Random Graph Processes with Dependencies

Lutz Peter Warnke
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy in Mathematics

June 2012

Abstract

Random graph processes are basic mathematical models for large-scale networks evolving over time. Their systematic study was pioneered by Erdős and Rényi around 1960, and one key feature of many 'classical' models is that the edges appear independently. While this makes them amenable to a rigorous analysis, it is desirable, both mathematically and in terms of applications, to understand more complicated situations. In this thesis the main goal is to improve our rigorous understanding of evolving random graphs with significant dependencies.

The first model we consider is known as an *Achlioptas process*: in each step two random edges are chosen, and using a given rule only one of them is selected and added to the evolving graph. Since 2000 a large class of 'complex' rules has eluded a rigorous analysis, and it was widely believed that these could give rise to a striking and unusual phenomenon. Making this explicit, Achlioptas, D'Souza and Spencer conjectured in *Science* that one such rule yields a very abrupt (discontinuous) percolation phase transition. We disprove this, showing that the transition is in fact continuous for *all* Achlioptas process. In addition, we give the first rigorous analysis of the more 'complex' rules, proving that certain key statistics are tightly concentrated (i) in the subcritical evolution, and (ii) also later on if an associated system of differential equations has a unique solution.

The second model we study is the *H-free process*, where random edges are added subject to the constraint that they do not complete a copy of some fixed graph H. The most important open question for such 'constrained' processes is due to Erdős, Suen and Winkler: in 1995 they asked what the typical final number of edges is. While Osthus and Taraz answered this in 2000 up to logarithmic factors for a large class of graphs H, more precise bounds are only known for a few special graphs. We close this gap for the cases where a cycle of fixed length is forbidden, determining the final number of edges up to constants. Our result not only establishes several conjectures, it is also the first which answers the more than 15-year old question of Erdős et. al. for a *class* of forbidden graphs H.

Acknowledgements

First of all, I would like to thank my advisor Oliver Riordan: for the continuous encouragement and support, for allowing me to pursue my research interests, for always insisting on simplifying proofs, for sharing mathematical intuition, and for giving me the opportunity to visit Memphis.

I am grateful to Béla Bollobás and Colin McDiarmid, who both kindly agreed to examine my thesis. In addition, I am indebted to Angelika Steger and Emo Welzl, who greatly shaped my interest in discrete mathematics and probabilistic combinatorics during my undergraduate studies in Zürich. I would also like to thank Svante Janson for writing numerous reference letters.

Finally, I wish to thank my family, friends and colleagues for their emotional and moral support. I also gratefully acknowledge the financial support of EPSRC and University of Oxford in form of a Research Studentship and Scatcherd European Scholarship, respectively.

Lutz Warnke
Oxford, 20.06.2012

Contents

Chapter 1

Introduction

Random graph processes are basic mathematical models for large-scale networks that evolve over time. Their systematic study was pioneered by Erdős and Rényi around 1960, who (implicitly) considered the 'classical' *random graph process* (later formalized by Bollobás). It starts with an empty graph on n vertices, and then, in each step, adds a new random edge to the graph. Historically, this and related random models were first used to solve problems in discrete mathematics, the main idea of the so-called *probabilistic method* being that deterministic results can be proved by probabilistic reasoning. However, it soon became clear that random graphs are not only important tools in combinatorics, but also interesting for their own sake. Their analysis combines elements from graph theory and probability theory in an intriguing way, and nowadays many of their evolving properties are very well understood. Here we restrict ourselves to two highlights in the theory of random graphs, briefly giving their flavour: (i) phase transitions and (ii) concentration of measure phenomena.

The concept of *phase transitions* is ubiquitous in physics, and the perhaps most prominent example is the transition from water into ice or steam. The basic idea is that a slight change of a key parameter (temperature) can completely change or transform the structural properties (state of matter). One of the major discoveries of Erdős and Rényi was that the evolution of the random graph process also exhibits phase transitions: up to a certain point it is very unlikely to have a certain property, whereas shortly afterwards it is extremely likely to have it. For example, for fixed $k \geqslant 2$, in the random graph process connected components of size k start appearing around $n^{1-1/(k-1)}$ steps; so the largest component has bounded size during the first $n^{1-\varepsilon}$ steps. Perhaps the most influential result of Erdős and Rényi concerns the discovery of the *percolation phase transition*: as the ratio of the number of edges to vertices increases past a certain critical point, the global structure changes radically, from only small components to a single macroscopic ('giant') component plus small ones. In other words, the size of the largest component exhibits a phase transition. One key difference to the 'classical' percolation theory of mathematical physics is that we are not studying infinite lattices but (very large) finite graphs on n vertices, where we then let $n \to \infty$. Despite this difference both areas of research have much in common, and one perhaps uniting basic theme is related to the following question:

Where is the phase transition located, and how 'fast' does the transition happen?

We will return to this question several times in this thesis. The (simplified) answer for the Erdős and Rényi process is that the percolation phase transition happens around $n/2$ steps, and that the transition is 'second order', i.e., that the size of the largest component grows in a 'continuous way'. Of course, this basic answer is not the end of the story: the finer details and dynamics of the Erdős–Rényi transition are extremely well understood (see e.g. [21, 24, 42, 43] and the many references therein). It serves as a natural reference point when studying a wide range of phase transitions in mathematics, computer science and statistical physics: identifying the similarities and differences is often key for understanding more complex and perhaps less tractable models.

Various *concentration of measure phenomena* are often crucial for understanding the typical properties of random graphs. The basic idea is perhaps best illustrated by considering fair coin flips. If we toss a single coin the outcome is completely random (heads or tails), but if we toss a large number of coins the situation changes: we 'know' (although there are many possible outcomes!) that typically around half of the coins show heads. This apparent paradox, that the result of such random experiments is sharply predictable, is a well-known instance of the concentration of measure phenomenon. To see why such results are very useful for studying random graphs, note that (after appropriate rescaling) they allow us to treat many properties as 'almost deterministic', which essentially answers the following important question:

What are the typical values of certain key statistics?

In this context the perhaps best-understood property of the Erdős–Rényi process is the largest component, whose size after m steps we denote by $L_1(m)$. After rescaling, $L_1(tn)/n$ converges in probability to a continuous (deterministic) function $\rho(t)$ as $n \to \infty$, with $\rho(t) = 0$ for $t \leqslant 1/2$ and $\rho(t) > 0$ for $t > 1/2$. One nice feature of such convergence is that analytic properties of $\rho(t)$ give us a qualitative understanding of how L_1 grows. We will encounter similar concentration and convergence several times in this thesis: these will be key to understanding the dynamics of various random processes. Mathematically, here one key difference to classical results from probability theory (such as the 'law of large numbers' or the 'central limit theorem') is that the variables of interest are typically complicated (random) functions, which often requires the delicate usage and development of suitable tools and techniques.

So far we have seen that the Erdős–Rényi random graph process is an abstract mathematical model for complex networks evolving over time, which despite its simplicity exhibits fascinating phenomena. In fact, intuition from its study is also the starting point for understanding more recent network models, such as the 'scale-free' models of Barabási and Albert [9] and many others. However, the Erdős–Rényi process is unlikely to be appropriate itself for any particular application. In this context one inherent drawback of the classical Erdős–Rényi model is that it results from a sequence of independent choices, leading to a uniform distribution on all possible graphs; this is in great contrast to many real world networks, which, for example, satisfy additional structural constraints. To underline this drawback even further, the resulting graph after m steps is easily seen to be essentially equivalent to the so-called *binomial random graph* model, where each possible edge is inserted independently with a suitable probability p. This connection is exploited by the tools and techniques used in many proofs: these crucially rely on the independence between the edges! For this reason it is desirable, both mathematically and in terms of applications, to be able to understand and analyze more involved situations and processes evolving over time – in particular those where there is significant dependence among the edges and choices in different rounds. In this thesis we shall focus on two major lines of research in this direction; their basic ideas can be summarized as follows:

- *Power of random choices*: this paradigm goes one step beyond the classical case. Instead of always adding a purely random edge, each time *two* potential random edges are chosen and using some rule only *one* of them is selected and added to the evolving graph.

- *Constrained evolution*: we consider *conditional* random choices, instead of uniform ones. Given a (non-trivial) monotone decreasing graph property $\mathcal{P}$ (such as not containing a copy of some fixed graph H), we always add edges, one at a time, each chosen uniformly at random only from those ones whose addition does not destroy $\mathcal{P}$.

In the next two sections we briefly discuss the corresponding previous work and key questions. Afterwards we present our own main contributions to these areas of research.

1.1 Achlioptas Processes

At a Fields Institute workshop in 2000, Dimitris Achlioptas suggested a class of variants of the Erdős–Rényi random graph process. Starting with an empty graph on n vertices, these proceed as follows: in each step two potential edges e_1 and e_2 are chosen independently and uniformly at random from all possible edges, and using a given rule one of them is selected and added to the evolving graph. Any possible rule or selection strategy gives rise to a random graph process, and the class of these is known as *Achlioptas processes.* Always adding the first edge e_1 gives the classical process, where the linear size giant component starts emerging after roughly $n/2$ steps.

Achlioptas processes were introduced in an attempt to create random graph processes with potentially different behaviour than the Erdős–Rényi one, and in contrast to the classical model there can indeed be significant dependencies between the edges added. The motivation for studying such variations is (i) to test and develop the methods for analyzing random processes with dependencies and (ii) to explore and improve our understanding of phase transition phenomena in evolving random structures. Much in this spirit, Achlioptas originally asked whether the 'power of choices' in each step can be used to substantially delay the appearance of the giant component (for the classical 'balls into bins' problem it had been much earlier observed [8] that a limited amount of random choices in each step can be used to substantially alter structural properties). Bollobás [23] suggested the *product rule* as a natural candidate to achieve this: it always picks the edge minimizing the product of the sizes of the components of its endvertices. Bohman and Frieze [14] answered Achlioptas' question affirmatively in 2001: they proved that a much simpler edge selection rule is able to push back the emergence of the giant component substantially (beyond $0.53n$ steps). Their work showed for the first time that the 'power of choices' could indeed alter the behaviour of the Erdős–Rényi process significantly, initiating an entire line of research [10, 16, 18, 19, 35, 45, 46, 49, 50, 55, 74] investigating Achlioptas processes in more detail.

During the last decade the evolution of certain 'simple' Achlioptas processes has received considerable attention, mainly for so-called *bounded-size* rules, see e.g. [11, 19, 45, 46, 74], which are closely related to the one considered by Bohman and Frieze [14]. These make their decisions based only on the sizes of the components containing the endvertices of e_1 and e_2, with the restriction that all sizes larger than some constant B are treated in the same way. For bounded-size rules a number of results have been established by Bohman and Kravitz [19] and Spencer and Wormald [74]. These mainly concern the number of vertices in components of fixed size, and the existence and location of the percolation phase transition, but several intriguing questions regarding the size of the largest component remain open. For one particular bounded-size rule (a variant of the process first considered by Bohman and Frieze [14]), some finer details of the percolation phase transition have very recently also been investigated [11, 45, 46]. The punchline of the work mentioned above is that bounded-size rules seem to have many qualitative similarities with the classical Erdős–Rényi random graph process. While this can be interpreted as saying that the Erdős–Rényi evolution is rather 'universal', it might also indicate that bounded-size rules are 'too simple' for altering the behaviour significantly.

In contrast, for more involved Achlioptas processes very few rigorous results are known, and it may well be that these exhibit a qualitatively different behaviour. This in particular applies to 'unbounded'-size rules, usually simply called *size rules*, whose choices depend only on the sizes of the four components containing the endvertices of the two offered edges. Indeed, in one line of research it was widely believed (based on extensive simulations dating back to at least 2003) that these could give rise to a striking and unusual phenomenon. Making this explicit, Achlioptas, D'Souza and Spencer conjectured 2009 in *Science* [1], based on 'conclusive numerical evidence', that the product rule yields a very abrupt (discontinuous) percolation phase transition, a phenomenon known as *explosive percolation*. This conjecture was supported by many physicists (see e.g. [5, 7, 28, 29, 36, 40, 52, 56, 60, 61, 82, 83]) on the basis of simulations and heuristics. However, so far no theorems have been proven, and Achlioptas, D'Souza and Spencer [1] even state that involved Achlioptas processes using e.g. the product rule 'seem beyond the reach of current mathematical techniques'.

Summarizing, for simple Achlioptas rules we have a rather good understanding, with mathematical proofs, and their behaviour is qualitatively similar to the classical Erdős–Rényi process. In contrast, during the last decade more complex Achlioptas rules have remained resistant to a mathematical analysis, and it is widely believed that these might exhibit a novel and surprising behaviour. This makes their rigorous understanding the most prominent challenge in this area of research.

1.2 H-free Processes

The *H-free process* was suggested by Bollobás and Erdős [22] at the '*Quo Vadis, Graph Theory?*' conference in 1990. Given some fixed graph H, it is a modification of the classical random graph process, where each new edge is chosen uniformly at random subject to the condition that no (not necessarily induced) copy of H is formed. It was first described in print in 1995 by Erdős, Suen and Winkler [34], who asked how many edges the final graph typically has (this also appears as a problem in [30]).

From the beginning one key motivation for studying the H-free process comes from its potential applications to Ramsey and Turán type problems, two central topics in extremal combinatorics. Both areas of research have a long history, and determining certain Ramsey and Turán numbers is notoriously difficult. In this context the analysis of the H-free process has produced several new results, which still give the best known estimates. For example, improved lower bounds on the Turán numbers of certain bipartite graphs and Ramsey numbers $R(s,t)$ with $s \geqslant 4$ have been established in [13, 17, 78], and Bohman [13] also reproved the famous lower bound for $R(3,t)$ obtained in 1995 by Kim [47].

Further motivation for investigating the H-free process lies in the development of tools and techniques that can be used to analyse random processes in which there are significant dependencies between different rounds. Historically, here the impact of the H-free process can be traced back till at least 1995, where Kim's breakthrough [47] concerning $R(3,t)$ was achieved by analyzing a 'semi-random' variant of the C_3-free process in a technical tour-de-force (using sophisticated martingale estimates). A major technical breakthrough came in 2009, when Bohman [13] managed to analyse the C_3-free process itself. His proof technique significantly develops Wormald's differential equation method [80, 81] from 1995, which is widely used when dealing with dependencies in stochastic processes.

The history of the H-free process is rather complex. The first results considered special graphs and determined the typical final number of edges up to logarithmic factors. It started with Erdős, Suen and Winkler [34], who studied $H = C_3$ in 1995; later, in 2000, Bollobás and Riordan [25] analyzed $H \in \{K_4, C_4\}$. To be totally precise, one (much simpler) special case had been studied beforehand (in disguise): in 1992 Ruciński and Wormald [68] analyzed the maximum degree d-process, which corresponds to $H = K_{1,d+1}$. The general H-free process was first studied independently by Bollobás and Riordan [25] and Osthus and Taraz [57]. For H satisfying a certain density condition (strictly 2-balanced), Osthus and Taraz determined in 2000 the typical final number of edges up to logarithmic factors and conjectured that the average degree in the final graph of the C_ℓ-free process is $\Theta((n \log n)^{1/(\ell-1)})$.

The next improvements came about ten years later. In a breakthrough in 2009, Bohman [13] obtained the first matching bounds: he proved that the C_3-free process typically ends with $\Theta(n^{3/2}\sqrt{\log n})$ edges, confirming a conjecture of Spencer [71]. Next, Wolfovitz [78] slightly improved the lower bound on the (expected) final number of edges for a range of graphs H. Very recently, Bohman and Keevash [17] obtained new lower bounds for the class of strictly 2-balanced graphs H, which they conjectured to be tight up to the constants. In fact, their conjecture is for the maximum degree: for the C_ℓ-free process they conjectured that the maximum degree is usually at most $D(n \log n)^{1/(\ell-1)}$ for some $D > 0$.

To summarize, the most important open question regarding H-free processes is to determine its likely final number of edges, which is a more than 15-year old problem of Erdős et. al. [34, 30]. So far we know the answer for a large class of graphs H up to logarithmic factors, but more precise results are only known in a few isolated special cases.

1.3 Summary of Results

In this section we briefly outline the main results of this thesis, which concern variations of the classical Erdős–Rényi random graph process; for a more detailed discussion we refer to Chapters 2 and 6. The results in Chapters 3–5 are joint work with Oliver Riordan.

Part I: Achlioptas processes

The first part of this thesis is devoted to Achlioptas processes. We focus on the two main challenges related to 'involved' Achlioptas rules: (i) can these significantly alter the phase transition, i.e., achieve 'explosive percolation', and (ii) is it possible to rigorously analyze how these evolve over time?

Chapter 3: Explosive percolation is continuous

In one line of research [1, 5, 7, 26, 27, 28, 29, 32, 36, 52, 56, 60, 61, 82, 83] it was widely believed that certain 'complex' Achlioptas processes could give rise to a very abrupt (discontinuous) phase transition, a striking and unusual phenomenon. Based on 'conclusive numerical evidence' Achlioptas, D'Souza and Spencer [1] conjectured in *Science* that for the product rule there exists $\delta > 0$ such that with high probability the size of the largest component 'jumps' from $o(n)$ to at least δn in $o(n)$ steps of the process, a phenomenon known as 'explosive percolation'.

In this chapter we give a simple proof that this is not the case, showing that the transition is in fact continuous for *all* Achlioptas process. In fact, our result applies to a more general class of processes where a fixed number of independent random vertices are chosen at each step, and (at least) one edge between these vertices is added to the current graph, according to any (online) rule. We also prove the existence and continuity of the limit of the rescaled size of the giant component in a class of such processes, settling a number of conjectures of Spencer and Wormald [74] and Spencer [73]. Furthermore, we show that if the number of vertices sampled in each step is allowed to grow with n (arbitrarily slowly), then explosive percolation is possible.

This chapter is based on [62] and [63], which has appeared in *Science* and is to appear in *Annals of Applied Probability*, respectively.

Chapter 4: Convergence via differential equations with unique solutions

From the beginning, the rigorous mathematical understanding of Achlioptas processes has been a key challenge. So far the main focus has been on rather simple rules, where all component sizes larger than some constant B are treated the same way. Indeed, the evolution of such bounded-size rules is nowadays well understood, mainly due to the work of Bohman and Kravitz [19] and Spencer and Wormald [74]. In contrast, for 'unbounded'-size rules such as the product rule very few rigorous results are known.

In this chapter we make progress in our mathematical understanding of complex size rules. We show that for these the limit of the rescaled size of the giant component exists and is continuous provided that a certain system of differential equations has a unique solution. Our result applies to a very large class of Achlioptas-like processes and generalizes previous convergence results in this area of research. Indeed, for bounded-size rules uniqueness is easy to establish, while for the product rule (and many other involved rules) we so far only know that at least one solution exists.

Our main contribution is a new approach for proving convergence, which relates the evolution of stochastic processes to an associated system of differential equations. Provided that the latter has a unique solution, our approach shows that certain discrete quantities converge (after appropriate rescaling) to this solution.

This chapter is based on the manuscripts [65] and [64].

Chapter 5: The evolution of subcritical Achlioptas processes

In this chapter we present the first rigorous results for complex Achlioptas rules such as the product rule, which had remained resistant to a mathematical analysis for more than 10 years. We show that several key statistics are tightly concentrated at least until the susceptibility (the expected size of the component containing a randomly chosen vertex) diverges. Our convergence result is most likely best possible for certain rules: in the later evolution the number of vertices in small components may not be concentrated. Furthermore, we believe that the critical time where the susceptibility 'blows up' coincides with the percolation threshold, where the linear size giant component starts emerging.

Our proof is based on a variant of the neighbourhood exploration process and relies on branching process (approximation) arguments. This is quite different from previous approaches in this area, which rely on Wormald's differential equation method [80, 81].

This chapter is based on the manuscript [66].

Part II: H-free processes

In the second part of this thesis we consider the H-free process. Recall that, given a fixed graph H, this process starts with the empty graph on n vertices and adds edges chosen uniformly at random, one at a time, subject to the condition that no copy of H is created. We address the main open question for the H-free process, which was asked in 1995 by Erdős, Suen and Winkler [34, 30]: with how many edges does it typically terminate?

Chapter 7: The C_ℓ-free process

We focus on the C_ℓ-free process, where we forbid a cycle on ℓ vertices. In a breakthrough in 2009, Bohman [13] determined the final number of edges for the special case $\ell = 3$ up to constants, but his subsequent joint work with Keevash [17] as well as the much earlier work of Osthus and Taraz [57] only give lower and upper bounds on the final number of edges that differ by logarithmic factors. We close this gap: for every $\ell \geqslant 4$ we show that the maximum degree in the C_ℓ-free process is typically at most $D(n \log n)^{1/(\ell-1)}$ for some $D > 0$, confirming conjectures of Osthus and Taraz [57] and Bohman and Keevash [17].

Combined with previous results this implies that the C_ℓ-free process typically terminates with $\Theta(n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)})$ edges, which answers the more than 15-year old question of Erdős et. al. [34] for the C_ℓ-free process. In fact, our result is the first one that determines the final number of edges for a non-trivial *class* of forbidden graphs H, rather than an isolated single case.

Our proof combines (a new variant of) the differential equation method with a tool that might be of independent interest: we establish a rigorous way to 'transfer' certain decreasing properties from the binomial random graph to the H-free process.

This chapter is based on [77] and [76], which are to appear in *Random Structures & Algorithms*.

Part I

Achlioptas processes

Chapter 2

Overview

2.1 Background and motivation

In Chapters 3–5 we study variants of the classical Erdős–Rényi random graph process which are known as *Achlioptas processes*. Starting with an empty graph on n vertices, these proceed as follows: in each step two potential edges e_1 and e_2 are chosen independently and uniformly at random from all possible edges, and using a given rule one of them is selected and added to the evolving graph. Note that this actually yields a *family* of random graph processes (one process for each rule) which includes the original Erdős–Rényi process (the rule which, say, always adds the first edge e_1).

These processes were introduced by Dimitris Achlioptas in 2000 in an attempt to create random graph processes with potentially different behaviour than the Erdős–Rényi one. The main motivation for studying such variations is twofold: (i) to explore and improve our understanding of phase transitions in random processes, and (ii) to test and develop the methods for analyzing random processes with dependencies. Closely related to these two points Achlioptas originally asked whether there is a rule which substantially delays the appearance of the linear size 'giant' component compared to the classical case (where it starts emerging after roughly $n/2$ steps).

Bollobás [23] suggested the *product rule* as a natural candidate for delaying the giant component: it always picks the edge minimizing the product of the sizes of the components of its endvertices. Bohman and Frieze [14] answered Achlioptas' question affirmatively in 2001 using a much simpler rule, which is whp (meaning with high probability, i.e., with probability tending to 1 as $n \to \infty$) able to delay the giant component beyond, say, $0.53n$ steps. The following variant of their rule is nowadays known as the 'Bohman–Frieze process': it adds the first edge e_1 if it connects two isolated vertices, otherwise it simply adds the second edge e_2 (no matter what it looks like). Having established that Achlioptas processes could indeed substantially differ from the Erdős–Rényi process, [14] initiated a line of research [10, 16, 18, 19, 35, 45, 46, 49, 50, 55, 74] investigating their behaviour in more detail.

The natural benchmark in this context is the Erdős–Rényi process. Writing $N_k(m)$ and $L_1(m)$ for the number of vertices in components of size k and the largest component after m steps, respectively, some of its most most basic and fundamental properties are:

- **Local convergence**

  There exist functions $(\rho_k)_{k\geqslant 1}$ with $\rho_k : [0,\infty) \to [0,1]$ such that for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have $N_k(\lfloor tn \rfloor)/n \xrightarrow{\text{p}} \rho_k(t)$ as $n \to \infty$, where $\xrightarrow{\text{p}}$ denotes convergence in probability.

- **Global convergence**

  For each fixed $t \geqslant 0$ we have $L_1(\lfloor tn \rfloor)/n \xrightarrow{\text{p}} \rho(t)$ as $n \to \infty$, where $\rho : [0,\infty) \to [0,1]$ satisfies $\rho(t) = 0$ and $\rho(t) > 0$ for $t \leqslant 1/2$ and $t > 1/2$, respectively.

- **Continuous phase transition**

  For each fixed $\alpha > 0$ there exists $t > 1/2$ such that whp $L_1(tn) \leqslant \alpha n$.

It is intriguing to explore which of these features are model-specific and which are universal, i.e., hold in a wide range of related processes. With this in mind certain 'simple' Achlioptas rules, which can be thought of as generalizations of the Bohman–Frieze process, have been widely studied [11, 19, 45, 46, 74]. For these rules our mathematical understanding is rather good, and it turns out that they share many similarities with the classical Erdős–Rényi process. By contrast certain 'complex' Achlioptas rules, which include the product rule, have eluded a rigorous analysis during the last decade: these have only been investigated via simulations [1, 5, 7, 26, 27, 28, 29, 31, 32, 36, 52, 56, 60, 61, 82, 83]. The delicacy of the difference between these two classes of rules ('simple' and 'complex') is further stressed by the fact that certain complex rules seem to show a fundamentally different (and surprising) behaviour in simulations (see e.g. [1, 32, 36, 61]). This makes their rigorous mathematical understanding an important challenge.

2.2 Different rules

In the following we briefly introduce the different kinds of Achlioptas rules we consider. First of all, there is nothing special about the choice between two edges in each round used in the original examples of Achlioptas. Indeed, we will usually consider the more general class of *ℓ-vertex rules*, where in each step $\ell \geqslant 2$ vertices $v_1, \ldots, v_\ell$ are chosen independently and uniformly at random and then at least (usually exactly) one edge between these vertices is added according to some rule $\mathcal{R}$. Informally, we can think of ℓ-vertex rules in terms of a game: a hypothetical purposeful agent is presented with the random sequence $(v_1, \ldots, v_\ell)$ of vertices, and must add one or more edges between them, according to any deterministic or random rule that depends only on the history. As one would expect, these definitions are robust with respect to small changes: for example, the difference between selecting the endpoints of two (distinct) edges (not already present) and four uniformly random vertices is negligible for our purposes. This allows us to treat the original examples of Achlioptas as 4-vertex rules where $\mathcal{R}$ always selects one of the pairs $\{v_1, v_2\}, \{v_3, v_4\}$.

The rules discussed so far have been very general, making choices between the given edges using any information about the current graph. The following natural smaller classes of ℓ-vertex rules, which both *limit the 'knowledge'* in each step, have been widely studied:

- **Size rules:** the decision which edge(s) to add depends only on the sizes $(c_1, \ldots, c_\ell)$ of the components containing the sampled vertices $(v_1, \ldots, v_\ell)$.

- **Bounded-size rules:** these are size rules for which all component sizes c_j larger than some constant B are treated the same way.

Note that bounded-size rules can only distinguish between a constant number of different component sizes, which at least on an intuitive level indicates that these have a much simpler structure than (unbounded) size rules. Perhaps it thus comes as no surprise that these two classes essentially coincide with the informal 'complex' versus 'simple' distinction we used earlier. Indeed, the product rule is an example of a size rule, which has only been studied via simulations [5, 7, 28, 29, 36, 40, 52, 56, 60, 61, 82, 83]. In contrast, the Bohman–Frieze process is perhaps the simplest example of a bounded-size rule (with $B = 1$), and for this class of rules we have a rather good mathematical understanding [11, 14, 19, 45, 46, 74].

For certain questions the classes defined above are too broad, and for this reason we also consider further *dynamic properties* of these rules. In this context we study the following two natural restrictions:

- **Merging rules:** whenever C, C' are distinct components with sizes $|C|, |C'| \geqslant \varepsilon n$, then in the next step the rule joins C to C' with probability at least ε^ℓ.
- **Well behaved rules:** informally, these are size rules whose decisions are robust with respect to small changes in the size of the largest component L_1. Namely, whenever all sampled vertices v_j outside L_1 are contained in different components (of much smaller size), these essentially can decide which edge(s) to add *without* knowing the exact size of L_1 (see Section 4.2.1 for the formal definition).

Intuitively, 'merging' ensures that 'large' components quickly merge, similar as in the Erdős–Rényi process. This constraint arises very naturally; it is satisfied by the original examples of Achlioptas and many other rules, including all 'edge-based' ones which can only join (at least) one of the pairs $\{v_1, v_2\}, \{v_3, v_4\}, \ldots, \{v_{\ell-1}, v_\ell\}$ in each round. The 'well-behaved' condition is also very mild and holds for essentially all Achlioptas processes previously studied, including all bounded-size rules as well as interesting 'complex' rules such as the product rule (see Section 4.2.1). It conveniently ensures that decisions are not affected by small variations in the size of the largest component (such random fluctuations occur even in the classical Erdős–Rényi case), which allows us to avoid certain technical issues.

2.3 Guide to main results

In this section we give an informal overview of our main results for Achlioptas Processes.

2.3.1 Explosive Percolation

In the Erdős–Rényi random graph process the linear size giant component starts emerging after roughly $n/2$ steps, and the phase transition is *continuous*: informally this means that 'adding a few edges only increases the largest component by a bit'. In contrast, in a *discontinuous transition* the size of the largest component 'jumps' almost instantaneously, growing from sublinear size to some fixed proportion of all vertices. This striking phenomenon is nowadays also known as *explosive percolation* [1, 12, 41, 62], which formally occurs if there is a positive δ such that whp L_1 grows from size $o(n)$ to size at least δn in $o(n)$ steps.

Spencer and Wormald [74] conjectured that all bounded-size rules have continuous phase transitions. In contrast simulations dating back to at least 2003 (see [74]) suggested that (unbounded) size rules, in particular the product rule, could give rise to an 'explosive' discontinuous transition. This behaviour, if true, would indeed be rather surprising since in a wide range of random graphs studied in mathematical physics (percolation models) continuity is considered to be a universal feature of such phase transitions. Furthermore, for a large family of random graph models, the so-called inhomogeneous random graph models [24] introduced by Bollobás, Janson and Riordan (which include the Erdős–Rényi one), we know that the transition is continuous. With this in mind, even the possibility that a 'small' modification of the classical Erdős–Rényi model (given here by the product rule) could yield such a strikingly different behaviour is really remarkable. For the above reasons it was a big surprise for many when Achlioptas, D'Souza and Spencer presented in *Science* [1] 'conclusive numerical evidence' for the conjecture that the product rule exhibits explosive percolation (in fact, they suggested that the transition is even more rapid, namely that the largest component grows from size at most $\sqrt{n}$ to size at least $n/2$ in at most $2n^{2/3}$ steps). This was supported by many follow-up papers in the physics literature (see e.g. [5, 7, 28, 29, 36, 40, 52, 56, 60, 61, 82, 83]) using heuristics and simulations. In a *Science* 'perspectives' commentary Bohman [12] also describes this conjecture as an important and intriguing mathematical question.

The main result of Chapter 3 disproves this widely believed conjecture in a very strong sense: it shows that whenever we sample a fixed number of edges (vertices) in each step, then no matter which rule we use, it is simply impossible to obtain explosive percolation. In informal language we show the following:

(A) **No Achlioptas process can exhibit explosive percolation** (Theorem 3.1.1). Given fixed $\ell \geqslant 2$, *all* ℓ-vertex rules have continuous phase transitions.

(B) **Infinite choice per step can yield discontinuous transitions** (Theorem 3.6.1). If $\ell = \ell(n) \to \infty$ as $n \to \infty$, then a particular simple ℓ-vertex rule exhibits explosive percolation.

In some sense (A) shows that continuity of the phase transition is such an essential feature of the Erdős–Rényi process that it survives under a wide range of 'fixed size deformations' in each step. In fact, (B) demonstrates that for this continuity property we even have a threshold phenomenon with respect to ℓ (the number of vertices sampled in each step): we can obtain a discontinuous transition if we allow for an 'infinite' amount of choices in each step.

The proof of Theorem 3.1.1 is surprisingly simple (given the prominence of the explosive percolation conjecture), and the arguments used yield further structural insights. For example, for merging ℓ-vertex rules rules we obtain that whp the second largest component always has $o(n)$ size (see the discussion following Theorem 3.1.2). This implies that there will be at most one giant component, settling Spencer's 'no two giants' conjecture [73] in a strong form. Here the merging assumption is in fact needed for $\ell \geqslant 3$, as certain ℓ-vertex rules create $\ell - 1$ linear size components of similar size (see Section 3.3).

2.3.2 Convergence properties

From the beginning one key question about Achlioptas processes has been which rules are 'convergent', but so far only few rigorous results are known. More precisely, we say that a rule $\mathcal{R}$ is *locally convergent* if there exist functions $\rho_k = \rho_k^{\mathcal{R}} : [0, \infty) \to [0, 1]$ such that, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have $N_k(\lfloor tn \rfloor)/n \xrightarrow{\text{p}} \rho_k(t)$ as $n \to \infty$. Similarly, a rule $\mathcal{R}$ is called *globally convergent* if there exists an increasing function $\rho = \rho^{\mathcal{R}} : [0, \infty) \to [0, 1]$ such that for any t at which ρ is continuous we have $L_1(\lfloor tn \rfloor)/n \xrightarrow{\text{p}} \rho(t)$ as $n \to \infty$, where $\rho = \rho^{\mathcal{R}}$ is called the scaling limit of $\mathcal{R}$. Previously, local convergence has only been established for all (edge-based) bounded-size rules [19, 74], whereas global convergence is not even known for any non-trivial rules. However, Spencer and Wormald [74] conjectured that all (edge-based) bounded-size rules are globally convergent.

The main results of Chapters 3–5 improve our rigorous understanding of (local and global) convergence in Achlioptas processes in several ways. Informally these can be summarized (in simplified form) as follows:

(A) **Local convergence $\Longrightarrow$ global convergence** (Theorem 3.1.3).

For merging ℓ-vertex rules $\mathcal{R}$ local convergence implies global convergence, where the scaling limit $\rho = \rho^{\mathcal{R}}$ is continuous and satisfies $\rho(t) = 1 - \sum_{k \geqslant 1} \rho_k(t)$.

(B) **Unique solution $\Longrightarrow$ local convergence** (Theorem 4.1.1).

For merging ℓ-vertex size rules $\mathcal{R}$ that are well behaved the following holds: if a certain associated system of differential equations has a unique solution $(\hat{\rho}_k(t))_{k \geqslant 1}$, then $\mathcal{R}$ is locally convergent with $\rho_k = \hat{\rho}_k$.

(C) **Bounded susceptibility $\Longrightarrow$ local convergence** (Theorem 5.1.1).

All ℓ-vertex size rules $\mathcal{R}$ are locally convergent (in a strong form) up to a critical time $t_{\mathrm{b}} = t_{\mathrm{b}}^{\mathcal{R}}$, where the susceptibility (the expected size of the component containing a randomly chosen vertex) starts diverging.

Although the conditional nature of (A) and (B) might seem unsatisfactory at first sight, they do apply to a wide range of Achlioptas processes. In fact, (A) establishes the first global convergence results for Achlioptas processes. To be more precise, since for bounded-size rules local convergence is well known [19, 74], (A) settles two conjectures of Spencer and Wormald [74] for bounded-size rules: the scaling limit ρ exists and is continuous. Furthermore, (B) generalizes and extends previous convergence results for Achlioptas processes, as the rules they apply to are all (i) merging and well-behaved and (ii) their associated system of differential equations has a unique solution (see Section 4.1). The main contribution of this result is a new approach for proving convergence (establishing a more direct connection between the random process and the differential equations than in previous work).

One drawback of (B) is that for complex size rules such as the product rule the associated system of differential equations has infinite size, which introduces technical difficulties (for the product rule we know that at least one solution exists, but uniqueness remains an open problem). Here (C) shows that up to the point where the susceptibility diverges we can remove the uniqueness assumption of (B) for all size rules. Again, the main contribution is of a technical nature: Theorem 5.1.1 is the first rigorous convergence result for unbounded size rules such as the product rule, which had previously remained resistant to a rigorous analysis for more than 10 years. One key difference to previous work is that we rely on branching process approximation arguments (avoiding differential equations completely).

Finally, in the evolution of Achlioptas processes intriguing phenomena seem to happen at the critical time t_{b} given by (C). First of all, based on the heuristics and simulations presented in [64] we believe that convergence up to t_{b} as in (C) is best possible for the class of ℓ-vertex size rules: there are certain natural size rules which seem to exhibit a non-convergent behaviour beyond t_{b}. Secondly, we believe that for size rules t_{b} coincides with the percolation threshold t_{c} where the linear size giant component starts emerging (Conjecture 5.1.2). We prove that this is indeed true for all bounded-size ℓ-vertex rules (Theorem 5.3.1), as well as many other size rules (Theorems 5.3.2 and 5.3.3), including the 'reverse product rule', for example. Furthermore, we show that this can not be extended to general ℓ-vertex rules: the critical point where the susceptibility blows up does not always coincide with the percolation threshold (Theorems 5.3.5 and 5.3.6).

Chapter 3

Explosive percolation is continuous

3.1 Main results

This chapter considers the 'explosive percolation' phenomenon in Achlioptas processes. In this context the *product rule* has been widely studied: given the choice between two potential edges, it always picks the one that minimizes the product of the component sizes of its endpoints. While originally suggested (by Bollobás [23]) as a candidate for delaying the appearance of the linear size giant component, by 2003 at the latest (see [74]), extensive simulations of D'Souza and others strongly suggested that the product rule shows a much more interesting behaviour. Indeed, as illustrated by Figure 3.1 it seems to undergo a discontinuous phase transition, a phenomenon is known as *explosive percolation*: there is $\delta > 0$ such that whp the rescaled size of the largest component L_1/n 'jumps' from $o(1)$ to at least δ in $o(n)$ steps. Making this earlier belief explicit, Achlioptas, D'Souza and Spencer recently presented in *Science* [1] 'conclusive numerical evidence' for the conjecture that the product rule indeed exhibits explosive percolation, suggesting indeed that the largest component grows from size at most $\sqrt{n}$ to size at least $n/2$ in at most $2n^{2/3} = o(n)$ steps. This explosive percolation conjecture was supported by many papers in the physics literature, see e.g. [5, 7, 28, 29, 36, 40, 52, 56, 60, 61, 82, 83]; Bohman [12] also describes it as an important and intriguing mathematical question.

Our main result disproves this widely believed conjecture in strong form, showing that the transition is in fact continuous for *all* Achlioptas process (not only for the product

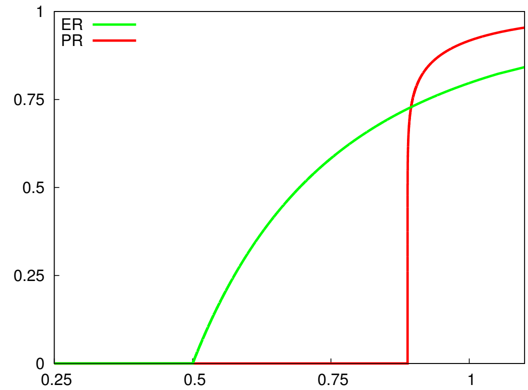

Figure 3.1: Evolution of the rescaled size of the largest component $L_1(tn)/n$ for the Erdős–Rényi process (ER) and the Product Rule (PR) as a function of t (simulation with $n = 10^{12}$). The PR curve appears to be discontinuous at the critical point, but Theorem 3.1.1 shows that this is not the case.

rule). In fact, it applies to a more general class of Achlioptas-like processes ('ℓ-vertex rules') defined in Section 3.2. A form of this result first appeared in *Science* [62] (together with an accompanying 'perspectives' commentary by Janson [41]), with more restrictive assumptions, and without full technical details.

Theorem 3.1.1. *Let $\mathcal{R}$ be an ℓ-vertex rule for some $\ell \geqslant 2$. For each n, let $(G(m))_{m\geqslant 0} = (G_n^{\mathcal{R}}(m))_{m\geqslant 0}$ be the random sequence of graphs on vertex set $\{1, 2, \dots, n\}$ associated to $\mathcal{R}$. Given any functions $h_L(n)$ and $h_m(n)$ that are $o(n)$, and any constant $\delta > 0$, the probability that there exist m_1 and m_2 with $L_1(G(m_1)) \leqslant h_L(n)$, $L_1(G(m_2)) \geqslant \delta n$ and $m_2 \leqslant m_1 + h_m(n)$ tends to 0 as $n \to \infty$.*

Let $N_k(G)$ denote the number of vertices of a graph G in components with k vertices, so $N_k(G)$ is k times the number of k-vertex components. Similarly, $N_{\leqslant k}(G)$ and $N_{\geqslant k}(G)$ denote the number of vertices in components with at most (at least) k vertices. Having a rule $\mathcal{R}$ in mind, and suppressing the dependence on n, we write $N_k(m)$ for the random quantity $N_k(G(m))$, and similarly $L_1(m)$ for $L_1(G(m))$.

Under a mild additional condition (which holds for all Achlioptas processes), a slight modification of the proof of Theorem 3.1.1 shows, roughly speaking, that the giant component is unique. In fact, we obtain much more: whp there is no time at which there are 'many' vertices in 'large' components but not in the single largest component. For the precise definition of a 'merging' rule see Section 3.3; any Achlioptas process is merging.

Theorem 3.1.2. *Let $\mathcal{R}$ be a merging ℓ-vertex rule for some $\ell \geqslant 2$. For each n, let $(G(m))_{m\geqslant 0} = (G_n^{\mathcal{R}}(m))_{m\geqslant 0}$ be the random sequence of graphs on $\{1,2,\dots,n\}$ associated to $\mathcal{R}$. For each $\varepsilon > 0$ there is a $K = K(\varepsilon,\ell)$ such that*

$$\mathbb{P}\big(\forall m : N_{\geqslant K}(m) < L_1(m) + \varepsilon n\big) \to 1$$

as $n \to \infty$.

With ℓ fixed, our proof gives a value for K of the form $\exp(\exp(c\varepsilon^{-(\ell-1)}))$ for some positive $c = c(\ell)$. Furthermore, we can allow ε to depend on n, as long as $\varepsilon = \varepsilon(n) \geqslant d/(\log\log n)^{1/(\ell-1)}$, where $d = d(\ell) > 0$.

For the classical random graph process it is well known that at any fixed time, whp there will be at most one 'giant' component. Indeed, the maximum size of the second largest component throughout the evolution of the process is whp $o(n)$: this can be read out of the original results of Erdős and Rényi [33] or (more easily) the more precise results of Bollobás [20]. Spencer's 'no two giants' conjecture [73] states that this should also hold for Achlioptas processes. Theorem 3.1.2 proves this conjecture for the larger class of merging ℓ-vertex rules; indeed it readily implies that, with high probability, the second largest component has size at most $\max\{K, \varepsilon n\} = \varepsilon n$. Allowing ε to vary with n as noted above, the bound we obtain is of the form $d(\ell)n/(\log\log n)^{1/(\ell-1)}$.

Before turning to the proofs of Theorems 3.1.1 and 3.1.2, let us discuss some related questions of convergence. Recall that the rule $\mathcal{R}$ is called *locally convergent* if there exist functions $\rho_k = \rho_k^{\mathcal{R}} : [0,\infty) \to [0,1]$ such that, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have

$$\frac{N_k(\lfloor tn \rfloor)}{n} \xrightarrow{\mathrm{p}} \rho_k(t) \tag{3.1}$$

as $n \to \infty$. Similarly, we say that the rule $\mathcal{R}$ is *globally convergent* if there exists an increasing function $\rho = \rho^{\mathcal{R}} : [0,\infty) \to [0,1]$ such that for any t at which ρ is continuous we have

$$\frac{L_1(\lfloor tn \rfloor)}{n} \xrightarrow{\mathrm{p}} \rho(t)$$

as $n \to \infty$.

Theorem 3.1.1 clearly implies that if a rule $\mathcal{R}$ is globally convergent, then the limiting function ρ is continuous at the critical point $t_{\mathrm{c}} = \inf\{t : \rho(t) > 0\}$. Using Theorem 3.1.2, it is not hard to establish continuity elsewhere for merging rules; see Theorem 3.3.2 and Corollary 3.3.3 in Section 3.3. Unfortunately, we cannot show that the product rule *is* globally convergent. However, as we shall see in Section 3.4, Theorem 3.1.2 implies the following result.

Theorem 3.1.3. *Let $\mathcal{R}$ be a merging ℓ-vertex rule for some $\ell \geqslant 2$. If $\mathcal{R}$ is locally convergent, then $\mathcal{R}$ is globally convergent, and the limiting function $\rho^{\mathcal{R}}$ is continuous and satisfies $\rho^{\mathcal{R}}(t) = 1 - \sum_{k \geqslant 1} \rho_k^{\mathcal{R}}(t)$.*

The conditional result above is of course rather unsatisfactory. However, for many Achlioptas processes, local convergence is well known; global convergence had not previously been established for any non-trivial rule. In particular, Theorem 3.1.3 settles two conjectures of Spencer and Wormald [74] concerning so-called 'bounded-size Achlioptas processes'; see Section 3.5.

Recently, in a paper in the physics literature, da Costa, Dorogovtsev, Goltsev and Mendes [31] announced a version of Theorem 3.1.1. However, their actual analysis concerned only one specific rule (not the product rule, though they claim that 'clearly' the product rule is less likely to have a discontinuous transition). More importantly, even the 'analytic' part of it is heuristic, and of a type that seems to us very hard (if at all possible) to make precise. Crucially, the starting point for their analysis is not only to assume convergence, but also to assume that the phase transition is continuous! From this, and some further assumptions, by solving approximations to certain equations they deduce certain 'self-consistent' behaviour, which apparently justifies the assumption of continuity. The argument (which is considerably more involved than the simple proof presented here) is certainly very interesting, and the conclusion is (as we now know) correct, but it seems to be very far from a mathematical proof.

Finally, it is clear that if one departs far enough from the Erdős–Rényi model, then a discontinuous transition is possible. For example, suppose that at every step the two smallest components in the whole graph are joined. With n a power of two, this corresponds to filling in the edges of a binary tree from the leaves to the root, so the largest component grows extremely quickly in the last few steps (see e.g. [36]). Theorem 3.1.1 implies that any rule based on picking a fixed number ℓ of random vertices gives a continuous transition (see e.g. Section 3.2 for the definition of ℓ-vertex rules); the previous example demonstrates that considering all n vertices can give a discontinuous one. Where is the cut-off? It turns out that the cut-off is at the lower end of the range: whenever $\ell \to \infty$ as $n \to \infty$ then the *SDC rule* 'pick ℓ vertices at random and join the two smallest distinct components selected' exhibits explosive percolation (see Theorem 3.6.1).

In the next section we prove Theorem 3.1.1. In Section 3.3, restricting the class of rules slightly, we prove Theorem 3.1.2 and deduce that 'jumps' in L_1 are also impossible after a giant component first emerges. Next, in Section 3.4, we prove Theorem 3.1.3. Afterwards, in Section 3.5 we consider more restrictive rules such as 'bounded-size' rules, and discuss the relationship of our results to earlier work. Finally, in Section 3.6 we show that explosive percolation is possible if we sample an 'infinite' amount of vertices per step.

3.2 Definitions and proof of Theorem 3.1.1

Throughout we fix an integer $\ell \geqslant 2$. For each n, let $(\underline{v}_1, \underline{v}_2, \ldots)$ be an i.i.d. sequence where each $\underline{v}_m$ is a sequence $(v_{m,1}, \ldots, v_{m,\ell})$ of ℓ vertices from $[n] = \{1, 2, \ldots, n\}$ chosen independently and uniformly at random. Suppressing the dependence on n, informally, an *ℓ-vertex rule* is a random sequence $(G(m))_{m \geqslant 0}$ of graphs on $[n]$ satisfying (i) $G(0)$ is the empty graph, (ii) for $m \geqslant 1$ $G(m)$ is formed from $G(m-1)$ by adding a (possibly empty) set E_m of edges, with all edges in E_m between vertices in $\underline{v}_m$, and (iii) if all ℓ vertices in $\underline{v}_m$ are in distinct components of $G(m-1)$, then $E_m \neq \emptyset$. The set E_m may be chosen according to any deterministic or random 'online' rule.

Formally, we assume the existence of a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ such that $\underline{v}_m$ is $\mathcal{F}_m$-measurable and independent of $\mathcal{F}_{m-1}$, and require E_m (and hence $G(m)$) to be $\mathcal{F}_m$-measurable.

In other words, the 'agent' is presented with the random list (set) $\underline{v}_m$ of vertices, and, unless two or more are already in the same component, must add one or more edges between them, according to any deterministic or random rule that depends only on the history. In the original examples of Achlioptas, the rule always adds either the edge $\{v_{m,1}, v_{m,2}\}$ or the edge $\{v_{m,3}, v_{m,4}\}$. Note that (for now) no connection between the 'algorithms' used for different n (or indeed at different steps m) is assumed.

The arguments that follow are robust to small changes in the definition, since they can be written to rely only on deterministic properties of $(G(m))$, plus bounds on the probabilities of certain events at each step. The latter always have $\Theta(1)$ elbow room. It follows that we may weaken the conditions on $(\underline{v}_m)$: it suffices if, for $m = O(n)$, say, the conditional distribution of $\underline{v}_m$ given the history (i.e., given $\mathcal{F}_{m-1}$) is close to (at total variation distance $o(1)$ from, as $n \to \infty$) that described above. This covers variations such as picking an ℓ-tuple of *distinct* vertices, or picking (the ends of) $\ell/2$ randomly selected (distinct) edges not already present in $G(m-1)$.

The proof of Theorem 3.1.1 is based on two observations, which we first present in heuristic form.

Observation 1: If at some time t (i.e., when $m \sim tn$) there are αn vertices in components of order at least k, then within time $\gamma = O(1/(\alpha^{\ell-1}k))$ a component of order at least $\alpha n/\ell^2 = \beta n$ will emerge. Indeed, fix a set W with $|W| \geqslant \alpha n$ consisting of components of order at least k. At every subsequent step we have probability at least α^ℓ of choosing only vertices in W, and if no component has order more than βn, it is likely that all these vertices are in different components, so the rule is *forced* to join two components meeting W. This cannot happen more than $|W|/k$ times.

(A form of Observation 1 appears in a paper of Friedman and Landsberg [36], as a key part of a heuristic argument *for* explosive percolation. It is not quite stated correctly, although this does not seem to be why the heuristic fails.)

Observation 2: Components of order k have a half-life that may be bounded below in terms of k: in an individual step, such a component disappears (by joining another component) with probability at most $k\ell/n$. Assuming (which we shall not assume in the actual proof), that the rule $\mathcal{R}$ is locally convergent, it follows easily that for all t_1, t_2 and k we have $\rho_k(t_1 + t_2) \geqslant \rho_k(t_1)e^{-k\ell t_2}$.

We place vertices into 'bins' corresponding to component sizes between 2^j and $2^{j+1}-1$, writing $\sigma_j(t)$ for $\sum_{2^j \leqslant k < 2^{j+1}} \rho_k(t)$. Let $\alpha > 0$ be constant, and suppose that $\sigma_j(t) \geqslant \alpha$ for some $t < t_c$. Writing $k = 2^j$, by Observation 1 we have $t_c - t = O(1/k)$, with the implicit constant depending on α, since the $\geqslant \alpha n$ vertices in components of size at least k will quickly form a giant component. Using Observation 2, it follows that $\sigma_j(t_c) \geqslant g(\alpha) > 0$, for some (explicit but irrelevant) function $g(\alpha)$.

Let $\sigma_j = \sup_{t \leqslant t_c} \sigma_j(t)$. If $\sigma_j > \alpha$, then $\sigma_j(t_c) \geqslant g(\alpha)$. Counting vertices, we have $\sum_j \sigma_j(t_c) \leqslant 1$. Hence, for each $\alpha > 0$, only a finite number of σ_j can exceed α. Thus $\sigma_j \to 0$ as $j \to \infty$. It follows that for any constant $B \geqslant 2$ and any $k = k(n) \to \infty$, at no $t = t(n) < t_c$ can there be $\Theta(n)$ vertices in components of size between k and Bk.

Using Observation 1, it is easy to deduce that there cannot be a discontinuous transition. Indeed, if $\lim_{t \to t_c^+} \rho(t) \geqslant \delta > 0$, then for any k, at time $t_k = t_c - \delta/(\ell^2 k)$ there must be at least $\delta n/2$ vertices in components of order at least k, so $\rho_{\geqslant k}(t_k) \geqslant \delta/2$, where $\rho_{\geqslant k} = 1 - \sum_{k' < k} \rho_{k'}$. For any constant $B \geqslant 2$, if k is large it follows that $\rho_{\geqslant Bk}(t_k) \geqslant \delta/3$. Taking B large enough, Observation 1 then implies that $t_c - t_k$ is much smaller than $\delta/(\ell^2 k)$.

We now make the above argument precise, without assuming convergence. This introduces some minor additional complications, but they are easily handled. We start with two lemmas corresponding to the two observations above.

Lemma 3.2.1. *Given $0 < \alpha \leqslant 1$, let $\mathcal{C}(\alpha)$ denote the event that for all $0 \leqslant m \leqslant n^2$ and $1 \leqslant k \leqslant \frac{\alpha}{16} \frac{n}{\log n}$ the following holds: $N_{\geqslant k}(m) \geqslant \alpha n$ implies $L_1(m + \Delta) > \frac{\alpha}{\ell^2} n$ for $\Delta = \lceil \frac{4}{\alpha^{\ell-1}} \frac{n}{k} \rceil$. Then $\mathbb{P}(\mathcal{C}(\alpha)) \geqslant 1 - n^{-1}$.*

Proof. It suffices to consider 'fixed' m and k and show that, conditional on $\mathcal{F}_m$, if $G(m)$ satisfies $N_{\geqslant k}(m) \geqslant \alpha n$, then we have $L_1(m + \Delta) > \frac{\alpha}{\ell^2} n$ with probability at least $1 - n^{-4}$.

Condition on $\mathcal{F}_m$. Let W be the union of all components with size at least k in $G(m)$, set $\tilde{\alpha} = |W|/n \geqslant \alpha$ and let $\beta = \tilde{\alpha}/\ell^2$. We now consider the next Δ steps.

We say that a step is *good* if (a) all ℓ randomly chosen vertices are in W and (b) all these vertices are in different components. Let X_j denote the indicator function of the event that step $m+j$ is good. Set $X = \sum_{1\leqslant j\leqslant \Delta} X_j$ and $Y = \sum_{1\leqslant j\leqslant \Delta} Y_j$, where

$$Y_j = \begin{cases} X_j, & \text{if } L_1(m+j-1) \leqslant \beta n, \\ 1, & \text{otherwise.} \end{cases}$$

Clearly, in each step (a) holds with probability $\tilde{\alpha}^\ell$. Furthermore, whenever $L_1(m+j-1) \leqslant \beta n$ holds, in step $m+j$ the probability that (a) holds and (b) fails is at most $\binom{\ell}{2}\tilde{\alpha}^{\ell-1}\beta < \tilde{\alpha}^\ell/2$ (there must be v_a and v_b with $1 \leqslant a < b \leqslant \ell$ such that v_b lies in the same component as v_a; all v_c must also be in W) and so in this case step $m+j$ is good with probability at least $\tilde{\alpha}^\ell/2$. Since otherwise $Y_j = 1$ by definition, we deduce that Y stochastically dominates a binomial random variable with mean $\Delta\tilde{\alpha}^\ell/2 \geqslant 2\tilde{\alpha}n/k$. Standard Chernoff bounds now imply that $\mathbb{P}(Y \leqslant \tilde{\alpha}n/k) \leqslant e^{-\tilde{\alpha}n/(4k)} \leqslant e^{-\alpha n/(4k)} \leqslant n^{-4}$.

Assume that $L_1(m+\Delta) \leqslant \beta n$. Then by monotonicity $L_1(m+j-1) \leqslant \beta n$ for every $1 \leqslant j \leqslant \Delta$, so $X = Y$. Note that W contains at most $|W|/k = \tilde{\alpha}n/k$ components in $G(m)$. Since every good step joins two components meeting W (at least one such edge must be added since by (a) all endpoints are in W and by (b) all endpoints are in distinct components) we deduce that $Y \leqslant \tilde{\alpha}n/k$. Hence $\mathbb{P}(L_1(m+\Delta) \leqslant \beta n) \leqslant \mathbb{P}(Y \leqslant \tilde{\alpha}n/k) \leqslant n^{-4}$, as required. □

Applying Lemma 3.2.1 with $m = 0$, $k = 1$ and $\alpha = 1$, we readily deduce that whp a giant component exists after at most $4n$ steps. In fact, it is easy to see that for any $\varepsilon > 0$, whp there is a giant component after at most $(1+\varepsilon)n$ steps (see the proof of Lemma 3.3.1).

Lemma 3.2.2. *Fix $0 < \alpha \leqslant 1$, $D > 0$, and an integer $B \geqslant 2$. Define $M_k^B(m) = N_{\geqslant k}(m) - N_{\geqslant Bk}(m)$. Let $\mathcal{L}(\alpha, B, D)$ denote the event that for all $0 \leqslant m \leqslant n^2$ and $1 \leqslant k \leqslant \min\{\frac{\alpha^2 e^{-4\ell BD}}{8\ell^2 B^2 D}\frac{n}{\log n}, \frac{n}{2B}\}$ the following holds: $M_k^B(m) \geqslant \alpha n$ implies $M_k^B(m+\Delta) > \frac{\alpha}{2B}e^{-2\ell BD}n$ for every $0 \leqslant \Delta \leqslant D\frac{n}{k}$. Then $\mathbb{P}(\mathcal{L}(\alpha, B, D)) \geqslant 1 - n^{-1}$.*

Proof. As in the proof of Lemma 3.2.1, it suffices to consider fixed m and k, and show that conditional on $\mathcal{F}_m$, if $G(m)$ satisfies $M_k^B(m) \geqslant \alpha n$, then with probability at least $1 - n^{-4}$ we have $M_k^B(m+\Delta) > \frac{\alpha}{2B}e^{-2\ell BD}n$ for every $0 \leqslant \Delta \leqslant \tilde{\Delta}$, where $\tilde{\Delta} = \lfloor Dn/k \rfloor$.

Condition on $\mathcal{F}_m$ and $M_k^B(m) \geqslant \alpha n$, and let $C_1, \ldots, C_r$ be the components of $G(m)$ with sizes between k and $Bk-1$. Note that $r \geqslant M_k^B(m)/(Bk) \geqslant \alpha n/(Bk)$.

Starting from $G(m)$, we now analyze the next $\tilde{\Delta}$ steps. We say that C_i is *safe* if in each of these steps none of the ℓ randomly chosen vertices is contained in C_i, and we denote by X the number of safe components. Using $|C_i| \leqslant Bk \leqslant n/2$, note that C_i is safe with probability

$$(1-|C_i|/n)^{\ell\tilde{\Delta}} > e^{-2\ell\tilde{\Delta}|C_i|/n} \geqslant e^{-2\ell BD},$$

which gives $\mathbb{E}X \geqslant re^{-2\ell BD}$. Clearly, the random variable X can be written as $X = f(\underline{v}_{m+1}, \ldots, \underline{v}_{m+\Delta})$, where the $\underline{v}_j$ denote the ℓ-tuples generated by the ℓ-vertex process in each step (uniformly and independently). The function f satisfies $|f(\omega)-f(\tilde{\omega})| \leqslant \ell$ whenever ω and $\tilde{\omega}$ differ in one coordinate. So, using $r \geqslant \alpha n/(Bk)$, McDiarmid's inequality [53] implies that $\mathbb{P}(X \leqslant re^{-2\ell BD}/2)$ is at most

$$\exp\left(-\frac{2[re^{-2\ell BD}/2]^2}{\tilde{\Delta}\ell^2}\right) \leqslant \exp\left(-\frac{\alpha^2 e^{-4\ell BD}}{2\ell^2 B^2 D}\frac{n}{k}\right) \leqslant n^{-4}.$$

Suppose that $X > re^{-2\ell BD}/2$. Since every safe component contributes at least k vertices to every $M_k^B(m+\Delta)$ with $0 \leqslant \Delta \leqslant \tilde{\Delta}$ (in each step all edges which can be added are disjoint from safe components), using $r \geqslant \alpha n/(Bk)$ we deduce that for all such Δ we have $M_k^B(m+\Delta) \geqslant kX > \alpha e^{-2\ell BD} n/(2B)$, completing the proof. □

Note that by considering instead the number Y of vertices in safe components one can prove the slightly stronger bound $M_k^B(m+\Delta) > (1-\varepsilon)\alpha e^{-2\ell BD} n$, for k not too large.

We are now ready to prove Theorem 3.1.1.

Proof of Theorem 3.1.1. Let $h_L(n)$ and $h_m(n)$ be non-negative functions satisfying $h_L(n) = o(n)$ and $h_m(n) = o(n)$, and let $\delta > 0$ be constant. Let $\mathcal{X} = \mathcal{X}_n(\delta, h_L, h_m)$ denote the event that there exist m_1 and m_2 satisfying $L_1(m_1) \leqslant h_L(n)$, $L_1(m_2) \geqslant \delta n$, and $m_2 \leqslant m_1 + h_m(n)$, so our aim is to show that $\mathbb{P}(\mathcal{X}) \to 0$ as $n \to \infty$. We shall define a 'good' event $\mathcal{G} = \mathcal{G}_n(\delta)$ such that $\mathbb{P}(\mathcal{G}) \to 1$ as $n \to \infty$ and show *deterministically* that there is some n_0 such that for $n \geqslant n_0$, when $\mathcal{G}$ holds, $\mathcal{X}$ does not.

To be totally explicit, set $\alpha = \delta/4$, $A = 5/\alpha^{\ell-1}$ and $D = 1$. Set $B = \lceil 2A\ell^2/\delta \rceil$, and let $\beta = \alpha e^{-2\ell B}/(2B) > 0$. Finally, let $K = B^{1+\lceil 1/\beta \rceil}$, noting that K does not depend on n.

Let $\mathcal{G}$ be the event that $\mathcal{C}(1)$, $\mathcal{C}(\delta/4)$ and $\mathcal{L}(\delta/4, B, D)$ all hold simultaneously. By Lemmas 3.2.1 and 3.2.2, $\mathbb{P}(\mathcal{G}) \geqslant 1 - 3n^{-1} = 1 - o(1)$. The definition of $\mathcal{G}$ ensures that if n is large enough (larger than some constant depending only on δ and ℓ), then for all $m \leqslant 5n$ and $k \leqslant K$ the following hold:

$$\text{(i) } N_{\geqslant k}(m) \geqslant \delta n/4 \text{ implies (ii) } L_1(m + \lfloor An/k \rfloor) \geqslant \delta n/(4\ell^2),$$

and

$$\text{(iii) } M_k^B(m) \geqslant \delta n/4 \text{ implies (iv) } M_k^B(m') \geqslant \beta n \text{ for all } m \leqslant m' \leqslant m + n/k.$$

Suppose that $\mathcal{G}$ holds, and that $m^- = \max\{m : L_1(m) \leqslant h_L(n)\}$ and $m^+ = \min\{m : L_1(m) \geqslant \delta n\}$ differ by at most $h_m(n)$. It suffices to show deterministically that if n is large enough, then this leads to a contradiction.

Since $N_1(0) = n$ and $\mathcal{C}(1)$ holds, we have $L_1(4n) \geqslant n/\ell^2$. If n is large enough, it follows that $m^- \leqslant 4n$, so $m^+ \leqslant 5n$.

For $k \leqslant K/B$ set $m_k = m^+ - \delta n/(\ell^2 k)$, which is easily seen to be positive; we ignore the irrelevant rounding to integers. Since at most $\binom{\ell}{2}(m^+ - m_k) < \ell^2(m^+ - m_k)/2$ edges are added passing from $G(m_k)$ to $G(m^+)$, the components of $G(m_k)$ with size at most k together contribute at most $k\ell^2(m^+ - m_k)/2 \leqslant \delta n/2$ vertices to any one component of $G(m^+)$. It follows that

$$N_{\geqslant k}(m_k) \geqslant L_1(m^+) - \delta n/2 \geqslant \delta n/2.$$

Suppose that $N_{\geqslant Bk}(m_k) \geqslant \delta n/4$. Then (i) holds at step m_k with $Bk \leqslant K$ in place of k, so (ii) tells us that by step

$$m^* = m_k + \lfloor An/(Bk) \rfloor \leqslant m_k + \delta n/(2\ell^2 k) = m^+ - \delta n/(2\ell^2 k) = m^+ - \Theta(n)$$

we have $L_1(m^*) > \delta n/(4\ell^2)$, which is larger than $h_L(n)$ if n is large enough. Since $m^+ - m^- \leqslant h_m(n) = o(n)$, if n is large enough we have $m^* < m^-$, contradicting the definition of m^-.

It follows that $M_k^B(m_k) = N_{\geqslant k}(m_k) - N_{\geqslant Bk}(m_k) \geqslant \delta n/4$. Using (iii) implies (iv), this gives $M_k^B(m^+) \geqslant \beta n$. Applying this for $k = 1, B, B^2, \ldots, B^{\lceil 1/\beta \rceil}$ shows that $G(m^+)$ has more than n vertices, a contradiction. □

Setting $D = 2\delta/\ell^2$ (instead of $D = 1$), the proof above shows that the number of steps between $m^- = \max\{m : L_1(m) \leqslant \delta/(4\ell^2)n\}$ and $m^+ = \min\{m : L_1(m) \geqslant \delta n\}$ is at least $\delta n/(2\ell^2 B^{\lceil 1/\beta \rceil}) = f(\delta)n$, where $f(\delta)$ essentially grows like the inverse of a double exponential in $\delta^{-(\ell-1)}$ for $\delta \to 0$.

3.3 Results for merging rules

Although Theorem 3.1.1 applies to any ℓ-vertex rule, for many questions, this class is too broad. Indeed, consider a rule which only joins two components when forced to (i.e., when presented with ℓ vertices from distinct components) and then joins the two smallest components presented. Such a rule will *never* join two of the $\ell - 1$ largest components, and it is not hard to see that during the process $\ell - 1$ 'giant' components (with order $\Theta(n)$) will emerge and grow simultaneously, with their sizes keeping roughly in step. In what follows we could replace 'the largest component' by 'the union of the $\ell - 1$ largest components' and work with arbitrary ℓ-vertex rules, but this seems rather unnatural.

By an *r-Achlioptas rule* we mean an ℓ-vertex rule with $\ell = 2r$ that always joins (at least) one of the pairs $\{v_1, v_2\}, \{v_3, v_4\}, \ldots, \{v_{\ell-1}, v_\ell\}$. (How we treat the case where one or more of these pairs is in fact a single vertex will not be relevant.) An *Achlioptas rule* is an r-Achlioptas rule for any $r \geqslant 1$. Taking $r = 2$ and insisting that only one edge is added gives the original class of rules suggested by Achlioptas.

Let us say that an ℓ-vertex rule is *merging* if whenever C, C' are distinct components with $|C|, |C'| \geqslant \varepsilon n$, then in the next step we have probability at least ε^ℓ of joining C to C'. This implies that the probability that they are *not* united after m further steps is at most $e^{-\varepsilon^\ell m}$. (We could replace ε^ℓ by any $f(\varepsilon) > 0$, and it suffices if the chance of merging in one of the next few steps, rather than the next step, is not too small.) Clearly, any Achlioptas rule is merging: with probability at least ε^ℓ all $r = \ell/2$ potential edges join C to C'. There are other interesting examples of merging rules; see Section 3.5.

For merging rules we have the following variant of Lemma 3.2.1. We write $V_{\geqslant k}(m)$ for the union of all components with size at least k in $G(m)$, so $|V_{\geqslant k}(m)| = N_{\geqslant k}(m)$.

Lemma 3.3.1. *Let $\mathcal{R}$ be a merging ℓ-vertex rule, let $\varepsilon > 0$, let $k \geqslant 1$ and m be integers, and set $\Delta = 2\lceil \frac{2^\ell}{\varepsilon^{\ell-1}} \frac{n}{k} \rceil$. Conditioned on $\mathcal{F}_m$, with probability at least $1 - \ell \exp(-cn/k)$ there is a component of $G(m + \Delta)$ containing at least $N_{\geqslant k}(m) - \varepsilon n$ vertices from $V_{\geqslant k}(m)$, where $c = c(\varepsilon, \ell) > 0$.*

Proof. Let $W = V_{\geqslant k}(m)$, so $|W| = N_{\geqslant k}(m)$. We may assume that $|W| - \varepsilon n \geqslant 0$. Let $\alpha = |W|/n \geqslant \varepsilon$. Until the point that there are $\ell - 1$ components between them containing at least $(\alpha - \varepsilon/2)n$ vertices from W, at each step we have probability at least $\alpha(\varepsilon/2)^{\ell-1}$ of choosing ℓ vertices of W in distinct components to form $\underline{v}_j$, in which case the number of components meeting W must decrease by (at least) one. As in the proof of Lemma 3.2.1, it follows that off an event whose probability is exponentially small in n/k, after $\Delta/2$ steps we do have $\ell - 1$ components $C_1, \ldots, C_{\ell-1}$ together containing at least $(\alpha - \varepsilon/2)n$ vertices of W. Ignoring any containing fewer than $\varepsilon n/(2\ell)$ vertices of W, using the property of merging rules noted above, the probability that some pair of the remaining C_i are not joined in the next $\Delta/2$ steps is exponentially small in n/k. □

It is easy to check that we may take $c(\varepsilon, \ell) = \varepsilon/\ell^\ell$. With this technical result in hand, we now prove Theorem 3.1.2.

Proof of Theorem 3.1.2. We outline the argument, much of which is very similar to the proof of Theorem 3.1.1 given in the previous section.

Let $\varepsilon > 0$ be given, and set $\delta = \varepsilon/5$. Lemma 3.3.1 implies that there is some $A = A(\delta, \ell)$ such that for any fixed k, it is very likely that (i) there is a component of $G(m + \lfloor An/k \rfloor)$ containing at least $N_{\geqslant k}(m) - \delta n$ vertices. By Lemma 3.2.2, for every fixed B there is some $\beta = \beta(\delta, \ell, B) > 0$ such that if (ii) $M_k^B(m) = N_{\geqslant k}(m) - N_{\geqslant Bk}(m) \geqslant \delta n$, then it is very likely that (iii) $M_k^B(m') \geqslant \beta n$ for all $m \leqslant m' \leqslant m + n/k$, say.

To be more precise, let $B = \lceil A\ell^2/\delta \rceil$ and $K = B^{1+\lceil 1/\beta \rceil}$. Then it follows easily from Lemma 3.2.2, Lemma 3.3.1 and the union bound that for n large enough there is a 'good' event $\mathcal{G} = \mathcal{G}_n(\delta)$ such that $\mathbb{P}(\mathcal{G}) \to 1$ and such that whenever $\mathcal{G}$ holds, then for all $m \leqslant n^2$ and $k \leqslant K$, (i) holds, and (ii) implies (iii).

Suppose that $\mathcal{G}$ holds, and that $m^+ = \min\{m : N_{\geqslant K}(m) \geqslant L_1(m) + \varepsilon n\}$ exists. It suffices to show deterministically that if n is large enough, then this leads to a contradiction. Since $\mathcal{G}$ holds, considering (i) with $m = 0$ and $k = 1$ shows that for some $C = C(\delta, \ell)$ we have $L_1(Cn) \geqslant (1-\delta)n > (1-\varepsilon)n$, so $m^+ \leqslant Cn$.

For $k \leqslant K/B$ set $m_k = m^+ - 2\delta n/(\ell^2 k)$. Recall that $V_{\geqslant k}(m)$ denotes the union of all components with size at least k in $G(m)$. Since at most $\binom{\ell}{2}(m^+ - m_k) < \delta n/k$ edges are added passing from $G(m_k)$ to $G(m^+)$, vertices outside of $V_{\geqslant k}(m_k)$ contribute at most $2\delta n$ vertices to $V_{\geqslant k}(m^+)$. Hence

$$N_{\geqslant k}(m_k) \geqslant N_{\geqslant k}(m^+) - 2\delta n \geqslant N_{\geqslant K}(m^+) - 2\delta n \geqslant L_1(m^+) + (\varepsilon - 2\delta)n.$$

Suppose that $N_{\geqslant Bk}(m_k) \geqslant N_{\geqslant k}(m_k) - \delta n$. Then (i) (with Bk in place of k) tells us that by step

$$m = m_k + \lfloor An/(Bk) \rfloor \leqslant m_k + \delta n/(\ell^2 k) = m^+ - \delta n/(\ell^2 k) < m^+$$

there exists a component of $G(m)$ containing at least

$$N_{\geqslant Bk}(m_k) - \delta n \geqslant N_{\geqslant k}(m_k) - 2\delta n \geqslant L_1(m^+) + (\varepsilon - 4\delta)n > L_1(m^+)$$

vertices, which contradicts $G(m^+) \supseteq G(m)$. It follows that $M_k^B(m_k) \geqslant \delta n$. Using (ii) implies (iii) we deduce that $M_k^B(m^+) \geqslant \beta n$. Applying this for $k = 1, B, B^2, \ldots, B^{\lceil 1/\beta \rceil}$ and counting vertices in $G(m^+)$ gives a contradiction. $\square$

Working through the conditions on the constants in the proof above, and using $D = 3\delta/\ell^2$ instead of $D = 1$ when applying Lemma 3.2.2, one can check that for some positive constants c and d depending only on ℓ the result holds for any $\varepsilon = \varepsilon(n) \geqslant d/(\log\log n)^{1/(\ell-1)}$, with $K = K(\varepsilon) \leqslant \exp(\exp(c\varepsilon^{-(\ell-1)}))$.

Theorem 3.3.2. *Let $\mathcal{R}$ be a merging ℓ-vertex rule. For each n, let $(G(m))_{m\geqslant 0} = (G_n^{\mathcal{R}}(m))_{m\geqslant 0}$ be the random sequence of graphs on $\{1, 2, \ldots, n\}$ associated to $\mathcal{R}$. Given any function $h_m(n)$ that is $o(n)$, and any constants $0 \leqslant a < b$, the probability that there exist m_1 and m_2 with $L_1(G(m_1)) \leqslant an$, $L_1(G(m_2)) \geqslant bn$ and $m_2 \leqslant m_1 + h_m(n)$ tends to 0 as $n \to \infty$.*

Note that for merging rules, Theorem 3.3.2 implies the conclusion of Theorem 3.1.1: a 'jump' from $o(n)$ to $\geqslant \delta n$ implies a 'jump' from $\leqslant \delta n/2$ to $\geqslant \delta n$.

Proof. Let $a < b$ be given, and set $\varepsilon = (b-a)/2$. Using Theorem 3.1.2 we may assume that there exists $K = K(\varepsilon, \ell)$ such that $N_{\geqslant K}(m) < L_1(m) + \varepsilon n$ for all m. Suppose that $m^- = \max\{m : L_1(m) \leqslant an\}$ and $m^+ = \min\{m : L_1(m) \geqslant bn\}$ differ by at most $h_m(n)$. Set $m^* = m^+ - \varepsilon n/(2\ell^2 K)$. As before, we have

$$N_{\geqslant K}(m^*) \geqslant L_1(m^+) - \ell^2 K(m^+ - m^*) > (b-\varepsilon)n = (a+\varepsilon)n. \tag{3.2}$$

If n is large enough, which we assume, then $m^+ \leqslant m^- + h_m(n)$ implies $m^* < m^-$. This gives $N_{\geqslant K}(m^*) \leqslant N_{\geqslant K}(m^-) < L_1(m^-) + \varepsilon n \leqslant (a+\varepsilon)n$, contradicting (3.2). □

Let us remark that Theorem 3.3.2 (which can be proved without first proving Theorem 3.1.2) gives an alternative proof of Spencer's 'no two giants' conjecture: if at any time there are two components with at least εn vertices, then in the step after the last such time, L_1 must increase by at least εn in a single step. Hence Theorem 3.3.2 implies that if $\mathcal{R}$ is merging, then for any $\varepsilon > 0$ we have $\max_m L_2(m) \leqslant \varepsilon n$ whp.

Corollary 3.3.3. *Let $\mathcal{R}$ be a merging ℓ-vertex rule. If $\mathcal{R}$ is globally convergent, then $\rho^{\mathcal{R}}$ is continuous on $[0, \infty)$.*

Proof. Let $\rho(t) = \rho^{\mathcal{R}}(t)$. We have $0 \leqslant \rho(t) \leqslant \binom{\ell}{2}t$, so ρ is continuous at 0. Suppose ρ is discontinuous at some $t > 0$. Since ρ is increasing, $\sup_{t'<t} \rho(t') < \inf_{t'>t} \rho(t')$, so we may pick $a < b$ with $\sup_{t'<t} \rho(t') < a < b < \inf_{t'>t} \rho(t')$. By definition of global convergence, for any fixed $\varepsilon > 0$,

$$\mathbb{P}\Big(L_1(\lfloor (t-\varepsilon)n \rfloor) \leqslant an \text{ and } L_1(\lfloor (t+\varepsilon)n \rfloor) \geqslant bn\Big) \geqslant 1 - \varepsilon \tag{3.3}$$

if n is large enough. It follows as usual that there is some $\varepsilon(n) \to 0$ such that (3.3) holds with $\varepsilon = \varepsilon(n)$. But this contradicts Theorem 3.3.2. □

3.4 Convergence considerations

From the beginning, a key question about Achlioptas processes has been which rules are globally convergent. In some cases, local convergence has been established, but as far as we are aware, global convergence has not been shown for any non-trivial rules.

We now turn to the proof of Theorem 3.1.3, that local convergence implies global convergence for merging rules (in particular, for Achlioptas rules). We comment further on local convergence below. Theorem 3.1.3 is easy to deduce from Theorem 3.1.2; we shall give a more direct proof that seems more informative.

Proof of Theorem 3.1.3. Suppose $\mathcal{R}$ is locally convergent. Then there exist functions $\rho_k : [0,\infty) \to [0,1]$ such that (3.1) holds for any fixed $k \geqslant 1$ and $t \geqslant 0$. Since N_k changes by at most $2k$ when an edge is added to a graph, it follows easily that each ρ_k is continuous (indeed Lipschitz). From monotonicity of the underlying process, it is easy to see that for each k the function $\rho_{\leqslant k}(t) = \sum_{j\leqslant k} \rho_j(t)$ is decreasing.

Define $\rho = \rho^{\mathcal{R}}$ by

$$\rho(t) = 1 - \sum_{k=1}^{\infty} \rho_k(t) = 1 - \lim_{k\to\infty} \rho_{\leqslant k}(t),$$

so $\rho : [0,\infty) \to [0,1]$ is increasing. We claim that for any fixed $t > 0$ and $\varepsilon > 0$, the probability that

$$\sup_{0\leqslant t'<t} \rho(t') - \varepsilon \leqslant \frac{L_1(\lfloor tn \rfloor)}{n} \leqslant \rho(t) + \varepsilon \tag{3.4}$$

tends to 1 as $n \to \infty$. This clearly implies that $L_1(\lfloor tn \rfloor)/n \xrightarrow{\mathrm{p}} \rho(t)$ whenever ρ is continuous at t, which is the definition of global convergence. Corollary 3.3.3 then implies that ρ is continuous.

The upper bound in (3.4) is immediate: by definition of ρ there is some K such that $\rho_{\leqslant K}(t) \geqslant 1 - \rho(t) - \varepsilon/4$. Summing (3.1) up to K gives $N_{\leqslant K}(\lfloor tn \rfloor)/n \geqslant 1 - \rho(t) - \varepsilon/2$ whp. When n is large enough, this bound implies $L_1(\lfloor tn \rfloor)/n \leqslant \rho(t) + \varepsilon$.

For the lower bound, we combine the 'sprinkling' argument of Erdős and Rényi [33] with Lemma 3.3.1. Choose $t' < t$ such that $\rho(t')$ is within $\varepsilon/2$ of the supremum, and let $m_1 = \lfloor t'n \rfloor$ and $m_2 = \lfloor tn \rfloor$, so $m_2 - m_1 = \Theta(n)$. It suffices to show that $L_1(m_2)/n \geqslant \rho(t') - \varepsilon/2$ holds whp. In doing so we may assume that $\rho(t') - \varepsilon/2 \geqslant 0$. For any constant K, whp we have $N_{\leqslant K}(m_1)/n \leqslant \rho_{\leqslant K}(t') + \varepsilon/4 \leqslant 1 - \rho(t') + \varepsilon/4$, so $N_{\geqslant K}(m_1)/n \geqslant \rho(t') - \varepsilon/4$ whp. If K is large enough (depending only on t' and ε), Lemma 3.3.1 then gives $L_1(m_2)/n \geqslant \rho(t') - \varepsilon/2$ whp, as required. □

Remark 3.4.1. *Since non-merging ℓ-vertex rules have received some attention (see e.g., [56]), let us spell out what our method gives for such rules. Lemma 3.3.1 applies in this case provided 'there is a component containing' is changed to 'there are $\ell - 1$ components together containing'. Let $L(m)$ denote the sum of the sizes of the $\ell - 1$ largest components in $G(m)$. With this modified Lemma 3.3.1, the proof of Theorem 3.1.2 goes through with L_1 replaced by L. The same is true of Theorem 3.3.2 (with an extra $-(\ell - 2)K$ in* (3.2)*, since the largest $\ell - 1$ components may not all be large). Finally, Corollary 3.3.3 and Theorem 3.1.3 similarly go through, now with ρ defined using L rather than L_1.*

3.5 Size rules

So far, even in the Achlioptas-rule case our rules have been very general, making choices between the given edges using any information about the current graph. There is a natural much smaller class (of vertex or Achlioptas rules) called *size rules*, where only the sequence $c_1, \ldots, c_\ell$ of the orders of the components containing the presented vertices $v_1, \ldots, v_\ell$ may be used to decide which edge(s) to add. (Here we suppress the dependence on the step m in the notation.) Note that the product rule is a size rule.

In fact, most past results concern *bounded-size rules*: here there is a constant B such that all sizes $c_i > B$ are treated the same way by the rule, so the rule only 'sees' the data $(\min\{c_i, B + 1\})_{i=1}^{\ell}$. Perhaps the simplest example is the 'Bohman–Frieze process', the bounded-size rule with $B = 1$ in which the edge v_1v_2 is added if $c_1 = c_2 = 1$, and otherwise v_3v_4 is added. Bohman and Frieze [14] showed that for a closely related rule there is no giant component when $m \sim 0.535n$. (The actual rule they used considered whether v_1 and

v_2 are isolated in the graph formed by all pairs *offered* to the rule, rather than the graph $G(m)$ formed by the pairs *accepted* so far.)

Considering for simplicity rules in which one edge is added at each step, a key property of bounded-size rules is that at each step, the expected change in N_k can be expressed as a simple function of $N_1, N_2, \ldots, N_{\max\{k,B\}}$. (It is clear that the rate of formation of k-vertex components can be so expressed; for the rate of destruction, consider separately the cases k joins to k' for each $k' \leqslant B$ and the case k joins to some $k' > B$.) Spencer and Wormald [74], who considered bounded-size Achlioptas rules, and Bohman and Kravitz [19], who considered a large subset of such rules, noted that in this case one can easily use Wormald's 'differential equation method' [81] to show that the rule is locally convergent, and that the $\rho_k(t)$ satisfy certain differential equations. This remark applies to all bounded-size ℓ-vertex rules.

Resolving a conjecture of Spencer [72], Spencer and Wormald [74] proved that any bounded-size 2-Achlioptas rule exhibits a phase transition: there is some t_c, depending on the rule, such that for $t < t_c$, whp $L_1(\lfloor tn \rfloor) = o(n)$ (in fact $O(\log n)$), while for $t > t_c$, $L_1(\lfloor tn \rfloor) = \Omega(n)$ whp. They conjectured that any bounded-size 2-Achlioptas rule is globally convergent, and that the phase transition is 'second order' (continuous). Theorem 3.1.3 establishes both these conjectures.

Very recently, Janson and Spencer [45] established bounds on the size of the giant component in the Bohman–Frieze process just above the (known) critical point t_c. They deduce that *if* it is globally convergent, then the right derivative of ρ at t_c has a certain specific value. The required 'if' part is established by Theorem 3.1.3.

Informally, let us call a size rule *nice* if there is some K such that, for each k, the expected change in N_k is a function of $N_1, N_2, \ldots, N_{\max\{k,K\}}$. (More precisely, the individual decisions whether to create or destroy a component of size k depend only on the data $(\min\{c_i, k'+1\})_{i=1}^{\ell}$ where $k' = \max\{k, K\}$ and c_i is the size of the component containing v_i.) Just as in the bounded-size case, using the differential equation method, it is easy to show that any nice rule is locally convergent. Hence, by Theorem 3.1.3, any nice *merging* rule is globally convergent with continuous phase transition; this applies to all *nice* Achlioptas rules.

The simplest examples of nice rules have $K = 1$, i.e., only compare component sizes. One example is 'join the two smallest'. For $\ell = 3$ this rule is mentioned briefly by Friedman and Landsberg [36] as another example of a rule that should be explosive, and discussed by D'Souza and Mitzenmacher [32], who 'established' the explosive nature of the transition for this and a related nice rule numerically; Theorem 3.1.1 contradicts these predictions.

Another nice rule is the following: join the smaller of C_1 and C_2 to the smaller of C_3 and C_4, where C_i is the component containing v_i. We call this the 'dCDGM' rule since it was introduced by da Costa, Dorogovtsev, Goltsev and Mendes [31]. Note that this is *not* an Achlioptas rule, but it *is* merging: if $|C|, |C'| \geqslant \varepsilon n$ then with probability at least ε^4 we choose $v_1, v_2 \in C$ and $v_3, v_4 \in C'$ and so join C to C'. Hence, the dCDGM rule, which is locally convergent by the differential equation method, is globally convergent and has a continuous phase transition. Da Costa, Dorogovtsev, Goltsev and Mendes [31] proposed this rule as simpler to analyze than the product rule, but at least as likely to have a discontinuous phase transition. For a brief discussion of their arguments, see the end of the introduction.

There are many open questions concerning the precise nature of the phase transitions in various Achlioptas and related processes. One of the most intriguing is the following: is the product rule globally convergent?

3.6 Explosive percolation when $\ell \to \infty$

In this final section we will allow ℓ, the number of random vertices sampled in each step, to depend on n. We focus on the *SDC rule*, which in each step picks ℓ random vertices and joins the two smallest distinct components. We prove that the SDC rule exhibits explosive percolation whenever $\ell \to \infty$ as $n \to \infty$, i.e., that the transition from $L_1 = o(n)$ to $L_1 = \Omega(n)$ happens in $o(n)$ steps. The next result also demonstrates that the restriction to fixed ℓ in Theorem 3.1.1 is necessary.

Theorem 3.6.1. *For the SDC rule there exists $C > 0$ such that for every $n \geqslant n_0$ and*

$$\varepsilon \geqslant \max\left\{\sqrt{\frac{C\log n}{n}}, \frac{C\log \ell}{\ell}\right\}$$

the following holds with probability at least $1 - n^{-1}$*: the number of steps between the last time* $L_1 \leqslant \varepsilon n$ *and the first time* $L_1 \geqslant (1-\varepsilon)n$ *is at most* $9\varepsilon n$*. Hence, the SDC rule exhibits explosive percolation if* $\ell \to \infty$ *as* $n \to \infty$*.*

The proof is based on two probabilistic statements (which are not best possible, but suffice for our purposes). The first intuitively states that if 'most' vertices are in components of size at least $1/\varepsilon$, then after at most $8\varepsilon n$ subsequent steps the giant component contains almost all vertices (since $2^{-1/(\ell-1)} \to 1$ as $\ell \to \infty$).

Lemma 3.6.2. *Let* $\mathcal{C}_\ell(\varepsilon)$ *denote the event that for all* $0 \leqslant m \leqslant n^2$ *and* $1/\varepsilon \leqslant k \leqslant n$ *the following holds:* $N_{\leqslant k}(m) \leqslant \varepsilon n$ *implies* $L_1(m+\Delta) > 2^{-1/(\ell-1)}n$ *for* $\Delta = \lceil 8\varepsilon n \rceil$*. If* $\varepsilon \geqslant 10(\log n)/n$*, then for the SDC rule we have* $\mathbb{P}(\mathcal{C}_\ell(\varepsilon)) \geqslant 1 - n^{-2}$*.*

Proof. Set $\beta = 2^{-1/(\ell-1)}$. Similar as in the proofs of Lemmas 3.2.1 and 3.3.1, we consider fixed m and k and show that, conditional on $\mathcal{F}_m$, if $G(m)$ satisfies $N_{\leqslant k}(m) \leqslant \varepsilon n$, then we have $L_1(m+\Delta) \leqslant \beta n$ with probability at most n^{-5}.

Condition on $\mathcal{F}_m$ and, starting from $G(m)$, consider the next Δ steps. We say that a step is *good* if at least two of the ℓ randomly chosen vertices are in different components, and let X_j denote the indicator function of the event that step $m+j$ is good. Set $X = \sum_{1 \leqslant j \leqslant \Delta} X_j$ and $Y = \sum_{1 \leqslant j \leqslant \Delta} Y_j$, where

$$Y_j = \begin{cases} X_j, & \text{if } L_1(m+j-1) \leqslant \beta n, \\ 1, & \text{otherwise.} \end{cases}$$

Whenever $L_1(m+j-1) \leqslant \beta n$ holds, the probability that step $m+j$ is good is at least $1 - \beta^{\ell-1} = 1/2$. Since otherwise $Y_j = 1$ by definition, we deduce that Y stochastically dominates a binomial random variable with mean $\Delta/2 \geqslant 4\varepsilon n$. Using $\varepsilon \geqslant 10(\log n)/n$ standard Chernoff bounds now imply that $\mathbb{P}(Y \leqslant 2\varepsilon n) \leqslant e^{-\varepsilon n/2} \leqslant n^{-5}$.

Assume that $L_1(m+\Delta) \leqslant \beta n$. Then by monotonicity $L_1(m+j-1) \leqslant \beta n$ for every $1 \leqslant j \leqslant \Delta$, so the number X of good steps satisfies $X = Y$. Since $N_{\leqslant k}(m) \leqslant \varepsilon n$ holds, note that $G(m)$ contains at most $n/k + \varepsilon n \leqslant 2\varepsilon n$ components. Since every good step joins two different components (the SDC rule always joins two components unless all ℓ vertices are in the same component), we deduce that $Y \leqslant 2\varepsilon n$. Hence $\mathbb{P}(L_1(m+\Delta) \leqslant \beta n) \leqslant \mathbb{P}(Y \leqslant 2\varepsilon n) \leqslant n^{-5}$, as required. □

The second technical lemma essentially implies that, throughout its initial evolution, the SDC rule ensures that most vertices are in components of similar size (considering the minimal k with $N_{\leqslant k}(m) \geqslant \varepsilon n$, the event $\mathcal{I}_\ell(\varepsilon)$ implies that all but at most $2\varepsilon n$ vertices are in components of size between k and $2k$).

Lemma 3.6.3. *Let $\mathcal{I}_\ell(\varepsilon)$ denote the event that for all $0 \leqslant m \leqslant 9n$ and $1 \leqslant k \leqslant \min\{2/\varepsilon, \varepsilon n/(120 \log n)\}$ the following holds: $N_{\leqslant k}(m) \geqslant \varepsilon n$ implies $N_{\geqslant 2k+1}(m) < \varepsilon n$. If $\varepsilon \geqslant 2\log(72\ell^3)/(\ell-1)$, then for the SDC rule we have $\mathbb{P}(\mathcal{I}_\ell(\varepsilon)) \geqslant 1 - n^{-2}$ for $n \geqslant n_0$.*

Proof. Clearly it suffices to consider the case $\varepsilon \leqslant 1$. We consider fixed m and k and show that the (unconditional) probability that $N_{\leqslant k}(m) \geqslant \varepsilon n$ and $N_{\leqslant 2k+1}(m) \geqslant \varepsilon n$ hold is at most n^{-5}.

Consider the first m steps of the process. We say that a step is *bad* if the SDC rule joins two (distinct) components C_1 and C_2 with $\min\{|C_1|, |C_2|\} \leqslant 2k$ and $|C_1| + |C_2| \geqslant 2k+1$. Let X_j denote the indicator function of the event that step j is bad. Set $X = \sum_{1 \leqslant j \leqslant m} X_j$ and $Y = \sum_{1 \leqslant j \leqslant m} Y_j$, where

$$Y_j = \begin{cases} X_j, & \text{if } N_{\leqslant k}(j-1) \geqslant \varepsilon n, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $|C_1| + |C_2| \geqslant 2k+1$ implies $\max\{|C_1|, |C_2|\} \geqslant k+1$. Note that if step j is bad, then there exists a component C (the smallest selected) of size at most $2k$ such that all ℓ vertices are either in C or in components of size at least k. Using that $2k/n \leqslant \varepsilon/2$ this has probability at most

$$\ell \left(\frac{N_{\geqslant k+1}(j-1) + 2k}{n} \right)^{\ell-1} \leqslant \ell \left(\frac{N_{\geqslant k+1}(j-1)}{n} + \frac{\varepsilon}{2} \right)^{\ell-1}.$$

Whenever $N_{\leqslant k}(j-1) \geqslant \varepsilon n$ holds we have $N_{\geqslant k+1}(j-1) \leqslant (1-\varepsilon)n$; so in this case, using

$\varepsilon(\ell-1)/2 \geqslant \log(72\ell^3)$ the probability that step j is bad is at most

$$\ell\left(\frac{N_{\geqslant k+1}(j-1)}{n}+\frac{\varepsilon}{2}\right)^{\ell-1} \leqslant \ell\,(1-\varepsilon/2)^{\ell-1} \leqslant \ell e^{-\varepsilon(\ell-1)/2} \leqslant \frac{1}{72\ell^2} \leqslant \frac{\varepsilon}{72k},$$

where the last inequality follows by noting that $\varepsilon\ell^2/k \geqslant \varepsilon^2\ell^2/2 \geqslant 1$. Since otherwise $Y_j = 0$ by definition, we deduce that Y is stochastically dominated by a binomial random variable with mean $m \cdot \varepsilon/(72k) \leqslant \varepsilon n/(8k)$. Using $k \leqslant \varepsilon n/(120 \log n)$ standard Chernoff bounds now imply that $\mathbb{P}(Y \geqslant \varepsilon n/(4k)) \leqslant e^{-\varepsilon n/(24k)} \leqslant n^{-5}$.

Suppose that $N_{\leqslant k}(m) \geqslant \varepsilon n$. Then by monotonicity $N_{\leqslant k}(j-1) \geqslant \varepsilon n$ for every $1 \leqslant j \leqslant m$, so $X = Y$. Clearly $N_{\geqslant 2k+1}(0) = 0$. Note that $N_{\geqslant 2k+1}$ only changes in bad steps, each time increasing by at most $4k$. Therefore $N_{\geqslant 2k+1}(m) \leqslant 4k \cdot Y$. Hence $\mathbb{P}(N_{\leqslant k}(m) \geqslant \varepsilon n \wedge N_{\geqslant 2k+1}(m) \geqslant \varepsilon n) \leqslant \mathbb{P}(Y \geqslant \varepsilon n/(4k)) \leqslant n^{-5}$, as required. $\square$

After these preparations we are now in the position to show that the SDC rule undergoes a discontinuous transition if $\ell \to \infty$ as $n \to \infty$.

Proof of Theorem 3.6.1. We may assume that $\varepsilon < 1$ holds (otherwise the claim is trivial). Furthermore, it suffices to consider

$$\varepsilon = \max\left\{\sqrt{\frac{240 \log n}{n}}, \frac{10 \log n}{n}, \frac{2\log(72\ell^3)}{\ell-1}, \frac{3 \log 2}{\ell-1}\right\}. \tag{3.5}$$

We define the 'good' event as $\mathcal{G} = \mathcal{C}_\ell(1) \wedge \mathcal{C}_\ell(\varepsilon) \wedge \mathcal{I}_\ell(\varepsilon)$. Using Lemmas 3.6.2 and 3.6.3 we have $\mathbb{P}(\mathcal{G}) \geqslant 1 - 3n^{-2} \geqslant 1 - n^{-1}$ for $n \geqslant n_0$. Note that (3.5) implies $2/\varepsilon \leqslant \varepsilon n/(120 \log n)$, so that $\mathcal{I}_\ell(\varepsilon)$ of Lemma 3.6.3 applies to $k = \lceil 1/\varepsilon \rceil$.

Suppose that $\mathcal{G}$ holds. Set $m^- = \max\{m : N_{\leqslant k}(m) \geqslant \varepsilon n\}$. Let $\Delta = \lceil 8\varepsilon n \rceil$ and $m^+ = m^- + 1 + \Delta$. To complete the proof it suffices to show deterministically that $L_1(m^-) \leqslant \varepsilon n$ and $L_1(m^+) \geqslant 2^{-1/(\ell-1)} n$ since $2^{-1/(\ell-1)} \geqslant 1 - (\log 2)/(\ell-1) > 1 - \varepsilon/2$ by (3.5).

By $\mathcal{C}_\ell(1)$ we have $L_1(8n) \geqslant 2^{-1/(\ell-1)} n \geqslant k+1$, where the last inequality follows from (3.5) by noting that $2^{-1/(\ell-1)} n > n/2 \geqslant 4/\varepsilon \geqslant 2k$. From this we readily infer $N_{\geqslant k+1}(8n) \geqslant L_1(8n) \geqslant 2^{-1/(\ell-1)} n > (1-\varepsilon)n$. Hence $m^- < 8n$ and so $\mathcal{I}_\ell(\varepsilon)$ gives $N_{\geqslant 2k+1}(m^-) \leqslant \varepsilon n$. Thus, using $2k+1 \leqslant 5/\varepsilon \leqslant \varepsilon n$, we obtain $L_1(m^-) \leqslant \max\{2k+1, N_{\geqslant 2k+1}(m^-)\} \leqslant \varepsilon n$, as claimed. Furthermore, by definition $N_{\leqslant k}(m^- + 1) < \varepsilon n$, and so $\mathcal{C}_\ell(\varepsilon)$ gives $L_1(m^+) \geqslant 2^{-1/(\ell-1)}n$, as required. $\square$

Chapter 4

Convergence via differential equations with unique solutions

4.1 Main result

In this chapter we address the convergence question for Achlioptas processes. As we have seen so far, much research has been devoted to understanding the evolution of the largest component in various Achlioptas processes. However, most of the previous work regarding explosive percolation (in particular in the physics literature, see e.g. [29, 31, 37, 51, 60, 61]) takes an important question for granted: does the deterministic scaling limit $\rho(t) = \lim_{n\to\infty} L_1(tn)/n$ even exist? More precisely, since the size of the largest component is random, the question is whether a rule $\mathcal{R}$ is *globally convergent*, i.e., if there exists an increasing function $\rho = \rho^{\mathcal{R}} : [0,\infty) \to [0,1]$, called the scaling limit of $\mathcal{R}$, such that for any t at which ρ is continuous we have $L_1(\lfloor tn \rfloor)/n \xrightarrow{\text{p}} \rho(t)$ as $n \to \infty$. In Chapter 3 we showed that global convergence follows if the rule $\mathcal{R}$ is *locally convergent*, i.e., if there exist functions $\rho_k = \rho_k^{\mathcal{R}} : [0,\infty) \to [0,1]$ such that, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have $N_k(\lfloor tn \rfloor)/n \xrightarrow{\text{p}} \rho_k(t)$ as $n \to \infty$. However, as discussed in Section 3.5 convergence is essentially only known for bounded-size rules; for 'unbounded'-size rules (such as the product rule) it remains an open question whether the corresponding scaling limits exist. In fact, according to Achlioptas, D'Souza and Spencer [1] these more involved rules even 'seem beyond the reach of current mathematical techniques'; Janson [41] also remarks that most likely new methods are needed for understanding the detailed behaviour of such rules.

Our main result shows that complex size rules such as the product rule are globally convergent provided that a certain associated system of differential equations (defined in Section 4.2.2) has a unique solution. In fact, our result applies to a very large class of Achlioptas-like processes, including essentially all Achlioptas processes studied so far. For the formal definitions of ℓ-vertex rule, merging and well-behaved see Section 4.2.

Theorem 4.1.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be a merging ℓ-vertex rule that is well behaved. Suppose the associated system of differential equations given by* (4.4)–(4.6) *has a unique solution $(\hat{\rho}_k(t))_{k\geqslant 1}$. Then $\mathcal{R}$ is locally and globally convergent. In particular, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have*

$$N_k(\lfloor tn \rfloor)/n \xrightarrow{\mathrm{p}} \hat{\rho}_k(t) \tag{4.1}$$

as $n \to \infty$. The scaling limit $\rho^{\mathcal{R}}$ is continuous and satisfies $\rho^{\mathcal{R}}(t) = 1 - \sum_{k\geqslant 1} \hat{\rho}_k(t)$.

The merging assumption seems needed in Theorem 4.1.1, since in [64] examples of 'natural' non-merging rules are given where in simulations $L_1(\lfloor tn \rfloor)/n$ does not converge (see also Figure 4.1), i.e., which are presumably *not* globally convergent. All rules for which convergence has been established are merging and well-behaved, including the classical Erdős–Rényi case [33], all bounded size rules [74] (such as the Bohman–Frieze rule [45]) as well as the dCDGM rule [31] and adjacent edge rule [32]. In fact, for all such rules there is a $K \geqslant 1$ such that each ρ'_k in (4.5) can be written as a function of $\rho_1, \ldots, \rho_{\max\{k,K\}}$. In this case the form of the differential equations (4.4)–(4.6) implies by standard results that its solution is unique. So Theorem 4.1.1 generalizes these previous convergence results.

The main contribution of this chapter is a new approach for proving convergence. Previous results in this area apply Wormald's 'differential equation method' [80, 81], which is nowadays widely used in probabilistic combinatorics. This shows that under certain conditions, suitable sequences of random variables converge to the solution of a system of differential equations. The key point is that Wormald's conditions imply that the differential equations have a unique solution, but are not implied by this. By establishing a more direct connection between the random process and the differential equations, we only need

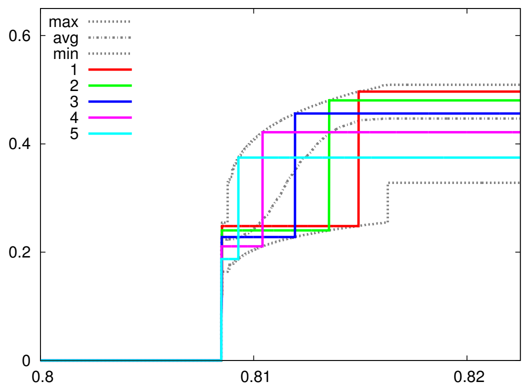

Figure 4.1: Simulation of $L_1(tn)/n$ as a function of t using the non-merging NG rule [64]: five sample runs together with min, max and average of 10^3 runs (with $n = 10^{11}$). On the basis of heuristics and simulations presented in [64] we believe that the size of the emerging giant component is not concentrated (so the NG rule is presumably not globally convergent).

to assume that the system of differential equations has a unique solution. Thus, our method is potentially applicable to a much larger class of Achlioptas processes. The general proof idea outlined in Section 4.3 might also be useful to establish convergence in other stochastic processes.

We see our main result as a first step towards resolving the convergence question in Achlioptas processes. In particular, further investigation of the system of differential equations (4.4)–(4.6) associated to the product rule (and other involved rules) seems to be needed: does it have a unique solution? When the equations do have a unique solution many questions remain, for example which conditions are needed to establish asymptotic normality as in [69].

In the next section we define the processes under consideration and state the system of differential equations associated to these. In Section 4.3 we first outline a general idea for proving convergence in stochastic processes, and then use this approach to establish our main result.

4.2 Preliminaries and notation

Our core argument will involve considering sequences of points ω_n in different probability spaces. For this reason we indicate the dependence on n explicitly in the notation. We now recall the relevant definitions from Chapter 3. Fix $\ell \geqslant 2$. Each ℓ-vertex rule $\mathcal{R}$ yields for each n a random sequence $(G_{n,m})_{m\geqslant 0}$ of graphs with vertex set $[n] = \{1, \ldots, n\}$, where $G_{n,0}$ is the empty graph. For each $m \geqslant 0$ we draw ℓ vertices $\underline{v}_{n,m+1} = (v_1, \ldots, v_\ell)$ from $[n]$ independently and uniformly at random, and then obtain $G_{n,m+1}$ by adding a (possibly empty) set of edges $E_{n,m+1}$ to $G_{n,m}$, where $\mathcal{R}$ selects $E_{n,m+1}$ as a subset of all pairs between vertices in $\underline{v}_{n,m+1}$. To avoid 'trivial' rules we require that $E_{n,m+1} \neq \emptyset$ if all ℓ vertices in $\underline{v}_{n,m+1}$ are in distinct components of $G_{n,m}$. Formally, we assume the existence of a sample space Ω_n and a filtration $\mathcal{F}_{n,0} \subseteq \mathcal{F}_{n,1} \subseteq \cdots$ such that $\underline{v}_{n,m+1}$ is $\mathcal{F}_{n,m+1}$-measurable and independent of $\mathcal{F}_{n,m}$, and require $E_{n,m+1}$ (and hence $G_{n,m+1}$) to be $\mathcal{F}_{n,m+1}$-measurable. For later usage we let $\underline{c}_{n,m+1} = (c_1, \ldots, c_\ell)$ denote the sizes of the components containing the chosen vertices $\underline{v}_{n,m+1} = (v_1, \ldots, v_\ell)$ in $G_{n,m}$. We write $N_{n,k,m}$ for the number of vertices of $G_{n,m}$ in components of size k, and let $N_{n,\leqslant k,m} = \sum_{1\leqslant j\leqslant k} N_{n,j,m}$. We define $N_{n,\geqslant k,m}$ in an analogous way.

For the purposes of this chapter these definitions are robust with respect to small changes, since our arguments have $o(1)$ elbow room in each step of the process. So we may weaken the conditions on $\underline{v}_{n,m+1}$: it suffices if, for $m = O(n)$, say, the conditional distribution of $\underline{v}_{n,m+1}$ given $\mathcal{F}_{n,m}$ is close to (at total variation distance $\alpha_n = o(1)$ from) the one defined above. This includes variations such as picking an ℓ-tuple of distinct vertices, or picking (the ends of) $\ell/2$ randomly selected (distinct) edges not already present in $G_{n,m}$, see Section 3.2. Hence we may treat the original examples of Achlioptas as 4-vertex rules where $\mathcal{R}$ always selects one of the pairs $\{v_1, v_2\}, \{v_3, v_4\}$; below we call such $\mathcal{R}$ *Achlioptas rules.*

We say that an ℓ-vertex rule is *merging* if whenever C, C' are distinct components with $|C|, |C'| \geqslant \varepsilon n$, then in the next step we have probability at least ε^ℓ of joining C to C' (this can be slightly weakened, see Section 3.3). In particular all Achlioptas rules are merging, since with probability at least ε^4 both potential pairs join C to C'.

4.2.1 Well behaved rules

We say that an ℓ-vertex rule $\mathcal{R}$ is *well behaved (at infinity)* if there are functions d_k and g such that the following conditions hold:

1. Whenever all vertices v_j are in different components we have

$$\mathbb{E}(N_{n,k,m+1} - N_{n,k,m} \mid \mathcal{F}_{n,m}, \underline{v}_{n,m+1}) = d_k(c_1, \ldots, c_\ell), \tag{4.2}$$

   where $\underline{c}_{n,m+1} = (c_1, \ldots, c_\ell)$ lists the sizes of the components containing the selected vertices.

2. Suppose there are $L \subseteq [\ell]$ and $S \geqslant k$ such that all v_j with $j \in L$ are in the same component of size $c_j > g(S)$, whereas all other vertices are in different components with sizes $c_j \leqslant S$. Whenever this holds we have

$$\mathbb{E}(N_{n,k,m+1} - N_{n,k,m} \mid \mathcal{F}_{n,m}, \underline{v}_{n,m+1}) = d_k(\tilde{c}_1, \ldots, \tilde{c}_\ell), \tag{4.3}$$

   where $\tilde{c}_j = \infty$ for $j \in L$ and $\tilde{c}_j = c_j$ otherwise.

In fact, taking $L = \emptyset$ in (4.3) gives (4.2), but we note (4.2) separately for clarity. As we shall discuss below, these conditions are very mild and hold for essentially all Achlioptas processes previously studied, including 'unbounded rules' such as the sum and product rules. All rules which have been considered so far are *size-rules*, which only use $\underline{c}_{n,m+1}$ to decide which edge(s) are added. For these the change of $N_{n,k,m}$ in (4.2) is deterministic given $\underline{c}_{n,m+1}$, but considering the conditional expected change is slightly more general (we can also allow for small deviations in (4.2) and (4.3), but leave this to the interested reader). Intuitively, the second condition ensures that whenever one component is significantly larger than all others, then we can decide which relevant pairs are joined *without* knowing its exact size (this fails, for example, if the change depends on the parity of $\max_{j\in[\ell]} c_j$). This mild assumption holds for a large class of rules; for example, $g(s) = \max\{K, s\}$, $g(s) = \max\{B, s\}$, $g(s) = s^2$ and $g(s) = 2s$ suffice for nice rules as defined in Section 3.5, bounded-size rules, the product rule and the sum rule, respectively. Note that since $N_{n,k,m}$ always changes by at most ℓk per step, we have $|d_k(\cdot)| \leqslant \ell k$.

4.2.2 An associated system of differential equations

Suppose that $\mathcal{R}$ is a well-behaved ℓ-vertex rule. In the following equations, each $\rho_k(t)$ is a function on $[0,\infty)$ satisfying

$$0 \leqslant \rho_k(t) \leqslant 1 \qquad \text{and} \qquad 0 \leqslant \sum_{k \geqslant 1} \rho_k(t) \leqslant 1. \tag{4.4}$$

The system of differential equations associated to $\mathcal{R}$ is given by

$$\rho_k'(t) = \sum_{c_1,\ldots,c_\ell \in \mathbb{N}\cup\{\infty\}} d_k(c_1,\ldots,c_\ell) \prod_{j\in[\ell]} \rho_{c_j}(t) \tag{4.5}$$

for all $k \geqslant 1$, where

$$\rho(t) = \rho_\infty(t) = 1 - \sum_{k\geqslant 1} \rho_k(t),$$

together with the initial conditions

$$\rho_k(0) \;=\; \begin{cases} 1, & \text{if } k = 1,\\ 0, & \text{otherwise.} \end{cases} \tag{4.6}$$

For $t = 0$, the derivative in (4.5) is taken to be the right-derivative. Note that for all $t \geqslant 0$ we have $|\rho_k'(t)| \leqslant \max_{\underline{c}} |d_k(\underline{c})| \leqslant \ell k$.

As a basic example, consider the Erdős–Rényi random graph process, for which we have $d_k(c_1, c_2) \in \{-2k, -k, 0, k\}$. It is not difficult to see that in this case (4.5) simplifies to

$$\rho_k'(t) = -2k\rho_k(t) + k \sum_{c_1+c_2=k} \rho_{c_1}(t)\rho_{c_2}(t), \tag{4.7}$$

which is a special case of *Smoluchowski's coagulation equations* in a form where sol-gel interaction is considered, see e.g. [3] and the references therein. Here uniqueness follows easily from standard results, since ρ_k' depends only on $\rho_1,\ldots,\rho_k$.

4.3 Proof of the main result

We starting by outlining a rather general idea for proving convergence to the unique solution of a system of differential equations, which we shall later use to establish Theorem 4.1.1. We consider a discrete stochastic process with sample space Ω_n and filtration $\mathcal{F}_{n,0} \subseteq \mathcal{F}_{n,1} \subseteq \cdots$. For each (discrete) step m we introduce (continuous) time $t = m/s_n$, where the scaling satisfies $s_n \to \infty$ as $n \to \infty$. Suppose our objective is to find a collection of random variables $X_{n,k,m}$ and (continuous) functions x_k together with scaling parameters $S_{n,k}$ such that for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have

$$X_{n,k,ts_n}/S_{n,k} \xrightarrow{\text{p}} x_k(t)$$

as $n \to \infty$, where we ignore the rounding to integers. The two main steps of our approach are as follows:

1. Defining the one-step change as $\Delta X_{n,k,m+1} = X_{n,k,m+1} - X_{n,k,m}$, we use martingale techniques (the Azuma–Hoeffding inequality together with an absolute bound on $|\Delta X_{n,k,m+1}|$) to show that, with probability tending to 1 as $n \to \infty$, the following holds: for each fixed k and all $m_1, m_2 \geqslant 0$ with $m_2 - m_1 = O(s_n)$ we have

$$X_{n,k,m_2} - X_{n,k,m_1} = \sum_{m_1 \leqslant m < m_2} \mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m}) + o(S_{n,k}). \tag{4.8}$$

2. Suppose we are given a sequence of sample points $\omega_n \in \Omega_n$, defined for some infinite subset of $\mathbb{N}$, for which (4.8) and some additional technical conditions hold. Proceeding as in the proof of Helley's selection theorem (see e.g. Theorem 5.8.1 in [38]), we pick a subsequence $(\omega_{\tilde{n}})$ such that for each $t \geqslant 0$ and $k \geqslant 1$, for some limiting value $x_k(t)$ we have

$$X_{\tilde{n},k,ts_{\tilde{n}}}(\omega_{\tilde{n}})/S_{\tilde{n},k} \to x_k(t) \tag{4.9}$$

as $\tilde{n} \to \infty$. Here we exploit that each $X_{\tilde{n},k,ts_{\tilde{n}}}(\omega_{\tilde{n}})/S_{\tilde{n},k}$ as a function of t satisfies a Lipschitz condition. Along this subsequence, we show that for all $t \geqslant 0$, $\varepsilon > 0$ and $k \geqslant 1$ there exists $\delta > 0$ such that for $\tilde{n}$ large enough the following holds: for each

$m \geqslant 0$ satisfying $|m - ts_{\tilde{n}}| \leqslant \delta s_{\tilde{n}}$ we have

$$\mathbb{E}(\Delta X_{\tilde{n},k,m+1} \mid \mathcal{F}_{\tilde{n},m})(\omega_{\tilde{n}}) = (f_k(t) \pm \varepsilon) S_{\tilde{n},k}/s_{\tilde{n}}, \tag{4.10}$$

where $f_k(t) = f_k(t, x_1, x_2, \ldots)$ is a function of the scaling limits along the selected subsequence. To establish (4.10) we combine coupling arguments with 'typical' properties of the underlying stochastic process.

Now, using (4.8)–(4.10) it is straightforward to show that for all $t \geqslant 0$, $\varepsilon > 0$ and $k \geqslant 1$ there exists $\delta > 0$ such that for all $0 < |h| \leqslant \delta$ with $t + h \geqslant 0$ we have

$$\left| \frac{x_k(t+h) - x_k(t)}{h} - f_k(t) \right| \leqslant \varepsilon$$

along the selected subsequence, i.e., the x_k satisfy the differential equation

$$x_k'(t) = f_k(t, x_1, x_2, \ldots).$$

But if the associated system of differential equations has a unique solution, then this implies that the limiting functions $x_k(t)$ in (4.9) do *not* depend on the selected subsequence, which establishes the desired convergence. Finally, let us remark that by comparison with the underlying process we can (typically) derive additional properties of the x_k; these might help in proving uniqueness of the solution to the system of differential equations.

In the remainder we use the above approach to establish Theorem 4.1.1. Aiming at $N_{n,k,tn}/n \xrightarrow{\mathrm{p}} \rho_k(t)$, we closely follow steps one and two in Sections 4.3.1 and 4.3.2, respectively, with $X_{n,k,m} = N_{n,k,m}$, $x_k(t) = \rho_k(t)$, $S_{n,k} = n$ and $s_n = n$.

4.3.1 Proof of Theorem 4.1.1

Our proof of Theorem 4.1.1 relies on a technical lemma which requires some preparation. Set $\eta = \eta(n) = (\log\log\log n)^{-1}$, say. Let $\mathcal{U}_n$ denote the event that at every step m there is at most one component of size at least ηn. Since $\mathcal{R}$ is merging, by the discussion following Theorem 3.1.2 in Chapter 3 we know that $\mathbb{P}(\mathcal{U}_n) \to 1$ as $n \to \infty$. In the rest of this chapter the particular form of $\eta = \eta(n)$ does not matter, only that $\eta \to 0$ as $n \to \infty$. By

Theorem 3.1.2 there exist functions $K(\gamma)$ and $\xi = \xi(n)$ with $\xi \to 0$ as $n \to \infty$ such that, defining $\mathcal{K}_n$ as the event that for all $m \geqslant 0$ we have

$$\forall \gamma \geqslant \xi : \ N_{n,\geqslant K(\gamma),m} < L_{1,n,m} + \gamma n, \tag{4.11}$$

it holds that $\mathbb{P}(\mathcal{K}_n) \to 1$ as $n \to \infty$. Here $L_{1,n,m}$ is the number of vertices in the largest component of $G_{n,m}$.

Fix $0 < \lambda < 1/4$, say $\lambda = 1/8$ for concreteness. For each $m \geqslant 0$ set $\Delta X_{n,k,m+1} = N_{n,k,m+1} - N_{n,k,m}$ and $Y_{n,k,m+1} = \Delta X_{n,k,m+1} - \mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})$. Set $Z_{n,k,j} = \sum_{0 \leqslant m < j} Y_{n,k,m+1}$. Let $\mathcal{D}_n = \mathcal{D}_n(\lambda)$ denote the event that for all $1 \leqslant k \leqslant n^\lambda$ and $1 \leqslant m_1 \leqslant m_2 \leqslant n^2$ with $m_2 - m_1 \leqslant n^{1+\lambda}$ we have $|Z_{n,k,m_2} - Z_{n,k,m_1}| < n^{1/2+2\lambda}$. Note that by rearranging terms, for all such k, m_1, m_2 the event $\mathcal{D}_n$ implies

$$N_{n,k,m_2} - N_{n,k,m_1} = \sum_{m_1 \leqslant m < m_2} \mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m}) \pm n^{1/2+2\lambda}. \tag{4.12}$$

Since the number of vertices in components of size k changes by at most ℓk per step, we have $|\Delta X_{n,k,m+1}| \leqslant \ell k$ and thus $|Z_{n,k,m+1} - Z_{n,k,m}| = |Y_{n,k,m+1}| \leqslant 2\ell k$. Furthermore $\mathbb{E}(Y_{n,k,m+1} \mid \mathcal{F}_{n,m}) = 0$, so $(Z_{n,k,j})_{j \geqslant m_1}$ is a martingale. Thus, for fixed k, m_1, m_2 satisfying the conditions above, by the Azuma–Hoeffding inequality we have

$$\mathbb{P}(|Z_{n,k,m_2} - Z_{n,k,m_1}| \geqslant n^{1/2+2\lambda}) \leqslant 2e^{-n^{3\lambda}/(8\ell^2 k^2)} \leqslant 2e^{-n^{\lambda}/(8\ell^2)} \leqslant n^{-9}$$

for n large enough. Taking a union bound (to account for all choices of k, m_1, m_2) yields $\mathbb{P}(\mathcal{D}_n) \to 1$ as $n \to \infty$.

Finally, define the 'good' event $\mathcal{G}_n = \mathcal{D}_n \cap \mathcal{K}_n \cap \mathcal{U}_n$. Our discussion yields that $\mathbb{P}(\mathcal{G}_n) \to 1$ as $n \to \infty$. We are now ready to state our main technical lemma. As usual, we ignore the irrelevant rounding to integers.

Lemma 4.3.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be a merging ℓ-vertex rule that is well behaved. Let (ω_n) with $\omega_n \in \mathcal{G}_n \subseteq \Omega_n$ be defined for some infinite subset of $\mathbb{N}$. There exists a subsequence $(\omega_{\tilde{n}})$ such that for each $t \geqslant 0$ and $k \geqslant 1$ we have $N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n} \to \rho_k(t)$, where the $(\rho_k(t))_{k \geqslant 1}$ are* *functions on* $\mathbb{R}^+$ *satisfying the system of differential equations* (4.4)–(4.6) *associated to* $\mathcal{R}$*.*

Note that Lemma 4.3.1 implies that the system of differential equations (4.4)–(4.6) has at least one solution. By comparison with the underlying process we can establish additional properties of the $\rho_k(t)$, e.g. that $\rho_{\leqslant k}(t) = \sum_{1\leqslant j\leqslant k}\rho_j(t)$ is monotone decreasing. Before giving the proof of Lemma 4.3.1, we first show how it implies Theorem 4.1.1. By Theorem 3.1.3 in Chapter 3 it suffices to establish (4.1), i.e., local convergence. Aiming at a contradiction, suppose there exists $\varepsilon > 0$, $t_0 \geqslant 0$, $k_0 \geqslant 1$ and an infinite subsequence $\bar{n}$ of $\mathbb{N}$ such that $|N_{\bar{n},k_0,t_0\bar{n}}/\bar{n} - \hat{\rho}_{k_0}(t_0)| > \varepsilon$ holds with probability at least ε, where $\hat{\rho}_{k_0}(t)$ is given by the (by assumption) unique solution to (4.4)–(4.6). Since $\mathbb{P}(\mathcal{G}_n) \to 1$ as $n \to \infty$, this implies (by discarding a finite number of elements in the beginning) that there exists an infinite subsequence of sample points $(\omega_{\bar{n}})$ with $\omega_{\bar{n}} \in \mathcal{G}_{\bar{n}} \subseteq \Omega_{\bar{n}}$ that satisfy

$$|N_{\bar{n},k_0,t_0\bar{n}}(\omega_{\bar{n}})/\bar{n} - \hat{\rho}_{k_0}(t_0)| > \varepsilon. \tag{4.13}$$

Now Lemma 4.3.1 gives a subsubsequence $(\omega_{\tilde{n}})$ satisfying $N_{\tilde{n},k,t\tilde{n}}(\omega_{\tilde{n}})/\tilde{n} \to \rho_k(t)$ for each $t \geqslant 0$ and $k \geqslant 1$, where the $(\rho_k(t))_{k\geqslant 1}$ solve (4.4)–(4.6) on $\mathbb{R}^+$. But now (4.13) implies $\rho_{k_0}(t_0) \neq \hat{\rho}_{k_0}(t_0)$, contradicting uniqueness.

4.3.2 Proof of Lemma 4.3.1

We start by selecting a 'nice' subsequence of (ω_n), and proceed as in the proof of Helley's selection theorem (see e.g. Theorem 5.8.1 in [38]). Define $F_n(k,t) = N_{n,k,tn}(\omega_n)/n$ if $1 \leqslant k \leqslant n$, otherwise set $F_n(k,t) = 0$. Clearly, $F_n(k,t) \in [0,1]$. Furthermore, $F_n(1,0) = 1$ and $F_n(k,0) = 0$ for $k \geqslant 2$. Let $(q_r)_{r\geqslant 1}$ be an enumeration of $\mathbb{Q}^+$. By using the Bolzano–Weierstrass theorem repeatedly, a standard diagonal argument yields a subsequence $(\omega_{\tilde{n}})$ such that for all $(k, q_r) \in \mathbb{N} \times \mathbb{Q}^+$ the value of $F_{\tilde{n}}(k, q_r)$ converges to some limit s_{k,q_r}. For each $k \in \mathbb{N}$ we now define $\rho_k(q_r) = s_{k,q_r}$ for all $q_r \in \mathbb{Q}^+$. Since $N_{n,k,m}$ changes by at most ℓk per step, as a function of t each $F_n(k,t)$ is Lipschitz on $\mathbb{Q}^+$ with constant ℓk, so ρ_k has this property on $\mathbb{Q}^+$. For each $k \in \mathbb{N}$ we thus can extend ρ_k to a Lipschitz continuous function on $\mathbb{R}^+$. Henceforth we always work with the subsequence selected above, but write

n instead of $\tilde{n}$ for ease of notation. In particular, for each $t \in \mathbb{R}^+$ and $k \in \mathbb{N}$ we thus have

$$N_{n,k,tn}(\omega_n)/n \to \rho_k(t). \tag{4.14}$$

Turning to some basic properties of the $\rho_k(t)$, clearly $0 \leqslant \rho_k(t) \leqslant 1$. Furthermore, the initial conditions $\rho_1(0) = 1$ and $\rho_k(0) = 0$ for $k \geqslant 2$ hold. Counting vertices we see that $0 \leqslant \sum_{k \geqslant 1} \rho_k(t) \leqslant 1$. So far we established (4.4) and (4.6).

The following *deterministic* claim is the main ingredient in the proof of Lemma 4.3.1. Recall that $\Delta X_{n,k,m+1} = N_{n,k,m+1} - N_{n,k,m}$. For brevity, we write $f_k(t) = f_k(t, \rho_1, \rho_2, \ldots)$ for the right hand side of (4.5).

Claim 4.3.2. *For all $t \geqslant 0$, $\varepsilon > 0$ and $k \geqslant 1$ there exists $0 < \delta \leqslant 1$ such that for n large enough the following holds: for each $m \geqslant 0$ satisfying $|m - tn| \leqslant \delta n$ we have*

$$\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = f_k(t) \pm \varepsilon/3. \tag{4.15}$$

Proof. Recall that $\omega_n \in \mathcal{K}_n \cap \mathcal{U}_n$ satisfies (4.14). Given $t \geqslant 0$, $\varepsilon > 0$ and $k \geqslant 1$, pick $\gamma \leqslant \varepsilon/(50\ell^2 k)$. Then choose $S \geqslant k$ such that $S \geqslant K = K(\gamma)$ and $\rho_{<S}(t) \geqslant 1 - \rho(t) - \gamma$ hold, which is possible since $1 - \rho(t) = \lim_{s\to\infty} \rho_{<s}(t)$, where $\rho_{<s}(t) = \sum_{1 \leqslant j < s} \rho_j(t)$ is increasing in s. Pick $\delta \leqslant \min\{\gamma/(3\ell S^2), 1\}$. By choice of S we have

$$\rho_{\geqslant S}(t) \leqslant \gamma. \tag{4.16}$$

Consider $m \geqslant 0$ satisfying $|m - tn| \leqslant \delta n$. Since $\omega_n \in \mathcal{U}_n$, we know that

$$N_{n,\geqslant \eta n,m}(\omega_n) > 0 \quad \text{implies} \quad N_{n,\geqslant \eta n,m}(\omega_n) = L_{1,n,m}(\omega_n). \tag{4.17}$$

We henceforth assume $\eta \leqslant \gamma$ and $\max\{S, g(S)\} < \eta n$, which both hold for n large enough (depending on γ, S, g). As $\omega_n \in \mathcal{K}_n$, $S \geqslant K$ and (4.11) imply $N_{n,\geqslant S,m}(\omega_n) \leqslant L_{1,n,m}(\omega_n) + \gamma n$ for n large enough (depending on γ). So, by distinguishing whether $L_{1,n,m}(\omega_n)$ is larger or smaller than ηn, we see that

$$N_{n,\geqslant S,m}(\omega_n) - N_{n,\geqslant \eta n,m}(\omega_n) \leqslant \eta n + \gamma n \leqslant 2\gamma n. \tag{4.18}$$

We shall now evaluate $\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n)$. For this we regard the graph $G_{n,m}(\omega_n)$ as fixed, and the vertices $\underline{v}_{n,m+1} = (v_1, \dots, v_\ell)$ as random. So, in the following all probabilities $\mathbb{P}(*)$ are shorthand for $\mathbb{P}(* \mid \mathcal{F}_{n,m})(\omega_n)$. Recall the definitions of $\underline{v}_{n,m+1} = (v_1, \dots, v_\ell)$ and $\underline{c}_{n,m+1} = (c_1, \dots, c_\ell)$: the vertices $v_1, \dots, v_\ell$ are chosen independently and uniformly at random from $[n]$, and each c_j denotes the size of the component in $G_{n,m}(\omega_n)$ containing v_j. So, for each $s \in [n]$ we have

$$\mathbb{P}(c_j = s) = N_{n,s,m}(\omega_n)/n.$$

We define $\mathcal{T}$ as the event that (a) all vertices v_j with $c_j \leqslant S$ are in different components, and (b) there are no vertices v_j with $S < c_j < \eta n$. Recall that $\max\{S, g(S)\} < \eta n$. Note that whenever $\mathcal{T}$ holds, by (4.17) all vertices in components of size larger than $g(S)$ are in the same component (the largest), so (4.2) or (4.3) applies giving

$$\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m}, \underline{v}_{n,m+1}) = d_k(\tilde{c}_1, \dots, \tilde{c}_\ell), \tag{4.19}$$

where $\tilde{c}_j = \infty$ for $c_j \geqslant \eta n$ and $\tilde{c}_j = c_j$ otherwise. In any case, the two sides of (4.19) are bounded by ℓk. Using (4.18) we see that $\mathbb{P}(\neg\mathcal{T}) \leqslant \ell^2 S/n + 2\ell\gamma$, and so by choice of γ we have

$$2\ell k \cdot \mathbb{P}(\neg\mathcal{T}) \leqslant 2\ell k \cdot (\ell^2 S/n + 2\ell\gamma) \leqslant \varepsilon/9.$$

Setting $\mathfrak{S} = [S] \cup \{s \in [n] : s \geqslant \eta n\}$, by taking expectations of both sides of (4.19), it follows that

$$\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \sum_{s_1,\dots,s_\ell \in \mathfrak{S}} d_k(\tilde{s}_1, \dots, \tilde{s}_\ell) \prod_{j \in [\ell]} \mathbb{P}\big(c_j = s_j\big) \pm \varepsilon/9, \tag{4.20}$$

where $\tilde{s}_j = \infty$ for $s_j \geqslant \eta n$ and $\tilde{s}_j = s_j$ otherwise.

Note that in the estimates above we had plenty of elbow room. So, if the conditional distribution of $\underline{v}_{n,m+1}$ is at total variation distance $\alpha_n = o(1)$ from the one used above, then a standard coupling argument shows that this only adds an additive error of at most $2\ell k\alpha_n$, which is negligible for n large enough (depending on k, ε), say, at most $\varepsilon/99$. So (4.20) is easily seen to still hold in such slight variations.

We define the random variable Y as follows:

$$\mathbb{P}\big(Y = \infty\big) \;=\; \begin{cases} L_{1,n,m}(\omega_n)/n, & \text{if } L_{1,n,m}(\omega_n) \geqslant \eta n, \\ 0, & \text{otherwise,} \end{cases}$$

and, for all $s \in \mathbb{N}$ we set

$$\mathbb{P}\big(Y = s\big) \;=\; \begin{cases} N_{n,s,m}(\omega_n)/n, & \text{if } s < \eta n, \\ 0, & \text{otherwise.} \end{cases}$$

Note that by (4.17) this yields a probability distribution. Let $Y_1, \ldots, Y_\ell$ be iid with distribution Y and observe that $\mathbb{P}(Y_j = s) = \mathbb{P}(c_j = s)$ for $s \leqslant S < \eta n$. Since by (4.17) there is at most one component of size at least ηn (and $\tilde{s}_j = \infty$ for $s_j \geqslant \eta n$), we see that (4.20) gives

$$\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \sum_{s_1,\ldots,s_\ell \in [S]\cup\{\infty\}} d_k(s_1, \ldots, s_\ell) \prod_{j\in[\ell]} \mathbb{P}\big(Y_j = s_j\big) \pm \varepsilon/9.$$

Now, using $|d_k(\cdot)| \leqslant \ell k$ and (4.18), we extend the sum to all $s_1, \ldots, s_\ell \in \mathbb{N} \cup \{\infty\}$ at the price of an additive error of $2\gamma\ell^2 k$. Since $2\gamma\ell^2 k \leqslant \varepsilon/20$ by choice of γ, this gives

$$\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \mathbb{E}(d_k(Y_1, \ldots, Y_\ell)) \pm \varepsilon/6. \tag{4.21}$$

For $s \leqslant S$ note that $N_{n,s,m}$ changes by at most $\ell s \leqslant \ell S$ in each step, so $|m - tn| \leqslant \delta n$ implies $|N_{n,s,m}(\omega_n) - N_{n,s,tn}(\omega_n)| \leqslant \ell S\delta n$. Hence, using the definition of δ and that (4.14) yields $N_{n,s,tn}(\omega_n)/n \to \rho_s(t)$, for $s \leqslant S$ we have, say,

$$|N_{n,s,m}(\omega_n)/n - \rho_s(t)| \leqslant \ell S\delta + \gamma/(2S) \leqslant \gamma/S \tag{4.22}$$

for n large enough (depending on γ, S). Using this observation we shall now show that the right hand side of (4.21) is essentially determined by the $(\rho_k(t))_{k\geqslant 1}$; this is key for our approach. To this end we consider the random variable Z which is defined by as follows for every $s \in \mathbb{N} \cup \{\infty\}$:

$$\mathbb{P}\big(Z = s\big) \;=\; \begin{cases} \rho(t), & \text{if } s = \infty, \\ \rho_s(t), & \text{otherwise.} \end{cases} \tag{4.23}$$

Claim 4.3.3. *For n large enough (depending on γ, S) we have*

$$\mathrm{d_{TV}}\big(Y\,, Z\big) \leqslant 4\gamma.$$

Proof. Recall that the total variation distance is given by

$$\mathrm{d_{TV}}\big(Y\,, Z\big) = \frac{1}{2} \sum_{s\in\mathbb{N}\cup\{\infty\}} \big|\mathbb{P}(Y = s) - \mathbb{P}(Z = s)\big|. \tag{4.24}$$

For $s \leqslant S$, note that (4.22) readily yields for n large enough (depending on γ, S)

$$\sum_{s\in[S]} \big|\mathbb{P}(Y = s) - \mathbb{P}(Z = s)\big| \leqslant \gamma.$$

Next, we consider the summands where $S < s < \infty$. By (4.16) and (4.18) we have $\mathbb{P}\big(S < Z < \infty\big) \leqslant \gamma$ and $\mathbb{P}\big(S < Y < \infty\big) \leqslant 2\gamma$, so

$$\sum_{S<s<\infty} \big|\mathbb{P}(Y = s) - \mathbb{P}(Z = s)\big| \leqslant 3\gamma.$$

Finally, since Y and Z are probability distributions, they differ on $s = \infty$ no more than the sum of the differences of the other values, i.e., by at most 4γ. This readily completes the proof (using the additional factor of 1/2 in (4.24)). □

Now, taking $Z_1, \ldots, Z_\ell$ iid with distribution Z, using Claim 4.3.3 the distribution of the Y_j and Z_j can be coupled such that they all agree with probability at least $1 - 4\ell\gamma$. So, since $|d_k(\cdot)| \leqslant \ell k$, in (4.21) we may replace all occurrences of Y_j by Z_j at the price of an additive error of $8\gamma\ell^2 k$. Since $8\gamma\ell^2 k \leqslant \varepsilon/6$ by choice of γ, our discussion yields

$$\mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) = \mathbb{E}(d_k(Z_1, \ldots, Z_\ell)) \pm \varepsilon/3.$$

Note that the first term on the right hand side equals $f_k(t) = f_k(t, \rho_1, \rho_2, \ldots)$ by definition of the Z_j, see (4.5) and (4.23). This establishes (4.15) and thus completes the proof of Claim 4.3.2. □

Finally, with Claim 4.3.2 in hand, we now complete the proof of Lemma 4.3.1. Given $t \geqslant 0$, $\varepsilon > 0$ and $k \geqslant 1$, pick $0 < \delta \leqslant 1$ as given by Claim 4.3.2. For each $0 < |h| \leqslant \delta$ with $t+h \geqslant 0$ write m_1, m_2 for the minimum and maximum of $\{(t+h)n, tn\}$, which satisfy $m_1 \geqslant 0$ and $0 < m_2 - m_1 < n^{1+\lambda}$. Recall that $\omega_n \in \mathcal{D}_n$, and note that $k \leqslant n^\lambda$ for n large enough (depending on k). Now, using (4.12) and (4.15) we see that for n large enough

$$\begin{aligned} N_{n,k,(t+h)n}(\omega_n) - N_{n,k,tn}(\omega_n) &= \operatorname{sgn}(h) \cdot \sum_{m_1 \leqslant m < m_2} \mathbb{E}(\Delta X_{n,k,m+1} \mid \mathcal{F}_{n,m})(\omega_n) \pm n^{1/2+2\lambda} \\ &= hn \cdot (f_k(t) \pm \varepsilon/3) \pm n^{1/2+2\lambda}. \end{aligned}$$

Rearranging terms, using (4.14) and $\lambda < 1/4$ we deduce that for n large enough (depending on ε, h) we have

$$\left| \frac{\rho_k(t+h) - \rho_k(t)}{h} - f_k(t) \right| \leqslant \varepsilon/2 + n^{-1/2+2\lambda}/|h| \leqslant \varepsilon. \tag{4.25}$$

To summarize, for all $t \geqslant 0$, $\varepsilon > 0$ and $k \geqslant 1$ there exists $\delta > 0$ such that for all $0 < |h| \leqslant \delta$ with $t + h \geqslant 0$ equation (4.25) holds along the selected subsequence (for n large enough). In other words, for $t > 0$ we have $\rho_k'(t) = f_k(t)$, which establishes (4.5). For $t = 0$ we only considered $0 < h \leqslant \delta$, so we proved the corresponding statement for the right derivative, and the proof of Lemma 4.3.1 is complete.

Chapter 5

The subcritical evolution

5.1 Main result

The goal of this chapter is to improve on the conditional convergence result for Achlioptas processes established in Chapter 4, and we aim at removing the extra analytic assumption of Theorem 4.1.1 (which made sure that a certain system of differential equations has a unique solution). To this end, given a rule $\mathcal{R}$, we analyze the evolution of the associated random sequence of graphs $(G_i)_{i \geqslant 1}$ in more detail. Here the *susceptibility* turns out to be a key parameter, which for a graph G is defined as $S(G) = \sum kN_k(G)/n$. In other words, the susceptibility is the expected size of the component containing a randomly chosen vertex. For bounded-size rules there is a close connection between the susceptibility and the *percolation threshold* $t_c = t_c^{\mathcal{R}}$, which satisfies the following: for $t < t_c$ whp $L_1(G_{tn}) = o(n)$ while for $t > t_c$ whp $L_1(G_{tn}) = \Omega(n)$. Indeed, the analysis for bounded-size rules in [19, 45, 74] uses $S(G_i)$ as well as $N_k(G_i)$ for fixed k as key statistics. In the 'subcritical' regime $t < t_c$, they use Wormald's differential equation method [80, 81] to establish the existence of functions $\rho_k = \rho_k^{\mathcal{R}}$, $s = s^{\mathcal{R}}$ such that $N_k(G_{tn}) \approx \rho_k(t)n$ and $S(G_{tn}) \approx s(t)$ hold whp. Based on this they show that $t_c = t_c^{\mathcal{R}}$ is given by the blow-up point of the susceptibility: $\lim_{t \nearrow t_c} s(t) = \infty$. As indicated by Spencer and Wormald [74], for general size rules the approximation of the key statistics using the differential equation method seems difficult.

Our main result establishes the first rigorous convergence result for Achlioptas processes using unbounded size rules such as the product rule: we show that the number of vertices in components of size $k \geqslant 1$ (and the susceptibility) is tightly concentrated until the susceptibility 'blows up', which happens at a critical time t_{b}. In fact, our result holds for a very large class of Achlioptas-like processes, including essentially all Achlioptas processes studied so far (see Section 5.2 for the formal definition of ℓ-vertex size rules). Here $S(G^{\mathcal{R}}_{tn}) \xrightarrow{\mathrm{p}} \infty$ as $n \to \infty$ means that for any $C > 0$ we have $\mathbb{P}(S(G^{\mathcal{R}}_{tn}) \leqslant C) \to 0$ as $n \to \infty$, and $N_{\geqslant k}(G)$ denotes the number of vertices of G in components of size at least k.

Theorem 5.1.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. There exist $t_{\mathrm{b}} = t^{\mathcal{R}}_{\mathrm{b}} \in [\frac{1}{\ell(\ell-1)}, 1]$ and functions $(\rho_k)_{k \geqslant 1}$ with $\rho_k = \rho^{\mathcal{R}}_k : [0, t_{\mathrm{b}}) \to [0, 1]$ such that the following holds. For every $t \geqslant t_{\mathrm{b}}$ we have*

$$S(G^{\mathcal{R}}_{tn}) \xrightarrow{\mathrm{p}} \infty \tag{5.1}$$

as $n \to \infty$. For every $t < t_{\mathrm{b}}$ we have $\sum_{k \geqslant 1} \rho_k(t) = 1$. Also, for every $t < t_{\mathrm{b}}$ there exist $a, A, C > 0$ (depending only on $\mathcal{R}, \ell, t$) such that for every $t' \in [0, t]$ we have $\rho_k(t') \leqslant Ae^{-ak}$ for all $k \geqslant 1$. In addition, for $n \geqslant n_0(\mathcal{R}, \ell, t)$ the following holds with probability at least $1 - n^{-99}$: for every $0 \leqslant i \leqslant tn$ we have

$$|N_k(G^{\mathcal{R}}_i) - \rho_k(i/n)n| \leqslant (\log n)^C n^{1/2} \quad \textit{for all } k \geqslant 1, \tag{5.2}$$

$$|S(G^{\mathcal{R}}_i) - \textstyle\sum_{k \geqslant 1} k\rho_k(i/n)| \leqslant (\log n)^C n^{-1/2}, \tag{5.3}$$

and $N_{\geqslant k}(G^{\mathcal{R}}_i) \leqslant Ae^{-ak}n$ for all $k \geqslant 1$.

To interpret this result, we think of the functions $\rho_k(t)$ as describing the 'scaling limit' of the component size distribution at 'time' $t < t_{\mathrm{b}}$, where time is the number of steps divided by n. A key aspect of the result is that this limit does not depend on n; in fact, most of our technical work is devoted to establishing this property – to show only that $N_k(G^{\mathcal{R}}_{tn})$ is concentrated around its expectation, simpler arguments would suffice. The tail bound on ρ_k given in Theorem 5.1.1 states that the idealized component size distribution has an exponential tail for $t < t_{\mathrm{b}}$, as one would expect in a strictly sub-critical random graph. It implies that $s(t) = \sum_{k\geqslant 1} k\rho_k(t) < \infty$ if $t < t_{\mathrm{b}}$, so (5.3) implies that that for $t < t_{\mathrm{b}}$ we have

$$S(G^{\mathcal{R}}_{tn}) \xrightarrow{\mathrm{p}} s(t) < \infty, \tag{5.4}$$

where $\xrightarrow{\mathrm{p}}$ denotes convergence in probability. The proof of Theorem 5.1.1 will show that $s(t_{\mathrm{b}} - \varepsilon) \geqslant (\ell(\ell-1)\varepsilon)^{-1}$, so the (idealized) susceptibility $s(t)$ blows up at t_{b}. The last statement of the theorem implies that for $t < t_{\mathrm{b}}$ we have $L_1(G^{\mathcal{R}}_{tn}) \leqslant B_t \log n$ whp, for some constant B_t that depends on t. Finally, we shall show in Section 5.4 that (unless $\mathcal{R}$ directly adds cycles to the graph), for $t < t_{\mathrm{b}}$ whp almost all components are trees, with the rest unicyclic.

Theorem 5.1.1 allows us to say something about what happens at time $t = t_{\mathrm{b}}$. Indeed, the definition of an ℓ-vertex rule ensures that in one step, at most ℓ components are destroyed and at most ℓ (in fact at most $\ell/2$) are created, so $|N_k(G^{\mathcal{R}}_{i+1}) - N_k(G^{\mathcal{R}}_{i})| \leqslant \ell k$. It follows that each ρ_k is Lipschitz continuous on $[0, t_{\mathrm{b}})$ with constant $k\ell$. Hence we can extend each ρ_k continuously to the point t_{b}, and (5.2) and the Lipschitz properties of N_k and ρ_k imply that

$$N_k(G^{\mathcal{R}}_{t_{\mathrm{b}}n})/n \xrightarrow{\mathrm{p}} \rho_k(t_{\mathrm{b}}). \tag{5.5}$$

Together with Theorem 5.1.1, the continuity results of Chapter 3 imply that $L_1(G_{t_{\mathrm{b}}n})/n \xrightarrow{\mathrm{p}} 0$, and that $\sum_k \rho_k(t_{\mathrm{b}}) = 1$, so the numbers $(\rho_k(t_{\mathrm{b}}))_{k\geqslant 1}$ do capture the asymptotic component size distribution of $G^{\mathcal{R}}_{t_{\mathrm{b}}n}$, although we do not have such tight error bounds as in (5.2).

The proof of Theorem 5.1.1 is based on a variant of the neighbourhood exploration process and relies on branching process (approximation) arguments. This is quite different from previous approaches in this area, which are based on the differential equation method; for certain (restricted) classes of rules these establish *local convergence*, i.e., that there exist functions $\rho_k = \rho_k^{\mathcal{R}} : \mathbb{R}^+ \to [0,1]$ such that, for each fixed $k \geqslant 1$ and $t \geqslant 0$, we have $N_k(G^{\mathcal{R}}_{tn})/n \xrightarrow{\mathrm{p}} \rho_k(t)$ as $n \to \infty$. The limitations of these approaches are that they (i) only apply to certain bounded-size rules [19, 74], or (ii) when applied to size rules need the additional assumption that certain systems of differential equations have unique solutions [65], which is not known to hold for the product rule, for example. So, Theorem 5.1.1 establishes for the first time (a strong form of) local convergence for unbounded size rules (such as the product rule) until the susceptibility diverges. We believe that this convergence result is best possible: on the basis of heuristics and simulations presented in [64] we believe that there are certain natural size rules for which beyond $t = t_{\mathrm{b}}$ a giant component emerges whose size is not concentrated (see also Figure 4.1). In these rules the numbers of vertices in components of each fixed size k are presumably also not concentrated after this point.

Theorem 5.1.1 has some analogies with 'classical' percolation theory on, for example, the infinite lattice $\mathbb{Z}^d$, where there are two *a priori* different critical probabilities p_{H} and p_{T}. Intuitively, these correspond to the thresholds for (i) having (with positive probability) an infinite cluster and (ii) the expected cluster size being infinite. For essentially all 'natural' lattices of interest it is nowadays known that $p_{\mathrm{H}} = p_{\mathrm{T}}$ (see e.g. [2, 54]), but this fact is not at all obvious! Note that in the finite setting of this thesis these two properties correspond to (i) having a linear size component, and (ii) diverging susceptibility. More formally, define $t_{\mathrm{c}} = t_{\mathrm{c}}^{\mathcal{R}}$ as the supremum of the set of $t \geqslant 0$ for which $L_1(G_{tn}^{\mathcal{R}})/n \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$, and $t_{\mathrm{b}} = t_{\mathrm{b}}^{\mathcal{R}}$ as the supremum of the set of $t \geqslant 0$ for which $S(G_{tn}^{\mathcal{R}})$ is bounded in probability. Note that $L_1(G) \leqslant \sqrt{nS(G)}$ implies $t_{\mathrm{b}} \leqslant t_{\mathrm{c}}$. The remark after Theorem 5.1.1 entails that for size rules t_{b} is equal to the infimum of the set of $t \geqslant 0$ for which (5.1) holds, and that $S(G_{t_{\mathrm{b}}n}^{\mathcal{R}}) \xrightarrow{\mathrm{p}} \infty$. In fact, we believe that both thresholds coincide for size rules (analogous to the 'classical' case).

Conjecture 5.1.2. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. Then $t_{\mathrm{b}}^{\mathcal{R}} = t_{\mathrm{c}}^{\mathcal{R}}$. More precisely, for any $t > t_{\mathrm{b}}^{\mathcal{R}}$ and $\varepsilon > 0$ there exist $\delta, n_0 > 0$ (depending only on $\mathcal{R}, \ell, t, \varepsilon$) such that $\mathbb{P}(L_1(G_{tn}^{\mathcal{R}}) \geqslant \delta n) \geqslant 1 - \varepsilon$ for $n \geqslant n_0$.*

Recall that Achlioptas processes (where the choice is between two edges) are a sub-class of 4-vertex rules. Conjecture 5.1.2 was proved for bounded-size Achlioptas processes by Spencer and Wormald [74], and for a subset of these processes by Bohman and Kravitz [19]. We shall show in Section 5.3.1 that it holds for all bounded-size ℓ-vertex rules, as well as many other size rules, including the 'reverse product rule', for example. However, it does not hold for general ℓ-vertex rules. Indeed, in Section 5.3.2 we show that modified size rules with one additional feature, namely that they may once switch their behaviour based on the number n of vertices and the number i of steps (or the value of the susceptibility), can delay the appearance of a linear size component for $\Omega(n)$ steps beyond the point where the susceptibility diverges.

5.2 Evolution of Achlioptas processes with an initial graph

In this chapter we consider the evolution of Achlioptas processes starting with an initial graph F with vertex set $V = [n]$; we restrict our attention to *ℓ-vertex size rules* $\mathcal{R}$, whose decisions depend only on the sizes of the components containing the randomly chosen vertices. More precisely, each such rule $\mathcal{R}$ yields a random sequence $(F_i^{\mathcal{R}})_{i\geqslant 0}$ of graphs on V with $F_0^{\mathcal{R}} = F$. For every $i \geqslant 0$ we draw ℓ vertices $\underline{v}_{i+1} = (v_1, \dots, v_\ell)$ from V independently and uniformly at random, and then, writing $\underline{c}_{i+1} = (c_1, \dots, c_\ell)$ for the sizes of the components containing $v_1, \dots, v_\ell$ in $F_i^{\mathcal{R}}$, we obtain obtain $F_{i+1}^{\mathcal{R}}$ by adding a non-empty set of edges E_{i+1} to $F_i^{\mathcal{R}}$, where $\mathcal{R}$ deterministically selects E_{i+1} as a subset of all pairs between vertices in $\underline{v}_{i+1}$ based only on the list of component sizes $\underline{c}_{i+1}$. Usually exactly one edge is added, but there is no reason to insist on this.

When $F = G_0$ is the empty graph on n vertices we obtain the 'standard' Achlioptas processes using ℓ-vertex size rules $\mathcal{R}$ as defined in Chapter 3. As usual, we can allow for small variations in the above definition; this includes, for example, each time picking an ℓ-tuple of *distinct* vertices, or picking (the ends of) $\ell/2$ randomly selected (distinct) edges not already present in $G_i^{\mathcal{R}}$, see also Section 3.2. For $\ell = 2$ we thus recover the 'classical' Erdős–Rényi random graph process by always adding the pair v_1v_2. In addition, our proofs can be written to allow $\mathcal{R}$ to make randomized decisions (with the probability of adding some set of edges depending only on $\underline{c}_{i+1}$), and, furthermore, to allow $\mathcal{R}$ to know which vertices in $\underline{v}_{i+1}$ are in the same components of $F_i^{\mathcal{R}}$ (for compatibility with Chapter 3 we then require $E_{i+1} \neq \emptyset$ whenever all v_j are in distinct components, although nothing in the proof of Theorem 5.1.1, except for the bound $t_{\mathrm{b}} \leqslant 1$, needs this).

One difficulty in the proof of Theorem 5.1.1 is that there is a complicated dependence between the decisions of $\mathcal{R}$ in each round (and their order is also important). Indeed, changes can 'propagate' throughout the process: if the sizes of a few components are modified (e.g.

by altering decisions of $\mathcal{R}$ or tuples $\underline{v} = (v_1, \dots, v_\ell)$ offered), then this might change many future decisions of $\mathcal{R}$, which in turn might alter further decisions, etc. To overcome this our proof proceeds by induction, always establishing concentration only for a small number of steps; this is also the reason why we study the more general evolution starting from an initial graph F. Each time we rely on a two-round exposure argument: in the first round we reveal which tuples are selected, and in the second we then expose their order. For size rules *not* all tuples and components of F 'influence' the size of the component in $F_i^{\mathcal{R}}$ containing v: only those which can be reached from v after adding *all* pairs of each ℓ-tuple to the graph (every rule only selects a subset of these pairs). The key observation is now that given the corresponding 'relevant' tuples and components of F of the first round, the order of these tuples (exposed in the second round) determines the size of the component containing v. It turns out that if we only consider σn rounds for σ sufficiently small, then an exploration process determining these relevant tuples and components in the first round can be closely approximated by a subcritical branching process $\mathfrak{X}_\sigma$ which is defined *without* reference to n. Since the outcome of the second round is a (random) function of the first one, it thus seems plausible that $\mathbb{E}N_k(F_{\sigma n}^{\mathcal{R}})/n$ is independent of n (up to small error terms). In addition, since the first round is subcritical, this means that there typically are not too many tuples and components which influence the size of the component containing v. At least on an intuitive level this makes it plausible that it should be possible to establish concentration of $N_k(F_i^{\mathcal{R}})$ around its expectation by applying McDiarmid's inequality.

The rest of this chapter is organized as follows. In the next section we state our main technical result (Theorem 5.2.1), and then show in Section 5.2.2 how it implies Theorem 5.1.1. Afterwards, in Section 5.2.3 we present some branching process preliminaries; these are used in Section 5.2.4, where we establish Theorem 5.2.1. In Section 5.3 we discuss Conjecture 5.1.2, giving examples of classes of size rules for which we can prove the conjecture, and examples of non-size rules for which it does not hold. Finally, in Section 5.4 we consider the cycle structure of Achlioptas processes.

5.2.1 Main technical result

Our main technical result establishes concentration during the evolution of Achlioptas processes starting with an initial graph F. The special case of an Erdős–Rényi evolution from an initial graph F (which can be seen as an evolving version of a special case of the inhomogeneous random graph model of Bollobás, Janson and Riordan [24]) has been previously studied by Spencer and Wormald [74] and Janson and Spencer [45], the main focus being on the size of the largest component. In this context the susceptibility turns out to be the key parameter, and both papers use in essential ways that the Erdős–Rényi evolution corresponds to the addition of *uniform* random edges (or pairs of vertices). In contrast, when studying the evolution of Achlioptas processes, we need to deal with intricate dependencies between the edges added.

Using susceptibility as a guide, we now briefly motivate the number of steps our result applies to. Suppose that, starting with F satisfying $S(F) = L$, we use the rule $\mathcal{I}$ which in each step joins all ℓ random vertices by edges. Set $s(t) = S(F^{\mathcal{I}}_{tn})$. If the sizes of the joined components are c_i, then (assuming that all components are distinct) the susceptibility changes by $(\sum c_i)^2/n - \sum c_i^2/n = \sum_{i \neq j} c_i c_j/n$. So, since the vertices of each tuple are chosen uniformly at random, it seems plausible that typicality we have $s'(t) \approx n\mathbb{E}(S(F^{\mathcal{I}}_{tn+1}) - S(F^{\mathcal{I}}_{tn})) \approx \ell(\ell-1)s(t)^2$. For $t < [\ell(\ell-1)L]^{-1} = t_{\mathrm{s}}$ this suggests $s(t) \approx [1/L - \ell(\ell-1)t]^{-1}$. Since in each step any rule $\mathcal{R}$ only adds a subset of all $\binom{\ell}{2}$ pairs to the graph, this indicates that the susceptibility does not 'blow up' as long as $t < t_{\mathrm{s}}$. The following result confirms this heuristic argument and shows that, under suitable conditions, for $t < t_{\mathrm{s}}$ the number of vertices in components of size $k \geqslant 1$ is also tightly concentrated (the function ρ intuitively results from an 'infinite' version of the rule $\mathcal{R}$). Here we set $\chi(\varphi) = \sum_{k \geqslant 1} k\varphi(k)$ and $\chi(\rho, t) = \sum_{k \geqslant 1} k\rho(k,t)$, and write $x = a \pm b$ as shorthand for $x \in [a-b, a+b]$.

Theorem 5.2.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. Suppose $\beta > 1$, $B > 0$,*

$L \geqslant 1$ and $\varphi : \mathbb{N} \to [0,1]$ satisfy

$$\textstyle\sum_{k\geqslant 1}\varphi(k) = 1, \tag{5.6}$$

$$\textstyle\sum_{k\geqslant 1}\varphi(k)\beta^k \leqslant B, \tag{5.7}$$

$$\chi(\varphi) \leqslant L. \tag{5.8}$$

There is a function $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ (depending only on $\varphi, \mathcal{R}, \ell$) such that for all $\sigma \geqslant 0$ satisfying

$$\sigma < [\ell(\ell-1)L]^{-1} \tag{5.9}$$

there exist $\tilde{\beta}, \tilde{B}, \tilde{L} > 1$ (depending only on $\ell, L, \sigma, \beta, B$) such that for every $t \in [0,\sigma]$ equations (5.6)–(5.8) hold when $\beta, B, L, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, \rho(\cdot, t)$. If in addition F is a graph on n vertices which for $C \geqslant 0$ satisfies

$$N_k(F) = \varphi(k)n \pm (\log n)^C n^{1/2} \quad \textit{for all } k \geqslant 1, \tag{5.10}$$

$$\textstyle\sum_{k\in[n]}N_k(F)\beta^k \leqslant Bn, \tag{5.11}$$

$$S(F) \leqslant L, \tag{5.12}$$

then, setting $\tilde{C} = C + 9$, for $n \geqslant n_0(\ell, L, \sigma, \beta, B, C)$ the following holds with probability at least $1 - n^{-200}$: for every $0 \leqslant i \leqslant \sigma n$ we have

$$S(F_i^{\mathcal{R}}) = \chi(\rho, i/n) \pm (\log n)^{\tilde{C}} n^{-1/2}, \tag{5.13}$$

and equations (5.10)–(5.12) hold when $\beta, B, L, C, F, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, \tilde{C}, F_i^{\mathcal{R}}, \rho(\cdot, i/n)$.

The proof of Theorem 5.2.1 is quite involved and is deferred to Section 5.2.4. It is useful to observe that since $\beta > 1$ holds, (5.7) and (5.11) imply the tail bounds $\max\{\sum_{j\geqslant k}\varphi(j), N_{\geqslant k}(F)/n\} \leqslant B\beta^{-k}$ for all $k \geqslant 1$, so $L_1(F) = O(\log n)$. By Theorem 5.2.1 analogous estimates also hold for $F_i^{\mathcal{R}}$ with $i \leqslant \sigma n$. In fact, for (5.11), (5.12) to hold with β, B, L, F replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, F_i^{\mathcal{R}}$, a minor modification of our proof shows that it suffices to assume (5.9), (5.11) and (5.12) only; for the special case $\ell = 2$ this was established by Spencer and Wormald [74] under similar conditions. However, the key point of Theorem 5.2.1 is (5.10), i.e., that we obtain concentration of number of vertices in components of size k.

Turning to the susceptibility, by combining (5.10) with the tail bounds following from (5.7) and (5.11), for each fixed j we readily obtain rather precise estimates for $S_j(F_i^{\mathcal{R}}) = \sum_{k\in[n]} k^j N_k(F_i^{\mathcal{R}})/n$ with $i \leqslant \sigma n$, similar to (5.13). Furthermore, since $L_1(F_{\sigma n}^{\mathcal{R}}) = O(\log n)$ whp, we can easily use the differential equation method to make our heuristic discussion regarding the susceptibility rigorous, which e.g. yields $\chi(\rho,\sigma) \leqslant [1/L - \ell(\ell-1)\sigma]^{-1}$ (for the special case $\ell = 2$ this was noted by Bohman et. al [15]; it is also implicit in [74]). However, this crude bound, which follows from always connecting all ℓ vertices by edges in each step, is generally far from the truth; for this reason it does not suffice in our inductive application of Theorem 5.2.1, where we use the 'correct' value given by (5.13).

5.2.2 Proof of Theorem 5.1.1

This section is devoted to the proof of Theorem 5.1.1, which we establish by an inductive application of Theorem 5.2.1: each time we show concentration during a small number of steps (and maintain certain technical conditions), where the lengths of these intervals decrease as the susceptibility increases. This is also the main idea of the following rather technical construction: as we shall see in the proof of Lemma 5.2.2, for each interval of length Δ_j it determines the scaling limits ρ_k (and certain tail bounds) in a way that does *not* depend on n.

We inductively define a sequence $(\beta_j, B_j, \rho_j, \Delta_j, L_j)_{j\geqslant 0}$ with $\beta_j > 1$, $B_j > 0$, $\Delta_j \geqslant 0$, $L_j \geqslant 1$ and $\rho_j : \mathbb{N}\times\mathbb{R}^+ \to [0,1]$, where the β_j are decreasing ($\beta_{j+1} \leqslant \beta_j$) and the B_j are increasing ($B_{j+1} \geqslant B_j$). In addition, for each $j \geqslant 0$ the sequence satisfies the following invariant: for every $t \in [0,\Delta_j]$ equations (5.6) and (5.7) hold with $\beta, B, \varphi(\cdot)$ replaced by $\beta_j, B_j, \rho_j(\cdot,t)$. We start by setting $\beta_0 = B_0 = L_0 = 2$, $\Delta_0 = 0$ and defining $\rho_0 : \mathbb{N}\times\mathbb{R}^+ \to [0,1]$ with $\rho_0(1,0) = 1$ and $\rho_0(k,t) = 0$ otherwise. Given $j \geqslant 1$, recall that $\chi(\rho_{j-1},t) = \sum_{k\geqslant 1} k\rho_{j-1}(k,t)$ and set

$$L_j = \chi(\rho_{j-1},\Delta_{j-1}) + 1 \quad \text{and} \quad \Delta_j = [\ell(\ell-1)\,(L_j+1)]^{-1}. \tag{5.14}$$

Applying the first part of Theorem 5.2.1 with $\beta = \beta_{j-1}$, $B = B_{j-1}$, $L = L_j$ and $\varphi(k) =$

$\rho_{j-1}(k, \Delta_{j-1})$, we use the resulting ρ to define $\rho_j = \rho$. Furthermore, by considering $\sigma = \Delta_j$ we obtain $\tilde{\beta}, \tilde{B}$ and set $\beta_j = \min\{\tilde{\beta}, \beta_{j-1}\}$ and $B_j = \max\{\tilde{B}, B_{j-1}\}$; by Theorem 5.2.1 these satisfy the required invariant. Furthermore, it is not difficult to see that the entire sequence $(\beta_j, B_j, \rho_j, \Delta_j, L_j)_{j \geqslant 0}$ depends only on $\mathcal{R}, \ell$.

Next we combine the ρ_j (each valid on an interval of length Δ_j) to form $\varphi(k,t)$, which will eventually be $\rho_k(t)$ in Theorem 5.1.1; this notation avoids confusion with the ρ_j used. For $t \geqslant 0$ we define r_t as the smallest r such that $t \leqslant \sum_{0 \leqslant j \leqslant r} \Delta_j$ and set $r = \infty$ if no such r exists. For all $(k,t) \in \mathbb{N} \times \mathbb{R}^+$ set

$$\varphi(k,t) = \begin{cases} \rho_{r_t}(k, t - \sum_{0 \leqslant j < r_t} \Delta_j), & \text{if } r_t < \infty, \\ 0, & \text{otherwise.} \end{cases} \tag{5.15}$$

Transferring this definition to the invariant of the sequence introduced above, for all $t \geqslant 0$ with $r_t < \infty$ it follows that

$$\textstyle\sum_{k \geqslant 1} \varphi(k,t) = 1, \tag{5.16}$$

and that for every $t' \in [0,t]$ we have

$$\textstyle\sum_{k \geqslant 1} \varphi(k,t') \beta_{r_t}^k \leqslant B_{r_t}. \tag{5.17}$$

Now we are ready to prove the following concentration result, which also implies that in the previous construction we always have $r_t < \infty$ if (5.1) fails.

Lemma 5.2.2. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an ℓ-vertex size rule. For every $t \geqslant 0$ for which* (5.1) *fails we have $r_t < \infty$, and there exist $a, A, C > 0$ (depending only on $\mathcal{R}, \ell, t$) such that the following holds for $n \geqslant n_0(\mathcal{R}, \ell, t)$ with probability at least $1 - n^{-99}$: for every $0 \leqslant i \leqslant tn$ we have*

$$N_k(G_i^{\mathcal{R}}) = \varphi(k, i/n) n \pm (\log n)^C n^{1/2} \quad \textit{for all } k \geqslant 1,$$
$$S(G_i^{\mathcal{R}}) = \textstyle\sum_{k \geqslant 1} k \varphi(k, i/n) \pm (\log n)^C n^{-1/2},$$

and $N_{\geqslant k}(G_i^{\mathcal{R}}) \leqslant A e^{-ak} n$ for all $k \geqslant 1$.

Proof. Given $t \geqslant 0$, if (5.1) fails there exists $\varepsilon > 0$ and an infinite subsequence $\bar{n}$ of $\mathbb{N}$ (depending only on $\mathcal{R}, \ell, t$) satisfying

$$\mathbb{P}(S(G^{\mathcal{R}}_{t\bar{n}}) \leqslant \varepsilon^{-1}) \geqslant \varepsilon. \tag{5.18}$$

Let $\overline{L} = \varepsilon^{-1} + 3$ and $K = \lceil t\ell(\ell-1)\overline{L}\rceil + 1$. Let $t_0 = 0$, and for $j \geqslant 1$ let

$$t_j = \begin{cases} t_{j-1}, & \text{if } t_{j-1} > t, \\ t_{j-1} + \Delta_j, & \text{otherwise.} \end{cases}$$

For $n \geqslant n_0(\mathcal{R}, \ell, t)$ we inductively show that for every $0 \leqslant j \leqslant K$, setting $C_j = 9j + 2$, with probability at least $1 - jn^{-200}$, for every $0 \leqslant i \leqslant t_j n$ we have

$$N_k(G^{\mathcal{R}}_i) = \varphi(k, i/n)n \pm (\log n)^{C_j} n^{1/2} \quad \text{for all } k \geqslant 1, \tag{5.19}$$

$$\textstyle\sum_{k\in[n]} N_k(G^{\mathcal{R}}_i)\beta_j^k \leqslant B_j n, \tag{5.20}$$

$$S(G^{\mathcal{R}}_i) = \chi(\varphi, i/n) \pm (\log n)^{C_j} n^{-1/2}, \tag{5.21}$$

and for every $0 \leqslant s \leqslant j$ with $t_s \leqslant t$ we have

$$\chi(\varphi, t_s) = \chi(\rho_s, \Delta_s) < \overline{L} - 2. \tag{5.22}$$

Note that if $t_K < t$, then substituting (5.22) into (5.14) yields $\Delta_j > [\ell(\ell-1)\overline{L}]^{-1}$ for all $1 \leqslant j \leqslant K$. From $K > t\ell(\ell-1)\overline{L}$ it thus follows that $t_K > t$, a contradiction. Thus (5.22) implies $t_K \geqslant t$, i.e., $r_t < \infty$. Recall that $(\beta_j, B_j)_{j\geqslant 1}$ and K depend only on $\mathcal{R}, \ell$ and on $\mathcal{R}, \ell, t$ respectively. Hence the induction hypothesis for $j = K$ implies Lemma 5.2.2, where the tail bounds follow from (5.20) as $\beta_K > 1$.

For the base case $j = 0$ we start with an empty graph on n vertices, and it is easy to see that (5.19)–(5.22) hold with $\beta_0 = B_0 = C_0 = 2$ and $\varphi(k, 0) = \rho_0(k, 0)$, as defined above (5.14).

Given $j \geqslant 1$, for the induction step we may assume that $t_{j-1} \leqslant t$ (otherwise $t_j = t_{j-1}$, and there is nothing to prove). We first assume that $G^{\mathcal{R}}_{t_{j-1}n}$ satisfies the induction hypothesis, i.e., (5.19)–(5.22) with j replaced by $j-1$. In particular, (5.16) and (5.17) hold for $t = t_{j-1}$ with $r_t = j-1$, and we have $S(G^{\mathcal{R}}_{t_{j-1}n}) \leqslant \chi(\varphi, t_{j-1})+1 = \chi(\rho_{j-1}, \Delta_{j-1})+1 = L_j$ for $n \geqslant n_0(C_{j-1})$. Now we condition on $G^{\mathcal{R}}_{t_{j-1}n} = F$ and, analogous as after (5.14), apply Theorem 5.2.1 with $\beta = \beta_{j-1}$, $B = B_{j-1}$, $L = L_j$, $\sigma = \Delta_j$, $C = C_{j-1}$ and $\varphi(k) = \rho_{j-1}(k, \Delta_{j-1}) = \varphi(k, t_{j-1})$, which is possible by the induction hypothesis (and the properties established above). So, for $n \geqslant n_0(\ell, L_j, \Delta_j, \beta_{j-1}, B_{j-1}, C_{j-1})$, with probability at least $1 - n^{-200}$, for every $0 \leqslant i \leqslant \Delta_j n$ the graph $F^{\mathcal{R}}_i$ satisfies (5.6)–(5.7), (5.10)–(5.11) and (5.13) when $\beta, B, C, F, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{C}, F^{\mathcal{R}}_i, \rho(\cdot, i/n)$, where $\tilde{C} = C_{j-1} + 9$. Note that for size rules $F^{\mathcal{R}}_{\Delta_j n}$ is exactly $G^{\mathcal{R}}_{t_j n}$ conditional on $G^{\mathcal{R}}_{t_{j-1}n} = F$. It is crucial that $\tilde{\beta}, \tilde{B}, \tilde{C}, \rho$ do *not* depend on the initial graph F, and that by construction $\rho = \rho_j$, $\beta_j \leqslant \tilde{\beta}$, $B_j \geqslant \tilde{B}$ and $C_j = \tilde{C}$. So, by appealing to the induction hypothesis and recalling (5.15), it follows that with probability at least $1 - (j-1)n^{-200} - n^{-200}$ equations (5.19)–(5.21) hold. It remains to show that (5.22) holds. To this end recall that (5.21) holds with probability at least $1 - jn^{-200} > 1 - \varepsilon$ for all $n \geqslant n_0(C_j, K, \varepsilon)$. So, using that susceptibility is monotone increasing, by (5.18) it follows that for all $0 \leqslant t' \leqslant \min\{t, t_j\}$ we have $\chi(\varphi, t') < \varepsilon^{-1} + 1 = \overline{L} - 2$, say. Now (5.22) follows by combining the previous estimate with the observation that for every $s \leqslant j$ with $t_s \leqslant t$ we have $\chi(\varphi, t_s) = \chi(\rho_s, \Delta_s)$. This completes the induction step.

Finally, to see that $n \geqslant n_0(\mathcal{R}, \ell, t)$ suffices note that in each of the K steps we only used $n \geqslant n_0(\ell, L_j, \Delta_j, \beta_{j-1}, B_{j-1}, C_{j-1}, C_j, K, \varepsilon)$, where $C_j = 9j + 2$ and $L_j, \Delta_j, \beta_{j-1}, B_{j-1}$ depend only on $\mathcal{R}, \ell$. This concludes the proof since ε (and thus K) only depends on $\mathcal{R}, \ell, t$. □

Now we define $t_{\mathrm{b}} = t^{\mathcal{R}}_{\mathrm{b}}$ as the infimum of the set of $t \geqslant 0$ for which (5.1) holds as $n \to \infty$; so (5.1) fails for $t < t_{\mathrm{b}}$. The remark after the proof of Lemma 3.2.1 in Chapter 3 implies that for $t > 1$ we whp have $L_1(G^{\mathcal{R}}_{tn}) \geqslant cn$ for $c = c(\ell, t) > 0$, yielding $S(G^{\mathcal{R}}_{tn}) \geqslant [L_1(G^{\mathcal{R}}_{tn})]^2/n \geqslant c^2 n$; so $t_{\mathrm{b}} \leqslant 1$. Furthermore, for $t < [\ell(\ell-1)]^{-1}$ an application of Theorem 5.2.1 to the empty graph $F = G^{\mathcal{R}}_0$ on n vertices with $\sigma = t$ and $L = 1$ (similar as in the proof of Lemma 5.2.2) readily shows $S(G^{\mathcal{R}}_{\sigma n}) \leqslant \tilde{L}$ whp, so $t_{\mathrm{b}} \geqslant [\ell(\ell-1)]^{-1}$. Now suppose that (5.1) fails for $t = t_{\mathrm{b}}$. The proof of Lemma 5.2.2 then shows that whp (5.19)–(5.21) hold for $i = t_{\mathrm{b}}n$, and that $\chi(\varphi, t_{\mathrm{b}}) < \overline{L} - 2$. It follows that we can apply Theorem 5.2.1 with $\sigma = [\ell^2\overline{L}]^{-1}$ and $L = \overline{L}$; this implies $S(G^{\mathcal{R}}_{(t_{\mathrm{b}}+\sigma)n}) \leqslant \tilde{L}$ whp, contradicting the definition of t_{b}. So, since the susceptibility is monotone increasing, it follows that (5.1) holds for all $t \geqslant t_{\mathrm{b}}$. Combining our findings, Lemma 5.2.2, (5.16) and (5.17) now yield Theorem 5.1.1 with $\rho_k(t) = \varphi(k, t)$.

5.2.3 Branching processes preliminaries

The following basic results for branching processes will be used in the proof of Theorem 5.2.1. They are similar to Theorems 3.2 and 3.3 in [74], where they are attributed to much earlier results of Crámer. Given a non-negative integer valued random variable X, let $F_X(z) = \mathbb{E}z^X$ denote the *(probability) generating function* of X. Note that $F_X(z)$ is convex and monotone increasing for $z \geqslant 0$.

The first lemma essentially states that a two-generation branching process has (uniform) exponential tails provided that the generating function of each offspring distribution has radius of convergence strictly larger than one (and thus also exhibits exponential decay).

Lemma 5.2.3. *Let $X, Y \geqslant 0$ be integer valued random variables with $F_X(\alpha) \leqslant A$ and $F_Y(\beta) \leqslant B$, where $\alpha, \beta > 1$. Let Z be the number of grandchildren in the two-generation branching process in which the root node has X children and then each child, independently, has Y children. There exists $a > 0$ (depending only on α, β, A, B) such that $\mathbb{P}(Z \geqslant s) \leqslant Ae^{-as}$ for all $s \geqslant 0$.*

Proof. Pick $C \geqslant \max\{B, 2\}$ such that $x = 1 + (\alpha-1)(\beta-1)/(C-1) \leqslant \beta$. Using $F_Y(1) = 1$ and $F_Y(\beta) \leqslant B \leqslant C$, convexity yields $F_Y(z) \leqslant [(z-1)C + (\beta - z)]/(\beta - 1)$ for all $z \in [1, \beta]$. So, by choice of x we have $F_Y(x) \leqslant \alpha$. Observing that $F_Z(z) = F_X(F_Y(z))$, using monotonicity we obtain $F_Z(x) = F_X(F_Y(x)) \leqslant F_X(\alpha) \leqslant A$. Since $x > 1$ implies $F_Z(x) \geqslant \mathbb{P}(Z \geqslant s)x^s$ for every $s \geqslant 0$, we deduce $\mathbb{P}(Z \geqslant s) \leqslant Ae^{-as}$ for $a = \log x > 0$, completing the proof. □

The second lemma is a standard result for subcritical Galton–Watson branching process: these exhibit (uniform) exponential decay if the offspring distribution itself has (uniform) exponential tails.

Lemma 5.2.4. *Let $Z \geqslant 0$ be an integer valued random variable with $\mathbb{E}Z \leqslant \mu < 1$ and $F_Z(\beta) \leqslant B$, where $\beta > 1$. Let T be the total size of the Galton–Watson branching process in which each node, independently, has Z children. There exist $\delta > 1$ and $D > 0$ (depending only on β, B, μ) such that $F_T(\delta) \leqslant D$.*

Proof. Let $f(t) = \mathbb{E}(e^{t(Z-1)})$. Observe that $f(0) = 1$ and $f'(0) = \mathbb{E}(Z-1) \leqslant \mu - 1$. As in the proof of Lemma 5.2.3, $F_Z(\beta) \leqslant B$ for $\beta > 1$ yields $\mathbb{P}(Z \geqslant s) \leqslant B\beta^{-s}$, which in turn readily implies that for some $C = C(\beta, B)$ we have $f''(t) = \mathbb{E}((Z-1)^2 e^{t(Z-1)}) \leqslant C$ for all $0 \leqslant t \leqslant (\log \beta)/2$, say. So, using Taylor's theorem, for $0 \leqslant t \leqslant (\log \beta)/2$ we deduce $f(t) \leqslant 1 + (\mu - 1)t + Ct^2/2 = h(t)$. Let $x = \min\{(\log \beta)/2, (1-\mu)/C\} > 0$, and observe that $c = \max\{h(x), 1/2\} > 0$ satisfies $f(x) \leqslant c < 1$. Exploring the branching process tree as usual in breadth-first search order, we see that $T > s$ implies $\sum_{i=1}^{s} Z_i \geqslant s$, where the Z_i are independent copies of Z (corresponding to the number of children of the i-th node). Now, using Markov's inequality and independence of the Z_i, for every $s \geqslant 0$ we obtain $\mathbb{P}(T > s) \leqslant \mathbb{E}(e^{x(\sum_{i=1}^{s} Z_i)})e^{-xs} = f(x)^s \leqslant c^s$. Finally, picking $1 < \delta < 1/c$, it follows that $F_T(\delta) \leqslant D = D(\delta, c)$, as claimed. □

5.2.4 Proof of Theorem 5.2.1

The proof of Theorem 5.2.1 relies on a two-round exposure argument: we first reveal the random tuples selected, and afterwards expose their order of appearance. It will be convenient to work with a continuous-time random graph model, where the n^ℓ tuples arrive according to independent Poisson processes with rates $1/n^{\ell-1}$. So tuples appear with rate n, and each tuple is chosen uniformly at random and independently of all previous choices. Let E_t denote the set of tuples which arrive in $[0, t]$; so $|E_t| \sim \mathrm{Po}(tn)$. Observe that for each tuple $\underline{u} \in [n]^\ell$ the number $A_{\underline{u}}(t)$ of its arrivals in $[0, t]$ satisfies $A_{\underline{u}}(t) \sim \mathrm{Po}(t/n^{\ell-1})$, and that these random variables are independent for different tuples. Furthermore, writing $x = t/n^{\ell-1}$ and using $e^{-x} \geqslant 1 - x$ twice, note that for $\ell \geqslant 2$ we have

$$\mathbb{P}(A_{\underline{u}}(t) \geqslant 2) = 1 - e^{-x} - xe^{-x} \leqslant x(1 - e^{-x}) \leqslant x^2 \leqslant t^2/n^\ell. \tag{5.23}$$

Similarly

$$\mathbb{P}(A_{\underline{u}}(t) \geqslant 1) = 1 - e^{-x} \leqslant x = t/n^{\ell-1}. \tag{5.24}$$

Starting with F, for each tuple $\underline{u} \in E_t$ we join all $\binom{\ell}{2}$ pairs of vertices by edges, and we denote the resulting graph by H_t. We define $H_t^{\mathcal{R}}$ as the graph which we obtain by starting with F, and then presenting the tuples to $\mathcal{R}$ (together with the component sizes of the vertices) in a random order, always updating the graph according to the decisions of $\mathcal{R}$ (adding the pairs selected by $\mathcal{R}$). Since conditioned on $|E_t| = i$ we have i tuples chosen independently and uniformly at random, it follows that

$$\mathbb{E}(N_k(H_t^{\mathcal{R}}) \mid |E_t| = i) = \mathbb{E}N_k(F_i^{\mathcal{R}}). \tag{5.25}$$

Furthermore, mimicking the proof of *Pittel's inequality* (see e.g. [21]) for $0 < tn < n^\ell$, a short calculation shows that for any graph property $\mathcal{Q}$ we have

$$\mathbb{P}(F_{tn}^{\mathcal{R}} \notin \mathcal{Q}) \leqslant 3\sqrt{tn} \cdot \mathbb{P}(H_t^{\mathcal{R}} \notin \mathcal{Q}). \tag{5.26}$$

In the following sections we always tacitly assume that the assumptions of Theorem 5.2.1 hold and consider $t = t(n)$ satisfying

$$0 \leqslant t \leqslant \sigma \leqslant 1, \tag{5.27}$$

where $\sigma \leqslant 1$ follows from (5.9). Furthermore, unless stated otherwise, we will use the continuous-time random graph models H_t and $H_t^{\mathcal{R}}$. For later usage let

$$U = (\log n)^{6/5}. \tag{5.28}$$

5.2.4.1 Component exploration process for $\ell = 2$

Our main ingredient for analyzing the first exposure round is a certain exploration process. Given a (random) vertex v, it finds all tuples in E_t and components of F that are 'relevant' in the second exposure round for determining $|C_v(H_t^{\mathcal{R}})|$, where we write $C_v(G)$ for the set of vertices of G that are in the same component as v. As certain details are rather technical for Achlioptas processes, here we first outline some of the basic ideas and techniques for the simpler case of an Erdős–Rényi evolution starting from an initial graph F (in this special case similar ideas were used by Spencer and Wormald [74]). This formally corresponds to the special case $\ell = 2$ and the rule which always adds the offered pair $v_1 v_2$ to the evolving graph; so $H_t = H_t^{\mathcal{R}}$.

One major difference to the Erdős–Rényi case (where we start with an empty graph on n vertices) is that here we have two sources of edges: (i) the initial graph F and (ii) the random pairs in E_t. As edges of type (i) are deterministic and those of type (ii) are random, our exploration process explicitly considers them separately. In the first round we start with a randomly chosen v and mark all $u \in C_v(F)$ as *reached*; all other vertices are *unreached*. In each later round we sequentially go through the vertices w reached in the previous round (the order does not matter here) and determine all its so far unreached neighbours u in E_t (corresponding to pairs $(u_1, u_2) \in E_t$ containing u and w), each time marking all $\tilde{u} \in C_u(F)$ as reached. Note that upon termination $C_v(H_t)$ equals the set of all reached vertices.

The previous procedure yields an associated 'exploration tree' $\mathcal{T}_v(H_t)$ in a rather natural way: loosely speaking, u is a child of w if u was 'reached' via w. With an eye to the upcoming analysis for size rules, here we already introduce different types of nodes: *vertex nodes*, *component nodes*, and *root nodes*. More precisely, we define $\mathcal{T}_v(H_t)$ inductively as follows: it has a root node v, whose children are vertex nodes $u \in C_v(F)$. Then, given any vertex node w, each of its so far unreached neighbours u in E_t yields a component node as a child, which in turn has vertex nodes $\tilde{u} \in C_u(F)$ as children. It follows that the set of all vertex nodes in $\mathcal{T}_v(H_t)$ equals $C_v(H_t)$. The main point is that, even after ignoring all labels, the *structure* of $\mathcal{T}_v(H_t)$ is enough to determine $|C_v(H_t)|$.

The key idea is now to approximate $\mathcal{T}_v(H_t)$ by an 'idealized' branching process, similar as in the 'classical' Erdős–Rényi case (exploiting, as usual, that by construction every edge is tested at most once). Recall that in $\mathcal{T}_v(H_t)$ already reached vertices are 'ignored'. So, noting that endpoints of random pairs in E_t correspond to random vertices, and that each edge gives rise to two *ordered* tuples, it seems plausible that $\mathcal{T}_v(H_t)$ is dominated by (may be regarded as a subset of) a branching process $\mathfrak{T}_{v,t}$ where (ignoring for simplicity the root and all labels) every vertex node, independently, has Po$(2t)$ component nodes as children, each of which in turn, independently, has N vertex node descendants, where $N \sim |C_u(F)|$ for a randomly chosen vertex u. Now, using (5.9) and (5.12) each vertex node has in expectation $2t \cdot S(F) \leqslant 2\sigma L < 1$ vertex nodes as grandchildren, so we expect that $\mathfrak{T}_{v,t}$ resembles a subcritical branching process which has $O(\log n)$ size with very high probability. From this it follows that $\mathcal{T}_v(H_t)$ and $\mathfrak{T}_{v,t}$ are both small and have similar offspring distribution (as not too many vertices are reached and thus ignored), so it seems plausible that we can couple them so that they agree whp. Note that $\mathfrak{T}_{v,t}$ still depends on n and the initial graph F. Define $\mathbb{P}(R = k) = \varphi(k)$, where φ is given by Theorem 5.2.1. The point is now that using (5.10) it follows that R is very close to N. So, denoting by $\mathfrak{X}_{\varphi,t}$ the 'idealized' version of $\mathfrak{T}_{v,t}$ where we use R instead of N, the former considerations suggest that there is a coupling such that whp $\mathfrak{T}_{v,t} \cong \mathfrak{X}_{\varphi,t}$ holds (ignoring the labels of the vertices). To summarize, we just outlined that using the 'intermediate' process $\mathfrak{T}_{v,t}$ we can couple $\mathcal{T}_v(H_t)$ and $\mathfrak{X}_{\varphi,t}$ so that they typically agree up to isomorphisms. Consequently, the distribution of $|C_v(H_t)|$ can be approximated using $\mathfrak{X}_{\varphi,t}$, which does *not* depend on n or F.

In the above construction and analysis we used in essential ways that in each round only one pair of vertices is chosen and connected by an edge. In contrast, when considering Achlioptas processes several vertices $\underline{v} = (v_1, \ldots, v_\ell)$ are chosen in each round, and only a subset of the edges between these vertices is added to the evolving graph. Furthermore, in the second exposure round the order in which the tuples $\underline{v}$ are presented matters (as well as the order of the vertices in each tuple). This motivates the more involved exploration processes used in the next section, whose associated exploration tree captures more detailed structural information (also using more types of nodes).

5.2.4.2 Component exploration process (the general case)

In this section we consider the first exposure round, where the selected set of tuples E_t is revealed. Note that this defines H_t, which we obtain by starting with F and then joining all ℓ vertices of each tuple in E_t by edges. Using a natural variant of the standard neighbourhood exploration process, for any vertex v we can determine $C_v(H_t)$ as follows. First we determine

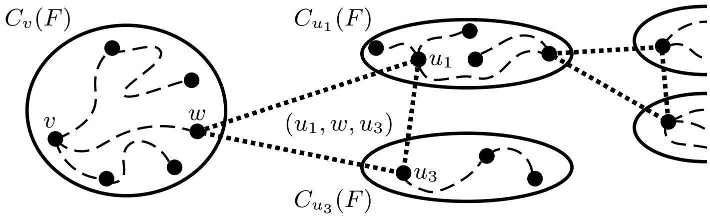

Figure 5.1: Example of the neighbourhood exploration process for $\ell = 3$. It determines $C_v(H_t)$ by first finding other vertices in the same component of F, then finding tuples containing them; afterwards it repeats this procedure for the new vertices in those tuples, and so on.

$C_v(F)$, i.e., find all other vertices which are in the same component of F as v. Then, for each $w \in C_v(F)$ we find all tuples $\underline{u} = (u_1, \ldots, u_\ell) \in E_t$ containing w, and repeat the same procedure (recursively) for each $u_j \neq w$, see Figure 5.1. Observe that for determining $C_v(H_t)$ it suffices to consider only those vertices $u_j \neq w$ which we have not already reached in some previous exploration step.

In the analysis it is easier to start with a *random* vertex v and break down the above exploration process into small steps, constructing an associated *exploration tree* $\mathcal{T}_{v,t} = \mathcal{T}_{v,t}(F)$. As we shall see, one key property of $\mathcal{T}_{v,t}$ is that we can (typically) reconstruct the vertices and components which have been reached, as well as the tuples which have been 'tested' so far. The vertices of each exploration tree have different types: *vertex nodes*, *component nodes* and *tuple nodes* will represent vertices, components of F and ℓ-tuples, respectively. For technical reasons we also have *root nodes* and *index nodes*. We denote the vertex nodes of $\mathcal{T}_{v,t}$ by $\mathcal{V}_{v,t}$.

As mentioned above, our exploration starts with a random vertex v, which serves as the root node of $\mathcal{T}_{v,t}$, see Figure 5.2. Next we (deterministically) find all vertices $w \in C_v(F)$ and then add the vertex nodes w as children of the root. In the following we sequentially traverse each level containing vertex nodes (which essentially corresponds to a breadth first search). Given a vertex node w, we add ℓ index nodes $w_1, \ldots, w_\ell$ as children, where w_j is an index node of type j. For each $j = 1, \ldots, \ell$ we sequentially test for the presence and multiplicity of all so far untested tuples $\underline{u} = (u_1, \ldots, u_\ell)$ with $u_j = w$; we denote the resulting multiset of found tuples by $S_{j,w}$. Now we sequentially traverse the $\underline{u} \in S_{j,w}$. For each such $\underline{u} = (u_1, \ldots, u_\ell)$ we add a tuple node $\underline{u}$ and traverse the u_i with $i \neq j$ sequentially.

For each $i \neq j$, we add a component node u_i of type $\lambda_j(i)$ as a child of $\underline{u}$, where $\lambda_j(i) = i$ for $i < j$ and $\lambda_j(i) = i - 1$ for $i > j$ (so that the component nodes $\{u_i\}$ with $i \neq j$ have types $1, \ldots, \ell - 1$). If u_i is already contained in $\mathcal{T}_{v,t}$ then we 'ignore' this component node. Otherwise we add vertex nodes $w \in C_{u_i}(F)$ as children of u_i, see Figure 5.2. Note that $C_v(H_t)$ consists exactly of the union of all vertex nodes of $\mathcal{T}_{v,t}$, so

$$C_v(H_t^{\mathcal{R}}) \subseteq C_v(H_t) = \mathcal{V}_{v,t}. \tag{5.29}$$

The main point is that whenever no component nodes are ignored, then from $\mathcal{T}_{v,t}$ we can reconstruct all explored tuples (in E_t) and components (of F), which for size rules are the only ones relevant for determining the size of $C_v(H_t^{\mathcal{R}})$. In fact, up to relabellings, we can reconstruct these tuples and the relevant component sizes of F *without* looking at the vertex labels (the tree structure, including the node types, is enough). Motivated by this we say that $S_{j,w}$ is *bad* if one of the following conditions hold:

- $S_{j,w}$ contains some tuple $\underline{u} = (u_1, \ldots, u_\ell)$ multiple times.
- $S_{j,w}$ contains a tuple $\underline{u} = (u_1, \ldots, u_\ell)$ where u_i with $i \neq j$ is already a vertex node of $\mathcal{T}_{v,t}$ constructed so far.
- $S_{j,w}$ contains a tuple $\underline{u} = (u_1, \ldots, u_\ell)$ where u_i and u_k with $i \neq k$ are in the same component of F (note that this holds for $u_i \in C_w(F)$ for $i \neq j$).
- $S_{j,w}$ contains tuples $\underline{u} = (u_1, \ldots, u_\ell)$ and $\underline{v} = (v_1, \ldots, v_\ell)$ for which u_i and v_k with $i, k \neq j$ are in the same component of F.

Otherwise $S_{j,w}$ is *good*; Observe that if $S_{j,w}$ is good, then in $\mathcal{T}_{v,t}$ none of w's component node descendants u_i with $\underline{u} = (u_1, \ldots, u_\ell) \in S_{j,w}$ are ignored. For this reason we call $\mathcal{T}_{v,t}$ *good* if every $S_{j,w}$ is good. In the following we estimate the probability that $S_{j,w}$ is *bad*. Clearly, there are at most $n^{\ell-1}$ different tuples with $u_j = w$. Recalling that $\mathcal{V}_{v,t}$ denotes the vertex nodes of $\mathcal{T}_{v,t}$, there are at most $\ell n^{\ell-2}|\mathcal{V}_{v,t}|$ different tuples satisfying the second condition, and at most $\ell^2 n^{\ell-2}|L_1(F)|$ tuples to which the third condition applies. Similarly, there are at most $\ell^2 n^{2(\ell-2)+1}|L_1(F)|$ pairs of tuples which satisfy the last condition. Recall that the

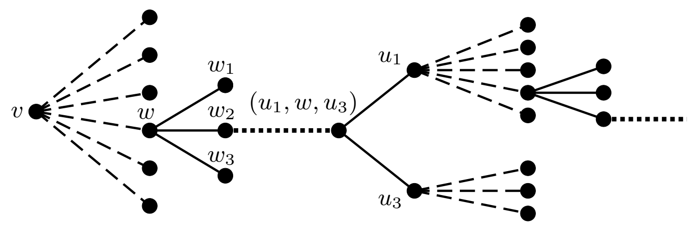

Figure 5.2: Example of the exploration tree $\mathcal{T}_{v,t}$ for $\ell = 3$. The children of the root vertex v are $w \in C_v(F)$ (vertex nodes), which in turn each have children w_1, w_2, w_3 (index nodes of types $1, 2, 3$). Every w_j has all (so far untested) tuples $\underline{u} = (u_1, u_2, u_3) \in E_t$ with $u_j = w$ as children (tuple nodes), whose descendants are component nodes u_i with $i \neq j$ (of types $1, 2$). If u_i is not already a vertex node of $\mathcal{T}_{v,t}$, then its children are $w \in C_{u_i}(F)$ (vertex nodes), for which we repeat the above construction.

random variables $A_{\underline{u}}(t)$, which count the number of times $\underline{u}$ is in E_t, are independent for different tuples $\underline{u}$. So, using (5.23), (5.24) and $t \leqslant 1$, whenever $\max\{|\mathcal{V}_{v,t}|, |L_1(F)|\} \leqslant U$ holds we see that the probability of $S_{j,w}$ being bad is at most

$$n^{\ell-1} \cdot t^2/n^{\ell} + 2\ell^2 n^{\ell-2} U \cdot t/n^{\ell-1} + \ell^2 n^{2\ell-3} U \cdot t^2/n^{2(\ell-1)} \leqslant 4\ell^2 U/n. \tag{5.30}$$

To understand the structural properties of $\mathcal{T}_{v,t}$ it will be useful to compare it with a closely related process that is simpler to analyze. Recall that when determining the $S_{j,w}$ we only consider so far untested tuples. Thus each $S_{j,w}$ is dominated (with respect to the subset relation) by $\tilde{S}_{j,w}$, where for each of the $n^{\ell-1}$ tuples $\underline{u} = (u_1, \ldots, u_\ell)$ with $u_j = w$, independently, the number of its arrivals is given by a $\mathrm{Po}(t/n^{\ell-1})$ distribution. There is a natural coupling between $S_{j,w}$ and $\tilde{S}_{j,w}$ which only fails if $\tilde{S}_{j,w}$ contains $\underline{u}$ which are forbidden for $S_{j,w}$. Since each of these 'bad' tuples contains at least one vertex from $\mathcal{V}_{v,t}$, there are at most $\ell|\mathcal{V}_{v,t}|n^{\ell-2}$ of them. So, with (5.24) and $t \leqslant 1$ in mind, by considering the probability that $\tilde{S}_{j,w}$ selects at least one of them, whenever $|\mathcal{V}_{v,t}| \leqslant U$ holds it follows that

$$\mathrm{d_{TV}}\big(S_{j,w}, \tilde{S}_{j,w}\big) \leqslant \ell|\mathcal{V}_{v,t}|n^{\ell-2} \cdot t/n^{\ell-1} \leqslant \ell U/n. \tag{5.31}$$

We now define $\mathfrak{T}_{v,t} = \mathfrak{T}_{v,t}(F)$ similarly to $\mathcal{T}_{v,t}$: we employ the same construction except that we use (independent copies of) $\tilde{S}_{j,w}$ instead of $S_{j,w}$ and always proceed as if $\tilde{S}_{j,w}$ is good. Since each $S_{j,w}$ is dominated by (may be regarded as a subset of) $\tilde{S}_{j,w}$, it follows that

$\mathcal{T}_{v,t}$ is dominated by $\mathfrak{T}_{v,t}$ with respect to the subgraph relation. Denoting the set of vertex nodes of $\mathfrak{T}_{v,t}$ by $\mathfrak{V}_{v,t}$, we see that $\mathcal{V}_{v,t}$ is dominated by $\mathfrak{V}_{v,t}$.

The next lemma states that the number of vertex nodes in $\mathcal{T}_{v,t}$ and $\mathfrak{T}_{v,t}$ have (uniform) exponential decay.

Lemma 5.2.5. *Suppose that* (5.9) *and* (5.11)–(5.12) *hold with* $\beta > 1$. *There exist* $a, A > 0$ *(depending only on* $\ell, L, \sigma, \beta, B$*) such that for all* $0 \leqslant t \leqslant \sigma$ *and* $s \geqslant 0$ *we have* $\mathbb{P}(|\mathcal{V}_{v,t}| \geqslant s) \leqslant \mathbb{P}(|\mathfrak{V}_{v,t}| \geqslant s) \leqslant Ae^{-as}$, $\mathbb{E}N_{\geqslant s}(H_t^{\mathcal{R}}) \leqslant Ae^{-as}n$ *and* $\mathbb{P}(L_1(H_t^{\mathcal{R}}) \geqslant s) \leqslant Ae^{-as}n$.

Before giving the proof of this result, which is based on branching process arguments, we use it to show that $\mathcal{T}_{v,t}$ and $\mathfrak{T}_{v,t}$ can be coupled so that they typically agree. Note that at distance $4i+1, 4i+2, 4i+3, 4i+4$ from the root $\mathcal{T}_{v,t}$ and $\mathfrak{T}_{v,t}$ always have vertex, index, tuple and component nodes.

Lemma 5.2.6. *Suppose* $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ *and that the assumptions of Theorem 5.2.1 as well as* (5.27) *hold. There exists a coupling of* $\mathcal{T}_{v,t}$ *and* $\mathfrak{T}_{v,t}$ *so that with probability at least* $1 - (\log n)^4/n$ *we have* $\mathcal{T}_{v,t} = \mathfrak{T}_{v,t}$ *and* $\mathcal{T}_{v,t}$ *is good.*

Proof. We write T^i for the *restriction* of a rooted tree T to all vertices within distance at most i from the root. Let $\mathcal{V}_{v,t}^i$ and $\mathfrak{V}_{v,t}^i$ denote the vertex nodes in $\mathcal{T}_{v,t}^i$ and $\mathfrak{T}_{v,t}^i$, respectively. Recall that $U = (\log n)^{6/5}$. Since $\beta > 1$, note that $L_1(F) \leqslant U$ follows from (5.11) for $n \geqslant n_0(B, \beta)$.

We inductively couple $\mathcal{T}_{v,t}^{4i+1}$ and $\mathfrak{T}_{v,t}^{4i+1}$ for $0 \leqslant i \leqslant U$ so that with probability at least $1 - i \cdot 5\ell^3U^2/n$ we have either $\max\{|\mathcal{V}_{v,t}^{4i+1}|, |\mathfrak{V}_{v,t}^{4i+1}|\} \geqslant U$, or $\mathcal{T}_{v,t}^{4i+1} = \mathfrak{T}_{v,t}^{4i+1}$ with all $S_{j,w}$ of $\mathcal{T}_{v,t}^{4i+1}$ being good. The base case $i = 0$ is straightforward, as both use the same procedure for generating the root and its children. Now suppose that we have constructed $\mathcal{T}_{v,t}^{4i+1}$ and $\mathfrak{T}_{v,t}^{4i+1}$ coupled as above. In the following we sequentially consider vertex nodes w at distance $4i+1$ from the root and extend the coupling to their descendants with distance up to $4(i+1)+1$; here we clearly may assume $|\mathcal{V}_{v,t}^{4i+1}| = |\mathfrak{V}_{v,t}^{4i+1}| < U$. For each vertex node w we create ℓ index nodes $w_1, \dots w_\ell$ (of types $1, \dots, \ell$). We abandon our coupling whenever we have found more than U vertex nodes (in which case we are done), so (5.31) holds. Thus we can couple $S_{j,w}$ and $\tilde{S}_{j,w}$ so that they agree with probability at least $1 - \ell U/n$.

Now we also abandon our coupling whenever $S_{j,w}$ is bad, which happens with probability at most $4\ell^2 U/n$ by (5.30). The point is that given good $S_{j,w} = \tilde{S}_{j,w}$, in both cases the same deterministic construction is used for generating the descendants of w_j with distance up to $4(i+1)+1$ from the root. So, by repeating this for $w_1, \dots, w_\ell$, with probability at least $1 - 5\ell^3 U/n$ we can couple the descendants of w with distance up to $4(i+1)+1$ from the root. Since we follow this argument for each of the at most U vertex nodes at distance $4i$ from the root, we see that we can extend our coupling to $\mathcal{T}_{v,t}^{4(i+1)+1}$ and $\mathfrak{T}_{v,t}^{4(i+1)+1}$ with probability at least $1 - 5\ell^3 U^2/n$, establishing the claim.

Finally, by Lemma 5.2.5 we know that $V = \max\{|\mathcal{V}_{v,t}|, |\mathfrak{V}_{v,t}|\} < U/10$ holds with probability at least, say, $1 - n^{-9}$ for $n \geqslant n_0(a, A)$. This together with the above coupling completes the proof (as there are no vertex nodes with distance larger than $4V + 1$ from the root). $\square$

We now introduce an idealized 'infinite' version $\mathfrak{X}_{\varphi,t}$ of the exploration tree that is defined *without* reference to n or F, and in which 'bad' things (such as 'ignored' component nodes) cannot happen by definition. Let R be the random variable with $\mathbb{P}(R = k) = \varphi(k)$ for each $k \geqslant 1$, where φ is given by Theorem 5.2.1. We start $\mathfrak{X}_{\varphi,t}$ with a root node and add R vertex nodes as children. Then, given any vertex node, we deterministically create ℓ children (index nodes of types $1, \dots, \ell$). Each of these, independently, has $Z \sim \mathrm{Po}(t)$ children (tuple nodes). For each of these grandchildren we assign again (deterministically) $\ell - 1$ children (component nodes of types $1, \dots, \ell - 1$). All of these, independently, give birth to R many descendants (vertex nodes).

For our subsequent analysis it will be key to observe that if we are only interested in equality up to isomorphisms, then we can generate $\mathfrak{T}_{v,t}$ in a more convenient way, similarly to $\mathfrak{X}_{\varphi,t}$. Indeed, using standard properties of Poisson processes and noting that selecting a uniform tuple $\underline{u} = (u_1, \dots, u_\ell)$ with $u_j = w$ is equivalent to picking $\ell - 1$ random vertices, we can generate the descendants of w_j constructed by $\tilde{S}_{j,w}$ using the following three-generation tree process: the root has $Z \sim \mathrm{Po}(t)$ children (tuple nodes); then for each of the resulting children we construct (deterministically) $\ell - 1$ grandchildren (component nodes of types $1, \dots, \ell - 1$), which each in turn give birth to N descendants (vertex nodes), where $N \sim |C_u(F)|$ for a uniformly and independently chosen vertex u. Comparing the resulting construction with $\mathfrak{X}_{\varphi,t}$, it follows that we can generate $\mathfrak{T}_{v,t}$ up to relabellings in the same way as $\mathfrak{X}_{\varphi,t}$, with the only difference that we use N instead of R.

Proof of Lemma 5.2.5. Since $\mathcal{V}_{v,t}$ is dominated by (may be regarded as a subset of) $\mathfrak{V}_{v,t}$, we have $\mathbb{P}(|\mathcal{V}_{v,t}| \geqslant s) \leqslant \mathbb{P}(|\mathfrak{V}_{v,t}| \geqslant s)$. Using this inequality, we claim that it is enough to prove existence of $a, A > 0$ (depending only on $\ell, L, \sigma, \beta, B$) satisfying

$$\mathbb{P}(|\mathfrak{V}_{v,t}| \geqslant s) \leqslant Ae^{-as} \quad \text{for all } s \geqslant 0. \tag{5.32}$$

Indeed, recall that v is chosen uniformly at random, so that $\mathbb{P}(|C_v(H_t^{\mathcal{R}})| \geqslant s \mid H_t^{\mathcal{R}} = G) = N_{\geqslant s}(G)/n$. Taking expectations, we see that $\mathbb{E}N_{\geqslant s}(H_t^{\mathcal{R}}) = n\mathbb{P}(|C_v(H_t^{\mathcal{R}})| \geqslant s)$. Using (5.29) we have $|C_v(H_t^{\mathcal{R}})| \leqslant |\mathcal{V}_{v,t}|$, so $\mathbb{P}(|C_v(H_t^{\mathcal{R}})| \geqslant s) \leqslant Ae^{-as}$ by (5.32). Now Markov's inequality gives $\mathbb{P}(L_1(H_t^{\mathcal{R}}) \geqslant s) \leqslant Ae^{-as}n$.

In the remainder we establish (5.32) using Lemmas 5.2.3 and 5.2.4. Let Z_j be independent copies of $Z \sim \text{Po}(t)$, and let $v_{j,r,k}$ be uniformly and independently chosen random vertices. We henceforth construct $\mathfrak{T}_{v,t}$ up to relabellings, as described in the paragraph proceeding this proof. Given a vertex node w with distance $4i+1$ from the root, in this tree construction it has

$$W = \sum_{1 \leqslant j \leqslant \ell} \sum_{1 \leqslant r \leqslant Z_j} \sum_{1 \leqslant k \leqslant \ell - 1} |C_{v_{j,r,k}}(F)|$$

vertex node descendants at distance $4(i+1)+1$ from the root, where $\mathbb{E}(W) = \ell t(\ell-1)S(F) \leqslant \ell\sigma(\ell-1)L < 1$ due to $t \leqslant \sigma$ and (5.9). Note that $F_W(z) = [F_Z([F_N(z)]^{\ell-1})]^{\ell}$, where $N \sim |C_u(F)|$ for a uniformly chosen vertex u. By (5.11) we have $[F_N(\beta)]^{\ell-1} \leqslant B^{\ell-1}$. Now, since $Z \sim \text{Po}(t)$ and $0 \leqslant t \leqslant \sigma$, it easily follows that $F_Z(z) = e^{t(z-1)} \leqslant e^{\sigma z}$ for $z \geqslant 0$, so $F_W(\beta) \leqslant \tilde{B} = \tilde{B}(\ell, \sigma, B)$. Let W^+ be the size of the Galton–Watson branching process in which each node, independently, has W children. Lemma 5.2.4 yields $F_{W^+}(\delta) \leqslant D$, where $\delta > 1$ and $D > 0$ depend only on $\ell, L, \sigma, \beta, \tilde{B}$. Since the distribution of W does not depend on the w or i considered above, it in particular follows that each vertex node with distance 1 from the root has W^+ vertex node descendants in $\mathfrak{T}_{v,t}$.

Finally, note that $\mathfrak{T}_{v,t}$ starts with a root vertex which gives birth to N vertex node children, each of whose vertex nodes descendants is given by independent copies of W^+. With this in mind $|\mathfrak{V}_{v,t}| \sim T$, where T is a two-generation branching process where the root has N children, and then each of these, independently, has W^+ children. Recall that $F_N(\beta) \leqslant \tilde{B}$ and $F_{W^+}(\delta) \leqslant D$ for $\beta, \delta > 1$ and $\tilde{B}, D > 0$. So, Lemma 5.2.3 yields (5.32) for $A = \tilde{B}$ and $a > 0$ depending only on $\beta, \delta, \tilde{B}, D$. As explained, this completes the proof. □

Recall that $\mathfrak{X}_{\varphi,t}$ uses the same construction as $\mathfrak{T}_{v,t}$, with the difference that it employs R instead of N. When establishing the exponential decay in the proof of Lemma 5.2.5, note that the only properties of N used are $\mathbb{E}N = S(F) \leqslant L$ and $F_N(\beta) \leqslant B$. Since $\mathbb{E}R = \chi(\varphi) \leqslant L$ and $F_R(\beta) \leqslant B$ by (5.6)–(5.8), the same argument thus carries over word-by-word when applied to the vertex nodes of $\mathfrak{X}_{\varphi,t}$, which we denote by $\mathfrak{V}_{\varphi,t}$.

Lemma 5.2.7. *Suppose that* (5.6)–(5.8) *and* (5.9) *hold with* $\beta > 1$. *There exist* $a, A > 0$ *(depending only on* $\ell, L, \sigma, \beta, B$*) such that for all* $0 \leqslant t \leqslant \sigma$ *and* $s \geqslant 0$ *we have* $\mathbb{P}(|\mathfrak{V}_{\varphi,t}| \geqslant s) \leqslant Ae^{-as}$, *where* a, A *are defined in the same way as in Lemma 5.2.5.* □

After these preparations, we are now ready to show that we can couple $\mathcal{T}_{v,t}$ and $\mathfrak{X}_{\varphi,t}$ so that they typically agree up to isomorphisms (by using $\mathfrak{T}_{v,t}$ as an 'intermediate' process).

Lemma 5.2.8. *Suppose* $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ *and that the assumptions of Theorem 5.2.1 as well as* (5.27) *hold. There exists a coupling of* $\mathcal{T}_{v,t}$ *and* $\mathfrak{X}_{\varphi,t}$ *so that with probability at least* $1 - (\log n)^{C+5} n^{-1/2}$ *we have* $\mathcal{T}_{v,t} \cong \mathfrak{X}_{\varphi,t}$ *and* $\mathcal{T}_{v,t}$ *is good.*

Proof. Recall that $U = (\log n)^{6/5}$. By Lemma 5.2.6 it suffices to couple $\mathfrak{T}_{v,t}$ and $\mathfrak{X}_{\varphi,t}$ so that with probability at least $1 - 4\ell^2 U^4 (\log n)^C n^{-1/2}$ we have $\mathfrak{T}_{v,t} \cong \mathfrak{X}_{\varphi,t}$. To this end we use a similar but simpler argument as in the proof of Lemma 5.2.6, inductively extending our coupling from distance $4i+1$ to $4(i+1)+1$ from the root. As before, using Lemma 5.2.5 and 5.2.7 we can safely abandon our coupling whenever we have seen at least U vertex nodes, or when we reach distance U from the root. In the inductive step, the only difference between $\mathfrak{X}_{\varphi,t}$ and $\mathfrak{T}_{v,t}$ is that $\mathfrak{X}_{\varphi,t}$ uses R whereas $\mathfrak{T}_{v,t}$ uses N. Recall that $\mathbb{P}(R = k) = \varphi(k)$ and $\mathbb{P}(N = k) = N_k(F)/n$. It is not difficult to see that (5.7) and (5.11) imply $\mathbb{P}(R \geqslant U) \leqslant n^{-2}$ and $\mathbb{P}(N \geqslant U) = 0$ for $n \geqslant n_0(\beta, B)$. Using these tail estimates together with (5.10), by distinguishing values smaller and larger than U we obtain

$$d_{\mathrm{TV}}\big(R, N\big) \leqslant U \cdot (\log n)^C n^{-1/2} + n^{-2} \leqslant 2U(\log n)^C n^{-1/2}. \tag{5.33}$$

We furthermore may safely abandon our coupling whenever some index node has $Z \geqslant U$ children, since (using $t \leqslant \sigma$) this occurs with probability at most n^{-9} for $n \geqslant n_0(\sigma)$. The point is that this ensures that we only need to couple R and N at most $\ell^2 U^2$ times when going from distance $4i + 1$ to $4(i + 1) + 1$. So, each time we can extend the coupling inductively with probability at least, say, $1 - 3\ell^2 U^3 (\log n)^C n^{-1/2}$. Arguing as in the proof of Lemma 5.2.6, this completes the coupling argument. □

5.2.4.3 Expected component sizes

After analyzing the tuple and component structure induced by E_t, we now consider the second exposure round, where the selected tuples are presented in random order to $\mathcal{R}$. Intuitively, the coupling given by Lemma 5.2.8 allows us to estimate $\mathbb{E}N_k(H_t^{\mathcal{R}})$ using $\mathfrak{X}_{\varphi,t}$. As we shall see, this also carries over to $\mathbb{E}N_k(F_{tn}^{\mathcal{R}})$.

Recall that if the exploration tree $\mathcal{T}_{v,t} \cong T$ is good, then during its construction no component nodes are ignored. As mentioned in Section 5.2.4.2, the key point is that if no nodes are ignored (i.e., all component nodes have at least one child), then from the structure of T (which includes the vertex types) we can reconstruct all tuples in E_t and component sizes of F (up to relabellings) which are relevant for determining $|C_v(H_t^{\mathcal{R}})|$. We denote the corresponding set of tuples and component sizes by $\mathcal{T}_T$ and $\mathcal{C}_T$, respectively. As the above 'reconstruction' procedure only uses the tree-structure of T, it in fact can be applied to any exploration tree in which each component node has at least one child; so, in particular, to $\mathfrak{X}_{\varphi,t} \cong T$. In the following we define $|C^{\mathcal{R}}(T)|$ for any exploration tree T, where we formally set $|C^{\mathcal{R}}(T)| = 0$ if T contains a component node with 0 descendants. Otherwise, we traverse in (uniform) random order the tuples in $\mathcal{T}_T$; for each tuple we present the component sizes of its vertices to $\mathcal{R}$ and update the list of components (and their sizes) according to the decisions of $\mathcal{R}$ (by adding the pairs selected by $\mathcal{R}$). Finally, we define $|C^{\mathcal{R}}(T)|$ as the size of the resulting component which contains the root vertex of T. Since the second exposure round of $H_t^{\mathcal{R}}$ presents the tuples in E_t to $\mathcal{R}$ in random order, a moment's thought reveals that conditional on $\mathcal{T}_{v,t} \cong T$ being good, both $|C_v(H_t^{\mathcal{R}})|$ and $|C^{\mathcal{R}}(\mathcal{T}_{v,t})|$ have exactly the same distribution for size rules. So, for all $k \geqslant 1$ we have

$$\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k \mid \mathcal{T}_{v,t} \cong T \text{ is good}) = \mathbb{P}(|C^{\mathcal{R}}(\mathcal{T}_{v,t})| = k \mid \mathcal{T}_{v,t} \cong T \text{ is good}). \qquad (5.34)$$

Before using this observation to estimate $\mathbb{E}N_k(H_t^{\mathcal{R}})$, we first collect some basic properties of the function ρ, where we set

$$\rho(k,t) = \mathbb{P}(|C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| = k) \quad \text{for all } (k,t) \in \mathbb{N} \times \mathbb{R}^+. \qquad (5.35)$$

Lemma 5.2.9. *Suppose that* (5.6)–(5.8) *and* (5.9) *hold with* $\beta > 1$. *The function* $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ *defined in* (5.35) *depends only on* $\varphi, \mathcal{R}, \ell$ *and satisfies* $\sum_{k \geqslant 1} \rho(k,t) = 1$ *for all* $0 \leqslant t \leqslant \sigma$. *Furthermore, there exist* $a, A > 0$ *(depending only on* $\ell, L, \sigma, \beta, B$*) such that for all* $0 \leqslant t \leqslant \sigma$ *and* $s \geqslant 0$ *we have* $\rho(s,t) \leqslant Ae^{-as}$*, where* a, A *are given by Lemma 5.2.7.*

Proof. The definitions of $C^{\mathcal{R}}(\cdot)$ and of $\mathfrak{X}_{\varphi,t}$ depend only on $\mathcal{R}, \ell$ and on $\varphi, \mathcal{R}, \ell, t$ respectively. So, from (5.35) we see that $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ depends only on $\varphi, \mathcal{R}, \ell$. Since the component containing the root vertex of $\mathfrak{X}_{\varphi,t}$ can only contain vertex nodes of $\mathfrak{X}_{\varphi,t}$, we see that $1 \leqslant |C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| \leqslant |\mathfrak{V}_{\varphi,t}|$ holds, from which $\rho(0,t) = 0$ follows. Furthermore, Lemma 5.2.7 implies $\rho(s,t) \leqslant \mathbb{P}(|\mathfrak{V}_{\varphi,t}| \geqslant s) \leqslant Ae^{-as}$ for all $s \geqslant 1$, where $a, A > 0$ depend only on $\ell, L, \sigma, \beta, B$. Similarly, for all $s \geqslant 0$ we have $\mathbb{P}(\mathfrak{X}_{\varphi,t} \text{ is infinite}) \leqslant \mathbb{P}(|\mathfrak{V}_{\varphi,t}| \geqslant s) \leqslant Ae^{-as}$. But $Ae^{-as} \to 0$ as $s \to \infty$, so $\mathbb{P}(\mathfrak{X}_{\varphi,t} \text{ is infinite}) = 0$, which in turn yields $\sum_{k \geqslant 1} \rho(k,t) = 1$. $\square$

Lemma 5.2.10. *Suppose* $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ *and that the assumptions of Theorem 5.2.1 as well as* (5.27) *hold. We have*

$$\mathbb{E}N_k(H_t^{\mathcal{R}}) = \rho(k,t)n \pm (\log n)^{C+6} n^{1/2} \quad \textit{for all } k \geqslant 1. \qquad (5.36)$$

Proof. Similar as in the proof of Lemma 5.2.5, since v is chosen uniformly at random we have $\mathbb{E}N_k(H_t^{\mathcal{R}}) = n\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k)$. To prove the claim it thus suffices to relate $\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k)$ and $\rho(k,t) = \mathbb{P}(|C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| = k)$. The coupling of Lemma 5.2.8 implies that $\mathfrak{X}_{\varphi,t} \cong \mathcal{T}_{v,t}$ holds with probability at least $1-(\log n)^{C+5}n^{-1/2}$ for $n \geqslant n_0(\ell, L, \sigma, \beta, B)$. Hence

$$\mathbb{P}(|C^{\mathcal{R}}(\mathcal{T}_{v,t})| = k) = \mathbb{P}(|C^{\mathcal{R}}(\mathfrak{X}_{\varphi,t})| = k) \pm 2(\log n)^{C+5}n^{-1/2}.$$

Since this coupling also implies that $\mathcal{T}_{v,t}$ is good, using (5.34) it follows that

$$\mathbb{P}(|C_v(H_t^{\mathcal{R}})| = k) = \mathbb{P}(|C^{\mathcal{R}}(\mathcal{T}_{v,t})| = k) \pm 2(\log n)^{C+5}n^{-1/2}.$$

Finally, combining our findings and recalling (5.35), we readily obtain (5.36). □

Now we relate $H_t^{\mathcal{R}}$ with $F_{tn}^{\mathcal{R}}$ by establishing that $\mathbb{E}N_k(H_t^{\mathcal{R}}) \approx \mathbb{E}N_k(F_{tn}^{\mathcal{R}})$.

Lemma 5.2.11. *Suppose that $0 \leqslant t \leqslant 1$. Then for $n \geqslant n_0(\ell)$ we have*

$$\mathbb{E}N_k(F_{tn}^{\mathcal{R}}) = \mathbb{E}N_k(H_t^{\mathcal{R}}) \pm k(\log n)n^{1/2} \quad \textit{for all } k \geqslant 1. \tag{5.37}$$

Proof. Observe that N_k changes by at most ℓk per step. So, for $r \leqslant s$ we have $\mathbb{E}(N_k(F_s^{\mathcal{R}}) \mid F_r^{\mathcal{R}} = G) = N_k(G) \pm (s-r)\ell k$. Taking expectations and restricting our attention to $r \in \{tn-i, tn\}$ shows that for each $i \geqslant 0$ we have

$$\mathbb{E}N_k(F_{tn\pm i}^{\mathcal{R}}) = \mathbb{E}N_k(F_{tn}^{\mathcal{R}}) \pm \ell k i. \tag{5.38}$$

Set $s = 3\sqrt{n \log n}$. Using $t \leqslant 1$, standard Chernoff bounds yield that $|E_t| = tn \pm s$ with probability at least $1-n^{-2}$ for $n \geqslant n_0$. Combining this with (5.25) and (5.38), we readily obtain

$$\mathbb{E}N_k(H_t^{\mathcal{R}}) = \mathbb{E}N_k(F_{tn}^{\mathcal{R}}) \pm \ell k s \pm n \cdot 2n^{-2},$$

which implies (5.37) for $n \geqslant n_0(\ell)$, with room to spare. □

5.2.4.4 Concentration of component sizes

In this section we establish concentration of $N_k(F_i^{\mathcal{R}})$ around its expected value. The main technical difficulty here is that few changes of the offered tuples might alter many decisions of size rules (as the component sizes observed in later rounds can change); as we shall see, the bounds for $L_1(\cdot)$ implied by Lemma 5.2.5 will be a crucial ingredient for showing that this is typically not the case.

Lemma 5.2.12. *Suppose $n \geqslant n_0(\ell, L, \sigma, \beta, B)$ and that the assumptions of Theorem 5.2.1 hold. With probability at least $1 - n^{-250}$, for every $0 \leqslant i \leqslant \sigma n$ we have*

$$N_k(F_i^{\mathcal{R}}) = \mathbb{E}N_k(F_i^{\mathcal{R}}) \pm (\log n)^2 n^{1/2} \quad \textit{for all } 1 \leqslant k \leqslant (\log n)^2. \tag{5.39}$$

Proof. We sequentially draw σn random tuples and consider two associated graph sequences $F_i^{\mathcal{R}}$ and $F_i^{\mathcal{I}}$, where the 'influence' rule $\mathcal{I}$ in each step simply joins all ℓ randomly chosen vertices by edges. Note that $F_i^{\mathcal{R}} \subseteq F_i^{\mathcal{I}}$ always holds. Let $\mathcal{L}$ denote the event that $L_1(F_{\sigma n}^{\mathcal{I}}) < U = (\log n)^{6/5}$, which by monotonicity implies $L_1(F_i^{\mathcal{I}}) < U$ for all $0 \leqslant i \leqslant \sigma n$. Combining Lemma 5.2.5 with (5.26), for $n \geqslant n_0(a, A, \sigma)$ we have, say,

$$\mathbb{P}(\neg\mathcal{L}) \leqslant 3\sqrt{\sigma n} \cdot \mathbb{P}(L_1(H_\sigma^{\mathcal{I}}) \geqslant U) \leqslant n^{-300}. \tag{5.40}$$

For every $1 \leqslant i \leqslant \sigma n$ let $X_{k,i}$ denote the number of vertices which satisfy $|C_v(F_i^{\mathcal{R}})| = k$ and $|C_v(F_i^{\mathcal{I}})| < U$. When $\mathcal{L}$ holds no vertices are 'ignored' due to $|C_v(F_i^{\mathcal{I}})| \geqslant U$, so we have $X_{k,i} = N_k(F_i^{\mathcal{R}})$. Together with (5.40) this readily gives, say, $\mathbb{E}X_{k,i} = \mathbb{E}N_k(F_i^{\mathcal{R}}) \pm n^{-1}$. So, for $\Delta = U^{3/2} n^{1/2}$ it follows that

$$\mathbb{P}(\{|N_k(F_i^{\mathcal{R}}) - \mathbb{E}N_k(F_i^{\mathcal{R}})| \geqslant 2\Delta\} \cap \mathcal{L}) \leqslant \mathbb{P}(|X_{k,i} - \mathbb{E}X_{k,i}| \geqslant \Delta). \tag{5.41}$$

Note that for every size rule $\mathcal{R}$ the random variable $X_{k,i}$ can be written as $X_{k,i} = f(\underline{v}_1, \ldots, \underline{v}_i)$, where the $\underline{v}_j$ denote the ℓ-tuples generated by the ℓ-vertex process in each step (uniformly and independently). We claim that the function f satisfies $|f(\omega) - f(\tilde{\omega})| \leqslant 4\ell U$ whenever ω and $\tilde{\omega}$ differ in one coordinate, i.e., in one tuple. Assuming that $(\underline{v}_1, \ldots, \underline{v}_i)$ yield $F_i^{\mathcal{R}}$ and $F_i^{\mathcal{I}}$, respectively, let $\tilde{F}_i^{\mathcal{R}}$ and $\tilde{F}_i^{\mathcal{I}}$ denote the graphs which result by changing $\underline{v}_j$ to $\tilde{\underline{v}}_j$. Since $F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{I}}$ only differ in the edges induced by $\underline{v}_j$ and $\tilde{\underline{v}}_j$, there is a set of vertices W containing at most 2ℓ components in each of $F_i^{\mathcal{I}}$ and $\tilde{F}_i^{\mathcal{I}}$ so that outside of W the component structure of both graphs is the same (to see this note that the order is irrelevant for $\mathcal{I}$, so we may assume $i = j$; then defining W as the union of the components containing the vertices of $\underline{v}_j$ and $\tilde{\underline{v}}_j$ in $F^{\mathcal{I}}_{i-1} = \tilde{F}^{\mathcal{I}}_{i-1}$ suffices). The key point is now that for size rules the decisions of $\mathcal{R}$ in $F^{\mathcal{R}}_i$ and $\tilde{F}^{\mathcal{R}}_i$ are the same for all tuples which contain no vertices from W (indeed, if a decision of $\mathcal{R}$ is modified then any changes of the resulting component sizes can only 'propagate' inside the components of $F^{\mathcal{I}}_i$ and $\tilde{F}^{\mathcal{I}}_i$; so only tuples containing vertices from W can be affected). It follows that the component structure outside of W is also the same in $F^{\mathcal{R}}_i$ and $\tilde{F}^{\mathcal{R}}_i$. Recall that W contains at most 2ℓ components in each of $F^{\mathcal{I}}_i$ and $\tilde{F}^{\mathcal{I}}_i$. So, since $X_{k,i}$ only counts those vertices v with $|C_v(F^{\mathcal{I}}_i)| < U$, we see that a change of one tuple can alter f by at most $2 \cdot 2\ell \cdot U$, as claimed. So, recalling that $1 \leqslant i \leqslant \sigma n$, for $n \geqslant n_0(\ell, \sigma)$ McDiarmid's inequality [53] implies

$$\mathbb{P}(|X_{k,i} - \mathbb{E}X_{k,i}| \geqslant \Delta) \leqslant \exp\left(-\frac{2\Delta^2}{i(4\ell U)^2}\right) \leqslant n^{-300}. \tag{5.42}$$

Finally, after combining (5.40)–(5.42), taking a union bound to account for all choices of $1 \leqslant i \leqslant \sigma n$ and $1 \leqslant k \leqslant (\log n)^2$ completes the proof (noting that the claim is trivial for $i = 0$). □

Using the main idea of the above proof we can directly show that $\mathbb{E}N_k(F^{\mathcal{R}}_i)$ is essentially independent of the initial graph $F^{\mathcal{R}}_0 = F$ for $i \leqslant \sigma n$: for any two graphs $F, \tilde{F}$ satisfying the assumptions of Theorem 5.2.1 their expected values can differ by at most, say, $(\log n)^{C+3} n^{1/2}$. The key point is that for such graphs we can construct a bijection Ψ between their vertex sets which, up to an exceptional set W of at most, say, $4U(\log n)^C n^{1/2}$ vertices, preserves the component structure of F and $\tilde{F}$, respectively. Now, using Ψ we couple $F^{\mathcal{R}}_i, F^{\mathcal{I}}_i$ and $\tilde{F}^{\mathcal{R}}_i, \tilde{F}^{\mathcal{I}}_i$ in a measure preserving way. Since changes can only propagate inside the components of the 'influence' graphs, only those vertices whose components in $F^{\mathcal{I}}_i$ or $\tilde{F}^{\mathcal{I}}_i$ contain vertices of W or $\Psi(W)$ can be 'spoiled'. Intuitively, since the components usually have size at most U, under this coupling N_k thus typically differs by at most $2|W| \cdot U$ for both graphs. Taking the error probability of $\max\{L_1(F^{\mathcal{I}}_{\sigma n}), L_1(\tilde{F}^{\mathcal{I}}_{\sigma n})\} < U$ into account, the claim now follows without much work.

5.2.4.5 Putting things together

In this section we combine our findings to prove Theorem 5.2.1. Lemma 5.2.9 easily implies the first part, i.e., existence of $\rho : \mathbb{N} \times \mathbb{R}^+ \to [0,1]$ with the desired properties. Indeed, it ensures that for every $\sigma \geqslant 0$ satisfying (5.9) there exist $a, A > 0$ (depending only on $\ell, L, \sigma, \beta, B$) such that for every $t \in [0,\sigma]$ we have $\sum_{k\geqslant 1} \rho(k,t) = 1$ and

$$\rho(s,t) \leqslant Ae^{-as} \quad \text{for all } s \geqslant 0. \tag{5.43}$$

For later usage we remark that Lemma 5.2.5 holds for the same a, A. Let $D = 300/a > 0$ and $\tilde{\beta} = \min\{e^{a/2}, e^{1/(4D)}\} > 1$. Now, using (5.43) we see that

$$\sum_{k\geqslant 1} \rho(k,t)\tilde{\beta}^k \leqslant A \sum_{k\geqslant 1} e^{-ak/2} = \tilde{B} - 1, \tag{5.44}$$

with $1 < \tilde{B} < \infty$ depending only on a, A. Similarly, we obtain

$$\chi(\rho,t) = \sum_{k\geqslant 1} k\rho(k,t) \leqslant A \sum_{k\geqslant 1} ke^{-ak} = \tilde{L} - 1, \tag{5.45}$$

with $1 < \tilde{L} < \infty$ depending only on a, A. Summarizing, equations (5.6)–(5.8) hold when $\beta, B, L, \varphi(\cdot)$ are replaced by $\tilde{\beta}, \tilde{B}, \tilde{L}, \rho(\cdot, t)$, with room to spare.

Turning to properties of $F_i^{\mathcal{R}}$, from Lemmas 5.2.10–5.2.12 it follows that with probability at least $1 - n^{-250}$, for every $0 \leqslant i \leqslant \sigma n$ (by considering $t = i/n \in [0,\sigma]$) we have

$$N_k(F_i^{\mathcal{R}}) = \rho(k,i/n)n \pm 3(\log n)^{C+6} n^{1/2} \quad \text{for all } 1 \leqslant k \leqslant (\log n)^2 \tag{5.46}$$

for $n \geqslant n_0(\ell, L, \sigma, \beta, B)$. Recall that Lemma 5.2.5 holds with the a, A chosen above. By definition of D it follows that, with probability at least $1 - n^{-250}$, we have

$$L_1(F_{\sigma n}^{\mathcal{R}}) \leqslant D \log n \tag{5.47}$$

for $n \geqslant n_0(A)$. In the remainder we assume that (5.46)–(5.47) hold. Recalling (5.43) and the definition of D, note that for all $k \geqslant D \log n$ and $0 \leqslant i \leqslant \sigma n$ we have, say, $|\rho(k,i/n)| \leqslant n^{-9}$

for $n \geqslant n_0(A)$. Using (5.47) it follows that

$$N_k(F_i^{\mathcal{R}}) = \rho(k, i/n)n \pm (\log n)^C n^{1/2} \quad \text{for all } k \geqslant D\log n \text{ and } 0 \leqslant i \leqslant \sigma n.$$

Together with (5.46), for every $0 \leqslant i \leqslant \sigma n$ this establishes (5.10) with $C, F, \varphi(\cdot)$ replaced by $\tilde{C}, F_i^{\mathcal{R}}, \rho(\cdot, i/n)$ for $n \geqslant n_0(D)$, where $\tilde{C} = C + 9$. Now, using (5.44) and (5.46)–(5.47) we see that for every $0 \leqslant i \leqslant \sigma n$ we have (as $i/n \in [0, \sigma]$)

$$\begin{aligned}\sum_{k\in[n]} N_k(F_i^{\mathcal{R}})\tilde{\beta}^k &\leqslant n \sum_{1\leqslant k\leqslant D\log n} \rho(k, i/n)\tilde{\beta}^k + 3D(\log n)^{C+6} n^{1/2} \sum_{1\leqslant k\leqslant D\log n} \tilde{\beta}^k \\ &\leqslant n \sum_{k\geqslant 1} \rho(k, i/n)\tilde{\beta}^k + 3D^2(\log n)^{C+7} n^{3/4} \leqslant \tilde{B}n\end{aligned}$$

for $n \geqslant n_0(C, D)$, which establishes (5.11) with β, B, F replaced by $\tilde{\beta}, \tilde{B}, F_i^{\mathcal{R}}$. It remains to show that (5.12)–(5.13) hold. Recall that $S(G) = \sum_{k\in[n]} kN_k(G)/n$. Now, assuming $n \geqslant n_0(a, A, D)$ and using (5.45)–(5.47) similarly as above, for every $0 \leqslant i \leqslant \sigma n$ we have

$$\begin{aligned}S(F_i^{\mathcal{R}}) &= \sum_{k\geqslant 1} k\rho(k, i/n) \pm A \sum_{k\geqslant D\log n} ke^{-ak} \pm 3D^2(\log n)^{C+8} n^{-1/2} \\ &= \chi(\rho, i/n) \pm 4D^2(\log n)^{C+8} n^{-1/2},\end{aligned}$$

which establishes (5.13) for $n \geqslant n_0(D)$. Finally, recalling (5.45), it follows that (5.12) holds with L, F replaced by $\tilde{L}, F_i^{\mathcal{R}}$ for $n \geqslant n_0(C, D)$, which completes the proof of Theorem 5.2.1.

5.3 When does $t_{\mathrm{b}} = t_{\mathrm{c}}$?

In this section we discuss Conjecture 5.1.2, first showing that it does hold for many size rules and then, in Section 5.3.2, showing that it cannot be extended to general ℓ-vertex rules, i.e., that the critical point where the susceptibility blows up need not always coincide with the percolation threshold.

5.3.1 Rules with uniform random edges

It is well known, and not hard to check, that under suitable assumptions the graph $F_{\theta n}$ given by adding θn independent and uniformly random edges to a given n-vertex initial graph F can be viewed as an instance of the inhomogeneous random graph model of Bollobás, Janson and Riordan [24]. To make this precise, consider instead the (multi-)graph $\widetilde{F}_\theta$ obtained from F by adding a Poisson number $\mathrm{Po}(2\theta/n)$ of copies of each of the $\binom{n}{2}$ possible edges, with these numbers independent; we may then ignore multiple edges, as we are only interested in the component structure. Since $\mathrm{Po}(\theta(n-1))$ edges are added in total, and there will be few multiple edges, $\widetilde{F}_\theta$ and $F_{\theta n}$ are essentially interchangeable (one may use domination arguments comparing them for different θ to make this precise). Given two components C_1 and C_2 of F, the number of edges between them in $\widetilde{F}_\theta$ is Poisson with mean $|C_1||C_2|2\theta/n$. Making (for a change) the n dependence explicit, let H_n be the random graph whose vertices are the components of F, with an edge between two vertices if these components are joined by an edge of $\widetilde{F}_\theta$. We say that a vertex of H_n has *type* k if the corresponding component of F has k vertices. Then the probability of an edge between a given type-i vertex and a given type-j vertex of H_n is $1 - e^{-2\theta ij/n}$, which is around $2\theta ij/n$ if i and j are not too big, and the events that different edges are present are independent.

More precisely, let $\kappa(i,j) = 2\theta ij$ for all positive integers i and j. Suppose that μ is a finite measure on $\mathbb{Z}^+$, i.e., that $\mu_k = \mu(\{k\}) \geqslant 0$ for all k and $0 < \sum_{k\geqslant 1}\mu_k < \infty$. Let $F = F_n$ be a random n-vertex starting graph. Suppose that, for each fixed $k \geqslant 1$,

$$\frac{N_k(F_n)}{kn} \xrightarrow{\mathrm{p}} \mu_k \tag{5.48}$$

as $n \to \infty$, i.e., that H_n has asymptotically $\mu_k n$ vertices of type k, and that

$$\sum_{k\geqslant 1} k\mu_k = 1. \tag{5.49}$$

Then one can use (5.49), the fact that F_n has n vertices and (5.48) to show that whenever $K(n) \to \infty$ we have

$$N_{\geqslant K(n)}(F_n)/n \xrightarrow{\mathrm{p}} 0, \tag{5.50}$$

and it follows that for any $A \subset \mathbb{Z}^+$ we have

$$\sum_{k \in A} \frac{N_k(F_n)}{kn} \xrightarrow{\mathrm{p}} \sum_{k \in A} \mu_k. \tag{5.51}$$

In the terminology of [24], this means that the (random) sets of vertices of the graphs H_n, together with their types, form a *generalized vertex space* on the *generalized ground space* $(\mathbb{Z}^+, \mu)$. Taking $A = \mathbb{Z}^+$ in (5.51), we have in particular that $|H_n|/n \xrightarrow{\mathrm{p}} \mu(\mathbb{Z}^+) \in (0, \infty)$. By (5.49) the function κ forms an integrable kernel on the ground space $(\mathbb{Z}^+, \mu)$, with integral 2θ. Finally, the technical 'graphicality' condition of [24] is met since $\widetilde{F}_\theta$ has asymptotically θn edges. It follows that under these assumptions, the results of [24] apply to H_n (see Remark 2.4 there). The most important of these results is [24, Theorem 3.1], which tells us that H_n will whp contain a giant component (one with $\Theta(n)$ vertices) if and only if $||T_\kappa|| > 1$, where T_κ is a certain integral operator associated to κ. In particular, if $||T_\kappa|| > 1$ then there is some constant $\alpha = \alpha(\kappa, \mu) > 0$ (anything smaller than the quantity $\rho(\kappa)$ in [24]) such that whp H_n has a component with at least αn vertices. For the particular κ considered here, which is 'rank 1', we have $||T_\kappa|| = \sum_k 2\theta k^2 \mu_k$; see (16.8) in [24]. Note that if H_n contains a component with at least αn vertices, then so does $\widetilde{F}_\theta$ – the union of the components of F corresponding to these vertices of H_n. So, in short, if (5.48) and (5.49) hold, then $\widetilde{F}_\theta$ will have a giant component (whp) if (and, one can check, only if) $\sum_k 2\theta k^2 \mu_k > 1$. Moreover, it is not hard to check that these conclusions remain true if we delete some subset of the components of F_n, and adjust μ, as long as (5.48) holds for the new graph and $\Theta(n)$ components remain; this is because (5.51) still holds, and the kernel is still graphical.

We shall apply the observations above with initial graph $F = F_n = G^{\mathcal{R}}_{t_{\mathrm{b}}n}$, where $\mathcal{R}$ is some ℓ-vertex size rule. By (5.5), the condition (5.48) holds with $\mu_k = \rho_k(t_{\mathrm{b}})/k$. Furthermore, as noted after (5.5), we have $\sum_k \rho_k(t_{\mathrm{b}}) = 1$, which gives (5.49). Finally, note that

$$||T_\kappa|| = \sum_k 2\theta k \rho_k(t_{\mathrm{b}}) = 2\theta s(t_{\mathrm{b}}) = \infty, \tag{5.52}$$

since $s(t) = \sum_k \rho_k(t)$ diverges at $t = t_{\mathrm{b}}$. So far this tells us only that if we run any size rule up to time $t = t_{\mathrm{b}}^{\mathcal{R}}$ and then switch to adding uniformly random edges, after any constant times n further edges a giant component will emerge. The key point is that variants of this argument can be used to study the further evolution of $G_i^{\mathcal{R}}$ for suitable rules $\mathcal{R}$. A related approach was taken in [74] and [45].

Theorem 5.3.1. *Let $\mathcal{R}$ be a bounded-size ℓ-vertex rule. Then the conclusion of Conjecture 5.1.2 holds for $\mathcal{R}$; in particular, $t_{\mathrm{c}}^{\mathcal{R}} = t_{\mathrm{b}}^{\mathcal{R}}$, and moreover for any $\varepsilon > 0$ there is an $\alpha > 0$ such that whp $L_1(G_{(t_{\mathrm{b}}+\varepsilon)n}^{\mathcal{R}}) \geqslant \alpha n$.*

Note that this result was proved for some bounded-size 4-vertex rules (ones in which either v_1v_2 or v_3v_4 is added) already by Spencer and Wormald [74].

Proof. By definition of bounded-size rules, there is a constant B such that $\mathcal{R}$ treats all components of size greater than B in the same way. Consider the graph $G_{t_{\mathrm{b}}n}^{\mathcal{R}}$ generated by the rule after $t_{\mathrm{b}}n$ steps. Let W be the set of vertices of this graph in components of size greater than B, and let $F = F_n$ be the subgraph of $G_{t_{\mathrm{b}}n}^{\mathcal{R}}$ induced by W. Noting that $s(t_{\mathrm{b}}) = \sum_k k\rho_k(t_{\mathrm{b}}) = \infty > B$, we have $\rho_k(t_{\mathrm{b}}) > 0$ for some $k > B$, and it follows that for some constant $\beta > 0$, we have $|W| \geqslant \beta n$ whp. From now on we assume that this is the case. In all subsequent steps of our original process $G_i^{\mathcal{R}}$, every vertex of W is in a component of size greater than B. Fix $\varepsilon > 0$. Let us call a step *good* if in this step all ℓ selected vertices are in W. Then each step is good with probability at least β^ℓ, and it follows that whp at least θn of the next εn steps are good, where $\theta = \varepsilon\beta^\ell/2$ is a positive constant. Again using the definition of a bounded-size rule, in each good step at least one edge is added and by symmetry it is chosen uniformly at random from all possible edges with ends in W. It follows that we may couple $G_{(t_{\mathrm{b}}+\varepsilon)n}^{\mathcal{R}}$ and $\widetilde{F}_\theta$ so that whp the former contains the latter as a subgraph. But F satisfies the assumptions above with $\mu_k = \rho_k(t_{\mathrm{b}})/k$ for $k > B$ and $\mu_k = 0$ for $k \leqslant B$. Since the sum in (5.52) remains infinite after removing the first B terms, Theorem 3.1 of [24] and the discussion above imply that for some positive α, whp $\widetilde{F}_\theta$ contains a component with at least αn vertices. $\square$

Our next result concerns a different generalization of the Bohman–Frieze process [14]. Let us call an Achlioptas rule $\mathcal{R}$ *take-it-or-leave-it* if, when presented with a choice of two edges e_1 and e_2, the rule decides which to select depending only on the current graph and on e_1. In other words, the rule first sees e_1 and must decide whether to take this edge of not; if not, it selects the uniformly random edge e_2. Bounded-size rules of this type were studied, for example, by Bohman and Kravitz [19]; here we do not assume that the rule is bounded-size.

Theorem 5.3.2. *Let $\mathcal{R}$ be a take-it-or-leave-it size rule. Then the conclusions of Conjecture 5.1.2 and Theorem 5.3.1 hold for $\mathcal{R}$.*

Proof. Consider the process $(G^{\mathcal{R}}_{t_{\mathrm{b}}n+i})_{i\geqslant 0}$, i.e., our Achlioptas process started at step $t_{\mathrm{b}}n$. As above, set $F = F_n = G^{\mathcal{R}}_{t_{\mathrm{b}}n}$. Since $\mathcal{R}$ is a take-it-or-leave-it rule, the further evolution may be described as follows. Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be independent lists of independent (potential) edges each chosen uniformly at random from all $\binom{n}{2}$ possibilities. In step i of our process (step $t_{\mathrm{b}}n + i$ of the original), take for e_1 the ith element of $\mathcal{L}_1$. The rule now decides whether to add this edge to the current graph. If not, take for e_2 the *next* edge from $\mathcal{L}_2$, and add that. Thus, the jth time that the rule declines the first edge, we take the jth edge from $\mathcal{L}_2$.

Since the edges in $\mathcal{L}_2$ are uniformly random, the discussion before Theorem 5.3.1 shows that for any constant $\delta > 0$, whp the first δn edges from $\mathcal{L}_2$ will, when added to $F = G^{\mathcal{R}}_{t_{\mathrm{b}}n}$, be enough to form a giant component. Fix $\varepsilon > 0$, and define as above a graph H_n whose vertices are the components of F, with edges corresponding to the first εn edges from $\mathcal{L}_1$. As noted before, this graph H_n may be viewed as an instance of the model studied in [24], and there is some $\alpha > 0$ such that whp H_n has a component with at least αn vertices. Furthermore, by the stability result [24, Theorem 3.9], there is some $\delta > 0$ such that whp H_n has the property that deleting any δn edges still leaves a component with at least $\alpha n/2$ vertices of H_n. Hence, whp $\mathcal{L}_1$ has the property that if we add any subset of at least $(\varepsilon - \delta)n$ of the first εn edges to F, we will create a component of size at least $\alpha n/2$, and whp $\mathcal{L}_2$ has the property that adding its first δn edges to F creates a component of size at least some constant times n. But when both properties hold, then whatever the rule does, $G^{\mathcal{R}}_{(t_{\mathrm{b}}+\varepsilon)n}$ will have a giant component. □

For our final result, let us call an Achlioptas rule *large-biased* if there exists some constant B such that if both endvertices of e_1 are isolated vertices (components of size one) and both endvertices of e_2 are in components of size greater than B, then the rule will select e_2. Perhaps the most interesting examples of such rules are the *reverse product rule*, where we select the (a if there is a tie) edge maximizing the product of the sizes of the components containing its endvertices, or the *reverse sum rule*, defined similarly but with product replaced by sum. Perhaps surprisingly (given the difficulty of analyzing the usual product rule), we can prove Conjecture 5.1.2 for such rules.

Theorem 5.3.3. *Let $\mathcal{R}$ be a large-biased size rule. Then the conclusions of Conjecture 5.1.2 and Theorem 5.3.1 hold for $\mathcal{R}$.*

Proof. The proof is very similar to that of Theorem 5.3.1. Indeed, as usual we start from $F = G^{\mathcal{R}}_{t_{\mathrm{b}}n}$. As before, let W be the set of vertices of F in components of size greater than B. Call a subsequent step *good* if e_1 joins two vertices in components of size one and e_2 joins two vertices in W. Since there are whp at least some constant times n isolated vertices in $G^{\mathcal{R}}_{(t_{\mathrm{b}}+1)n}$, and (as before), W whp has size at least a constant times n, off an event of small probability the conditional probability (given the history) that the next step is good is always at least some positive constant. Furthermore, when a step is good, the added edge is uniformly random among all possible edges inside W. The remainder of the argument is as for Theorem 5.3.1; we omit the details. $\square$

The results above all illustrate the idea that if we can find a reasonable number of uniformly random edges among the edges selected by our process, then the process will be 'well behaved' (will have $t_{\mathrm{c}} = t_{\mathrm{b}}$). This approach can be used to prove Conjecture 5.1.2 for other special classes of size rules, but it seems that additional ideas are needed for the general case.

5.3.2 Examples of delayed percolation

Having given several partial results supporting our belief in Conjecture 5.1.2, in this section we show that the conjecture cannot be extended to arbitrary ℓ-vertex rules. More concretely, we give examples of simple rules that can delay the appearance of linear size components for $\Omega(n)$ steps beyond the point where the susceptibility diverges. The rules we use behave like size rules almost all the time.

We start by introducing the *r-sum rule* $\mathcal{S}_r$, which is a $2r$-vertex size rule. Given vertices $(v_1, \ldots, v_\ell)$ and the corresponding list of component sizes $(c_1, \ldots, c_\ell)$, the r-sum rule adds the pair $v_{2j-1}v_{2j}$ with the (smallest, if there are ties) $j \in [r]$ that minimizes the sum $c_{2j-1} + c_{2j}$ of the component sizes. Recall the definition of $F_i^{\mathcal{R}}$ given in Section 5.2: informally it denotes the graph that we obtain by starting with the initial graph $F_0^{\mathcal{R}} = F$ and then following i steps of an Achlioptas process using the rule $\mathcal{R}$ (to decide which edges to add in each step). Intuitively, the next lemma states that the r-sum rule does not substantially change (uniform) polynomial tails for $N_{\geqslant k}$ during some δn steps (here we use $\mathcal{S}_r$ for concreteness; other size rules exhibit similar behaviour).

Lemma 5.3.4. *Let F be a graph on n vertices. Suppose there are $x, C > 0$ and $K = K(n) \geqslant 1$ such that for all $1 \leqslant k \leqslant K$ we have*

$$N_{\geqslant k}(F) \leqslant Ck^{-x}n. \tag{5.53}$$

Given $r \geqslant 1 + 1/x$ there exists $\delta = \delta(x, C, r) > 0$ such if n is large enough then, with probability at least $1 - n^{-99}$, for all $1 \leqslant k \leqslant K' = \min\{K, n^{1/[2(1+x)]}\}$ we have

$$N_{\geqslant k}(F_{\delta n}^{\mathcal{S}_r}) \leqslant 2Ck^{-x}n. \tag{5.54}$$

Proof. Set $\delta = 2^{-[(2+x)r+3]}C^{-(r-1)}$. Let $\mathcal{E}_{i',k'}$ denote the event that for all $0 \leqslant i \leqslant i'$ and $1 \leqslant k \leqslant k'$ we have

$$N_{\geqslant k}(F_i^{\mathcal{S}_r}) \leqslant 2Ck^{-x}n. \tag{5.55}$$

Observe that it suffices to show that $\mathcal{E}_{\delta n, K'}$ fails with probability at most n^{-99}. For $k \leqslant K'$ let $X_{k,i}$ denote the indicator function of the event $N_{\geqslant k}(F_i^{\mathcal{S}_r}) \neq N_{\geqslant k}(F_{i-1}^{\mathcal{S}_r})$. Set $X_k = \sum_{1 \leqslant i \leqslant \delta n} X_{k,i}$ and $Y_k = \sum_{1 \leqslant i \leqslant \delta n} Y_{k,i}$, where

$$Y_{k,i} = \begin{cases} X_{k,i}, & \text{if } \mathcal{E}_{i-1,k-1} \text{ holds,} \\ 0, & \text{otherwise.} \end{cases}$$

Note that in each step a new component of size at least k is only created by $\mathcal{S}_r$ if in each pair $v_{2j-1}v_{2j}$ at least one vertex is in a component of size at least $\lceil k/2 \rceil$. So, whenever $\mathcal{E}_{i-1,k-1}$ holds, using (5.55) and $r \geqslant 1 + 1/x$, we see that the probability that $X_{k,i} = 1$ is at most

$$\left(\frac{2N_{\geqslant \lceil k/2 \rceil}(F^{\mathcal{S}_r}_{i-1})}{n}\right)^r \leqslant \left(\frac{4C}{(k/2)^x}\right)^r = \frac{\left(2^{2+x}C\right)^r}{k^{rx}} \leqslant \frac{\left(2^{2+x}C\right)^r}{k^{1+x}} = \xi_k.$$

Since $Y_{k,i} = 0$ whenever $\mathcal{E}_{i-1,k-1}$ fails, it follows that Y_k is stochastically dominated by a binomial random variable with δn trials and success probability ξ_k. Note that (5.53) implies $C \geqslant 1$. Now, using $k \leqslant K'$ we have $\delta n \xi_k \geqslant C/8 \cdot n^{1/2} \geqslant 600 \log n$ for $n \geqslant n_0$, so standard Chernoff bounds yield

$$\mathbb{P}(Y_k \geqslant 2\delta n \xi_k) \leqslant e^{-\delta n \xi_k/3} \leqslant n^{-200}. \tag{5.56}$$

Next we claim that $\mathcal{E}_{\delta n,k-1}$ and $Y_k < 2\delta n \xi_k$ together imply $\mathcal{E}_{\delta n,k}$, so that $\mathbb{P}(\neg\mathcal{E}_{\delta n,k}) \leqslant \mathbb{P}(\neg\mathcal{E}_{\delta n,k-1}) + \mathbb{P}(Y_k \geqslant 2\delta n \xi_k)$. Indeed, by monotonicity $\mathcal{E}_{\delta n,k-1}$ implies $\mathcal{E}_{i,k-1}$ for every $0 \leqslant i \leqslant \delta n$, so $Y_k = X_k$. Now, since $N_{\geqslant k}$ increases by at most $2k$ per step, by choice of δ it follows that

$$N_{\geqslant k}(F^{\mathcal{S}_r}_{\delta n}) - N_{\geqslant k}(F^{\mathcal{S}_r}_0) \leqslant 2kY_k \leqslant 4\delta \left(2^{2+x}C\right)^r k^{-x} n \leqslant Ck^{-x} n,$$

which together with $N_{\geqslant k}(F^{\mathcal{S}_r}_0) = N_{\geqslant k}(F) \leqslant Ck^{-x}n$ implies $N_{\geqslant k}(F^{\mathcal{S}_r}_{\delta n}) \leqslant 2Ck^{-x}n$, as claimed. Iterating the above argument for $k \leqslant K'$ and noting that $\mathcal{E}_{\delta n,1}$ always holds due to $C \geqslant 1$, using (5.56) we obtain

$$\mathbb{P}(\neg\mathcal{E}_{\delta n,K'}) \leqslant \sum_{2 \leqslant k \leqslant K'} \mathbb{P}(Y_k \geqslant 2\delta n \xi_k) \leqslant n^{-99},$$

and the proof is complete. □

Let $\mathcal{D}_r$ denote the rule which always adds the pair v_1v_2 during the first $n/2$ steps (corresponding to an Erdős–Rényi evolution with $\ell = 2$); afterwards it 'switches' and uses the r-sum rule $\mathcal{S}_r$. The point is that many properties of the 'critical' Erdős–Rényi random graph $G_{n,n/2}$ are well known: there exist constants $C, \alpha > 0$ and a function $K = K(n)$ with $K \to \infty$ as $n \to \infty$ such that whp $S(G_{n,n/2}) \geqslant n^\alpha$ and $N_{\geqslant k}(G_{n,n/2}) \leqslant Ck^{-1/2}n$ for all $1 \leqslant k \leqslant K$. So, by conditioning on these properties and then using Lemma 5.3.4, we immediately deduce the main result of this section. Indeed, using the rule $\mathcal{D}_r$ for $r \geqslant 3$ we whp have diverging susceptibility after $n/2$ steps, but in δn subsequent steps whp no linear size components appear (in fact, in this case $t_{\mathrm{b}} = 1/2 < t_{\mathrm{c}}$ holds).

Corollary 5.3.5. *For every $r \geqslant 3$ there exists $\delta = \delta(r) > 0$ such that we have whp $S(G_{n/2}^{\mathcal{D}_r}) = \omega(1)$ and $L_1(G_{n/2+\delta n}^{\mathcal{D}_r}) = o(n)$.*

Alternatively, using essentially the same line of reasoning, we obtain a similar result by switching after the first step where the susceptibility is at least $L = L(n) = \omega(1)$, for L not too large. Furthermore, we can replace $\mathcal{S}_r$ by other suitable size rules. For example, the rule $\mathcal{M}_\ell$, which always connects two vertices with the two smallest component sizes c_j, satisfies an analogue of Lemma 5.3.4 for $\ell \geqslant 2 + 1/x$. So the rule $\mathcal{C}_\ell$, which switches from an Erdős–Rényi evolution (always adding v_1v_2) to $\mathcal{M}_\ell$ after $n/2$ steps, yields another example with $t_{\mathrm{b}} < t_{\mathrm{c}}$.

Corollary 5.3.6. *For every $\ell \geqslant 4$ there exists $\delta = \delta(\ell) > 0$ such that we have whp $S(G_{n/2}^{\mathcal{C}_\ell}) = \omega(1)$ and $L_1(G_{n/2+\delta n}^{\mathcal{C}_\ell}) = o(n)$.*

Note that the examples given in Corollary 5.3.5 and 5.3.6 always behave like size rules except that once between two steps they change the rule used (by only querying natural parameters such as the number of vertices and steps, or the susceptibility). So, one can argue that Conjecture 5.1.2 already fails for a rather restricted superset of size rules.

5.4 Cycle structure

In this final section we show that, as long as an Achlioptas process contains only small components, it will have a very simple cycle structure: most components will be trees, some will be unicyclic, and there will (whp) be no 'complex' components, i.e., ones containing more than one cycle. In fact, we prove the result for the more general class of ℓ-vertex rules. However, here we need an additional assumption: in each round the set of added edges is a forest. Allowing slightly greater generality, we call a rule *acyclic* if the edges between $v_1, \ldots, v_\ell$ added in a single step correspond to a forest on $1, \ldots, \ell$. This in particular holds if in each step at most two edges are added. Note that such an assumption is in fact necessary for $\ell \geqslant 3$, since always connecting all vertices in each step quickly creates many cycles and complex components.

Lemma 5.4.1. *Let $\ell \geqslant 2$ and let $\mathcal{R}$ be an acyclic ℓ-vertex rule. Given $0 < \delta < 1/4$ and $U = U(n)$, suppose that for $n \geqslant n_0(\ell, \delta)$ we have $1 \leqslant U \leqslant n^{1/4-\delta}$. For $n \geqslant n_0(\ell, \delta)$, with probability at least $1 - n^{-\delta/2}$ the following holds for every $0 \leqslant i \leqslant n^{1+\delta}$: in $G_i^{\mathcal{R}}$ there are no components of size at most U which contain at least two cycles, and the number of vertices in components of size at most U with exactly one cycle is at most $U^2 n^{2\delta}$.*

Proof. Set $m = n^{1+\delta}$. Let $\mathcal{B}_{1,i}$ denote the event that in step i one of the following happens: (a) at least three of the ℓ randomly chosen vertices are contained in the same component of size at most U, or (b) there are randomly chosen vertices w_1, w_2, w_3, w_4 and two components C_1, C_2 of size at most U such that $w_1, w_2 \in C_1$ and $w_3, w_4 \in C_2$. It is easy to see that $\mathcal{B}_{1,i}$ holds with probability at most $\ell^3 U^2/n^2 + \ell^4 U^2/n^2$. So, denoting by $\mathcal{B}_1$ the event that $\mathcal{B}_{1,i}$ holds for some $i \leqslant m$, we see that

$$\mathbb{P}(\mathcal{B}_1) \leqslant m \cdot 2\ell^4 U^2/n^2 \leqslant 2\ell^4 n^{-1/2-\delta}.$$

Let $S_{i,U}$ denote the set of vertices of $G_i^{\mathcal{R}}$ which are in components of size at most U containing exactly one cycle, and let $\mathcal{B}_2$ be the event that $S_{i,U}$ contains at least $2\ell^3 Um/n$ components for some $i \leqslant m$. Then $\neg\mathcal{B}_2$ implies $|S_{i,U}| \leqslant 2\ell^3 U^2 m/n = R$ for every $i \leqslant m$, where $R \leqslant U^2 n^{2\delta}$ for $n \geqslant n_0(\ell, \delta)$. Since in each step the number of components in $S_{i,U}$ changes by at most ℓ, when $\mathcal{B}_2$ holds there are at least $2\ell^2 Um/n$ steps in which the number of components in $S_{i,U}$ increases. This can only happen if at least two randomly chosen vertices are in the same component of size at most U. Since in each step this has probability at most $\ell^2 U/n$, the expected number of such steps is bounded by $\lambda = \ell^2 Um/n$. Using standard Chernoff bounds (and stochastic domination) it follows that

$$\mathbb{P}(\mathcal{B}_2) \leqslant e^{-\lambda/3} \leqslant e^{-n^{\delta}}.$$

Let $\mathcal{B}_3$ denote the event that in some $G_i^{\mathcal{R}}$ with $i \leqslant m$ there exists a component of size at most U which contains at least two cycles. In each step where $\mathcal{B}_{1,i}$ fails, note that a new complex component of size at most U can only be created if (a) at least two randomly chosen vertices are in $S_{i-1,U}$ or (b) one randomly chosen vertex lies in $S_{i-1,U}$, and two randomly chosen vertices are in the same tree component of size at most U. So, by considering the probability that this happens for some $i \leqslant m$ (assuming $|S_{i-1,U}| \leqslant R$), we see that

$$\mathbb{P}(\neg\mathcal{B}_1 \cap \neg\mathcal{B}_2 \cap \mathcal{B}_3) \leqslant m \cdot (\ell^2 R^2/n^2 + \ell^3 RU/n^2) \leqslant 5\ell^8 U^4 m^3/n^4 \leqslant 5\ell^8 n^{-\delta},$$

completing the proof for $n \geqslant n_0(\delta, \ell)$. □

Theorem 5.1.1 tells us that for size rules, for any fixed $t < t_{\mathrm{b}}$, the largest component of $G_{tn}^{\mathcal{R}}$ whp has size at most $O(\log n)$. Taking $U = (\log n)^2$, say, we see that if $\mathcal{R}$ is acyclic, then whp there are no complex components, and at most $n^{o(1)}$ vertices on cyclic components – in other words, almost all components are trees. (We have not tried to optimize the bound here – the method actually gives some power of $\log n$.)

Part II

H-free processes

Chapter 6

Overview

6.1 Background and motivation

We consider a 'constrained' variant of the Erdős–Rényi random graph process that was suggested in 1990 by Bollobás and Erdős [22]. Given some fixed graph H, the *H-free process* starts with an empty graph on n vertices and then adds new edges, one at a time, where each edge is chosen uniformly at random subject to the condition that no copy of H is completed. It was first described in print in 1995 by Erdős, Suen and Winkler [34], who asked how many edges the final graph typically has (this also appears as a problem in [30]). The main difficulty when analysing this process is that there is a complicated dependence among the edges; the order in which they are inserted is also relevant.

Two main motivations for studying the evolution of H-free processes are (i) its potential applications to Ramsey and Turán type problems and (ii) the development of tools and techniques for analyzing random processes with significant dependencies between different rounds. Related to both the analysis of the H-free process (and closely related processes) has significantly advanced the state of knowledge in recent years. Indeed, with respect to problems in extremal combinatorics, (i) improved lower bounds on the Turán numbers of certain bipartite graphs and Ramsey numbers $R(s,t)$ with $s \geqslant 4$ have been established in [13, 17, 78]; Bohman [13] also reproved the famous lower bound for $R(3,t)$ first obtained (using a 'semi-random' variant of the C_3-free process) by Kim [47] in 1995. Turning to the technical advances, (ii) a major breakthrough was made in 2009 by Bohman [13] while analyzing the C_3-free process: he significantly improved Wormald's differential equation method [80, 81] from 1995 (which is a widely used tool for dealing with dependencies in random processes).

Historically, the first results for the H-free process determined the typical final number of edges up to logarithmic factors. Here the special cases $H = C_3$ and $H \in \{K_4, C_4\}$ were considered by Erdős, Suen and Winkler [34] and Bollobás and Riordan [25] in 1995 and 2000, respectively. To get the complex history right, the earliest result was obtained (in disguise) by Ruciński and Wormald [68] for the (much simpler) case of a forbidden star $H = K_{1,d+1}$ (corresponding to the maximum degree d-process): in 1992 they showed that it ends whp with $\lfloor nd/2 \rfloor$ edges. The study of the general H-free process was initiated independently by Bollobás and Riordan [25] and Osthus and Taraz [57] around 2000: they considered the class of strictly 2-balanced graphs (see e.g. [43, 57] for the precise definition), which includes many interesting graphs such as cycles, complete graphs, complete r-partite graphs and the d-dimensional cube. For these graphs Osthus and Taraz determined the likely final number of edges up to logarithmic factors.

The first non-trivial matching bound was obtained almost a decade later: in a breakthrough in 2009, Bohman [13] proved that the C_3-free process ends whp with $\Theta(n^{3/2}\sqrt{\log n})$ edges, confirming a conjecture of Spencer [71]. Shortly afterwards, Wolfovitz [78] slightly improved the lower bound on the expected final number of edges for a range of graphs H. Bohman and Keevash [17] subsequently obtained new lower bounds that hold whp for the class of strictly 2-balanced graphs H, which they conjectured to be tight up to constant factors. One intriguing consequence of their analysis is that (at least during some substantial part of its initial evolution) the H-free process seems to be similar to the (uniform) Erdős–Rényi random graph process with the same number of edges: for example, it has comparable subgraph statistics.

Summarizing, the 15-year old question of Erdős, Suen and Winkler [34] concerning the typical final number of edges in the H-free process has attracted much attention, and for a large class of graphs H interesting bounds have been established. However, not much progress has been made in improving the known upper bounds. One can argue (since the upper bound of $\lfloor nd/2 \rfloor$ for the maximum degree d process [68] is immediate) that non-trivial matching upper bounds have not been determined for any infinite *class* of graphs. Closing this gap is one of the main open questions regarding H-free processes.

6.2 Heuristic for the final number of edges

In this section we give some intuition for what the typical final number of edges of the H-free process will roughly be. To this end we compare (couple) the Erdős–Rényi process with the H-free process, and we do so by considering a random permutation $e_1, \ldots, e_{\binom{n}{2}}$ of the edges. On the one hand, the union of the first m edges gives the graph $G_{n,m}$ produced by the Erdős–Rényi process. On the other hand, a moment's thought reveals that we recover the H-free process by sequentially traversing $e_1, e_2, \ldots$, each time adding only those edges which do not complete a copy of H (together with the edges added so far). Now, the main point is that the H-free process accepts most edges as long as the number of H-copies in $G_{n,m}$ is small compared to the number of edges m; this holds up to $m \approx n^{2-(v_H-2)/(e_H-1)}$ in expectation, where v_H and e_H denote the number of vertices and edges of H. Furthermore, it seems plausible that the rate of acceptance quickly drops as soon as in $G_{n,m}$ every pair of vertices uv can be 'extended' to a copy of H, i.e., adding the edge uv creates a (new) copy of H. By a result of Spencer [70] this happens after $m = \Theta(n^{2-(v_H-2)/(e_H-1)}(\log n)^{1/(e_H-1)})$ steps if H is strictly 2-balanced, which therefore (for this class of graphs) may seem like a natural first guess for the final number of edges.

Although the above heuristic might seem overly optimistic at first sight, it actually coincides (up to constants) with the rigorous lower bound obtained recently by Bohman and Keevash [17], which they conjectured to be tight (up to constant factors). Their proof is of course much more involved, analyzing the accepted and rejected edges rather precisely. In contrast, the heuristic leading to the lower bound of order $n^{2-(v_H-2)/(e_H-1)}$ can easily be converted into a rigorous proof for a large family of graphs H, as noted in [25, 57]. The main observation here is that, writing $X_{m,H}$ for the number of H-copies in $G_{n,m}$, at least $m - e_H X_{m,H}$ edges are accepted by the H-free process (since every rejected edge in $e_1, \ldots, e_m$ must lie in an H-copy in $G_{n,m}$). Now, since $X_{m,H}$ is concentrated around its expected value $\Theta(n^{v_H}(m/n^2)^{e_H})$ under very mild conditions on H (see e.g. [21, 43]), a suitable choice of $\delta = \delta(H) > 0$ ensures that for $m = \delta n^{2-(v_H-2)/(e_H-1)}$ whp at least $m - e_H X_{m,H} \geqslant m/2$ edges are accepted, establishing the claimed lower bound.

6.3 Guide to main results

In this section we give an informal overview of our main results for the H-free process.

6.3.1 Final number of edges

The main open question for H-free processes is due to Erdős, Suen and Winkler [34, 30], asking what the final number of edges typically is. While for many graphs we know the answer up to logarithmic factors due to Osthus and Taraz [57], more precise results (up to constants) were known only for the special case $H = C_3$ due to a recent breakthrough of Bohman [13]. In fact, for the C_ℓ-free process, where we forbid a cycle of fixed length $\ell \geqslant 3$, Osthus and Taraz [57] conjectured that the final graph has average degree $\Theta((n \log n)^{1/(\ell-1)})$, which, prior to our work, was still open for $\ell \geqslant 4$. Indeed, the best known lower and upper bounds on the final number of edges were $\Omega(n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)})$ and $O(n^{\ell/(\ell-1)} \log n)$, respectively. Here the lower bound is due to Bohman and Keevash [17], who also conjectured that it is tight up to constants (in fact, their conjecture is for the maximum degree).

In Chapter 7 we prove the conjectures of Osthus and Taraz [57] and Bohman and Keevash [17] for the C_ℓ-free process, thus closing the gap between the known lower and upper bounds. In particular, we establish the following result:

(A) **Matching bound for the C_ℓ-free process** (Corollary 7.1.2).
For every $\ell \geqslant 4$ the C_ℓ-free process ends whp with $\Theta(n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)})$ edges.

This is a natural extension of the main result of Bohman [13] for the C_3-free process, and answers the more than 15-year old question of Erdős, Suen and Winkler [34] for the C_ℓ-free process. Furthermore, it is the first result that determines (up to constants) the final number of edges for a *class* of forbidden graphs H, rather than an isolated single case.

The proof of (A) is based on an extension of some ideas that we have developed earlier [76] in order to establish matching bounds for the final number of edges in the K_4-free process. These will be crucial for overcoming certain technical difficulties that do not arise in the C_3-free process. In fact, our proof yields additional structural information concerning the final graph of the C_ℓ-free process: in particular, it whp has independence number $\Theta((n \log n)^{(\ell-2)/(\ell-1)})$ and is 'nearly regular', i.e., every vertex has degree $\Theta((n \log n)^{1/(\ell-1)})$. From the results we established in [75] it furthermore follows that every small set of vertices is 'sparse', so that the final graph also contains no sufficiently dense small subgraphs.

6.3.2 Two technical contributions

From the beginning, the complicated dependencies among the edges of the H-free process has been one of the main technical difficulties. In this context the major breakthrough is due to Bohman [13]: he managed to overcome many of these by developing Wormald's differential equation method [80, 81] significantly. A general formulation of this improved method can be found in his joint work with Keevash [17] for the H-free process. It requires, as usual for applications of the differential equation method, that we have good 'control' over the step-by-step changes of the random variables we are interested in. When applying it to the H-free process one key difficulty is to show that certain 'bad' events do not happen (such as very large one-step changes); in [13, 17] these are overcome using somewhat ad-hoc density arguments.

Here we briefly discuss two of our technical contributions, which include an improved version of the differential equations method and a general tool for overcoming certain technical difficulties in the H-free process. Informally, these can be described as follows:

(A) **Improved differential equation method** (Lemmas 7.3.1 and 7.3.4).
Strengthened variant of the differential equation method of Bohman and Keevash [17].

(B) **Transfer theorem** (Theorem 7.6.2).
A rigorous way to prove certain results for the H-free process using the *much* simpler Erdős–Rényi process.

The variant of the differential equation method which we establish in (A) develops the one proposed by Bohman and Keevash [17], which in turn improves upon Wormald's original variant [80, 81]. It has several advantages that make it easier to apply in many different situations. This includes a simpler formulation of the 'approximation error' and certain technical features that increase the range of potential applications (also introducing new parameters). In particular, it allows us to weaken certain technical assumptions. For these reasons we believe that it will be useful when analyzing other stochastic processes.

The 'transfer theorem' of (B) establishes a new connection between the H-free process and the Erdős–Rényi random graph process. It allows us to prove that certain (decreasing) events do not happen in the H-free process by showing that these do not happen in the (much easier to analyze) Erdős–Rényi random graph process (with a slightly larger number of edges). This is a key ingredient in our analysis of the C_ℓ-free process in Chapter 7 (allowing us establish certain properties which otherwise seem difficult to derive), and we believe that this connection will also be helpful in the analysis of other H-free processes.

Chapter 7

The C_ℓ-free process

7.1 Main result

In this chapter we consider the C_ℓ-free process, and we prove a new upper bound on its final number of edges. In fact, we give a new upper bound for the maximum degree, which confirms a conjecture of Bohman and Keevash [17] and improves previous upper bounds by Osthus and Taraz [57].

Theorem 7.1.1. *For every $\ell \geqslant 4$ there exists $D > 0$ such that whp the maximum degree in the final graph of the C_ℓ-free process is at most $D(n \log n)^{1/(\ell-1)}$.*

Up to the constant our upper bound is best possible, since the results of Bohman and Keevash [17] imply that for some $c > 0$, whp the minimum degree is at least $c(n \log n)^{1/(\ell-1)}$. The special case $\ell = 4$ was proved independently by Picollelli [58]; since a manuscript of this chapter was submitted Picollelli [59] has independently also proved the case $\ell \geqslant 4$. So, combining our findings with [17], we not only verify the mentioned conjecture of Osthus and Taraz [57], but establish the following stronger result.

Corollary 7.1.2. *For every $\ell \geqslant 4$ there exist $c, D > 0$ such that in the final graph of the C_ℓ-free process whp the number of edges is between $cn^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)}$ and $Dn^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)}$, and whp the degree of every vertex is between $c(n \log n)^{1/(\ell-1)}$ and $D(n \log n)^{1/(\ell-1)}$.* □

This is a natural extension of the main result of Bohman [13] for the C_3-free process, and answers a question of Erdős, Suen and Winkler for the C_ℓ-free process (see [34, 30]):

whp the final graph has $\Theta(n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)})$ edges. Since this question was asked for the H-free process in 1995, this is the first result that determines (up to constants) the final number of edges for a *class* of graphs.

We also obtain a new lower bound on the independence number of the C_ℓ-free process. Indeed, as pointed out to us by Picollelli, using Corollary 2.4 of Alon, Krivelevich and Sudakov [4], Corollary 7.1.2 implies the following bound conjectured in an earlier manuscript (together with a proof of a weaker bound).

Corollary 7.1.3. *For every $\ell \geqslant 4$ there exists $c > 0$ such that whp the independence number in the final graph of the C_ℓ-free process is at least $c(n \log n)^{(\ell-2)/(\ell-1)}$.* □

Up to the constant this matches the upper bound established by Bohman and Keevash [17]. We infer that whp the independence number in the final graph of the C_ℓ-free process is $\Theta((n \log n)^{(\ell-2)/(\ell-1)})$.

7.1.1 Comparison with previous work

The results of Bohman and Keevash [17] only apply to some initial $m \approx \delta n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)}$ steps of the C_ℓ-free process; the later behaviour is so far not well understood. To overcome this obstacle we use a similar idea as in [34, 25, 57] to bound the maximum degree: we show that already after these m steps, i.e., in the 'early' evolution, every pair $(\tilde{v}, U)$ with $\tilde{v} \notin U$ and $|U| = D(n \log n)^{1/(\ell-1)}$ has some property that prevents $U \subseteq \Gamma(\tilde{v})$ in the final graph of the C_ℓ-free process. Osthus and Taraz [57] establish their $O(n^{1/(\ell-1)} \log n)$ bound for the maximum degree using a 'static' point of view: they couple the C_ℓ-free process (or more generally the H-free process) with the classical random graph process and then show that even after deleting all edges contained in a copy of C_ℓ, every $(\tilde{v}, U)$ has the desired property. By contrast, we obtain the better $O((n \log n)^{1/(\ell-1)})$ bound by tracking the step-by-step effects of each edge added in the C_ℓ-free process, and our main tool is the variant of the differential equation method we introduced in [76].

Our argument relates to the proof of Bohman for the C_3-free process as follows. In [13] it is shown that every large set of vertices contains at least one edge, which implies a bound on the maximum degree, since the neighbourhood of each vertex is an independent set. In other words, the upper bound follows from a bound on the independence number. For the C_ℓ-free process, $\ell \geqslant 4$, the maximum degree is a separate question. In particular, we need to consider a more involved event, and thus must study the combinatorial structure of large sets more precisely.

To this end we track several random variables for every $(\tilde{v}, U)$. But, when applying the differential equation method, there are significant technical difficulties, and a simple refinement of the approach that we used in [76] for the K_4-free process does not suffice to overcome them. Here one crucial ingredient is a new connection between the H-free process and the Erdős–Rényi random graph, which might be of independent interest. More precisely, we develop a 'transfer theorem', which enables us to prove certain results for the H-free process using the *much* simpler binomial random graph model. This is a key tool for establishing properties of the C_ℓ-free process which otherwise seem difficult to derive. We believe that it will also aid in proving new upper bounds for the H-free process.

7.1.2 Organization of the chapter

We start by collecting the relevant properties of the C_ℓ-free process in Section 7.2, where we also state several probabilistic tools. In Section 7.3 we then introduce our improved variant of the differential equation method. Section 7.4 is devoted to the proof of Theorem 7.1.1. Our argument relies on two key statements, whose proofs are deferred to Sections 7.5 and 7.8. We apply the differential equation method in Section 7.5, and introduce the 'transfer theorem' in Section 7.6. Next, in Section 7.7 we collect properties of the binomial random graph, which are then used to complete the proof in Section 7.8.

7.2 Preliminaries and notation

In this section we first introduce some notation and briefly review properties of the C_ℓ-free process needed in our argument (we closely follow [17] and the reader familiar with the results of Bohman and Keevash may wish to skip Sections 7.2.1–7.2.3). Afterwards we state several probabilistic tools that we will use in our argument.

7.2.1 Terminology and notation

Let $G(i)$ denote the graph with vertex set $[n] = \{1, \ldots, n\}$ after i steps of the C_ℓ-free process. Its edge set $E(i)$ contains i edges; we partition the remaining non-edges $\binom{[n]}{2} \setminus E(i)$ into two sets, $O(i)$ and $C(i)$, which we call *open* and *closed* pairs, respectively. We say that a pair uv of vertices is *open* in $G(i)$ if $G(i) \cup \{uv\}$ contains no copy of C_ℓ. So, the C_ℓ-free process always chooses the next edge e_{i+1} uniformly at random from $O(i)$. In addition, for $uv \in O(i) \cup C(i)$ we write $C_{uv}(i)$ for the set of pairs $xy \in O(i)$ such that adding uv and xy to $G(i)$ creates a copy of C_ℓ containing both uv and xy. Note that $uv \in O(i)$ would become closed, i.e., belong to $C(i+1)$, if $e_{i+1} \in C_{uv}(i)$.

With a given graph in mind, we denote the *neighbourhood* of a vertex v by $\Gamma(v)$, where, as usual, $\Gamma(v)$ does not include v. For $S \subseteq [n]$ we define $\Gamma(S) = \bigcup_{v \in S} \Gamma(v)$. Furthermore, for $A, B \subseteq [n]$, let $e(A, B)$ denote the number of edges that have one endpoint in A and the other in B, where an edge with both ends in $A \cap B$ is counted once. If the graph under consideration is $G(i)$ we simply write $\Gamma_i(\cdot)$, but usually we omit the subscript if the corresponding i is clear from the context. Given a set S and an integer $k \geqslant 0$, we write $\binom{S}{k}$ for the set of all k-element subsets of S.

We use the symbol $\pm$ in two different ways, following [13, 17]. First, we denote by $a \pm b$ the interval $\{a + xb : -1 \leqslant x \leqslant 1\}$. Multiple occurrences are treated independently; for example, $\sum_{i \in [j]}(a_i \pm b_i)$ and $\prod_{i \in [j]}(a_i \pm b_i)$ mean $\{\sum_{i \in [j]}(a_i + x_i b_i) : -1 \leqslant x_1, \ldots, x_j \leqslant 1\}$ and $\{\prod_{i \in [j]}(a_i + x_i b_i) : -1 \leqslant x_1, \ldots, x_j \leqslant 1\}$, respectively. Slightly abusing notation, we also use the convention that $x = a \pm b$ means $x \in a \pm b$. Second, when considering pairs of random variables and functions, e.g. Y^+, Y^- and y^+, y^-, we use the superscript $\pm$ to denote two different statements: one with $\pm$ replaced by $+$, and the other with $\pm$ replaced by $-$. For example, $Y^\pm(i) = y^\pm(t)$ means $Y^+(i) = y^+(t)$ and $Y^-(i) = y^-(t)$. Finally, combinations of both ways are treated independently; for example, $Y^\pm(i) = y^\pm(t) \pm b$ means $Y^+(i) = y^+(t) \pm b$ and $Y^-(i) = y^-(t) \pm b$.

7.2.2 Parameters, functions and constants

In the remainder of this chapter we fix $\ell \geqslant 4$. Following [17], we introduce constants ε, μ and W. We choose W sufficiently large and afterwards ε and μ small enough such that, in addition to the constraints implicit in [17] for $H = C_\ell$, we have

$$W \geqslant \ell^2 2^{\ell+1} \geqslant 50, \qquad \varepsilon \leqslant 1/\big(2^{15}\ell^3\big) \qquad \text{and} \qquad 2W\mu^{\ell-1} \leqslant \varepsilon. \tag{7.1}$$

Since the additional constraints in [17] only depend on $H = C_\ell$, we deduce that μ is an absolute constant (depending only on ℓ). Next, similar as in [17] we set

$$p = n^{-1+1/(\ell-1)}, \quad t_{\max} = \mu(\log n)^{1/(\ell-1)} \quad \text{and} \quad m = n^2 p t_{\max} = \mu n^{\ell/(\ell-1)}(\log n)^{1/(\ell-1)}. \tag{7.2}$$

Formally, m (a number of steps) should be defined as $\lfloor n^2 p t_{\max} \rfloor$, say, but, as usual, we will henceforth ignore the irrelevant rounding to integers. For every step i we define $t = t(i) = i/(n^2 p)$, where, for the sake of brevity, we simply write t if the corresponding i is clear from the context. Next we introduce the functions

$$q(t) = e^{-(2t)^{\ell-1}} \qquad \text{and} \qquad f(t) = e^{(t^{\ell-1}+t)W}. \tag{7.3}$$

Now, using (7.1), for every $0 \leqslant t \leqslant t_{\max}$, for n large enough we readily obtain

$$1 \geqslant q(t) \geqslant n^{-\varepsilon/4} \qquad \text{and} \qquad 1 \leqslant f(t)q(t)^\ell \leqslant f(t) \leqslant n^\varepsilon. \tag{7.4}$$

7.2.3 Previous results for the C_ℓ-free process

The results of Bohman and Keevash [17] imply that a wide range of random variables are dynamically concentrated throughout the first m steps of the C_ℓ-free process. For our argument the key properties are estimates on the number of open pairs as well as bounds for the degree and certain closed pairs. So, for the reader's convenience we state their results here in a simplified form.

Theorem 7.2.1. [17] *Set* $s_e = n^{1/(2\ell)-\varepsilon}$. *Let* $\mathcal{T}_j$ *denote the event that for every* $0 \leqslant i \leqslant j$, *we have* $|O(i)| > 0$ *as well as*

$$|O(i)| = (1 \pm 3f(t)/s_e)\, q(t)n^2/2 \qquad \text{and} \tag{7.5}$$

$$|\Gamma_i(v)| \leqslant 3npt_{\max} \qquad \text{for all vertices } v \in [n]. \tag{7.6}$$

Let $\mathcal{J}_j$ *denote the event that for every* $0 \leqslant i \leqslant j$ *we have*

$$|C_{uv}(i)| = \Big((\ell-1)(2t)^{\ell-2}q(t) \pm 7\ell f(t)/s_e\Big)\, p^{-1} \quad \text{for all } uv \in O(i) \cup C(i) \text{ and} \tag{7.7}$$

$$|C_{u'v'}(i) \cap C_{u''v''}(i)| \leqslant n^{-1/\ell}p^{-1} \qquad \text{for all distinct } u'v', u''v'' \in O(i). \tag{7.8}$$

Then $\mathcal{J}_m \cap \mathcal{T}_m$ *holds whp in the* C_ℓ*-free process.* □

After some simple estimates, both (7.5) and (7.6) follow directly from Theorem 1.4 in [17]. Now, using $\mathrm{aut}(C_\ell) = 2\ell$ and $(2t)^{\ell-2}q(t) \leqslant 1$, which follow from elementary considerations, Corollary 6.2 and Lemma 8.4 in [17] imply (7.7) and (7.8). (Because the 'high probability events' of [17] in fact hold with probability at least $1-n^{-\omega(1)}$, we may take the union bound over all steps and pairs.) We remark that there is a factor of 2 difference in (7.7) since we use unordered instead of ordered pairs.

In our argument we use two additional properties of the C_ℓ-free process. The next lemma follows from Lemmas 4.2 and 4.3 in [76], which in turn are based on Lemmas 4.1–4.3 in [17].

Lemma 7.2.2. [76] Let $\mathcal{K}_i$ denote the event that for all $a, b \geqslant 1$ and every $A, B \subseteq [n]$ with $|A| = a$ and $|B| = b$, in $G(i)$ we have $e(A,B) < \max\{4\varepsilon^{-1}(a+b), pabn^{2\varepsilon}\}$. Let $\mathcal{L}_i$ denote the event that for all $a \geqslant 1$ and $d \geqslant \max\{16\varepsilon^{-1}, 2apn^{2\varepsilon}\}$, for every $A \subseteq [n]$ with $|A| = a$ we have $|D_{A,d}(i)| < 16\varepsilon^{-1}d^{-1}a$, where $D_{A,d}(i) \subseteq [n]$ contains all vertices $v \in [n]$ with $|\Gamma(v) \cap A| \geqslant d$ in $G(i)$. Then the probability that $\mathcal{T}_m$ holds and $\mathcal{K}_m \cap \mathcal{L}_m$ fails is $o(1)$. □

7.2.4 Concentration inequalities

The following Chernoff bounds, see e.g. Section 2.1 of [43], provide estimates for the probability that a sum of independent indicator variables deviates substantially from its expected value.

Lemma 7.2.3 ('Chernoff bounds')**.** *Let* $X = \sum_{i\in[n]} X_i$*, where the* X_i*'s are independent Bernoulli-distributed random variables. Set* $\mu = \mathbb{E}[X]$*. Then for all* $t \geqslant 0$ *we have*

$$\mathbb{P}[X \leqslant \mu - t] \leqslant e^{-t^2/(2\mu)}. \tag{7.9}$$

Furthermore, for all $t \geqslant 7\mu$ *we have*

$$\mathbb{P}[X \geqslant t] \leqslant e^{-t}. \tag{7.10}$$

In our argument we need to estimate the probability that in the binomial random graph $G_{n,p}$ some subset of the vertices contains 'too many' copies of a certain graph. Rödl and Ruciński [67] showed that exponential upper-tail bounds can be obtained if we allow for deleting a few edges; this is usually referred to as the Deletion Lemma [44].

Lemma 7.2.4 ('Deletion Lemma')**.** *Suppose* $0 < p < 1$ *and that* $\mathcal{S}$ *is a family of subsets from* $\binom{[n]}{2}$*. We say that a graph* G contains $\alpha \in \mathcal{S}$ *if all the edges of* α *are present in* G*. Let* μ *denote the expected number of elements in* $\mathcal{S}$ *that are contained in* $G_{n,p}$*. Let* $\mathcal{DL}(b, k, \mathcal{S})$ *denote the event that there exists* $\mathcal{I}_0 \subseteq \mathcal{S}$ *with* $|\mathcal{I}_0| \leqslant b$ *such that, setting* $E_0 = \bigcup_{\alpha\in\mathcal{I}_0} \alpha$*,* $G(n, p) \setminus E_0$ *contains at most* $\mu + k$ *elements from* $\mathcal{S}$*. Then for every* $b, k > 0$ *the probability that* $\mathcal{DL}(b, k, \mathcal{S})$ *fails is at most*

$$\left(1 + \frac{k}{\mu}\right)^{-b} \leqslant \exp\left\{-\frac{bk}{\mu + k}\right\}.$$

In [76] we proved a slightly weaker variant of the above lemma for the H-free process, where H is strictly 2-balanced. The results of Section 7.6 will shed some light on this intriguing phenomenon.

7.3 Differential equation method

A crucial ingredient of our analysis is the differential equation method, which was developed by Wormald [80, 81] to show that in certain discrete stochastic processes a collection $\mathcal{V}$ of random variables is whp approximated by the solution of a suitably defined system of differential equations. Developing ideas of Bohman and Keevash [17], we introduced the following variant in [76]. It will be an important tool for showing that certain random variables are dynamically concentrated throughout the evolution of the C_ℓ-free process.

Lemma 7.3.1 ('Differential Equation Method')**.** *Suppose that $m = m(n)$ and $s = s(n)$ are positive parameters. Let $\mathcal{C} = \mathcal{C}(n)$ and $\mathcal{V} = \mathcal{V}(n)$ be sets. For every $0 \leqslant i \leqslant m$ set $t = t(i) = i/s$. Suppose we have a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ and random variables $X_\sigma(i)$ and $Y_\sigma^{\pm}(i)$ which satisfy the following conditions. Assume that for all $\sigma \in \mathcal{C} \times \mathcal{V}$ the random variables $X_\sigma(i)$ are non-negative and $\mathcal{F}_i$-measurable for all $0 \leqslant i \leqslant m$, and that for all $0 \leqslant i < m$ the random variables $Y_\sigma^{\pm}(i)$ are non-negative, $\mathcal{F}_{i+1}$-measurable and satisfy*

$$X_\sigma(i+1) - X_\sigma(i) = Y_\sigma^{+}(i) - Y_\sigma^{-}(i). \tag{7.11}$$

Furthermore, suppose that for all $0 \leqslant i \leqslant m$ and $\Sigma \in \mathcal{C}$ we have an event $\mathcal{B}_i(\Sigma) \in \mathcal{F}_i$. Then, for all $0 \leqslant i \leqslant m$ we define $\mathcal{B}_{\leqslant i}(\Sigma) = \bigcup_{0 \leqslant j \leqslant i} \mathcal{B}_j(\Sigma)$. In addition, suppose that for each $\sigma \in \mathcal{C} \times \mathcal{V}$ we have positive parameters $u_\sigma = u_\sigma(n)$, $\lambda_\sigma = \lambda_\sigma(n)$, $\beta_\sigma = \beta_\sigma(n)$, $\tau_\sigma = \tau_\sigma(n)$, $s_\sigma = s_\sigma(n)$ and $S_\sigma = S_\sigma(n)$, as well as functions $x_\sigma(t)$ and $f_\sigma(t)$ that are smooth and non-negative for $t \geqslant 0$. For all $0 \leqslant i^* \leqslant m$ and $\Sigma \in \mathcal{C}$, let $\mathcal{G}_{i^*}(\Sigma)$ denote the event that for every $0 \leqslant i \leqslant i^*$ and $\sigma = (\Sigma, j)$ with $j \in \mathcal{V}$ we have

$$X_\sigma(i) = \left(x_\sigma(t) \pm \frac{f_\sigma(t)}{s_\sigma}\right) S_\sigma. \tag{7.12}$$

Next, for all $0 \leqslant i^* \leqslant m$ let $\mathcal{E}_{i^*}$ denote the event that for every $0 \leqslant i \leqslant i^*$ and $\Sigma \in \mathcal{C}$ the event $\mathcal{B}_{\leqslant i-1}(\Sigma) \cup \mathcal{G}_i(\Sigma)$ holds. Moreover, assume that we have an event $\mathcal{H}_i \in \mathcal{F}_i$ for all $0 \leqslant i \leqslant m$ with $\mathcal{H}_{i+1} \subseteq \mathcal{H}_i$ for all $0 \leqslant i < m$. Finally, suppose that the following conditions hold:

1. *(Trend hypothesis) For all $0 \leqslant i < m$ and $\sigma = (\Sigma, j) \in \mathcal{C} \times \mathcal{V}$, whenever $\mathcal{E}_i \cap \neg\mathcal{B}_{\leqslant i}(\Sigma) \cap \mathcal{H}_i$ holds we have*

$$\mathbb{E}\big[Y_\sigma^{\pm}(i) \mid \mathcal{F}_i\big] = \left(y_\sigma^{\pm}(t) \pm \frac{h_\sigma(t)}{s_\sigma}\right) \frac{S_\sigma}{s}, \tag{7.13}$$

where $y_\sigma^\pm(t)$ *and* $h_\sigma(t)$ *are smooth non-negative functions such that*

$$x'_\sigma(t) = y_\sigma^+(t) - y_\sigma^-(t) \qquad \text{and} \qquad f_\sigma(t) \geqslant 2\int_0^t h_\sigma(\tau)\, d\tau + \beta_\sigma. \tag{7.14}$$

2. *(Boundedness hypothesis) For all* $0 \leqslant i < m$ *and* $\sigma = (\Sigma, j) \in \mathcal{C} \times \mathcal{V}$*, whenever* $\mathcal{E}_i \cap \neg\mathcal{B}_{\leqslant i}(\Sigma) \cap \mathcal{H}_i$ *holds we have*

$$Y_\sigma^\pm(i) \leqslant \frac{\beta_\sigma^2}{s_\sigma^2 \lambda_\sigma \tau_\sigma} \cdot \frac{S_\sigma}{u_\sigma}. \tag{7.15}$$

3. *(Initial conditions) For all* $\sigma \in \mathcal{C} \times \mathcal{V}$ *we have*

$$X_\sigma(0) = \left(x_\sigma(0) \pm \frac{\beta_\sigma}{3s_\sigma}\right) S_\sigma. \tag{7.16}$$

4. *(Bounded number of configurations and variables) We have*

$$\max\{|\mathcal{C}|, |\mathcal{V}|\} \leqslant \min_{\sigma \in \mathcal{C}\times\mathcal{V}} e^{u_\sigma}. \tag{7.17}$$

5. *(Additional technical assumptions) For all* $\sigma \in \mathcal{C} \times \mathcal{V}$ *we have*

$$s \geqslant \max\{15 u_\sigma \tau_\sigma (s_\sigma \lambda_\sigma/\beta_\sigma)^2, 9 s_\sigma \lambda_\sigma/\beta_\sigma\}, \qquad s/(18 s_\sigma \lambda_\sigma/\beta_\sigma) < m \leqslant s \cdot \tau_\sigma/1944, \tag{7.18}$$

$$\sup_{0\leqslant t\leqslant m/s} y_\sigma^\pm(t) \leqslant \lambda_\sigma, \qquad \int_0^{m/s} |x''_\sigma(t)|\, dt \leqslant \lambda_\sigma, \tag{7.19}$$

$$h_\sigma(0) \leqslant s_\sigma \lambda_\sigma \qquad \text{and} \qquad \int_0^{m/s} |h'_\sigma(t)|\, dt \leqslant s_\sigma \lambda_\sigma. \tag{7.20}$$

Then we have

$$\mathbb{P}[\neg\mathcal{E}_m \cap \mathcal{H}_m] \leqslant 4 \max_{\sigma\in\mathcal{C}\times\mathcal{V}} e^{-u_\sigma}.$$

An important feature of Lemma 7.3.1 is that the variables in $\mathcal{V}$ are tracked for every *configuration* $\Sigma \in \mathcal{C}$. However, it only gives approximation guarantees for the variables that 'belong' to Σ as long as the 'local' *bad event* $\mathcal{B}_{\leqslant i}(\Sigma)$ fails. For more details we refer to

Section 5.3 and Appendix A.1 in [76]. Here we just remark that if the above conditions 1–5 are satisfied for n large enough, $\mathcal{H}_m$ holds whp and $u_\sigma = \omega(1)$ for all $\sigma \in \mathcal{C} \times \mathcal{V}$, then Lemma 7.3.1 implies that $\mathcal{E}_m$ holds whp.

7.3.1 Proof of Lemma 7.3.1

The proof of Lemma 7.3.1 is based on large deviation inequalities. More precisely, suppose we have a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ and a sequence $X_0, X_1, \ldots$ of random variables where each X_i is $\mathcal{F}_i$-measurable. Then $X_0, X_1, \ldots$ is a *supermartingale* if $\mathbb{E}[X_{i+1}|\mathcal{F}_i] \leqslant X_i$ for all i and a *submartingale* if $\mathbb{E}[X_{i+1}|\mathcal{F}_i] \geqslant X_i$ for all i. Furthermore, we say that $X_0, X_1, \ldots$ is (M,N)*-bounded* if for all i we have

$$-M \leqslant X_{i+1} - X_i \leqslant N.$$

The following martingale inequalities are due to Bohman [13] and follow from the original martingale inequality of Hoeffding [39].

Lemma 7.3.2. [13, Lemma 7] *Suppose* $0 \equiv X_0, X_1, \ldots$ *is an* (M,N)*-bounded supermartingale with* $M \leqslant N/10$*. Then for any* $m \geqslant 1$ *and* $0 < a < mM$ *we have*

$$\mathbb{P}[X_m \geqslant a] \leqslant e^{-\frac{a^2}{3mMN}}.$$

Lemma 7.3.3. [13, Lemma 6] *Suppose* $0 \equiv X_0, X_1, \ldots$ *is an* (M,N)*-bounded submartingale with* $M \leqslant N/2$*. Then for any* $m \geqslant 1$ *and* $0 < a < mM$ *we have*

$$\mathbb{P}[X_m \leqslant -a] \leqslant e^{-\frac{a^2}{3mMN}}.$$

Observe that for supermartingales we have $\mathbb{E}[X_i] \leqslant \mathbb{E}[X_0]$ and for submartingales $\mathbb{E}[X_i] \geqslant \mathbb{E}[X_0]$. Intuitively, the former inequalities thus give (one sided) exponential error probabilities for deviating 'too much' from the expected value.

Proof of Lemma 7.3.1. The proof is similar to the proof of Lemma 7.3 in [17]. The important differences here are the more involved definition of the desired event $\mathcal{E}_m$, the modified error functions $f_\sigma(t)$ and $h_\sigma(t)$, as well as the new parameters λ_σ, β_σ and τ_σ. The main new ingredients are the configurations Σ and bad events $\mathcal{B}_i(\Sigma)$. Let us briefly outline the main ideas of the proof. First, using (7.13) we add and subtract appropriate functions from $Y_\sigma^\pm(i)$ in order to construct super-/submartingales with an initial value of 0. Suppose $\mathcal{H}_m$ holds and that $\mathcal{E}_m$ fails for the first time at step i. Roughly speaking, it suffices to consider the case when $\mathcal{G}_i(\Sigma)$ fails. But, if (7.12) is violated, then, as we shall see, this implies that at least one of our super-/submartingales deviates substantially from 0. By Lemma 7.3.2 and 7.3.3 these large deviations are very unlikely, and it turns out that even after using a union bound over all such events the resulting error probability is negligible.

First, we derive some additional inequalities that our functions satisfy. Using (7.20) we see that

$$\sup_{0 \leqslant t \leqslant m/s} h_\sigma(t) \leqslant h_\sigma(0) + \int_0^{m/s} |h'_\sigma(t)| \, dt \leqslant 2 s_\sigma \lambda_\sigma. \tag{7.21}$$

We claim that the following estimates hold for all $0 \leqslant i^* \leqslant m$, writing $t(i) = i/s$ and $t^* = i^*/s$:

$$\frac{1}{s} \sum_{i=0}^{i^*-1} x'_\sigma\big(t(i)\big) \cdot S_\sigma = \left(x_\sigma(t^*) - x_\sigma(0) \pm \frac{\lambda_\sigma}{s} \right) S_\sigma \qquad \text{and} \tag{7.22}$$

$$\frac{1}{s} \sum_{i=0}^{i^*-1} h_\sigma\big(t(i)\big) \cdot \frac{2S_\sigma}{s_\sigma} = \left(2\int_0^{t^*} h_\sigma(t) \, dt \pm \frac{2 s_\sigma \lambda_\sigma}{s} \right) \frac{S_\sigma}{s_\sigma}. \tag{7.23}$$

Both bounds are obtained with very similar calculations as in the proof of Lemma 7.3 in [17], using the Euler-Maclaurin summation formula (see e.g. [6]) and then estimating the approximation error with the additional technical assumptions (7.19) and (7.20). For concreteness, we use the following Euler-Maclaurin variant which is e.g. implicit in [6]: given integers $a < b$ and a function f with a continuous derivative on the interval $[a, b]$ we have

$$\left| \int_a^b f(x) dx - \sum_{k=a}^{b-1} f(k) \right| \leqslant \int_a^b |f'(x)| dx. \tag{7.24}$$

We start by proving (7.22). Elementary calculus shows

$$\int_0^{i^*} x'_\sigma\big(t(i)\big) \, di = s \int_0^{t^*} x'_\sigma(t) \, dt = s\big[x_\sigma(t^*) - x_\sigma(0)\big]. \tag{7.25}$$

Furthermore, using the Euler-Maclaurin summation formula (7.24) and $t(i) = i/s$, we see that

$$\left| \int_0^{i^*} x'_\sigma\big(t(i)\big)\ di - \sum_{i=0}^{i^*-1} x'_\sigma\big(t(i)\big) \right| \leqslant \frac{1}{s} \int_0^{i^*} |x''_\sigma\big(t(i)\big)|\ di = \int_0^{t^*} |x''_\sigma(t)|\ dt. \tag{7.26}$$

Now, combining (7.25) and (7.26) with the additional technical assumptions (7.19), we deduce

$$\left| x_\sigma(t^*) - x_\sigma(0) - \frac{1}{s} \sum_{i=0}^{i^*-1} x'_\sigma\big(t(i)\big) \right| \leqslant \frac{1}{s} \int_0^{t^*} |x''_\sigma(t)|\ dt \leqslant \frac{\lambda_\sigma}{s},$$

which in turn implies (7.22). Using (7.20), with a similar calculation we obtain

$$\left| \int_0^{t^*} h_\sigma(t)\ dt - \frac{1}{s} \sum_{i=0}^{i^*-1} h_\sigma\big(t(i)\big) \right| \leqslant \frac{1}{s} \int_0^{t^*} |h'_\sigma(t)|\ dt \leqslant \frac{s_\sigma \lambda_\sigma}{s},$$

from which (7.23) readily follows.

Second, we define several random variables and start with $Y_\sigma^{\pm_1\pm_2}$ (recall that this is an abbreviation for four different variables, one for each way of choosing $\pm_1$ and $\pm_2$). For all $(\Sigma, j) = \sigma \in \mathcal{C} \times \mathcal{V}$, if $\mathcal{E}_i \cap \neg\mathcal{B}_{\leqslant i}(\Sigma) \cap \mathcal{H}_i$ holds we set

$$Y_\sigma^{\pm_1\pm_2}(i) := Y_\sigma^{\pm_1}(i) - \left(y_\sigma^{\pm_1}(t) \mp_2 \frac{h_\sigma(t)}{s_\sigma} \right) \frac{S_\sigma}{s}, \tag{7.27}$$

and, otherwise (i.e., if $\neg\mathcal{E}_i \cup \mathcal{B}_{\leqslant i}(\Sigma) \cup \neg\mathcal{H}_i$ holds) we define $Y_\sigma^{\pm_1\pm_2}(i)$ to be 0. Note that in this case $Y_\sigma^{\pm_1\pm_2}(i') = 0$ for all $i' \geqslant i$. Next, we define

$$Z_\sigma^{\pm_1\pm_2}(i) := \sum_{i'=0}^{i-1} Y_\sigma^{\pm_1\pm_2}(i'), \qquad M_\sigma := \frac{3\lambda_\sigma S_\sigma}{s} \qquad \text{and} \qquad N_\sigma := \frac{2\beta_\sigma^2}{s_\sigma^2 \lambda_\sigma \tau_\sigma} \cdot \frac{S_\sigma}{u_\sigma} \tag{7.28}$$

We claim that $Z_\sigma^{\pm-}(i)$ and $Z_\sigma^{\pm+}(i)$ are (M_σ, N_σ)-bounded super-/submartingales with $Z_\sigma^{\pm_1\pm_2}(0) = 0$ and $M_\sigma \leqslant N_\sigma/10$. Clearly, $Z_\sigma^{\pm_1\pm_2}(0) = 0$ holds, and, furthermore, (7.18) implies $M_\sigma \leqslant N_\sigma/10$. For bounding the maximum change $Z_\sigma^{\pm_1\pm_2}(i+1) - Z_\sigma^{\pm_1\pm_2}(i)$ we may assume that $\mathcal{E}_i \cap \neg\mathcal{B}_{\leqslant i}(\Sigma) \cap \mathcal{H}_i$ holds (otherwise the difference is by definition equal to 0). In that case

$$Z_\sigma^{\pm_1\pm_2}(i+1) - Z_\sigma^{\pm_1\pm_2}(i) \;=\; Y_\sigma^{\pm_1\pm_2}(i) \;=\; Y_\sigma^{\pm_1}(i) - \left(y_\sigma^{\pm_1}(t) \mp_2 \frac{h_\sigma(t)}{s_\sigma}\right)\frac{S_\sigma}{s}. \qquad (7.29)$$

Now, using the boundedness hypothesis (7.15) as well as $y_\sigma^{\pm}(t) \geqslant 0$ and (7.21), i.e., $h_\sigma(t) \leqslant 2s_\sigma\lambda_\sigma$, we see that (7.29) is bounded from above by

$$\frac{\beta_\sigma^2}{s_\sigma^2\lambda_\sigma\tau_\sigma}\frac{S_\sigma}{u_\sigma} + \frac{h_\sigma(t)}{s_\sigma}\frac{S_\sigma}{s} \leqslant \frac{N_\sigma}{2} + \frac{2\lambda_\sigma S_\sigma}{s} \leqslant N_\sigma,$$

where we used (7.18), i.e., $s \geqslant 15u_\sigma\tau_\sigma(s_\sigma\lambda_\sigma/\beta_\sigma)^2$, and (7.28) for the last inequality. Similarly, using $Y_\sigma^{\pm}(i), h_\sigma(t) \geqslant 0$ and (7.19), i.e., $y_\sigma^{\pm}(t) \leqslant \lambda_\sigma$, we see that (7.29) is bounded from below by

$$-\left(\lambda_\sigma + \frac{h_\sigma(t)}{s_\sigma}\right)\frac{S_\sigma}{s} \geqslant -\frac{3\lambda_\sigma S_\sigma}{s} = -M_\sigma.$$

For checking the super-/submartingale property we may again assume that $\mathcal{E}_i \cap \neg\mathcal{B}_{\leqslant i}(\Sigma) \cap \mathcal{H}_i$ holds (otherwise the value of $Z_\sigma^{\pm_1\pm_2}(i)$ remains unchanged). Now by combining (7.29) with the trend hypothesis (7.13), it is easy to see that $Z_\sigma^{\pm -}(i)$ is a supermartingale and $Z_\sigma^{\pm +}(i)$ a submartingale.

In the following we estimate the probability of the event $\neg\mathcal{E}_m \cap \mathcal{H}_m$. Loosely speaking, we focus on the first step $i^* \leqslant m$ where $\mathcal{E}_i$ fails, and, in particular, on the $\Sigma \in \mathcal{C}$ for which $\mathcal{B}_{\leqslant i^*-1}(\Sigma) \cup \mathcal{G}_{i^*}(\Sigma)$ fails. Note that (7.14) and (7.16) ensure that $\mathcal{G}_0(\Sigma)$ holds for all $\Sigma \in \mathcal{C}$, and thus $\mathcal{E}_0$ holds. So, considering all $i^* \leqslant m$, $\Sigma \in \mathcal{C}$ and using $\mathcal{H}_m \subseteq \mathcal{H}_{i^*-1}$, we have

$$\begin{aligned}\neg\mathcal{E}_m \cap \mathcal{H}_m &\subseteq \bigcup_{1\leqslant i^*\leqslant m} \left[\mathcal{H}_m \cap \mathcal{E}_{i^*-1} \cap \neg\mathcal{E}_{i^*}\right] \\ &\subseteq \bigcup_{1\leqslant i^*\leqslant m}\bigcup_{\Sigma\in\mathcal{C}} \left[\mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\big(\mathcal{B}_{\leqslant i^*-1}(\Sigma) \cup \mathcal{G}_{i^*}(\Sigma)\big)\right].\end{aligned} \qquad (7.30)$$

Henceforth we fix $1 \leqslant i^* \leqslant m$ and $\Sigma \in \mathcal{C}$. Using that $\mathcal{E}_i \cap \neg\mathcal{B}_{\leqslant i}(\Sigma)$ implies $\mathcal{G}_i(\Sigma)$, we see that

$$\mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\big(\mathcal{B}_{\leqslant i^*-1}(\Sigma) \cup \mathcal{G}_{i^*}(\Sigma)\big) \;=\; \mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\mathcal{B}_{\leqslant i^*-1}(\Sigma) \cap \mathcal{G}_{i^*-1}(\Sigma) \cap \neg\mathcal{G}_{i^*}(\Sigma).$$

Observe that when $\mathcal{G}_{i^*-1}(\Sigma)$ holds, the event $\mathcal{G}_{i^*}(\Sigma)$ can only fail if $X_\sigma(i^*)$ violates (7.12) for some $\sigma = (\Sigma, j)$ with $j \in \mathcal{V}$, and for the following calculations we fix such a $\sigma = (\Sigma, j)$.

Suppose that $\mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\mathcal{B}_{\leqslant i^*-1}(\Sigma)$ holds and $X_\sigma(i^*)$ fails to satisfy (7.12) because $X_\sigma(i^*) > (x_\sigma(t^*) + f_\sigma(t^*)/s_\sigma)S_\sigma$. With a virtually identical calculation as in the proof of the Lemma 7.3 in [17], whenever $\mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\mathcal{B}_{\leqslant i^*-1}(\Sigma)$ holds we now establish

$$Z_\sigma^{+\pm_2}(i^*) - Z_\sigma^{-\mp_2}(i^*) = X_\sigma(i^*) - X_\sigma(0) - \frac{1}{s}\sum_{i=0}^{i^*-1} x'_\sigma\big(t(i)\big) \cdot S_\sigma \pm_2 \frac{1}{s}\sum_{i=0}^{i^*-1} h_\sigma\big(t(i)\big) \cdot \frac{2S_\sigma}{s_\sigma}. \quad (7.31)$$

Indeed, using (7.11) and (7.14) as well as the definition of $Y_\sigma^{\pm}(i)$ and $Z_\sigma^{+\pm_2}(i)$, i.e. (7.27) and (7.28), because $\mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\mathcal{B}_{\leqslant i^*-1}(\Sigma)$ holds we have

$$\begin{aligned}
X_\sigma(i^*) - X_\sigma(0) - \sum_{i=0}^{i^*-1} x'_\sigma\big(t(i)\big) \cdot \frac{S_\sigma}{s} &= \sum_{i=0}^{i^*-1}\left(X_\sigma(i+1) - X_\sigma(i) - x'_\sigma\big(t(i)\big)\frac{S_\sigma}{s}\right)\\
&= \sum_{i=0}^{i^*-1}\left(Y_\sigma^{+}(i) - y_\sigma^{+}\big(t(i)\big)\frac{S_\sigma}{s} - Y_\sigma^{-}(i) + y_\sigma^{-}\big(t(i)\big)\frac{S_\sigma}{s}\right)\\
&= \sum_{i=0}^{i^*-1} Y_\sigma^{+\pm_2}(i) - \sum_{i=0}^{i^*-1} Y_\sigma^{-\mp_2}(i) \mp_2 2\sum_{i=0}^{i^*-1}\frac{h_\sigma\big(t(i)\big)}{s_\sigma}\cdot\frac{S_\sigma}{s}\\
&= Z_\sigma^{+\pm_2}(i^*) - Z_\sigma^{-\mp_2}(i^*) \mp_2 \frac{1}{s}\sum_{i=0}^{i^*-1} h_\sigma\big(t(i)\big)\cdot\frac{2S_\sigma}{s_\sigma}.
\end{aligned}$$

Rearranging gives (7.31), which in particular implies

$$Z_\sigma^{+-}(i^*) - Z_\sigma^{-+}(i^*) = X_\sigma(i^*) - X_\sigma(0) - \frac{1}{s}\sum_{i=0}^{i^*-1} x'_\sigma\big(t(i)\big) \cdot S_\sigma - \frac{1}{s}\sum_{i=0}^{i^*-1} h_\sigma\big(t(i)\big) \cdot \frac{2S_\sigma}{s_\sigma}, \quad (7.32)$$

With this in hand, using the lower bound on $X_\sigma(i^*)$, the initial condition (7.16) as well as (7.22) and (7.23), we deduce

$$Z_\sigma^{+-}(i^*) - Z_\sigma^{-+}(i^*) > \left(f_\sigma(t^*) - 2\int_0^{t^*} h_\sigma(t)\ dt - \frac{\beta_\sigma}{3} - \frac{3s_\sigma\lambda_\sigma}{s}\right)\frac{S_\sigma}{s_\sigma} \geqslant \frac{\beta_\sigma}{3}\frac{S_\sigma}{s_\sigma},$$

where we used (7.14) and (7.18), i.e., $s \geqslant 9s_\sigma\lambda_\sigma/\beta_\sigma$, for the last inequality. This readily implies

$$Z_\sigma^{+-}(i^*) \geqslant \frac{\beta_\sigma}{6}\frac{S_\sigma}{s_\sigma} =: a \qquad \text{or} \qquad Z_\sigma^{-+}(i^*) \leqslant -\frac{\beta_\sigma}{6}\frac{S_\sigma}{s_\sigma} = -a. \quad (7.33)$$

Since the variables $Z_\sigma^{\pm_1\pm_2}(i)$ are 'frozen' once $X_\sigma(i)$ leaves the allowed range (7.12), we deduce that $Z_\sigma^{+-}(m) \geqslant a$ or $Z_\sigma^{-+}(m) \leqslant -a$ holds.

Similarly, if $\mathcal{H}_{i^*-1} \cap \mathcal{E}_{i^*-1} \cap \neg\mathcal{B}_{\leqslant i^*-1}(\Sigma)$ holds and $X_\sigma(i^*)$ fails to satisfy (7.12) because $X_\sigma(i^*) < (x_\sigma(t^*) - f_\sigma(t^*)/s_\sigma)S_\sigma$, with calculations completely analogous to those of the previous case, we deduce that $Z_\sigma^{--}(m) \geqslant a$ or $Z_\sigma^{++}(m) \leqslant -a$ holds.

Plugging our findings into (7.30), we obtain

$$\neg\mathcal{E}_m \cap \mathcal{H}_m \subseteq \bigcup_{\sigma \in \mathcal{C}\times\mathcal{V}} \big[\{Z_\sigma^{+-}(m) \geqslant a\} \cup \{Z_\sigma^{--}(m) \geqslant a\} \cup \{Z_\sigma^{++}(m) \leqslant -a\} \cup \{Z_\sigma^{-+}(m) \leqslant -a\}\big]. \tag{7.34}$$

Recall that $Z_\sigma^{\pm-}(i)$ and $Z_\sigma^{\pm+}(i)$ are (M_σ, N_σ)-bounded super-/submartingales with $M_\sigma \leqslant N_\sigma/10$ and initial values of 0. Note that (7.18), i.e., $m > s/(18 s_\sigma \lambda_\sigma/\beta_\sigma)$, implies $a < mM_\sigma$. Therefore, using Lemmas 7.3.2 and 7.3.3 as well as the definition of a, M_σ and N_σ, we deduce that the probabilities of $Z_\sigma^{\pm-}(m) \geqslant a$ and $Z_\sigma^{\pm+}(m) \leqslant -a$ are each bounded by

$$\exp\left\{-\frac{a^2}{3mM_\sigma N_\sigma}\right\} = \exp\left\{-\frac{1}{648}\frac{s}{m}\tau_\sigma u_\sigma\right\} \leqslant \exp\{-3u_\sigma\}, \tag{7.35}$$

where we used (7.18), i.e., $m \leqslant s \cdot \tau_\sigma/1944$, for the last inequality. Finally, we estimate (7.34) with a union bound argument. Using (7.17) and (7.35) we deduce

$$\mathbb{P}[\neg\mathcal{E}_m \cap \mathcal{H}_m] \leqslant \sum_{\sigma \in \mathcal{C}\times\mathcal{V}} 4e^{-3u_\sigma} \leqslant 4 \max_{\sigma \in \mathcal{C}\times\mathcal{V}} e^{-u_\sigma},$$

and the proof is complete. □

7.3.2 Improved variant for the setup of Bohman and Keevash

Now we formulate the improved version of Lemma 7.3 in [17], which can be obtained by adapting the ideas/modifications we used in the proof of Lemma 7.3.1 back to the original setup used in [17] by Bohman and Keevash. Intuitively, there are different 'types' $j \in \mathcal{V}$ of random variables, where $\sigma \in \mathcal{I}_j$ denotes particular 'instances', which can e.g. take into account different 'positions' in a graph.

Lemma 7.3.4 ('Differential Equation Method')**.** *Suppose that* $m = m(n)$ *and* $s = s(n)$

are positive parameters. Let $\mathcal{V} = \mathcal{V}(n)$ *be a set, and* $\{\mathcal{I}_j\}_{j\in\mathcal{V}}$ *be a family of sets, where* $\mathcal{I}_j = \mathcal{I}_j(n)$*. For every* $0 \leqslant i \leqslant m$ *set* $t = t(i) := i/s$*. Suppose we have a filtration* $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ *and random variables* $X_\sigma(i)$ *and* $Y_\sigma^\pm(i)$ *which satisfy the following conditions. Assume that for all* $j \in \mathcal{V}$ *and* $\sigma \in \mathcal{I}_j$ *the random variables* $X_\sigma(i)$ *are non-negative and* $\mathcal{F}_i$*-measurable for all* $0 \leqslant i \leqslant m$*, and that for all* $0 \leqslant i < m$ *the random variables* $Y_\sigma^\pm(i)$ *are non-negative,* $\mathcal{F}_{i+1}$*-measurable and satisfy*

$$X_\sigma(i+1) - X_\sigma(i) = Y_\sigma^+(i) - Y_\sigma^-(i). \tag{7.36}$$

In addition, suppose that for each $j \in \mathcal{V}$ *and* $\sigma \in \mathcal{I}_j$ *we have positive parameters* $u_\sigma = u_\sigma(n)$*,* $\lambda_\sigma = \lambda_\sigma(n)$*,* $\beta_\sigma = \beta_\sigma(n)$*,* $\tau_\sigma = \tau_\sigma(n)$*,* $s_\sigma = s_\sigma(n)$ *and* $S_\sigma = S_\sigma(n)$*, as well as functions* $x_\sigma(t)$ *and* $f_\sigma(t)$ *that are smooth and non-negative for* $t \geqslant 0$*. For all* $0 \leqslant i^* \leqslant m$*, let* $\mathcal{G}_{i^*}$ *denote the event that for all* $0 \leqslant i \leqslant i^*$*,* $j \in \mathcal{V}$ *and* $\sigma \in \mathcal{I}_j$*, we have*

$$X_\sigma(i) = \left(x_\sigma(t) \pm \frac{f_\sigma(t)}{s_\sigma}\right) S_\sigma. \tag{7.37}$$

Moreover, assume that we have an event $\mathcal{H}_i \in \mathcal{F}_i$ *for all* $0 \leqslant i \leqslant m$ *with* $\mathcal{H}_{i+1} \subseteq \mathcal{H}_i$ *for all* $0 \leqslant i < m$*. Finally, suppose that for* n *large enough the following conditions hold:*

1. *(Trend hypothesis) For all* $0 \leqslant i < m$*,* $j \in \mathcal{V}$ *and* $\sigma \in \mathcal{I}_j$*, whenever* $\mathcal{G}_i \cap \mathcal{H}_i$ *holds we have*
$$\mathbb{E}\big[Y_\sigma^{\pm_1}(i) \mid \mathcal{F}_i\big] = \left(y_\sigma^{\pm_1}(t) \pm \frac{h_\sigma(t)}{s_\sigma}\right) \frac{S_\sigma}{s}, \tag{7.38}$$
where $y_\sigma^\pm(t)$ *and* $h_\sigma(t)$ *are smooth non-negative functions such that*
$$x'_\sigma(t) = y_\sigma^+(t) - y_\sigma^-(t) \qquad \text{and} \qquad f_\sigma(t) \geqslant 2\int_0^t h_\sigma(\tau)\, d\tau + \beta_\sigma. \tag{7.39}$$

2. *(Boundedness hypothesis) For all* $0 \leqslant i < m$*,* $j \in \mathcal{V}$ *and* $\sigma \in \mathcal{I}_j$*, whenever* $\mathcal{E}_i \cap \mathcal{H}_i$ *holds we have*
$$Y_\sigma^\pm(i) \leqslant \frac{\beta_\sigma^2}{s_\sigma^2 \lambda_\sigma \tau_\sigma} \cdot \frac{S_\sigma}{u_\sigma}. \tag{7.40}$$

3. (Initial conditions) For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have

$$X_\sigma(0) = \left(x_\sigma(0) \pm \frac{\beta_\sigma}{3s_\sigma}\right) S_\sigma. \tag{7.41}$$

4. (Bounded number of variables) For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have

$$\max\{|\mathcal{V}|, |\mathcal{I}_j|\} \leqslant e^{u_\sigma}. \tag{7.42}$$

5. (High probability event) The event $\mathcal{H}_i$ satisfies

$$\mathbb{P}[\exists i \leqslant m : \ \mathcal{G}_i \cap \neg\mathcal{H}_i] = o(1). \tag{7.43}$$

6. (Additional technical assumptions) For all $j \in \mathcal{V}$ and $\sigma \in \mathcal{I}_j$ we have $u_\sigma = \omega(1)$ as well as

$$s \geqslant \max\{15u_\sigma\tau_\sigma(s_\sigma\lambda_\sigma/\beta_\sigma)^2, 9s_\sigma\lambda_\sigma/\beta_\sigma\}, \qquad s/(18s_\sigma\lambda_\sigma/\beta_\sigma) < m \leqslant s \cdot \tau_\sigma/1944, \tag{7.44}$$

$$\sup_{0\leqslant t\leqslant m/s} y_\sigma^{\pm}(t) \leqslant \lambda_\sigma, \qquad \int_0^{m/s} |x''_\sigma(t)| \, dt \leqslant \lambda_\sigma, \tag{7.45}$$

$$h_\sigma(0) \leqslant s_\sigma\lambda_\sigma \qquad \textit{and} \qquad \int_0^{m/s} |h'_\sigma(t)| \, dt \leqslant s_\sigma\lambda_\sigma. \tag{7.46}$$

Then $\mathcal{G}_m \cap \mathcal{H}_m$ holds with high probability.

Compared to Lemma 7.3 in [17], one important advantage of Lemma 7.3.4 is that we state the approximation error in a simpler form and allow for more freedom in choosing the corresponding error functions; this should make it easier to apply our variant in other contexts. If possible, it is often convenient to choose the same parametrization for all $\sigma \in \mathcal{I}_j$, e.g. $x_\sigma(t) = x_j(t)$, since they typically correspond to different 'instances' of the same type of random variables. We point out that by using this simplification, then choosing $s_j \geqslant n^\varepsilon$ as well as $u_j = 2k_j \log n$, and afterwards setting $\tau_j = 1944n^{\varepsilon/2}$, $\beta_j^{-1} := \lambda_j := n^{\varepsilon/7}$, $h_j(t) := (e_j \cdot x_j + \gamma_j)'(t)/4$ and $f_j(t) := e_j(t)x_j(t) - \theta_j(t)e_j(t)/s_j + \theta_j(t)$, this not only implies Lemma 7.3 in [17], but also weakens certain assumptions significantly. For example, in the additional technical assumptions we relax $y_j^{\pm}(t) = O(1)$ to $y_j^{\pm}(t) \leqslant n^{\varepsilon/7}$, their c from $\Omega(1)$ to $\Omega(n^{-\varepsilon/7})$, and also weaken the lower bound on m from $m > s$ to, say, $m \geqslant sn^{-\varepsilon}$. Furthermore, we e.g. improve and simplify the initial conditions by allowing for a small error in the initial value $X_\sigma(0)$ and removing the requirement that $e_j(0) = \gamma_j(0) = 0$. At first sight our assumption that $\mathcal{H}_i$ satisfies (7.43) seems to be more restrictive, however, due to an oversight in the proof given by Bohman and Keevash in [17], their Lemma 7.3 also needs this additional assumption, which of course holds in their application. Another new ingredient is the introduction of the parameters λ_σ, β_σ and τ_σ, which allows for a trade-off between the approximation error, the boundedness hypothesis and the additional technical assumptions. For example, in certain applications this might allow for larger one-step changes in (7.40), since in contrast to Lemma 7.3 in [17], we do not rule out the possibility that our parameters are small, say, $o(n^\varepsilon)$. Finally, as noted in [17], compared to Wormald's formulation of the differential equation method [80, 81], if applicable, Lemma 7.3.4 has the advantage that in certain applications much weaker estimates on the one-step changes suffice.

7.4 Bounding the maximum degree

In this section we prove our main result, namely that whp the maximum degree in the final graph of the C_ℓ-free process is $O((n \log n)^{1/(\ell-1)})$. In Sections 7.4.1 and 7.4.2 we first discuss the main proof ideas and introduce the formal setup used. Section 7.4.3 is then devoted to the proof of Theorem 7.1.1, which in turn relies on two involved statements that are proved in subsequent sections.

7.4.1 Sketch of the proof

The following definition plays a crucial role in our proof. Given $(\tilde{v}, U)$, where $\tilde{v} \in [n]$ and $U \subseteq [n] \setminus \{\tilde{v}\}$, a *$C_\ell$-extension for $(\tilde{v}, U)$* is a path on $\ell - 1$ vertices whose end vertices are in U and whose remaining vertices are disjoint from $U \cup \{\tilde{v}\}$. Clearly, for every vertex $\tilde{v} \in [n]$, in the final graph of the C_ℓ-free process $(\tilde{v}, \Gamma(\tilde{v}))$ must not have a C_ℓ-extension. Set

$$\delta = \frac{1}{60^2 \ell! \ell^\ell}, \qquad \gamma = \max\left\{\frac{3^{\ell+1}}{\delta \mu^{\ell-1}}, 180\right\} \qquad \text{and} \qquad u = \gamma n p t_{\max} = \gamma \mu (n \log n)^{1/(\ell-1)}, \tag{7.47}$$

again ignoring the irrelevant rounding to integers in the definition of u. In order to bound the maximum degree by $u = D(n \log n)^{1/(\ell-1)}$, where $D = \gamma\mu$, it is enough to prove that whp every $(\tilde{v}, U) \in [n] \times \binom{[n]}{u}$ with $\tilde{v} \notin U$ has at least one C_ℓ-extension after the first m steps. The same basic idea was used in [57], but our proof takes a different route, inspired by our earlier analysis of the K_4-free process [76]. After i steps, we denote by $O_{\tilde{v},U}(i)$ the set of open pairs which would complete a C_ℓ-extension for $(\tilde{v}, U)$ if chosen as the next edge. It seems plausible that in order prove Theorem 7.1.1, it suffices to show that, after some initial number of steps, $|O_{\tilde{v},U}(i)|$ is always not too small. Indeed, this implies a reasonable probability of completing such an extension in each step, which in turn suggests that the probability of avoiding a C_ℓ-extension in all of the first m steps is very small.

We now illustrate our approach for establishing a good lower bound on $|O_{\tilde{v},U}(i)|$ for the case when $\ell = 5$. For ease of exposition, we ignore n^ε factors whenever these are not crucial and also assume that the number of steps i is large. So, in our rough calculations we will e.g. ignore whether an edge is open or not, since $|O(i)| = \omega(n^{2-\varepsilon})$ by (7.4) and (7.5). Note that in this case we have $p = n^{-3/4}$, $m \approx n^{5/4}$, $|C_{xy}(i)| \approx p^{-1}$ and $|U| \approx np = n^{1/4}$ by (7.2), (7.7) and (7.47).

7.4.1.1 The random variables used

We define $O'_{\tilde{v},U}(i)$ as the set of pairs $xy \in O_{\tilde{v},U}(i)$ with $x \in U$ and $y \notin U \cup \{\tilde{v}\}$. Observe that for every $xy \in O'_{\tilde{v},U}(i)$ there exists a path $v_0 v_1 v_2 = y$ with $v_0 \in U \setminus \{x\}$ and $v_1 \notin U \cup \{\tilde{v}, x, y\}$, cf. Figure 7.1. The 'last' edge completing a C_5-extension for $(\tilde{v}, U)$ could be any one of the edges of the path, so we expect that $O'_{\tilde{v},U}(i)$ contains constant proportion of $O_{\tilde{v},U}(i)$.

Let $Z_{\tilde{v},U}(i)$ contain all quadruples $(v_0, v_1, v_2, v_3) \in U \times [n]^2 \times U$ with $\{v_0v_1, v_1v_2\} \subseteq E(i)$, $v_2 v_3 \in O(i)$ and $\{v_1, v_2\} \cap (U \cup \{\tilde{v}\}) = \emptyset$. Using random graphs as a guide, we expect that $G(i)$ shares many properties with the binomial random graph $G_{n,p}$, since its edge density is roughly $2tp \approx n^{-3/4} = p$. So, given y, the expected number of $v_0 \in U$ for which there

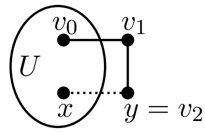

Figure 7.1: A pair $xy \in O'_{\tilde{v},U}(i)$. Solid lines represent edges and dotted lines open pairs.

exists a path $v_0v_1v_2 = y$ should be roughly $n|U|p^2 = o(1)$. Hence on average $xy \in O'_{\tilde{v},U}(i)$ is contained in only one such path ending in U, which suggests that up to constants $|Z_{\tilde{v},U}(i)| \approx |O'_{\tilde{v},U}(i)|$. To sum up, our discussion indicates that a reasonable lower bound for $|Z_{\tilde{v},U}(i)|$ suffices to prove that $|O_{\tilde{v},U}(i)|$ is large. For this we intend to use the differential equation method and so we introduce additional variables in order to control the one-step changes of $|Z_{\tilde{v},U}(i)|$. To this end let $Y_{\tilde{v},U}(i)$ be the set of all $(v_0, v_1, v_2, v_3) \in U \times [n]^2 \times U$ with $\{v_1, v_2\} \cap (U \cup \{\tilde{v}\}) = \emptyset$ that satisfy $v_0v_1 \in E(i)$, $\{v_1v_2, v_2v_3\} \subseteq O(i)$, and, similarly, let $X_{\tilde{v},U}(i)$ contain all such quadruples with $\{v_0v_1, v_1v_2, v_2v_3\} \subseteq O(i)$.

7.4.1.2 Technical difficulties

One of the main problems with the approach described above is the bound on the one-step changes. It can happen that in one step up to p^{-1} quadruples are removed from $Z_{\tilde{v},U}(i)$, which turns out to be too large for applying the differential equation method directly. Indeed, pick $\tilde{v}, U$ such that $\{v_0\} \cup \Gamma_i(w) \subseteq U$, $|\Gamma_i(w)| \approx |U|$ and $\tilde{v} \notin \{w\} \cup U \cup \Gamma_i(U)$; taking the random graph $G_{n,p}$ as a guide, for $e_{i+1} = wv_0$ it is easy to see that about $(np)^2|U| \approx p^{-1}$ quadruples (v_0, v_1, v_2, v_3) with $v_3 \in \Gamma_i(w)$ are removed from $Z_{\tilde{v},U}(i)$. For the C_4-free process this can be resolved using ad-hoc arguments (e.g. exploiting that every $v \neq \tilde{v}$ satisfies $|\Gamma_i(v) \cap U| \leqslant 1$ if no C_4-extension for $(\tilde{v}, U)$ exists), but for larger cycles the situation is more delicate. To overcome this issue, we consider a different random variable $T_{\tilde{v},U}(i)$, which is an approximation of $Z_{\tilde{v},U}(i)$ and is defined in such a way that the one-step changes are automatically not too large. Roughly speaking, this can be achieved by 'ignoring' the steps where the one-step changes would be too large; similar ideas have been used e.g. in [13, 17, 48, 76]. Clearly, this introduces a new difficulty: we need to ensure that we do not ignore 'too much', so that on the one hand the expected one-step changes are still 'correct', and on the other hand $|Z_{\tilde{v},U}(i)| \approx |T_{\tilde{v},U}(i)|$ holds. Consequently, we refine the tracked variables and use more sophisticated rules for ignoring tuples.

There is another significant obstacle when applying the differential equation method: adding $e_{i+1} = v_1v_2$ to $(v_0, v_1, v_2, v_3) \in Y_{\tilde{v},U}(i)$ does *not* always result in an element of $Z_{\tilde{v},U}(i+1)$, since $e_{i+1} = v_1v_2$ closes v_2v_3 whenever $v_2v_3 \in C_{v_1v_2}(i)$ holds. This is an important difference to the C_ℓ-free process with $\ell \leqslant 4$, where this does not cause any problems when bounding the maximum degree. For example, whenever this happens for $\ell = 4$, it is not difficult to deduce that at least one C_4-extension for $(\tilde{v}, U)$ already exists. Returning to the case $\ell = 5$, using our random graph intuition we expect that $|Y_{\tilde{v},U}(i)| \approx |U|^2n^2p \approx n^{7/4}$. Similar calculations suggest that the expected number of quadruples in $Y_{\tilde{v},U}(i)$ with $v_2v_3 \in C_{v_1v_2}(i)$ should be negligible compared to $|Y_{\tilde{v},U}(i)|$. However, if we pick U such that $\Gamma_i(w) \subseteq U$ and $|\Gamma_i(w)| \approx |U|$, for $\tilde{v} \notin \{w\} \cup U \cup \Gamma_i(U)$, it certainly can happen that there are $|U|^2 \cdot np \cdot n \approx |Y_{\tilde{v},U}(i)|$ quadruples in $Y_{\tilde{v},U}(i)$ with $v_2v_3 \in C_{v_1v_2}(i)$. In other words, it is simply *not* true that for all $(\tilde{v}, U)$ the effect of these 'bad' quadruples is negligible. This is a new difficulty in comparison to the variables tracked in the analysis of the H-free process [17]. To deal with this issue, we substantially refine the tracked random variables, developing ideas we used in [76]. Intuitively, we show that for every $(\tilde{v}, U)$ there exists a slightly altered set of random variables where the above extreme example (and other difficulties) can be avoided. Here the new 'transfer theorem' (Theorem 7.6.2) is an important ingredient, which allows us to use the *much* more tractable binomial random graph model for certain calculations (see Section 7.7).

7.4.2 Formal setup

We now introduce the formal setup used in our argument. In the following it is useful to keep in mind that we intend to apply the differential equation method (Lemma 7.3.1).

7.4.2.1 Preliminaries: neighbourhoods and partitions

Recall that by (7.47) we have $u = \gamma npt_{\max} = \gamma\mu(n \log n)^{1/(\ell-1)}$. We set

$$k = u/60 = \gamma/60 \cdot npt_{\max} = \gamma\mu/60 \cdot (n \log n)^{1/(\ell-1)} \qquad \text{and} \qquad r = \lfloor n/(\ell-3) \rfloor. \quad (7.48)$$

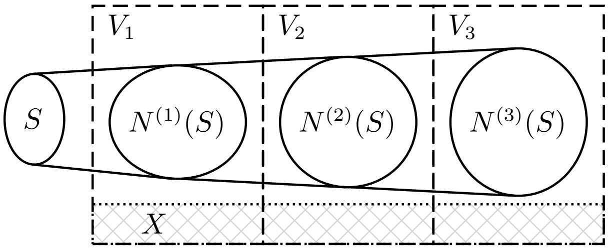

Figure 7.2: The neighbourhoods $N^{(j)}(S) = N^{(j)}(S, X)$ for $j \in [3]$, where S may also intersect with X and the vertex classes, i.e., with $X \cup V_1 \cup V_2 \cup V_3$. Furthermore, $S \cap N^{(j)}(S) \neq \emptyset$ is also possible.

Given $X \subseteq [n]$, we partition $\{1, \ldots, (\ell-3)r\} \setminus X$ as follows: for every $1 \leqslant j \leqslant \ell - 3$ we set

$$V_j = V_j(X) = \{v \in [n] \setminus X \ : \ (j-1)r < v \leqslant jr\}. \tag{7.49}$$

With a given graph in mind, which will later be $G(i)$ or the binomial random graph, for every $S \subseteq [n]$ we define its *neighbourhoods wrt.* X as

$$N^{(0)}(S, X) = S \qquad \text{and} \qquad N^{(j+1)}(S, X) = \Gamma\big(N^{(j)}(S, X)\big) \cap V_{j+1}(X),$$

see also Figure 7.2. Observe that all $N^{(j)}(S, X)$ are disjoint if $S \subseteq X$. Furthermore, $X \subseteq Y$ implies

$$V_j(Y) \subseteq V_j(X) \qquad \text{and} \qquad N^{(j)}(S, Y) \subseteq N^{(j)}(S, X). \tag{7.50}$$

Finally, for the sake of brevity we define $N^{(\leqslant j)}(S, X) = \bigcup_{0 \leqslant j' \leqslant j} N^{(j')}(S, X)$.

7.4.2.2 Configurations

We define the set $\mathcal{C}$ of configurations to be the set of all $\Sigma = (\tilde{v}, U, A, B, R)$ with $\tilde{v} \in [n]$, $U \in \binom{[n] \setminus \{\tilde{v}\}}{u}$, disjoint $A, B \in \binom{U}{k}$, and $R \subseteq [n]$ with $\{\tilde{v}\} \cup U \subseteq R$ and $|R| \leqslant kn^{10\ell\varepsilon}$. Given $\Sigma \in \mathcal{C}$, we then set $T_\Sigma = A \times V_1 \times \cdots \times V_{\ell-3} \times B$, where each $V_j = V_j(R)$ is given by (7.49).

Given $\Sigma \in \mathcal{C}$, distinct $x, y \in [n]$ and $j \in [\ell-1]$, let $C_{x,y,\Sigma}(i, j)$ contain all pairs $bw \in B \times N^{(\ell-3)}(A, R)$ for which there exist disjoint paths $b = w_1 \cdots w_j = x$ and $y = w_{j+1} \cdots w_\ell = w$ in $G(i)$. Note that adding xy and bw completes a copy of C_ℓ containing both xy and bw. Furthermore, observe that $C_{x,y,\Sigma}(i, j)$ and $C_{y,x,\Sigma}(i, j)$ may differ. So, for all $xy \in$

$O(i) \cup C(i)$ we see that the intersection of $C_{xy}(i)$ with $B \times N^{(\ell-3)}(A,R)$ is contained in $\bigcup_{j\in[\ell-1]} \big[C_{x,y,\Sigma}(i,j) \cup C_{y,x,\Sigma}(i,j)\big]$. Finally, note that by monotonicity we have $C_{x,y,\Sigma}(i,j) \subseteq C_{x,y,\Sigma}(i+1,j)$.

7.4.2.3 Random variables

For every $\Sigma \in \mathcal{C}$ we track the sizes of several sets throughout the evolution of the C_ℓ-free process. For brevity, given $(v_0,\dots,v_{\ell-2}) \in T_\Sigma$, we set $f_j = v_{j-1}v_j$ for all $1 \leqslant j \leqslant \ell-2$. For every $0 \leqslant j \leqslant \ell-3$ we introduce sets $T_{\Sigma,j}(i)$, which for $0 \leqslant j < \ell-3$ will satisfy

$$T_{\Sigma,j}(i) \subseteq \big\{(v_0,\dots,v_{\ell-2}) \in T_\Sigma \ : \ \{f_1,\dots,f_j\} \subseteq E(i) \ \wedge \ \{f_{j+1},\dots,f_{\ell-2}\} \subseteq O(i)\big\}, \quad (7.51)$$

and for the special case $j = \ell-3$ we will have

$$T_{\Sigma,\ell-3}(i) \subseteq \big\{(v_0,\dots,v_{\ell-2}) \in T_\Sigma \ : \ \{f_1,\dots,f_{\ell-3}\} \subseteq E(i) \ \wedge \ f_{\ell-2} \in O(i) \cup C(i)\big\}, \quad (7.52)$$

see also Figure 7.3. Note that $f_{\ell-2}$ can be in $O(i)$ or $C(i)$ for $T_{\Sigma,\ell-3}(i)$, but we will see later that the number of tuples with pairs in $C(i)$ is negligible. In the following we define the $T_{\Sigma,j}(i)$ inductively, starting with $T_{\Sigma,j}(0) = \emptyset$ for $j > 0$ and $T_{\Sigma,0}(0) = T_\Sigma$. Now suppose the process chooses $e_{i+1} = xy \in O(i)$ as the next edge in step $i+1$. For $j > 0$ a tuple $(v_0,\dots,v_{\ell-2}) \in T_{\Sigma,j-1}(i)$ is *added* to $T_{\Sigma,j}(i+1)$, i.e., is in $T_{\Sigma,j}(i+1)$, if $f_j = e_{i+1}$, $\{f_{j+1},\dots,f_{\ell-2}\} \cap C_{f_j}(i) = \emptyset$, and in $G(i)$ there is no path $w_0 \cdots w_j = v_j$ with $w_0 \in A$. Furthermore, for $j < \ell-3$ a tuple $(v_0,\dots,v_{\ell-2}) \in T_{\Sigma,j}(i)$ is *removed*, i.e., not in $T_{\Sigma,j}(i+1)$, if $e_{i+1} \in \{f_{j+1},\dots,f_{\ell-2}\}$ or $e_{i+1} \in C_{f_{j+1}}(i) \cup \cdots \cup C_{f_{\ell-2}}(i)$. For the special case $j = \ell-3$, a tuple $(v_0,\dots,v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$ is *removed*, i.e., not in $T_{\Sigma,\ell-3}(i+1)$, or *ignored*, i.e., remains in $T_{\Sigma,\ell-3}(i+1)$, according to the following rules:

Case 1. If $f_{\ell-2} = e_{i+1}$, then the tuple $(v_0,\dots,v_{\ell-2})$ is removed,

Case 2. If $e_{i+1} \in C_{f_{\ell-2}}(i)$, then the tuple $(v_0,\dots,v_{\ell-2})$ is

(R2) removed if there exists $j \in [\ell-1]$ and $x,y \in [n]$ such that $e_{i+1} = xy$, $f_{\ell-2} \in C_{x,y,\Sigma}(i,j)$ and $|C_{x,y,\Sigma}(i,j)| \leqslant p^{-1}n^{-30\ell\varepsilon}$, and

(I2) ignored otherwise.

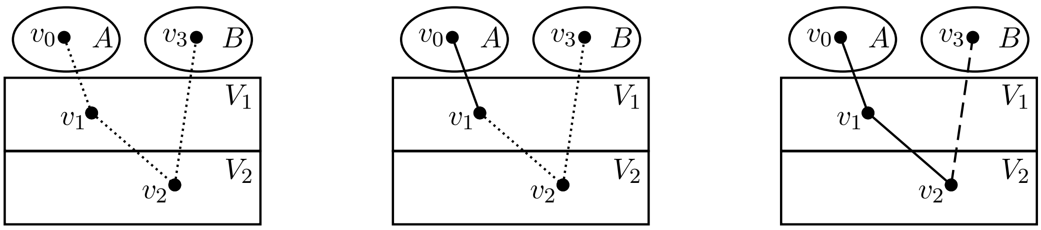

Figure 7.3: Tuples (v_0, v_1, v_2, v_3) in $T_{\Sigma,0}(i)$, $T_{\Sigma,1}(i)$ and $T_{\Sigma,2}(i)$ for $\ell = 5$, where $\Sigma = (\tilde{v}, U, A, B, R)$. Solid lines represent edges, dotted lines open pairs and dashed lines pairs that are open or closed. For the other pairs there is no restriction, i.e., they may be open, closed or an edge.

The above definition clearly satisfies (7.51) and (7.52). Intuitively, the rules for removing tuples from $T_{\Sigma,\ell-3}(i)$ ensure that the one-step changes are 'by definition' not too large. Furthermore, the way in which the tuples are added yields the following *extension property* $\mathcal{U}_T$.

Lemma 7.4.1. *Given $i \geqslant 0$, let $\mathcal{U}_T(i)$ denote the property that for all $\Sigma \in \mathcal{C}$ and $1 \leqslant j \leqslant \ell - 3$, for every $(v_j, \dots, v_{\ell-2}) \in V_j \times \cdots V_{\ell-3} \times B$ there exists at most one $(v_0, \dots, v_{j-1}) \in A \times V_1 \times \cdots \times V_{j-1}$ such that $(v_0, \dots, v_{\ell-2}) \in \bigcup_{i' \leqslant i} T_{\Sigma,j}(i')$. Then $\mathcal{U}_T = \mathcal{U}_T(i)$ holds for every $i \geqslant 0$.* □

The proof proceeds by induction on i and j; we leave the straightforward details to the reader (it is helpful to observe that after $(v_0, \dots, v_{\ell-2}) \in T_{\Sigma,j-1}(i)$ is added to $T_{\Sigma,j}(i+1)$, no further tuples containing v_j can be added due to the $v_0 \cdots v_j$ path). Note that by $\mathcal{U}_T$ every $(v_j, \dots, v_{\ell-2}) \in V_j \times \cdots V_{\ell-3} \times B$ is contained in at most one tuple in $\bigcup_{i' \leqslant i} T_{\Sigma,j}(i')$. This is an important ingredient of our argument, which refines a simpler variant of the extension property we used in [76].

Recall that our goal is to show that there are many open pairs whose addition would complete a C_ℓ-extension for $(\tilde{v}, U)$. Given $\Sigma = (\tilde{v}, U, A, B, R)$, note that for every $(v_0, \dots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$, if $f_{\ell-2} \in O(i)$, then adding $f_{\ell-2}$ to $G(i)$ would complete such a C_ℓ-extension. Now, since $\mathcal{U}_T$ implies that every pair $f_{\ell-2} = xy$ with $x \in V_{\ell-3}$ and $y \in B$ is contained in at most one such tuple in $T_{\Sigma,\ell-3}(i)$, our aim is to obtain a lower bound on the size of

$$Z_{\Sigma,\ell-3}(i) = \big\{(v_0, \dots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i) \ : \ f_{\ell-2} \in O(i)\big\}. \tag{7.53}$$

7.4.2.4 Bad events

The following bad event $\mathcal{B}_i(\Sigma)$ is crucial for our argument: it addresses the two main technical difficulties outlined in Section 7.4.1.2. For all $0 \leqslant i \leqslant m$ and $\Sigma \in \mathcal{C}$ we define $\mathcal{B}_i(\Sigma) = \mathcal{B}_{1,i}(\Sigma) \cup \mathcal{B}_{2,i}(\Sigma)$, where

$$\mathcal{B}_{1,i}(\Sigma) = \text{ in } G(i) \text{ there are more than } k^2(np)^{\ell-4}n^{-9\varepsilon} \text{ pairs } (b,w) \in B \times N^{(\ell-4)}(A,R) \text{ for which there exists a path } b = w_0 \cdots w_{\ell-2} = w, \text{ and}$$

$$\mathcal{B}_{2,i}(\Sigma) = \text{ in } G(i) \text{ we have } |L_\Sigma(i)| \geqslant p^{-1}n^{-1/(2\ell)}, \text{ where } L_\Sigma(i) \text{ contains all } xy \in \binom{[n]}{2} \text{ with } \max_{j\in[\ell-1]}\{|C_{x,y,\Sigma}(i,j)|, |C_{y,x,\Sigma}(i,j)|\} \geqslant p^{-1}n^{-30\ell\varepsilon}.$$

Clearly, $\mathcal{B}_i(\Sigma)$ depends only on the first i steps and is increasing, i.e., $\mathcal{B}_i(\Sigma) \subseteq \mathcal{B}_{i+1}(\Sigma)$ holds.

We now briefly give some intuition for $\mathcal{B}_{1,i}(\Sigma)$ and $\mathcal{B}_{2,i}(\Sigma)$, which are important ingredients for estimating the number of tuples added to $T_{\Sigma,\ell-3}(i+1)$ and removed from $T_{\Sigma,\ell-3}(i)$. First, recall that $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,\ell-4}(i)$ can not be added to $T_{\Sigma,\ell-3}(i+1)$ if $f_{\ell-2} \in C_{f_{\ell-3}}(i)$. For such 'useless' tuples there exists a path $v_{\ell-2} = w_0 \cdots w_{\ell-2} = v_{\ell-4}$ with $(v_{\ell-2}, v_{\ell-4}) \in B \times N^{(\ell-4)}(A,R)$ in $G(i)$, and whenever $\neg\mathcal{B}_{1,i}(\Sigma)$ holds there can not be 'too many' such pairs. As we shall see, from this we can deduce (using the extension property $\mathcal{U}_T$) that the number of 'useless' tuples is small compared to $|T_{\Sigma,\ell-4}(i)|$. Second, recall that not all tuples $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$ are removed if $e_{i+1} \in C_{f_{\ell-2}}(i)$: some are are ignored. Here the key point is that $e_{i+1} \in C_{f_{\ell-2}}(i) \setminus L_\Sigma(i)$ is a sufficient condition for being removed, and, with (7.7) in mind, that $\neg\mathcal{B}_{2,i}(\Sigma)$ essentially implies that $|L_\Sigma(i)|$ is small compared to $|C_{f_{\ell-2}}(i)|$. Intuitively, this will allow us to show that the ignored tuples have negligible impact, i.e., that $|Z_{\Sigma,\ell-3}(i)| \approx |T_{\Sigma,\ell-3}(i)|$.

7.4.3 Proof of Theorem 7.1.1

In this section we prove Theorem 7.1.1 assuming the following two statements. Intuitively, the first lemma ensures that for 'good' configurations Σ the variables $|T_{\Sigma,j}(i)|$ are dynamically concentrated, and the second lemma essentially guarantees that for every $(\tilde{v}, U)$ there exists a good $\Sigma^* = (\tilde{v}, U, A, B, R)$ for which $|T_{\Sigma^*,\ell-3}(i)| \approx |Z_{\Sigma^*,\ell-3}(i)|$. Now we give some intuition for the trajectories our variables follow. Using (7.5), we see that the proportion of pairs which are open or an edge in $G(i)$ roughly equals $q(t)$ or $2tp$, respectively, where $t = i/(n^2p)$. So, using random graphs as a guide, it seems plausible to expect $|T_{\Sigma,j}(i)| \approx c_j(2tp)^j q(t)^{\ell-2-j} k^2 r^{\ell-3}$, where the factor $c_j = 1/j!$ takes into account that we only count tuples created in a certain order. In the following results the functions $q(t)$, $f(t)$ and parameters k, m, p, r, u are defined by (7.2), (7.3), (7.47) and (7.48).

Lemma 7.4.2. *For all $0 \leqslant i^* \leqslant m$ and $\Sigma \in \mathcal{C}$, let $\mathcal{G}_{i^*}(\Sigma)$ denote the event that for every $0 \leqslant i \leqslant i^*$ and all $0 \leqslant j \leqslant \ell - 3$ we have*

$$|T_{\Sigma,j}(i)| = \Big((2t)^j q(t)^{\ell-2-j}/j! \ \pm\ f(t)q(t)^{\ell-3-j}/n^{2\varepsilon}\Big)\, k^2 r^{\ell-3} p^j, \tag{7.54}$$

and let $\mathcal{E}_j$ denote the event that for all $0 \leqslant i \leqslant j$ and $\Sigma \in \mathcal{C}$ the event $\mathcal{B}_{i-1}(\Sigma) \cup \mathcal{G}_i(\Sigma)$ holds. Then $\mathcal{E}_m$ holds whp in the C_ℓ-free process.

Lemma 7.4.3. *Let $\mathcal{R}_j$ denote the event that for all $0 \leqslant i \leqslant j$, for every $(\tilde{v}, U) \in [n] \times \binom{[n]}{u}$ with $\tilde{v} \notin U$ there exists $\Sigma^* = (\tilde{v}, U, A, B, R) \in \mathcal{C}$ such that $\neg\mathcal{B}_{i-1}(\Sigma^*)$ holds and*

$$|T_{\Sigma^*,\ell-3}(i) \setminus Z_{\Sigma^*,\ell-3}(i)| \leqslant k^2 (rp)^{\ell-3} n^{-9\varepsilon}. \tag{7.55}$$

Then $\mathcal{R}_m$ holds whp in the C_ℓ-free process.

The proofs of these lemmas are rather involved and therefore deferred to Sections 7.5 and 7.8. With these results in hand, we are now ready to establish our main result.

Proof of Theorem 7.1.1. For the sake of concreteness, we prove the theorem with $D = \gamma\mu$. Given $\tilde{v} \in [n]$, $U \subseteq [n] \setminus \{\tilde{v}\}$ and $i \leqslant m$, let $\mathcal{X}_{\tilde{v},U,i}$ denote the event that up to step i, there is no C_ℓ-extension for $(\tilde{v}, U)$ in the C_ℓ-free process. By $\mathcal{X}_m$ we denote the event that there exists $(\tilde{v}, U) \in [n] \times \binom{[n]}{u}$ with $\tilde{v} \notin U$ for which $\mathcal{X}_{\tilde{v},U,m}$ holds. Furthermore, for every $i \leqslant m$ we set $\mathcal{A}_i = \mathcal{E}_i \cap \mathcal{R}_i \cap \mathcal{T}_i$, where $\mathcal{T}_i$ is defined as in Theorem 7.2.1 and $\mathcal{E}_i$, $\mathcal{R}_i$ as in Lemmas 7.4.2 and 7.4.3. If $\mathcal{X}_m$ fails, then, as discussed in Section 7.4.1, the C_ℓ-free process has maximum degree at most $u = D(n \log n)^{1/(\ell-1)}$. So, since $\mathcal{A}_m$ holds whp by Theorem 7.2.1 and Lemmas 7.4.2 and 7.4.3, to complete the proof it suffices to show

$$\mathbb{P}[\mathcal{X}_m \cap \mathcal{A}_m] = o(1). \tag{7.56}$$

Suppose that for $m/2 \leqslant i \leqslant m$ the event $\mathcal{A}_i = \mathcal{E}_i \cap \mathcal{R}_i \cap \mathcal{T}_i$ holds. Observe that $\mathcal{E}_i \cap \neg\mathcal{B}_{i-1}(\Sigma^*)$ implies $\mathcal{G}_i(\Sigma)$, which is defined as in Lemma 7.4.2. Using (7.2) we see that $m/2 \leqslant i \leqslant m$ implies $t = i/(n^2p) = \omega(1)$, so for $j = \ell - 3$ the main term in the brackets of (7.54) is $(2t)^{\ell-3}q(t)/(\ell-3)!$ since $f(t)/[n^{2\varepsilon}q(t)] = o(1)$ by (7.4). Thus, whenever $\mathcal{E}_i \cap \mathcal{R}_i$ holds, using (7.54), (7.55) and $q(t) \geqslant n^{-\varepsilon/4}$, it follows that for every $(\tilde{v}, U)$ with $U \in \binom{[n]\setminus\{\tilde{v}\}}{u}$ there exists $\Sigma^* = (\tilde{v}, U, A, B, R) \in \mathcal{C}$ satisfying

$$|T_{\Sigma^*,\ell-3}(i)| \geqslant k^2(2tpr)^{\ell-3}q(t)/(\ell-1)! \quad \text{and} \quad |T_{\Sigma^*,\ell-3}(i)\setminus Z_{\Sigma^*,\ell-3}(i)| \leqslant k^2(2tpr)^{\ell-3}q(t)n^{-7\varepsilon}.$$

Note that $\mathcal{T}_i$ gives $q(t) \geqslant |O(i)|/n^2$ by (7.4) and (7.5). So, combining our findings with $Z_{\Sigma^*,\ell-3}(i) \subseteq T_{\Sigma^*,\ell-3}(i)$, using $k = u/60$, $r \geqslant n/\ell$, (7.47) and $t = i/(n^2p)$ we see that for such Σ^* we crudely have

$$\begin{aligned} |Z_{\Sigma^*,\ell-3}(i)| = |T_{\Sigma^*,\ell-3}(i)| - |T_{\Sigma^*,\ell-3}(i) \setminus Z_{\Sigma^*,\ell-3}(i)| &\geqslant k^2(2tpr)^{\ell-3}q(t)/\ell! \\ \geqslant \delta u^2(tpn)^{\ell-3}q(t) = \delta\frac{u^2 i^{\ell-3}}{n^{\ell-3}}q(t) &\geqslant \delta\frac{u^2 i^{\ell-3}}{n^{\ell-1}}|O(i)|. \end{aligned} \tag{7.57}$$

Recall that $O_{\tilde{v},U}(i) \subseteq O(i)$ denotes the set of open pairs which would complete a C_ℓ-extension for $(\tilde{v}, U)$ if chosen as the next edge e_{i+1}. Let $O_{\Sigma^*}(i)$ be the set of all $xy \in O(i)$ for which there exists $(v_0, \ldots, v_{\ell-2}) \in Z_{\Sigma^*,\ell-3}(i)$ with $f_{\ell-2} = xy$. As already discussed in Section 7.4.2.3, by construction we have $O_{\Sigma^*}(i) \subseteq O_{\tilde{v},U}(i)$, and $\mathcal{U}_T$ implies $|O_{\Sigma^*}(i)| = |Z_{\Sigma^*,\ell-3}(i)|$. Together with (7.57) this establishes

$$|O_{\tilde{v},U}(i)| \geqslant \delta\frac{u^2 i^{\ell-3}}{n^{\ell-1}}|O(i)|. \tag{7.58}$$

Using this estimate, we now prove (7.56). To this end fix $(\tilde{v}, U) \in [n] \times \binom{[n]}{u}$ with $\tilde{v} \notin U$. We see that

$$\begin{aligned} \mathbb{P}[\mathcal{X}_{\tilde{v},U,m} \cap \mathcal{A}_m] &= \mathbb{P}[\mathcal{X}_{\tilde{v},U,m/2} \cap \mathcal{A}_{m/2}] \prod_{m/2\leqslant i\leqslant m-1} \mathbb{P}[\mathcal{X}_{\tilde{v},U,i+1} \cap \mathcal{A}_{i+1} \mid \mathcal{X}_{\tilde{v},U,i} \cap \mathcal{A}_i] \\ &\leqslant \prod_{m/2\leqslant i\leqslant m-1} \mathbb{P}[e_{i+1} \notin O_{\tilde{v},U}(i) \mid \mathcal{X}_{\tilde{v},U,i} \cap \mathcal{A}_i]. \end{aligned} \tag{7.59}$$

Note that $\mathcal{X}_{\tilde{v},U,i} \cap \mathcal{A}_i$ depends only on the first i steps of the process, so given this, the process fails to choose e_{i+1} from $O_{\tilde{v},U}(i)$ with probability $1-|O_{\tilde{v},U}(i)|/|O(i)|$. Now from (7.58) and (7.59) as well as the inequality $1-x \leqslant e^{-x}$ we deduce, with room to spare,

$$\mathbb{P}[\mathcal{X}_{\tilde{v},U,m} \cap \mathcal{A}_m] \leqslant \exp\left\{-\delta \frac{u^2}{n^{\ell-1}} \sum_{m/2 \leqslant i \leqslant m-1} i^{\ell-3}\right\} \leqslant \exp\left\{-\frac{\delta}{2^\ell}\frac{u^2 m^{\ell-2}}{n^{\ell-1}}\right\}. \tag{7.60}$$

Substituting the definitions of m, u, p and $t_{\max}$ into (7.60) we obtain

$$\mathbb{P}[\mathcal{X}_{\tilde{v},U,m} \cap \mathcal{A}_m] \leqslant \exp\left\{-\frac{\delta\gamma}{2^\ell} n^{\ell-2} p^{\ell-1} t_{\max}^{\ell-1} u\right\} = \exp\left\{-\gamma \frac{\delta \mu^{\ell-1}}{2^\ell} u \log n\right\} \leqslant n^{-2u},$$

where the last inequality follows from (7.47), i.e., the definition of γ. Finally, taking the union bound over all choices of $(\tilde{v}, U)$ implies (7.56), which, as explained, completes the proof. □

7.5 Trajectory verification

This section is devoted to the proof of Lemma 7.4.2. Henceforth we work with the 'natural' filtration given by the C_ℓ-free process, where $\mathcal{F}_i$ corresponds to the first i steps, and tacitly assume that n is sufficiently large whenever necessary. For every $0 \leqslant i \leqslant m$ we set $\mathcal{H}_i = \mathcal{J}_i \cap \mathcal{T}_i$, where $\mathcal{J}_i$, $\mathcal{T}_i$ are defined as in Theorem 7.2.1. Clearly, $\mathcal{H}_m$ holds whp. Furthermore $\mathcal{H}_{i+1} \subseteq \mathcal{H}_i$ and $\mathcal{H}_i \in \mathcal{F}_i$, since $\mathcal{H}_i$ is monotone decreasing and depends only on the first i steps. We set $s = n^2 p$ and apply the differential equation method (Lemma 7.3.1) with $\mathcal{V} = \{0, \ldots, \ell-3\}$. Recalling that $\mathcal{B}_i(\Sigma)$ is monotone increasing, we see that $\mathcal{B}_i(\Sigma) = \mathcal{B}_{\leqslant i}(\Sigma)$. For all $\sigma \in \mathcal{C} \times \mathcal{V}$ we define

$$u_\sigma = kn^{15\ell\varepsilon} = \omega(1), \qquad \lambda_\sigma = \tau_\sigma = n^\varepsilon, \qquad \beta_\sigma = 1, \qquad \text{and} \qquad s_\sigma = s_o = n^{2\varepsilon}. \tag{7.61}$$

Formally, for all $\sigma = (\Sigma, j) \in \mathcal{C} \times \mathcal{V}$ we set $X_\sigma(i) = |T_{\Sigma,j}(i)|$ and $Y_\sigma^\pm(i) = |T_{\Sigma,j}^\pm(i)|$, where $T_{\Sigma,j}^+(i) = T_{\Sigma,j}(i+1) \setminus T_{\Sigma,j}(i)$ and $T_{\Sigma,j}^-(i) = T_{\Sigma,j}(i) \setminus T_{\Sigma,j}(i+1)$. But, for the sake of clarity, we will henceforth just use $|T_{\Sigma,j}(i)|$ and $|T_{\Sigma,j}^\pm(i)|$. Now, for every $\sigma = (\Sigma, j) \in \mathcal{C} \times \mathcal{V}$ we set

$x_\sigma(t) = x_j(t)$, $y_\sigma^\pm(t) = x_j^\pm(t)$, $S_\sigma = S_j$, $f_\sigma(t) = f_j(t)$ and $h_\sigma(t) = h_j(t)$, where

$$x_j(t) = 1/j! \cdot (2t)^j q(t)^{\ell-2-j}, \qquad S_j = k^2 r^{\ell-3} p^j, \tag{7.62}$$

$$x_j^+(t) = 2j/j! \cdot (2t)^{j-1} q(t)^{\ell-2-j}, \qquad f_j(t) = f(t) q(t)^{\ell-3-j}, \tag{7.63}$$

$$x_j^-(t) = 2(\ell-2-j)(\ell-1)(2t)^{\ell-2} x_j(t), \qquad h_j(t) = f_j'(t)/2. \tag{7.64}$$

The definition of $x_j^+(t)$ might seem overly complicated, but it conveniently ensures $x_0^+(t) = 0$ and $x_j^+(t) = 2x_{j-1}(t)/q(t)$ for $j > 0$. With the above parametrization we can restate (7.54) as

$$|T_{\Sigma,j}(i)| = (x_j(t) \pm f_j(t)/s_o)\, k^2 r^{\ell-3} p^j. \tag{7.65}$$

The remainder of this section is organized as follows. First, in Section 7.5.1 we verify the trend hypothesis of Lemma 7.3.1, and, next, the boundedness hypothesis in Section 7.5.2. Finally, in Section 7.5.3 we check the remaining conditions of the differential equation method.

7.5.1 Trend hypothesis

In order to establish (7.13), whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds, for every $j \in \mathcal{V}$ we have to prove

$$\mathbb{E}[|T_{\Sigma,j}^\pm(i)| \mid \mathcal{F}_i] = \left(x_j^\pm(t) \pm \frac{h_j(t)}{s_o}\right) \frac{k^2 r^{\ell-3} p^j}{n^2 p}. \tag{7.66}$$

7.5.1.1 Basic estimates

The following inequalities can easily be verified using elementary calculus. Recall that $a \pm b$ denotes the interval $\{a + xb : -1 \leqslant x \leqslant 1\}$, where multiple occurrences of $\pm$ are treated independently (see Section 7.2.1).

Lemma 7.5.1. *Suppose* $0 \leqslant x \leqslant 1/2$. *Then*

$$(1 \pm x)^{-1} \subseteq 1 \pm 2x. \tag{7.67}$$

The next lemma provides estimates for products of a special form.

Lemma 7.5.2. *Suppose $x, y, f_x, f_y, g, h \geqslant 0$ and $g \leqslant 1$. Then $f_x + xg \leqslant h/2$ implies*

$$(1 \pm g)(x \pm f_x) \subseteq x \pm h. \tag{7.68}$$

Furthermore, $xf_y + yf_x + f_xf_y + xyg \leqslant h/2$ implies

$$(1 \pm g)(x \pm f_x)(y \pm f_y) \subseteq xy \pm h. \tag{7.69}$$

Proof. Using $x, y, f_x, f_y \geqslant 0$ we see that

$$(x \pm f_x)(y \pm f_y) \subseteq xy \pm (xf_y + yf_x + f_xf_y). \tag{7.70}$$

Plugging $y = 1$ and $f_y = g$ into (7.70), and using $g \leqslant 1$, $xg \geqslant 0$ as well as $f_x + xg \leqslant h/2$, we obtain

$$(x \pm f_x)(1 \pm g) \subseteq x \pm (xg + f_x + f_xg) \subseteq x \pm 2(f_x + xg) \subseteq x \pm h,$$

which establishes (7.68). Finally, using (7.70) and plugging $x' = xy$ and $f'_x = xf_y + yf_x + f_xf_y$ together with $f'_x + x'g \leqslant h/2$ into (7.68) gives (7.69), which completes the proof. □

7.5.1.2 Triples added in one step.

In this section we verify (7.66) for $T^+_{\Sigma,j}(i)$.

The case $j = 0$. Clearly, adding an edge to $G(i)$ can not create new open tuples in $T_{\Sigma,0}(i)$. Thus we always have $|T^+_{\Sigma,0}(i)| = 0 = x^+_0(t)$, which settles this case.

The case $j > 0$. Recall that $e_{i+1} \in O(i)$ is added to $G(i)$. Let $P_{\Sigma,j-1}(i)$ contain all $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j-1}(i)$ for which there exists a path $w_0 \ldots w_j = v_j$ with $w_0 \in A$ in $G(i)$. Similarly, $D_{\Sigma,j-1}(i) \subseteq T_{\Sigma,j-1}(i)$ contains all tuples with $\{f_{j+1}, \ldots, f_{\ell-2}\} \cap C_{f_j}(i) \neq \emptyset$, where $f_{j'} = v_{j'-1}v_{j'}$. With these definitions in hand, note that $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j-1}(i)$ is added to $T_{\Sigma,j}(i+1)$, i.e., is in $T_{\Sigma,j}(i+1)$, if and only if $f_j = e_{i+1}$ and $(v_0, \ldots, v_{\ell-2}) \notin P_{\Sigma,j-1}(i) \cup D_{\Sigma,j-1}(i)$, see Section 7.4.2.3. Since the C_ℓ-free process chooses e_{i+1} uniformly

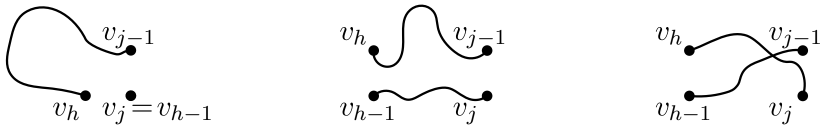

Figure 7.4: The solid lines represent paths such that adding both $f_j = v_{j-1}v_j$ and $f_h = v_{h-1}v_h$ completes a copy of C_ℓ consisting of those paths. In other words, adding f_j closes f_h, i.e., $f_h \in C_{f_j}(i)$.

at random from $O(i)$, whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds we have

$$\mathbb{E}[|T^+_{\Sigma,j}(i)| \mid \mathcal{F}_i] = \sum_{(v_0,\ldots,v_{\ell-2})\in T_{\Sigma,j-1}(i)\setminus[P_{\Sigma,j-1}(i)\cup D_{\Sigma,j-1}(i)]} \frac{1}{|O(i)|}. \tag{7.71}$$

We now bound the size of $P_{\Sigma,j-1}(i)$. Since $\mathcal{H}_i$ implies (7.6), the degree of every vertex is bounded by, say, npn^ε. So, using $|A| = k \leqslant npn^\varepsilon$, $j \leqslant \ell - 3$, $(np)^{\ell-2} = n^{1-1/(\ell-1)}$ and $r \geqslant n/\ell$, in $G(i)$ the number of w_j for which there exists a path $w_0 \ldots w_j$ with $w_0 \in A$ is at most

$$|A| \cdot (npn^\varepsilon)^j \leqslant (npn^\varepsilon)^{\ell-2} \leqslant n^{1+\ell\varepsilon-1/(\ell-1)} \leqslant rn^{-1/(2\ell)}. \tag{7.72}$$

Given w_j, we now bound the number of $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j-1}(i)$ with $w_j = v_j$. Observe that there are at most $k(npn^\varepsilon)^{j-1}$ choices for such $v_1, \ldots, v_{j-1}$, and at most $r^{\ell-j-3}k$ choices for $v_{j+1}, \ldots, v_{\ell-2}$. Putting things together, we deduce that

$$|P_{\Sigma,j-1}(i)| \leqslant rn^{-1/(2\ell)} \cdot k(npn^\varepsilon)^{j-1} \cdot r^{\ell-j-3}k \leqslant k^2 r^{\ell-3} p^{j-1} n^{-1/(3\ell)}. \tag{7.73}$$

Turning to $D_{\Sigma,j-1}(i)$, we first consider the case where $0 < j < \ell - 3$. Suppose that $f_h \in C_{f_j}(i)$. Depending on whether $h = j + 1$ or $h > j + 1$, there exists either a path $v_{j-1} = w_1 \cdots w_{\ell-1} = v_h$ with $j < h < \ell - 2$, or a path $w_1 \cdots w_\kappa = v_{h-1}$ with $w_1 \in \{v_j, v_{j-1}\}$, $1 < \kappa \leqslant \ell - 2$ and $j < h - 1 < \ell - 2$, cf. Figure 7.4. So, in both cases, there exists a path $w_1 \cdots w_\kappa = v_x$ with $w_1 \in \{v_j, v_{j-1}\}$, $1 < \kappa \leqslant \ell - 1$ and $j < x < \ell - 2$. With this observations in hand, we are now ready to estimate the number of tuples $(v_0, \ldots, v_{\ell-2}) \in D_{\Sigma,j-1}(i)$. Recall that by $\mathcal{H}_i$ the degree of every vertex is at most npn^ε. It follows that there are at most $k(npn^\varepsilon)^{j-1}r$ choices for $v_0, \ldots, v_j$, and at most ℓ^2 choices for h and x. Given $v_0, \ldots, v_j$ as well as h and x, there are at most $2\ell(npn^\varepsilon)^{\ell-2} \leqslant rn^{-1/(3\ell)}$ choices for v_x by (7.72). Since we already picked v_x with $j < x < \ell - 2$, for the remaining vertices among $v_{j+1}, \dots, v_{\ell-2}$ we have at most $r^{\ell-j-4}k$ choices. Putting things together, we see that for $0 < j < \ell - 3$ we have

$$|D_{\Sigma,j-1}(i)| \leqslant k(npn^{\varepsilon})^{j-1} \cdot r \cdot \ell^2 \cdot rn^{-1/(3\ell)} \cdot r^{\ell-j-4}k \leqslant k^2 r^{\ell-3} p^{j-1} n^{-1/(4\ell)}. \tag{7.74}$$

Now we bound $|D_{\Sigma,j-1}(i)|$ for the remaining case $j = \ell - 3$. Recall that $f_{\ell-3} = v_{\ell-4}v_{\ell-3}$. If $f_{\ell-2} = v_{\ell-3}v_{\ell-2} \in C_{f_{\ell-3}}(i)$, then, with a similar reasoning as in the previous case, there exists a path $v_{\ell-4} = w_0 \cdots w_{\ell-2} = v_{\ell-2}$, where $v_{\ell-4} \in N^{(\ell-4)}(A, R)$ and $v_{\ell-2} \in B$. Since $\neg\mathcal{B}_i(\Sigma)$ holds, by $\neg\mathcal{B}_{1,i}(\Sigma)$ there are at most $k^2(np)^{\ell-4}n^{-9\varepsilon}$ such pairs $(v_{\ell-2}, v_{\ell-4}) \in B \times N^{(\ell-4)}(A, R)$ in $G(i)$. Recall that by the extension property $\mathcal{U}_T$ (cf. Lemma 7.4.1) every triple $(v_{\ell-4}, v_{\ell-3}, v_{\ell-2})$ is contained in at most one tuple in $T_{\Sigma,\ell-4}(i)$. So, since there are at most $k^2(np)^{\ell-4}n^{-9\varepsilon}$ choices for $v_{\ell-4}, v_{\ell-2}$, and at most r choices for $v_{\ell-3} \in V_{\ell-3}$, using $\mathcal{U}_T$ we deduce that for $j = \ell - 3$ we have

$$|D_{\Sigma,j-1}(i)| \leqslant k^2(np)^{\ell-4}n^{-9\varepsilon} \cdot r \leqslant k^2 r^{\ell-3} p^{\ell-4} n^{-8\varepsilon} = k^2 r^{\ell-3} p^{j-1} n^{-8\varepsilon}. \tag{7.75}$$

After these preparations, we now estimate (7.71) whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds. Observe that $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma)$ implies $\mathcal{G}_i(\Sigma)$, and so $|T_{\Sigma,j-1}(i)|$ satisfies (7.65). Furthermore, since $\mathcal{H}_i$ holds, this implies that $|O(i)|$ satisfies (7.5). In addition, note that $s_e = n^{1/(2\ell)-\varepsilon}$ and (7.4) imply $f(t)/s_e = o(1)$ and $f_{j-1}(t) \geqslant 1$. Substituting the former estimates and (7.73)–(7.75) into (7.71), using $n^{1/(3\ell)} \geqslant n^{8\varepsilon} = \omega(s_o)$, (7.67), $x_j^+(t) = 2x_{j-1}(t)/q(t)$ and $f_j(t) = f_{j-1}(t)/q(t)$, we deduce that

$$\begin{aligned}
\mathbb{E}[|T^+_{\Sigma,j}(i)| \mid \mathcal{F}_i] &= \frac{(x_{j-1}(t) \pm f_{j-1}(t)/s_o)k^2 r^{\ell-3} p^{j-1} \pm 2k^2 r^{\ell-3} p^{j-1} n^{-8\varepsilon}}{(1 \pm 3f(t)/s_e)q(t)n^2/2} \\
&\subseteq \frac{(x_{j-1}(t) \pm 2f_{j-1}(t)/s_o)k^2 r^{\ell-3} p^{j-1}}{(1 \pm 3f(t)/s_e)q(t)n^2/2} \\
&\subseteq (1 \pm 6f(t)/s_e) \cdot (x_j^+(t) \pm 4f_j(t)/s_o) \cdot k^2 r^{\ell-3} p^j/(n^2 p).
\end{aligned}$$

Therefore the desired bound, i.e., (7.66) for $T^+_{\Sigma,j}(i)$, follows if

$$(1 \pm 6f(t)/s_e) \cdot (x^+_j(t) \pm 4f_j(t)/s_o) \subseteq x^+_j(t) \pm h_j(t)/s_o. \tag{7.76}$$

Now, using $f(t) = o(s_e)$ and Lemma 7.5.2, by writing down the assumptions of (7.68) and multiplying both sides with $2s_o$, observe that (7.76) follows from

$$8f_j(t) + 12x^+_j(t)f(t)s_o/s_e \leqslant h_j(t).$$

Using (7.4) and (7.63) we see that the second term on the left hand side is $o(1)$. So, it suffices if

$$8f_j(t) + 1 \leqslant h_j(t),$$

which is easily seen to be true, since $h_j(t) \geqslant W/4 \cdot (f_j(t) + 1)$ and $W \geqslant 50$ by (7.1), (7.4) and (7.64).

7.5.1.3 Triples removed in one step

Next, we prove (7.66) for $T^-_{\Sigma,j}(i)$. Since the rules for removing tuples from $T_{\Sigma,j}(i)$ are different for $j < \ell - 3$ and $j = \ell - 3$, we use a case distinction.

The case $j < \ell - 3$. Recall that a tuple $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j}(i)$ is removed, i.e., not in $T_{\Sigma,j}(i+1)$, if $e_{i+1} \in \{f_{j+1}, \ldots, f_{\ell-2}\}$ or $e_{i+1} \in C_{f_{j+1}}(i) \cup \cdots \cup C_{f_{\ell-2}}(i)$. Since the edge e_{i+1} is chosen uniformly at random from $O(i)$, whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds, using $|\{f_{j+1}, \ldots, f_{\ell-2}\}| \leqslant \ell$ we have

$$\mathbb{E}[|T^-_{\Sigma,j}(i)| \mid \mathcal{F}_i] = \sum_{(v_0,\ldots,v_{\ell-2}) \in T_{\Sigma,j}(i)} \frac{|C_{f_{j+1}}(i) \cup \cdots \cup C_{f_{\ell-2}}(i)| \pm \ell}{|O(i)|}. \tag{7.77}$$

Note that $\mathcal{H}_i$ implies that the inequalities (7.5), (7.7) and (7.8) hold. In particular, using

$n^{1/\ell} = \omega(s_e)$, $n^{-1/\ell}p^{-1} = \omega(1)$ and $f(t) \geqslant 1$, this yields

$$|C_{f_{j+1}}(i) \cup \cdots \cup C_{f_{\ell-2}}(i)| \pm \ell \subseteq (\ell - j - 2)[(\ell-1)(2t)^{\ell-2}q(t) \pm 7\ell f(t)/s_e]p^{-1} \pm \ell^2 n^{-1/\ell}p^{-1} \pm \ell$$
$$\subseteq (\ell - j - 2)[(\ell-1)(2t)^{\ell-2}q(t) \pm 9\ell f(t)/s_e]p^{-1}. \tag{7.78}$$

Since $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma)$ implies $\mathcal{G}_i(\Sigma)$, it follows that $|T_{\Sigma,j}(i)|$ satisfies (7.65). In addition, as in Section 7.5.1.2, $f(t)/s_e = o(1)$ holds and $|O(i)|$ satisfies (7.5) by $\mathcal{H}_i$. Substituting the former estimates into (7.77), and using (7.67) as well as $x_j^-(t)/x_j(t) = 2(\ell-j-2)(\ell-1)(2t)^{\ell-2}$, we obtain

$$\mathbb{E}[|T^-_{\Sigma,j}(i)| \mid \mathcal{F}_i] = \frac{(x_j(t) \pm f_j(t)/s_o)k^2r^{\ell-3}p^j \cdot (\ell-j-2)[(\ell-1)(2t)^{\ell-2}q(t) \pm 9\ell f(t)/s_e]p^{-1}}{(1 \pm 3f(t)/s_e)q(t)n^2/2}$$
$$\subseteq (1 \pm 6f(t)/s_e) \cdot (x_j(t) \pm f_j(t)/s_o) \cdot [x_j^-(t)/x_j(t) \pm 20\ell^2 f(t)/(q(t)s_e)] \cdot k^2r^{\ell-3}p^j/(n^2p).$$

Therefore the desired bound, i.e., (7.66) for $T^-_{\Sigma,j}(i)$, follows if

$$(1 \pm 6f(t)/s_e) \cdot (x_j(t) \pm f_j(t)/s_o) \cdot [x_j^-(t)/x_j(t) \pm 20\ell^2 f(t)/(q(t)s_e)] \subseteq x_j^-(t) \pm h_j(t)/s_o. \tag{7.79}$$

We now show (7.79) using Lemma 7.5.2. Similar as for the added tuples, by writing down the assumptions of (7.69), multiplying with $2s_o$ and then noticing that all terms containing s_e contribute $o(1)$, we see that it suffices if

$$(\ell - 2 - j)(\ell-1)2^\ell t^{\ell-2} f_j(t) + 1 \leqslant h_j(t),$$

which is easily seen to be true, since $h_j(t) \geqslant W/2 \cdot (t^{\ell-2}f_j(t) + 1)$ and $W/2 \geqslant \ell^2 2^\ell$ by (7.1) and (7.64).

The case $j = \ell - 3$. Recall that a tuple $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$ is removed, i.e., not in $T_{\Sigma,\ell-3}(i+1)$, if $e_{i+1} = f_{\ell-2}$, or in addition to $e_{i+1} \in C_{f_{\ell-2}}(i)$ it is not ignored. A moment's thought reveals that for every $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$ with $e_{i+1} \in C_{f_{\ell-2}}(i)$, if $e_{i+1} \notin L_\Sigma(i)$ then (R2) holds, where $L_\Sigma(i)$ is as in the definition of $\mathcal{B}_{2,i}(\Sigma)$. In other words, for every $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$ we see that $e_{i+1} \in C_{f_{\ell-2}}(i) \setminus L_\Sigma(i)$ is a sufficient condition for being removed. Clearly, a necessary condition for being removed is $e_{i+1} \in \{f_{\ell-2}\} \cup C_{f_{\ell-2}}(i)$. Combining our previous findings and using that e_{i+1} is chosen uniformly at random from $O(i)$, whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds we deduce that

$$\mathbb{E}[|T^-_{\Sigma,\ell-3}(i)| \mid \mathcal{F}_i] = \sum_{(v_0,\dots,v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)} \frac{|C_{f_{\ell-2}}(i)| \pm |L_\Sigma(i)| \pm 1}{|O(i)|}.$$

Recall that $\mathcal{H}_i$ implies the inequalities (7.7) and (7.8). Furthermore, since $\neg\mathcal{B}_{2,i}(\Sigma)$ holds, we have $|L_\Sigma(i)| \leqslant p^{-1}n^{-1/(2\ell)}$. So, similar as in the previous case, using $n^{1/(2\ell)} = \omega(s_e)$, $n^{-1/(2\ell)}p^{-1} = \omega(1)$ and $f(t) \geqslant 1$, we obtain

$$\begin{aligned} |C_{f_{\ell-2}}(i)| \pm |L_\Sigma(i)| \pm 1 &\subseteq [(\ell-1)(2t)^{\ell-2}q(t) \pm 7\ell f(t)/s_e]p^{-1} \pm p^{-1}n^{-1/(2\ell)} \pm 1 \\ &\subseteq [(\ell-1)(2t)^{\ell-2}q(t) \pm 9\ell f(t)/s_e]p^{-1}, \end{aligned}$$

where the final estimate equals that of (7.78) for $j = \ell - 3$. It is not difficult to see that the remaining calculations of the case $j < \ell - 3$ carry over word by word, which yields (7.66) for $T^-_{\Sigma,\ell-3}(i)$. To summarize, we have verified the trend hypothesis (7.66).

7.5.2 Boundedness hypothesis

Observe that in order to verify the boundedness hypothesis (7.15), using (7.61) it suffices to show that whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds, for every $j \in \mathcal{V}$ we have

$$|T^\pm_{\Sigma,j}(i)| \leqslant kr^{\ell-3}p^j n^{-20\ell\varepsilon}. \tag{7.80}$$

7.5.2.1 Triples added in one step.

In this section we verify (7.80) for $T^+_{\Sigma,j}(i)$. Recall that $e_{i+1} \in O(i)$ is added to $G(i)$. By construction we always have $|T^+_{\Sigma,0}(i)| = 0$, and thus we henceforth consider the case $j > 0$. Note that a necessary condition for $(v_0, \dots, v_{\ell-2}) \in T_{\Sigma,j-1}(i)$ being added to $T_{\Sigma,j}(i+1)$ is $f_j = e_{i+1}$. Observe that there are at most $kr^{\ell-3-j}$ choices for $(v_{j+1}, \dots, v_{\ell-2}) \in V_{j+1} \times \cdots \times V_{\ell-3} \times B$. So, using the extension property $\mathcal{U}_T$ (cf. Lemma 7.4.1), we deduce that for each e_{i+1} there are at most $kr^{\ell-3-j}$ tuples in $T_{\Sigma,j-1}(i)$ with $f_j = e_{i+1}$. Together with (7.2), (7.4), (7.48) and $j \geqslant 1$ this implies

$$|T^{+}_{\Sigma,j}(i)| \leqslant kr^{\ell-3-j} = kr^{\ell-3}p^j \cdot (rp)^{-j} = o(kr^{\ell-3}p^j n^{-20\ell\varepsilon}), \tag{7.81}$$

as desired.

7.5.2.2 Triples removed in one step

Next we use case distinction to establish (7.80) for $T^{-}_{\Sigma,j}(i)$.

The case $j < \ell - 3$. We claim that whenever $\mathcal{E}_i \cap \neg\mathcal{B}_i(\Sigma) \cap \mathcal{H}_i$ holds, for all $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j}(i)$ and every $xy \in \{f_{j+1}, \ldots, f_{\ell-2}\}$, the number of tuples in $T_{\Sigma,j}(i)$ containing xy is bounded by

$$kr^{\ell-4}p^j n^{\ell\varepsilon}. \tag{7.82}$$

First suppose that $xy = f_{j+1}$. For $(v_{j+2}, \ldots, v_{\ell-2}) \in V_{j+2} \times \cdots \times V_{\ell-3} \times B$ there are at most $kr^{\ell-4-j} \leqslant kr^{\ell-4}p^j$ choices, and so (7.82) follows using the extension property $\mathcal{U}_T$ (cf. Lemma 7.4.1).

Next we consider the case $xy = f_{\ell-2}$. As usual, whenever $\mathcal{H}_i$ holds, the degree of every vertex is bounded by, say, npn^{ε}. Since for every $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j}(i)$ the vertices $v_0, \ldots, v_j$ form a path starting in A, we deduce that there are at most $k(npn^{\varepsilon})^j$ choices for such $v_0, \ldots, v_j$. Furthermore, there are most $r^{\ell-4-j}$ choices for $(v_{j+1}, \ldots, v_{\ell-4}) \in V_{j+1} \times \cdots \times V_{\ell-4}$. Therefore the number of tuples in $T_{\Sigma,j}(i)$ with $xy = f_{\ell-2}$ is bounded by $k(npn^{\varepsilon})^j \cdot r^{\ell-4-j} \leqslant kr^{\ell-4}p^j n^{\ell\varepsilon}$, as claimed by (7.82).

Finally we consider the case where $xy = f_h$ with $j + 1 < h < \ell - 2$. With a similar reasoning as in the previous case, there are at most $k(npn^{\varepsilon})^j$ choices for $v_0, \ldots, v_j$, at most r^{h-j-2} choices for $v_{j+1}, \ldots, v_{h-2}$ and at most $kr^{\ell-h-3}$ choices for $v_{h+1}, \ldots, v_{\ell-2}$. To summarize, there are at most

$$k(npn^{\varepsilon})^j \cdot r^{h-j-2} \cdot kr^{\ell-h-3} \leqslant k^2 r^{\ell-5} p^j n^{\ell\varepsilon} \leqslant kr^{\ell-4}p^j$$

tuples in $T_{\Sigma,j}(i)$ with $xy = f_h$, which establishes (7.82), with room to spare.

With the above estimate in hand, we are now ready to bound $|T^{-}_{\Sigma,j}(i)|$. Recall that

$(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,j}(i)$ is removed, i.e., not in $T_{\Sigma,j}(i+1)$, if $e_{i+1} \in \{f_{j+1}, \ldots, f_{\ell-2}\}$ or $e_{i+1} \in C_{f_{j+1}}(i) \cup \cdots \cup C_{f_{\ell-2}}(i)$, which is equivalent to $\{f_{j+1}, \ldots, f_{\ell-2}\} \cap C_{e_{i+1}}(i) \neq \emptyset$. In other words, such a tuple is removed if for some $j+1 \leqslant h \leqslant \ell-2$ we have $f_h = e_{i+1}$ or $f_h \in C_{e_{i+1}}(i)$. Recall that whenever $\mathcal{H}_i$ holds, by (7.7) we have, say, $|C_{e_{i+1}}(i)| \leqslant p^{-1}n^{\varepsilon}$. So, using that (7.82) gives an upper bound for the number of tuples in $T_{\Sigma,j}(i)$ which contain f_h, we deduce that

$$|T^{-}_{\Sigma,j}(i)| \leqslant (\ell + |C_{e_{i+1}}(i)|) \cdot kr^{\ell-4}p^j n^{\ell\varepsilon} \leqslant kr^{\ell-4}p^{j-1}n^{2\ell\varepsilon} \leqslant kr^{\ell-3}p^j \cdot n^{2\ell\varepsilon}/(rp),$$

which, with a similar reasoning as in (7.81), establishes (7.80) for $T^{-}_{\Sigma,j}(i)$ with $j < \ell-3$.

The case $j = \ell - 3$. Recall that a tuple $(v_0, \ldots, v_{\ell-2}) \in T_{\Sigma,\ell-3}(i)$ is removed, i.e., not in $T_{\Sigma,\ell-3}(i+1)$, according to different rules. In the following we bound the total number of tuples removed in one step by each rule, which were called cases 1 and 2 in Section 7.4.2.3. In case 1 we have $f_{\ell-2} = e_{i+1}$ and so, given e_{i+1}, using $\mathcal{U}_T$ we deduce that at most one tuple is removed under case 1.

Turning to case 2, given $e_{i+1} = xy$, note that a necessary condition for being removed by (R2) is that for some $j \in [\ell-1]$ we have $f_{\ell-2} \in C_{x,y,\Sigma}(i,j)$ or $f_{\ell-2} \in C_{y,x,\Sigma}(i,j)$. Recall that by $\mathcal{U}_T$ every such pair $f_{\ell-2}$ is contained in at most one tuple in $T_{\Sigma,\ell-3}(i)$. So, since a tuple is only removed if the corresponding $C_{x,y,\Sigma}(i,j)$ or $C_{y,x,\Sigma}(i,j)$ has size at most $p^{-1}n^{-30\ell\varepsilon}$, we deduce that at most $2\ell \cdot p^{-1}n^{-30\ell\varepsilon}$ tuples are removed in one step by (R2).

Putting it all together, using $p^{-1} = (np)^{\ell-2}$ and $np \leqslant k$, for $j = \ell-3$ we obtain

$$|T^{-}_{\Sigma,\ell-3}(i)| \leqslant 1 + 2\ell p^{-1}n^{-30\ell\varepsilon} \leqslant (np)^{\ell-2}n^{-25\ell\varepsilon} \leqslant k(np)^{\ell-3}n^{-25\ell\varepsilon},$$

which readily establishes the boundedness hypothesis (7.80).

7.5.3 Finishing the trajectory verification

In this section we verify the remaining conditions of the differential equation method (Lemma 7.3.1).

Initial conditions. Using (7.62), for $j > 0$ we clearly have $|T_{\Sigma,j}(0)| = 0 = x_j(0)$, which settles these cases. For the remaining case $j=0$ we crudely have

$$|T_{\Sigma,0}(0)| = |T_\Sigma| = k^2(r \pm kn^{10\ell\varepsilon})^{\ell-3} = (1 \pm kn^{10\ell\varepsilon}/r)^{\ell-3}k^2r^{\ell-3} \subseteq (1 \pm o(1)/s_o)k^2r^{\ell-3},$$

which together with $x_0(0)=1$, $S_0 = k^2r^{\ell-3}$ and $\beta_\sigma = 1$ establishes (7.16).

Bounded number of configurations and variables. Using $k = u/60$ and (7.61) we obtain

$$|\mathcal{C}| \leqslant n \cdot \binom{n}{u} \cdot 3^u \cdot \sum_{r \leqslant kn^{10\ell\varepsilon}} \binom{n}{r} \leqslant n^{2u+kn^{10\ell\varepsilon}} < e^{kn^{15\ell\varepsilon}} = e^{u_\sigma},$$

which together with $|\mathcal{V}| \leqslant \ell$ clearly establishes (7.17).

Additional technical assumptions and the function $f_\sigma(t)$. Using $s = n^2p$ as well as (7.2), (7.47) and (7.61), straightforward calculations show that (7.18) holds, with room to spare; we leave the details to the reader. Recall that by (7.2) we have $t_{\max} = m/s = \Theta((\log n)^{1/(\ell-1)})$. Furthermore, using (7.62)–(7.64), elementary calculus yields $x_j^\pm(t) = O(t_{\max}^{\ell+j-2})$ and $|x_j''(t)| = O(t_{\max}^{2\ell+j-4})$ for $t \leqslant t_{\max}$. Thus, since for all $\sigma = (\Sigma, j) \in \mathcal{C} \times \mathcal{V}$ we have $x_\sigma(t) = x_j(t)$ and $y_\sigma^\pm(t) = x_j^\pm(t)$, it follows that

$$\sup_{0 \leqslant t \leqslant m/s} y_\sigma^\pm(t) = O(\log^2 n) \leqslant n^\varepsilon = \lambda_\sigma \qquad \text{and} \qquad \int_0^{m/s} |x_\sigma''(t)|\ dt = O(\log n \cdot \log^3 n) \leqslant \lambda_\sigma.$$

Recall that for all $\sigma \in \mathcal{C} \times \mathcal{V}$ we have $h_\sigma(t) = f_\sigma'(t)/2$ and $f_\sigma(t) = f(t)q(t)^\iota$, where $\iota \in \{0, \ldots, \ell-3\}$. Hence, using $f_\sigma(0) = 1 = \beta_\sigma$, we see that

$$f_\sigma(t) = 2\int_0^t h_\sigma(\tau)\ d\tau + f_\sigma(0) = 2\int_0^t h_\sigma(\tau)\ d\tau + \beta_\sigma.$$

Note that $h_\sigma(0) = O(1) \leqslant n^{3\varepsilon} = s_\sigma\lambda_\sigma$ and $h_\sigma'(t) \geqslant 0$. Pick $t^* = t^*(\ell) \geqslant 1$ large enough such that for all $t \geqslant t^*$ we have $t^{2\ell} \leqslant f(t)$. Observe that $h_\sigma'(t)$ is bounded by some constant for $t \leqslant t^*$, and note that for larger t we have, say, $h_\sigma'(t) \leqslant W^3 f(t)^2$. Putting things together, using (7.2) and (7.4), i.e., $m/s = O(\log n)$ and $f(t) \leqslant n^\varepsilon$, we readily obtain

$$\int_0^{m/s} |h_\sigma'(t)|\ dt \leqslant \int_0^{t^*} h_\sigma'(t)\ dt + \int_{t^*}^{m/s} W^3 f(t)^2\ dt \leqslant O(1) + O(\log n \cdot n^{2\varepsilon}) \leqslant n^{3\varepsilon} = s_\sigma\lambda_\sigma.$$

To summarize, we showed that (7.14) as well as the additional technical assumptions (7.18)–(7.20) hold, and this completes the proof of Lemma 7.4.2. □

7.6 A 'transfer theorem' for the H-free process

In the H-free process there is a complicated dependency among the edges, and thus standard concentration inequalities are not directly applicable. In this section we show how to overcome this problem for decreasing properties by establishing a 'transfer theorem'. Roughly speaking, this allows us to 'transfer' results for decreasing properties from the binomial random graph model to the H-free process, at the cost of only slightly increasing the 'expected' edge density. In our argument this will be a crucial tool for establishing Lemma 7.4.3.

7.6.1 Relating the H-free process with the uniform random graph

We start by relating the H-free process with the more familiar uniform random graph. In the H-free process the set of open pairs $O(i)$ is defined in the obvious way: it contains all pairs $xy \in \binom{[n]}{2} \setminus E(i)$ for which $G(i) \cup \{xy\}$ remains H-free. The following estimate is not best possible, but it suffices for our purposes and keeps the formulas simple.

Lemma 7.6.1. *Suppose $\mathcal{Q}$ is a decreasing graph property and that $\lambda = \lambda(n) \geqslant 2$ is a parameter. Then for every $1 \leqslant i \leqslant \binom{n}{2}/\lambda$, setting $M = i\lambda$, we have*

$$\mathbb{P}[G(i) \notin \mathcal{Q} \text{ and } |O(i)| \geqslant n^2/\lambda] \leqslant \mathbb{P}[G_{n,M} \notin \mathcal{Q}] + e^{-i/4}, \tag{7.83}$$

where $G(i)$ denotes the graph produced by the H-free process after the first i steps.

Proof. We sequentially generate the edges $e_1, e_2, \ldots$, where each edge e_{j+1} is chosen uniformly at random from $E(K_n) \setminus \{e_1, e_2, \ldots, e_j\}$. On the one hand, the edge-set $\{e_1, e_2, \ldots, e_M\}$ clearly gives $G_{n,M}$. On the other hand, we obtain the graph produced by the H-free process by sequentially traversing the e_j and only adding those edges which do not complete a copy of H. First, for every $1 \leqslant j \leqslant M$ we define the indicator variable X_j for the event that e_j

is added to the graph of the H-free process, and, furthermore, define the random variable

$$X^j = \sum_{1 \leqslant j' \leqslant j} X_{j'},$$

which counts the number of edges in the graph produced by the H-free process after traversing $e_1, \ldots, e_j$. Next, for every $1 \leqslant j \leqslant M$ we define

$$Y_j = \begin{cases} 1, & \text{if } |O(X^{j-1})| < n^2/\lambda, \\ X_j, & \text{otherwise,} \end{cases} \qquad \text{and} \qquad Y^j = \sum_{1 \leqslant j' \leqslant j} Y_{j'}.$$

If $|O(X^{j-1})| \geqslant n^2/\lambda$ holds, we have $Y_j = X_j$ by construction. In this case the next edge is added to the graph of the H-free process with probability at least $|O(X^{j-1})|/\binom{n}{2} \geqslant 2/\lambda$. Otherwise $Y_j = 1$ holds, and so we conclude that $\mathbb{P}[Y_j = 1 \mid Y_1, \ldots, Y_{j-1}] \geqslant 2/\lambda$, which implies that Y^M stochastically dominates a binomial random variable with M trials and success probability $2/\lambda$. With this in mind, standard Chernoff bounds, see e.g. (7.9) of Lemma 7.2.3, give

$$\mathbb{P}[Y^M \leqslant 2i - t] \leqslant e^{-t^2/(4i)}. \tag{7.84}$$

In the remainder we prove (7.83). To this end first observe that

$$\mathbb{P}[G(i) \notin \mathcal{Q} \text{ and } |O(i)| \geqslant n^2/\lambda] \leqslant \mathbb{P}[G(i) \notin \mathcal{Q} \text{ and } X^M \geqslant i] + \mathbb{P}[|O(i)| \geqslant n^2/\lambda \text{ and } X^M < i]. \tag{7.85}$$

Note that by construction $X^M \geqslant i$ implies $G(i) \subseteq G_{n,M}$, and, since $\mathcal{Q}$ is a decreasing graph property, in this case $G(i) \notin \mathcal{Q}$ implies $G_{n,M} \notin \mathcal{Q}$. It follows that

$$\mathbb{P}[G(i) \notin \mathcal{Q} \text{ and } X^M \geqslant i] \leqslant \mathbb{P}[G_{n,M} \notin \mathcal{Q}].$$

Furthermore, since $O(i)$ is decreasing, if both $|O(i)| \geqslant n^2/\lambda$ and $X^M < i$ hold, then this implies $Y^M = X^M < i$. So, by (7.84) we have

$$\mathbb{P}[|O(i)| \geqslant n^2/\lambda \text{ and } X^M < i] \leqslant \mathbb{P}[Y^M < i] \leqslant e^{-i/4}.$$

Substituting these bounds into (7.85) gives (7.83), completing the proof. □

If we relax the additive error in Lemma 7.6.1 to $o(1)$, then for $|O(i)| \geqslant \binom{n}{2}/\lambda$ a slight modification of the above proof works with $M = i\lambda + \omega(1)\lambda\sqrt{i}$; we leave these details to the interested reader.

7.6.2 A 'transfer theorem' for decreasing properties

Using Theorem 7.2.1 and (7.4), we see that $|O(m)| \geqslant n^{2-\varepsilon/2}$ holds whp in the C_ℓ-free process. So, setting $\lambda = \lambda(n) = n^{\varepsilon/2}$ and using the 'asymptotic equivalence' of the uniform and the binomial random graph for monotone graph properties (see e.g. Section 1.4 of [43]), Lemma 7.6.1 readily gives the next theorem. A similar idea is used in [79] for $H = K_4$. Observe that the edge-density of $G(m)$ is roughly $2pt_{\max} = \Theta(p(\log n)^{1/(\ell-1)})$ in the C_ℓ-free process. Intuitively, the following theorem thus states that for decreasing properties, $G(m)$ is 'comparable' with the binomial random graph with only slightly larger edge density pn^ε.

Theorem 7.6.2 ('Transfer Theorem')**.** *Define $m = m(n)$ and $p = p(n)$ as in (7.2). Suppose that ε is chosen as in (7.1) and that $\mathcal{Q}$ is a decreasing graph property. Then for the C_ℓ-free process we have*

$$\mathbb{P}[G(m) \notin \mathcal{Q}] \leqslant \mathbb{P}[G_{n,pn^\varepsilon} \notin \mathcal{Q}] + o(1).$$

In fact, this result also holds for the H-free process, where H is strictly 2-balanced, if m, p and ε are chosen as in Sections 1.2 and 1.3 of [17], since then $|O(m)| \geqslant n^{2-\varepsilon/2}$, with room to spare. We believe that the above 'transfer theorem' will significantly aid in the future analysis of the H-free process, since for decreasing properties it often allows us to work with the *much* easier binomial random graph model, which has been extensively studied and for which e.g. sophisticated concentration inequalities are available.

7.7 Properties of random graphs

In this section we introduce several decreasing graph properties, which are key ingredients in our proof of Lemma 7.4.3. Using the 'transfer theorem' of Section 7.6, it suffices to prove that they hold whp for the binomial random graph $G_{n,p'}$ with $p' = pn^\varepsilon$, where p is defined as in (7.2) and ε is chosen as in (7.1). We remark that essentially all results in this section are not best possible, but suffice for our purposes. For example, in an attempt to keep the formulas simple, we have not optimized the multiplicative n^{ε} factors involved (their contribution in our later arguments will be negligible).

7.7.1 Basic properties

Lemma 7.7.1. *Let $\mathcal{N}$ denote the event that for all pairs of distinct vertices $x, y \in [n]$ we have $|\Gamma(x) \cap \Gamma(y)| \leqslant 9$. Then $\mathcal{N}$ holds whp in $G_{n,p'}$.*

Proof. Using $\ell \geqslant 4$, (7.1) and (7.2), i.e., $p = n^{-1+1/(\ell-1)} \leqslant n^{-2/3}$ and $\varepsilon \leqslant 1/20$, we deduce that

$$\mathbb{P}[\neg\mathcal{N}] \leqslant \binom{n}{2}\binom{n-2}{10}(pn^{\varepsilon})^{20} \leqslant n^2(np^2n^{2\varepsilon})^{10} \leqslant n^2(n^{-1/3+2\varepsilon})^{10} = o(1),$$

as claimed. □

The following result states that every set of size at most u contains a large independent subset. A similar argument was used by Bollobás and Riordan in [25].

Lemma 7.7.2. *Let $\mathcal{I}$ denote the event that for every $U \subseteq [n]$ with $|U| \leqslant u$ there exists an independent set $S \subseteq U$ with $|S| \geqslant |U|/6$. Then $\mathcal{I}$ holds whp in $G_{n,p'}$.*

Proof. Let $\mathcal{E}$ denote the event that every $U \subseteq [n]$ with $|U| \leqslant u$ spans less than $3|U|$ edges. We have

$$\mathbb{P}[\neg\mathcal{E}] \leqslant \sum_{1\leqslant x\leqslant u}\binom{n}{x}\binom{\binom{x}{2}}{3x}(pn^{\varepsilon})^{3x} \leqslant \sum_{1\leqslant x\leqslant u}\left(\frac{ne}{x}\right)^{x}\left(\frac{xe}{6}\right)^{3x}(pn^{\varepsilon})^{3x} \leqslant \sum_{x\geqslant 1}\left(nu^2p^3n^{3\varepsilon}\right)^{x}.$$

Using $\ell \geqslant 4$, (7.1), (7.2) and (7.47), i.e., $u \leqslant npn^{\varepsilon}$, $p = n^{-1+1/(\ell-1)} \leqslant n^{-2/3}$ and $\varepsilon \leqslant 1/60$, we see that

$$nu^2p^3n^{3\varepsilon} \leqslant n^3p^5n^{5\varepsilon} \leqslant n^{-1/3+5\varepsilon} \leqslant n^{-1/4},$$

which implies $\mathbb{P}[\neg\mathcal{E}] = o(1)$. Suppose that $\mathcal{E}$ holds. Then every set of at most u vertices induces a graph with minimum degree less than six. Given $U \subseteq [n]$ with $|U| \leqslant u$, we set $W = U$. Now, by iteratively selecting a vertex $v \in W$ with at most five neighbours in $G[W]$

and removing $\{v\} \cup \Gamma(v)$ from W, we obtain an independent set with at least $|U|/6$ vertices, and the proof is complete. $\square$

7.7.2 Bounding the numbers of certain paths

The results in this section give estimates for the numbers of certain paths. Their statements will contain certain exceptions, and, as we shall see, many of these complications are in fact necessary.

7.7.2.1 Preliminaries: the size of certain neighbourhoods

The following crude upper bound on the degree of every vertex readily follows from standard Chernoff bounds (Lemma 7.2.3) – we omit the straightforward details.

Lemma 7.7.3. *Let $\mathcal{D}$ denote the event that for every $v \in [n]$ we have $|\Gamma(v)| \leqslant npn^{2\varepsilon}$. Then $\mathcal{D}$ holds whp in $G_{n,p'}$.* $\square$

With similar reasoning it is also not difficult to see that whp for all large sets S, in $G_{n,p'}$ we have, say, $|\Gamma(S)| \geqslant |S|np$, which is much larger than $|S|$. Intuitively, the next lemma thus implies that for most reasonable sized $A \subseteq [n]$, only a small proportion of $\Gamma(S)$ is contained in $N^{(\leqslant \ell-3)}(A, S \cup A)$.

Lemma 7.7.4. *Let $\mathcal{M}$ denote the event that for all disjoint $A, S \subseteq [n]$ with $|A|, |S| \leqslant kn^{5\varepsilon}$ we have*

$$e\big(S,\ N^{(\leqslant \ell-3)}(A, S \cup A)\big) \leqslant kn^{4\ell\varepsilon}. \tag{7.86}$$

Then $\mathcal{M}$ holds whp in $G_{n,p'}$.

Proof. Let Ψ contain all pairs (A, S) with disjoint $A, S \subseteq [n]$ satisfying $|A|, |S| \leqslant kn^{5\varepsilon}$. Given $\psi = (A, S) \in \Psi$, let $\mathcal{M}_\psi$ denote the event that (7.86) holds, and let $\mathcal{Y}_\psi$ contain all $Y \subseteq A \cup \bigcup_{1 \leqslant d \leqslant \ell-3} V_d(S \cup A)$ with $|Y| \leqslant (npn^{2\varepsilon})^{\ell-2} n^{5\varepsilon}$. Given $\psi = (A, S) \in \Psi$ and $Y \in \mathcal{Y}_\psi$, let $\mathcal{N}_{\psi,Y}$ denote the event that $N^{(\leqslant \ell-3)}(A, S \cup A) = Y$. Using $k \leqslant npn^{\varepsilon}$, it is not difficult to see that whenever $\mathcal{D}$ holds, then for every $\psi \in \Psi$ some $\mathcal{N}_{\psi,Y}$ with $Y \in \mathcal{Y}_\psi$ holds.

Furthermore, $\neg\mathcal{M}$ clearly implies that some $\mathcal{M}_\psi$ with $\psi \in \Psi$ fails. So, we obtain

$$\mathbb{P}[\neg\mathcal{M}] \leqslant \mathbb{P}[\neg\mathcal{D}] + \sum_{\psi=(A,S)\in\Psi} \sum_{Y\in\mathcal{Y}_\psi} \mathbb{P}[\neg\mathcal{M}_\psi \cap \mathcal{N}_{\psi,Y}].$$

Note that for every $\psi = (A,S) \in \Psi$ the events $\mathcal{N}_{\psi,Y}$ are mutually exclusive. So, using $|\Psi| \leqslant n^{2kn^{5\varepsilon}}$ and that $\mathcal{D}$ holds whp by Lemma 7.7.3, to finish the proof it is enough to show that for every $\psi = (A,S) \in \Psi$ and $Y \in \mathcal{Y}_\psi$ we have

$$\mathbb{P}[\neg\mathcal{M}_\psi \mid \mathcal{N}_{\psi,Y}] \leqslant n^{-\omega(kn^{5\varepsilon})}. \tag{7.87}$$

Observe that we can find $Y = N^{(\leqslant \ell-3)}(A, S\cup A)$ by starting with $N^{(0)}(A, S\cup A) = A$, and then iteratively testing vertices in $V_d(S\cup A)$ to see whether they are adjacent to $N^{(d-1)}(A, S\cup A)$, up to $d = \ell - 3$. Since S is disjoint from A and all $V_d(S\cup A)$ with $1 \leqslant d \leqslant \ell - 3$, this exploration has not revealed any pairs between S and Y. We deduce that, conditioned on $\mathcal{N}_{\psi,Y}$, all edges between S and $Y = N^{(\leqslant \ell-3)}(A, S\cup A)$ are included independently with probability $p' = pn^\varepsilon$. Now, using $(np)^{\ell-2} = p^{-1}$ and $\ell \geqslant 4$, the expected number of these edges is bounded by

$$|S| \cdot |Y| \cdot p' \leqslant kn^{5\varepsilon} \cdot (npn^{2\varepsilon})^{\ell-2} n^{5\varepsilon} \cdot pn^\varepsilon = kn^{(2\ell+7)\varepsilon} \leqslant kn^{(4\ell-1)\varepsilon}.$$

Thus standard Chernoff bounds, see e.g. (7.10) of Lemma 7.2.3, imply (7.87), completing the proof. □

7.7.2.2 Paths ending in the neighbourhood of another set

We start with a technical lemma, which will be used in the subsequent proofs of Lemmas 7.7.6 and 7.7.8.

Lemma 7.7.5. *Let $\mathcal{Q}_1$ denote the event that for all $v \in [n]$ and $A, X \subseteq [n]$ with $A \subseteq X$ and $|A|, |X| \leqslant kn^{5\ell\varepsilon}$, for every $2 \leqslant j \leqslant \ell - 1$ and $0 \leqslant d \leqslant \ell - 3$ there are at most at most $(np)^{j-1} n^{9\ell\varepsilon}$ vertices $w \in N^{(\leqslant d)}(A, X)$ for which there exists a path*

$$v = w_0 \cdots w_j = w \qquad \textit{with} \qquad \{w_0, \ldots, w_{j-1}\} \cap N^{(\leqslant d)}(A, X) = \emptyset. \tag{7.88}$$

Then $\mathcal{Q}_1$ holds whp in $G_{n,p'}$.

Proof. Let Ψ contain all tuples (v, A, X, j, d) with $v \in [n]$, $A, X \subseteq [n]$, $2 \leqslant j \leqslant \ell - 1$ and $0 \leqslant d \leqslant \ell - 3$ satisfying $A \subseteq X$ and $|A|, |X| \leqslant kn^{5\ell\varepsilon}$. Given $\psi = (v, A, X, j, d) \in \Psi$, by $\mathcal{Q}_\psi$ we denote the event that there are at most $(np)^{j-1}n^{9\ell\varepsilon}$ vertices $w \in N^{(\leqslant d)}(A, X)$ for which there exists a path satisfying (7.88). Clearly, $\neg\mathcal{Q}_1$ implies that some $\mathcal{Q}_\psi$ with $\psi \in \Psi$ fails.

Next, given $\psi = (v, A, X, j, d) \in \Psi$, we denote by $\mathcal{Y}_\psi$ the set of pairs (Y, Z) with $Y \subseteq A \cup \bigcup_{1 \leqslant d' \leqslant d} V_{d'}(X)$ and $Z \subseteq [n] \setminus Y$ satisfying $|Y| \leqslant (npn^{2\varepsilon})^{d+1}n^{5\ell\varepsilon}$ and $|Z| \leqslant (npn^{2\varepsilon})^{j-1}$. Furthermore, for every $Y \subseteq [n]$ and $v \in [n]$ we inductively define

$$\Gamma^{(0)}(v, Y) = \{v\} \setminus Y \qquad \text{and} \qquad \Gamma^{(i+1)}(v, Y) = \Gamma(\Gamma^{(i)}(v, Y)) \setminus Y. \tag{7.89}$$

Given $\psi = (v, A, X, j, d) \in \Psi$ and $\phi = (Y, Z) \in \mathcal{Y}_\psi$, let $\mathcal{N}_{\psi,\phi}$ be the event that $N^{(\leqslant d)}(A, X) = Y$ and $\Gamma^{(j-1)}(v, Y) = Z$. Whenever $\mathcal{D}$ holds, using $k \leqslant npn^\varepsilon$ it is easy to see that for every $\psi \in \Psi$ some $\mathcal{N}_{\psi,\phi}$ with $\phi \in \mathcal{Y}_\psi$ holds. Putting things together, we obtain

$$\mathbb{P}[\neg\mathcal{Q}_1] \leqslant \mathbb{P}[\neg\mathcal{D}] + \sum_{\psi=(v,A,X,j,d)\in\Psi} \; \sum_{\phi=(Y,Z)\in\mathcal{Y}_\psi} \mathbb{P}[\neg\mathcal{Q}_\psi \cap \mathcal{N}_{\psi,\phi}].$$

Since $\mathcal{D}$ holds whp by Lemma 7.7.3, using $|\Psi| \leqslant n^{3kn^{5\ell\varepsilon}}$ and that for every $\psi \in \Psi$ the events $\mathcal{N}_{\psi,\phi}$ are mutually exclusive, to complete the proof it suffices to show that for every $\psi = (v, A, X, j, d) \in \Psi$ and $\phi = (Y, Z) \in \mathcal{Y}_\psi$ we have

$$\mathbb{P}[\neg\mathcal{Q}_\psi \mid \mathcal{N}_{\psi,\phi}] \leqslant n^{-\omega(kn^{5\ell\varepsilon})}. \tag{7.90}$$

Recall that on $\mathcal{N}_{\psi,\phi}$ we have $Y = N^{(\leqslant d)}(A, X)$ and $Z = \Gamma^{(j-1)}(v, Y)$. Every $w \in Y$ for which there exists a path satisfying (7.88) is contained in $\Gamma(Z)$, and so whenever $\mathcal{Q}_\psi$ fails we deduce $|\Gamma(Z) \cap Y| \geqslant (np)^{j-1}n^{9\ell\varepsilon}$, which in turn implies

$$e(Y, Z) \geqslant (np)^{j-1}n^{9\ell\varepsilon}. \tag{7.91}$$

Next we analyse the distribution of the edges between Y and Z conditional on $\mathcal{N}_{\psi,\phi}$. We can iteratively determine $Y = N^{(\leqslant d)}(A, X)$ as in the proof of Lemma 7.7.4. Then, given Y, we can similarly find $Z = \Gamma^{(j-1)}(v, Y)$; by (7.89) this can clearly be done without testing any pairs between Y and Z. It certainly can happen that during the first exploration, i.e., when determining Y, we have already revealed some pairs between Y and Z, consider e.g. the case where $Z \cap V_1(X) \neq \emptyset$. However, by construction all such pairs are *non-edges*. Therefore the number of edges between Y and Z is stochastically dominated by a binomial distribution with $|Y| \cdot |Z|$ trials and success probability $p' = pn^{\varepsilon}$. Using $d \leqslant \ell - 3$ and $j \leqslant \ell - 1$ as well as $(np)^{d+1} \leqslant (np)^{\ell-2} = p^{-1}$, the expected value of the corresponding binomial random variable is at most

$$|Y| \cdot |Z| \cdot p' \leqslant (npn^{2\varepsilon})^{d+1} n^{5\ell\varepsilon} \cdot (npn^{2\varepsilon})^{j-1} \cdot pn^{\varepsilon} \leqslant (np)^{j-1} n^{(5\ell+2d+2j+1)\varepsilon} \leqslant (np)^{j-1} n^{(9\ell-1)\varepsilon}.$$

So, since $j \geqslant 2$ and $k \leqslant npn^{\varepsilon}$, standard Chernoff bounds show that (7.91) holds with probability at most $e^{-kn^{8\ell\varepsilon}}$, see e.g. (7.10) of Lemma 7.2.3. This establishes (7.90) and thus completes the proof. □

Given a vertex $v \in [n]$, we expect that roughly $(np')^{\ell-2}$ vertices $w \in [n]$ are endpoints of a path $v = w_0 \cdots w_{\ell-2} = w$. Loosely speaking, the next lemma states that there are significantly fewer such vertices w if we only count endpoints in a certain restricted set and forbid some exceptional paths. For the argument of Section 7.8 it is important to observe that $\mathcal{P}_1$ is monotone decreasing.

Lemma 7.7.6. *Let $\mathcal{P}_1$ denote the event that for all disjoint $A, S \subseteq [n]$ with $|A|, |S| \leqslant k$ there exists $X \subseteq [n]$ with $|X| \leqslant kn^{5\ell\varepsilon}$, such that for every $v \in S$ there are at most $(np)^{\ell-3} n^{15\ell\varepsilon}$ vertices $w \in N^{(\ell-3)}(A, X)$ for which there exists a path*

$$v = w_0 \cdots w_{\ell-2} = w \qquad \text{with} \qquad w_1 \notin A. \tag{7.92}$$

Then $\mathcal{P}_1$ holds whp in $G_{n,p'}$.

Proof. By Lemmas 7.7.3, 7.7.4 and 7.7.5 the event $\mathcal{D} \cap \mathcal{M} \cap \mathcal{Q}_1$ holds whp. In the following we are going to argue that for every graph G satisfying those properties, $\mathcal{P}_1$ holds as well. As this claim is purely deterministic, it suffices to prove it for fixed disjoint $A, S \subseteq [n]$ with

$|A|, |S| \leqslant k$. By $\mathcal{M}$ there are at most $kn^{4\ell\varepsilon}$ edges between S and $N^{(\leqslant \ell-3)}(A, S \cup A)$. Let $V_{S,A}$ contain the endpoints of those edges and define

$$X = A \cup S \cup V_{S,A}. \tag{7.93}$$

Note that $|X| \leqslant kn^{5\ell\varepsilon}$. Given $v \in [n]$, by W_v we denote the set of $w \in N^{(\ell-3)}(A, X)$ for which there exists a path satisfying (7.92). To finish the proof, it suffices to show that for every $v \in S$ we have

$$|W_v| \leqslant (np)^{\ell-3} n^{10\ell\varepsilon}. \tag{7.94}$$

Fix $v \in S \subseteq X$. Since $S \cap A = \emptyset$, for every path $v = w_0 \cdots w_{\ell-2} = w$ with $w \in N^{(\ell-3)}(A, X)$ there exists $1 \leqslant j \leqslant \ell - 2$ such that

$$\{w_0, \ldots, w_{j-1}\} \cap N^{(\leqslant \ell-3)}(A, X) = \emptyset \qquad \text{and} \qquad w_j \in N^{(\leqslant \ell-3)}(A, X). \tag{7.95}$$

Recall that by assumption $w_1 \notin A$. So, by (7.50) and (7.93) we may restrict our attention to the case $j \geqslant 2$, since S has no neighbours in $N^{(\leqslant \ell-3)}(A, X) \setminus A$. Now, as $\mathcal{Q}_1$ holds, considering $d \leftarrow \ell - 3$, for every $2 \leqslant j \leqslant \ell - 2$ we deduce that there are at most $(np)^{j-1} n^{9\ell\varepsilon}$ vertices $w_j \in N^{(\leqslant \ell-3)}(A, X)$ for which there exists a path $v = w_0 \cdots w_j$ satisfying (7.95). Recall that the degree of every vertex is at most $npn^{2\varepsilon}$ by $\mathcal{D}$. So, given w_j, there are at most $(npn^{2\varepsilon})^{\ell-j-2}$ vertices $w \in N^{(\ell-3)}(A, X)$ for which there exists a path $w_j \cdots w_{\ell-2} = w$. Putting things together, we deduce that

$$|W_v| \leqslant \sum_{2 \leqslant j \leqslant \ell-2} (np)^{j-1} n^{9\ell\varepsilon} \cdot (npn^{2\varepsilon})^{\ell-j-2} \leqslant (np)^{\ell-3} n^{15\ell\varepsilon}.$$

As explained, this implies $\mathcal{P}_1$, and the proof is complete. $\square$

Note that in Lemma 7.7.6 a condition of the form $w_1 \notin A$ is necessary. Indeed, standard Chernoff bounds imply that whp every vertex has degree $\Omega(np')$. Furthermore, e.g. with a similar argument as in the proof of Lemma 10.6 in [21], one can show that whp for all choices of A, S, X, for all $Z \subseteq A$ with $|Z| \geqslant np$ we have, say, $|N^{(\ell-3)}(Z, X)| \geqslant |Z|(np)^{\ell-3} \geqslant (np)^{\ell-2}$. So, by picking $A \in \binom{[n]}{k}$ such that it contains at least $np = o(k)$ neighbours of some

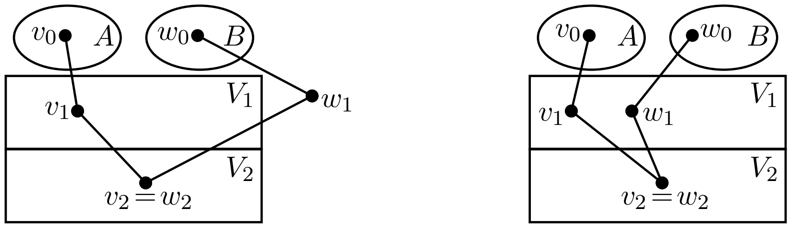

Figure 7.5: Examples of $(2, 2)$-paths for $\ell = 5$. As usual, solid lines represent edges; for the other pairs there are no restrictions. Note that w_1 may be in $A \cup B$ or the vertex classes $V_1 \cup V_2$.

vertex v^*, we have at least $(np)^{\ell-2}$ vertices $w \in N^{(\ell-3)}(A, X)$ which are endpoints of paths $v^* = w_0 \cdots w_{\ell-2} = w$ with $w_1 \in A$, violating the claimed bound.

7.7.2.3 Paths connecting two sets

Given $A, B, X \subseteq [n]$, for every $j \geqslant 1$ and $0 \leqslant d \leqslant \ell - 3$, we say that $w_0 \cdots w_j = v_d \cdots v_0$ is a *(j, d)-path* wrt. (A, B, X) if $v_0 \in A$, $w_0 \in B$ and $v_{d'} \in V_{d'}(X)$ for all $1 \leqslant d' \leqslant d$, cf. Figure 7.5. Intuitively, the next technical result states that the number of (j, d)-paths is not 'too large' if we allow for deleting a few edges.

Lemma 7.7.7. *Let $\mathcal{Q}_2$ denote the event that for all $A, B \subseteq [n]$ with $|A|, |B| \leqslant k$ there exists $F \subseteq \binom{[n]}{2}$ with $|F| \leqslant kn^{2\varepsilon}$, such that for every $1 \leqslant j \leqslant \ell - 1$ and $0 \leqslant d \leqslant \ell - 4$ the number of (j, d)-paths wrt. $(A, B, A \cup B)$ that are edge disjoint from F is bounded by $k^2(np)^{j-3}n^{4\ell\varepsilon}$. Then $\mathcal{Q}_2$ holds whp in $G_{n,p'}$.*

Proof. Fix $A, B \subseteq [n]$ with $|A|, |B| \leqslant k$. Given j and d, we denote by $\mathcal{S}_{j,d} = \mathcal{S}_{j,d}(A, B)$ the family of edge-sets of all possible (j, d)-paths wrt. $(A, B, A \cup B)$. Clearly, $|V_{d'}(A \cup B)| \leqslant n$ for all $1 \leqslant d' \leqslant d$. So, using $p = (np)^{-(\ell-2)}$, $j \leqslant \ell - 1$ and $d \leqslant \ell - 4$, the expected number $\mu_{j,d}$ of such (j, d)-paths satisfies

$$\mu_{j,d} \leqslant k^2 n^{j+d-1} (pn^{\varepsilon})^{d+j} \leqslant k^2 (np)^{j+d-1} pn^{2\ell\varepsilon} = k^2 (np)^{j+d+1-\ell} n^{2\ell\varepsilon} \leqslant k^2 (np)^{j-3} n^{2\ell\varepsilon}.$$

Set $\kappa_j = k^2(np)^{j-3}n^{3\ell\varepsilon}$ and $b = kn^{\varepsilon}$. Using the Deletion Lemma (cf. Lemma 7.2.4) the probability that $\mathcal{DL}(b, \kappa_j, \mathcal{S}_{j,d})$ fails for some $1 \leqslant j \leqslant \ell - 1$ and $0 \leqslant d \leqslant \ell - 4$ is bounded by

$$\sum_{1 \leqslant j \leqslant \ell} \sum_{0 \leqslant d \leqslant \ell-4} (1 + \kappa_j/\mu_{j,d})^{-b} \leqslant \ell^2 \cdot n^{-\ell\varepsilon b} = n^{-\omega(k)},$$

with room to spare. Whenever $\mathcal{DL}(b, \kappa_j, \mathcal{S}_{j,d})$ holds, we denote by $F_{j,d}$ the corresponding 'ignored' edge set E_0 as in Lemma 7.2.4. If all $\mathcal{DL}(b, \kappa_j, \mathcal{S}_{j,d})$ with $1 \leqslant j \leqslant \ell - 1$ and $0 \leqslant d \leqslant \ell - 4$ hold simultaneously, then defining F as the union of all edge sets $F_{j,d}$ has the required properties. Finally, taking the union bound over all choices of A and B completes the proof. $\square$

For most large sets B and W, we expect that the number of $(b, w) \in B \times W$ for which there exists a path $b = w_0 \cdots w_{\ell-2} = w$ should be roughly $|B||W|n^{\ell-3}p'^{\ell-2} = |B||W|n^{(\ell-2)\varepsilon}/(np)$. Loosely speaking, the next lemma suggests that for most reasonable sized $A, B \subseteq [n]$, this upper bound holds for $W = N^{(\ell-4)}(A, X)$ if we forbid certain exceptional paths, as in this case $|W| \approx |A|(np')^{\ell-4}$.

Lemma 7.7.8. *Let $\mathcal{P}_2$ denote the event that for all disjoint $A, B \subseteq [n]$ with $|A|, |B| \leqslant k$ there exists $X \subseteq [n]$ and $F \subseteq \binom{[n]}{2}$ with $|X| \leqslant kn^{5\ell\varepsilon}$ and $|F| \leqslant kn^{2\varepsilon}$, such that the number of pairs $(b, w) \in B \times N^{(\ell-4)}(A, X)$ for which there exists a path $b = w_0 \cdots w_{\ell-2} = w$ with*

$$w_1 \notin A \qquad \textit{and} \qquad \big(w_2 \notin A \;\; \textit{or} \;\; \{w_0w_1, w_1w_2\} \cap F = \emptyset\big) \tag{7.96}$$

is at most $k^2(np)^{\ell-5}n^{15\ell\varepsilon}$. Then $\mathcal{P}_2$ holds whp in $G_{n,p'}$.

Before turning to the proof, note that $\mathcal{P}_2$ is monotone decreasing.

Proof of Lemma 7.7.8. By Lemmas 7.7.3, 7.7.4, 7.7.5 and 7.7.7 it is enough to show that $\mathcal{P}_2$ holds for every graph G satisfying $\mathcal{D} \cap \mathcal{M} \cap \mathcal{Q}_1 \cap \mathcal{Q}_2$. As this claim is purely deterministic, it suffices to prove it for fixed disjoint $A, B \subseteq [n]$ with $|A|, |B| \leqslant k$. Given $X \subseteq [n]$ and $F \subseteq \binom{[n]}{2}$, we denote by $P_{j,d}(X, F)$ the set of (j, d)-paths wrt. (A, B, X) that are edge disjoint from F. By $\mathcal{Q}_2$ there exists $F \subseteq \binom{[n]}{2}$ with $|F| \leqslant kn^{2\varepsilon}$ such that for all $1 \leqslant j \leqslant \ell - 2$

and $0 \leqslant d \leqslant \ell - 4$ we have

$$|P_{j,d}(A \cup B, F)| \leqslant k^2 (np)^{j-3} n^{4\ell\varepsilon}. \tag{7.97}$$

Let V_F contain all vertices outside A that are endpoints of edges in F. Note that $|V_F| \leqslant 2kn^{2\varepsilon}$. Considering $S \leftarrow B \cup V_F$, by $\mathcal{M}$ there are at most $kn^{4\ell\varepsilon}$ edges between $B \cup V_F$ and $N^{(\leqslant \ell-3)}(A, B \cup V_F \cup A)$. Let $V_{B,F}$ contain the endpoints of all those edges and set

$$X = A \cup B \cup V_F \cup V_{B,F}. \tag{7.98}$$

Observe that, say, $|X| \leqslant kn^{5\ell\varepsilon}$. Furthermore, using (7.50) we see that

$$V_F \cap \bigcup_{1 \leqslant \kappa \leqslant \ell-4} V_\kappa(X) = \emptyset \qquad \text{and} \qquad \Gamma\big(V_F \cup B\big) \cap \big(N^{(\leqslant \ell-4)}(A, X) \setminus A\big) = \emptyset. \tag{7.99}$$

For every $1 \leqslant j \leqslant \ell - 2$ we define W_j as the set of all pairs $(b, y) \in B \times N^{(\leqslant \ell-4)}(A, X)$ for which there exists a path $b = w_0 \cdots w_j = y$ satisfying (7.96) and

$$\{w_0, \ldots, w_{j-1}\} \cap N^{(\leqslant \ell-4)}(A, X) = \emptyset \qquad \text{and} \qquad w_j \in N^{(\leqslant \ell-4)}(A, X). \tag{7.100}$$

We claim that in order to complete the proof, it suffices to show that for all $1 \leqslant j \leqslant \ell - 2$ we have

$$|W_j| \leqslant k^2 (np)^{j-3} n^{10\ell\varepsilon}. \tag{7.101}$$

Indeed, let W contain all pairs $(b, w) \in B \times N^{(\ell-4)}(A, X)$ for which there exists a path $b = w_0 \cdots w_{\ell-2} = w$ satisfying (7.96). Note that for every such $b = w_0 \cdots w_{\ell-2} = w$ there exists $1 \leqslant j \leqslant \ell - 2$ such that $b = w_0 \cdots w_j$ satisfies (7.100). Recall that by $\mathcal{D}$ the degree is bounded by $npn^{2\varepsilon}$. So, given w_j, there are at most $(npn^{2\varepsilon})^{\ell-j-2}$ vertices $w \in N^{(\ell-4)}(A, X)$ for which there exists a path $w_j \cdots w_{\ell-2} = w$. Putting things together, assuming (7.101) we obtain

$$|W| \leqslant \sum_{1 \leqslant j \leqslant \ell-2} |W_j| \cdot (npn^{2\varepsilon})^{\ell-j-2} \leqslant k^2 (np)^{\ell-5} n^{15\ell\varepsilon},$$

and so $\mathcal{P}_2$ holds, as claimed.

We shall now prove (7.101). Observe that for $j = 1$ we need to consider paths w_0w_1 with $w_0 \in B$ and $w_1 \in N^{(\leqslant \ell-4)}(A, X) \setminus A$. Now, using the second part of (7.99) we see that $w_1 \in \Gamma(w_0) \cap (N^{(\leqslant \ell-4)}(A, X) \setminus A)$ is impossible. This implies $|W_1| = 0$, which clearly establishes (7.101) for $j = 1$.

For $j \geqslant 2$ we first consider $W_{j,F} \subseteq W_j$, which contains all pairs $(b, y) \in W_j$ for which there exists a path $b = w_0 \cdots w_j = y$ satisfying (7.100) and

$$\{w_0w_1, \ldots, w_{j-1}w_j\} \cap F = \emptyset. \tag{7.102}$$

Clearly, for every $(b, y) \in W_{j,F}$ there exists $0 \leqslant d \leqslant \ell - 4$ such that at least one (j, d)-path wrt. (A, B, X) with $b = w_0$ and $w_j = y$ satisfies (7.102). We claim that the corresponding (j, d)-path $w_0 \cdots w_j = v_d \cdots v_0$ is edge-disjoint from F, i.e., contained in $P_{j,d}(X, F)$. To see this, observe that every $f \in \{v_dv_{d-1}, \cdots, v_1v_0\} \cap F$ has at least one vertex outside of A, say $v_\kappa \in V_\kappa(X)$ with $1 \leqslant \kappa \leqslant d$, which contradicts (7.99), since by construction $v_\kappa \in V_F$. In addition, by (7.50) and (7.98) we see that $P_{j,d}(X, F) \subseteq P_{j,d}(A \cup B, F)$. Putting things together, using (7.97) our discussion yields

$$|W_{j,F}| \leqslant \sum_{0 \leqslant d \leqslant \ell-4} |P_{j,d}(X, F)| \leqslant \sum_{0 \leqslant d \leqslant \ell-4} |P_{j,d}(A \cup B, F)| \leqslant k^2(np)^{j-3}n^{5\ell\varepsilon}. \tag{7.103}$$

It remains to estimate the number of pairs in $W^*_{j,F} = W_j \setminus W_{j,F}$, where the corresponding paths intersect with F. We start with the special case $j = 2$, i.e., paths $b = w_0w_1w_2 = y$ with $(b, y) \in W^*_{2,F}$ satisfying (7.96). Observe that every $f \in \{w_0w_1, w_1w_2\} \cap F$ contains $w_1 \in V_F$, since $w_1 \notin A$ by (7.96). Note that $w_2 \in A$ contradicts the second part of (7.96), and that $w_2 \in \Gamma(w_1) \cap (N^{(\leqslant \ell-4)}(A, X) \setminus A)$ is impossible by (7.99). To sum up, $|W^*_{2,F}| = 0$, which together with (7.103) implies (7.101) for $j = 2$.

Turning to $j \geqslant 3$, for every $1 \leqslant \varsigma \leqslant j$ we denote by $W^*_{j,F,\varsigma} \subseteq W^*_{j,F}$ the set of pairs $(b, y) \in W^*_{j,F}$ with $y \notin A$ where the corresponding path $b = w_0 \cdots w_j = y$ satisfies $w_{\varsigma-1}w_\varsigma \in F$ and (7.100). We claim that it is enough to show that for every $1 \leqslant \varsigma \leqslant j$ we have

$$|W^*_{j,F,\varsigma}| \leqslant k^2(np)^{j-3}n^{8\ell\varepsilon}. \tag{7.104}$$

Indeed, since there are at most $|B| \cdot |A| \leqslant k^2 \leqslant k^2(np)^{j-3}$ pairs $(b, y) \in W^*_{j,F}$ with $y \in A$, we obtain

$$|W^*_{j,F}| \leqslant k^2(np)^{j-3} + \sum_{1 \leqslant \varsigma \leqslant j} |W^*_{j,F,\varsigma}| \leqslant k^2(np)^{j-3}n^{9\ell\varepsilon},$$

which together with (7.103) establishes (7.101), as claimed.

In the following we verify (7.104). First we show that $|W^*_{j,F,\varsigma}| = 0$ for $\varsigma \in \{j-1, j\}$. If $w_{j-1}w_j \in F$, then $w_j \notin A$ implies $w_j \in V_F$, but the remaining possibility $w_j \in N^{(\leqslant \ell-4)}(A, X) \setminus A$ contradicts (7.99). If $w_{j-2}w_{j-1} \in F$, then by (7.100) we have $w_{j-1} \notin N^{(\leqslant \ell-4)}(A, X)$ and so $w_{j-1} \in V_F$. Since by assumption $w_j \notin A$ we must have $w_j \in \Gamma(w_{j-1}) \cap (N^{(\leqslant \ell-4)}(A, X) \setminus A)$, which is impossible by (7.99).

Now, suppose that $w_{\varsigma-1}w_\varsigma \in F$ with $1 \leqslant \varsigma \leqslant j-2$. Considering $v \leftarrow w_\varsigma$ and $d \leftarrow \ell - 4$, by $\mathcal{Q}_1$ there are at most $(np)^{j-\varsigma-1}n^{6\ell\varepsilon}$ vertices $w_j \in N^{(\leqslant \ell-4)}(A, X)$ for which there exists a path $w_\varsigma = w'_0 \cdots w'_{j-\varsigma} = w_j$ with $\{w_\varsigma, \ldots, w_{j-1}\} \cap N^{(\leqslant \ell-4)}(A, X) = \emptyset$. So, using $|F| \leqslant kn^{2\varepsilon}$, since there are at most $|B| = k$ choices for $b \in B$, for $\varsigma \geqslant 2$ we deduce that

$$|W^*_{j,F,\varsigma}| \leqslant |B| \cdot 2|F| \cdot (np)^{j-\varsigma-1}n^{6\ell\varepsilon} \leqslant k^2(np)^{j-\varsigma-1}n^{(6\ell+3)\varepsilon} \leqslant k^2(np)^{j-3}n^{8\ell\varepsilon},$$

as claimed. Note that for the remaining case $\varsigma = 1$ each (ordered) edge $w_0w_1 \in F$ also determines the vertex $b = w_0 \in B$. So, compared to the estimate above we win a factor of $|B|$, and a virtually identical calculation yields that (7.104) also holds in this case, which completes the proof. □

With very similar reasoning as for Lemma 7.7.6, one can argue that an extra condition for the case $w_2 \in A$ is needed in Lemma 7.7.8: this time we can otherwise violate the claimed bound whp by fixing some vertex v^* and then choosing disjoint $A, B \subseteq [n]$ such that each contains at least np vertices from $\Gamma(v^*)$; we leave the details to the interested reader.

7.8 Very good configurations exist

In this section we prove Lemma 7.4.3. Given a graph property $\mathcal{Y}$, let $\mathcal{Y}_i$ denote the event $G(i) \in \mathcal{Y}$, i.e., that $G(i)$ satisfies $\mathcal{Y}$. Now, for every $0 \leqslant i \leqslant m$ we set

$$\mathcal{W}_i = \mathcal{I}_i \cap \mathcal{K}_i \cap \mathcal{L}_i \cap \mathcal{N}_i \cap \mathcal{P}_{1,i} \cap \mathcal{P}_{2,i} \cap \mathcal{T}_i,$$

where $\mathcal{K}_i$, $\mathcal{L}_i$, $\mathcal{T}_i$ are defined as in Theorem 7.2.1 and Lemma 7.2.2, and $\mathcal{I}$, $\mathcal{N}$, $\mathcal{P}_1$, $\mathcal{P}_2$ are defined as in Lemmas 7.7.1, 7.7.2, 7.7.6 and 7.7.8. It is not difficult to see that $\mathcal{W}_i$ is monotone decreasing and, using the 'transfer theorem' (Theorem 7.6.2), that $\mathcal{W}_m$ holds whp. Observe that by monotonicity $\mathcal{W}_m$ implies $\mathcal{W}_i$ for every $i \leqslant m$, and that $\neg\mathcal{B}_i(\Sigma)$ implies $\neg\mathcal{B}_{i-1}(\Sigma)$. So, to complete the proof it suffices to consider fixed $G(i)$ satisfying $\mathcal{W}_i$ and show that for every $(\tilde{v}, U)$ with $U \in \binom{[n]\setminus\{\tilde{v}\}}{u}$ there exists $\Sigma^* = (\tilde{v}, U, A, B, R) \in \mathcal{C}$ satisfying $\neg\mathcal{B}_i(\Sigma^*)$ and (7.55). In fact, since the above claim is purely deterministic, it is enough to also consider fixed $(\tilde{v}, U)$. Our proof proceeds in several steps and we tacitly assume that n is sufficiently large whenever necessary. First, in Section 7.8.1 we choose a 'special' configuration $\Sigma^* = (\tilde{v}, U, A, B, R)$ and collect some of its basic properties. In the remaining sections we verify that Σ^* has the properties claimed by Lemma 7.4.3. More precisely, in Section 7.8.2 we show that $\neg\mathcal{B}_i(\Sigma^*)$ holds, and in Section 7.8.3 we establish (7.55).

7.8.1 Finding $\Sigma^* = (\tilde{v}, U, A, B, R)$

In the following we show how we pick $\Sigma^* = (\tilde{v}, U, A, B, R)$. Along the way, we furthermore collect some immediate properties of the resulting Σ^*. We set

$$\tau = 40\ell \qquad \text{and} \qquad \vartheta = 20\ell\tau = 800\ell^2. \tag{7.105}$$

For the main steps of our argument it is useful to keep in mind that $\vartheta \gg \tau \gg \ell$ and $\vartheta\varepsilon \ll 1/\ell$. First, we choose $S \subseteq U$ such that

$$S \text{ is an independent set} \qquad \text{and} \qquad |S| \geqslant u/6, \tag{7.106}$$

which is possible since $\mathcal{I}_i$ holds. Henceforth we assume that $v_1, \ldots, v_n \in [n]$ are ordered so that

$$|\Gamma(v_1) \cap S| \geqslant |\Gamma(v_2) \cap S| \geqslant \cdots \geqslant |\Gamma(v_j) \cap S| \geqslant \cdots \geqslant |\Gamma(v_n) \cap S|. \tag{7.107}$$

We greedily choose first ℓ_A, and afterwards ℓ_B, such that they are the smallest indices for which

$$N_A = \bigcup_{1 \leqslant j \leqslant \ell_A} \big(\Gamma(v_j) \cap S\big) \qquad \text{and} \qquad N_B = \bigcup_{\ell_A < j \leqslant \ell_B} \big(\Gamma(v_j) \cap S\big) \setminus N_A$$

each have cardinality at least $2k$, where we set the corresponding index to ∞ if this is not possible. Recall that $k = \gamma/60 \cdot npt_{\max}$ by (7.48) and $\gamma \geqslant 180$ by (7.47). So, since $\mathcal{T}_i$ holds, by (7.6) the maximum degree is at most $3npt_{\max} \leqslant k$. Using $k = u/60$, we deduce that

$$|N_A \cup N_B| \leqslant 6k \leqslant u/10. \tag{7.108}$$

7.8.1.1 Picking A, B

If $\ell_B = \infty$ or $\ell_B > n^{2\vartheta\varepsilon}$, we choose arbitrary disjoint sets, each of size $k = u/60$, satisfying

$$A, B \subseteq S \setminus (N_A \cup N_B),$$

which is possible by (7.106) and (7.108). For later usage, we furthermore set $I_A = \emptyset$ and $I_B = \emptyset$.

If $\ell_B \leqslant n^{2\vartheta\varepsilon} = o(k)$, we set $I_A = \{v_1, \ldots, v_{\ell_A}\}$ and $I_B = \{v_{\ell_A+1}, \ldots, v_{\ell_B}\}$. Since $G(i)$ satisfies $\mathcal{N}_i$, the codegrees are all bounded by nine, and thus

$$|\Gamma(I_B) \cap N_A| \leqslant |\Gamma(I_B) \cap \Gamma(I_A)| \leqslant 9 \cdot \ell_B \cdot \ell_A \leqslant 9n^{4\vartheta\varepsilon} = o(k). \tag{7.109}$$

Now we choose arbitrary sets, each of size k, satisfying

$$A \subseteq N_A \setminus \big(I_B \cup \Gamma(I_B)\big) \qquad \text{and} \qquad B \subseteq N_B,$$

which is possible by (7.109). Clearly, A and $B \cup I_B$ are disjoint.

Next we estimate the size of certain neighbourhoods (using a similar argument as in [76]).

Lemma 7.8.1. *We have* $\Gamma(I_A) \cap B = \emptyset$ *and* $\Gamma(I_B) \cap A = \emptyset$*. Given* $Y \in \{A, B\}$*, every* $v \notin I_Y$ *satisfies*

$$|\Gamma(v) \cap Y| \leqslant npn^{-\vartheta\varepsilon}. \tag{7.110}$$

Proof. If $\ell_B = \infty$, then all vertices $v \in [n]$ satisfy the stronger bound $|\Gamma(v) \cap (A \cup B)| = 0$.

Next, we consider the case $n^{2\vartheta\varepsilon} < \ell_B < \infty$, where $I_A = I_B = \emptyset$. Since all vertices $v \in \{v_1, \ldots, v_{\ell_B}\}$ satisfy $|\Gamma(v) \cap (A \cup B)| = 0$, using (7.107) it is not difficult to see that in order to prove (7.110), it suffices to show $|\Gamma(v_x) \cap S| \leqslant npn^{-\vartheta\varepsilon}$ for $x = n^{2\vartheta\varepsilon}$. Set $H = \{v_1, \ldots, v_x\}$. On the one hand, using (7.107) we have $2e(H, S) \geqslant x|\Gamma(v_x) \cap S|$. On the other hand, since $G(i)$ satisfies $\mathcal{K}_i$, using $|H| = n^{2\vartheta\varepsilon}$ and $|S| \leqslant npn^{\varepsilon}$, we have, say, $e(H, S) \leqslant npn^{2\varepsilon}$. So, we deduce $|\Gamma(v_x) \cap S| \leqslant npn^{-\vartheta\varepsilon}$, as claimed.

Finally, suppose that $\ell_B \leqslant n^{2\vartheta\varepsilon}$. Observe that $\Gamma(I_A) \cap B = \emptyset$ and $\Gamma(I_B) \cap A = \emptyset$ hold by construction. Fix $Y \in \{A, B\}$. Since by $\mathcal{N}_i$ all codegrees are at most nine, for every $v \notin I_Y$ we have $|\Gamma(v) \cap Y| \leqslant |\Gamma(v) \cap \Gamma(I_Y)| \leqslant 9\ell_B$, which readily establishes (7.110), and thus completes the proof. □

7.8.1.2 Choosing R

Observe that $|I_B| \leqslant n^{2\vartheta\varepsilon}$. Considering A and $S \leftarrow I_B$, we denote by X_1 the set X whose existence is guaranteed by $\mathcal{P}_{1,i}$. Similarly, let X_2 and F denote the sets X and F whose existence is guaranteed by $\mathcal{P}_{2,i}$ when considering A and B. We have $|X_1|, |X_2| \leqslant kn^{5\ell\varepsilon}$ and $|F| \leqslant kn^{2\varepsilon}$. Now we set

$$R = \{\tilde{v}\} \cup U \cup X_1 \cup X_2. \tag{7.111}$$

Clearly, $|R| \leqslant kn^{10\ell\varepsilon}$ holds, with room to spare. Next we collect several structural properties. By (7.50) and (7.111) and have $N^{(j)}(A, R) \subseteq N^{(j)}(A, X_1) \cap N^{(j)}(A, X_2)$. So, using $(\Gamma(I_B) \cup I_B) \cap A = \emptyset$, we immediately obtain the following statement:

Lemma 7.8.2. *We have* $|I_B| \leqslant n^{2\vartheta\varepsilon}$*, and for every* $v \in I_B$ *there are at most* $(np)^{\ell-3} n^{15\ell\varepsilon}$ *vertices* $w \in N^{(\ell-3)}(A, R)$ *for which there exists a path* $v = w_0 \cdots w_{\ell-2} = w$*.* □

In addition, using that $A \cup B$ is an independent set, we readily deduce the following result:

Lemma 7.8.3. *We have $|F| \leqslant kn^{2\varepsilon}$, and there are at most $k^2(np)^{\ell-5}n^{15\ell\varepsilon}$ pairs $(b,w) \in B \times N^{(\ell-4)}(A,R)$ for which there exists a path $b = w_0 \cdots w_{\ell-2} = w$ satisfying $w_2 \notin A$ or $\{w_0w_1, w_1w_2\} \cap F = \emptyset$.* □

In the subsequent sections, the construction of A and B is irrelevant; all that we use is that A, B are disjoint subsets of U with size k, and there are sets F, I_A, I_B, R such that the conclusions of Lemmas 7.8.1–7.8.3 hold in $G(i)$.

7.8.2 The configuration Σ^* is good

In this section we show that $\neg\mathcal{B}_i(\Sigma^*) = \neg\mathcal{B}_{1,i}(\Sigma^*) \cap \neg\mathcal{B}_{2,i}(\Sigma^*)$ holds.

7.8.2.1 The bad event $\mathcal{B}_{1,i}(\Sigma^*)$

In order to prove that $\mathcal{B}_{1,i}(\Sigma^*)$ fails, using Lemma 7.8.3 it suffices to show that there are at most $k^2(np)^{\ell-4}n^{-10\varepsilon}$ paths $w_0 \cdots w_{\ell-2}$ with $(w_0, w_2) \in B \times A$ satisfying $w_0w_1 \in F$ or $w_1w_2 \in F$. Let P_{Σ^*} denote all such paths. For every $w_0w_1 \in F \cap E(i)$ with $w_0 \in B$, using Lemma 7.8.1 we see that $w_1 \notin I_A$, which by (7.110) implies that there are at most $npn^{-\vartheta\varepsilon}$ choices for $w_2 \in \Gamma(w_1) \cap A$. With a similar argument, for every $w_1w_2 \in F \cap E(i)$ with $w_2 \in A$ we have at most $npn^{-\vartheta\varepsilon}$ choices for $w_0 \in \Gamma(w_1) \cap B$. Furthermore, since the degree is bounded by npn^{ε}, given $w_2 \in A$ there are at most $(npn^{\varepsilon})^{\ell-4}$ paths $w_2 \cdots w_{\ell-2}$. So, using $np \leqslant k$, $|F| \leqslant kn^{2\varepsilon}$ and (7.105), i.e., $\vartheta \geqslant 20\ell$, we deduce that

$$|P_{\Sigma^*}| \leqslant npn^{-\vartheta\varepsilon} \cdot 2|F| \cdot (npn^{\varepsilon})^{\ell-4} \leqslant k^2(np)^{\ell-4}n^{(\ell-\vartheta)\varepsilon} < k^2(np)^{\ell-4}n^{-10\varepsilon},$$

which, as explained, establishes $\neg\mathcal{B}_{1,i}(\Sigma^*)$.

7.8.2.2 The bad event $\mathcal{B}_{2,i}(\Sigma^*)$

In anticipation of the estimates in Section 7.8.3, here we analyse the combinatorial structure of $L_{\Sigma^*}(i)$ much more precisely than needed. To this end we introduce the sets $L_{\Sigma^*}(i,j)$, where for every $j \in [\ell - 1]$ we denote by $L_{\Sigma^*}(i,j)$ the set of all *ordered* pairs xy with distinct $x, y \in [n]$ such that $|C_{x,y,\Sigma^*}(i,j)| \geqslant p^{-1}n^{-30\ell\varepsilon}$. We start by showing that we may restrict our attention to the case $j \in \{1,2\}$. Recall that $C_{x,y,\Sigma^*}(i,j)$ contains all pairs $bw \in B \times N^{(\ell-3)}(A,R)$ for which there exist disjoint paths $b = w_1 \cdots w_j = x$ and $y = w_{j+1} \cdots w_\ell = w$ in $G(i)$. Fix $x \neq y$. Since the degree is at most npn^ε by (7.6), for $j \geqslant 3$ the number of choices for w is at most $(npn^\varepsilon)^{\ell-j-1} \leqslant (npn^\varepsilon)^{\ell-4}$. Now, as there are at most $|B| \leqslant k \leqslant npn^\varepsilon$ ways to pick $b \in B$, using $(np)^{\ell-2} = p^{-1}$ we crudely have

$$|C_{x,y,\Sigma^*}(i,j)| \leqslant npn^\varepsilon \cdot (npn^\varepsilon)^{\ell-4} \leqslant p^{-1}n^{\ell\varepsilon}/(np) < p^{-1}n^{-30\ell\varepsilon}, \tag{7.112}$$

which implies $xy \notin L_{\Sigma^*}(i,j)$. Therefore $L_{\Sigma^*}(i,j) = \emptyset$ for $j \geqslant 3$, so

$$|L_{\Sigma^*}(i)| \leqslant |L_{\Sigma^*}(i,1)| + |L_{\Sigma^*}(i,2)|. \tag{7.113}$$

With foresight, for all $j \geqslant 1$ we define $M^{(j)}(A)$ as the set of $v \in [n]$ with $|W^{(j)}(v,A)| \geqslant (np)^j n^{-\tau\varepsilon}$, where $W^{(j)}(v,A)$ contains all vertices $w \in N^{(\ell-3)}(A,R)$ for which there exists a path $v = w_0 \cdots w_j = w$ in $G(i)$. Now we claim that

$$L_{\Sigma^*}(i,2) \subseteq \Big\{ xy \;:\; x \in I_B \;\wedge\; y \in M^{(\ell-3)}(A) \Big\}. \tag{7.114}$$

Note that $C_{x,y,\Sigma^*}(i,2)$ contains only pairs $bw \in B \times N^{(\ell-3)}(A,R)$ for which there exists paths $b = w_1w_2 = x$ and $y = w_3 \cdots w_\ell = w$ in $G(i)$. First suppose that $x \notin I_B$. Using Lemma 7.8.1, by (7.110) we have at most $npn^{-\vartheta\varepsilon}$ choices for $b \in \Gamma(x) \cap B$. Since the degree is at most npn^ε, we have at most $(npn^\varepsilon)^{\ell-3}$ choices for w. So, using $(np)^{\ell-2} = p^{-1}$ and (7.105), i.e., $\vartheta \geqslant 40\ell$, we deduce that

$$|C_{x,y,\Sigma^*}(i,2)| \leqslant npn^{-\vartheta\varepsilon} \cdot (npn^\varepsilon)^{\ell-3} \leqslant p^{-1}n^{(\ell-\vartheta)\varepsilon} < p^{-1}n^{-30\ell\varepsilon},$$

which implies $xy \notin L_{\Sigma^*}(i,2)$. Next, we consider the case where $y \notin M^{(\ell-3)}(A)$. With a very similar reasoning as above, this time using $|W^{(\ell-3)}(y,A)| \leqslant (np)^{\ell-3}n^{-\tau\varepsilon}$ and (7.105), i.e., $\tau = 40\ell$, we obtain

$$|C_{x,y,\Sigma^*}(i,2)| \leqslant npn^{\varepsilon} \cdot (np)^{\ell-3} n^{-\tau\varepsilon} \leqslant p^{-1} n^{(1-\tau)\varepsilon} < p^{-1} n^{-30\ell\varepsilon},$$

which implies $xy \notin L_{\Sigma^*}(i,2)$. This completes the proof of (7.114).

By a similar but simpler argument we furthermore see that

$$L_{\Sigma^*}(i,1) \subseteq \left\{ xy \;:\; x \in B \;\wedge\; y \in M^{(\ell-2)}(A) \right\}. \tag{7.115}$$

Next we estimate the cardinality of $M^{(j)}(A)$. A similar argument is implicit in [17].

Lemma 7.8.4. *For every* $1 \leqslant j \leqslant \ell - 2$ *we have* $|M^{(j)}(A)| \leqslant (np)^{\ell-2-j} n^{2\ell\tau\varepsilon}$.

Proof. Set $H^{(0)}(A) = N^{(\ell-3)}(A,R)$, and for every $j \geqslant 1$ we let $H^{(j)}(A)$ contain all $v \in [n]$ with $|\Gamma(v) \cap H^{(j-1)}(A)| \geqslant npn^{-2\tau\varepsilon}$. First, we claim that for all $1 \leqslant j \leqslant \ell - 2$ we have

$$M^{(j)}(A) \subseteq H^{(j)}(A). \tag{7.116}$$

Since $\tau \geqslant 2\ell$ by (7.105), it clearly suffices to show that for all $1 \leqslant j \leqslant \ell - 2$, for every $v \notin H^{(j)}(A)$ we have $|W^{(j)}(v,A)| \leqslant j(npn^{\varepsilon})^j n^{-2\tau\varepsilon}$. We proceed by induction on j. For the base case $j = 1$ the claim is trivial, since $H^{(1)}(A)$ contains all vertices $v \in [n]$ with $|\Gamma(v) \cap N^{(\ell-3)}(A,R)| \geqslant npn^{-2\tau\varepsilon}$. Turning to $j \geqslant 2$, fix $v \notin H^{(j)}(A)$. By distinguishing between the neighbours of v inside and outside of $H^{(j-1)}(A)$, using the induction hypothesis and that the degree is bounded by npn^{ε}, we obtain

$$|W^{(j)}(v,A)| \leqslant npn^{-2\tau\varepsilon} \cdot (npn^{\varepsilon})^{j-1} + npn^{\varepsilon} \cdot (j-1)(npn^{\varepsilon})^{j-1} n^{-2\tau\varepsilon} \leqslant j(npn^{\varepsilon})^j n^{-2\tau\varepsilon},$$

which, as explained, establishes (7.116).

To finish the proof, again using $\tau \geqslant 2\ell$, it suffices to show that for all $0 \leqslant j \leqslant \ell - 2$ we have

$$|H^{(j)}(A)| \leqslant (np)^{\ell-2-j} n^{(2j\tau+\ell+j)\varepsilon}. \tag{7.117}$$

As before, we proceed by induction on j. Using $|A| \leqslant k \leqslant npn^{\varepsilon}$ and that the degree is bounded by npn^ε, we establish the base case $j = 0$ by observing that $|H^{(0)}(A)| \leqslant |\Gamma^{(\ell-3)}(A)| \leqslant (npn^\varepsilon)^{\ell-2}$. Suppose $j \geqslant 1$. Recall that $(np)^{\ell-2} = p^{-1}$. Since $\mathcal{L}_i$ holds, using the induction hypothesis we obtain

$$|H^{(j)}(A)| \leqslant 16\varepsilon^{-1}(np)^{\ell-2-j}n^{(2j\tau+\ell+j-1)\varepsilon} \leqslant (np)^{\ell-2-j}n^{(2j\tau+\ell+j)\varepsilon},$$

completing the proof. □

With Lemma 7.8.4 in hand, combing (7.113)–(7.115) with $|B| = k \leqslant npn^\varepsilon$ as well as $|I_B| \leqslant n^{2\vartheta\varepsilon}$, and then using (7.1), (7.105) as well as $\ell \geqslant 4$, $np = n^{1/(\ell-1)}$ and $(np)^2 \leqslant (np)^{\ell-2} = p^{-1}$, we deduce that

$$|L_{\Sigma^*}(i)| \leqslant npn^\varepsilon \cdot n^{2\ell\tau\varepsilon} + n^{2\vartheta\varepsilon} \cdot npn^{2\ell\tau\varepsilon} \leqslant npn^{5\vartheta\varepsilon} < (np)^2 n^{-1/(2\ell)} \leqslant p^{-1}n^{-1/(2\ell)},$$

which establishes $\neg\mathcal{B}_{2,i}(\Sigma^*)$.

7.8.3 Few tuples are ignored for Σ^*

In this section we estimate the size of $T_{\Sigma^*,\ell-3}(i) \setminus Z_{\Sigma^*,\ell-3}(i)$. Let $Q_{\Sigma^*}(i)$ contain all pairs $(w_1, w_\ell) \in B \times N^{(\ell-3)}(A, R)$ for which there exists a path $w_1 \cdots w_\ell$ with $w_2 \in I_B \cup M^{(\ell-2)}(A)$. We claim that

$$|T_{\Sigma^*,\ell-3}(i) \setminus Z_{\Sigma^*,\ell-3}(i)| \leqslant |Q_{\Sigma^*}(i)|. \tag{7.118}$$

Every tuple $(v_0, \dots, v_{\ell-2}) \in T_{\Sigma^*,\ell-3}(i) \setminus Z_{\Sigma^*,\ell-3}(i)$ was ignored in one of the first i steps because (R2) failed. Recall that $C_{x,y,\Sigma^*}(i,j)$ contains all pairs $bw \in B \times N^{(\ell-3)}(A, R)$ for which there exist disjoint paths $b = w_1 \cdots w_j = x$ and $y = w_{j+1} \cdots w_\ell = w$ in $G(i)$. Observe that for every ignored tuple there exists $i' < i$, distinct $x, y \in [n]$ and $j \in [\ell - 1]$ with $e_{i'+1} = xy$, $f_{\ell-2} \in C_{x,y,\Sigma}(i', j)$ and $|C_{x,y,\Sigma}(i', j)| > p^{-1}n^{-30\ell\varepsilon}$. So, since $e_{i'+1} = xy$ was added, for every such tuple there exists a path $v_{\ell-2} = w_1 \cdots w_j w_{j+1} \cdots w_\ell = v_{\ell-3}$ with $w_j = x$ and $w_{j+1} = y$ in $G(i'+1) \subseteq G(i)$. Note that by monotonicity we have $C_{x,y,\Sigma^*}(i', j) \subseteq C_{x,y,\Sigma^*}(i, j)$, and therefore all such 'bad' pairs xy satisfy $|C_{x,y,\Sigma^*}(i, j)| > p^{-1}n^{-30\ell\varepsilon}$. By the findings of Section 7.8.2.2 it thus suffices to consider $C_{x,y,\Sigma^*}(i, j)$ for $xy \in L_{\Sigma^*}(i, j)$ with

$j \in \{1,2\}$, since for all others (7.112) holds. Now, using (7.114) and (7.115), it is not difficult to see that the corresponding paths $v_{\ell-2} = w_1 \cdots w_\ell = v_{\ell-3}$ satisfy $w_1 \in B$, $w_2 \in I_B \cup M^{(\ell-2)}(A)$ and $w_\ell \in N^{(\ell-3)}(A,R)$. Putting things together, the extension property $\mathcal{U}_T$ (cf. Lemma 7.4.1) implies (7.118), since every $(v_0,\ldots,v_{\ell-2}) \in T_{\Sigma^*,\ell-3}(i) \setminus Z_{\Sigma^*,\ell-3}(i)$ is uniquely determined by the pair $f_{\ell-2} = v_{\ell-3}v_{\ell-2}$.

Let $Q_{\Sigma^*,I}(i)$ and $Q_{\Sigma^*,M}(i)$ contain all pairs $(w_1,w_\ell) \in Q_{\Sigma^*}(i)$ where at least one corresponding path $w_1 \cdots w_\ell$ satisfies $w_2 \in I_B$ and $w_2 \in M^{(\ell-2)}(A) \setminus I_B$, respectively. Now, using (7.48) and (7.118), to establish (7.55), it suffices to prove, say,

$$\max\{|Q_{\Sigma^*,I}(i)|, |Q_{\Sigma^*,M}(i)|\} \leqslant (np)^{\ell-1} n^{-15\varepsilon}. \tag{7.119}$$

Using Lemma 7.8.2, $|I_B| \leqslant n^{2\vartheta\varepsilon}$ and that the degree is at most npn^ε, we obtain, with room to spare,

$$|Q_{\Sigma^*,I}(i)| \leqslant npn^\varepsilon \cdot |I_B| \cdot (np)^{\ell-3} n^{15\ell\varepsilon} \leqslant (np)^{\ell-2} n^{(15\ell+2\vartheta+1)\varepsilon} \leqslant (np)^{\ell-1} n^{-15\varepsilon}.$$

Turning to $Q_{\Sigma^*,M}(i)$, note that for every $w_2 \in M^{(\ell-2)}(A) \setminus I_B$ we have $|\Gamma(w_2) \cap B| \leqslant npn^{-\vartheta\varepsilon}$ by (7.110). With a similar argument as above, using Lemma 7.8.4, i.e., $|M^{(\ell-2)}(A)| \leqslant n^{2\ell\tau\varepsilon}$, we see that

$$|Q_{\Sigma^*,M}(i)| \leqslant npn^{-\vartheta\varepsilon} \cdot |M^{(\ell-2)}(A)| \cdot (npn^\varepsilon)^{\ell-2} \leqslant (np)^{\ell-1} n^{(2\ell\tau+\ell-\vartheta)\varepsilon} \leqslant (np)^{\ell-1} n^{-15\varepsilon},$$

where the last inequality follows from (7.105), i.e., $\vartheta = 20\ell\tau$. This establishes (7.119), which, as explained, completes the proof of Lemma 7.4.3. $\square$

Bibliography

[1] D. Achlioptas, R.M. D'Souza, and J. Spencer. Explosive percolation in random networks. *Science* **323** (2009), 1453–1455.

[2] M. Aizenman and D.J. Barsky Sharpness of the phase transition in percolation models. *Comm. Math. Phys.* **108** (1987), 489–526.

[3] D. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5** (1999), 3–48.

[4] N. Alon, M. Krivelevich, and B. Sudakov. Coloring graphs with sparse neighborhoods. *J. Combin. Theory Ser. B* **77** (1999), 73–82.

[5] J.S. Andrade, H.J. Herrmann, A.A. Moreira, and C.L.N. Oliveira. Transport on exploding percolation clusters. *Phys. Rev. E* **83** (2011), 031133.

[6] T.M. Apostol. An elementary view of Euler's summation formula. *Amer. Math. Monthly* **106** (1999), 409–418.

[7] N.A.M. Araújo, J.S. Andrade, R.M. Ziff, and H.J. Herrmann. Tricritical point in explosive percolation. *Phys. Rev. Lett.* **106** (2011), 095703.

[8] Y. Azar, A.Z. Broder, A.R. Karlin, and E. Upfal. Balanced allocations. *SIAM J. Comput.* **29** (1999), 180–200.

[9] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science* **286** (1999), 509–512.

[10] A. Beveridge, T. Bohman, A. Frieze, and O. Pikhurko. Memoryless rules for Achlioptas processes. *SIAM J. Discrete Math.* **23** (2009), 993–1008.

[11] S. Bhamidi, A. Budhiraja, and X. Wang. Bohman–Frieze processes at criticality and emergence of the giant component. Preprint (2011). arXiv:1106.1022.

[12] T. Bohman. Emergence of connectivity in networks. *Science* **323** (2009), 1438–1439.

[13] T. Bohman. The triangle-free process. *Adv. Math.* **221** (2009), 1653–1677.

[14] T. Bohman and A. Frieze. Avoiding a giant component. *Random Struct. Alg.* **19** (2001), 75–85.

[15] T. Bohman, A. Frieze, M. Krivelevich, P.-S. Loh, and B. Sudakov. Ramsey games with giants *Random Struct. Alg.* **38** (2011), 1–32.

[16] T. Bohman, A. Frieze, and N.C. Wormald. Avoidance of a giant component in half the edge set of a random graph. *Random Struct. Alg.* **25** (2004), 432–449.

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

[18] T. Bohman and J.H. Kim. A phase transition for avoiding a giant component. *Random Struct. Alg.* **28** (2006), 195–214.

[19] T. Bohman and D. Kravitz. Creating a giant component. *Combin. Probab. Comput.* **15** (2006), 489–511.

[20] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** (1984), 257–274.

[21] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge Univ. Press (2001).

[22] B. Bollobás. Personal communication (2010).

[23] B. Bollobás. Personal communication (2011).

[24] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** (2007), 3–122.

[25] B. Bollobás and O. Riordan. Constrained graph processes. *Electron. J. Combin.* **7** (2000), #R18.

[26] W. Chen and R.M. D'Souza. Explosive percolation with multiple giant components. *Phys. Rev. Lett.* **106** (2011), 115701.

[27] Y.S. Cho and B. Kahng. Suppression effect on explosive percolation. *Phys. Rev. Lett.* **107** (2011), 275703.

[28] Y.S. Cho, J. S. Kim, J. Park, B. Kahng, and D. Kim. Percolation transitions in scale-free networks under the achlioptas process. *Phys. Rev. Lett.* **103** (2009), 135702.

[29] Y.S. Cho, S.W. Kim, J.D. Noh, B. Kahng, and D. Kim. Finite-size scaling theory for explosive percolation transitions. *Phys. Rev. E* **82** (2010), 042102.

[30] F. Chung and R. Graham. *Erdős on Graphs: His Legacy of Unsolved Problems.* A K Peters, Ltd., Massachusetts (1998).

[31] R.A. da Costa, S.N. Dorogovtsev, A.V. Goltsev, and J.F.F. Mendes. Explosive percolation transition is actually continuous. *Phys. Rev. Lett.* **105** (2010), 255701.

[32] R.M. D'Souza and M. Mitzenmacher. Local cluster aggregation models of explosive percolation. *Phys. Rev. Lett.* **104** (2010), 195702.

[33] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960), 17–61.

[34] P. Erdős, S. Suen, and P. Winkler. On the size of a random maximal graph. *Random Struct. Alg.* **6** (1995), 309–318.

[35] A.D. Flaxman, D. Gamarnik, and G.B. Sorkin. Embracing the giant component. *Random Struct. Alg.* **27** (2005), 277–289.

[36] E.J. Friedman and A.S. Landsberg. Construction and analysis of random networks with explosive percolation. *Phys. Rev. Lett.* **103** (2009), 255701.

[37] P. Grassberger, C. Christensen, G. Bizhani, S.-W. Son, and M. Paczuski. Explosive percolation is continuous, but with unusual finite size behavior. *Phys. Rev. Lett.* **106** (2011), 225701.

[38] A. Gut. *Probability: A Graduate Course.* Springer, New York (2005).

[39] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** (1963), 13–30.

[40] H. Hooyberghs and B. Van Schaeybroeck. Criterion for explosive percolation transitions on complex networks. *Phys. Rev. E* **83** (2011), 032101.

[41] S. Janson. Networking – Smoothly does it. *Science* **333** (2011), 298–299.

[42] S. Janson, D.E. Knuth, T. Łuczak and B. Pittel. The birth of the giant component. *Random Struct. Alg.* **4** (1993), 231–358.

[43] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs.* Wiley-Interscience, New York (2000).

[44] S. Janson and A. Ruciński. The deletion method for upper tail estimates. *Combinatorica* **24** (2004), 615–640.

[45] S. Janson and J. Spencer. Phase transitions for modified Erdős-Rényi processes. *Ark. Math.*, to appear. `arXiv:1005.4494`.

[46] M. Kang, W. Perkins, and J. Spencer. The Bohman–Frieze process near criticality. *Random Struct. Alg.*, to appear. `arXiv:1106.0484`.

[47] J.H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Struct. Alg.* **7** (1995), 173–207.

[48] J.H. Kim and V.H. Vu. Divide and conquer martingales and the number of triangles in a random graph. *Random Struct. Alg.* **24** (2004), 166–174.

[49] M. Krivelevich, P.-S. Loh, and B. Sudakov. Avoiding small subgraphs in Achlioptas processes. *Random Struct. Alg.* **34** (2009), 165–195.

[50] M. Krivelevich, E. Lubetzky, and B. Sudakov. Hamiltonicity thresholds in Achlioptas processes. *Random Struct. Alg.* **37** (2010), 1–24.

[51] H.K. Lee, B.J. Kim, and H. Park. Continuity of the explosive percolation transition. *Phys. Rev. E* **84** (2011), 020101.

[52] S. Manna and A. Chatterjee. A new route to explosive percolation. *Physica A* **390** (2011), 177–182.

[53] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics*, London Math. Soc. Lecture Note Ser. **101**, pp. 148–188. Cambridge Univ. Press (1989).

[54] M.V. Men'shikov, Coincidence of critical points in percolation problems. *Dokl. Akad. Nauk SSSR* **288** (1986), 1308–1311.

[55] T. Mütze, R. Spöhel, and H. Thomas. Small subgraphs in random graphs and the power of multiple choices. *J. Combin. Theory Ser. B* **101** (2011), 237–268.

[56] J. Nagler, A. Levina, and M. Timme. Impact of single links in competitive percolation. *Nature Physics* **7** (2011), 265–270.

[57] D. Osthus and A. Taraz. Random maximal H-free graphs. *Random Struct. Alg.* **18** (2001), 61–82.

[58] M.E. Picollelli. The final size of the C_4-free process. *Combin. Probab. Comput.* **20** (2011), 939–955.

[59] M.E. Picollelli. The final size of the C_ℓ-free process. Preprint (2011). Available at `http://sites.google.com/site/mepicollelli/`.

[60] F. Radicchi and S. Fortunato. Explosive percolation in scale-free networks. *Phys. Rev. Lett.* **103** (2009), 168701.

[61] F. Radicchi and S. Fortunato. Explosive percolation: A numerical analysis. *Phys. Rev. E* **81** (2010), 036110.

[62] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science* **333** (2011), 322–324.

[63] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.*, to appear. `arXiv:1102.5306`.

[64] O. Riordan and L. Warnke. Achlioptas processes are not always self-averaging. Preprint (2011). `arXiv:1111.6177`.

[65] O. Riordan and L. Warnke. Convergence of achlioptas processes via differential equations with unique solutions. Preprint (2011). `arXiv:1111.6179`.

[66] O. Riordan and L. Warnke. The evolution of subcritical achlioptas processes. Preprint (2012). `arXiv:1204.5068`.

[67] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.* **8** (1995), 917–942.

[68] A. Ruciński and N.C. Wormald. Random graph processes with degree restrictions. *Combin. Probab. Comput.* **1** (1992), 169–180.

[69] T.G. Seierstad. A central limit theorem via differential equations *Ann. Appl. Probab.* **19** (2009), 661–675.

[70] J. Spencer. Counting extensions. *J. Combin. Theory Ser. A* **55** (1990), 247–255.

[71] J. Spencer. Maximal triangle-free graphs and Ramsey $R(3,t)$. Unpublished manuscript (1995). Available at `http://cs.nyu.edu/spencer/papers/ramsey3k.pdf`.

[72] J. Spencer. Potpourri. *J. Combinatorics* **1** (2010), 237–264.

[73] J. Spencer. Personal communication (2011).

[74] J. Spencer and N.C. Wormald. Birth control for giants. *Combinatorica* **27** (2007), 587–628.

[75] L. Warnke. Dense subgraphs in the H-free process. *Discrete Math.* **311** (2011), 2703–2707.

[76] L. Warnke. When does the K_4-free process stop? *Random Struct. Alg.*, to appear. `arXiv:1007.3037`.

[77] L. Warnke. The C_ℓ-free process. *Random Struct. Alg.*, to appear. arXiv:1101.0693.

[78] G. Wolfovitz. Lower bounds for the size of random maximal H-free graphs. *Electron. J. Combin.* **1** (2009), #R16.

[79] G. Wolfovitz. The K_4-free process. Preprint (2010). arXiv:1008.4044.

[80] N.C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** (1995), 1217–1235.

[81] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on approximation and randomized algorithms*, pp. 73–155. PWN, Warsaw (1999).

[82] R.M. Ziff. Explosive growth in biased dynamic percolation on two-dimensional regular lattice networks. *Phys. Rev. Lett.* **103** (2009), 045701.

[83] R. M. Ziff. Scaling behavior of explosive percolation on the square lattice. *Phys. Rev. E* **82** (2010), 051105.