 4.6, we then prove Theorem 4.2. In Section 4.7 we show how the proof of Theorem 4.2 may be modified to prove Theorem 4.3. In the final section, we discuss the open problem of forbidding an arbitrary graph H.

4.2 Tools and notation

Let $G_{n,p}$ be the graph which has as its edges all those edges whose birthtime is at most p and let $G_{n,p}(H)$ be the subgraph of $G_{n,p}$ that is obtained by deleting *all* edges of $G_{n,p}$ that lie in a copy of H. We emphasize the immediate but important fact that the edge set of $G_{n,p}(H)$ is contained in that of $M_{n,p}(H)$. Thus

$$G_{n,p}(H) \subseteq M_{n,p}(H) \subseteq G_{n,p}. \tag{4.1}$$

For a graph G, the vertex set of G is denoted by $V(G)$. Furthermore we shall omit floors and ceilings throughout, as these will not affect the proof. Also, we shall always assume that n is sufficiently large for our estimates to hold.

We shall make use of two bounds on the number of disjoint subgraphs of a random graph. A slightly more general form of these inequalities, together with their proofs, may be found in Chapter 8 of [6].

Let $\mathcal{S}$ be a set of subsets of the edge set of the complete graph K_n on n vertices and for $\alpha \in \mathcal{S}$ let I_α be the corresponding indicator variable which equals one if and only if the edges of α are contained in the random graph $G_{n,p}$. Let $X = \sum_{\alpha \in \mathcal{S}} I_\alpha$ and $\mu = \mathbb{E}[X]$. As in Section 3.2, define

$$\Delta = \sum_{\alpha \sim \alpha'} \mathbb{E}[I_\alpha I_{\alpha'}],$$

where the summation is over all ordered pairs $\alpha \sim \alpha'$ in $\mathcal{S}$.

Let Y denote the size of a largest edge disjoint family of elements of $\mathcal{S}$ in $G_{n,p}$. Erdős and Tetali [29] proved the following simple but very useful upper tail bound on Y. (It is also proved in Chapter 8 of [6] and is a special case of Lemma 2.46 in [41].) For every $k \in \mathbb{N}$,

$$\mathbb{P}[Y \geq k] \leq \left(\frac{\mathrm{e}\mu}{k}\right)^k. \tag{4.2}$$

We will also need a lower tail bound on Y. Let $\eta = \max_{\alpha \in \mathcal{S}} \mathbb{E}[I_\alpha]$ and

$$\nu = \max_{\alpha \in \mathcal{S}} \sum_{\alpha' \sim \alpha} \mathbb{E}[I_{\alpha'}].$$

Then Lemma 4.2 in Chapter 8 of [6] states that for any $s \in \mathbb{N}$,

$$\mathbb{P}[Y = s] \leq \frac{\mu^s}{s!} \exp\left\{-\mu + s\nu + \frac{\Delta}{2(1-\eta)}\right\}. \tag{4.3}$$

Thus for every $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{P}[Y \leq (1-\varepsilon)\mu] &\leq \exp\left\{\frac{\Delta}{2(1-\eta)}\right\} \sum_{s \leq (1-\varepsilon)\mu} \frac{\mu^s}{s!}\mathrm{e}^{-\mu+s\nu} \\
&\leq \exp\left\{\frac{\Delta}{2(1-\eta)} + (1-\varepsilon)\mu\nu\right\} \sum_{s \leq (1-\varepsilon)\mu} \frac{\mu^s}{s!}\mathrm{e}^{-\mu} \\
&= \exp\left\{\frac{\Delta}{2(1-\eta)} + (1-\varepsilon)\mu\nu\right\} \mathbb{P}[P \leq (1-\varepsilon)\mu],
\end{aligned}$$

where P is a Poisson random variable with mean μ. But Theorem A.15 in [6] states that

$$\mathbb{P}[P \leq (1-\varepsilon)\mu] \leq \mathrm{e}^{-\varepsilon^2\mu/2}$$

which implies that

$$\mathbb{P}[Y \leq (1-\varepsilon)\mu] \leq \exp\left\{(1-\varepsilon)\mu\nu + \frac{\Delta}{2(1-\eta)} - \varepsilon^2\mu/2\right\}. \tag{4.4}$$

The above derivation of (4.4) from (4.3) is taken from [55], we have included it here for completeness.

Alternatively, we could have applied the following inequality, which follows from Talagrand's inequality using the same argument as in the proof of inequality (2.43) in Janson, Łuczak, and Ruciński [41]. For $t \geq 0$ and some constant γ depending only on the maximum number of edges of an element in $\mathcal{S}$, we have

$$\mathbb{P}[|Y - \mathbb{E}[Y]| \geq t] \leq 4\mathrm{e}^{-\gamma t^2/(\mathbb{E}[Y]+t)}.$$

4.3 Graphs with large girth

The idea of the proof is to show that deleting some maximal collection of edge-disjoint cycles from $G_{n,p}$, where p is suitably chosen, yields a graph with no large independent sets. An analogous idea was used by Krivelevich [55] to give lower bounds on Ramsey numbers.

Proof of Theorem 2.3. Let $p = c_0 n^{-\beta}$, where $\beta = (\ell-2)/(\ell-1)$. Let $G'_{n,p}$ be the graph obtained from $G_{n,p}$ by deleting some maximal collection of edge-disjoint cycles of length at most ℓ. Thus $G'_{n,p}$ has girth greater than ℓ. The theorem will follow if we can show that almost surely $G'_{n,p}$ has no independent set T of size t, where $t = |T| = Cn^{\beta}\log n$, where C is some sufficiently large constant depending on ℓ (see below). So for given T, let Y denote the number of edges in $G_{n,p}$ with both endpoints in T. Note that Y is binomially distributed with mean $\mu = \mathbb{E}[Y] = \binom{t}{2}p$. Hence a simple Chernoff tail estimate (e.g. Theorem A.13 in [6] with $a = \mathbb{E}[Y]/2$ or Theorem 2.1 in [41] with $t = \mathbb{E}[Y]/2$) yields

$$\mathbb{P}[Y \le \mu/2] \le \mathrm{e}^{-\mu/8}.$$

Let X' be the number of cycles in $G_{n,p}$ of length at most ℓ with at least one edge in T and let Y' be the maximum number of edge-disjoint cycles in $G_{n,p}$ of length at most ℓ with at least one edge in T. Then the number of edges of $G'_{n,p}$ which lie in T is at least $Y - \ell Y'$, so we are done if we can find a sufficiently sharp upper tail bound on Y'. If we choose the constant c_0 sufficiently small, then whatever the value of C we have that

$$\mathbb{E}[X'] \le \sum_{j=3}^{\ell} t^2 n^{j-2} p^j \le \ell\, t^2 n^{\ell-2} p^{\ell} \le \mu/(\mathrm{e}\, 4\ell\, \mathrm{e}^{\ell/2}).$$

Now denote by $\mathcal{S}$ the set of all cycles of length at most ℓ with at least one edge in T and apply inequality (4.2) to see that

$$\mathbb{P}[Y' \ge \mu/(4\ell)] \le \left(\frac{\mathrm{e}\, 4\ell\, \mathbb{E}[X']}{\mu}\right)^{\mu/(4\ell)} \le \mathrm{e}^{-\mu/8}.$$

Putting the above together, the probability that $G'_{n,p}$ has an independent set of size t is at most

$$\binom{n}{t} 2\mathrm{e}^{-\mu/8} \le 2\mathrm{e}^{t\log n - \mu/8} = 2\mathrm{e}^{t(\log n - (t-1)p/16)},$$

which tends to zero provided that C is sufficiently large. □

4.4 Theorem 4.1 – the lower bound

The lower bound is an immediate consequence of the following well-known result of Erdős and Rényi (see Theorem 4.2 in Chapter 4 of [6] or Theorem 3.4 and Remarks 3.7 and 3.8 in [41]) which we state here in a form convenient to us.

Theorem 4.5 (Erdős and Rényi) *Let H be a balanced graph and let X denote the number of copies of H in $G_{n,p}$. If $pn^{v(H)/e(H)} \to \infty$, then almost surely $X \leq 2\mathbb{E}[X]$.*

Proof of Theorem 4.1. Let $p = c_0 n^{-(v-2)/(e-1)}$, where c_0 satisfies $c_0^{e-1}\, e = 1/8$, which is certainly large enough to be able to apply Theorem 4.5. Clearly, $e(G_{n,p}) - eX$ is a lower bound for the number of edges in $G_{n,p}(H)$ and thus also in $M_{n,p}(H)$ or $M_n(H)$. But, crudely,

$$e\,\mathbb{E}[X] \leq e\, n^v p^e \leq e\, n^2 p\, c_0^{e-1} = pn^2/8,$$

and the result follows. □

4.5 Forbidding a K_4 – the upper bound

To illustrate the ideas and techniques involved in the proof of Theorem 4.2, we now give a detailed proof of Theorem 4.2 for the case when the forbidden graph H is K_4. Let $p = n^{-2/5}/4$ and let

$$t = 10^4 n^{3/5}\sqrt{\log n}. \tag{4.5}$$

Suppose that we have a vertex y in the final random greedy graph $M_n(K_4)$ whose degree is at least t. Consider any set T of vertices which are adjacent to y in $M_n(K_4)$. Since $M_n(K_4)$ contains no copy of K_4 and since $M_{n,p}(K_4) \subseteq M_n(K_4)$, the graph $M_{n,p}(K_4)$ constructed by the random greedy process up to time p contains no copy of a K_4, the subgraph of $M_{n,p}(K_4)$ induced by T is triangle-free. By (4.1), it follows that the subgraph of $G_{n,p}(K_4)$ induced by T is also triangle-free. Thus it suffices to prove that almost surely , for all $T \subseteq V(K_n)$ of size t, $G_{n,p}(K_4)[T]$ contains a triangle – since this would imply that almost surely , $M_n(K_4)$ has maximum degree less than t, as stated in Theorem 4.2.

Given an edge f in K_n and $k \in \mathbb{N}$ we say that a sequence $H_1, \ldots, H_k$ of copies of K_4 forms a *(k, f)-cluster* in K_n if they all contain f and have the property that for all i with $1 \leq i \leq k$, H_i contains an edge which is not contained in any of the other H_j with $j < i$. Let $\mathcal{A}$ denote the event that $G_{n,p}$ does not contain a $(\log n, f)$-cluster for any edge f. Below, we will prove that $\mathbb{P}[\mathcal{A}] \to 1$. Also, we will prove that for any fixed $T \subseteq V(K_n)$ with $|T| = t$, we have

$$\mathbb{P}[\{G_{n,p}(K_4)[T] \text{ is } K_3\text{-free}\} \cap \mathcal{A}] \leq n^{-2t}. \tag{4.6}$$

These two results imply that the probability that there exists a set of t vertices which is triangle-free in $G_{n,p}(K_4)$ is at most

$$\mathbb{P}[\mathcal{A}^c] + \binom{n}{t}\mathbb{P}[\{G_{n,p}(K_4)[T] \text{ is } K_3\text{-free}\} \cap \mathcal{A}] \leq o(1) + n^t\, n^{-2t} = o(1),$$

which proves the theorem.

It remains to prove that $\mathbb{P}[\mathcal{A}] \to 1$ and (4.6). First we consider $\mathbb{P}[\mathcal{A}]$. For a fixed edge f in K_n and $k \in \mathbb{N}$ with $k \leq \log n$, let $Z_{k,f}$ denote the number of (k,f)-clusters in $G_{n,p}$. Clearly, a $(1,f)$-cluster is a copy of a K_4 containing f, and thus

$$\mathbb{E}[Z_{1,f}] \leq n^2 p^6 = 4^{-5} p \leq \mathrm{e}^{-4}. \tag{4.7}$$

Note that every $(k+1,f)$-cluster can be viewed as the union of a (k,f)-cluster $C_{k,f}$ and a $(1,f)$-cluster which is not contained in $C_{k,f}$. Fix some (k,f)-cluster $C_{k,f}$ and let $Z_{1,f,C_{k,f}}$ denote the number of those $(1,f)$-clusters in $G_{n,p}$ which are not contained in $C_{k,f}$. By separately considering those $(1,f)$-clusters having zero, one, or two vertices which are not contained in $C_{k,f}$, we have (recalling $k \leq \log n$)

$$\mathbb{E}[Z_{1,f,C_{k,f}} \mid C_{k,f} \subseteq G_{n,p}] \leq \mathcal{O}(k^2 p) + \mathcal{O}(knp^3) + n^2 p^5 = o(1) + 4^{-5} \leq \mathrm{e}^{-4}.$$

Thus, if we let $\sum_{C_{k,f}}$ denote the sum over all (k,f)-clusters in K_n, we have

$$\begin{aligned}\mathbb{E}[Z_{k+1,f}] &\leq \sum_{C_{k,f}} \sum_{C_{1,f} \not\subseteq C_{k,f}} \mathbb{P}[\{C_{k,f} \subseteq G_{n,p}\} \cap \{C_{1,f} \subseteq G_{n,p}\}] \\ &= \sum_{C_{k,f}} \mathbb{P}[C_{k,f} \subseteq G_{n,p}]\, \mathbb{E}[Z_{1,f,C_{k,f}} \mid C_{k,f} \subseteq G_{n,p}] \\ &\leq \sum_{C_{k,f}} \mathbb{P}[C_{k,f} \subseteq G_{n,p}]\, \mathrm{e}^{-4} = \mathbb{E}[Z_{k,f}]\, \mathrm{e}^{-4}.\end{aligned}$$

Together with (4.7) this shows that $\mathbb{E}[Z_{k,f}] \leq \mathrm{e}^{-4k}$ and thus that $\mathbb{E}[Z_{\log n,f}] \leq n^{-4}$. Summing over all edges of the complete graph, we have that

$$\mathbb{P}[\mathcal{A}^c] \leq \binom{n}{2} n^{-4} = o(1).$$

Now we prove (4.6), which is done in two steps. In the first step, we show that p is sufficiently large to ensure that with sufficiently high probability, T contains a large set of *edge-disjoint* triangles in $G_{n,p}$. In the second step, we will then show that p is sufficiently small to ensure that if $\mathcal{A}$ holds, then with sufficiently high probability, at least one of these triangles is also contained in $G_{n,p}(K_4)[T]$.

Turning to the first step, let X denote the number of triangles in $G_{n,p}[T]$ and let $\mu = \mathbb{E}[X]$. Note that

$$\mu = \binom{t}{3} p^3 \overset{(4.5)}{=} \frac{(1+o(1))t}{6} \left(10^4 n^{3/5} \sqrt{\log n}\right)^2 \left(n^{-2/5}/4\right)^3 \geq 10^5 t \log n. \tag{4.8}$$

Let Y denote the size of a largest family of edge-disjoint triangles in $G_{n,p}[T]$. We want to use (4.4) to give a lower bound on Y. Thus, let $\mathcal{S}$ be the set of triples of edges forming a triangle in $K_n[T]$. Δ, as defined in Section 4.2, is then the expected number of ordered pairs of triangles in $G_{n,p}[T]$ having exactly an edge in common. Thus

$$\Delta = \mathcal{O}\left(t^4 p^5\right) = \mathcal{O}\left(\mu t p^2\right) \leq \mu n^{-1/5+o(1)} = o(\mu).$$

Also, ν is the expected number of triangles in $G_{n,p}[T]$ containing an edge of a fixed triangle and thus $\nu \leq 3tp^2 = n^{-1/5+o(1)} = o(1)$. η is the probability that a fixed triangle is contained in $G_{n,p}[T]$ and thus $\eta = p^3 = o(1)$. Thus by (4.4), with $\varepsilon = 1/2$, we have

$$\mathbb{P}[Y \leq \mu/2] \leq \mathrm{e}^{-(1+o(1))\mu/8} \overset{(4.8)}{\leq} n^{-3t}. \tag{4.9}$$

Now we turn to the second step. We define an auxiliary graph G_{K_4} on $G_{n,p}$ as follows. Let $\mathcal{S}' \subseteq \mathcal{S}$ be the lexicographically first set of edge-disjoint triangles in $G_{n,p}[T]$ of maximum size (i.e. $|\mathcal{S}'| = Y$). We say that a subgraph of K_n is a 1-*cluster* if it is the union of a K_3 in T and a K_4 which share at least an edge. For each triangle in $\mathcal{S}'$ which has an edge in common with a K_4 in $G_{n,p}$, pick the lexicographically first of these K_4's and let the 1-cluster which is the union of the triangle and this K_4 be a vertex of G_{K_4}. Two (distinct) vertices in G_{K_4} are joined by an edge if the corresponding 1-clusters share at least an edge. Since the number of vertices of G_{K_4} is equal to the number of elements of $\mathcal{S}'$ which share at least an edge with a K_4 in $G_{n,p}$, (4.6) will then follow if we can show that if $\mathcal{A}$ holds, then the probability that G_{K_4} has $|\mathcal{S}'| = Y$ vertices is sufficiently small.

To bound the number of vertices of G_{K_4}, we will use the fact that for any graph G, we have $v(G) \leq \alpha(G) + 2\Delta(G)\gamma(G)$, where $\alpha(G)$ denotes the size of a largest independent set of vertices in G, and $\gamma(G)$ denotes the size of a largest induced matching in G.

First we bound $\Delta(G_{K_4})$. For this, we will use the fact that any collection of size $6k$ of 1-clusters where the K_4's all contain a fixed edge f must contain a (k, f)-cluster. (To see this, consider the 1-clusters as given in a sequence $F_1, \ldots, F_{6k}$ and denote by e_i the (lexicographically first) edge in F_i that is shared by the K_3 and the K_4. Now for each i (in increasing order) delete those F_j with $j > i$ from the current sequence for which e_j lies H_i, provided F_i was not previously deleted. Since the triangles in different F_j are edge-disjoint and $e(K_4) = 6$, a subsequence containing at least k of the F_i remains, and it is clear that for each such F_i the edge e_i is not contained in the K_4 of any of the F_j in that subsequence with $j < i$.) Also a 1-cluster has at most 8 edges. So if $\mathcal{A}$ holds, we have $\Delta(G_{K_4}) \leq 48 \log n$.

Next we bound $\alpha(G_{K_4})$. Let X_1 denote the number of 1-clusters in $G_{n,p}$ and let $\mu_1 = \mathbb{E}[X_1]$. A 1-cluster is either the union of a K_3 in T and a K_4 which have exactly an edge in common, or a K_4 with at least three vertices in T. Thus

$$\mu_1 \leq \binom{t}{3} n^2 p^8 + \binom{t}{3} n p^6 = \mu\left(n^2 p^5 + n p^3\right) = \mu\left(4^{-5} + \mathcal{O}(n^{-1/5})\right) \leq \mu/80.$$

Let Y_1 denote the size of a largest family of edge-disjoint 1-clusters in $G_{n,p}$. Note that $\alpha(G_{K_4}) \leq Y_1$. Then we have

$$\mathbb{P}[\alpha(G_{K_4}) \geq \mu/8] \leq \mathbb{P}[Y_1 \geq \mu/8] \overset{(4.2)}{\leq} (\mathrm{e}/10)^{\mu/8} \leq \mathrm{e}^{-\mu/8} \overset{(4.8)}{\leq} n^{-3t}.$$

Now we bound $\gamma(G_{K_4})$. The details of the argument we give here are tailored towards the K_4 case and thus slightly different than in the general case. Let $\mathcal{G}$ be the set of graphs $G_1, \ldots, G_4$ which are shown in Figure 1. We say that a copy of G_i in K_n is a T-*copy* of G_i if the black vertices are all contained in T (the white vertices

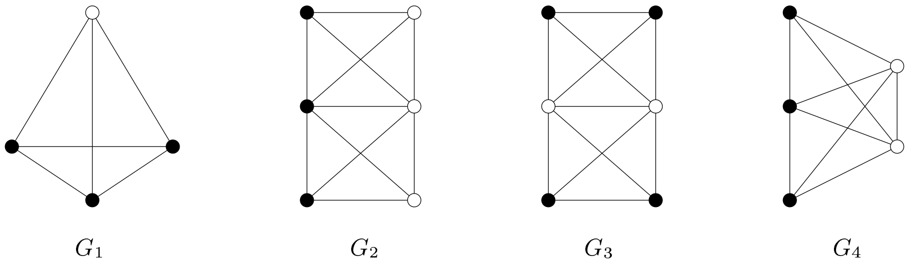

Figure 4.1: Minimal structures in a 2-cluster

are allowed to be anywhere). Let $X_{\mathcal{G}}$ denote the number of T-copies of graphs in $\mathcal{G}$ in $G_{n,p}$ and let $\mu_2 = \mathbb{E}[X_{\mathcal{G}}]$. Then

$$\mu_2 \le t^3np^6 + t^3n^3p^{11} + t^4n^2p^{11} + t^3n^2p^9 = tn^{-1/5+o(1)} = \mu n^{-1/5+o(1)}. \tag{4.10}$$

We say that a 2-*cluster* is the union of two 1-clusters which share at least an edge and which have been obtained from two edge-disjoint triangles in T. Let Y_2 denote the size of a largest family of edge-disjoint 2-clusters in $G_{n,p}$ and let Y_2' denote the size of a largest family in $G_{n,p}$ of edge-disjoint T-copies of graphs from $\mathcal{G}$. Now one may check that every 2-cluster must contain a T-copy of a graph from $\mathcal{G}$. Thus $Y_2 \le Y_2'$. Also note that $\gamma(G_{K_4}) \le Y_2$. Let $y_2 = \mu/(2000 \log n)$. Then

$$\mathbb{P}[\gamma(G_{K_4}) \ge y_2] \le \mathbb{P}[Y_2' \ge y_2] \overset{(4.2)}{\le} \left(\frac{\mathrm{e}\mu_2}{y_2}\right)^{y_2} \overset{(4.10)}{\le} n^{-(1+o(1))y_2/5} \overset{(4.8)}{\le} n^{-3t}.$$

Since for $i = 1, 2$, $\mathbb{P}[\{Y_i \ge y_i\} \cap \mathcal{A}] \le \mathbb{P}[Y_i \ge y_i]$, we have thus proved that if $\mathcal{A}$ holds, then with probability at least $1 - 2n^{-3t}$ we have $v(G_{K_4}) \le \mu/8 + 2\,(48 \log n) y_2 \le \mu/4$. Combining this with (4.9), this implies (4.6).

4.6 Theorem 4.2 – the upper bound

Let H be the forbidden graph, and suppose that H is strictly 2-balanced. Let x be a vertex of maximum degree in H, and let $\Gamma(x)$ denote the neighbourhood of x in H. Let $H \backslash x$ be the subgraph of H induced by $V(H) \backslash \{x\}$. Let $H_{\Gamma(x)}$ be the subgraph of $H \backslash x$ induced by $\Gamma(x)$. We may assume that $H \backslash x$ contains an edge (and thus also that $\Delta(H) > 1$), since otherwise H is a star, and hence not strictly 2-balanced.

We now define a generalized notion of the (non)-independence of a set. Let H and x be chosen as above. Suppose we are given a graph G together with a subset T of its vertices and a vertex $y \in V(G) \backslash T$. Then we say that (T, y) is (H, x)-*extendible in* G if G contains a copy of $H \backslash x$, which we call an *extension (of* (T, y) *in* G*)*, so that $H_{\Gamma(x)}$ is contained in T and all other vertices are contained in $V(G) \backslash \{T \cup \{y\}\}$. In other words, if y is connected to all vertices of T and (T, y) is (H, x)-extendible, then y may be extended to a copy of H where x is mapped to y, the neighbours of x are mapped into T, and all other vertices are mapped outside $T \cup \{y\}$. Thus for example (T, y) is (K_3, x)-extendible in G if and only if T is not an independent set in G. (T, y) is (K_4, x)-extendible in G if and only if T contains a triangle, and (T, y) is (C_ℓ, x)-extendible in G if and only if G contains a path on $\ell - 1$ vertices whose endvertices are contained in T and whose remaining vertices are not in $T \cup \{y\}$. Finally, let $\mathcal{S} = \mathcal{S}_{T,y}$ denote the set of all extensions of (T, y) in K_n. Let

$$\beta = (v - 2)/(e - 1)$$

and note that $0 < \beta < 1$. Let

$$t = Cn^{1-\beta} (\log n)^{1/(\Delta(H)-1)}, \tag{4.11}$$

where C is some sufficiently large constant (see Proposition 4.9 below) depending only on H. Also let $p = c_0 n^{-\beta}$, where c_0 satisfies $c_0^{e-1} e^2 = 1/100$. So $t = pn^{1+o(1)}$ and

$$n^{v-2} p^{e-1} = 1/(100e^2). \tag{4.12}$$

The proof of Theorem 4.2 relies on Lemmas 4.6 and 4.7. Lemma 4.6 asserts that almost surely the copies of H in $G_{n,p}$ do not cluster too much. Lemma 4.7 implies that for the above choice of t and p, almost surely any set of size t will be extendible in $G_{n,p}(H)$, and thus in $M_{n,p}(H)$. Given an edge f in K_n and $k \in \mathbb{N}$ we say that a sequence $H_1, \dots, H_k$ of copies of H forms a (k, f, H)-*cluster* in K_n if they all contain f and have the property that for all i with $1 \leq i \leq k$, H_i contains an edge which is not contained in any of the other H_j with $j < i$. Let $\mathcal{A}$ denote the event that $G_{n,p}$ does not contain a $(\log n, f, H)$-cluster for any edge f.

Lemma 4.6 *$\mathcal{A}$ holds almost surely .*

We defer the proof of this lemma, and first show how it may be used to prove the theorem. Given a subset T of vertices of K_n with $|T| = t$ and a vertex $y \in V(K_n) \setminus T$, let $X_{M(H)}$ be the number of extensions of (T, y) in $M_{n,p}(H)$, let $X_{G(H)}$ denote their number in $G_{n,p}(H)$, and let X_G denote their number in $G_{n,p}$. Let $\mu = \mathbb{E}[X_G]$. (We will suppress the dependence of our random variables on T and y throughout). Note that by (4.1), we have

$$X_{G(H)} \leq X_{M(H)} \leq X_G.$$

The second (and main) ingredient in the proof of the upper bound is the following lemma.

Lemma 4.7 *For a fixed pair (T, y) as above, we have*

$$\mathbb{P}[\{X_{G(H)} = 0\} \cap \mathcal{A}] \leq n^{-2t},$$

provided n is sufficiently large.

The proof of this lemma is also deferred to the end of this section.

Proof of Theorem 4.2 (modulo Lemmas 4.6 and 4.7). Suppose we have a vertex y in $M_n(H)$ whose degree is at least t. Let T be a set of neighbours of y with $|T| = t$. As $M_n(H)$ contains no copies of H and $M_{n,p}(H) \subseteq M_n(H)$, $M_{n,p}(H)$ contains no copies of H, and so (T, y) is not extendible in $M_{n,p}(H)$. But since $X_{M(H)} \geq X_{G(H)}$, Lemma 4.7 tells us that this happens with probability tending to zero: there are (very crudely) at most n^{t+1} choices for (T, y) and thus the probability that there exists a pair (T, y) which is not extendible is at most

$$1 - \mathbb{P}[\mathcal{A}] + n^{t+1}\, n^{-2t}.$$

Applying Lemma 4.6 completes the proof of the theorem. □

We now prove Lemma 4.6 and Lemma 4.7. In the proofs we will repeatedly make use of the following three propositions. Throughout, for a graph F, we write Z_F for the number of copies of F in $G_{n,p}$.

Proposition 4.8 *There exists a constant $\delta > 0$, depending only on H, so that for any proper subgraph H' of H with $e(H') \geq 1$ and $v(H') \geq 3$, we have*

$$\mathbb{E}[Z_{H'}] \geq n^{2+\delta+o(1)}p = n^{\delta+o(1)}\mathbb{E}[Z_H].$$

Moreover, $\mathbb{E}[Z_H] \leq n^2p/(100e^2)$.

Proof. Let

$$\delta = \min_{H' \subset H, e(H') \geq 2} \frac{v(H') - 2}{e(H') - 1} - \beta, \tag{4.13}$$

where the minimum is over all proper subgraphs H' of H with $e(H') \geq 2$. Note that the fact that H is strictly 2-balanced implies that $0 < \delta < 1$.

Let H' be a proper subgraph of H with $e(H') \geq 1$ and $v(H') \geq 3$. First suppose that $e(H') = 1$. Then for some sufficiently small constant $c > 0$, we have

$$\mathbb{E}[Z_{H'}] \geq cn^{v(H')}p \geq cn^{2+\delta}p.$$

Now suppose that $e(H') > 1$. Then for some sufficiently small constant $c > 0$, we have (recalling that $p = c_0n^{-\beta}$)

$$\begin{aligned}
\mathbb{E}[Z_{H'}] &\geq cn^{v(H')}p^{e(H')} \\
&= cn^2p\left(n^{(v(H')-2)/(e(H')-1)}p\right)^{e(H')-1} \\
&\geq cn^2p\left(n^{\delta+\beta}c_0n^{-\beta}\right)^{e(H')-1} \\
&\geq c\,c_0\,n^{2+\delta}p.
\end{aligned}$$

The right hand side and the “moreover” part follow from the fact that for some constant $c' \leq 1$, we have

$$\mathbb{E}[Z_H] = (1 + o(1))c'n^vp^e \overset{(4.12)}{=} (1 + o(1))c'n^2p/(100e^2),$$

where the $o(1)$ term is negative. □

For the remainder of this section we fix δ as in (4.13). Let

$$c_2 = \min\{\delta, \beta, 1 - \beta\}, \tag{4.14}$$

and note that $c_2 < 1$. Recall that μ is the expected number of extensions of (T, y) in $G_{n,p}$.

Proposition 4.9

$$\mu \geq \frac{80e^2}{c_2}\, t \log n.$$

Moreover, $\mu = n^{1+o(1)}p$.

Proof. For some sufficiently small c_1 depending on H, we have

$$\begin{aligned}
\mu &= |\mathcal{S}|\, p^{e-\Delta(H)} && (4.15)\\
&\geq c_1 t^{\Delta(H)}\, (n-t-1)^{v-\Delta(H)-1}\, p^{e-\Delta(H)} \\
&\overset{(4.11)}{\geq} c_1 C^{\Delta(H)-1}\, t\log n\, (n-t-1)^{v-2}\, p^{e-1} \\
&\overset{(4.12)}{=} \frac{c_1 C^{\Delta(H)-1}}{100e^2}\, t \log n \left(\frac{n-t-1}{n}\right)^{v-2}. && (4.16)
\end{aligned}$$

Recalling that $\Delta(H) > 1$ and $t = o(n)$, the assertion then follows if we choose C sufficiently large. The proof of the "moreover" part is similar. □

The following proposition will imply that intersecting copies of H will typically only have an edge in common.

Proposition 4.10 *Given a copy* F *of a subgraph in* K_n *with* $e(F) = \mathcal{O}(\log n)$ *and a proper subgraph* H' *of* H *with* $e(H') \geq 1$, *let* $Z'_{H \cap F = H'}$ *denote the number of copies of* H *in* $G_{n,p} \cup F$ *so that their intersection with* F *is isomorphic to* H'. *If* H' *is an edge, then*

$$\mathbb{E}[Z'_{H\cap F=H'}] \leq \frac{e(F)}{100e}.$$

If H' *is not an edge, then*

$$\mathbb{E}[Z'_{H\cap F=H'}] \leq n^{-\delta+o(1)}.$$

Proof. Suppose first that H' is an edge. Thus there are $e(F)$ possibilities for choosing H' in F and e possibilities for choosing H' in H. Thus in this case

$$\mathbb{E}[Z'_{H\cap F=H'}] \leq e(F) e n^{v-2} p^{e-1} \overset{(4.12)}{\leq} \frac{e(F)}{100e}.$$

Now suppose that H' is not an edge. Then

$$\mathbb{E}[Z'_{H\cap F=H'}] = \mathcal{O}\left(e(F)^{e(H')} n^{v-v(H')} p^{e-e(H')}\right) = \frac{\mathcal{O}\left(e(F)^{e(H')}\mathbb{E}[Z_H]\right)}{\mathbb{E}[Z_{H'}]}.$$

The result now follows from Proposition 4.8. □

Proof of Lemma 4.6. Let $Z_{k,f,H}$ denote the number of (k,f,H)-clusters in $G_{n,p}$. By applying Proposition 4.10 with $F = f$, we have,

$$\mathbb{E}[Z_{1,f,H}] \leq 1/(100e) \leq \mathrm{e}^{-4}. \qquad (4.17)$$

Note that every $(k+1, f, H)$-cluster can be viewed as the union of a (k, f, H)-cluster $C_{k,f}$ and a $(1, f, H)$-cluster which is not contained in $C_{k,f}$. For a fixed (k, f, H)-cluster $C_{k,f}$ in K_n (with $k < \log n$), the expected number of $(1, f, H)$-clusters in $G_{n,p} \cup C_{k,f}$ intersecting $C_{k,f}$ only in f is at most

$$en^{v-2}p^{e-1} \overset{(4.12)}{=} 1/(100e).$$

By Proposition 4.10, the expected number of $(1, f, H)$-clusters in $G_{n,p} \cup C_{k,f}$ intersecting $C_{k,f}$ in some proper subgraph H' of H which is not an edge is at most $n^{-\delta+o(1)}$. Let $Z_{1,f,C_{k,f}}$ denote the number of $(1, f, H)$-clusters in $G_{n,p}$ which are not contained in $C_{k,f}$. Then by the above,

$$\mathbb{E}[Z_{1,f,C_{k,f}} \mid C_{k,f} \subseteq G_{n,p}] \leq 1/(99e) \leq \mathrm{e}^{-4}.$$

Thus, if we let $\sum_{C_{k,f}}$ denote the sum over all (k, f, H)-clusters in K_n, we have

$$\begin{aligned}\mathbb{E}[Z_{k+1,f,H}] &\leq \sum_{C_{k,f}} \mathbb{P}[C_{k,f} \subseteq G_{n,p}]\, \mathbb{E}[Z_{1,f,C_{k,f}} \mid C_{k,f} \subseteq G_{n,p}] \\ &\leq \sum_{C_{k,f}} \mathbb{P}[C_{k,f} \subseteq G_{n,p}]\, \mathrm{e}^{-4} \\ &= \mathbb{E}[Z_{k,f,H}]\, \mathrm{e}^{-4}.\end{aligned}$$

Together with (4.17) this shows that $\mathbb{E}[Z_{k,f,H}] \leq \mathrm{e}^{-4k}$ and thus that $\mathbb{E}[Z_{\log n,f,H}] \leq n^{-4}$. Summing over all edges of the complete graph, the result now follows. □

It now remains to prove Lemma 4.7. To this end, we first show that for our choice of p and of the size of T, there should be many *edge-disjoint* extensions of (T, y) in $G_{n,p}$ for a fixed pair (T, y). The next lemma will be the key towards proving this fact.

Lemma 4.11 *Consider a given copy F of a graph with at most $3e$ edges in K_n. Let $Z^*_{F\cup H\setminus x}$ denote the number of extensions in $G_{n,p} \cup F$ having at least an edge in common with F and which are not contained in F. Then*

$$\mathbb{E}[Z^*_{F\cup H\setminus x}] \leq n^{-\delta+o(1)}.$$

Proof. The main problem is that we have to take care of how many vertices of the intersection of an extension and F lie in T. Fix a proper subgraph H' of $H\setminus x$. Let $Z'_{H'}$ denote the number of those copies of H' in $G_{n,p}$ for which the vertices corresponding to vertices in $V(H') \cap \Gamma(x)$ are contained in T and the vertices not corresponding to vertices in $V(H') \cap \Gamma(x)$ do not lie in T. Let $\phi = |V(H') \cap \Gamma(x)|$. Then the expected number of extensions in $G_{n,p} \cup F$ whose intersection with F is H' (and thus exactly ϕ vertices of the intersection are contained in T) is

$$\begin{aligned}&\mathcal{O}\left(t^{\Delta(H)-\phi}\, n^{v-\Delta(H)-1-(v(H')-\phi)}\, p^{e-\Delta(H)-e(H')}\right) \\ \overset{(4.15)}{=}\, &\mathcal{O}\left(\mu t^{-\phi}\, n^{-(v(H')-\phi)}\, p^{-e(H')}\right) \\ =\, &\mathcal{O}\left(\mu/\mathbb{E}[Z'_{H'}]\right).\end{aligned}$$

Let $\sum_{H'}$ denote the summation over all subgraphs H' of $H\backslash x$. Then by the above,

$$\mathbb{E}[Z^*_{F\cup H\backslash x}] = \sum_{H'} \frac{\mathcal{O}(\mu)}{\mathbb{E}[Z'_{H'}]}. \tag{4.18}$$

Since $t = n^{1+o(1)}p$,

$$\mathbb{E}[Z'_{H'}] = \mathcal{O}\left(\mathbb{E}[Z_{H'}](t/n)^{\phi}\right) = \mathbb{E}[Z_{H'}]p^{\phi}n^{o(1)} = \mathbb{E}[Z_{H'\cup x}]n^{-1+o(1)},$$

where $H' \cup x$ is the subgraph of H obtained from H' by adding a vertex x and connecting it to those vertices of H' that are contained in $\Gamma(x)$. Since H' is a proper subgraph of $H\backslash x$ containing at least one edge, $H' \cup x$ is a proper subgraph of H which is not an edge. Proposition 4.8 then implies that $\mathbb{E}[Z_{H'\cup x}] \geq n^{2+\delta+o(1)}p$ and thus that $\mathbb{E}[Z'_{H'}] \geq n^{1+\delta+o(1)}p$. Combining this with equation (4.18), and recalling that $\mu = n^{1+o(1)}p$ (Proposition 4.9), we have

$$\mathbb{E}[Z^*_{F\cup H\backslash x}] \leq \frac{\mu}{n^{1+\delta+o(1)}p} \leq n^{-\delta+o(1)}.$$

□

Let Y be the maximum number of edge-disjoint extensions of (T, y) in $G_{n,p}$ and recall that μ is the expected number of extensions of (T, y) in $G_{n,p}$.

Lemma 4.12 *For n sufficiently large,*

$$\mathbb{P}[Y \leq \mu/2] \leq n^{-3t}.$$

Proof. We shall apply inequality (4.4) to $\mathcal{S}$, the set of extensions of (T, y) in K_n. In order to be able to do this, we will apply Lemma 4.11 to show that η, ν and Δ (as defined in Section 4.2) are sufficiently small.

We first consider Δ. Since the extensions are symmetric (more formally, for all α, $\alpha' \in \mathcal{S}$ there is an automorphism of the underlying probability space which sends the event I_α to $I_{\alpha'}$), we have $\Delta = \mu\Delta^*$, where for some fixed $\alpha \in \mathcal{S}$,

$$\Delta^* = \sum_{\alpha' \sim \alpha} \mathbb{E}[I_{\alpha'} | I_\alpha = 1].$$

But, defining $Z^*_{\alpha\cup H\backslash x}$ as in Lemma 4.11 (with α playing the role of F), one may observe that $\Delta^* = \mathbb{E}[Z^*_{\alpha\cup H\backslash x}]$, and thus by Lemma 4.11, we have $\Delta^* \leq n^{-\delta+o(1)} = o(1)$. Thus $\Delta = o(\mu)$.

We now bound ν. Fix some $\alpha \in \mathcal{S}$ again. Note that, since $\alpha' \subseteq G_{n,p}$ certainly implies that $\alpha' \subseteq G_{n,p} \cup \alpha$, we have

$$\nu = \sum_{\alpha' \sim \alpha} \mathbb{E}[I_{\alpha'}] \leq \mathbb{E}[Z^*_{\alpha\cup H\backslash x}].$$

Thus by Lemma 4.11 again, we have $\nu \leq n^{-\delta+o(1)} = o(1)$.

Finally, since an extension contains at least an edge, we have $\eta = \mathbb{E}[I_\alpha] = p^{e-\Delta(H)} = o(1)$, and thus

$$\mathbb{P}[Y \leq \mu/2] \overset{(4.4)}{\leq} \mathrm{e}^{-(1+o(1))\mu/8} \leq n^{-3t},$$

where the final inequality follows from Proposition 4.9. □

We now show that at least some of the extensions which are present in $G_{n,p}$ are also present in $G_{n,p}(H)$. In what follows, we will say that a graph $G \subset K_n$ can be obtained by *merging* two graphs G_1 and G_2 if G_1 and G_2 share at least an edge and if $G = G_1 \cup G_2$. We say that a subgraph F of K_n is a 1-*cluster* $E \cup H_1$ if it can be obtained by merging an extension E with a copy H_1 of H. A subgraph F of K_n is a 2-*cluster* if it can be obtained by merging two 1-clusters $E_1 \cup H_1$ and $E_2 \cup H_2$ where E_1 and E_2 are edge-disjoint extensions. We say that a 2-cluster is a *minimal* 2-*cluster* if it contains no 2-cluster as a proper subgraph.

For $a = 1, 2$, let Y_a be the maximum size of a family of pairwise edge-disjoint a-clusters in $G_{n,p}$. Let Y_2' denote the maximum size of a family of pairwise edge-disjoint *minimal* 2-clusters in $G_{n,p}$. Note that since every 2-cluster contains a minimal 2-cluster, we have $Y_2' = Y_2$. Let μ_1 denote the expected number of 1-clusters in $G_{n,p}$ and let μ_2 denote the expected number of minimal 2-clusters in $G_{n,p}$. Let $y_1 = \mu/8$ and $y_2 = \mu/(20e^2 \log n)$.

Lemma 4.13

$$\mathbb{P}\left[Y_1 \geq y_1\right] \leq n^{-3t}.$$

Proof. Fix an extension F in K_n and consider a subgraph H' (which contains at least an edge) of $H \backslash x$. As in Proposition 4.10, let $Z'_{H \cap F = H'}$ denote the number of copies of H in $G_{n,p} \cup F$ whose intersection with F is isomorphic to H'. Then, since a 1-cluster is the union of an extension and a copy of H sharing at least an edge, we have

$$\mu_1 \leq \mu \sum_{H' \subseteq H \backslash x} \mathbb{E}[Z'_{H \cap F = H'}],$$

where the summation is over all subgraphs H' of $H \backslash x$ with $e(H') \geq 1$. Thus Proposition 4.10 implies that

$$\mu_1 \leq \mu \left(1/100 + n^{-\delta + o(1)}\right) \leq \mu/80. \tag{4.19}$$

Combining this with inequality (4.2), we have

$$\mathbb{P}\left[Y_1 \geq y_1\right] \leq \left(\frac{\mathrm{e}\mu_1}{y_1}\right)^{y_1} \leq \mathrm{e}^{-\mu/8}.$$

The result now follows from Proposition 4.9. □

Lemma 4.14

$$\mathbb{P}\left[Y_2 \geq y_2\right] \leq n^{-3t}.$$

Proof. We will actually show that $\mathbb{P}[Y_2' \geq y_2] \leq n^{-3t}$, which implies the lemma by our previous observation that $Y_2' = Y_2$. We first show that

$$\mu_2 \leq \mu n^{-c_2+o(1)}, \tag{4.20}$$

where c_2 is defined as in (4.14). To this end, note that a (minimal) 2-cluster may be obtained from a 1-cluster F_1 by first merging F_1 with a copy H_2 of H (call the resulting graph an *augmented* 1-cluster), and then merging the augmented 1-cluster with an extension E. Write $\mu_2 = \mu_2' + \mu_2''$, where μ_2' denotes the expected number of minimal 2-clusters where the extension E is not contained in the augmented 1-cluster and μ_2'' denotes the expected number of minimal 2-clusters where it is (see Fig. 4.6). First we bound μ_2'. By (4.19), the expected number of 1-clusters is $\mathcal{O}(\mu)$. We claim that also the expected number of augmented 1-clusters is $\mathcal{O}(\mu)$. Indeed, the expected number of augmented 1-clusters with $H_2 \subset F_1$ is certainly $\mathcal{O}(\mu)$. So consider those with $H_2 \not\subset F_1$. Since a 1-cluster has less than $2e$ edges, applying Proposition 4.10 to all proper subgraphs H' of H (with the 1-cluster playing the role of F in the statement of the proposition) shows that the expected number of augmented 1-clusters is also $\mathcal{O}(\mu)$. Thus by Lemma 4.11 (with the augmented 1-cluster playing the role of F in the statement of the lemma), we have

$$\mu_2' \leq \mu n^{-\delta+o(1)}.$$

It remains to estimate μ_2''. Unfortunately, in this case we cannot use Lemma 4.11 (since here we have $E \subset F$), so we have to work harder. Note that the augmented 1-clusters are obtained by merging a 1-cluster F_1 (which by definition must contain a copy H_1 of H) with a copy H_2 of H which may be contained in F_1.

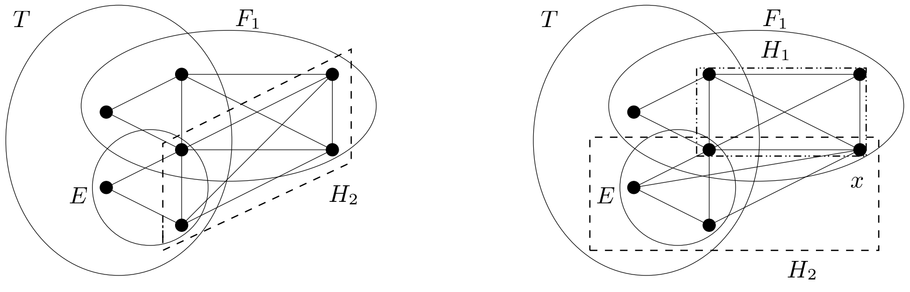

Figure 4.2: Minimal 2-clusters when $H = K_4$: the one on the left illustrates the estimation of μ_2' and the one on the right illustrates the estimation of μ_2'' for the case when $H_2 \not\subseteq F_1$.

Suppose first that $H_2 \not\subseteq F_1$. By definition of μ_2'', the augmented 1-cluster must contain the extension E. Since we are considering only minimal 2-clusters, the extension E must be edge-disjoint from H_1, otherwise $F_1 \cup E$ would already be a 2-cluster. As $H_2 \not\subseteq F_1$, this would contradict minimality. Thus $E \subset H_2$. Now $V(H_2) \backslash V(E)$ contains only a single vertex x (which is the same vertex as in the definition of an extension) and so the fact that E is edge-disjoint from H_1 implies that $x \in F_1$ and that $F_1 \cap H_2$ consists only of edges running between x and E and hence is a star with i edges, where $1 \leq i \leq \Delta(H)$. Since for the extension E the vertices belonging to $\Gamma_{H_2}(x)$ must be contained in T, the vertices of E which are adjacent to the centre x must be mapped to T, while the remaining $v - \Delta(H) - 1$ vertices of E can be mapped anywhere. Since by (4.19), the expected number of 1-clusters is $\mathcal{O}(\mu)$, the contribution to μ_2'' from the case $H_2 \not\subseteq F_1$ is thus

$$\sum_{i=1}^{\Delta(H)} \mathcal{O}\left(\mu\, t^{\Delta(H)-i} n^{v-\Delta(H)-1} p^{e-i}\right). \tag{4.21}$$

Now consider the case when $H_2 \subseteq F_1$. Then the minimal 2-cluster is actually a 1-cluster containing two edge-disjoint extensions. (This can occur e.g. if H is a triangle and F_1 is a triangle in T.) Such a minimal 2-cluster can be obtained by merging an extension E with a copy H_1 of H which contains an extension. Again, by the edge-disjointness of the extensions the intersection $E \cap H_1$ is a star with at most $\Delta(H)$ edges. As the expected number of extensions is μ, arguing as above but starting from the extension E rather than the 1-cluster F_1, we find that the contribution to μ_2'' from this case is also

$$\sum_{i=1}^{\Delta(H)} \mathcal{O}\left(\mu\, t^{\Delta(H)-i} n^{v-\Delta(H)-1} p^{e-i}\right). \tag{4.22}$$

Using the facts that $\Delta(H) > 1$ and that $\mu = n^{1+o(1)}p = tn^{o(1)}$ (see Proposition 4.9), (4.21) and (4.22) imply that

$$\begin{aligned}\mu_2'' &\overset{(4.12)}{=} \sum_{i=1}^{\Delta(H)} \mathcal{O}\left(\mu\, t^{\Delta(H)-i} n^{1-\Delta(H)} p^{1-i}\right)\\ &\le \mu\, n^{o(1)} \sum_{i=1}^{\Delta(H)} t^{1-i} p^{\Delta(H)-i} \le \mu\, n^{o(1)} \left(n^{-\beta} + n^{-(1-\beta)}\right),\end{aligned}$$

where the first summand comes from the case $i = 1$ and the second one from the case $i = \Delta(H)$ (note that if $\Delta(H) > 2$ the terms with $1 < i < \Delta(H)$ are dominated by the ones with $i = 1$ and $i = \Delta(H)$). Thus (4.20) follows by the definition of c_2 in (4.14).

Combining inequalities (4.2) and (4.20) this time, we have

$$\begin{aligned}\mathbb{P}\left[Y_2 \ge y_2\right] &\le \left(\frac{\mathrm{e}\mu_2}{y_2}\right)^{y_2}\\ &\le n^{-(1+o(1))c_2\mu/(20e^2 \log n)}\\ &\le n^{-(1+o(1))4t},\end{aligned}$$

where the last line follows by Proposition 4.9. □

Proof of Lemma 4.7. Let $\mathcal{S}' \subseteq \mathcal{S}$ be the lexicographically first set of edge-disjoint extensions in $G_{n,p}$ of maximum size (so that $|\mathcal{S}'| = Y$). We now define an auxiliary graph G_H on $G_{n,p}$ as follows. For each extension in $\mathcal{S}'$ which has an edge in common with a copy of H in $G_{n,p}$ (and thus lies in a 1-cluster), pick the lexicographically first of these copies of H and let the 1-cluster which is the union of the extension and this copy of H be a vertex of G_H. Two (distinct) vertices are connected by an edge if the 1-clusters have at least an edge in common. Since the number of vertices of G_H is equal to the number of elements of $\mathcal{S}'$ which share at least an edge with a copy of H, the lemma will then follow if the probability that G_H has $|\mathcal{S}'| = Y$ vertices is sufficiently small.

To bound the number of vertices of G_H, we will use the fact that for any graph G, we have $v(G) \leq \alpha(G) + 2\Delta(G)\gamma(G)$, where $\alpha(G)$ denotes the size of a largest independent set of vertices in G and $\gamma(G)$ denotes the size of a largest induced matching in G.

First we bound $\Delta(G_H)$. For this, we will use the fact that any collection of size ek of 1-clusters where the copies of H all contain a fixed edge f must contain a (k, f, H)-cluster. (To see this, consider the 1-clusters as given in a sequence $F_1, \dots, F_{ek}$, where $F_i = E_i \cup H_i$ is obtained by merging an extension E_i with a copy H_i of H. Also denote by e_i the (lexicographically first) edge in $E_i \cap H_i$. For each i (in increasing order) delete those F_j with $j > i$ from the current sequence for which e_j lies H_i, provided F_i was not previously deleted. Since the E_j are edge-disjoint, a subsequence $\mathcal{F}$ containing at least k of the F_i remains, and for each $F_i \in \mathcal{F}$ the edge e_i is not contained in any of the H_j with $j < i$ and $F_j \in \mathcal{F}$.) Also, a 1-cluster has at most $2e$ edges. So if $\mathcal{A}$ holds, we have $\Delta(G_H) \leq 2e^2 \log n$.

Furthermore, it is easily verified that the definitions of Y_1, Y_2, and G_H imply that $\alpha(G_H) \leq Y_1$ and $\gamma(G_H) \leq Y_2$. By Lemmas 4.13 and 4.14, $\mathbb{P}[\{Y_a \geq y_a\} \cap \mathcal{A}] \leq \mathbb{P}[Y_a \geq y_a] \leq n^{-3t}$ for $a = 1, 2$. Thus if $\mathcal{A}$ holds, with probability at least $1 - 2n^{-3t}$ our auxiliary graph has at most

$$y_1 + 2\,(2e^2 \log n)\, y_2 \leq (1/8 + 4/20)\mu \leq 2\mu/5 \tag{4.23}$$

vertices. On the other hand, by Lemma 4.12, we have

$$\mathbb{P}[\{Y \leq \mu/2\} \cap \mathcal{A}] \leq \mathbb{P}[Y \leq \mu/2] \leq n^{-3t}. \tag{4.24}$$

Since $X_{G(H)} \geq Y - v(G_H)$, this completes the proof of the lemma. □

4.7 Forbidding several strictly 2-balanced graphs

Here we indicate how the proofs of Sections 4.4 and 4.6 can easily be modified in order to establish Theorem 4.3. Denote by v_i and e_i the number of vertices and edges of H_i respectively and let $\beta = (v_0 - 2)/(e_0 - 1)$.

For the lower bound, denote by X_i the number of copies of H_i in $G_{n,p}$ and let $X = \sum_{i=0}^{s-1} e_i X_i$. Let $p = c_0 n^{-\beta}$, where c_0 satisfies $\max_i\{c_0^{e_i - 1} e_i\} = \frac{1}{8s}$. Then an upper bound on $\mathbb{E}[X]$, very similarly to the proof of Theorem 4.1, implies the result.

As for the proof of the upper bound, let $p = c_0 n^{-\beta}$, where c_0 now satisfies

$$\max_i\{c_0^{e_i - 1} e_0 e_i\} = \frac{1}{100s}.$$

In the definitions of (H, x)-*extendible*, *extension*, $X_{G(H)}$, $X_{M(H)}$ and X_G we substitute H_0 for H. Instead of (k, f, H)-clusters we consider $(k, f, \mathcal{H})$-*clusters*, where a $(k, f, \mathcal{H})$-cluster is now the union of k copies $H^1, \ldots, H^k$ of possibly different elements in $\mathcal{H}$ which all contain the edge f and which have the property that for all i with $1 \leq i \leq k$, H^i contains an edge which is not contained in any of the other H^j with $j < i$. It is then clear that the theorem follows from Lemmas 4.6 and 4.7 as before. In the proof of (the analogue of) Lemma 4.7 and within this proof Lemmas 4.13 and 4.14 in particular, the definition of a cluster is modified as follows: a 1-cluster is obtained by merging an extension and a copy of a graph in $\mathcal{H}$. A 2-cluster is again obtained by merging two 1-clusters, where we require disjointness of the extensions as before. Now it suffices to check (4.19) and (4.20).

With regard to (4.19), the main contribution again is from the case when each copy of an element H_i of $\mathcal{H}$ has exactly an edge in common with the extension. Note that each H_i contributes $\mu n^{v_i-2}p^{e_i-1}e_0e_i$, which is less than $\mu/(100s)$ due to the choice of p. Hence the same bound as in (4.19) holds. (4.20) is also proven using the same arguments as in Section 4.6.

4.8 Forbidding an arbitrary subgraph – an open question

We remark that lower bounds on $e(M_n(H))$ when H is an arbitrary graph may be obtained by choosing p so that approximately half of the edges of $G_{n,p}$ lie in a copy of H. Deleting these edges as in Section 4.4 yields the lower bound. This value of p may be determined using the methods introduced by Spencer in [81]. We do not give any details however, as there are graphs for which the lower bounds obtained in this way are far from tight, as the example in the following paragraph will show.

Concerning upper bounds on the number of edges in $M_n(H)$, we have seen that if H is strictly 2-balanced, then $e(M_n(H))$ is not much larger that $e(G_{n,p})$, where p was chosen so that the number of copies of H in $G_{n,p}$ is roughly $e(G_{n,p})$. However, this is not true in general, as the following example shows. Let H be the graph which is obtained from K_4 by adding a new vertex and joining it to one of the vertices of the K_4. Then it is easily seen that almost surely $e(M_n(H))$ equals $e(M_n(K_4))$. (Indeed, suppose not and consider the first edge which is admitted in one process and not in the other.) The proof of Theorem 4.2 implies that $e(M_n(K_4))$ is almost surely close to $e(G_{n,p})$ for $p = n^{-2/5}$. However the number of copies of H in $G_{n,p}$ is close to $e(G_{n,p})$ only if p is close to $n^{-1/2}$.

Alternatively, one might think that for all graphs H with $v(H) \geq 3$ and $e(H) \geq 2$, for any $\varepsilon > 0$ almost surely we have

$$n^{2-1/m_2(H)-\varepsilon} \leq e(M_n(H)) \leq n^{2-1/m_2(H)+\varepsilon}, \tag{4.25}$$

where for a graph H with $v(H) \geq 3$ and $e(H) \geq 2$ its 2-*density* $m_2(H)$ is defined by

$$m_2(H) = \max_{H' \subseteq H,\ v(H')>2} \frac{e(H')-1}{v(H')-2}.$$

The following example shows that this too need not be true. Furthermore, it shows that $e(M_n(H))$ need not necessarily be concentrated on a single interval. Let H be the disjoint union of a triangle and a cycle of length four. Then with probability bounded away from zero, a triangle appears in the random greedy process before a four-cycle, and thus one may check that the proof of Theorem 4.2 shows that in this case, almost surely $e(M_n(H))$ will be close to $n^{4/3}$. With probability bounded away from zero, it will be the other way round, and then almost surely $e(M_n(H))$ will be close to $n^{3/2}$. Thus it appears to be impossible to give good bounds on $e(M_n(H))$ in general. Nevertheless, we conjecture that for an arbitrary graph H with $v(H) \geq 3$ and any $\varepsilon > 0$, almost surely we have

$$e(M_n(H)) \leq n^{2-1/m_2(H)+\varepsilon}.$$

Chapter 5

The evolution of triangle-free graphs

5.1 Results

The main result of this chapter is the following theorem, which gives a "double threshold" for the property that almost all C_ℓ-free graphs are bipartite (for fixed odd ℓ). Note that Theorem 2.6 is exactly the case $\ell = 3$.

Theorem 5.1 *Given an odd integer ℓ, let*

$$t_\ell = t_\ell(n) = \left(\frac{\ell}{\ell-1}\left(\frac{n}{2}\right)^\ell \log n\right)^{1/(\ell-1)}.$$

Then for all $\varepsilon > 0$,

$$\frac{\mathbb{P}\left[G_{n,m} \textit{ is bipartite}\right]}{\mathbb{P}\left[G_{n,m} \textit{ is } C_\ell\textit{-free}\right]} \to \begin{cases} 1 & \textit{if} \quad m = o(n) \\ 0 & \begin{array}{l}\textit{if} \quad m \geq n/2 \\ \textit{and } m \leq (1-\varepsilon)t_\ell\end{array} \\ 1 & \textit{if} \quad m \geq (1+\varepsilon)t_\ell. \end{cases}$$

As already mentioned in Section 2.2, Erdős, Kleitman and Rothschild [26] proved that almost all triangle-free graphs are bipartite. This was extended by Lamken and Rothschild [56], who proved that for an odd integer ℓ, almost all C_ℓ-free graphs are bipartite. Prömel and Steger [66] generalized this further by determining the class of forbidden subgraphs for which an analogous result holds: they are those graphs whose chromatic number is at least three and whose chromatic number can be decreased by deleting an edge.

We will now briefly mention some results related to the $\ell = 3$ case of Theorem 5.1. Firstly, a result of Bollobás (see Chapter X of Bollobás [8] or Spencer [82]) implies that

$$m \geq \left(\frac{n^3}{2}\log(\omega n)\right)^{1/2}$$

is a necessary and sufficient condition for $G_{n,m}$ to have diameter two almost surely, where ω is some function tending to infinity arbitrarily slowly. The proof of the latter also shows that the threshold function for the property that almost surely every edge of $G_{n,m}$ lies on a triangle is $\Theta(t_3)$.

Secondly, to prove his lower bound on the Ramsey number $R(3,s)$, Kim [47] showed that there exists a triangle-free graph whose independent sets have size $\mathcal{O}(t_3/n)$. This graph is certainly "far from being bipartite" and since the neighbourhood of every vertex in a triangle-free graph is an independent set, it thus has $\mathcal{O}(t_3)$ edges (see also Section 1.2.1).

Furthermore, it is worth noting that Theorems 2.5 and 2.6 have a deterministic analogue: Erdős et al. [25] proved that a graph in $\mathcal{T}(n,m)$ must exhibit bipartite-like behaviour (i.e. it contains a large induced bipartite graph) if and only if $m/n^{3/2} \to \infty$. (Recall that $\mathcal{T}(n,m)$ denotes the set of triangle-free graphs with n vertices and m edges).

Clearly, in the range of the 0-statement of Theorem 5.1 almost every graph in $\mathcal{T}(n,m)$ has chromatic number at least three. Our next result improves this bound as long as m is not too close to n or t_3 (analogous results hold also for the case of arbitrary odd ℓ). Note that before Erdős proved his result on the existence of graphs of arbitrarily high girth and arbitrarily high chromatic number, Zykov [96] (and independently several others) already proved that there are triangle-free graphs of arbitrarily high chromatic number. Theorem 5.2 is then a strong form of this result – it implies that *almost all* triangle-free graphs with n vertices and m edges have high chromatic number, provided that m lies in a certain range.

Theorem 5.2 *If $m = \mathcal{O}(n^{5/4})$, then almost all graphs in $\mathcal{T}(n,m)$ have chromatic number at least*

$$\frac{m}{2n\log(2m/n)}. \tag{5.1}$$

If $n^{5/4} \le m = o(n^{3/2})$, then almost all graphs in $\mathcal{T}(n,m)$ have chromatic number at least

$$\frac{n^{3/2}}{9m\log n}. \tag{5.2}$$

Note that the bound (5.1) is useless unless $m \ge 40n$, say. Substituting $m = n^{5/4}$ in (5.1) we obtain that for such m, the chromatic number of almost all graphs in $\mathcal{T}(n,m)$ is at least $n^{1/4+o(1)}$. We also remark that with more detailed analysis, the proof shows that the constant in (5.1) as well as the exponent of the $\log n$-factor in (5.2) can be reduced. Note that for $m = \mathcal{O}(n^{5/4})$, the bounds are of the same order of magnitude as those which are known for $G_{n,p}$, where $p = m/\binom{n}{2}$. In fact, Łuczak [58] proved that if $pn \to \infty$, then the chromatic number of $G_{n,p}$ is almost surely

$$(1+o(1))\frac{pn}{2\log(pn)}.$$

This chapter is organized as follows. In Section 5.2, we sketch the proof of the second 1-statement of the $\ell = 3$ case of Theorem 5.1. In Section 5.3, we prove Theorem 2.7, giving tight asymptotic bounds on the number of bipartite graphs with n vertices and m edges for a wide range of m. Furthermore, we prove some facts about bipartite graphs which we shall need in Sections 5.4 and 5.5. In Section 5.4, we prove the 0-statement of the $\ell = 3$ case of Theorem 5.1. Sections 5.5–5.7 are devoted to the proof of the second 1-statement of the $\ell = 3$ case of Theorem 5.1. In Section 5.8, we show how the proof of the $\ell = 3$ case may be extended to yield a proof of the general case of Theorem 5.1. Finally, in Section 5.9, we prove Theorem 5.2.

5.2 Sketch of the proof of the triangle-free case

In this section, we outline the proof of the second 1-statement of the $\ell = 3$ case of Theorem 5.1. We stress that most of the notions and statements are only made precise in the later sections. In contrast to [70] and [59], the main tool used in our proof of the second 1-statement are the correlation inequalities (3.5) and (3.6), which are from Janson, Łuczak and Ruciński [40]. In [40], these were applied to give an exponential upper bound on the probability that a random graph $G_{n,p}$ (and thus $G_{n,m}$) is H-free for some fixed graph H, where $G_{n,p}$ denotes a random graph with n vertices and edge probability p. For $m \geq t_3$ and H a triangle, the bound obtained in this way is asymptotically about $\mathrm{e}^{-m/9}$ (note that $\mathrm{e}^{1/9} \approx 1.12$). However, the probability that a random graph $G_{n,m}$ is bipartite (see Corollary 5.5) is asymptotically only about 2^{-m}. Roughly speaking, the reason for the gap between the two bounds is that the variance of the number of triangles in $G_{n,p}$ is too large.

The outline of our approach is as follows. We say that a graph is *k-bipartite* if k is the minimal number of edges that have to be deleted to make it bipartite. In Lemma 5.9 we show that the ratio of k-bipartite graphs in $\mathcal{G}(n,m)$ to bipartite graphs in $\mathcal{G}(n,m)$ is at most $f(k)$, where the function $f(k)$ is not much larger than $\binom{m}{k}$. Since every graph is k-bipartite for some $k \leq m/2$, we would thus be done if we could show that for all k with $1 \leq k \leq m/2$,

$$\mathbb{P}[G_{n,m} \text{ is } K_3\text{-free} \mid G_{n,m} \text{ is } k\text{-bipartite}] \leq \frac{o(1)}{k^2 f(k)}. \tag{5.3}$$

Indeed, (5.3) would imply that

$$\begin{aligned}\mathbb{P}[G_{n,m} \text{ is } K_3\text{-free}] &= \sum_{k=0}^{m/2} \mathbb{P}[G_{n,m} \text{ is } K_3\text{-free} \mid G_{n,m} \text{ is } k\text{-bip.}]\,\mathbb{P}[G_{n,m} \text{ is } k\text{-bip.}] \\ &\leq \mathbb{P}[G_{n,m} \text{ is bipartite}] + \sum_{k\geq 1} \frac{o(1)}{k^2 f(k)} \mathbb{P}[G_{n,m} \text{ is } k\text{-bip.}] \\ &= (1+o(1))\mathbb{P}[G_{n,m} \text{ is bipartite}].\end{aligned}$$

In other words, the hope would be that by conditioning on the number of edges needed to make $G_{n,m}$ bipartite, one can reduce the variance of the number of triangles and thus obtain the desired result.

We now illustrate this idea for the case when $k = 1$. In Proposition 5.6 we show that the 1-bipartite graphs outnumber the bipartite graphs by a factor of about m. Now consider a random graph chosen as follows. Fix an almost equitable bipartition into classes A and B, where $a = |A| = \lfloor n/2 \rfloor$ and $b = |B| = \lceil n/2 \rceil$. Fix one edge e with both endpoints in A, and include the edges between A and B independently at random with probability $p = (m-1)/(ab) \sim 4m/n^2$. What is the probability that the edge e can be extended to a triangle using a vertex in B? For each vertex in B the probability that there is a triangle extension using this vertex is p^2. But the triangle extensions are independent and thus the probability that the edge has no extensions is

$$(1-p^2)^b \sim \mathrm{e}^{-bp^2} \sim \mathrm{e}^{-8m^2/n^3} \begin{cases} \gg 1/m & \text{if } m \leq (1-\varepsilon)\, t_3 \\ \ll 1/m & \text{if } m \geq (1+\varepsilon)\, t_3. \end{cases}$$

Here we write $f(n) \ll g(n)$ if $f(n)/g(n) \to 0$ and $f(n) \sim g(n)$ if $f(n)/g(n) \to 1$. Note that this does not yet imply (5.3) for $k = 1$, as the probability model is different. We will also encounter this problem later on, where we will usually prove results in some binomial probability model like the one described above, as these have the advantage of greater independence and tractability, and then use some technical arguments to show that essentially these results are true for the corresponding k-bipartite graphs in $\mathcal{G}(n, m)$.

When $k > 1$, we can no longer assume independence between extensions, so we apply the correlation inequalities (3.5) and (3.6). As in the case $k = 1$, we fix some vertex bipartition A, B of the set of n vertices, and fix some graphs G_A and G_B, where the vertex set of G_A is A and that of G_B is B and where these graphs together have k edges. We then apply the correlation inequalities to the bipartite graph between A and B with edge probability $p \sim (m-k)/(ab)$.

However, this yields good upper bounds on the probability that such a graph is triangle-free only if G_A and G_B are close to being regular graphs, as otherwise the variance of the number of triangles is too large. Thus we classify all possible pairs (G_A, G_B) according to how far they are from being regular. This is done by considering the size of the "torsos" – the largest subgraphs of G_A and G_B of sensible maximum degree, see Section 5.5. In Lemma 5.9 we prove that the further the pair (G_A, G_B) is from being regular, the less k-bipartite graphs there are whose subgraphs induced by A and B equal G_A and G_B. Accordingly, the bounds which we need on the probability that such a graph is triangle-free are allowed to get larger the more irregular G_A and/or G_B are. This tradeoff is modelled by the function $f(k, r)$ introduced in Section 5.5. Thus we can show in the proofs of Lemmas 5.13 and 5.14 that the correlation inequalities are strong enough to cope with moderately irregular graphs too.

However, this scheme does break down when G_A or G_B is much too far from being regular. We deal with these graphs in Section 5.7, where we will make use of the fact that there are not too many such graphs. Surprisingly, it also turns out that we can actually take advantage of the strong irregularity of G_A, say. Namely, it implies that G_A contains many vertices of high degree (see Proposition 5.8). It is easily seen that for any k-bipartite graph and any $x \in A$, the degree $d_A(x)$ of x in A is at most the degree $d_B(x)$ in B. This fact implies that A contains many vertices x with large neighbourhood $\Gamma_B(x)$ in B. We then show (see Lemma 5.18) that with high probability at least half of these neighbourhoods "expand", in the sense that almost all vertices in A are adjacent to some vertex in $\Gamma_B(x)$, i.e. $|\Gamma_A(\Gamma_B(x))| \sim a$. But then the whole graph will contain a triangle unless $\Gamma_A(\Gamma_B(x))$ and $\Gamma_A(x)$ are disjoint, which happens very rarely, since the former set is so large (see the proof of Lemma 5.15).

Finally, we should mention that our methods fail if $k \geq \delta m$, where δ is some small constant, as we can then no longer control the variance of the number of triangles. Fortunately however, the result of Łuczak (Theorem 2.5) means that we can ignore this case. Our methods also fail if $m/n^2 \not\to 0$. However, this case is already covered in Theorem 2.4 by Prömel and Steger, so we can ignore this possibility too.

For clarity, we will often omit floors and ceilings whenever this does not affect the proof. We will assume throughout that n is sufficiently large for our estimates to hold.

5.3 Counting bipartite graphs

In this section we prove Theorem 2.7 and some facts about bipartite graphs which we will need later on. Similar results are also implicit in Prömel and Steger [68]. As the proofs are short, we include them here for completeness.

First we need some definitions. We say that a bipartite graph with n vertices and m edges together with an admissible bipartition of its vertex set is a *two-coloured* graph. We denote the number of these by $\mathrm{Col2}_{n,m}$ and the number of bipartite graphs by $\mathrm{Bip}_{n,m}$. We say that a vertex bipartition is *almost equitable* if the sizes of the vertex classes differ by at most τ, where

$$\tau = \frac{n(\log n)^{1/4}}{\sqrt{m}}.$$

We denote by $\sum_{A,B}^{eq}$ the sum over all almost equitable bipartitions of n vertices and denote by a and b the size of the corresponding vertex classes. For a fixed bipartition with vertex classes A and B, we call any bipartite graph with vertex classes A and B an *AB-graph.* In what follows, we shall make the following inequality (see page 5 of [8]). For $0 \leq \eta \leq 1/2$,

$$1 - \eta \geq \mathrm{e}^{-(\eta+\eta^2)}. \tag{5.4}$$

Lemma 5.3 *For $m \geq 20n \log n$,*

$$\mathrm{Col2}_{n,m} = (1 + o(1)) \sum_{A,B}^{eq} \binom{ab}{m}.$$

Proof. By inequality (3.1), the number of two-coloured graphs not counted by the above sum is

$$\begin{aligned}
&\sum_{x>\tau/2}^{\lfloor n/2\rfloor} \binom{n}{\lfloor n/2\rfloor - x}\binom{(\lfloor n/2\rfloor - x)(\lceil n/2\rceil + x)}{m} \\
\leq &\binom{n}{\lfloor n/2\rfloor} \sum_{x>\tau/2}^{\lfloor n/2\rfloor} \binom{\lfloor n^2/4\rfloor - x^2}{m} \\
\overset{(3.1)}{\leq} &\binom{n}{\lfloor n/2\rfloor}\binom{\lfloor n^2/4\rfloor}{m} \sum_{x>\tau/2}^{\lfloor n/2\rfloor} \mathrm{e}^{-4mx^2/n^2}.
\end{aligned}$$

The result now follows from the fact that the last line is $o(.)$ of the number of 2-coloured graphs with partition size as equal as possible. □

Lemma 5.4 *Suppose that* $m \geq 20n \log n$. *Fix an almost equitable bipartition with vertex classes* A *and* B. *The probability that an* AB*-graph with* m *edges chosen uniformly at random is disconnected is* $o(1/m^2)$.

Proof. Let $a = |A|$ and $b = |B|$. Let $\sum_{a',b'}$ denote the sum over all ordered pairs of integers a' and b' with $a' + b' \geq 1$ and $a' \leq a/2$ and $b' \leq b/2$. Then the probability that a random AB-graph as above is disconnected is at most

$$\begin{aligned}
& \sum_{a',b'} \binom{a}{a'} \binom{b}{b'} \binom{a'b' + (a-a')(b-b')}{m} \Big/ \binom{ab}{m} \\
\leq \ & \sum_{a',b'} n^{a'+b'} \binom{ab - a'(b-b') - b'(a-a')}{m} \Big/ \binom{ab}{m} \\
\overset{(3.1)}{\leq} \ & \sum_{a',b'} \exp\left\{ a' \left(\log n - \frac{m}{ab}(b-b') \right) + b' \left(\log n - \frac{m}{ab}(a-a') \right) \right\} \\
\leq \ & \sum_{a',b'} \exp\left\{ -a'm/(4a) - b'm/(4b) \right\}.
\end{aligned}$$

□

Corollary 5.5 *For* $m \geq 20n \log n$,

$$\mathrm{Bip}_{n,m} = (1 + o(1)) \sum_{A,B}^{eq} \binom{ab}{m}. \tag{5.5}$$

Moreover, almost all bipartite graphs with n *vertices and* m *edges are uniquely two-colourable.*

Proof. Certainly $\mathrm{Bip}_{n,m} \leq \mathrm{Col2}_{n,m}$. To prove the lower bound, note that a bipartite graph is uniquely two-colourable exactly when it is connected. But by Lemma 5.4, the proportion of two-coloured graphs on n vertices and m edges which do not correspond to a uniquely two-colourable graph is $o(1)$, with room to spare. Hence the result now follows by Lemma 5.3. □

Proof of Theorem 2.7. Note that Corollary 5.5 immediately implies the first part of Theorem 2.7. To prove the "moreover" part, note that Corollary 5.5 states that it suffices to count the number of 2-coloured graphs with an almost equitable bipartition.

First assume that $m = o(n^2)$. Then the upper bound on the number of these follows from the proof of Lemma 5.3, using the fact that

$$\begin{aligned}\sum_{x=0}^{\tau/2} \mathrm{e}^{-4mx^2/n^2} &= (1+o(1)) \int_0^{\tau/2} \mathrm{e}^{-4mx^2/n^2} dx \\ &= (1+o(1)) \frac{n}{\sqrt{8m}} \int_0^{\sqrt{2}(\log n)^{1/4}} \mathrm{e}^{-u^2/2} du \\ &= (1+o(1)) \frac{n}{\sqrt{8m}} \sqrt{\pi/2}.\end{aligned} \tag{5.6}$$

Here we substituted $u = x\sqrt{8m}/n$ to obtain the density function for the normal distribution. To prove the lower bound, we proceed in the same way as in the proof of the upper bound, except that we now make use of the following two inequalities. Fix some δ with $0 < \delta < 1$ and suppose that $x \leq \tau$. Then firstly we have (for $m = o(n^2)$ and using $(\lfloor n/2 \rfloor - x)(\lceil n/2 \rceil + x) \geq \lfloor n^2/4 \rfloor - x^2 - x$) that

$$\begin{aligned}\binom{(\lfloor n/2 \rfloor - x)(\lceil n/2 \rceil + x)}{m} \Big/ \binom{\lfloor n^2/4 \rfloor}{m} &\overset{(3.1)}{\geq} \left(1 - \frac{m}{\lfloor n^2/4 \rfloor - x^2 - x}\right)^{x^2+x} \\ &\overset{(5.4)}{\geq} \mathrm{e}^{-4(1+\delta)m(x^2+x)/n^2} \\ &\geq (1-\delta)\mathrm{e}^{-4(1+\delta)mx^2/n^2},\end{aligned}$$

and secondly we have

$$\binom{n}{\lfloor n/2 \rfloor - x} \Big/ \binom{n}{\lfloor n/2 \rfloor} \geq \frac{(\lfloor n/2 \rfloor)_x}{(\lceil n/2 \rceil + x)_x} \geq \left(1 - \frac{x+1}{\lceil n/2 \rceil}\right)^x \geq 1 - \delta.$$

Here the last inequality follows since $m \geq 20n \log n$ implies that $\tau = o\left(\sqrt{n}\right)$.

If $m/n^2 \not\to 0$, for the upper bound it suffices to note that (5.6) still holds with the $1 + o(1)$ factor replaced by a sufficiently large constant. For the lower bound, it suffices to note that the number of 2-coloured graphs whose bipartition is as equal as possible is already sufficiently large. □

5.4 Proof of the 0-statement of the triangle case

We now consider those m which are below the threshold for bipartiteness, but which are not covered by Theorem 2.4. First observe that the case when $n/2 \leq m \leq c_1 n$, where c_1 is some fixed constant, follows from results of Erdős and Rényi. Indeed, they showed that in this range, the probability that a random graph $G_{n,m}$ is triangle-free is bounded away from zero (see e.g. Chapter IV in [8]). On the other hand, they noted that for $m \geq n/2$, the probability that $G_{n,m}$ is bipartite tends to zero (see page 57 of [27]). So to prove the 0-statement of the $\ell = 3$ case of Theorem 5.1, by Theorem 2.4 it suffices to consider those m with $20n \log n \leq m \leq (1 - \varepsilon)t_3$, where without loss of generality we will assume that $\varepsilon \leq 10^{-6}$.

A $(2, 1)$-*coloured graph* is a graph together with a two-colouring of its vertex set with the property that there is exactly one edge which lies within a colour class and a $(2,1)$-*colourable graph* is one which can be $(2,1)$-coloured. Let $\mathrm{Bip}^{+1}_{n,m}$ denote the number of $(2,1)$-colourable graphs on n vertices with m edges, and let $\mathrm{Col2}^{+1}_{n,m}$ denote the number of $(2,1)$-coloured graphs on n vertices with m edges. Recall that $\mathrm{Bip}_{n,m}$ denotes the number of bipartite graphs on n vertices with m edges.

Throughout, $\sum_{A,B}$ denotes the sum over all bipartitions of n vertices, and we write $a = |A|$ and $b = |B|$.

Proposition 5.6 *If* $20n \log n \leq m = o(n^2)$, *then*

$$\mathrm{Col2}^{+1}_{n,m} = (1 + o(1))m\,\mathrm{Bip}_{n,m}.$$

Proof.

$$\begin{aligned}\mathrm{Col2}^{+1}_{n,m} &= \sum_{A,B} \binom{ab}{m-1}\left(\binom{a}{2} + \binom{b}{2}\right)\\ &= \frac{m}{2}\sum_{A,B} \frac{a(a-1)+b(b-1)}{ab-m+1}\binom{ab}{m}\\ &= (1+o(1))m\sum_{A,B}^{eq}\binom{ab}{m}. \qquad (5.7)\end{aligned}$$

The last line follows by separately considering those bipartitions which are almost equitable and those which are not. Then the proof of Lemma 5.3 shows that the contribution of the latter is $o(1)$ of that of the almost equitable ones. Now Corollary 5.5 implies the result. □

Proposition 5.7 *If* $20n \log n \leq m = o(n^2)$, *then the ratio of the number of* $(2,1)$-*coloured graphs which do not correspond to a uniquely* $(2,1)$-*colourable graph to the number of bipartite graphs tends to zero.*

Proof. Consider the set $\mathcal{C}$ of $(2,1)$-coloured graphs which do not correspond to a uniquely $(2,1)$-colourable graph. The subset of $\mathcal{C}$ where the corresponding $(2,1)$-colourable graph is in fact bipartite contains $o(\mathrm{Bip}_{n,m})$ elements by the “moreover” part of Corollary 5.5.

Now consider the subset of $\mathcal{C}$ where the corresponding $(2,1)$-colourable graph is not bipartite. Then it is easy to see that either the bipartite subgraph obtained by deleting the edge in the colour class is not uniquely two-colourable or that there exists an edge e between A and B so that deleting this edge and the one within the colour class yields a bipartite graph which is not uniquely two-colourable. Lemma 5.4 now implies that the number of elements of $\mathcal{C}$ where this happens, and where the vertex bipartition into A and B is almost equitable, is $o(\mathrm{Col2}^{+1}_{n,m}/m)$. Since the proof of Lemma 5.3 shows that the number of $(2,1)$-coloured graphs with a vertex bipartition which is not almost equitable is also $o(\mathrm{Col2}^{+1}_{n,m}/m)$, the assertion now follows from Proposition 5.6 (since it implies that we may replace the $o(\mathrm{Col2}^{+1}_{n,m}/m)$ by $o(\mathrm{Bip}_{n,m})$.

□

Proof of the 0*-statement of Theorem 5.1 for* $\ell = 3$. As remarked at the beginning of this section, we may assume that $20n \log n \leq m \leq (1-\varepsilon)t_3$. Fix an almost equitable bipartition into classes A and B, fix an edge e in A and let $p = (1+\varepsilon^2)4m/n^2$. Then the probability that a random AB-graph with edge probability p forms no triangle together with the edge e is

$$\left(1-p^2\right)^b \overset{(5.4)}{\geq} \mathrm{e}^{-(p^2+p^4)b} \geq \mathrm{e}^{-(1+\varepsilon^2)^3 8m^2/n^3} =: \nu.$$

Note that $m\nu \geq (1-\varepsilon)t_3\mathrm{e}^{-(1-\varepsilon)8(t_3)^2/n^3} \to \infty$. The same result holds if we fix an edge in B. Now (3.3) implies that the number of edges in the bipartite graph is less than $m-1$ with probability at most $\mathrm{e}^{-cm} = o(1/m)$, for some constant c depending only on ε. Since the probability of being triangle-free is monotone decreasing with m, this implies that the probability that an AB-graph, with $m-1$ edges chosen uniformly at random, forms no triangle together with e is certainly at least $\nu/2$. Thus as in the proof of (5.7), the number of $(2,1)$-coloured graphs not containing a triangle is at least

$$\begin{aligned}\sum_{A,B}^{eq} \binom{ab}{m-1}\left(\binom{a}{2}+\binom{b}{2}\right)\frac{\nu}{2} &= (1+o(1))\frac{\nu}{2}m\sum_{A,B}^{eq}\binom{ab}{m} \\ &\overset{(5.5)}{=} (1+o(1))\frac{\nu}{2}m\,\mathrm{Bip}_{n,m}.\end{aligned}$$

The result now follows from Proposition 5.7 and the fact that $\nu m \to \infty$. □

5.5 Counting almost bipartite graphs

In this section, we reduce the proof of the second 1-statement of the $\ell = 3$ case of Theorem 5.1 to proving that certain graphs (which are close to being bipartite) are triangle-free with sufficiently high probability. Throughout this section, we assume that $m \geq (1+\varepsilon)t_3$, where without loss of generality, we assume that $\varepsilon \leq 10^{-6}$.

As a preliminary step, we partition the set of graphs with m edges and n vertices. Given any graph G, one of its vertices x and a set of its vertices S, we denote by $d_S(x)$ the number of neighbours of x in S. Given a bipartition of the vertices of G into A and B, we say that the bipartite subgraph spanned by A and B *dominates* the graphs induced by A and B if for each vertex $x \in A$ we have $d_A(x) \leq d_B(x)$ and for each vertex $y \in B$ we have $d_B(y) \leq d_A(y)$. We now say that a (k_A, k_B, A, B)*-graph* is a graph together with a vertex bipartition into classes A and B so that the graph spanned by A contains exactly k_A edges and that the one spanned by B contains exactly k_B edges, and furthermore that the bipartite graph spanned by A and B dominates the graph induced by A and B.

We say that a graph G is (k_A, k_B, A, B)*-bipartite* if it forms a (k_A, k_B, A, B)-graph together with the bipartition into A and B. We say that a graph G is k*-bipartite* if k is the minimal number of edges needed to make it bipartite. Note that, for $k = k_A + k_B$, every k-bipartite graph is (k_A, k_B, A, B)-bipartite for some A, B, but that the converse is not necessarily true. Note also that a $(0,0,A,B)$-bipartite graph is bipartite. We will usually write (k_A, k_B) instead of (k_A, k_B, A, B). Let $a = |A|$

and $b = |B|$ again. We will refer to the subgraph of a (k_A, k_B)-graph induced by A as its *k_A-graph.* Equipped with these definitions, we can now write

$$\mathbb{P}[G_{n,m} \text{ is } K_3\text{-free}] \le \sum_{A,B} \sum_{k_A+k_B=0}^{m/2} \mathbb{P}[G_{n,m} \text{ is } K_3\text{-free and } (k_A, k_B)\text{-bip.}]. \tag{5.8}$$

Recall that the sum $\sum_{A,B}$ is over all bipartitions of n vertices. We do not have equality in (5.8) since a graph may well be (k_A, k_B)-bipartite for several different pairs (k_A, k_B). On the other hand, note that for a fixed bipartition into A and B, there is a one-to-one correspondence between (k_A, k_B)-graphs and (k_A, k_B)-bipartite graphs.

Now we classify the (k_A, k_B)-bipartite graphs according to the properties of the subgraphs induced by A and B. For all $r \ge 0$, let

$$k_{A,r} = (1-\varepsilon)^r k_A,$$

let

$$D_A = \frac{\varepsilon^4 m}{n \log(m/k_A)},$$

and define r_A by

$$(1-\varepsilon)^{r_A} = \frac{4 \log\log(m/k_A)}{\varepsilon^2 \log(m/k_A)}. \tag{5.9}$$

We denote by T_A a largest spanning subgraph of the k_A-graph of maximum degree at most D_A and call it the *torso* of the k_A-graph. For $1 \le r < r_A$, we say that a k_A-graph is a *$k_{A,r}$-graph* if its torso has more than $k_{A,r}$ and at most $k_{A,r-1}$ edges. Note that since D_A is larger than the average degree of a k_A-graph, a regular k_A-graph is a $k_{A,1}$-graph and is equal to its torso. A graph has a small torso if it is “far” from being regular. Thus in a sense, the parameter r measures how far a k_A-graph is from being a regular graph. Furthermore note that if $k_A \le (1+\varepsilon)D_A$, then any k_A-graph is a $k_{A,1}$-graph. We call a k_A-graph a *k_{A,r_A}-graph* if it is not a $k_{A,r}$-graph for any $r < r_A$. We define $k_{B,r}$, D_B, T_B and r_B similarly and then extend these definitions to k_B-graphs, (k_A, k_B)-bipartite graphs and (k_A, k_B)-graphs in the obvious way.

For later reference, we note the main consequences of our definitions in the following proposition.

Proposition 5.8 *For $1 < r \le r_A$, consider a $k_{A,r}$-graph and one of its torsos, T_A say. Then A contains a set S, which has at most*

$$s = \frac{2k_{A,r-1}}{D_A} \tag{5.10}$$

vertices, so that each edge not in T_A is adjacent to a vertex in S. Moreover, the number of $k_{A,r}$-graphs is at most

$$\binom{a}{s}\binom{as}{k_A - k_{A,r-1}}\binom{\binom{a}{2}}{k_{A,r-1}}. \tag{5.11}$$

We call the set S a *spine* of the $k_{A,r}$-graph (corresponding to T_A). In what follows, we will often use the fact that (crudely) for $a \geq n/4$,

$$\frac{k_A}{\varepsilon^2 m} \leq \frac{k_A}{\varepsilon^2 n D_A \log(m/k_A)} \leq \frac{s}{a} \leq \frac{8k_A}{nD_A} = \frac{8k_A \log(m/k_A)}{\varepsilon^4 m}. \tag{5.12}$$

In some cases it will suffice to use that the right hand inequality in turn implies that

$$s/a \leq 1/16 \quad \text{and} \quad s \log a \leq k_A. \tag{5.13}$$

Note also that $s \geq 1$.

Proof of Proposition 5.8. Since by definition T_A has at most $k_{A,r-1}$ edges and has maximal degree D_A, there are at most s vertices in A whose degree in T_A is exactly D_A. Let S be the set of these vertices. Every edge not contained in T_A has at least one endvertex in S, since otherwise it would have been included in T_A by the edge-maximality of T_A. To prove (5.11), note that the above implies that every $k_{A,r}$-graph can be constructed by first choosing the vertices of S, then $k_A - k_{A,r-1}$ edges with at least one vertex in S, and then the remaining edges. □

For $k \geq 1$, let

$$f(k,1) = \mathrm{e}^{k \log(20m/k)}.$$

Here and below for convenience we interpret $k \log m/k$ as zero if $k = 0$. For $1 < r \leq r_A$, let

$$f(k,r) = \mathrm{e}^{(1+\varepsilon^2)(1-\varepsilon)^{r-1} k \log(m/k)} = \mathrm{e}^{(1+\varepsilon^2) k_{r-1} \log(m/k)}.$$

Lemma 5.9 *If a, $b \geq n/4$, $m = o(n^2)$, and $k_A + k_B \leq \varepsilon^5 m$,*

$$\mathbb{P}[G_{n,m} \textit{ is } (k_{A,r_1}, k_{B,r_2})\textit{-bip.}] \leq f(k_A, r_1)\, f(k_B, r_2)\, \mathbb{P}[G_{n,m} \textit{ is } (0,0,A,B)\textit{-bip.}]$$

Proof. Consider the case $r_1 = r_2 = 1$ first. We bound the number of $(k_{A,1}, k_{B,1})$-bipartite graphs by the total number of (k_A, k_B)-bipartite graphs. Since the colour classes are fixed, the ratio of (k_A, k_B)-bipartite graphs to $(0,0,A,B)$-bipartite graphs may be bounded by

$$\begin{aligned}
&\binom{\binom{a}{2}}{k_A}\binom{\binom{b}{2}}{k_B}\binom{ab}{m-k_A-k_B}\Big/\binom{ab}{m} \\
&\leq \frac{(a^2/2)^{k_A}}{k_A!}\,\frac{(b^2/2)^{k_B}}{k_B!}\,\frac{(ab)_{m-k_A-k_B}}{(m-k_A-k_B)!}\,\frac{m!}{(ab)_m} \\
&\leq \binom{m}{k_A; k_B}(a^2/2)^{k_A}(b^2/2)^{k_B}\frac{1}{(ab-m)_{k_A+k_B}} \\
&\leq \binom{m}{k_A; k_B}\left(\frac{a}{b}\right)^{k_A}\left(\frac{b}{a}\right)^{k_B} && (5.14)\\
&\leq f(k_A,1)\, f(k_B,1). && (5.15)
\end{aligned}$$

Here (5.15) follows from (5.14) since for the multinomial coefficient we have

$$\binom{m}{k_A; k_B} = \binom{m}{k_A}\binom{m-k_A}{k_B} \leq \binom{m}{k_A}\binom{m}{k_B} \overset{(3.2)}{\leq} \left(\frac{\mathrm{e}m}{k_A}\right)^{k_A}\left(\frac{\mathrm{e}m}{k_B}\right)^{k_B}. \tag{5.16}$$

Now (5.15) follows by substituting this into (5.14) and using a, $b \geq n/4$.

Now we turn to the case $r_1 > 1$ and $r_2 = 1$. For simplicity, let $k = k_A$ and $k' = k_{A,r_1-1}$. By (5.11), the ratio of the number of $(k_{A,r_1}, k_{B,1})$-bipartite graphs to $(0, 0, A, B)$-bipartite graphs may be bounded by

$$\begin{aligned}
& \binom{a}{s}\binom{as}{k-k'}\binom{\binom{a}{2}}{k'}\binom{\binom{b}{2}}{k_B}\binom{ab}{m-k-k_B}\Big/\binom{ab}{m} \\
\overset{(5.15)}{\leq} \quad & f(k,1)f(k_B,1)\binom{a}{s}\binom{as}{k-k'}\binom{\binom{a}{2}}{k'}\Big/\binom{\binom{a}{2}}{k} \\
\leq \quad & f(k,1)f(k_B,1)\, a^s \frac{(as)^{k-k'}}{(k-k')!}\,\frac{a^{2k'}}{k'!}\,\frac{k!}{(a^2/4)^k} \\
\leq \quad & f(k,1)f(k_B,1)\, a^s\, 8^k \left(\frac{s}{a}\right)^{k-k'}, \qquad (5.17)
\end{aligned}$$

where in the last line we used $k!/((k-k')!k'!) \leq 2^k$. Now use (5.12) and also the fact that $s \log a \leq k$ (see (5.13)) to see that (5.17) is at most

$$\begin{aligned}
& f(k,1)f(k_B,1)\,(8\mathrm{e})^k \exp\left\{(k-k')\log\left(\frac{8k\log(m/k)}{\varepsilon^4 m}\right)\right\} \\
\leq \quad & f(k_B,1)\, \exp\left\{2k\log\log(m/k) + k'\log(m/k)\right\} \\
= \quad & f(k,r_1)f(k_B,1)\, \exp\left\{2k\log\log(m/k) - \varepsilon^2 k'\log(m/k)\right\} \\
\leq \quad & f(k,r_1)f(k_B,1). \qquad (5.18)
\end{aligned}$$

In the last line we used the fact that $r_1 \leq r_A$ and (5.9) imply that k/k' is not too large.

Now one may use (5.18) to prove the general case when $r_1 > 1$ and $r_2 > 1$ in the same way. □

We now prove a lemma and a corollary which we shall need in Sections 5.6 and 5.7. They will imply that the bound of Lemma 5.9 is not too far from the truth. The proof of Lemma 5.10 implies that it also holds with A replaced by B. For each k_A, let r_1^* denote the largest r_1 so that the class of k_{A,r_1}-graphs is nonempty, and for k_B, define r_2^* similarly.

Lemma 5.10 *Suppose that $a \geq n/4$, $k_A \leq \varepsilon^5 m$, and $r_1 \leq r_1^*$. With probability at least $(f(k_A, r_1))^{-2\varepsilon^2}$, a k_{A,r_1}-graph chosen uniformly at random has maximum degree at most m/n.*

The probability bound in the lemma is far from best possible, but it suffices for our purposes.

Proof. The case $r_1 = 1$ is straightforward. Indeed, we are immediately done if $k_A \leq m/n$. If not, consider a k_A-graph chosen uniformly at random and for a fixed vertex x in A, let X denote the degree of x in this random graph. Then $\mathbb{E}[X] = 2k_A/a \leq D_A/2$ and moreover, for $0 \leq j < a$,

$$\mathbb{P}[X = j] = \binom{a-1}{j}\binom{\binom{a}{2}-(a-1)}{k_A - j}\Big/\binom{\binom{a}{2}}{k_A}.$$

In other words, X is hypergeometrically distributed. Thus we can apply Theorem 2.10. in [41], which says that the Chernoff bounds (3.3) and (3.4) hold also for such X. Let Z denote the number of vertices of degree at least D_A. Then applying (3.4) with $(1+\delta)\mathbb{E}[X] = D_A$ shows that

$$\mathbb{P}[Z > 0] \le \mathbb{E}[Z] \le a\mathbb{P}[X \ge D_A] \le a\eta^{D_A} \to 0,$$

for some fixed $\eta < 1$. Thus almost all k_A-graphs have maximum degree at most D_A. Any such graph is a $k_{A,1}$-graph and since $D_A \le m/n$, the statement follows.

Now consider the case when $r_1 > 1$. Note that this implies that $k_A \ge (1+\varepsilon)D_A$. As before, we write $k = k_A$ and $k' = k_{A,r_1-1}$ for simplicity. First we prove a lower bound on the number of k_{A,r_1}-graphs with maximum degree at most m/n. Fix a set S' of $s' = \lceil \varepsilon^2 k'/D_A \rceil$ vertices in A. Let $\mathcal{H}_1$ be the set of all graphs with exactly $k-(1-\varepsilon^2)k'$ edges such that each of its edges has one vertex in S' and the other in $A\backslash S'$. Let $\mathcal{H}_1'$ be the set of graphs in $\mathcal{H}_1$ where the vertices in S' all have degree at least D_A and at most m/n, and where the vertices in $A\backslash S'$ all have degree at most $D_A/2$. Consider a graph chosen uniformly at random from $\mathcal{H}_1$. Then the expected degree of a vertex in S' is $(k-(1-\varepsilon^2)k')/s'$, and it is easily seen that

$$(1+\varepsilon/2)D_A \le \frac{k-(1-\varepsilon^2)k'}{s'} \le \frac{kD_A}{\varepsilon^2 k'} \le \frac{m}{n}\frac{\varepsilon^2}{(1-\varepsilon)^{r_A}\log(m/k)} \overset{(5.9)}{\le} \frac{m}{2n}.$$

The expected degree of a vertex in $A\backslash S'$ is at most $k/|A\backslash S'| \le D_A/4$. Similarly to the case when $r_1 = 1$, by counting the expected number of vertices of too large or too small degree, one sees that $|\mathcal{H}_1'| = (1+o(1))|\mathcal{H}_1|$. Now let $\mathcal{H}_2$ be the set of all graphs with vertex set $A\backslash S'$ and exactly $(1-\varepsilon^2)k'$ edges, and let $\mathcal{H}_2'$ be the set of all graphs in $\mathcal{H}_2$ with maximum degree at most $D_A/2$. Again, one can show as before that $|\mathcal{H}_2'| = (1+o(1))|\mathcal{H}_2|$. It is easily seen that all graphs which are the union of a graph in $\mathcal{H}_1'$ and one in $\mathcal{H}_2'$ are k_{A,r_1}-graphs – any torso of such a graph is constructed by taking all edges from the graph in $\mathcal{H}_2'$ and by taking, for each vertex $x \in S'$, D_A of the edges adjacent to x in the graph in $\mathcal{H}_1'$. Thus the above implies that the number of k_{A,r_1}-graphs (with maximum degree at most m/n) is at least

$$|\mathcal{H}_1'|\,|\mathcal{H}_2'| = (1+o(1))\binom{s'(a-s')}{k-(1-\varepsilon^2)k'}\binom{\binom{a-s'}{2}}{(1-\varepsilon^2)k'}. \tag{5.19}$$

The result now follows by comparing this lower bound with the upper bound on the number of k_{A,r_1}-graphs in (5.11). The calculation is similar to the proof of the second part of Lemma 5.9. Note that $s' \ge \varepsilon^2 s/2$.

$$\begin{aligned}
&|\mathcal{H}_1'|\,|\mathcal{H}_2'| \Big/ \binom{a}{s}\binom{as}{k-k'}\binom{\binom{a}{2}}{k'} \\
&\ge \frac{(s'a/2)^{k-(1-\varepsilon^2)k'}\,(a^2/4)^{(1-\varepsilon^2)k'}}{a^{s+2k'}(as)^{k-k'}}\binom{k}{(1-\varepsilon^2)k'}\Big/\binom{k}{k'} \\
&\ge a^{-s}\,(\varepsilon^2/4)^k\,(2s'/a)^{\varepsilon^2 k'}\,2^{-k} \\
&\overset{(5.13)}{\ge} (\varepsilon^2/(8\mathrm{e}))^k(\varepsilon^2 s/a)^{\varepsilon^2 k'} \\
&\ge (k/m)^{3\varepsilon^2 k'/2},
\end{aligned}$$

where the last line follows since (5.9) implies that $k \leq k' \log(m/k)/\log\log(m/k)$ and furthermore by (5.12) we have that $\varepsilon^2 s/a \geq k/m$. □

Corollary 5.11 *Suppose that $|a-b| \leq \varepsilon^2 n$ and that $k_A + k_B \leq \varepsilon^5 m$. The number of (k_{A,r_1}, k_{B,r_2})-graphs is at least $(1+o(1))(f(k_A, r_1) f(k_B, r_2))^{-2\varepsilon^2}$ multiplied with the product of the number of k_{A,r_1}-graphs, the number of k_{B,r_2}-graphs and the number of AB-graphs with $m_1 = m - k_A - k_B$ edges.*

Proof. We count only those (k_{A,r_1}, k_{B,r_2})-graphs where the AB-graph has minimum degree at least m/n and where its k_{A,r_1}-graph and its k_{B,r_2}-graph have maximum degree at most m/n. In an AB-graph with m_1 edges chosen uniformly at random, the vertices in A have average degree $m_1/a \geq (1-\varepsilon)2m/n$ and the vertices in B have average degree $m_1/b \geq (1-\varepsilon)2m/n$. Thus, by counting the expected number of vertices of degree at most m/n, one easily sees that almost all AB-graphs with m_1 edges have minimum degree at least m/n. By Lemma 5.10, the proportion of k_{A,r_1}-graphs with maximum degree m/n is at least $f(k_A, r_1)^{-2\varepsilon^2}$ and also the proportion of k_{B,r_2}-graphs with maximum degree m/n is at least $f(k_B, r_2)^{-2\varepsilon^2}$. □

Lemma 5.12 *The ratio of the number of (k_A, k_B)-bipartite graphs with $k_A + k_B \leq \varepsilon^5 m$, $|a-b| \geq \varepsilon^2 n$, and m edges to the total number of bipartite graphs with m edges tends to zero as n tends to infinity.*

Proof. The proof is similar to the first part of the proof of Lemma 5.9. The number of (k_A, k_B)-bipartite graphs as above is certainly at most

$$\begin{aligned}
&\sum_{A,B\colon |a-b| \geq \varepsilon^2 n} \; \sum_{k_A+k_B \geq 0}^{\varepsilon^5 m} \binom{\binom{a}{2}}{k_A} \binom{\binom{b}{2}}{k_B} \binom{ab}{m - k_A - k_B} \\
&\leq 2^n \sum_{k_A+k_B \geq 0}^{\varepsilon^5 m} \binom{m}{k_A; k_B} n^{2(k_A+k_B)} \frac{\left((1-\varepsilon^4/4)n^2/4\right)_{m-k_A-k_B}}{\left(\lfloor n^2/4 \rfloor\right)_m} \binom{\lfloor n^2/4 \rfloor}{m} \\
&\leq 2^n \binom{\lfloor n^2/4 \rfloor}{m} \sum_{k_A+k_B \geq 0}^{\varepsilon^5 m} \binom{m}{k_A; k_B} n^{2(k_A+k_B)} \frac{\left(1-\varepsilon^4/4\right)^{m-k_A-k_B}}{\left(\lfloor n^2/4 \rfloor - m\right)^{k_A+k_B}} \\
&\overset{(5.16)}{\leq} 2^n \binom{\lfloor n^2/4 \rfloor}{m} \left(\mathrm{e}/\varepsilon^5\right)^{2\varepsilon^5 m} \left(1-\varepsilon^4/4\right)^{m-\varepsilon^5 m} \sum_{k_A+k_B \geq 0}^{\varepsilon^5 m} \frac{n^{2(k_A+k_B)}}{\left(n^2/5\right)^{k_A+k_B}} \\
&\leq 2^n m^2 \left\{ \left(5\mathrm{e}^2/\varepsilon^{10}\right)^{\varepsilon^5} \left(1-\varepsilon^4/4\right)^{1-\varepsilon^5} \right\}^m \binom{\lfloor n^2/4 \rfloor}{m}.
\end{aligned}$$

Since $(1-\varepsilon^4/4)^{1-\varepsilon^5} \leq \mathrm{e}^{-\varepsilon^4/8}$, the term in the curly brackets is strictly less than one. The result now follows since the total number of bipartite graphs with m edges is at least $\binom{\lfloor n^2/4 \rfloor}{m}$. □

Lemma 5.9 and Lemma 5.12 will enable us to deduce the $\ell = 3$ case of Theorem 5.1 once we have proven the following three lemmas. In each of the following lemmas, we assume that $|a-b| \leq \varepsilon^2 n$, $k_A + k_B \leq \varepsilon^{\varepsilon^{-5}} m$, $m = o(n^2)$, $r_1 \leq r_1^*$, and $r_2 \leq r_2^*$.

Lemma 5.13 *If $r_1 < r_A$ and $f(k_A, r_1) \geq f(k_B, r_2)$, then*

$$\mathbb{P}[G_{n,m} \textit{ is } K_3\textit{-free} \mid G_{n,m} \textit{ is } (k_{A,r_1}, k_{B,r_2})\textit{-bipartite}] \leq (f(k_A, r_1))^{-1-\varepsilon^2}.$$

Lemma 5.14 *If $r_1 < r_A$ and $r_2 < r_B$, then*

$$\mathbb{P}[G_{n,m} \textit{ is } K_3\textit{-free} \mid G_{n,m} \textit{ is } (k_{A,r_1}, k_{B,r_2})\textit{-bip.}] \leq (f(k_A, r_1) f(k_B, r_2))^{-1-\varepsilon^2}.$$

Lemma 5.15 *If $r_1 = r_A$ and $f(k_A, r_1) \geq (f(k_B, r_2))^{\varepsilon^2/2}$, then*

$$\mathbb{P}[G_{n,m} \textit{ is } K_3\textit{-free} \mid G_{n,m} \textit{ is } (k_{A,r_1}, k_{B,r_2})\textit{-bipartite}] \leq (f(k_A, r_1))^{-4/\varepsilon^2}.$$

Below, we will use the fact that the proofs of Lemmas 5.13 and 5.15 imply that they also hold with A and B etc. interchanged. *Proof of the second 1-statement of the $\ell = 3$ case of Theorem 5.1 (modulo Lemmas 5.13, 5.14 and 5.15).* We will assume that $m = o(n^2)$, since the result was proven already by Prömel and Steger [70] for larger m. Furthermore, by Theorem 2.5 by Łuczak [59] we have that

$$\begin{aligned} &\mathbb{P}[G_{n,m} \text{ is } K_3\text{-free}] \\ \leq &(1 + o(1))\mathbb{P}[G_{n,m} \text{ is } K_3\text{-free and } k\text{-bipartite for some } k \leq \varepsilon^{\varepsilon^{-5}} m\ ], \end{aligned} \tag{5.20}$$

where without loss of generality we assume that $\varepsilon \leq 10^{-6}$. Since every k-bipartite graph is (k_{A,r_1}, k_{B,r_2})-bipartite for some A, B, r_1, r_2, and some k_A and k_B adding up to k, (5.20) is at most (see also (5.8))

$$\begin{aligned} (1 + o(1)) \sum_{A,B} \sum_{k_A+k_B=0}^{\varepsilon^{\varepsilon^{-5}} m} \sum_{r_1,r_2 \geq 1}^{r_1^*,r_2^*} \mathbb{P}[G_{n,m} \text{ is } (k_{A,r_1}, k_{B,r_2})\text{-bipartite}] \times \\ \times \mathbb{P}[G_{n,m} \text{ is } K_3\text{-free} \mid G_{n,m} \text{ is } (k_{A,r_1}, k_{B,r_2})\text{-bipartite}]. \end{aligned} \tag{5.21}$$

Let $\sum^*_{A,B}$ denote the sum over all bipartitions with $|a - b| \leq \varepsilon^2 n$. By applying Lemma 5.12 and then Lemma 5.9, one sees that (5.21) is at most

$$\begin{aligned} (1 + o(1)) \Bigg\{ \mathbb{P}[G_{n,m} \text{ is bipartite}] + \sum_{A,B}^{*} \sum_{k_A+k_B=1}^{\varepsilon^{\varepsilon^{-5}} m} \sum_{r_1,r_2 \geq 1}^{r_1^*,r_2^*} f(k_A, r_1) f(k_B, r_2) \times \\ \times \mathbb{P}[G_{n,m} \text{ is } K_3\text{-free} \mid G_{n,m} \text{ is } (k_{A,r_1}, k_{B,r_2})\text{-bip.}]\, \mathbb{P}[G_{n,m} \text{ is } (0, 0, A, B)\text{-bip.}] \Bigg\}. \end{aligned} \tag{5.22}$$

For those summands with $r_1 < r_A$ and $r_2 < r_B$ we now apply Lemma 5.14. If $r_1 = r_A$ and $r_2 = r_B$ we apply Lemma 5.15 if $f(k_A, r_1) \geq f(k_B, r_2)$ and Lemma 5.15 with A replaced by B otherwise. For those summands with $r_1 = r_A$ and $r_2 < r_B$ we apply Lemma 5.13 with A replaced by B if $f(k_A, r_1) \leq f(k_B, r_2)^{\varepsilon^2/2}$ and otherwise Lemma 5.15. Similarly, if $r_1 < r_A$ and $r_2 = r_B$, we apply Lemma 5.13 if

$f(k_B, r_2) \leq f(k_A, r_1)^{\varepsilon^2/2}$ and otherwise Lemma 5.15, with A replaced by B. Thus we see that (5.22) is (crudely) at most

$$\begin{aligned}&(1+o(1))\Bigg\{\mathbb{P}[G_{n,m} \text{ is bipartite}] + \\ &+ \sum_{A,B}^{*} \mathbb{P}[G_{n,m} \text{ is } (0,0,A,B)\text{-bip.}] \sum_{k_A+k_B=1}^{\varepsilon^{\varepsilon^{-5}}m} \sum_{r_1,r_2\geq 1}^{r_1^*,r_2^*} \left(f(k_A,r_1)f(k_B,r_2)\right)^{-\varepsilon^2/3}\Bigg\}.\end{aligned}$$

Since it is easily checked that the double sum is $o(1)$, an application of Lemma 5.3 and Corollary 5.5 now completes the proof. □

5.6 Large torsos – Poisson behaviour

This section is devoted to the proofs of Lemmas 5.13 and 5.14. Let

$$m_1 = m - k_A - k_B$$

and note that $m_1 \geq (1-\varepsilon^5)m$.

Lemma 5.16 *Suppose that $|a-b| \leq \varepsilon^2 n$, $k_A + k_B \leq \varepsilon^5 m$, and that $r_1 < r_A$. Fix a k_{A,r_1}-graph G_A and a k_{B,r_2}-graph G_B. Let*

$$p = \frac{(1-\varepsilon^2)m_1}{ab}.$$

Consider a random graph $G_{G_A,G_B,p}$ which is obtained by setting $G_{G_A,G_B,p}[A] = G_A$, $G_{G_A,G_B,p}[B] = G_B$ and including the edges between A and B with probability p independently. Then the probability that $G_{G_A,G_B,p}$ is triangle-free is at most

$$\left(f(k_A,r_1)\right)^{-1-6\varepsilon^2}.$$

Lemma 5.17 *If in addition to the conditions of the previous lemma, we also have $r_2 < r_B$, then the probability that $G_{G_A,G_B,p}$ is triangle-free is at most*

$$\eta = \left(f(k_A,r_1)f(k_B,r_2)\right)^{-1-4\varepsilon^2}.$$

Before we prove Lemmas 5.16 and 5.17, we show how they imply Lemmas 5.13 and 5.14 respectively.

Proof of Lemma 5.14. We first transform from the binomial $G_{G_A,G_B,p}$-model to one with a fixed number of edges. Fix a k_{A,r_1}-graph G_A and a k_{B,r_2}-graph G_B. Let G_{G_A,G_B,m_1} be a random graph chosen like $G_{G_A,G_B,p}$, but with the difference that the AB-graph is chosen uniformly at random from the set of all AB-graphs with exactly m_1 edges. Alternatively, one may view G_{G_A,G_B,m_1} as a random graph obtained by picking a graph $G_{n,m}$ in $\mathcal{G}(n,m)$ uniformly at random, conditional on $G_{n,m}[A] = G_A$ and $G_{n,m}[B] = G_B$. Let $\mathcal{G}_{G_A,G_B,m_1}$ be the set of such graphs.

By the Chernoff bound (3.4), the probability that in $G_{G_A,G_B,p}$, the AB-graph has more than m_1 edges is at most e^{-cm_1}, for some constant c depending only on ε. The probability of being triangle-free is monotone decreasing with m_1, and thus

$$\begin{aligned}\mathbb{P}[G_{G_A,G_B,m_1} \text{ is } K_3\text{-free}] &\leq \mathbb{P}[G_{G_A,G_B,p} \text{ is } K_3\text{-free}] + \mathrm{e}^{-cm_1}\\ &\leq \mathbb{P}[G_{G_A,G_B,p} \text{ is } K_3\text{-free}] + \eta\\ &\leq 2\eta. \end{aligned}\tag{5.23}$$

In the second inequality we used that k is small compared to m_1 and the last inequality follows from Lemma 5.17.

Let $\mathcal{A}$ denote the event that a (k_{A,r_1}, k_{B,r_2})-graph (chosen uniformly at random) is triangle-free. Also, for every pair G_A, G_B (where G_A is a k_{A,r_1}-graph and G_B is a k_{B,r_2}-graph), let $\mathcal{F}_{G_A,G_B}$ denote the event that a random (k_{A,r_1}, k_{B,r_2})-graph G satisfies $G[A] = G_A$ and $G[B] = G_B$. Let $\mathcal{D}_{G_A,G_B}$ denote the event that G_{G_A,G_B,m_1} is a (k_{A,r_1}, k_{B,r_2})-graph. Note that $\mathcal{D}_{G_A,G_B}$ is exactly the event that the AB-graph in G_{G_A,G_B,m_1} dominates G_A and G_B. For fixed A and B, there is a one-to-one correspondence between (k_{A,r_1}, k_{B,r_2})-graphs and (k_{A,r_1}, k_{B,r_2})-bipartite graphs in $\mathcal{G}(n,m)$. This implies that it makes sense to consider the events defined above in the uniform probability measure on $\mathcal{G}(n,m)$. Also note that it implies that $\mathbb{P}[\mathcal{A}]$ is exactly the probability we are aiming to bound. Then

$$\begin{aligned}\mathbb{P}[\mathcal{A}] &= \sum_{G_A,G_B} \mathbb{P}[\mathcal{A} \mid \mathcal{F}_{G_A,G_B}]\,\mathbb{P}[\mathcal{F}_{G_A,G_B}]\\ &= \sum_{G_A,G_B} \mathbb{P}[G_{G_A,G_B,m_1} \text{ is } K_3\text{-free} \mid \mathcal{D}_{G_A,G_B}]\,\mathbb{P}[\mathcal{F}_{G_A,G_B}]\\ &\leq \sum_{G_A,G_B} \frac{\mathbb{P}[G_{G_A,G_B,m_1} \text{ is } K_3\text{-free}]}{\mathbb{P}[\mathcal{D}_{G_A,G_B}]}\,\mathbb{P}[\mathcal{F}_{G_A,G_B}]\\ &\overset{(5.23)}{\leq} 2\eta \sum_{G_A,G_B} \frac{\mathbb{P}[\mathcal{F}_{G_A,G_B}]}{\mathbb{P}[\mathcal{D}_{G_A,G_B}]}.\end{aligned}$$

But $\mathbb{P}[\mathcal{D}_{G_A,G_B}]$ is exactly the number of graphs in $\mathcal{G}(n,m)$ dominating G_A and G_B divided by $|\mathcal{G}_{G_A,G_B,m_1}|$. On the other hand, $\mathbb{P}[\mathcal{F}_{G_A,G_B}]$ is exactly the number of graphs in $\mathcal{G}(n,m)$ dominating G_A and G_B divided by the total number of (k_{A,r_1}, k_{B,r_2})-graphs. Thus, writing $\#\{k_{A,r_1}, k_{B,r_2}\}$ for the number of these, Corollary 5.11 implies that

$$\mathbb{P}[\mathcal{A}] \leq 2\eta \sum_{G_A,G_B} \frac{|\mathcal{G}_{G_A,G_B,m_1}|}{\#\{k_{A,r_1}, k_{B,r_2}\}} \leq 2\eta(1+o(1))(f(k_A,r_1)f(k_B,r_2))^{2\varepsilon^2},$$

as required. □

The proof of Lemma 5.13 is the same (except that in the end, we use the assumption that $f(k_A,r_1) \geq f(k_B,r_2)$), and is therefore omitted.

The large deviation inequalities stated in Section 3.2 will be an essential ingredient in the proofs of Lemmas 5.16 and 5.17.

Proof of Lemma 5.16. The proof here is similar to that of Lemma 5.17. It is simplified by the fact that we count only triangles with an edge in A. Thus we can ignore all terms involving k_{B,r_2}. □

Proof of Lemma 5.17. Fix torsos T_A and T_B of G_A and G_B and consider the setting of Section 3.2. In our case, $\mathcal{J}$ is the set of edges between A and B, and $\mathcal{S}$ consists of all those pairs of adjacent edges xz', yz' between A and B so that xy is contained either in T_A or T_B. Thus X is the number of triangles in $G_{G_A,G_B,p}$ with one edge in a torso and the opposite vertex in the opposite vertex class. For any $\alpha \in \mathcal{S}$, we have $\mathbb{E}[I_\alpha] = p^2$. Certainly $G_{G_A,G_B,p}$ is triangle-free only if X is zero. Furthermore, we have

$$\mu = \mathbb{E}[X] \geq (bk_{A,r_1} + ak_{B,r_2})p^2 \geq (k_{A,r_1} + k_{B,r_2})\frac{n(1-\varepsilon^2)p^2}{2} \tag{5.24}$$

$$\geq (k_{A,r_1} + k_{B,r_2})\frac{8(1-\varepsilon^2)^9 m^2}{n^3}, \tag{5.25}$$

and

$$\Delta = \sum_{xy,yz}\sum_{z'} \mathbb{P}[xz',\, yz',\, zz' \in G_{G_A,G_B,p}] + \sum_{xy,x'y'} 8\,\mathbb{P}[xx',\, yy',\, yx' \in G_{G_A,G_B,p}].$$

Here $\sum_{xy,yz}$ denotes the sum over the set of ordered pairs of edges in the torsos T_A and T_B which are adjacent. Note that the number of summands is at most $k_{A,r_1-1}D_A + k_{B,r_2-1}D_B$. $\sum_{xy,x'y'}$ denotes the sum over the set of pairs of edges with xy in T_A and $x'y'$ in T_B. The number of summands here is $k_{A,r_1-1}k_{B,r_2-1} \leq k_{A,r_1-1}nD_B$. The factor 8 is due to the fact that there are 8 possible ways of extending two given opposite edges into two ordered adjacent triangles. Without loss of generality we will assume in the remainder of the proof that $k_A \geq k_B$ (which implies that $D_A \geq D_B$). Thus

$$\Delta = \left(\sum_{xy,yz} n + \sum_{xy,x'y'} 8\right)p^3 \leq (k_{A,r_1-1}D_A + k_{B,r_2-1}D_B + 8k_{A,r_1-1}D_B)np^3$$

$$\leq 9(k_{A,r_1} + k_{B,r_2})D_A np^3 \overset{(5.24)}{\leq} 25\mu D_A p. \tag{5.26}$$

Furthermore note that by Lemma 5.16 (with A replaced by B) we may assume that

$$\log f(k_A, r_1) \geq \frac{\varepsilon^2}{2}\log f(k_B, r_2). \tag{5.27}$$

Suppose first that $\Delta > \mu$. Crudely, we have

$$\begin{aligned}\frac{\mu^2}{2\Delta} &\overset{(5.26)}{\geq} \frac{\mu}{50D_A p} \overset{(5.24)}{\geq} (k_{A,r_1} + k_{B,r_2})\frac{np}{101D_A} \geq k_{A,r_1}\frac{n\log(m/k_A)}{102\varepsilon^4 m}\frac{4m}{n} \\ &\geq \frac{4}{\varepsilon^2}\log f(k_A, r_1) \overset{(5.27)}{\geq} (1+4\varepsilon^2)(\log f(k_A, r_1) + \log f(k_B, r_2)),\end{aligned}$$

and so we are done by correlation inequality (3.6) in this case. So henceforth assume that $\Delta \leq \mu$. Thus by correlation inequality (3.5) we will be done whenever we can show that

$$\mu/2 \geq (1+4\varepsilon^2)(\log f(k_A, r_1) + \log f(k_B, r_2)).$$

Now consider the case when $k_A \geq n^{-3\varepsilon^2/16}m$. This implies

$$\begin{aligned} \mu &\overset{(5.25)}{\geq} k_{A,r_1}(1-\varepsilon^2)^9(1+\varepsilon)^2\log(n^{3/2}) \geq k_{A,r_1-1}\log(n^{3/2}) \\ &\geq k_{A,r_1-1}\frac{8}{\varepsilon^2}\log(m/k_A) \geq \frac{8}{\varepsilon^2(1+\varepsilon^2)}\log f(k_A,r_1) \\ &\overset{(5.27)}{\geq} 2(1+4\varepsilon^2)(\log f(k_A,r_1)+\log f(k_B,r_2)). \end{aligned}$$

So henceforth assume that $k_A \leq n^{-3\varepsilon^2/16}m$. Now suppose additionally that $m \leq 3t_3$. Then

$$\frac{\Delta}{\mu} \overset{(5.26)}{\leq} \frac{25\varepsilon^4 m}{n\log(m/k_A)}\frac{4m}{n^2} \leq 100\varepsilon^4\frac{9\cdot 16(t_3)^2}{3\varepsilon^2 n^3\log n} \leq 1000\varepsilon^2. \tag{5.28}$$

On the other hand, we have that, using $m \geq (1+\varepsilon)t_3$ in the first line and $m \leq 3t_3$ in the third line,

$$\begin{aligned} \mu &\overset{(5.25)}{\geq} (k_{A,r_1}+k_{B,r_2})(1-\varepsilon^2)^5(1+\varepsilon)^2\log(n^{3/2}) \\ &\geq (1+2001\varepsilon^2)(k_{A,r_1-1}+k_{B,r_2-1})\log(n^{3/2}) \\ &\geq (1+2000\varepsilon^2)(\log f(k_A,r_1)+\log f(k_B,r_2)), \end{aligned} \tag{5.29}$$

Hence in this case, inequality (5.28) and correlation inequality (3.5) give the desired result.

If $m \geq 3t_3$, using (5.25), it is straightforward to check that this implies

$$\mu \geq 4(\log f(k_A,r_1)+\log f(k_B,r_2)).$$

Since we assumed that $\Delta \leq \mu$, we are again done by correlation inequality (3.5) as observed above. ☐

We remark that the only place in the entire proof where we made essential use of the condition $m \geq (1+\varepsilon)t_3$ was to prove (5.29) when $k_A + k_B$ is small. The larger $k_A + k_B$ is, the more the condition on m can be relaxed. Indeed, the proof can be extended to show that for any fixed ε, $\delta > 0$, if $k_0 \leq n^{3/2-\delta}$ and

$$m \geq (1+\varepsilon)\frac{\sqrt{3}}{4}n^{3/2}\left(\log n - \frac{2}{3}\log k_0\right)^{1/2} = (1+\varepsilon)\left((n/2)^3\log\left(n^{3/2}/k_0\right)\right)^{1/2}, \tag{5.30}$$

then almost all graphs in $\mathcal{T}(n,m)$ can be made bipartite by deleting at most k_0 edges.

5.7 Small torsos – expanding neighbourhoods

The aim of this section is to prove Lemma 5.15. Its proof hinges on the fact that in a (k_{A,r_A}, k_{B,r_2})-graph, sufficiently large sets in B will have neighbourhoods of size almost $|A|$ in A with very large probability. We first prove this in a more tractable probability model (Lemma 5.18).

Let $k = k_A$ and let A' be a given subset of A with $|A'| \geq a - s$, where $s = 2k_{A,r_A-1}/D_A$ as in (5.10). Consider a random $A'B$-graph with edge probability $p_0 = m_0/(ab)$, where $m_0 \geq m/4$. We say that a set $V \subset B$ *expands* if its neighbourhood $\Gamma_{A'}(V)$ in A' contains at least $a - y$ vertices, where

$$y = a/\log^{400/\varepsilon^4}(m/k). \tag{5.31}$$

Note that this definition makes sense, since one may check (using $k/m \leq \varepsilon^{\varepsilon^{-5}}$ and $\varepsilon \leq 10^{-6}$) that $s \leq y/2$. Furthermore we let

$$D'_A = \frac{\varepsilon^6 m}{40n \log\log(m/k)}.$$

For a sequence $V_1, \dots, V_h$ of subsets of B we define its *total weight* to be equal to

$$\sum_{i=1}^{h} |V_i|.$$

Lemma 5.18 *Consider the setting of the previous paragraph. Given $h \leq k$, for $1 \leq i \leq h$ let v_i satisfy $v_i \geq D'_A$. Furthermore, suppose that $\sum_{i=1}^{h} v_i \geq k/4$. Let $V_1, \dots, V_h$ with $|V_i| = v_i$ be chosen uniformly and independently at random in B. Then*

$$\mathbb{P}[\textit{the expanding } V_i \textit{ have total weight at most } k/8] \leq f(k, r_A)^{-18/\varepsilon^2}.$$

In later calculations we shall make use of the fact that

$$\log f(k, r_A) < \frac{5}{\varepsilon^2} k \log\log(m/k). \tag{5.32}$$

The basic idea of the proof of this lemma is based on that of Lemma 11 in Kohayakawa, Łuczak and Rödl [52]. Consider selecting h' subsets W_i of B sequentially. Inductively, we now define whether some W_i is *useful* or not. For $i \geq 1$, let $\mathcal{U}_i$ be the set of useful W_j with $j < i$ and let

$$u_i = |\bigcup_{j:\, W_j \in \mathcal{U}_i} W_j|.$$

We then say that W_i is *useful* if at least one of the following holds.

- $|u_i| \geq y'$, where $y' = \frac{b}{\log^{4000/\varepsilon^4}(m/k)}$.
- $|W_i \setminus \bigcup_{j:\, W_j \in \mathcal{U}_i} W_j| \geq |W_i|/2$.

Otherwise call it *not useful.* Thus in particular, W_1 is always useful.

Lemma 5.19 *Given $h' \leq k$ and a sequence of integers $w_i > 0$ with $1 \leq i \leq h'$, suppose that $\sum_{i=1}^{h'} w_i \geq k/8$. Let $W_1, \dots, W_{h'}$ with $|W_i| = w_i$ be sets chosen uniformly and independently at random in B. Then*

$$\mathbb{P}[\textit{the useful } W_i \textit{ have total weight at most } k/16] \leq f(k, r_A)^{-19/\varepsilon^2}.$$

Proof. Fix $i \geq 1$ and let

$$\eta_i = \left(\frac{\log^{4000/\varepsilon^4}(m/k)}{2\mathrm{e}} \right)^{-|W_i|/2}.$$

Then for any fixed u^* and i, we claim that

$$\mathbb{P}[W_i \text{ is not useful} \,|\, u_i = u^*] \leq \eta_i. \tag{5.33}$$

Indeed, if $u^* \geq y'$, this follows immediately from the definition. If $u^* < y'$, we apply the Chernoff bound (3.4), with $X = |W_i \cap \bigcup_{j\colon W_j \in \mathcal{U}_i} W_j|$, $p' = u^*/b$, $N = |W_i|$, and $1{+}\delta = b/(2u^*)$, noting that the right hand side of (3.4) is at most $((1{+}\delta)/\mathrm{e})^{-(1+\delta)Np'}$. Since the event that W_i is useful depends only on u_i, (5.33) in turn implies that for any event $\mathcal{A}$ depending only on $W_1, \dots, W_{i-1}$, we have

$$\mathbb{P}[W_i \text{ is not useful} \,|\, \mathcal{A}] \leq \eta_i.$$

Thus for any fixed subsequence $\mathcal{S}$ of the W_i which has total weight at least $k/16$, the probability that none of the sets in $\mathcal{S}$ are useful is at most

$$\prod_{i\colon W_i \in \mathcal{S}} \eta_i \leq \left(\frac{\log^{4000/\varepsilon^4}(m/k)}{2\mathrm{e}} \right)^{-k/32} \overset{(5.32)}{\leq} 2^{-k} f(k, r_A)^{-19/\varepsilon^2}.$$

Now note that if the useful W_i have total weight at most $k/16$, then there must exist a subsequence $\mathcal{S}$ of the W_i, which has total weight at least $k/16$ and where each W_i in $\mathcal{S}$ is not useful. Multiplying by $2^{h'} \leq 2^k$ (to account for all possible choices of $\mathcal{S}$) yields the result. □

Corollary 5.20 *Given $h' \leq k$ and a sequence of integers w_i with $1 \leq i \leq h'$, suppose that $\sum_{i=1}^{h'} w_i \geq k/8$. Let $W_1, \dots, W_{h'}$ with $|W_i| = w_i$ be sets chosen uniformly and independently at random in B. Let*

$$u = \min\left\{k/16\,,\, y'\right\}.$$

Then with probability at least $1 - f(k, r_A)^{-19/\varepsilon^2}$ there exist (after relabelling) disjoint sets $W''_1, \dots, W''_{h''}$ of total weight at least $u/2$ so that $W''_i \subseteq W_i$ and $|W''_i| \geq |W_i|/2$ for all i.

Proof. By the previous lemma, with sufficiently high probability the sequence $\mathcal{W} = (W_1, \dots, W_{h'})$ contains a sequence $\mathcal{U} = (U_1, \dots, U_{h^*})$ of useful sets of total weight at least $k/16$. If $k/16 \leq y'$, set $h'' = h^*$. Otherwise, let h'' be the smallest number satisfying

$$|\bigcup_{j=1}^{h''} U_j| \geq y'.$$

For each $i \leq h''$, let $W_i'' = U_i \setminus \bigcup_{j=1}^{i-1} U_j$ and observe that by the definition of "useful", the W_i'' have the desired properties. □

Proof of Lemma 5.18. First we calculate the probability that a fixed subsequence $W_1, \dots, W_{h'}$ (of total weight at least $k/8$) of the sequence of sets $V_1, \dots, V_h$ contains no expanding sets. By the previous corollary, we may condition on the fact that the W_i contain a sequence $W_1'', \dots, W_{h''}''$ of sets as in the statement of that corollary. Also, for $a' = |A'|$,

$$\mathbb{P}[W_i'' \text{ is not expanding}] \leq \binom{a'}{a-y}(1-p_0)^{(a'-(a-y))|W_i''|}.$$

But the definition of y in (5.31) and the fact that $a' \geq a - y/2$ imply that $y \leq a'/2$. Thus

$$\binom{a'}{a-y} = \binom{a'}{a'-(a-y)} \leq \binom{a'}{y} \leq \binom{a}{y}.$$

Hence we have, using $a' \geq a - y/2$ again and (3.2),

$$\begin{aligned} \mathbb{P}[W_i'' \text{ is not expanding}] &\leq \exp\{y(\log(\mathrm{e}a/y) - p_0|W_i''|/2)\} \\ &\overset{(5.31)}{\leq} \exp\{y((400/\varepsilon^4)(1+\log\log(m/k)) - p_0|W_i''|/2)\} \\ &\leq \exp\{-yp_0|W_i''|/4\}, \end{aligned}$$

where we used in the last line that $|W_i''| \geq D_A'/2$. The disjointness of the W_i'' then implies that

$$\begin{aligned} \mathbb{P}[\text{none of the } W_i'' \text{ expand}] &\leq \mathrm{e}^{-uyp_0/2^3} \leq \mathrm{e}^{-uym/(2^5 ab)} \\ &\overset{(5.31)}{\leq} \mathrm{e}^{-\frac{um}{2^5 b \log^{400/\varepsilon^4}(m/k)}} \\ &\leq \max\left\{\mathrm{e}^{-\frac{km}{2^9 b \log^{400/\varepsilon^4}(m/k)}}, \mathrm{e}^{-\frac{m}{2^5 \log^{4400/\varepsilon^4}(m/k)}}\right\} \\ &\overset{(5.32)}{\leq} f(k, r_A)^{-19/\varepsilon^2}. \end{aligned}$$

In the last line we used the fact that $k/m \leq \varepsilon^{\varepsilon^{-5}}$ and $\varepsilon \leq 10^{-6}$. Thus in particular

$$\mathbb{P}\left[\text{none of the } W_i \text{ expand}\right] \leq f(k, r_A)^{-19/\varepsilon^2}.$$

As in the proof of Lemma 5.19, the result now follows if we sum over all possibilities of choosing $\mathcal{W}'$ and thus multiplying the above probability by $2^h \leq 2^k$. □

Now we derive Lemma 5.15 from Lemma 5.18. For convenience let $k = k_A$ and $k' = k_{A,r_A-1}$ again. We say that a k_{A,r_A}-graph G_A is *friendly* if with probability at least $f(k, r_A)^{-6/\varepsilon^2}$ it is dominated by an AB-graph with $m_1 = m - k - k_B$ edges chosen uniformly at random.

For a given set $S' \subset A$, we let $A' = A \setminus S'$. We say that a vertex x *expands* if

$$\Gamma_{A'}(\Gamma_B(x)) \geq a - y.$$

We say that a set of vertices *expands well* if each of its vertices expands.

Lemma 5.21 *Fix a friendly k_{A,r_A}-graph G_A and choose an AB-graph uniformly at random from the set of AB-graphs with m_1 edges dominating G_A. Let S be the lexicographically first spine of G_A. With probability at least $1 - f(k, r_A)^{-11/\varepsilon^2}$, S contains a well-expanding set S' with*

$$\sum_{x \in S'} d_A(x) \geq k/8. \tag{5.34}$$

Proof. By Proposition 5.8, the vertices of S are adjacent to at least $k - k' \geq k/2$ edges of G_A. Recall also that $|S| \leq s = 2k'/D_A$. Thus the average degree in G_A of the vertices in S is at least

$$\frac{k}{2s} \overset{(5.10)}{\geq} \frac{D_A}{5(1-\varepsilon)^{r_A}} \overset{(5.9)}{=} \frac{\varepsilon^6 m}{20n \log\log(m/k)} = 2D'_A. \tag{5.35}$$

Let S' be the subset of S which contains all those vertices whose degree in G_A is at least D'_A. Then by (5.35), the vertices in $S \backslash S'$ can be adjacent to at most $|S \backslash S'| D'_A \leq s D'_A \leq k/4$ edges in G_A, and thus the vertices in S' are still adjacent to at least $k/4$ edges in G_A. For any AB-graph with m_1 edges, let

$$s_B = \sum_{x \in S'} d_B(x)$$

and let $s' = |S'|$. Note that by definition, we have $s_B \leq s'b$. Then the proportion of AB-graphs with $s_B \geq m_1/2$ is at most

$$\begin{aligned} \binom{s'b}{s_B}\binom{(a-s')b}{m_1 - s_B}\Big/\binom{ab}{m_1} &\leq \binom{m_1}{s_B}\frac{(s'b)^{s_B}(ab)^{m_1 - s_B}}{(ab/2)^{m_1}} \\ &\leq 4^{m_1}(s'/a)^{s_B} \overset{(5.13)}{\leq} 4^{-m_1} \leq 2^{-m} \\ &\overset{(5.32)}{\leq} f(k, r_A)^{-18/\varepsilon^2}. \end{aligned}$$

Since for any two events $\mathcal{D}$ and $\mathcal{E}$, one has $\mathbb{P}[\mathcal{D}|\mathcal{E}] \leq \mathbb{P}[\mathcal{D}]/\mathbb{P}[\mathcal{E}]$, by the definition of "friendly", the above implies that the proportion of AB-graphs as in the statement of the Lemma with $s_B \geq m_1/2$ is at most $f(k, r_A)^{-12/\varepsilon^2}$. Now consider some fixed sequence of degrees $d_B(x)$ for all $x \in S'$, where we require that $d_B(x) \geq d_A(x)$ and $s_B \leq m_1/2$. Let $\mathbb{P}_{m_2,S'}$ denote the probability measure on the set of all AB-graphs where the $S'B$-graph is uniformly chosen from all $S'B$-graphs which conform to the above degree sequence and the $A'B$-graph is uniformly chosen from the set of all $A'B$-graphs with m_2 edges, where

$$m_2 = m - k - k_B - s_B.$$

Note that $m_2 \geq m/4$. Correspondingly, let $\mathbb{P}_{p_2,S'}$ denote the probability measure on the set of all AB-graphs where we choose the $S'B$-graph as above and where the $A'B$-graph has edge probability p_2, where $p_2 = m_2/(ab)$. Let S'' denote the set of expanding vertices in S'. The lemma will follow once we have shown that

$$\mathbb{P}_{m_2,S'}[\sum_{x \in S''} d_A(x) \leq k/8 \mid AB\text{-graph dominates } G_A] \leq m_2 f(k, r_A)^{-12/\varepsilon^2}. \tag{5.36}$$

But by definition, the left hand side can be bounded by

$$\begin{aligned}\mathbb{P}_{m_2,S'}[\sum_{x\in S''} d_A(x) \leq k/8] \,/\, \mathbb{P}_{m_2,S'}[AB\text{-graph dominates } G_A] \\ \leq \mathbb{P}_{m_2,S'}[\sum_{x\in S''} d_A(x) \leq k/8]\, f(k,r_A)^{6/\varepsilon^2}.\end{aligned}$$

Furthermore, the proof of Pittel's inequality (see pages 4 and 35 of [8]) implies that

$$\begin{aligned}\mathbb{P}_{p_2,S'}[\sum_{x\in S''} d_A(x) \leq k/8] \\ \geq \binom{\lfloor n^2/4\rfloor}{m_2} p_2^{m_2}(1-p_2)^{\lfloor n^2/4\rfloor - m_2}\mathbb{P}_{m_2,S'}[\sum_{x\in S''} d_A(x) \leq k/8] \\ \geq \frac{1}{m_2}\mathbb{P}_{m_2,S'}[\sum_{x\in S''} d_A(x) \leq k/8].\end{aligned} \tag{5.37}$$

Write $S' = \{x_1, \dots, x_h\}$. To bound (5.37) from above, we apply Lemma 5.18 with $A' = A\backslash S'$, $h = |S'| \leq k$ and where for all i, the V_i are subsets of $\Gamma_B(x_i)$ of size exactly $d_A(x_i)$, chosen uniformly at random in $\Gamma_B(x_i)$. Recall that $d_B(x_i) \geq d_A(x_i) \geq D'_A$ and $\sum_{x\in S'} d_B(x) \geq k/4$. Furthermore note that the definition of $\mathbb{P}_{p_2,S'}$ implies that the sets $\Gamma_B(x)$ are independently and uniformly distributed in B. This in turn implies that the V_i are then also uniformly distributed in B. Thus the conditions of Lemma 5.18 are satisfied and inequality (5.36) follows. □

Given a friendly k_{A,r_A}-graph G_A, we say that a k_{B,r_2}-graph G_B is *friendly with respect to* G_A if with probability at least $f(k,r_A)^{-6/\varepsilon^2}$ it is dominated by an AB-graph chosen uniformly at random from the set of AB-graphs with m_1 edges dominating G_A. In this case, we say that G_A and G_B form a *friendly pair* and that any (k_{A,r_A}, k_{B,r_2})-graph G with $G[A] = G_A$ and $G[B] = G_B$ is *friendly.* Otherwise we call it an *unfriendly pair* and call any corresponding (k_{A,r_A}, k_{B,r_2})-graph G *unfriendly.*

Corollary 5.22 *Choose a friendly (k_{A,r_A}, k_{B,r_2})-graph uniformly at random and let S be the lexicographically first spine of the k_{A,r_A}-graph. With probability at least $1 - f(k,r_A)^{-5/\varepsilon^2}$, S contains a well-expanding set S'' with*

$$\sum_{x\in S''} d_A(x) \geq k/8. \tag{5.38}$$

Proof. Fix a friendly pair G_A and G_B. Consider choosing an AB-graph uniformly at random from the set of AB graphs with m_1 edges dominating G_A. Let $\mathcal{D}$ be the event that it also dominates G_B. Let $\mathcal{E}$ be the event that S contains a set S'' satisfying (5.38). Let $\mathcal{E}'$ be the event $\mathcal{E}$, conditional on $\mathcal{D}$ and note that we are done once we have shown that $\mathbb{P}[\mathcal{E}'] \leq f(k,r_A)^{-5/\varepsilon^2}$. But $\mathbb{P}[\mathcal{E}'] = \mathbb{P}[\mathcal{E} \mid \mathcal{D}] \leq \mathbb{P}[\mathcal{E}]/\mathbb{P}[\mathcal{D}]$, and thus we are done by Lemma 5.21 □

Note that the reason why we introduced the notion of a dominating AB-graph in the definition of a (k_{A,r_A}, k_{B,r_2})-graph was that it makes it easier to prove a result like Corollary 5.22 – if for instance the AB-graph is chosen uniformly at random from the set of AB-graphs with m_1 edges and if S is a fixed set of vertices in A with $|S| \leq \varepsilon^3 s$, then one may check that even the probability that S is isolated is larger than $f(k, r_A)^{-5/\varepsilon^2}$. Finally, we are in a position to deal with (k_{A,r_A}, k_{B,r_2})-graphs.

Proof of Lemma 5.15. We first consider the (k_{A,r_A}, k_{B,r_2})-graphs which are unfriendly. Denoting the sum over all unfriendly pairs by $\sum_{G_A,G_B}^{unf}$, the number of these is

$$\sum_{G_A,G_B}^{unf} \|G_{AB} \colon G_{AB} \text{ is an } AB\text{-graph with } m_1 \text{ edges dominating } G_A \text{ and } G_B\|.$$

We crudely bound the number of summands by the product of the total number of k_{A,r_A}-graphs and k_{B,r_2}-graphs. By definition, each summand is at most the number of AB-graphs multiplied by $f(k, r_A)^{-6/\varepsilon^2}$. Comparing the resulting upper bound with the lower bound on the total number of (k_{A,r_A}, k_{B,r_2})-graphs in Corollary 5.11, one readily sees that the probability that a (k_{A,r_A}, k_{B,r_2})-graph (chosen uniformly at random) is unfriendly is at most

$$(1 + o(1)) f(k, r_A)^{-6/\varepsilon^2 + 2\varepsilon^2} f(k_B, r_2)^{2\varepsilon^2} \leq f(k, r_A)^{-6/\varepsilon^2 + 5},$$

where the final inequality comes from our assumption made in the statement of the lemma.

Thus it suffices to show that the probability that a friendly (k_{A,r_A}, k_{B,r_2})-graph is triangle-free is at most $2f(k, r_A)^{-5/\varepsilon^2}$. For a (k_{A,r_A}, k_{B,r_2})-graph G, let S be the lexicographically first spine in $G[A]$. We say that the AB-graph in G is S-*good* if S contains a well-expanding subset S'' satisfying (5.38). By Corollary 5.22, the probability that in a friendly (k_{A,r_A}, k_{B,r_2})-graph, the AB-graph is not S-good is at most $f(k, r_A)^{-5/\varepsilon^2}$.

Now suppose that in a friendly (k_{A,r_A}, k_{B,r_2})-graph the AB-graph is S-good. Consider first a fixed set S'' in A and a fixed AB-graph which is S-good (for some S containing S'') for at least one k_{A,r_A}-graph and so that S'' is well expanding and satisfies (5.38). Write $S'' = \{x_1, \dots, x_{s''}\}$. But then, for a vertex $x_i \in S''$, there are at most y possibilities for neighbours of x_i in A which do not produce a triangle together with the AB-graph. For $1 \leq i \leq s''$, let $d_A(x_i)$ be a sequence of nonnegative integers satisfying $d_A(x_i) \leq y$ and adding up to exactly d_A, where $d_A \geq k/8$. Let $\sum^{d_A}$ denote the sum over all such sequences. Thus, similarly to the proof of Proposition 5.8, the number of k_{A,r_A}-graphs not producing a triangle together with the AB-graph is at most

$$\begin{aligned}
U &:= \sum_{d_A = k/8}^{s''y} \sum^{d_A} \left(\prod_{i=1}^{s''} \binom{y}{d_A(x_i)} \right) \binom{sa}{k - k' - d_A} \binom{\binom{a}{2}}{k'} \\
&\leq \sum_{d_A = k/8}^{s''y} \binom{s''y}{d_A} \binom{sa}{k - k' - d_A} \binom{\binom{a}{2}}{k'} \\
&\leq k \binom{sy}{k/8} \binom{sa}{7k/8 - k'} \binom{\binom{a}{2}}{k'}. \qquad (5.39)
\end{aligned}$$

In the second line we used ($s''-1$ times) the fact that $\sum_{j=0}^{d}\binom{t}{j}\binom{v}{d-j}=\binom{t+v}{d}$. In the third line we used $s''\le s$ and that the largest summand is the one with $d_A=k/8$. Now let $\#\{k_B,r_2\}$ denote the total number of k_{B,r_2}-graphs. By (5.13) there are at most $\sum_{s''=1}^{s}\binom{a}{s''}\le s\binom{a}{s}$ choices for S''. Thus the above implies that the total number of triangle-free friendly (k_{A,r_A},k_{B,r_2})-graphs where the AB-graph is S-good for some S is at most

$$s\binom{a}{s}\,U\,\#\{k_B,r_2\}\binom{ab}{m_1}.$$

Similarly as in the "unfriendly" case, we compare this with a lower bound on the total number of friendly (k_{A,r_A},k_{B,r_2})-graphs where the AB-graph is S-good. To this end, note that all k_{A,r_A}-graphs with maximum degree at most m/n are friendly. Indeed, this follows since (as already noted in the proof of Corollary 5.11) almost all AB-graphs with m_1 edges have minimum degree at least m/n. This also implies that any k_{B,r_2}-graph with maximum degree at most m/n forms a friendly pair together with any k_{A,r_A}-graph with maximum degree at most m/n. Furthermore, note that by Corollary 5.22, certainly almost all AB-graphs with m_1 edges dominating a friendly pair are S-good. Together with (5.19), the proof of Corollary 5.22 then shows that the number of friendly (k_{A,r_A},k_{B,r_2})-graphs where the AB-graph is S-good is at least

$$(1+o(1))|\mathcal{H}_1'||\mathcal{H}_2'|\#\{k_B,r_2\}f(k_B,r_2)^{-2\varepsilon^2}\binom{ab}{m_1}.$$

Thus the probability that a random friendly (k_{A,r_A},k_{B,r_2})-graph where the AB-graph is S-good is triangle-free is at most

$$\begin{aligned}
&(1+o(1))s\binom{a}{s}U\Big/\binom{s(a-s)}{k-(1-\varepsilon^2)k'}\binom{\binom{a-s}{2}}{(1-\varepsilon^2)k'}f(k_B,r_2)^{-2\varepsilon^2}\\
\overset{(5.13)}{\le}\;& a^s\,U\,(sa/2)^{-k+(1-\varepsilon^2)k'}\,\left(a^2/4\right)^{-(1-\varepsilon^2)k'}\,k!\,f(k_B,r_2)^{2\varepsilon^2}\\
\le\;& a^s\,k\,(sy)^{k/8}(sa)^{7k/8-k'}a^{2k'}\left(\frac{sa}{2}\right)^{(1-\varepsilon^2)k'-k}\left(\frac{a^2}{4}\right)^{-(1-\varepsilon^2)k'}\binom{k}{k/8;\,k'}f(k_B,r_2)^{2\varepsilon^2}\\
\le\;& a^s\,k\,2^{k'}(a/(2s))^{\varepsilon^2k'}\,2^k\,(y/a)^{k/8}\,3^k\,f(k_B,r_2)^{2\varepsilon^2}. \qquad (5.40)
\end{aligned}$$

Now use that $s\log a\le k$ (see (5.13)), that by (5.10), $a/s\le m/k$, and furthermore that $k'\le k$, to see that (5.40) is at most

$$\begin{aligned}
&k(m/k)^{\varepsilon^2k'}(12\mathrm{e})^k(y/a)^{k/8}f(k_B,r_2)^{2\varepsilon^2}\\
\overset{(5.9)}{\le}\;&\exp\left\{k\left(5\log\log(m/k)+\log(12\mathrm{e})+\frac{1}{8}\log(y/a)\right)\right\}f(k_B,r_2)^{2\varepsilon^2}\\
\overset{(5.31)}{\le}\;&\exp\left\{-\frac{49}{\varepsilon^4}k\log\log(m/k)\right\}f(k_B,r_2)^{2\varepsilon^2}\\
\overset{(5.32)}{\le}\;&f(k,r_A)^{-6/\varepsilon^2}f(k_B,r_2)^{2\varepsilon^2}\\
\le\;&f(k,r_A)^{-5/\varepsilon^2},
\end{aligned}$$

where in the last line we used the assumption made in the statement of the lemma.

□

5.8 Odd cycles

5.8.1 The sparse case: $m/n^2 \to 0$

The proof of the general case of Theorem 5.1 is quite similar to the one for triangles. However, some of the details are more complicated, which is the reason for presenting a sketch of the necessary modifications to the proof of the triangle case separately. We assume without loss of generality that $\varepsilon \leq 10^{-6}/\ell^2$.

First consider the proof of the 0-statement. For $n/2 \leq m \leq c_1 n$, where c_1 is fixed, the probability that a random graph $G_{n,m}$ is C_ℓ-free is bounded away from zero, so for these m, we may apply the same reasoning as for triangles. If $c_1 n \leq m \leq 20n \log n$, the result follows immediately by noting that the number of bipartite graphs on n vertices and m edges is at most $2^n\binom{\lfloor n^2/4\rfloor}{m}$, whereas the following special case of a more general result of Prömel and Steger (see [69]) shows that the number of C_ℓ-free graphs on n vertices and m edges is much larger. We omit the calculations, since they are very similar to the ones for the triangle case in [70]. Let X_ℓ denote the number of ℓ-cycles in $G_{n,m}$. (Note that $\mathbb{E}[X_\ell] = \Theta(m^\ell/n^\ell)$).

Theorem 5.23 (Prömel and Steger [69]) *If $m = o(n^{1+1/(\ell-1)})$, then there exists a positive constant c so that*

$$\mathbb{P}[G_{n,m} \textit{ is } C_\ell\textit{-free}] \geq \mathrm{e}^{-c\mathbb{E}[X_\ell]}.$$

So assume that $20n \log n \leq m \leq (1-\varepsilon)t_\ell$. As in the case of triangles, we fix an almost equitable bipartition into classes A and B, fix an edge e in A and consider a random AB graph with edge probability $p = (1+\varepsilon^2)4m/n^2$. Note that $a = (1+o(1))b = (1+o(1))n/2$. We need a lower bound on the probability that the union of such an AB-graph and e contains no ℓ-cycle. But the existence of such cycles is positively correlated, and thus the FKG-inequality (see e.g. Theorem 2.12 in [41]) implies that this probability is at least

$$\left(1 - p^{\ell-1}\right)^{(1+o(1))(n/2)^{\ell-2}} \overset{(5.4)}{\geq} \mathrm{e}^{-(1+\varepsilon^2)p^{\ell-1}(n/2)^{\ell-2}} := \nu_\ell.$$

By the definition of t_ℓ, we have $m\nu_\ell \geq (1-\varepsilon)t_\ell \mathrm{e}^{-(1-\varepsilon)(t_\ell)^{\ell-1}(2/n)^\ell} \to \infty$, and the rest of the proof goes through as before.

To prove the second 1-statement, we will make use of Theorem 5.24 below. Łuczak [59] proved that it follows from a well-known probabilistic conjecture (see also [41]). Roughly speaking, this conjecture asserts that for a fixed graph H, only a superexponentially small proportion of a certain class of graphs $\mathcal{F}$ is not H-free. For odd cycles, this conjecture was proven by Kreuter [53] for a smaller class of graphs (which was sufficient for his purposes). However, by adapting the notion of an $(\ell-2)$-expanding set which is used later on in this section, one can show that all but a superexponentially small proportion of the graphs in $\mathcal{F}$ have a subgraph which is almost as large as the original graph and which satisfies the Kreuter's additional conditions, and which implies Theorem 5.24.

Theorem 5.24 *Given $\delta > 0$, there exists a constant $C > 0$ so that almost all C_ℓ-free graphs with n vertices and $Cn^{\ell/(\ell-1)} \leq m \leq n^2/C$ edges can be made bipartite by deleting at most δm edges.*

The proof of the second 1-statement is then the same as for triangles, except that of course we have to prove Lemmas 5.13, 5.14 and 5.15 for ℓ-cycles. Lemmas 5.13 and 5.14 are derived from Lemmas 5.16 and 5.17 as before. We now show how to modify the proof of Lemma 5.16 so that it applies to ℓ-cycles instead of triangles. $\mathcal{S}$ is now the set of those paths on $\ell-1$ edges (where the edges all have one endpoint in A and one in B) which form an ℓ-cycle together with some edge in T_A or T_B. Thus X now counts the number of ℓ-cycles in $G_{G_A,G_B,p}$ with one edge contained in some torso and the others in the AB-graph. Then, similarly to (5.24), we have

$$\mu \geq (k_{A,r_1}+k_{B,r_2})((1-\varepsilon^2)n/2)^{\ell-2}p^{\ell-1}.$$

One difference is that Δ now counts the expected number of pairs of ℓ-cycles with at least an edge in common, where the edges in the torso are not necessarily adjacent. Thus

$$\Delta = \sum_{e,e'}\sum_{i=1}^{\ell-2}\mathbb{E}[X_{e,e',i}].$$

Here $\sum_{e,e'}$ denotes the sum over all pairs of edges $\{e,e'\}$ in $T_A\cup T_B$ and $X_{e,e',i}$ denotes the number of ordered pairs of paths α, $\alpha'\in\mathcal{S}$ in $G_{G_A,G_B,p}$ so that $\alpha\cup e$ and $\alpha'\cup e'$ are ℓ-cycles and so that α and α' have exactly i common edges. Now note that if α and α' have i common edges, they must have at least $i+1$ common vertices. Thus one can show that the fact that $pn\to\infty$ implies that the contribution from those terms with $i>1$ is $o(1)$ of those with $i=1$. Distinguishing between the cases where e and e' have an endvertex in common or not as in the derivation of (5.26), one may now check that (assuming without loss of generality that $k_A\geq k_B$ as before and also with the same summation as before)

$$\begin{aligned}\Delta &\leq (1+\varepsilon)p^{2(\ell-1)-1}\Bigg(\sum_{xy,yz}(n/2)^{2(\ell-2)-1}+\sum_{xy,x'y'}8\ell(n/2)^{2(\ell-2)-2}\Bigg)\\ &\leq 25\ell\mu D_A n^{\ell-3}p^{\ell-2}.\end{aligned}$$

Using this, it is easily checked that $\frac{\mu^2}{2\Delta}\geq\frac{4}{\varepsilon^2}\log f(k_A,r_1)$ as before. In the final case analysis, we now distinguish whether $k_A/m\leq n^{-\varepsilon^2\ell/(8(\ell-1))}$ and whether $(m/t_\ell)^{\ell-1}\leq 9$. The analogue of (5.28) is then that $\Delta/\mu\leq 1000\ell\varepsilon^2$. One can compensate this by showing that (using $\varepsilon\leq 10^{-6}/\ell^2$) in (5.29) the term $2000\varepsilon^2$ can be replaced by $2000\ell\varepsilon^2$.

The proof of Lemma 5.15 is also similar, except that instead of Lemma 5.18 we need a lemma which is concerned with $(\ell-2)$-*expanding sets* instead of expanding sets. As in Lemma 5.18, in what follows, we are given a vertex bipartition into A and B and a subset $A'\subset A$ with $|A'|\geq a-s$. We then consider a random (bipartite) $A'B$-graph with edge probability $p_0=m_0/(ab)$, where $m_0\geq m/4$. We say that a set $V\subset B$ is $(\ell-2)$-*expanding* if its $(\ell-2)$th neighbourhood

$$\Gamma_{A'}(\Gamma_B\Gamma_{A'})^{(\ell-2)/2}(V)$$

contains at least $a-y$ vertices. Thus we are done if we can prove

Lemma 5.25 *Consider the setting above. Given $h \leq k$, for $1 \leq i \leq h$ let v_i satisfy $v_i \geq D'_A$. Furthermore, suppose that $\sum_{i=1}^{h} v_i \geq k/4$. Let $V_1, \dots, V_h$ with $|V_i| = v_i$ be chosen uniformly and independently at random in B. Then*

$$\mathbb{P}[\textit{the } (\ell-2)\textit{-expanding } V_i \textit{ have total weight at most } k/8] \leq f(k, r_A)^{-18/\varepsilon^2}.$$

To prove this, we need an auxiliary lemma. For $1 \leq d \leq \ell - 1$, let

$$w_d = \left(\frac{\varepsilon^6 t_\ell}{80 n \log\log(m/k)} \right)^d .$$

Note that $w_1 = D'_a/$ and that $w_d \leq n^{1-1/(\ell-1)+o(1)}$ if $d \leq \ell - 2$. Let

$$u' = \min\left\{ k/32, y'/2 \right\}.$$

Lemma 5.26 *Let d with $2 \leq d \leq \ell - 2$ be an integer. Let $W_{1,d-1}, \dots W_{h',d-1}$ be disjoint sets, all in the same vertex class, satisfying $|W_{i,d-1}| \geq w_{d-1}$ and $\sum_{i=1}^{h'} W_{i,d-1} \geq u'$. Consider an $A'B$-graph with edge probability $p_\ell \geq p_0/\ell$. Then with probability at least $1 - f(k, r_A)^{-19/\varepsilon^2}$, there exist (after relabelling) disjoint $W_{1,d}, \dots, W_{h'',d}$, satisfying $|W_{i,d}| \geq w_d$, $\sum_{i=1}^{h''} W_{i,d} \geq u'$, and $W_{i,d} \subseteq \Gamma(W_{i,d-1})$.*

The idea of the proof is similar to that of Lemma 5.19 and Corollary 5.20. The details are more complicated, but on the other hand, there is more room to spare in the calculations.

Proof of Lemma 5.26. Since $a' = |A'|$ and $b = |B|$ are approximately equal, without loss of generality we may assume that the $W_{i,d-1}$ are contained in B. Let $W'_{i,d} = \Gamma(W_{j,d-1})$. Inductively, we now define a notion called *d-usefulness*. For all i, let $\mathcal{U}_{i,d}$ be the set of those $W'_{j,d}$ with $j < i$ which are d-useful and let

$$u_i = |\bigcup_{j\colon W_{j,d} \in \mathcal{U}_{i,d}} W_{j,d}|. \tag{5.41}$$

Then we say that $W'_{i,d}$ is *d-useful* if at least one of the following two possibilities hold.

- $|u_i| \geq a'/2$.
- $|W'_{i,d} \backslash \bigcup_{j\colon W_{j,d} \in \mathcal{U}_{i,d}} W_{j,d}| \geq \mathbb{E}[|W'_{i,d}|]/2^5$.

Now let $W' = \bigcup_{j\colon W_{j,d} \in \mathcal{U}_{i,d}} W_{j,d}$. Then the probability that $W'_{i,d}$ is d-useful is equal to one if $|W'| \geq a'/2$. So suppose that $|W'| < a'/2$. Suppose also that $p_\ell |W_{i,d-1}| \leq 4$. Then for a vertex $x \in A' \backslash W'$, we have

$$\begin{aligned} p' := \mathbb{P}[x \notin W'_{i,d}] &= (1 - p_\ell)^{|W_{i,d-1}|} \leq (1 - p_\ell)^{|W_{i,d-1}|/4} \\ &\leq 1 - \frac{p_\ell |W_{i,d-1}|}{4} \left(1 - \frac{p_\ell |W_{i,d-1}|}{8} \right) \leq 1 - \frac{p_\ell |W_{i,d-1}|}{8}. \end{aligned}$$

Here we used that for $\tau j \leq 1$, we have $(1-\tau)^j \leq 1-\tau j+(\tau j)^2/2$. Furthermore, for any x, $y \in A'$ with $x \neq y$, the events that $x \in W'_{i,d}$ and $y \in W'_{i,d}$ are independent. Thus $|W'_{i,d}\backslash W'|$ is binomially distributed with mean

$$\mathbb{E}[|W'_{i,d}\backslash W'|] = |A'\backslash W'|(1-p') \geq a' p_\ell |W_{i,d-1}|/2^4. \tag{5.42}$$

On the other hand, $\mathbb{E}[|W'_{i,d}|] = a'(1-p') \leq a' p_\ell |W_{i,d-1}|$. Indeed, the second inequality is trivial if $1 \leq p_\ell |W_{i,d-1}| \leq 4$. If $p_\ell |W_{i,d-1}| \leq 1$, the second inequality follows from the fact that for $\tau j \leq 1$, we have $(1-\tau)^j \geq 1-\tau j$.

Thus by applying the Chernoff bound (3.3) with $\delta = 1/2$, we have

$$\mathbb{P}[W'_{i,d} \text{ is not useful} \mid |W'| < a'/2] \leq \exp\left\{-a' p_\ell |W_{i,d-1}|/2^7\right\}.$$

Now consider the case $p_\ell |W_{i,d-1}| \geq 4$. If $W'_{i,d}$ is not useful, the trivial inequality $\mathbb{E}[|W'_{i,d}|] \leq a'$ implies that there must be a set V of at least $a'/4$ vertices in A' which are not adjacent to any vertices in $W_{i,d-1}$. Since there are at most $2^{a'}$ choices for V, in this case we have

$$\begin{aligned}\mathbb{P}[W'_{i,d} \text{ is not useful} \mid |W'| < a'/2] &\leq 2^{a'}(1-p_\ell)^{a'|W_{i,d-1}|/4} \\ &\leq (\mathrm{e}/2)^{-p_\ell a' |W_{i,d-1}|/4} \\ &\leq \mathrm{e}^{-p_\ell a' |W_{i,d-1}|/8}.\end{aligned}$$

Thus for any fixed subsequence $\mathcal{W}_{d-1}$ of the $W_{i,d-1}$ which has total weight at least $u'/2$, the probability that none of the corresponding neighbouring sets $W'_{i,d}$ are d-useful is at most

$$\begin{aligned}\exp\left\{-\sum_{i\colon W_{i,d-1}\in\mathcal{W}_{d-1}} a' p_\ell |W_{i,d-1}|/2^7\right\} &\leq \exp\left\{-a' p_0 u'/(2^8\ell)\right\} \\ &\leq 2^{-k} f(k, r_A)^{-19/\varepsilon^2},\end{aligned} \tag{5.43}$$

where we used that $\varepsilon \leq 10^{-6}/\ell^2$ to get rid of the $1/\ell$ factor in the exponent. The details of the intermediate calculations are similar to those of the proof of Lemma 5.18.

Now note that if those $W_{i,d-1}$ whose neighbouring sets $W'_{i,d}$ are useful have total weight at most $u'/2$, then there must exist a subsequence $\mathcal{W}_{d-1}$ of the $W_{i,d-1}$ which has total weight at least $u'/2$ and where none of $W'_{i,d}$ which correspond to some $W_{i,d-1} \in \mathcal{W}_{d-1}$ is useful. Multiplying by $2^{h'} \leq 2^k$, (5.43) implies that with probability at least $1 - f(k, r_A)^{-19/\varepsilon^2}$, no such subsequence $\mathcal{W}_{d-1}$ exists, and so the $W_{i,d-1}$ whose neighbouring sets $W'_{i,d}$ are useful have total weight at least $u'/2$.

Let $\mathcal{W}'_d$ be the set of these useful $W'_{i,d}$. Without loss of generality, assume that the sets are numbered so that $\mathcal{W}'_d = \{W'_{1,d}, \ldots, W'_{h''',d}\}$ for some $h''' \leq h'$. Let $\mathcal{W}''_d = \{W'_{1,d}, \ldots, W'_{h^*,d}\}$, where $h^* = h'''$ if $u_{h'''} \leq a'/2$ (where u_i is defined as in (5.41)), and otherwise h^* is the smallest number satisfying

$$\left|\bigcup_{i=1}^{h^*} W'_{i,d}\right| \geq a'/2.$$

Suppose that $\mathcal{W}''_d$ contains a set $W'_{j,d}$ so that $p_\ell|W_{j,d-1}| \geq 4$ for this j. Then

$$\mathbb{E}[|W'_{j,d}|] = a'(1-p_\ell)^{|W_{j,d-1}|} \geq a'(1-\mathrm{e}^{-p_\ell|W_{j,d-1}|}) \geq a'(1-\mathrm{e}^{-4}) \geq 2^5 y'.$$

So setting $h''=1$ and $W_{1,d}=W'_{j,d}$ satisfies the requirements of the lemma. So we may assume that $p_\ell|W_{i,d-1}| \leq 4$ for all $i \leq h^*$. We claim that in this case, the sets in $\mathcal{W}''_d$ have total weight at least $2^5u'$ and satisfy $|W'_{i,d}| \geq 2^5 w_d$ for all $i \leq h^*$. Indeed, inequality (5.42) implies that

$$\mathbb{E}[|W'_{i,d}|] \geq \mathbb{E}[|W'_{i,d} \backslash W'|] \geq a'p_\ell w_{d-1}/2^4 \geq 2^6 w_d.$$

Similarly, $\mathbb{E}[|W'_{i,d}|] \geq 2^7|W_{i,d-1}|$, and so the claim follows from the fact that those $W_{i,d-1}$ whose neighbouring sets $W'_{i,d}$ are useful have total weight at least $u'/2$ and from the definition of d-usefulness. Now we choose disjoint $W_{i,d}$ from the $W'_{i,d}$ in $\mathcal{W}''_d$ satisfying the assertion of the lemma as in the proof of Corollary 5.20. □

Proof of Lemma 5.25. This is similar to the proof of Lemma 5.18. Again, we first calculate the probability that a fixed subsequence $\mathcal{W}=(W_1,\dots,W_{h'})$ of total weight at least $k/8$ of the sequence of sets $V_1,\dots,V_h$ contains no $(\ell-2)$-expanding sets. Since for all i, $|W_i|/2 \geq D'_A/2 \geq w_1$, an application of Corollary 5.20 tells us that with probability at least $1-f(k,r_A)^{-19/\varepsilon^2}$, we can find a sequence of sets satisfying the conditions of Lemma 5.26 with $d=2$. Now define p_ℓ by $(1-p_\ell)^\ell = 1-p_0$ and note that an $A'B$-graph with edge probability p_0 can be considered as the union of ℓ independent $A'B$-graphs with edge probability p_ℓ. Note also that $p_\ell \geq p_0/\ell$. Now apply Lemma 5.26 with $d=2$ to the sets obtained from Corollary 5.20 and the first of the ℓ random $A'B$-graphs with edge probability p_ℓ. Then with probability at least $1-f(k,r_A)^{-19/\varepsilon^2}$, we obtain sets $W_{i,2}$ which satisfy the conditions of Lemma 5.26 for $d=3$. Now continue this process, until after $\ell-3$ applications of Lemma 5.26, with probability at least $1-(\ell-2)f(k,r_A)^{-19/\varepsilon^2}$, we have a sequence of disjoint sets $W_{1,\ell-2},\dots,W_{h'',\ell-2}$ as in the assertion of Lemma 5.26 for $d=\ell-2$. The remainder of the proof is now the same as that of Lemma 5.18, except that we consider a random $A'B$-graph with edge probability p_ℓ again. For all i with $1 \leq i \leq h''$ we have (using $W_{i,\ell-2} \geq w_{\ell-2}$)

$$\begin{aligned}\mathbb{P}[\,|\Gamma_{A'}(W_{i,\ell-2})| \leq a-y\,] &\leq \binom{a'}{a-y}(1-p_\ell)^{(a'-(a-y))|W_{i,\ell-2}|}\\ &\leq \exp\{-yp_\ell|W_{i,\ell-2}|/4\}.\end{aligned}$$

The disjointness of the $W_{i,\ell-2}$ then implies that

$$\mathbb{P}\Big[\bigcap_{1\leq i\leq h''}\{|\Gamma_{A'}(W_{i,\ell-2})| \leq a-y\}\Big] \leq \mathrm{e}^{-yp_\ell u'/4} \leq f(k,r_A)^{-19/\varepsilon^2}.$$

Putting the above together, the probability that in the union of the above graphs (and thus in the random $A'B$-graph with edge probability p_0) none of the W_i is $(\ell-2)$-expanding is at most $\ell f(k,r_A)^{-19/\varepsilon^2}$. Now note that if the $(\ell-2)$-expanding V_i have total weight at most $k/8$, then there must exist a subsequence $\mathcal{W}=(W_1,\dots,W_{h'})$ of the V_i, where the W_i have total weight at least $k/8$ and where each W_i is not $(\ell-2)$-expanding. The result now follows if we sum over all possibilities of choosing $\mathcal{W}$ and thus multiplying the above probability by $2^h \le 2^k$. $\square$

5.8.2 The dense case: $m/n^2 \not\to 0$

In this section, we show how one can deduce the result of Theorem 5.1 for the case when $m/n^2 \not\to 0$ from the results and techniques in Lamken and Rothschild [56] and Prömel and Steger [70]. The proof will proceed by induction on ℓ and n. For this section only, all logarithms are base 2.

Let $\mathcal{F}_{n,m}(C_\ell)$ denote the set of C_ℓ-free graph with n vertices and m edges and let $F_{n,m}(\ell) = |\mathcal{F}_{n,m}(C_\ell)|$. Recall that $\mathrm{Bip}_{n,m}$ denotes the number of bipartite graphs with n vertices and m edges. Note that for sufficiently large n and if $e \le m/4$,

$$\mathrm{Bip}_{n-1,m-e} \le 2^{n-1}\binom{(n-1)^2/4}{m-e} \overset{(3.1)}{\le} \mathrm{e}^{-m/n}\, 2^n \binom{n^2/4 - 4n^{3/4}}{m-e}. \tag{5.44}$$

Let $\mathcal{A}(n,m)$ denote the set of graphs in $\mathcal{F}_{n,m}(C_\ell)$ which contain a vertex of degree less than $4n^{3/4}$.

Let $\mathcal{B}(n,m)$ denote the set of graphs in $\mathcal{F}_{n,m}(C_\ell)$ which contain a set Q of size $\sqrt{n}$ such that $|\Gamma(Q)| \le n/2 - n^{3/4}$.

Let $\mathcal{E}(n,m)$ denote the set of graphs in $\mathcal{F}_{n,m}(C_\ell)$ which are not contained in $\mathcal{A}(n,m)$ or $\mathcal{B}(n,m)$ and which contain a cycle of length $\ell - 2$.

Lemma 5.27

$$|\mathcal{A}(n,m)| \le 2^{5n^{3/4}\log n} \sum_{e=0}^{4n^{3/4}-1} F_{n-1,m-e}(\ell).$$

Proof. Construct all graphs in $\mathcal{A}(n,m)$ as follows. First choose a vertex v (for which there are $n = 2^{\log n}$ possibilities). For all $0 \le e < 4n^{3/4}$ then choose a C_ℓ-free graph with $m-e$ edges on the remaining vertices (for which there are $F_{n-1,m-e}(\ell)$ choices). Finally, choose the e neighbours of v (for which there are $\binom{n-1}{e} \le 2^{4n^{3/4}\log n}$ choices). $\square$

Lemma 5.28 *For $n \ge 10\ell$,*

$$|\mathcal{E}(n,m)| \le n(n-1)^2\, 3 \sum_{e=4n^{3/4}}^{n} \binom{n/2 - n^{3/4}}{e} F_{n-1,m-e}(\ell). \tag{5.45}$$

Proof. All graphs in $\mathcal{E}(n,m)$ are obtained as follows: first choose a vertex v (n possibilities). Then for all $e \ge 4n^{3/4}$, chose a C_ℓ-free graph with $n-1$ vertices and $m-e$ edges (for which there are $F_{n-1,m-e}(\ell)$ choices). Then connect v to two vertices a and b which are joined by a path P on $\ell - 4$ vertices in order to form an $(\ell-2)$-cycle ($(n-1)^2$ choices), and then choose the remaining neighbours of v. If v has a neighbour in $\Gamma(a)$, one may verify (see the proof of Lemma 3 in [56]) that the set of possible neighbours of v (i.e. those which can be connected to v without producing an ℓ-cycle) is at most

$$n - (|\Gamma(a)| - \ell) - (|\Gamma(\Gamma(b))| - \ell) \leq n - (4n^{3/4} - \ell) - (n/2 - n^{3/4} - \ell) \leq n/2 - n^{3/4}.$$

One obtains the same bound if v has a neighbour in $\Gamma(b)$, and a smaller one if v is not adjacent to $\Gamma(a) \cup \Gamma(b)$. Thus for each e, there are $3(n-1)^2\binom{n/2-n^{3/4}}{e}$ choices for the neighbourhood of v, which implies the result. □

Let

$$C_{n,m} = 2^{\log n + \sqrt{n}\log n} \sum_{e=0}^{n} \binom{n/2 - n^{3/4}}{e} F_{n-1,m-e}(\ell) \tag{5.46}$$

and note that Lemma 5.28 implies that $|\mathcal{E}(n,m)| \leq C_{n,m}$.

The following theorem now immediately implies the $m/n^2 \not\to 0$ case of Theorem 5.1.

Theorem 5.29 *For any odd ℓ and any $\varepsilon > 0$, there exist constants $a > 0$, $a' > 0$ and $c > 0$, so that for all $n \in \mathbb{N}$ and all $0 \leq m \leq \binom{n}{2}$, we have*

$$F_{n,m}(\ell) \leq (1 + (\ell - 2))(\varepsilon + g_{n,m} + h_{n,m}))\,\mathrm{Bip}_{n,m}, \tag{5.47}$$

where

$$g_{n,m} = a2^{-\sqrt{n}} + 2^{\left(2 - \frac{m}{14cn^{7/4}\log n}\right)cn^{3/2}\log n}$$

and

$$h_{n,m} = a'2^{-\sqrt{n}} + 2^{n^{11/6} - m}.$$

Proof. Without loss of generality, assume that $\varepsilon \leq 1/10$. The proof is by induction on ℓ (actually on $k = (\ell-1)/2$). Choose the constants a and c to be the same as those in Theorem 2.13 in [70]. For $\ell = 3$, the assertion of Theorem 5.29 is then weaker than Theorem 2.13 in [70], so we may assume that ℓ is an odd integer greater than three and that (5.47) holds for all odd cycles whose length is smaller than ℓ. Obviously, for all n and $0 \leq m \leq \binom{n}{2}$ we have

$$\begin{aligned} F_{n,m}(\ell) &\leq |\mathcal{A}(n,m)| + |\mathcal{B}(n,m)| + |\mathcal{E}(n,m)| + F_{n,m}(\ell-2). \\ &\overset{(5.46)}{\leq} |\mathcal{A}(n,m)| + |\mathcal{B}(n,m)| + C_{n,m} + F_{n,m}(\ell-2). \end{aligned}$$

In their proof of Theorem 2.13, Prömel and Steger showed that

$$|\mathcal{B}(n,m)| + C_{n,m} \leq g_{n,m}\mathrm{Bip}_{n,m}. \tag{5.48}$$

(In their case, $C_{n,m}$ was actually a bound on the size of a different class of graphs – see Lemma 2.6 in [70] – but they did not use this fact in their bounds on $C_{n,m}$.) Thus, combining (5.48) with our induction hypothesis, we know that for all n and all $0 \leq m \leq \binom{n}{2}$,

$$F_{n,m}(\ell) \leq |\mathcal{A}(n,m)| + (1 + (\ell-3)(\varepsilon + g_{n,m} + h_{n,m})\mathrm{Bip}_{n,m}. \tag{5.49}$$

Thus it suffices to prove that for all n and all $0 \le m \le \binom{n}{2}$, we have

$$|\mathcal{A}(n,m)| \le (\varepsilon + h_{n,m})\mathrm{Bip}_{n,m}. \tag{5.50}$$

We prove (5.50) by induction on n. Choose n_0 sufficiently large so that for all $n \ge n_0$, we have

(i) $n/\log n \ge 32n^{4/5}$ and $n \ge 10\ell$;

(ii) $g_{n,m} \le 1$ if $m \ge n^2/\log n$;

(iii) (5.44) holds;

(iv) the proportion of bipartite graphs with n vertices and m edges which have a vertex of degree less than $4n^{3/4}$ is at most $\varepsilon/2$ if $m \ge n^{9/5}$;

(v) $F_{n,m}(\ell) \le (1+\varepsilon/2)\mathrm{Bip}_{n,m}$ for all m with $n^{9/5} \le m \le n^2/\log n$.

The proof that the condition (iv) can be satisfied is elementary, and therefore omitted. The fact that the condition (v) can be satisfied follows from the result on the $m/n^2 \to 0$ case proven in the previous chapter. Choose a' (in the definition of $h_{n,m}$) sufficiently large so that (5.50) is true for all $n \le n_0$. Now assume that we are given $n \in \mathbb{N}$ and that (5.50) holds for $n-1$ (and for all possible values of m). We will show that it then also holds for n and all $0 \le m \le \binom{n}{2}$.

Suppose first that $m \le n^{9/5}$. Then

$$|\mathcal{A}(n,m)| \le \binom{\binom{n}{2}}{m} \overset{(3.1)}{\le} \mathrm{e}^{\frac{mn^2/4}{n^2/4-m}} \binom{n^2/4}{m} \le \mathrm{e}^{2m}\binom{n^2/4}{m} \le \mathrm{e}^{2m}\mathrm{Bip}_{n,m} \le h_{n,m}\mathrm{Bip}_{n,m}.$$

So we may assume that $m \ge n^{9/5}$. If we also have $m \le n^2/\log n$, then (iv), (v) and the fact that $F_{n,m}(\ell) \ge \mathrm{Bip}_{n,m}$ imply that $|\mathcal{A}(n,m)| \le \varepsilon\mathrm{Bip}_{n,m}$.

So suppose that $m \ge n^2/\log n$. Recall that in this case $g_{n,m} \le 1$. Using this, substituting the induction hypothesis (5.50) for $n-1$ into (5.49) tells us that (5.47) holds for $n-1$. Thus, if $e \le 4n^{3/4}$,

$$\begin{aligned}
F_{n-1,m-e}(\ell) &\le (2+\varepsilon+\ell h_{n-1,m-e})\mathrm{Bip}_{n-1,m-e} \\
&\le 4\ell\left(a' + 2^{n^{11/6}-(m-e)}\right)\mathrm{Bip}_{n-1,m-e} \\
&\overset{(5.44)}{\le} 4\ell\left(a' + 2^{n^{11/6}-(m-4n^{3/4})}\right)\mathrm{e}^{-m/n}\,2^n\binom{n^2/4-4n^{3/4}}{m-e} \\
&\le \mathrm{e}^{-m/(2n)}\,h_{n,m}\,2^n\binom{n^2/4-4n^{3/4}}{m-e}.
\end{aligned}$$

To estimate $|\mathcal{A}(n,m)|$, we shall make use of the following elementary inequality. For $0 \le v \le k \le u$, we have

$$\sum_{e=0}^{v}\binom{u}{k-e} \le \sum_{e=0}^{v}\binom{u}{k-e}\binom{v}{e} = \binom{u+v}{k}. \tag{5.51}$$

Then by Lemma 5.27, we have

$$\begin{aligned}
|\mathcal{A}(n,m)| &\leq 2^{5n^{3/4}\log n} \sum_{e=0}^{4n^{3/4}-1} F_{n-1,m-e}(\ell) \\
&\leq 2^{5n^{3/4}\log n} \mathrm{e}^{-m/(2n)}\, h_{n,m}\, 2^n \sum_{e=0}^{4n^{3/4}-1} \binom{n^2/4-4n^{3/4}}{m-e} \\
&\overset{(5.51)}{\leq} 2^{5n^{3/4}\log n} \mathrm{e}^{-m/(2n)}\, h_{n,m}\, 2^n \binom{n^2/4}{m} \\
&\leq \mathrm{e}^{-m/(4n)}\, h_{n,m}\, 2^n \binom{n^2/4}{m}.
\end{aligned}$$

Inequality (5.50) now follows from Theorem 2.7. □

5.9 Proof of Theorem 5.2

Throughout this section, we set $p = m/\binom{n}{2}$. Using the fact that $\chi(G) \geq n/\alpha(G)$ for any graph G on n vertices, Theorem 5.2 will follow immediately from the following lemma (for a graph G, $\alpha(G)$ denotes the number of vertices in the largest set of independent vertices in G).

Lemma 5.30 *If* $p = \mathcal{O}(n^{-3/4})$, *then*

$$\mathbb{P}[\,\alpha(G_{n,m}) \geq \frac{4}{p}\log(np) \mid G_{n,m} \textit{ is } K_3\textit{-free}\,] = o(1)). \tag{5.52}$$

If $n^{-3/4} \leq p = o(n^{-1/2})$, *then*

$$\mathbb{P}[\,\alpha(G_{n,m}) \geq 4pn^{3/2}\log n \mid G_{n,m} \textit{ is } K_3\textit{-free}\,] = o(1). \tag{5.53}$$

To prove Lemma 5.30, we shall need Pittel's inequality (see page 35 in [8]), which states that if $\mathcal{Q}$ is any property and $0 < p = m/\binom{n}{2} < 1$, then

$$\mathbb{P}[G_{n,m} \text{ has } \mathcal{Q}] \leq 3\sqrt{m}\,\mathbb{P}[G_{n,p} \text{ has } \mathcal{Q}]. \tag{5.54}$$

The proof of Lemma 5.30 now uses the fact that, for suitable choices of parameters, the probability that there is a "large" independent set in $G_{n,m}$ is much smaller than the probability that $G_{n,m}$ is triangle-free.

Proof of Lemma 5.30. First note that for any $r = r(n)$ with $r \to \infty$,

$$\mathbb{P}[\alpha(G_{n,p}) \geq r] \leq \binom{n}{r}(1-p)^{\binom{r}{2}} \leq (\mathrm{e}n/r)^r \mathrm{e}^{-pr(r-1)/2} = \mathrm{e}^{-(1+o(1))\phi},$$

where for convenience we write

$$\phi = r(pr/2 - \log(n/r)).$$

Then by Theorem 5.23 and Pittel's inequality (5.54), there is a constant $c > 0$ so that

$$\begin{aligned}\mathbb{P}[\,\alpha(G_{n,m}) \geq r \mid G_{n,m} \text{ is } K_3\text{-free}\,] &\leq \frac{\mathbb{P}[\,\alpha(G_{n,m}) \geq r\,]}{\mathbb{P}[\,G_{n,m} \text{ is } K_3\text{-free}\,]} \\ &\leq 3\sqrt{m}\,\mathbb{P}[\,\alpha(G_{n,p}) \geq r\,]\,\mathrm{e}^{c\mathbb{E}[X_3]} \\ &\leq 3\sqrt{m}\,\mathrm{e}^{-(1+o(1))(\phi - c\mathbb{E}[X_3])}.\end{aligned}$$

(Recall that X_3 denotes the number of triangles in $G_{n,m}$.) Thus to prove (5.52), it suffices to prove that if $r = \frac{4}{p}\log(np)$ and $p = \mathcal{O}(n^{-3/4})$, then $\phi \geq 4\log n$ and $\phi/\mathbb{E}[X_3] \geq 2c$. Note that our choice of r implies that $\log(1/p) = (1+o(1))\log r$. This in turn implies that $pr/4 = (1+o(1))\log(n/r)$ and thus that

$$\phi = (1+o(1))pr^2/4 = \frac{4}{p}\,(\log(np))^2 \geq 4\log n.$$

Also

$$\frac{\phi}{\mathbb{E}[X_3]} = (1+o(1))\frac{p}{4}\left(\frac{4}{p}\log(np)\right)^2\frac{6}{n^3p^3} = (1+o(1))\frac{24(\log(np))^2}{n^3p^4} \geq 2c,$$

as required.

The case when $n^{-3/4} \leq p = o(n^{-1/2})$ is dealt with in a similar way. Indeed, setting $r = 4pn^{3/2}\log n$ gives $pr/4 \geq \log(n/r)$ (with room to spare) and thus

$$\phi \geq pr^2/4 = (pn)^{3+o(1)} \geq 4\log n.$$

Also, we have

$$\frac{\phi}{\mathbb{E}[X_3]} = (1+o(1))\frac{p}{4}\left(4pn^{3/2}\log n\right)^2\frac{6}{n^3p^3} = (1+o(1))24(\log n)^2 \geq 2c,$$

as required. □

Chapter 6

The length of random subsets of a finite set

As outlined in Section 2.3, in this chapter we will study the length of a random subset $\mathcal{P}(n,p)$ of $\mathcal{P}(n)$. Recall that $\mathcal{P}(n,p)$ is obtained by selecting the elements of $\mathcal{P}(n)$ independently with probability p, and that if we order the elements of $\mathcal{P}(n,p)$ by inclusion, we obtain a random partially ordered set (also called a *poset*).

Brightwell [17] gives an overview of other models of random posets. The *uniform model* was first investigated by Kleitman and Rothschild, who in [49] asymptotically determined the number and typical structure of partially ordered sets on n points. Recently, Prömel, Steger, and Taraz [73, 71] gave an asymptotic for the number of partially ordered sets on n points with fixed density, thus solving an open problem of Dhar, Kleitman, and Rothschild. As another example, several authors (see e.g. Bollobás and Brightwell [12]) studied the length of the random poset generated by picking n random points in the k-dimensional hypercube. However this model, although superficially similar to ours, produces different results and requires other proof techniques, whereas the proofs of this chapter are based on the techniques developed by Fill and Pemantle [31], where they prove the analogue of Theorem 6.4 in the context of oriented bond percolation on the hypercube (see also the remark after Lemma 6.7). Our methods are also related to those used in the past to investigate the diameter and other parameters of random subgraphs of the n-dimensional hypercube (see e.g. Bollobás, Kohayakawa and Łuczak [13, 14]).

As described in Section 2.3, the first result on $\mathcal{P}(n,p)$ is due to Rényi [75]. In [74, 76], he also proved results on $\mathcal{P}(n,p)$ which have applications in information theory. Weber (see e.g. [93, 94]) was motivated by the fact that $\mathcal{P}(n,p)$ can be viewed as a random Boolean function on n variables and established asymptotic bounds on the likely number of conjunctions which are needed to represent such a function in disjunctive normal form. More recent results on $\mathcal{P}(n,p)$ were proven by Kohayakawa and Kreuter [50] (see Corollary 7.2) and Kreuter [54].

Throughout, we refer to the subset $\mathcal{P}_t(n) \subset \mathcal{P}(n)$ whose elements have cardinality t as the tth *level* of $\mathcal{P}(n)$. The tth level of $\mathcal{P}(n,p)$ is defined in the same way. The (Hamming) distance $d_H(x,y)$ of x, $y \in \mathcal{P}(n)$ is given by $|(x \setminus y) \cup (y \setminus x)|$.

This chapter is organized as follows. In the next section, we state our main results, namely Theorems 6.1 and 6.4. Our main technical tool, Lemma 6.7, is stated and proved in Section 6.2. Theorems 6.1 and 6.4 are then proved in Sections 6.3 and 6.4.

6.1 Results

Amongst others, Kreuter [54] proved that the length of $\mathcal{P}(n,p)$ is constant when p decreases exponentially. The following theorem tells us about the length of $\mathcal{P}(n,p)$ when p decays at a rate slower than exponential:

Theorem 6.1 *Let $k = k(n) \leq n{-}1$ be an integer valued function tending to infinity. Then*

$$\mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p) \geq k{-}1\,] \to \begin{cases} 0 & \textit{if } p\,(k{+}1)^{n/k} \to 0, \\ 1 & \textit{if } p\,(k{+}1)^{n/k}(k/n)^2 \to \infty. \end{cases} \tag{6.1}$$

At the end of Section 6.2 we indicate how one may remove the factor $(k/n)^2$ at the expense of more extensive calculations. One consequence of Theorem 6.1 is that when viewed as a poset, $\mathcal{P}(n,p)$ is in a sense close to being a ranked poset. More precisely, given a constant ν with $0 < \nu < 1$ and an element $x \in \mathcal{P}_{\nu n}(n,p)$ one can apply Theorem 6.1 to $\mathcal{P}(\nu n)$ to obtain bounds on the length of the longest chain in $\mathcal{P}(n,p)$ whose maximal element is x.

Another immediate corollary of the above theorem is that when p decreases polynomially in n, the degree of the polynomial turns out to be the average distance between two elements of a typical maximum chain.

Corollary 6.2 *Let $r \geq 1$ be a fixed constant. Then*

$$\mathbb{P}\left[\,\mathrm{length}\,\mathcal{P}(n,p) \geq \frac{n}{r} - 2\,\right] \to \begin{cases} 0 & \textit{if } pn^r \to 0 \\ 1 & \textit{if } pn^r \to \infty. \end{cases} \tag{6.2}$$

We remark that the above threshold could be sharpened slightly by the use of branching processes and indicate how this may be done at the end of the final section. Another immediate corollary of Theorem 6.1 is the following asymptotic formula for the length of $\mathcal{P}(n,p)$.

Corollary 6.3 *If $c^n = o(p)$ for any $0 < c < 1$ and $p = \mathcal{O}(1/n)$, then almost surely*

$$\mathrm{length}\,\mathcal{P}(n,p) = (1 + o(1))\,\frac{n}{\log 1/p}\,(\log n - \log\log 1/p)\,. \tag{6.3}$$

The following theorem (which we stated already in Section 2.3) improves Corollary 6.2 for the special case when $r = 1$ and gives a detailed picture of the final stage of the evolution of the length of $\mathcal{P}(n,p)$.

Theorem 6.4 *Let $c \geq \mathrm{e}$ be a given constant. Define $p(n) = c/n$. Let $\eta = \eta(c)$ satisfy $\eta = \mathrm{e}^{c(\eta-1)}$ and $0 < \eta < 1$. Then*

$$\begin{aligned} \mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p) = n{-}2\,] &\to (1-\eta)^2, \\ \mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p) = n{-}3\,] &\to 2\eta(1-\eta), \\ \mathbb{P}[\,\mathrm{length}\,\mathcal{P}(n,p) = n{-}4\,] &\to \eta^2. \end{aligned}$$

Recall that η is the extinction probability of a branching process whose family size has Poisson distribution with parameter c. Since the top and bottom levels of $\mathcal{P}(n)$ contain only a single element, almost surely we have that $\operatorname{length}\mathcal{P}(n,p) < n-1$ as long as $p \to 0$. Finally, the following proposition shows that the previous theorem is essentially best possible in the sense that if $c < \mathrm{e}$, then $\mathcal{P}(n,p)$ becomes significantly shorter. Indeed, if $p = p(n) = c/n$, Theorem 6.1 and Proposition 6.5 imply that, very crudely, almost surely we have

$$n - \operatorname{length}\mathcal{P}(n,p) \quad \begin{cases} = \mathcal{O}(1) & \text{if } c \geq \mathrm{e} \\ \to \infty. & \text{if } c < \mathrm{e} \end{cases} \tag{6.4}$$

Proposition 6.5 *Let $c < \mathrm{e}$ be a given constant. Define $p(n) = c/n$. Then almost surely*

$$\operatorname{length}\mathcal{P}(n,p) \leq \left(1 - \frac{\log(\mathrm{e}/c)}{3\log n}\right) n.$$

Proof. Let

$$a = \left\lceil \frac{\log(\mathrm{e}/c)}{3\log n} n \right\rceil .$$

For $\mathcal{Q} \subset \mathcal{P}(n)$ denote by $Y(\mathcal{Q})$ the number of $(n-a)$-element chains in $\mathcal{Q}$. A chain on $n-a$ elements corresponds to an ordered partition of $[n]$ into $n-a$ nonempty sets. To bound the number of those, first choose $n-a$ elements, one for each partition class. There are $(n)_{n-a} = n!/a!$ ways of doing this. Then order the remaining a elements and distribute them among the $n-a$ classes, with $(n-a)^a$ possibilities for this second step. Hence by Stirling's formula

$$Y(\mathcal{P}(n)) \leq n^a n! = \mathcal{O}(n^{a+1/2}(n/\mathrm{e})^n).$$

Thus the expected value of $Y = Y(\mathcal{P}(n,p))$ is

$$\mathbb{E}[Y] = \mathcal{O}(n^{a+1/2}(n/\mathrm{e})^n p^{n-a}) = \mathcal{O}(n^{2a+1/2}(c/\mathrm{e})^{n-a}) = o(1)$$

by the definition of a. $\square$

6.2 Regular chains

The aim of this section is to prove Lemma 6.7 below. This lemma implies that the number of certain "regular" chains is not too far from its expected value provided that p is large enough. Once we have this result, Theorem 6.1 will follow immediately, while we still have to do some additional work for Theorem 6.4. These "regular" chains are defined as follows: suppose x, $y \in \mathcal{P}(n)$ are given with $x \subset y$. Given any r and k, we say $(x_1, \dots, x_k)$ is an (x,y,r,k)-*chain* if $x \subset x_1 \subset \cdots \subset x_k \subset y$ with the additional condition that $|x_i| = |x| + \lceil ir \rceil$ for all i. Since r and k are always related in what follows, (x,y,r,k)-chains are simply referred to as (x,y,r)-chains.

Throughout this section we assume that $x \in \mathcal{P}_t(n)$ and $y \in \mathcal{P}_{n-t}(n)$ for some fixed integer t. Also k, the number of elements in a chain, is a given integer function of n with $k \le n-2t-1$. As a measure of the average distance between two chain elements we define

$$r = \begin{cases} 1 & \text{if } k = n-2t-1 \\ (n-2t-1)/(k+1) & \text{if } k < n-2t-1. \end{cases} \tag{6.5}$$

The extra minus one in the fraction is a technical artifice to ensure that $\mathcal{P}_{|x_i|}(n,p)$ is large enough even when r is not an integer and i is close to k. We separate the two cases since we want to achieve the best possible result when $r = 1$. Now, for given $m \in \mathbb{N}$, let

$$\psi(m) = \begin{cases} m+1 & \text{if } k = n-2t-1 \\ \lfloor (m+1)r \rfloor + 1 & \text{if } k < n-2t-1. \end{cases} \tag{6.6}$$

Writing $x_0 = x$ and $x_{k+1} = y$, the Hamming distance between chain elements x_j and $x_{j'}$ with $j' > j$ is then at most $\psi(j'-j-1)$. We shall often make use of the fact that $\psi(j') - \psi(j) \le (j'-j)r + 1$. Finally we assume that

$$p(n) = d\,\psi(0)! \left(\frac{\mathrm{e}}{(n-2t)(k+1)^{1/k}} \right)^r \tag{6.7}$$

with $d = d(n) \ge 1$. The following form of Stirling's formula (valid for all $n > 0$) will suffice for our purposes:

$$\sqrt{n} \left(\frac{n}{\mathrm{e}} \right)^n < n! < 3\sqrt{n} \left(\frac{n}{\mathrm{e}} \right)^n. \tag{6.8}$$

Let $\mathcal{S}$ be a family of subsets of $\mathcal{P}(n)$ (in what follows, $\mathcal{S}$ will usually consist of various chains in $\mathcal{P}(n)$) and define $X_{\mathcal{S}}$ to be the number of elements of $\mathcal{S}$ in $\mathcal{P}(n,p)$ as in Section 3.2. Also define $\Delta(\mathcal{S})$ as in Section 3.2.

Proposition 6.6 *Let $t \in \mathbb{N}$, $x \in \mathcal{P}_t(n)$ and $y \in \mathcal{P}_{n-t}(n)$ with $x \subset y$ be given. Let $\mathcal{S} = \mathcal{S}(x,y,r)$ be the set of all (x,y,r)-chains in $\mathcal{P}(n)$. Then*

$$\lim_{n\to\infty} \mathbb{E}[X_{\mathcal{S}}] = \infty.$$

Proof. Let $R_k = |\mathcal{S}|$ denote the number of (x,y,r)-chains in $\mathcal{P}(n)$. By definition of r and ψ we have

$$\begin{aligned} R_k &= \binom{n-2t}{\lceil r \rceil, \lceil 2r \rceil - \lceil r \rceil, \ldots, n-2t-\lceil kr \rceil} \\ &\ge \frac{(n-2t)!}{\psi(0)!^{k+1}}. \end{aligned} \tag{6.9}$$

Thus

$$\begin{aligned} \mathbb{E}[X_{\mathcal{S}}] = R_k\, p^k &\overset{(6.8)}{\ge} \left(\frac{n-2t}{\mathrm{e}} \right)^{n-2t-kr} \frac{\sqrt{n-2t}}{(k+1)^r \psi(0)!} \\ &\ge \left(\frac{n-2t}{\mathrm{e}} \right)^{\psi(0)} \frac{\sqrt{n-2t}}{(k+1)^{\psi(0)} \psi(0)!}. \end{aligned}$$

The result now follows by Stirling's formula (noting that $(\psi(0)/r)^{\psi(0)}$ is bounded by some constant) and the fact that $(n-2t)/r \to \infty$. $\square$

Lemma 6.7 *Let* $t \in \mathbb{N}$, $x \in \mathcal{P}_t(n)$ *and* $y \in \mathcal{P}_{n-t}(n)$ *with* $x \subset y$ *be given. Let* $\mathcal{S} = \mathcal{S}(x, y, r)$ *be the set of all* (x, y, r)*-chains in* $\mathcal{P}(n)$*. Then*

$$\Delta(\mathcal{S}) \leq \frac{9}{d}\,\mathbb{E}[X_{\mathcal{S}}]^2,$$

provided n *is large enough.*

There are several results on random subgraphs of the hypercube which rely on a similar lemma, essentially giving bounds on the number of pairs of various kinds of paths in the hypercube. In fact, case (a) of the proof of Lemma 6.7 is similar to Lemma 2.2 in Fill and Pemantle [31], but is sketched for completeness. However, in general their bounds do not apply to our model (even for the simplest case $r = 1$) owing to the differences in the geometry of the paths involved.

Proof of Lemma 6.7. Consider any ordered pair (α, β) of (x, y, r)-chains with $\alpha \sim \beta$. Let $i = i(\alpha, \beta) = |\alpha \cap \beta|$ be the number of common elements. Note that $0 < i < k$.

A *disjoint interval* γ is a subset of $\alpha \setminus \beta$ with the property that, whenever we have w, $w' \in \gamma$, $v \in \alpha$, and $w \subset v \subset w'$, we must have $v \in \gamma$. We then let l be the maximum cardinality (or *size*) of a disjoint interval and let s be the number of disjoint intervals. Note that $0 < l \leq k - i$ and $0 < s \leq k - i$.

Now write $\sum_{i,l,s}$ for the sum over all ordered pairs of (x, y, r)-chains α and β with $i(\alpha, \beta) = i$, $l(\alpha, \beta) = l$, and $s(\alpha, \beta) = s$. Let

$$\Delta_{ils} = \sum_{i,l,s} \mathbb{E}[I_\alpha I_\beta]$$

and define Δ_i, Δ_{il}, and Δ_{is} similarly. We can then write $\Delta = \Delta(\mathcal{S})$ as

$$\Delta = \sum_{i=1}^{k-1} \Delta_i, \quad \text{where} \quad \Delta_i = \sum_{l=1}^{k-i} \Delta_{il} = \sum_{s=1}^{k-i} \Delta_{is} \quad \text{and} \quad \Delta_{is} = \sum_{l=1}^{k-i} \Delta_{ils}. \tag{6.10}$$

Let now i, l and $s > 0$ with $i + l \leq k$ and $i + s \leq k$ be given. Then an *admissible* (i, l, s)*-configuration* is a set $A \subset [k]$ with $|A| = i$ such that the longest interval of consecutive elements in $[k] \setminus A$ contains exactly l elements and such that the number of intervals of consecutive elements is exactly s. The set of all admissible (i, l, s)-configurations, denoted $\mathcal{A}_{ils}$, corresponds to the possible intersection patterns of two (x, y, r)-chains (indeed, the configurations give the positions in which two chains may intersect). Let $C_{ils} = |\mathcal{A}_{ils}|$. Let $C_{il} = \sum_s C_{ils}$, $C_{is} = \sum_l C_{ils}$, and $C_i = \sum_l C_{il}$. Note the trivial bounds $C_{ils} \leq C_{il} \leq C_i \leq \binom{k}{i}$.

Now we investigate the number of ways in which a given configuration can be *extended*. Let i, l and s be given and let $\alpha = (x_1 \subset x_2 \subset \cdots \subset x_k)$ be a fixed (x, y, r)-chain and $A \in \mathcal{A}_{ils}$. Let $E(A)$ be the number of possible extensions corresponding to the intersection pattern given by A: i.e., $E(A)$ is the number of β that are (x, y, r)-chains in $\mathcal{P}(n)$ with $\alpha \cap \beta = \{x_i : i \in A\}$. Let

$$E_{ils} = \max_{A \in \mathcal{A}_{ils}} E(A).$$

This time, define $E_{il} = \max_s E_{ils}$ and $E_{is} = \max_l E_{ils}$. To estimate E_{ils}, note that each disjoint interval, between x_j and $x_{j'}$ with $j' > j$ say, contributes a factor of

$$\binom{\lceil j'r \rceil - \lceil jr \rceil}{\lceil (j+1)r \rceil - \lceil jr \rceil, \ldots, \lceil j'r \rceil - \lceil (j'-1)r \rceil}. \tag{6.11}$$

By the log-convexity of factorials, $E(A)$ is then maximized if A has as few disjoint intervals as possible and if as many of these intervals as possible have maximum size. Instead of bounding E_{ils} directly, it turns out to be more convenient to consider E_{ils}/R_k, where R_k is defined to be the number of (x, y, r)-chains in $\mathcal{P}(n)$ as in the proof of Proposition 6.6. This has the advantage that some of the factorials involving $\lceil r \rceil$ and $\lfloor r \rfloor$ cancel straightaway. The same also applies to E_{il} and E_{is}: for instance, the above argument and the formula for R_k given in the proof of Proposition 6.6 imply that

$$\frac{E_{il}}{R_k} \leq \frac{\psi(0)!^{i-1}}{(n-2t)!} \psi(l)! \, \psi(k-i-l)!. \tag{6.12}$$

Indeed, note that we must have at least one disjoint interval of size l and that in the worst case, all the remaining disjoint vertices are contained in a single disjoint interval of size $k - i - l$. Furthermore, for each disjoint interval of size l', the $l' + 1$ factorials in the denominator of (6.11) will cancel with the corresponding factorials in (6.9). We remark that the above inequality also shows that one may interpret E_{il}/R_k as an upper bound on the fraction of all $(x, y, 1)$-chains whose intersection with a given (x, y, r)-chain is an admissible (i, l)-configuration.

For notational convenience let $\delta_{ils} = \Delta_{ils}/\mathbb{E}[X_{\mathcal{S}}]^2$ and define the δ_i etc. similarly. Recalling that $\mathbb{E}[X_{\mathcal{S}}] = R_k \, p^k$, we have

$$\delta_{ils} = \frac{\Delta_{ils}}{\mathbb{E}[X_{\mathcal{S}}]^2} \leq \frac{R_k \, p^{2k-i} \, C_{ils} \, E_{ils}}{\mathbb{E}[X_{\mathcal{S}}]^2} \leq \frac{C_{ils} \, E_{ils}}{R_k \, p^i}. \tag{6.13}$$

We will bound the δ_{ils} by applying different estimates for different ranges of i, l and s. To fix these ranges we need to define the following constants:

$$\lambda = 1/40 \quad \text{and} \quad \rho = 1/40\mathrm{e}. \tag{6.14}$$

Choose $\kappa > 0$ sufficiently small so that

$$(1 - \lambda)(1 + \kappa)^2 < 1. \tag{6.15}$$

Choose $\sigma > 0$ sufficiently small so that

$$(\mathrm{e}/\sigma)^{2\sigma}(1 - \lambda)^{\rho}(1 + \kappa)^2 < 1. \tag{6.16}$$

Finally choose $l_0 \in \mathbb{N}$ large so that for any $l \geq l_0$,

$$2\left(\frac{4\mathrm{e}}{l}\right)^{\sigma} < 1 \quad \text{and} \quad \psi(l)!^{1/(l-1)} \leq \left(\frac{lr(1+\kappa)}{\mathrm{e}}\right)^{r}. \tag{6.17}$$

We assume that n is large enough whenever needed. We also make use of the following inequality, which follows from Stirling's formula (6.8) and the definition of r and p in (6.5) and (6.7):

$$\frac{\psi(0)!^{i}}{p^{i}\,(n-2t)!} \leq \frac{(k+1)^{ir/k}}{d^{i}\sqrt{kr}}\left(\frac{\mathrm{e}}{(k+1)r}\right)^{\psi(k-i)}. \tag{6.18}$$

(a) $i < \lambda k$.

It turns out that the only significant contribution of Δ_{ils} to Δ comes from the case when i is small and $l = k - i$ (which implies that $s = 1$): in the case (a1), when l is small, we have only a trivial bound on the number of configurations but this is compensated by the fact that, in this case, the number of extensions is severely limited. In the case (a2), when l is large, the situation is reversed. Overall this allows us to show that δ_i is a constant multiple of the worst case contribution $\delta_{i,k-i}$.

(a1) $l \leq k - 4i$.

We have $\sum_{l=1}^{k-4i} C_{il} \leq C_i \leq \binom{k}{i}$ and by (6.12),

$$\frac{E_{il}}{R_k} \leq \frac{\psi(0)!^{i}}{(n-2t)!}\,\psi(k-4i)!\,\psi(3i)!.$$

Thus, using $r \geq 1$ and $\lambda = 1/40$, we have

$$\begin{aligned}
\sum_{l\leq k-4i} C_{il}\frac{E_{il}}{R_k} &\leq \frac{\psi(0)!^{i}}{(n-2t)!}\,\psi(k-i)!\,\frac{\psi(k-4i)!}{\psi(k-i)!}\,\psi(3i)!\binom{k}{i} \\
&\overset{(3.2)}{\leq} \frac{\psi(0)!^{i}}{(n-2t)!}\,\psi(k-i)!\left(\frac{(4ir)^3}{((k-4i)r)^3}\,\frac{\mathrm{e}k}{i}\right)^{\lfloor ir\rfloor} \\
&\leq \frac{\psi(0)!^{i}}{(n-2t)!}\,\psi(k-i)!\left(\frac{64\mathrm{e}\lambda^2}{(1-4\lambda)^3}\right)^{\lfloor ir\rfloor} \\
&\leq \frac{\psi(0)!^{i}}{(n-2t)!}\,\psi(k-i)!\,\frac{1}{4}.
\end{aligned}$$

(a2) $l > k - 4i$.

As there are $k - l + 1$ choices for picking the position of the lowest disjoint interval of length l and at most $\binom{k-l-1}{i-1}$ choices for picking the remaining elements we have

$$C_{il} \leq (k-l+1)\binom{k-l-1}{i-1}. \tag{6.19}$$

Define M_{il} to be the product of the right-hand sides of inequalities (6.19) and (6.12). Thus $C_{il}\, E_{il}/R_k \leq M_{il}$. Furthermore, since $r \geq 1$,

$$\begin{aligned} \frac{M_{i,l+1}}{M_{i,l}} &\geq \frac{k-l}{k-l+1}\,\frac{k-i-l}{k-l-1}\left(\frac{(l+1)r}{(k-i-l+2)r}\right)^{r} \\ &\geq \frac{k-i-l}{k-i-l+2}\;\frac{k-4i}{4i} \geq \frac{1}{3}\;\frac{1-4\lambda}{4\lambda} = 3. \end{aligned}$$

Since $\sum_{j=0}^{\infty} 3^{-j} = 3/2$, the total contribution of the M_{il} for $l > k-4i$ is at most $3/2$ times that of the maximum term $l = k-i$.

Putting both (a1) and (a2) together gives

$$\delta_i \leq \sum_{l=1}^{k-i} \delta_{il} \leq 2\,\frac{M_{i,k-i}}{p^i} \leq 2\,(i+1)\,\frac{\psi(0)!^i}{(n-2t)!p^i}\,\psi(k-i)!.$$

But $n - 2t - \psi(k-i) \geq ir$ and $i \leq \lambda k$ imply that

$$\frac{\psi(k-i)!}{(n-2t)!} \leq ((1-2\lambda)(n-2t))^{-ir}.$$

Thus

$$\delta_i \leq\; 2(i+1)\,d^{-i}\left(\frac{(k+1)^{1/k}}{\mathrm{e}(1-2\lambda)}\right)^{ir} \leq\; 2(i+1)\left(\frac{1}{2d}\right)^{i},$$

where the second line follows since the fraction in brackets is less than 1/2. Thus, using this time that $d \geq 1$,

$$\sum_{1 \leq i \leq \lambda k} \delta_i \leq 2\left(\left(1-\frac{1}{2d}\right)^{-2} - 1\right) \leq \frac{8}{d}.$$

(b) $\lambda k \leq i < (1-\rho)k$.

(b1) $l < l_0$.

We have $C_{il} \leq \binom{k}{i} \leq 2^k$. To estimate E_{il}/R_k, note that chain elements in the intersection which lie at either end of a disjoint stretch have distance at most $\psi(l)$. Thus for each disjoint interval (of size $\bar{l}$ say) we have a contribution of a factor of $\psi(\bar{l})! \leq \psi(0)!\,(2\psi(l))^{\bar{l}r}$. Using the fact that there are $k-i$ disjoint vertices and at most $i+1$ disjoint intervals yields the following bound:

$$\frac{E_{il}}{R_k} \leq \frac{\psi(0)!^{i+1}}{(n-2t)!}\,(2\psi(l))^{(k-i)r} \leq \frac{\psi(0)!^{i}}{(n-2t)!}\,(2\psi(l))^{(k-i+1)r}.$$

Hence, by inequality (6.18), and noting (crudely) that $(k+1)^{ir/k} \leq (k+1)^r < \mathrm{e}^{\rho k r}$, we have

$$\begin{aligned}
\delta_{il} &\leq \frac{\psi(0)!^i}{p^i(n-2t)!}\,(2\psi(l))^{(k-i+1)r}\,2^k \\
&\leq \frac{e^{\rho kr}}{d}\left(\frac{\mathrm{e}}{(k+1)r}\right)^{(k-i+1)r}\quad(2(l_0+1)r)^{(k-i+1)r}\,2^k \\
&\leq \frac{1}{d}\left(\frac{2^{1+1/\rho}(l_0+1)\mathrm{e}^2}{k}\right)^{\rho kr} \\
&\leq \frac{1}{dk^3}.
\end{aligned}$$

In the last line we used the fact that $k \to \infty$.

(b2) $l > l_0$.

Fix a chain α and any configuration with parameters i, l, and s. Denote the size of the disjoint intervals by l_j, for $1 \leq j \leq s$. Then there are at most

$$\psi(0)^{i-(s-1)}\prod_{j=1}^{s}\psi(l_j)!$$

$(x, y, 1)$-chains which meet α exactly at the i chosen points. However, this product is maximized if we have as many disjoint intervals of size l as possible, say q of them. The remaining disjoint vertices will be distributed among one interval of size $\bar{l}$, with $1 \leq \bar{l} < l$, and $s-q-1$ unit size intervals. Then noting that $ql+\bar{l}+(s-q-1) = k-i$ and $s-q-1 < s$ gives

$$\begin{aligned}
\frac{E_{ils}}{R_k} &\leq \frac{\psi(0)!^{i-(s-1)}}{(n-2t)!}\,\psi(l)!^{k-i-s-\bar{l}/l-1}\,\psi(\bar{l})!\,\psi(1)!^{s-1} \\
&\leq \frac{\psi(0)!^i}{(n-2t)!}\,\psi(l)!^{k-i-s/l-1}\,(4r)^{sr}.
\end{aligned}$$

Now use $l \leq k-i < (1-\lambda)k$ and the definition of l in (6.17) to obtain

$$\begin{aligned}
\frac{E_{ils}}{R_k} &\overset{(3.2)}{\leq} \psi(0)!^i\left(\frac{\mathrm{e}}{n-2t}\right)^{n-2t}\left(\frac{lr(1+\kappa)}{\mathrm{e}}\right)^{(k-i)r}\left(\frac{4\mathrm{e}r}{(l+1)r}\right)^{sr} \\
&\leq \psi(0)!^i\left(\frac{\mathrm{e}}{n-2t}\right)^{ir}((1-\lambda)(1+\kappa))^{(k-i)r}\left(\frac{4\mathrm{e}}{l}\right)^{s}.
\end{aligned}$$

By the definition of p in (6.7), and again crudely estimating $(k+1)^{ir/k} \leq (k+1)^r \leq (1+\kappa)^{\rho kr}$, we have

$$\delta_{il} \leq \frac{1}{d}\sum_{s=1}^{k} C_{ils}\left((1-\lambda)(1+\kappa)^2\right)^{\rho kr}\left(\frac{4\mathrm{e}}{l}\right)^{s}.$$

When bounding C_{ils} we distinguish two cases. First suppose that $s \leq \sigma k$. Note that any admissible configuration may be defined in the following way: we first fix the lowest point of each of the s disjoint intervals. $\binom{k}{s}$ is a crude bound on the number of choices. This first step fixes the positions of at least $s-1$ common vertices, since unless a disjoint interval is right at the top, it must obviously be below a common vertex. We now fix a configuration uniquely by successively adding a further $i-s$ common vertices on top of the existing common intervals. In other words, we are now drawing $i-s+1$ elements (with replacement) from a set of s distinct elements, giving us

$$\binom{s+(i-s+1)-1}{i-s+1} = \binom{i}{s-1}$$

choices on the second step. Thus

$$C_{ils} \le \binom{k}{s}\binom{i}{s-1} \overset{(3.2)}{\le} \left(\frac{\mathrm{e}k}{s}\right)^{s}\left(\frac{\mathrm{e}i}{s-1}\right)^{s-1} \le \left(\frac{\mathrm{e}k}{s}\right)^{2s} \le \left(\frac{\mathrm{e}}{\sigma}\right)^{2\sigma k}.$$

For $s > \sigma k$ we just use the trivial bound $C_{ils} \le 2^k$. Plugging in and using (6.16) and (6.17) this yields

$$\begin{aligned}\delta_{il} &\le \frac{1}{d}\left(\sum_{1\le s\le\sigma k}\left(\frac{\mathrm{e}}{\sigma}\right)^{2\sigma k}((1-\lambda)(1+\kappa)^2)^{\rho k r} + \sum_{\sigma k<s\le k} 2^k\left(\frac{4\mathrm{e}}{l}\right)^{s}\right)\\ &\le \frac{1}{d}\left(k^{-4} + k2^k\left(\frac{4\mathrm{e}}{l_0}\right)^{\sigma k}\right)\\ &\le \frac{1}{dk^3}. \end{aligned} \tag{6.20}$$

(c) $i \ge (1-\rho)k$.

There are at most $\binom{i+s}{s}$ ways of choosing the relative order of i common vertices and the s disjoint intervals. We now fix the size of the intervals, by distributing the $k-i-s$ remaining "disjoint" vertices among the s disjoint intervals. We are choosing these elements (with replacement) from s "urns", giving at most $\binom{k-i-1}{k-i-s}$ possibilities for this step. Hence, by applying Stirling's inequality and then that $i+s \le k$ and $m^{-m} \le 3(m+1)^{-m}$ for all $m \ge 0$,

$$\begin{aligned}C_{is} &\le \binom{i+s}{s}\binom{k-i-1}{k-i-s}\\ &\le \left(\frac{\mathrm{e}(i+s)}{s}\right)^{s}\left(\frac{k-i}{\mathrm{e}}\right)^{k-i}\left(\frac{k-i-s}{\mathrm{e}}\right)^{s-(k-i)}\left(\frac{\mathrm{e}}{s-1}\right)^{(s-1)}\\ &\le 3^{16}\left(\frac{\mathrm{e}k}{s^2}\right)^{s}(k-i)^{k-i+1}\,(k-i-s+15)^{s-(k-i)}. \end{aligned} \tag{6.21}$$

Similarly to case (b2), but without the constraint on the size of the longest disjoint interval, the extremal configuration for given s is now attained if we have $s-1$ disjoint intervals of unit size and one interval containing all the other disjoint vertices. Thus

$$\begin{aligned}\frac{E_{is}}{R_k} &\le \frac{\psi(0)!^{i-(s-1)}}{(n-2t)!}\,\psi(k-i-s+1)!\,\psi(1)!^{s-1}\\ &\le \frac{3\psi(0)!^{i}}{(n-2t)!}\left(\frac{(k-i-s+15)r}{\mathrm{e}}\right)^{\psi(k-i-s+1)+1/2}(4r)^{(s-1)r}. \end{aligned}$$

The second line follows by Stirling's formula (6.8). By the definition of h, the fact that $(k-i-s+15) \leq 3^3(k-i)$, and inequality (6.18), we now have

$$\begin{aligned} \frac{E_{is}}{R_k\, p^i} &\leq \frac{3^4}{d\sqrt{k}}(k-i)^{3/2}(k+1)^r \times \\ &\quad \times \left(\frac{\mathrm{e}}{(k+1)r}\right)^{\psi(k-i)} \left(\frac{(k-i-s+15)r}{\mathrm{e}}\right)^{\psi(k-i)-(s-1)r} (5r)^{(s-1)r} \\ &\leq \frac{3^7}{d\sqrt{k}}(k-i)^{3/2} \left(\frac{k-i-s+15}{(k+1)^{1-1/(k-i+1)}}\right)^{\psi(k-i)} \left(\frac{5\mathrm{e}}{k-i-s+15}\right)^{(s-1)r} \\ &\leq \frac{3^{10}}{d\sqrt{k}}(k-i)^{5/2} \left(\frac{k-i-s+15}{k}\right)^{k-i} \left(\frac{5\mathrm{e}}{k-i-s+15}\right)^{s-1}. \qquad (6.22) \end{aligned}$$

The last line follows since one may check that both of the bracketed terms in the second line are smaller than one. Combining inequalities (6.21) and (6.22), we have

$$\delta_{is} \leq \frac{3^{30}}{d\sqrt{k}}(k-i)^4 \left(\frac{3\mathrm{e}^2 k}{s^2}\right)^s \left(\frac{k-i}{k}\right)^{k-i}.$$

It is easily verified that for fixed k and i, the expression $\phi(s) = \left(3\mathrm{e}^2 k/s^2\right)^s$ is maximized at $s = \sqrt{3k}$ if $i \leq k - \sqrt{3k}$ and is maximized at $s = k - i$ if $i \geq k - \sqrt{3k}$. Thus if $i \leq k - \sqrt{3k}$, we have

$$\delta_{is} \leq \frac{3^{30}}{d\sqrt{k}}(k-i)^4 \mathrm{e}^{2\sqrt{3k}} \rho^{k-i} \leq \frac{3^{30}}{d\sqrt{k}} k^4 (\rho\mathrm{e})^{2\sqrt{3k}},$$

and if $i \geq k - \sqrt{2k}$, we have

$$\delta_{is} \leq \frac{3^{30}}{d\sqrt{k}}(k-i)^4 \left(\frac{3\mathrm{e}^2}{k-i}\right)^{k-i}.$$

Putting both the above together, we have

$$\begin{aligned} \sum_{(1-\rho)k \leq i < k} \delta_i &\leq \frac{3^{30}}{d\sqrt{k}} \left(\sum_{(1-\rho)k \leq i \leq k-\sqrt{3k}} k^5 (\rho\mathrm{e})^{2\sqrt{3k}} + \sum_{k-\sqrt{3k} \leq i < k} (k-i)^5 \left(\frac{3\mathrm{e}^2}{k-i}\right)^{k-i} \right) \\ &\leq \frac{3^{30}}{d\sqrt{k}} \left(k^6 (\rho\mathrm{e}^2)^{\sqrt{3k}} + K\right) \leq \frac{1}{dk^{1/4}}, \end{aligned}$$

where K is some fixed constant.

Summing up the bounds in (a)–(c) now gives the desired result:

$$\frac{\Delta}{\mathbb{E}[X_{\mathcal{S}}]^2} \leq \frac{8}{d} + k^2 \frac{1}{dk^3} + \frac{1}{dk^{1/4}} \leq \frac{9}{d}.$$

□

We remark that the above result is best possible up to a constant factor. This may be seen by noting that the contribution to $\Delta/\mathbb{E}[X_{\mathcal{S}}]^2$ of those pairs of chains which have only the lowest element in common is asymptotically equal to $1/\mathrm{e}d$.

In the previous section, we mentioned that the threshold in Theorem 6.1 could be tightened by a factor of $(n/k)^2$: this can be done by considering not just (x, y, r)-chains (where the level of the ith element x_i is fixed at $|x_i| = |x| + \lceil ir \rceil$), but allowing the level of x_i to deviate from this by a factor of $\sqrt{r}$. The calculations are then more cumbersome, but the techniques of this section still apply.

Finally, we state a slight variant of the previous lemma, which will be needed in Chapter 7, the difference being that the regular chains under consideration are now "squeezed": let k be given as before and let $m = m(n)$ and $\overline{m} = \overline{m}(n)$ be given integer functions which are $o(n)$. In the definitions of r, ψ, and p in (6.5), (6.6), and (6.7) respectively, replace the term $2t$ by $m + \overline{m}$. Thus, for instance,

$$r = \begin{cases} 1 & \text{if } k = n - m - \overline{m} - 1 \\ (n-m-\overline{m}-1)/(k+1) & \text{if } k < n - m - \overline{m} - 1. \end{cases} \tag{6.23}$$

Then we say that $(x_1 \subset \cdots \subset x_k)$ is an $[m, \overline{m}, r]$-*chain* if $|x_i| = m + \lceil ir \rceil$ for all i.

Lemma 6.8 *Let $\mathcal{S}$ be the set of all $[m, \overline{m}, r]$-chains in $\mathcal{P}(n)$. Then for large enough n, and*

$$p(n) \geq d\,\psi(0)! \left(\frac{\mathrm{e}}{(n - m - \overline{m})(k+1)^{1/k}} \right)^r, \tag{6.24}$$

where $d = d(n) \geq 1$, we have

$$\Delta(\mathcal{S}) \leq \frac{9}{d}\,\mathbb{E}[X_{\mathcal{S}}]^2. \tag{6.25}$$

The proof of Lemma 6.7 goes through as before: we may first of all assume without loss of generality that $m \geq \overline{m}$; it turns out that the only difference arises when bounding the number of extensions: in the extremal configuration the largest disjoint interval is now situated at the bottom, which contributes an extra factor of at most $(\psi(l) + m)_m$ to the numerator of the E_{ils}/R_k. Similarly, the second largest interval contributes an extra factor of at most $(\psi(l) + \overline{m})_{\overline{m}}$. This is compensated by the additional $(n)_{m+\overline{m}}$ in the denominator. In fact the only reason for stating Lemma 6.8 separately was to avoid additional notation in the previous lemma and its proof.

We further remark that the above bound on $\Delta(\mathcal{S})$ is only required in Chapter 7. For the purposes of this chapter, it would have been slightly more convenient to bound $\mathbb{E}[X_{\mathcal{S}}^2]$ directly instead of $\Delta(\mathcal{S})$.

6.3 Proof of Theorem 6.1

The technically difficult part in the proof of Theorem 6.1 has been dealt with in Lemma 6.7, and hence this theorem may now be proved easily.

Proof of Theorem 6.1. For the first part assume that $p\,(k+1)^{n/k} \to 0$. Let Y_k be the number of k-element chains in $\mathcal{P}(n, p)$. The number of k-element chains in $\mathcal{P}(n)$ is at most $(k+1)^n$. To see this, note that we may associate with each k-element chain α a $(k+1)$-colouring of $\{1, \ldots, n\}$ as follows. Suppose $\alpha = (x_1 \subset x_2 \cdots \subset x_k)$. Assign colour 1 to all elements in x_1, colour j to all elements in $x_j \setminus x_{j-1}$ for $1 < j \leq k$ and colour $k{+}1$ to $[n]\setminus x_k$. Hence

$$\mathbb{P}[Y_k > 0] \leq \mathbb{E}[Y_k] \leq (k{+}1)^n \, p^k = o(1),$$

which implies the first statement.

To prove the second statement, suppose $p\,(k{+}1)^{n/k}(k/n)^2 \to \infty$. Let r be given by (6.5). It is easily checked that this implies that p satisfies (6.7) with $d = d(n) \to \infty$. Now let $\mathcal{S}$ be the set of $(\emptyset, [n], r)$-chains, so that $X_{\mathcal{S}}$ is the number of $(\emptyset, [n], r)$-chains in $\mathcal{P}(n,p)$. Since $\mathrm{Var}[X_{\mathcal{S}}] \leq \Delta(\mathcal{S}) + \mathbb{E}[X_{\mathcal{S}}]$ (see e.g. [6]), the desired result follows immediately from Chebyshev's inequality after applying Proposition 6.6 and Lemma 6.7 with $t = 0$. □

6.4 Proof of Theorem 6.4

We shall use some results on branching processes in the proof of Theorem 6.4 to "amplify" the results of the previous section. This is a natural approach also exploited in [31]. We first present a few basic concepts and some notation.

Let X be a nonnegative integer valued random variable. A *Galton-Watson branching process* with family size X is a discrete Markov chain $(X_t)_{t=0}^{\infty}$ with the following properties: at time $t = 0$ we have one "ancestor", so we set $X_0 = 1$. For $t \geq 1$ define X_t to be the sum of X_{t-1} independent random variables, each having the same distribution as X (see e.g. [38] for details). As we shall only be concerned with the case when X has Poisson distribution with parameter c, in what follows $(X_t)_{t=0}^{\infty}$ will always refer to a branching process with this property. Let η be the *extinction probability* of this process:

$$\eta = \eta(c) = \mathbb{P}[X_t = 0 \text{ for some } t]. \tag{6.26}$$

A basic result in branching process theory says that if $c > 1$ then η is the (unique) root of $s = \mathbb{E}[s^X] = \mathrm{e}^{c(s-1)}$ in $0 < s < 1$. Another basic result (see [38] again) is that for any constant $k > 0$, we have

$$\lim_{t\to\infty} \mathbb{P}[X_t = k] = 0. \tag{6.27}$$

We now define branching processes in $\mathcal{P}(n)$. Consider any $x \in \mathcal{P}(n)$ at level m, say. An *upward* x-process $(Z_t^n)_{t=0}^{n-m} = (Z_t^n(x))_{t=0}^{n-m}$ is defined by letting Z_t^n denote the number of $z \in P_{m+t}(n,p)$ that are connected to x via an $(x,z,1)$-chain in $\mathcal{P}(n,p)$. A downward y-process is defined similarly. For simplicity, we will henceforth call an $(x,y,1)$-chain an xy-chain.

The following proposition will help us determine the number of descendants we will need at each end to link the "top" and "bottom" of $\mathcal{P}(n,p)$:

Proposition 6.9 *For any given $c \geq \mathrm{e}$ define $\eta = \eta(c)$ as in (6.26). Given $0 < \varepsilon < 1-\eta$, there exists $a = a(c,\varepsilon) \in \mathbb{N}$ with $(43/44)^a \leq \varepsilon$ satisfying the following: for a binomially distributed random variable $Z \sim \mathrm{Bin}(\lceil 2a/(1-\eta-\varepsilon)\rceil, 1-\eta-\varepsilon)$, we have $\mathbb{P}[Z < a] < \varepsilon$.*

Proof. Certainly $\mathbb{P}[Z > a] \leq \mathbb{P}[|Z - \mathbb{E}[Z]| \geq a]$. Thus the result follows from the Chernoff bounds in Section 3.1. $\square$

For positive integers m and n with $n > 2m$, denote by $\mathcal{P}_{m\ldots n-m}(n,p)$ the family of all sets from $\mathcal{P}(n,p)$ of cardinality at least m and at most $n-m$. Let $X \subseteq \mathcal{P}_m(n)$ and $Y \subseteq \mathcal{P}_{n-m}(n)$ be given. We say that X and Y *can be linked* if there are $x \in X$ and $y \in Y$ such that there is an xy-chain contained in $\mathcal{P}_{m+1\ldots n-m-1}(n,p)$.

Lemma 6.10 *Let $\varepsilon > 0$, $c \geq \mathrm{e}$ and $m \in \mathbb{N}$ be given. Let $p = c/n$ and define $\eta = \eta(c)$ as in (6.26). Let $a = a(c,\varepsilon)$ be as in Proposition 6.9. Let*

$$b = \lceil 2a/(1-\eta-\varepsilon) \rceil .$$

Furthermore let $X \subseteq \mathcal{P}_m(n)$ with $|X| = b$ and $Y \subseteq \mathcal{P}_{n-m}(n)$ with $|Y| = b$ and $x \subset y$ for any $x \in X$ and $y \in Y$ be given. Then for large enough n

$$\mathbb{P}[X \text{ and } Y \text{ can be linked}] \geq 1 - \varepsilon.$$

Lemma 6.7 implies that a single pair $x \subset y$ is linked with probability bounded away from zero. Hence, to prove Lemma 6.10, we would like to consider several pairs at the same time to ensure that with arbitrarily high probability we can find at least one that is linked by a chain. The following lemma tells us that this is indeed possible, provided the elements at each level are far enough apart.

Given x, $y \in \mathcal{P}(n)$ with $x \subset y$, denote by $\mathcal{P}^{xy}$ the subset of $\mathcal{P}(n)$ consisting of all elements (including x and y) which are both subset of y and superset of x.

Lemma 6.11 *Let t, $h \in \mathbb{N}$ be given. Suppose that x_1, $x_2 \in \mathcal{P}_t(n)$ and y_1, $y_2 \in \mathcal{P}_{n-t}(n)$ are such that $x_1 \subset y_1$, $x_2 \subset y_2$, $d_H(x_1,x_2) \geq h$ and $d_H(y_1,y_2) \geq h$. Then the proportion of all x_1y_1-chains that meet $\mathcal{P}^{x_2y_2}$ is at most $2\mathrm{e}^{-h/2}$, provided n is large enough.*

Proof. Let $\bar{x} = x_1 \cup x_2$ and $\bar{y} = y_1 \cap y_2$. Then one may easily verify that $\mathcal{P}^{x_1y_1} \cap \mathcal{P}^{x_2y_2} = \mathcal{P}^{\bar{x}\bar{y}}$. Furthermore $\bar{x}$ has cardinality at least $t+h$ and $\bar{y}$ has cardinality at most $n-(t+h)$. The crucial observation now is that any x_1y_1-chain which meets $\mathcal{P}^{x_2y_2}$ must be *contained* in $\mathcal{P}^{x_1\bar{y}} \cup \mathcal{P}^{\bar{x}y_1}$. Indeed, suppose an x_1y_1-chain contains a $z \in \mathcal{P}^{x_2y_2}$. Then all elements in that chain below z are contained in $\mathcal{P}^{x_1\bar{y}}$ and all elements above z are contained in $\mathcal{P}^{\bar{x}y_1}$. It follows that any x_1y_1-chain which meets $\mathcal{P}^{x_2y_2}$ intersects $(\mathcal{P}^{x_1\bar{y}} \cup \mathcal{P}^{\bar{x}y_1}) \cap \mathcal{P}_{\lfloor n/2 \rfloor}(n)$. However, the fraction of such chains among all x_1y_1-chains is at most

$$\frac{|(\mathcal{P}^{x_1\bar{y}} \cup \mathcal{P}^{\bar{x}y_1}) \cap \mathcal{P}_{\lfloor n/2\rfloor}(n)|}{|\mathcal{P}^{x_1y_1} \cap \mathcal{P}_{\lfloor n/2\rfloor}(n)|} \leq 2\binom{n-2t-h}{\lfloor (n-2t)/2 \rfloor}\Big/\binom{n-2t}{\lfloor (n-2t)/2\rfloor} \overset{(3.1)}{\leq} 2\mathrm{e}^{-h/2}$$

and the lemma follows. $\square$

We are now ready to prove the Lemma 6.10. In what follows, we will make use of the following form of the second moment inequality:

$$\mathbb{P}[X_{\mathcal{S}} = 0] \leq 1 - \frac{\mathbb{E}[X_{\mathcal{S}}]^2}{\mathbb{E}[X_{\mathcal{S}}^2]}. \tag{6.28}$$

A proof of this elementary inequality may be found in [8].

Proof of Lemma 6.10. Let a be as in Proposition 6.9 and set $h = \log 4a$. For each $x \in X$ consider the upward x-branching process. For each such process the probability of survival during the next h steps is at least $1 - \eta - \varepsilon$. We may assume that n is large enough to ensure that the probability that these processes intersect is at most ε. Thus, by the definition of b in Lemma 6.10 and by Proposition 6.9, with probability at least $1-2\varepsilon$ there are a processes which survive until level $m+h$. Pick one descendant $x_i \in \mathcal{P}_{m+h}(n)$ of each, where $1 \leq i \leq a$. Let $X^{'} = \{x_1, \ldots, x_a\}$. Note that we may find $\bar{X} \subset X$, $\bar{X} = \{\bar{x}_1, \ldots, \bar{x}_a\}$, so that there is an $\bar{x}_i x_i$-chain in $\mathcal{P}(n,p)$ for all i. Analogously, with probability $1 - 2\varepsilon$, we may find a set $Y^{'} = \{y_1, \ldots, y_a\}$ with $Y^{'} \subset \mathcal{P}_{n-m-h}(n)$ and which is linked to Y in a similar way. Now assume that n is large enough to ensure that with probability at least $1 - \varepsilon$ these branching processes have the following property: $(x_i \backslash \bar{x}_i) \cap (x_j \backslash \bar{x}_j) = \emptyset$ for all pairs $i \neq j$, and similarly for the y_i. The latter implies that $d_H(x_i, x_j) \geq 2h$ and $d_H(y_i, y_j) \geq 2h$ for $i \neq j$.

Now let $\mathcal{S}_i$ be the set of $x_i y_i$-chains and let $\mathcal{T}_i \subseteq \mathcal{S}_i$ be the set of $x_i y_i$-chains not meeting $\bigcup_{j \neq i} \mathcal{P}^{x_j y_j}$. We let $Y_i = X_{\mathcal{S}_i}$ and $Y_i^{'} = X_{\mathcal{T}_i}$, where the notation is as in Section 3.2. (Thus Y_i denotes the number of $x_i y_i$-chains in $\mathcal{P}(n,p)$ and $Y_i^{'}$ denotes the number of such chains not meeting $\bigcup_{j \neq i} \mathcal{P}^{x_j y_j}$.) We also let $\Delta_i = \Delta(\mathcal{S}_i)$ and $\Delta_i^{'} = \Delta(\mathcal{T}_i)$, using the notation of Section 3.2 again. Note that Proposition 6.6 certainly implies that $\mathbb{E}[Y_i] \leq \mathbb{E}[Y_i]^2$. Thus by the definition of $\Delta_{\mathcal{S}_i}$ and Lemma 6.7 (with $t = m + h$ and $d = 1$) we have

$$\mathbb{E}[Y_i^2] \leq \Delta_i + \mathbb{E}[Y_i] + \mathbb{E}[Y_i]^2 \leq 11\mathbb{E}[Y_i]^2.$$

Now $Y_i^{'} \leq Y_i$ implies that $\mathbb{E}[Y_i^{'2}] \leq \mathbb{E}[Y_i^2]$. But since $2a \exp(-h) \leq 1/2$, Lemma 6.11 shows that $\mathbb{E}[Y_i^{'}] \geq \mathbb{E}[Y_i]/2$. Altogether this implies that

$$\frac{\mathbb{E}[Y_i^{'}]^2}{\mathbb{E}[Y_i^{'2}]} \geq \frac{\mathbb{E}[Y_i^2]}{4\mathbb{E}[Y_i^2]} \geq \frac{1}{44}.$$

Thus applying in turn the independence of the $Y_i^{'}$, the second moment inequality (6.28), and the definition of a in Proposition 6.9, we have (for n sufficiently large)

$$\mathbb{P}\left[\bigcap_{i=1}^{a} Y_i^{'} = 0\right] = \prod_{i=1}^{a} \mathbb{P}\left[Y_i^{'} = 0\right] \leq \left(\frac{43}{44}\right)^a \leq \varepsilon$$

and the lemma follows. □

In the remainder of this section we shall define $(X_t^n)_{t=0}^n$ as the Markov chain of an upward $\emptyset$-process and $(Y_t^n)_{t=0}^n$ as that of a downward $[n]$-process. We need a final preparation for the proof of Theorem 6.4:

Proposition 6.12 *For a given* $c \geq \mathrm{e}$ *define* $\eta = \eta(c)$ *as in* (6.26) *and let* $p = c/n$. *Let* $0 < \varepsilon < 1 - \eta$ *and* $b \in \mathbb{N}$ *be given. Then there exist* m_0, $n_0 \in \mathbb{N}$ *so that for* $m \in \{m_0, m_0 + 1\}$ *and all* $n \geq n_0$, *we have*

$$\mathbb{P}[X_m^n = 0] \geq \eta - \varepsilon, \tag{6.29}$$
$$\mathbb{P}[X_m^n \geq b] \geq 1 - \eta - \varepsilon. \tag{6.30}$$

Proof. Let $(X_m)_{m=0}^{\infty}$ be a Poisson branching process with parameter c. Note that $\mathbb{P}[X_m = 0] \uparrow \eta$. Thus by (6.27) there is an m_0 such that for $m \in \{m_0, m_0+1\}$ we have

$$\begin{aligned} \mathbb{P}[X_m = 0] &\geq \eta - \varepsilon/2, \\ \mathbb{P}[X_m \leq b \mid X_m > 0] &\leq \varepsilon/2. \end{aligned}$$

It is straightforward to show that for fixed m, the random variables X_m^n tend to X_m in distribution as $n \to \infty$. This implies that (6.29) and (6.30) hold for sufficiently large n. □

Proof of Theorem 6.4. Let $\varepsilon > 0$ be given. We now fix an $a = a(c, \varepsilon)$ satisfying the conditions in Proposition 6.9 and let $b = \lceil 2a/(1-\eta-\varepsilon) \rceil$. Then we fix $m_0 = m_0(b, \varepsilon)$ as in Proposition 6.12 and let $m = m_0 + 1$. We set n large enough to satisfy all the requirements below. Consider the branching processes $(X_t^n)_{t=0}^n$ and $(Y_t^n)_{t=0}^n$ defined earlier. We will distinguish several cases according to their (non)-extinction during the first m generations. Note that almost surely $\mathcal{P}(n,p)$ contains neither the bottom element $\emptyset$ nor the top element $[n]$. For the remainder of the proof we will condition on this event. Thus a necessary condition for the length of $\mathcal{P}(n,p)$ to equal $n-2$ is that $X_m^n > 0$ and $Y_m^n > 0$. By Proposition 6.12 this happens with probability $(1-\eta)^2 + \mathcal{O}(\varepsilon)$. But the same proposition also tells us that in this case with probability $1-\varepsilon$ we have sets $X \subseteq \mathcal{P}_m(n,p)$ and $Y \subseteq \mathcal{P}_{n-m}(n,p)$ with $|X| = |Y| = b$ whose elements are the descendants of the respective branching processes. As m and b are fixed, with probability $1-\varepsilon$ each element of X will be contained in every element of Y. Since the event that X and Y can be linked depends only on $\mathcal{P}_{m+1\ldots n-m-1}(n,p)$, Lemma 6.10 then implies that

$$\mathbb{P}[\operatorname{length} \mathcal{P}(n,p) = n-2] = (1-\eta)^2 + \mathcal{O}(\varepsilon).$$

Now consider the case when $X_m^n = 0$ and $Y_m^n > 0$, which by Proposition 6.12 occurs with probability $\eta(1-\eta) + \mathcal{O}(\varepsilon)$. Since we are assuming that $\emptyset$, $[n] \notin \mathcal{P}(n,p)$, in this case the length of $\mathcal{P}(n,p)$ is at most $n-3$. Let k satisfy $(\eta+\varepsilon)^k \leq \varepsilon$ and let $x_i = \{i\}$ for $i = 1, \ldots, k$. For each x_i, consider the upward x_i-branching process. Note that with probability $1-\varepsilon$ these branching processes are disjoint up to level m, i.e., $x_i' \neq x_j'$ whenever $i \neq j$, $x_i \subset x_i'$, $x_j \subset x_j'$, and x_i', $x_j' \in \mathcal{P}_{1\ldots m}(n,p)$. Another application of Proposition 6.12 tells us that with probability $1-\varepsilon$ at least one of the branching processes has (at least) b descendants in $\mathcal{P}_m(n,p)$. This time we pick any $X' \subseteq \mathcal{P}_m(n,p)$ with $|X'| = b$ from the set of descendants of this branching process and we pick any $Y \subseteq \mathcal{P}_{n-m}(n,p)$ as above. Again it is easy to see that with probability close to one, X' and Y satisfy the requirements of Lemma 6.10 and thus in this case $\mathcal{P}(n,p)$ has length exactly $n-3$. By symmetry the same applies if $X_m^n > 0$ and $Y_m^n = 0$ and thus, putting all of the above together, we have

$$\mathbb{P}[\operatorname{length} \mathcal{P}(n,p) = n-3] = 2\eta(1-\eta) + \mathcal{O}(\varepsilon).$$

Finally suppose that $X_m^n = 0$ and $Y_m^n = 0$, which occurs with probability $\eta^2 + \mathcal{O}(\varepsilon)$. As above, with sufficiently high probability one can find suitable $X' \subseteq \mathcal{P}_m(n,p)$ and $Y' \subseteq \mathcal{P}_{n-m}(n,p)$ by considering k branching processes starting in $\mathcal{P}_1(n,p)$ and

$\mathcal{P}_{n-1}(n,p)$, respectively, so that in this case Lemma 6.10 implies that $\mathcal{P}(n,p)$ has length $n-4$. $\square$

We remark that it is easy to strengthen Corollary 6.2 by replacing the condition $pn^r \to 0$ by $pn^r \leq c$ for some c depending only on the constant r. The methods of this section would also enable us to replace the condition $pn^r \to \infty$ by $pn^r \geq C$ for some C depending only on r. However, the analysis is rather cumbersome if r is not an integer, so we have chosen to omit it.

Chapter 7

The width of random subsets of a finite set

7.1 Results

The main result of this chapter is the following general theorem, which gives upper and lower bounds on the width of $\mathcal{P}(n,p)$ for any given p. Note that it contains the main part of Theorem 2.9 as a special case (corresponding to $r = s = 1$). The "moreover" part of Theorem 2.9 follows from the fact that by Chernoff inequalities in Section 3.1, the number of elements of $\mathcal{P}(n,p)$ whose cardinality is $\lfloor n/2 \rfloor$ is almost surely sharply concentrated about its expected value $p\binom{n}{\lfloor n/2 \rfloor}$.

Theorem 7.1 *Given $p = p(n)$, let $r = r(n)$ be an integer function which is large enough to satisfy*

$$p\left(\frac{n}{r}\right)^r / \log n \to \infty \tag{7.1}$$

Then almost surely we have

$$\operatorname{width}\mathcal{P}(n,p) \leq (1+o(1))\, p \sum_{-r/2 \leq j < r/2} \binom{n}{\lfloor n/2 \rfloor + j}.$$

If $p \to 0$, let $s = s(n) \geq 1$ be an integer function which is small enough to satisfy

$$p\left(\frac{n}{s}\right)^{s-1} \to 0. \tag{7.2}$$

If $p \not\to 0$, let $s = 1$. Then almost surely we have

$$\operatorname{width}\mathcal{P}(n,p) \geq (1+o(1))\, p \sum_{-s/2 \leq j < s/2} \binom{n}{\lfloor n/2 \rfloor + j}.$$

The upper bound in Theorem 7.1 will be obtained by a double counting argument applied to chains in $\mathcal{P}(n,p)$, following the elegant proof of Sperner's theorem by Lubell [57, 6]. To make this argument work, we make heavy use of Lemma 6.8, which implies that condition (7.1) essentially guarantees the existence of many chains in $\mathcal{P}(n,p)$ whose cardinalities are close to n/r.

The lower bound is obtained simply by showing that the expected number of relations between the elements contained in

$$A' = \bigcup_{-\overline{s}/2 \le j < \overline{s}/2} \mathcal{P}_{\lfloor n/2 \rfloor + j}(n,p)$$

is small, where $\overline{s} = (1+o(1))s$. (Recall that we write $\mathcal{P}_t(n,p)$ for the set of all those elements of $\mathcal{P}(n,p)$ which have cardinality t and call it the tth *level* of $\mathcal{P}(n,p)$.) This implies that for given p we may always find an antichain A in $\mathcal{P}(n,p)$ which has at least $(1+o(1))|A'|$ elements, by discarding from A' any elements which are contained in another element of A'.

How good are the bounds in Theorem 7.1? Suppose r is a minimal integer function satisfying (7.1) in the sense that r satisfies (7.1) but $r-1$ does not. Similarly suppose s is a maximal integer function satisfying (7.2) in the sense that s satisfies (7.2) but $s+1$ does not. Then it is easily checked that $0 \le r - s \le 1$, and thus the width of $\mathcal{P}(n,p)$ is determined up to a factor of $1+1/r$. If $r \to \infty$, or equivalently, if p decays faster than any polynomial in n, this gives matching (up to first order) upper and lower bounds on the width of $\mathcal{P}(n,p)$:

Corollary 7.2 *Given $p = p(n)$, let $r = r(n)$ be defined by $p = (r/n)^r$. Then almost surely we have*

$$\operatorname{width} \mathcal{P}(n,p) = (1 + \mathcal{O}(1/r) + o(1))\ p \sum_{-r/2 \le j < r/2} \binom{n}{\lfloor n/2 \rfloor + j}.$$

Using quite different methods, Corollary 7.2 was also obtained by Kohayakawa and Kreuter [50] if r is of the form $r = c\sqrt{n}$, for any constant $c > 0$. Observe that this is an especially interesting period in the evolution of $\mathcal{P}(n,p)$: Corollary 7.2 implies that

$$\frac{\operatorname{width} \mathcal{P}(n,p)}{|\mathcal{P}(n,p)|} \to \begin{cases} 0 & \text{if } r/\sqrt{n} \to 0 \\ 1 & \text{if } r/\sqrt{n} \to \infty. \end{cases} \tag{7.3}$$

If $r = o(\sqrt{n})$ however, the matching argument used in [50] provides no further information beyond (7.3) about the width of $\mathcal{P}(n,p)$.

Finally, we point out that it would be very interesting to know whether the condition $pn/\log n \to \infty$ in Theorem 2.9 could be replaced by $pn \to \infty$. The following proposition shows that the latter condition is certainly necessary (similar remarks hold also for the factor $\log n$ in Theorem 7.1).

Proposition 7.3 *Suppose $p = c/n$ where $c > 0$ is a fixed constant. Then*

$$\operatorname{width} \mathcal{P}(n,p) \ge (1 + \mathrm{e}^{-c/2} + o(1))\, p \binom{n}{\lfloor n/2 \rfloor}.$$

Proof. Consider the bipartite graph whose vertex classes are $V_1 = \mathcal{P}_{\lfloor n/2 \rfloor}(n,p)$ and $V_2 = \mathcal{P}_{\lfloor n/2 \rfloor + 1}(n,p)$, with an edge between $v_1 \in V_1$ and $v_2 \in V_2$ if $v_1 \subset v_2$. The probability that a vertex $v_2 \in V_2$ is isolated is equal to $(1-p)^{\lfloor n/2 \rfloor + 1}$. Let Z be the number of isolated vertices in V_2. Then

$$\mathbb{E}[Z] = (1+o(1))\mathrm{e}^{-c/2}\mathbb{E}[|V_2|] = (1+o(1))\mathrm{e}^{-c/2} p \binom{n}{\lfloor n/2 \rfloor}.$$

Using Chebyshev's inequality, it is easily verified that almost surely $Z = (1+o(1))\mathbb{E}[Z]$. The union of V_1 and the isolated vertices in V_2 is thus almost surely an antichain whose size satisfies the requirements. $\square$

The remainder of this chapter is organized as follows: in Section 7.2 we derive properties of "regular" chains in $\mathcal{P}(n,p)$, which will enable us to prove Theorem 7.1 in Section 7.3.

7.2 The chain cover

To prove the upper bound in Theorem 7.1 we need a covering of $\mathcal{P}(n,p)$ with chains which is well behaved in the sense that for most elements $a \in \mathcal{P}(n,p)$, the number of chains meeting a is large and very close to its expected value. It turns out that a cover whose chain elements are equally spaced and sufficiently far apart will do: we will use Lemma 6.8, which gives sufficient conditions on p and the average distance r between chain elements for the desired concentration result to hold. We shall then deduce two corollaries which will be of crucial importance in the proof of the main theorem.

The chain cover of $\mathcal{P}(n)$ is defined as follows. Given integer functions $r = r(n)$ and $j = j(n)$ so that r is $o(n)$ and $j < r$, let

$$k = \begin{cases} n-1 & \text{if } r = 1 \\ \lfloor n/r \rfloor - 2 & \text{if } r > 1. \end{cases} \tag{7.4}$$

We remark that it is straightforward to check that this is consistent with (6.23), setting $m = j$ and $\overline{m} = n - 1 - (\lfloor n/r \rfloor - 1)r - j$. We say that $(x_1 \subset \cdots \subset x_k)$ is a $[j,r]$-chain if $|x_i| = j + ir$ for all i. Lemma 6.8 then immediately implies the following:

Corollary 7.4 *Let $\varepsilon > 0$ be given. Suppose that*

$$p(n)\left(\frac{n}{6r}\right)^r / \log n \to \infty. \tag{7.5}$$

Let $\mathcal{S}$ be the set of $[j,r]$-chains in $\mathcal{P}(n)$ and let $X_{\mathcal{S}}$ denote the number of $[j,r]$-chains in $\mathcal{P}(n,p)$. Then

$$\mathbb{P}\left[X_{\mathcal{S}} \geq (1+\varepsilon)\mathbb{E}[X_{\mathcal{S}}]\right] = o(1/r).$$

Proof. It is easily checked that if p satisfies (7.5), then it also satisfies (6.7) with $d = r \log n$, say (note that $\psi(0) \leq (r+1)!$). A straightforward application of of a weak form of Stirling's formula (i.e. that $n! \geq (n/\mathrm{e})^n$) yields that, for large enough

n, we have

$$\begin{aligned}\mathbb{E}\left[X_{\mathcal{S}}\right] &= \binom{n}{j+r, r, \ldots, r, n-kr-j} p^k \\ &\geq \left(\frac{n}{\mathrm{e}}\right)^{n-kr} \frac{d^k (r+1)^k}{(j+r)_j (n-kr-j)!} \\ &\geq \left(\frac{n}{\mathrm{e}}\right)^{n-kr} \frac{d}{(j+r)_j (n-kr-j)!} \geq d. \end{aligned} \tag{7.6}$$

The result now follows immediately from Chebyshev's inequality using the fact that $\operatorname{Var}\left[X_{\mathcal{S}}\right] \leq \Delta(\mathcal{S}) + \mathbb{E}\left[X_{\mathcal{S}}\right]$. For a proof of the latter inequality see e.g. [6]. $\square$

To obtain our second corollary of Lemma 6.8, we need the following version of Janson's inequality (for a proof see e.g. [39]). Consider the setting of Section 3.2. Then given $\tau > 0$, we have

$$\mathbb{P}\left[X_{\mathcal{S}} \leq (1-\tau)\mathbb{E}[X_{\mathcal{S}}]\right] \leq \exp\left(-\frac{\tau^2 \mathbb{E}\left[X_{\mathcal{S}}\right]^2}{2(\mathbb{E}\left[X_{\mathcal{S}}\right] + \Delta(\mathcal{S}))}\right). \tag{7.7}$$

Let X_a denote the number of $[j, r]$-chains (for any $j < r$) which meet a and which are contained in $\mathcal{P}(n, p) \cup \{a\}$. Suppressing the dependence on r, let $\mathcal{P}^j(n)$ be the set of elements in $\mathcal{P}(n)$ which intersect some $[j, r]$-chain. In other words, $\mathcal{P}^j(n)$ consists of every rth level of $\mathcal{P}(n)$, starting with the $(j+r)$th.

Corollary 7.5 *Let $\varepsilon > 0$ be given and fix $a \in \mathcal{P}(n)$ with $n/3 \leq |a| \leq 2n/3$. Then if p satisfies (7.5) and n is sufficiently large, we have*

$$\mathbb{P}\left[X_a \leq (1-\varepsilon)\mathbb{E}[X_a]\right] \leq 2n^{-2}.$$

Proof. This time we apply Lemma 6.8 not to the whole of $\mathcal{P}(n)$ but to $\mathcal{P}(|a|)$ and $\mathcal{P}(n-|a|)$ separately, with

$$d(n) = \frac{80}{\varepsilon^2} \log n. \tag{7.8}$$

Consider $\mathcal{P}(|a|)$ first: fix $j < r$ so that $a \in \mathcal{P}^j(n)$. Then let $m' = 0$, $\overline{m}' = j-1$, and $k' = \frac{|a|-j}{r} - 1$. Let $\mathcal{S}_1$ be the set of all chains $(x_1 \supset x_2 \supset \cdots \supset x_{k'})$ in $\mathcal{P}(|a|)$ with $|x_i| = |a| - ir$ for all i. Let $X_{\mathcal{S}_1}$ be the number of such chains in $\mathcal{P}(n, p)$. Similarly to (7.6) in the proof of the previous corollary, for large enough n we have

$$\mathbb{E}\left[X_{\mathcal{S}_1}\right] = \binom{|a|}{r, \ldots, r, r+j} p^{k'} \geq d(n).$$

It is also easily checked that if $p(n)$ satisfies (7.5), then $p(|a|)$ satisfies (6.7) with n replaced by $|a|$; k, m, and $\overline{m}$ replaced by k', m', and $\overline{m}'$ respectively; but keeping $r(n)$ as before and $d(n)$ as in (7.8). Thus the conclusion (6.25) of Lemma 6.8 holds for the set $\mathcal{S}_1$ and $d(n)$ as in (7.8). Now apply Janson's inequality (7.7) to obtain

$$\mathbb{P}\left[X_{\mathcal{S}_1} \leq (1-\tau)\mathbb{E}[X_{\mathcal{S}_1}]\right] \leq \mathrm{e}^{-\tau^2 d/20} = n^{-4\tau^2/\varepsilon^2}.$$

For $\mathcal{P}(n-|a|)$ we define $\mathcal{S}_2$ similarly and apply Janson's inequality in the same way. Note that $X_a = X_{\mathcal{S}_1} X_{\mathcal{S}_2}$. By independence we then have

$$\begin{aligned}\mathbb{P}\left[X_a \leq (1-\varepsilon)\mathbb{E}[X_a]\right] &= \mathbb{P}\left[X_{\mathcal{S}_1} X_{\mathcal{S}_2} \leq (1-\varepsilon)\mathbb{E}[X_{\mathcal{S}_1}]\mathbb{E}[X_{\mathcal{S}_2}]\right] \\ &\leq \sum_{i=1,2} \mathbb{P}\left[X_{\mathcal{S}_i} \leq \sqrt{1-\varepsilon}\,\mathbb{E}[X_{\mathcal{S}_i}]\right]\end{aligned}$$

Since $\sqrt{1-\varepsilon} \leq 1-\varepsilon/2$, the proof now follows by setting $\tau = \varepsilon/2$. □

Finally, let $\mathcal{P}^j_{max}(n)$ be a level of $\mathcal{P}^j(n)$ which is closest to the middle and thus contains the most elements.

Proposition 7.6 *Suppose that for some j, we have $a \in \mathcal{P}^j(n)$ and furthermore that $a' \in \mathcal{P}^j_{max}(n)$. Then $\mathbb{E}[X_a] \geq \mathbb{E}[X_{a'}]$.*

Proof.

$$\mathbb{E}[X_a] = \binom{|a|}{j+r, r, \ldots, r}\binom{n-|a|}{r, \ldots, r, n-kr-j} p^{k-1},$$

and the result follows by the log-convexity of factorials. □

7.3 Proof of Theorem 7.1

With the results of the previous section at hand, the proof of the upper bound of Theorem 7.1 proceeds as follows: first we show that we may assume that $\mathcal{P}(n,p)$ does not contain many more "regular" chains than we would expect. Now suppose we have a large antichain A in $\mathcal{P}(n,p)$. Since A intersects any chain at most once, this means that A (and thus $\mathcal{P}(n,p)$) contains many "bad" elements in the sense that these elements lie in fewer "regular" chains than one would expect. But again this is extremely unlikely.

Proof of Theorem 7.1. We first prove the lower bound. As remarked at the beginning of this chapter, the case $s=1$ follows immediately from the Chernoff inequalities in Section 3.1. So suppose that p and s satisfy (7.2), where $p \to 0$. Since the assertion implies that almost surely width $\mathcal{P}(n,p) \geq (1+o(1))|\mathcal{P}(n,p)|$ whenever $s/\sqrt{n} \to \infty$, we may also assume that $s = o(n)$. Thus one may define $\overline{s} \leq s$ so that $\overline{s} = (1+o(1))s$ and

$$p\,\overline{s}^2\left(\frac{en}{\overline{s}-1}\right)^{\overline{s}-1} \to 0. \tag{7.9}$$

Now consider the expected number of edges $\mathbb{E}[|E(G)|]$ in the $\overline{s}$-partite graph G whose vertex classes are $V_j = \mathcal{P}_{\lfloor n/2\rfloor + j}(n,p)$ for all j satisfying $\overline{s}/2 \leq j < \overline{s}/2$, with an edge between $v_i \in V_i$ and $v_j \in V_j$ for $i \neq j$ if $v_i \subset v_j$ or $v_j \subset v_i$. Then Stirling's formula and our assumption (7.9) on $\overline{s}$ imply that

$$\begin{aligned}
\mathbb{E}\left[|E(G)|\right] &= \sum_{-\overline{s}/2 \le \ell < \ell' < \overline{s}/2} \binom{n}{\lfloor n/2 \rfloor + l} \binom{\lfloor n/2 \rfloor - \ell}{\ell' - \ell} p^2 \\
&\le \sum_{-\overline{s}/2 \le \ell < \ell' < \overline{s}/2} \binom{n}{\lfloor n/2 \rfloor} \left(\frac{\mathrm{e} n}{\ell' - \ell} \right)^{\ell' - \ell} p^2 \\
&\le p \binom{n}{\lfloor n/2 \rfloor} p \, \overline{s}^2 \left(\frac{\mathrm{e} n}{\overline{s} - 1} \right)^{\overline{s} - 1} \\
&= o\left(p \binom{n}{\lfloor n/2 \rfloor} \right).
\end{aligned}$$

This proves the lower bound since it implies that an antichain of the required size may be obtained by picking any maximal independent set in G.

We now prove the upper bound. Suppose first that p and r satisfy (7.5), so that we may apply Corollaries 7.4 and 7.5. Furthermore note that we may assume that $r = \mathcal{O}(\sqrt{n})$. Indeed, if this is not the case, then a simple application of Stirling's formula tells us that in that case Theorem 7.1 just states that the width of $\mathcal{P}(n,p)$ is almost surely at most $(1 + o(1))|\mathcal{P}(n,p)|$, which is of course trivial. Thus, defining $\mathcal{P}^j_{max}(n,p) = \mathcal{P}^j_{max}(n) \cap \mathcal{P}(n,p)$, another application of Stirling's formula implies that we may assume that

$$\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right] = p\,|\mathcal{P}^j_{max}(n)| \ge p\,2^n/n, \tag{7.10}$$

which will be convenient later on. For given $\varepsilon > 0$ define the event

$$\mathcal{A}_\varepsilon = \left\{ \exists \text{ an antichain } A \subset \mathcal{P}(n,p) \text{ with } |A| \ge (1+\varepsilon) p \sum_{-r/2 \le j < r/2} \binom{n}{\lfloor n/2 \rfloor + j} \right\}.$$

For all j with $0 \le j < r$ define the event

$$\mathcal{A}^j_\varepsilon = \left\{ \mathcal{P}(n,p) \text{ contains an antichain } A \text{ with } |A^j| \ge (1+\varepsilon)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right] \right\},$$

where A^j is defined by $A^j = A \cap \mathcal{P}^j(n)$. Given ε_1, $\nu > 0$, our goal is to show that $\mathbb{P}\left[\mathcal{A}_{\varepsilon_1}\right] \le 3\nu$ for large enough n. Clearly, if $\mathcal{A}_{\varepsilon_1}$ holds, then we must have $\mathcal{A}^j_{\varepsilon_1}$ for some j. Thus,

$$\mathbb{P}\left[\mathcal{A}_{\varepsilon_1}\right] \le \mathbb{P}\left[\bigcup_j \mathcal{A}^j_{\varepsilon_1} \right] \le \sum_j \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\right]. \tag{7.11}$$

Fix some j and let X be equal to the number of $[j,r]$-chains contained in $\mathcal{P}(n,p)$. Set $\varepsilon_2 = \varepsilon_1/5$ and define the event

$$\mathcal{B}_{\varepsilon_2} = \{X < (1+\varepsilon_2)\mathbb{E}[X]\}.$$

Then for large enough n, Corollary 7.4 implies that

$$\begin{aligned}\mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\right] &= \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\cap\mathcal{B}_{\varepsilon_2}\right] + \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\cap\mathcal{B}^c_{\varepsilon_2}\right]\\ &\le \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\cap\mathcal{B}_{\varepsilon_2}\right] + \mathbb{P}\left[\mathcal{B}^c_{\varepsilon_2}\right]\\ &\le \mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\cap\mathcal{B}_{\varepsilon_2}\right] + \nu/r.\end{aligned}$$

Now define the event

$$\begin{aligned}\mathcal{C}_{\varepsilon_1\varepsilon_2} = \Big\{&\mathcal{P}(n,p) \text{ contains an antichain } A \text{ with } |A^j| \ge (1+\varepsilon_1)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\\ &\text{and } \sum_{a\in A^j} X_a < (1+\varepsilon_2)\mathbb{E}[X]\Big\}.\end{aligned}$$

Note that $\sum_{a\in A^j} X_a \le X$ implies that

$$\mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\cap\mathcal{B}_{\varepsilon_2}\right] \le \mathbb{P}\left[\mathcal{C}_{\varepsilon_1\varepsilon_2}\right].$$

For a given $\varepsilon > 0$ we say that $a \in \mathcal{P}(n)$ is *ε-bad* if $X_a \le (1-\varepsilon)\mathbb{E}[X_a]$. Setting $\mathcal{P}^j(n,p) = \mathcal{P}^j(n)\cap\mathcal{P}(n,p)$, we then let Y_ε be the number of ε-bad elements in $\mathcal{P}^j(n,p)$. Define

$$\mathcal{D}_{\varepsilon_2} = \left\{Y_{\varepsilon_2} \ge \varepsilon_2\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\right\}.$$

It is in the proof of the following claim that we will apply the double counting argument that Lubell used in his proof of Sperner's theorem.

Claim. We have $\mathcal{C}_{\varepsilon_1\varepsilon_2} \subset \mathcal{D}_{\varepsilon_2}$, and hence $\mathbb{P}\left[\mathcal{C}_{\varepsilon_1\varepsilon_2}\right] \le \mathbb{P}\left[\mathcal{D}_{\varepsilon_2}\right]$.

Proof. Suppose to the contrary that $\mathcal{C}_{\varepsilon_1\varepsilon_2}\cap\mathcal{D}^c_{\varepsilon_2}$ holds and consider an antichain A guaranteed by $\mathcal{C}_{\varepsilon_1\varepsilon_2}$. Since we assumed that $\mathcal{D}_{\varepsilon_2}$ is false, it follows that

$$\left|\{a \in A^j \colon a \text{ is not } \varepsilon_2\text{-bad}\}\right| \ge (1+\varepsilon_1-\varepsilon_2)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right].$$

The crucial point now is that, by double counting chains in $\mathcal{P}^j(n,p)$,

$$X = \sum_{a\in\mathcal{P}^j_{max}(n,p)} X_a,$$

and furthermore X_a is independent of the event that $a \in \mathcal{P}(n,p)$. Thus, by linearity of expectation, we have (for any $a' \in \mathcal{P}^j_{max}(n)$)

$$\mathbb{E}[X] = \mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\,\mathbb{E}\left[X_{a'}\right].$$

Combining all this with Proposition 7.6, we have that (again for any $a' \in \mathcal{P}^j_{max}(n)$)

$$\begin{aligned}\sum_{a\in A^j} X_a &\ge (1+\varepsilon_1-\varepsilon_2)\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]\,(1-\varepsilon_2)\mathbb{E}\left[X_{a'}\right]\\ &= (1+\varepsilon_1-\varepsilon_2)(1-\varepsilon_2)\mathbb{E}[X]\\ &\ge (1+2\varepsilon_2)\mathbb{E}[X],\end{aligned}$$

since $\varepsilon_2 = \varepsilon_1/5$. But this contradicts the definition of $\mathcal{C}_{\varepsilon_1\varepsilon_2}$. □

Thus it remains to show that $\mathbb{P}\left[\mathcal{D}_{\varepsilon_2}\right]$ is small. But Corollary 7.5 tells us that if $n/3 \le |a| \le 2n/3$, then

$$\mathbb{P}\left[a \in \mathcal{P}(n) \text{ is } \varepsilon_2\text{-bad}\right] \le 2n^{-2} \le \frac{\nu\varepsilon_2}{r} \frac{\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right]}{p\,2^n}. \tag{7.12}$$

The last inequality follows from our assumption that $r = \mathcal{O}(\sqrt{n})$, which gives the lower bound (7.10) on $\mathbb{E}[|\mathcal{P}^j_{max}(n,p)|]$. Inequality (7.12) immediately yields an upper bound on the expected number of bad elements in $\mathcal{P}^j(n,p)$. Indeed, writing $\sum'_a$ for the sum over $a \in \mathcal{P}(n)$ with $n/3 \le |a| \le 2n/3$ and writing $\sum''_a$ for the sum over $a \in \mathcal{P}(n)$ with $|a| < n/3$ or $|a| > 2n/3$, we have

$$\begin{aligned}\mathbb{E}\left[Y_{\varepsilon_2}\right] &= \sum_{a \in \mathcal{P}^j(n)} p\,\mathbb{P}\left[a \text{ is } \varepsilon_2\text{-bad}\right] \le \sum_{a \in \mathcal{P}(n)} p\,\mathbb{P}\left[a \text{ is } \varepsilon_2\text{-bad}\right] \\ &\le {\sum}'_a p\,\mathbb{P}\left[a \text{ is } \varepsilon_2\text{-bad}\right] + {\sum}''_a p \\ &\le \frac{2\nu\varepsilon_2}{r}\,\mathbb{E}\left[|\mathcal{P}^j_{max}(n,p)|\right].\end{aligned}$$

In the last line we made use of the fact that the bottom and top third of $\mathcal{P}(n)$ contain (very crudely) at most $2^n/n^4$ elements. Markov's inequality now implies that $\mathbb{P}[\mathcal{D}_{\varepsilon_2}] \le \mathbb{P}[Y_{\varepsilon_2} \ge \frac{r}{2\nu}\mathbb{E}[Y_{\varepsilon_2}]] \le 2\nu/r$, and thus

$$\mathbb{P}\left[\mathcal{A}^j_{\varepsilon_1}\right] \le \mathbb{P}\left[\mathcal{C}_{\varepsilon_1\varepsilon_2}\right] + \nu/r \le \mathbb{P}\left[\mathcal{D}_{\varepsilon_2}\right] + \nu/r \le 3\nu/r,$$

as required. By (7.11), this completes the proof of Theorem 7.1 except for the fact that for the proof we assumed that p and r satisfy (7.5) instead of the original condition (7.1) in the statement of Theorem 7.1. This can be remedied immediately however, by noting that if p and r satisfy (7.1), then one may find an integer function $r' \ge r$ with $r' = (1 + o(1))r$ so that p and r' satisfy (7.5). □

Note that the methods of Section 7.2 show that if $pn \to \infty$, then almost surely the fraction of elements in $\mathcal{P}(n,p)$ that are ε-bad tends to zero for any $\varepsilon > 0$. However, for our argument to work, we need the number of ε-bad elements in $\mathcal{P}(n,p)$ to be much less than the size of the middle level of $\mathcal{P}(n,p)$, which is the reason why we need the additional $\log n$ factor.

Finally, we remark that by weighting the elements $a \in \mathcal{P}(n)$ according to the number of $[j,r]$-chains meeting a, the above proof may easily be modified to yield probabilistic versions of the LYM-inequality (see [6]). The following result is the analogue of Theorem 2.9.

Theorem 7.7 *Let $\varepsilon > 0$ be given. Suppose that $pn/\log n \to \infty$ and that $A \subseteq \mathcal{P}(n,p)$ is an antichain with $\varepsilon n \le |a| \le (1-\varepsilon)n$ for all $a \in A$. Then almost surely*

$$\sum_{a \in A} \frac{1}{\binom{n}{|a|}} \le (1 + o(1))\,p.$$

Note that it is necessary to impose some condition on $|a|$, since for instance if p is not much larger than the above threshold, the proof of Proposition 7.3 shows that almost surely one may pick an antichain consisting of all elements in the first level and almost all of the elements in the second level of $\mathcal{P}(n,p)$, and whose weight is $(2+o(1))p$.

Bibliography

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29:354–360, 1980.

[2] R. Albert, A.L. Barabási, and H. Jeong. Diameter of the world wide web. *Nature*, 401:130–131, 1999.

[3] R. Albert, A.L. Barabási, and H. Jeong. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.

[4] D. Aldous. A random tree model associated with random graphs. *Random Struct. Algorithms*, 1:383–402, 1990.

[5] N. Alon, J.H. Kim, and J.H. Spencer. Nearly perfect matchings in regular simple hypergraphs. *Israel J. Math.*, 100:171–187, 1997.

[6] N. Alon and J.H. Spencer. *The Probabilistic Method.* Wiley-Interscience, New York, 1992.

[7] B. Bollobás. *Extremal Graph Theory.* Academic Press, London, 1978.

[8] B. Bollobás. *Random Graphs.* Academic Press, London, 1985.

[9] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8:49–56, 1988.

[10] B. Bollobás. Extremal graph theory. In R.L. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, volume 2, pages 1231–1292. Elsevier, 1995.

[11] B. Bollobás. *Modern Graph Theory.* Springer, New York, 1998.

[12] B. Bollobás and G. Brightwell. Random high-dimensional orders. *Adv. Appl. Prob.*, 27:161–184, 1995.

[13] B. Bollobás, Y. Kohayakawa, and T. Łuczak. On the diameter and radius of random subgraphs of the cube. *Random Struct. Algorithms*, 5:627–648, 1994.

[14] B. Bollobás, Y. Kohayakawa, and T. Łuczak. Connectivity properties of random subgraphs of the cube. *Random Struct. Algorithms*, 6:221–230, 1995.

[15] B. Bollobás and O.M. Riordan. Constrained graph processes. *Electronic Journal of Combinatorics*, 7:R18, 2000.

[16] B. Bollobás and A. Thomason. Threshold functions. *Combinatorica*, 7:35–38, 1987.

[17] G. Brightwell. Models of random partial orders. In K. Walker, editor, *Surveys in Combinatorics 1993*, pages 53–83. Cambridge University Press, 1993.

[18] F. Chung, Z. Füredi, R.L. Graham, and P.D. Seymour. On induced subgraphs of the cube. *J. Combin. Theory Ser. A*, 49:180–187, 1988.

[19] F. Chung and R.L. Graham. *Erdős on Graphs: His Legacy of Unsolved Problems.* A. K. Peters, Wellesley, 1998.

[20] R. Diestel. *Graph Theory.* Springer, New York, 1997.

[21] M.E. Dyer and A.M. Frieze. The solution of some random NP-hard problems in polynomial expected time. *J. Algorithms*, 10:451–489, 1989.

[22] K. Engel. *Sperner Theory.* Cambridge University Press, 1997.

[23] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[24] P. Erdős. Graph theory and probability. *Canad. J. Mathematics*, 11:34–38, 1959.

[25] P. Erdős, S. Janson, T. Łuczak, and J.H. Spencer. A note on triangle-free graphs. In D. Aldous and R. Pemantle, editors, *Random discrete structures*, volume 76 of *IMA*. Springer, 1996.

[26] P. Erdős, D.J. Kleitman, and B.L. Rothschild. Asymptotic enumeration of K_n-free graphs. In *International Colloquium on Combinatorial theory, Atti dei Convegni Lincei 17*, volume 2, pages 19–27, Rome, 1976.

[27] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kut. Int. Közl.*, 5:17–61, 1960.

[28] P. Erdős, S. Suen, and P. Winkler. On the size of a random maximal graph. *Random Struct. Algorithms*, 6:309–318, 1995.

[29] P. Erdős and P. Tetali. Representation of integers as the sum of k terms. *Random Struct. Algorithms*, 1:245–261, 1990.

[30] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Proc. Eleventh Ann. IEEE Conf. on Comp. Complexity*, pages 278–287, 1996.

[31] J.A. Fill and R. Pemantle. Percolation, first-passage percolation and covering times for Richardson's model of the n-cube. *Annals Appl. Probab.*, 3:593–629, 1993.

[32] A. Frieze. On the value of a minimum spanning tree problem. *Discrete App. Math.*, 10:47–56, 1985.

[33] A. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Struct. Algorithms*, 10:5–42, 1997.

[34] D.A. Grable. On random greedy triangle packing. *Electronic J. Combin.*, 4, 1997.

[35] D.A. Grable. More-than-nearly-perfect packings and partial designs. *Combinatorica*, 19:221–239, 1999.

[36] G. Grimmett. *Percolation.* Springer, New York, 1992.

[37] G. Grimmett and C. McDiarmid. On colouring random graphs. *Math. Proc. Cam. Phil. Soc.*, 77:313–324., 1975.

[38] T.E. Harris. *The Theory of Branching Processes.* Springer, Berlin, 1963.

[39] S. Janson. Poisson approximation for large deviations. *Random Struct. Algorithms*, 1:221–229, 1990.

[40] S. Janson, T. Łuczak, and A. Ruciński. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In M. Karoński, J. Jaworski, and A. Ruciński, editors, *Random Graphs '87*, pages 73–87. John Wiley & Sons, 1987.

[41] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs.* Wiley-Interscience, 2000.

[42] T.R. Jensen and B. Toft. *Graph Coloring Problems.* Wiley-Interscience, New York, 1995.

[43] J. Kahn. Asymptotically good list-colorings. *J. Combin. Theory Ser. A*, 73:1–59, 1996.

[44] M. Karoński. Random graphs. In R.L. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, volume 1, pages 351–380. Elsevier, 1995.

[45] M. Karoński and A. Ruciński. The origins of the theory of random graphs. In R.L. Graham and J. Nešetřil, editors, *The Mathematics of Paul Erdős*, volume 1, pages 311–336. Springer, New York, 1997.

[46] R.M. Karp. The probabilistic analysis of combinatorial optimization algorithms. In *Proc. Int. Congress of Mathematicians 1984*, pages 1601–1609, Warszawa, 1989.

[47] J.H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Struct. Algorithms*, 7:173–207, 1995.

[48] J. Kleinberg. The small-world phenomenon: an algorithmic perspective. In *Proc. 32nd ACM STOC*, 2000.

[49] D.J. Kleitman and B.L. Rothschild. Asymptotic enumeration of partial orders on a finite set. *Trans. Amer. Math. Soc.*, 205:205–220, 1975.

[50] Y. Kohayakawa and B. Kreuter. The width of random subsets of Boolean lattices. *submitted*, 1998.

[51] Y. Kohayakawa, B. Kreuter, and D. Osthus. The length of random subsets of Boolean lattices. *Random Struct. Algorithms*, to appear.

[52] Y. Kohayakawa, T. Łuczak, and V. Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arithmetica*, 75:133–163, 1996.

[53] B. Kreuter. *Probabilistic versions of Ramsey's and Turán's theorems.* PhD thesis, Institut für Informatik, Humboldt-Universität Berlin, 1997.

[54] B. Kreuter. Small sublattices in random subsets of Boolean lattices. *Random Struct. Algorithms*, 13:383–407., 1998.

[55] M. Krivelevich. Bounding Ramsey numbers through large deviation inequalities. *Random Struct. Algorithms*, 7:145–155, 1995.

[56] E. Lamken and B.L. Rothschild. The numbers of odd-cycle-free graphs. In A. Hajnal, L. Lovász, and V.T. Sós, editors, *Finite and infinite sets*, volume 37 of *Colloq. Math. Soc. János Bolyai*, pages 547–553. North-Holland, Amsterdam, 1984.

[57] D. Lubell. A short proof of Sperner's lemma. *J. Combin. Theory*, 1:299, 1966.

[58] T. Łuczak. The chromatic number of random graphs. *Combinatorica*, 11:45–54, 1991.

[59] T. Łuczak. On triangle-free random graphs. *Random Struct. Algorithms*, to appear.

[60] W. Mantel. Problem 28. *Wiskundige opgaven*, 10:60–61, 1906.

[61] D. Osthus. Notes on the "small world" phenomenon. *manuscript*, 1999.

[62] D. Osthus. Maximum antichains in random subsets of a finite set. *J. Combin. Theory Ser. A*, to appear.

[63] D. Osthus, H.J. Prömel, and A. Taraz. On the evolution of triangle-free graphs. *submitted*, 1999.

[64] D. Osthus and A. Taraz. Random maximal H-free graphs. *Random Struct. Algorithms*, to appear.

[65] N. Pippenger and J. Spencer. Asymptotic behaviour of the chromatic index for hypergraphs. *J. Combin. Theory Ser. A*, 51:24–42, 1989.

[66] H.J. Prömel and A. Steger. The asymptotic number of graphs not containing a fixed color-critical subgraph. *Combinatorica*, 12:463–473, 1992.

[67] H.J. Prömel and A. Steger. Coloring clique free graphs in linear time. *Random Struct. Algorithms*, 3:375–402, 1992.

[68] H.J. Prömel and A. Steger. Random l-colourable graphs. *Random Struct. Algorithms*, 6:21–37, 1995.

[69] H.J. Prömel and A. Steger. Counting H-free graphs. *Discr. Math.*, 154:311–315, 1996.

[70] H.J. Prömel and A. Steger. On the asymptotic structure of sparse triangle free graphs. *Journal of Graph Theory*, 21:137–151, 1996.

[71] H.J. Prömel, A. Steger, and A. Taraz. Phase transitions in the evolution of partial orders. *submitted*, 1999.

[72] H.J. Prömel, A. Steger, and A. Taraz. Asymptotic enumeration, global structure and constrained evolution. *Discr. Math.*, to appear.

[73] H.J. Prömel, A. Steger, and A. Taraz. Counting partial orders with a fixed number of comparable pairs. *Comb. Probab. Comput.*, to appear.

[74] A. Rényi. On random generating elements of a finite Boolean algebra. *Acta Sci. Math. Szeged*, 22:75–81, 1961.

[75] A. Rényi. On random subsets of a finite set. *Mathematica (Cluj)*, 3:355–362, 1961.

[76] A. Rényi. On statistical laws of accumulation of information. *Mat. Kut. Int. Közl.*, 12:15–33, 1962.

[77] V. Rödl. On a packing and covering problem. *Europ. J. Combin.*, 5:69–78, 1985.

[78] V. Rödl and L. Thoma. Asymptotic packing and the random greedy algorithm. *Random Struct. Algorithms*, 8:161–177, 1996.

[79] A. Ruciński and N.C. Wormald. Random graph processes with degree restrictions. *Combin. Probab. Comput.*, 1:169–180, 1992.

[80] M. Simonovits. Paul Erdős' influence on extremal graph theory. In R.L. Graham and J. Nešetřil, editors, *The Mathematics of Paul Erdős*, volume 2, pages 148–192. Springer, New York, 1997.

[81] J.H. Spencer. Counting extensions. *J. Combin. Theory Ser. A*, 53:247–255, 1990.

[82] J.H. Spencer. Threshold functions for extension statements. *J. Combin. Theory Ser. A*, 53:286–305, 1990.

[83] J.H. Spencer. Asymptotic packing via a branching process. *Random Struct. Algorithms*, 7:167–172, 1995.

[84] J.H. Spencer. Maximal triangle-free graphs and Ramsey $R(3,t)$. *unpublished manuscript*, 1995.

[85] E. Sperner. Ein Satz über Untermengen einer endlichen Menge. *Math. Z.*, 27:544–548, 1928.

[86] A. Steger. unpublished manuscript, 1999.

[87] A. Steger and N.C. Wormald. Generating random regular graphs quickly. *Comb. Probab. Comput.*, 8:377–396, 1999.

[88] S. Strogatz and D. Watts. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[89] P. Turán. On an extremal problem in graph theory. *Mat. Fitz. Lapok*, 48:436–452, 1941.

[90] J.S. Turner. Almost all k-colourable graphs are easy to color. *J. Algorithms*, 9:63–82, 1988.

[91] L.G. Valiant. *Circuits of the Mind.* Oxford University Press, 1994.

[92] V.H. Vu. On some simple degree conditions that guarantee the upper bound on the chromatic (choice) number of random graphs. *Journal of Graph Theory*, 31:201–226, 1999.

[93] K. Weber. The length of random Boolean functions. *J. Information Processing and Cybernetics*, 12:659–668, 1982.

[94] K. Weber. Subcube coverings of random graphs in the n-cube. In M. Karoński and A. Ruciński, editors, *Random Graphs '83*, pages 319–336. North-Holland, 1985.

[95] N.C. Wormald. The perturbation method and triangle-free random graphs. *Random Struct. Algorithms*, 9:253–270, 1996.

[96] A.A. Zykov. On some properties of linear complexes (in Russian). *Mat. Sbornik N.S.*, 24:163–188, 1949. English translation in: *Amer. Math. Soc. Transl.*, volume 79, 1972.