M.Sc. B.Sc. Christoph Koch

Phase transition phenomena in random graphs and hypergraphs

DOCTORAL THESIS

to achieve the university degree of

Doktor der Naturwissenschaften

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Mihyun Kang, Ph.D.

Institute of Discrete Mathematics

Graz, September 2016

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date

Signature

Contents

List of publications

Journal articles

[1] B. Bollobás, O. Cooley, M. Kang, and C. Koch. 'Jigsaw percolation on random hypergraphs'. *Submitted* (2016), 15 pages. arXiv: 1603.07883.

[2] O. Cooley, M. Kang, and C. Koch. 'The size of the giant component in random hypergraphs'. *Submitted* (2015), 48 pages. arXiv: 1501.07835.

[3] O. Cooley, M. Kang, and C. Koch. 'Threshold and hitting time for high-order connectivity in random hypergraphs'. *Electron. J. Combin.* **23** (2016), no. 2, P48.

[4] N. Fountoulakis, M. Kang, C. Koch, and T. Makai. 'A phase transition on the evolution of bootstrap processes in inhomogeneous random graphs'. *Submitted* (2016), 44 pages. arXiv: 1609.08892.

[5] M. Kang, M. Karoński, C. Koch, and T. Makai. 'Properties of stochastic Kronecker graphs'. *J. Comb.* **6** (2015), no. 4, 395–432.

[6] M. Kang, C. Koch, and A. Pachón. 'The phase transition in multitype binomial random graphs'. *SIAM J. Discrete Math.* **29** (2015), no. 2, 1042–1064.

[7] C. Koch and J. Lengler. 'Bootstrap percolation on geometric inhomogeneous random graphs'. *Submitted* (2016), 32 pages. arXiv: 1603.02057.

Conference proceedings

[8] O. Cooley, M. Kang, and C. Koch. 'Evolution of high-order connected components in random hypergraphs'. *Electronic Notes in Discrete Mathematics* **49** (2015), 569–575.

[9] M. Kang, C. Koch, and T. Makai. 'Bootstrap percolation in random k-uniform hypergraphs'. *Electronic Notes in Discrete Mathematics* **49** (2015), 595–601.

[10] C. Koch and J. Lengler. 'Bootstrap percolation on geometric inhomogeneous random graphs'. In: *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*. Vol. 55. LIPIcs. Schloss Dagstuhl. Leibniz-Zent. Inform., 2016, 147:1–147:15.

CHAPTER 1

Introduction

The first few pages of this thesis provide a gentle and informal introduction to its topic targeted at a general audience. No background in mathematics or any related fields is assumed, and statements in Section 1.1 are deliberately kept vague aiming at a more intuitive understanding.

This being said, experts in related scientific areas might want to skip Section 1.1 and proceed directly to Section 1.2. There we highlight the main results obtained in this thesis, followed by a concise description of key techniques and discussions of related open problems in Sections 1.3 and 1.4, respectively. We then conclude Chapter 1 with an outline of the remainder of the thesis.

1.1. An even briefer history of random graphs[1]

> Dear reader,
>
> please find a comfortable seat, help yourself to a drink of your choice – alcoholic or not – and imagine yourself at a bar or party asking the notorious question: 'So tell me, what is the topic of your Ph.D. studies?' If you are addressing me and are willing to spend a couple of minutes of your precious time, the answer might somewhat resemble the explanation on the following few pages. That is, of course, except for a lot of smiling, which unfortunately cannot be adequately transferred to the medium in your hands.

So what is a *graph*? Well, let us modify the question slightly and ask: what is a *social network*? A social network consists of a group of users, and pairs of them who decided to become 'friends' so that they can share the happy (and sad) moments of their lives with one another. Once we forget about everything else, and represent each user by a *vertex* (or a point), and each pair of friends by an *edge* (or a line) we obtain a *graph*.[2] Clearly this concept of graphs is very abstract, and thus there are various other real-world networks which we also regard as graphs. For example the network of all actors where any two of them are linked if they both play a role in the same film; a financial market when mapping the participants and their dependencies; or the brain when looking at neurons and synapses.

[1] Inspired by the book title 'A Briefer History of Time' by Stephen Hawking and Leonard Mlodinow [69].

[2] So in particular the graphs in this thesis are not to be confused with the *graph of a function*, which you may be familiar with from high-school.

Classical graph theory would then ask questions such as: is the graph connected, i.e. starting at a given vertex can we reach any other vertex by moving along edges? Or, is it possible to colour each vertex either red or blue such that all edges contain precisely one vertex of each colour? What is the maximal number of vertices such that any two of them form an edge? Is it possible to draw the graph on a piece of paper without having two edges intersect (apart from possibly their endpoints)? Under which conditions can we match every vertex to one of its neighbours so that every vertex gets matched precisely once?

Investigating any of these will provide us with a (slightly) better understanding of the graph and its structure. However, giving actual answers might take a long time if the graph in question has a huge number of vertices. To make matters even worse, we typically would have to perform an immense amount of such calculations in order to be able to tell our users 'You might also know these people: [...]' followed by a stunningly accurate list of people you actually know but might have lost touch with. But a thorough understanding of the general structure of the network may help simplify and speed up our answers. In other words, we would like to have a model for the network which is comparatively easy to analyse (mathematically) yet retains the general properties of the network.

Now finally, this is where random graphs join the fray. Arguably the simplest model of a random graph is the following: we fix a set of (distinguishable) vertices, and then for each two of them we toss a coin; if it comes up heads the edge is present, if it comes up tails the edge is not present. After having tossed all coins we obtain a graph that was chosen randomly.[3]

But what use is something random to make any predictions? If we roll a die (in a fair way), we cannot forecast what number it will show, so is this not very chaotic? As a matter of fact it is not: assuming that we throw a large enough number of dice, we can confidently predict that the number of ones will be very close to one sixth of all die rolls.

In fact this observation is somewhat a paradigm for random graphs in general: graph parameters – both basic ones such as the number of edges or significantly more involved ones like the minimal number of colours we need such that there is no edge whose end points have the same colour – are often close to what we should expect them to be. Hence the art of analysing random graphs lies in finding the right intuition of what should be true, and then rigorously proving that the paradigm actually holds in the particular case at hand.

Phase transitions. So what happens with properties that can either be satisfied or unsatisfied, such as for example the property of the graph being connected? If we use a fair coin (in the previously described model), then the answer is actually quite simple: it is extremely likely to be connected, the reason being that the number of edges it contains is very large.

[3]from the class of all graphs having this set of vertices

However, once we replace our coin with one which is heavily biased towards tails (say it comes up heads with some tiny probability $p > 0$), we observe a far more intriguing behaviour! For this let us think about the largest component, i.e. the biggest set of vertices so that for every pair of them we may walk from one to the other along edges. It turns out that there is a critical value $\hat{p}$ with the following properties: if p is smaller than $\hat{p}$ it is very likely that the largest component is insignificantly small, but if p is larger than $\hat{p}$ typically there is a unique component which contains a substantial fraction of all vertices. This is actually a fundamental result from 1960 due to two famous Hungarian mathematicians, Paul Erdős and Alfréd Rényi. In fact, they published a series of papers on similar topics, which are widely seen as the foundation of random graph theory.[4]

We call this type of phenomenon a *phase transition*, where the terminology is inspired by different phases of matter known in physics: solid, liquid, gas, and plasma. (Incidentally this is also where random graphs play an important role in statistical physics as models for systems of interacting particles.)

Component structure of random graphs. With this we have now reached one of the core topics of this thesis. The first half of the results describe key features of the component structure of some more general random graph models. At first we study a model where not every vertex looks the same in the sense that some of them are more likely to have a larger number of neighbours, this is often modelled by each vertex having its individual *weight*. We describe precisely how fast the largest component grows once it is *born*, i.e. immediately after the phase transition. Whereas this is certainly not the strongest result in this thesis, it is still of profound importance as a preparation for following results.

Mastering the methods used in its proof allows us to prove the second result: we determine the size of the largest component of random *hypergraphs* in the very delicate regime where the giant component is born. A hypergraph in itself is similar to a graph, it consist of a set of vertices and edges. However now edges may also consist of more than two vertices. At first glance, this may not sound like a big deal, but most importantly there is a large number of 'natural' yet genuinely different notions of connectedness in hypergraphs. In many cases studying the component structure becomes very challenging and profoundly novel concepts are necessary.

In our results the crucial observation was the following: components grow 'smoothly'. This concept is arguably the most fundamental contribution of this thesis, and will presumably pave the way for many results on random hypergraphs in future. One example is our third result: we characterise precisely when the entire hypergraph becomes connected by exploiting the smoothness of a substantial part of the hypergraph. So far our results were largely of intrinsic mathematical interest, in the second half of the thesis we return to problems more directly motivated by real-world networks.

[4]So random graphs were already of deep mathematical interest long before computers became available to the general public, let alone social networks, in the modern sense, existed.

Percolation processes. Imagine for instance that we would like to model the spread of some infection (or a rumour, or some neuronal activity) in a real-world network. Mathematically we think of these as so-called *percolation processes* on the components of a random graph: whenever the neighbourhood of a given vertex satisfies some condition, for instance, it contains at least $r \geq 2$ infected vertices, then the vertex also becomes infected.[5] We keep iterating until no more vertices become infected. There may be two different reasons for this: either there is a *global outbreak* meaning that every vertex became infected, or none of the still uninfected vertices has sufficiently many infected neighbours to become infected as well.

It turns out that, more often than not, these processes have a threshold behaviour similar to the phase transition phenomena we studied earlier: by a tiny increase in some parameter of the random graph model the probability of an outbreak rises from being almost negligible to being very close to one.

While for classical random graph models many variations of percolation processes are well-studied, things become very difficult on random graphs with properties similar to typical real-world networks, even for the most basic percolation processes as the one described previously. By now it is well-known that many of these networks share some key features, one of them being *clustering*: people are much more likely to know a friend of their friends than to know an average person with a similar profile. Random graphs with an underlying geometry, i.e. every vertex also has a position and the further two vertices are apart the less likely they form an edge, tend to exhibit this feature. Alas, studying these kind of models is very challenging and to date rather little is known about them.

We investigate the spread of an infection (modelled as bootstrap percolation) on a recently introduced geometric random graph. In particular, we describe in detail under which circumstances and how an infection starting within a small local region will gradually spread along the geometry and in the end cause a global outbreak. Interestingly, viewing things from a more applied side, this information could be used to locally isolate the infection and thereby save most of the individuals of the network.

This result should be seen as one of many steps which will be necessary to bridge the gap between mathematically tractable graph and percolation models on the one hand, and those which are actually close enough to real-world networks but are notoriously difficult to analyse rigorously, on the other.

There are also interesting percolation processes having a slightly different flavour, one example being the recent model of *jigsaw percolation* on a pair of graphs. We think of a group of individuals, each having a certain (unique) piece of a jigsaw puzzle. The first graph has a vertex for each individual and two individuals meet if and only if they are adjacent, while the second graph encodes the information whether the puzzle pieces of any two individuals fit together. Whenever two individuals meet and have compatible puzzle pieces they share these. The main

[5] In the literature, this example is called *bootstrap percolation*.

question is whether in the end the puzzle has been completed. The general idea is that this process could help us understand how co-operative efforts of large groups of individuals give rise to great and ingenious collective ideas.

However, also from a purely mathematical perspective this process is of profound interest. It may be seen as a notion of several graphs being 'simultaneously connected'. With this perspective it is very natural to study the process on a pair of random graphs, and it was shown that this percolation process also exhibits a phase transition phenomenon. We demonstrate how this behaviour generalises in two directions: first of all for various notions of connectedness in hypergraphs, and secondly for larger numbers of (hyper-)graphs.

Summary. Random graphs are relevant to various scientific areas ranging from social networks to biomedical/neurological applications and also statistical physics. Studying properties of random graphs on an abstract level is a challenging task which is full of profound mathematical insights. The thesis covers four main topics centred on phase transition phenomena, which we investigate in great detail. The focus lies on the component structure of random graphs and hypergraphs and the behaviour of percolation processes in a random setting.

1.2. Main results

How do structural properties of random graphs and hypergraphs change under small alterations of their parameters? In this thesis we investigate this question and primarily focus on regimes where seemingly negligible alterations cause drastic differences in the behaviour of the model. Loosely speaking, these regimes are characterised by threshold functions: if the parameter is 'smaller' than the threshold, then with high probability (*whp* for short)[6] the random (hyper-)graph does not yet have a certain property; while if the parameter is 'larger' than the threshold, then whp the random (hyper-)graph has this property. Depending on the meaning of 'small' and 'large' we call the threshold either sharp or coarse.

This type of phenomenon is also known as a phase transition: one of the first and arguably the most famous phase transition is that of the giant component. In 1960 Erdős and Rényi [58] showed that the *size* of the largest component, i.e. its number of vertices, in the binomial random graph $G(n,p)$ exhibits a 'double jump' from logarithmic, to polynomial (but still sub-linear), and then linear order.[7] However, in contrast to initial beliefs, this was only a rather incomplete picture of what later became known as the *birth of the giant component*, a term coined by Janson, Łuczak, Knuth, and Pittel [73], who extended results by Bollobás [31] and Łuczak [86]. When the parameter p passes the (sharp) threshold n^{-1}, a unique largest component forms, which then starts growing gradually and very soon reaches linear size, turning into a giant component. It then continues to merge with further

[6]Meaning with probability tending to one as $n \to \infty$, with n being the number of vertices in the random (hyper-)graph. Unspecified asymptotics are with respect to $n \to \infty$.

[7]Their proof is given for the uniformly chosen random graph $G(n,m)$. However due to the *asymptotic equivalence* of these two models (Section 1.4 in [74]), the stated result follows.

components, until p reaches the (sharp) threshold for connectedness $n^{-1}\log n$. In fact, the binomial random graph $G(n,p)$ – seen as a random graph process[8] – becomes connected precisely when the last isolated vertex disappears, as was proven by Bollobás and Thomason [38].

In the thesis we determine the thresholds for a number of properties related to the component structure of various random (hyper-)graphs: from the birth of the giant component via connectedness to percolation processes describing a strengthened or simultaneous notion of connectedness. In the process we develop new concepts and techniques providing a detailed picture of these random hypergraphs from a close-up perspective (Section 1.3). Naturally, these insights open up novel directions for research and give rise to intriguing open problems (Section 1.4).

1.2.1. **Giant component in multi-type random graphs.** In recent years the study of random graphs is largely motivated by their applications as models of real-world networks. It is hardly surprising that the most basic models such as $G(n,p)$ and $G(n,m)$ fall short in this task: they are entirely homogeneous and contain very few vertices of large degrees, which would model 'hubs' of the real-world network properly. Thus, we investigate a random graph model, denoted by $G(\mathbf{n},P)$, exhibiting a certain degree of inhomogeneity between its vertices.

In $G(\mathbf{n},P)$ there are n labelled vertices, where each vertex has one of two types, and any two vertices of types i and j form an edge independently with probability $p_{i,j}$ (encoded as an entry of the (symmetric) parameter matrix $P \in (0,1]^{2\times 2}$). The number of vertices of type i is denoted by n_i, i.e. the i-th coordinate of the parameter vector $\mathbf{n}$, and we assume that $n_1 \geq n_2 \to \infty$. Our main contribution lies in taking a very close look at the delicate regime in which the giant component gradually emerges.

Theorem 1.1 (Theorems 2.1 and 2.4)**.** *Let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon = o(1)$. Then the following holds whp:*

(a) if $\varepsilon^3 n_2 \min\{1, \varepsilon^{-1}p_{2,1}n_1\} \to \infty$ and

$$p_{j,1}n_1 + p_{j,2}n_2 = 1 + \varepsilon \pm o(\varepsilon), \quad \forall\, j \in \{1,2\}, \tag{1.1}$$

then the largest component contains $(2 \pm o(1))\varepsilon n_i$ vertices of type i, while all other components contain only $o(\varepsilon n_i)$ such vertices for any $i \in \{1,2\}$;

(b) if $\varepsilon^3 n_2 \to \infty$ and

$$p_{j,1}n_1 + p_{j,2}n_2 = 1 - \varepsilon \pm o(\varepsilon), \quad \forall\, j \in \{1,2\}, \tag{1.2}$$

then the largest component contains at most $o(n^{2/3})$ vertices.

Note that even a very small number of vertices of type 2 can significantly alter the behaviour of the model (when $\varepsilon = o(1)$): let $n_2 = \sqrt{\varepsilon}n_1$, $\varepsilon^3 n_2 \to \infty$, $p_{2,1}n_1 = 1$, and thus $p_{1,1}n_1 = 1 - \sqrt{\varepsilon} + \varepsilon \pm o(\varepsilon)$, so that (1.1) is satisfied. Then

[8] We associate each edge with a *birth time* chosen independently and uniformly from $[0,1]$. Then the edges are added one by one to the initially empty graph in their (unique) order of birth times.

$G(\mathbf{n}, P)$ has a unique largest component containing approximately $2\varepsilon n_1 = \omega(n_1^{5/7})$ vertices of type 1, but after removing all vertices of type 2 all components have size at most $o(n_1^{2/3})$. Qualitatively, this setting allows us to discriminate between hubs and ordinary vertices in a way which decisively influences the random graph. Notably, a similar behaviour has not yet been observed within the closely related framework of inhomogeneous random graphs introduced by Bollobás, Janson, and Riordan [33]. For more details see Section 2.1.4.

Observe that the conditions in (1.1) and (1.2) are more restrictive than we would wish for. While the condition in (1.1) implies that the Frobenius eigenvalue λ of the matrix $M := P \cdot \mathrm{diag}(\mathbf{n})$, which encodes the expected degrees of the vertices, satisfies $\lambda = 1 + \varepsilon \pm o(\varepsilon)$, the converse is clearly not true. However, the branching processes heuristic suggests that, even when $\varepsilon = o(1)$, the sub- and supercritical regimes should be characterised by λ. We will pick up on this line of thought (and a generalisation to an arbitrary number of types) in Section 1.4.1.

The proof is largely based on coupling a component exploration process with a multi-type Galton-Watson branching process. We highlight some of these coupling techniques in Section 1.3.1

1.2.2. **High-order connected components in random hypergraphs.** A large part of this thesis concentrates on investigating the component structure of random k-uniform hypergraphs. Despite this being one of the aspects in which hypergraphs demonstrate their richness in structure and appealing beauty, as yet this topic has only been studied for the most simple notions of connectedness.

In the following we investigate an entire class of these notions called *high-order connectedness*: given an integer $k \geq 2$ and an integer j satisfying $1 \leq j \leq k-1$, we consider j-sets (j-tuples of distinct vertices), and say that two distinct j-sets J and J' are *j-connected* if there is a sequence of edges $e_1, \ldots, e_m$ such that

- $J \subset e_1$ and $J' \subset e_m$,
- $|e_i \cap e_{i+1}| \geq j$ for all $1 \leq i \leq m-1$.

Any j-set is j-connected to itself. This defines an equivalence relation on the set of all j-sets, and we refer to its equivalence classes *j-components*.

High-order connected components exhibit a significantly more diverse behaviour than their vertex-connected counterparts ($j = 1$). For instance, the collection of j-components of a k-uniform hypergraph does not induce a partition of its vertex set (nor on the set of ℓ-sets for any $1 \leq \ell \leq j-1$). In particular this means that for a pair of components C_1 and C_2 the number of k-sets containing at least one j-set in C_1 and C_2, each, is highly non-trivial for $j \geq 2$. In fact, this turns out to be the major challenge when analysing the size of the largest high-order connected component in the binomial random k-uniform hypergraph $\mathcal{H}^k(n,p)$ using a branching process approach.

In the process we obtain a remarkable insight: the larger a component grows, the more evenly it is distributed over all ℓ-sets, for any $1 \leq \ell \leq j-1$. In other

words, with exponentially high probability, each ℓ-set is contained in approximately the same number of j-sets of any sufficiently large j-component in $\mathcal{H}^k(n,p)$. This result is called the *smooth boundary lemma* (Lemma 1.9 in Section 1.3.2) and it has already proven to be a powerful tool. The underlying concept of *smoothness* will be discuss in detail in Section 1.3.2.

With the help of the smooth boundary lemma we analyse the size of the largest high-order connected component from a close-up perspective.

Theorem 1.2 (Theorem 3.2)**.** *For any $1 \le j \le k-1$ we set*

$$\hat{p}_{\mathrm{g}} = \hat{p}_{\mathrm{g}}(n,k,j) := \tfrac{1}{\binom{k}{j}-1}\tfrac{1}{\binom{n}{k-j}}, \tag{1.3}$$

and let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n^{1-\delta} \to \infty$, for some constant $\delta > 0$.

(a) If $p = (1+\varepsilon)\hat{p}_{\mathrm{g}}$, then whp the size of the largest j-component of $\mathcal{H}^k(n,p)$ satisfies $(1 \pm o(1))\frac{2\varepsilon}{\binom{k}{j}-1}\binom{n}{j}$ while all other j-components have size at most $o(\varepsilon n^j)$.

(b) If $p = (1-\varepsilon)\hat{p}_{\mathrm{g}}$, then whp all j-components of $\mathcal{H}^k(n,p)$ have size at most $O(\varepsilon^{-2}\log n)$.

Previously it was already shown that $\hat{p}_{\mathrm{g}}$ defined in (1.3) is the threshold for the appearance of the giant j-component in $\mathcal{H}^k(n,p)$ by Cooley, Kang, and Person [51] and independently by Lu and Peng [84]. While the first group already determined the size of the largest j-component up to a multiplicative constant even when $\varepsilon = o(1)$, the second group studied only the simpler regime when $\varepsilon > 0$ is a constant, although they also provide the leading constant of the size of the largest j-component.

While the proof for the supercritical regime of Theorem 1.2 is based on the smooth boundary lemma (Lemma 1.9), the analysis of the subcritical regime is a simple application of an idea by Krivelevich and Sudakov [82].

Even though we develop the concept of smoothness in the regime barely above the threshold $\hat{p}_{\mathrm{g}}$, it proves to be a powerful tool at later stages of the evolution of $\mathcal{H}^k(n,p)$ and also for the random k-uniform hypergraph process $\{\mathcal{H}^k(n,M)\}_M$. We write τ_c for the hitting time for j-connectedness of $\{\mathcal{H}^k(n,M)\}_M$, and denote the moment when the last isolated j-set disappears by τ_i. We prove that whp both hitting times, τ_c and τ_i, coincide, thus extending the classical result by Bollobás and Thomason [38].

Theorem 1.3 (Theorem 4.1)**.** *For any $1 \le j \le k-1$ whp we have $\tau_c = \tau_i$.*

Special cases of Theorem 1.3 were already proved by Poole [96] for $j = 1$, and by Kahle and Pittel [77] for $j = k-1$.

Furthermore, by coupling $\mathcal{H}^k(n,p)$ and $\{\mathcal{H}^k(n,M)\}_M$ via birth times, we use Theorem 1.3 to prove Theorem 4.3 stating that the property of $\mathcal{H}^k(n,p)$ being j-connected has a sharp threshold given by

$$\hat{p}_{\mathrm{c}} = \hat{p}_{\mathrm{c}}(n,k,j) := \frac{j\log n}{\binom{n}{k-j}}. \tag{1.4}$$

As part of Section 1.3.2, which is dedicated to the notion of smoothness, we will highlight the importance of this concept for the proofs of Theorems 1.2 and 1.3. As mentioned earlier, many classical questions concerning the j-component structure of random hypergraphs are still widely open. We briefly discuss some of these in Section 1.4.2. We expect the notion of smoothness to prove to be a widely-applicable and important tool for future advances in the theory of random hypergraphs, as in the case of Theorem 1.3.

1.2.3. **Jigsaw percolation: simultaneous connectedness.** Recently, Brummitt, Chatterjee, Dey, and Sivakoff [43] proposed a mathematical model, called *jigsaw percolation*, which aims to explain how a collective creative process of the individuals in a social network can achieve extraordinary results, such as a major scientific breakthrough. The scenario is the following: each individual has a (unique) piece of a jigsaw puzzle, and whenever two groups of individuals meet and have compatible puzzle pieces they share these. The question is whether the individuals can collaboratively reconstruct the complete puzzle.

We take a more purely mathematical view: consider two random graphs on a common vertex set (one has red edges, the other blue), we would like to know whether they are 'simultaneously connected'. More formally, the process keeps track of a partition of the vertices into clusters. At the beginning each vertex forms its own cluster. Then clusters merge if they are connected by both a red edge and a blue edge. (The merging is done in rounds: first we compute an auxiliary graph of merge-able partition classes, and then the components of this graph correspond to the new partition classes.)

We investigate an extension of jigsaw percolation for hypergraphs based on high-order connectedness. If the process stops with all j-sets in a single cluster we say that the pair of hypergraphs *j-percolates*. Our first result shows that the property that a pair $\mathcal{H} = (\mathcal{H}^k(n,p_1), \mathcal{H}^k(n,p_2))$ of k-uniform binomial hypergraphs j-percolates has a threshold $\hat{p}_{\mathrm{jp}} = \hat{p}_{\mathrm{jp}}(n,k,j)$ in terms of the product p_1p_2, and this threshold satisfies

$$\hat{p}_{\mathrm{jp}} = \Theta\left(\frac{1}{n^{2k-2j-1}\log n}\right). \tag{1.5}$$

Theorem 1.4 (Theorem 5.2)**.** *For $1 \le j < k$ there exists a constant $c = c(k,j) > 0$ such that*

(a) if $p_1p_2 \ge \frac{c}{n^{2k-2j-1}\log n}$ and $\min\{p_1,p_2\} \ge \frac{c\log n}{n^{k-j}}$, then whp $\mathcal{H}$ j-percolates;
(b) if $p_1p_2 \le \frac{1}{cn^{2k-2j-1}\log n}$, then whp $\mathcal{H}$ does not j-percolate.

In the graph case ($k = 2$ and $j = 1$) this result was proven by Bollobás, Riordan, Slivken, and Smith [37]. Note that a necessary condition for j-percolation is that both hypergraphs are j-connected. In the supercritical regime this holds whp because the additional condition $\min\{p_1,p_2\} \ge \frac{c\log n}{n^{k-j}}$ (for large $c > 0$) guarantees that both p_1 and p_2 exceed the threshold $\hat{p}_{\mathrm{c}}$ for j-connectedness given in (1.4). We observe that none of the constants in this result have been optimised, and indeed

the question whether the threshold is sharp or coarse is still open. We discuss this and related issues in Section 1.4.3.

To derive an upper bound on the threshold we provide a very concise reduction argument guaranteeing percolation. There are two steps: firstly, in a pair of k-uniform hypergraphs on n vertices which is supercritical with respect to j-jigsaw percolation, the pair of *link-hypergraphs* of any vertex (each being a $(k-1)$-uniform hypergraph on $n-1$ vertices) is itself supercritical with respect to $(j-1)$-jigsaw percolation. From this fact we then deduce j-jigsaw percolation of the original pair of hypergraphs. By iterating it remains to consider a pair of hypergraphs being supercritical with respect to vertex-jigsaw percolation. In this case, we split the vertex set in half and only consider edges with precisely two vertices in one of the sides. The resulting pairs of auxiliary graphs are supercritical with respect to (vertex)-jigsaw percolation, and a union bound completes the argument.

Interestingly, in contrast to the case of graphs, providing an asymptotically matching lower bound is far from trivial. For this reason we dedicate Section 1.3.3 to the methods for achieving the lower bound.

We strengthen the above result by investigating joint connectedness of an arbitrary (but fixed) number of k-uniform hypergraphs/colours. In other words, for any integer $s \geq 2$ let $\mathcal{H}_s = (\mathcal{H}^k(n,p_1),\dots,\mathcal{H}^k(n,p_s))$, then the threshold $\hat{p}_{\mathrm{jp},s} = \hat{p}_{\mathrm{jp},s}(n,k,j)$ for j-percolation satisfies

$$\hat{p}_{\mathrm{jp},s} = \Theta\left(\frac{1}{n^{s(k-j-1)+1}(\log n)^{s-1}}\right). \tag{1.6}$$

Theorem 1.5 (Theorem 5.12)**.** *For $1 \leq j \leq k-1$ and $s \geq 2$ there exists a constant $c = c(k,j,s) > 0$ such that*

(a) if $\prod_{i=1}^{s} p_i \geq \frac{1}{cn^{s(k-j-1)+1}(\log n)^{s-1}}$ and $\min\{p_1,\dots,p_s\} \geq \frac{c\log n}{n^{k-j}}$, then whp $\mathcal{H}_s$ j-percolates;

(b) if $\prod_{i=1}^{s} p_i \geq \frac{c}{n^{s(k-j-1)+1}(\log n)^{s-1}}$, then whp $\mathcal{H}_s$ does not j-percolate.

Note that $\hat{p}_{\mathrm{jp},2} = \hat{p}_{\mathrm{jp}}$ (cf. (1.5) and (1.6)). The special case $k=2$, $j=1$ and $s \geq 2$ of Theorem 1.5 was proved by Gutiérrez Sánchez [67] using an adaptation of the approach in [37]. Our proof for the supercritical regime relies on the result in [67] as the base case for an inductive argument.

In Section 1.3.3 we focus on the subcritical regime of Theorems 1.4 and 1.5. The focus of Section 1.4.3 lies on analysing the threshold for jigsaw percolation in more detail.

1.2.4. **Bootstrap percolation: localised infections.** The fourth and final topic is largely motivated by applications in real-world networks: percolation processes, such as bootstrap percolation (where in each round any vertex having at least $r \geq 2$ infected neighbours also becomes infected, and remains so forever), are used to model the spread of an infection (or some other form of activity within the network). Therefore their study is specifically focused on random graph models exhibiting typical properties such as power-law degree sequence, small diameter,

and strong clustering. Very few rigorous results are known for this type of models, since their analysis is very challenging.

One of the few exceptions is the model of *geometric inhomogeneous random graphs* (GIRGs) which was recently introduced by Bringmann, Keusch, and Lengler [41]. There the desired clustering properties originate from an underlying geometry. Compared to the binomial random graph model $G(n,p)$ there are two major differences.

First of all, the vertex set is determined by a Poisson point process with intensity n on the underlying geometric space (a torus $\mathbb{T}^d$ of dimension $d \geq 1$), meaning that every vertex has a *position* chosen uniformly within this space. The probability of a pair of vertices forming an edge decreases as their distance increases.

Secondly, every vertex has a *weight* chosen according to a power-law with exponent $2 < \beta < 3$, and the greater the weight of a vertex the more likely it is to form an edge with any other vertex. The model has an additional parameter $\alpha > 1$ determining how rapidly the edge probabilities decay once vertices are 'far' apart.[9]

More formally, each pair of vertices u and v of weights w_u, w_v and with positions x_u, x_v, forms an edge independently with some probability $p = p(w_u, w_v, x_u, x_v)$. and the function p satisfies

$$p(w_u, w_v, x_u, x_v) = \Theta(1) \cdot \min\left\{\left(\frac{w_u w_v}{\|x_u - x_v\|^d}\right)^\alpha, 1\right\}.$$

The underlying geometry provides the opportunity to investigate a *localised* variant of bootstrap percolation. Suppose that an infection starts within a local region B_0 called the *origin of the infection* (to be thought of as very small but still such that $\nu = \nu(n) := n\mathrm{Vol}(B_0) \to \infty$), and every vertex in B_0 becomes infected according to some *initial infection rate* $\rho = \rho(n) \in [0,1]$ independently. We take a close look at how this infection then spreads in both time and space. The first result shows that the threshold with respect to ρ for a *(linear-sized) outbreak* (meaning that $\Omega(n)$ vertices become infected eventually) is

$$\hat{\rho} = \hat{\rho}(\nu, \beta) := \nu^{-\frac{1}{\beta-1}}. \tag{1.7}$$

Theorem 1.6 (Theorem 6.1)**.** *Consider localised bootstrap percolation with initial infection rate $\rho = \rho(n) \in [0,1]$.*

(a) If $\rho = \omega(\hat{\rho})$, then whp there is an outbreak.
(b) If $\rho = o(\hat{\rho})$, then whp no additional vertices become infected.
(c) If $\rho = \Theta(\hat{\rho})$, then the probability of an outbreak is $\Omega(1)$, but the probability that no additional vertices become infected is also $\Omega(1)$.

Note that hyperbolic random graphs are an instance of GIRGs (this was shown in [41]). Therefore our result is a strengthening of the recent result on bootstrap percolation on random hyperbolic graphs by Candellero and Fountoulakis [44].

[9]The model and our results also extend to $\alpha = \infty$ with minor modifications, see Chapter 6.

Our proof is based on a detailed analysis of how a given vertex will typically become infected depending on its position and weight. We present the key ideas of this analysis in Section 1.3.4. As a consequence of this thorough analysis we determine when the infection reaches linear size. We denote the hitting time for this property by τ_o (i.e. τ_o is a random variables taking values in $\mathbb{N} \cup \{\infty\}$) and set

$$i_\infty = i_\infty(n, \nu, \beta) := \frac{\log\log_\nu n + \log\log n}{|\log(\beta - 2)|}.$$

Theorem 1.7 (Theorem 6.2)**.** *For any $\varepsilon > 0$, we have*

$$\mathbb{P}(\tau_o \le (1+\varepsilon) i_\infty) = \begin{cases} 1 - o(1) & \text{if } \rho = \omega(\hat{\rho}), \\ \Omega(1) & \text{if } \rho = \Theta(\hat{\rho}); \end{cases}$$

and furthermore, if $\alpha < \beta - 1$ and $\nu = n^{o(1)}$, then $\mathbb{P}(\tau_o \le (1-\varepsilon) i_\infty) = o(1)$.

In the supercritical regime, we can also accurately predict when each individual vertex will become infected depending on its weight and position. Given a vertex v we write L_v for its *infection time*, i.e. $L_v \in \mathbb{N} \cup \{\infty\}$ is the hitting time of v becoming infected, and for any $x \in \mathbb{T}^d \setminus B_0$ and $w \in \mathbb{R}_{>0}$ we define

$$\Lambda(x, w) := \begin{cases} \max\left\{0, \ \frac{\log\log_\nu\left(\|x\|^d n / w\right)}{|\log(\beta-2)|}\right\}, & \text{if } w > (\|x\|^d n)^{1/(\beta-1)}, \\ \\ \frac{2\log\log_\nu(\|x\|^d n) - \log\log_\nu w}{|\log(\beta-2)|}, & \text{if } w \le (\|x\|^d n)^{1/(\beta-1)}. \end{cases}$$

Theorem 1.8 (Theorem 6.4)**.** *Suppose that $\rho = \omega(\hat{\rho})$, for $\hat{\rho}$ as in (1.7). Consider a vertex $v = (x_v, w_v)$ with $x_v \in \mathbb{T}^d \setminus B_0$, $w_v = \omega(1)$, and $\Lambda(x_v, w_v) \le \log_2(\nu^{-2/(\beta-2)} \|x_v\|^d n)$. Then whp we have*

$$L_v \le (1 + o(1))\Lambda(x_v, w_v) + O(1).$$

If additionally $\alpha > \beta - 1$, then whp we also have

$$L_v \ge (1 - o(1))\Lambda(x_v, w_v) - O(1).$$

Summing up, we characterise under which conditions a localised initial infection causes a global, linear-sized outbreak. Moreover, we determine how quickly the infection spreads, and also provide the infection time of individual vertices. In Section 1.3.4 we sketch the typical evolution of bootstrap percolation on GIRGs, and how Theorems 1.6, 1.7 and 1.8 follow from these insights. More elaborate infection models are discussed in Section 1.4.4

1.3. Key techniques

This section briefly summarises the concepts and methods on which the proofs in this thesis are based. These techniques form the major original mathematical contribution of our work, in addition to the results highlighted in the previous section.

1.3.1. **Coupling methods and branching processes.** In this section we describe variations of a branching process approach to the giant component problem which was originally developed by Bollobás and Riordan [35]. The general idea is to (partially) explore components using a breadth-first-search algorithm, and then analyse these exploration processes using stochastic domination.

We use a multi-type variant of this approach to prove Theorem 1.1. For any vertex, we tightly couple its exploration processes with a 2-type Galton-Watson branching process. Roughly speaking, vertices in large components are represented by branching processes which survive, and those lying in small components correspond to branching processes which die out quickly. Hence, by a rather delicate second moment argument the number of vertices in large components is approximately linear in the survival probability vector of these branching processes. The argument is then completed by a standard sprinkling argument: any two large components would be very likely to have an edge between them, and thus the probability of two of them coexisting is negligible.

1.3.2. **Smoothness.** Our analysis of the threshold of the giant high-order connected component in random hypergraphs (Theorem 1.2) is based on a more elaborate adaptation of the branching process method described in Section 1.3.1. However, there is one major obstacle when it comes to the second moment argument.

In order to upper bound the number of j-sets in large components we have to control how two j-component exploration processes interact. Thus we explore one component partially (just long enough so that we know if it is likely to be large), denote the partial component by C_1, and call its last generation the *boundary* ∂C_1.

Then for exploring the second component we forbid to query any k-sets containing a j-set of C_1, and call the resulting partial component C_2. Because we use a breadth-first search algorithm for the exploration, the only way that both C_1 and C_2 are actually contained in the same component (and therefore the two exploration processes are positively correlated) is if there is successful query of a k-set containing a j-set of both the boundary ∂C_1 and C_2.

Whereas in the graph case this is simple since both ∂C_1 and C_2 are just sets of vertices, for high-order connectedness ($j \geq 2$) things are far more complicated: the number of k-sets containing at least one j-set from ∂C_1 and C_2 depends sensitively on how the individual j-sets in ∂C_1 and C_2 intersect. We overcome this challenge by proving that ∂C_1 is *smooth*, in the sense that for any $0 \leq \ell \leq j-1$ every ℓ-set of vertices is contained in approximately the same number of j-sets of ∂C_1. The concept of *smoothness* is arguably one of the deepest concepts in this thesis.

Smooth boundary lemma. More formally, we write $\partial C_J(i)$ for the i-th generation of the breadth-first search process starting in a j-set J. For any $\ell = 1, \dots, j-1$ let $i_0(\ell)$ be the first round i for which $\partial C_J(i)$ is significantly larger than n^ℓ and for an ℓ-set L let $d_L(\partial C_J(i))$ be the number of j-sets of $\partial C_J(i)$ that contain L. Furthermore, let i_1 denote the generation at which the search process hits one of three stopping conditions (which will be specified in Section 3.3).

Lemma 1.9 (Smooth boundary lemma – Lemmas 3.3 (simplified) and 3.8)**.** *Let $1 \le j \le k-1$ and $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n^{1-\delta} \to \infty$, for some constant $\delta > 0$. Then with probability at least $1 - \exp\left(-n^{\Theta(1)}\right)$, for all J, ℓ, L, i such that*

- *J is a j-set of vertices;*
- *$0 \le \ell \le j-1$;*
- *L is an ℓ-set of vertices;*
- *$i_0(\ell) + \Theta(\log n) \le i \le i_1$,*

the following holds:

$$d_L(\partial C_J(i)) = (1 \pm o(1)) \frac{|\partial C_J(i)|}{\binom{n}{j}} \binom{n}{j-\ell}.$$

The proof is based on closely tracking the degrees $d_L(\partial C_J(i))$ of an ℓ-set L during the exploration process. The major challenge lies in both **a)** obtaining the precise leading constant and **b)** investigating the degree within a single generation and not the entire currently discovered component.

From the perspective of L there are two types of queries (of k-sets) which contribute to its degree: queries from j-sets not containing L, called *jumps*; and queries from j-sets containing L, called *pivots*. This perspective was previously used by Cooley, Kang, and Person in [51] for bounding from above the degree of L within the entire currently discovered component. We use a strengthened form, Lemma 3.13, of their upper bound at various points of our proof. With its help the analysis of both jumps and pivots boils down to studying sums of independent indicator random variables (each of them representing a query). Thus the Chernoff bounds are applicable and provide concentration with exponentially small error probabilities provided that the previous generation is not too small.

As we consider a supercritical regime, generations have a tendency to grow in size (Lemma 3.16). So as soon as they reach a reasonable size, the concentration arguments remain valid for long enough to ensure that the disparity between the degrees of different ℓ-sets has evened out. Once generations are smooth they remain so at least until one of the stopping conditions is reached. Even though the idea itself is not overly complicated, the details of its proof are rather involved and form a major part of Chapter 3.

Applications. In fact, in order to complete the proof of Theorem 1.2 by applying Lemma 1.9 we need to guarantee some properties of the supercritical exploration process: firstly, a typical component has a reasonably sized boundary before it grows too large (Lemma 3.10), or in other words, the smoothing process starts early. Secondly, this implies that whp smoothness is already reached by the time the exploration process stops (Lemma 3.11).

Lemmas 1.9, 3.10 and 3.11 enable us to not only obtain smooth generations, but also larger smooth sets, for instance by taking the union of smooth generations. Actually, this fact is extremely useful for determining the hitting time for high-order connectedness in random hypergraphs (Theorem 1.3).

Lemma 1.10 (Smooth subset lemma – Lemma 4.7)**.** *Let $1 \le j \le k-1$, let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n^{1-\delta} \to \infty$, for some constant $\delta > 0$, and set $p^* = \frac{1+\varepsilon}{(\binom{k}{j}-1)\binom{n}{k-j}}$. Then whp there is a j-component of $\mathcal{H}^k(n,p^*)$ with a subset S of at least $\varepsilon^3 n^j$ many j-sets satisfying the following property:*

Each $(j-1)$-set in $\mathcal{H}^k(n,p^*)$ is contained in $(1\pm o(1))\frac{|S|}{\binom{n}{j}}n$ many j-sets of S.

Applying the Chen-Stein method we show that the last isolated j-set in $\mathcal{H}^k(n,p)$ most likely disappears when $p = (1\pm o(1))\frac{j\log n}{\binom{n}{k-j}}$. Now using a two-round exposure once again, Lemma 1.10 implies that there must be a large component C containing a fairly large smooth subset. We use this information and demonstrate that thus all other non-trivial components merge with C very quickly. In this regime, we typically see a giant j-component being the only non-trivial j-component, i.e. all other j-components are isolated j-sets. This structure transfers to the hypergraph process $\{\mathcal{H}^k(n,M)\}_M$, due to the asymptotic equivalence of the two models. Hence j-connectedness is reached precisely when the last isolated j-set disappears as stated in Theorem 1.3.

1.3.3. **Traversable triples.** Next we take a closer look at jigsaw percolation: it is surprisingly challenging to prove that complete percolation does not take place in the subcritical regime of jigsaw percolation on hypergraphs (Theorem 1.4). Our proof is based on investigating the number of *traversable triples* contained in a pair of random hypergraphs $(\mathcal{H}^k(n,p_1),\mathcal{H}^k(n,p_2))$.

We define this purely combinatorial concept in two steps: firstly, we say that a collection $\mathcal{J}$ of j-sets is *traversable* (in a hypergraph H) if for every two distinct j-sets $J, J' \in \mathcal{J}$, $J' \neq J$ there is a sequence of edges $e_1,\dots,e_m$ such that

- $J \subset e_1$ and $J' \subset e_m$;
- $\forall 1 \le i \le m-1$: some j-set $J_i \in \mathcal{J}$ is contained in $e_i \cap e_{i+1}$.

In other words, we may walk from J to J' using edges such that the intersection of two consecutive edges contains at least one j-set from $\mathcal{J}$.[10]

Then given a vertex set V we call a triple $T = (\mathcal{J}_0,\mathcal{E}_1,\mathcal{E}_2)$ *traversable* if $\mathcal{J}_0$ is traversable in both the red and blue hypergraphs, $(V,\mathcal{E}_1)$ and $(V,\mathcal{E}_2)$, respectively. Furthermore, we write $\mathcal{T}_{\ell,r,b}$ for the set of all (edge-minimal) traversable triples with ℓ many j-sets, and the number of red/blue edges being r and b respectively.

We provide an algorithm which associates any triple $T \in \mathcal{T}_{\ell,r,b}$ with a *blueprint* $\pi(T)$, which contains the structural information of T, but not how it is embedded into the vertex set V. Moreover we denote by $\mathcal{M}_{a,m}$ the set of $\ell \times (\binom{k}{j}+1)$ matrices having non-negative integer entries and satisfying two linear relations depending on the parameters $a, m \in \mathbb{N}$. Then the blueprint is a pair of such matrices.

Lemma 1.11 (Lemma 5.5)**.** *For integers $\ell > r, b \ge 0$: $\pi(\mathcal{T}_{\ell,r,b}) \subset \mathcal{M}_{\ell,r} \times \mathcal{M}_{\ell,b}$.*

[10]Note that this is *not* equivalent to $\mathcal{J}$ being contained in a single j-component in the hypergraph spanned by all edges (from H) containing at least one j-set from $\mathcal{J}$. The reason is that we insist that each intersection (of consecutive edges) contains at least one j-set of $\mathcal{J}$.

In a second step, from any blueprint we algorithmically reconstruct all traversable triples (embedded into a pair of hypergraphs (V, E_1) and (V, E_2)) having this blueprint. We denote the total number of distinct outputs of this reconstruction algorithm (with input $(M_1, M_2) \in \mathcal{M}_{\ell,r} \times \mathcal{M}_{\ell,b}$) by $\mathcal{Q}_{\ell,r,b}$.

Lemma 1.12 (Lemma 5.7)**.** *For integers $\ell > r, b \geq 0$ we have $\mathcal{T}_{\ell,r,b} \subset \mathcal{Q}_{\ell,r,b}$.*

Furthermore, there is a constant $C > 0$ (independent of ℓ, r, b) such that

$$|\mathcal{Q}_{\ell,r,b}| \leq |V|^j C^{\ell-1} \left(|V|^{k-j}\right)^r \left(\ell |V|^{k-j-1}\right)^b.$$

Together, Lemmas 1.11 and 1.12 provide an upper bound on the number of traversable triples containing $\Theta(\log n)$ many j-sets in a pair of complete hypergraphs on a (joint) set of n labelled vertices. Thus, it follows by a first moment argument that whp none of these triples exist in the pair $(\mathcal{H}^k(n, p_1), \mathcal{H}^k(n, p_2))$ in the subcritical regime of Theorem 1.4, i.e. if $p_1 p_2 \leq 1/(cn^{2k-2j-1} \log n)$ for a sufficiently large constant $c > 1$. As a matter of fact, this already suffices to prove that percolation does not take place due to the following observation.

Let us consider the jigsaw percolation process on a finer time scale, revealing edges one by one (no matter their colour), and merging clusters immediately if possible. Note that each edge can merge at most $\binom{k}{j}$ clusters. Thus, in order for percolation to take place, there must be a structure containing $\Theta(\log n)$ many j-sets, and sets of red and blue edges certifying that it would percolate on its own. (These correspond to *internally spanned sets* in [37].) However, each of these must in turn contain an (edge-minimal) traversable triple $(\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2)$, with $|\mathcal{J}_0| = \Theta(\log n)$. Yet we just proved that the existence of such a traversable triple is unlikely, completing the argument.

1.3.4. **Geometric spread of infection.** In the following we investigate the evolution of a localised infection process on GIRGs, modelled by bootstrap percolation (with parameter $r \geq 2$). We aim to sketch the proofs of Theorems 1.6, 1.7, and 1.8. The question of whether there is an outbreak or not is decided in the earliest stages of the process: typically the process either dies out immediately not infecting any additional vertex at all, or infects a linear number of vertices.

In the intermediate critical regime each of these events happens with constant probability: if at least r vertices of 'large' weight in the origin B_0 of the infection become infected in the first round, then all of them do and hence the process behaves as if it was supercritical. On the other hand, the probability of infecting no additional vertex in the first round is still bounded away from 0.

The proof for dying out immediately is essentially a first moment argument. Thus we focus on the evolution of the process in the supercritical regime for the remainder of this section.

So now suppose we are in the case where all heavy vertices close to the origin B_0 of the infection are infected. Then we define two sequences of nested *inner/outer infection regions* $\{B_i\}_i$ and $\{\widetilde{B}_i\}_i$ and use them to provide a detailed description of

the process evolving in both time and along the geometry. Certain classes of *heavy* vertices – those lying in the inner infection region B_i and having weight at least $w_{i,0}$ (the 'typical maximal weight' in B_i) – play a key role.

The general idea is to proceed inductively as follows (where all statements hold whp): we show that in round i the heaviest vertices in the inner infection region B_i, i.e. those of weight at least $w_{i,0}$, are all infected (Theorem 6.17(b)), and similarly we prove that there are no vertices outside of the outer infection region $\widetilde{B}_i$ which have already become infected (Theorem 6.20(c)). The set of heavy vertices, once infected, will then start a *cascade of infection* which reaches vertices of subsequently lower weights in each round.

We write $\mathcal{A}(i)$ for the event that in each round $i' \leq i$ the heavy vertices in the inner infection region B_i have become infected, and introduce a sequences of additional *weight-bounds* $\{w_{i,\ell}\}_\ell$, where for each i we have $\log_\nu(w_{i,\ell}) \to 0$ exponentially quickly as $\ell \to \infty$. Then Theorem 6.17(c) provides lower bound on the probability (conditional on $\mathcal{A}(i)$) for any given vertex in the inner infection region B_i of weight at least $w_{i,\ell}$ to be infected at time $i + \ell$.

We then use this bound to show that any given vertex of weight at least (some large constant) $C_0 > 0$ will eventually (in fact by round $(1+\varepsilon)i_\infty$) become infected with probability bounded away from 0. Hence in expectation a linear number of vertices has already become infected at this point of time. The whp statement follows with a little additional effort, proving that a linear size outbreak occurs by round $(1+\varepsilon)i_\infty$ at the latest and thus completing the proof of Theorem 1.6.

To prove Theorem 1.7 it remains to establish the corresponding lower bound on the hitting time τ_o for the infection reaching linear size. Actually the process is far from as clear-cut as our idealised description in the previous paragraphs. However we can control the influence of additional vertices of lower degrees becoming infected if edge probabilities decay sufficiently rapidly once the 'maximal distance' is reached, i.e. $\alpha > \beta - 1$ (which is only a minor restriction since $\beta < 3$).

We prove that 'not too many' additional vertices are infected in round i (Theorem 6.20(e)), and denote this event by $\widetilde{\mathcal{A}}(i)$. Then we provide an upper bound on the probability (conditional on $\widetilde{\mathcal{A}}(i)$) of a given vertex being infected in round $i+1$ depending on how far it is from the outer infection regions $\widetilde{B}_0, \dots, \widetilde{B}_i$ 'relative' to its weight (Theorem 6.20(f)). From here the proof of Theorem 1.7 is completed by a first moment argument.

Similarly, Theorems 6.17 and 6.20 provide all the information for determining the infection times of individual vertices (Theorem 1.8). However, the actual proof splits into several cases and requires quite a bit of calculation.

To conclude this section let us briefly return to our motivation of modelling an infection process in a real-world network. From this perspective we would like to not only understand when an outbreak occurs, but also how to save as many individuals as possible. Our description of the process suggest such a *quarantine*

strategy: because Theorem 6.20(e) shows that whp in round i the infection is still completely contained within the outer infection region $\widetilde{B}_i$, it suffices to remove all edges which connect this region to the outside ($\mathbb{T}^d \setminus \widetilde{B}_i$) before round i. Based on a proof in [41] we show that the expected number of these edges is very small.

1.4. Discussion

Every scientific result leads to a variety of intriguing and challenging open problems and directions for future research. In this section we will briefly outline a selection of such 'loose ends' related to the topics covered in the thesis. Following the previous pattern we also split this part into four subsections, one for each of the core topics.

1.4.1. **Young giant in inhomogeneous random graphs.** Recall that we analysed the emergence of the giant component in $G(\mathbf{n}, P)$ using 2-type branching processes (cf. Sections 1.2.1 and 1.3.1). We identified the threshold in terms of the matrix $M = P \cdot \mathrm{diag}(\mathbf{n})$ based on certain row-sum conditions (cf. (1.1) and (1.2)).

However, drawing our intuition from the branching process approximation, we suspect the threshold to be most naturally characterised by the Perron-Frobenius eigenvalue λ (and the corresponding normalised left-eigenvector $\mathbf{v}$) of M satisfying $\lambda = 1 + \varepsilon$ or $\lambda = 1 - \varepsilon$, where $\varepsilon \to 0$ not too quickly. Moreover, we know from the Perron-Frobenius theory (cf. Section V.6 in [19]) that, after appropriate rescaling, the vector describing the expected offspring in the i-th generation of the associated branching process converges to $\mathbf{v}$ as $i \to \infty$, assuming that the process survives. In other words we would expect that there is a function $\beta_\varepsilon = \beta_\varepsilon(\mathbf{n}) = o(1)$ such that the vector of survival probabilities satisfies $\boldsymbol{\rho} = (1 \pm o(1))\beta_\varepsilon \mathbf{v}$ (component-wise).

A result of this form describing the growth of the young giant component would then generalise readily to more sophisticated inhomogeneous random graphs with an arbitrary number of types. In other words, this would close the gap between the results for constant $\varepsilon > 0$ by Bollobás, Janson, and Riordan [33] and those inside the critical window due to Bhamidi, Broutin, Sen, and Wang [27].

1.4.2. **Structure and distribution of high-order connected components.** The focus of Sections 1.2.2 and 1.3.2 was on two of the most characteristic regimes in the evolution of high-order connected components in random hypergraphs: firstly, we determined the size of the giant j-component right after its emergence, and secondly, we proved that the random hypergraph process becomes j-connected precisely when its minimal obstruction, i.e. the last isolated j-set, disappears. In order to obtain these key results we have developed the concept of smoothness and provided a powerful tool called the smooth boundary lemma.

A very interesting open problem is to pin down the limiting distribution of the size of the giant j-component. The existing central and local limit theorems for the case of graphs [85, 95] and more generally vertex-connectedness [25, 36, 79] indicate that the most likely candidate is also a normal distribution. However, based on our

experience with high-order connectedness there will be various obstacles, some of which may be overcome by the smooth boundary lemma or a variation thereof, while others will require substantial new ideas.

Closely related is the question of the structure of 'small' components. It is well-known that for graphs whp these are all either trees or unicyclic components. But what is the 'right' notion of a tree or a unicyclic component in hypergraphs? In this context it seems natural to call a j-component with the minimal number of j-sets a 'tree', i.e. a j-component with e edges and $(\binom{k}{j}-1)e+1$ many j-sets. Similarly, a j-component with e edges and $(\binom{k}{j}-1)e$ many j-sets would be 'unicyclic'. Is it then true that whp all j-components except for the giant j-component are either trees or unicyclic?

An affirmative and (very) precise answer to this question might prove to be the first step towards a local limit theorem for the size of the giant j-component, as was the case for vertex-connectedness [79]. Also for the approach in [25] controlling the interaction between small components and the giant component is key, even though the way in which this is done is quite different from that in [79]. At this point knowing that the giant j-component contains a large smooth subset (Lemma 1.10, or more generally Lemma 1.9) might be just the right property to handle how likely small j-components are to merge with the giant j-component.

1.4.3. **Critical window for jigsaw percolation.** While we proved the existence of a threshold for jigsaw percolation on random hypergraphs in Sections 1.2.3 and 1.3.3, there is still a gap of a large but constant multiplicative factor c^2 between the super- and subcritical regimes. Even though this gap has not been optimised (either by us or in the graph case [37]), this already suggests that this property has a sharp threshold, and raises the question of determining its asymptotics exactly.

Moreover, recall that there is the additional condition $\min\{p_1,p_2\} \geq \frac{c\log n}{n^{k-j}}$ for the supercritical regime, guaranteeing that the necessary condition of both the red/blue hypergraphs being j-connected is met whp. However, we determined the sharp threshold for j-connectedness to be $\hat{p}_c = \frac{j(k-j)!\log n}{n^{k-j}}$ in (1.4). Thus it is natural to conjecture that the condition can be relaxed to $\min\{p_1,p_2\} \geq C\hat{p}_c$ for any constant $C > 1$. But also if $p_1 = \hat{p}_c$, then the probability of the red hypergraph being j-connected is bounded away from both 0 and 1. We conjecture that conditioned on it being j-connected whp percolation takes place (as whp the blue hypergraph is also j-connected).

1.4.4. **Localised SIRS-infections.** Our results concerning bootstrap percolation on GIRGs provide a very detailed picture of how this process evolves both geometrically and in time (cf. Sections 1.2.4 and 1.3.4). We demonstrated not only how an infection originating in a small local region can cause an outbreak infecting a linear number of vertices, but also how fast the infection will spread in this scenario. We also proved that this kind of information is invaluable for saving the majority of the population from an imminent outbreak.

We should emphasise that the importance of our contribution is not merely limited to the mathematical statements about bootstrap percolation on GIRGs but also, and maybe foremost, a proof of concept: we illustrate a way of rigorously modelling an infection process in a geometric setting.

This being said, it is up to us to question whether this is a good model. While GIRGs are certainly a state-of-the-art model for real-world networks, modelling the spread of an infection, a belief, or some neuronal activity via bootstrap percolation is certainly not optimal. Naturally, a large variety of models have been proposed for these scenarios, amongst them for instance SIR-and SIRS-models (*susceptible-infected-recovering*) for epidemiological applications, where infected individuals can recover from an infection. Investigating the influence of an underlying geometry on these processes promises to be a challenging task, and certainly one from which there is much to learn.

1.5. Outline

Chapters 2–6 each consist of a research paper, with modifications, from the publication list preceding this introduction. Some results have also been published in conference proceedings. More precisely we have the following correspondence:

Chapter 2	↔	[6]	(with M. Kang and A. Pachón)
Chapter 3	↔	[2] and [8]	(with O. Cooley and M. Kang)
Chapter 4	↔	[3] and [8]	(with O. Cooley and M. Kang)
Chapter 5	↔	[1]	(with B. Bollobás, O. Cooley, and M. Kang)
Chapter 6	↔	[7] and [10]	(with J. Lengler)

The research leading to [7] was conducted during a 3-month research stay at the ETH Zürich, within the doctoral program ‘Discrete Mathematics’.

CHAPTER 2

Emergence of the giant component in a multi-type random graph

2.1. Introduction and main results

The theory of random graphs was founded by Erdős and Rényi in the late 1950s. One of their most striking results concerned the *phase transition* of the size of the largest component – adding a few additional edges to a random graph can drastically alter the size of its largest component. In [58] they considered the random graph $G(n, m)$ obtained by choosing a graph uniformly at random amongst all graphs on n (labelled) vertices containing precisely m edges and proved the following result: Let $c \geq 0$ be any constant. If $c < 1$, then *with high probability* (*whp* for short, meaning with probability tending to one as $n \to \infty$) all components in $G(n, cn/2)$ have size $O(\log n)$, while if $c = 1$, whp the largest component is of size $\Theta(n^{2/3})$, and if $c > 1$, then whp there is a component of size $\Theta(n)$, called the 'giant component', and all other components are of size $O(\log n)$.

Bollobás [31] investigated this phenomenon further and described in detail the behaviour of $G(n, m)$ when m is close to $n/2$, i.e. $m = (1 \pm \varepsilon)n/2$ for some $\varepsilon = \varepsilon(n) > 0$ satisfying $\varepsilon \to 0$ as $n \to \infty$. His initial results were then improved by Łuczak [86]. In particular, if in addition $\varepsilon^3 n \to \infty$, whp the largest component in $G(n, (1-\varepsilon)n/2)$ has size $o(n^{2/3})$, whereas the largest component in $G(n, (1+\varepsilon)n/2)$ contains asymptotically $2\varepsilon n$ vertices and all other components are of size $o(\varepsilon n)$. For a comprehensive account of the results see [15, 30, 74].

In the meantime many of these results have been reproved and strengthened using various modern techniques such as martingales [89], partial differential equations [103], and search algorithms [33, 82]. Furthermore, more complicated discrete structures like random hypergraphs have been studied [26, 36, 79].

Over the last years, random graphs have proved to have wide-ranging applications in neurobiology, statistical physics, and the modelling of complex networks [90, 104]. Frequently some properties of real-world networks are already empirically 'known' and have motivated the definition of more sophisticated random graph models [48, 49, 102]. In particular, applicable random graph models should allow for different types of vertices having different degree distributions, i.e. some level of *inhomogeneity*. A general theory of inhomogeneous random graphs was developed by Bollobás, Janson, and Riordan [33] providing a unified framework for a large number of previously studied random graph models [34, 42, 91]. For

example they analysed the degree distribution, the number of paths and cycles, and the phase transition for the giant component. The behaviour at the *critical point* (corresponding to $G(n, cn/2)$ for $c = 1$) has been studied by van der Hofstad in the so-called *rank one* case [70]. Recently Bhamidi, Broutin, Sen, and Wang studied the general inhomogeneous random graph with a bounded number of types inside the *critical window* (corresponding to $G(n, (1 \pm \varepsilon)n/2)$ for some $\varepsilon = \varepsilon(n) > 0$ satisfying $\varepsilon^3 n \to C$, $0 \leq C < \infty$) and described the joint distribution of the largest components using Brownian motion [27].

In this chapter, we study an inhomogeneous random graph model in which there are n vertices, each vertex has one of two types, and an edge between a pair of vertices of types i and j is present with probability $p_{i,j}$ independently of all other pairs. The focus lies on the *weakly* supercritical regime, i.e. when the distance to the critical point of phase transition approaches zero as $n \to \infty$. In this regime the behaviour of the random graph depends very sensitively on the parameters and could not be studied using the parametrisation in [33]. We determine the size of the largest component in this regime (Theorem 2.1).

In order to derive the main results, we apply a simple breadth-first search approach to construct a rooted spanning tree of a component and couple it with a multi-type branching process with binomial offspring distributions, which is viewed as a random rooted tree. In addition, the width and the dual of that random rooted tree play important roles in the second moment analysis.

The results of this chapter are indeed not surprising and the techniques used in the chapter may look familiar. The main contribution of this chapter is that it shows how a simple branching process approach combined with the concepts of tree width and dual processes can be applied nicely to a multi-type random graph *all the way through* the supercritical regime.

2.1.1. **Model and notation.** In this section we will define multi-type binomial random graphs and associate them with branching processes.

Multi-type binomial random graph. Let $k \in \mathbb{N}$ be fixed. Every vertex is associated with a *type* $i \in \{1, \dots, k\}$ and we denote by V_i the set of all vertices of type $i \in \{1, \dots, k\}$. Given an arbitrary vector $\mathbf{n} = (n_1, \dots, n_k) \in \mathbb{N}^k$ and a symmetric matrix of probabilities $P = (p_{i,j})_{i,j=1,\dots,k} \in [0,1]^{k \times k}$ we consider the *k-type binomial random graph* $G_k(\mathbf{n}, P)$ on n_l vertices of type l, for $l \in \{1, \dots .k\}$, with the following edge set: for each pair $\{u, v\}$, where u is of type i and v of type j, we include the edge $\{u, v\}$ independently of any other pair with probability $p_{i,j}$ and exclude it with probability $1 - p_{i,j}$. We write $M = (\mu_{i,j})_{i,j \in \{1,2\}}$ for the matrix of the expected number of neighbours $\mu_{i,j} = p_{i,j} n_j$ of type $j \in \{1, \dots, k\}$ for a vertex of type $i \in \{1, \dots, k\}$.

Associated branching process. Next we associate a binomial branching process in which each individual has a type $i \in \{1, \dots, k\}$ with the random graph $G_k(\mathbf{n}, P)$. Fix a time $t \in \mathbb{N}_0$ and let I_t be a set of individuals (i.e. the population) at time t, which we also call the t-th generation of individuals. Then, with each

individual $v \in I_t$ of type $j' \in \{1,\dots,k\}$, we associate a random vector $\mathbf{X}^v = (X_1^v,\dots,X_k^v)$, where for each $j \in \{1,\dots,k\}$ the random variable X_j^v is independent and binomially distributed with parameters n_j and $p_{j',j}$ and thus with mean $\mu_{j',j}$. Then the population I_{t+1} at time $t+1$ will be a set containing exactly $\sum_{v\in I_t} X_j^v$ *new* individuals of type j, for each $j \in \{1,\dots,k\}$. In other words, the random variable X_j^v represents the number of children of type j that are born from the individual v. A k-type binomial branching process starting with an initial population I_0 is a sequence of random vectors $(Z_t(1),\dots,Z_t(k))_{t\in\mathbb{N}_0}$ generated by iterating the construction described above, where $Z_t(j)$ is the random variable describing the number of individuals of type j in the t-th generation for each $j \in \{1,\dots,k\}$ and $t \in \mathbb{N}_0$. For $i \in \{1,\dots,k\}$ we denote by $\mathcal{T}^i_{\mathbf{n},P}$ a k-type binomial branching process starting with a single vertex of type i. We may also use $\mathcal{T}^i_{\mathbf{n},P}$ to denote the rooted (possibly infinite) tree created by an instance of the branching process. The context will always clarify the notation. Furthermore, if a statement is independent of the starting type we simply write $\mathcal{T}_{\mathbf{n},P}$, for instance, we refer to the matrix M as *offspring expectation matrix* of the branching process $\mathcal{T}_{\mathbf{n},P}$.

Observe that for $k = 1$ we obtain the classical binomial random graph $G(n,p)$ where $n = n_1$ and $p = p_{1,1}$ and the corresponding binomial branching process.

Setup. Throughout the chapter we focus on the case $k = 2$ and for simplicity we write $G(\mathbf{n},P) := G_2(\mathbf{n},P)$. We denote by $n = n_1+n_2$ the total number of vertices in $G(\mathbf{n},P)$ and without loss of generality we assume that $n_1 \geq n_2$. Furthermore, unless specified explicitly, all asymptotic statements are to be understood in terms of n_1 and n_2 being *large enough* yet fixed and we use the notation $\min\{n_1,n_2\} = n_2 \to \infty$ for this. Note that in general $\eta_{1,2} \neq \eta_{2,1}$ and it is possible that $\eta_{2,1}/\eta_{1,2} \to \infty$, even though $p_{1,2} = p_{2,1}$.

Notation. Given a graph $\mathcal{G}$ with components $C_1,\dots,C_r$ ordered by size such that $|C_1| \geq |C_2| \geq \cdots \geq |C_r|$ we denote by $\mathcal{L}_i(\mathcal{G}) = C_i$ the i-th largest component of $\mathcal{G}$ and its size by $L_i(\mathcal{G}) = |\mathcal{L}_i(\mathcal{G})| = |C_i|$, for any $i \in \{1,\dots,r\}$, and set $\mathcal{L}_i(\mathcal{G}) = \emptyset$ and $L_i(\mathcal{G}) = 0$ if $i > r$. Moreover, we will use the following standard notation to describe asymptotic statements: For any real functions $f = f(n_1,n_2)$ and $g = g(n_1,n_2)$ we write: $f = O(g)$ if $\exists c > 0, n_0$ such that $f(n_1,n_2) \leq c|g(n_1,n_2)|$ for all $n_1 \geq n_2 \geq n_0$; $f = o(g)$ if $\forall c > 0 : \exists n_0$ such that $f(n_1,n_2) \leq c|g(n_1,n_2)|$ for all $n_1 \geq n_2 \geq n_0$; $f = \Omega(g)$ if $g = O(f)$; $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$ and $f \sim g$ if $f - g = o(g)$.

2.1.2. **Main results.** We show that $G(\mathbf{n},P)$ exhibits a phase transition in the size of the largest component. In particular, we show that in the *weakly* supercritical regime there is a unique largest component containing asymptotically $2\varepsilon n$ vertices. In fact, we prove a stronger result.

Theorem 2.1. *For $n_1 \in \mathbb{N}$ and $n_2 \in \mathbb{N}$ with $n_1 \geq n_2$, let $n = n_1 + n_2$ and let $\varepsilon = \varepsilon(n_1,n_2) > 0$ with $\varepsilon = o(1)$. Furthermore, let*

$$P = (p_{i,j})_{i,j\in\{1,2\}} \in (0,1]^{2\times 2}$$

be a symmetric matrix of probabilities satisfying the following conditions:

$$\varepsilon^3 n_2 \min\{1, \varepsilon^{-1}\mu_{2,1}\} \to \infty, \tag{2.1}$$

$$\mu_{\iota,1} + \mu_{\iota,2} = 1 + \varepsilon \pm o(\varepsilon), \text{ for any } \iota \in \{1,2\}, \tag{2.2}$$

where $\mu_{i,j} = p_{i,j} n_j$ *for every pair* $(i,j) \in \{1,2\}^2$. *Then, whp the following holds for all integers* $r \geq 2$ *and* $i \in \{1,2\}$:

$$|\mathcal{L}_1(G(\mathbf{n},P)) \cap V_i| = (2 \pm o(1))\varepsilon n_i \quad \text{and} \quad |\mathcal{L}_r(G(\mathbf{n},P)) \cap V_i| = o(\varepsilon n_i);$$

therefore, in particular,

$$L_1(G(\mathbf{n},P)) = (2 \pm o(1))\varepsilon n \quad \text{and} \quad L_r(G(\mathbf{n},P)) = o(\varepsilon n).$$

Remark 2.2. Observe that, up to the term $\min\{1, \varepsilon^{-1}\mu_{2,1}\}$, the condition in (2.1) mirrors the condition $\varepsilon^3 n \to \infty$ that is necessary and sufficient for the existence of a unique largest component in $G(n,(1+\varepsilon)/n)$. In $G(n,(1+\varepsilon)/n)$ the average degree is $1+\varepsilon = \Theta(1)$ and therefore it does not influence the asymptotic statement in (2.1). In $G(\mathbf{n},P)$ however, P can be such that we are close to criticality but the average number $\mu_{2,1}$ (respectively $\mu_{1,2}$) of neighbours of the opposite type for a given vertex is still $o(1)$. Roughly speaking, it is reasonable that if $\mu_{2,1}$ is 'very small', then the random graph $G(\mathbf{n},P)$ may have two largest components, one of each type, that coexist independently since the probability of adding any edge between them is negligible. In particular, this would happen in case probability $p_{1,2}$ was equal to zero and therefore $\mu_{1,2} = \mu_{2,1} = 0$.

On the other hand, in the special case $n_2 = o(n)$, the condition in (2.1) differs by an additional factor of $n_2/n = o(1)$ from that in $G(n,(1+\varepsilon)/n)$. This factor is, for instance, necessary to show that the number of vertices of type 2 in the largest components is concentrated around its mean. Therefore it is not avoidable with this method, even though it might not be optimal.

Remark 2.3. The symmetry of P simply reflects the fact that $G(\mathbf{n},P)$ is an *undirected* random graph.

Note that the parameter $\varepsilon > 0$ describes the *distance* to the critical point for the emergence of the giant component in a sense that we will explain now. Roughly speaking, for some time, the breadth-first exploration process of a component in $G(\mathbf{n},P)$ looks like a 2-type binomial branching process $\mathcal{T}_{\mathbf{n},P}$. This can be described by a coupling of the two processes. If the branching process dies out its total population should be rather 'small'. Thus, by the coupling, the explored component is also 'small'. It is well-known that for a 2-type binomial branching process the property of survival has a threshold and that the critical point is characterised by the Perron-Frobenius eigenvalue

$$\lambda = \frac{\mu_{1,1}+\mu_{2,2}}{2} + \frac{1}{2}\sqrt{(\mu_{1,1}+\mu_{2,2})^2 + 4(\mu_{1,2}\mu_{2,1} - \mu_{1,1}\mu_{2,2})} \tag{2.3}$$

of its offspring expectation matrix $M = (\mu_{i,j})_{i,j\in\{1,2\}}$. If $\lambda > 1$, the process has a positive probability of survival, while if $\lambda \leq 1$, it dies out with probability 1.

Next, let us compute λ for the 2-type binomial branching process $\mathcal{T}_{\mathbf{n},P}$ with parameters as in Theorem 2.1. The condition in (2.2) states that for every constant $\delta \in (0,1)$ there is an $n_0 = n_0(\delta)$ such that we have $|\mu_{i,1} + \mu_{i,2} - (1+\varepsilon)| \leq \delta\varepsilon$, for $i \in \{1,2\}$ and all $n_1 \geq n_2 \geq n_0$. This implies

$$\mu_{1,2}\mu_{2,1} - \mu_{1,1}\mu_{2,2} \leq (1+\varepsilon+\delta\varepsilon)^2 - (\mu_{1,1}+\mu_{2,2})(1+\varepsilon+\delta\varepsilon)$$

and similarly

$$\mu_{1,2}\mu_{2,1} - \mu_{1,1}\mu_{2,2} \geq (1+\varepsilon-\delta\varepsilon)^2 - (\mu_{1,1}+\mu_{2,2})(1+\varepsilon-\delta\varepsilon).$$

Therefore we can bound the argument of the square root in (2.3) from above by

$$(\mu_{1,1}+\mu_{2,2})^2 + 4(\mu_{1,2}\mu_{2,1} - \mu_{1,1}\mu_{2,2}) \leq (2(1+\varepsilon+\delta\varepsilon) - (\mu_{1,1}+\mu_{2,2}))^2$$

and from below by

$$(\mu_{1,1}+\mu_{2,2})^2 + 4(\mu_{1,2}\mu_{2,1} - \mu_{1,1}\mu_{2,2}) \geq (2(1+\varepsilon-\delta\varepsilon) - (\mu_{1,1}+\mu_{2,2}))^2.$$

Thus, by (2.3) and since δ was arbitrary, we obtain the following asymptotic estimate for the Perron-Frobenius eigenvalue

$$\lambda = 1 + \varepsilon \pm o(\varepsilon). \tag{2.4}$$

In other words, ε describes how close λ is to 1.

Our next result concerns the weakly *subcritical* regime.

Theorem 2.4. *For $n_1 \in \mathbb{N}$ and $n_2 \in \mathbb{N}$ with $n_1 \geq n_2$, let $n = n_1 + n_2$ and let $\varepsilon = \varepsilon(n_1, n_2) > 0$ with $\varepsilon = o(1)$. Furthermore, let*

$$P = (p_{i,j})_{i,j\in\{1,2\}} \in (0,1]^{2\times 2}$$

be a symmetric matrix of probabilities satisfying the following conditions:

$$\varepsilon^3 n_2 \to \infty, \tag{2.5}$$

$$\mu_{\iota,1} + \mu_{\iota,2} = 1 - \varepsilon \pm o(\varepsilon), \text{ for any } \iota \in \{1,2\}, \tag{2.6}$$

where $\mu_{i,j} = p_{i,j}n_j$ for every pair $(i,j) \in \{1,2\}^2$. Then we have whp

$$L_1(G(\mathbf{n},P)) = o(n^{2/3}).$$

Note that analogously to Theorem 2.1 the parameter $\varepsilon > 0$ describes the *distance* to the critical point from below. In other words, by (2.6), we know that the Perron-Frobenius eigenvalue λ of the offspring expectation matrix M of a 2-type binomial branching process $\mathcal{T}_{\mathbf{n},P}$ with parameters as in Theorem 2.4 satisfies

$$\lambda = 1 - \varepsilon \pm o(\varepsilon). \tag{2.7}$$

We will dedicate most of this chapter to the more delicate weakly supercritical regime. A sketch of the proof of Theorem 2.1 is given in Subsection 2.1.3, properties of supercritical branching processes will be analysed in Section 2.2, and the actual proof of Theorem 2.1 is provided in Section 2.3. The weakly subcritical regime follows in Section 2.4 with the proof of Theorem 2.4.

In Section 2.5 we also consider the size of the largest component in the regimes where the distance to the critical value is a *constant* (independent of n_1 and n_2). In the supercritical regime the largest component will already be a *giant component*, i.e. it is unique and of linear size. Similarly, we get a stronger upper bound on the size of all components in the subcritical regime. These results can also be proved using the general framework in [33], however, we give alternative simple proofs.

2.1.3. **Proof outline.** We extend the method employed by Bollobás and Riordan in [35] to study the weakly supercritical regime of $G(n,p)$.

To prove Theorem 2.1 we consider the set S of vertices in 'large' components. The first goal is to show that the size of S is concentrated around $2\varepsilon n$ by applying Chebyshev's inequality. We calculate asymptotically matching upper and lower bounds for the expected size of S by coupling the breadth-first component exploration process from below and above with 2-type branching processes. Once this is done, using a more refined version of this idea, we show that the square of this expectation is an upper bound for the second moment of the size of S, therefore the variance of the size of S is indeed 'small' compared to the square of the expectation and concentration follows by Chebyshev's inequality.

So now we know that whp the appropriate number of vertices lie in 'large' components, but there might be several distinct such components all of which may also be much smaller than claimed in Theorem 2.1. However, we can construct a random graph via a two-round exposure. In the first round we reduce the probability of including some edges by a tiny bit and note that the above arguments will still hold in this setting. In the second round we once again look at each pair not yet connected by an edge and 'sprinkle' an edge with a tiny probability independently for each such pair.

By choosing the magnitude of these probabilities appropriately we can ensure that the resulting random graph has the same distribution as $G(\mathbf{n},P)$ and thus we can identify both random graphs by a coupling argument. Analysing the probability that 'large' components are connected by at least one edge and we use a union bound to show that whp almost all vertices from S lie in a single component of $G(\mathbf{n},P)$.

2.1.4. **Related work.** The general inhomogeneous random graph model $G(n,c\kappa_n)$ studied by Bollobás, Janson, and Riordan [33] is closely related to the model $G(\mathbf{n},P)$, and defined as follows. For any $n \in \mathbb{N}$ consider a random sequence $\mathbf{x}_n = (x_1,\dots,x_n)$ of points from a separable metric space $\mathcal{S}$ equipped with a Borel probability measure ν and let ν_n be the empirical distribution of $\mathbf{x}_n$. Assume that ν_n converges in probability to ν, then the triple $(\mathcal{S},\nu,(\mathbf{x}_n)_{n\geq 1})$ is called a *vertex space*. Furthermore let $\{\kappa_n\}$ be a sequence of symmetric non-negative ν-measurable functions on $\mathcal{S}\times\mathcal{S}$, which converges to a limit κ, and let $c>0$ be a constant. Then the random graph $G(n,c\kappa_n)$ is a graph with vertex set $[n]$, where each pair of vertices $\{k,l\}$ is connected by an edge with probability $p_{k,l} := \min\{1, c\kappa_n(x_k,x_l)/n\}$ independently of all other pairs.

It is proved that with respect to the parameter c there is a phase transition concerning the size of the largest component. In particular, the existence and uniqueness of the giant component in $G(n, c\kappa_n)$ in the supercritical regime are proved using an appropriate multi-type branching process and analysing an integral operator T_κ. The critical point of the phase transition is characterised by $c_0 := ||T_\kappa||^{-1}$: if $c \leq c_0$, then the random graph contains only small components, but if $c > c_0$, then there is a giant component which contains asymptotically $\rho_c n$ vertices, where ρ_c is independent of n and grows linearly in $c - c_0 > 0$.

By contrast the focus of this chapter lies on the *weakly* supercritical regime (Theorem 2.1), i.e. the distance $\varepsilon = \varepsilon(\mathbf{n})$ from the critical point of the phase transition tends to zero as the number of vertices increases. The analysis in this regime is quite sophisticated and in particular much more delicate than in the supercritical regime for $\varepsilon > 0$ being a constant independent of $\mathbf{n}$.

In general it is not sufficient to only scale the edge probabilities multiplicatively as in $G(n, c\kappa_n)$, since even if $\varepsilon \to 0$ the *spectral gap* of the operator $(1+\varepsilon)c_0 T_\kappa$ is always bounded away from 0. In contrast to this, the spectral gap of the offspring expectation matrix in $G(\mathbf{n}, P)$ is given by $\mu_{1,2} + \mu_{2,1}$ and thus may *tend to zero arbitrarily quickly*. Similarly, if one type has significantly fewer vertices than the others, it will not influence the behaviour of $G(n, c\kappa_n)$; however for $G(\mathbf{n}, P)$ in the weakly supercritical regime these vertices may well be crucial for the component structure and ignoring them may even result in a subcritical random graph.

For $\varepsilon \to 0$, an example is given by $n_2 = \sqrt{\varepsilon} n_1$, $\varepsilon^3 n_2 \to \infty$, $\mu_{2,1} = p_{2,1} n_1 = 1$, and thus $\mu_{1,1} = p_{1,1} n_1 = 1 - \sqrt{\varepsilon} + \varepsilon \pm o(\varepsilon)$, so that (2.1) is satisfied. Then $G(\mathbf{n}, P)$ has a unique largest component containing $(2 \pm o(1))\varepsilon n_1 = \omega(n_1^{5/7})$ vertices of type 1, but after removing all vertices of type 2 all components have size at most $o(n_1^{2/3})$.

2.2. Multi-type binomial branching process

Later we will study the component sizes of the random graph $G(\mathbf{n}, P)$ by investigating 2-type binomial branching processes. In this section we investigate some of their most important properties. We start with a key result concerning the survival probability of a general multi-type Galton-Watson branching processes.

Lemma 2.5 (e.g. [68], simplified)**.** *Let $\mathcal{T}_{\mathbf{n},P}$ be a 2-type binomial branching process with parameters $n_1 \in \mathbb{N}$ and $n_2 \in \mathbb{N}$, with $n_1 \geq n_2$, and $P = (p_{i,j})_{i,j\in\{1,2\}} \in (0,1]^{2\times 2}$. Let $\lambda > 0$ be the Perron-Frobenius eigenvalue of its offspring expectation matrix $M = (\mu_{i,j})_{i,j\in\{1,2\}}$, where $\mu_{i,j} = p_{i,j} n_j$, and let (ρ_1, ρ_2) be the pair of survival probabilities. Then the following holds:*

- *if $\lambda \leq 1$, we have $\rho_1 = \rho_2 = 0$;*
- *if $\lambda > 1$, then (ρ_1, ρ_2) is the unique positive solution of*

$$F_1(\rho_1, \rho_2) = F_2(\rho_1, \rho_2) = 0, \tag{2.8}$$

where $$F_i(\rho_1, \rho_2) := 1 - \rho_i - \left(1 - \frac{\mu_{i,1}\rho_1}{n_1}\right)^{n_1} \left(1 - \frac{\mu_{i,2}\rho_2}{n_2}\right)^{n_2}, \text{ for } i \in \{1,2\}. \tag{2.9}$$

We call a branching process that has a positive survival probability *supercritical* and otherwise we call it *subcritical*.

Remark 2.6. There is a very simple way to see that the survival probabilities must satisfy these equations: we consider the extinction probabilities before and after the *first* step of the process and apply the Binomial Theorem.

2.2.1. **Asymptotic survival probability.** Because the conditions of Theorem 2.1 imply that the Perron-Frobenius eigenvalue of the offspring expectation matrix M is strictly larger than 1, as seen in (2.4), the associated branching process will have a positive survival probability that is given implicitly by (2.8).

It is sufficient for us to extract some information about the asymptotic behaviour of the unique positive solution from these equations. However, even trying to solve these equations only asymptotically we have to be very careful with cancellation and take into account higher order terms: this is a major reason why the weakly supercritical regime is significantly harder to analyse than the other regimes.

Lemma 2.7. *Under the conditions as in Theorem 2.1 the survival probabilities of the 2-type binomial branching process $\mathcal{T}_{\mathbf{n},P}$ satisfy*

$$\rho_1 \sim \rho_2 \sim 2\varepsilon.$$

Proof. The key idea is to find suitable bounding functions for the F_i's defined in (2.9), for which the asymptotic values of the zeros can be computed easily, and then to observe that these coincide for the upper and lower bound.

First observe the following fact: if $\rho_1 \geq \rho_2$, we have $F_1(\rho_1, \rho_2) \leq F_1(\rho_1, \rho_1)$ and $F_2(\rho_1, \rho_2) \geq F_2(\rho_2, \rho_2)$; analogously, if $\rho_1 < \rho_2$, then $F_2(\rho_1, \rho_2) < F_2(\rho_2, \rho_2)$ and $F_1(\rho_1, \rho_2) > F_1(\rho_1, \rho_1)$. Thus, without loss of generality due to the Subsubsequence Principle (e.g. [74]), we assume $\rho_1 \geq \rho_2$ and consider the bounding functions $F_i(\rho_i, \rho_i)$, for $i \in \{1, 2\}$:

$$\begin{aligned} F_i(\rho_i, \rho_i) &= 1 - \rho_i - \left(1 - \frac{\mu_{i,1}\rho_i}{n_1}\right)^{n_1} \left(1 - \frac{\mu_{i,2}\rho_i}{n_2}\right)^{n_2} \\ &= 1 - \rho_i - \exp\left(-(\mu_{i,1} + \mu_{i,2})\rho_i - O\left(\frac{\mu_{i,1}^2\rho_i^2}{n_1} + \frac{\mu_{i,2}^2\rho_i^2}{n_2}\right)\right), \end{aligned}$$

by the Taylor-expansion of the natural logarithm around 1. Since $\mu_{i,1} \leq 2$ and $\mu_{i,2} \leq 2$, by the conditions of Theorem 2.1 and the fact that $\rho_i \leq 1$ (since it is a probability), we have

$$\begin{aligned} F_i(\rho_i, \rho_i) &= 1 - \rho_i - \exp\left(-\left(\mu_{i,1} + \mu_{i,2} \pm O\left(n_2^{-1}\right)\right)\rho_i\right) \\ &= 1 - \rho_i - \exp\left(-(1 + \varepsilon_i)\rho_i\right), \end{aligned}$$

where $\varepsilon_i = \mu_{i,1} + \mu_{i,2} - 1 \pm O\left(n_2^{-1}\right) \sim \varepsilon$, by (2.1), and (2.2).

We define $f_i(\rho_i) := 1 - \rho_i - \exp\left(-(1 + \varepsilon_i)\rho_i\right)$ and note that solving $f_i(\rho_i^*) = 0$ asymptotically is a well-known problem that turns up when calculating the asymptotic value of the survival probability for a single-type Poisson branching process.

Using the Taylor-expansion of the natural logarithm we get

$$\varepsilon_i = \frac{-\log(1-\rho_i^*) - \rho_i^*}{\rho_i^*} = \sum_{m=1}^{\infty} \frac{(\rho_i^*)^m}{m+1}.$$

Since the coefficients in this series are all positive and $\varepsilon_i \to 0$, this shows that $\rho_i^* \to 0$ and thus $\varepsilon_i = \frac{\rho_i^*}{2} + O((\rho_i^*)^2)$. Having established the asymptotic behaviour of ρ_1^* and ρ_2^* it remains to show that $\rho_2^* \le \rho_2$ and $\rho_1 \le \rho_1^*$, since this together with $\rho_2 \le \rho_1$ and $\varepsilon_2 \sim \varepsilon_1 \sim \varepsilon$ implies $\rho_2 \sim \rho_1 \sim 2\varepsilon$.

For this last step, assume towards contradiction that $\rho_1 > \rho_1^*$ and observe that f_1 is negative on the interval $(\rho_1^*, 1]$. Since (ρ_1, ρ_2) is by definition a solution of (2.8) we have

$$0 = F_1(\rho_1, \rho_2) \le f_1(\rho_1) < 0,$$

a contradiction. Analogously, $\rho_2 < \rho_2^*$ leads to a contradiction since f_2 is positive on $(0, \rho_2^*)$, completing the proof. □

2.2.2. **Dual process.** In the proof of Theorem 2.1 we consider the supercritical branching process $\mathcal{T}_{\mathbf{n},P}$ associated with $G(\mathbf{n}, P)$ and we will need a good upper bound on the probability its total number of offspring of type $j \in \{1,2\}$ is at least l_j, for carefully chosen real functions l_1 and l_2. Since this probability is 1 if the process survives, this reduces to analysing the conditional probability given the event $\mathcal{D}$ that the process dies out. We call the resulting 2-type binomial branching process the *dual process* and we can describe its offspring distributions as follows.

We need to know, for a vertex v of type $i \in \{1,2\}$ born in generation I_t, for some integer $t \ge 0$, and a potential child u of type $j \in \{1,2\}$, whether the edge $e = \{u, v\}$ is present in the dual process, i.e. conditioned on $\mathcal{D}$. Let $\mathcal{A}_e$ be the event that u is a child of v in $\mathcal{T}_{\mathbf{n},P}$ and note that conditioning on $\mathcal{A}_e$ will decrease the probability of $\mathcal{D}$. More precisely, let $\mathbf{Y} = (Y_1, Y_2)$ denote the vector of the number of individuals of each type in generation I_{t+1}. Since Y_1 is a independent binomially distributed random variables, calculating $\mathbb{P}(\mathcal{D}|\mathcal{A}_e)$ by conditioning on $\mathbf{Y}$ leads to

$$\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e) = \sum_{r_j=0}^{n_j-1} \mathbb{P}(Y_j = r_j + 1 \mid \mathcal{A}_e)(1-\rho_j)^{r_j+1} \cdot \sum_{r_{3-j}=0}^{n_{3-j}} \mathbb{P}(Y_{3-j} = r_{3-j} \mid \mathcal{A}_e)(1-\rho_{3-j})^{r_{3-j}}.$$

Next we observe that by definition we have $\mathbb{P}(Y_j = r_j + 1 \mid \mathcal{A}_e) = \mathbb{P}(Y_j = r_j \mid \neg\mathcal{A}_e)$, for all $r_j = 0, \dots, n_j - 1$. Moreover, Y_{3-j} is independent of $\mathcal{A}_e$, and thus we obtain a similar expression for $\mathbb{P}(\mathcal{D}|\neg\mathcal{A}_e)$. More precisely, we have $\frac{\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)}{\mathbb{P}(\mathcal{D} \mid \neg\mathcal{A}_e)} = 1 - \rho_j$. Therefore we get

$$\mathbb{P}(\mathcal{A}_e|\mathcal{D}) = \frac{\mathbb{P}(\mathcal{D}|\mathcal{A}_e)\,\mathbb{P}(\mathcal{A}_e)}{\mathbb{P}(\mathcal{D}|\mathcal{A}_e)\,\mathbb{P}(\mathcal{A}_e) + \mathbb{P}(\mathcal{D}|\neg\mathcal{A}_e)\,\mathbb{P}(\neg\mathcal{A}_e)} = \frac{p_{i,j}(1-\rho_j)}{1-\rho_j p_{i,j}} =: \pi_{i,j},$$

uniformly for all edges e (with one end point of type i and the other of type j). An analogous calculation shows that the presence of e does not depend on any other

edges, i.e. the dual process is also a 2-type binomial branching process. Hence, we write $\Pi = (\pi_{i,j})_{i,j\in\{1,2\}}$, $H = (h_{i,j})_{i,j\in\{1,2\}}$, where $h_{i,j} := \pi_{i,j}n_j$, and denote the dual process of $\mathcal{T}_{\mathbf{n},P}$ by $\mathcal{T}_{\mathbf{n},\Pi}$.

Intuitively it is obvious that the dual process of any supercritical process is subcritical. For completeness we give a short proof for the processes that we use. First observe that for each pair $(i,j) \in \{1,2\}^2$ we have $p_{i,j} = O(n_j^{-1})$ by (2.2), and thus

$$\pi_{i,j} = p_{i,j}(1-\rho_j)\left(1 + O(n_j^{-2}\rho_j)\right). \tag{2.10}$$

Lemma 2.8. *Let $\mathcal{T}_{\mathbf{n},P}$ be a 2-type binomial branching process satisfying the conditions of Theorem 2.1. Then the offspring expectation matrix $H = (h_{i,j})_{i,j\in\{1,2\}}$ of the dual process $\mathcal{T}_{\mathbf{n},\Pi}$ satisfies*

$$h_{\iota,1} + h_{\iota,2} = 1 - \varepsilon \pm o(\varepsilon), \text{ for } \iota \in \{1,2\}, \tag{2.11}$$

and thus we have $\lambda = 1 - \varepsilon \pm o(\varepsilon)$ for the Perron-Frobenius eigenvalue λ of H.

Proof. By (2.2), (2.10), and Lemma 2.7 we get

$$h_{\iota,1} + h_{\iota,2} = (\mu_{\iota,1} + \mu_{\iota,2})(1-2\varepsilon) = 1 - \varepsilon \pm o(\varepsilon), \text{ for } \iota \in \{1,2\}.$$

The second statement follows analogously to (2.7). □

The benefit of using the *subcritical* dual process $\mathcal{T}_{\mathbf{n},\Pi}$ is that we can bound the expected total number of offspring of each type.

Lemma 2.9. *For $i \in \{1,2\}$ let $\mathcal{T}^i_{\mathbf{n},P}$ be a 2-type binomial branching process satisfying the conditions of Theorem 2.1. Then the associated dual process $\mathcal{T}^i_{\mathbf{n},\Pi}$ satisfies*

$$\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_j\right|\right) \le \varepsilon^{-1}, \text{ for } j \in \{1,2\}. \tag{2.12}$$

Moreover, for any real functions l_1 and l_2, this implies that

$$\mathbb{P}\left(\left|\mathcal{T}^i_{\mathbf{n},P} \cap V_1\right| \ge l_1 \vee \left|\mathcal{T}^i_{\mathbf{n},P} \cap V_2\right| \ge l_2\right) \le 2\varepsilon + \varepsilon^{-1}l_1^{-1} + \varepsilon^{-1}l_2^{-1} \pm o(\varepsilon),$$

and in particular

$$\mathbb{P}\left(\left|\mathcal{T}^i_{\mathbf{n},P} \cap V_1\right| \ge l_1 \vee \left|\mathcal{T}^i_{\mathbf{n},P} \cap V_2\right| \ge l_2\right) \le (2 \pm o(1))\varepsilon, \tag{2.13}$$

if $\varepsilon^2 l_1 \to \infty$ and $\varepsilon^2 l_2 \to \infty$.

Proof. Consider the dual process $\mathcal{T}_{\mathbf{n},\Pi}$. We associate a vertex born in generation I_t, for integer $t \ge 1$, with its *line of ancestry*, i.e. the string $\sigma \in \Sigma_t := \{1,2\}^{t+1}$ which is the finite sequence of types of all its ancestors (starting with the root of $\mathcal{T}_{\mathbf{n},\Pi}$ and including itself). Set $\Sigma := \bigcup_{t\ge1}\Sigma_t$ and denote by Ξ^* the set of all finite strings over the alphabet $\Xi := \{(1,1),(1,2),(2,1),(2,2)\}$.

We consider the injective function $f\colon \Sigma \to \Xi^*$ defined by

$$f\big|_{\Sigma_t}\colon \Sigma_t \to \Xi^*,\ \sigma \mapsto \Big(\big(\sigma(0),\sigma(1)\big),\big(\sigma(1),\sigma(2)\big),\dots,\big(\sigma(t-1),\sigma(t)\big)\Big),$$

for $t \ge 1$. A string $\tau \in \Xi^*$ is called *admissible* if $\tau \in f(\Sigma)$ and we denote the set of admissible strings by $\Xi^{ad} := f(\Sigma)$. Observe that, for every pair $(i,j) \in \{1,2\}^2$,

the function f can be seen as a bijection that maps the subset $\Sigma_{i,j} \subset \Sigma$ of lines of ancestry starting with i and ending in j to the subset $\Xi^{ad}_{i,j} \subset \Xi^{ad}$ of admissible strings starting with $(i,1)$ or $(i,2)$ and ending with $(1,j)$ or $(2,j)$.

Now let $\tilde{g}\colon \Xi \to \mathbb{R}_{>0}$, $(i,j) \mapsto h_{i,j}$ and note $\tilde{g}$ canonically extends to a function

$$g\colon \Xi^* \to \mathbb{R}_{>0},\ \tau \mapsto \prod_{r=0,\dots,t-1} \tilde{g}(\tau(r)).$$

This allows us to compute the expected number of offspring with a fixed line of ancestry $\sigma \in \Sigma$, it is precisely $g(f(\sigma))$. Hence, for $i \in \{1,2\}$, we obtain

$$\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_i\right|\right) = 1 + \sum_{\tau \in \Xi^{ad}_{i,i}} g(\tau) = 1 + G_{i,i}\left(h_{1,1}, h_{1,2}, h_{2,1}, h_{2,2}\right),$$
$$\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_{3-i}\right|\right) = \sum_{\tau \in \Xi^{ad}_{i,3-i}} g(\tau) = G_{i,3-i}\left(h_{1,1}, h_{1,2}, h_{2,1}, h_{2,2}\right),$$

where, for all $(j,j') \in \{1,2\}$, $G_{j,j'}$ is the four-variate ordinary generating function of $\Xi^{ad}_{j,j'}$ (marking the occurrences of $(1,1)$, $(1,2)$, $(2,1)$, and $(2,2)$ respectively).

Using the 'symbolic method' (e.g. [59]) we compute the closed forms of these generating functions providing

$$\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_i\right|\right) = \frac{1 - h_{3-i,3-i}}{d} \qquad \text{and} \qquad \mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_{3-i}\right|\right) = \frac{h_{i,3-i}}{d},$$

for $i \in \{1,2\}$, if the denominator

$$d = 1 - h_{1,1} - h_{2,2} + h_{1,1}h_{2,2} - h_{1,2}h_{2,1}$$

is positive. We now establish stronger lower bounds, which allow us to prove statement (2.12). By (2.11), for any constant $\delta \in (0,1)$ there exists an integer $n_* = n_*(\delta)$ such that for all $n_1 \geq n_2 \geq n_*$ we have

$$\begin{aligned} d &\geq 1 - h_{1,1} - h_{2,2} + h_{1,1}h_{2,2} - (1 - \varepsilon(1-\delta) - h_{1,1})(1 - \varepsilon(1-\delta) - h_{2,2}) \\ &= \varepsilon(1-\delta)\left(2 - h_{1,1} - h_{2,2} - \varepsilon(1-\delta)\right) \\ &\geq \varepsilon(1-\delta)(2 - (1 - \varepsilon(1-\delta)) - h_{3-i,3-i} - \varepsilon(1-\delta)) \\ &= \varepsilon(1-\delta)(1 - h_{3-i,3-i}) \\ &\geq \varepsilon(1-\delta)h_{i,3-i} > 0, \end{aligned}$$

and thus, letting $\delta \to 0$, we obtain

$$\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_i\right|\right) \leq \varepsilon^{-1} \qquad \text{and} \qquad \mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_{3-i}\right|\right) \leq \varepsilon^{-1}.$$

For the second statement we distinguish the two cases of whether $\mathcal{D}_i$ (the event that the primal process $\mathcal{T}^i_{\mathbf{n},P}$ dies out) holds or not. we obtain

$$\begin{aligned} &\mathbb{P}\big(\left|\mathcal{T}^i_{\mathbf{n},P} \cap V_1\right| \geq l_1 \vee \left|\mathcal{T}^i_{\mathbf{n},P} \cap V_2\right| \geq l_2\big) \\ &\qquad\qquad \leq \rho_i + \mathbb{P}\big(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_1\right| \geq l_1\big) + \mathbb{P}\big(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_2\right| \geq l_2\big). \end{aligned}$$

Applying Markov's inequality to the right-hand side we deduce

$$\mathbb{P}\big(\left|\mathcal{T}^i_{\mathbf{n},P} \cap V_1\right| \geq l_1 \vee \left|\mathcal{T}^i_{\mathbf{n},P} \cap V_2\right| \geq l_2\big) \leq 2\varepsilon + \varepsilon^{-1}l_1^{-1} + \varepsilon^{-1}l_2^{-1} \pm o(\varepsilon). \qquad \square$$

2.2.3. **Width of a tree.** The last tool that we need for the proof is the concept of the width of a rooted tree. The *width* $w(\mathcal{T})$ of a rooted tree $\mathcal{T}$ is defined as the supremum of the sizes of all its generations. In this context we will interpret any branching process as a potentially infinite random rooted tree.

Lemma 2.10. *Let $\mathcal{T}_{\mathbf{n},P}$ be a 2-type branching process satisfying the conditions of Theorem 2.1 and denote by $\mathcal{D}$ the event that this process dies out. Then for any real function $m = m(n)$ such that $\varepsilon m \to \infty$ we have*

$$\mathbb{P}\left(\{w\left(\mathcal{T}_{\mathbf{n},P}\right) \geq m\} \cap \mathcal{D}\right) = o(\varepsilon).$$

Proof. Denote $\mathcal{W}_m = \{w\left(\mathcal{T}_{\mathbf{n},P}\right) \geq m\}$ and let us construct $\mathcal{T}_{\mathbf{n},P}$ generation by generation and stop as soon as we see the first generation of size at least m if there is one. Then we have m_1 vertices of type 1 and m_2 vertices of type 2 where $m_1 + m_2 \geq m$. Since each of the vertices of this generation starts an independent copy of $\mathcal{T}^1_{\mathbf{n},P}$ (respectively $\mathcal{T}^2_{\mathbf{n},P}$) we get for the probability of dying out given that $\mathcal{W}_m$ holds

$$\mathbb{P}(\mathcal{D}|\mathcal{W}_m) = (1-\rho_1)^{m_1}(1-\rho_2)^{m_2} \leq e^{-(\rho_1 m_1 + \rho_2 m_2)} = O(\exp(-2\varepsilon m)) = o(1),$$

where the asymptotic statements hold due to Lemma 2.7 and since $\varepsilon m \to \infty$. Hence, we obtain

$$\mathbb{P}(\neg\mathcal{D}|\mathcal{W}_m) = 1 - o(1),$$

and by the law of conditional probability and Lemma 2.7 we have

$$\mathbb{P}\left(\mathcal{W}_m \wedge \mathcal{D}\right) = \mathbb{P}\left(\mathcal{W}_m \wedge \neg\mathcal{D}\right) \cdot \frac{\mathbb{P}(\mathcal{D}|\mathcal{W}_m)}{\mathbb{P}(\neg\mathcal{D}|\mathcal{W}_m)} \leq o(\mathbb{P}(\neg\mathcal{D})) = o(\varepsilon),$$

proving Lemma 2.10. □

2.3. Supercritical regime: proof of Theorem 2.1

The main idea of the proof is to couple the component exploration process in $G(\mathbf{n}, P)$ with instances of the 2-type binomial branching process $\mathcal{T}_{\mathbf{n},P}$.

2.3.1. **Coupling.** Given a vertex v of type i in $G(\mathbf{n}, P)$ we denote its component by $\mathcal{C}_v$. Furthermore let $\mathcal{T}_v$ be the random spanning-tree rooted at v constructed by exploring new neighbours in $\mathcal{C}_v$ via a breadth-first search. Again we interpret branching processes as potentially infinite random rooted trees.

Lemma 2.11. *Given any vector $\mathbf{n} \in \mathbb{N}^2$, any symmetric matrix $P \in [0,1]^{2\times 2}$, and any vertex v of type $i \in \{1,2\}$, the following two statements hold.*

(i) There is a coupling of the random rooted trees $\mathcal{T}_v$ and $\mathcal{T}^i_{\mathbf{n},P}$ such that $\mathcal{T}_v \subset \mathcal{T}^i_{\mathbf{n},P}$. In particular, $|\mathcal{C}_v \cap V_j| \leq \left|\mathcal{T}^i_{\mathbf{n},P} \cap V_j\right|$, for $j \in \{1,2\}$.

(ii) For any vector $\mathbf{m} = (m_1, m_2) \in \mathbb{N}^2$ satisfying $\mathbf{m} \leq \mathbf{n}$, there is a coupling of the random rooted trees $\mathcal{T}_v$ and $\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P}$ such that $\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P} \subset \mathcal{T}_v$ or both trees contain at least m_1 vertices of type 1 or m_2 vertices of type 2. In particular, either $|\mathcal{C}_v \cap V_j| \geq \left|\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P} \cap V_j\right|$, for $j \in \{1,2\}$, or the total number of vertices of type r in $\mathcal{C}_v$ and $\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P}$ is at least m_r for some $r \in \{1,2\}$.

Proof. For the first statement, we generate $\mathcal{T}_v$ and $\mathcal{T}^i_{\mathbf{n},P}$ simultaneously, restoring the set of potential neighbours in the breadth-first search by adding fictional vertices of the same type to the vertex set of $G(\mathbf{n}, P)$ for each vertex already added as a neighbour. Any offspring of a fictional vertex is automatically fictional. In this way, for each type $j \in \{1, 2\}$, we always have n_j potential new neighbours of this type each chosen independently with probability $p_{j',j}$ according to the type $j' \in \{1, 2\}$ of the current vertex. After removal of all fictional vertices from $\mathcal{T}^i_{\mathbf{n},P}$ we obtain $\mathcal{T}_v$. Therefore, we have $(\mathcal{T}_v \cap V_j) \subset \left(\mathcal{T}^i_{\mathbf{n},P} \cap V_j\right)$ and since $|\mathcal{C}_v \cap V_j| = |\mathcal{T}_v \cap V_j|$ the first statement holds.

For the second statement we proceed as before with the slight change that in each step we choose for any type $j \in \{1, 2\}$ exactly $n_j - m_j$ neighbours from all possible new neighbours of type j and only add those independently with probability $p_{j',j}$, where $j' \in \{1, 2\}$ is the type of the current vertex, and ignore all other vertices.

Until we have encountered a total of at least m_r vertices of type r in $\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P}$ for some $r \in \{1, 2\}$ there are always enough vertices of each type to choose from. Assuming that this does not happen for any $r \in \{1, 2\}$, we thus have $\left(\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P} \cap V_j\right) \subset (\mathcal{T}_v \cap V_j) \subset (\mathcal{C}_v \cap V_j)$ and the claim follows. □

2.3.2. **Total size of large components.** Using Lemma 2.9 and Lemma 2.11 we can now establish the expectation of the number of vertices in 'large' components.

For any type $i \in \{1, 2\}$, we denote by $S_{i,\mathbf{L}} = S_{i,\mathbf{L}}(G(\mathbf{n}, P))$ the set of all vertices of type i in components that contain at least l_j vertices of type j, for some $j \in \{1, 2\}$ and a properly chosen pair $\mathbf{L} = (l_1, l_2)$ of real functions. Moreover, we denote by $s_{i,\mathbf{L}} = |S_{i,\mathbf{L}}|$ the cardinality of this set.

Lemma 2.12. *Let l_j be a real function satisfying $\varepsilon^2 l_j \to \infty$, for $j \in \{1, 2\}$. Then*

$$\mathbb{E}(s_{i,\mathbf{L}}) \leq (2 \pm o(1))\varepsilon n_i, \text{ for } i \in \{1, 2\}.$$

Proof. For $i \in \{1, 2\}$, by Lemma 2.11(i) and linearity of expectation, we have

$$\begin{aligned}\mathbb{E}(s_{i,\mathbf{L}}) &= \sum_{v \in V_i} \mathbb{P}(|\mathcal{C}_v \cap V_1| \geq l_1 \vee |\mathcal{C}_v \cap V_2| \geq l_2) \\ &\leq n_i \mathbb{P}\left(\left|\mathcal{T}^i_{\mathbf{n},P} \cap V_1\right| \geq l_1 \vee \left|\mathcal{T}^i_{\mathbf{n},P} \cap V_2\right| \geq l_2\right) \sim 2\varepsilon n_i,\end{aligned}$$

where the last step holds by equation (2.13) in Lemma 2.9. □

Lemma 2.13. *Let l_j be a real function satisfying $l_j = o(\varepsilon n_j)$, for $j \in \{1, 2\}$. Then*

$$\mathbb{E}(s_{i,\mathbf{L}}) \geq (2 \pm o(1))\varepsilon n_i, \text{ for } i \in \{1, 2\}.$$

Proof. We apply Lemma 2.11(ii) with $\mathbf{m} = \mathbf{L} = (l_1, l_2)$, since $l_j = o(\varepsilon n_j)$, for $j \in \{1, 2\}$, and note that the parameters of the coupling branching process satisfy (2.1) and (2.2). Hence, for $i \in \{1, 2\}$, this yields by linearity of expectation

$$\begin{aligned}\mathbb{E}(s_{i,\mathbf{L}}) &\geq n_i \mathbb{P}\left(\left|\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P} \cap V_1\right| \geq l_1 \vee \left|\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P} \cap V_2\right| \geq l_2\right) \\ &\geq n_i \mathbb{P}\left(\mathcal{T}^i_{\mathbf{n}-\mathbf{m},P} \text{ survives}\right) \sim 2\varepsilon n_i,\end{aligned}$$

where the last step holds due to Lemma 2.7. □

In the next lemma we will show that $s_{i,\mathbf{L}}(G(\mathbf{n},P))$, i.e. the number of vertices of type i in large components, is concentrated around its expectation.

Lemma 2.14. *Let l_j be a real function satisfying $\varepsilon^2 l_j \to \infty$ and $l_j = o(\varepsilon n_j)$, for $j \in \{1,2\}$. Then whp*

$$s_{i,\mathbf{L}}(G(\mathbf{n},P)) = (2 \pm o(1))\varepsilon n_i, \text{ for } i \in \{1,2\}.$$

Proof. Lemmas 2.12 and 2.13 show that $\mathbb{E}(s_{i,\mathbf{L}}) \sim 2\varepsilon n_i$, hence it is sufficient to derive the upper bound

$$\mathbb{E}(s_{i,\mathbf{L}}^2) \leq (4 \pm o(1))\varepsilon^2 n_i^2 \sim \mathbb{E}(s_{i,\mathbf{L}})^2 \quad \text{for } i \in \{1,2\}. \tag{2.14}$$

The reason for this is the classical 'second moment method' (e.g. [15, 74]): equation (2.14) implies that for the random variable $s_{i,\mathbf{L}}$ the variance is of smaller order than the square of the expectation, i.e.

$$\mathbb{V}(s_{i,\mathbf{L}}) = \mathbb{E}\left(s_{i,\mathbf{L}}^2\right) - \mathbb{E}(s_{i,\mathbf{L}})^2 \leq o\left(\mathbb{E}(s_{i,\mathbf{L}})^2\right), \text{ for } i \in \{1,2\},$$

which provides concentration by Chebyshev's inequality.

Without loss of generality fix a type $i \in \{1,2\}$ for the rest of the proof. Furthermore, fix a vertex v of type i in $G(\mathbf{n},P)$. Once again we explore the component $\mathcal{C}_v$ of that vertex in a breadth-first search generating a tree $\mathcal{T}_v' \subset \mathcal{C}_v$. However, we will stop the exploration immediately, even midway through revealing the neighbours of one particular vertex, if one of the following two events occurs:

(i) we have already reached a total of l_j vertices of type j for some $j \in \{1,2\}$;
(ii) there are εl_2 vertices that have been reached (i.e. children of earlier vertices) but not yet fully explored (flipped a coin for each possible neighbour).

Note that for stopping condition (ii) we do not distinguish the types of vertices.

Any vertex that has been reached but not fully explored is called *boundary vertex*. Observe that this process will create at most $\varepsilon l_2 + 1 \leq 2\varepsilon l_2$ boundary vertices. Furthermore, denote by $\mathcal{A}$ the event that the process stops due to (i) or (ii), rather than because it has revealed the whole component $\mathcal{C}_v$. Note that

$$\{|\mathcal{C}_v \cap V_1| \geq l_1 \vee |\mathcal{C}_v \cap V_2| \geq l_2\} \implies \mathcal{A}, \tag{2.15}$$

a fact that we will use later on.

Now we estimate the probability that $\mathcal{A}$ holds: by the coupling in Lemma 2.11(i) we may assume that $\mathcal{T}_v' \subset \mathcal{T}_v \subset \mathcal{T}_{\mathbf{n},P}^i$ and, since we proceed in a breadth-first manner, at every point of time all the boundary vertices are contained in at most two consecutive generations. Hence if $\mathcal{A}$ holds, either $\left|\mathcal{T}_{\mathbf{n},P}^i \cap V_j\right| \geq l_j$, for some $j \in \{1,2\}$, or the total number of offspring of the process $\mathcal{T}_{\mathbf{n},P}^i$ is finite and $w(\mathcal{T}_{\mathbf{n},P}^i) \geq \varepsilon l_2/2$. As calculated in Lemma 2.9 the probability that the first case occurs is asymptotically at most 2ε, while for the second case we calculated in Lemma 2.10 that the probability of having large width but still dying out is $o(\varepsilon)$, hence

$$\mathbb{P}(\mathcal{A}) \leq (2 \pm o(1))\varepsilon. \tag{2.16}$$

We use this to relate the second moment to the expectation on the conditional probability space, where we condition on $\mathcal{A}$ holding. We replace $s_{i,\mathbf{L}}$ by a sum of indicator random variables and obtain

$$\mathbb{E}\left(s_{i,\mathbf{L}}^2\right) = \sum_{v\in V_i} \mathbb{E}\left(s_{i,\mathbf{L}} \cdot \mathbb{1}_{\{|\mathcal{C}_v\cap V_1|\geq l_1 \vee |\mathcal{C}_v\cap V_2|\geq l_2\}}\right) \overset{(2.14)}{\leq} n_i \mathbb{P}(\mathcal{A})\mathbb{E}\left(s_{i,\mathbf{L}} | \mathcal{A}\right).$$

Using (2.15) this implies

$$\mathbb{E}\left(s_{i,\mathbf{L}}^2\right) \leq (2\pm o(1))\varepsilon n_i \mathbb{E}\left(s_{i,\mathbf{L}} \mid \mathcal{A}\right). \tag{2.17}$$

For the remainder of this proof we will compute an asymptotic upper bound for the conditional expectation $\mathbb{E}\left(s_{i,\mathbf{L}} \mid \mathcal{A}\right)$. Now for any vertex $u \notin \mathcal{T}'_v$ of type i we reveal its component as before in a breadth-first manner but ignore any vertices that are in $\mathcal{T}'_v$. In other words, we explore in $G' = G(\mathbf{n}, P)\backslash V(\mathcal{T}'_v)$ until we have revealed the whole component in this subgraph, and couple the generated tree $\mathcal{T}''_u$ with $\mathcal{T}^i_{\mathbf{n},P}$ such that $\mathcal{T}''_u \subset \mathcal{T}^i_{\mathbf{n},P}$. We denote by $\mathcal{D}_i$ the event that this instance of $\mathcal{T}^i_{\mathbf{n},P}$ dies out and note, in particular, that $\mathcal{D}_i$ is independent of the event $\mathcal{A}$, hence

$$\mathbb{P}\left(\neg\mathcal{D}_i | \mathcal{A}\right) = \mathbb{P}(\neg\mathcal{D}_i) \overset{L.2.7}{=} (2\pm o(1))\varepsilon. \tag{2.18}$$

Let us further observe that $|\mathcal{T}''_u| \leq |\mathcal{C}_u|$ and furthermore that equality holds unless $G(\mathbf{n}, P)$ contains an edge connecting a boundary vertex to a vertex of $\mathcal{T}''_u$. Therefore, for any given $r \in \mathbb{N}$, we have

$$\mathbb{P}(|\mathcal{C}_u| \neq |\mathcal{T}''_u| \mid \mathcal{D}_i \wedge \mathcal{A} \wedge \{|\mathcal{T}''_u| = r\}) \leq 2\varepsilon l_2 r \max\left\{p_{j,j'} \mid j,j' \in \{1,2\}\right\},$$

by the union bound, as there are at most $2\varepsilon l_2$ boundary vertices. Note that by (2.2) we have $\max\left\{p_{j,j'} \mid j,j' \in \{1,2\}\right\} \leq (1+\varepsilon)n_2^{-1} \leq 2n_2^{-1}$. Hence, by the law of total probability,

$$\mathbb{P}(|\mathcal{C}_u| \neq |\mathcal{T}''_u| \mid \mathcal{D}_i \wedge \mathcal{A}) \leq 4\varepsilon l_2 n_2^{-1}\mathbb{E}\left(|\mathcal{T}''_u| \mid \mathcal{D}_i\right) = 4\varepsilon l_2 n_2^{-1}\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi}\right|\right).$$

In order to simplify notation we will write

$$\mathcal{X}_{u,\mathbf{L}} = \{|\mathcal{C}_u \cap V_1| \geq l_1 \vee |\mathcal{C}_u \cap V_2| \geq l_2\}$$

for the event that the component $\mathcal{C}_u$ is large. Hence it follows that

$$\begin{aligned}
\mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \mid \mathcal{A}) &\leq \mathbb{P}(\neg\mathcal{D}_i) + \mathbb{P}(\mathcal{D}_i)\mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \mid \mathcal{D}_i \wedge \mathcal{A}) \\
&\leq \mathbb{P}(\neg\mathcal{D}_i) + \mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \wedge \{|\mathcal{C}_u| = |\mathcal{T}''_u|\} \mid \mathcal{D}_i \wedge \mathcal{A}) \\
&\qquad + \mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \wedge \{|\mathcal{C}_u| \neq |\mathcal{T}''_u|\} \mid \mathcal{D}_i \wedge \mathcal{A}) \\
&\leq \mathbb{P}(\neg\mathcal{D}_i) + \mathbb{P}(|\mathcal{T}''_u \cap V_1| \geq l_1 \vee |\mathcal{T}''_u \cap V_2| \geq l_2 \mid \mathcal{D}_i) \\
&\qquad + \mathbb{P}(|\mathcal{C}_u| \neq |\mathcal{T}''_u| \mid \mathcal{D}_i \wedge \mathcal{A})
\end{aligned}$$

Using Markov's inequality we express the right-hand side in terms of expectations

$$\mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \mid \mathcal{A}) \leq \mathbb{P}(\neg\mathcal{D}_i) + \frac{\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_1\right|\right)}{l_1} + \frac{\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi} \cap V_2\right|\right)}{l_2} + 4\varepsilon l_2 n_2^{-1}\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},\Pi}\right|\right),$$

and observe that these expectations are all of order $O\left(\varepsilon^{-1}\right)$ by the bound (2.12) in Lemma 2.9. Additionally, by our assumptions on $\mathbf{L}$, the coefficients l_1^{-1}, l_2^{-1} and $4\varepsilon l_2 n_2^{-1}$ are all of order $o(\varepsilon)$ and therefore using (2.18) we get

$$\mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \mid \mathcal{A}) \leq (2 \pm o(1))\varepsilon.$$

This bound applies for at least $n_i - l_i$ vertices of type i and thus we have

$$\mathbb{E}\left(s_{i,\mathbf{L}} \mid \mathcal{A}\right) \leq l_i + (n_i - l_i)\mathbb{P}(\mathcal{X}_{u,\mathbf{L}} \mid \mathcal{A}) \leq l_i + (2 \pm o(1))\varepsilon n_i = (2 \pm o(1))\varepsilon n_i.$$

Inserting this into inequality (2.17) and then applying Chebyshev's inequality completes the proof of Lemma 2.14. □

2.3.3. **Sprinkling.** It remains to prove that almost all vertices in large components lie indeed in a single component. Essentially we exploit the fact that in order for two large components to coexist, a large number of edges cannot be present in the random graph. This however is very unlikely. Due to vertices having two different types the argument is slightly more involved than in the homogeneous setting.

Proof of Theorem 2.1. Let us first introduce some further notation. We write $\alpha := \min\{1, \varepsilon^{-1}\mu_{2,1}\}$ and set

$$\xi := \alpha\varepsilon^3 n_2. \tag{2.19}$$

Note that $\alpha > 0$ and $\xi \to \infty$ by the assumptions of Theorem 2.1. We set

$$l_j := \frac{\varepsilon n_j}{\log \xi} = o(\varepsilon n_j), \text{ for } j \in \{1, 2\}, \tag{2.20}$$

and note that we could replace $\log \xi$ by any function $\hat{\xi}$ such that $\hat{\xi} \to \infty$ but growing very slowly compared to ξ. Moreover, observe that since $\alpha \leq 1$ we have

$$\varepsilon^2 l_j = \frac{\varepsilon^3 n_j}{\log \xi} \geq \frac{\xi}{\log \xi} \to \infty, \text{ for } j \in \{1, 2\}. \tag{2.21}$$

Essentially, we know so far that the random graph $G(\mathbf{n}, P)$ satisfying the conditions of Theorem 2.1 contains the 'right' number of vertices in large components. It only remains to show that all these components are connected, and thus form a single component, if we 'sprinkle' some more edges.

Formally we define a symmetric probability matrix P^b by setting

$$p^b_{1,2} := \frac{\alpha\varepsilon}{n_1 \log \xi} = \min\left\{\frac{\varepsilon}{n_1 \log \xi}, \frac{p_{1,2}}{\log \xi}\right\},$$

and $p^b_{1,1} := p^b_{2,2} := 0$. Then let P^a be the symmetric probability matrix whose off-diagonal entries satisfy $p^a_{1,2} + p^b_{1,2} - p^a_{1,2}p^b_{1,2} = p_{1,2}$ and set $p^a_{i,i} := p_{i,i}$, for $i \in \{1, 2\}$. Then we construct $G(n, P^a)$ and $G(n, P^b)$ independently and couple them in such a way that we have

$$G(\mathbf{n}, P^a) \cup G(\mathbf{n}, P^b) = G(\mathbf{n}, P).$$

Since $p^b_{1,2} \leq p_{1,2}/\log \xi$ we have $p^a_{1,2} \geq p_{1,2}(1 - 1/\log \xi)$ implying that the entries of P^a are all positive for large enough n_1 and n_2. Furthermore, as $p^b_{1,2} = o(\varepsilon/n_1)$, we have $p^a_{1,2} n_i = \mu_{3-i,i} \pm o(\varepsilon)$, for $i \in \{1, 2\}$ and therefore $G(n, P^a)$ also meets all

requirements of Theorem 2.1. Moreover, we have calculated in (2.20) and (2.21) that the further conditions of Lemma 2.14 are also satisfied for $\mathbf{L} = (l_1, l_2)$. Let us denote by $S^a_{i,\mathbf{L}} = S_{i,\mathbf{L}}(G(\mathbf{n}, P^a))$ the set of vertices of type $i \in \{1,2\}$ in large components of $G(\mathbf{n}, P^a)$, i.e. components containing at least l_j vertices of type j for some $j \in \{1,2\}$. Then, by Lemma 2.14, we have whp

$$\left|S^a_{i,\mathbf{L}}\right| = 2\varepsilon n_i + \zeta_i^a$$

for some real function ζ_i^a satisfying $|\zeta_i^a| = o(\varepsilon n_i)$. We assume that this event holds.

Let $\mathcal{U}$ denote the set of all large components in $G(\mathbf{n}, P^a)$. Then for any component $\mathcal{C} \in \mathcal{U}$ we say that the type $j \in \{1,2\}$ is a *witness* for $\mathcal{C}$ being large if $|\mathcal{C} \cap V_j| > \frac{1}{2} l_j$. Observe that having a witness is a necessary condition for any component to be large, hence each large component $\mathcal{C} \in \mathcal{U}$ has at least one witness, yet it is not a sufficient condition.

For $j \in \{1,2\}$ we define the set $\mathcal{U}^j \subset \mathcal{U}$ of large components for that type j is a witness and write $\mathcal{U}^j = \left\{U^j_1, \dots, U^j_{r_j}\right\}$, for some integer $r_j \geq 0$. Intuitively, it should neither of these sets should be empty. We prove this by a counting argument.

Claim 2.15. *$\mathcal{U}^1$ and $\mathcal{U}^2$ are not empty, i.e. $r_1 > 0$ and $r_2 > 0$.*

Proof. Without loss of generality assume towards contradiction that $r_1 = 0$, and thus clearly $r_2 > 0$. Observe that this implies that

$$\left|U^2_\iota \cap V_1\right| \leq \frac{1}{2} l_1 = \frac{\varepsilon n_1}{2 \log \xi} \quad \text{and} \quad \left|U^2_\iota \cap V_2\right| \geq l_2 = \frac{\varepsilon n_2}{\log \xi},$$

for $\iota \in \{1, \dots, r_2\}$. Counting vertices of both types separately, we therefore get

$$2\varepsilon n_1 + \zeta_1^a = \left|S^a_{1,\mathbf{L}}\right| = \sum_{\iota=1}^{r_2} \left|U^2_\iota \cap V_1\right| \leq \frac{r_2 \varepsilon n_1}{2 \log \xi}$$

and

$$2\varepsilon n_2 + \zeta_2^a = \left|S^a_{2,\mathbf{L}}\right| = \sum_{\iota=1}^{r_2} \left|U^2_\iota \cap V_2\right| \geq \frac{r_2 \varepsilon n_2}{\log \xi}.$$

This shows

$$\left(4 + \frac{2\zeta_1^a}{\varepsilon n_1}\right) \log \xi \leq r_2 \leq \left(2 + \frac{\zeta_2^a}{\varepsilon n_2}\right) \log \xi,$$

a contradiction for large enough n_1 and n_2, since $|\zeta_i^a| \varepsilon^{-1} n_i^{-1} = o(1)$, for $i \in \{1,2\}$. Hence Claim 2.15 holds. □

We continue the proof of Theorem 2.1. Observe that by the definition of witnesses we have

$$\left|U^j_\iota\right| \geq \left|U^j_\iota \cap V_j\right| > \frac{1}{2} l_j = \frac{\varepsilon n_j}{2 \log \xi},$$

for $j \in \{1,2\}$ and $\iota \in \{1, \dots, r_j\}$. Hence, if we estimate the number of vertices of type j by only summing over the components in $\mathcal{U}^j$ we get

$$\frac{r_j \varepsilon n_j}{2 \log \xi} < \sum_{\iota=1}^{r_j} \left|U^j_\iota \cap V_j\right| \leq (2 \pm o(1)) \varepsilon n_j,$$

and consequently $r_j \leq 5 \log \xi$, for $j \in \{1,2\}$.

Let $U^1 \in \mathcal{U}^1$ and $U^2 \in \mathcal{U}^2$ be any two large components, i.e. they satisfy $\left|U^1 \cap V_1\right| \geq \frac{1}{2} l_1 = \frac{\varepsilon n_1}{2 \log \xi}$ and $\left|U^2 \cap V_2\right| \geq \frac{1}{2} l_2 = \frac{\varepsilon n_2}{2 \log \xi}$. Then the probability that in $G(\mathbf{n}, P^b)$ there is no edge between U^1 and U^2 is at most

$$(1-p_{1,2}^b)^{|U^1 \cap V_1||U^2 \cap V_2|} \leq \exp\left(-\frac{\alpha\varepsilon}{n_1 \log \xi} \cdot \frac{\varepsilon n_1}{2 \log \xi} \cdot \frac{\varepsilon n_2}{2 \log \xi}\right) \overset{(2.19)}{=} \exp\left(-\frac{\xi}{4 \log^3 \xi}\right).$$

Taking the union bound for (up to) $r_1 + r_2 - 1$ of these events shows that the probability that in $G(\mathbf{n}, P)$ all components that were large in $G(\mathbf{n}, P^a)$ are connected is at least

$$1 - (r_1 + r_2 - 1) \exp\left(-\frac{\xi}{4 \log^3 \xi}\right) \geq 1 - 10 \log \xi \exp\left(-\frac{\xi}{4 \log^3 \xi}\right) = 1 - o(1).$$

Thus, whp there is a component $\mathcal{C}^*$ in $G(\mathbf{n}, P^a) \cup G(\mathbf{n}, P^b) = G(\mathbf{n}, P)$ which contains $S_{i,\mathbf{L}}^a$, for $i \in \{1, 2\}$.

On the other hand, writing $S_{i,\mathbf{L}} = S_{i,\mathbf{L}}\left(G(\mathbf{n}, P)\right)$ for the set of vertices of type $i \in \{1, 2\}$ in large components of $G(\mathbf{n}, P)$ we get $(\mathcal{C}^* \cap V_i) \subset S_{i,\mathbf{L}}$ due to the coupling. Hence we have

$$S_{i,\mathbf{L}}^a \subset (\mathcal{C}^* \cap V_i) \subset S_{i,\mathbf{L}}.$$

Furthermore, applying Lemma 2.14, with the same choice of $\mathbf{L}$, directly to $G(\mathbf{n}, P)$ we obtain whp

$$|S_{i,\mathbf{L}}| = 2\varepsilon n_i + \zeta_i$$

for some real function ζ_i satisfying $|\zeta_i| = o(\varepsilon n_i)$. Thus the number of vertices of type i in the component $\mathcal{C}^*$ satisfies

$$\left| |\mathcal{C}^* \cap V_i| - 2\varepsilon n_i \right| \leq |\zeta_i| + |\zeta_i^a| = o(\varepsilon n_i). \tag{2.22}$$

Moreover, any other large component $\mathcal{C}$ in $G(\mathbf{n}, P)$ may at most contain all the vertices from $S_{1,\mathbf{L}} \setminus S_{1,\mathbf{L}}^a$ and $S_{2,\mathbf{L}} \setminus S_{2,\mathbf{L}}^a$, and therefore satisfies

$$|\mathcal{C} \cap V_i| \leq |S_{i,\mathbf{L}}| - \left|S_{i,\mathbf{L}}^a\right| \leq |\zeta_i| + |\zeta_i^a| = o(\varepsilon n_i), \text{ for } i \in \{1, 2\}.$$

In particular, summing over both types, we have

$$|\mathcal{C}| \leq |\mathcal{C}^*|,$$

for large enough n_1 and n_2. Consequently, $\mathcal{C}^*$ is already the largest component $\mathcal{L}_1\left(G(\mathbf{n}, P)\right)$ and satisfies the required asymptotics by (2.22). □

2.4. Subcritical regime: proof of Theorem 2.4

Most of the work for this regime has already been done in Section 2.2.2, since the associated branching process has essentially the same distribution as the dual process in the weakly supercritical regime (cf. Section 2.2.2).

Proof of Theorem 2.4. Let the conditions be as in Theorem 2.4. Then, analogously to the proof of Lemma 2.9, we calculate the expected total size of the 2-type

binomial branching process $\mathcal{T}^i_{\mathbf{n},P}$, for $i \in \{1,2\}$, and get

$$\mathbb{E}\left(\left|\mathcal{T}^i_{\mathbf{n},P}\right|\right) = \frac{1+\mu_{i,3-i}-\mu_{3-i,3-i}}{1-(\mu_{1,1}+\mu_{2,2}-\mu_{1,1}\mu_{2,2}+\mu_{1,2}\mu_{2,1})} \sim \varepsilon^{-1}.$$

Now let $L = \delta n^{2/3}$, for any fixed constant $\delta > 0$ and write S_L for the set of vertices in components of size at least L and $s_l = |S_L|$. Then with the coupling as in Lemma 2.11(i) we get, by applying Markov's inequality twice and linearity of expectation,

$$\begin{aligned}\mathbb{P}(s_L \geq L) &\leq L^{-1}\mathbb{E}(s_L)\\ &\leq L^{-1}\left(\sum_{v\in V_1}\mathbb{P}(|\mathcal{C}_v|\geq L)+\sum_{v\in V_2}\mathbb{P}(|\mathcal{C}_v|\geq L)\right)\\ &\leq L^{-1}\left(n_1\mathbb{P}\left(\left|\mathcal{T}^1_{\mathbf{n},P}\right|\geq L\right)+n_2\mathbb{P}\left(\left|\mathcal{T}^2_{\mathbf{n},P}\right|\geq L\right)\right)\\ &\leq \varepsilon^{-1}L^{-2}n = (\delta^2\varepsilon n^{1/3})^{-1}\to 0,\end{aligned}$$

since $\varepsilon^3 n \to \infty$ by (2.5). Hence, since $\delta > 0$ was arbitrary, whp all components are of size $o(n^{2/3})$. □

Remark 2.16. This result can be slightly strengthened: let $\xi := \varepsilon^3 n \to \infty$ and replace L by $\hat{L} = \delta n^{2/3}\xi^{-1/6+c}$ for any $0 < c < 1/6$.

2.5. Constant distance from threshold

We complement our close-up analysis of the threshold for the giant component by investigating the behaviour of $G(\mathbf{n},P)$ when the distance ε from the threshold is bounded away from 0.

2.5.1. **Above the threshold.** In the supercritical regime, when the distance from the critical point is a constant, $G(\mathbf{n},P)$ whp has a giant component. The proof is essentially the same as in Section 2.3 except for some of the arguments used for calculating the survival probabilities.

Theorem 2.17. *For $n_1 \in \mathbb{N}$ and $n_2 \in \mathbb{N}$ with $n_1 \geq n_2$, let $n = n_1 + n_2$ and let $\varepsilon > 0$ be a fixed constant. Furthermore, let*

$$P = (p_{i,j})_{i,j\in\{1,2\}} \in (0,1]^{2\times 2}$$

be a symmetric matrix of probabilities satisfying the following conditions:

$$n_2\mu_{2,1} \to \infty, \tag{2.23}$$

$$\mu_{\iota,1}+\mu_{\iota,2} = 1+\varepsilon\pm o(1), \text{ for any } \iota\in\{1,2\}, \tag{2.24}$$

where $\mu_{i,j} = p_{i,j}n_j$ for every pair $(i,j) \in \{1,2\}^2$. Let ρ_ε be the unique positive solution of the equation

$$1-\rho_\varepsilon-\exp(-(1+\varepsilon)\rho_\varepsilon) = 0.$$

Then, whp the following holds for every integer $r \geq 2$ and $i \in \{1,2\}$:

$$|\mathcal{L}_1(G(\mathbf{n},P))\cap V_i| = (\rho_\varepsilon \pm o(1))\,n_i \quad \text{and} \quad |\mathcal{L}_r(G(\mathbf{n},P))\cap V_i| = o(n_i).$$

Therefore, in particular,

$$L_1\left(G(\mathbf{n},P)\right) = (\rho_\varepsilon \pm o(1))n \quad \text{and} \quad L_r\left(G(\mathbf{n},P)\right) = o(n).$$

Proof. Let the conditions be as in Theorem 2.17. We will only show the computation for the survival probabilities, which is very similar to the proof of Lemma 2.7. For the F_i's defined in (2.9) we use the same bounding functions as before.

As in the proof of Lemma 2.7 we assume without loss of generality that $\rho_1 \geq \rho_2$ and thus we have $F_2(\rho_1,\rho_2) < F_2(\rho_2,\rho_2)$ and $F_1(\rho_1,\rho_2) > F_1(\rho_1,\rho_1)$. Fix $i \in \{1,2\}$. We consider the bounding functions $F_i(\rho_i,\rho_i)$:

$$\begin{aligned} F_i(\rho_i,\rho_i) &= 1-\rho_i - \left(1-\frac{\mu_{i,1}\rho_i}{n_1}\right)^{n_1}\left(1-\frac{\mu_{i,2}\rho_i}{n_2}\right)^{n_2} \\ &= 1-\rho_i - \exp\left(-(\mu_{i,1}+\mu_{i,2})\rho_i - O\left(\frac{\mu_{i,1}^2\rho_i^2}{n_1}+\frac{\mu_{i,2}^2\rho_i^2}{n_2}\right)\right), \end{aligned}$$

by the Taylor-expansion of the natural logarithm around 1. Since $\mu_{i,1} \leq 1+2\varepsilon$ and $\mu_{i,2} \leq 1+2\varepsilon$, by the conditions of Theorem 2.17 and the fact that $\rho_i \leq 1$ (since it is a probability), we have

$$\begin{aligned} F_i(\rho_i,\rho_i) &= 1-\rho_i - \exp\left(-\left(\mu_{i,1}+\mu_{i,2} \pm O\left(n_2^{-1}\right)\right)\rho_i\right) \\ &= 1-\rho_i - \exp\left(-(1+\varepsilon_i)\rho_i\right), \end{aligned}$$

where $\varepsilon_i = \mu_{i,1}+\mu_{i,2}-1 \pm O\left(n_2^{-1}\right) \sim \varepsilon$, by (2.23) and (2.24). We set $D = \mathbb{R}_{>0}\times(0,1)$ and a real function f on D by setting

$$f(x,\rho) = 1-\rho-\exp(-x\rho),$$

for $(x,\rho) \in D$. Note that we have $F_i(\rho_i,\rho_i) = f(\varepsilon_i,\rho_i)$.

It is well-known that $f(x,\rho) = 0$ has exactly one solution for any fixed $x > 0$. Furthermore note that the partial derivative with respect to the variable x of f does not vanish on D, therefore we can apply the classical implicit function theorem in $\mathbb{R}^2$. We consider $x = \varepsilon$ and denote by $(\varepsilon,\rho_\varepsilon)$ the corresponding solution of $f = 0$. Hence, there is an open set U with $\varepsilon \in U$ and an open set V with $\rho_\varepsilon \in V$ such that

$$\{(u,g(u)) \mid u \in U\} = \{(u,v) \in U \times V \mid f(u,v) = 0\},$$

where g is a continuous function on U with $\rho_\varepsilon = g(\varepsilon)$. Let $i \in \{1,2\}$. Because $|\varepsilon_i - \varepsilon| = o(1)$, we know that $\varepsilon_i \in U$ for large enough n_1 and n_2, and this implies that $f(\varepsilon_i, g(\varepsilon_i)) = 0$. Since g is continuous we have

$$g(\varepsilon_1) \sim g(\varepsilon_2) \sim \rho_\varepsilon,$$

and it is sufficient to show that $\rho_1 \leq g(\varepsilon_1)$ and $\rho_2 \geq g(\varepsilon_2)$.

For this last step, assume towards contradiction that $\rho_1 > g(\varepsilon_1)$ and observe that $f(\varepsilon_1,\rho) < 0$ for all $\rho \in (g(\varepsilon_1),1]$. Since (ρ_1,ρ_2) is by definition a solution of (2.8) we have $0 = F_1(\rho_1,\rho_2) \leq F_1(\rho_1,\rho_1) = f(\varepsilon_1,\rho_1) < 0$, a contradiction. Analogously, $\rho_2 < g(\varepsilon_2)$ contradicts $f(\varepsilon_2,\rho) > 0$ for all $\rho \in (0,g(\varepsilon_2))$.

Thus we have

$$\rho_1 \sim \rho_2 \sim \rho_\varepsilon, \tag{2.25}$$

and the remainder of the proof follows the lines of the proof of Theorem 2.1 in Sections 2.2 and 2.3, by replacing $\rho_1 \sim \rho_2 \sim 2\varepsilon$ with statement (2.25). □

2.5.2. **Below the threshold.** In the subcritical regime, where the distance to the critical point is a constant, one can obtain a strong upper bound on the size of all components by a standard application of large deviation inequalities.

Theorem 2.18. *For $n_1 \in \mathbb{N}$ and $n_2 \in \mathbb{N}$ with $n_1 \geq n_2$, let $n = n_1 + n_2$ and let $1 > \varepsilon > 0$ be a fixed constant. Furthermore, let*

$$P = (p_{i,j})_{i,j\in\{1,2\}} \in (0,1]^{2\times 2}$$

be a symmetric matrix of probabilities satisfying the following conditions:

$$\mu_{\iota,1} + \mu_{\iota,2} = 1 - \varepsilon \pm o(1), \text{ for any } \iota \in \{1,2\}, \tag{2.26}$$

where $\mu_{i,j} = p_{i,j}n_j$ for every pair $(i,j) \in \{1,2\}^2$. Then we have whp

$$L_1\left(G(\mathbf{n},P)\right) = O(\log n).$$

Proof. Let the conditions be as in Theorem 2.18. We fix a vertex v and explore its component $\mathcal{C}_v$ in $G(\mathbf{n},P)$. Denote the resulting spanning tree by $\mathcal{T}_v$ and couple this process with a 2-type branching process $\mathcal{T}_{\mathbf{n},P}$ as in Lemma 2.11(i) such that $\mathcal{T}_v \subset \mathcal{T}_{\mathbf{n},P}$. Let $\mathcal{S}_L$ be the event that $G(\mathbf{n},P)$ contains a component of size at least L for some appropriately chosen real function L. We want to show that $\mathbb{P}\left(\mathcal{S}_L\right) = o(1)$.

Let us denote the (possibly infinite) sequence of vertices born in $\mathcal{T}_{\mathbf{n},P}$, with respect to the breadth-first exploration, by $\sigma = (v_1, v_2, v_3, \dots)$, where $v_1 = v$.

For any vertex $u \in V_1 \cup V_2$ let X_u be the random variable that counts the number of children of u and has a distribution $\mathrm{Bin}(n_1, p_{j,1}) + \mathrm{Bin}(n_2, p_{j,2})$, where $j \in \{1,2\}$ is the type of u. Then consider the random variables

$$X_{v,L} := \sum_{r=1}^{\min\{L,|\sigma|\}} X_{v_r} \leq \sum_{r=1}^{L} X_{v_r} =: X^*_{v,L},$$

where $\left\{v_{|\sigma|+1}, \dots, v_L\right\}$ is an arbitrary sequence of distinct additional vertices. Notice that $X^*_{v,L}$ is a sum of independent Bernoulli random variables satisfying

$$\left|\mathbb{E}\left(X^*_{v,L}\right) - L(1-\varepsilon)\right| \leq \gamma, \tag{2.27}$$

for some $\gamma = \gamma(n) = o(L)$, by (2.26). Using a Chernoff bound (e.g. [74]) yields

$$\begin{aligned}
\mathbb{P}\left(X^*_{v,L} \geq L-1\right) &\overset{(2.27)}{\leq} \mathbb{P}\left(X^*_{v,L} \geq \mathbb{E}\left(X^*_{v,L}\right) + \varepsilon L - 1 - \gamma\right) \\
&\overset{(2.27)}{\leq} \exp\left(-\frac{(\varepsilon L - 1 - \gamma)^2}{2\left(L\left(1-\varepsilon\right) + \gamma + \left(\varepsilon L - 1 - \gamma\right)/3\right)}\right) \\
&\overset{\gamma=o(L)}{\leq} \exp\left(-\frac{\varepsilon^2}{2 - \frac{4\varepsilon}{3}} L(1 \pm o(1))\right),
\end{aligned} \tag{2.28}$$

uniformly for all vertices $v \in V_1 \cup V_2$.

In order to complete the proof we observe that the event $|\mathcal{T}_{\mathbf{n},P}| \geq L$ implies the event $X_{v,L} \geq L-1$ and therefore we get by application of the union bound

$$\begin{aligned}\mathbb{P}(\mathcal{S}_L) &\leq \sum_{v\in V_1\cup V_2} \mathbb{P}(|\mathcal{C}_v| \geq L) \leq \sum_{v\in V_1\cup V_2} \mathbb{P}(|\mathcal{T}_{\mathbf{n},P}| \geq L)\\ &\leq \sum_{v\in V_1\cup V_2} \mathbb{P}(X_{v,L} \geq L-1) \leq \sum_{v\in V_1\cup V_2} \mathbb{P}\left(X^*_{v,L} \geq L-1\right)\\ &\overset{(2.28)}{\leq} \exp\left(\log n - \frac{\varepsilon^2}{2-4\varepsilon/3} L(1\pm o(1))\right) = o(1),\end{aligned}$$

for any $L > 3\varepsilon^{-2}\log n$, completing the proof. □

Remark 2.19. In fact this proof also works in the weakly subcritical regime and shows that whp $L_1(G(\mathbf{n},P)) \leq 3\varepsilon^{-2}\log n$. Note that if $\varepsilon^3 n(\log n)^{-2} \to \infty$, then this bound is stronger than the one in Theorem 2.4 (even with the strengthening from Remark 2.16).

2.6. Concluding remarks

In the previous sections we showed that the emergence of the giant component in the 2-type random graph $G(\mathbf{n},P)$ is very similar to the behaviour of the binomial random graph $G(n,p)$, at least when each row of the expectation matrix is scaled similarly.

In theory one therefore could study $G_k(\mathbf{n},P)$, the k-type version of $G(\mathbf{n},P)$, assuming that each row of the expectation matrix sums up to approximately $1+\varepsilon$. It is to be expected that in this case we would have $\rho_1 \sim \cdots \sim \rho_k \sim 2\varepsilon$ and thus also a unique largest component of size $L_1(G_k(\mathbf{n},P)) \sim 2\varepsilon n$. Proving this for all $k \geq 3$ would be cumbersome at best, since for instance in our proof the bound on the total expected number of offspring of the dual process relies on explicitly calculating a set of generating functions.

Therefore let us take another perspective: imposing the row-sum conditions ensures that the Perron-Frobenius eigenvalue of the offspring expectation matrix M is roughly $1+\varepsilon$, however it also implies that the corresponding normalised left-eigenvector is not necessarily equal but close to $k^{-1}(1,\dots,1)$. In this spirit we could consider $G_k(\mathbf{n},P)$ for offspring expectation matrices M whose Perron-Frobenius eigenvalue is $1+\varepsilon$ with the corresponding normalised positive left-eigenvalue $\mathbf{v}$ and study how the survival probabilities behave asymptotically.

The Perron-Frobenius theory (Chapter V.6 in [19]) provides a heuristic for this since the properly rescaled offspring vector of generation t of the corresponding branching process converges almost surely to $\mathbf{v}$ as $t \to \infty$, under the assumption that it survives. Thus it would be interesting to know whether in this case it is true that $(\rho_1,\dots,\rho_k) \sim \beta_\varepsilon \mathbf{v}$ for some real function $\beta_\varepsilon = o(1)$.

CHAPTER 3

Emergence of the giant component in random hypergraphs

3.1. Introduction and main results

Random graph theory was founded by Erdős and Rényi in a seminal series of papers from 1959-1968. One of the earliest, and perhaps the most important, result concerned the *phase transition* in the size of the largest component - a very small change in the number of edges present in the random graph dramatically alters the size of the largest component giving birth to the *giant component* [58]. Over the years, this result has been refined and improved, and the properties of the largest component at or near the critical threshold are now very well understood. With modern terminology, and incorporating the strengthenings due to Bollobás [31] and Łuczak [86] we may state the result as follows.

Let $G(n,p)$ denote the random graph on n vertices in which each pair of vertices forms an edge with probability p independently. We consider the asymptotic properties of $G(n,p)$ as n tends to infinity. By the phrase *with high probability* (or *whp*) we mean with probability tending to 1 as n tends to infinity.

Theorem 3.1 ([31, 58, 73, 86])**.** *Let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$.*

(a) If $p = \frac{1+\varepsilon}{n}$, then whp the largest component of $G(n,p)$ has size $(1 \pm o(1))2\varepsilon n$, while all other components have size $O(\varepsilon^{-2}\log(\varepsilon^3 n))$.
(b) If $p = \frac{1-\varepsilon}{n}$, then whp all components of $G(n,p)$ have size $O(\varepsilon^{-2}\log(\varepsilon^3 n))$.

The focus of this chapter lies in a generalisation of this result to hypergraphs. While Theorem 3.1 (and much more) has been known for several decades, for hypergraphs relatively little was known until recently.

Given an integer $k \geq 2$ a k-uniform hypergraph consists of a set V of vertices and a set $E \subset \binom{V}{k}$ of (hyper-)edges, each of which consists of k vertices. (The case $k = 2$ corresponds to a graph.) For hypergraphs there are several meaningful and natural definitions for connectedness. In the following we concentrate on the notion of 'high-order connectedness'. Given an integer $1 \leq j \leq k-1$, two distinct j-sets of vertices (i.e. j-element subsets of the vertex set V) $J_1 \neq J_2$ are said to be *j-connected* if there is a sequence $e_1, \ldots, e_m \in E$ of edges such that

- $J_1 \subset e_1$ and $J_2 \subset e_m$;
- $|e_i \cap e_{i+1}| \geq j$ for all $1 \leq i \leq m-1$.

In other words, we may walk from J_1 to J_2 using edges which consecutively intersect in at least j vertices. Any j-set is always j-connected to itself. This forms an equivalence relation on the set $\binom{V}{j}$ of j-sets. A *j-component* is an equivalence class of this relation, i.e. a maximal set of pairwise j-connected j-sets. The case $j = 1$ is also known as *vertex-connectedness*, and for $j \geq 2$ we use the term *high-order connectedness*.[1]

The case of vertex-connectedness is by far the most studied, not necessarily because it is a more natural definition, but because it is usually substantially easier to understand and analyse.

3.1.1. **Main results.** We consider the natural analogue of the $G(n,p)$ model: let $\mathcal{H}^k(n,p)$ be a random k-uniform hypergraph with vertex set $V = [n]$ in which each k-tuple of vertices is an edge with probability p independently.

Recently, Cooley, Kang, and Person [51], and independently Lu and Peng [84], proved that for all $1 \leq j \leq k-1$, the phase transition for the largest j-component in $\mathcal{H}^k(n,p)$ occurs at the critical probability threshold of

$$\hat{p}_{\mathrm{g}} = \hat{p}_{\mathrm{g}}(n,k,j) := \tfrac{1}{\binom{k}{j}-1}\tfrac{1}{\binom{n}{k-j}}.$$

While the first group already determined the size of the largest j-component up to a multiplicative constant even when $\varepsilon = o(1)$, the second group studied only the simpler regime when $\varepsilon > 0$ is a constant, although they also provide the leading constant of the size of the largest j-component.

We extend these results and determine the size of the largest j-component up to lower order terms immediately after the phase transition, and also show that it is unique in the sense that the size of the second largest j-component is of smaller order. Once we know that it is unique, we refer to it as the *giant component*.

Theorem 3.2. *Let $1 \leq j \leq k-1$ and let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n^{1-2\delta} \to \infty$, for some constant $\delta > 0$.*

(a) If $p = (1+\varepsilon)\hat{p}_{\mathrm{g}}$, then whp the size of the largest j-component of $\mathcal{H}^k(n,p)$ is $(1 \pm o(1))\frac{2\varepsilon}{\binom{k}{j}-1}\binom{n}{j}$ while all other j-components have size at most $o(\varepsilon n^j)$.

(b) If $p = (1-\varepsilon)\hat{p}_{\mathrm{g}}$, then whp all j-components of $\mathcal{H}^k(n,p)$ have size at most $O(\varepsilon^{-2}\log n)$.

The case $k = 2$ (then automatically $j = 1$) is simply Theorem 3.1. The case $j = 1$ for general k was already proved by Schmidt-Pruzan and Shamir [99]. Indeed much more is known for that case and will be illustrated in Section 3.5.

Our proof works for all $k \geq 2$ and $1 \leq j \leq k-1$ (i.e. all permissible pairs j,k). For $j \geq 2$ the condition $\varepsilon^3 n^{1-2\delta}$ arises from our proof method and is probably not optimal. By contrast, for $j = 1$ we do not need any of the machinery which we develop in this chapter, so our proof works under the slightly weaker, optimal assumption $\varepsilon^3 n \to \infty$. Thus Sections 3.2 and 3.4 provide a new, short proof of

[1] This notion is not to be confused with the *(vertex-)connectivity* of a (hyper-)graph H measuring the size of the smallest *vertex-separator* in H.

the result in [99]. Most likely, for general $j \geq 2$, the optimal condition would be $\varepsilon^3 n^j \to \infty$ for which the bounds from the super-critical case ($\Theta(\varepsilon n^j)$) and the sub-critical case ($O(\varepsilon^{-2} \log n)$) match up to the $\log n$ term, suggesting that we have a smooth transition. We discuss the critical window in more detail in Section 3.5.

3.1.2. **Key lemma.** For general j, in the *supercritical regime* (i.e. $p = (1+\varepsilon)\hat{p}_{\mathrm{g}}$), we need an additional, fundamentally new tool called the *smooth boundary lemma* (Lemma 3.8 in Section 3.3), which is the main original contribution of this chapter. In this introduction we will be deliberately vague about definitions and state a simplified version of the smooth boundary lemma (Lemma 3.3 below).

The rest of the argument is in essence very similar to a recent proof of the graph case by Bollobás and Riordan [35]. We will prove Theorem 3.2 using a breadth-first search algorithm to explore the component containing an initial j-set. We refer to the collection of j-sets at fixed distance from the initial j-set as a *generation*. Roughly speaking, the smooth boundary lemma says that during the (supercritical) search process, most generations are 'smooth' in the sense that any set L of size at most $j-1$ lies in approximately the 'right' number of j-sets of the generation.

For any given j-set J we explore its component C_J via a breadth-first search algorithm. Let $\partial C_J(i)$ denote the i-th generation of this process, which we sometimes refer to as the *boundary* of the component after round i. For any $\ell = 1, \dots, j-1$ let $i_0(\ell)$ be the first generation i for which $\partial C_J(i)$ is significantly larger than n^ℓ and for an ℓ-set L let $d_L(\partial C_J(i))$ be the number of j-sets of $\partial C_J(i)$ that contain L.

Let i_1 denote the generation at which the search process hits one of three stopping conditions (S1), (S2), and (S3), which will be stated explicitly in Section 3.3.1.

Lemma 3.3 (Smooth boundary lemma – simplified form)**.** *Let ε, p be as in Theorem 3.2*(a)*. With probability at least $1 - \exp(-n^{\Theta(1)})$, for all J, ℓ, L, i such that*

- *J is a j-set of vertices;*
- *$0 \leq \ell \leq j-1$;*
- *L is an ℓ-set of vertices;*
- *$i_0(\ell) + \Theta(\log n) \leq i \leq i_1$,*

the following holds:

$$d_L(\partial C_J(i)) = (1 \pm o(1)) \frac{|\partial C_J(i)|}{\binom{n}{j}} \binom{n}{j-\ell}.$$

Lemma 3.3 (or its more explicit form, Lemma 3.8) is an interesting result in itself, giving valuable information about the structure of (large) components.

In Chapter 4 we will demonstrate how this lemma also plays a key role in establishing the threshold for j-connectedness in $\mathcal{H}^k(n,p)$. Indeed we derive a stronger result: the hitting time for j-connectedness in the random k-uniform hypergraph process coincides with the moment when the last isolated j-set disappears. Most notably the proof of this result – except for Lemma 3.8 itself – is largely elementary. Thus, we believe that the smooth boundary lemma will prove useful for many other applications in the field of random hypergraphs.

3.1.3. **Intuition and proof outline.** In this section, we relate the study of 'large' components to branching processes, and provide a sketch of the proof of Theorem 3.2(a) using the smooth boundary lemma. We will be informal about quantifications and use terms like 'large' and 'many' without properly defining them. The rigorous definitions can be found in the actual proofs. From now on, whenever we refer to a 'component' we mean a j-component.

Intuition: branching processes. As is the case for graphs, there is a very simple heuristic argument which suggests at what threshold we expect the phase transition to occur and how large we expect the largest component to be. We explore a component of $\mathcal{H}^k(n,p)$ via a breadth-first search process: we begin with one 'active' j-set, then find all edges containing that j-set, thus 'discovering' any further j-sets that they contain, then from each of these new 'active' j-sets in turn we look for any more edges containing them and so on.

The first active j-set is contained in $\binom{n-j}{k-j}$ many k-sets, all of which could potentially be edges. Later on when considering an active j-set we may already have queried some of the k-sets containing it. However, early in the process, $\binom{n}{k-j}$ is a good approximation (and certainly an upper bound) for the number of queries which we make from any j-set. Each of these queries results in an edge with probability p, and for each edge we generally discover $\binom{k}{j}-1$ new j-sets (it could be fewer if some of these j-sets were already discovered, but intuitively this should not happen often).

We may therefore approximate the search process by a branching process $\mathcal{T}^*$: here we represent j-sets by individuals and for each individual its number of children is given by a random variable with distribution $\mathrm{Bin}\big(\binom{n}{k-j},p\big)$ multiplied by $\binom{k}{j}-1$.

The expected number of children of each individual is $\big(\binom{k}{j}-1\big)\binom{n}{k-j}p$. If this number is smaller than 1, then the process always has a tendency to shrink and will therefore die out with probability 1. This roughly corresponds to the initial j-set being in a small component. On the other hand, if the expected number of children is larger than 1, the process has a tendency to grow and therefore has a certain positive probability of surviving indefinitely, corresponding to the initial j-set being in a large component.

The expected number of children is exactly 1 when $p=\hat{p}_{\mathrm{g}}=(\binom{k}{j}-1)^{-1}\binom{n}{k-j}^{-1}$, which is why we expect the threshold to be located there. Furthermore, for $p=(1+\varepsilon)\hat{p}_{\mathrm{g}}$ the survival probability ϱ is asymptotically $\frac{2\varepsilon}{\binom{k}{j}-1}$. This tells us that we should expect about $\varrho\binom{n}{j}$ of the j-sets of $\mathcal{H}^k(n,p)$ to be contained in large components. Moreover, large components should intuitively merge quickly and thus form a unique giant component.

Proof outline: motivating smoothness. As is often the case in such theorems, one half of Theorem 3.2 is easy to prove. Specifically, part (b) can easily be proved using the ideas implemented by Krivelevich and Sudakov [82] for the graph case. We now give an outline of the proof of Theorem 3.2(a). Much of this is similar to

the argument for graphs given by Bollobás and Riordan [35] – we will highlight the points at which new ideas are required.

The main difficulty is to calculate the number X of j-sets in large components. We first compute its expectation $\mathbb{E}(X)$. Now if we can prove that X is well-concentrated around $\mathbb{E}(X)$, then we can show that almost all of these j-sets lie in one large component using a standard sprinkling argument.

The first issue that we encounter is, if we find an edge, how many new j-sets do we discover? It could be as many as $\binom{k}{j}-1$, but on the other hand some of these j-sets may already have been discovered, so perhaps only one of these is genuinely new. (Any k-set which would give no new j-set should not be queried at all.) This is important for two reasons: firstly, it is important for the survival probability of the branching process approximation; and secondly, it may have a significant impact on the size of the component we discover.

Early on in the exploration process (since a very small proportion of j-sets have already been discovered) we should discover $\binom{k}{j}-1$ new j-sets for almost every edge. For an upper coupling, we will simply assume that we discover $\binom{k}{j}-1$ new j-sets for each edge, and couple this process with the branching process $\mathcal{T}^*$. In order to construct a lower coupling we have to be a bit more careful; we only query a k-set if it contains $\binom{k}{j}-1$ previously undiscovered j-sets, thus ensuring that we can define a lower coupling with a branching process $\mathcal{T}_*$ (defined in Section 3.1.4) which has a structure similar to $\mathcal{T}^*$. It follows from the bounded degree lemma in [51] (we use a strengthening of this result: Lemma 3.13) that these two couplings have essentially the same behaviour. For this overview we therefore consider only $\mathcal{T}^*$.

The probability that a j-set lies in a large component, and therefore contributes to X, is approximately the survival probability of the branching process $\mathcal{T}^*$, which is approximately $\frac{2\varepsilon}{\binom{k}{j}-1}$. Thus we have $\mathbb{E}(X) \sim \frac{2\varepsilon}{\binom{k}{j}-1}\binom{n}{j}$. For the second moment we need to consider the probability that two (not necessarily distinct) j-sets J_1 and J_2 are both in large components. We grow a (partial) component C_{J_1} from the j-set J_1 until one of the following three stopping conditions is reached.

(i) the component is fully explored;
(ii) 'many' j-sets have been discovered;
(iii) 'fairly many' j-sets are currently active (i.e. are in the 'boundary' ∂C_{J_1}).

Since we are interested in the probability that both j-sets lie in large components we may assume that we do not stop due to condition (i). Next, note that if stopping condition (iii) is applied, then with high probability J_1 lies in a large component. This is not hard to prove using the branching process approximation – if we have fairly many active individuals, then it is highly probable that the branching process survives. Hence stopping conditions (ii) or (iii) are essentially only applied if the component of J_1 is large. This happens with probability roughly $\frac{2\varepsilon}{\binom{k}{j}-1}$.

We then delete all of the j-sets contained in C_{J_1} from $\mathcal{H}^k(n,p)$ and begin growing a component C_{J_2} from the second j-set J_2 (assuming J_2 itself has not been

deleted), where any k-set containing a deleted j-set can now no longer be queried. Deleting these j-sets ensures that the new search process is independent of the old process, albeit in a restricted random hypergraph. Furthermore, it still follows from the bounded degree lemma (Lemma 3.13) that $\mathcal{T}_*$ will be a lower coupling, and trivially $\mathcal{T}^*$ is still an upper coupling. Once again, we stop the process if C_{J_2} is fully explored or it becomes large, and the probability that it becomes large is approximately $\frac{2\varepsilon}{\binom{k}{j}-1}$.

However, since we deleted some j-sets, it might happen that the search process for J_2 stays small even though the component of J_2 in $\mathcal{H}^k(n,p)$ is large. This can only happen if there is an edge in $\mathcal{H}^k(n,p)$ containing both a j-set of the boundary $\partial C_{J_1} \subset C_{J_1}$ (i.e. it was active when we deleted it), and a j-set of C_{J_2}. We would like to show that the expected number of such edges is $o(1)$, or equivalently, that the number of k-sets containing two j-sets as above is $o(1/p)$, and thus with high probability no such edge exists.

This is the point at which the proof requires new ideas not needed in the graph case. For it is not enough simply to count the number of pairs of j-sets, one from ∂C_{J_1} and one from C_{J_2}, for the following reason: given two j-sets, how many k-sets contain both of them? The answer is heavily dependent on the size of their intersection. Increasing the size of the intersection by just one may lead to an additional factor of n in the number of such k-sets.

We therefore need to know that ∂C_{J_1} and C_{J_2} behave approximately as expected with respect to the size of intersections of j-sets chosen one from each. For this we prove the *smooth stop lemma* (Lemma 3.28). It states that, with (exponentially) high probability, every set L of size $1 \leq \ell < j$ lies in approximately the 'right' number of j-sets of ∂C_{J_1}.

Yet we do not prove Lemma 3.28 directly, but instead we establish a much more powerful tool – the *smooth boundary lemma* (Lemma 3.3, or rather its more precise form Lemma 3.8) – which provides a much better understanding of the way in which components grow. Thus it is certainly also very interesting in its own right and has already proven to be applicable in a much broader context (see Chapter 4).

This allows us to complete the proof of Theorem 3.2(a): we apply Lemma 3.28 to each j-set J of C_{J_2} and thereby determine the number of j-sets of ∂C_{J_1} which intersect J in some subset L up to some small error. As a consequence we obtain a close approximation of the total number of k-sets which we have not queried because of deletions. This shows that with high probability we have not missed any edges from C_{J_2}, and thus the probability that J_1 and J_2 both lie in large components is approximately $\left(\frac{2\varepsilon}{\binom{k}{j}-1}\right)^2$, which multiplied by the number of pairs (J_1, J_2) shows that the second moment is small enough to apply Chebyshev's inequality and deduce that X is concentrated around its expectation.

Note that Lemma 3.3 (respectively Lemmas 3.8 and 3.28) is trivial in the case $j = 1$. In this case the proof of Theorem 3.2 therefore becomes substantially shorter. This suggests a concrete mathematical reason why the case of vertex-connectedness

is genuinely easier than the more general case and not simply easier to visualise. Essentially it comes down to the following: if we restrict our attention to a set of vertices of a given size, then every such set behaves identically. However for sets of j-sets this is far from true.

Structure of the chapter. The chapter is organised as follows.

- In Section 3.2 we quickly prove Theorem 3.2(b).
- In Section 3.3 we investigate the breadth-first search process (with stopping conditions) in detail and prove the smooth boundary lemma (Lemmas 3.3 and 3.8).
- In Section 3.4 we prove Theorem 3.2(a) using the smooth stop lemma (Lemma 3.28), which is based on Lemma 3.8.
- In Section 3.5 we conclude the chapter with a discussion of open questions.

3.1.4. **Notation and setup.** Throughout the chapter we fix integers $k \geq 2$ and $1 \leq j \leq k-1$. We omit floors and ceilings when they do not significantly affect the argument. We use log to denote the natural logarithm (i.e. base e).

All asymptotics in the chapter will be as $n \to \infty$. In particular, we use the phrase *with high probability*, or *whp*, to mean with probability tending to 1 as $n \to \infty$. For functions $f = f(n)$, $g = g(n) > 0$, and $h = h(n) \geq 0$ we write $f = (1 \pm h)g$ to mean $(1 - h(n))g(n) \leq f(n) \leq (1 + h(n))g(n)$ for all sufficiently large $n \in \mathbb{N}$, and similarly with slight abuse of notation we write $f \neq (1 \pm h)g$ to mean that for all sufficiently large $n \in \mathbb{N}$ either $f(n) < (1 - h(n))g(n)$ or $f(n) > (1 + h(n))g(n)$. Furthermore, we also use the notation $f \sim g$ to mean $f = (1 \pm o(1))g$ (or in other words $|f - g| = o(g)$). By the notation $f \ll g$ we mean that $f = o(g)$.

By the notation $c \cdot X$, for a constant $c \in \mathbb{N}$ and an integer-valued random variable X, we mean a random variable Y with distribution given by $\Pr(Y = ci) = \Pr(X = i)$, for any $i \in \mathbb{N}$. (Alternatively $c \cdot X$ may be considered as consisting of c *identical* copies of X – note that it does *not* consist of c *independent* copies of X.)

Given two (real-valued) random variables X, Y, we say that X *is stochastically dominated by* Y if $\Pr(X \geq z) \leq \Pr(Y \geq z)$ for all $z \in \mathbb{R}$.

3.1.4.1. *Parameters.* We will fix a certain set of parameters (in addition to j and k) from now on for the remainder of the chapter. Their relation to one another (and asymptotic behaviour) will be used without explicit reference, since these would have a negative impact on the overall readability.

First of all, for convenience, we define $c_\ell := \binom{k-\ell}{j-\ell} - 1$ for every $0 \leq \ell \leq j-1$. We fix a constant $0 < \delta < 1/6$, and think of it as an arbitrarily small constant (in general our results become stronger for smaller δ). We will have various further parameters throughout the chapter satisfying the following hierarchies:

$$n^{-1/3+2\delta/3} \ll \lambda \ll \varepsilon_* \ll \varepsilon \ll 1$$

and

$$\varepsilon_*, n^{-\delta/24} \ll \eta \ll \gamma_0 (\log n)^{-1} \ll (\log n)^{-1}.$$

We further define $\gamma_\ell := 8^\ell \gamma_0$ for any $1 \le \ell \le j-1$.

Note in particular that for any ε satisfying the conditions of Theorem 3.2 we can choose the remaining parameters such that this hierarchy is satisfied.

3.1.4.2. *Set-degrees.* Given a hypergraph $\mathcal{H}$ and a set L of vertices of $\mathcal{H}$, the *degree* of L in $\mathcal{H}$, denoted by $d_L(\mathcal{H})$, is the number of edges of $\mathcal{H}$ which contain L. For a natural number ℓ, the *maximum ℓ-degree* of $\mathcal{H}$, denoted by $\Delta_\ell(\mathcal{H})$, is the maximum of $d_L(\mathcal{H})$ over all sets L of ℓ vertices of $\mathcal{H}$. When $\ell = 0$, this is simply the number of edges of $\mathcal{H}$.

With slight abuse of notation we will regard a collection $\mathcal{J} \subset \binom{[n]}{j}$ of j-sets (for instance a j-component) as a j-uniform hypergraph $\mathcal{H}_\mathcal{J}$ with vertex set $[n]$ and edge set $\mathcal{J}$. In particular we use $d_L(\mathcal{J})$ and $\Delta_\ell(\mathcal{J})$ instead of $d_L(\mathcal{H}_\mathcal{J})$ and $\Delta_\ell(\mathcal{H}_\mathcal{J})$, respectively.

3.1.4.3. *Exploration process.* We will explore components in $\mathcal{H}^k(n,p)$ via a breadth-first search algorithm: loosely speaking we begin with a j-set J and query all k-sets which contain J to determine whether they form edges. For any that do, the further j-sets they contain are neighbours of J, and for each of these in turn we query k-sets containing them to discover whether they form an edge, and so on.

During this process we denote a j-set as:

- **neutral** if it has not yet been visited by the search process;
- **active** if it has been visited, but not yet fully queried;
- **explored** if it has been fully queried.

We refer to *discovered* j-sets to mean j-sets that are either active or explored, but not neutral.

More formally, we fix a pair $\sigma = (\sigma_j, \sigma_k)$ of total orders, where σ_j is an order on $\binom{[n]}{j}$ and σ_k on $\binom{[n]}{k}$. Given a j-set J we will consider a standard breadth-first search algorithm BFS(J) which, in round i, keeps track of a list $\mathcal{Q}_i$ of discovered j-sets. Initially, the list $\mathcal{Q}_0$ consists of the active j-set J, and all other $\mathcal{Q}_i$ are empty.

In round $i \ge 0$, for each j-set $J' \in \mathcal{Q}_i$ in turn, we query all k-sets containing J' and at least one (still) neutral j-set (one by one, according to σ_k). If a queried k-set K is indeed an edge, then all (still) neutral j-sets $J'' \subset K$ are added at the end of $\mathcal{Q}_{i+1}$ (in order according to σ_j). Then J' is marked as explored, and we proceed with the next j-set in $\mathcal{Q}_i$. Once all j-sets in $\mathcal{Q}_i$ are explored, round i ends. Now if $\mathcal{Q}_{i+1}$ is non-empty, then we proceed with round $i+1$. Otherwise, the process stops.

For $i \ge 0$ we write $C_J(i)$ for the *(partial) component* discovered up to the beginning of round i, i.e. it contains all j-sets from $\mathcal{Q}_0, \dots, \mathcal{Q}_i$ (without order). Similarly, we denote by $\partial C_J(i)$ the *i-th generation* of the process, which is the set of j-sets which are active at the beginning of round i, i.e. the set of all j-sets in $\mathcal{Q}_i$ (without order).

However, for some arguments we need to consider the process on a finer time scale. When speaking of the process BFS(J) at *time t* we mean after the t-th query has taken place, and write $B_J(t)$ for the set of j-sets discovered at that point.

Additionally, we denote the time at which the last query of round $i-1$ takes place by ϑ_i, so in particular we have $B_J(\vartheta_i) = C_J(i)$ for any $i \geq 0$.

For purely technical reasons (see Lemmas 3.13 and 3.15) we want $B_J(t)$ to be defined for all $0 \leq t \leq \binom{n}{k}$. Therefore, once the process has terminated, we query all k-sets which have not been queried yet in an arbitrary order (however no j-sets are considered active any more). Similarly, if the context clarifies the initially active j-set J we usually write BFS instead of BFS(J) for convenience.

3.1.4.4. *Branching processes.* We will want to approximate this search process by branching processes (where individuals represent j-sets):

- $\mathcal{T}^*$ is a branching process in which for each individual the number of children is given by $c_0 \cdot \mathrm{Bin}\big(\binom{n}{k-j}, p\big)$ independently;
- $\mathcal{T}_*$ is a branching process in which for each individual the number of children is given by $c_0 \cdot \mathrm{Bin}\big((1-\varepsilon_*)\binom{n}{k-j}, p\big)$ independently.

It is clear that $\mathcal{T}^*$ forms an upper coupling for BFS. It is less obvious that $\mathcal{T}_*$ is a lower coupling (whp), but this fact will be proved later (see Lemma 3.17).

Remark 3.4. It is for this reason that we need $\varepsilon_* \ll \varepsilon$ – then in the lower coupling $\mathcal{T}_*$ the expected number of children is still approximately $1+\varepsilon$, i.e. the lower coupling and upper coupling are still very similar.

3.1.4.5. *Concentration inequality.* We will use the following form of a Chernoff bound for sums of independent Bernoulli random variables.

Theorem 3.5 (e.g. [74])**.** *Let X be the sum of finitely many i.i.d. Bernoulli random variables. Then for any $\zeta > 0$,*

$$\mathbb{P}\left(X \geq \mathbb{E}(X) + \zeta\right) \leq \exp\left(-\frac{\zeta^2}{2\left(\mathbb{E}(X) + \zeta/3\right)}\right)$$
$$\mathbb{P}\left(X \leq \mathbb{E}(X) - \zeta\right) \leq \exp\left(-\frac{\zeta^2}{2\,\mathbb{E}(X)}\right).$$

We also borrow a lemma from [51], which states that we may pick out a subsequence of queries and treat it like an interval of the search process. To this end we denote by X_t, for each $t = 1, \ldots, \binom{n}{k}$, the indicator random variable associated to the t-th query of the search process.

We will be considering a random subsequence $t_1, t_2, \ldots, t_s$ from $[\binom{n}{k}]$. We say *t_i is determined by the values of $X_1, \ldots, X_{t_i-1}$* to mean the following: for any $a \in [\binom{n}{k}]$, the indicator function of the event $\{t_i = a\}$ is a deterministic function of (the values of) $X_1, \ldots, X_{a-1}$. In particular this means that t_i is chosen before X_{t_i} is revealed.

Lemma 3.6 ([51])**.** *Let $S = (t_1, t_2, \ldots, t_s)$ be a (random, ordered) index set chosen according to some criterion such that*

- *t_i is determined by the values of $X_1, \ldots, X_{t_i-1}$;*
- *with probability 1 we have $1 \leq t_1 < t_2 < \ldots < t_s \leq \binom{n}{k}$.*

Then $(X_{t_1},\dots,X_{t_s})$ is distributed as $(Y_1,\dots,Y_s)$, where $Y_1,\dots,Y_s$ are independent Bernoulli random variables with mean p. In particular, we may apply the Chernoff bound (Theorem 3.5) to $\sum_{i\in S} X_i$.

This lemma applies in particular to subsequences given by the set of queries to k-sets containing a particular ℓ-set while exploring any component, and we therefore have the following corollary.

Corollary 3.7. *With probability at least $1-\exp(-\Theta(n^{\delta/2}))$ for every $1\le \ell \le j-1$, for every ℓ-set L and for every component $\mathcal{C}$, if the number x of queries from j-sets in $\mathcal{C}$ to a k-set containing L satisfies $x \ge n^{k-j+\delta}$, then the number of edges in $\mathcal{C}$ containing L is $(1\pm o(1))px$.*

Proof. First fix ℓ, L and $\mathcal{C}$. By Lemma 3.6 we may apply the Chernoff bound (Theorem 3.5) to the set of queries to k-sets containing L within $\mathcal{C}$. Then the probability that the number of edges in $\mathcal{C}$ containing L is not in $(1\pm\zeta)px$, for some constant $\zeta>0$, is at most

$$2\exp\left(-\frac{(px\zeta)^2}{3px}\right) = 2\exp\left(-px\zeta^2/3\right) \le 2\exp\left(-n^\delta\zeta^2/3\right).$$

We may now apply a union bound over all choices of ℓ and L (of which there are at most n^j) and all choices of $\mathcal{C}$ (of which there are at most n^j) and deduce that the probability that any one of these choices goes wrong is at most

$$2n^{2j}\exp\left(-\zeta^2 n^\delta/3\right) \le \exp\left(-n^{\delta/2}\right) \qquad \square$$

3.2. Subcritical regime: proof of Theorem 3.2(b)

The proof idea is the same as that of the subcritical graph case as proved by Krivelevich and Sudakov [82].

Proof of Theorem 3.2(b). We observe that a component of size m must have at least m/c_0 edges, which were found within an interval of length at most $m\binom{n}{k-j}$ of the search process. Let us consider the probability that an interval of this length contains so many edges. By Lemma 3.6 we may apply the Chernoff bound (Theorem 3.5) and obtain

$$\begin{aligned}\mathbb{P}\left(\mathrm{Bin}\left(m\binom{n}{k-j},(1-\varepsilon)\hat{p}_{\mathrm{g}}\right)\ge m/c_0\right) &\le \exp\left(-\frac{\varepsilon^2 m^2/{c_0}^2}{2\left((1-\varepsilon)m/c_0+\varepsilon m/3c_0\right)}\right)\\ &\le \exp\left(-\frac{\varepsilon^2 m}{2c_0}\right).\end{aligned}$$

If $m \ge 3c_0k\varepsilon^{-2}\log n$, then this probability is at most $n^{-3k/2}=o(n^{-k})$, and therefore we may take a union bound over all possible starting points for the interval, of which there are at most $\binom{n}{k}<n^k$, and still have a probability of $o(1)$. In other words, whp no such interval exists, and therefore no component of size $m\ge 3c_0k\varepsilon^{-2}\log n$ exists. Note that we were not concerned about optimising the bound on m. $\square$

3.3. Smooth exploration

In Section 3.1.3 we illustrated why a thorough understanding of the exploration process BFS (with additional stopping conditions) is crucial to the proof of Theorem 3.2(a) in Section 3.4. We use BFS to grow the component of some j-set J until at the beginning of some round $i \in \mathbb{N}$ one of the following three stopping conditions is reached:

(S1) the component of J is fully explored (i.e. $C_J(i) = C_J(i-1)$);
(S2) the (partial) component $C_J(i)$ has reached size at least λn^j;
(S3) the i-th generation $\partial C_J(i)$ has reached size at least $\lambda^2 n^j$.

Moreover, we denote the (first) round in which any these stopping conditions is invoked[2] by

$$i_1 = i_1(J, \lambda) := \min_{i \in \mathbb{N}} \{(\text{S1}) \vee (\text{S2}) \vee (\text{S3}) \text{ holds in round } i\}. \tag{3.1}$$

Note that this stopping time is crucially influenced by the choice of λ (see Section 3.1.4.1). Our choice for the asymptotic properties of λ will be motivated in the proof of Theorem 3.2 (a) in Section 3.4. Recall that we denoted the time of the last query performed before the start of round i_1 by ϑ_{i_1}.

A technical point on both the second and the third stopping conditions is that we check them only at the beginning of each round, i.e. we do not stop in the middle of a generation. This is in contrast to the stopping procedure in the graph case in [35] – in the hypergraph case, stopping immediately would lead to significant technical difficulties due to the set of active j-sets being spread over two generations. We will show later that the process does not expand too quickly (Lemma 3.16), and thus neither the set of active j-sets nor the whole (partial) component ends up being significantly bigger than the threshold for the stopping condition. Therefore our convention is much more convenient than the one in [35].

We will focus primarily on 'large' components, and thus we define the event $\mathcal{E}$ of stopping due to stopping condition (S2) or (S3) by

$$\mathcal{E} = \mathcal{E}(J, \lambda) := \{(\text{S2}) \vee (\text{S3}) \text{ holds in round } i_1\}. \tag{3.2}$$

3.3.1. **Smooth boundaries.** Our goal is to show that the generations of BFS will eventually become 'smooth'.

The argument is based on various concentration results, and in order for these to be valid we need the average degree of ℓ-sets to be reasonably large. Therefore we will define 'starting rounds' $i_0(\ell)$, prior to which we have no information about the degrees of ℓ-sets, so they may be very 'non-smooth' (in the sense that there is a large disparity between maximum and minimum degree). Over time, though, any disparity will tend to even itself out. We will set $i^*(\ell)$ to be the number of rounds necessary for this process to be complete, and then from round $i_1(\ell)$ onwards, the degrees of ℓ-sets should be smooth.

[2]This is well-defined since BFS always terminates in finite time.

Formally we define these integers as follows: for any $\ell = 1, \dots, j-1$ let

$$i_0(\ell) := \min_{i \in \mathbb{N}} \left\{ |\partial C_J(i)| \geq n^{\ell+\delta} \right\},$$

and set $r_\ell := \frac{c_\ell}{c_0} = \frac{\binom{k-\ell}{j-\ell}-1}{\binom{k}{j}-1}$. Moreover we set $i^*(\ell) := \left\lceil \frac{(j-\ell)\log n}{-\log((1+2\eta+2\varepsilon)r_\ell)} \right\rceil$, and

$$i_1(\ell) := i_0(\ell) + i^*(\ell).$$

We observe that $0 < i^*(\ell) = \Theta(\log n)$ since $r_\ell < 1$ is bounded away from 1. Lastly, we set $i_1(0) := i^*(0) := i_0(0) := 0$.

Lemma 3.8 (Smooth boundary lemma)**.** *Let ε, p be as in Theorem 3.2(a). With probability at least $1 - \exp(-\Theta(n^{\delta/4}))$, using* BFS($J$) *with stopping conditions* (S1)*,* (S2)*, and* (S3)*, for every J, ℓ, L, i such that*

- *J is a j-set of vertices;*
- *$0 \leq \ell \leq j-1$;*
- *L is an ℓ-set of vertices;*
- *$i_1(\ell) \leq i \leq i_1$*

the following holds:

$$d_L(\partial C_J(i)) = (1 \pm \gamma_\ell) \frac{|\partial C_J(i)|}{\binom{n}{j}} \binom{n}{j-\ell}.$$

Furthermore, we have $i_0(0) \leq i_1(0) \leq i_0(1) \leq i_1(1) \leq \cdots \leq i_0(j-1) \leq i_1(j-1)$.

Recall that the parameters γ_ℓ, for $0 \leq \ell \leq j-1$, which we use as error-terms in Lemma 3.8, were defined in Section 3.1.4.1. Note that we may choose $\gamma_\ell \ll \varepsilon$, for $0 \leq \ell < j$, leading to very sharp bounds on the set degrees. However, this assumption is not necessary for our proof.

Lemma 3.8 is a very deep result and one of our major contributions. In order to prove Lemma 3.8 we develop various tools in Sections 3.3.2 to 3.3.6, followed by its proof in Section 3.3.7.

Remark 3.9. Throughout the chapter we will have various claims and lemmas stating that a certain good event holds with very high probability, generally $1 - \exp(-\Theta(n^{\delta/2}))$ (though sometimes also $1 - \exp(-\Theta(n^{\delta/4}))$). Without explicitly stating so, we will subsequently assume that the good event always holds.

More formally, we introduce a new stopping condition for each lemma, and terminate the process if the corresponding good event does not hold. By a union bound over all bad events, as long as there are not too many of them, with very high probability no such stopping condition is ever invoked (note that $P(n) \cdot \exp(-\Theta(n^{\delta/2})) = \exp(-\Theta(n^{\delta/2}))$ for any polynomial P).

A priori it is not obvious that the statement of Lemma 3.8 is not empty, because we do not yet know that $i_1 \geq i_1(\ell)$, and in fact this knowledge is also used in its proof. To this end we show that initially the component grows fast enough that the portion of the component discovered before the start of the smoothing process (for $(j-1)$-sets) is negligible.

Lemma 3.10. *With probability at least* $1 - \exp(-\Theta(n^{\delta/2}))$*, for every* $s = 1, \dots, j$ *and every* $\ell = 0, \dots, s-1$*, if* $i_0(s-1)$ *exists, then we have*

$$\Delta_\ell \left(C_J \left(i_0(s-1) \right) \right) = o(\lambda n^{s-\ell}).$$

This result is surprisingly tricky to obtain, but we prove Lemma 3.10 in Section 3.3.5. Subsequently we then apply Lemma 3.10 to show that $i_1 \geq i_1(\ell)$ for all $0 \leq \ell < j$ conditional on the event $\mathcal{E}$, i.e. in particular when exploring a 'large' component.

Lemma 3.11. *Conditioned on* $\mathcal{E}$*, with probability at least* $1 - \exp(-\Theta(n^{\delta/2}))$*, we have* $i_1 - i_0(j-1) \geq \frac{\delta}{2}\varepsilon^{-1} \log n$

The proof of Lemma 3.11 concludes Section 3.3.5 and is based on Lemma 3.16 providing an upper bound on the expansion of the search process BFS.

3.3.2. **Bounded degrees.** Recall that $B_J(t)$ denotes the set of j-sets which have been discovered by BFS(J) up to time t, i.e. having made precisely t queries so far. We observe that the number of edges $e_{t,J}$ which we have found after t queries is well-concentrated. In particular, this allows us to switch between the time scales of queries and rounds with negligible errors.

Lemma 3.12. *With probability at least* $1 - \exp(-\Theta(n^{\delta/2}))$*, for every* j*-set* J *and every* $t = 0, \dots, \binom{n}{k}$*, the number of edges* $e_{t,J}$ *found in the process* BFS(J) *up to time* t *satisfies*

$$\begin{cases} e_{t,J} = (1 \pm \eta)pt & \text{if } pt \geq n^\delta, \\ e_{t,J} \leq (1+\eta)n^\delta & \text{otherwise.} \end{cases}$$

Proof. Note that $e_{t;J}$ has distribution $\mathrm{Bin}(t, p)$. Let us first consider the case $t \geq n^\delta/p$. Then by the Chernoff Bound (Theorem 3.5) we have

$$\mathbb{P}\left(e_{t,J} \neq (1 \pm \eta)pt \right) \leq 2\exp\left(-\eta^2 n^\delta / 3 \right) \leq 2\exp(-n^{\delta/2}),$$

since $\eta \geq \sqrt{3}n^{-\delta/4}$. Similarly if $t < n^\delta/p$ we have

$$\mathbb{P}\left(e_{t,J} > (1+\eta)n^\delta \right) \leq \mathbb{P}\left(\mathrm{Bin}\left(p^{-1}n^\delta, p \right) > (1+\eta)n^\delta \right) \leq \exp(-n^{\delta/2}).$$

Finally we apply a union bound over all j-sets J and times t to bound from above the probability that does not hold by

$$2\binom{n}{k}\exp(-n^{\delta/2}) \leq \exp(-n^{\delta/2}/2) = \exp(-\Theta(n^{\delta/2}))$$

as required. □

Observe that $e_{t,J}$ can also be seen as a set-degree for $\ell = 0$, in fact we have $e_{t,J} = d_\emptyset(B_J(t)) = \Delta_0(B_J(t))$. The following lemma provides a crude upper bound on the set degrees in $B_J(t)$ for general $0 \leq \ell < j$. It is a more widely applicable strengthening of Lemma 12 from [51].

Lemma 3.13. *With probability at least* $1-\exp(-\Theta(n^{\delta/2}))$*, for every* j*-set* J*, every* $\ell = 0, \dots, j-1$ *and every* $t = 0, 1, \dots, \binom{n}{k}$ *the process* $\mathrm{BFS}(J)$ *with stopping conditions* (S1)*,* (S2)*, and* (S3) *has the following property:*

$$\Delta_\ell(B_J(t)) = O(tpn^{-\ell} + n^\delta).$$

In particular, if for some $i = 0, \dots, i_1$ *and* $\ell = 0, \dots, j-1$ *we have* $|C_J(i)| \geq n^{\ell+\delta}$*, then*

$$\Delta_\ell(C_J(i)) = O(|C_J(i)|/n^\ell).$$

The proof of Lemma 3.13 is based on analysing the degree $d_L(B_J(t))$ by splitting it into two parts.

- A *jump* to L occurs when we query a k-set containing L from a j-set which did not contain L and the k-set forms an edge of $\mathcal{H}^k(n,p)$. Such an edge contributes at most $\binom{k-\ell}{j-\ell}$ to $d_L(B_J(t))$.
- A *pivot* at L occurs when we query any k-set from a j-set containing L and it forms an edge. Such an edge contributes at most $\binom{k-\ell}{j-\ell} - 1$ to $d_L(B_J(t))$.

The concepts of jumps and pivots have already been used in [51]. A significantly more precise investigation of jumps and pivots will prove crucial in the proof of Lemma 3.8 in Section 3.3.7.

Proof of Lemma 3.13. Clearly the contribution of J to the degree of any ℓ-set is at most one and thus negligible, so we ignore it. We prove the statement by induction on ℓ. For $\ell = 0$ the statement follows from Lemma 3.12, so consider any $\ell \geq 0$ and assume that the statement holds for all $0, \dots, \ell-1$.

Fix $0 \leq \ell < j$ and an ℓ-set L, and consider its degree $d_L(B_J(t))$, i.e. the number of j-sets from $B_J(t)$ which contain L. Let $d_L^{(\mathrm{jp})}(B_J(t))$ denote the contribution to $d_L(B_J(t))$ made by jumps to L and $d_L^{(\mathrm{pv})}(B_J(t))$ denote the contribution made by pivots at L.

We start with the contribution $d_L^{(\mathrm{jp})}(B_J(t))$ made by jumps, and first consider the case when $t \geq p^{-1}n^{\ell+\delta}$. In other words, we look at an arbitrary j-set J' and distinguish the size $0 \leq w \leq \ell-1$ of its intersection $W := J' \cap L$ with L. By the induction hypothesis, with probability at least $1-\exp(-\Theta(n^{\delta/2}))$, for any particular W the number of j-sets J' with this intersection is at most

$$\Delta_w(B_J(t)) = O(tpn^{-w}),$$

and each set $J' \cup L$ can be extended to at most $\binom{n}{k-j-\ell+w}$ many k-sets. Thus the number of queries which could potentially lead to a jump to L is at most

$$\sum_{w=0}^{\ell-1} \binom{\ell}{w} O(tpn^{-w}) \binom{n}{k-j-\ell+w} = O(tpn^{k-j-\ell}).$$

Since we are only interested in an upper bound on $d_L^{(\mathrm{jp})}(B_J(t))$ we may assume that this bound is asymptotically tight. By Lemma 3.6 the corresponding

subsequence of successful queries has distribution $\mathrm{Bin}(\Theta(tpn^{k-j-\ell}), p)$, where in particular the mean is of order $\Theta(tpn^{-\ell}) = \Omega(n^\delta)$, as $t \geq p^{-1}n^{\ell+\delta}$. Therefore the Chernoff Bound (Theorem 3.5) implies that with probability at least $1-\exp(\Omega(n^\delta))$ the contribution to $d_L(B_J(t))$ made by jumps is at most

$$O(tpn^{-\ell}). \tag{3.3}$$

On the other hand, if $t < p^{-1}n^{\ell+\delta}$, we simply observe that

$$d_L(B_J(t)) \leq d_L(B_J(p^{-1}n^{\ell+\delta})) = O(n^\delta),$$

where the equality holds with probability $1 - \exp(\Omega(n^\delta))$ by the arguments above.

Thus we have shown that $d_L^{(\mathrm{jp})}(B_J(t)) = O\left(tpn^{-\ell} + n^\delta\right)$ and now aim to show that $d_L^{(\mathrm{pv})}(B_J(t)) = O(d_L^{(\mathrm{jp})}(B_J(t)))$ with probability at least $1 - \exp(\Omega(n^\delta))$.

From every j-set containing L, we make at most $\binom{n}{k-j}$ queries, and each time we discover an edge in this way, it contributes at most $c_\ell = \binom{k-\ell}{j-\ell} - 1$ to $d_L(B_J(t))$. From each of these newly discovered j-sets containing L we make at most $\binom{n}{k-j}$ queries, which could also lead to pivots at L, and this scheme iterates. Thus for an upper bound on $d_L^{(\mathrm{pv})}(B_J(t))$ we may count the total number of vertices contained in $d_L^{(\mathrm{jp})}(B_J(t))$ abstract branching processes each starting at a single vertex and where each vertex has a number of children distributed as $c_\ell \cdot \mathrm{Bin}\big(\binom{n}{k-j} d_L(B_J(t)), p\big)$.

For an upper bound, we also assume $d_L^{(\mathrm{jp})}(B_J(t)) \geq n^\delta$. The following argument appeared in [51] based on ideas in [82], but we repeat it here for completeness.

We think of the branching processes as being subtrees of the infinite $c_\ell N$-ary tree, where $N = \binom{n}{k-j}$ and in which the children of a vertex are partitioned into clusters of size c_ℓ and the edges to such a cluster are all present with probability p, or all absent with probability $1-p$, independently for each cluster. Then we are interested in the size of the subtrees containing the roots, and consider exploring these via a breadth-first search process, always querying a cluster-child of the infinite tree to check whether it is present.

If a tree containing a root has at least s vertices, then we must have found at least $\frac{s-1}{c_\ell}$ cluster-children after having made at most sN queries. Thus if the total size of the $d_L^{(\mathrm{jp})}(B_J(t))$ branching processes is at least $Cd_L^{(\mathrm{jp})}(B_J(t))$, for some constant C, then we must have found at least $(C-1){c_\ell}^{-1}d_L^{(\mathrm{jp})}(B_J(t))$ edges after having made at most $CNd_L^{(\mathrm{jp})}(B_J(t))$ queries.

On the other hand, the number of edges we find in the first $Cd_L^{(\mathrm{jp})}(B_J(t))N$ queries is distributed as $\mathrm{Bin}(CNd_L^{(\mathrm{jp})}(B_J(t)), p)$, which has expectation

$$CNpd_L^{(\mathrm{jp})}(B_J(t)) = C{c_0}^{-1}d_L^{(\mathrm{jp})}(B_J(t)).$$

Note that $c_\ell < c_0$ and therefore for sufficiently large C, by the Chernoff bound (Theorem 3.5) the probability we actually find at least $(C-1){c_\ell}^{-1}d_L^{(\mathrm{jp})}(B_J(t))$ cluster children is at most $\exp(-\Theta(d_L^{(\mathrm{jp})}(B_J(t)))) \leq \exp(-\Theta(n^\delta))$ as required. This proves the first statement.

The second statement is a direct consequence of the first statement: recall that ϑ_i denotes the last query performed in round $i-1$, and so we have $B_J(\vartheta_i) = C_J(i)$. Then by Lemma 3.12 we have with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$ that $\vartheta_i \geq p^{-1}|C_J(i)|/2$. We apply the first statement with $t = \vartheta_i$, and the second claim follows, since $|C_J(i)| \geq n^{\ell+\delta}$ implies $\vartheta_i p n^{-\ell} = \Omega(n^\delta)$. □

To provide a lower bound on the ℓ-degrees in large components, a less sophisticated argument works just fine.

Lemma 3.14. *With probability at least $1 - \exp(-\Theta(n^{\delta/2}))$, for each $0 \leq \ell \leq j-1$ and every component $\mathcal{C}$ of size at least Λ we have the following property: each ℓ-set L is contained in $\Omega(|\mathcal{C}|n^{-\ell})$ j-sets of $\mathcal{C}$.*

Proof. We assume that the high probability event from Corollary 3.7 holds. Consider a component $\mathcal{C}$ with at least Λ many j-sets. We prove by induction on ℓ that any ℓ-set is contained in at least $\left(2\binom{k}{j}^2\right)^{-\ell}|\mathcal{C}|n^{-\ell}$ many j-sets of $\mathcal{C}$. The case $\ell = 0$ simply reasserts the fact that $\mathcal{C}$ has size $|\mathcal{C}|$, so assume $\ell \geq 1$.

Now let $L' \subset L$ be a set of $\ell - 1$ vertices, which by the inductive hypothesis lies in $\left(2\binom{k}{j}^2\right)^{1-\ell}|\mathcal{C}|n^{-\ell+1}$ many j-sets of $\mathcal{C}$. Then for each such j-set J there are at least $\binom{n-j-1}{k-j-1}$ many k-sets containing $J \cup L$ (more if $L \subset J$), and each of these k-sets will be counted in this way at most $\binom{k}{j}$ times. Thus the number of queries from j-sets in $\mathcal{C}$ to a k-set containing L is at least

$$\frac{\binom{n-j-1}{k-j-1}|\mathcal{C}|n^{-\ell+1}}{\left(2\binom{k}{j}^2\right)^{\ell-1}\binom{k}{j}} \geq (1-o(1))\frac{|\mathcal{C}|n^{k-j-\ell}}{2^{\ell-1}\binom{k}{j}^{2\ell-1}(k-j-1)!}.$$

Since $|\mathcal{C}|n^{k-j-\ell} \geq \Lambda n^{k-2j+1} = \lambda n^{k-j+1} \geq n^{k-j+\delta}$, we may apply Corollary 3.7 and thus the number of edges we discover is at least

$$\frac{(1-o(1))p|\mathcal{C}|n^{k-j-\ell}}{2^{\ell-1}\binom{k}{j}^{2\ell-1}(k-j-1)!} \geq \frac{(k-j)!n^{j-k}}{\binom{k}{j}-1} \cdot \frac{|\mathcal{C}|n^{k-j-\ell}}{2^{\ell}\binom{k}{j}^{2\ell-1}(k-j-1)!} \geq \frac{|\mathcal{C}|n^{-\ell}}{2^{\ell}\binom{k}{j}^{2\ell}}.$$

Since each edge contains at least one still neutral j-set, which then becomes active, and furthermore L and $\mathcal{C}$ were arbitrary, this proves the inductive step. □

We will use Lemma 3.14 in the sprinkling argument in the proof of Theorem 3.2(a) in Section 3.4.3. Note that a substantially stronger form of Lemma 3.14 follows (with a little care) from Lemmas 3.8, 3.10, and 3.11. We gave this proof here because it is much more elementary and does not rely on the heavy machinery of the smooth boundary lemma.

3.3.3. **Coupling.** As indicated previously, we aim to use $\mathcal{T}_*$ and $\mathcal{T}^*$ as lower and upper couplings on BFS. Let us describe more precisely what we mean by this.

In our search process we will certainly always make at most $\binom{n}{k-j}$ queries from any j-set. If the actual number is $x \leq \binom{n}{k-j}$, then we identify these with the first x queries from an individual of $\mathcal{T}^*$ (which we view as a subgraph of the infinite $\binom{n}{k-j}$-ary tree in which each edge is present with probability p, and we consider the

subtree containing the root). The remaining $\binom{n}{k-j} - x$ queries in $\mathcal{T}^*$ are in effect 'dummy queries' which do not exist in the search process, but making extra queries is permissible for an upper bound. Thus if we are at time t in the search process, then $\mathcal{T}^*$ may have made more than t queries. However, since we will generally be considering generations of the search process rather than the exact time, this difference does not affect anything.

Similarly we can couple the search process with $\mathcal{T}_*$ by ignoring some queries which the search process makes, and considering only those queries which would give $c_0 = \binom{k}{j} - 1$ new j-sets, and of these consider only the first $(1-\varepsilon_*)\binom{n}{k-j}$. Of course, this requires that there are at least this many such queries in the search process, which we will prove using Lemma 3.13. We will denote this coupling of processes (when it holds) by $\mathcal{T}_* \prec \mathrm{BFS} \prec \mathcal{T}^*$.

Lemma 3.15. *With probability at least* $1 - \exp(-\Theta(n^{\delta/2}))$*, for every* j*-set* J *and every time* t *satisfying* $t \ll \varepsilon_* n^k$ *the process* $\mathrm{BFS}(J)$ *satisfies*

$$\mathcal{T}_* \prec \mathrm{BFS}(J) \prec \mathcal{T}^*.$$

Proof. The upper coupling is immediate from the definitions. The lower coupling follows from Lemma 3.13. More precisely, from any j-set J and time t we can bound from above the number of discovered j-sets which intersect J in ℓ vertices by $\binom{j}{\ell}\Delta_\ell(B_J(t))$. For each such discovered j-set, the number of k-sets containing its union with J is at most $\binom{n}{k-2j+\ell}$. Thus by Lemma 3.13 the number of queries from J which would give fewer than c_0 new j-sets is at most

$$\sum_{\ell=0}^{j-1} \binom{j}{\ell} \Delta_\ell(B_J(t)) \binom{n}{k-2j+\ell} = \sum_{\ell=0}^{j-1} O\left((ptn^{-\ell} + n^\delta) n^{k-2j+\ell}\right) = o(\varepsilon_* n^{k-j}),$$

since $ptn^{-j}, n^{-1+\delta} \ll \varepsilon_*$. Thus the number of k-sets which may still be queried from J and which would give c_0 new j-sets is at least $(1-\varepsilon_*)\binom{n}{k-j}$, establishing the lower coupling. □

Note that the lower coupling in Lemma 3.15 only holds early in the process. However, it follows from stopping conditions (S1), (S2), and (S3) that, with exponentially high probability, we stop after having performed at most $O(\lambda n^k) = o(\varepsilon_* n^k)$ queries (see Lemma 3.17).

3.3.4. **Bounded expansion.** Next we prove that the expansion of the search process is approximately as fast as we expect (once the generations become large). Note that for small generations there is only an upper bound, which is also much weaker than for larger generations.

Lemma 3.16. *For a* j*-set* J *and a round* i *such that* $\mathcal{T}_* \prec \mathrm{BFS}(J) \prec \mathcal{T}^*$*, conditioned on* $|\partial C_J(i)| = x \in \mathbb{N}$*, with probability at least* $1 - \exp(-\Theta(n^{\delta/2}))$ *we have*

$$|\partial C_J(i+1)| \begin{cases} = (1 \pm 2\varepsilon_*)(1+\varepsilon)x & \text{if } x \geq n^{1-\delta} \\ \leq 2\max(x, n^\delta) & \text{otherwise.} \end{cases}$$

Proof. Note that since $\mathcal{T}^*$ is an upper coupling, conditioned on $|\partial C_J(i)| = x$, the size of the next generation $|\partial C_J(i+1)|$ is stochastically dominated by a random variable Y_{i+1} with distribution $c_0 \cdot \mathrm{Bin}\big(x\binom{n}{k-j}, p\big)$. We have

$$\mathbb{E}(Y_{i+1}) = c_0 x \binom{n}{k-j} p = (1+\varepsilon)x.$$

Furthermore, by the Chernoff bound (Theorem 3.5) we have

$$\begin{aligned}
\mathbb{P}(|\partial C_J(i+1)| \geq (1+2\varepsilon_*)(1+\varepsilon)x) &\leq \mathbb{P}(Y_{i+1} \geq (1+2\varepsilon_*)(1+\varepsilon)x) \\
&= \mathbb{P}\left(\tfrac{Y_{i+1}}{c_0} \geq (1+2\varepsilon_*)\mathbb{E}\left(\tfrac{Y_{i+1}}{c_0}\right)\right) \\
&\leq \exp\left(\frac{-(2\varepsilon_*)^2\mathbb{E}(Y_{i+1})}{3c_0}\right) \\
&= \exp(-\Theta({\varepsilon_*}^2 x)) \\
&\leq \exp(-\Theta(n^{-1+2\delta}n^{1-\delta})) \\
&\leq \exp(-\Theta(n^{\delta/2}))
\end{aligned}$$

where for the penultimate inequality we used the fact that $\varepsilon_* \geq n^{-1/2+\delta}$. This concludes the proof of the upper bound for the first half of the statement.

For the lower bound, we note that since we are in the range where $\mathcal{T}_*$ will give us a lower bound, $|\partial C_J(i+1)|$ stochastically dominates a random variable Z_{i+1} with the distribution of $c_0 \cdot \mathrm{Bin}\big(x(1-\varepsilon_*)\binom{n}{k-j}, p\big)$. Thus we have, again by the Chernoff bound (Theorem 3.5),

$$\begin{aligned}
\mathbb{P}\left(|\partial C_J(i+1)| \leq (1-2\varepsilon_*)(1+\varepsilon)x\right) &\leq \mathbb{P}\left(Z_{i+1} \leq (1-2\varepsilon_*)(1+\varepsilon)x\right) \\
&\leq \mathbb{P}\left(\tfrac{Z_{i+1}}{c_0} \leq (1-\varepsilon_*)\mathbb{E}\left(\tfrac{Z_{i+1}}{c_0}\right)\right) \\
&\leq \exp\left(\frac{-(\varepsilon_*)^2\mathbb{E}(Z_{i+1})}{2c_0}\right) \\
&= \exp\left(-\Theta\left({\varepsilon_*}^2 x\right)\right) \\
&\leq \exp(-\Theta(n^{\delta/2}))
\end{aligned}$$

provided that $x \geq n^{1-\delta}$.

For the second half, the calculation is very similar. $|\partial C_J(i+1)|$ is dominated by a random variable Y_{i+1} with distribution $c_0 \cdot \mathrm{Bin}\big(n^\delta\binom{n}{k-j}, p\big)$, and so

$$\begin{aligned}
\mathbb{P}\left(|\partial C_J(i+1)| \geq 2n^\delta\right) &\leq \mathbb{P}\left(Y_{i+1} \geq 2n^\delta\right) \\
&\leq \exp\left(\frac{-(1-\varepsilon)^2\mathbb{E}(Y_{i+1})}{3c_0}\right) \\
&= \exp\left(-\Theta\left(n^\delta\right)\right) \\
&\leq \exp(-\Theta(n^{\delta/2})).
\end{aligned}$$

□

Observe that in round $i_0(\ell)$ we have $|\partial C_J(i_0(\ell))| \geq n^{\ell+\delta} \geq n^{1-\delta}$, for $\ell \geq 1$. This means that, by Lemma 3.16 the size of the generations will never decrease

from that moment onwards for as long as $\mathcal{T}_*$ remains a valid lower coupling. This will be important for various concentration results.

Recall that ϑ_{i_1} denotes the time of the last query before round i_1 started, and in this round one of the stopping conditions is hit.

Lemma 3.17. *With probability at least* $1 - \exp(-\Theta(n^{\delta/2}))$*, for any* j*-set* J *using* BFS(J) *with stopping conditions* (S1)*,* (S2)*, and* (S3) *we have*

$$\vartheta_{i_1} = O(\lambda n^k).$$

Thus, in particular, we have

- *for all* $i = 0, \dots, i_1$*:*
$$\mathcal{T}_* \prec \mathrm{BFS}(J) \prec \mathcal{T}^*;$$
- *for all* $\ell = 1, \dots, j-1$ *and* $i = i_0(\ell), \dots, i_1$*:*
$$|\partial C_J(i)| \geq n^{\ell+\delta};$$
- $|C_J(i_1)| \leq 2\lambda n^j$ *and* $|\partial C_J(i_1)| \leq 2\lambda^2 n^j$*.*

Proof. We observe that the upper bound in Lemma 3.16 only relies on the upper coupling BFS $\prec \mathcal{T}^*$ which is always satisfied. Thus $|C_J(i_1)| \leq 2\lambda n^j$ (as well as $|\partial C_J(i_1)| \leq 2\lambda^2 n^j$), and consequently by Lemma 3.12 we have $\vartheta_{i_1} = O(\lambda n^k)$.

Now this immediately implies that the lower coupling also holds for all $0 \leq i \leq i_1$ by Lemma 3.15, proving the second assertion. The remaining statements follow from Lemma 3.16. □

3.3.5. **Initial component growth.** We first aim to prove Lemma 3.10, which states that if round $i_0(s-1)$ exists, for some $2 \leq s \leq j$, then the maximal ℓ-degrees in the partial component $C_J(i_0(s-1))$ are still small, and in particular, BFS(J) has not yet found a significant number of j-sets. The critical tool for proving Lemma 3.10 is Lemma 3.18, which gives a lower bound on the probability that the process will become large within a certain number of rounds.

Recall that $\mathcal{T}_*$ is a branching process starting with one individual in the 0-th generation and whose offspring is given by the random variable

$$c_0 \cdot \mathrm{Bin}\left((1-\varepsilon_*)\binom{n}{k-j}, p\right).$$

We have also seen in Lemma 3.17 that this provides a lower coupling $\mathcal{T}_* \prec \mathrm{BFS}$ as long as we have not reached any stopping condition (which happens in round i_1).

Let $c > 0$ be some constant and let

$$\tau := (s - 1 + \delta + c)\varepsilon^{-1} \log n. \tag{3.4}$$

For any $i \in \mathbb{N}$ we denote by $|\partial_*(i)|$ the size of the i-th generation of $\mathcal{T}_*$.

Lemma 3.18. *With probability at least* εn^{-c} *we have* $|\partial_*(\tau)| \geq n^{s-1+\delta}$*.*

We aim to imitate a proof of Markov's inequality. However, since we have no upper limit on the size of each generation, we will need to limit the contribution that the upper tail makes to the expected of $|\partial_*(i)|$.

In order to do this we need to bound the upper tail probabilities, preferably exponentially. For this we imitate a proof of a Chernoff bound, applying Markov's inequality to the random variable $e^{\eta|\partial_*(\tau)|}$ for some well chosen η.

We therefore need to calculate $\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right)$. We define

$$\eta := \frac{\varepsilon}{(1+\varepsilon)^\tau n^{c/2}} \tag{3.5}$$

set $C_i := 1 + \delta + i\varepsilon$ for $0 \le i \le \tau$, and observe that

$$C_i \le 1 + \delta + \tau\varepsilon \overset{(3.4)}{=} O(\log n). \tag{3.6}$$

Claim 3.19. *For all integers $0 \le i \le \tau$ we have*

$$\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right) \le 1 + C_i\,(1+\varepsilon)^i\,\eta\,,$$

for sufficiently large n.

Proof. Since $\mathcal{T}_*$ is a Galton-Watson branching process, each subtree rooted at the first generation has the same distribution and thus we obtain the following recursion

$$\begin{aligned}\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right) &= \sum_{t=0}^{\infty} \Pr\left(|\partial_*(1)| = t\right) \mathbb{E}\left(e^{\eta|\partial_*(i)|} \mid |\partial_*(1)| = t\right) \\ &= \sum_{t=0}^{\infty} \Pr\left(|\partial_*(1)| = t\right) \mathbb{E}\left(e^{\eta|\partial_*(i-1)|}\right)^t \end{aligned} \tag{3.7}$$

We prove the claim by induction on i. For the case $i = 0$ the statement becomes $e^\eta \le 1 + C_0\eta$ which is certainly true for sufficiently large n since $\eta = o(1)$ and $C_0 > 1$ is bounded away from 1. We therefore assume that the result holds for $i - 1$.

To simplify notation we define $x_i := (1+\varepsilon)^i\eta$. Observe that $x_i \le x_\tau = \varepsilon/n^{c/2}$ and consequently, by (3.6), for all $0 \le i \le \tau$, we have

$$C_i x_i = o(\varepsilon n^{-c/4}). \tag{3.8}$$

Also, let

$$p_i := \left(1 + C_{i-1}x_{i-1}\right)^{c_0} p$$

and note that for $0 \le i \le \tau$ we have $p \le p_i \le p_\tau = (1 + o(1))p$ by (3.8). This guarantees in particular that

$$p_i \ll \varepsilon n^{-c/4} \ll 1. \tag{3.9}$$

We now have

$$\begin{aligned}\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right) &\overset{(3.7)}{\le} \sum_s \binom{\binom{n}{k-j}}{s} p^s (1-p)^{\binom{n}{k-j}-s}\, \mathbb{E}\left(e^{\eta|\partial_*(i-1)|}\right)^{c_0 s} \\ &\le \sum_s \binom{\binom{n}{k-j}}{s} p^s (1-p)^{\binom{n}{k-j}-s}\, \left(1 + C_{i-1}x_{i-1}\right)^{c_0 s}\,, \end{aligned}$$

and furthermore obtain

$$\begin{aligned}\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right) &\leq \sum_s \binom{\binom{n}{k-j}}{s} p_i^s (1-p)^{\binom{n}{k-j}-s} \\ &\leq \left(\frac{1-p}{1-p_i}\right)^{\binom{n}{k-j}} \sum_s \binom{\binom{n}{k-j}}{s} p_i^s (1-p_i)^{\binom{n}{k-j}-s}.\end{aligned}$$

The terms in the sum are simply binomial probability expressions, now with a slightly different probability, and therefore their sum is 1. Thus we obtain

$$\begin{aligned}\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right) &\leq \left(\frac{1-p}{1-p_i}\right)^{\binom{n}{k-j}} \\ &= \left(1+(p_i-p)(1+p_i+p_i^2+\ldots)\right)^{\binom{n}{k-j}} \\ &\leq 1+\binom{n}{k-j} p \left(\frac{p_i}{p}-1\right)(1+O(p_i)) + O\left(\left(\frac{p_i}{p}-1\right)^2\right).\end{aligned}$$

Now

$$\frac{p_i}{p} - 1 = c_0 C_{i-1} x_{i-1} + O({C_{i-1}}^2 {x_{i-1}}^2)$$

so we have

$$\begin{aligned}\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right) &\leq 1 + (1+\varepsilon)\left(C_{i-1}x_{i-1} + O({C_{i-1}}^2 {x_{i-1}}^2)\right)\left(1 + O(p_i + {C_{i-1}}^2 {x_{i-1}}^2)\right) \\ &\leq 1 + x_i\left(C_{i-1} + O({C_{i-1}}^2 x_{i-1}) + O(C_{i-1}p_i)\right) \\ &\overset{(3.6),(3.8),(3.9)}{=} 1 + x_i\left(C_{i-1} + o(\varepsilon n^{-c/4}\log n)\right) \\ &\leq 1 + C_i x_i\end{aligned}$$

for any sufficiently large $n \in \mathbb{N}$. □

Having obtained a suitable upper bound on $\mathbb{E}\left(e^{\eta|\partial_*(i)|}\right)$ we return to the proof of Lemma 3.18.

Proof of Lemma 3.18. We imitate a proof of Markov's inequality. For any real $a \geq 0$, let

$$q_a := \mathbb{P}(|\partial_*(\tau)| \geq a)$$

We set $z := n^{s-1+\delta}$ and $y_i := (1+\varepsilon)^\tau n^{ic}/(2\varepsilon)$ for every $i = 1, 2, \ldots$ and note that

$$\begin{aligned}y_1 &= \varepsilon^{-1} n^{(s-1+\delta+c)\varepsilon^{-1}\log(1+\varepsilon)} n^c \\ &\geq \varepsilon^{-1} n^{s-1+\delta+3c/2} \\ &> z.\end{aligned}$$

Our main aim is to find a lower bound on q_z. We observe that

$$\mathbb{E}(|\partial_*(\tau)|) \leq (1-q_z)z + q_z y_1 + \sum_{i=1}^{\infty} q_{y_i} y_{i+1}. \tag{3.10}$$

Observe that we have

$$C_\tau \tau (1+\varepsilon)^\tau \eta \overset{(3.5)}{=} C_\tau \tau \varepsilon n^{-c/2} \overset{(3.4),(3.6)}{=} o(1)$$

and thus together with Markov's inequality we obtain

$$q_{y_i} \leq \frac{\mathbb{E}\left(e^{\eta|\partial_*(\tau)|}\right)}{e^{\eta y_i}} \overset{C.3.19}{\leq} \frac{1 + C_\tau \tau \left(1+\varepsilon\right)^\tau \eta}{e^{n^{(i-1/2)c}/2}} \leq e^{-n^{ic/3}}$$

for n sufficiently large. Hence

$$\sum_{i=1}^{\infty} q_{y_i} y_{i+1} \leq \frac{(1+\varepsilon)^\tau}{2\varepsilon} \sum_{i=1}^{\infty} e^{-n^{ic/3}} n^{(i+1)c},$$

and since $a^{-1}\log(1+a) \leq 1$ for any real number a we further obtain

$$\sum_{i=1}^{\infty} q_{y_i} y_{i+1} \overset{(3.4)}{\leq} \varepsilon^{-1} n^{s-1+\delta+c} e^{-n^{c/4}} = o(e^{-n^{c/5}}).$$

We have thus shown that in (3.10) the contribution to $\mathbb{E}(|\partial_*(\tau)|)$ from the upper tail is negligible. Now (3.10) tells us that

$$q_z y_1 - q_z z \geq \mathbb{E}(|\partial_*(\tau)|) - z - o(e^{-n^{c/5}})$$

and therefore

$$q_z \geq \frac{(1+\varepsilon)^\tau - n^{s-1+\delta} - o(e^{-n^{c/5}})}{(1+\varepsilon)^\tau n^c/(2\varepsilon) - n^{s-1+\delta}} \geq \frac{\varepsilon(1+\varepsilon)^\tau(1 \pm o(1))}{2n^c(1+\varepsilon)^\tau(1 \pm o(1))} \geq \varepsilon n^{-c},$$

where we used the fact that $n^{s-1+\delta} = o((1+\varepsilon)^\tau)$. This completes the proof of Lemma 3.18. □

We now show how to deduce Lemma 3.10 from Lemma 3.18.

Proof of Lemma 3.10. Note that if $s = 1$, then $i_0(s-1) = 0$ and thus there is nothing to show. So assume $2 \leq s \leq j$, assume $i_0(s-1)$ exists, and we first consider the case $\ell = 0$. In this case we aim to prove that with exponentially high probability we have $|C_J\left(i_0(s-1)\right)| = o(\lambda n^s)$.

We first observe that if the statement does not hold, then we certainly have $|C_J\left(i_0(s-1)\right)| \geq \lambda n^s/\xi$, for any $\xi \to \infty$, but each generation up to time $i_0(s-1)$ only has size at most $n^{s-1+\delta}$. We consider two cases for possibilities of how the desired conclusion might fail, and show that each of these is very unlikely.

Case 1: There is a generation, say i, of size at least $n^{1/2-\delta/2+c}$. In this case, using $\mathcal{T}_*$ as a lower coupling for the search process, we begin $y = n^{1/2-\delta/2+c}$ independent new processes at round i. By Lemma 3.18, each of these has a probability of at least ε/n^c of reaching size at least $n^{s-1+\delta}$ within τ rounds. The probability that $i_0(s-1) > i + \tau$ is therefore at most

$$\begin{aligned}(1-\varepsilon/n^c)^y &\leq \exp(-\varepsilon y/n^c)\\ &\leq \exp(-n^{-1/3}n^{1/2-\delta/2+c}/n^c)\\ &= \exp(-n^{\delta/2}).\end{aligned}$$

We may therefore assume that Case 1 does not occur for $i \leq i_0(s-1) - \tau$.

Case 2: $i_0(s-1) \geq (\tau+1)n^{1/2-\delta/2+c}$. To show that this is unlikely we once again aim to consider y independent processes, but in order to do this we take one

individual from each of the generations $(\tau+1)\iota$, where $0 \le \iota \le n^{1/2-\delta/2+c}-1$ and consider the lower coupling $\mathcal{T}_*$ for the search process. Technically these are not independent processes, since one may be a subprocess of the other, but for this reason we consider $\mathcal{T}_*(\tau)$, the process which is cut off after τ rounds. These processes are now independent, and the same calculation as above shows that the probability that none of them become large enough after τ rounds is very small. We may therefore assume that Case 2 does not occur.

However, if neither case occurs, then we have

$$\begin{aligned}|C_J(i_0(s-1))| &\le (\tau+1)n^{1/2-\delta/2+c}n^{1/2-\delta/2+c} + \tau n^{s-1+\delta}\\ &\le 2\tau(n^{1-\delta+2c}+n^{s-1+\delta})\\ &\overset{s\ge 2}{\le} n^{s-1+2\delta}/\varepsilon\\ &= o(\lambda n^s).\end{aligned}$$

This shows that Lemma 3.10 holds for $\ell=0$. However, now we know that up to round $i_0(s-1)$ we have found few edges, which intuitively should mean that we are early in the process, and that all maximum degrees will be small. More precisely, since $|C_J(i_0(s-1))| \ge n^{s+\delta}$ by definition of $i_0(s-1)$, Lemma 3.13 implies

$$\Delta_\ell(C_J(i_0(s-1))) = O(|C_J(i_0(s-1))|n^{-\ell}) = o(\lambda n^{s-\ell})$$

for every $\ell = 1,\dots,s-1$ as required. This completes the proof of Lemma 3.10. □

Moreover, with the help of Lemmas 3.10 and 3.16 we now prove Lemma 3.11 which states that (conditional $\mathcal{E}$) the generations reach size $n^{j-1+\delta}$ – which happens in round $i_0(j-1)$ – long before round i_1 when the process is stopped.

Proof of Lemma 3.11. The key property is the following: the expansion from one generation to the next is at most a multiplicative factor of $(1+2\varepsilon_*)(1+\varepsilon) \le 1+2\varepsilon$ by Lemma 3.16. Let $x := i_1 - i_0(j-1)$, we aim to show that conditional on $\mathcal{E}$ with probability at least $1-\exp(-\Theta(n^{\delta/2}))$ we have $x \ge \frac{\delta}{2}\varepsilon^{-1}\log n$.

Suppose first that we hit the stopping condition (S3), i.e. we have $|\partial C_J(i_1)| \ge \lambda^2 n^j$. Since $|\partial C_J(i_0(j-1))| \le 2n^{j-1+\delta}$, trivially if $j=1$ and otherwise by Lemma 3.16, we then have $(1+2\varepsilon)^x \ge \lambda^2 n^{1-\delta}/2$. This yields

$$x \ge \frac{\log(\lambda^2 n^{1-\delta}/2)}{\log(1+2\varepsilon)} \ge \frac{\log(n^\delta)}{2\varepsilon} = \frac{\delta}{2}\varepsilon^{-1}\log n.$$

On the other hand suppose we hit the stopping condition (S2), i.e. we have $|C_J(i_1)| \ge \lambda n^j$. Then Lemma 3.10 (for $\ell=0$) implies that up to time $i_0(j-1)$ we have only seen $o(\lambda n^j)$ j-sets in total. Furthermore, Lemma 3.16 implies that $|\partial C_J(i_0(j-1))| \le 2n^{j-1+\delta}$, and in each subsequent round $i \in [i_0(j-1)+1, i_1]$ the size of the i-th generation increases at most by a factor of $(1+2\varepsilon)$ compared to the previous generation. Thus we have

$$\sum_{i=0}^{x}(1+2\varepsilon)^i 2n^{j-1+\delta} \ge |C_J(i_1) \setminus C_J(i_0(j-1)-1)| \ge (1-o(1))\lambda n^j,$$

which implies

$$\frac{1-(1+2\varepsilon)^{x+1}}{1-(1+2\varepsilon)} \geq (1-o(1))\lambda n^{1-\delta}/2,$$

and furthermore, we obtain $(1+2\varepsilon)^x \geq (1-o(1))\varepsilon\lambda n^{1-\delta}$. Because the right-hand side is at least n^δ, this we conclude $x\log(1+2\varepsilon) \geq \delta \log n$ and so

$$x \geq \frac{\delta}{2}\varepsilon^{-1}\log n$$

as required. Note that in all the Lemmas and Corollaries that we used here, the 'good' event holds with probability $1-\exp(-\Theta(n^{\delta/2}))$, and a union bound over all failure probabilities still leaves a probability of at least $1-\exp(-\Theta(n^{\delta/2}))$. □

Furthermore, Lemmas 3.10 and 3.16 imply that the partial component can never be too big compared to the current generation.

Claim 3.20. *With probability at least* $1-\exp(-\Theta(n^{\delta/2}))$*, for any* j*-set* J *using* BFS(J) *with stopping conditions* (S1)*,* (S2)*, and* (S3)*, the following holds: if for any* $i=0,\dots,i_1$ *we have* $|\partial C_J(i)| \geq \varepsilon\lambda n^2$*, then*

$$|C_J(i)| \leq 3\varepsilon^{-1}|\partial C_J(i)|.$$

Proof. We have $|C_J(i)| = |C_J(i_0(1))| + \sum_{i'=i_0(1)+1}^{i} |\partial C_J(i')|$ and Lemma 3.16 is applicable for all $i' = i_0(1)+1,\dots,i$. Hence we obtain

$$\begin{aligned}|C_J(i)| &\leq |C_J(i_0(1))| + |\partial C_J(i)| \sum_{s=0}^{i-i_0(1)-1} ((1-2\varepsilon_*)(1+\varepsilon))^{-s} \\ &\leq 2\varepsilon^{-1}|\partial C_J(i)| + |C_J(i_0(1))|,\end{aligned}$$

and the claim follows since by Lemma 3.10 (for $\ell = 0$ and $s = 2$) we have $|C_J(i_0(1))| = o(\lambda n^2)$. □

3.3.6. **Bipeds.** In order to control the contribution of jumps very precisely it turns out we need to investigate a certain type of structure called 'biped': given two integers $1 \leq m_1 \leq j-2$ and $m_1 \leq m_2 \leq j-1$, an m_1-set M_1, and a (distinct) m_2-set M_2,[3] it will be necessary to control how many pairs of j-sets they can be extended to which intersect in $j-m_2$ vertices (not including $M_1 \cap M_2$), and might be 'seen' while exploring a given generation.

More formally, given any (distinct) M_1 and M_2 and any round $i=0,1,\dots$, we define an (M_1,M_2,i)-*biped* to be a pair (X,Y) where X is a set of m_2-m_1 distinct vertices and Y is a set of $j-m_2$ distinct vertices such that

- $X \cap (M_1 \cup M_2) = \emptyset$;
- $Y \cap (M_1 \cup M_2 \cup X) = \emptyset$;
- $I_1 := M_1 \cup X \cup Y \in \partial C_J(i)$;
- $I_2 := M_2 \cup Y \in C_J(i)$.

[3]Distinct meaning that if $m_1 = m_2$ we have $M_2 \neq M_1$, and no restriction if $m_1 < m_2$.

We call the j-sets $I_1 = M_1 \cup X \cup Y$ and $I_2 = M_2 \cup Y$ the *first leg* and the *second leg* of (X, Y), respectively. Observe that the notion of bipeds is not symmetric in the parameters M_1 and M_2. We write $\Xi := \binom{[n]}{m_2 - m_1} \times \binom{[n]}{j - m_2}$. Then the *biped degree* of M_1 and M_2 in round i is defined as

$$d_{M_1,M_2}(i) := |\{(X, Y) \in \Xi \mid (X, Y) \text{ is an } (M_1, M_2, i)\text{-biped}\}|$$

and we aim to show that whp these biped degrees are small.

Lemma 3.21. *With probability at least* $1 - \exp(-\Theta(n^{\delta/4}))$*, for any integers* $1 \leq m_1 \leq j - 2$ *and* $m_1 \leq m_2 \leq j - 1$*, for any* m_1*-set* M_1 *and (distinct)* m_2*-set* M_2*, and any* $i = 1, 2, \ldots, i_1$*, the following assertion holds: suppose that*

$$|\partial C_J(i-1)| \geq n^{|M_1 \cup M_2| + \delta} \tag{3.11}$$

and, if $i_1(j-1) < i \leq i_1$*, additionally assume*

$$\Delta_\ell(\partial C_J(i-1)) = O(|\partial C_J(i-1)| n^{-\ell}) \quad \forall 1 \leq \ell \leq j-1, \tag{3.12}$$

then we have

$$d_{M_1,M_2}(i) = O(\lambda |\partial C_J(i)| n^{-m_1}). \tag{3.13}$$

Note that the additional assumption (3.12) is in particular satisfied if the statement of Lemma 3.8 holds at time $i-1$. So when proving Lemma 3.8 this condition will always be satisfied at the times when we apply Lemma 3.21 because we will be proving Lemma 3.8 inductively.

Proof of Lemma 3.21. Fix a pair (M_1, M_2) and a round $i \geq 2$, since $\partial C_J(0) = \{J\}$ and thus (3.11) is not satisfied. Moreover note that we may assume that $|M_1 \cup M_2| \leq j - 1$, otherwise (3.11) is violated since the total number of j-sets is only $\binom{n}{j}$. We want to provide an upper bound on the number of (M_1, M_2, i)-bipeds.

First we take care of bipeds created in a single query. For this to happen the k-set to be queried needs to contain a j-set $J' \in \partial C_J(i-1)$ as well as $M_1 \cup M_2$. We condition on the size w of their intersection $W := J' \cap (M_1 \cup M_2)$. Clearly

$$|C_J(i-1)| \geq |\partial C_J(i-1)| \geq n^{|M_1 \cup M_2| + \delta} \geq n^{w + \delta}$$

and so Lemma 3.13 is applicable with $\ell = w$ and thus the number of k-sets containing a j-set $J' \in \partial C_J(i-1) \subset C_J(i-1)$ and $M_1 \cup M_2$ is at most

$$\sum_{w = k - j - |M_1 \cup M_2|}^{|M_1 \cup M_2|} n^{k - j - |M_1 \cup M_2| + w} O(\Delta_w(C_J(i))) = O(n^{k - j - |M_1 \cup M_2|} |C_J(i)|).$$

Moreover, each query succeeds with probability p independently and if it does, then this creates at most $\binom{k}{j}$ new (M_1, M_2, i)-bipeds. Because $p n^{k - j - |M_1 \cup M_2|} |C_J(i)| \geq n^\delta$ by (3.11), the Chernoff Bound (Theorem 3.5) thus implies that with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$ their contribution is at most

$$O(n^{-|M_1 \cup M_2|} |C_J(i)|) \overset{C.\ 3.20}{=} O(\varepsilon^{-1} n^{-|M_1 \cup M_2|} |\partial C_J(i)|) = O\left(\frac{\varepsilon^{-1}}{n} |\partial C_J(i)| n^{-m_1}\right)$$

since $|M_1 \cup M_2| \geq m_1 + 1$. In other words, it is negligible compared to the claimed upper bound. So now let us call k-sets not containing $M_1 \cup M_2$ *good*, and observe that it remains to estimate the number of (M_1, M_2, i)-bipeds created when querying good k-sets.

Now we imagine that throughout the entire process BFS before querying any good k-set K we pause briefly and colour K red if it satisfies all of the following conditions:

- M_1 is contained in K;
- there exists a set $X \subset K \setminus M_1$ of size $|X| = m_2 - m_1$ and a still neutral j-set I_1 with $M_1 \cup X \subset I_1 \subset K$;
- the j-set $I_2 = (I_1 \setminus (M_1 \cup X)) \cup M_2$ is not neutral any more.

Similarly we colour K blue if all of the following conditions hold:

- M_2 is contained in K;
- there is a still neutral j-set I_2 with $M_2 \subset I_2 \subset K$;
- there exists a set $X \subset [n] \setminus (I_2 \cup M_1)$ of size $|X| = m_2 - m_1$ such that the j-set $I_1 = (I_2 \setminus M_2) \cup M_1 \cup X$ is not neutral any more.

Then no matter if K was coloured red, blue, or not at all, we perform its query. We do this to guarantee that **a)** the sequence of queries corresponding to marked k-sets satisfies the conditions of Lemma 3.6 and **b)** for any query that might result in an (M_1, M_2, i)-biped which we have not seen before, the corresponding k-set is marked.

Note that we need the blue case only to cover the possibility that I_2 is also in $\partial C_J(i)$ and happens to be discovered *after* I_1. We will significantly overcount such cases by allowing $I_1 \in C_J(i)$ rather than the more restrictive $I_1 \in \partial C_J(i)$, but this is permissible for an upper bound.

First we bound the number of k-sets marked red. In this case there must be a j-set which includes M_2 and is not neutral any more, and therefore is certainly contained in $C_J(i+1)$. By Lemma 3.13 with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$ this number is at most $O(|C_J(i)|n^{-m_2})$.

Given such a j-set I_2, this does not yet uniquely define the second leg: we choose a set X consisting of $m_2 - m_1$ vertices (disjoint from M_1, at most $n^{m_2 - m_1}$ choices), and obtain the j-set $I_1 := (I_2 \setminus M_2) \cup M_1 \cup X$ as the second leg. Combining these two steps the number of choices for the legs I_1 and I_2 is at most

$$O(|C_J(i)|n^{-m_1}). \tag{3.14}$$

Now, if I_1 is responsible for marking a k-set K red, this means that I_1 is still neutral, and therefore K also has to contain a j-set $I_3 \neq I_1$ which is currently being explored. In particular we have $I_3 \in \partial C_J(i-1)$, since we only consider queries taking place in round i. By distinguishing the size ℓ of the intersection $I_3 \cap I_1$, and noting that $\ell \leq j - 1$ since $I_3 \neq I_1$, the number of red k-sets containing

some active I_3 (and the already fixed I_1) is at most

$$\sum_{\ell=0}^{j-1} O\left(\Delta_\ell(\partial C_J(i-1))\right) n^{k-2j+\ell}. \tag{3.15}$$

We distinguish two cases: first assume that $i \leq i_1(j-1)$. Then we have

$$|\partial C_J(i)| \leq |\partial C_J(i_1(j-1))| \leq |\partial C_J(i_0(j-1))|(1+2\varepsilon)^{i^*(j-1)} \leq n^{j-1+2\delta}$$

by Lemma 3.16 and furthermore, Claim 3.20 implies

$$|C_J(i)| = O(\varepsilon^{-1}|\partial C_J(i)|).$$

Consequently we obtain

$$\Delta_\ell(\partial C_J(i-1)) \leq \Delta_\ell(C_J(i)) \overset{L.3.13}{=} O(|C_J(i)|n^{-\ell}) = O(\varepsilon^{-1}n^{j-\ell-1+2\delta}),$$

and in particular all summands in (3.15) are of the same order. Moreover, the term in (3.14) is at most

$$O(\varepsilon^{-1}|\partial C_J(i)|n^{-m_1}),$$

and thus the total number of red k-sets is upper bounded by

$$O(\varepsilon^{-2}|\partial C_J(i)|n^{k-j-m_1-1+2\delta}) = O(\lambda|\partial C_J(i)|n^{k-j-m_1}), \tag{3.16}$$

where the last step holds since $\lambda\varepsilon^2 n^{1-2\delta} \to \infty$.

On the other hand, if $i_1(j-1) < i \leq i_1$, then by the additional assumption (3.12) we have $\Delta_\ell(\partial C_J(i-1)) = \Theta(|\partial C_J(i-1)|n^{-\ell})$ for all $0 \leq \ell \leq j-1$ and thus once again all summands in (3.15) are of the same order

$$O(|\partial C_J(i-1)|n^{k-2j}) \overset{L.3.16}{=} O(|\partial C_J(i)|n^{k-2j}).$$

For the term in (3.14) we use that $|C_J(i)| \leq |C_J(i_1)| \leq 2\lambda n^j$ by Lemma 3.17. Multiplying shows that the total number of red k-sets is upper bounded by

$$O(\lambda|\partial C_J(i)|n^{k-j-m_1}). \tag{3.17}$$

Since we are only interested in an upper bound we assume that the upper bounds in (3.16) and (3.17) are asymptotically tight. Moreover note that by (3.11) we have

$$\Theta(p\lambda|\partial C_J(i)|n^{k-j-m_1}) = \Omega(n^{\delta/2}).$$

Thus the Chernoff Bound (Theorem 3.5) implies that with probability at least $1-\exp(-\Theta(n^{\delta/2}))$ the number of successful queries of red k-sets is at most

$$O\left(\lambda|\partial C_J(i)|n^{-m_1}\right). \tag{3.18}$$

The argument providing an upper bound on the number of successful queries of blue k-sets is very similar: we establish counterparts for (3.14) and (3.15), and from there on the calculations are identical. This time there has to be a j-set I_1 which includes M_1 and is not neutral any more, so by Lemma 3.13 their number is upper bounded by the term in (3.14).

Once this first leg is fixed, there is only a constant number of choices for $X \subset I_1 \setminus M_1$, and then the second leg $I_2 := (I_1 \setminus (M_1 \cup X)) \cup M_2$ is also fixed. Again I_2 must still be neutral in order to be responsible for marking a k-set K blue, hence K must contain a j-set $I_3 \neq I_2$ which is currently being queried, so in particular $I_3 \in \partial C_J(i-1)$. Distinguishing the size ℓ of the intersection $I_3 \cap I_2$, where $\ell \leq j-1$ holds since $I_3 \neq I_2$, we can upper bound the number of blue k-sets by (3.15).

Therefore with probability at least $1-\exp(-\Theta(n^{\delta/4}))$, the total number of successful queries of k-sets which were marked (either red or blue) is upper bounded by the term in (3.18). Furthermore, because a single query can create at most a constant number of bipeds, this is also an upper bound on the number of (M_1, M_2, i)-bipeds. Now the claim follows by a union bound over all choices for M_1 and M_2 and all rounds, since

$$\sum_{m_1,m_2 \leq j} n^{m_1+m_2} \cdot n^k (1-\exp(-\Theta(n^{\delta/4}))) = 1-\exp(-\Theta(n^{\delta/4})). \qquad \square$$

3.3.7. **Jumps and pivots.** In a way, the smooth boundary lemma (Lemma 3.8) is a much more detailed and precise statement of our (strengthened) bounded degree lemma (Lemma 3.13). The basic idea of the proof is similar: we investigate how the degree $d_L(\partial C_J(i+1))$ of a set L increases relative to $d_L(\partial C_J(i))$ by distinguishing the contributions made by jumps and pivots separately. Recall that these were defined in Section 3.3.2.

- A *jump* to L occurs when we query a k-set containing L from a j-set which did not contain L and the k-set forms an edge of $\mathcal{H}^k(n,p)$. Such an edge contributes at most $\binom{k-\ell}{j-\ell}$ to $d_L(\partial C_J(i+1))$.
- A *pivot* at L occurs when we query any k-set from a j-set containing L and it forms an edge. Such an edge contributes at most $\binom{k-\ell}{j-\ell}-1$ to $d_L(\partial C_J(i+1))$.

The statement of Lemma 3.8 says that with probability $1-\exp(-\Theta(n^{\delta/4}))$, a certain property must hold for *every* initial j-set J. We note that it is enough to show this for a single initial j-set – then the full generality is implied by a union bound, since $\binom{n}{j}\exp(-\Theta(n^{\delta/4})) = \exp(-\Theta(n^{\delta/4}))$. Therefore from now on we fix an initial j-set J_0.

For simplicity we denote $\partial C_{J_0}(i)$ by $\partial(i)$ for each i, and similarly, for any ℓ'-set L' we write $d_{L'}(i)$ instead of $d_{L'}(\partial(i))$, for each i.

Definition 3.22. *For all $0 \leq \ell \leq j-1$ and $0 \leq i \leq i_1$ we write $S(\ell,i)$ for the property that for all ℓ-sets L we have*

$$d_L(i) = (1 \pm \gamma_\ell)\frac{|\partial(i)|}{\binom{n}{j}}\binom{n}{j-\ell}. \tag{3.19}$$

In other words, $S(\ell,i)$ states that all ℓ-degrees within the i-th generation are essentially the same.

We aim to prove the following four claims. Firstly, the contribution $d_L^{(\mathrm{jp})}(i+1)$ to $d_L(i+1)$ made by jumps to L is approximately the same for each L.

Claim 3.23 (Smooth jumps)**.** *With probability at least* $1-\exp(-\Theta(n^{\delta/4}))$*, for all* $0 \le \ell \le j$*,* ℓ*-sets* L*, and* $i_0(\ell) \le i \le i_1$ *the following holds. Suppose that*

- $S(\ell', i)$ *holds for all* $0 \le \ell' \le \ell - 1$*,*
- *and, if* $i \ge i_1(j-1)$*, additionally* $S(\ell', i)$ *holds for all* $\ell \le \ell' \le j-1$*,*

then we have

$$d_L^{(jp)}(i+1) = (1 \pm 2\gamma_{\ell-1})(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell} - \binom{j}{\ell}\right)}{c_0}\frac{|\partial(i)|}{\binom{n}{\ell}} - O\left(\frac{d_L(i)}{n}\right).$$

Note that the additional condition for the case $i \ge i_1(j-1)$ is necessary since the proof of Claim 3.23 uses Lemma 3.21.

Secondly, the contribution $d_L^{(\mathrm{pv})}(i+1)$ to $d_L(i+1)$ made by pivots at L tends to be smaller than $d_L(i)$. We recall that $r_\ell = \frac{c_\ell}{c_0} = \frac{\binom{k-\ell}{j-\ell}-1}{\binom{k}{j}-1}$ and $r_\ell < 1$ for all $\ell \ge 1$.

Claim 3.24 (Pivots contract)**.** *With probability at least* $1-\exp(-\Theta(n^{\delta/4}))$*, for all* $0 \le \ell \le j-1$*,* ℓ*-sets* L*, and* $i_0(\ell) \le i \le i_1$ *the following holds. Suppose that*

- $S(\ell', i)$ *holds for all* $0 \le \ell' \le \ell - 1$*,*

then we have

$$d_L^{(pv)}(i+1)\begin{cases} = (1\pm\eta)(1+\varepsilon)r_\ell d_L(i) & \text{if } d_L(i) \ge n^{\delta/3}; \\ \le (1+\eta)(1+\varepsilon)r_\ell n^{\delta/3} & \text{otherwise.}\end{cases}$$

Claims 3.23 and 3.24 are essentially concentration of probability arguments, which can only hold whp once generations are large. This is the reason why we introduced starting times $i_0(\ell)$.

Thirdly, using Claims 3.23 and 3.24 we prove that ℓ-degrees will become smooth after $i^*(\ell)$ rounds.

Claim 3.25 (Smooth degrees)**.** *With probability at least* $1-\exp(-\Theta(n^{\delta/4}))$*, for all* $0 \le \ell \le j-1$ *the following holds. Suppose that*

- $S(\ell', i')$ *holds for all* $0 \le \ell' \le \ell - 1$ *and* $i_0(\ell) \le i' \le i_1(\ell)$*,*

then $S(\ell, i_1(\ell))$ *holds.*

Lastly, we prove that once ℓ-degrees are smooth, they remain smooth (at least) until the round i_1 when one of the stopping conditions (S1), (S2), or (S3) is hit.

Claim 3.26 (Smoothness inheritance)**.** *With probability at least* $1-\exp(-\Theta(n^{\delta/4}))$*, for all* $0 \le \ell \le j-1$ *and* $i_0(\ell) \le i < i_1$ *the following holds. Suppose that*

- $S(\ell', i)$ *holds for all* $0 \le \ell' \le \ell$*,*
- *and, if* $i \ge i_1(j-1)$*, additionally* $S(\ell', i)$ *holds for all* $\ell + 1 \le \ell' \le j-1$*,*

then so does $S(\ell, i+1)$*.*

Once again from round $i_1(j-1)$ onwards we need to assume some additional conditions, since we need to apply Claim 3.23 also in this regime.

Before we provide the proofs for these four claims, we use them to prove Lemma 3.8. In the remainder of this section we will always assume that all 'good' events (with exponentially small error probabilities) happen without explicitly mentioning this.

Proof of Lemma 3.8. Clearly $i_1(0) \le i_0(1)$ since $i_1(0) = 0$. We will now show that

$$i_1(\ell) \le i_0(\ell+1) \tag{3.20}$$

for each $\ell = 1, \dots, j-2$, i.e. we finish the smoothing process for the ℓ-sets before we start the smoothing process for the $(\ell+1)$-sets. (This is important because of the inductive nature of the remainder of the proof.)

Recall that by the lower bound of Lemma 3.16, from round $i_0(1)$ onwards, subsequent generations remain at least as large as the previous generation. Thus by Lemma 3.16 (applying the tight upper bound iteratively for all $i_0(\ell) < i \le i_1(\ell)$ and the cruder upper bound for $i = i_0(\ell)$) we have

$$\begin{aligned}
|\partial(i_1(\ell))| &\le |\partial(i_0(\ell))| \left((1+\varepsilon)(1+2\varepsilon_*)\right)^{i^*(\ell)} \\
&\le 2n^{\ell+\delta} \exp\left(2\varepsilon i^*(\ell)\right) \\
&\le 2n^{\ell+\delta} \exp(\Theta(\varepsilon \log n)) \\
&\le n^{\ell+2\delta} < n^{\ell+1},
\end{aligned}$$

and therefore we have not yet reached generation $i_0(\ell+1)$. This proves the second assertion of Lemma 3.8.

To prove the first assertion of Lemma 3.8 we use induction on ℓ: we first prove that $S(\ell, i)$ holds for all $0 \le \ell < j$ and $i_1(\ell) \le i \le i_1(j-1)$.

Note that $S(0, i)$ is trivially true for all i, since the empty set is contained in all edges of the i-th generation. So now assume that $\ell \ge 1$ and $S(\ell', i)$ holds for all $0 \le \ell' < \ell$ and $i_1(\ell') \le i \le i_1(j-1)$. Thus Claim 3.25 is applicable and therefore $S(\ell, i_1(\ell))$ holds. The induction step is completed by iteratively applying Claim 3.26 for $i = i_1(\ell), \dots, i_1(j-1)$.

Now the proof of Lemma 3.8 is completed by iteratively applying Claim 3.26 for rounds $i = i_1(j-1), \dots, i_1$ (and all $0 \le \ell < j$), and it remains only to prove our auxiliary results (Claims 3.23, 3.24, 3.25, and 3.26). □

Smooth jumps. Before we prove Claim 3.23, we provide a brief heuristic argument for why the main term should intuitively be the correct contribution from jumps to the degree of L. To see this, we consider the number of jumps we expect to *all* ℓ-sets in generation $i+1$. We have $|\partial(i)|$ many j-sets in generation i, from each of which we may query approximately $\binom{n}{k-j}$ many k-sets, and each forms an edge with probability p. Thus we expect to find about

$$p\binom{n}{k-j}|\partial(i)| = \frac{1+\varepsilon}{c_0}|\partial(i)|$$

edges. Furthermore, an edge results in a jump to $\binom{k}{\ell} - \binom{j}{\ell}$ many ℓ-sets (any which are contained in the edge but not in the j-set from which we made the query). Since there are $\binom{n}{\ell}$ many ℓ-sets in total, the average number of jumps to an ℓ-set 'should' be about

$$\frac{1+\varepsilon}{c_0}|\partial(i)|\frac{\binom{k}{\ell}-\binom{j}{\ell}}{\binom{n}{\ell}}.$$

Finally, for most jumps to L, the number of j-sets containing L which become active is $\binom{k-\ell}{j-\ell}$, and thus we obtain the main term in Claim 3.23.

Proof of Claim 3.23. The statement for $\ell = 0$ is trivially true, so we fix some $1 \le \ell < j$, an ℓ-set L and a round $i_0(\ell) \le i \le i_1$. (In the end we will take a union bound over all L and i.) Moreover we note that by Lemma 3.16 we have $|\partial(i)| \ge n^{\delta}$.

We consider from how many j-sets $J \in \partial(i)$ we we might jump to L. We make a case distinction on the number $0 \le m \le \ell - 1$ of vertices in the intersection of $J \cap L$. Clearly there cannot be a k-set containing $J \cup L$ if $k < j + \ell - m$, so we assume

$$k \ge j + \ell - m.$$

Now let $M \subsetneq L$ be an m-set, then there are $d_M(i) - d_L(i)$ many j-sets containing M in $\partial(i)$ from which we might jump to L. However, this may include some j-sets which actually have a larger intersection with L than M. The number of j-sets in $\partial(i)$ which have intersection exactly M with L is also at least

$$d_M(i) - \sum_{M \subsetneq \widetilde{M} \subset L} d_{\widetilde{M}}(i).$$

Note that by $S(m,i)$, the degree $d_M(i)$ is of order $|\partial(i)|n^{-m}$, and similarly for each $M \subsetneq \widetilde{M} \subsetneq L$. Furthermore $d_L(i) \le d_{\widetilde{M}}(i)$ by definition. Therefore the number of j-sets in $\partial(i)$ which intersect L in exactly M is

$$d^*_{M,L}(i) = \begin{cases} d_M(i) - d_L(i) & \text{if } m = \ell - 1 \\ (1 - O(1/n))d_M(i) & \text{otherwise.} \end{cases} \tag{3.21}$$

Moreover, note that we may assume

$$d^*_{M,L}(i) = \Omega(d_M(i)). \tag{3.22}$$

For $0 \le m \le \ell - 2$ this is immediate from $S(m,i)$. On the other hand, for $m = \ell - 1$, if the assumption (3.22) is not true, then $d^*_{M,L}(i) = o(d_M(i))$ which implies $d_L(i) = \Theta(|\partial(i)|n^{-\ell+1})$ by $S(\ell-1,i)$, and therefore the $O(d_L(i)/n)$ error term in the statement of Claim 3.23 is as large as the main term. The lower bound is therefore automatic and we only need to prove an upper bound, which will only become harder if we increase $d^*_{M,L}(i)$ artificially.

Next observe that the total number of k-sets which would lead to a jump to L from a j-set intersecting L in M is is given by

$$\binom{n-j-\ell+m}{k-j-\ell+m} d^*_{M,L}(i). \tag{3.23}$$

However some of them might not contain $c_0 = \binom{k}{j} - 1$ still neutral j-sets. We call any k-set containing L, a j-set $J \in \partial(i)$ with $M = J \cap L$, and another j-set $J' \in C_{J_0}(i)$ *tainted* and show that the total number of tainted k-sets is at most

$$O\left(\lambda n^{k-j-\ell+m} d^*_{M,L}(i)\right). \tag{3.24}$$

We distinguish cases based on the size w of the intersection $W := J' \cap (J \cup L)$. We need only consider the cases $\max\{0, 2j+\ell-m-k\} \le w \le j$ (the lower bound comes from the fact that if it is violated, there is no k-set containing $J' \cup J \cup L$). Furthermore assume for now that $w \le j-1$. Then the number of non-neutral j-sets J' with this intersection is at most $\Delta_w(C_{J_0}(i+1)) = O(\lambda n^{j-w})$ by Lemma 3.13. Moreover, the number of k-sets containing L, J and J' is of order $n^{k-j-\ell+m-(j-w)}$ (and this quantity is $\Omega(1)$ in our range of w). Thus the number of non-permissible k-sets is at most

$$O(\lambda n^{k-j-\ell+m} d^*_{M,L}(i)).$$

Note that this is independent of w. Furthermore, the number of possible choices for W is at most $\sum_{w=\max\{0,2j+\ell-m-k\}}^{j-1} \binom{j+\ell-m}{w} = O(1)$. This shows that the total number of tainted k-sets of this type can be bounded as in (3.24).

It remains to consider the case $w = j$, or in other words $J' \subset J \cup L$. Then writing $M' := J' \cap L$ and $m' := |M'|$ we observe that $(J \setminus (J' \cup L), (J \cap J') \setminus L)$ is an (M, M', i)-biped, where $1 \le m \le \ell - 1 \le j - 2$ and $m \le m' \le \ell \le j - 1$. Furthermore, we have $|\partial(i)| \ge n^{\ell+\delta} \ge n^{|M \cup M'|+\delta}$, and if $i \ge i_1(j-1)$, then (3.12) is satisfied since $S(0,i), \dots, S(j-1,i)$ hold by the assumption of Claim 3.23. Thus Lemma 3.21 is applicable for all i and implies that the number of k-sets containing such a configuration of L, J, and J' is at most

$$d_{M,M'}(i) n^{k-j-\ell+m} = O(\lambda n^{k-j-\ell} |\partial(i)|).$$

By $S(m,i)$ we have $d_M(i) = \Theta(|\partial(i)| n^{-m})$, and hence also in the case when $w = j$, the term in (3.24) is an upper bound on the number of tainted queries, by the assumption in (3.22), as required.

Now let N denote the total number of queries we make from j-sets intersecting L in precisely M to k-sets containing L and exactly c_0 neutral j-sets. Together (3.23) and (3.24) imply that

$$N = (1 - O(\lambda)) \binom{n}{k-j-\ell+m} d^*_{M,L}(i). \tag{3.25}$$

Furthermore, all of these queries are independent and succeed with probability p, thus the number of edges resulting in jumps from M to L (and which cause its degree to increase by $\binom{k-\ell}{j-\ell}$) has distribution $\mathrm{Bin}(N,p)$. Now recall that by our assumptions on m we have $\binom{n}{k-j-\ell+m} = \Omega(1)$. Note that the main term in (3.25) is also an upper bound on the total number (permissible and non-permissible) of queries which would result in jumps from M to L. Since we will only need the main term anyway, we therefore slightly abuse notation by using N for both the lower and upper bound.

Moreover, note that by (3.22) and $S(m,i)$, we have

$$N = \Omega(n^{k-j-\ell+m} d_M(i)) = \Omega(n^{k-j-\ell}|\partial(i)|) = \Omega(n^{k-j+\delta}).$$

Thus, we may apply the Chernoff bound (Theorem 3.5) to show that the probability that the number of such edges is not in $(1\pm\eta)pN$ is at most

$$2\exp(-\eta^2 Np/3) \le \exp(-\Theta(\eta^2 n^\delta)) \le \exp(-\Theta(n^{\delta/2})).$$

If this unlikely event does not occur, then the contribution to $d_L(i)$ made by jumps from M to L is

$$\binom{k-\ell}{j-\ell}(1\pm\eta)N\frac{1+\varepsilon}{c_0\binom{n}{k-j}} \overset{(3.25)}{=} (1\pm 2\eta)(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}}{c_0}\frac{(k-j)!}{(k-j-\ell+m)!}\frac{d^*_{M,L}(i)}{n^{\ell-m}},$$

using $\lambda \ll \eta$. We now sum over all such sets $M \subsetneq L$ and use the fact that

$$\frac{d^*_{M,L}(i)}{n^{\ell-m}} = (1-O(1/n))\frac{d_M(i)}{n^{\ell-m}} - O\left(\frac{d_L(i)}{n}\right)$$

in all cases. Moreover, by $S(m,i)$ we have

$$d_M(i) = (1\pm\gamma_m)\frac{|\partial(i)|}{\binom{n}{j}}\binom{n}{j-m},$$

and consequently the contribution to $d_L(i+1)$ made by jumps to L is

$$\begin{aligned}
&\sum_{m=0}^{\ell-1}\binom{\ell}{m}(1\pm 2\eta)(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}}{c_0}\frac{(k-j)!}{(k-j-\ell+m)!}\frac{d^*_{M,L}(i)}{n^{\ell-m}}\\
&\quad= (1\pm 3\eta)(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}}{c_0}\sum_{m=0}^{\ell-1}\binom{\ell}{m}\frac{(k-j)!}{(k-j-\ell+m)!}\frac{d_M(i)}{n^{\ell-m}} - O\left(\frac{d_L(i)}{n}\right)\\
&\quad= (1\pm\frac{3}{2}\gamma_{\ell-1})(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}}{c_0}\sum_{m=0}^{\ell-1}\frac{\binom{\ell}{m}(k-j)!j!}{(k-j-\ell+m)!(j-m)!}\frac{|\partial(i)|}{n^\ell} - O\left(\frac{d_L(i)}{n}\right)\\
&\quad= (1\pm 2\gamma_{\ell-1})(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}}{c_0}\frac{|\partial(i)|}{\binom{n}{\ell}}f - O\left(\frac{d_L(i)}{n}\right),
\end{aligned}$$

where

$$\begin{aligned}
f = f(j,k,\ell) :=& \frac{(k-j)!j!}{\ell!}\sum_{m=0}^{\ell-1}\frac{\binom{\ell}{m}}{(j-m)!(k-j-\ell+m)!}\\
=& \frac{(k-j)!j!}{\ell!}\sum_{m=0}^{\ell-1}\binom{\ell}{m}\frac{\binom{k-\ell}{j-m}}{(k-\ell)!}\\
=& \frac{(k-j)!j!}{(k-\ell)!\ell!}\left(\sum_{m=0}^{\ell}\binom{\ell}{m}\binom{k-\ell}{j-m} - \binom{k-\ell}{j-\ell}\right)\\
=& \frac{(k-j)!j!}{(k-\ell)!\ell!}\left(\binom{k}{j} - \binom{k-\ell}{j-\ell}\right).
\end{aligned}$$

It remains to observe that the right-hand side simplifies to

$$\frac{k!}{\ell!(k-\ell)!} - \frac{j!}{\ell!(j-\ell)!} = \binom{k}{\ell} - \binom{j}{\ell}. \qquad \square$$

Note that the effect that $d_L(i)$ has on the number of jumps to L in generation $i+1$ is very small (the $O(d_L(i)/n)$ term).

Pivots contract. Claim 3.24 shows that the effect that $d_L(i)$ has on the number of pivots at L in generation $i+1$, while significantly larger, is still likely to be smaller than $d_L(i)$.

Proof of Claim 3.24. The statement for $\ell = 0$ is trivially true, so we fix some $1 \le \ell < j$, an ℓ-set L and a round $i_0(\ell) \le i \le i_1$. (In the end we will take a union bound over all L and i.) Moreover we note that since $i \ge i_0(\ell)$ we have $|\partial(i)| \ge n^\delta$ by Lemma 3.16.

Recall that $d_L^{(\mathrm{pv})}(i+1)$ denote the contribution to $d_L(i+1)$ made by pivots at L. We first prove the upper bound. From every j-set containing L, we make at most $\binom{n}{k-j}$ queries, and each time we discover an edge in this way, it contributes at most $c_\ell = \binom{k-\ell}{j-\ell} - 1$ to $d_L(i+1)$. Thus $d_L^{(\mathrm{pv})}(i+1)$ is stochastically dominated by a random variable Z^* with distribution $c_\ell \cdot \mathrm{Bin}\big(\binom{n}{k-j} d_L(i), p\big)$, which has expectation

$$\begin{aligned}\mathbb{E}(Z^*) = c_\ell \binom{n}{k-j} d_L(i) p &= (1+\varepsilon)\frac{c_\ell}{c_0} d_L(i) \\ &= (1+\varepsilon) r_\ell d_L(i).\end{aligned}$$

Thus when $d_L(i) \ge n^{\delta/3}$, by the Chernoff bound (Theorem 3.5) applied to Z^* we have

$$\begin{aligned}\mathbb{P}\left(d_L^{(\mathrm{pv})}(i+1) \ge (1+\eta)(1+\varepsilon) r_\ell d_L(i)\right) &\le \mathbb{P}(Z^* \ge (1+\eta)(1+\varepsilon) r_\ell d_L(i)) \\ &\le \exp\left(-\eta^2 (1+\varepsilon) r_\ell d_L(i)/3\right) \\ &\le \exp(-\Theta(\eta^2 d_L(i))) \\ &\le \exp(-\Theta(n^{\delta/4}))\end{aligned}$$

since $\eta \ge n^{-\delta/24}$. This proves the upper bound in the case that $d_L(i) \ge n^{\delta/3}$. In the second case, we simply take $c_\ell \cdot \mathrm{Bin}\big(\binom{n}{k-j} n^{\delta/3}, p\big)$ as a dominating variable and a similar calculation holds. This proves the upper bound in both cases.

For the lower bound, we observe that since $\mathcal{T}_* \prec \mathrm{BFS}$, by Lemma 3.17, the number of queries that would result in exactly c_ℓ pivots that we make from each j-set is at least $(1-\varepsilon_*)\binom{n}{k-j}$. Thus $d_L^{(\mathrm{pv})}(i+1)$ dominates a random variable Z with distribution $c_\ell \cdot \mathrm{Bin}\big((1-\varepsilon_*)\binom{n}{k-j} d_L(i), p\big)$. A similar calculation shows that

$$\begin{aligned}\mathbb{P}\left(d_L^{(\mathrm{pv})}(i+1) \le (1-\eta)(1+\varepsilon) r_\ell d_L(i)\right) &\le \mathbb{P}(Z \le (1-\eta)(1+\varepsilon) r_\ell d_L(i)) \\ &\overset{\text{Thm 3.5}}{\le} \exp\left(-(\eta-\varepsilon_*)^2 (1+\varepsilon)(1-\varepsilon_*) r_\ell d_L(i)/2\right) \\ &\le \exp(-\Theta(\eta^2 d_L(i))) \\ &\le \exp(-\Theta(n^{\delta/4})).\end{aligned}$$

□

Smooth degrees. Next we prove Claim 3.25 by showing that in each further round the degree of any ℓ-set L is more and more concentrated around its mean based on Claims 3.23 and 3.24.

Proof of Claim 3.25. The statement for $\ell = 0$ is trivially true, so we fix some $1 \le \ell < j$ and an ℓ-set L. Note that Claims 3.23 and 3.24 are applicable for all $i_0(\ell) \le i \le i_1(\ell)$ by the assumptions of Claim 3.25.

For $s \in [0, i^*(\ell)]$, we set $d_s := d_L(i_0(\ell)+s)$ and set $d'_s := \max\{d_s, n^{\delta/3}\}$. Then by Claims 3.23 and 3.24 we have

$$\begin{aligned} d_s &\le (1+\eta)(1+\varepsilon) r_\ell d'_{s-1} + (1+2\gamma_{\ell-1})(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell}-\binom{j}{\ell}\right)}{c_0}\frac{|\partial(i_0(\ell)+s-1)|}{\binom{n}{\ell}} \\ &= (1+\eta)(1+\varepsilon) r_\ell d'_{s-1} + (1+2\gamma_{\ell-1}) g\,|\partial(i_0(\ell)+s-1)| \end{aligned}$$

where

$$g = g(k,j,\ell,n,\varepsilon) := (1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell}-\binom{j}{\ell}\right)}{c_0}\binom{n}{\ell}^{-1}.$$

Note that g is not dependent on s.

If $d'_{s-1} = d_{s-1}$, we apply the same inequality with an index shift, and keep iterating until we have some $d_{s^*} \le n^{\delta/3}$ or else we reach $s^* = 0$. In each iteration, say $s' = 1, \dots, s - s^*$, we obtain one additional summand depending on $|\partial(i_0(\ell)+s-s')|$, and there is exactly one summand depending only on $d'_{s-s'}$. For $s' = s - s^*$, this latter term can be bounded by $((1+\eta)(1+\varepsilon)r_\ell)^s d_0 + n^{\delta/3}$, since $(1+\eta)(1+\varepsilon)r_\ell < 1$. Hence we obtain

$$\begin{aligned} d_s &\le ((1+\eta)(1+\varepsilon)r_\ell)^s d_0 + n^{\delta/3} \\ &\qquad + (1+2\gamma_{\ell-1}) g \sum_{s'=1}^{s-s^*} ((1+\eta)(1+\varepsilon)r_\ell)^{s'} |\partial(i_0(\ell)+s-s')| \\ &\le (1+\eta)^s \left((1+\varepsilon)^s r_\ell{}^s d_0 + (1+2\gamma_{\ell-1}) g \sum_{s'=1}^{s} (1+\varepsilon)^{s'} r_\ell{}^{s'} |\partial(i_0(\ell)+s-s')| \right) + n^{\delta/3}. \end{aligned}$$

To calculate the corresponding lower bound we cannot use $(1-\eta)(1+\varepsilon)r_\ell d_{s-1}$ from Claim 3.24, since it may be that $d_{s-1} < n^{\delta/3}$, in which case that result does not give us any lower bound. Instead we use the lower bound (in all cases) of $(1-\eta)(1+\varepsilon)r_\ell d_{s-1} - n^{\delta/3}$, which may be negative but which will turn out to be good enough. Thus we have

$$\begin{aligned} d_s &\ge (1-\eta)(1+\varepsilon) r_\ell d_{s-1} - n^{\delta/3} \\ &\qquad + (1-2\gamma_{\ell-1})\,(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell}-\binom{j}{\ell}\right)}{c_0}\frac{|\partial(i_0(\ell)+s-1)|}{\binom{n}{\ell}} - O\left(\frac{d_{s-1}}{n}\right) \\ &\ge (1-2\eta)(1+\varepsilon) r_\ell d_{s-1} + (1-3\gamma_{\ell-1}) g\,|\partial(i_0(\ell)+s-1)|. \end{aligned}$$

Note that we absorbed the $-n^{\delta/3}$ term by modifying the $\gamma_{\ell-1}$ term. This is permissible since for any $i \in [i_0(\ell), i_1]$ we have $|\partial(i)| \ge n^{\ell+\delta}$, and so $\gamma_{\ell-1}|\partial(i)|/n^\ell \gg n^{\delta/3}$.

This time there is no need to stop prior to $s' = s$, so iterating gives

$$\begin{aligned} d_s &\geq ((1-2\eta)(1+\varepsilon)r_\ell)^s d_0 \\ &\qquad + (1-3\gamma_{\ell-1})g\sum_{s'=1}^{s}((1-2\eta)(1+\varepsilon)r_\ell)^{s'}|\partial(i_0(\ell)+s-s')| \\ &\geq (1-2\eta)^s\left((1+\varepsilon)^s {r_\ell}^s d_0 + (1-3\gamma_{\ell-1})g\sum_{s'=1}^{s}(1+\varepsilon)^{s'}{r_\ell}^{s'}|\partial(i_0(\ell)+s-s')|\right). \end{aligned}$$

Observe that only the first term of both the upper and lower bound on d_s depend on L (via d_0). In particular for $s = i^*(\ell)$, we have

$$\begin{aligned} 0 \leq ((1-2\eta)(1+\varepsilon)r_\ell)^{i^*(\ell)} d_0 &\leq ((1+\eta)(1+\varepsilon)r_\ell)^{i^*(\ell)} d_0 \\ &\leq \left((1+2\eta+2\varepsilon)\, r_\ell\right)^{i^*(\ell)} n^{j-\ell} \leq 1 \end{aligned}$$

since we chose $i^*(\ell) = \left\lceil \frac{(j-\ell)\log n}{-\log((1+2\eta+2\varepsilon)r_\ell)} \right\rceil$. In other words, d_0, the degree of L at time $i_0(\ell)$ only has an influence of at most one by time $i_1(\ell) = i_0(\ell) + i^*(\ell)$, which will not affect calculations significantly, so we ignore it.

The remaining upper and lower bounds do not depend on L. Furthermore, observing that $g = \Theta(n^{-\ell})$ and $|\partial(i)| \geq \Theta(n^{\ell+\delta})$, we have $n^{\delta/3} = O(n^{-2\delta/3} g|\partial(i)|)$, and so the remaining upper and lower bounds differ by a multiplicative factor of

$$\frac{(1+O(n^{-2\delta/3}))(1+\eta)^s(1+2\gamma_{\ell-1})}{(1-2\eta)^s(1-3\gamma_{\ell-1})} \leq (1+5\eta)^s(1+6\gamma_{\ell-1}) \leq (1+7\gamma_{\ell-1}).$$

By definition we have

$$1 + 7\gamma_{\ell-1} \leq 1 + \gamma_\ell.$$

Taking a union bound over all sets L of size ℓ, we may say that with probability at least $1 - \exp(-\Theta(n^{\delta/4}))$, all ℓ-sets have asymptotically the same degree in $\partial(i)$. More precisely this is stated as follows. We have

$$\sum_L d_L(i) = \binom{j}{\ell}|\partial(i)|$$

which implies that

$$\frac{1}{\binom{n}{\ell}}\sum_L d_L(i) = \frac{\binom{n-\ell}{j-\ell}}{\binom{n}{j}}|\partial(i)|.$$

We further have, for any particular ℓ-set L_0, that by the arguments above

$$\frac{1}{1+\gamma_\ell}\frac{1}{\binom{n}{\ell}}\sum_L d_L(i) \leq d_{L_0}(i) \leq (1+\gamma_\ell)\frac{1}{\binom{n}{\ell}}\sum_L d_L(i)$$

and since $1/(1+\gamma_\ell) \geq 1-\gamma_\ell$, the conclusion of Claim 3.25 follows. □

Smoothness inheritance. It remains to prove Claim 3.26 which, roughly speaking, states that once generations are smooth, they remain smooth.

Proof of Claim 3.26. The statement for $\ell = 0$ is trivially true, so we fix some $1 \leq \ell < j$, $i_0(\ell) \leq i < i_1$, and an ℓ-set L. Note that Claims 3.23 and 3.24 are applicable for i by the assumptions of Claim 3.26.

From $S(\ell,i)$ we deduce that the error-term $O(d_L(i)n^{-1}) = O(|\partial(i)|n^{-\ell-1})$ in Claim 3.23 is negligible compared to the main order term, which has order $\Theta(|\partial(i)|n^{-\ell})$. Similarly, property $S(\ell,i)$ implies that $d_L(i) = \Theta(|\partial(i)|n^{-\ell}) = \Omega(n^{\delta})$ by Lemma 3.16 since $i \geq i_0(\ell)$, and thus Claim 3.24 provides both an upper and a lower bound on the contribution of pivots at L to $d_L(i+1)$. Therefore, by Claims 3.23 and 3.24 we have

$$d_L(i+1) = (1\pm 3\gamma_{\ell-1})(1+\varepsilon)\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell}-\binom{j}{\ell}\right)}{c_0}\frac{|\partial(i)|}{\binom{n}{\ell}} + (1\pm 2\eta)(1+\varepsilon)r_\ell d_L(i).$$

Using $S(\ell,i)$, this becomes

$$d_L(i+1) = (1+\varepsilon)\frac{|\partial(i)|}{\binom{n}{j}}\binom{n}{j-\ell}\left((1\pm 4\gamma_{\ell-1})\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell}-\binom{j}{\ell}\right)}{\binom{j}{\ell}\left(\binom{k}{j}-1\right)} + (1\pm 2\eta)r_\ell\right).$$

Moreover, we have $|\partial(i+1)| = (1\pm 2\varepsilon_*)(1+\varepsilon)|\partial(i)|$ by Lemma 3.16, and recall that $\varepsilon_* \ll \eta \ll \gamma_0$ and $\gamma_\ell = 8\gamma_{\ell-1} \geq \gamma_0$. Lastly observe that

$$\frac{\binom{k-\ell}{j-\ell}\left(\binom{k}{\ell}-\binom{j}{\ell}\right)}{\binom{j}{\ell}} = \binom{k}{j} - \binom{k-\ell}{j-\ell}$$

and

$$\left(\binom{k}{j}-1\right)r_\ell = \binom{k-\ell}{j-\ell} - 1.$$

Thus we obtain

$$d_L(i+1) = (1\pm\gamma_\ell)\frac{|\partial(i+1)|}{\binom{n}{j}}\binom{n}{j-\ell},$$

as claimed. □

3.4. Supercritical regime

We now use Lemma 3.8 to prove Theorem 3.2(a), which is substantially harder than the proof of Theorem 3.2(b) even having proved Lemma 3.8 already.

As mentioned earlier we will consider the random variable X counting the number of j-sets in 'large' components. In a first step we compute its mean, and then in the second step we establish concentration around its mean by a second moment argument. Once the intermediate goal of X being concentrated around its mean is achieved, we complete the proof using a sprinkling argument to show that almost all vertices in 'large' components in fact lie in a single giant component.

But first we briefly highlight two important properties of the supercritical branching processes, which we use to couple exploration processes using BFS.

3.4.1. **Properties of branching processes.** We will also need some results on branching processes. The arguments here are similar to standard and well-known arguments, but the actual processes may have an unfamiliar distribution, so for completeness we give the full arguments in Appendix 3.A.

Survival probability. Consider the branching process $\mathcal{T}^*$, in which the number of children has distribution $c_0 \cdot \mathrm{Bin}\big(\binom{n}{k-j}, p\big)$. For the supercritical case we have

$p = (1+\varepsilon)\hat{p}_{\mathrm{g}}$. By setting up a suitable recursion (see Appendix 3.A), we may deduce that the survival probability ϱ of this process is the unique positive solution to the equation

$$(1-\varrho)^{c_0} = \sum_{i=0}^{c_0\binom{n}{k-j}} \mathbb{P}\left(\mathrm{Bin}\left(c_0\binom{n}{k-j}, p\right) = i\right)(1-\varrho)^{c_0 i}$$

which is asymptotically

$$\varrho \sim 2\varepsilon/c_0. \tag{3.26}$$

Likewise the lower coupling process $\mathcal{T}_*$, in which the number of children has distribution $c_0 \cdot \mathrm{Bin}\big((1-\varepsilon_*)\binom{n}{k-j}, p\big)$, has survival probability ϱ_* with

$$\varrho_* \sim 2\varepsilon/c_0. \tag{3.27}$$

Dual process. When we explore a component and the lower coupling $\mathcal{T}_*$ survives indefinitely we can conclude that the component must be large, since the search process certainly survives until $\mathcal{T}_*$ is no longer a lower coupling. However if the upper coupling $\mathcal{T}^*$ dies out we need some information on the total size of all its generations in order to decide whether this implies that the component must be small.

Recall that $\mathcal{T}^*$ has the offspring distribution

$$c_0 \cdot \mathrm{Bin}\left(\binom{n}{k-j}, p\right)$$

where $p = (1+\varepsilon)p_o$. We denote the event that $\mathcal{T}^*$ dies out by $\mathcal{D}$ and condition on it. This defines a conditional branching process $\mathcal{T}_{\mathcal{D}}$, called the *dual process*, with offspring distribution

$$c_0 \cdot \mathrm{Bin}\left(\binom{n}{k-j}, p_{\mathcal{D}}\right),$$

where

$$p_{\mathcal{D}} = (1-\varepsilon+o(\varepsilon))\hat{p}_{\mathrm{g}},$$

(see Appendix 3.A for a proof). Thus the expected number of children of any individual in $\mathcal{T}_{\mathcal{D}}$ is given by

$$c_0\binom{n}{k-j}p_{\mathcal{D}} = 1-\varepsilon \pm o(\varepsilon) < 1$$

and hence in particular the dual process is a subcritical branching process.

Therefore we can give an asymptotic estimate on the expected total size of $\mathcal{T}_{\mathcal{D}}$ by standard techniques (see [68]):

$$\mathbb{E}\left(|\mathcal{T}_{\mathcal{D}}|\right) = \sum_{i=0}^{\infty}\left(c_0\binom{n}{k-j}p_{\mathcal{D}}\right)^i = \frac{1}{1-c_0\binom{n}{k-j}p_{\mathcal{D}}} \sim \varepsilon^{-1}. \tag{3.28}$$

Consequently we can bound the probability of the process $\mathcal{T}^*$ being larger than some $\Lambda = \Lambda(n)$ by conditioning on $\mathcal{D}$ and applying Markov's inequality

$$\begin{aligned}\mathbb{P}\left(|\mathcal{T}^*| \geq \Lambda\right) &= \mathbb{P}\left(\neg\mathcal{D}\right) \cdot 1 + \mathbb{P}(\mathcal{D}) \cdot \mathbb{P}\left(|\mathcal{T}^*| \geq \Lambda \mid \mathcal{D}\right) \\ &\leq \mathbb{P}\left(\neg\mathcal{D}\right) + 1 \cdot \mathbb{P}\left(|\mathcal{T}_{\mathcal{D}}| \geq \Lambda\right) \\ &\overset{(3.28)}{\leq} \varrho + (1 \pm o(1))\varepsilon^{-1}\Lambda^{-1} \\ &\overset{(3.26)}{\sim} \frac{2\varepsilon}{c_0}, \end{aligned} \tag{3.29}$$

as long as $\varepsilon^2\Lambda \to \infty$.

3.4.2. **Total size of large components.** Let X denote the number of j-sets in components of size at least

$$\Lambda := \lambda n^j$$

and observe that $\varepsilon^2\Lambda \to \infty$ is satisfied (see Section 3.1.4.1). While calculating the expectation $\mathbb{E}(X)$ is easy, proving that X is concentrated around $\mathbb{E}(X)$ is the main challenge of the proof of Theorem 3.2(a). We will show the following claim

Claim 3.27. *Whp we have*

$$X = (1 \pm o(1))\frac{2\varepsilon}{\binom{k}{j} - 1}\binom{n}{j}. \tag{3.30}$$

First moment. We first determine $\mathbb{E}(X)$ by (partially) growing the component from an arbitrary j-set J using BFS and verifying whether it has size at least Λ. By Lemma 3.12 we only need $O(\lambda n^k)$ queries to do so. Therefore the coupling $\mathcal{T}_* \prec \text{BFS} \prec \mathcal{T}^*$ holds by Lemma 3.15 (until time $t = O(\lambda n^k)$) and we obtain

$$\mathbb{E}\left(X\right) \leq \binom{n}{j}\mathbb{P}\left(|\mathcal{T}^*| \geq \Lambda\right) \overset{(3.29)}{\sim} \frac{2\varepsilon}{c_0}\binom{n}{j}$$

and similarly

$$\mathbb{E}\left(X\right) \geq \binom{n}{j}\mathbb{P}\left(|\mathcal{T}_*| = \infty\right) = \varrho_*\binom{n}{j} \overset{(3.27)}{\sim} \frac{2\varepsilon}{c_0}\binom{n}{j}.$$

Hence we have

$$\mathbb{E}(X) = (1 \pm o(1))\frac{2\varepsilon}{c_0}\binom{n}{j}. \tag{3.31}$$

Second moment. Let $\mathcal{L}$ denote the union of all *large components*, i.e. components of size at least Λ. In order to apply Chebyshev's inequality to prove that X is concentrated around its expectation, we need to show that $\mathbb{E}(X^2) \sim \mathbb{E}(X)^2$. We may interpret X^2 as the number of ordered pairs of j-sets in large components (formally we may pick the same j-set twice in such a pair) and thus we can write its expectation as

$$\mathbb{E}\left(X^2\right) = \sum_{J_1 \in \binom{V}{j},\, J_2 \in C_{J_1}} \mathbb{P}\left(J_1, J_2 \in \mathcal{L}\right) + \sum_{J_1 \in \binom{V}{j},\, J_2 \in \binom{n}{j} \setminus C_{J_1}} \mathbb{P}\left(J_1, J_2 \in \mathcal{L}\right). \tag{3.32}$$

Fix an arbitrary j-set J_1. We start growing its component using BFS, and recall that $C_{J_1}(i)$ denotes its partial component at the beginning of round i and

$\partial C_{J_1}(i)$ the i-th generation. We denote the upper coupling branching process for the exploration by $\mathcal{T}^{J_1}$ (i.e. $\mathcal{T}^{J_1}$ is a particular instance of $\mathcal{T}^*$). We continue to grow the component until at the beginning of some round $i \in \mathbb{N}$ one of the following three stopping conditions is reached[4]:

(S1) the component of J_1 is fully explored (i.e. no j-sets are still active);
(S2) the (partial) component $C_{J_1}(i)$ has reached size at least $\Lambda = \lambda n^j$;
(S3) the i-th generation $\partial C_{J_1}(i)$ has reached size at least $\lambda\Lambda = \lambda^2 n^j$.

Moreover, the (first) round in which any these stopping conditions is invoked was denoted by

$$i_1 = \min_{i\in\mathbb{N}}\{\text{(S1)} \vee \text{(S2)} \vee \text{(S3) holds in round } i\}.$$

With slight abuse of notation we write C_{J_1} for $C_{J_1}(i_1)$. Similarly, we write ∂C_{J_1} instead of $\partial C_{J_1}(i_1)$ and call ∂C_{J_1} the *boundary* of C_{J_1}.

Note that the choice of λ (see Section 3.1.4.1) was tailored for this application: on the one hand, we want to stop as early as possible so that the coupling still holds; on the other, we want to be sure, that we can use the information which stopping condition was invoked to reliably distinguish between large components and those which are not large. While we already proved that the coupling is valid for the entire stopped exploration process (Lemma 3.17), we will now demonstrate that this setup also allows us to detect large components accurately.

We start by comparing the reasons why the search processes stop with the events that the corresponding components are large. First, observe that if the exploration C_{J_1} stopped because the component was fully explored, (S1), then it never reached size Λ and therefore J_1 does not lie in a large component. Consequently it will not contribute to X^2. Hence we are interested in the case when the exploration of the component of J_1 stops due to stopping condition (S2) or (S3) and we recall that this event was denoted by

$$\mathcal{E} = \{\text{(S2)} \vee \text{(S3) holds in round } i_1\}.$$

From the previous observation it is immediate that $\{J_1 \in \mathcal{L}\} \implies \mathcal{E}$ and thus

$$\mathbb{P}(J_1,\, J_2 \in \mathcal{L}) \leq \mathbb{P}(\mathcal{E} \wedge \{J_2 \in \mathcal{L}\}) = \mathbb{P}(\mathcal{E})\,\mathbb{P}(J_2 \in \mathcal{L} \mid \mathcal{E})\ . \tag{3.33}$$

In order to give a suitable upper bound for $\mathbb{P}(\mathcal{E})$ we have to analyse the upper coupling $\mathcal{T}^{J_1}$ if we stop due to the second or third stopping condition, (S2) or (S3) respectively. For technical reasons we distinguish two cases in a way that might seem a little awkward: if stopping condition (S2) was implemented, then clearly J_1 does lie in a large component and therefore $\mathcal{T}^{J_1}$ will contain at least Λ many j-sets since we already reached that size when we stopped the exploration. This motivates a case distinction according to the following implication

$$\mathcal{E} \implies \left\{\left|\mathcal{T}^{J_1}\right| \geq \Lambda\right\} \vee \left(\mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_1}\right| < \Lambda\right\}\right)\ . \tag{3.34}$$

[4]Recall that we already saw these stopping conditions defined for general exploration processes (using BFS) in Section 3.3

Let $\mathcal{W}$ be the event that the generation of $\mathcal{T}^{J_1}$ at which we stopped the exploration process is larger than $\lambda\Lambda$. Observe that

$$\mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_1}\right| < \Lambda\right\} \implies \mathcal{W} \wedge \left\{\left|\mathcal{T}^{J_1}\right| < \Lambda\right\}, \tag{3.35}$$

since the event $\mathcal{E}$ can only hold (subject to $\left\{\left|\mathcal{T}^{J_1}\right| < \Lambda\right\}$) if the boundary ∂C_{J_1} of the explored component was at least of size $\lambda\Lambda$ and hence the corresponding generation of the upper coupling must also have been large, i.e. $\mathcal{W}$ needs to hold. Because

$$\left\{\left|\mathcal{T}^{J_1}\right| < \Lambda\right\} \implies \left\{\left|\mathcal{T}^{J_1}\right| < \infty\right\}, \tag{3.36}$$

we want to consider the event that $\mathcal{T}^{J_1}$ dies out after having had a large generation. Intuitively we would imagine that the chances of this happening are very small so let us make this intuition more precise.

Assume that $\mathcal{W}$ holds. Then there is a generation of $\mathcal{T}^{J_1}$ with at least $\lambda\Lambda$ j-sets in the boundary, and from each of these we start an independent copy of $\mathcal{T}^*$, all of which need to die out in order for $\mathcal{T}^{J_1}$ to die out. Since the survival probability of $\mathcal{T}^*$ is $(1 \pm o(1))\frac{2\varepsilon}{\binom{k}{j}-1} \geq \varepsilon/2^k$ we thus have

$$\mathbb{P}\left(\left|\mathcal{T}^{J_1}\right| < \infty \mid \mathcal{W}\right) \leq \mathbb{P}\left(|\mathcal{T}^*| < \infty\right)^{\lambda\Lambda} \leq \left(1 - \frac{\varepsilon}{2^k}\right)^{\lambda\Lambda} \leq \exp\left(-\frac{\varepsilon\lambda\Lambda}{2^k}\right), \tag{3.37}$$

and the right-hand side is $o(1)$. This implies

$$\begin{aligned}
\mathbb{P}\left(\mathcal{W} \wedge \left\{\left|\mathcal{T}^{J_1}\right| < \infty\right\}\right) &= \mathbb{P}(\mathcal{W})\mathbb{P}\left(\left|\mathcal{T}^{J_1}\right| < \infty \mid \mathcal{W}\right) \\
&= \mathbb{P}\left(\mathcal{W} \wedge \left\{\left|\mathcal{T}^{J_1}\right| = \infty\right\}\right) \cdot \frac{\mathbb{P}\left(\left|\mathcal{T}^{J_1}\right| < \infty \mid \mathcal{W}\right)}{\mathbb{P}\left(|\mathcal{T}^{J_1}| = \infty \mid \mathcal{W}\right)} \\
&\overset{(3.37)}{=} \mathbb{P}\left(\left|\mathcal{T}^{J_1}\right| = \infty\right) \cdot o(1) \\
&\overset{(3.26)}{=} o(\varepsilon).
\end{aligned} \tag{3.38}$$

Putting things together, from (3.34), (3.35), and (3.36) we obtain

$$\mathbb{P}\left(\mathcal{E}\right) \leq \mathbb{P}\left(\left|\mathcal{T}^{J_1}\right| \geq \Lambda\right) + \mathbb{P}\left(\left\{\left|\mathcal{T}^{J_1}\right| < \infty\right\} \wedge \mathcal{W}\right) \leq \frac{2\varepsilon}{c_0} + o\left(\varepsilon\right), \tag{3.39}$$

where the second inequality holds by (3.29) and (3.38).

Now let us first consider the contribution to $\mathbb{E}(X^2)$ that comes from j-sets J_2 which lie inside C_{J_1}. By Lemma 3.17, there are at most 2Λ such j-sets, and therefore the contribution is at most

$$\sum_{J_1 \in \binom{V}{j},\, J_2 \in C_{J_1}} \Pr(J_1, J_2 \in \mathcal{L}) \overset{(3.29)}{\leq} \binom{n}{j} \cdot 2\Lambda \cdot O(\varepsilon) = O(\varepsilon\lambda n^{2j}) = o(\varepsilon^2 n^{2j}).$$

Since $\mathbb{E}(X^2) \geq \mathbb{E}(X)^2 = \Theta(\varepsilon^2 n^{2j})$, this contribution is negligible, and therefore the expression in (3.32) simplifies and we obtain

$$\mathbb{E}(X^2) = (1 \pm o(1)) \sum_{J_1 \in \binom{V}{j},\, J_2 \in \binom{n}{j} \setminus C_{J_1}} \Pr\left(J_1, J_2 \in \mathcal{L}\right).$$

Consequently, by (3.33) and (3.39), we have

$$\mathbb{E}\left(X^2\right) \leq (1+o(1))\binom{n}{j}\mathbb{P}\left(\mathcal{E}\right) \sum_{J_2 \in \binom{V}{j}\setminus C_{J_1}} \mathbb{P}\left(J_2 \in \mathcal{L} \mid \mathcal{E}\right).$$
$$\leq (1+o(1))\frac{2\varepsilon}{c_0}\binom{n}{j} \sum_{J_2 \in \binom{V}{j}\setminus C_{J_1}} \mathbb{P}\left(J_2 \in \mathcal{L} \mid \mathcal{E}\right). \tag{3.40}$$

Therefore assume that J_2 lies outside C_{J_1} and fix it for the remainder of the proof. We delete all the j-sets of C_{J_1} from $\mathcal{H}^k(n,p)$ – any k-sets containing them may now no longer be queried.

We start a new BFS process from J_2 growing a component in this restricted hypergraph which we denote by $\mathcal{H}'$. The partial component obtained at the beginning of round i is denoted by $C_{J_2}(i)$. Similarly as before, we denote the upper coupling branching process for this exploration by $\mathcal{T}^{J_2}$ (which is also an independent instance of $\mathcal{T}^*$). We continue to grow the component until one of the following two stopping conditions is reached:

(T1) the component of J_2 (in $\mathcal{H}'$) is fully explored;
(T2) the (partial) component $C_{J_2}(i)$ has reached size $\Lambda = \lambda n^j$.

Again, we only stop at the beginning of a round, and denote the corresponding stopping time by

$$i_2 = \min_{i\in\mathbb{N}}\{(\text{T1}) \vee (\text{T2}) \text{ holds in round } i\}.$$

Moreover, with slight abuse of notation write C_{J_2} instead of $C_{J_2}(i_2)$.

If C_{J_2} has size at least Λ, then certainly the component containing J_2 in $\mathcal{H}^k(n,p)$ has size at least Λ. On the other hand, if C_{J_2} stops because of (T1), then it may be that in fact the whole component is large, but we missed some of it because of the j-sets of C_{J_1} which we deleted. We will show that the number of k-sets we were forbidden to query is small enough that whp none of them would have resulted in an edge of the hypergraph.

We first observe that such queries can only occur between C_{J_2} and the boundary ∂C_{J_1} of C_{J_1} (any j-sets of C_{J_1} not in ∂C_{J_1} were already fully explored). The intuition is that C_{J_2} remains small, while the boundary of C_{J_1} is very small, so the number of pairs of j-sets, one from each side, should still be small. We might therefore expect that there are very few k-sets containing such pairs.

The problem with this naive argument is that the number of k-sets containing a pair of j-sets is heavily dependent on the size of their intersection. While on the whole most pairs of j-sets do not intersect, those which do carry disproportionately large weight because there are many more k-sets containing both of them.

However, we already proved the smooth boundary lemma (Lemma 3.8) allows us to overcome this obstacle: it implies that ∂C_{J_1} is 'smooth' in the sense that for any $0 \leq \ell \leq j-1$ and for any ℓ-set L, the number of j-sets in ∂C_{J_1} which contain L is about the 'right' number, and in particular almost the same regardless

of the choice of L (though dependent on $|L| = \ell$). This statement is formalised in Lemma 3.28. Recall (from Section 3.1.4.2) that $d_L(\partial C_{J_1})$ denotes the number of j-sets of the boundary ∂C_{J_1} containing the set L.

Lemma 3.28. *Conditioned on $\mathcal{E}$, with probability at least $1 - \exp(-\Theta(n^{\delta/4}))$, for every $0 \le \ell \le j-1$ and ℓ-set L the following holds (at the beginning of round i_1):*

$$d_L(\partial C_{J_1}) = (1 \pm o(1)) \frac{|\partial C_{J_1}|}{\binom{n}{j}} \binom{n}{j-\ell}.$$

Proof. Recall (from Section 3.3.1) that $i_0(j-1)$ was defined to be the round when the generations reach size $n^{j-1+\delta}$ for the first time, and after round $i_1(j-1) = i_0(j-1) + \Theta(\log n)$ generations are smooth by Lemma 3.8. Hence it only remains to prove that $i_1(j-1) \le i_1$. This follows directly from Lemma 3.11 stating that conditional on $\mathcal{E}$ we have $i_1 - i_0(j-1) \ge \varepsilon^{-1} \log n \gg \log n$. □

We note that Lemma 3.28 is already much weaker than Lemma 3.8 itself, but sill considerably stronger than we would need for the proof of Theorem 3.2(a), for which concentration within a constant multiplicative factor would be sufficient.

Next we demonstrate how to employ Lemma 3.28 to provide an upper bound on the probability of the (partial) component C_{J_2} being large conditioned on $\mathcal{E}$. We denote the upper coupling process of C_{J_2} by $\mathcal{T}^{J_2}$ (being an instance of $\mathcal{T}^*$). Conditioned on $\mathcal{E}$ and $\left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}$, we need to analyse the event $\mathcal{F}$ of J_2 being a (potential) *false negative*, i.e. C_{J_2} contains fewer than Λ j-sets but the component of J_2 in $\mathcal{H}^k(n,p)$ is larger than C_{J_2}. (In fact, a genuine false negative would require the component to be larger than Λ, but bounding the probability of this weaker event will be sufficient.)

We know, by Lemma 3.17, that the boundary ∂C_{J_1} has size at most $2\lambda\Lambda$ and each ℓ-set is contained in $O(\lambda\Lambda/n^\ell)$ sets of ∂C_{J_1}, by Lemma 3.28. Thus for a j-set of C_{J_2}, the number of k-sets which we did not query because they contained j-sets of C_{J_1} is

$$\sum_{\ell=0}^{j-1} O(\lambda\Lambda/n^\ell) \binom{j}{\ell} \binom{n-j}{k-2j+\ell} = O(\lambda\Lambda n^{k-2j}).$$

Therefore, if we assume C_{J_2} contains precisely $r \in \mathbb{N}$ j-sets, the expected number of edges within these disallowed k-sets is at most $O\left(\lambda\Lambda n^{k-2j} rp\right) = O\left(\lambda^2 r\right)$, and thus the probability that we have overlooked at least one edge is $O\left(\lambda^2 r\right)$. Hence, by the law of total probability we obtain

$$\begin{aligned}
\mathbb{P}\left(\mathcal{F} \mid \mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}\right) &= O\left(\lambda^2 \sum_{r=0}^{\infty} r \mathbb{P}\left(|C_{J_2}| = r \mid \mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}\right)\right) \\
&= O\left(\lambda^2 \mathbb{E}\left(|C_{J_2}| \mid \mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}\right)\right) \\
&= o\left(\varepsilon^2 \mathbb{E}\left(\left|\mathcal{T}^{J_2}\right| \mid \mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}\right)\right).
\end{aligned}$$

Now since the process $\mathcal{T}^{J_2}$ is independent of $\mathcal{E}$ and is distributed as $\mathcal{T}^*$ we obtain

$$\mathbb{P}\left(\mathcal{F} \mid \mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}\right) = o\left(\varepsilon^2 \mathbb{E}\left(|\mathcal{T}_{\mathcal{D}}|\right)\right) \overset{(3.28)}{=} o(\varepsilon). \tag{3.41}$$

In other words, false negatives are very unlikely. This solves the main difficulty in the proof of Claim 3.27.

Since we only need an upper bound for the probability that the component of J_2 in $\mathcal{H}^k(n,p)$ is large, we do not need to care about *false positives*. Therefore we consider J_2 to be large in the following three cases:

- $\mathcal{T}^{J_2}$ survives;
- $\mathcal{T}^{J_2}$ dies out and J_2 is a false negative;
- $\mathcal{T}^{J_2}$ dies out, J_2 is not a false negative and its component in $\mathcal{H}^k(n,p)$ is large.

Observe that every j-set $J_2 \in \mathcal{L}$ will satisfy exactly one of these conditions.

Still assuming that $\mathcal{E}$ holds, we calculate the probabilities of these events. For the first case we obtain

$$\mathbb{P}\left(\left|\mathcal{T}^{J_2}\right| = \infty \mid \mathcal{E}\right) = \mathbb{P}\left(\left|\mathcal{T}^{J_2}\right| = \infty\right) \overset{(3.26)}{\sim} \frac{2\varepsilon}{c_0}, \tag{3.42}$$

since the branching process $\mathcal{T}^{J_2}$ is independent of $\mathcal{E}$. Moreover, estimate (3.41) immediately shows

$$\mathbb{P}\left(\left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\} \wedge \mathcal{F} \mid \mathcal{E}\right) \leq \mathbb{P}\left(\mathcal{F} \mid \mathcal{E} \wedge \left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\}\right) \overset{(3.41)}{=} o(\varepsilon). \tag{3.43}$$

For the last case let us first note that

$$\{\neg\mathcal{F} \wedge \{J_2 \in \mathcal{L}\}\} \implies \{|C_{J_2}| \geq \Lambda\} \implies \left\{\left|\mathcal{T}^{J_2}\right| \geq \Lambda\right\}$$

and $\mathcal{T}^{J_2}$ is independent of the event $\mathcal{E}$. Thus we have

$$\begin{aligned}
\mathbb{P}\left(\left\{\left|\mathcal{T}^{J_2}\right| < \infty\right\} \wedge \neg\mathcal{F} \wedge \{J_2 \in \mathcal{L}\} \mid \mathcal{E}\right) &\leq \mathbb{P}\left(\Lambda \leq \left|\mathcal{T}^{J_2}\right| < \infty\right) \\
&\leq \mathbb{P}\left(\left|\mathcal{T}^{J_2}\right| \geq \Lambda \mid \left|\mathcal{T}^{J_2}\right| < \infty\right) \\
&= \mathbb{P}\left(|\mathcal{T}_{\mathcal{D}}| \geq \Lambda\right) \\
&\overset{(3.28)}{\leq} (1+o(1))\varepsilon^{-1}/\Lambda = o(\varepsilon),
\end{aligned} \tag{3.44}$$

by Markov's inequality and since $\varepsilon^2\Lambda \to \infty$. Consequently, by estimates (3.42), (3.43) and (3.44), we obtain

$$\mathbb{P}\left(J_2 \in \mathcal{L} \mid \mathcal{E}\right) \leq \frac{2\varepsilon}{c_0} + o(\varepsilon). \tag{3.45}$$

This yields a good enough upper bound on the second moment:

$$\begin{aligned}
\mathbb{E}\left(X^2\right) &\overset{(3.40)}{\leq} (1+o(1))\binom{n}{j}\frac{2\varepsilon}{c_0} \sum_{J_2 \in \binom{V}{j}\setminus C_{J_1}} \mathbb{P}\left(J_2 \in \mathcal{L} \mid \mathcal{E}\right) \\
&\overset{(3.45)}{\leq} (1+o(1))\left(\frac{2\varepsilon}{c_0}\binom{n}{j}\right)^2 \\
&\overset{(3.31)}{=} (1+o(1))\mathbb{E}\left(X\right)^2,
\end{aligned}$$

and therefore we have for the variance

$$\mathbb{V}\left(X\right) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = o(\mathbb{E}(X)^2).$$

Hence Chebyshev's inequality tells us that for any constant $\zeta > 0$,

$$\mathbb{P}(X \neq (1 \pm \zeta)\mathbb{E}(X)) \leq \frac{2\,\mathbb{V}(X)}{\zeta^2\,\mathbb{E}(X)^2} = o(1/\zeta^2) = o(1),$$

and so whp $X = (1 \pm o(1))\mathbb{E}(X)$. We have thus shown that the number of j-sets in large components is approximately as expected, and consequently Claim 3.27 follows from (3.31). □

3.4.3. **Sprinkling.** In order to complete the proof of Theorem 3.2(a) based on Claim 3.27 we also need to know that all (or at least almost all) j-sets of $\mathcal{L}$ lie in the same component. To prove this, we use a standard sprinkling argument.

Proof of Theorem 3.2(a). Let $p_2 := \frac{\varepsilon}{\lambda^3 n^k}$ and p_1 be such that

$$p_1 + p_2 - p_1 p_2 = (1+\varepsilon)\hat{p}_{\mathrm{g}} = p.$$

Recall that $\lambda^3 n^j \to \infty$, thus $p_2 = o(\varepsilon \hat{p}_{\mathrm{g}})$, and consequently $p_1 = (1 + \varepsilon + o(\varepsilon))\hat{p}_{\mathrm{g}}$. We expose the edges of $\mathcal{H}^k(n,p)$ in two rounds and couple the random hypergraphs so that we have $\mathcal{H}^k(n,p_1) \cup \mathcal{H}^k(n,p_2) = \mathcal{H}^k(n,p)$.

Now observe that after the first round of exposure the number of j-sets in large components already satisfies the asymptotics in Claim 3.27. In other words, $\mathcal{H}^k(n,p_1)$ has $(1 \pm o(1))\frac{2\varepsilon}{\binom{k}{j}-1}\binom{n}{j}$ j-sets in large components. Let us denote the large components in $\mathcal{H}^k(n,p_1)$ by $\mathcal{C}_1, \ldots, \mathcal{C}_s$, where $s \leq (1 \pm o(1))\varepsilon\binom{n}{j}/\Lambda = O(\varepsilon/\lambda)$.

Furthermore assume that $s \geq 2$, i.e. there are at least two large components, since otherwise there is nothing to prove. We will merge all of them with $\mathcal{C}_1$ by using a union bound, let us therefore concentrate on merging $\mathcal{C}_1$ with $\mathcal{C}_2$. By Lemma 3.14 each $(j-1)$-set lies in $\Omega(\lambda n)$ many j-sets of any large component of $\mathcal{H}^k(n,p_1)$.

Pick a $(j-1)$-set J' and consider the $(k-j+1)$-uniform link hypergraph of J', i.e. the hypergraph on $[n] \setminus J'$ whose edges are all $(k-j+1)$-sets which, together with J', form an edge of $\mathcal{H}^k(n,p)$. This has two distinct vertex-components $\mathcal{C}_1^{(J')}$ and $\mathcal{C}_2^{(J')}$ containing $\Omega(\lambda n)$ vertices each, corresponding to $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. There are therefore $\Omega(\lambda^2 n^{k-j+1})$ possible $(k-j+1)$-sets which intersect both $\mathcal{C}_1^{(J')}$ and $\mathcal{C}_2^{(J')}$. (These correspond to k-sets containing J', which also contain a j-sets from both $\mathcal{C}_1$ and $\mathcal{C}_2$.)

Moreover, we may do the same for any $(j-1)$-set J' (of which there are $\binom{n}{j-1}$ in total). In doing so we may count k-sets multiple times, but certainly at most $\binom{k}{j} = O(1)$ times. Hence there are at least $\binom{n}{j-1}\binom{k}{j}^{-1}\Omega(\lambda^2 n^{k-j+1}) = \Omega(\lambda^2 n^k)$ k-sets which, if they form an edge in $\mathcal{H}^k(n,p_2)$, will merge the components $\mathcal{C}_1$ and $\mathcal{C}_2$. Thus the probability that these two components do not merge is at most

$$(1-p_2)^{\Omega(\lambda^2 n^k)} \leq \exp\left(-\frac{\varepsilon}{\lambda^3 n^k}\,\Omega(\lambda^2 n^k)\right) \leq \exp\left(-\Omega(\varepsilon/\lambda)\right).$$

Consequently the probability that at least one of $\mathcal{C}_2, \ldots, \mathcal{C}_s$ does not merge with $\mathcal{C}_1$ is at most

$$O(\varepsilon/\lambda)\exp\left(-\Omega(\varepsilon/\lambda)\right) = o(1).$$

Hence, in $\mathcal{H}^k(n,p)$, whp there is a single component of size $(1 \pm o(1))\frac{2\varepsilon}{\binom{k}{j}-1}\binom{n}{j}$, while this is also the number of j-sets in large components by Claim 3.27. Thus, even though sprinkling the edges with probability p_2 may have created more large components, they can only have total size $o(\varepsilon n^j)$, completing the proof. □

3.5. Concluding remarks

Since random graphs have been so extensively studied, various further questions immediately suggest themselves, regarding whether we can also prove similar results for hypergraphs. Before we discuss these topics we briefly comment on a very interesting alteration of the breadth-first search exploration process for hypergraphs.

Hypertrees. In this chapter, we used the fact that Lemma 3.13 can be applied to the search process BFS to show that $\mathcal{T}_* \prec \text{BFS} \prec \mathcal{T}^*$. However, this also applies to the following variant of the search process: we define BFS2 to be the corresponding breadth-first search process in which a k-set may only be queried if it contains c_0 neutral j-sets (as opposed to at least one for BFS).

BFS2 is a search algorithm specifically looking for a tree, where we define a tree to be a component with e edges and $c_0 e + 1$ many j-sets. Note that this algorithm will not necessarily reveal all of a component; however all the arguments involving BFS which we use in this chapter also hold for BFS2. Thus we deduce that the number of j-sets contained in large trees is approximately $\frac{2\varepsilon}{c_0}\binom{n}{j}$, and indeed these (rooted) trees are smooth in the sense of the smooth boundary lemma.

Critical window. The natural candidate for the optimal lower bound on ε in Theorem 3.2 would be $\varepsilon^3 n^j \to \infty$, for which the bounds from the super-critical case ($\Theta(\varepsilon n^j)$) and the sub-critical case ($O(\varepsilon^{-2}\log n)$) match up to the $\log n$ term, suggesting that we have a smooth transition. In particular, this condition is also sufficient for the sprinkling argument in Section 3.4.3 to work.

However, for $j \geq 2$, our proof method requires additional conditions. Thus for these cases, our range of ε is probably not best possible. On the other hand, for $j = 1$ most of our machinery is not needed, and the proofs simplify and actually work under the optimal assumption $\varepsilon^3 n \to \infty$.

The additional condition arises in the proof because we need the smooth boundary lemma. The boundary has size up to $\lambda^2 n^j = o(\varepsilon^2 n^j)$, and the typical degree within the boundary of an ℓ-set is $o(\varepsilon^2 n^{j-\ell})$. If we are to show that these degrees are concentrated, we need this typical degree to be large, which in the case of $\ell = j-1$ means $\varepsilon^2 n$ must be large, thus leading to our condition on ε. However, currently there is another bottleneck in the proof of Lemma 3.21, which provides an upper bound on the biped degrees. In a nutshell we lose an factor of λ by estimating $\Delta_\ell(\partial C_J(i)) \leq \Delta_\ell(C_J(i))$, while i is still too small to expect smoothness of ℓ-sets for all $1 \leq \ell \leq j-1$. We use this estimate for instance to derive (3.16). Unfortunately

this error compounds similarly as in the previous obstruction meaning that $\varepsilon^3 n^{1-2\delta}$ must be large.

If we were to attempt to do away with this condition, we would need to have some control over degrees which may be very small. Presumably we would need to determine more precisely what the probability distribution of such degrees is.

Limiting distribution for giant component. The size of the giant component shortly after the phase transition is a random variable whose mean we have (asymptotically) determined in this chapter, and we have further shown that it is concentrated around its mean. However, we have not proved what the actual distribution of this random variable is.

The most likely candidate would be a normal distribution as is the case for graphs (as proved by Pittel and Wormald [95] and by Luczak and Łuczak [85]). More generally for the 1-component in k-uniform hypergraphs, analogous results were proved for various ranges of ε by Karoński and Łuczak [79] and by Behrisch, Coja-Oghlan and Kang [25], and for the whole of the supercritical regime ($\varepsilon \gg n^{-1/3}$) by Bollobás and Riordan [36]. For these cases, central and local limit theorems are known. It would be interesting to prove similar results for the size of the largest j-component.

Structure of components. For graphs, it is well known that shortly after phase transition, all components are whp trees or at most unicyclic. It would be interesting to know if something similar holds in hypergraphs. (A natural generalisation of unicyclic would be the most tree-like connected non-tree structure, i.e. a component with e edges and $c_0 e$ many j-sets. In this case, when discovering the component, exactly one j-set would be seen twice.) This would be particularly interesting when aiming for a local limit theorem for the size of the largest j-component, because in both [25, 79] ($j = 1$) it proved crucial to closely investigate the interactions of small components with the giant component.

Cores. The study of cores in random graphs was initiated by Bollobás [32]. The ℓ-core is the (unique) maximal subgraph of minimum degree at least ℓ. The threshold for the existence of a giant ℓ-core for $\ell \geq 3$ was determined by Pittel, Spencer, and Wormald [94].

For hypergraphs, the degree has many possible generalisations just as connectedness does. For each of these, we may then define the ℓ-core and consider its properties. For vertex-degree, Molloy [87] determined the critical threshold for the existence of a non-empty ℓ-core and its asymptotic size, but in general this is still an open question.

3.A. Appendix: branching processes

Recall that we aim to calculate the asymptotic value of the survival probability ϱ of the branching process $\mathcal{T}^*$, in which the number of children has distribution $c_0 \cdot \mathrm{Bin}\big(\binom{n}{k-j}, p\big)$, where $p = (1+\varepsilon)\hat{p}_{\mathrm{g}}$. For technical reasons, it is slightly easier

first to calculate the probability ϱ_0 that at least one of c_0 independent branching processes (each an instance of $\mathcal{T}^*$) survives. In this case the number of children in the first generation has distribution $\mathrm{Bin}\big(c_0\binom{n}{k-j}, p\big)$.

By standard results for branching processes (see [68]) we know that $\varrho > 0$. We note that the processes all die out if and only if all subprocesses starting at children in the first generation die out. Making a case distinction on the number of such children, we obtain

$$\begin{aligned} 1-\varrho_0 &= \sum_{i=0}^{c_0\binom{n}{k-j}} \mathbb{P}\left(\mathrm{Bin}\left(c_0\binom{n}{k-j}, p\right) = i\right)(1-\varrho_0)^i \\ &= \sum_{i=0}^{c_0\binom{n}{k-j}} \binom{c_0\binom{n}{k-j}}{i} p^i (1-p)^{c_0\binom{n}{k-j}-i}(1-\varrho_0)^i, \end{aligned}$$

and thus

$$1-\varrho_0 = (p(1-\varrho_0)+1-p)^{c_0\binom{n}{k-j}} = (1-p\varrho_0)^{c_0\binom{n}{k-j}} = (1-(1+\varepsilon)\hat{p}_{\mathrm{g}}\varrho_0)^{\hat{p}_{\mathrm{g}}^{-1}}.$$

Solving this equation for ε yields

$$\varepsilon = \frac{1-(1-\varrho_0)^{\hat{p}_{\mathrm{g}}}}{\varrho_0\hat{p}_{\mathrm{g}}} - 1 = \frac{f(\varrho_0)}{\varrho_0\hat{p}_{\mathrm{g}}}, \tag{3.46}$$

for

$$\begin{aligned} f(\varrho_0) &= 1-\varrho_0\hat{p}_{\mathrm{g}} - (1-\varrho_0)^{\hat{p}_{\mathrm{g}}} \\ &= \frac{\hat{p}_{\mathrm{g}}(1-\hat{p}_{\mathrm{g}})}{2}\varrho_0^2 + \frac{\hat{p}_{\mathrm{g}}(1-\hat{p}_{\mathrm{g}})(2-\hat{p}_{\mathrm{g}})}{6}\varrho_0^3 + \dots . \end{aligned}$$

Since this expression has only non-negative coefficients, (3.46) implies that

$$\varepsilon \geq \frac{\frac{\hat{p}_{\mathrm{g}}(1-\hat{p}_{\mathrm{g}})}{2}\varrho_0^2}{\varrho_0\hat{p}_{\mathrm{g}}} = \frac{\varrho_0}{2} - O(\hat{p}_{\mathrm{g}}\varrho_0)$$

and hence $\varrho_0 = o(1)$ since $\varepsilon = o(1)$ and $\hat{p}_{\mathrm{g}} = o(1)$. Consequently we derive the asymptotic estimate

$$\varepsilon = \frac{\varrho_0}{2} + O\left(\varrho_0\hat{p}_{\mathrm{g}} + \varrho_0^2\right) \sim \frac{\varrho_0}{2} \tag{3.47}$$

from (3.46). Since $1-\varrho_0 = (1-\varrho)^{c_0}$, we have

$$\varrho = 1-(1-\varrho_0)^{\frac{1}{c_0}} = \frac{\varrho_0}{c_0} + O\left(\varrho_0^2\right) \overset{(3.47)}{\sim} \frac{2\varepsilon}{c_0}. \tag{3.48}$$

Remark 3.29. Note that an almost identical calculation shows that the survival probability ϱ_* of the lower coupling process $\mathcal{T}_*$ also satisfies

$$\varrho_* \sim \frac{2\varepsilon}{\binom{k}{j}-1}, \tag{3.49}$$

because each individual has $(1-\varepsilon_*)(1+\varepsilon) = 1+\varepsilon \pm o(\varepsilon)$ children in expectation.

Recall that $\mathcal{T}_{\mathcal{D}}$ is the dual process, i.e. the process $\mathcal{T}^*$ conditioned on $\mathcal{D}$, the event that $\mathcal{T}^*$ dies. We consider the children of an individual as being grouped into *litters*, each containing c_0 children, where the number of litters has distribution

$\mathrm{Bin}\big(\binom{n}{k-j}, p\big)$ in $\mathcal{T}^*$. We want to show that $\mathcal{T}_{\mathcal{D}}$ is a branching process in which each individual has a number of litters of children which has binomial distribution.

To analyse the dual process we consider the change of the probability of $\mathcal{D}$ subject to the presence $\mathcal{A}_e$ of a litter of children of an individual J, in other words $\mathbb{P}(\mathcal{A}_e) = p$ and $\mathcal{A}_e$ is independent from the rest of the process. We denote by $\partial(J)$ the set of individuals in the same generation (of $\mathcal{T}^*$) as J and calculate the probabilities $\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)$ and $\Pr(\mathcal{D} \mid \neg\mathcal{A}_e)$ by conditioning on the number of litters (of children) $s(\partial(J))$ of individuals in $\partial(J)$. Denoting $M := |\partial(J)|\binom{n}{k-j}$ we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e) &= \sum_{i=0}^{M-1} \mathbb{P}(s(\partial(J)) = i+1 \mid \mathcal{A}_e)\,(1-\varrho)^{(i+1)c_0} \\ &= \sum_{i=0}^{M-1} \mathbb{P}\big(\mathrm{Bin}(M-1,p) = i\big)(1-\varrho)^{(i+1)c_0}\end{aligned}$$

and similarly

$$\begin{aligned}\mathbb{P}(\mathcal{D} \mid \neg\mathcal{A}_e) &= \sum_{i=0}^{M-1} \mathbb{P}(s(\partial(J)) = i \mid \neg\mathcal{A}_e)\,(1-\varrho)^{ic_0} \\ &= \sum_{i=0}^{M-1} \mathbb{P}\big(\mathrm{Bin}(M-1,p) = i\big)(1-\varrho)^{ic_0}.\end{aligned}$$

Consequently we have

$$\frac{\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)}{\mathbb{P}(\mathcal{D} \mid \neg\mathcal{A}_e)} = (1-\varrho)^{c_0}, \tag{3.50}$$

and thus obtain

$$\begin{aligned}\mathbb{P}(\mathcal{A}_e \mid \mathcal{D}) &= \frac{\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)\,\mathbb{P}(\mathcal{A}_e)}{\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)\,\mathbb{P}(\mathcal{A}_e) + \mathbb{P}(\mathcal{D} \mid \neg\mathcal{A}_e)\,\mathbb{P}(\neg\mathcal{A}_e)} \\ &= \frac{\frac{\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)}{\mathbb{P}(\mathcal{D} \mid \neg\mathcal{A}_e)} \cdot \mathbb{P}(\mathcal{A}_e)}{\frac{\mathbb{P}(\mathcal{D} \mid \mathcal{A}_e)}{\mathbb{P}(\mathcal{D} \mid \neg\mathcal{A}_e)} \cdot \mathbb{P}(\mathcal{A}_e) + \mathbb{P}(\neg\mathcal{A}_e)} \\ &\overset{(3.50)}{=} \frac{(1-\varrho)^{c_0} p}{1 - p\,(1-(1-\varrho)^{c_0})},\end{aligned}$$

In particular, note that this probability is independent of the choice of e and J, hence we denote it by $p_{\mathcal{D}}$. Moreover we obtain the estimate

$$p_{\mathcal{D}} = \frac{(1-\varrho)^{c_0} p}{1 - O(p\varrho)} = p\,(1-\varrho)^{c_0} + O(p^2\varrho) \overset{(3.48)}{=} (1 - \varepsilon \pm o(\varepsilon))\hat{p}_{\mathrm{g}}\,.$$

Furthermore a very similar calculation shows that conditioned on $\mathcal{D}$ the presence of e is still independent of all other edges. Hence the dual process $\mathcal{T}_{\mathcal{D}}$ is a branching process whose offspring distribution is given by

$$c_0 \cdot \mathrm{Bin}\left(\binom{n}{k-j}, p_{\mathcal{D}}\right).$$

CHAPTER 4

Hitting time for connectedness in random hypergraphs

4.1. Introduction and main results

In the study of random graphs, one of the most well-known results concerns the *hitting time* for connectedness. More precisely, if we add randomly chosen edges one by one to an initially empty graph on n vertices, then at the moment the last isolated vertex gains its first edge, the whole graph will also become connected (this classical result was first proved by Bollobás and Thomason in [38]). This interplay between local and global properties is an example of the common phenomenon relating graph properties with their smallest obstruction; the graph can certainly not be connected while an isolated vertex still exists, but this smallest obstruction is also the critical one which is last to disappear.

In this chapter we generalise the result of Bollobás and Thomason to random k-uniform hypergraphs. For an integer $k \geq 2$, a k-uniform hypergraph consists of a set V of vertices together with a set E of (hyper-)edges, each consisting of k vertices. We need to define the notion of connectedness, for which there is a whole family of possible definitions. For any $1 \leq j \leq k-1$, we say that two distinct j-sets of vertices (i.e. j-element subsets of the vertex set V) $J_1 \neq J_2$ are *j-connected* if there is a sequence $e_1, \dots, e_m \in E$ of edges such that

- $J_1 \subset e_1$ and $J_2 \subset e_m$;
- $|e_i \cap e_{i+1}| \geq j$ for all $1 \leq i \leq m-1$.

In other words, we may walk from J_1 to J_2 using edges which consecutively intersect in at least j vertices. Any j-set is always j-connected to itself. This forms an equivalence relation on the set $\binom{V}{j}$ of j-sets. A *j-component* is an equivalence class of this relation, i.e. a maximal set of pairwise j-connected j-sets. The case $j = 1$ is also known as *vertex-connectedness*, and for $j \geq 2$ we use the term *high-order connectedness*.[1]

Note that in the case $k = 2, j = 1$ this is simply the usual definition of connectedness for graphs. More generally, for arbitrary $k \geq 2$ the case $j = 1$ is by far the most well-studied. The definition for general j is also entirely natural, albeit harder to visualise and often requires more complex analysis. In this chapter we will be interested in arbitrary $1 \leq j \leq k-1$ and $k \geq 3$.

[1]This notion is not to be confused with the *(vertex-)connectivity* of a (hyper-)graph H measuring the size of the smallest *vertex-separator* in H.

4.1.1. **Main results.** We first define the *uniform model*, the counterpart of the uniform model of Erdős and Rényi for graphs: given any natural numbers k, M, n such that $M \leq \binom{n}{k}$, the random hypergraph $\mathcal{H}^k(n, M)$ is a hypergraph chosen uniformly at random from all k-uniform hypergraphs on vertex set $\{1, \dots, n\}$ which have M edges. This is closely related to the *random hypergraph process* $\{\mathcal{H}^k(n, M)\}_M$ which is defined as follows:

- $\mathcal{H}^k(n, 0)$ is the hypergraph on vertex set $\{1, \dots, n\}$ with no edges;
- For $1 \leq M \leq \binom{n}{k}$, $\mathcal{H}^k(n, M)$ is obtained from $\mathcal{H}^k(n, M-1)$ by adding an edge chosen uniformly at random from among those k-sets which do not already form an edge.

Note that the random hypergraph obtained in the M-th step of the process has the same distribution as in the uniformly chosen random hypergraph $\mathcal{H}^k(n, M)$, so the notation is consistent.

We consider asymptotic properties of random hypergraphs and throughout this chapter any asymptotics are as $n \to \infty$. In particular we say *with high probability* (or *whp*) to mean with probability tending to 1 as $n \to \infty$.

We say that a j-set is *isolated* if it is not contained in any edges. It is trivial to see that if a hypergraph contains isolated j-sets, then it is not j-connected (assuming it has more than j vertices). Our main result is that this trivial smallest obstruction is also the critical one in a random hypergraph.

Let τ_c denote the time step in the hypergraph process $\{\mathcal{H}^k(n, M)\}_M$ at which the hypergraph becomes j-connected. Similarly, let τ_i denote the time at which the last isolated j-set disappears. Note that the properties of being j-connected and of having no isolated j-set are monotone increasing properties, so these two variables are well-defined. Furthermore, as noted above, $\tau_i \leq \tau_c$ holds deterministically.

Theorem 4.1. *For any $1 \leq j \leq k-1$ and $k \geq 3$, whp in the random hypergraph process $\{\mathcal{H}^k(n, M)\}_M$ we have $\tau_c = \tau_i$.*

Special cases of Theorem 4.1 were already proved by Poole [96] for $j = 1$, and by Kahle and Pittel [77] for $j = k - 1$.

The uniform model and the associated hypergraph process allow us to formulate exact hitting time results such as Theorem 4.1, which we prove in Section 4.4. However, the drawback is that the analysis of the model can become tricky due to the fact that the presence of different edges is not independent (the total number is fixed). For this reason, it is often easier to analyse the *binomial model*: $\mathcal{H}^k(n, p)$ is a random k-uniform hypergraph on vertex set $\{1, \dots, n\}$ in which each k-set is an edge with probability p independently of all other k-sets. In Section 4.2 we will show that if $p = M/\binom{n}{k}$, then the two models are very similar and we can transfer results from one model to the other.

For the proof of Theorem 4.1 we will also make use of the following result (Theorem 4.2), which is interesting in itself and is therefore stated in a significantly more general form than we need for Theorem 4.1. For integer valued random

variables Z and Z' we denote their *total variation distance* by $d_{TV}(Z, Z')$, i.e.

$$d_{TV}(Z, Z') = \frac{1}{2} \sum_{i \in \mathbb{Z}} |\mathbb{P}(Z = i) - \mathbb{P}(Z' = i)|.$$

For integer-valued random variables X_n and Y, we say X_n *converges in distribution to* Y, denoted by $X_n \xrightarrow{d} Y$, if for every integer i we have $\mathbb{P}(X_n = i) \to \mathbb{P}(Y = i)$.

Theorem 4.2. *For any $k \geq 3$ and $1 \leq j \leq k-1$ and for any integer $s \geq 0$, let*

$$p_s = p_s(n, k, j) = \frac{j \log n + s \log\log n + c_n}{\binom{n}{k-j}},$$

where $|c_n| = o(\log n)$, and let D_s be the number of j-sets of degree precisely s in $\mathcal{H}^k(n, p_s)$ (i.e. which lie in s edges). Then we have

$$d_{TV}\left(D_s, \mathrm{Po}\left(\mathbb{E}\left(D_s\right)\right)\right) = O(n^{-j}(\log n)^{s+1}). \tag{4.1}$$

In particular, D_s satisfies the following:

$$\begin{array}{lll} (i) & D_s = 0 \text{ whp} & \text{if } c_n \to \infty; \\ (ii) & D_s \xrightarrow{d} \mathrm{Po}\left(\frac{j^s e^{-c}}{j! s!}\right) & \text{if } c_n \to c \quad \text{for any } c \in \mathbb{R}; \\ (iii) & D_s \to \infty \text{ whp} & \text{if } c_n \to -\infty. \end{array}$$

We prove Theorem 4.2 in Section 4.3. These two theorems together give the following corollary, proved in Section 4.3.

Theorem 4.3. *Let $k \geq 3$ and $1 \leq j \leq k-1$, and let $p_0 = \frac{j \log n + c_n}{\binom{n}{k-j}}$.*

(a) If $c_n \to \infty$, then whp $\mathcal{H}^k(n, p)$ is j-connected (and therefore contains no isolated j-sets).

(b) If $c_n \to -\infty$, then whp $\mathcal{H}^k(n, p)$ contains isolated j-sets (and is therefore not j-connected).

In other words, the properties of being j-connected and having no isolated j-sets both undergo a (sharp) phase transition at threshold $\hat{p}_c$, given by

$$\hat{p}_c = \hat{p}_c(n, k, j) := \frac{j \log n}{\binom{n}{k-j}}.$$

4.1.2. **Methods.** The main contribution of this chapter is to deduce Theorem 4.1 from Theorem 4.2. Attempting to prove this directly using standard techniques generalised from the graph case does not work because j-components in a hypergraph may be strangely and non-intuitively distributed. To overcome this problem we quote a powerful result from Chapter 3, which guarantees one component with a large subset which is in some sense *smoothly distributed*. We then show that whp all non-trivial components are connected to this smooth subset.

4.1.3. **Notation and definitions.** We introduce a few more definitions before we proceed with the proofs. We fix $k \geq 3$ and $1 \leq j \leq k-1$ for the remainder of the chapter. The *order* $|H|$ of a hypergraph H is the number of vertices it contains, while its *size* $e(H)$ is the number of edges. Since a j-component consists of j-sets of vertices, we may view it as a j-uniform hypergraph in which the edges are the

j-sets in the component. In particular, the size of a j-component is the number of j-sets it contains. From now on we will use *component* to mean j-component.

We will sometimes need to relate the j-sets of a component to the edges of the hypergraph which connect them. To allow us to do this, for a k-uniform hypergraph H we define the *j-size* of H to be the number of j-sets contained in edges of H. We ignore floors and ceilings whenever they do not significantly affect the argument.

4.2. Asymptotic equivalence of $\mathcal{H}^k(n,M)$ and $\mathcal{H}^k(n,p)$

We need to know that $\mathcal{H}^k(n,p)$ and $\mathcal{H}^k(n,M)$ are roughly equivalent, which is a generalisation of a standard fact about the corresponding graph models (see [30, 74]). In fact, [74] considers a more general setting than we require here, but what we state is an immediate corollary of the results there (see [74], Corollary 1.16).

Let $N = \binom{n}{k}$ and to ease notation, for some property Q we will denote by $\mathbb{P}_M(Q)$ the probability that $\mathcal{H}^k(n,M)$ has property Q. $\mathbb{P}_p(Q)$ is defined similarly.

Lemma 4.4. *Let Q be some monotone increasing property of k-uniform hypergraphs and let $M = Np \to \infty$. Then*

(a) $\mathbb{P}_p(Q) \to 1$ implies $\mathbb{P}_M(Q) \to 1$;
(b) $\mathbb{P}_p(Q) \to 0$ implies $\mathbb{P}_M(Q) \to 0$.

This lemma allows us to transfer properties from $\mathcal{H}^k(n,p)$ to $\mathcal{H}^k(n,M)$ (transferring in the other direction is also possible, with some small modifications, but we will not need to do this here). However, this only works for monotonically increasing properties. This is fine for the properties of being j-connected or of having no isolated j-sets, but in the proof of Theorem 4.1 we will need to consider the probability of having a component of size r, for various fixed r. This property is not even convex (and nor is its complement) and so for this case we will need some more careful arguments.

The following standard argument allows us to transfer properties from the binomial to the uniform model provided that the failure probability is small enough.

Lemma 4.5. *Let Q be an arbitrary property, and suppose that $M \to \infty$ and $p = M/N \to 0$. Then*

$$\mathbb{P}_M(Q) \leq \frac{\mathbb{P}_p(Q)}{\mathbb{P}(e(\mathcal{H}^k(n,p)) = M)} = \Theta(M^{1/2})\mathbb{P}_p(Q).$$

Proof. The inequality follows from the fact that

$$\mathbb{P}_p(Q) = \sum_{m=0}^{N} \mathbb{P}_m(Q)\mathbb{P}(e(\mathcal{H}^k(n,p)) = m) \geq \mathbb{P}_M(Q)\mathbb{P}(e(\mathcal{H}^k(n,p)) = M).$$

For the equality we note that using Stirling's approximation we have

$$\binom{N}{M}p^M(1-p)^{N-M} = \Theta(1)\sqrt{\frac{N}{M(N-M)}}\frac{N^N}{M^M(N-M)^{N-M}}p^M(1-p)^{N-M},$$

and thus obtain $\mathbb{P}\left(e(\mathcal{H}^k(n,p)) = M\right) = \Theta(M^{-1/2})$. □

4.3. Degree distribution

Next we use the Chen-Stein method to approximate the distribution of the number of j-sets with a given degree.

Theorem 4.6 (Theorem 1.B in [23])**.** *Given a finite index set $\mathcal{I}$ and a random variable $W = \sum_{i\in\mathcal{I}} Z_i$, where Z_i is a Bernoulli random variable with parameter $p_i \in [0,1]$, denote by $\lambda = \sum_{i\in\mathcal{I}} p_i$ the expectation of W. Assume that for each $i \in \mathcal{I}$ there is a pair of coupled random variables (U_i, V_i) such that U_i has the distribution of W and $V_i + 1$ has the distribution of W conditioned on $\{Z_i = 1\}$. Then we have*

$$d_{TV}\big(W, \mathrm{Po}(\lambda)\big) \le \min\{1, \lambda^{-1}\} \sum_{i\in\mathcal{I}} p_i \mathbb{E}\left(|U_i - V_i|\right).$$

Now we use Theorem 4.6 to prove Theorem 4.2 stating that the number of j-sets of a given degree in $\mathcal{H}^k(n,p)$ is asymptotically Poisson distributed when p is close to $\frac{j\log n + s\log\log n}{\binom{n}{k-j}}$.

Proof of Theorem 4.2. Let $C = \binom{k}{j} - 1$. Fix an integer $s \ge 0$ and suppose $p = p_s = \frac{j\log n + s\log\log n + c_n}{\binom{n}{k-j}}$, where $|c_n| = o(\log n)$. Then the expected number of j-sets of degree s in $\mathcal{H}^k(n,p)$ satisfies

$$\begin{aligned}
\mathbb{E}(D_s) &= \binom{n}{j}\binom{\binom{n-j}{k-j}}{s} p^s (1-p)^{\binom{n-j}{k-j}-s} \\
&= (1\pm o(1))\frac{n^j}{j!}\frac{\left(\frac{n^{k-j}}{(k-j)!}\right)^s}{s!} p^s \exp\left(-p\binom{n}{k-j}\right) \\
&= (1\pm o(1))\frac{1}{j!s!} e^{s(k-j)\log n - s\log((k-j)!) + s\log p - s\log\log n - c_n} \\
&= (1\pm o(1))\frac{j^s}{j!s!} e^{-c_n}, \qquad (4.2)
\end{aligned}$$

since

$$\log p = -(k-j)\log n + \log\log n + \log(j(k-j)!) + O\left(\frac{\log\log n + |c_n|}{\log n} + \frac{1}{n}\right).$$

For any j-set J we denote its degree in $\mathcal{H}^k(n,p)$ by $\deg(J)$ and analyse how D_s changes by conditioning on the event $\{\deg(J_0) = s\}$ for an arbitrary j-set J_0.

First we construct $\mathcal{H}^k(n,p)$ and denote by E_0 the set of edges containing J_0, then we distinguish three cases:

(a) If $\deg(J_0) < s$, add $s - \deg(J_0)$ distinct k-sets chosen uniformly at random from $\left\{K \in \binom{V}{k} \;\middle|\; J_0 \subset K\right\} \setminus E_0$ to the hypergraph;
(b) If $\deg(J_0) = s$, do nothing;
(c) If $\deg(J_0) > s$, delete a set of $\deg(J_0) - s$ edges chosen uniformly at random from E_0.

We denote the resulting hypergraph by $\mathcal{H}^* = \mathcal{H}^*(J_0)$. For any j-set J we write $\deg^*(J)$ for its degree in $\mathcal{H}^*$ and $D_s^*(J_0)$ for the number of j-sets $J \ne J_0$ such

that $\deg^*(J) = s$. Furthermore observe that this construction provides a coupling of $\mathcal{H}^k(n,p)$ and $\mathcal{H}^*$ such that removing all edges containing J_0 in either one of them yields the same random hypergraph $\mathcal{H}^- = \mathcal{H}^-(J_0)$. For any j-set J we write $\deg^-(J)$ for its degree in $\mathcal{H}^-$.

Aiming to apply Theorem 4.6, we let $\mathcal{I}$ be the set of all j-sets and for all J let $Z_J = \mathbb{1}_{\{\deg(J)=S\}}$, $p_J = \mathbb{P}(\deg(J) = s)$, $U_J = W = D_s$ and $V_J = D_s^*(J_0)$. We observe that $\mathcal{H}^*$ has the same distribution as $\mathcal{H}$ conditioned on the event $\{\deg(J_0) = s\}$, so V_J has the same distribution as W conditioned on $\{Z_{J_0} = 1\}$.

Now applying Theorem 4.6 and using $\min\left\{1, \frac{1}{\mathbb{E}(D_s)}\right\} \le \frac{1}{\mathbb{E}(D_s)}$, we obtain

$$\begin{aligned} d_{TV}\big(D_s, \mathrm{Po}(\mathbb{E}(D_s))\big) &\le \frac{\sum_J \mathbb{P}(\deg(J) = s)\,\mathbb{E}\left(|D_s - D_s^*(J_0)|\right)}{\mathbb{E}(D_s)} \\ &= \mathbb{E}\left(|D_s - D_s^*(J_0)|\right). \end{aligned} \tag{4.3}$$

Hence it suffices to estimate the random variable $|D_s - D_s^*(J_0)|$. We observe that

$$\begin{aligned} |D_s - D_s^*(J_0)| &= \mathbb{1}_{\{\deg(J_0)=s\}} + \sum_{J\neq J_0} |\mathbb{1}_{\{\deg(J)=s\}} - \mathbb{1}_{\{\deg^*(J)=s\}}| \\ &\le \mathbb{1}_{\{\deg(J_0)=s\}} + \sum_{t=1}^{s} \sum_{\substack{J\neq J_0 \\ \deg^*(J)>\deg(J)}} \mathbb{1}_{\{\deg(J_0)=s-t\}} \\ &\qquad + \sum_{t=1}^{\binom{n-j}{k-j}-s} \sum_{\substack{J\neq J_0 \\ \deg^*(J)<\deg(J) \\ \deg^-(J)\le s}} \mathbb{1}_{\{\deg(J_0)=s+t\}}. \end{aligned}$$

To justify the inequality, first note that if $\deg(J_0) = s$, then $\mathcal{H} = \mathcal{H}^*$ and only the first term contributes. Furthermore, if $\deg(J_0) < s$, say $\deg(J_0) = s - t$ for some $t \in [1, s]$, then the only contribution to $|D_s - D_s^*(J_0)|$ comes from j-sets $J \neq J_0$ whose degree increased, i.e. $\deg^*(J) > \deg(J)$. Similarly, if $\deg(J_0) = s+t$ for some $t \in \left[1, \binom{n-j}{k-j} - s\right]$, observe that for a j-set J to contribute it is necessary to have either $\deg(J) = s$ or $\deg^*(J) = s$. Note that these cannot hold unless $\deg^-(J) \le s$, and we will simply bound the probability of this (more likely) event.

Moreover, each inner sum has at most Ct terms, since we certainly only sum over j-sets J whose degree has changed, and adding or deleting an edge influences the degree of at most C j-sets (other than J_0).

Note also that $\deg^-(J)$ has distribution

$$\mathrm{Bin}\left(\binom{n-j}{k-j} - \binom{n-|J_0\cup J|}{k-|J_0\cup J|}, p\right),$$

independently of $\deg(J_0)$, and the probability that $\deg^-(J) \le s$ is maximised when $|J_0 \cup J|$ is minimised. Hence for an upper bound we will assume that $|J_0 \cup J| = j+1$. With this motivation we write $\widetilde{N} := \binom{n-j}{k-j} - \binom{n-j-1}{k-j-1} = (1 \pm o(1))\binom{n}{k-j}$ and define

$$q = \mathbb{P}\left(\mathrm{Bin}\left(\widetilde{N}, p\right) \le s\right).$$

Combining these two arguments we obtain the upper bound

$$|D_s - D_s^*(J_0)| \leq \mathbb{1}_{\{\deg(J_0)=s\}} + \sum_{t=1}^{s} \mathbb{1}_{\{\deg(J_0)=s-t\}}\, Ct + \sum_{t=1}^{\binom{n-j}{k-j}-s} \mathbb{1}_{\{\deg(J_0)=s+t\}} \mathrm{Bin}(Ct, q).$$

Therefore, using the notation $x^+ = \max\{x, 0\}$ for any $x \in \mathbb{R}$, we have

$$\mathbb{E}\left(|D_s - D_s^*(J_0)|\right) \leq \mathbb{P}\left(\deg(J_0) = s\right) + C\mathbb{E}\left((s - \deg(J_0))^+\right) + Cq\mathbb{E}\left((\deg(J_0) - s)^+\right), \tag{4.4}$$

We can estimate both probabilities in (4.4) using

$$\begin{aligned}\mathbb{P}\left(\deg(J_0) = s\right) \leq q &= \sum_{i=0}^{s} \binom{\widetilde{N}}{i} p^i (1-p)^{\widetilde{N}-i} \\ &\leq O(1) \cdot \left(\widetilde{N}p\right)^s \exp\left(-\widetilde{N}p\right) \\ &= O((\log n)^s n^{-j}),\end{aligned}$$

where the second and third lines follow because s is bounded. Moreover, we have

$$\mathbb{E}\left(s - \deg(J_0)\right)^+ \leq s\,\mathbb{P}\left(\deg(J_0) \leq s\right) \leq sq = O((\log n)^s n^{-j})$$

and furthermore

$$\mathbb{E}\left(\deg(J_0) - s\right)^+ \leq \mathbb{E}(\deg(J_0)) + s = O(\log n).$$

Therefore (4.3) and (4.4) provide (4.1), i.e.

$$d_{TV}\left(D_s, \mathrm{Po}\left(\mathbb{E}\left(D_s\right)\right)\right) = O(n^{-j}(\log n)^{s+1}). \tag{4.5}$$

Now assume $\lim_{n\to\infty} c_n = c$. By (4.2) we know that $\mathbb{E}\left(D_s\right) \to \frac{j^s e^{-c}}{j!s!}$ and by the continuity in λ of the function $\mathbb{P}(\mathrm{Po}(\lambda) = i)$ for each i

$$\mathrm{Po}\left(\mathbb{E}\left(D_s\right)\right) \xrightarrow{d} \mathrm{Po}\left(\frac{j^s e^{-c}}{j!s!}\right),$$

hence by the triangle inequality and (4.1), case (ii) in the second claim follows. Cases (i) and (iii) can be easily deduced from case (ii). □

4.4. Hitting time for connectedness

The proof which we present is largely elementary except for the use of Theorem 4.2 and the concept of smoothness which we introduced and investigated in detail in Chapter 3. Note that in Chapter 3 we focused on $\mathcal{H}^k(n,p)$ for much smaller probabilities, and thus we are now not in the optimal range for the application of these methods. But nevertheless, our results will turn out to be strong enough for proving Theorem 4.1.

We provide the appropriate tool (Lemma 4.7) in Section 4.4.1. After deriving a second preliminary result in Section 4.4.2, we prove Theorem 4.1 in Section 4.4.3.

4.4.1. **Smooth subset.** We use the smooth boundary lemma (Lemma 3.8) to prove the existence of a component containing a 'reasonably large' subset which is 'smooth' in the sense that all $(j-1)$-sets are in about the 'right' number of j-sets of S (see Lemma 4.7 below). More precisely, we say that a set S of j-sets is *smooth* if every $(j-1)$-set is contained in $(1\pm o(1))\frac{|S|}{\binom{n}{j}}n$ j-sets of S.

Lemma 4.7. *Let $1 \le j \le k-1$, let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n^{1-\delta} \to \infty$, for some constant $\delta > 0$, and set $p^* = \frac{1+\varepsilon}{(\binom{k}{j}-1)\binom{n}{k-j}}$. Then whp there is a component of $\mathcal{H}^k(n,p^*)$ with a subset S of at least $\varepsilon^3 n^j$ many j-sets which is smooth:*

Each $(j-1)$-set in $\mathcal{H}^k(n,p^*)$ is contained in $(1\pm o(1))\frac{|S|}{\binom{n}{j}}n$ many j-sets of S.

We note that Lemma 4.7 is not stated explicitly in this form in Chapter 3, but is implicit in the proofs in Section 3.3. We therefore give a brief outline of how it can be deduced from our previous results.

Proof of Lemma 4.7. Starting from some j-set J, we explore the component containing J using a breadth-first search process $\mathrm{BFS}(J)$. This partitions the j-sets of the component into *generations*, which can be numbered according to the order they were discovered in.

We fix a starting j-set J which lies in the largest component of $\mathcal{H}^k(n,p)$, let ∂C_g denote the g-th generation of this search process $\mathrm{BFS}(J)$, and $C_g = \cup_{g' \le g} \partial C_{g'}$. Then there are generations g_0 and g_1 such that the following statements hold whp.

(1) Either $|\partial C_{g_1}| \ge \varepsilon^3 n^j$ or $|C_{g_1}| \ge \varepsilon^{3/2} n^j$;
(2) $|C_{g_0}| = o(|C_{g_1}|)$ (and in particular $g_0 < g_1$);
(3) Every generation ∂C_g with $g_0 \le g \le g_1$ is smooth.

We set $g_0 = i_1(j-1)$ and $g_1 = i_1$, where $i_1(j-1)$ and i_1 are defined in Section 3.3. Recall that i_1 is the round at which one of three stopping conditions (S1), (S2) or (S3) is invoked, and these stopping conditions contain a parameter λ, which we choose to be $\lambda = \varepsilon^{3/2}$.

Property (1) follows from these stopping conditions. We use here the fact that J is in the largest component of $\mathcal{H}^k(n,p)$, which is a giant component whp by Theorem 3.2(a), therefore whp either (S2) or (S3) is invoked at time i_1 (as (S2) would be invoked before (S1)).

Property (2) follows from Lemmas 3.10, 3.11 and 3.16. More precisely, whp $g_0 = i_1(j-1) = i_0(j-1) + O(\log n)$, while $g_1 = i_1 \ge i_0(j-1) + \Omega(\varepsilon^{-1}\log n)$ by Lemma 3.11. Furthermore, whp $|C_J(i_0(j-1))| = o(\varepsilon^{3/2} n^j)$ by Lemma 3.10 applied with $s = j$ and $\ell = 0$. Finally, by Lemma 3.16, whp the generations between g_0 and g_1 are at least as large as those between $i_0(j-1)$ and g_0, and there are significantly more of them ($\Omega(\varepsilon^{-1}\log n)$ compared to $O(\log n)$).

Finally, property (3) is given by Lemma 3.8 (with $\ell = j-1$ and using the fact that $i_1 \ge i_1(j-1)$.

We now use these three properties to prove the existence of the set S. We make a case distinction based on (1). If $|\partial C_{g_1}| \ge \varepsilon^3 n^j$, then we simply set $S =$

∂C_{g_1}, and S is smooth by (3). On the other hand, if $|C_{g_1}| \geq \varepsilon^{3/2} n^j$, then we let $S = C_{g_1} \setminus C_{g_0-1}$. Then since every generation from g_0 to g_1 is smooth, and since a union of smooth sets is also smooth, we have that S is smooth. Furthermore, $|S| = (1 - o(1))|C_{g_1}| \geq \varepsilon^3 n^j$. □

Lemma 4.7 has the following corollary which we will apply later.

Corollary 4.8. *Let $1 \leq j \leq k-1$, let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n^{1-\delta} \to \infty$, for some constant $\delta > 0$, and suppose $p \geq \frac{1+\varepsilon}{(\binom{k}{j}-1)\binom{n}{k-j}}$. Then whp there is a component of $\mathcal{H}^k(n,p)$ with a subset S of at least $\varepsilon^3 n^j$ many j-sets which is smooth:*

Each $(j-1)$-set in $\mathcal{H}^k(n,p^*)$ is contained in $(1 \pm o(1))\frac{|S|}{\binom{n}{j}} n$ many j-sets of S.

Proof. We set $p^* = \frac{1+\varepsilon}{(\binom{k}{j}-1)\binom{n}{k-j}}$ and $p' = \frac{p-p^*}{1-p^*}$ and let $\mathcal{H}_1 = \mathcal{H}(n,p^*)$ and $\mathcal{H}_2 = \mathcal{H}(n,p')$ independently. Observe that we may couple in such a way that $\mathcal{H}^k(n,p) = \mathcal{H}_1 \cup \mathcal{H}_2$. Furthermore, by Lemma 4.7, whp $\mathcal{H}_1$ has a component containing a smooth set S of the appropriate size. In $\mathcal{H}^k(n,p)$ this component may be bigger than in $\mathcal{H}_1$, but certainly still contains S. □

4.4.2. **Well-constructed hypergraphs.** We will also use the following proposition. We say that a hypergraph is *well-constructed* if it can be generated from an initial j-set via a search process, i.e. by successively adding edges such that each edge contains at least one previously discovered j-set, and such that each edge also contains at least one previously undiscovered j-set.

Proposition 4.9. *Up to isomorphism, the number of well-constructed k-uniform hypergraphs of j-size s is at most 2^{ks^2}.*

Proof. We explore the hypergraph by adding the edges one by one in the order in which it is well-constructed. The resulting hypergraph is uniquely determined, up to isomorphism, by the intersection of each edge with the previous vertices (though we will multiple count the isomorphism classes, this is permissible for an upper bound). When adding the i-th edge, we certainly have at most $(i-1)k$ vertices so far, and so the number of possible intersections is at most $2^{(i-1)k}$. Multiplying over all edges, of which there are certainly at most s (each edge gives at least one new j-set), we have that the number of such hypergraphs is at most $2^{\sum_{i=1}^{s}(i-1)k} \leq 2^{ks^2}$. □

4.4.3. **Critical obstruction for connectedness.** We now prove Theorem 4.1 using Corollary 4.8 and Proposition 4.9.

Proof of Theorem 4.1. Let us consider any p, M satisfying

$$\frac{j \log n - \xi}{\binom{n}{k-j}} \leq p = M/N \leq \frac{j \log n + \xi}{\binom{n}{k-j}} \tag{4.6}$$

where $\xi = \log\log n$. We apply Theorem 4.2 (with $s = 0$ and $c_n = \pm\xi$) and Lemma 4.4 to observe that in both $\mathcal{H}^k(n,p)$ and $\mathcal{H}^k(n,M)$, whp there are isolated j-sets at the lower end of this range but not at the upper end. We will now prove

that other than these isolated j-sets, there is just one very large component whp. We certainly know by Corollary 4.8 that there is a component containing a large smooth set S. For the rest of the proof we fix this component and the set S.

We consider the possibility that there is a second non-trivial component containing r j-sets, and make a case distinction on the size of r. Note that for any component of size r in a k-uniform hypergraph, there is a well-constructed subhypergraph of every (up to a constant $\binom{k}{j}$ error) j-size up to r. (More precisely, for any integer $r' \le r$ there exists an integer r'' with $|r'' - r'| \le \binom{k}{j}$ and a well-constructed subhypergraph of j-size r''.)

Let us set $r_0 = \log\log n$ and distinguish two cases.

Case 1: $2 \le r \le r_0$. We first observe that in a component of size $r \ge 2$ we must have at least one edge, and therefore at least $\binom{k}{j} \ge k \ge 3$ j-sets, i.e. we automatically have $r \ge 3$.

We show that the expected number of components of size r is very small and apply Markov's inequality. Any component of size r can be associated with a well-constructed hypergraph H of j-size r which is isolated from the remaining j-sets of $\mathcal{H}^k(n,p)$. Then $e(H) \le r$ and furthermore $|H| \le j + (k-j)e(H)$, since each new edge of H gives at most $k-j$ new vertices. For each j-set of H, we have at least $\binom{n-j}{k-j} - r\binom{n-j-1}{k-j-1}$ non-edges (any k-set containing this j-set but no other j-sets of H). Thus the expected number of isolated copies of H in $\mathcal{H}^k(n,p)$ satisfies

$$\mathbb{E}(X_H) \le n^{j+(k-j)e(H)} p^{e(H)} (1-p)^{r\left(\binom{n-j}{k-j} - r\binom{n-j-1}{k-j-1}\right)} \tag{4.7}$$

and so

$$\begin{aligned}\log(\mathbb{E}(X_H)) \le\ & (j + (k-j)e(H))\log n + O(r\log\log n) \\ & - (k-j)e(H)\log n - (1 - O(r/n) - O(\xi/\log n))rj\log n \\ =\ & (1 - r \pm o(1))j\log n \le (-3rj/5)\log n.\end{aligned}$$

Note that this bound does not depend on the specific structure of H, only on the number of j-sets r. Let X_r be the number of components of size r. Then by Proposition 4.9 we have

$$\mathbb{E}(X_r) \le 2^{kr^2} n^{-3rj/5} \le n^{-4rj/7}$$

where for the last inequality we use the fact that $r \le r_0 = o(\log n)$. By taking a union bound over all $3 \le r \le r_0$, we conclude that with probability at least $1 - 2n^{-12j/7}$ there are no components of this size.

Case 2: $r \ge r_0$. In this case, rather than looking at the full component we look at a well-constructed subgraph H of j-size r_0. Such a subgraph certainly exists up to an additive $\binom{k}{j}$ error term in the j-size, which will not affect calculations significantly. Most of the calculations which lead to (4.7) are still valid, replacing r by r_0. However, since we are no longer considering a full component, we must be more careful about the number of non-edges.

At this point we make use of the set S of j-sets guaranteed by Corollary 4.8, which lie in a different component to H. For each of the r_0 many j-sets of H, pick an arbitrary $(j-1)$-set within it and by Corollary 4.8, this $(j-1)$-set is contained in $(1 \pm o(1))\varepsilon^3 n$ many j-sets of S. For each such pair of j-sets intersecting in $j-1$ vertices, there are $\binom{n-j-1}{k-j-1}$ k-sets containing both of them, all of which must be non-edges, since the j-sets lie in different components.

It may be that we multiple count the non-edges in this way. However, each k-set may only be counted from a pair of j-sets it contains, and therefore the number of times it is counted is certainly at most $\binom{k}{j}(k-j)$. Thus in total the number of non-edges is at least

$$\frac{r_0\varepsilon^3 n}{2\binom{k}{j}(k-j)}\binom{n}{k-j-1} = \Theta\left(r_0\varepsilon^3 n^{k-j}\right).$$

Similarly a in (4.7) we bound the expected number of such structures H by

$$\mathbb{E}(X_H) \leq n^{j+(k-j)e(H)}p^{e(H)}(1-p)^{\Theta(r_0\varepsilon^3 n^{k-j})}.$$

Then, writing Y for the number of such well-constructed hypergraphs of j-size r_0 which are not in the same component as S, we obtain

$$\log(\mathbb{E}(Y)) \leq k{r_0}^2\log 2 + j\log n + O\left(r_0\log\log n\right) - \Theta\left(r_0\varepsilon^3\log n\right).$$

Now observe that in Corollary 4.8 we may choose any $\varepsilon = o(1)$ such that $\varepsilon^3 n^{1-\delta} \to \infty$. Choosing $\varepsilon^3 = \frac{1}{\log\log\log n}$, we have $r_0\varepsilon^3 \to \infty$ and the last term in the above inequality dominates, and we have $\log(\mathbb{E}(Y)) \leq -C\log n$ for any constant C. In particular, choosing $C = 12j/7$, we have $\mathbb{E}(Y) \leq n^{-12j/7}$. By Markov's inequality, this implies that with probability at least $1 - n^{-12j/7}$ we have $Y = 0$ and therefore no further components of size r.

Combining the two cases, this tells us that with probability at least $1-3n^{-12j/7}$, $\mathcal{H}^k(n,p)$ only has one non-trivial component. Finally note that $M = pN = \Theta(n^j\log n)$. Thus by Lemma 4.5 we conclude that with probability at least $1 - 3n^{-12j/7}\sqrt{M} = 1 - o(n^{-8j/7})$, $\mathcal{H}^k(n,M)$ also has only one non-trivial component.

We now take a union bound over all possible M satisfying (4.6), of which there are at most $\frac{2\xi}{\binom{n}{k-j}}\binom{n}{k} = O(\xi n^j)$, and deduce that the probability that there is ever a second non-trivial within this time period is at most

$$O(\xi n^j)n^{-8j/7} = O(\xi n^{-j/7}) = o(1)$$

as required. □

4.5. Threshold for connectedness

Theorem 4.3 follows almost immediately from Theorems 4.1 and 4.6. In order to apply Theorem 4.1 in the binomial model, we apply the standard trick of birth times: to each k-tuple we assign a number (the *birth time*) between 0 and 1 uniformly at random and independently of all other k-tuples. Then the hypergraph process $\{\mathcal{H}^k(n,M)\}_M$ can be obtained by adding edges in increasing order

of birth time (with probability 1 no two edges have the same birth time), while the hypergraph obtained by taking all edges with birth time at most p is distributed as $\mathcal{H}^k(n,p)$.

Proof of Theorem 4.3. Theorem 4.6 (with $s = 0$) tells us that if $c_n \to \infty$, then whp there are no isolated j-sets, and therefore Theorem 4.1 tells us that whp the hypergraph is j-connected. This proves part (a). Part (b) is simply an application of Theorem 4.6 with $s = 0$. □

4.6. Concluding remarks

In [96], it is determined for the case $j = 1$ that the hitting time for *d-strong* vertex-connectedness, i.e. the time at which the hypergraph first has the property that deleting any set of less than d vertices still leaves a vertex-connected hypergraph, is the same as the hitting time for having no vertices of degree less than d whp. It would be interesting to generalise this result to d-strong j-connectedness (removing fewer than d many j-sets still leaves a j-connected hypergraph), which is presumably attained whp when every j-set has degree at least d. However, this would present significant additional difficulties, not least that Lemma 4.7 would no longer give the substructure which we require.

CHAPTER 5

Jigsaw percolation on random hypergraphs

5.1. Introduction and main results

Jigsaw percolation on graphs was introduced by Brummitt, Chatterjee, Dey, and Sivakoff [43] as a model for interactions within a social network. It was inspired by the idea of collectively solving a puzzle. The premise is that each of n people has a piece of a puzzle which must be combined in a certain way to solve the puzzle.

In the model there are two possibly overlapping sets of edges coloured red and blue defined on a common set of vertices. (In particular, any pair of vertices may form both a red and a blue edge at the same time.) Jigsaw percolation is a deterministic process on *clusters* of vertices that evolves in discrete time. Initially, each vertex forms its own cluster and in each subsequent time-step two clusters merge if they are joined by at least one edge of each colour. The process stops once no two clusters can be merged. Our jigsaw process *percolates* if we end in a single cluster. In particular, the process cannot percolate if either of the graphs given by blue or red edges is not connected.

More generally, given integers $1 \leq r \leq s$, define (r, s)-*jigsaw percolation* as follows. Let $G_1, \ldots, G_s$ be graphs on the same vertex set V. At each discrete time $t = 0, 1, \ldots$, we have a partition of V into *clusters*. At time $t = 0$ this is the finest partition: every vertex forms its own cluster. At time t, let $\widetilde{G}_t$ be the graph whose vertices are the clusters, with two vertices joined by an edge if the corresponding clusters are joined by an edge in at least r of the graphs G_i. The clusters of our jigsaw process at time $t + 1$ are the unions of the clusters that belong to the same component of $\widetilde{G}_t$. The process *percolates* if eventually we arrive at a single cluster. Note that a $(1, s)$-process percolates if and only if the union of the graphs G_i is connected. On the other hand, if we have (s, s)-percolation then each G_i must be connected. However, the connectedness of each G_i is far from sufficient for (s, s)-percolation. So far, only $(2, 2)$-jigsaw percolation has been considered: in most of this chapter, we shall do the same (and drop $(2, 2)$ from the notation), but we shall consider larger s in Section 5.4.

Returning to the motivation, the blue graph may represent a *puzzle graph* of how the pieces of the puzzle may be combined to reach a solution, while the red graph may represent the *people graph*, modelling friendships between the people who hold the puzzle pieces.

Brummitt, Chatterjee, Dey, and Sivakoff [43] studied the model when the red graph is the binomial random graph and with various deterministic possibilities for the blue graph, including a Hamilton cycle, or other connected graphs of bounded maximum degree, and provided upper and lower bounds for the percolation threshold probabilities.

Gravner and Sivakoff [65] improved on these results for many different puzzle graphs of bounded degree, and also introduced a generalised process with redundancy parameters in the number of neighbours required for clusters to merge.

The setting in which both graphs are binomial random graphs was studied by Bollobás, Riordan, Slivken, and Smith [37], who determined the asymptotic order of the threshold for percolation in terms of the product of the two associated probabilities.

Before we state their result, we set the scene. Let $G(n,p_1)$ and $G(n,p_2)$ denote the pair of random graphs on the (common) vertex set $[n]$ (for $a \in \mathbb{N}$, we define $[a] := \{1,\dots,a\}$), where each edge is present independently with probability p_1 or p_2 respectively. Throughout the chapter any *unspecified* asymptotic is with respect to $n \to \infty$ and in particular we use the phrase *with high probability*, abbreviated to *whp*, to mean with probability tending to 1 as $n \to \infty$. With this notation the main result in [37] is the following.

Theorem 5.1 ([37])**.** *Let $0 \le p_1, p_2 \le 1$ and let $G_1 = G(n,p_1)$ and $G_2 = G(n,p_2)$. Then there exists a constant $c > 0$ such that*

(a) if $p_1p_2 \ge \frac{c}{n\log n}$ and $\min\{p_1,p_2\} \ge \frac{c\log n}{n}$, then whp (G_1,G_2) percolates;
(b) if $p_1p_2 \le \frac{1}{cn\log n}$, then whp (G_1,G_2) does not percolate.

Our main aim in this chapter is to extend the percolation process above and this result to a hypergraph setting.

5.1.1. **Setup.** We denote by $\binom{V}{i}$ the set of i-element subsets of a set V and call its elements *i-sets*.

Hypergraphs and high-order connectedness. Given an integer $k \ge 2$, a *k-uniform hypergraph* (or *k-graph*) H consists of a set $V = V(H)$ of *vertices* and a set $E = E(H)$ of *edges*, where $E \subset \binom{V}{k}$. (Thus for $k = 2$ this defines a graph.) There are several natural possibilities for the concept of connectedness of a hypergraph: here we define some of them. Given an integer $1 \le j < k$, we say that two distinct j-sets of vertices (i.e. j-element subsets of the vertex set V) $J_1 \ne J_2$ are *j-connected* if there is a sequence $e_1,\dots,e_m \in E$ of edges such that

- $J_1 \subset e_1$ and $J_2 \subset e_m$;
- $|e_i \cap e_{i+1}| \ge j$ for all $1 \le i \le m-1$.

In other words, we may walk from J_1 to J_2 using edges which consecutively intersect in at least j vertices. Any j-set is always j-connected to itself. This forms an equivalence relation on the set $\binom{V}{j}$ of j-sets. A *j-component* is an equivalence class of this relation, i.e. a maximal set of pairwise j-connected j-sets. The case $j = 1$

is also known as *vertex-connectedness*, and for $j \geq 2$ we use the term *high-order connectedness*.[1]

High-order jigsaw percolation. A *double k-graph* (V, E_1, E_2) consists of a set $E_1 \subset \binom{V}{k}$ of *red* edges and a set $E_2 \subset \binom{V}{k}$ of *blue* edges on a common set V of vertices. We call the k-graphs (V, E_1) and (V, E_2) the *red k-graph* and the *blue k-graph* respectively.

Let (V, E_1, E_2) be a double k-graph, $\mathcal{J} \subset \binom{V}{j}$, and let $\mathcal{C}_0$ be a partition of $\mathcal{J}$ into $\kappa_0 := |\mathcal{C}_0|$ partition classes. (Unless stated otherwise, $\mathcal{C}_0$ will be the partition in which every j-set forms its own partition class and so $\kappa_0 = |\mathcal{J}|$.) Then *j-jigsaw percolation* is a deterministic, discrete time process which is characterised by a sequence of partitions $\{\mathcal{C}_t\}_{t=0,1,\dots}$ of $\mathcal{J}$. We write $\kappa_t := |\mathcal{C}_t|$ for the number of partition classes at time t and denote the partition classes by $\mathcal{C}_{t,1}, \dots, \mathcal{C}_{t,\kappa_t}$. The partition $\mathcal{C}_{t+1}$ is obtained from $\mathcal{C}_t$ as follows:

(I) We define an auxiliary graph $\widetilde{G}_t$ on the vertex set $[\kappa_t]$, where the edge $\{i, i'\}$ is present if and only if there are j-sets $J_{i,1}, J_{i,2} \in \mathcal{C}_{t,i}$ and $J_{i',1}, J_{i',2} \in \mathcal{C}_{t,i'}$ and edges $e_1 \in E_1$ and $e_2 \in E_2$ such that $J_{i,1} \cup J_{i',1} \subset e_1$ and $J_{i,2} \cup J_{i',2} \subset e_2$.

(II) If $\widetilde{G}_t$ is an empty graph, then STOP.

(III) Otherwise we set κ_{t+1} to be the number of components of $\widetilde{G}_t$ and denote these components by $C_1, \dots, C_{\kappa_{t+1}}$. The partition $\mathcal{C}_{t+1}$ is then given by $\mathcal{C}_{t+1,i} = \bigcup_{w \in C_i} \mathcal{C}_{t,w}$, for $1 \leq i \leq \kappa_{t+1}$.

(IV) If $\kappa_{t+1} = 1$, then STOP.

(V) Otherwise we proceed to time $t + 1$.

If the process stopped in step (IV), then $\mathcal{J}$ *percolates on* (V, E_1, E_2); otherwise, i.e. the process stopped in step (II), we say that $\mathcal{J}$ *does not percolate on* (V, E_1, E_2). We refer to the partition classes $\mathcal{C}_{t,i}$ as *clusters*.

The main results in this chapter deal with the setting when $\mathcal{J} = \binom{V}{j}$. In this case we simply say (V, E_1, E_2) *j-percolates* or (V, E_1, E_2) *does not j-percolate*. The generalised version is required in the proofs of our results.

Throughout the chapter we will ignore floors and ceilings whenever this does not significantly affect the argument.

5.1.2. **Main results.** In this chapter we consider the *random binomial double k-graph* $\mathcal{H}^k(n, p_1, p_2)$ on the vertex set $[n]$ where every k-set is present as a red edge with probability p_1 and present as a blue edge with probability p_2 independently of each other and of all other k-sets.

As we have already noted, a necessary condition for j-jigsaw percolation on a double k-graph is that both the red and the blue k-graphs are j-connected. In this setting the red/blue k-graph is a copy of the *binomial random k-graph* $\mathcal{H}^k(n, p_1)$ or $\mathcal{H}^k(n, p_2)$ respectively, i.e. the vertex set is $[n]$ and each k-set is present independently with probability p_1 or p_2 respectively. It was shown in [3] that the (sharp)

[1] This notion is not to be confused with the *(vertex-)connectivity* of a (hyper-)graph H measuring the size of the smallest *vertex-separator* in H.

threshold for j-connectedness in $\mathcal{H}^k(n,p)$ is

$$\hat{p}_c = \hat{p}_c(n,k,j) := \frac{j \log n}{\binom{n}{k-j}}. \tag{5.1}$$

As in [37], we impose a slightly stronger condition on $\min\{p_1,p_2\}$ to ensure the j-connectedness of both random hypergraphs in the supercritical regime. We prove the following extension of Theorem 5.1.

Theorem 5.2. *For integers $1 \le j < k$ let $0 \le p_1, p_2 \le 1$ and let $\mathcal{H} = \mathcal{H}^k(n,p_1,p_2)$. Then there exists a constant $c = c(k,j) > 0$ such that*

(a) if $p_1p_2 \ge \frac{c}{n^{2k-2j-1}\log n}$ and $\min\{p_1,p_2\} \ge \frac{c\log n}{n^{k-j}}$, then whp $\mathcal{H}$ j-percolates;
(b) if $p_1p_2 \le \frac{1}{cn^{2k-2j-1}\log n}$, then whp $\mathcal{H}$ does not j-percolate.

In other words, the threshold $\hat{p}_{jp} = \hat{p}_{jp}(n,k,j)$ for j-jigsaw percolation on $\mathcal{H}^k(n,p_1,p_2)$ in terms of the product $p = p_1p_2$ is of order

$$\hat{p}_{jp} = \Theta\left(\frac{1}{n^{2k-2j-1}\log n}\right).$$

Somewhat surprisingly, the result in the subcritical regime (Theorem 5.2(b)) does not seem to be easy to prove. The corresponding subcritical case for graphs (i.e. $k=2$ and $j=1$) was almost trivial, but the general case requires significantly more involved analysis.

On the other hand, the supercritical case (Theorem 5.2(a)) becomes much easier since we prove a neat reduction to the graph case.

The methods that we apply in this chapter can also be used, with minimal additional work, to prove some generalisations of and related results to Theorem 5.2. We give these results, and outlines of the proofs, in Section 5.4.

5.2. Subcritical regime

Let $1 \le j < k$ be integers. As a first step we develop some necessary conditions for any double k-graph H with vertex set V to j-percolate. Then, in the second step, we show that if H is distributed as $\mathcal{H}^k(n,p_1,p_2)$ then whp it fails to satisfy the weakest of these necessary conditions.

5.2.1. **Necessary conditions for complete percolation.** The initial idea is to modify the algorithm (given by (I)–(V) in Section 5.1.1) slightly so that for each cluster it additionally keeps track of two sets of edges, one red and one blue, which allow us to obtain this cluster by a sequence of merges.

In this spirit we consider triples of the form $(\mathcal{J}_0,\mathcal{E}_1,\mathcal{E}_2)$ where $\mathcal{J}_0 \subset \binom{V}{j}$, $\mathcal{E}_1 \subset \binom{V}{k}$, and $\mathcal{E}_2 \subset \binom{V}{k}$, i.e. embedded into the complete double k-graph $\left(V,\binom{V}{k},\binom{V}{k}\right)$. We call the edges in $\mathcal{E}_1$ *red* and those in $\mathcal{E}_2$ *blue*; note that edges may be red and blue at the same time. The *size* of a triple is given by $|\mathcal{J}_0|$. Furthermore, we call $(\mathcal{J}_0,\mathcal{E}_1,\mathcal{E}_2)$ *internally spanned* if $\mathcal{J}_0$ percolates on the double k-graph $(V,\mathcal{E}_1,\mathcal{E}_2)$. Note that these triples play a crucial role for j-jigsaw percolation on double k-graphs (comparable to *internally spanned sets* in [37]).

Claim 5.3. *For every positive integer $N \leq \binom{|V|}{j}/\binom{k}{j}$, if a k-graph H j-percolates, then it must contain an internally spanned triple of size ℓ for some $N \leq \ell \leq \binom{k}{j}N$.*

Proof. We shall think of the edges of H arriving one at a time: whenever an edge arrives, we check whether it can be used to merge some clusters. Note that each edge e that arrives can be used to merge a cluster C with some others if e contains some j-set of C, so with the arrival of an edge at most $\binom{k}{j}$ clusters will be merged. Therefore, when the first cluster of size at least N forms, it was created by merging at most $\binom{k}{j}$ clusters each of size at most $N-1$, and therefore has size at most $\binom{k}{j}N$. □

Constructing all internally spanned triples of a given size does not seem easy, so instead of doing this, we consider a relaxation of internally spanned triples. To do so we first define a strengthened notion of j-connectedness on k-graphs. Let (V,E) be a k-graph and let $\mathcal{J} \subset \binom{V}{j}$ be a collection of j-sets. We say that a subset $\mathcal{J}^* \subset \mathcal{J}$ is *$\mathcal{J}$-traversable* (on (V,E)) if for every two distinct j-sets $J, J' \in \mathcal{J}^*$, $J' \neq J$, there is a sequence of edges $e_1, \dots, e_m \in E$ such that

- $J \subset e_1$ and $J' \subset e_m$;
- for all $1 \leq i \leq m-1$ some j-set $J_i \in \mathcal{J}$ is contained in $e_i \cap e_{i+1}$.

In other words, we may walk from J to J' using edges such that the intersection of two consecutive edges contains at least one j-set from $\mathcal{J}$. Furthermore any singleton $\{J\} \subset \mathcal{J}$ is $\mathcal{J}$-traversable. To shorten notation, a collection $\mathcal{J}$ is called *traversable* if it is $\mathcal{J}$-traversable. Note that $\binom{V}{j}$-traversable collections of j-sets are precisely j-connected collections. Now we say that a triple $(\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_1)$ is *traversable* if $\mathcal{J}_0$ is traversable in both the red and blue k-graphs, $(V, \mathcal{E}_1)$ and $(V, \mathcal{E}_2)$ respectively.

Fact 5.4. *Every internally spanned triple (of size $\ell \leq \binom{|V|}{j}$) is traversable and contains an edge-minimal traversable triple (with at most $\ell-1$ edges of each colour).*

Let us denote the set of all edge-minimal traversable triples (within the complete double k-graph) of size ℓ by $\mathcal{T}_\ell$. To give a bound on this set, we partition it according to the number of red and blue edges: for all integers $0 \leq r, b \leq \ell - 1$ we define

$$\mathcal{T}_{\ell,r,b} := \{(\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2) \in \mathcal{T}_\ell \mid |\mathcal{E}_1| = r \wedge |\mathcal{E}_2| = b\}.$$

Next, in order to extract further structural information from these triples, we fix an arbitrary pair $\sigma = (\sigma_j, \sigma_k)$ of total orders. Here σ_j is an order on $\binom{V}{j}$ and σ_k on $\binom{V}{k}$. Clearly, σ induces a total order on each of the sets $\mathcal{J}_0$, $\mathcal{E}_1$ and $\mathcal{E}_2$.

Given a triple $T = (\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2) \in \mathcal{T}_{\ell,r,b}$ we explore $\mathcal{J}_0$ in a breadth-first search using only red edges and colouring j-sets white once they have been discovered. More precisely, we study the following exploration process starting from the minimal element $J_{(1)}$ of $\mathcal{J}_0$ with respect to σ. Initially we colour $J_{(1)}$ white and make it active. Then in each step there is an active j-set, say J, and we consider all edges of $\mathcal{E}_1$ containing J which have not been considered previously. For each of these edges in turn (according to σ), we consider all the j-sets of $\mathcal{J}_0$ that it contains and

colour them white, if they are not yet coloured white, in the order given by σ. We have then finished exploring J and move on to the next j-set which was coloured white in this process. Since $\mathcal{J}_0$ is traversable in $(V, \mathcal{E}_1)$, this induces a new total order on $\mathcal{J}_0$ that we call the *BFS-order* of T (with respect to σ) and denote by $\tau_\sigma = \tau_\sigma(T) := (J_{(1)}, \dots, J_{(|\mathcal{J}_0|)})$. (Note that τ_σ will in general be different from the order induced by σ on $\mathcal{J}_0$.)

Additionally, we introduce *marks* on the j-sets of $\mathcal{J}_0$. As we shall see, marking a j-set in the blue process is similar to colouring it white in the red process. Initially $J_{(1)}$ is marked and all other j-sets are unmarked. We then go through $\mathcal{J}_0$ according to τ_σ. In the i-th step, i.e. when $J_{(i)}$ is active, we reveal all blue edges that contain $J_{(i)}$ and none of $\{J_{(i+1)}, \dots, J_{(|\mathcal{J}_0|)}\}$ one by one according to σ. Whenever we reveal a blue edge in this way we mark all the still unmarked j-sets in $\mathcal{J}_0$ that it contains.

The reason for colouring and marking the j-sets in $\mathcal{J}_0$ is that it allows us to additionally keep track of two sets of parameters, $r_{i,z}$ and $b_{i,z}$, for $1 \le i \le |\mathcal{J}_0|$ and $0 \le z \le \binom{k}{j}$. We say that a red (respectively blue) edge performs *z-duty* if we colour (respectively mark) precisely z j-sets when it is revealed. Then $r_{i,z}$ denotes the number of red z-duty edges that were revealed while $J_{(i)}$ was active. Similarly $b_{i,z}$ denotes the number of blue z-duty edges that were revealed when $J_{(i)}$ was coloured white. We store this information in matrices $R := (r_{i,z})_{i,z}$ and $B := (b_{i,z})_{i,z}$ and call the pair $\pi_\sigma(T) := (R, B)$ the *blueprint* of T with respect to the pair of orders σ.

For our upcoming arguments we introduce some notation for such matrices. Given a positive integer a and an $a \times \left(\binom{k}{j} + 1\right)$ matrix $M = (m_{i,z})_{i,z}$ with non-negative integer entries, we define

$$f(M) := \sum_{i=1}^{a} \sum_{z=0}^{\binom{k}{j}} z m_{i,z} \quad \text{and} \quad g(M) := \sum_{i=1}^{a} \sum_{z=0}^{\binom{k}{j}} m_{i,z}.$$

Furthermore, for a non-negative integer m, we define

$$\mathcal{M}_{a,m} := \left\{ M \in \mathbb{Z}_{\ge 0}^{a \times \left(\binom{k}{j}+1\right)} \;\middle|\; f(M) = a - 1 \wedge g(M) = m \right\}.$$

(We shall only ever use these definitions with $a = \ell$ and m either b or r.)

The intuition behind these definitions is the following: f allows us to reconstruct the total number of white (respectively marked) j-sets (apart from $J_{(1)}$) from the blueprint of a triple while g provides the number of red (respectively blue) edges. More formally this is stated in the following lemma.

Lemma 5.5. *Let $\ell > r, b \ge 0$ be integers and σ be a pair of total orders on $\binom{V}{j}$ and $\binom{V}{k}$. For any $T = (\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2) \in \mathcal{T}_{\ell,r,b}$, the blueprint $\pi_\sigma(T) = (R, B)$ satisfies*

- $f(R) = f(B) = |\mathcal{J}_0| - 1 = \ell - 1$;
- $g(R) = |\mathcal{E}_1| = r$;
- $g(B) = |\mathcal{E}_2| = b$.

In other words $\pi_\sigma(\mathcal{T}_{\ell,r,b}) \subset \mathcal{M}_{\ell,r} \times \mathcal{M}_{\ell,b}$.

Proof. Let $T = (\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2)$ and recall that $\mathcal{J}_0$ contains precisely $\ell - 1$ j-sets excluding $J_{(1)}$ (which is already initially both white and marked) and is traversable in both the red and blue k-graphs $(V, \mathcal{E}_1)$ and $(V, \mathcal{E}_2)$ respectively. Therefore, each j-set in $\mathcal{J}_0 \setminus \{J_{(1)}\}$ coloured white by at least one red edge and likewise receives a mark from at least one blue edge. On the other hand, no j-set is coloured/marked more than once. Since $f(R)$ counts the number of j-sets coloured white by red edges and $f(B)$ counts the number of marks given by blue edges, the first statement follows.

On the other hand, recall that T contains r red edges and b blue edges. Since T is edge-minimal, any edge, either red or blue, was revealed in the process and thus was counted in precisely one of the $r_{i,z}$ and one of the $b_{i,z}$. Thus the second and third statements follow. □

Next we give a (crude) upper bound on the number of such matrices.

Claim 5.6. *Let $a > m \geq 0$ be integers. If $\mathcal{M}_{a,m} \neq \emptyset$, then*

$$m \geq \binom{k}{j}^{-1}(a-1).$$

Furthermore, there is a constant $C' > 0$ (independent of a, m) such that

$$|\mathcal{M}_{a,m}| \leq (C')^{a-1}.$$

Note that for minimal traversable triples, certainly the number of j-sets is larger than the number of edges, since we start from one j-set and each edge should give rise to at least one new j-set, and so the condition $a > m$ is fulfilled in these cases.

Proof. For the first statement let M be any matrix in $\mathcal{M}_{a,m}$. Then

$$a - 1 = f(M) = \sum_{i=1}^{a}\sum_{z=0}^{\binom{k}{j}} z m_{i,z} \leq \binom{k}{j}\sum_{i=1}^{a}\sum_{z=0}^{\binom{k}{j}} m_{i,z} = \binom{k}{j} m.$$

For the second statement, first note that all entries of a matrix in $\mathcal{M}_{a,0}$ must be zero, hence $|\mathcal{M}_{a,0}| \leq 1$. Therefore let us assume $m > 0$. Choosing an arbitrary $a \times \left(\binom{k}{j} + 1\right)$ matrix M with non-negative integer entries satisfying $h(M) = m$ can be seen as having a set of m elements and partitioning it into $a\left(\binom{k}{j} + 1\right) \leq 2m\binom{k}{j}^2$ potentially empty partition classes. Thus, (using the fact that there are $\binom{t+s-1}{t-1} = \binom{t+s-1}{s}$ ways of partitioning s indistinguishable elements into t distinguishable classes) we certainly obtain the upper bound

$$|\mathcal{M}_{a,m}| \leq \binom{2\binom{k}{j}^2 m + m - 1}{m} \leq \left(\frac{3e\binom{k}{j}^2 m}{m}\right)^m \leq \left(9\binom{k}{j}^2\right)^{a-1}. \qquad \square$$

Recall that, given ℓ, r and b, we wanted to provide an upper bound on $|\mathcal{T}_{\ell,r,b}|$. We do this by constructing a superset of $\mathcal{T}_{\ell,r,b}$. To this end we first choose a blueprint $(R, B) \in \mathcal{M}_{\ell,r} \times \mathcal{M}_{\ell,b}$ and then construct all possible triples $(\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2)$ (within the complete double k-graph) such that its collection of white j-sets $\mathcal{J}_0$ is traversable in the red k-graph $(V, \mathcal{E}_1)$. We use the following procedure that

terminates after ℓ steps, or if at any point we do not have a valid choice; in the latter case, the current instance of the procedure is discarded. The input is a blueprint (R, B); for $0 \le i \le \ell$, we keep track of an integer ν_i, a sequentially ordered collection $\mathcal{J}_0(i) := (J_{(1)}, \dots, J_{(\nu_i)})$ of white j-sets (embedded into V) and two collections, $\mathcal{E}_1(i)$ and $\mathcal{E}_2(i)$, of red and blue edges respectively (embedded into V). Furthermore, we keep track of marks on the white j-sets in every step. For any instance that does not get discarded, the output is the triple $(\mathcal{J}_0(\ell), \mathcal{E}_1(\ell), \mathcal{E}_2(\ell))$.

Initially we colour an arbitrary j-set $J_{(1)}$ white and mark it. No edges are coloured yet. At time $1 \le i \le \ell$, we consider the i-th white j-set $J_{(i)}$ (if it does not exist, the instance is discarded immediately). For each $1 \le z \le \binom{k}{j}$ we colour $r_{i,z}$ distinct edges containing $J_{(i)}$ red one by one and for each of them we perform z-duty, i.e. colour precisely z of the non-white j-sets it contains white, one by one. Whenever a j-set $J_{(i)}$ is coloured white, for every $0 \le z \le \binom{k}{j}$ we colour $b_{i,z}$ distinct edges blue one by one such that each edge satisfies the following three conditions:

- it contains $J_{(i)}$ and some white j-set $J_{(i^*)}$ with $i^* < i$;
- it does not contain any (already) white j-set $J_{(i')}$ with $i' > i$;[2]
- it is a z-duty edge, i.e. it contains precisely z (still) unmarked j-sets from $(J_{(0)}, \dots, J_{(i)})$ (this may include $J_{(i)}$ or not); these are now marked.

Now let $\mathcal{Q}_{\ell,r,b}$ denote the set of all outputs of instances of this procedure that did not get discarded, i.e. every one of our choices in the procedure was valid.

Lemma 5.7. *For all integers $\ell > r, b \ge 0$ we have $\mathcal{T}_{\ell,r,b} \subset \mathcal{Q}_{\ell,r,b}$; furthermore, there is a constant $C > 0$ independent of ℓ, r, b such that*

$$|\mathcal{Q}_{\ell,r,b}| \le |V|^j C^{\ell-1} \left(|V|^{k-j}\right)^r \left(\ell |V|^{k-j-1}\right)^b.$$

Proof. The first assertion is simple. Fix a pair σ of total orders on $\binom{V}{j}$ and $\binom{V}{k}$. Now consider a triple $T \in \mathcal{T}_{\ell,r,b}$ and note that by Lemma 5.5 its blueprint satisfies $\pi_\sigma(T) = (R, B)$ for some $(R, B) \in \mathcal{M}_{\ell,r} \times \mathcal{M}_{\ell,b}$. Knowing T, σ, and therefore also $\tau_\sigma(T)$, it is straightforward to provide an instance of the above procedure with input (R, B) and output T that does not get discarded.

For the upper bound on the total number of outputs let us first fix any particular blueprint $(R, B) \in \mathcal{M}_{\ell,r} \times \mathcal{M}_{\ell,b}$. For the red edges, as far as an upper bound is concerned, in each step and for each z we may choose $r_{i,z}$ elements from a set of at most $|V|^{k-j}$ with replacement and then there are at most $\binom{k}{j}^z$ ways to choose the white j-sets within any red edge. (Of course these bounds are quite crude in general, but sufficient for our result.) Since we have at most $|V|^j$ choices for the initial white j-set we have at most

$$|V|^j \prod_{i,z} \left(\binom{k}{j}^z |V|^{k-j} \right)^{r_{i,z}} = |V|^j \binom{k}{j}^{f(R)} \left(|V|^{k-j}\right)^{g(R)} = |V|^j \binom{k}{j}^{\ell-1} \left(|V|^{k-j}\right)^r$$

[2]Note that we could insist that an edge contains no j-sets with $J_{(i')}$ at all with $i' > i$, as is indeed the case for traversable triples. However, substituting this weaker condition is valid since we seek an upper bound on the number of structures, and will be sufficient for our purposes.

instances which can be distinguished by their combined white and red structure.

Now consider the number of ways of choosing a blue z-duty edge together with the z j-sets to be marked. A blue edge contains at least $j+1$ already embedded vertices, for instance those in $J_{(i)} \cup J_{(i^*)}$, and there are at most ℓ choices for $J_{(i^*)}$. Thus we have at most $\ell|V|^{k-j-1}$ choices for this edge. Once it is fixed, there are at most $\binom{k}{j}^z$ choices for which j-sets should be marked. Therefore the number of choices for the blue edges and marks is at most

$$\prod_{i,z} \binom{k}{j}^{b_{i,z}} (\ell|V|^{k-j-1})^{b_{i,z}} = \binom{k}{j}^{f(B)} (\ell|V|^{k-j-1})^{g(B)} = \binom{k}{j}^{\ell-1} (\ell|V|^{k-j-1})^{b}.$$

Finally, we have already counted the number of matrices $R \in \mathcal{M}_{\ell,r}$ and $B \in \mathcal{M}_{\ell,b}$ in Claim 5.6: there is a constant $C' > 0$ such that we have $|\mathcal{M}_{\ell,r}| \le (C')^{\ell-1}$ and $|\mathcal{M}_{\ell,b}| \le (C')^{\ell-1}$. Combining this with the previous calculations provides the desired upper bound with $C = \left(C'\binom{k}{j}\right)^2$ □

5.2.2. **No percolation.** We now prove Theorem 5.2(b) stating that in the subcritical regime the binomial random double k-graph $\mathcal{H}^k(n, p_1, p_2)$ whp does not j-percolate, because it does not contain an edge-minimal traversable triple of size roughly $\log n$.

Proof of Theorem 5.2(b). Let $c > 0$ be a sufficiently large constant and suppose

$$p = p_1 p_2 = \frac{1}{cn^{2k-2j-1}\log n}.$$

As mentioned earlier, it is a necessary condition for j-percolation of $\mathcal{H}^k(n, p_1, p_2)$ that both its red and blue k-graphs are j-connected. Therefore, by (5.1), we may without loss of generality assume

$$\frac{j(k-j)!\log n}{2n^{k-j}} \le p_1, p_2 \le \frac{2}{j(k-j)!cn^{k-j-1}(\log n)^2},$$

where the upper bound follows directly from the lower bound and the assumption on the product p. We will prove that H distributed as $\mathcal{H}^k(n, p_1, p_2)$ does not j-percolate by showing that there is a bottleneck in the process.

By Claim 5.3 and Fact 5.4 we have the upper bound

$$\mathbb{P}(H \text{ } j\text{-percolates}) \le \sum_{\ell,r,b} \sum_{T \in \mathcal{T}_{\ell,r,b}} \mathbb{P}(T \subset H),$$

where (a priori) the first sum ranges over all integers ℓ, r and b satisfying $\log n \le \ell \le \binom{k}{j}\log n$ and $0 \le r, b \le \ell - 1$. Furthermore, Lemma 5.5 and Claim 5.6 imply that $\binom{k}{j}^{-1}(\ell-1) \le r, b$ whenever $\mathcal{T}_{\ell,r,b} \neq \emptyset$.

It will be convenient to split the summation into two parts, depending on whether the edge-minimal traversable triple $T = (\mathcal{J}_0, \mathcal{E}_1, \mathcal{E}_2)$ contains more red or blue edges. So let us first consider the term

$$q_1 := \sum_{r \le b < \ell} \sum_{T \in \mathcal{T}_{\ell,r,b}} \mathbb{P}(T \subset H).$$

Since any triple $T \in \mathcal{T}_{\ell,r,b}$ contains precisely r red edges and b blue edges, the probability that it is contained in H is $p_1^r p_2^b$ and in particular this probability depends only on the parameters r and b. Furthermore note that by Lemma 5.7 we have

$$|\mathcal{T}_{\ell,r,b}| \leq n^j C^{\ell-1} \left(n^{k-j}\right)^r \left(\ell n^{k-j-1}\right)^b$$

for some positive constant $C > 0$ and thus

$$\begin{aligned}\sum_{T \in \mathcal{T}_{\ell,r,b}} \mathbb{P}\left(T \subset H\right) &\leq n^j C^{\ell-1} \left(n^{k-j} p_1\right)^r \left(\ell n^{k-j-1} p_2\right)^b \\ &\leq n^j C^{\ell-1} \left(\ell n^{2k-2j-1} p\right)^r, \end{aligned} \tag{5.2}$$

since $r \leq b$ and $\ell n^{k-j-1} p_2 = o(1)$. Now recall that $\ell n^{2k-2j-1} p \leq \binom{k}{j}/c < 1$ for c sufficiently large, and thus the term is maximal when r is minimal, i.e. $r = \binom{k}{j}^{-1}(\ell - 1)$. This provides the upper bound

$$q_1 \leq \sum_{r \leq b < \ell} n^j \left(\frac{\binom{k}{j} C^{\binom{k}{j}}}{c}\right)^{\frac{\ell-1}{\binom{k}{j}}} \leq \left(\binom{k}{j} \log n\right)^3 n^j \left(\frac{\binom{k}{j} C^{\binom{k}{j}}}{c}\right)^{(\log n - 1)/\binom{k}{j}} = o(1)$$

for any sufficiently large constant c dependent on k and j.

To complete the proof consider

$$q_2 := \sum_{b \leq r < \ell} \sum_{T \in \mathcal{T}_{\ell,r,b}} \mathbb{P}\left(T \subset H\right).$$

Swapping the roles of red and blue we can use precisely the same argument to show that $q_2 = o(1)$ (using $p_1' := p_2$, $p_2' := p_1$, $\ell' := \ell$, $r' := b$ and $b' := r$). Thus

$$\mathbb{P}\left(H \text{ } j\text{-percolates}\right) \leq q_1 + q_2 = o(1),$$

in other words the binomial random double k-graph $\mathcal{H}^k(n, p_1, p_2)$ whp does not j-percolate. □

5.3. Supercritical regime

As we shall see, the supercritical regime is easier to prove, since we simply reduce to the graph case $k = 2$ and $j = 1$. In the first step we provide the reduction for any $k \geq 3$ but $j = 1$. We then show how to obtain the result for arbitrary pairs (k, j) from the setting with $(k-1, j-1)$ as long as $j \geq 2$.

Proof of Theorem 5.2(a). Given integers $1 \leq j < k$, we define a statement $S(k, j)$ as follows.

> $S(k, j)$: There exists a constant $c = c(k, j) > 0$ such that for any functions $p_1 = p_1(n)$ and $p_2 = p_2(n)$ satisfying $\frac{c \log n}{n^{k-j}} \leq p_1, p_2 \leq 1$ and $p_1 p_2 \geq \frac{c}{n^{2k-2j-1} \log n}$ for sufficiently large n, the double k-graph $\mathcal{H}^k(n, p_1, p_2)$ j-percolates whp.

Let us observe that showing that $S(k, j)$ holds for all pairs of integers $1 \leq j < k$ proves Theorem 5.2 (ii). We proceed inductively, with the base case being the result on graphs, proved in Theorem 5.1 (i).

Remark 5.8. $S(2,1)$ holds.

We split the induction step into two parts.

Claim 5.9. *Let $k \geq 3$, then $S(2,1)$ implies $S(k,1)$.*

Proof. Let $k \geq 3$ be an integer. Assume $S(2,1)$ holds and let $c' = c(2,1)$. Set $c = c(k,1) := 3(2k)^{2k}c'$. By monotonicity, in order to show $S(k,1)$ it suffices to prove that H distributed as $\mathcal{H}^k(n,p_1,p_2)$ 1-percolates for probabilities $p_1 = p_1(n)$ and $p_2 = p_2(n)$ satisfying

$$p = p_1p_2 = \frac{c}{n^{2k-3}\log n}$$

and

$$\frac{c\log n}{n^{k-1}} \leq p_1, p_2 \leq \frac{1}{n^{k-2}(\log n)^2},$$

where the upper bound is an immediate consequence of the lower bound and the first condition.

We use the following construction to reduce to the graph case. Split the vertex set $[n]$ into two disjoint sets, say $Q := [n/2]$ and $Q^* := [n] \setminus [n/2]$. Let H' be the double graph on Q whose edges are any pair in Q contained in an edge of H whose remaining $k-2$ vertices are all in Q^*. Note that if H' 1-percolates then all vertices of Q must lie in a single cluster $\mathcal{C}$ of the final partition $\mathcal{C}_\infty$ of the 1-jigsaw percolation process in H. Notice that H' is distributed as $\mathcal{H}^2(n',p_1',p_2')$ with $n' := n/2$ and the edge probabilities $p_1 = p_1(n')$ and $p_2 = p_2(n')$ satisfy

$$p_i' := 1-(1-p_i)^{\binom{n/2}{k-2}} \geq 1-\exp\left(-p_i(n/(2k-4))^{k-2}\right) \geq \frac{p_in^{k-2}}{(2k)^k} \tag{5.3}$$

for $i \in \{1,2\}$ and sufficiently large n, since $p_i = o(n^{-(k-2)})$. In particular we have

$$p_1'p_2' \geq \frac{p_1p_2n^{2k-4}}{(2k)^{2k}} \geq \frac{c'}{n'\log n'}$$

and also

$$p_i' \geq \frac{p_in^{k-2}}{(2k)^k} \geq \frac{c\log n}{(2k)^k n} \geq \frac{c'\log n'}{n'},$$

for $i \in \{1,2\}$. Thus, by the choice of c', H' does indeed 1-percolate whp.

Similarly, reversing the roles of Q and Q^*, i.e. considering only edges with precisely 2 vertices from Q^*, we deduce that whp there is a cluster $\mathcal{C}^* \in \mathcal{C}_\infty$ such that $Q^* \subset \mathcal{C}^*$. Even though these two events may not be independent (at least for $k=4$) applying a union bound guarantees that they happen simultaneously whp.

Furthermore, there are at least two edges, one red and one blue, present in H'. The edges of H which gave rise to these (which may not be uniquely determined) each contain at least one vertex from Q and Q^* (since $k \geq 3$). Thus the clusters $\mathcal{C}$ and $\mathcal{C}^*$ must coincide and contain all vertices, i.e. $\mathcal{C} = \mathcal{C}^* = [n]$. In other words, H 1-percolates whp. □

Claim 5.10. *Let $k > j \geq 2$, then $S(k-1,j-1)$ implies $S(k,j)$.*

Proof. Fix integers $k > j \geq 2$. Assume $S(k-1, j-1)$ holds and let $c' = c(k-1, j-1)$. Set $c := \max\{5c', (2k)^k\}$. Again, by monotonicity, in order to show $S(k, j)$ it suffices to prove that H distributed as $\mathcal{H}^k(n, p_1, p_2)$ j-percolates for probabilities $p_1 = p_1(n)$ and $p_2 = p_2(n)$ satisfying

$$p = p_1 p_2 = \frac{c}{n^{2k-2j-1} \log n} \qquad \text{and} \qquad p_1, p_2 \geq \frac{c \log n}{n^{k-j}}.$$

We expose the edges of H in two rounds, i.e. for $i = 1, 2$ we set $p_i' := 1 - \sqrt{1-p_i} \geq p_i/2$. Note that $p_i = 2p_i' - (p_i')^2$ and thus we obtain H as the union of two independent copies H_α and H_β of $\mathcal{H}^k(n, p_1', p_2')$. (In particular we obtain the final partition of j-jigsaw percolation on H by running j-jigsaw percolation on H_α and using its final partition, denoted by $\mathcal{C}_\infty[H_\alpha]$, as the initial partition for j-jigsaw percolation on H_β.)

In the following we will consider certain *link double* $(k-1)$*-graphs* associated to H_α. They are constructed as follows. Given a vertex $v \in [n]$ we first delete from H_α all edges (k-sets) that do not contain the vertex v. Then we delete v from the vertex set and replace every remaining edge e with the $(k-1)$-set $e \setminus v$. We denote by $H_{\alpha,v}$ the resulting random double $(k-1)$-graph on the vertex set $[n] \setminus \{v\}$, and call it the *link double* $(k-1)$*-graph of* v.

Now note that there is a natural bijection mapping the set of $(j-1)$-sets (respectively $(k-1)$-sets) in $H_{\alpha,v}$ to the set of j-sets (respectively k-sets) containing v in H_α. Therefore any cluster in the final partition of $(j-1)$-jigsaw percolation on $H_{\alpha,v}$ corresponds to a collection of j-sets (once we have added v to each) which must be contained in a cluster of $\mathcal{C}_\infty[H_\alpha]$. Therefore, whenever $H_{\alpha,v}$ $(j-1)$-percolates, there is a unique cluster in $\mathcal{C}_\infty[H_\alpha]$ which contains all j-sets which include v, and thus we call it *the corresponding cluster to* v. We call a vertex v *good* if $H_{\alpha,v}$ $(j-1)$-percolates; vertices that are not good are called *exceptional*. This notion is motivated by the following observation. The corresponding clusters of any two good vertices u and u' overlap in all j-sets containing both u and u', thus indeed they must coincide (since $j \geq 2$). In other words, the final partition $\mathcal{C}_\infty[H_\alpha]$ contains a cluster $\mathcal{C}$ which includes every j-set with at least one good vertex.

Hence it remains to study j-sets containing only exceptional vertices. For this we observe that $H_{\alpha,v}$ is distributed as $\mathcal{H}^{k-1}(n', p_1', p_2')$, where $n' := n-1$ and the probabilities $p_1' = p_1'(n')$ and $p_2' = p_2'(n')$ satisfy

$$p_1' p_2' \geq \frac{c}{4n^{2k-2j-1} \log n} \geq \frac{c'}{(n')^{2(k-1)-2(j-1)-1} \log(n')},$$

$$p_i' \geq \frac{c \log n}{2n^{k-j}} \geq \frac{c' \log(n')}{(n')^{(k-1)-(j-1)}}$$

for $i \in \{1, 2\}$ and sufficiently large n. Consequently, by the choice of c', $H_{\alpha,v}$ $(j-1)$-percolates whp and therefore the expected number of exceptional vertices is $o(n)$, say n/ξ, for some function $\xi \to \infty$. Thus whp there are at most $n/\sqrt{\xi}$ exceptional vertices, by Markov's inequality.

Now we run j-jigsaw percolation on H_β with initial partition $\mathcal{C}_0[H_\beta] := \mathcal{C}_\infty[H_\alpha]$ and show that whp $\mathcal{C}_1[H_\beta] = \mathcal{C}_\infty[H_\beta] = [n]$, i.e. there is percolation in a single step. For this it is sufficient to show that whp for every j-set $J = \{u_{(1)}, \dots, u_{(j)}\}$ containing *only* exceptional vertices there are edges e_1 (red) and e_2 (blue) in H_β each containing J and a good vertex v_1 and v_2 respectively, since $J' := \{u_{(1)}, \dots, u_{(j-1)}, v_i\}$ satisfies $J' \in \mathcal{C}_\infty[H_\alpha]$ and $J' \subset e_i$. For any $i \in \{1,2\}$ and any such j-set J, the probability that no edge e_i exists is at most

$$\begin{aligned}(1-p_i')^{(1-1/\sqrt{\xi})n\binom{n-j-1}{k-j-1}} &\le \exp\left(-p_i' n^{k-j}/(2k)^{k-j}\right)\\ &\le \exp\left(-(2k)^{j-k} c \log n\right) \le n^{-(2k)^j} = o\left(n^{-j}\right),\end{aligned}$$

for sufficiently large n and by the choice of c. Hence we may take the union bound over $i \in \{1,2\}$ and all (such) j-sets. Therefore $\mathcal{H}^k(n, p_1, p_2)$ j-percolates whp, completing the proof. □

Proposition 5.8 and Claims 5.9 and 5.10 imply that $S(k,j)$ holds for all pairs of integers $1 \le j < k$ and this proves Theorem 5.2(a). □

5.4. Related models

With some small alterations these methods can also be applied for other models, for instance line graphs or in a setting with any fixed number of colours.

5.4.1. **Line graphs.** We consider the following random double graph $\mathcal{L}(n, p_1, p_2)$ that has a vertex for every pair of elements from the set $[n]$, i.e. $V = \binom{[n]}{2}$, and any two vertices that intersect form a red/blue edge with probabilities p_1 and p_2, respectively, independently of each other and of all pairs of vertices; disjoint vertices cannot form an edge. Note that (graph-)jigsaw percolation on this model is closely related to 2-jigsaw percolation on $\mathcal{H}^3(n, p_1, p_2)$. Following the lines of our proof for $j = 2$ and $k = 3$ we obtain the following result.

Theorem 5.11. *Let $0 \le p_1, p_2 \le 1$ and let $\mathcal{L} = \mathcal{L}(n, p_1, p_2)$. Then there is a constant $c > 0$ such that*

(a) if $p_1 p_2 \le \frac{1}{cn \log n}$, then whp $\mathcal{L}$ does not j-percolate;
(b) if $p_1 p_2 \ge \frac{c}{n \log n}$ and $\min\{p_1, p_2\} \ge \frac{c \log n}{n}$, then whp $\mathcal{L}$ j-percolates.

In other words, the threshold $\hat{p}_{\mathrm{jp},\mathcal{L}} = \hat{p}_{\mathrm{jp},\mathcal{L}}(n)$ for jigsaw percolation on the random double graph $\mathcal{L}(n, p_1, p_2)$ in terms of the product $p = p_1 p_2$ is of the order

$$\hat{p}_{\mathrm{jp},\mathcal{L}} = \Theta\left(\frac{1}{n \log n}\right).$$

In fact, the proof for the subcritical regime will be simplified since there cannot be any edges doing multiple duty (neither in red nor in blue). For the supercritical regime one reduction step is enough, since the link graph of a vertex in the line graph is a binomial random double graph and thus the result from [37] applies.

5.4.2. **Multiple colours.** Returning to the original motivation of jigsaw percolation, one might ask whether the social network is able to collectively solve multiple puzzles simultaneously. In this spirit we define a binomial random s-fold k-graph $\mathcal{H}^k(n, p_1, \dots, p_s)$ in which the vertex set is $[n]$ and any k-set forms an i-edge with probability p_i independently for all $1 \le i \le s$ and all other k-sets.

Now in the process of j-jigsaw percolation clusters merge if for each colour there is at least one edge connecting them. Hence, based on the same intuition, we analyse internally spanned $(s+1)$-tuples (the s-coloured analogue of internally spanned triples) and observe that the sequence of white j-sets is already determined by the set of edges of the first colour, say red. Now any further colour essentially behaves like blue and in particular independently of the other colours, given the sequence of white j-sets. With this intuition we obtain the following generalisation of Theorem 5.2.

Theorem 5.12. *For integers $1 \le j < k$ and $s \ge 2$ let $0 \le p_1, \dots, p_s \le 1$ and let $\mathcal{H}_s = \mathcal{H}^k(n, p_1, \dots, p_s)$. Then there is a constant $c = c(s,k,j) > 0$ such that*

(a) if $\prod_{i=1}^s p_i \le \frac{1}{cn^{s(k-j-1)+1}(\log n)^{s-1}}$, then whp $\mathcal{H}_s$ does not j-percolate;

(b) if $\prod_{i=1}^s p_i \ge \frac{c}{n^{s(k-j-1)+1}(\log n)^{s-1}}$ and $\min\{p_1, \dots, p_s\} \ge \frac{c \log n}{n^{k-j}}$, then whp $\mathcal{H}_s$ j-percolates.

In other words, the threshold $\hat{p}_{\mathrm{jp},s} = \hat{p}_{\mathrm{jp},s}(n,k,j)$ for j-jigsaw percolation on the s-fold k-graph $\mathcal{H}^k(n, p_1, \dots, p_s)$ in terms of $p = \prod_{i=1}^s p_i$ is of order

$$\hat{p}_{\mathrm{jp},s} = \Theta\left(\frac{1}{n^{s(k-j-1)+1}(\log n)^{s-1}}\right).$$

Proof outline. It turns out that there is a minor technical obstacle when determining the upper bound on the probabilities p_i for all $i \in [s]$ from some necessary conditions for j-percolation of an s-fold k-graph. As before the k-graph of (any) colour $i \in [s]$ has to be j-connected, i.e. by (5.1) we may assume

$$p_i = \Omega\left(\frac{\log n}{n^{k-j}}\right).$$

However, this alone will not yield useful upper bounds. Instead we observe that additionally, for any proper subset $I \subsetneq \{1, \dots, s\}$ of size at least two, it is necessary that the $|I|$-fold k-graph $\mathcal{H}^k(n, (p_i)_{i \in I})$ j-percolates. Therefore, Theorem 5.12 is proved by induction over s with Theorem 5.2 providing the base case. Hence assume $\prod_{j \in [s]: j \ne i} p_j = \Omega(\hat{p}_{\mathrm{jp},s-1})$ for all $i \in [s]$. Additionally, by monotonicity, we may also assume $\prod_{j \in [s]} p_j = \Theta(\hat{p}_{\mathrm{jp},s})$ and thus obtain the following upper bounds on the probabilities p_i:

$$p_i = \frac{\prod_{j \in [s]} p_j}{\prod_{j \in [s]: j \ne i} p_j} = O\left(\frac{\hat{p}_{\mathrm{jp},s}}{\hat{p}_{\mathrm{jp},s-1}}\right) = O\left(\frac{1}{n^{k-j-1} \log n}\right). \tag{5.4}$$

Even though this upper bound is slightly weaker than in the two-colour case (where we had another factor of $1/\log n$ at our disposal) it turns out to be sufficient for adapting our proof.

More precisely, the bound in (5.4) is used in the following two arguments:

- In the subcritical regime, in order to derive (5.2), and there we only need that $\ell n^{k-j-1}p_i \leq 1$ for all $\ell \leq \binom{k}{j}\log n$ and all colours $i \in [s]$. This holds if the constant $c(s,k,j)$ is chosen sufficiently large compared to $c(s-1,k,j)$.
- In the supercritical regime, for the last estimate in (5.2), but the argument doesn't change at all since having a single $(1/\log)$-factor is already enough here.

Apart from these occurrences the upper bound is only used to reprove the statement (in the supercritical regime) for graphs [37] in the setting of multiple colours. There the upper bound allowed us to use several asymptotic approximations, which may not be useful any more if we only have the weaker bound. However, these technical issues can be dealt with and the details can be found in [67].

5.5. Concluding remarks

Theorem 5.2 holds for a large enough constant c, which we made no attempt to optimise (nor was any such attempt made for graphs in [37]). It would be interesting to know the exact threshold, and in particular whether it is sharp, i.e. the upper and lower thresholds are asymptotically the same.

Note also that in the supercritical regime there was an extra condition on $\min\{p_1,p_2\}$ which contained a factor of c. However, there is no intrinsic reason why this c should be the same as the c in the product. Indeed, the reason for this condition is to ensure that each hypergraph is j-connected whp, but as mentioned in the introduction, the asymptotic threshold for this was determined in [3] to be $\frac{j(k-j)!\log n}{n^{k-j}}$. It is therefore natural to conjecture that this condition can be replaced by $\min\{p_1,p_2\} \geq \frac{c'\log n}{n^{k-j}}$ for any constant c' such that $j(k-j)! < c'$. In fact, even if, say $p_1 = \frac{j(k-j)!\log n}{n^{k-j}}$, there is a certain probability, bounded away from 0 and 1, that the red k-graph is j-connected. It is then natural to conjecture that, conditioned on it being j-connected (note that whp the blue k-graph will also be j-connected), whp the double k-graph $\mathcal{H}^k(n,p_1,p_2)$ j-percolates.

CHAPTER 6

Bootstrap percolation on geometric inhomogeneous random graphs

6.1. Introduction and main results

One of the most challenging and intriguing questions about large real-world networks is how activity spreads through the network. 'Activity' in this context can mean many things, including infections in a population network, opinions and rumours in social networks, viruses in computer networks, action potentials in neural networks, and many more. While all these networks seem very different, in the last two decades there was growing evidence that most of them share fundamental properties [14, 54]. The most famous property is that the networks are *scale-free*, i.e. the degrees follow a power-law distribution $\mathbb{P}(\deg(v) \geq d) \approx d^{1-\beta}$, typically for some $2 < \beta < 3$. Other properties include a large connected component which is a small world (poly-logarithmic diameter) and an ultra-small world (constant or poly-loglog average distance), that the networks have small separators and a large clustering coefficient. We refer the reader to [41] for more detailed discussions.

Classical models for random graphs fail to have these common properties. For example, Erdős-Rényi graphs or Watts-Strogatz graphs do not have power-law degrees, while Chung-Lu graphs and preferential attachment (PA) graphs fail to have large clustering coefficients or small separators. The latter properties typically arise in real-world networks from an underlying geometry, either spatial or more abstract, e.g., two nodes in a social networks might be considered 'close' if they share similar professions or hobbies. It is well-known that in real-world networks the spread of activity (of the flu, of viral marketing, ...) is crucially governed by the underlying spatial or abstract geometry [83]. For this reason, the explanatory power of classical models is limited in this context.

In recent years models were developed in order to overcome the previously mentioned limitations, most notably *hyperbolic random graphs* (HypRGs) [22, 28, 29, 93] and in more general[1] *geometric inhomogeneous random graphs* (GIRGs) [41],[2] and *spatial preferential attachment* (SPA) models [12, 52, 72]. Apart from the power-law exponent β, these models come with a parameter $\alpha > 1$, which models how strongly the edges are predicted by their distance. There are very few theoretical results on impact of the geometry on the spread of activity in such networks.

[1] It is non-obvious that GIRGs are a generalisation of HypRGs, see Theorem 6.3 in [41].

[2] Other than in [41] we do not condition on the number of vertices to be exactly n, which leads to slightly less technical proofs.

In this chapter we make a first step by analysing a specific process, *bootstrap percolation* [47], on the recent and very general GIRG model. In this process, an initial set of *infected* (or *active*) vertices (the bootstrap) iteratively recruits all vertices which have at least k infected neighbours, where $k \geq 2$ is a parameter. It was originally developed to model various physical phenomena (see [11] for a short review), but has by now also become an established model for the spreading of activity in networks, for example for the spreading of beliefs [55, 64, 97, 101], behaviour [62, 63], or viral marketing [78] in social networks (see also [46]), of contagion in economic networks [17], of failures in physical networks of infrastructure [108] or computer architecture [60, 80], of action potentials in neuronal networks (e.g. [16, 50, 56, 57, 92, 100, 105, 106], see also [81] for a review), and of infections in populations [55].

6.1.1. **Our contribution.** We investigate bootstrap percolation on GIRGs with an expected number of n vertices. We fix a ball B in the underlying geometric space, and we form the bootstrap by infecting each vertex in B independently with probability ρ. In this way, we model that an infection (a rumour, an opinion, ...) often starts in some local region, and from there spreads to larger parts of the network.

In Theorem 6.1 we determine a threshold $\hat{\rho}$ such that in the *supercritical case* $\rho = \omega(\hat{\rho})$ *with high probability*[3] (*whp* for short) a linear fraction of the graph is infected eventually, and in the *subcritical case* $\rho = o(\hat{\rho})$ infection ceases immediately. In the *critical case* $\rho = \Theta(\hat{\rho})$ both options occur with non-vanishing probability: if there are enough (at least k) 'local hubs' in the starting region, i.e. vertices of relatively large expected degree, then they become infected and facilitate the process. Without such local hubs the initial infection is not dense enough, and comes to a halt.

For the supercritical case, we show that it only takes $O(\log\log n)$ rounds until a constant fraction of all vertices is infected, and we determine the number of rounds until this happens up to a factor $1 \pm o(1)$ in Theorem 6.2. For the matching lower bound in this result, we need the technical condition $\alpha > \beta - 1$, i.e. edge-formation may not depend too weakly on the geometry. Notably, if the starting region B is sufficiently small then the number of rounds agrees (up to minor terms) with the average distance in the network. In particular, it does not depend on the infection rate ρ, as long as ρ is supercritical.

Finally we demonstrate that the way the infection spreads is strongly governed by the geometry of the process, again under the assumption $\alpha > \beta - 1$. Starting from B, the infection is carried most quickly by local hubs. Once the local hubs in a region are infected, they pass on their infection (a) to other hubs that are even further away, and (b) locally to nodes of increasingly lower degree, until a constant fraction of all vertices the region is infected.

[3]Meaning with probability tending to 1 as $n \to \infty$.

Indeed, given a vertex v (i.e. given its expected degree and its distance from B), and assuming that v is not too close to B, we can predict whp in which round it will become infected (Theorem 6.4), again up to a factor $1 \pm o(1)$.

In real applications such knowledge is invaluable: for example, assume that a policy-maker only knows initial time and place of the infection, i.e. she knows the region B and the current round i. In particular, she does not know ρ, she does not know the graph, and she has no detailed knowledge about who is infected. Then we show that she is able to identify a region B' in which the infection can be quarantined. In other words, by removing (from round i onwards) all edges crossing the boundary of B' whp the infection remains contained in B'.

The number of edges to be deleted is relatively small (Theorem 6.7): it can be much smaller than n (in fact, any function $f(n) = \omega(1)$ can be an upper bound, if i and $\mathrm{Vol}(B)$ are sufficiently small), and it is even much smaller than the number of edges *inside* of B', as was already noted in [41].

6.1.2. **Related work.** The GIRG model was introduced in [41], and we rely on many results from this paper. The average distance of a GIRG (which, as we show, agrees with the time until the bootstrap percolation process has infected a constant portion of all vertices) was determined in [40] in a much more general setup.

Bootstrap percolation has been intensively studied theoretically and experimentally on a multitude of networks, including trees [21, 47], lattices [13, 20], Erdős-Rényi graphs [75], various geometric graphs [39, 61, 88, 105], and scale-free networks [18, 24, 53, 78]. On geometric scale-free networks there are some experimental results [45], but little is known theoretically.

Recently, Candellero and Fountoulakis [44] determined the threshold for bootstrap percolation on HypRGs (in the threshold case $\alpha = \infty$, cf. below). However, they assumed that the initial infection takes place *globally*, i.e. whether any vertex is infected initially is independent of its position, and not *locally* as in our setting, where no vertex outside of a certain geometric region is infected initially. We emphasise that this idea does not apply for non-geometric random graph models. Using a localised initial infection has two major consequences.

Firstly, in the global setting, the (expected) number of initially infected vertices needs to be polynomial in n in order for the infection to start spreading significantly; while in our setting every ball containing an expected number of $\omega(1)$ vertices can initiate a large infection whp.

Secondly, using our knowledge about how the process evolves in time with respect to the geometry, we show that the infection time of any vertex is mainly governed by its geometric position and its weight. By contrast, using a global initial infection the infection times only depend on the expected degrees. Note that we do not encode these expected degrees as geometric information (in contrast to [44]), but rather in the weights.

While there is plenty of experimental literature and also some mean-field heuristics on other activity spreading processes on geometric scale-free networks (e.g. [66,

71, 98, 107, 109, 110]), rigorous mathematical treatments are non-existent with the notable exception of [76], where rumour spreading is analysed in an SPA model with a push and a push&pull protocol.

6.1.3. **Model and notation.** In this section we first define the random graph model that we will discuss in this chapter. Afterwards we formally introduce bootstrap percolation. In the last part of this section we then collect all notation used in the chapter for global reference.

6.1.3.1. *Geometric inhomogeneous random graphs.* A GIRG is a graph $G = (V, E)$ where both the vertex set V and the edge set E are random. Each vertex v is represented by a pair (x_v, w_v) consisting of a *position* x_v (in some *ground space*) and a *weight* $w_v \in \mathbb{R}_{>0}$.

Ground space and positions. We fix a (constant) dimension $d \geq 1$ and consider the d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ as the ground space. We usually think of it as the d-dimensional cube $[0,1]^d$ where opposite boundaries are identified and measure distances by the ∞-norm on $\mathbb{T}^d$, i.e. for $x, y \in [0,1]^d$ we define

$$\|x - y\| := \max_{1 \leq i \leq d} \min\{|x_i - y_i|, 1 - |x_i - y_i|\}.$$

The set of vertices and their positions are given by a homogeneous Poisson point process on $\mathbb{T}^d$ with intensity $n \in \mathbb{N}$. More formally, for any (Lebesgue-)measurable set $B \subset \mathbb{T}^d$, let $V \cap B$ denote (with slight abuse of notation) the set of vertices with positions in B. Then $|V \cap B|$ is Poisson distributed with mean $n\mathrm{Vol}(B)$, i.e. for any integer $m \geq 0$ we have

$$\mathbb{P}\left(|V \cap B| = m\right) = \mathbb{P}(\mathrm{Po}(n\mathrm{Vol}(B)) = m) = \frac{(n\mathrm{Vol}(B))^m \exp(-n\mathrm{Vol}(B))}{m!},$$

and if B and B' are disjoint measurable subsets of $\mathbb{T}^d$ then $|V \cap B|$ and $|V \cap B'|$ are independent. Note in particular that the total number of vertices $|V|$ is Poisson distributed with mean n, i.e. it is also random. An important property of this process is the following: given a random vertex[4] $v = (x_v, w_v)$, if we condition on $x_v \in B$, where B is some measurable subset of $[0,1]^d$, then the position x_v is uniformly distributed in B.

Weights. For each vertex, we draw independently a weight from some distribution $\mathcal{D}$ on $\mathbb{R}_{>0}$. We say that the weights follow a *weak power-law* for some exponent $\beta \in (2,3)$ if a $\mathcal{D}$-distributed random variable D satisfies the following two conditions: there is a constant $w_{\min} \in \mathbb{R}_{>0}$ such that $\mathbb{P}\left(D \geq w_{\min}\right) = 1$, and for every constant $\gamma > 0$ there are constants $0 < c_1 \leq c_2$ such that

$$c_1 w^{1-\beta-\gamma} \leq \mathbb{P}\left(D \geq w\right) \leq c_2 w^{1-\beta+\gamma} \tag{6.1}$$

for all $w \geq w_{\min}$. If this condition is also satisfied for $\gamma = 0$, then we say that the weights follow a *strong power-law*.

[4] By abuse of notation, x_v and w_v may either denote random variables or values.

Edges. Next we fix an $\alpha \in \mathbb{R}_{>1} \cup \{\infty\}$. Then two distinct vertices $u = (x_u, w_u)$ and $v = (x_v, w_v)$ form an edge independently of all other pairs with probability $p(x_u, x_v, w_u, w_v)$, where the function p satisfies

$$p(x_u, x_v, w_u, w_v) = \Theta(1) \min\left\{ \left(\frac{w_u w_v}{\|x_u - x_v\|^d n} \right)^{\alpha}, 1 \right\}, \tag{6.2}$$

if $\alpha < \infty$. In the *threshold model* $\alpha = \infty$ we instead require that p satisfies

$$p(x_u, x_v, w_u, w_v) = \begin{cases} \Omega(1) & \text{if } \|x_u - x_v\| \le C_1 \left(\frac{w_u w_v}{n}\right)^{1/d} \\ 0 & \text{if } \|x_u - x_v\| > C_2 \left(\frac{w_u w_v}{n}\right)^{1/d} \end{cases} \tag{6.3}$$

for some constants $0 < C_1 \le C_2$. Note that for $C_1 \neq C_2$ the edge probability may be arbitrary in the interval $\left(C_1 \left(\frac{w_u w_v}{n}\right)^{1/d}, C_2 \left(\frac{w_u w_v}{n}\right)^{1/d}\right)$.

6.1.3.2. *Bootstrap percolation.* Let $k \ge 2$ be a constant, let $B_0 \subset \mathbb{T}^d$ be measurable, and let $0 \le \rho = \rho(n) \le 1$. Then bootstrap percolation with *threshold* k, *starting region* B_0, and *initial infection rate* ρ is the following process. For every integer $i \ge 0$ there is a set $V^{\le i} \subset V$ of vertices which are *infected* (or *active*) at time i. The process starts with a random set $V^{\le 0} \subset V$ which contains each vertex in $V \cap B_0$ independently with probability ρ, and which contains no other vertices. Then we define iteratively

$$V^{\le i+1} := V^{\le i} \cup \left\{ v \in V \;\middle|\; v \text{ has at least } k \text{ neighbours in } V^{\le i} \right\}$$

for all $i \ge 0$. Moreover, we set $V^{\le \infty} := \bigcup_{i \in \mathbb{N}} V^{\le i}$.

We denote by $\nu = \nu(n) := n\mathrm{Vol}(B_0)$ the expected number of vertices in B_0. Throughout the chapter we will assume that B_0 is a ball centred at 0, without loss of generality due to symmetry of $\mathbb{T}^d$. Moreover, we will assume that $\nu = \omega(1)$.

6.1.3.3. *Further notation.* We often consider subsets of the vertex sets which satisfy some restrictions on their weights, positions, or whether they are infected at a given point of time. We use the following (slightly abusive) notation throughout the chapter: for a weight $w \in \mathbb{R}_{>0}$, a measurable set $B \subset \mathbb{T}^d$, and a time $i \ge 0$ we set

$$V^{\le i}_{\ge w} \cap B := \left\{ u = (x_u, w_u) \in V \;\middle|\; w_u \ge w, x_u \in B, u \in V^{\le i} \right\}.$$

All three types of restrictions are optional.

Moreover, we use the superscript "$(= i)$" to describe vertices which become infected (precisely) in round i, i.e. $V^{=i} := V^{\le i} \setminus V^{\le i-1}$ and $V^{=0} := V^{\le 0}$ etc. Furthermore, the index "$\ge w$" may be replaced by "$< w$" or "$\in [w, w')$", with the obvious meaning.

Additionally, we denote the neighbourhood of a vertex $v \in V$ by $N(v) := \{u \in V \mid \{u, v\} \in E\}$ and this notation may be modified by the same three types of restrictions, i.e.

$$N^{\le i}_{\ge w}(v) \cap B := N(v) \cap (V^{\le i}_{\ge w} \cap B).$$

For any two sets of vertices U_1 and U_2, we denote the set of edges between them by $E(U_1, U_2) := \{e = \{u_1, u_2\} \mid u_1 \in U_1, u_2 \in U_2\}$.

For a vertex $v \in V$, we define its *infection time* as $L_v := \inf\left\{i \geq 0 \;\middle|\; v \in V^{\leq i}\right\}$ and $L_v := \infty$ if the infimum does not exist.

For any $\lambda \geq 0$ and any closed ball $B \subset \mathbb{T}^d$ of radius $r \geq 0$ centred at 0 we denote by λB the closed ball of radius λr around 0.

Since it occurs frequently in our proofs, we abbreviate $\zeta := 1/(\beta-2) > 1$. For all $0 < \varepsilon < \zeta$ and all $i \geq 0$, we set

$$\nu_0 := \nu \qquad \text{and} \qquad \nu_i = \nu_i(\varepsilon) := \nu_0^{(\zeta-\varepsilon)^i},$$
$$\widetilde{\nu}_0 = \widetilde{\nu}_0(\varepsilon) := \nu^{(\beta-1)/(\beta-2)+\varepsilon} \qquad \text{and} \qquad \widetilde{\nu}_i = \widetilde{\nu}_i(\varepsilon) := \widetilde{\nu}_0^{(\zeta+\varepsilon)^i}$$

We define $B_i = B_i(\varepsilon)$ and $\widetilde{B}_i = \widetilde{B}_i(\varepsilon)$ to be the closed ball centred around 0 of volume $\min\{\nu_i(\varepsilon)/n, 1\}$ and $\min\{\widetilde{\nu}_i(\varepsilon)/n, 1\}$, respectively. Note that $B_i(\varepsilon) \subset \widetilde{B}_i(\varepsilon')$ for all $i \geq 0$ and all $0 < \varepsilon, \varepsilon' < \zeta$. The balls B_i and $\widetilde{B}_i$ will play a crucial role in describing how fast the infection spreads, cf. Theorem 6.20 and Theorem 6.17.

Given any constant $\eta > 0$, we define the functions $\hat{w}_-^{(\eta)}, \hat{w}_+^{(\eta)} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by

$$\hat{w}_-(\mu) = \hat{w}_-^{(\eta)}(\mu) := \mu^{1/(\beta-1+\eta)} \qquad \text{and} \qquad \hat{w}_+(\mu) = \hat{w}_+^{(\eta)}(\mu) := \mu^{1/(\beta-1-\eta)}.$$

As we will see in Lemma 6.15, in a region with an expected number of $\mu = \omega(1)$ vertices, whp the largest weight that occurs in this region is in $[\hat{w}_-(\mu), \hat{w}_+(\mu)]$.

In general we will be interested in results for large values of n, and in particular we use the phrase *with high probability* (whp) to mean with probability tending to 1 as $n \to \infty$. Moreover, all unspecified limits and asymptotics will be with respect to $n \to \infty$, and all constants hidden by Landau-notation are positive. For example, for a function $f = f(n)$ the notation $f = O(1)$ means that there is $n_0 > 0$ and an absolute constant $C > 0$ that depends only the constant parameters $\alpha, \beta, d, w_{\min}, \mathcal{D}$, and k of the model, such that $f(n) \leq C$ for all $n \geq n_0$. Similarly, $f = \omega(1)$ means $\lim_{n\to\infty} f(n) = \infty$ etc.

Similarly, when our proofs involve parameters $\varepsilon, \eta > 0$ then by the notation $O(\varepsilon), O(\eta)$ etc. we implicitly mean that the (positive) hidden constants only depend on the parameters $d, \alpha, \beta, w_{\min}, \mathcal{D}$, and k of the model, but not on ε or η. To enhance readability, in all proofs we stick to the convention that if ε and η occur together, then $\eta = \eta(\varepsilon) > 0$ is chosen so small that $C\eta < c\varepsilon$ for all concrete constants C and c in our proofs that depend only on the model parameters. In particular, the expression $\Omega(\varepsilon) - O(\eta)$ will be positive for all hidden constants that appear in our proofs.

In the proofs, for the sake of readability, we will not state each time when we use inequalities that only hold for sufficiently large n. For example, we will use inequalities like $\nu > 2$ without further comment although they are only true for sufficiently large n.

Throughout the chapter we will ignore all events of probability 0. For example, we will always assume that V is a finite set, and that all vertices in V have different positions. Furthermore, whenever it does not affect the argument, we omit floors and ceilings.

6.1.4. **Main results.** First of all we show that bootstrap percolation on a *GIRG* has at threshold with respect to the initial infection rate ρ. Since hyperbolic random graphs are a special instance of GIRGs, this contains in particular the result of [44] on (threshold) hyperbolic random graphs, where the case $\nu = n$ was studied.

Theorem 6.1. *Consider a bootstrap percolation process on a GIRG $G = (V, E)$ with initial infection rate $\rho = \rho(n) \in [0, 1]$. We set*

$$\hat{\rho} = \hat{\rho}(\nu, \beta) := \nu^{-\frac{1}{\beta-1}}.$$

If the weights follow a strong power-law, then the following holds.

(i) If $\rho = \omega(\hat{\rho})$, then $|V^{\leq\infty}| = \Theta(n)$ whp.
(ii) If $\rho = \Theta(\hat{\rho})$, then $|V^{\leq\infty}| = \Theta(n)$ with probability $\Omega(1)$, but also $V^{\leq\infty} = V^{\leq 0}$ with probability $\Omega(1)$.
(iii) If $\rho = o(\hat{\rho})$, then $V^{\leq\infty} = V^{\leq 0}$ whp.

If the weights follow a weak power-law, then the following holds.

(iv) If there is a constant $\delta > 0$ such that $\rho \geq \hat{\rho}^{1-\delta}$, then $|V^{\leq\infty}| = \Theta(n)$ whp.
(v) If there is a constant $\delta > 0$ such that $\rho \leq \hat{\rho}^{1+\delta}$, then $V^{\leq\infty} = V^{\leq 0}$ whp.

Whenever we refer to the *supercritical* regime we mean case (i) and (iv). Similarly, (iii) and (v) form the *subcritical* regime and (ii) is the *critical* regime. Note in particular that there is a supercritical regime regardless of how small the expected number ν of vertices in the starting region is, provided that $\nu = \omega(1)$. This is in sharp contrast to non-geometric graphs like Chung-Lu graphs, where the expected size of the bootstrap must be polynomial in n (if the bootstrap is chosen at random).

In fact, the proof of Theorem 6.1 will grant a deeper insight into the evolution of the process, in particular in the supercritical and critical regimes.

We consider the hitting time τ_o for the increasing property of a constant fraction of all vertices are infected, i.e. τ_o is a random variables taking values in $\mathbb{N} \cup \{\infty\}$, and show a doubly logarithmic upper bound on τ_o in the supercritical and critical regimes. Furthermore, we prove that this bound is tight up to minor order terms if the influence of the underlying geometry on the random graphs is sufficiently strong ($\alpha > \beta - 1$).

Theorem 6.2. *In the situation of Theorem 6.1, set*

$$i_\infty = i_\infty(n, \nu, \beta) := \frac{\log\log_\nu n + \log\log n}{|\log(\beta - 2)|},$$

then for any $\varepsilon > 0$ we have

$$\mathbb{P}(\tau_o \leq (1+\varepsilon) i_\infty) = \begin{cases} 1 - o(1) & \text{if } (i) \text{ or } (iv) \text{ holds,} \\ \Omega(1) & \text{if } (ii) \text{ holds;} \end{cases}$$

and furthermore, if $\alpha < \beta - 1$ and $\nu = n^{o(1)}$, then $\mathbb{P}(\tau_o \leq (1-\varepsilon) i_\infty) = o(1)$.

Remarkably, the bounds do not depend on the initial infection rate ρ, as long as ρ is supercritical/critical. Moreover, if the expected number ν of vertices in the starting region is sufficiently small (if $\log\log\nu = o(\log\log n)$), then the bound coincides with the average distance in the graph, again up to minor order terms.

In fact, we can still refine the statement of Theorem 6.2 tremendously, at least in the case $\alpha > \beta - 1$. In the following, we determine for every fixed vertex v its infection time L_v, up to minor order terms (with the restriction that v may not be too close to the starting region). We will show that it is given by the following expression (see also Remark 6.6 below).

Definition 6.3. *For any $x \in \mathbb{T}^d \setminus B_0$ and $w \in \mathbb{R}_{>0}$ we define*

$$\Lambda(x,w) := \begin{cases} \max\left\{0,\ \frac{\log\log_\nu\left(\|x\|^d n/w\right)}{|\log(\beta-2)|}\right\}, & \text{if } w > (\|x\|^d n)^{1/(\beta-1)}, \\ \frac{2\log\log_\nu(\|x\|^d n) - \log\log_\nu w}{|\log(\beta-2)|}, & \text{if } w \le (\|x\|^d n)^{1/(\beta-1)}. \end{cases} \tag{6.4}$$

In the first case we use the convention that the second term is $-\infty$ if $\|x\|^d n/w < 1$, and thus does not contribute to the maximum.

Note that in the second case, the sign of $\log\log_\nu w$ may be either positive or negative. However, then we have the lower bound $\Lambda(x,w) \ge \log\log_\nu(\|x\|^d n)/|\log(\beta-2)| + O(1)$ due to the upper bound of w and thus, in particular $\Lambda(x,w) \ge 0$, since $x \in \mathbb{T}^d \setminus B_0$.

Theorem 6.4. *Assume we are in the situation of Theorem 6.1 and (i) or (iv) holds. Let $v = (x_v, w_v)$ be any fixed vertex such that $x_v \in \mathbb{T}^d \setminus B_0$, $w_v = \omega(1)$ and $\Lambda(x_v, w_v) \le \log_2(\|x_v\|^d n/\widetilde{\nu}_0)$. Then whp*

$$L_v \le (1 + o(1))\Lambda(x_v, w_v) + O(1).$$

If additionally $\alpha > \beta - 1$ then we also have whp

$$L_v \ge (1 - o(1))\Lambda(x_v, w_v) - O(1).$$

As in Theorem 6.2, the bounds do not depend on the initial infection rate ρ, as long as it is supercritical.

Remark 6.5. The technical restrictions in Theorem 6.4 are necessary: if a vertex has weight $w_v = O(1)$ then the number of neighbours is Poisson distributed with mean $\Theta(w_v)$ (see Lemma 6.12), so v is even isolated with probability $\Omega(1)$. In particular, we cannot expected that whp v is ever infected.

The restriction $\Lambda(x_v, w_v) \le \log_2(\|x_v\|^d n/\widetilde{\nu}_0)$ ensures that v is not too close to the starting region. If v is too close, then it may have neighbours inside of B_0, and in this case it does depend on ρ when they are infected. (And of course, this process iterates.)

The term $\log_2(\|x_v\|^d n/\widetilde{\nu}_0)$ is not tight and could be improved at the cost of more technical proofs. However, there are already rather few vertices that violate

the condition $\Lambda(x_v, w_v) \le \log_2(\|x_v\|^d n/\widetilde{\nu}_0)$. For example, recall that it only takes $O(\log\log n)$ steps until a constant fraction of all vertices are infected. At this time, we only exclude vertices which satisfy $\|x_v\|^d n \le \widetilde{\nu}_0 \cdot (\log n)^{O(1)}$, so the expected number of affected vertices is also at most $\widetilde{\nu}_0 \cdot (\log n)^{O(1)}$. Even this is a gross overestimate, since the vertices close to the origin have much smaller infection times L_v, and thus only very few of them are affected by the condition.

Remark 6.6. The first case in Definition 6.3 is not needed if we restrict ourselves to vertices as they typically appear in GIRGs. More precisely, as we will see in Lemma 6.15, Section 6.2.2, whp all vertices in $v = (x_v, w_v) \in V \cap (\mathbb{T}^d \setminus B_0)$ satisfy $w_v \le (\|x_v\|^d n)^{1/(\beta-1-\eta)}$ where $\eta > 0$ is an arbitrary constant. In the border case $(\|x_v\|^d n)^{1/(\beta-1)} \le w_v \le (\|x_v\|^d n)^{1/(\beta-1-\eta)}$ both expressions in (6.4) agree up to additive constants, i.e.

$$\Lambda(x_v, w_v) = \frac{2\log\log_\nu(\|x_v\|^d n) - \log\log_\nu w_v}{|\log(\beta-2)|} \pm O(1). \tag{6.5}$$

Therefore, we could also use (6.5) as definition for Λ if we would exclude vertices which are unlikely to exist in Theorem 6.4 .

Finally, we provide a strategy how to contain the infection within a certain region when only the starting set and the current round a known, but not the set of infected vertices. Note that the number of edges that need to be removed is substantially smaller than the expected number of vertices $\widetilde{\nu}_i$ in the containment area $\widetilde{B}_i$.

Theorem 6.7. *Assume that we are in the situation of Theorem 6.1, and that $\alpha > \beta - 1$. If the starting region B_0 is known, then by removing all edges crossing the boundary of $\widetilde{B}_i$ before round $i+1$, whp the infection is contained in $\widetilde{B}_i$. The expected number of edges crossing the boundary of $\widetilde{B}_i$ is $\widetilde{\nu}_i^{\max\{3-\beta,1-1/d\}\pm o(1)}$.*

Before we proceed with an outline of the proofs, we briefly describe the organisation of this chapter. In Section 6.2.1 we collect some tools and derive basic properties of GIRGs. Afterwards, in Section 6.3, we investigate the evolution of bootstrap percolation on GIRGs in great detail, followed by the proof of Theorem 6.7. Using this description we then derive Theorem 6.4 in Section 6.4, and prove Theorems 6.1 and 6.2 simultaneously in Section 6.5. We conclude this chapter with a brief discussion of open problems in Section 6.6.

6.1.5. **Intuition and proof outline.** In this section we give an intuitive description of how the process evolves, and at the same time a very rough outline of the proofs. We warn the reader that some statements in this section are not literally true, but they are only true if appropriate error margins are taken into account.[5]

[5]This applies in particular to the definition of the balls B_i the quantities ν_i, and the weights that will appear in the course of the argument. The exact statements are rather technical and are given in Sections 6.3.2 and 6.3.1.

Initial phase. For the subcritical regime, we distinguish between high-weight vertices ($w_v = \omega(w_0)$, where $w_0 = \nu^{1/(\beta-1)}$) and low-weight vertices ($w_v = O(w_0)$). By an easy computation, the expected number of low-weight vertices in B_0 that are infected in round 1 is $o(1)$, so by Markov's inequality no low-weight vertex becomes infected whp. On the other hand, whp no high-weight vertex exists in B_0, and the expected number of infected vertices outside of B_0 is also $o(1)$ because they are too far away from infected vertices. In order words, whp no vertex is infected in the first round, and thus the percolation process stops immediately.

In the critical regime, the calculation is comparable, but if there exist vertices of weight $\Theta(w_0)$ then these vertices are infected with probability $\Omega(1)$. The number of vertices of weight $\Theta(w_0)$ is Poisson distributed with mean $\Theta(1)$, so it may happen (each with probability $\Omega(1)$) that either no such vertex exists (so percolation process stops) or that there are at least k such vertices, and that all of them are infected.

Similarly, in the supercritical regime, whp k vertices of weight (slightly less than) w_0 are infected. Then these k vertices whp infect all other vertices of similar weight in at most two more rounds.

Cascade of infection. If all vertices of weight $\Theta(w_0)$ are infected (in either the critical or supercritical regime), they start a cascade of infection which evolves in two 'main directions': it spreads *along the geometry*, i.e. to larger regions; and to vertices of subsequently smaller weights, which is necessary in order to become a *linear-sized outbreak*. For this recall that we defined $\zeta = 1/(\beta-2) > 1$, and consider the sequence B_i of nested balls of volume ν_i/n centred at 0, where $\nu_i \approx \nu^{\zeta^i}$.

Geometric spread. We inductively show the following: if in the i-th round all vertices of weight $w_i \approx \nu_i^{1/(\beta-1)}$ in B_i are infected, then in the next round, whp the vertices of weight w_i in B_i infect all vertices of weight w_{i+1} in B_{i+1}, thus spreading the infection to new regions. Note that in order to prove this, for vertices in B_{i+1} it suffices to count the number of neighbours of a certain weight in B_i, which is a Poisson distributed random variable. This gives a lower bound on how fast the infection spreads *geometrically*. It can not spread faster since whp there are no edges from B_i to $\mathbb{T}^d \setminus B_{i+1}$. This latter fact already allows us to execute a quarantine strategy (Theorem 6.7).

Linear-sized outbreak. The crucial observation is the following: if in round j every vertex of weight w in some region has a large probability to be infected, then in round $j+1$ every vertex of weight at least $w' \approx w^{1/\zeta}$ in this region has a large (though slightly smaller) probability to be infected.

To prove this formally, we consider a vertex of weight w'. Such a vertex (but not vertices of smaller weight) has at least w^δ neighbours of weight w, with probability at least $1 - \exp(-w^\delta)$. So we pick k such neighbours, and bound the probability that at least one of them is *not* infected by a union bound. In this way, we lose a factor of k in each round, but going through the proof details it turns out that this factor is still negligible compared to the error term $\exp(-w^\delta)$ (see Theorem 6.17).

This describes an 'infection pathway' providing a lower bound on the speed of the outbreak reaching linear size. It is the most challenging and technical part of the proof to complement this infection pathway by a matching upper bound.

Since in round $i-1$ there is no infected vertex in $B_i \setminus B_{i-1}$ it is not hard to argue that in round i only vertices of large weight in $\mathbb{T}^d \setminus B_{i-1}$ are infected. However, in subsequent rounds it does happen that vertices of very small weight in $\mathbb{T}^d \setminus B_{i-1}$ become infected. Fortunately, this only happens with rather small probability, which we explicitly bound (Theorem 6.20(f)) as a function of the weight.

Once we have such a bound in some round, we use that whp no vertex in $\mathbb{T}^d \setminus B_{i-1}$ (not too close to the boundary) has strictly more than one neighbour in B_{i-1}. Therefore, in order for a vertex v in $\mathbb{T}^d \setminus B_{i-1}$ to be infected, at least one of its neighbours in $\mathbb{T}^d \setminus B_{i-1}$ must have been infected in the previous round, and we can bound the probability of this event by the expected number of previously infected neighbours in $\mathbb{T}^d \setminus B_{i-1}$. It turns out that this simple bound is sufficient to provide the desired matching upper bound.

We remark that it is in this last step where we use the assumption $\alpha > \beta - 1$ in Theorems 6.2 and 6.4, since otherwise there do exist vertices in $\mathbb{T}^d \setminus B_{i-1}$ that have several neighbours in B_{i-1}, and these vertices exist in a substantial part of B_i. Even worse, for $\alpha < \beta - 1$, in some (large) subregion of B_i the number of infections in round $i+1$ that come from neighbours in B_{i-1} dominates the number of infections that come from neighbours in B_i. For investigating the case $\alpha < \beta - 1$ (which is not done in this thesis), it will no longer be possible to use a bound on the infection probability that is uniform within $\mathbb{T}^d \setminus B_{i-1}$, or within $B_i \setminus B_{i-1}$.

Together Theorems 6.17 and 6.20 allow us to simultaneously prove Theorems 6.1 and 6.2 without further preparation, and provide almost complete control over the process. As a consequence, for a each vertex v with fixed weight and position (outside of the starting region B_0), and for each round j we obtain lower and upper bounds for the probability that v is infected before round j. We can thus compute rounds j_1, j_2 for which the probability is at most $o(1)$ and at least $1-o(1)$, respectively, and we demonstrate that these rounds coincide up to lower order terms (Theorem 6.4). While it is still rather complicated to perform the calculations of j_1 and j_2 (for technical reasons), no further knowledge about the infection process is required.

6.2. Preliminaries

In Section 6.2.1 we collect basic tools and observations that we use throughout the chapter. Afterwards, we provide some basic properties of GIRGs in Section 6.2.2.

6.2.1. **Tools.** We start with a fact which often allows us to treat the case $\alpha = \infty$ along with the case of finite α without case distinction.

Observation 6.8. *For every function p satisfying (6.3) and every $\alpha \in \mathbb{R}_{>1}$, there is a function $\bar{p}$ satisfying (6.2) such that $\bar{p}(x_1, x_2, w_1, w_2) \geq p(x_1, x_2, w_1, w_2)$ for all $x_1, x_2 \in \mathbb{T}^d$ and all $w_1, w_2 \geq w_{min}$.*

In other words, GIRGs in the threshold case $\alpha = \infty$ are dominated by GIRGs with finite α. In particular, whenever we prove an upper bound on the number of active vertices that holds for all GIRGs with finite α, the same upper bound also holds for threshold GIRGs.

In [41], GIRGs were defined with a fixed number of vertices, while we assume the set of vertices to be given by a homogeneous Poisson point process. Our choice allows for slightly less technical proofs. In particular, one of the benefits of the Poisson point process is the following elementary fact.

Fact 6.9. *Let $\lambda \in \mathbb{R}_{\geq 0}$ and let X be a Poisson distributed random variable with mean λ. Furthermore, given some $0 \leq q \leq 1$, let Y be a random variable which conditioned on $\{X = x\}$, for any $x \in \mathbb{N}_0$, is the sum of x independent Bernoulli random variables with mean q. Then Y is Poisson distributed with mean $q\lambda$.*

This means that for instance $|N_{\geq w}(v) \cap B|$ is Poisson distributed with mean nq, where q denotes the probability that a vertex u with random position x_u and random weight w_u satisfies $w_u \geq w$ and $x_u \in B$, and is a neighbours of v. We will apply this observation throughout the chapter without giving explicit reference.

Many relevant quantities can be expressed by summing (some function) over all vertices whose weights lie in a given interval, the following lemma provides an easy way of evaluating these.

Lemma 6.10 (Lemma 4.1 in [40])**.** *Let $0 \leq w_0 < w_1$, and let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a piecewise continuously differentiable function. Then in any finite set V of weighted vertices, we have*

$$\sum_{v \in V, w_0 \leq w_v < w_1} f(w_v) \;=\; f(w_0)|V_{\geq w_0}| \;-\; f(w_1)|V_{\geq w_1}| \;+\; \int_{w_0}^{w_1} |V_{\geq w}| \frac{d}{dw} f(w) dw.$$

Note in particular that if $f(0) = 0$, then, by using $w_0 = 0$ and $w_1 > \max\{w_v \mid v \in V\}$, we have

$$\sum_{v \in V} f(w_v) = \int_0^{w_1} |V_{\geq w}| \frac{d}{dw} f(w) dw = \int_0^{\infty} |V_{\geq w}| \frac{d}{dw} f(w) dw.$$

The lemma is actually formulated slightly more general than Lemma 4.1 in [40]. First, in [40] the lemma was only formulated if V is the vertex set of a GIRG, but their proof does not use this restriction. Second, f was assumed to be everywhere continuously differentiable. However, our version follows immediately by applying Lemma 4.1 in [40] to every continuously differentiable piece, and by noting that all the intermediate terms cancel out.[6]

[6]Formally speaking, the term $\frac{d}{dw} f(w)$ is a distribution and not a function if f is not differentiable everywhere, and the integral is then to be understood as evaluation of the distribution. However, in our simple applications we may just ignore the values where f is not differentiable.

The next lemma spells out an almost trivial calculation, but since it is ubiquitous in our proofs, we state it as a lemma nevertheless. In our applications, $g(w)$ is typically the number of vertices of weight at least w (possibly with additional restrictions), and f is the probability that such a vertex has a certain property (e.g., that it forms an edge with some fixed v). After application of Lemma 6.10, this almost always leads to an integral as given in (6.6) below.

Lemma 6.11. *Let $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a non-zero monomial, and let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be continuous and piecewise a non-zero monomial with non-negative exponent, for a finite number of pieces. Moreover, assume that there is $\tilde{w}$ such that the exponent of w in $f(w)g(w)$ is strictly larger than 0 for $w < \tilde{w}$, and strictly smaller than 0 for $w > \tilde{w}$. Then for every $w_0 \leq \tilde{w} \leq w_1$,*

$$\int_{w_0}^{w_1} g(w)\frac{d}{dw}f(w)dw = O(f(\tilde{w})g(\tilde{w})). \tag{6.6}$$

Moreover, assume that the exponent of f is non-zero in an interval

(a) $[(1-\Omega(1))\tilde{w}, \tilde{w}] \subset [w_0, \tilde{w}]$, or
(b) $[\tilde{w}, (1+\Omega(1))\tilde{w}] \subset [\tilde{w}, w_1]$,

then the $O(\cdot)$ in (6.6) may be replaced by $\Theta(\cdot)$.

Proof. Let $g(w) = Cw^r$. Let us first assume that, by continuity, f satisfies $f(w) = C_0 w^{s_0}$ for $w \leq \tilde{w}$ and $f(w) = C_1 w^{s_1}$ for $w \geq \tilde{w}$, i.e. that f consists of only two pieces. Then by assumption $r + s_0 > 0 > r + s_1$. We first consider the lower part of the integral. If $s_0 = 0$ then $(df/dw)(w) = 0$ for $w \leq \tilde{w}$, and the integral from w_0 to $\tilde{w}$ vanishes. So assume that $s_0 > 0$. Then $(df/dw)(w) = C_0 s_0 w^{s_0-1}$, and the antiderivative of $g(df/dw)$ is $CC_0 s_0/(r+s_0)w^{r+s_0}$. Since $r + s_0 > 0$, this function is increasing in w, and

$$\int_{w_0}^{\tilde{w}} g(w)\frac{d}{dw}f(w)dw = \frac{CC_0 s_0}{r+s_0}\left(\tilde{w}^{r+s_0} - w_0^{r+s_0}\right) = \Theta(f(\tilde{w})g(\tilde{w}) - f(w_0)g(w_0)). \tag{6.7}$$

Note that if $w_0 \leq (1-\Omega(1))\tilde{w}$, then $f(\tilde{w})g(\tilde{w}) - f(w_0)g(w_0) = \Omega(f(\tilde{w})g(\tilde{w}))$ since fg is a polynomial with positive exponent $r+s_0$ in $[w_0, \tilde{w}]$, which proves the additional statement (a).

For the upper part of the integral, we may assume $s_1 > 0$, since otherwise this part of the integral vanishes. Then $(df/dw)(w) = C_1 s_1 w^{s_1-1}$, and the antiderivative of $g(df/dw)$ is $CC_1 s_1/(r+s_1)w^{r+s_1}$. Note crucially that the sign of this function is negative since $r + s_1 < 0$. Hence,

$$\int_{\tilde{w}}^{w_1} g(w)\frac{d}{dw}f(w)dw = \frac{CC_1 s_1}{-(r+s_1)}\left(\tilde{w}^{r+s_1} - w_1^{r+s_1}\right) = \Theta(f(\tilde{w})g(\tilde{w}) - f(w_1)g(w_1)). \tag{6.8}$$

Similarly to the first part, if $w_1 \geq (1+\Omega(1))\tilde{w}$, then $f(\tilde{w})g(\tilde{w}) - f(w_1)g(w_1) = \Omega(f(\tilde{w})g(\tilde{w}))$, which proves the additional statement (b). On the other hand, (6.6) follows immediately from (6.7) and (6.8) by leaving out the negative terms. This proves the lemma in the case that f consists of only two pieces.

For the case of several pieces, the additional statement follows by restricting the integral to the two pieces bounded by $\tilde{w}$. For the upper piece, assume that $w_0 = w_0^{(1)} < \ldots < w_0^{(\lambda)} = \tilde{w}$ are the endpoints of the different pieces below $\tilde{w}$. Then in the same way as (6.7), we get

$$\int_{w_0}^{\tilde{w}} g(w)\frac{d}{dw}f(w)dw = O(\sum_{i=1}^{\lambda} f(w_0^{(i)})g(w_0^{(i)})) \leq O(\lambda f(\tilde{w})g(\tilde{w})) = O(f(\tilde{w})g(\tilde{w})),$$

since fg is an increasing function in $[w_0, \tilde{w}]$. Analogously we prove (b). □

6.2.2. **Basic properties of GIRGs.** In this section we list briefly some basic properties of GIRGs. The first lemma, taken from [40], tells us that the expected degree of a vertex equals its weight, up to constant factors. Moreover, it gives the marginal probability that two vertices u, v of fixed weights but random positions in $\mathbb{T}^d$ are adjacent. This probability remains the same if the position of one (but not both) of the vertices is fixed. An expert reader may recognise that it is the same marginal probability as in Chung-Lu random graphs, cf. [40] for a discussion in depth.

Lemma 6.12 (Lemma 4.4 and Theorem 7.3 in [40])**.** *Let $v = (x_v, w_v)$ be a vertex with fixed weight and position. Then*

$$\mathbb{E}(\deg(v)) = \Theta(w_v). \tag{6.9}$$

Moreover, if $u = (x_u, w_u)$ is a vertex with fixed weight, but with random position $x_u \in \mathbb{T}^d$. Then

$$\mathbb{P}\left(\{u, v\} \in E \mid w_u, w_v, x_v\right) = \Theta\left(\min\left\{\frac{w_u w_v}{n}, 1\right\}\right). \tag{6.10}$$

Note in particular that the right hand side of (6.10) is independent of x_v, so the same formula still applies if also the position x_v of v is randomised.

The next lemma gives a bound on the expected number of neighbours of a fixed vertex of large weight.

Lemma 6.13. *Let $\eta > 0$ be a constant and consider a vertex $v = (x_v, w_v)$ with fixed weight and position. Then for every $w \geq w_{min}$ we have*

(a) $\mathbb{E}(|N_{\geq w}(v)|) = O(\min\{w_v w^{2-\beta+\eta}, nw^{1-\beta+\eta}\})$.
In particular, for a random vertex u we have, independently of x_v and w_v,

$$\mathbb{P}(w_u \geq w \mid \{u, v\} \in E) = O(w^{2-\beta+\eta});$$

(b) $\mathbb{E}(|N_{\geq w}(v)|) = \Omega(\min\{w_v w^{2-\beta-\eta}, nw^{1-\beta-\eta}\})$.

Proof. (a) By Lemma 6.12, the probability that a vertex u with fixed weight w_u and random position $x_u \in \mathbb{T}^d$ is adjacent to v is $\Theta(1)\min\{w_u w_v/n, 1\}$. The expected number of vertices of weight at least w is at most $O(nw^{1-\beta+\eta})$ by the power-law condition (6.1). We distinguish two cases. If $ww_v \geq n$ then the probability to connect to any vertex of weight w is $\Theta(1)$, so $\mathbb{E}(|N_{\geq w}(v)|) = \Theta(\mathbb{E}(|V_{\geq w}|))$, and the

claim follows. So assume $ww_v \le n$. Then by Lemma 6.10 we can compute the expectation as the following integral, which we then evaluate using Lemma 6.11.

$$\begin{aligned}\mathbb{E}(|N_{\ge w}(v)|) &\le O\left(nw^{1-\beta+\eta}\frac{ww_v}{n} + \int_w^\infty nw_u^{1-\beta+\eta}\frac{d}{dw_u}\min\left\{\frac{w_vw_u}{n},1\right\}dw_u\right)\\ &\le O\left(w_vw^{2-\beta+\eta}\right).\end{aligned}$$

Combining both cases we obtain $\mathbb{E}(|N_{\ge w}(v)|) \le O(\min\{w_vw^{2-\beta+\eta}, nw^{1-\beta+\eta}\})$.

The second statement follows because the expected total number of neighbours of v is $\Theta(w_v)$. Therefore, the probability that a random neighbour of v has weight at least w is $\mathbb{P}(\{u,v\} \in E \wedge w_u \ge w)/\mathbb{P}(\{u,v\} \in E) = O(w^{2-\beta-\eta})$, as required.

(b) This follows completely analogously to (a), except that we use that the expected number of vertices of weight at least w is at least $\Omega(nw^{1-\beta-\eta})$ by the lower bound in the power-law condition (6.1). □

We often need to bound the expected number of neighbours of a given vertex in some geometric region, which we may do by the following lemma.

Lemma 6.14. *Let $\eta > 0$ and $C > 1$ be constants, define $m := \min\{\alpha, \beta-1-\eta\}$ and consider a closed ball $B \subset \mathbb{T}^d$ of radius $r > 0$ centred at 0. Let $v = (x_v, w_v)$ be a vertex with fixed weight and position. Then*

$$\mathbb{E}\left(|N(v)\cap B|\right) = O(n\mathrm{Vol}(B)) \cdot \begin{cases}\min\left\{\frac{w_v}{n\mathrm{Vol}(B)},1\right\}, & \text{if } \|x_v\| \le Cr,\\ \min\left\{\left(\frac{w_v}{\|x_v\|^d n}\right)^m,1\right\} & \text{if } \|x_v\| \ge Cr.\end{cases}$$

Proof. In the first case $\|x_v\| \le Cr$, the expected number of vertices in B is $n\mathrm{Vol}(B)$, so $\mu \le n\mathrm{Vol}(B)$. On the other hand, the expected number of neighbours of v is $O(w_v)$, so $\mu = O(1)\cdot\min\{w_v, n\mathrm{Vol}(B)\}$.

For the second case, as before $\mu \le n\mathrm{Vol}(B)$. This proves the claim in the case $w_v \ge \|x_v\|^d n$, so assume otherwise. Observe that every vertex in B has distance $\Theta(\|x_v\|)$ from v, and that the expected number of vertices in B of weight at least w is $O(n\mathrm{Vol}(B)w^{1-\beta+\eta})$.

Consider first the case $\alpha < \beta-1-\eta$. Then by Lemma 6.10,

$$\begin{aligned}\mathbb{E}\left(|N(v)\cap B|\right) \le O\Bigg(&n\mathrm{Vol}(B)\min\left\{\left(\frac{w_v}{\|x_v\|^d n}\right)^\alpha,1\right\}\\ &+\int_{w_{\min}}^\infty n\mathrm{Vol}(B)w^{1-\beta+\eta}\frac{d}{dw}\min\left\{\left(\frac{ww_v}{\|x_v\|^d n}\right)^\alpha,1\right\}dw\Bigg).\end{aligned}$$

Note that the exponent of w in the integrand is always negative, no matter which value the minimum attains. Moreover, recall that we assumed $w_v < \|x_v\|^d n$ and hence for $w = w_{\min}$ the minimum is $O((w_v/(\|x_v\|^d n))^\alpha)$. Thus by applying Lemma 6.11 (with $\tilde{w} = w_{\min} = \Theta(1)$), the integral also evaluates to $O(n\mathrm{Vol}(B)\cdot\min\{(w_v/(\|x_v\|^d n))^\alpha, 1\})$, as required.

On the other hand, if $\alpha+1-\beta+\eta \ge 0$, then by Observation 6.8 we may restrict ourselves to $\alpha < \infty$. Moreover, using Lemma 6.10 (with lower bound 0), we obtain

the following estimate

$$\mathbb{E}\left(|N(v)\cap B|\right) = O\left(\int_0^\infty n\mathrm{Vol}(B)w^{1-\beta+\eta}\frac{d}{dw}\min\left\{\left(\frac{ww_v}{\|x_v\|^d n}\right)^\alpha, 1\right\}dw\right).$$

This integral evaluates to $O(n\mathrm{Vol}(B)\cdot(w_v/(\|x_v\|^d n))^{\beta-1-\eta})$, by Lemma 6.11 (with $\tilde{w} = \|x_v\|^d n/w_v$). Since we have already shown that $\mathbb{E}\left(|N(v)\cap B|\right) \le n\mathrm{Vol}(B)$, this proves the claim. □

In the last lemma of this section we show that whp there are no vertices whose weight is much larger than their distance from the origin.

Lemma 6.15. *Let $\eta > 0$ be a constant and consider a closed ball $B \subset \mathbb{T}^d$ of radius $r > 0$ centred at 0, satisfying $n\mathrm{Vol}(B) = \Theta(r^d n) = \omega(1)$. Then with probability $1 - (r^d n)^{-\Omega(\eta)}$ there is no vertex $v = (x_v, w_v)$ with $x_v \in \mathbb{T}^d \setminus B$ and $w_v \ge \hat{w}_+(\|x_v\|^d n)$.*

Proof. Let $\hat{n}$ be the number of such vertices. Let $r' > r$, then the probability density to find a vertex $v = (x_v, w_v)$ with $\|x_v\| = r'$ is equal to the volume of an r'-sphere around 0 that is intersected with $\mathbb{T}^d$. By ignoring the intersection, we can only make the volume larger, so it is at most $O((r')^{d-1}n)$. Moreover, the probability that a vertex has weight at least w is at most $O(w^{1-\beta+\eta/2})$ by the power-law condition (6.1) (using $\gamma = \eta/2$). Hence, by Lemma 6.10 and Lemma 6.11 we obtain

$$\mathbb{E}\left(\hat{n}\right) = O(1)\cdot\int_r^\infty (r')^{d-1}n((r')^d n)^{(1-\beta+\eta/2)/(\beta-1-\eta)}dr' = (r^d n)^{-\Omega(\eta)},$$

and the statement follows by Markov's inequality. □

6.3. Evolution of the process

In this section we will prove two theorems which describe the geometrical evolution of the process in detail. First we show that in the supercritical regime the process will reach certain regions whp in a given time, yielding a *lower bound on its speed*. This lower bound also applies in the critical regime if in the first step sufficiently many heavy vertices were activated, an event which holds with at least constant probability. Afterwards, we show that certain regions cannot be reached too early in the process, providing an *upper bound on its speed*. From this we then derive Theorem 6.7 in Section 6.3.3.

6.3.1. **Lower bound on the speed.** In this section we show lower bounds for the probability that a vertex in a specific region and with a specific weight will be active in some round, provided that we start in the supercritical case. Recall that the supercritical case is defined by $\rho = \omega(\hat{\rho})$ if the weight follow a strong power-law, and $\rho \ge \hat{\rho}^{1-\delta}$ for some constant $\delta > 0$ otherwise. The same bounds also hold in the critical case if at least k 'heavy' vertices are activated in the first round, which happens with probability $\Omega(1)$.

The key idea is that the infection spreads in two ways: firstly, from 'heavy' vertices in one region to 'heavy' vertices in the next (surrounding) region, where the volume of the region increases by an exponent of $\zeta-\varepsilon$ in each step, and secondly,

from vertices of weight w to nearby vertices of weight $w^{1/(\zeta-\varepsilon)}$. We formalise the occurring weights with the following definitions.

Definition 6.16. *Let $\varepsilon > 0$ be a constant and let $\eta = \eta(\varepsilon) > 0$ be a constant which is sufficiently small compared to ε, cf. Section 6.1.3.3. For any integers $i \geq 0$ and $\ell \geq 0$, we define*

$$w_{i,\ell} := \hat{w}_-(\nu_i)^{(\zeta-\varepsilon)^{-\ell}} = \nu_i^{(\zeta-\varepsilon)^{-\ell}/(\beta-1+\eta)} = \nu^{(\zeta-\varepsilon)^{i-\ell}/(\beta-1+\eta)}.$$

We abbreviate the set of all vertices in B_i of weight at least at least $w_{i,0}$ by $U_i := V_{\geq w_{i,0}} \cap B_i$, and we call such vertices heavy *if the value of i is clear from the context. As usual, superscripts in the notation denote active vertices, so $U_i^{\leq j} := V_{\geq w_{i,0}}^{\leq j} \cap B_i$. Furthermore we denote by $\mathcal{H}(i)$ the event that in round $i+3$ all vertices in U_i are active.*

The following theorem gives lower bounds on the probability that a vertex is active in some round. It agrees with the above intuition in the following sense: if j is the first round in which a vertex has, say, probability $1/2$ to be active according to the bound in Theorem 6.17(c), then j agrees with the round that is predicted by the above intuition, up to additive constants. We will see in Section 6.4 that Theorem 6.20 provides matching lower bounds on j, up to minor order terms. So in this sense, Theorem 6.17 is tight.

Theorem 6.17. *Let ε and η be as in Definition 6.16. Assume furthermore that we are in the supercritical case, or instead that $|U_i^{\leq 1}| \geq k$. Then the following is true:*

(a) Whp $|U_i| = \nu_i^{\Omega(\eta)}$ and $|U_i| = O(\nu_i)$ uniformly for all $i \geq 0$.

(b) Whp all the events $\mathcal{H}(i)$ occur.

(c) There exist constants $C_0, C_1, C_2 > 0$ such that the following holds: Let $v = (x_v, w_v)$ be any vertex with fixed position and weight and let $i, \ell \geq 0$ be such that $x_v \in B_i$ and $w_v \geq \max\{w_{i,\ell}, C_0\}$. Then for sufficiently large $n \in \mathbb{N}$,

$$\mathbb{P}\left(v \in V^{\leq i+3+\ell} \;\middle|\; \mathcal{H}(0), \dots, \mathcal{H}(i)\right) \geq 1 - \exp\left(-C_1 \nu_i^{C_2(\zeta-\varepsilon)^{-\ell}}\right).$$

Before we prove Theorem 6.17, we remark the following.

Remark 6.18. Our proof will in fact show that (c) still holds if we replace B_i by an arbitrary ball of the same volume, and that it suffices if only a constant fraction of all heavy vertices is active. More precisely, let B be any ball, and let $\mathcal{H}(i', B)$ be the event that in round i' at least half of the vertices in B of weight at least $\hat{w}_-(n\mathrm{Vol}(B))$ are active. Then there are constants $C_0, C_1, C_2 > 0$ such that for any vertex $v = (x_v, w_v)$ with fixed position $x_v \in B$, and with fixed weight $w_v \geq \max\left\{(n\mathrm{Vol}(B))^{(\zeta-\varepsilon)^{-\ell}/(\beta-1+\eta)}, C_0\right\}$ we have

$$\mathbb{P}\left(v \in V^{\leq i'+\ell} \;\middle|\; \mathcal{H}(i', B)\right) \geq 1 - \exp\left(-C_1 (n\mathrm{Vol}(B))^{C_2(\zeta-\varepsilon)^{-\ell}}\right).$$

For the sake of readability, we omit the details and prove Theorem 6.17 only in the case $B = B_i$.

Proof of Theorem 6.17. (a) By definition of U_i,

$$\mathbb{E}(|U_i|) \geq \nu_i w_{i,0}^{1-\beta-\eta/2} = \nu_i^{\Omega(\eta)}.$$

The random variable $|U_i|$ is Poisson distributed, and thus $\mathbb{P}(|U_i| \leq \mathbb{E}(|U_i|)/2) \leq \exp(-\Omega(\mathbb{E}(|U_i|)))$. By a union bound over all i, we see that whp $|U_i| \geq \nu_i^{\Omega(\eta)}$ for all $i \geq 0$. Similarly, the upper bound follows since $0 \ll \mathbb{E}(|U_i|) \leq \mathbb{E}(|B_i|) = O(\nu_i)$.

(b) We first show that in the supercritical case for weak power-law weights, whp $|U_0^{\leq 1}| \geq k$. Let $v = (x_v, w_v)$ be a vertex in U_0. Then we claim that in round 1, such a vertex will be active with at least constant probability. We may restrict ourselves to the case $w_v \leq \nu$, since larger weights make it only easier to become active.

Consider a ball around v with the property that every vertex of weight at least $w_{\min}$ (so *all* vertices) in this ball have probability $\Omega(1)$ to connect to v. Observe that by condition (6.2) and (6.3) on the edge probabilities we may choose the ball to have volume $\Omega(w_v/n)$. (For $\alpha < \infty$ we may choose the volume to be exactly w_v/n, for $\alpha = \infty$ we may have to choose it smaller by at most a constant factor.) Since $w_v \leq \nu$, at least a constant fraction of this ball lies in B_0. Hence, $\mathbb{E}(|N^{\leq 0}(v)|) = \Omega(\rho w_v) = \omega(1)$. Since $|N^{\leq 0}(v)|$ is a Poisson distributed random variable, v becomes active whp. In particular, if we fix any k vertices in U_0 then whp all k of them will be active in round 1. This implies that whp $|U_0^{\leq 1}| \geq k$, as claimed.

Next we show that $|U_0^{\leq 1}| \geq k$ implies that whp at least an $\Omega(1)$ fraction of all vertices in U_0 is active in round 2. For $\alpha < \infty$, any two vertices in U_0 have probability at least $\Omega(\min\{w_0^2/\nu, 1\}) = \Omega(1)$ to be adjacent, independently of each other. For $\alpha = \infty$, the probability is also $\Omega(1)$ since $w_0^2/\nu = \omega(1)$. Hence, by the Chernoff bound, whp $|U_0^{\leq 2}| = \Omega(|U_0|)$.

We show that the same holds in the supercritical case for strong power-law weights, i.e. whp at least an $\Omega(1)$ fraction of U_0 is active in round 2. Recall that $\rho = \omega(\nu^{-1/(\beta-1)})$ since we are supercritical. Let ρ' be a function with the properties $\rho' = o(\rho)$, $\rho' = o(1/w_0) = o(\nu^{-1/(\beta-1+\eta)})$, and $\rho' = \omega(\nu^{-1/(\beta-1)})$, and let $w' := 1/\rho'$. Note that $\mathbb{E}(|V_{\geq w'} \cap B_0|) = \Omega((w')^{1-\beta}\nu) = \omega(1)$.

As for weak power-laws, for a vertex v of weight at least w', we consider a ball B around v of volume w'/n. In the case $\alpha < \infty$, every vertex in B has probability $\Omega(\min\{w_{\min}w'/w', 1\}) = \Omega(1)$ to connect to v. In the case $\alpha = \infty$, we may achieve the same by shrinking the ball B by at most a constant factor. In either case, the expected number of vertices in $V^{\leq 0} \cap B$ is $\rho n \mathrm{Vol}(B) = \Omega(\rho/\rho') = \omega(1)$. Hence, every vertex in $V_{\geq w'} \cap B_0$ is in $V^{\leq 1}$ with probability $1 - o(1)$. By Markov's inequality, whp the number of vertices in $V_{\geq w'} \cap B_0$ that are *not* in $V^{\leq 1}$ is $o(\mathbb{E}(|V_{\geq w'} \cap B_0|))$. In particular, whp $|V_{\geq w'}^{\leq 1} \cap B_0| = \omega(1)$.

Finally, for any two vertices $u \in U_0$ and $v \in V_{\geq w'} \cap B_0$, the probability that u and v are adjacent is $\Omega(1)$, since $w_0 w'/\nu \geq w_0^2/\nu = \omega(1)$ with room to spare. The claim now follows as for the other cases.

So we have shown that in all cases whp an $\Omega(1)$ fraction of all vertices in U_0 is active in round 2, so let us assume this. To show that $\mathcal{H}(0)$ holds whp, recall that any two vertices in U_0 have probability $\Omega(1)$ to be connected. Therefore, the probability that a vertex in U_0 does not become active in round 3 is at most $\mathbb{P}\left(\mathrm{Bin}(|U_0|, \Omega(1)) < k\right) = \exp(-\Omega(|U_0|)) = o(1/|U_0|)$ by a Chernoff bound. Hence, by the union bound whp all vertices in U_0 are active in round 3. This proves that $\mathcal{H}(0)$ holds whp.

It remains to prove that the statement holds uniformly for all $i \geq 1$. By (a) we may assume that for all $i \geq 0$ the set $|U_i|$ satisfies $|U_i| = \nu_i^{\Omega(\eta)}$ and $|U_i| = O(\nu_i)$.

We claim that any two vertices $v_{i-1} \in U_{i-1}$ and $v_i \in U_i$ with fixed position and weight form an edge with probability $\Omega(1)$. Indeed, this follows immediately since their distance is at most $(\nu_i/n)^{1/d}$, and hence

$$\frac{w_{v_{i-1}} w_{v_i}}{\|x_{v_i} - x_{v_{i-1}}\|^d n} \geq \frac{w_{i-1,0} w_{i,0}}{\nu_i} = \nu_i^{\frac{(\zeta-\varepsilon)^{-1}+1}{\beta-1+\eta}-1} = \nu_i^{\Omega(\varepsilon)-O(\eta)} = \omega(1).$$

Therefore, the number of edges from a vertex $v_i \in U_i$ into U_{i-1} is lower bounded by a binomially distributed random variable $\mathrm{Bin}(|U_{i-1}|, \Omega(1))$. By the Chernoff bound, the probability that v_i has less than k neighbours in U_{i-1} is at most $\exp(-\Omega(|U_i|)) = \exp\left(-\nu_{i-1}^{\Omega(\eta)}\right)$. A union bound over all vertices in U_{i-1} shows that still with probability at least $1 - \exp\left(-\nu_{i-1}^{\Omega(\eta)}\right)$ every vertex in U_i has at least k neighbours in U_{i-1}. A union bound over all $i \geq 1$ shows that whp the same is still true for all $i \geq 1$ simultaneously. Hence, by a simple induction, all the events $\mathcal{H}(i)$ occur, as required.

(c) We only give the proof in the case $\alpha < \infty$, and explain in the end the changes that are necessary for $\alpha = \infty$. For $\alpha < \infty$, we prove the statement for $C_0 := (8k)^{2d/(\varepsilon^2(\beta-2))}$, $C_1 := 4^{-d/\varepsilon}$ and $C_2 := (\varepsilon(\beta-2)/2)(\zeta-\varepsilon)/(\beta-1+\eta)$, where we assume that $\varepsilon > 0$ is sufficiently small. We use induction on ℓ. If $w_v \geq w_{i,0}$ then $\mathcal{H}(i)$ implies that $v \in A(i+3)$, so for $\ell = 0$ there is nothing to show.

So let $\ell \geq 1$. Before we start with the inductive step, note that we may assume

$$\nu_i^{(\zeta-\varepsilon)^{-\ell+1}} \geq C_0^{\beta-1+\eta} \geq (8k)^{2d(\beta-1+\eta)/(\varepsilon^2(\beta-2))}, \tag{6.11}$$

since otherwise both the statements for ℓ and $\ell - 1$ concern only vertices of weight at least C_0, and thus the case ℓ follows trivially from the statement for $\ell - 1$.

Let v be a vertex with position $x_v \in B_i$ and with weight $w_v \geq w_{i,\ell}$. We claim that every vertex in distance at most $r_\ell := \left(\nu_i^{(\zeta-\varepsilon)^{-\ell+1}}/n\right)^{1/d}$ with weight at least $w_{i,\ell-1}$ has probability $\Omega(1)$ to connect to v. Indeed, this follows from

$$\begin{aligned}\frac{w_{i,\ell-1} w_{i,\ell}}{r_\ell^d n} &= \nu_i^{(2-\beta-\eta+(\zeta-\varepsilon)^{-1})(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)} \\ &\geq \nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)} \overset{(6.11)}{\geq} (8k)^{d/\varepsilon} \geq 1 \end{aligned} \tag{6.12}$$

Since $\ell \geq 1$, we have $r_\ell \leq (\nu_i/n)^{1/d}$, which is the radius corresponding to the ball B_i. Hence, if we consider a ball around v with radius r_ℓ, then at least a 2^{-d}

proportion of this ball falls into B_i. Therefore, the expected size of $N_{\ell-1}(v)$ satisfies

$$\begin{aligned}\mathbb{E}(|N_{\ell-1}(v)|) &= \Omega(1)\cdot 2^{-d} r_\ell^d n w_{i,\ell-1}^{1-\beta+\eta} \nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)} \\ &\geq c2^{-d}\nu_i^{(\varepsilon(\beta-2)/2+2\eta)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)}\end{aligned}$$

for some constant $c > 0$ and any sufficiently large n. Furthermore, if the constant $\varepsilon > 0$ is sufficiently small we obtain

$$\mathbb{E}(|N_{\ell-1}(v)|) \geq 2^{-d/\varepsilon}\nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)} \overset{(6.11)}{\geq} (4k)^{d/\varepsilon} \geq 8k.$$

Recall that $|N_{\ell-1}(v)|$ is a Poisson distributed random variable and thus we have

$$\begin{aligned}\mathbb{P}(|N_{\ell-1}(v)| < k) &\leq \exp(-\mathbb{E}(|N_{\ell-1}(v)|)/4)\cdot\exp(-\mathbb{E}(|N_{\ell-1}(v)|)/4) \\ &\leq \exp(-2k)\exp\left(-4^{-d/\varepsilon}\nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)}\right).\end{aligned}$$

So assume that $|N_{\ell-1}(v)| \geq k$, and pick any k vertices $v_1,\dots,v_k \in N_{\ell-1}(v)$. By induction hypothesis,

$$\begin{aligned}\mathbb{P}(v_1 \notin A(i+\ell+1)) &\leq \exp\left(-4^{-d/\varepsilon}\nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+2}/(\beta-1+\eta)}\right) \\ &\overset{(6.11)}{\leq} \exp\left(-2\cdot 4^{-d/\varepsilon}\nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)}\right) \\ &\overset{(6.11)}{\leq} \exp(-k^d)\exp\left(-4^{-d/\varepsilon}\nu_i^{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}/(\beta-1+\eta)}\right)\end{aligned}$$

where the second inequality holds since $\zeta-\varepsilon \geq 1+\varepsilon$ for any sufficiently small $\varepsilon > 0$.

The same bound applies to the other v_j. By a simple union bound,

$$\mathbb{P}(v_1,\dots,v_k \in A(i+\ell+1)) \geq 1 - k\exp(-k^d)\exp\left(-4^{-d/\varepsilon}\nu_i^{\frac{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}}{\beta-1+\eta}}\right).$$

Hence,

$$\begin{aligned}\mathbb{P}\left(v \notin A(i+\ell+2)\right) &\leq \mathbb{P}\left(|N_{\ell-1}(v)| < k\right) + \mathbb{P}\left(\{v_1,\dots,v_k\} \not\subset A(i+\ell+1)\right) \\ &\leq \left(\exp(-2k) + k\exp(-k^d)\right)\exp\left(-4^{-d/\varepsilon}\nu_i^{\frac{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}}{\beta-1+\eta}}\right) \\ &\leq \exp\left(-4^{-d/\varepsilon}\nu_i^{\frac{(\varepsilon(\beta-2)/2)(\zeta-\varepsilon)^{-\ell+1}}{\beta-1+\eta}}\right) \\ &= \exp\left(-C_1\nu_i^{C_2(\zeta-\varepsilon)^{-\ell}}\right)\end{aligned}$$

as required.

For $\alpha = \infty$, equation (6.12) does not imply that the corresponding vertices connect with probability $\Omega(1)$, but it suffices to decrease r_ℓ by at most a constant factor to ensure this property. This can be compensated by changing (for example) C_1. We omit the details. □

6.3.2. **Upper bound on the speed.** In this section we show upper bounds for the probability that a vertex in a specific region and with a specific weight will be active in some round (Theorem 6.20(f)). To bound the probability, we need to condition

on the event that the process does not infect too many vertices in certain regions and rounds, which we show to hold whp in Theorem 6.20(e). More precisely, we consider the following families of events.

Definition 6.19. *Let $\varepsilon > 0$ be a constant and let $\eta = \eta(\varepsilon) > 0$ be a constant which is sufficiently small compared to ε, cf. Section 6.1.3.3. Moreover, let $h = h(n)$ be a function satisfying $h(n) = \omega(1)$, $h(n) = o(\log n)$, and $h(n) = \nu^{o(1)}$. Then we define the following families of events:*

- *For all $i \geq 0$*

$$\mathcal{E}(i) := \{V^{\leq i} \cap (\mathbb{T}^d \setminus \widetilde{B}_i) = \emptyset\},$$

  in other words, no vertex outside of $\widetilde{B}_i$ is activated by time i;

- *For all $\ell \geq 0$ and $w \geq w_{min}$*

$$\mathcal{F}(\ell, w) := \mathcal{F}_{\varepsilon,\eta,h}(\ell, w) := \left\{ \left| V^{\leq \ell}_{\geq w} \cap 2^\ell \widetilde{B}_0 \right| \leq h^\ell w^{2-\beta+\eta} \widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}(\beta-1)^{-1}} \right\},$$

  and

$$\mathcal{F}(\ell) := \mathcal{F}_{\varepsilon,\eta,h}(\ell) := \bigcap_{w' \geq w_{min}} \mathcal{F}(\ell, w'),$$

  i.e. the number of vertices in $2^\ell \widetilde{B}_0$ being activated by time ℓ is not 'too large';

- *For all $j \geq 0$*

$$\mathcal{G}(j) := \mathcal{G}_{\varepsilon,\eta,h}(j) := \bigcap_{j'=0}^{j} (\mathcal{E}(j') \cap \mathcal{F}(j')),$$

  in other words, it is all 'good' events up to time j hold.

Theorem 6.20. *Let ε, η, h be given as in Definition 6.19 and assume $\alpha > \beta - 1$. Then, for sufficiently large n,*

(a) $\mathcal{E}(0)$ is always satisfied;

(b) $\mathbb{P}(\mathcal{F}(0)) \geq 1 - O(h^{-1})$;

(c) For all $i \geq 1$ we have

$$\mathbb{P}\left(\mathcal{E}(i) \mid \mathcal{G}(i-1)\right) \geq 1 - h^{-\Omega(i)};$$

(d) For all $\ell \geq 1$ we have

$$\mathbb{P}\left(\mathcal{F}(\ell) \mid \mathcal{G}(\ell-1)\right) \geq 1 - h^{-\Omega(\ell)};$$

(e) Whp, the events $\mathcal{G}(j)$ hold for all $j \geq 0$;

(f) For all $i \geq 1$ and $\ell \geq 0$, and for every fixed vertex $v = (x_v, w_v)$ such that $x_v \in \mathbb{T}^d \setminus 2^{\ell+1}\widetilde{B}_{i-1}$ and $w_v \geq w_{min}$ we have

$$\mathbb{P}\left(v \in V^{\leq i+\ell} \,\middle|\, \mathcal{G}(i+\ell-1)\right) \leq w_v 2^{\ell d} \widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)}.$$

Proof. First note that all statements only become easier if the edge probabilities are decreased. Hence, by Observation 6.8 we may restrict ourselves to the case $\alpha < \infty$, since this case dominates the case $\alpha = \infty$.

To prove (c), (d), and (f), we will use induction on $i+\ell$, where we set $\ell = 0$ and $i = 0$ in (c) and (d), respectively. In particular, in order to prove (c) and (d) for i, ℓ, we will assume statement (f) for i', ℓ' as long as $i' + \ell' \le i + \ell - 1$. Throughout the proof, we will mutually assume that n is sufficiently large; for example, we will use that h and $\widetilde{\nu}_0$ are larger than any fixed constant without further comment.

(a) This statement is trivial as only vertices in $B_0 \subset \widetilde{B}_0$ are active at time 0.

(b) Fix a weight $w \ge w_{\min}$ and note that

$$\left|V_{\ge w}^{\le 0} \cap \widetilde{B}_0\right| \le \left|V_{\ge w} \cap B_0\right|$$

since initially activation only occurs within B_0. Furthermore, the right-hand side is a Poisson distributed random variable and we have

$$\mathbb{E}(|V_{\ge w} \cap B_0|) \le O(\nu w^{1-\beta+\eta}) = O(w^{2-\beta} \widetilde{\nu}_0^{\frac{\beta-2}{\beta-1} - \Omega(\varepsilon)} w^{-1+\eta}) \le O(w^{2-\beta} \widetilde{\nu}_0^{\frac{\beta-2}{\beta-1} - \Omega(\varepsilon)}), \tag{6.13}$$

where in the last step we used that $w^{-1+\eta} = O(1)$. Let $\bar{w}$ be the weight that satisfies $\bar{w}^{2-\beta} \widetilde{\nu}_0^{(\beta-2)/(\beta-1)} = h$. Then by Markov's inequality, $\left|V_{\ge \bar{w}}^{\le 0} \cap \widetilde{B}_0\right| = 0$ with probability $1 - O(\widetilde{\nu}_0^{-\Omega(\varepsilon)} h) = 1 - O(h^{-1})$ since $h = \widetilde{\nu}_0^{o(1)}$. Note that this implies (b) for all $w \ge \bar{w}$.

For smaller w, observe in (6.13) that $\widetilde{\nu}_0^{-\Omega(\varepsilon)}$ dominates every $O(1)$-term for sufficiently large n. Let $\mathcal{F}_*(0, w)$ be the event that $\left|V_{\ge w} \cap B_0\right| \le (2w)^{2-\beta} \widetilde{\nu}_0^{(\beta-2)/(\beta-1)}$ and note that

$$\mathbb{P}(\mathcal{F}_*(0, w)) \ge 1 - \exp\left(-\Omega(1) \cdot (2w)^{2-\beta} \widetilde{\nu}_0^{(\beta-2)/(\beta-1)}\right) \tag{6.14}$$

by (6.13) and a Chernoff bound. The exponent $(\beta-2)/(\beta-1)$ in (6.14) equals the exponent $1-(\beta-1)^{-1}$ of $\widetilde{\nu}_0$ in $\mathcal{F}(0, w)$. Hence, if $\mathcal{F}_*(0, w')$ holds for some $w' \ge w_{\min}$, then $\mathcal{F}(0, w)$ holds for all $w \in [w', 2w']$. Therefore, it remains to prove $\mathcal{F}_*(0, 2^s w_{\min})$ for all $s \in \{0, \dots, \log_2(\bar{w}/w_{\min}) - 1\}$. A union bound over all such s using (6.14) shows that all these events hold with probability $1 - \exp\{-\Omega(h)\} = 1 - O(h^{-1})$. This concludes the proof of (b).

(c) We will show that with sufficiently large probability, no vertex in $\mathbb{T}^d \setminus \widetilde{B}_i$ has a neighbour in $\widetilde{B}_{i-1}$. This will imply the statement, since we assumed $\mathcal{G}(i-1)$, which means in particular that all active vertices in round $i-1$ are in $\widetilde{B}_{i-1}$.

By Lemma 6.15, with probability $1 - \widetilde{\nu}_i^{-\Omega(\eta)} \ge 1 - h^{-\Omega(i)}$ there is no vertex $v = (x_v, w_v)$ such that $x_v \in \mathbb{T}^d \setminus \widetilde{B}_i$ and $w_v \ge \hat{w}_+(\|x_v\|^d n)$. So let $v = (x_v, w_v)$ be a vertex satisfying $x_v \in \mathbb{T}^d \setminus \widetilde{B}_i$ and $w_v \le \hat{w}_+(\|x_v\|^d n)$, and note in particular that $\|x_v\| \ge (\widetilde{\nu}_i/n)^{1/d} \ge 2(\widetilde{\nu}_{i-1}/n)^{1/d}$. Hence, due to Markov's inequality, the probability of v having a neighbour in $\widetilde{B}_{i-1}$ is at most

$$\mathbb{E}(N(v) \cap \widetilde{B}_{i-1}) = O(1) \cdot \widetilde{\nu}_{i-1} \left(\frac{w_v}{\|x_v\|^d n}\right)^{\beta-1-\eta}$$

by Lemma 6.14.

We call any such vertex $v = (x_v, w_v)$ *bad*, i.e. v is bad if it satisfies $x_v \in \mathbb{T}^d \setminus \widetilde{B}_i$ and $w_v \le \hat{w}_+(\|x_v\|^d n)$, and if v has at least one neighbour in $\widetilde{B}_{i-1}$. Integrating over $r_v := \|x_v\|$ and using Lemma 6.10, we can thus bound the expected number $\mathbb{E}(n_{\mathrm{bad}})$ of bad vertices by

$$O(1)\int_{(\widetilde{\nu}_i/n)^{1/d}}^{\infty} r_v^{d-1} n \int_{w_{\min}}^{\hat{w}_+(r_v^d n)} w_v^{1-\beta+2\eta} \frac{d}{dw_v}\widetilde{\nu}_{i-1}(r_v^d n)^{1-\beta+\eta} w_v^{\beta-1-\eta} dw_v dr_v.$$

Hence, since $(\hat{w}_+(r_v^d n))^\eta = (r_v^d n)^{\eta/(\beta-1-\eta)} \le (r_v^d n)^\eta$, we have

$$\mathbb{E}(n_{\mathrm{bad}}) = O(1)\int_{(\widetilde{\nu}_i/n)^{1/d}}^{\infty} r_v^{-1}(r^d n)^{2-\beta+2\eta}\widetilde{\nu}_{i-1} dr_v = O(1)\cdot \widetilde{\nu}_i^{2-\beta+2\eta}\widetilde{\nu}_{i-1} = \widetilde{\nu}_i^{2\eta-\Omega(\varepsilon)}.$$

Thus by Markov's inequality, with probability at least $1 - \widetilde{\nu}_i^{2\eta-\Omega(\varepsilon)} \ge 1 - h^{-\Omega(i)}$ there is no such vertex. Statement (c) follows.

(d) We distinguish two cases. For $w \ge \widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell}(1+3\eta)/(\beta-1)}$, we consider the upper bound $\left|V_{\ge w}^{\le \ell} \cap 2^\ell \widetilde{B}_0\right| \le \left|V_{\ge w} \cap 2^\ell \widetilde{B}_0\right|$. Since we have

$$\mathbb{E}\left(\left|V_{\ge w} \cap 2^\ell \widetilde{B}_0\right|\right) = O(1)\cdot 2^{\ell d}\widetilde{\nu}_0 w^{1-\beta+\eta} = O(1)\cdot w^{2-\beta} 2^{\ell d}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}(1+\eta)/(\beta-1)},$$

and since this number is Poisson distributed, we obtain

$$\mathbb{P}\left(\left|V_{\ge w} \cap 2^\ell \widetilde{B}_0\right| \le 2(2w)^{2-\beta} 2^{\ell d}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}\right) \tag{6.15}$$
$$\ge 1 - \exp\left(-\Omega(1)\cdot w^{2-\beta} 2^{\ell d}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}\right)$$

by a Chernoff bound.

Similarly as in the proof of (b), it suffices to establish the bound in (6.15) only for weights of the form $2^s w_{\min}$ for $s \in \{0, \dots, \log_2(\bar{w}_\ell/w_{\min}) - 1\}$, where $\bar{w}_\ell$ is defined by $\bar{w}_\ell^{2-\beta} 2^{\ell d}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)} = h^\ell$. A union bound over all such s proves that $\mathcal{F}(\ell, w)$ holds for all $w \ge \widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell}(1+3\eta)/(\beta-1)}$ with probability $1 - \exp(-\Omega(h^\ell)) = 1 - h^{-\Omega(\ell)}$.

For the second case assume that $w \le \widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell}(1+3\eta)/(\beta-1)}$. We claim that it suffices to restrict ourselves to vertices of weight at most $\hat{w} := \hat{w}_+^0(\widetilde{\nu}_0) = \widetilde{\nu}_0^{1/(\beta-1)}$. More precisely, we will show that with probability at least $1 - h^{-\Omega(\ell)}$, for every $w \le \widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell}(1+3\eta)/(\beta-1)}$ we have

$$|U(w)| \le h^\ell w^{2-\beta+\eta}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}, \tag{6.16}$$

where $U(w) := V_{\in[w,\hat{w}]}^{\le \ell} \cap 2^\ell \widetilde{B}_0$.

This suffices because by the first case there are sufficiently few other vertices active: we have proved that with probability at least $1 - h^{-\Omega(\ell)}$ we have

$$\left|V_{\ge \hat{w}}^{\le \ell} \cap 2^\ell \widetilde{B}_0\right| \le h^\ell \hat{w}^{2-\beta+\eta}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)} \le \frac{1}{2} h^\ell w^{2-\beta+\eta}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)},$$

for every $w \le \widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell}(\beta-1)^{-1}(1+3\eta)}$, since then $w = o(\hat{w})$ as $\ell \ge 1$.

Thus we want to bound $\mathbb{E}(|U|)$ by calculating the expected number of edges having one endpoint in $V^{\le \ell-1}$ and the other in $V_{\in[w,\hat{w}]} \cap 2^\ell \widetilde{B}_0$, i.e. we set

$$M(w) := E\left(V^{\le \ell-1}, V_{\in[w,\hat{w}]} \cap 2^\ell \widetilde{B}_0\right).$$

Furthermore we observe that each edge in $M(w)$ is also contained in at least one of the following two edge-sets: let

$$M_*(w) := E\left(V^{\leq \ell-1} \cap 2^{\ell+1}\widetilde{B}_0, V_{\in[w,\hat{w}]}\right),$$

and

$$M^*(w) := E\left(V \cap (\mathbb{T}^d \setminus 2^{\ell+1}\widetilde{B}_0), V_{\in[w,\hat{w}]} \cap 2^{\ell}\widetilde{B}_0\right),$$

then we have $M(w) \subset M_*(w) \cup M^*(w)$. It will turn out that the bound on $|U(w)| \leq |M_*(w)| + |M^*(w)|$ obtained this way strong enough to prove (6.16).

We start by estimating the summand $|M_*(w)|$. As a preparation, we first bound $\left|V_{\geq w'}^{\leq \ell-1} \cap 2^{\ell+1}\widetilde{B}_0\right|$, i.e. the number of vertices in a slightly larger region that were already active in the previous round. Since we assumed that $\mathcal{F}(\ell-1)$ holds, for those vertices which are also contained in the slightly smaller region $2^{\ell-1}\widetilde{B}_0$ we already know that

$$\left|V_{\geq w'}^{\leq \ell-1} \cap 2^{\ell-1}\widetilde{B}_0\right| \leq 2^{(\ell-1)d}(w')^{2-\beta+\eta}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell+1}/(\beta-1)}. \tag{6.17}$$

Now if $\ell = 1$, then no other vertices were active in round $\ell - 1 = 0$ by (a). For $\ell \geq 2$, we need to examine the remaining region $2^{\ell+1}\widetilde{B}_0 \setminus 2^{\ell-1}\widetilde{B}_0$. Note that this area is contained in $2^{\ell}\widetilde{B}_1$. Hence, we may apply (f) with $i' = 1$ and $\ell' = \ell - 2$, and thus

$$\begin{aligned}\mathbb{E}\left(\left|V_{\geq w'}^{\leq \ell-1} \cap (2^{\ell+1}\widetilde{B}_0 \setminus 2^{\ell-1}\widetilde{B}_0)\right|\right) &= O(1) \cdot 2^{(\ell+1)d}\widetilde{\nu}_0(w')^{1-\beta+\eta} w' 2^{(\ell-2)d}\widetilde{\nu}_1^{-\frac{(\zeta+\varepsilon)^{-\ell}}{\beta-1}} \\ &= O(1) \cdot h^{\ell}(w')^{2-\beta+\eta}\widetilde{\nu}_0^{1-\frac{(\zeta+\varepsilon)^{-\ell+1}}{\beta-1}}. \end{aligned} \tag{6.18}$$

Combining equations (6.17) and (6.18) we obtain

$$\begin{aligned}&\mathbb{E}\left(\left|V_{\geq w'}^{\leq \ell-1} \cap 2^{\ell+1}\widetilde{B}_0\right|\right) \\ &\qquad = O(1) \cdot \min\left\{h^{\ell}(w')^{2-\beta+\eta}\widetilde{\nu}_0^{1-\frac{(\zeta+\varepsilon)^{-\ell+1}}{\beta-1}}, 2^{(\ell+1)d}\widetilde{\nu}_0(w')^{1-\beta+\eta}\right\}, \end{aligned} \tag{6.19}$$

where the second term arises from dropping the condition on being active in round $i-1$. Now we denote by $\tilde{w} = \Theta(1)\widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell+1}(\beta-1)^{-1}} 2^{(\ell+1)d}h^{-\ell}$ the weight for which the two expressions in (6.19) coincide.

Recall that for any vertex $u = (x_u, w_u)$ of fixed weight (and independently of its position) we have $\mathbb{E}(|N(u)|) = \Theta(w_u)$ by Lemma 6.12. Moreover, by Lemma 6.13, the probability $q(w)$ for a random neighbour of u to have weight at least w satisfies $q(w) = O(w^{2-\beta+\eta/2})$, independently of u. Therefore we have

$$\mathbb{E}(|M_*(w)|) = O(q(w)) \sum_{u \in V_{\geq 0}^{\leq \ell-1} \cap\, 2^{\ell+1}\widetilde{B}_0} \mathbb{E}(|N(u)|),$$

and by Lemma 6.10 the sum evaluates to

$$O(1) \cdot \int_0^{\infty} \min\left\{h^{\ell} w_u^{2-\beta+\eta}\widetilde{\nu}_0^{1-\frac{(\zeta+\varepsilon)^{-\ell+1}}{\beta-1}}, 2^{(\ell+1)d}\widetilde{\nu}_0 w_u^{1-\beta+\eta}\right\}\left(\frac{d}{dw_u} w_u\right) dw_u.$$

Hence, by Lemma 6.11 with $\tilde{w}$ we obtain

$$\begin{aligned}\mathbb{E}(|M_*(w)|) &= O(1)\cdot w^{2-\beta+\eta/2}2^{(\ell+1)d}\widetilde{\nu}_0\tilde{w}^{2-\beta+\eta}\\ &= O(1)\cdot w^{2-\beta+\eta/2}2^{(\ell+1)d(3-\beta+\eta)}h^{\ell(\beta-2-\eta)}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell+1}(\beta-2+\eta)/(\beta-1)}\\ &\le w^{2-\beta+\eta/2}h^{\ell(\beta-2-2\eta)}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}(1+\Omega(\varepsilon)-O(\eta))/(\beta-1)}\\ &\le w^{2-\beta+\eta/2}h^{\ell(\beta-2-2\eta)}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}. \end{aligned}\tag{6.20}$$

Next we turn to the edges in $M^*(w)$. If $\ell = 1$, then $V^{\le\ell-1}\cap(\mathbb{T}^d\setminus 2^{\ell+1}\widetilde{B}_0)$ is empty by (a), hence also $M^*(w)$. So assume $\ell\ge 2$. Fix a vertex $v=(x_v,w_v)$ such that $x_v\in 2^\ell\widetilde{B}_0$ and $w_v\le\hat{w}$ and denote by $M^*(w,v):=\{e\in M^*(w)\mid v\in e\}$ the subset of $M^*(w)$ consisting of all edges incident with v. Now note that every edge in $M^*(w,v)$ must bridge a distance of at least $\tilde{r}_\ell := 2^\ell(\widetilde{\nu}_0/n)^{1/d}$ and hence Lemma 6.10 and Lemma 6.11 imply

$$\begin{aligned}\mathbb{E}\left(|M^*(w,v)|\right) &= O(1)\cdot\int_{r_\ell}^\infty r^{d-1}n\int_0^\infty w_*^{1-\beta+\eta}\frac{d}{dw_*}\min\left\{\left(\frac{w_*w_v}{r^dn}\right)^\alpha,1\right\}dw_*dr\\ &= O(1)\cdot\int_{\tilde{r}_\ell}^\infty r^{d-1}n\left(\frac{r^dn}{w_v}\right)^{1-\beta+\eta}dr = O(1)\cdot(\tilde{r}_\ell^dn)^{2-\beta+\eta}w_v^{\beta-1+\eta}\\ &= O(1)\cdot\left(\frac{2^{d\ell}\widetilde{\nu}_0}{w_v}\right)^{2-\beta+\eta}w_v = O(1)\cdot\widetilde{\nu}_0^{-\frac{(\zeta+\varepsilon)^{-\ell}}{\beta-1}}w_v,\end{aligned}$$

where the last step follows from $(\beta-2)(\beta-2-\eta)(\zeta+\varepsilon)^2\ge 1$ since we assumed $w_v\le\hat{w}=\widetilde{\nu}_0^{1/(\beta-1)}$ and $\ell\ge 2$. Hence,

$$\begin{aligned}\mathbb{E}(|M^*(w)|) &= O(1)\cdot\widetilde{\nu}_0\int_w^\infty w_v^{1-\beta+\eta/2}\frac{d}{dw_v}\left(\widetilde{\nu}_0^{-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}w_v\right)dw_v\\ &\le O(1)\cdot w^{2-\beta+\eta/2}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}.\end{aligned}$$

Together with (6.20), this shows that the expected number of vertices in $U(w)$ is also bounded by

$$\mathbb{E}(|U(w)|)\le 2w^{2-\beta+\eta/2}h^{\ell(\beta-2-2\eta)}\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)},$$

and therefore, by Markov's inequality, we have

$$\mathbb{P}\left(|U(w)|\ge\frac{1}{2}(2w)^{2-\beta+\eta}h^\ell\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)}\right)=w^{-\eta/2}h^{-\Omega(\ell)}.$$

As in the proof of (b), by a union bound over all weights of the form $2^sw_{\min}$ for $s\in\{0,\dots,\log_2(\hat{w}/w_{\min})-1\}$, we find that with probability $1-h^{-\Omega(\ell)}$, for all $w\ge w_{\min}$ we have

$$|U(w)|\le\frac{1}{2}w^{2-\beta+\eta}h^\ell\widetilde{\nu}_0^{1-(\zeta+\varepsilon)^{-\ell}/(\beta-1)},\tag{6.21}$$

concluding the proof of (d).

(e) This follows from (a)–(d) by a simple union bound.

(f) Fix a vertex $v=(x_v,w_v)$ such that $x_v\in\mathbb{T}^d\setminus 2^{\ell+1}\widetilde{B}_{i-1}$. The statement (f) is trivial if $w_v\ge\widetilde{\nu}_i^{(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)}$, so assume the contrary. We first estimate the

number of neighbours in $2^\ell \widetilde{B}_{i-1}$. Observe that every such vertex has distance at least $\tilde{r}_{i-1,\ell} := 2^\ell(\widetilde{\nu}_{i-1}/n)^{1/d}$ of v. Therefore, using

$$w_v \le \widetilde{\nu}_i^{(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)} \le \widetilde{\nu}_i^{(\zeta+\varepsilon)^{-2}/(\beta-1)},$$

we obtain

$$\begin{aligned}\mathbb{E}\left(\left|N(v)\cap 2^\ell \widetilde{B}_{i-1}\right|\right) &= O(1)\cdot 2^{\ell d}\widetilde{\nu}_{i-1}\int_0^\infty w^{1-\beta+\eta}\frac{d}{dw}\min\left\{\left(\frac{ww_v}{2^{\ell d}\widetilde{\nu}_{i-1}}\right)^\alpha, 1\right\} dw\\ &\overset{L.\ 6.11,\alpha>\beta-1}{=} O(1)\cdot w_v\left(\frac{2^{\ell d}\widetilde{\nu}_{i-1}}{w_v}\right)^{2-\beta+\eta}\\ &= O(1)\cdot w_v 2^{\ell d}\left(\widetilde{\nu}_i^{(\zeta+\varepsilon)^{-1}-(\zeta+\varepsilon)^{-2}/(\beta-1)}\right)^{2-\beta+\eta}.\end{aligned}$$

Since $(2-\beta+\eta)\left((\zeta+\varepsilon)^{-1}-(\zeta+\varepsilon)^{-2}/(\beta-1)\right) = -(\zeta+\varepsilon)^{-2}(\beta-1)^{-1}(1+\Omega(\varepsilon))$ we deduce

$$\begin{aligned}\mathbb{E}\left(\left|N(v)\cap 2^\ell \widetilde{B}_{i-1}\right|\right) &= O(1)\cdot w_v 2^{\ell d}\widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-2}/(\beta-1)-\Omega(\varepsilon)}\\ &\le \frac{1}{2}w_v 2^{\ell d}\widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)}. \qquad (6.22)\end{aligned}$$

In the case $\ell = 0$, this already proves the assertion since in round $i-1$ no vertex outside of $\widetilde{B}_{i-1}$ is active by $\mathcal{E}(i-1)$, and thus $\mathbb{P}(v \in V^{=i}) \le \mathbb{E}(N(v)\cap 2^\ell \widetilde{B}_{i-1})$.

So assume $\ell \ge 1$. Set $U_* := \left|N^{\le i+\ell-1}(v)\cap(\mathbb{T}^d \setminus 2^\ell \widetilde{B}_{i-1})\right|$. In this case, we can use the induction hypothesis of statement (f) for $i' = i$ and $\ell' = \ell - 1$ to estimate

$$\mathbb{E}(|U_*|) =$$

$$O(1)\int_0^\infty r^{d-1}n\int_0^\infty w^{1-\beta+\eta}\frac{d}{dw}\left(\min\{wh^\ell\widetilde{\nu}_i^{-\frac{(\zeta+\varepsilon)^{-\ell-1}}{\beta-1}}, 1\}\min\{\left(\frac{ww_v}{r^d n}\right)^\alpha, 1\}\right) dwdr.$$

To compute this integral, note that whenever the second minimum is attained by 1, the inner integral runs either over a polynomial in w with exponent $1-\beta+\eta < -1$, or over the zero function. On the other hand, whenever the second minimum is is attained by the expression $(ww_v/(r^d n))^\alpha$, then the inner integral runs over a a polynomial in w with exponent larger than -1 (either with exponent $\alpha-\beta+\eta > -1$, or even with exponent $\alpha-\beta+\eta+1$). Therefore, by Lemma 6.11 for $\tilde{w} = r^d n/w_v$, we obtain in all cases

$$\mathbb{E}(|U_*|) = O(1)\int_0^\infty r^{d-1}n\left(\frac{r^d n}{w_v}\right)^{1-\beta+\eta}\min\left\{\left(\frac{r^d n}{w_v}\right)h^\ell\widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-\ell-1}/(\beta-1)}, 1\right\} dr.$$

Similarly, let r_* be defined by $r_*^d n/w_v = h^{-\ell}\widetilde{\nu}_i^{(\zeta+\varepsilon)^{-\ell-1}/(\beta-1)}$, then the exponent of r in the antiderivative of the integrand is positive for all $r < r_*$ and negative for all $r > r_*$. Hence,

$$\mathbb{E}(|U_*|) = O(1)\cdot w_v\left(h^\ell\widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-\ell-1}/(\beta-1)}\right)^{\beta-2-\eta} \le \frac{1}{2}w_v 2^{\ell d}\widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-\ell-2}/(\beta-2)},$$

since $(\zeta+\varepsilon)(\beta-2-\eta) \ge 1$ and $h = \omega(1)$. Combined with equation (6.22) this proves the claim. □

6.3.3. **Isolation strategies.** In this section we prove Theorem 6.7. As outlined in Section 6.1.5, the corollary is a rather straightforward consequence of Theorem 6.20(e).

Proof of Theorem 6.7. By Theorem 6.20, whp there is no vertex outside of $\widetilde{B}_i$ which is active in round i. Therefore, it suffices to (permanently) remove by the end of round i all edges that cross the boundary of $\widetilde{B}_i$, i.e. all edges in $E(\widetilde{B}_i, \mathbb{T}^d \setminus \widetilde{B}_i)$.

This is very similar to [41, Lemma 7.1 and Theorem 7.2], where the number of edges cutting a grid is considered. It does not follow directly from this lemma since the error terms in [41] are too large for our purposes. However, what *does* follow directly from their proof is that among those edges that are completely contained in $2\widetilde{B}_i$, the number of edges that cross a fixed axis-parallel hyperplane is at most $\widetilde{\nu}_i^{\max\{3-\beta,1-1/d\}\pm o(1)}$. Since the boundary of $\widetilde{B}_i$ consists of a constant number of faces, this proves the bound for those edges which have both endpoints in $2\widetilde{B}_i$.

So it remains to consider the set $M_i := E(\widetilde{B}_i, \mathbb{T}^d \setminus 2\widetilde{B}_i)$. Let $\eta > 0$ be any constant, and let $v = (x_v, w_v)$ be a vertex such that $x_v \in \mathbb{T}^d \setminus 2\widetilde{B}_i$. Then by Lemma 6.14 (in the case $\alpha > \beta - 1$) the expected number of neighbours of v inside of $\widetilde{B}_i$ is

$$\mathbb{E}(N(v) \cap \widetilde{B}_i) = O(\widetilde{\nu}_i) \min\left\{ \left(\frac{w_v}{\|x_v\|^d n} \right)^{\beta-1-\eta}, 1 \right\}.$$

Note that v has distance at least $r_i := (\widetilde{\nu}_i/n)^{1/d}$ from the origin. Thus we may use Lemma 6.10 and Lemma 6.11 to estimate

$$\begin{aligned}\mathbb{E}(|M_i|) &= O(1) \int_{r_i}^{\infty} r^{d-1} n \int_0^{\infty} w_v^{1-\beta+2\eta} \frac{d}{dw_v} \widetilde{\nu}_i \min\left\{ \left(\frac{w_v}{r^d n} \right)^{\beta-1-\eta}, 1 \right\} dw_v dr \\ &= O(\widetilde{\nu}_i) \int_{r_i}^{\infty} r^{-1} (r^d n)^{2-\beta+2\eta} dr = O(\widetilde{\nu}_i^{3-\beta+2\eta}).\end{aligned}$$

Since $\widetilde{\nu}_i = \omega(1)$, we can deduce that $\mathbb{E}(|M_i|) \le \widetilde{\nu}_i^{3-\beta+3\eta}$ for sufficiently large n. Since this holds for all $\eta > 0$, the claim follows. □

6.4. Infection times

In this section we prove Theorem 6.4, which gives a precise formula for the infection time of an individual vertex. As outlined in Section 6.1.5, Theorem 6.4 is a straightforward consequence of the upper and lower bounds for the probability to be infected that are given in Theorem 6.20 and 6.17. However, due to the rather technical nature of these theorems, the proof is still a rather tedious calculation. We distinguish several cases as in the definition of $\Lambda(x_v, w_v)$, see Definition 6.3, the most relevant one being Case (III), cf. Remark 6.6.

Proof of Theorem 6.4. Let $v = (x_v, w_v)$ be an fixed vertex that satisfies the conditions in Theorem 6.4. Let $\varepsilon > 0$ be a constant and let $\eta = \eta(\varepsilon) > 0$ be a constant which is sufficiently small compared to ε, cf. Section 6.1.3.3.

First we remark that it suffices to show that for every sufficiently small $\varepsilon > 0$, whp $L_v = (1 \pm O(\varepsilon))\Lambda(x_v, w_v) \pm O(1)$, where the hidden constants are both independent of ε. Then by a standard diagonalizing argument, we also have whp $L_v = (1 \pm o(1))\Lambda(x_v, w_v) \pm O(1)$. We split the proof in three parts (I), (II) and (III), 'typical' vertices are treated in (III):

(I): Assume that $w_v > (\|x_v\|^d n)^{1/(\beta-1)}$ and the maximum in (6.4) is 0, i.e. we also have $w_v \geq \|x_v\|^d n/\nu$. In this case, the lower bound on L_v is trivial, so we show the upper bound. Since then the second term in the maximum must be non-positive, we have $\|x_v\|^d n \leq w_v \nu$. Let $i \geq 1$ be so large that $(\zeta-\varepsilon)^i/(\beta-1+\eta) > 1$, but observe that we may still choose $i = O(1)$.

Assume first $\|x_v\|^d n \leq \nu_i$, so $x_v \in B_i$. By the technical condition in Theorem 6.4 we have $\|x_v\|^d n \geq \widetilde{\nu}_0$, and hence $w_v \geq \widetilde{\nu}_0^{1/(\beta-1)} = \nu^{\Omega(1)} = \nu_i^{\Omega(1)}$. Hence, we may choose $\ell = O(1)$ such that $w_v \geq w_{i,\ell} = \nu_i^{(\zeta-\varepsilon)^{-\ell}/(\beta-1+\eta)}$, and it follows directly from part (c) of Theorem 6.17 that whp v is active after $i + \ell + 3 = O(1)$ rounds, as required.

On the other hand, if $\|x_v\|^d n \geq \nu_i$ then every vertex in B_i has distance at most $3\|x_v\|^d n$ from v. Let $w_i := \nu_i^{1/(\beta-1+\eta)}$. By Theorem 6.17, after $i+3$ rounds all vertices in $V_{\geq w_i} \cap B_i$ are active whp, and there are $\nu_i^{\Omega(\eta)} = \omega(1)$ many such vertices. Note that any such vertex has probability $\Omega(1)$ to form an edge with v, since $w_v w_i/(3^d \|x_v\|^d n) \geq w_i/(3^d \nu) = \Omega(1)$. Therefore, whp v is active in round $i + 4 = O(1)$, again as required.

(II): Assume that still $w_v > (\|x_v\|^d n)^{1/(\beta-1)}$, but that the maximum in (6.4) is attained by the second term, i.e. $\|x_v\|^d n/w_v \geq \nu$. We need to show an upper and a lower bound on L_v. For the upper bound, choose $i \geq 0$ minimal such that

$$\|x_v\|^d n/w_v \leq w_i, \tag{6.23}$$

where $w_i = \nu_i^{1/(\beta-1+\eta)} = \nu^{(\zeta-\varepsilon)^i/(\beta-1+\eta)}$ as before. Observe that this i satisfies

$$i = \log\log_\nu(\|x_v\|^d n/w_v)/\log(\zeta-\varepsilon) + O(1) \leq (1+O(\varepsilon))\Lambda(x_v, w_v) + O(1).$$

By Theorem 6.17, whp all vertices in $V_{\geq w_i} \cap B_i$ are active in round $i+3$, and there are $\omega(1)$ of them. As in (I), we discriminate two sub-cases.

Either $\|x_v\|^d n \geq \nu_i$. In this case, the distance from v to any point in B_i is at most $3\|x_v\|$, and v has probability $\Omega(1)$ to form an edge with each vertex in $V_{\geq w_i} \cap B_i$ by (6.23). By Theorem 6.17, whp all these vertices are active in round $i+3$, and there are $\omega(1)$ of them, so whp v will be active in round $i+4$.

Or $\|x_v\|^d n \leq \nu_i$, hence $x_v \in B_i$. Furthermore we have $w_v > (\|x_v\|^d n)^{1/(\beta-1)} \geq (w_{i-1} w_v)^{1/(\beta-1)}$ by minimality of i in (6.23), implying $w_v > w_{i-1}^{1/(\beta-2)} \geq w_i$. Thus by Theorem 6.17 whp v is active in round $i+3$. In either case, whp v is active in round $i + O(1)$, i.e. $L_v \leq i + O(1) \leq (1+O(\varepsilon))\Lambda(x_v, w_v) + O(1)$, as required.

For the lower bound, if $\|x_v\|^d n/w_v \leq \widetilde{\nu}_0 = \nu^{(\beta-1)/(\beta-2)+\varepsilon}$ then $\Lambda(x_v, w_v) = O(1)$, so there is nothing to show. Otherwise, $\|x_v\|^d n/w_v \geq \widetilde{\nu}_0 \geq \widetilde{\nu}_0^{1/(\beta-1-\eta)}$, so we

may choose $i \geq 0$ to be maximal such that

$$\|x_v\|^d n/w_v \geq \tilde{w}_i, \tag{6.24}$$

where $\tilde{w}_i = \widetilde{\nu}_i^{1/(\beta-1-\eta)} = \widetilde{\nu}_0^{(\zeta+\varepsilon)^i/(\beta-1-\eta)}$ as before. Note that this i satisfies

$$i = \log\log_\nu(\|x_v\|^d n/w_v)/\log(\zeta+\varepsilon) - O(1) \geq (1-O(\varepsilon))\Lambda(x_v, w_v) - O(1).$$

Let $\ell = O(1)$ be sufficiently large such that $\zeta^\ell > (\beta-1-\eta)$. If $i \leq \ell$, then $i = O(1)$, and there is nothing to show. Otherwise, (6.24) implies in particular

$$\|x_v\|^d n \geq \Omega(1)\cdot \widetilde{\nu}_i^{1/(\beta-1-\eta)} = \Omega(1)\cdot \widetilde{\nu}_{i-\ell}^{(\zeta+\varepsilon)^\ell/(\beta-1-\eta)} > 2\widetilde{\nu}_{i-\ell}.$$

Hence, by Lemma 6.14 (with $C = 2^{1/d} > 1$) we obtain

$$\mathbb{E}(N(v)\cap \widetilde{B}_{i-\ell}) = O(1)\widetilde{\nu}_{i-\ell}\left(\|x_v\|^d n/w_v\right)^{1-\beta+\eta}.$$

Using (6.24), we may continue

$$\mathbb{E}(N(v)\cap \widetilde{B}_{i-\ell}) = O(1)\widetilde{\nu}_{i-\ell}/\widetilde{\nu}_i = \widetilde{\nu}_i^{-\Omega(1)},$$

where the last step holds for any $\ell \geq 1$. By Markov's inequality, whp v has no neighbours in $\widetilde{B}_{i-\ell}$. On the other hand, by Theorem 6.20 whp there is no active vertex outside of $\widetilde{B}_{i-\ell}$ in round $i-\ell$. Therefore, whp v is not active in round $i-\ell+1$, i.e. $L_v > i-\ell+1 \geq (1-O(\varepsilon))\Lambda(x_v,w_v) - O(1)$, as required.

(III): Assume $w_v \leq (\|x_v\|^d n)^{1/(\beta-1)}$. We bound L_v from above and below.

For the *upper bound*, let $i \geq 0$ be minimal with the property that $x_v \in B_i$, i.e.

$$\|x_v\|^d n \leq \nu_i = \nu^{(\zeta-\varepsilon)^i} \tag{6.25}$$

Observe that $i \geq 1$ because $\|x_v\|^d n \geq \widetilde{\nu}_0 > \nu$ due to the technical assumption in Theorem 6.4 and since the function Λ is non-negative. Thus, we have $\|x_v\|^d n \geq \nu_{i-1} = \nu_i^{1/(\zeta-\varepsilon)}$ by minimality of i.

Let $\ell \geq 0$ be minimal with the property that

$$w_v > (\|x_v\|^d n)^{(\zeta-\varepsilon)^{-\ell}/(\beta-1+\eta)} \tag{6.26}$$

Since we are in the case $w_v \leq (\|x_v\|^d n)^{1/(\beta-1)}$, we have $\ell \geq 1$, and thus (6.26) is false if we replace ℓ by $\ell-1$. By minimality of i, the right hand side of (6.26) is at least $\nu_i^{(\zeta-\varepsilon)^{-\ell-1}/(\beta-1+\eta)}$ and recall that we only consider weights $w_v = \omega(1)$. Hence, Theorem 6.17 applies for $\ell+1$, and, if we condition on events that hold whp, tells us that v is active in round $i+\ell+4$ with probability

$$1 - \exp\left(-C_1\nu_i^{C_2(\zeta-\varepsilon)^{-\ell-1}}\right) = 1 - o(1),$$

where the last inequality holds due to the following estimate

$$\nu_i^{(\zeta-\varepsilon)^{-\ell-1}} \overset{(6.25)}{\geq} (\|x_v\|^d n)^{(\zeta-\varepsilon)^{-\ell-1}} \overset{(6.26),\ell-1}{\geq} w_v^{(\beta-1+\eta)/(\zeta-\varepsilon)^2} = \omega(1).$$

It remains to note that by choice of i and ℓ we have

$$i = \log\log_\nu(\|x_v\|^d n)/\log(\zeta-\varepsilon) + O(1)$$

and $\ell = i + \log\log_\nu w_v/\log(\zeta-\varepsilon) + O(1)$. Hence, whp

$$L_v \le i + \ell + 4 \le (1+O(\varepsilon))\Lambda(x_v, w_v) + O(1),$$

as required.

For the *lower bound*, we distinguish yet two more sub-cases. Let $\ell \ge 0$ be the smallest non-negative integer that satisfies

$$w_v^2 \ge 2^{-(\ell+1)d}(\|x_v\|^d n)^{(\zeta+\varepsilon)^{-\ell-3}/(\beta-1)}. \tag{6.27}$$

(IIIa) Assume first that $\|x_v\|^d n \ge 2^{\ell+1}\widetilde{\nu}_0$. In this case, let $i \ge 1$ be maximal with the property

$$\|x_v\|^d n \ge 2^{\ell+1}\widetilde{\nu}_{i-1} = 2^{\ell+1}\widetilde{\nu}_0^{(\zeta+\varepsilon)^{i-1}}. \tag{6.28}$$

It is easy to check (e.g., by using the very generous estimate $2 < \widetilde{\nu}_0^{(\zeta+\varepsilon)^i}$) that i satisfies

$$i \ge \log\log_\nu(\|x_v\|^d n)/\log(\zeta+\varepsilon) - O(\log(\ell+1)). \tag{6.29}$$

If $\ell = O(1)$ then $w_v = (\|x_v\|^d n)^{\Theta(1)}$ and therefore

$$\Lambda(x_v, w_v) = \log\log_\nu(\|x_v\|^d n)/|\log(\beta-2)| \pm O(1).$$

Since by Theorem 6.20(c) whp no vertex outside of $\widetilde{B}_{i-1}$ is active in round $i-1$ and $x_v \notin \widetilde{B}_{i-1}$, it follows then whp $L_v > i-1 \ge (1-O(\varepsilon))\Lambda(x_v,w_v) - O(1)$, as required. This settles the case $\ell = O(1)$.

Next observe that by minimality of i there exists $0 \le j \le O(\log(\ell+1))$ such that $\|x_v\|^d n \le \widetilde{\nu}_{i+j}$. In particular, if $\ell > C$ for some sufficiently large constant $C > 0$ then $j < \ell$. Since we have already treated the case $\ell = O(1)$, we may henceforth assume that $\ell > C$. Then $\ell - j > 0$, and by (6.28) the requirements of Theorem 6.20(f) are met for i and $\ell - j$. Since in particular $\ell \ge 1$, by the choice of ℓ, we have

$$\begin{aligned} w_v^2 &\le 2^{-\ell d}(\|x_v\|^d n)^{(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)} \le 2^{-\ell d}\widetilde{\nu}_{i+j}^{(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)} \\ &= 2^{-\ell d}\widetilde{\nu}_i^{(\zeta+\varepsilon)^{-(\ell-j)-2}/(\beta-1)}, \end{aligned} \tag{6.30}$$

and therefore Theorem 6.20(f) yields that v is not active in round $i+\ell-j$ with probability at least

$$1 - w_v 2^{(\ell-j)d}\widetilde{\nu}_i^{-(\zeta+\varepsilon)^{-(\ell-j)-2}/(\beta-1)} \overset{(6.30)}{\ge} 1 - w_v^{-1}2^{-jd} = 1 - o(1).$$

In order to relate $i+\ell-j$ with $\Lambda(x_v,w_v)$, we derive $\|x_v\|^d n \ge \widetilde{\nu}_0^{(\zeta+\varepsilon)^{i-1}}$ from (6.28), and plug it into (6.27) to obtain

$$w_v^2 \ge 2^{-(\ell+1)d}\widetilde{\nu}_0^{(\zeta+\varepsilon)^{i-\ell-4}/(\beta-1)} = 2^{-(\ell+1)d}\nu^{\Theta(1)(\zeta+\varepsilon)^{i-\ell}}.$$

Hence, taking logarithms on both sides,

$$\Theta(1)(\zeta+\varepsilon)^{i-\ell}\log\nu \le 2\log w_v + (\ell+1)d\log 2 \le 4d\max\{\log w_v, \ell\}. \tag{6.31}$$

If the maximum is attained by $\log w_v$, then (6.31) gives $\ell \ge i - \log\log_\nu w_v/\log(\zeta+\varepsilon) - O(1)$, and together with (6.29) and $j = O(\log(1+\ell))$, we conclude $i+\ell-j \ge$

$(1 - O(\varepsilon))\Lambda(x_v, w_v) - O(1)$, as required. On the other hand, if the maximum in (6.31) is attained by ℓ, then (6.31) yields

$$\ell + \frac{\log \ell}{\log(\zeta+\varepsilon)} \geq i + \frac{\log\log \nu}{\log(\zeta+\varepsilon)} - O(1) \geq i - \frac{\log\log_\nu w_v}{\log(\zeta+\varepsilon)} - O(1),$$

where the second inequality comes from $w_v = \omega(1)$. Thus we obtain again $i+\ell-j \geq (1-O(\varepsilon))\Lambda(x_v, w_v) - O(1)$, as required. This concludes the proof of the lower bound in the case $\|x_v\|^d n \geq 2^{\ell+1}\widetilde{\nu}_0$.

(IIIb) Assume $\|x_v\|^d n \leq 2^{\ell+1}\widetilde{\nu}_0$. It remains to show the lower bound on L_v in this case. We want to apply Theorem 6.20(f) for $i = 0$, but we need to change the definition of ℓ slightly. Let $\ell' \geq 0$ be the smallest non-negative integer satisfying

$$w_v^2 \geq 2^{-(\ell'-1)d}\widetilde{\nu}_0^{(\zeta+\varepsilon)^{-\ell'-1}/(\beta-1)}. \tag{6.32}$$

Similar as in (IIIa), this definition implies

$$\ell' + \log(\ell'+1)/\log(\zeta+\varepsilon) \geq -\log\log_\nu w_v/\log(\zeta+\varepsilon) - O(1). \tag{6.33}$$

If $\ell' \leq 1$ then $w_v = \widetilde{\nu}_0^{\Omega(1)}$. In this case, since $\|x_v\|^d n \leq 2^{\ell+1}\widetilde{\nu}_0$, a sufficient condition for ℓ to satisfy (6.27) is

$$\widetilde{\nu}_0^{\Omega(1)} \geq 2^{-(\ell+1)d}(2^{\ell+1}\widetilde{\nu}_0)^{(\zeta+\varepsilon)^{-\ell-3}/(\beta-1)}.$$

Since this is already satisfied for some large enough constant, by the definition of ℓ, this implies $\ell = O(1)$ and thus $\|x_v\|^d n = \widetilde{\nu}_0^{O(1)}$, and the lower bound is trivial, because $\Lambda(x_v, w_v) = O(1)$.

So assume instead that $\ell' > 1$. Let $\ell^* := \min\{\ell' - 1, \Lambda(x_v, w_v)\}$. Then by the assumption in Theorem 6.4, we have $\ell^* \leq \log_2(\|x_v\|^d n/\widetilde{\nu}_0)$, and hence $\|x_v\|^d n \geq 2^{\ell^*}\widetilde{\nu}_0$. Since $\ell^* < \ell'$, the reverse of (6.32) holds for ℓ^*. These two properties allow us to apply Theorem 6.20(f) with $i = 0$ and $\ell^* - 1$, which tells us that v is not active in round $\ell^* - 1$ with probability at least

$$1 - w_v 2^{(\ell^*-1)d}\widetilde{\nu}_0^{-(\zeta+\varepsilon)^{\ell^*-1}/(\beta-1)} \geq 1 - w_v^{-1} = 1 - o(1).$$

It remains to show the minimum in the definition of ℓ^* is attained by the second term, more precisely $\ell^* = \Lambda(x_v, w_v) + O(1)$. Next observe that by (6.33) it is sufficient to deduce $\log\log_\nu(\|x_v\|^d n) = o(\ell') + O(1)$ in order for this claim to hold. Since $\|x_v\|^d n \leq 2^{\ell+1}\widetilde{\nu}_0$, we have $\log_\nu(\|x_v\|^d n) \leq o(\ell) + O(1)$. However, by the choice of ℓ, we have

$$w_v^2 \leq 2^{-\ell d}(\|x_v\|^d n)^{(\zeta+\varepsilon)^{-\ell-2}/(\beta-1-\eta)} \leq (2^{\ell+1}\widetilde{\nu}_0)^{(\zeta+\varepsilon)^{-\ell-2}/(\beta-1)},$$

and similar as for (6.28) it can be easily deduced that $\ell = O(1 + |\log\log_\nu w_v|) = O(\ell')$. This concludes the proof for the case $\|x_v\|^d n \leq 2^\ell \widetilde{\nu}_0$. □

6.5. Threshold and speed of the process

We prove Theorem 6.1 and Theorem 6.2 together. Let $\varepsilon > 0$ be a constant and let $\eta = \eta(\varepsilon) > 0$ be a constant which is sufficiently small compared to ε, cf. Section 6.1.3.3.

We first show the second statement of Theorem 6.2.

Claim 6.21. *Assume that $\alpha > \beta - 1$ and $\nu = n^{o(1)}$, then $|V^{\leq (1-\varepsilon)i_\infty}| = o(n)$ whp.*

Proof. Let i_0 be (somewhat arbitrary) the largest integer such that $\widetilde{\nu}_{i_0-1} \leq n/\log^2 n$ and note that then $i_0 \geq (1 - O(\varepsilon))(\log\log_\nu n)/|\log(\beta-2)|$. Moreover, for $i_1 := (1-\varepsilon)(\log\log n)/|\log(\beta-2)|$ we have $2^{i_1}\widetilde{\nu}_{i_0-1} = o(n)$, so whp there are $o(n)$ vertices in $2^{i_1}\widetilde{B}_{i_0-1}$.

Next consider the vertices outside of $2^{i_1}\widetilde{B}_{i_0-1}$. By Theorem 6.20 each such vertex of weight at most $\log\log n$ has probability $o(1)$ to be in $V^{\leq i_0+i_1}$. Therefore, the expected number of vertices of weight at most $\log\log n$ in $V^{\leq i_0+i_1}$ is $o(n)$. On the other hand, the total expected number of vertices of weight larger than $\log\log n$ is also $o(n)$. Altogether, this shows $\mathbb{E}(|V^{\leq i_0+i_1}|) = o(n)$, and the statement follows from Markov's inequality. □

6.5.1. **Subcritical regime: (iii), (v).** We will indeed show that whp the process does not infect any vertices in the first step.

Claim 6.22. *$V^{\leq 1} = V^{\leq 0}$ whp.*

Proof. For any vertex $v = (x_v, w_v)$ with fixed weight and position we denote by $\mu_v := \mathbb{E}(|N(v) \cap B_0|)$ its expected number of neighbours in B_0. We have shown in Lemma 6.14 that for any constant $C > 1$,

$$\mu_v = O(\nu) \cdot \begin{cases} \min\{w_v/\nu, 1\}, & \text{if } \|x_v\| \leq C(\nu/n)^{1/d}/2, \\ \min\left\{\left(w_v/(\|x_v\|^d n)\right)^m, 1\right\} & \text{if } \|x_v\| \geq C(\nu/n)^{1/d}/2, \end{cases} \tag{6.34}$$

where $m = \min\{\alpha, \beta - 1 - \eta\} > 1$. Since initially only vertices in B_0 are activated, recall that the number $N^{\leq 0}(v)$ of *initially active* neighbours of v is Poisson distributed with mean $\rho\mu_v$. In particular, $\mathbb{P}(|N^{\leq 0}(v)| \geq k) = \mathbb{P}(\mathrm{Po}(\rho\mu_v) \geq k) = O(1) \cdot \min\{(\rho\mu_v)^k, 1\}$. Clearly, we can bound the number $|V^{=1}|$ of vertices that turn active in round 1 by the number of vertices that have at least k neighbours in $V^{\leq 0}$. (It is only an upper bound since the latter also counts vertices which were already in $V^{\leq 0}$.)

So let us first consider the contribution $n_{\mathrm{in}} := |V^{=1} \cap 2B_0|$ of vertices $v = (x_v, w_v)$ inside of $2B_0$. By (6.34) these satisfy $\mu_v = O(w_v)$, and thus by Lemmas 6.10 and 6.11 we obtain

$$\mathbb{E}(n_{\mathrm{in}}) = O(1)\int_0^\infty \nu w^{1-\beta+\gamma}\frac{d}{dw}\min\{(\rho w)^k, 1\}dw = O(\nu\rho^{\beta-1-\gamma}) = o(1), \tag{6.35}$$

where $\gamma = 0$ for strong power-laws, and otherwise $\gamma > 0$ is an arbitrary constant.

On the other hand, to estimate the contribution $n_{\mathrm{out}} := |V^{=1} \cap (\mathbb{T}^d \setminus 2B_0)|$ of vertices $v = (x_v, w_v)$ outside of $2B_0$, we may use $\mu_v = O(\nu)(w_v/(\|x_v\|^d n))^m$

by (6.34). Furthermore, since each such vertex has distance at least $(\nu/n)^{1/d}$ from the origin, Lemma 6.10 and Lemma 6.11 imply

$$\begin{aligned}\mathbb{E}(n_{\text{out}}) &= O(1)\int_{(\nu/n)^{1/d}}^{\infty} r^{d-1}n\int_0^{\infty} w^{1-\beta+\eta}\frac{d}{dw}\min\left\{\left(\rho\nu\left(\frac{w}{r^d n}\right)^m\right)^k,1\right\}dwdr\\ &= O(1)\int_{(\nu/n)^{1/d}}^{\infty} r^{d-1}n\left(\frac{r^d n}{(\rho\nu)^{1/m}}\right)^{1-\beta+\eta}dr\\ &= O(1)\cdot\nu^{2-\beta+\eta}(\rho\nu)^{(\beta-1+\eta)/m}\end{aligned}$$

Now we use that $\rho = O(\nu^{-1/(\beta-1)})$. Observe that this bound holds both in case (iii) and (v), and that it even holds for the critical case (ii). We derive $\rho\nu = O(\nu^{(\beta-2)/(\beta-1)})$, and hence $\mathbb{E}[n_{\text{out}}] = \nu^{-(\beta-2)(1-1/m)+O(\eta)}$. Thus, since $m > 1$, if $\eta > 0$ is small enough we have

$$\mathbb{E}(n_{\text{out}}) = o(1). \tag{6.36}$$

We will later use the fact that this also holds in the critical regime (ii).

Together (6.35) and (6.36) show that $\mathbb{E}(|V^{=1}|) = o(1)$, and thus by Markov's inequality whp no vertices turned active in round 1, as claimed. □

6.5.2. **Critical regime: (ii).** We first show that with constant probability no further vertices are ever activated.

Claim 6.23. *$V^{\le 1} = V^{\le 0}$ with probability $\Omega(1)$.*

Proof. First observe that (6.36) also holds in this regime, i.e. by Markov's inequality whp no vertex outside of $2B_0$ is active in round 1. Furthermore, let $c > 0$ be a (small) constant, to be determined later, and let $w_0 := \nu^{1/(\beta-1)}$. Moreover, note that $|V_{\ge cw_0}\cap 2B_0|$ is Poisson distributed with mean $O(\nu(cw_0)^{1-\beta}) = O(1)$, since $c = \Omega(1)$. Therefore the event $\mathcal{A} := \{V_{\ge cw_0}\cap 2B_0 = \emptyset\}$ occurs with probability at least $\exp(-O(1)) = \Omega(1)$. Consequently it suffices to show that $V^{\le 1}_{\le cw_0}\cap 2B_0 = V^{\le 0}$ with probability $\Omega(1)$ if we condition on the event $\mathcal{A}$.

Since every vertex $v = (x_v, w_v)$ satisfies $\mathbb{E}(|N(v)\cap B_0|) \le \mathbb{E}(|N(v)|) = O(w_v)$, by (6.9), the number of neighbours in $V^{\le 0}$ is a Poisson distributed random variable with mean at most $O(\rho w_v)$. Observe that this upper bound remains valid if we condition on the event $\mathcal{A}$, since this can only decrease the expected degree of v. Therefore we obtain

$$\mathbb{P}\left(|N^{\le 0}(v)| \ge k \;\middle|\; \mathcal{A}\right) = O(1)\cdot\min\{(\rho w_v)^k, 1\} = O((\rho w_v)^k),$$

and by Lemma 6.10 it follows that

$$\begin{aligned}\mathbb{E}\left(|V^{=1}_{\le cw_0}\cap 2B_0| \;\middle|\; \mathcal{A}\right) &= O\left(\nu(cw_0)^{1-\beta}(\rho cw_0)^k + \int_0^{cw_0}\nu w^{1-\beta}\frac{d}{dw}(\rho w)^k dw\right)\\ &= O(1)\cdot\nu(cw_0)^{1-\beta}(\rho cw_0)^k = O(c^{k+1-\beta}),\end{aligned}$$

where all the hidden constants are independent of c.

Now note that we may choose the constant $c > 0$ small enough such that $\mathbb{E}\left(|V^{=1}_{\le cw_0}\cap 2B_0| \;\middle|\; \mathcal{A}\right) \le 1/2$, and then by Markov's inequality $|V^{=1}_{\le cw_0}\cap 2B_0| = 0$

with conditional probability at least $1/2$. Thus $V^{\leq 1} = V^{\leq 0}$ with probability $\Omega(1)$, and the claim follows. □

Next we show that with constant probability at least k heavy vertices will be activated in the first round. Afterwards, the remaining steps will be identical with the supercritical regime, so we prove them together, cf. below.

Claim 6.24. *$|V^{=1}_{\geq w_0} \cap B_0| \geq k$ with probability $\Omega(1)$.*

Proof. We first consider the case $\alpha < \infty$. Note that the number of vertices in B_0 of weight at least $w_0 = \nu^{1/(\beta-1)}$ is Poisson distributed with mean $\Theta(\nu w_0^{\beta-1}) = \Theta(1)$. In particular, with probability $\Omega(1)$ there are at least k such vertices. So assume this event holds, and let $v_1, \dots, v_k$ be k distinct such vertices.

For each $1 \leq i \leq k$, denote by K_i be the intersection of B_0 with the ball of volume $\nu^{1/(\beta-1)}/n$ around v_i. Note that $n\mathrm{Vol}(K_i) = \Omega(\nu^{1/(\beta-1)})$. The number of vertices in $V^{\leq 0} \cap K_i$ is Poisson distributed with mean $\rho n \mathrm{Vol}(K_i)$, so in particular $\mathbb{P}(|V^{\leq 0} \cap K_i| \geq k) = \Omega(1)$. Note that the events $\mathcal{K}(i) := \{|V^{\leq 0} \cap K_i| \geq k\}$ are positively associated for different i, i.e, conditioning on the events $\mathcal{K}(i_1), \dots, \mathcal{K}(i_s)$ does not decrease the probability of $\mathcal{K}(i)$ for any subset of (distinct) indices $i_1, \dots, i_s, i$. Hence,

$$\mathbb{P}\left(\forall i \in \{1, \dots k\} : \mathcal{K}(i)\right) \geq \prod_{i=1}^{k} \mathbb{P}(\mathcal{K}(i)) = \Omega(1)$$

by the law of conditional probability.

On the other hand, for each vertex v with fixed weight and fixed position in K_i we have $\mathbb{P}(\{v, v_i\} \in E) = \Omega(w_0/\nu^{1/(\beta-1)}) = \Omega(1)$, and this lower bound is independent for any two such vertices. So conditioned on the events $\mathcal{K}(i)$, we have $\mathbb{P}(v_i \in V^{\leq 1} \mid \mathcal{K}(1), \dots, \mathcal{K}(k)) = \Omega(1)$, and this lower bound is independent for all i. Altogether, we have shown that with probability $\Omega(1)$ we have $\{v_1, \dots, v_k\} \subset V^{=1}_{\geq w_0} \cap B_0$, proving the claim in the case of $\alpha > \infty$.

The case $\alpha = \infty$ is completely analogous, except that it may be necessary to shrink the balls around $v_1, \dots, v_k$ be at most a constant factor, so that still every vertex in the i-th ball has probability $\Omega(1)$ to connect to v_i. Since this only decreases the expected number of (active) vertices in each ball by constant factors, the remaining proof stays the same. We omit the details. □

6.5.3. **Supercritical regime: (i), (iv).** In this proof we also include the critical regime (ii), provided that at least k heavy vertices got activated in the first round, i.e. $|V^{=1}_{\geq w_0} \cap B_0| \geq k$, where as before $w_0 = \nu^{1/(\beta-1)}$. Let $\varepsilon > 0$ be constant.

Claim 6.25. *$\mathbb{E}(|V^{\leq (1+\varepsilon) i_\infty}|) = \Omega(n)$.*

Proof. Let $i \geq 0$ be the smallest index such that $\nu_i \geq n$ (and thus $B_i = \mathbb{T}^d$), and note that $i \leq (1+\varepsilon)(\log\log_\nu n)/|\log(\beta-2)|$ if n is sufficiently large. Then there exists $\ell \leq (1+\varepsilon)(\log\log n)/|\log(\beta-2)|$ such that $w_{i,\ell} = O(1)$. Theorem 6.17 tells us immediately that every vertex of weight at least C has probability $\Omega(1)$ to be in

$A(i+3+\ell)$, for some sufficiently large constant $C > 0$. Since the expected number of vertices of weight at least C is $\Omega(n)$, this already shows the claim. □

It remains to proof the corresponding whp statement.

Claim 6.26. $|V^{\leq(1+\varepsilon)i_\infty}| = \Omega(n)$ *whp.*

Proof. Let $h = h(n) = \omega(1)$ be a function with $\log\log h = o(\log\log n)$. Then by the same argument as before, every vertex of weight at least $w_h := \hat{w}_-(h) = h^{1/(\beta-1+\eta)}$ has probability $1 - h^{-\Omega(1)}$ to be in $V^{\leq(1+\varepsilon/2)i_\infty}$. Now decompose the torus $\mathbb{T}^d$ into balls $Q_1, \dots, Q_s$ of volume $\Theta(h/n)$.[7] Fix any such ball Q, and call Q *good* if in round $j := (1+\varepsilon/2)i_\infty$ at least half of the vertices in $V_{\geq w_h} \cap Q$ are active, and *bad* otherwise. Recall that in expectation only a $o(1)$ fraction of the vertices in $V_{\geq w_h} \cap Q$ are inactive in round j, so by Markov's inequality the probability that Q is bad is $o(1)$. So in expectation only a $o(1)$ fraction of the sets $Q_1, \dots, Q_s$ are bad, and again by Markov's inequality, whp at least half of them are good.

Now let $C > 0$ be a sufficiently large constant, and assume Q is good. Then we may apply Remark 6.18 (for some suitably chosen $0 \leq \ell = O(\log\log h)$) to deduce that an expected 2/3-fraction of the vertices in $V_{\geq C} \cap Q$ are active in round $j+\ell$, if $C > 0$ is sufficiently large. In formula, $\mathbb{E}(|V_{\geq C}^{\leq j+\ell} \cap Q|) \geq 2\mathbb{E}(|V_{\geq C} \cap Q|)/3$, and thus by Markov's inequality,

$$\mathbb{P}\left(\left|V_{\geq C}^{\leq j+\ell} \cap Q\right| \geq \mathbb{E}\left(\left|V_{\geq C} \cap Q\right|\right)/2\right) = \Omega(1),$$

and this bound holds independently of the activity in any other ball. Therefore, by a Chernoff bound, whp an $\Omega(1)$ fraction of the balls $Q_1, \dots, Q_s$ satisfy $|V_{\geq C}^{\leq j+\ell} \cap Q| \geq \mathbb{E}\left(\left|V_{\geq C} \cap Q\right|\right)/2 = \Omega(h)$, and thus whp $|V^{\leq j+\ell}| = \Omega(s \cdot h) = \Omega(n)$. Since $j + \ell \leq (1+\varepsilon)i_\infty$ for sufficiently large n, the claim follows. □

Proof of Theorem 6.1 and Theorem 6.2. Theorem 6.1 follows immediately by combining Claims 6.22, 6.23, 6.24 and 6.26, while Theorem 6.2 is proven by Claims 6.21 and 6.26. □

6.6. Concluding remarks

We have shown that in the GIRG model for scale-free networks with underlying geometry, even a small region can cause an infection that spreads through a linear part of the population. We have analysed the process in great detail, and we have determined its metastability threshold, its speed, and the time at which individual vertices becomes infected. Moreover, we have shown how a policy-maker can utilise this knowledge to enforce a successful quarantine strategy.

We want to emphasise that the latter result is only a proof of concept, intended to illustrate the possibilities that come from a thorough understanding of the role of the underlying geometry in infection processes. In particular, we want to remind the reader that bootstrap percolation is not a perfect model for viral infections

[7] This is possible since we use the ∞-norm. It would also suffice to consider any disjoint balls with total volume $\Omega(1)$.

(though it has been used to this end), but is more adequate for processes in which the probability of transmission grows more than proportional if more than one neighbours is active, like believes spreading through a social network ('What I tell you three times is true.'), or action potential spreading through a neuronal network.

Therefore, this chapter is only a first step. There are many other models for the spread of an infection, most notably SIR and SIRS models for epidemiological applications, and we have much yet to learn from analysing these models in geometric power-law networks like GIRGs. From a technical point of view, it is unsatisfactory that our analysis does not include the case $\alpha \leq \beta - 1$. We believe that also in this case, the bootstrap percolation process is essentially governed by the geometry of the underlying space, only in a more complex way. Understanding this case would probably also add to our toolbox for analysing less 'clear-cut' processes.

Acknowledgements

First and foremost, I want to thank Mihyun Kang for supervising the first step of my scientific journey wholeheartedly – by constantly challenging me to approach new open problems and always finding time for extensive mathematical arguments. Besides, I highly appreciate her providing me with honest and constructive criticism while simultaneously showing an uninterrupted confidence in my capabilities concerning both research and teaching.

Moreover, I am grateful for the mathematical insights obtained in the intense discussions with my collaborators: from each of them I learnt a lot and some of these ideas eventually led to the results contained in this thesis.

Furthermore, I want to thank my colleagues and friends at TU Graz and within the doctoral program 'Discrete Mathematics' for creating an enjoyable yet productive scientific environment. In the last years, these institutions provided great opportunities for me, such as a 3-month research stay at ETH Zürich and travelling to a number of renowned international conferences in the field.

Above all, I am deeply grateful to my friends and family for supporting me in pursuing my dreams and backing all of my decisions without the slightest hesitation.

Bibliography

[11] J. Adler and U. Lev. 'Bootstrap percolation: visualizations and applications'. *Brazilian Journal of Physics* **33** (2003), no. 3, 641–644.

[12] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat. 'A spatial web graph model with local influence regions'. *Internet Mathematics* **5** (2008), no. 1-2, 175–196.

[13] M. Aizenman and J. Lebowitz. 'Metastability effects in bootstrap percolation'. *Journal of Physics A: Mathematical and General* **21** (1988), no. 19, P3801.

[14] R. Albert and A.-L. Barabási. 'Statistical mechanics of complex networks'. *Reviews of modern physics* **74** (2002), no. 1, P47.

[15] N. Alon and J. Spencer. *The probabilistic method*. Third. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2008, 352 pages. ISBN: 9780470170205.

[16] H. Amini. 'Bootstrap percolation in living neural networks'. *Journal of Statistical Physics* **141** (2010), no. 3, 459–475.

[17] H. Amini, R. Cont, and A. Minca. 'Resilience to contagion in financial networks'. *Mathematical Finance* **26** (2016), no. 2, 329–365.

[18] H. Amini and N. Fountoulakis. 'Bootstrap percolation in power-law random graphs'. *Journal of Statistical Physics* **155** (2014), no. 1, 72–92.

[19] K. Athreya and P. Ney. *Branching processes*. Dover Publications, Inc., Mineola, NY, 2004, 287 pages. ISBN: 9780486434742.

[20] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. 'The sharp threshold for bootstrap percolation in all dimensions'. *Trans. Amer. Math. Soc.* **364** (2012), no. 5, 2667–2701.

[21] J. Balogh, Y. Peres, and G. Pete. 'Bootstrap percolation on infinite trees and non-amenable groups'. *Combinatorics, Probability and Computing* **15** (2006), no. 05, 715–730.

[22] A.-L. Barabási. 'Network science: luck or reason'. *Nature* **489** (2012), no. 7417, 507–508.

[23] A. Barbour, L. Holst, and S. Janson. *Poisson approximation*. Vol. 2. Oxford Studies in Probability. The Clarendon Press, Oxford University Press, New York, 1992, 277 pages. ISBN: 9780198522355.

[24] G. Baxter, S. Dorogovtsev, A. Goltsev, and J. Mendes. 'Bootstrap percolation on complex networks'. *Physical Review E* **82** (2010), no. 1, P011103.

[25] M. Behrisch, A. Coja-Oghlan, and M. Kang. 'Local limit theorems for the giant component of random hypergraphs'. *Combin. Probab. Comput.* **23** (2014), no. 3, 331–366.

[26] M. Behrisch, A. Coja-Oghlan, and M. Kang. 'The order of the giant component of random hypergraphs'. *Random Struct. & Alg.* **36** (2010), no. 2, 149–184.

[27] S. Bhamidi, N. Broutin, S. Sen, and X. Wang. 'Scaling limits of random graph models at criticality: universality and the basin of attraction of the Erdős-Rényi random graph'. *Preprint* (2014), arXiv: 1411.3417.

[28] M. Boguñá, D. Krioukov, and K. Claffy. 'Navigability of complex networks'. *Nature Physics* **5** (2009), no. 1, 74–80.

[29] M. Boguñá, F. Papadopoulos, and D. Krioukov. 'Sustaining the internet with hyperbolic mapping'. *Nature communications* **1** (2010), P62.

[30] B. Bollobás. *Random graphs*. Second. Vol. 73. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2001, 498 pages. ISBN: 9780511814068.

[31] B. Bollobás. 'The evolution of random graphs'. *Trans. Amer. Math. Soc.* **286** (1984), no. 1, 257–274.

[32] B. Bollobás. 'The evolution of sparse graphs'. In: *Graph theory and combinatorics (Cambridge, 1983)*. Academic Press, London, 1984, 35–57.

[33] B. Bollobás, S. Janson, and O. Riordan. 'The phase transition in inhomogeneous random graphs'. *Random Struct. & Alg.* **31** (2007), no. 1, 3–122.

[34] B. Bollobás, S. Janson, and O. Riordan. 'The phase transition in the uniformly grown random graph has infinite order'. *Random Struct. & Alg.* **26** (2005), no. 1-2, 1–36.

[35] B. Bollobás and O. Riordan. 'A simple branching process approach to the phase transition in $G_{n,p}$'. *Electron. J. Combin.* **19** (2012), no. 4, P21.

[36] B. Bollobás and O. Riordan. 'Asymptotic normality of the size of the giant component in a random hypergraph'. *Random Struct. & Alg.* **41** (2012), no. 4, 441–450.

[37] B. Bollobás, O. Riordan, E. Slivken, and P. Smith. 'The threshold for jigsaw percolation on random graphs'. *Preprint* (2015), arXiv: 1503.05186.

[38] B. Bollobás and A. Thomason. 'Random graphs of small order'. In: *Random graphs '83 (Poznań, 1983)*. Vol. 118. North-Holland Math. Stud. North-Holland, Amsterdam, 1985, 47–97.

[39] M. Bradonjić and I. Saniee. 'Bootstrap percolation on random geometric graphs'. *Probability in the Engineering and Informational Sciences* **28** (2014), no. 2, 169–181.

[40] K. Bringmann, R. Keusch, and J. Lengler. 'Average distance in a general class of scale-free graphs with underlying geometry'. *Preprint* (2016), arXiv: 1602.05712.

[41] K. Bringmann, R. Keusch, and J. Lengler. 'Geometric inhomogeneous random graphs'. *Preprint* (2015), arXiv: 1511.00576.

[42] T. Britton, M. Deijfen, and A. Martin-Löf. 'Generating simple random graphs with prescribed degree distribution'. *J. Stat. Phys.* **124** (2006), no. 6, 1377–1397.

[43] C. Brummitt, S. Chatterjee, P. Dey, and D. Sivakoff. 'Jigsaw percolation: what social networks can collaboratively solve a puzzle?' *Ann. Appl. Probab.* **25** (2015), no. 4, 2013–2038.

[44] E. Candellero and N. Fountoulakis. 'Bootstrap percolation and the geometry of complex networks'. *Stochastic Processes and their Applications* **126** (2016), no. 1, 234–264.

[45] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir. 'A model of internet topology using k-shell decomposition'. *Proceedings of the National Academy of Sciences* **104** (2007), no. 27, 11150–11154.

[46] D. Centola. 'The spread of behavior in an online social network experiment'. *Science* **329** (2010), no. 5996, 1194–1197.

[47] J. Chalupa, P. Leath, and G. Reich. 'Bootstrap percolation on a Bethe lattice'. *Journal of Physics C: Solid State Physics* **12** (1979), no. 1, L31–L35.

[48] F. Chung and L. Lu. 'Connected components in random graphs with given expected degree sequences'. *Ann. Comb.* **6** (2002), no. 2, 125–145.

[49] F. Chung and L. Lu. 'The volume of the giant component of a random graph with given expected degrees'. *SIAM J. Discrete Math.* **20** (2006), no. 2, 395–411.

[50] O. Cohen, A. Keselman, E. Moses, R. Martínez, J. Soriano, and T. Tlusty. 'Quorum percolation in living neural networks'. *EPL (Europhys. Lett.)* **89** (2010), no. 1, P18008.

[51] O. Cooley, M. Kang, and Y. Person. 'Largest components in random hypergraphs'. *Preprint* (2014), arXiv: 1412.6366.

[52] C. Cooper, A. Frieze, and P. Prałat. 'Some typical properties of the spatial preferred attachment model'. *Internet Math.* **10** (2014), no. 1-2, 116–136.

[53] S. Dorogovtsev, A. Goltsev, and J. Mendes. 'K-core organization of complex networks'. *Physical review letters* **96** (2006), no. 4, P040601.

[54] S. Dorogovtsev and J. Mendes. 'Evolution of networks'. *Advances in Physics* **51** (2002), no. 4, 1079–1187.

[55] P. Dreyer and F. Roberts. 'Irreversible k-threshold processes: graph theoretical threshold models of the spread of disease and of opinion'. *Discrete Applied Mathematics* **157** (2009), no. 7, 1615–1627.

[56] H. Einarsson, J. Lengler, K. Panagiotou, F. Mousset, and A. Steger. 'Bootstrap percolation with inhibition'. *Preprint* (2014), arXiv: 1410.3291.

[57] H. Einarsson, J. Lengler, and A. Steger. 'A high-capacity model for one shot association learning in the brain'. *Front. Comput. Neurosci.* **8** (2014), no. 140.

[58] P. Erdős and A. Rényi. 'On the evolution of random graphs'. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61.

[59] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge: Cambridge University Press, 2009, 810 pages. ISBN: 9780521898065.

[60] P. Flocchini, E. Lodi, F. Luccio, L. Pagli, and N. Santoro. 'Dynamic monopolies in tori'. *Discrete applied mathematics* **137** (2004), no. 2, 197–212.

[61] J. Gao, T. Zhou, and Y. Hu. 'Bootstrap percolation on spatial networks'. *Scientific Reports* **5** (2015), P14662.

[62] A. Gersho and D. Mitra. 'A simple growth model for the diffusion of a new communication service'. *IEEE Trans. Systems, Man, and Cybernetics* **5** (1975), no. 2, 209–216.

[63] M. Granovetter. 'Threshold models of collective behavior'. *Am. J. Sociol.* **83** (1978), no. 6, 1420–1443.

[64] M. S. Granovetter. 'The strength of weak ties'. *Am. J. Sociol.* **78** (1973), no. 6, 1360–1380.

[65] J. Gravner and D. Sivakoff. 'Nucleation scaling in jigsaw percolation'. *Preprint* (2013), arXiv: 1310.2194.

[66] R. Griffin and C. Nunn. 'Community structure and the spread of infectious disease in primate social networks'. *Evolutionary Ecology* **26** (2012), no. 4, 779–800.

[67] A. Gutiérrez Sánchez. 'Multi-colored jigsaw percolation on random graphs'. MA thesis. Graz University of Technology, 2016.

[68] T. Harris. *The theory of branching processes*. Dover phoenix editions. Dover Publications, 2002, 230 pages. ISBN: 9780486495088.

[69] S. Hawking and L. Mlodinow. *A briefer history of time*. Bantam Books, 2005, 176 pages. ISBN: 9780553804362.

[70] R. van der Hofstad. 'Critical behavior in inhomogeneous random graphs'. *Random Struct. & Alg.* **42** (2013), no. 4, 480–508.

[71] W. Huang and C. Li. 'Epidemic spreading in scale-free networks with community structure'. *Journal of Statistical Mechanics: Theory and Experiment* (2007), no. 1, P01014.

[72] E. Jacob and P. Mörters. 'Spatial preferential attachment networks: power laws and clustering coefficients'. *The Annals of Applied Probability* **25** (2015), no. 2, 632–662.

[73] S. Janson, D. Knuth, T. Łuczak, and B. Pittel. 'The birth of the giant component'. *Random Struct. & Alg.* **4** (1993), no. 3, 231–358.

[74] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000, 333 pages. ISBN: 9780471175414.

[75] S. Janson, T. Łuczak, T. Turova, and T. Vallier. 'Bootstrap percolation on the random graph $G_{n,p}$'. *Ann. Appl. Probab.* **22** (2012), no. 5, 1989–2047.

[76] J. Janssen and A. Mehrabian. 'Rumours spread slowly in a small world spatial network'. In: *Algorithms and models for the web graph*. Vol. 9479. Lecture Notes in Computer Science. Springer, Cham, 2015, 107–118.

[77] M. Kahle and B. Pittel. 'Inside the critical window for cohomology of random k-complexes'. *Random Struct. & Alg.* **48** (2016), no. 1, 102–124.

[78] A. Karbasi, J. Lengler, and A. Steger. 'Normalization phenomena in asynchronous networks'. In: *Automata, languages, and programming. Part II*. Vol. 9135. Lecture Notes in Computer Science. Springer, Heidelberg, 2015, 688–700.

[79] M. Karoński and T. Łuczak. 'The phase transition in a random hypergraph'. *J. Comput. Appl. Math.* **142** (2002), no. 1, 125–135.

[80] S. Kirkpatrick, W. Wilcke, R. Garner, and H. Huels. 'Percolation in dense storage arrays'. *Physica A: Statistical Mechanics and its Applications* **314** (2002), no. 1, 220–229.

[81] R. Kozma. 'Neuropercolation'. *Scholarpedia* **2** (2007), no. 8, 1360.

[82] M. Krivelevich and B. Sudakov. 'The phase transition in random graphs: a simple proof'. *Random Struct. & Alg.* **43** (2013), no. 2, 131–138.

[83] R. Lambiotte and P. Panzarasa. 'Communities, knowledge creation, and information diffusion'. *Journal of Informetrics* **3** (2009), no. 3, 180–190.

[84] L. Lu and X. Peng. 'High-order Phase Transition in Random Hypergraphs'. *Preprint* (2014), arXiv: 1409.1174.

[85] M. Luczak and T. Łuczak. 'The phase transition in the cluster-scaled model of a random graph'. *Random Struct. & Alg.* **28** (2006), no. 2, 215–246.

[86] T. Łuczak. 'Component behavior near the critical point of the random graph process'. *Random Struct. & Alg.* **1** (1990), no. 3, 287–310.

[87] M. Molloy. 'Cores in random hypergraphs and Boolean formulas'. *Random Struct. & Alg.* **27** (2005), no. 1, 124–135.

[88] C. Moukarzel and T. Sokolowski. 'Long-range k-core percolation'. *Journal of Physics: Conference Series* **246** (2010), no. 1, P012019.

[89] A. Nachmias and Y. Peres. 'The critical random graph, with martingales'. *Israel J. Math.* **176** (2010), 29–41.

[90] M. Newman. 'Random graphs as models of networks'. In: *Handbook of graphs and networks*. Wiley-VCH, Weinheim, 2003, 35–68.

[91] I. Norros and H. Reittu. 'On a conditionally Poissonian graph process'. *Adv. in Appl. Probab.* **38** (2006), no. 1, 59–75.

[92] J. Orlandi, J. Soriano, E. Alvarez-Lacalle, S. Teller, and J. Casademunt. 'Noise focusing and the emergence of coherent activity in neuronal cultures'. *Nature Physics* **9** (2013), no. 9, 582–590.

[93] F. Papadopoulos, M. Kitsak, Á. Serrano, M. Boguñá, and D. Krioukov. 'Popularity versus similarity in growing networks'. *Nature* **489** (2012), no. 7417, 537–540.

[94] B. Pittel, J. Spencer, and N. Wormald. 'Sudden emergence of a giant k-core in a random graph'. *J. Combin. Theory Ser. B* **67** (1996), no. 1, 111–151.

[95] B. Pittel and N. C. Wormald. 'Counting connected graphs inside-out'. *J. Combin. Theory Ser. B* **93** (2005), no. 2, 127–172.

[96] D. Poole. 'On the strength of connectedness of a random hypergraph'. *Electron. J. Combin.* **22** (2015), no. 1, P69.

[97] M. Ramos, J. Shao, S. Reis, C. Anteneodo, J. Andrade, S. Havlin, and H. Makse. 'How does public opinion become extreme?' *Scientific reports* **5** (2015), P10032.

[98] C. Schmeltzer, J. Soriano, I. Sokolov, and S. Rüdiger. 'Percolation of spatially constrained Erdős-Rényi networks with degree correlations'. *Physical Review E* **89** (2014), no. 1, P012116.

[99] J. Schmidt-Pruzan and E. Shamir. 'Component structure in the evolution of random hypergraphs'. *Combinatorica* **5** (1985), no. 1, 81–94.

[100] F. Schwenker, F. Sommer, and G. Palm. 'Iterative retrieval of sparsely coded associative memory patterns'. *Neural Networks* **9** (1996), no. 3, 445–455.

[101] M. Shrestha and C. Moore. 'Message-passing approach for threshold models of behavior in networks'. *Physical Review E* **89** (2014), no. 2, P022805.

[102] B. Söderberg. 'General formalism for inhomogeneous random graphs'. *Phys. Rev. E* **66** (2002), no. 6, P066121.

[103] J. Spencer and N. Wormald. 'Birth control for giants'. *Combinatorica* **27** (2007), no. 5, 587–628.

[104] S. Strogatz. 'Exploring complex networks'. *Nature* **410** (2001), no. 6825, 268–276.

[105] T. Tlusty and J.-P. Eckmann. 'Remarks on bootstrap percolation in metric networks'. *Journal of Physics A: Mathematical and Theoretical* **42** (2009), no. 20, P205004.

[106] T. Turova. 'The emergence of connectivity in neuronal networks: from bootstrap percolation to auto-associative memory'. *Brain research* **1434** (2012), 277–284.

[107] C. Warren, L. Sander, and I. Sokolov. 'Geography in a scale-free network model'. *Physical Review E* **66** (2002), no. 5, P056105.

[108] D. Watts. 'A simple model of global cascades on random networks'. *Proceedings of the National Academy of Sciences* **99** (2002), no. 9, 5766–5771.

[109] X.-J. Xu, X. Zhang, and J. Mendes. 'Impacts of preference and geography on epidemic spreading'. *Physical Review E* **76** (2007), no. 5, P056109.

[110] Z.-D. Zhao, Y. Liu, and M. Tang. 'Epidemic variability in hierarchical geographical networks with human activity patterns'. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22** (2012), no. 2, P023150.