# Large Random Hypergraphs

**Christina Anna Goldschmidt**
**New Hall and Statistical Laboratory**
**University of Cambridge**

**A dissertation submitted for the**
**degree of Doctor of Philosophy**
**July 2003**

# Abstract

In this thesis, we use stochastic process methods to study certain aspects of large random hypergraphs. Consider a random hypergraph on a set of $N$ vertices in which, for $1 \leq k \leq N$, a Poisson($N\beta_k$) number of hyperedges is scattered randomly over all the subsets of size $k$. We collapse the hypergraph by running the following algorithm to exhaustion: pick a vertex having a 1-edge (or *patch*) and remove it; collapse the hyperedges over that vertex onto their remaining vertices; repeat until there are no patches left. We call the vertices removed in this process *identifiable*. Darling and Norris introduced this model and studied the limiting proportion of identifiable vertices.

Chapter 3 concerns the robustness of the Poisson model. We show that, under conditions, altering the initial hyperedge distributions (but retaining the means) gives the same limiting behaviour. To show this, we must prove a fluid limit theorem for an infinite-dimensional Markov process. The appropriate technical machinery is developed in Chapter 2, via truncation.

In Chapter 4, we seek to understand the process of identification better. We use the fluid limit methods of Chapter 2 to investigate the limiting distribution of the patches over the hypergraph at all stages of the process of identification.

In Chapter 5, we investigate the structure of the set of identifiable vertices. Define an *essential* hyperedge to be one whose removal would reduce the number of identifiable vertices. We establish the limiting number of essential edges per vertex.

The subject of Chapter 6 is a phase transition in the proportion of identifiable vertices, which was first discovered by Darling, Levin and Norris. It occurs when the underlying random graph becomes critical and takes various forms depending on the values of $\beta_2, \beta_3, \ldots$. There is a sequence of possible critical scalings and the phase transition may

be continuous, as in a random graph, or discontinuous. We are particularly interested in the latter case and give precise asymptotics.

# Preface

It has been a great pleasure to work under the guidance of James Norris. I am very grateful to him for his generosity, encouragement and sheer enjoyment of mathematics. My thanks also to Richard Darling for his continued interest in my work.

I would like to thank Lara Jamieson, Susie Goldschmidt, Olly Johnson and, especially, Julia Blackwell and Tim Deegan for their invaluable help with typing. I am greatly indebted to Alan Bain for a very careful reading and his comments.

I would like to thank the members of the Statslab, past and present, for making it such an enjoyable place to work. My thanks, in particular, to Richard Samworth for being a wonderful office-mate and to Lara Jamieson for her support.

The argument in Section 5.3 leading to the expected number of essential edges is due to James Norris. Apart from that, this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration with others.

*Christina Goldschmidt*
*New Hall*
*3rd July, 2003*

I would like to add my thanks to my examiners Imre Leader and Terry Lyons for their very thorough reading of this thesis and for their insightful comments.

*Christina Goldschmidt*
*9th December, 2003*

# Contents

# Chapter 1

# Introduction

## 1.1 Themes of the thesis

The primary theme of this thesis is the use of stochastic process methods to study combinatorial problems. In recent years, deep links have emerged between combinatorics and stochastic processes (see Pitman's survey [25]), some of which we are able to exploit. We are partially motivated by David Aldous' challenge to probabilists in [1]:

> "... for readers with a background in modern process theory, we should emphasize that our results open up challenging problems to even re-derive, via process techniques, existing random graph asymptotics formulas."

We take random hypergraphs as our main subject but propose that the collection of techniques used here may be much more widely applicable.

There is much current research interest in the interface between combinatorics, probability and computer science. One cause of this is the fascinating phenomenon of phase transition (a concept borrowed from statistical mechanics). Roughly speaking, a phase transition is a large change in the qualitative behaviour of a system caused by a small alteration in some parameter of the system. Łuczak [24] provides an overview of phase transitions in random discrete structures, taking the emergence of the giant component in a random graph as his prototype. The analysis of phase transitions forms our secondary theme.

## 1.2 Notation

We will denote the power set of a set $S$ by $\mathcal{P}(S)$ and the indicator function of $S$ by $\mathbb{1}_S$. The abbreviation a.a.s. will be used for the phrase 'asymptotically almost surely'. We say that the sequence of events $(E_n)$ occurs a.a.s. if $\mathbb{P}(E_n) \to 1$ as $n \to \infty$.

## 1.3 Background

We commence with an introduction to hypergraphs. A hypergraph $H = (V, E)$ is a set of vertices $V$ and a collection $E$ of subsets of $V$ called hyperedges. We will allow $E$ to contain repeated subsets of $V$. Because of this, we will find it convenient to specify a hypergraph by a map $\Lambda : \mathcal{P}(V) \to \mathbb{Z}^+$, where $\Lambda(A)$ gives the number of times that $A$ appears in $E$. Thus, we allow hyperedges over singleton sets, which we refer to as *patches*. For simplicity, we will not allow hyperedges over the empty set. Whenever necessary, we will draw a hypergraph as in Figure 1.1, with patches being represented by circles around their vertices (the bottom left vertex has two patches) and 2-edges being represented by ordinary graph edges.

Hypergraphs are a very general sort of combinatorial set system. They generalise random graphs but possess a duality property that random graphs do not: identify each edge in the hypergraph with a vertex in the dual and each vertex with an edge in the dual, preserving incidence, so that the size of an edge in the dual is its degree as a vertex in the original graph; then the dual of the hypergraph is another hypergraph (which may definitely contain multiple edges). Hypergraphs provide much scope for us to unify combinatorial concepts. Moreover, they have proved to have many applications, particularly in optimization and computer science, for example as a natural setting for satisfiability problems in combinatorial logic, or for information retrieval systems in relational databases. For a comprehensive guide to applications, see Part IV of Volume 2 of Graham, Grötschel and Lovász [18].

Many combinatorial properties of these structures are investigated in, for example, Duchet [14], Berge [5] and Bollobás [7]. We are interested here in an analogue of graph connectedness called *identifiability*, first introduced in Darling and Norris [12] and further studied in Darling, Levin and Norris [11], whose driving force is the patches on the hypergraph. We define a process of collapse by which we find all of the identifiable vertices in the

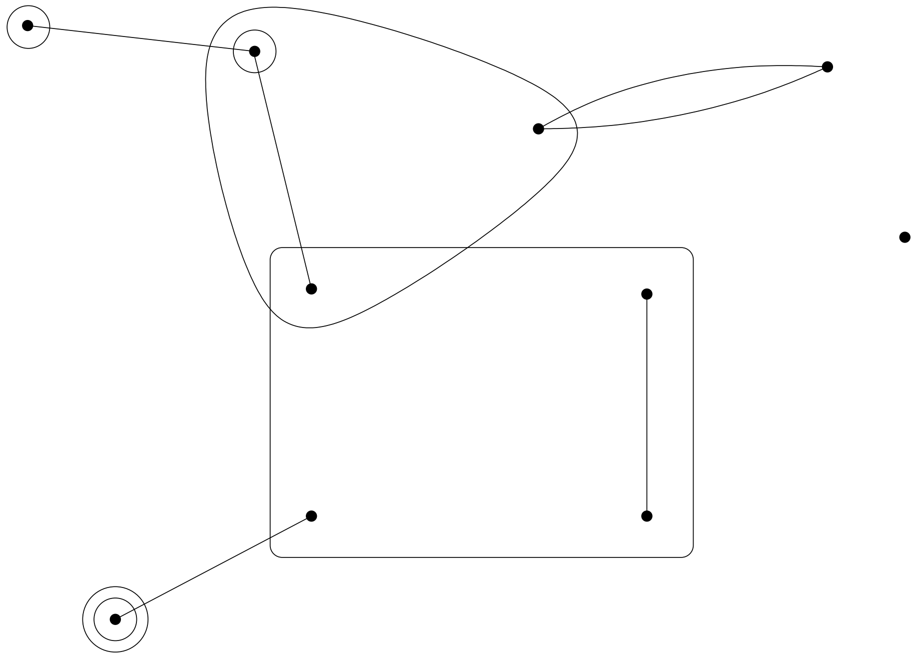

Figure 1.1: A hypergraph on 10 vertices.

hypergraph. Any vertex which has a patch on it is identifiable and we treat patches as roots of 'components' as follows. Pick a patch at random and delete it and the vertex $v$ underneath it. We say that $v$ has been identified. Collapse all of the other hyperedges over $v$ down onto their remaining vertices (so that a 3-edge over $\{u, v, w\}$ becomes a 2-edge over $\{u, w\}$, for example). Repeat until there are no more patches on the hypergraph. The identifiable vertices are those we have deleted. Figure 1.2 shows the process of collapse and Figure 1.3 shows the identifiable vertices in the hypergraph of Figure 1.1. It turns out that the order of this collapse does not affect the ultimate set of identifiable vertices. Note, however, that identifiability from a vertex (that is, whether putting a patch on $v$ enables us to identify $w$) is not a symmetric property. Imagine removing the patches from Figure 1.3 but leaving all other edges as they are. Then $w$ is identifiable from $v$ but $v$ is not identifiable from $w$. Because of this lack of symmetry, there is no obvious definition of a component. However, each vertex has a *domain*, namely the set of vertices identifiable if there is a patch on that vertex and on no other. $w$ may be in the domain of $v$ without $v$ being in the domain of $w$ but the domains of vertices in the same 2-edge component will necessarily be the same. A hyperedge is said to be identifiable if all of its vertices are identifiable.

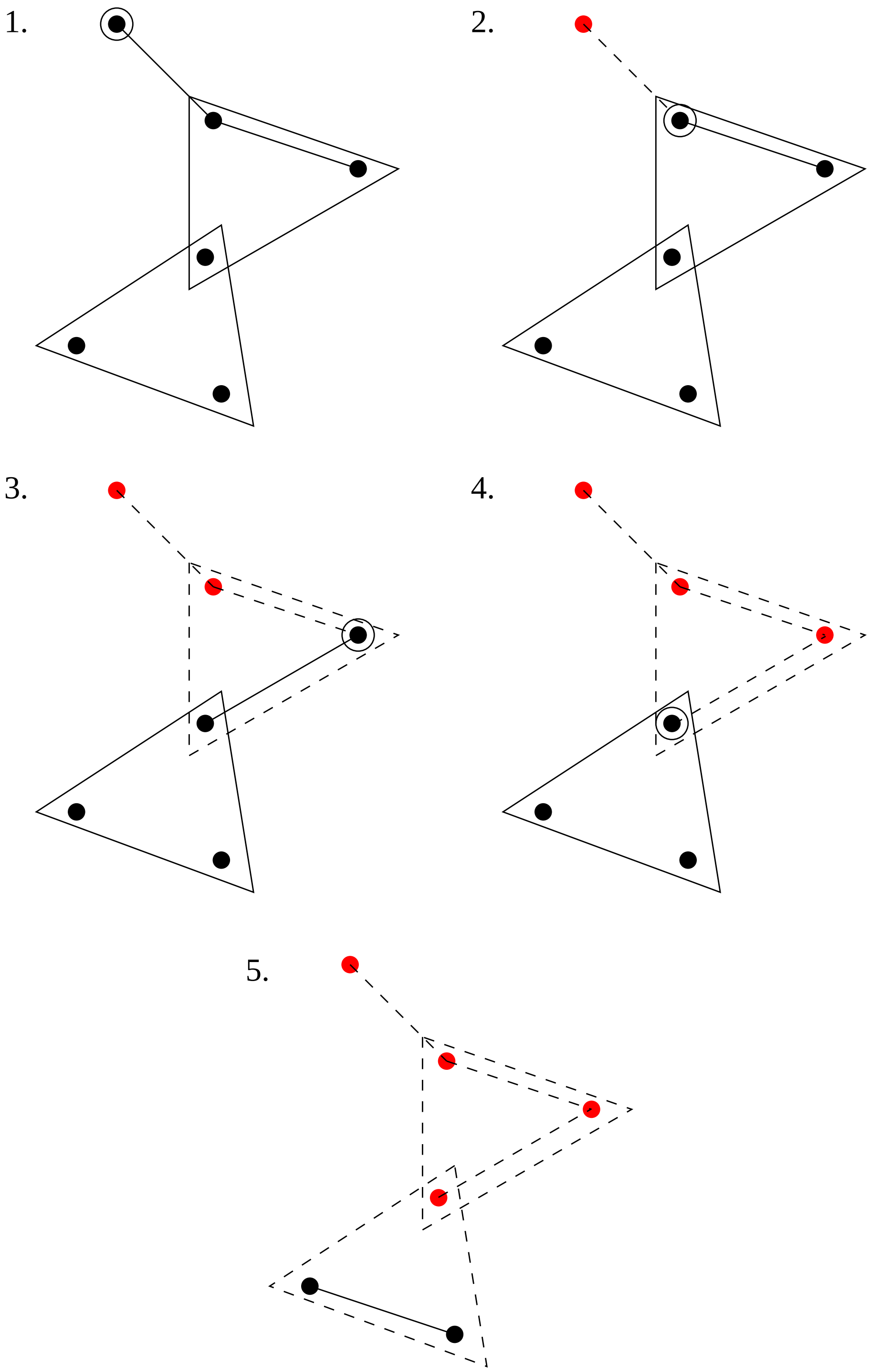


Figure 1.2: Identification, step by step. Vertices which have been identified are shown in red; hyperedges which have been collapsed are drawn with dashed lines.

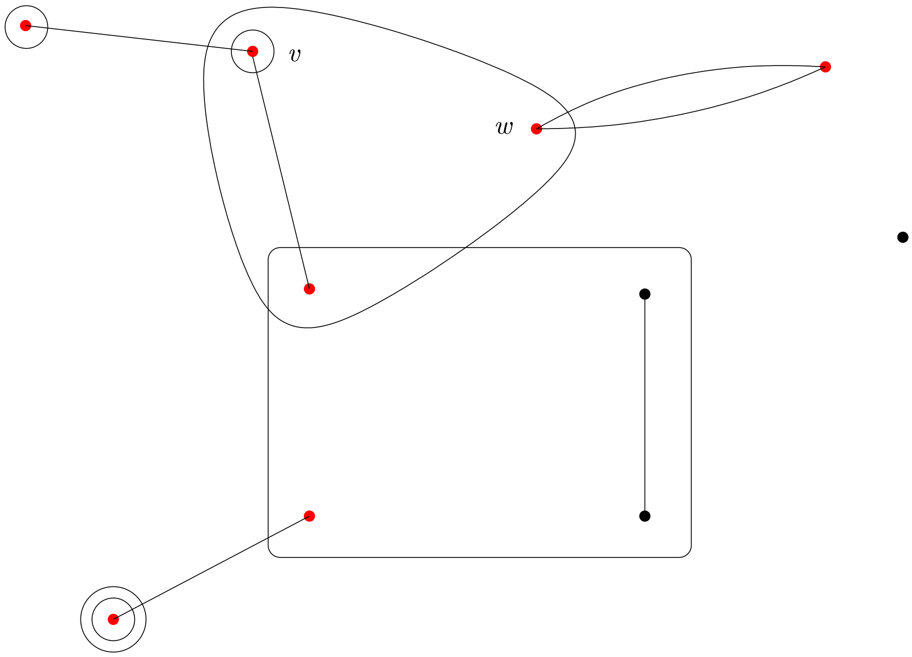


Figure 1.3: The identifiable vertices in this hypergraph are shown in red.

In a random graph (with patches), the identifiable vertices are those which are in the same component as a vertex with a patch on it. Thus, the patches 'pick out' some of the components. Another relationship to graphs comes via the core (the maximal subgraph with minimal vertex degree 2). Take a graph and find the hypergraph dual to it. Vertices of degree 1 in the graph correspond to patches in the dual and so if we collapse the dual, we are left with the dual of the core of the original graph. Indeed, the process of collapse in the dual hypergraph corresponds exactly to the algorithm given in Pittel, Spencer and Wormald [27] for finding the core of the original graph. They proceed as follows. Call a vertex *light* if it has degree at most 1 and *heavy* otherwise. Take a list of all non-isolated light vertices. Select a vertex $v$ uniformly at random from this list and delete all edges incident to it. Some vertices which were previously heavy may now have become light. Repeat this procedure until no non-isolated light vertices remain. The core exists if and only if the final set of heavy vertices is non-empty, in which case they and the remaining edges constitute the core. (See also Subsection 1.5.4.)

The process of collapse to find the identifiable vertices is an example of a greedy algorithm. Greedy algorithms in random structures are often amenable to analysis by Markov chains because they only use information from the current state, rather than past states. Thus, identifiability is likely to engender tractable problems to study.

## 1.4 Summary

Chapters 3 and 4 present variations on, and extensions of, the work of Darling, Levin and Norris. The main results of this thesis may be found in Chapters 5 and 6.

Darling and Norris [12] studied the limiting proportion of identifiable vertices in a particular model of random hypergraphs (called Poisson random hypergraphs; see Section 1.6 for a precise definition). Chapter 3 concerns the robustness of the Poisson model: if we alter the initial distributions of the numbers of hyperedges (but retain the means), does the same limiting behaviour occur? We show that, under certain conditions, this is indeed the case, although the analysis is more complicated than in the Poisson case. Our method involves tracking the process of identification in the hypergraph with a Markov chain and then proving a fluid limit (functional law of large numbers) for it. Considerable technical difficulty is introduced by the fact that this Markov chain is necessarily infinite-dimensional. The appropriate machinery to deal with this is developed in Chapter 2, by use of truncation.

In Chapter 4, we seek to understand the process of identification better. We define an infinite-dimensional Markov process whose co-ordinates are the numbers of vertices with different numbers of patches. We then use the infinite-dimensional fluid limit methods of Chapter 2 to investigate the limiting distribution of the patches over the hypergraph at all stages of the process of identification. This aspect of the hypergraph may be translated into a generic model of coloured balls in an urn with a scheme of deletion and replacement, which provides additional motivation for its study.

In Chapter 5, we investigate the structure of the set of identifiable vertices. We define an *essential* hyperedge to be one whose removal would reduce the number of identifiable vertices. We establish the limiting number of essential edges per vertex, which has a particularly striking form. We also discuss, in the case of a random graph, the relation of essential edges to the core of the graph.

The obvious generalisation of the graph notion of a component to hypergraphs yields a phase transition very similar to that in a random graph (see Schmidt-Pruzan and Shamir [29] and Karoński and Łuczak [21]). The subject of Chapter 6 is a phase transition in our alternative notion of a component, namely the number of identifiable vertices. This was first discovered by Darling, Levin and Norris [11]. It occurs when the underlying random graph becomes critical and takes various forms depending on the values of the parameters controlling the numbers of the different sorts of hyperedge. There is a sequence of possible critical scalings, which is not seen in the random graph. The phase transition may be continuous, as in a random graph (in fact, when there are no higher-order edges, it is precisely the emergence of the giant component), or discontinuous. We are particularly interested in the latter case and are able to exhibit precise asymptotics.

Our method extends a celebrated result of Aldous on component sizes in a random graph. We associate a certain random walk with the process of identification in a random hypergraph. Excursions of this random walk above its past-minima correspond to domains in the hypergraph. We first prove a diffusion limit for an initial section of the random walk, which enables us to show that, in the discontinuous case, we eventually reach a domain of size $\Theta(N)$. We can then prove a fluid limit for the whole random walk and also an associated Functional Central Limit Theorem which extends the initial diffusion limit.

The remainder of this introductory chapter is devoted to the exposition of known results on random graphs, random hypergraphs and fluid limits.

## 1.5 Random graphs background

We will not attempt to give a comprehensive introduction to the theory of random graphs, which has a voluminous literature. Instead, we will discuss some concepts which we will need in the body of the thesis.

### 1.5.1 Models

Suppose that we take a vertex set of size $n$. Then the two most commonly-used models of a random graph are $\mathbb{G}(n, M)$ and $\mathbb{G}(n, p)$. $\mathbb{G}(n, M)$ is the probability space consisting of all graphs on $n$ vertices with $M$ edges such that each graph is equally likely to be chosen. $\mathbb{G}(n, p)$ is the probability space consisting of graphs on $n$ vertices such that each edge is present with probability $p$ and absent otherwise, independently of all other edges. These models are asymptotically equivalent as long as $M \approx \binom{n}{2} p$. For a detailed introduction to the theory of random graphs, see Bollobás [8] and Janson, Łuczak and Ruciński [20];

the asymptotic equivalence of the two models is Proposition 1.13 of the latter.

Later, we will wish to work in the framework of Poisson random hypergraphs and, if we specialise to the case with no higher-order edges or patches, Poisson random graphs. In the Poisson random graph, each pair of vertices has a Poisson$(2\beta_2/(n-1))$ number of edges, independently of all other pairs of vertices. Technically, this gives a multigraph. However, for properties such as the size of components (in which we will be particularly interested) the multiple edges are hardly even an inconvenience. It is easy to see that the parameter correspondence between $\mathbb{G}(n,p)$ and the Poisson random graph is $p = 1 - \exp(-2\beta_2/(n-1)) \approx 2\beta_2/n$ for large $n$. Thus, later, when we consider the problem of the size of the largest component for $p = \lambda/n$ and $\lambda$ varying, we need only set $\lambda = 2\beta_2$ and we will have asymptotically the same problem.

The Poisson structure is probabilistically natural and considerable power is gained from the splitting and addition properties of Poisson random variables.

By a *random graph process*, we will mean the graph which evolves in time as follows. At time 0, it has no edges. Each pair of vertices $\{v,w\}$ has a U$[0,1]$ random variable $U_{vw}$ associated with it, independently for different pairs. At time $t$, the graph has edges given by $\{\{v,w\} : U_{vw} < t\}$. Thus, we can unambiguously denote the random graph process by $(\mathbb{G}(n,t))_{0\le t\le 1}$. Note that this is a particular coupling of the random graphs $\mathbb{G}(n,p)$ for all $p \in [0,1]$.

### 1.5.2 The size of the largest component

One of the most studied quantities in a random graph is the size of the largest component. The emergence of the *giant component* (i.e. a component containing a positive proportion of the vertices) has been the subject of many famous papers; see, in particular, Erdős and Rényi [15], Bollobás [6] and Łuczak [23]. For convenience, we will work with $\mathbb{G}(n,p)$. Firstly we will define a random walk associated with a graph, called the breadth-first walk.

Suppose we have a graph on $n$ vertices, with some arbitrary labelling of those vertices. We will first define the *breadth-first ordering*, which is simply a re-ordering of the vertex labels. Take the lowest-numbered vertex and call it $v(1)$. (We will think of $v(1)$ as the root of the first component.) Define the children of a vertex to be those vertices which have not yet been re-labelled and which are connected to the parent vertex by an edge. Suppose vertex $v(1)$ has $c(1)$ children. Label them $v(2),\ldots,v(c(1)+1)$, retaining the original order of the vertex labels. Now consider the children of $v(2)$ and call them $v(c(1)+2), \dots, v(c(1)+c(2)+1)$, as always retaining the original order. Repeat this process. If the last vertex in the list has no children, find the lowest-numbered vertex (in the old ordering) that we have not yet considered. This will be the root of the next component. The process terminates when we reach the last vertex. We have effectively constructed a rooted ordered spanning forest of the original graph.

Now let $z(0) = 0$ and, for $1 \leq i \leq n$,

$$z(i) = z(i-1) + c(i) - 1.$$

Then $(z(i))_{0 \leq i \leq n}$ is the *breadth-first walk*, which summarises certain information about the underlying graph. In particular, every time we reach the end of a component, the breadth-first walk hits a new minimum and so we can read off the component sizes as the distances between successive minima of the walk.

Define $m(i)$ to be the number of *marked* vertices at stage $i$; that is, the number of vertices which we have re-labelled in the breadth-first ordering by the time we have labelled the children of $v(i)$ but which are not contained in the set $\{v(1), \dots, v(i)\}$.

For a *random* graph, we obtain a random walk $(Z(i))_{0 \leq i \leq n}$. Let $C(i)$ be the number of children of the $i$th vertex in the breadth-first ordering and $M(i)$ the number of marked vertices by stage $i$, which is some function of $(Z(0), Z(1), \dots, Z(i))$. We assume that each edge is present independently with probability $p$. Then given $Z(0), Z(1), \dots, Z(i)$,

$$Z(i+1) = Z(i) + C(i+1) - 1$$

and

$$C(i+1) \sim \mathrm{Bin}(n - i - M(i), p).$$

$M(i)$ is bounded above by the number of vertices in two levels in the underlying tree and turns out to be negligible compared to $i$ (at least in the asymptotic regimes we will consider).

Suppose now that we take $p = \lambda/n$ for some constant $\lambda > 0$ and let $n \to \infty$. Then, by the Law of Small Numbers (see, for example, Barbour, Holst and Janson [4]), we have that the number of children of any vertex is asymptotically Poisson($\lambda$) and so the breadth-first walk looks like that on the tree produced by a Galton–Watson branching process with Poisson($\lambda$) offspring distribution. (For more information on branching processes, see

Harris [19] or Athreya and Ney [3].) The behaviour of this branching process is determined by the value of the parameter $\lambda$. If $\lambda \leq 1$, the branching process becomes extinct almost surely. If $\lambda > 1$, the branching process has probability $q$ of becoming extinct, where $q < 1$ is the minimal non-negative solution to

$$q = e^{-\lambda(1-q)}.$$

This argument, made rigorous, allows us to prove the following theorem.

**Theorem 1.1.** *Let $p = \lambda/n$ where $\lambda > 0$ and consider the sizes of the components for large $n$. Then a.a.s.,*

- *if $\lambda < 1$, all components are $\mathcal{O}(\log n)$;*
- *if $\lambda = 1$, the largest and second-largest components are $\Theta(n^{2/3})$;*
- *if $\lambda > 1$, the largest component contains a proportion $1 - q$ of the vertices; all other components are $\mathcal{O}(\log n)$.*

The dramatic change at $\lambda = 1$ is often referred to as the phase transition in a random graph. Considerably more detailed information is available about the sizes of the components but we shall not need it here. We give a proof (from Janson, Łuczak and Ruciński [20]) of the first and third statements now and come back to the critical case in the next section.

First we need a technical lemma.

**Lemma 1.2 (Chernoff's inequality).** *Suppose $X \sim \text{Bin}(n,p)$ with $p = \lambda/n$. Then for $t \geq 0$,*

$$\mathbb{P}(X \geq \lambda + t) \leq \exp\left(-\frac{t^2}{2(\lambda + t/3)}\right) \tag{1.1}$$

$$\mathbb{P}(X \leq \lambda - t) \leq \exp\left(-\frac{t^2}{2\lambda}\right). \tag{1.2}$$

*Proof.* The proof is an application of an optimised version of Chebyshev's inequality and may be found in Janson, Łuczak and Ruciński [20]. □

*Proof of Theorem 1.1.* Suppose first that $\lambda < 1$ and that $k \geq 3(\log n)/(1-\lambda)^2$. Then the probability that $v(1)$ is contained in a component of size at least $k$ is bounded above by

$$\mathbb{P}(C(1) + \cdots + C(k) \geq k - 1).$$

Each of $C(1), \ldots, C(k)$ is independently stochastically dominated by $\mathrm{Bin}(n, \lambda/n)$. Thus, the probability that the graph contains a component of size at least $k$ is bounded by

$$\begin{aligned}
n\mathbb{P}\left(C(1) + \cdots + C(k) \geq k-1\right) &\leq n\mathbb{P}\left(\mathrm{Bin}(nk, \lambda/n) \geq k-1\right) \\
&= n\mathbb{P}\left(\mathrm{Bin}(nk, \lambda/n) \geq \lambda k + (1-\lambda)k - 1\right) \\
&\leq n \exp\left(-\frac{((1-\lambda)k-1)^2}{2(\lambda k + ((1-\lambda)k-1)/3)}\right) \\
&\leq n \exp(-(1-\lambda)^2 k/2) \\
&\leq n^{-1/2}.
\end{aligned}$$

The third line follows from Chernoff's inequality (1.1). Clearly, this quantity converges to 0 as $n \to \infty$ and so we have proved the first statement of the theorem.

Now suppose that $\lambda > 1$. Let $k_1 = 16\lambda(\log n)/(\lambda-1)^2$ and $k_2 = n^{2/3}$. Take any vertex $v$ and imagine starting the breath-first ordering with $v(1) = v$. We shall show that a.a.s. any such vertex $v$ either lies in a component of size less than $k_1$ or, for all $k_1 \leq k \leq k_2$, $M(k) \geq (\lambda-1)k/2$. Thus, no component can have size between $k_1$ and $k_2$.

The probability that there exists a vertex in a component of size greater than $k_1$, but with $M(k) < (\lambda-1)k/2$ for some $k_1 \leq k \leq k_2$, is bounded above by

$$\begin{aligned}
&n \sum_{k=k_1}^{k_2} \mathbb{P}\left(C(1) + \cdots + C(k) \leq k - 1 + (\lambda-1)k/2\right) \\
&\qquad = n \sum_{k=k_1}^{k_2} \mathbb{P}\left(C(1) + \cdots + C(k) \leq (\lambda+1)k/2\right)
\end{aligned}$$

But on the set $\{C(1)+\cdots+C(k) \leq (\lambda+1)k/2\}$ we have that $C(i)$ stochastically dominates $\mathrm{Bin}(n-(\lambda+1)k/2, \lambda/n)$, independently for $1 \leq i \leq k$, and so the above formula is less than

$$\begin{aligned}
n \sum_{k=k_1}^{k_2} \mathbb{P}\left(\mathrm{Bin}(nk - (\lambda+1)k/2, \lambda/n) \leq (\lambda+1)k/2\right) &\leq n \sum_{k=k_1}^{k_2} \exp\left(-\frac{(\lambda-1)^2 k^2}{9\lambda k}\right) \\
&\leq nk_2 \exp\left(-\frac{(\lambda-1)^2 k_1}{9\lambda}\right) \\
&= n^{-1/9}
\end{aligned}$$

where we use Chernoff's inequality (1.2) on the first line.

Suppose now that we have vertices $v$ and $w$, both lying in components of size at least $k_2$. We wish to show that they must lie in the same component. Imagine running the breadth-first ordering algorithm starting from $v$ for $k_2$ steps. Then we must end up with at least $(\lambda-1)k_2/2$ marked vertices. Now imagine that we run the breadth-first ordering starting from $w$ for $k_2$ steps and that this results in a distinct set of vertices. Then the probability that $v$ and $w$ are in different components is bounded above by the probability that none of the at least $(\lambda-1)k_2/2$ marked vertices in each component are connected to one another with an edge. This is

$$(1-\lambda/n)^{((\lambda-1)k_2/2)^2} \leq \exp(-(\lambda-1)^2\lambda n^{1/3}/4)$$

which converges to 0 as $n \to \infty$. So, a.a.s. there is a unique largest component, which contains more than $k_2$ vertices.

We will complete the proof by finding the number of vertices in components of size at most $k_1$. The probability $p_{\text{small}}(n, \lambda/n)$ that a vertex is in one of these small components is bounded above by the extinction probability $q(n-k_1, \lambda/n)$ of a branching process with $\text{Bin}(n-k_1, \lambda/n)$ offspring distribution. It is bounded below by the extinction probability $q(n, \lambda/n)$ of a branching process with $\text{Bin}(n, \lambda/n)$, minus the probability of the same process dying out after $k_1$ steps, which is asymptotically negligible. Both $q(n-k_1, \lambda/n)$ and $q(n, \lambda/n)$ converge to the extinction probability, $q$, of a branching process with Poisson($\lambda$) offspring.

If $S$ is the number of vertices in components of size at most $k_1$ then we have just shown that $\mathbb{E}[S/n] \to q$ as $n \to \infty$. $S(S-1)$ is the number of ways of picking an ordered pair of vertices from small components. The second vertex is either in the same small component as the first or a different one and so

$$\frac{1}{n^2}\mathbb{E}[S(S-1)] \leq \frac{1}{n^2} n p_{\text{small}}(n, \lambda/n)(k_1 + n p_{\text{small}}(n - \mathcal{O}(k_1), \lambda/n)) \to q^2$$

as $n \to \infty$. Hence,

$$\frac{1}{n^2}\text{var}(S) = \frac{1}{n^2}\mathbb{E}[S(S-1)] + \frac{1}{n^2}\mathbb{E}[S] - \frac{1}{n^2}\mathbb{E}[S]^2 \to 0$$

as $n \to \infty$ and so, by Chebyshev's inequality,

$$\frac{1}{n}S \xrightarrow{p} q.$$

Thus, the proportion of vertices in the largest component converges to $1-q$. □

### 1.5.3 The critical random graph

One of the most fruitful applications of stochastic process theory to the study of random graphs has been Aldous' work on the breadth-first walk in Aldous [1]. We will give here a summary based on that paper and also on Section 5.3 of Pitman [25].

**Theorem 1.3 (Aldous).** *Let $(Z^r(i))_{0\leq i\leq n}$ be the breadth-first walk on $\mathbb{G}(n, n^{-1}+rn^{-4/3})$, where $r \in \mathbb{R}$. Let $(B(t))_{t\geq 0}$ be a standard Brownian motion and let $B^r(t) = B(t)+rt-t^2/2$. Then, defining*

$$\bar{Z}^r(t) = n^{-1/3} Z^r(\lfloor n^{2/3} t\rfloor),$$

*we have that $\bar{Z}^r \xrightarrow{d} B^r$ as $n \to \infty$ in $D[0, \infty)$.*

*Sketch of proof.* In the statement, we have used the floor function to interpolate between integer-valued time points. With a slightly different interpolation (which is easier to deal with and leads to the same result), we can express the breadth-first walk as the sum of a martingale and a bounded variation process. The limit of the bounded variation process is the expected drift. We can apply the martingale central limit theorem to the martingale, whose behaviour is essentially determined by the variance of the increments of the process. Now,

$$\begin{aligned}
&\mathbb{E}[Z^r(i+1) - Z^r(i)|Z^r(0), Z^r(1), \dots Z^r(i)] \\
&\qquad = (n - i - M(i))(n^{-1} + rn^{-4/3}) - 1 \\
&\qquad \approx rn^{-1/3} - in^{-1} \\
&\mathrm{var}(Z^r(i+1) - Z^r(i)|Z^r(0), Z^r(1), \dots Z^r(i)) \\
&\qquad = (n - i - M(i))(n^{-1} + rn^{-4/3})(1 - n^{-1} - rn^{-4/3}) \\
&\qquad \approx 1.
\end{aligned}$$

Putting $i = n^{2/3}t$ we obtain a conditional mean drift rate of

$$\mathbb{E}\Big[Z^r(n^{2/3}t + 1) - Z^r(n^{2/3}t)|Z^r(0), Z^r(1), \dots Z^r(n^{2/3}t)\Big] \approx (r - t)n^{-1/3}$$

and a conditional variance rate which is 1. Thus, on integrating, the bounded variation process converges to $rt - t^2/2$ and the martingale converges to Brownian motion. □

In the breadth-first walk, the components come in size-biased order ($v(1)$ is more likely to be in a larger component than in a smaller one) and a small amount of stochastic calculus is enough to show that the largest component of the graph must be of size $\Theta(n^{2/3})$. Of course, Theorem 1.3 tells us much more precise distributional information as well.

Aldous goes on to prove a fascinating and important link between random graphs and the *multiplicative coalescent* (see Aldous [2] for a wide-ranging survey of coalescence). He defines a standard multiplicative coalescent by using re-scaled component sizes of a certain random graph process.

### 1.5.4 An aspect of the structure of the giant component

The structure of the giant component is known in great detail. Here we wish to bring attention to just one aspect of this structure, so that later we may refer back to it.

Define the $k$-core of a graph $G$ to be the (unique) maximal subgraph $H$ of $G$ with minimal degree $k$. The 2-core (or *core*) consists of all cycles and paths between cycles. In a random graph process, the first core is just the first cycle. The core only contains an asymptotically positive proportion of the vertices for supercritical random graphs (that is, if $p = \lambda/n$ with $\lambda > 1$). Here the random graph consists of a giant component and small components of size $\mathcal{O}(\log N)$. These small components tend to be tree-like (consider their relation to branching process family trees) and so most of the core is in the giant component.

Suppose $q$ is the extinction probability of a Galton–Watson branching process with Poisson($\lambda$) offspring distribution, i.e. $q$ is the smallest positive root of $q = e^{\lambda(q-1)}$, and let $\theta = 1 - q$, the asymptotic proportion of vertices in the giant component. The following theorem is from Pittel [26].

**Theorem 1.4 (Pittel).** *Let $\phi$ be the smaller root of $\phi e^{-\phi} = \lambda e^{-\lambda}$. Then the giant component consists of a 2-vertex-connected subgraph, $H$, (that is, removal of no single vertex disconnects $H$) of size close to $n\theta(1 - \phi)$ and a collection of trees, and possibly a few unicyclic components, each sprouting from a different vertex of $H$.*

Thus, the core grows with the giant component and a vertex in the giant component has approximate probability $1 - \phi$ of being in the core.

The influential paper of Pittel, Spencer and Wormald [27] explores the point at which the $k$-core becomes giant, for $k \geq 3$. Unlike the 2-core, when the $k$-core appears it is already giant and contains a strictly positive proportion of the vertices.

The paper is also notable for the techniques employed to obtain the results, which are expounded in a more general form in Wormald [33]. The authors proceed by defining an edge-deletion algorithm, which is tracked by a certain Markov process. This Markov process is then shown to possess a fluid limit which gives the required information about the behaviour of order $n$ structures in the graph. This is much the same paradigm as is used in Darling and Norris [12] and large parts of the present thesis.

## 1.6 Poisson random hypergraphs

We now specify precisely the Poisson random hypergraph model which we will use hereafter. Suppose we are given a set of vertices $V$ of size $N$. We define a Poisson random hypergraph with parameters $\{\beta_k : k \geq 1\}$ by a random map $\Lambda : \mathcal{P}(V) \to \mathbb{Z}^+$ such that

$$\Lambda(A) \sim \text{Poisson}\left(N\beta_k / \binom{N}{k}\right)$$

whenever $|A| = k$. $\Lambda(A)$ is the number of hyperedges over the set $A$. The numbers of hyperedges over distinct subsets of the vertex set are independent and the distribution of $\Lambda(A)$ depends only on $|A|$. Define a generating function, $\beta(t) = \sum_{k=1}^{\infty} \beta_k t^k$. Throughout this thesis, we shall assume that $\beta(1) < \infty$ for simplicity of exposition (this condition, for example, enables us to differentiate $\beta(t)$ for $0 \leq t < 1$).

Suppose we take a Poisson random hypergraph with parameters $\beta_1, \beta_2, \ldots$. Now remove a set $S$ of size $n$ from $V$ and collapse the hyperedges over sets intersecting $S$ down onto their remaining vertices. Let $\Lambda'$ be the new hypergraph obtained. Then for $A \subseteq V \setminus S$,

$$\Lambda'(A) = \sum_{\substack{B \supseteq A: \\ B \setminus S = A}} \Lambda(B).$$

Suppose now that $|A| = k$. There are $\binom{n}{i}$ sets $B$ of size $k + i$ such that $A \subseteq B$ and $B \setminus S = A$. Thus, because the random variables $\Lambda(B)$ are independent for different sets $B$,

$$\Lambda'(A) \sim \text{Poisson}(\lambda_k(N, n)),$$

where

$$\lambda_k(N, n) = N \sum_{i=0}^{n} \beta_{i+k} \binom{n}{i} / \binom{N}{i+k}.$$

Furthermore, the random variables $(\Lambda'(A) : A \subseteq V \setminus S)$ inherit independence from the original hypergraph. Thus, $\Lambda'$ is another Poisson random hypergraph.

The following lemma concerning the 2-edge parameter of $\Lambda'$ is proved in Darling and Norris [12].

**Lemma 1.5 (Darling and Norris).** *Take $r < 1$. Then, uniformly for $n \in [0, rN] \cap \mathbb{Z}$,*

$$\left|N\lambda_2(N,n) - \beta''\left(\tfrac{n}{N}\right)\right| \leq \frac{C(\log N)^2}{N}.$$

Let $X_N$ be the number of identifiable vertices and $H_N$ the number of identifiable hyperedges in a Poisson random hypergraph. We will review here some asymptotic results for $X_N$ and $H_N$.

**Theorem 1.6 (Darling and Norris).** *Define $t^* = \inf\{y \geq 0 : \beta'(y) + \log(1-y) < 0\} \wedge 1$ and suppose that $t^* < 1$ and that there are no zeros of $\beta'(y) + \log(1-y)$ in $[0, t^*)$. Then for all $\epsilon > 0$,*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\left|N^{-1}X_N - t^*\right| > \epsilon\right) < 0$$

*and*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\left|N^{-1}H_N - \beta(t^*) + (1-t^*)\log(1-t^*)\right| > \epsilon\right) < 0.$$

Thus, $N^{-1}X_N$ and $N^{-1}H_N$ have limits in probability which are attained exponentially fast.

*Sketch proof.* The analysis rests on the observation that $(P_n, D_n)_{n\geq 0}$ is a Markov chain, where $P_n$ is the number of patches on the hypergraph and $D_n$ is the number of patches we have deleted so far, after $n$ vertex deletions. Specifically, given $P_n$ and $D_n$,

$$\begin{aligned} P_{n+1} &= P_n - 1 - S_{n+1} - A_{n+1} \\ D_{n+1} &= D_n + 1 + S_{n+1} \end{aligned}$$

where $S_{n+1}$ is the number of patches sharing the same vertex as the deleted one and $A_{n+1}$ is the number of patches added to the hypergraph, coming from 2-edges over the deleted vertex. We have $S_{n+1} \sim \mathrm{Bin}(P_n - 1, 1/(N-n))$ and $A_{n+1} \sim \mathrm{Poisson}((N-n-1)\lambda_2(N,n))$, independently. Thus, provided $P_n > 0$,

$$\begin{aligned} \mathbb{E}[P_{n+1} - P_n | P_n, D_n] &= (N-n-1)\lambda_2(N,n) - \frac{1}{(N-n)}(P_n - 1) - 1 \\ \mathbb{E}[D_{n+1} - D_n | P_n, D_n] &= 1 + \frac{1}{(N-n)}(P_n - 1). \end{aligned}$$

We proceed by deriving a fluid limit for the process $(P_n, D_n)_{n\geq 0}$; that is, we find deterministic functions $p$ and $d$ such that $N^{-1}(P_{\lfloor Nt\rfloor}, D_{\lfloor Nt\rfloor})_{t\geq 0}$ converges in probability (in some appropriate sense) to $(p(t), d(t))_{t\geq 0}$, at least until $T^N = \inf\{t \geq 0 : P_{\lfloor Nt\rfloor} \leq 0\}$, the first time that $P$ hits 0. The limit in probability of $T^N$ gives the limiting proportion of identifiable vertices.

We have

$$\begin{aligned}\mathbb{E}\left[P_{\lfloor Nt\rfloor+1} - P_{\lfloor Nt\rfloor} | P_{\lfloor Nt\rfloor}, D_{\lfloor Nt\rfloor}\right] &\approx (1-t-1/N)N\lambda_2(N,Nt) \\ &\quad - \frac{1}{(1-t)}N^{-1}(P_{\lfloor Nt\rfloor}-1) - 1 \\ \mathbb{E}\left[D_{\lfloor Nt\rfloor+1} - D_{\lfloor Nt\rfloor} | P_{\lfloor Nt\rfloor}, D_{\lfloor Nt\rfloor}\right] &\approx 1 + \frac{1}{(1-t)}N^{-1}(P_{\lfloor Nt\rfloor}-1).\end{aligned}$$

and so, heuristically, the limit must satisfy

$$\begin{aligned}\dot{p}(t) &= (1-t)\beta''(t) - \frac{1}{(1-t)}p(t) - 1 \\ \dot{d}(t) &= 1 + \frac{1}{(1-t)}p(t).\end{aligned}$$

The unique solution to this system is $p(t) = (1-t)(\beta'(t) + \log(1-t))$ and $d(t) = \beta(t) - (1-t)\log(1-t)$. Thus, the limit in probability for the proportion of identifiable vertices is $t^* = \inf\{t \geq 0 : p(t) < 0\} = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$ and that for the number of identifiable hyperedges is $d(t^*)$. $\square$

We can illustrate this result with some examples of the curve $p(t)$. Let $\beta'(t) = 6.5\gamma(0.1 + 0.9t)^5$ and consider the two cases $\gamma = 140$ and $\gamma = 143$. Figure 1.4 shows how in the $\gamma = 140$ case $p(t)$ hits zero very rapidly. In the $\gamma = 143$ case, $p(t)$ just manages to escape 0; Figure 1.5 shows how it then takes off and eventually yields all of the vertices. The fact that a small change in parameter values gives such a difference in behaviour (in one case, a proportion roughly 0.023 of the vertices is identifiable; in the other, all of the vertices are identifiable) indicates that there is some underlying phase transition in this model. This idea will be made concrete later.

Darling and Norris also proved a version of this result when there are zeros of $\beta'(t) + \log(1-t)$ in $[0, t^*)$. In this case, at any point where the limit function grazes 0, the process has a chance to step over the boundary. Thus, we obtain a random limit in distribution for $T^N$. One particular case is studied in Darling, Levin and Norris [11], namely that of a random hypergraph with no patches (i.e. $\beta_1 = 0$). Let $X_N$ be the number of vertices identified when a single patch is placed on the hypergraph, say on vertex $v$. Thus, $X_N$ is the size of the domain of $v$.

**Theorem 1.7 (Darling, Levin and Norris).** *Suppose that $\beta_1 = 0$ and $\beta_2 > 0$. Let $t^* = \inf\{y \geq 0 : \beta'(y) + \log(1-y) < 0\} \wedge 1$ and suppose that $t^* < 1$ and there are no zeros*

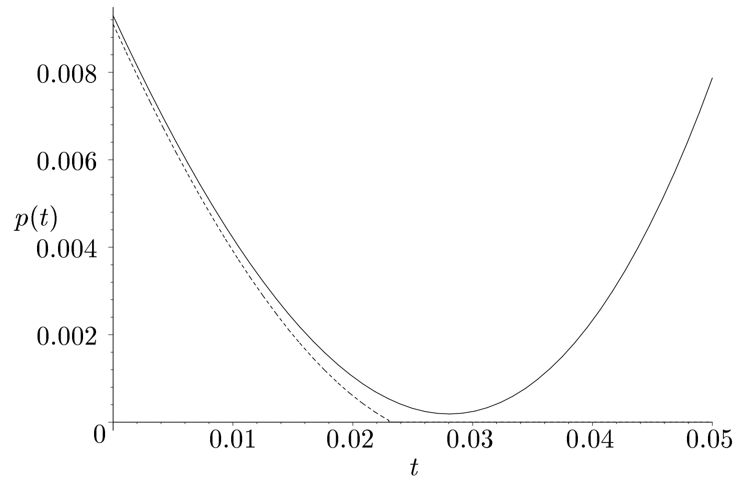


Figure 1.4: The leftmost section of the graph of $p(t)$ against $t$. The dashed line is the case $\gamma = 140$ and the solid line is the case $\gamma = 143$.

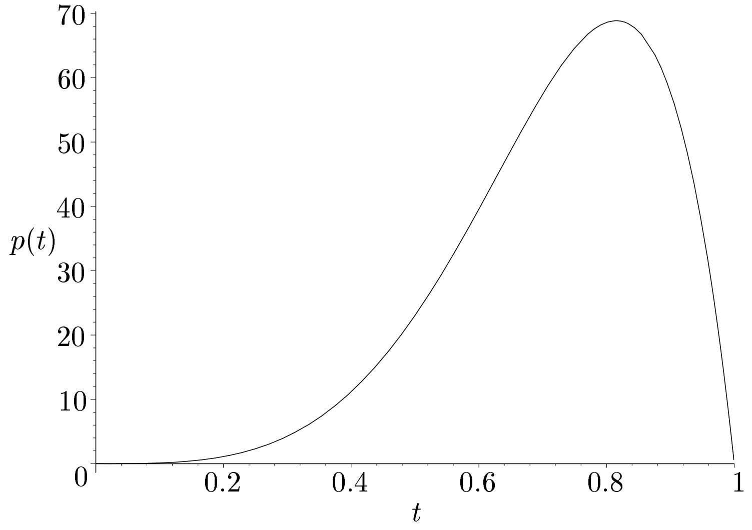


Figure 1.5: The whole graph of $p(t)$ against $t$ for $\gamma = 143$.

*of $\beta'(y)+\log(1-y)$ in $(0,t^*)$. Then,*

$$X_N \xrightarrow{d} D \tag{1.3}$$

*and*

$$\frac{1}{N}X_N \xrightarrow{d} t^*\mathbb{1}_{\{D=\infty\}} \tag{1.4}$$

*as $N\to\infty$, where $D$ has the Borel$(2\beta_2)$ distribution. That is, $D$ has the distribution of the total population of a Galton–Watson branching process with Poisson$(2\beta_2)$ offspring distribution:*

$$\begin{aligned}\mathbb{P}\left(D=n\right) &= e^{-2\beta_2 n}(2\beta_2 n)^{n-1}/n!, \qquad n\geq 1\\ \mathbb{P}\left(D=\infty\right) &= 1-q\end{aligned}$$

*with $q$ the extinction probability which is the smallest root in $[0,1]$ of $2\beta_2(1-x)+\log x=0$. (Note that $q=1$ for $\beta_2\leq 1/2$ and $q<1$ otherwise.)*

If $\beta_k=0$ for $k\geq 3$ then (1.3) is a more explicit statement of the branching process ideas used in Theorem 1.1: the domain of $v$ is the component containing $v$ and small components look microscopically like branching process family trees. Moreover, the result is true even in the presence of higher-order edges and so, interestingly, we cannot feel their effect on this scale.

If we look at the proportion of vertices in the domain of $v$, (1.4), we see that this can only be positive if $\beta_2>1/2$. In this case, $D=\infty$ with probability $1-q$, which is the probability that the vertex $v$ is in the giant component of the underlying random graph. However, the proportion of vertices identified does depend on the higher-order edges.

Darling, Levin and Norris also considered a Poisson random hypergraph process $(\Lambda_t)_{t\geq 0}$ such that for a set $A$ of size $k$, $(\Lambda_t(A))_{t\geq 0}$ is a Poisson process of rate $N\beta_k/\binom{N}{k}$, these Poisson processes being independent for distinct sets $A$. Let $X_N(t)$ be the number of identifiable vertices in the Poisson random hypergraph process at time $t$. Let $x(t)=\inf\{y\geq 0: t\beta'(y)+\log(1-y)<0\}$. Then Theorem 1.6 leads us to suspect that $(N^{-1}X_N(t))_{t\geq 0}\to(x(t))_{t\geq 0}$ in some sense. A complication occurs if the function $x$ is discontinuous. Let $Z=\{t\geq 0: x(t-)\neq x(t)\}$, the discontinuity set of $x$.

**Theorem 1.8 (Darling, Levin and Norris).** *Suppose $\beta_1>0$ and that there are no zeros of $t\beta'(y)+\log(1-y)$ (considered as a function of $y$) between $x(t-)$ and $x(t)$ (which* *are, themselves, both zeros of that function). Then*

$$\left(N^{-1}X_N(t)\right)_{t\geq 0} \xrightarrow{d} (X(t))_{t\geq 0}$$

*in the sense of convergence of finite-dimensional distributions, where*

$$X(t) = \begin{cases} x(t) & t \notin Z \\ x(t-) + (x(t) - x(t-))I_t & t \in Z \end{cases}$$

*and* $(I_t)_{t\in Z}$ *is a family of independent Bernoulli(1/2) random variables. Furthermore, for any compact interval* $K \subseteq \mathbb{R}^+ \setminus Z$,

$$\sup_{t\in K}\left|N^{-1}X_N(t) - x(t)\right| \xrightarrow{p} 0$$

*as* $N \to \infty$.

In other words, the expected convergence holds, except at discontinuities, where the limit takes the upper and lower values each with probability 1/2.

A particular example of such a discontinuity will be the subject of Chapter 6.

## 1.7 Fluid limit methodology

Fluid limits are the equivalent of the law of large numbers for stochastic processes. Suppose that we have a process $(X_t^N)_{t\geq 0}$ which depends in some way on the parameter $N$, for example, a Poisson process of rate $N\lambda$. Then its fluid limit is a deterministic function $(x(t))_{t\geq 0}$ such that if $t_0$ is a constant,

$$\left(N^{-1}X_{Nt}^N\right)_{0\leq t\leq t_0} \xrightarrow{p} (x(t))_{0\leq t\leq t_0}$$

in some appropriate sense (e.g. the supremum norm of the difference between $N^{-1}X_{Nt}^N$ and $x(t)$ converges to 0 in probability). We will refer to this scaling of a process as the *fluid scaling*. In the case of the Poisson process, the deterministic function is the line $x(t) = \lambda t$. Chapter 11 of Ethier and Kurtz [16] gives a general theory of these stochastic process limits; Chapters 5–8 of Shwartz and Weiss [30] give a discussion of fluid limits in the context of large deviations for queueing networks; Darling [10] is an accessible practical guide to proving fluid limit theorems.

Fluid limits have been used to great effect in the analysis of stochastic networks (see, for example, Chen and Yao [9] or Whitt [32]). They can be used to investigate the macroscopic behaviour of many other large random systems, although their use in the study of random graphs is relatively recent (see Wormald [33] and the discussion in Subsection 1.5.4).

In this thesis, we will use the fluid limit techniques expounded in Darling and Norris [12] and used there to prove Theorem 1.6. By imposing conditions on the Laplace transforms of the Lévy kernels, it is possible to obtain large deviations-style results (i.e. convergence exponentially fast in probability). Moreover, these techniques allow us to deal with the situation where the process leaves some set $S$ of interest. In this case, we would like to show that the first exit time of the process from $S$, suitably rescaled, converges in probability to the first exit time of the deterministic limit function. We present here a 'recipe' for proving such a fluid limit result.

Suppose that $(X_t^N)_{t\geq 0}$ is a time-homogeneous pure jump Markov process taking values in $I^N = \frac{1}{N}\mathbb{Z}^d \subseteq \mathbb{R}^d$. Let $K^N(x, dy)$ be the Lévy kernel of $(X_t^N)_{t\geq 0}$. Thus, the jumps of $(X_t^N)_{t\geq 0}$ are governed in the following way: if $T$ is the time of the first jump then

$$\mathbb{P}\left(T \in dt, X_T^N - X_{T-}^N \in dy \mid T > t, X_0^N = x\right) = K^N(x, dy)dt.$$

Define the Laplace transform of this Lévy kernel by

$$m^N(x, \theta) = \int_{\mathbb{R}^d} e^{\langle \theta, y\rangle} K^N(x, dy),$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{R}^d$. Let $S$ be an open subset of $\mathbb{R}^d$ and define $S^N = I^N \cap S$. We are interested in the behaviour of $(X_t^N)_{t\geq 0}$ up until the first time it leaves $S$ (for example, this may mean that one of its co-ordinates hits 0). With this in mind, define a stopping time

$$T^N = \inf\{t \geq 0 : X_t^N \notin S\} \wedge t_0$$

where $t_0 > 0$ is a constant.

**Theorem 1.9 (Darling and Norris).** *Assume the following conditions:*

*(C1) There exists a limit kernel $K(x, dy)$, defined for $x \in S$, and a constant $\eta_0 > 0$ such that $m(x, \theta) < \infty$ for all $x \in S$ and $\|\theta\| \leq \eta_0$.*

*(C2) We have*

$$\sup_{x\in S^N} \sup_{\|\theta\|\leq \eta_0} \left|N^{-1} m^N(x, N\theta) - m(x, \theta)\right| \to 0$$

*as $N \to \infty$.*

*(C3) Let $b(x) = m'(x, 0)$, where $m'(x, \theta)$ is the vector of partial derivatives in components of $\theta$. Then assume that $b$ is Lipschitz on $S$.*

*(C4) $S$ has a Lipschitz boundary.*

*(C5) For all $\delta > 0$,*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\|X_0^N - x_0\| > \delta\right) < 0$$

*for some constant $x_0 \in \bar{S}$.*

*Denote by $(x(t))_{t\geq 0}$ the unique solution to the differential equation $\dot{x}(t) = \tilde{b}(x(t))$ with initial condition $x(0) = x_0$, where $\tilde{b}$ is a Lipschitz extension of $b$ to $\mathbb{R}^d$. Then, for all $\delta > 0$,*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\sup_{0\leq t\leq T^N} \|X_t^N - x(t)\| > \delta\right) < 0.$$

In simple cases, where $x(t)$ does not graze the boundary of $S$ before crossing it, it is straightforward to show that $T^N$ converges in probability to the first exit time of $x(t)$ from $S$.

In the next chapter, we will present an extension of Theorem 1.9, so we refer the reader there for a proof.

# Chapter 2

# Fluid Limits in Infinite Dimensions

The purpose of this chapter is to develop an extension of the fluid limit methods in the Introduction to infinite dimensions. One approach to this is given in Luczak [22], but here we apply the simpler idea of truncation.

Suppose that $(X_t)_{t\geq 0}$ is a pure jump Markov process taking values in a subset $I$ of $\mathbb{R}^{\mathbb{N}}$. Let $K(x,dy)$ be its Lévy kernel, $\mu$ its jump measure and $\nu$ the compensator of its jump measure, so that

$$\begin{aligned}\mu(dt,dy) &= \epsilon_{(t,X_t - X_{t-})}(dt,dy)\\ \nu(dt,dy) &= K(X_{t-},dy)\,dt,\end{aligned}$$

where $\epsilon_x$ denotes unit mass at $x$. Let $a : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^d$ be such that $a_k(y) = y_k$ for $1 \leq k \leq d$. Assume that for all $t < \infty$,

$$\sup_{1\leq k\leq d} \int_0^t \int_{\mathbb{R}^{\mathbb{N}}} |a_k(y)|\nu(ds,dy) < \infty.$$

Let

$$A_t = \int_0^t \int_{\mathbb{R}^{\mathbb{N}}} a(y)(\mu - \nu)(ds,dy). \tag{2.1}$$

Then $(A_t)_{t\geq 0}$ is a $(\mathcal{F}_t^X)_{t\geq 0}$-martingale. For $\theta \in (\mathbb{R}^d)^*$, define

$$m(x,\theta) = \int_{\mathbb{R}^{\mathbb{N}}} e^{\langle\theta,a(y)\rangle} K(x,dy),$$

where $\langle x, y\rangle = \sum_{i=1}^{d} x_i y_i$. Take $\|\cdot\|_d$ to be the supremum norm such that $\|x\|_d = \sup_{1\leq k\leq d}|x_k|$.

The following lemma is adapted from Proposition 4.1 of Darling and Norris [12].

**Lemma 2.1.** *Suppose that there exists $\eta_0 > 0$ such that*

$$\sup_{x\in I}\sup_{\|\theta\|_d\leq\eta_0} m(x,\theta)\leq C$$

*for some constant $C<\infty$. Then for any $\delta\in(0, C't_0\eta_0 d)$,*

$$\mathbb{P}\left(\sup_{0\leq t\leq t_0}\|A_t\|_d>\delta\right)\leq(2d)e^{-\delta^2/2C'dt_0},$$

*where $t_0$ and $C'$ are constant.*

*Proof.* Suppose $M$ is a $d\times d$ matrix. Then, when applied to a matrix, define $\|\cdot\|$ to be the operator norm such that

$$\|M\| = \sup_{1\leq i\leq d}\sum_{j=1}^{d}|M_{ij}|.$$

Fix $\eta\in(0,\eta_0)$. Then there exists $C'<\infty$ such that

$$\|m''(x,\theta)\|\leq C' \quad \text{for all } x\in I,\ \|\theta\|_d\leq\eta,$$

where $m''(x,\theta)$ is the matrix of second-order partial derivatives $\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}m(x,\theta)\right)_{i,j}$. Define for $\theta\in(\mathbb{R}^d)^*$

$$\phi(x,\theta)=\int_{\mathbb{R}^{\mathbb{N}}}\left\{e^{\langle\theta,a(y)\rangle}-1-\langle\theta,a(y)\rangle\right\}K(x,dy).$$

Then $\phi\geq 0$, and, for $\|\theta\|_d\leq\eta$, by the second-order mean value theorem,

$$\phi(x,\theta)=\int_0^1\langle\theta,m''(x,r\theta)\theta\rangle(1-r)dr$$

and so

$$\phi(x,\theta)\leq\frac{1}{2}dC'\|\theta\|_d^2, \quad \text{for all } x\in I,\ \|\theta\|_d\leq\eta.$$

It can be shown that

$$Z_t^\theta=\exp\left(\langle\theta,A_t\rangle-\int_0^t\phi(X_s,\theta)ds\right)$$

is a martingale. Now fix $\theta\in(\mathbb{R}^d)^*$ such that $\|\theta\|_d=1$ and consider the stopping time

$$T=\inf\{t\geq 0:\langle\theta,A_t\rangle>\delta\}.$$

For $0 < \epsilon < \eta$, $(Z_t^{\epsilon\theta})_{t\geq 0}$ is a martingale. On the set $\{T \leq t_0\}$, $Z_T^{\epsilon\theta} \geq \exp(\delta\epsilon - C'dt_0\epsilon^2/2)$. By the Optional Stopping Theorem, $\mathbb{E}\left[Z_{T\wedge t_0}^{\epsilon\theta}\right] = \mathbb{E}\left[Z_0^{\epsilon\theta}\right] = 1$. Hence

$$\mathbb{P}\left(\sup_{0\leq t\leq t_0}\langle\theta, A_t\rangle > \delta\right) = \mathbb{P}\left(T \leq t_0\right) \leq \exp(-\delta\epsilon + C'dt_0\epsilon^2/2).$$

When $\delta < C'dt_0\eta$, we can take $\epsilon = \delta/C'dt_0$ to obtain

$$\mathbb{P}\left(\sup_{0\leq t\leq t_0}\langle\theta, A_t\rangle > \delta\right) \leq \exp(-\delta^2/2C'dt_0).$$

Finally, if $\sup_{0\leq t\leq t_0}\|A_t\|_d > \delta$ then $\sup_{0\leq t\leq t_0}\langle\theta, A_t\rangle > \delta$ for one of $\theta = \pm e_1, \ldots, \pm e_d$. The result follows. $\square$

Now suppose we have a sequence $(X_t^N)_{t\geq 0}$ of pure Markov jump processes, as in the Introduction, but now taking values in $I^N = \frac{1}{N}\mathbb{Z}^{\mathbb{N}} \subseteq \mathbb{R}^{\mathbb{N}}$. For $x \in I^N$, the Lévy kernel of $(X_t^N)_{t\geq 0}$ is given by $K^N(x, dy)$. Let $S$ be an open subset of $\mathbb{R}^{\mathbb{N}}$ (in the topology generated by the $\ell^\infty$-norm $\|\cdot\|$) and let $S^N = I^N \cap S$. Take

$$T^N = \inf\{t \geq 0 : X_t^N \notin S\} \wedge t_0,$$

where $t_0$ is a fixed constant. We commence by imposing some conditions similar to those in the finite-dimensional case. The following must hold for all $d$.

1. There exists a limit kernel $K(x, dy)$, defined for $x \in S$. For $\theta \in (\mathbb{R}^d)^*$, define

$$m(x, \theta) = \int_{\mathbb{R}^{\mathbb{N}}} e^{\langle\theta, a(y)\rangle} K(x, dy)$$

and assume that there exists $\eta_0 > 0$ such that

$$\sup_{x\in S}\sup_{\|\theta\|_d\leq\eta_0} m(x, \theta) < \infty.$$

2. Let

$$m^N(x, \theta) = \int_{\mathbb{R}^{\mathbb{N}}} e^{\langle\theta, a(y)\rangle} K^N(x, dy).$$

Then we assume that

$$\sup_{x\in S^N}\sup_{\|\theta\|_d\leq\eta_0}\left|\frac{m^N(x, N\theta)}{N} - m(x, \theta)\right| \to 0$$

as $N \to \infty$.

3. Let

$$b^N(x) = \int_{\mathbb{R}^{\mathbb{N}}} y\, K^N(x, dy)$$
$$b(x) = \int_{\mathbb{R}^{\mathbb{N}}} y\, K(x, dy).$$

Assume that $b_k^N(x)$ and $b_k(x)$ are finite for all $k \geq 1$ and all $x \in S$ and that for each $k \geq 1$,

$$b_k^N \to b_k$$

as $N \to \infty$, pointwise for $x \in S$. Assume also that there exists a unique solution to the differential equation

$$\dot{x}(t) = b(x(t))$$

in $S$, with given initial condition $x(0)$.

In practice, we will check Condition 3 by finding a metric $d_S$ such that $(S, d_S)$ is a complete metric space and $b$ is Lipschitz in $d_S$. Then we will be able to use the Contraction Mapping Theorem (see Simmons [31], p.338) to prove the existence and uniqueness of a solution.

Now split $x(t)$ by $x(t) = (\hat{x}(t), \bar{x}(t))$ where

$$\hat{x}(t) = (x_1(t), \dots, x_d(t)) \in \mathbb{R}^d$$
$$\bar{x}(t) = (x_{d+1}(t), x_{d+2}(t), \dots) \in \mathbb{R}^{\mathbb{N}}$$

and likewise for $X_t^N = (\hat{X}_t^N, \bar{X}_t^N)$. We have

$$\hat{x}(t) = \hat{x}(0) + \int_0^t \hat{b}(\hat{x}(s), \bar{x}(s))ds$$

and

$$\hat{X}_t^N = \hat{X}_0^N + M_t^N + \int_0^t \hat{b}^N(\hat{X}_s^N, \bar{X}_s^N)ds,$$

where $\hat{b}, \hat{b}^N : \mathbb{R}^d \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^d$ are such that for all $1 \leq k \leq d$,

$$\hat{b}_k(\hat{x}(s), \bar{x}(s)) = b_k(x(s))$$
$$\hat{b}_k^N(\hat{X}_s^N, \bar{X}_s^N) = b_k^N(X_s^N).$$

The process $(M_t^N)_{t \geq 0}$ is defined by

$$M_t^N = \int_0^t \int_{\mathbb{R}^{\mathbb{N}}} a(y)(\mu^N - \nu^N)(ds, dy)$$

and is a $d$-dimensional martingale with respect to the natural filtration of $(X_t^N)_{t\geq 0}$, as at (2.1). Let

$$\begin{aligned} E_t^N &= \hat{b}_N(\hat{X}_t^N, \bar{X}_t^N) - \hat{b}_N(\hat{X}_t^N, 0) \\ e(t) &= \hat{b}(\hat{x}(t), \bar{x}(t)) - \hat{b}(\hat{x}(t), 0), \end{aligned}$$

the errors we make by truncating at the $d$th co-ordinate. Note that for any $x \in S$ and $1 \leq k \leq d$,

$$\begin{aligned} \hat{b}_k^N(\hat{x}, \bar{x}) &= \frac{\partial}{\partial \theta_k} m^N(x, \theta)\Big|_{\theta=0} \\ \hat{b}_k(\hat{x}, \bar{x}) &= \frac{\partial}{\partial \theta_k} m(x, \theta)\Big|_{\theta=0}. \end{aligned} \tag{2.2}$$

Now assume the following additional conditions:

4. For all $\delta > 0$ and all $d > 0$,

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\|\hat{X}_0^N - \hat{x}(0)\|_d > \delta\right) < 0.$$

5. For all $\delta > 0$, there exists a $d$ sufficiently large that

$$\sup_{0\leq t\leq t_0} \sup_{k\geq d+1} |x_k(t)| < \delta.$$

6. For each $\delta > 0$, there exists a $d$ sufficiently large that

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\sup_{0\leq t\leq T^N} \sup_{k\geq d+1} \left|X_t^{k,N}\right| > \delta\right) < 0.$$

7. $\hat{b}(x, 0)$ is Lipschitz in its first argument i.e. there exists a constant $L$ such that for $x, y \in \mathbb{R}^d$,

$$\|\hat{b}(x, 0) - \hat{b}(y, 0)\|_d \leq L\|x - y\|_d.$$

8. For each $\delta > 0$, there exists a $d > 0$ such that

$$\sup_{0\leq t\leq t_0} \|e(t)\|_d < \delta.$$

9. For each $\delta > 0$, there exists a $d > 0$ such that

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\sup_{0\leq t\leq T^N} \|E_t^N\|_d > \delta\right) < 0.$$

**Theorem 2.2.** *Recall that $\|\cdot\|$ is the $\ell^\infty$-norm on $\mathbb{R}^{\mathbb{N}}$. Assume Conditions 1 to 9. Then, for each $\delta > 0$,*

$$\limsup_{N\to\infty} N^{-1}\log\mathbb{P}\left(\sup_{0\le t\le T^N}\|X_t^N - x(t)\| > \delta\right) < 0.$$

*Proof.* By 5 and 6, for any $\delta > 0$ there exists $d_0$ such that for all $d \ge d_0$,

$$\limsup_{N\to\infty} N^{-1}\log\mathbb{P}\left(\sup_{0\le t\le T^N}\sup_{k\ge d+1}\left|X_t^{k,N} - x^k(t)\right| > \delta/2\right) < 0.$$

Hence, it is sufficient to show that for $d \ge d_0$,

$$\limsup_{N\to\infty} N^{-1}\log\mathbb{P}\left(\sup_{0\le t\le T^N}\|\hat{X}_t^N - \hat{x}(t)\|_d > \delta/2\right) < 0.$$

Now,

$$\begin{aligned}\|\hat{X}_t^N - \hat{x}(t)\|_d \le \|\hat{X}_0^N - \hat{x}(0)\|_d + \|M_t^N\|_d \\ + \int_0^t \|\hat{b}^N(\hat{X}_s^N, 0) - \hat{b}(\hat{x}(s), 0)\|_d \, ds + \int_0^t (\|E_s^N\|_d + \|e(s)\|_d) ds.\end{aligned}$$

Let $\epsilon = \delta e^{-Lt_0}/10$. By Condition 2 and (2.2), there exists $N_0$ such that

$$\|\hat{b}^N(x,0) - \hat{b}(x,0)\|_d < \epsilon/t_0$$

for all $x \in S^N$ and all $N \ge N_0$. Thus, by the Lipschitz property of $\hat{b}$ (Condition 7), we have

$$\|\hat{b}^N(\hat{X}_s^N, 0) - \hat{b}(\hat{x}(s), 0)\|_d \le L\,\|\hat{X}_s^N - \hat{x}(s)\|_d + \epsilon/t_0.$$

Now for $(M_t^N)_{t\ge 0}$. Conditions 1 and 2 imply that

$$\sup_N \sup_{x\in S^N} \sup_{\|\theta\|_d \le \eta_0} m^N(x, N\theta)/N < \infty.$$

Fix $\eta \in (0, \eta_0)$. Then there exists $C' < \infty$ such that, for all $N$,

$$\left\|m^{N''}(x,\theta)\right\| \le C'/N, \quad x \in S^N, \quad \|\theta\|_d \le N\eta.$$

Thus, applying Lemma 2.1, we can find constants $\epsilon_0 > 0$ and $C_0 < \infty$ depending only on $C', \eta, d$ and $t_0$ such that for all $N$ and all $\epsilon \in (0, \epsilon_0]$,

$$\mathbb{P}\left(\sup_{0\le t\le T^N}\|M_t^N\|_d > \epsilon\right) \le C_0 e^{-N\epsilon^2/C_0}. \tag{2.3}$$

Let

$$\Omega_N = \Bigg\{ \|\hat{X}_0^N - \hat{x}(0)\|_d \le \epsilon, \sup_{0\le t\le T^N} \|M_t^N\|_d \le \epsilon,$$

$$\sup_{0\le t\le T^N} \sup_{k\ge d+1} \left|X_t^{k,N}\right| \le \epsilon/Ct_0, \sup_{0\le t\le T^N} \|E_t^N\|_d \le \epsilon/t_0 \Bigg\}.$$

Then by (2.3), 4, 6 and 9, there exists $d_1$ such that for all $d \ge d_1$,

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\Omega_N^c\right) < 0.$$

Also, by 8, there exists $d_2$ such that for all $d \ge d_2$,

$$\sup_{0\le t\le t_0} \|e(t)\|_d \le \epsilon/t_0.$$

Thus, for $N \ge N_0$ and $d \ge \max\{d_0, d_1, d_2\}$ on $\Omega_N$,

$$\|\hat{X}_t^N - \hat{x}(t)\|_d \le 5\epsilon + L\int_0^t \|\hat{X}_s^N - \hat{x}(s)\|_d \, ds$$

and so, by Gronwall's lemma (p.543 of Revuz and Yor [28]),

$$\sup_{0\le t\le T^N} \|\hat{X}_t^N - \hat{x}(t)\|_d \le 5\epsilon e^{Lt_0} = \delta/2.$$

□

It may seem that we are unlikely to be able to fulfil all of these conditions in applications. However, most of them are quite natural. The most important new conditions are 5 and 6, which ensure that we do not make too much difference by truncating at a finite level. Chapters 3 and 4 give examples of situations where these conditions are satisfied, in one case by a deterministic upper bound and in the other by the use of a coupling and stochastic domination.

# Chapter 3

# Robustness of the Poisson Model

In this chapter, we aim to demonstrate that asymptotic results for the Poisson random hypergraph remain true if we alter the original distributions of the hyperedges. More precisely, we will investigate what happens if we take $Z_0^{k,N}$ $k$-edges and distribute them uniformly over the $k$-subsets of the vertex set, where $Z_0^{k,N}$ has mean $N\beta_k$ (in the Poisson case, we had $Z_0^{k,N} \sim \text{Poisson}(N\beta_k)$). Note that, in general, this means that we lose the independence between numbers of hyperedges over distinct sets of the same size which was a feature of the Poisson model. Instead, we have the weaker property of exchangeability.

We will concentrate on proving that, under certain conditions, the asymptotic proportion of identifiable vertices is the same as in the Poisson model. In doing so, we will learn more about the process of identification. Recall that we have assumed that $\beta(1) < \infty$. To avoid unnecessary complication, we shall restrict attention to the case where the function $\beta'(t) + \log(1-t)$ has no zeros in $[0, t^*)$, where $t^* = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$ and $t^* < 1$.

## 3.1 Multinomial starting conditions

Suppose we start with a deterministic number $\lfloor N\beta_k \rfloor$ of $k$-edges, for $k \geq 1$, each of which is equally likely to be over each of the $\binom{N}{k}$ $k$-subsets of the vertex set. Thus, the vector giving the number of hyperedges over all $\binom{N}{k}$ sets of size $k$ has distribution

$$\text{Multinomial}\left(\lfloor N\beta_k \rfloor \; ; \; 1/\binom{N}{k}, \; 1/\binom{N}{k}, \ldots, 1/\binom{N}{k}\right).$$

The number of hyperedges over sets of different sizes are, initially, independent.

Now consider the process of identification, as in Section 1.6 of the Introduction. Let $Z_n^{k,N}$ be the number of $k$-edges on the graph after $n$ vertex deletions, for $k \geq 1$. We have assumed $Z_0^{k,N} = \lfloor N\beta_k \rfloor$ for $k \geq 1$. Given $\mathcal{F}_n^N = \sigma(Z_m^N : m \leq n)$ and, provided $Z_n^{1,N} > 0$, we have

$$Z_{n+1}^{1,N} = Z_n^{1,N} - 1 - Y_{n+1}^{1,N} + Y_{n+1}^{2,N}$$

and for $k \geq 2$,

$$Z_{n+1}^{k,N} = Z_n^{k,N} - Y_{n+1}^{k,N} + Y_{n+1}^{k+1,N},$$

where $Y_{n+1}^{1,N} \sim \mathrm{Bin}(Z_n^{1,N} - 1,\ 1/(N-n))$ and, for $k \geq 2$, $Y_{n+1}^{k,N} \sim \mathrm{Bin}(Z_n^{k,N},\ k/(N-n))$. This shows that $(Z_n^N)_{n\geq 0}$ is a Markov chain. We stop the process when $Z_n^{1,N}$ hits 0 (i.e. when there are no patches left). The distributions of the $Y_{n+1}^{k,N}$ arise as follows. For $k = 1$, we must delete at least one patch and so $Y_{n+1}^{1,N}$ is the number of other patches sharing the same vertex, which also get deleted. After $n$ deletions, we had $Z_n^{1,N}$ patches on the hypergraph, which were uniformly scattered across the remaining $N - n$ vertices. Thus, once we have picked a patch, the numbers of other patches on the different vertices are distributed as Multinomial$(Z_n^{1,N} - 1; 1/(N-n), \dots, 1/(N-n))$. So the number on a specific vertex is $\mathrm{Bin}(Z_n^{1,N} - 1, 1/(N-n))$ and this is the distribution of $Y_{n+1}^{1,N}$. The distributions of $Y_{n+1}^{k,N}$, $k \geq 2$, arise in a similar way: there are $X_n^{k,N}$ $k$-edges on the hypergraph after the $n$th vertex deletion, which are uniformly scattered across the remaining $k$-sets, of which there are $\binom{N-n}{k}$. When we delete a vertex, that shrinks $\binom{N-n-1}{k-1}$ of these sets, thus causing a $\mathrm{Bin}(Z_n^{k,N}, k/(N-n))$ number of the $k$-edges to become $(k-1)$-edges.

Thus,

$$\mathbb{E}\left[Z_{n+1}^{1,N} - Z_n^{1,N} | Z_n^N\right] = \frac{2}{(N-n)} Z_n^{2,N} - \frac{1}{(N-n)}(Z_n^{1,N} - 1) - 1$$

and for $k \geq 2$,

$$\mathbb{E}\left[Z_{n+1}^{k,N} - Z_n^{k,N} | Z_n^N\right] = \frac{(k+1)}{(N-n)} Z_n^{k+1,N} - \frac{k}{(N-n)} Z_n^{k,N}.$$

Unlike in the Poisson case (see the proof of Theorem 1.6), we cannot now de-couple $(Z_n^{1,N})_{n\geq 0}$ from the rest of the process, so in order to prove a fluid limit for it, we will have to prove one for the whole infinite-dimensional system. We are, therefore, looking for a deterministic function $z : \mathbb{R}^+ \to \mathbb{R}^{\mathbb{N}}$ such that $(N^{-1} Z_{\lfloor Nt \rfloor}^N)_{t\geq 0} \xrightarrow{p} (z(t))_{t\geq 0}$ (in an appropriate sense). Supposing such a limit exists, heuristically it must satisfy the following system of differential equations:

$$\dot{z}_1(t) = \frac{1}{(1-t)}(2z_2(t) - z_1(t)) - 1$$

and for $k \geq 2$,

$$\dot{z}_k(t) = \frac{1}{(1-t)}((k+1)z_{k+1}(t) - kz_k(t)),$$

with initial conditions $z_k(0) = \beta_k$ for $k \geq 1$. This system has solution

$$\begin{aligned} z_1(t) &= (1-t)(\beta'(t) + \log(1-t)) \\ z_k(t) &= \frac{1}{k!}(1-t)^k \beta^{(k)}(t) \quad \text{for } k \geq 2. \end{aligned} \tag{3.1}$$

Note that if we define $y_k(t) = \frac{k!}{(1-t)^k} z_k(t)$ and $F(t, y(t)) = \dot{y}(t)$ then

$$\begin{aligned} F_1(t, y(t)) &= y_2(t) - \frac{1}{1-t} \\ F_k(t, y(t)) &= y_{k+1}(t) \quad \text{for } k \geq 2. \end{aligned}$$

It is clear that $F$ is Lipschitz (in its second argument), with constant 1, in the $\ell^\infty-$norm $\|x\| = \sup_{k\geq 1} |x_k|$. Thus, by a standard argument using the Contraction Mapping Theorem, the differential equation $\dot{y}(t) = F(t, y(t))$ has a unique solution. Thus $z(t)$ as defined at (3.1) is the unique solution to the original system of equations in $\ell^\infty$.

Because the system is infinite dimensional, we will use the fluid limit methodology developed in Chapter 2. It is plausible that the truncation idea will work here because the total number of hyperedges of size greater than $k$ is monotonically decreasing in time. We will find it convenient to work with a Poissonized version of $Z^N$, so that we have a pure jump Markov process. Let $(\nu_t^N)_{t\geq 0}$ be a Poisson process of rate $N$. Let $X_t^{0,N} = N^{-1}\nu_t^N$ and, for $k \geq 1$, $X_t^{k,N} = N^{-1}Z_{\nu_t^N}^{k,N}$. Let $x_0(t) = t$, $x_1(t) = (1-t)(\beta'(t) + \log(1-t))$ and, for $k \geq 2$, $x_k(t) = \frac{1}{k!}(1-t)^k\beta^{(k)}(t)$. Let $T^N = \inf\{t \geq 0 : X_t^{1,N} \leq 0\}$.

**Theorem 3.1.** *For any $\delta > 0$,*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left( \sup_{0\leq t\leq T^N \wedge t^*} \|X_t^N - x(t)\| > \delta \right) < 0.$$

*Proof.* We must check Conditions 1 to 9 of Theorem 2.2. The Lévy kernel $K^N(x, dy)$ of $(X_t^N)_{t\geq 0}$ is such that $N^{-1}K^N(x, \cdot)$ is the law of the random variable

$$N^{-1}(1,\ Y_2 - Y_1 - 1,\ Y_3 - Y_2,\ Y_4 - Y_3, \dots)$$

where $Y_1 \sim \mathrm{Bin}(Nx_1 - 1, 1/(N - Nx_0))$ and, for $k \geq 2$, $Y_k \sim \mathrm{Bin}(Nx_k, k/(N - Nx_0))$. Thus, $Y_1 \xrightarrow{d} \mathrm{Poisson}\,(x_1/(1-x_0))$ and, for $k \geq 2$, $Y_k \xrightarrow{d} \mathrm{Poisson}\,(kx_k/(1-x_0))$, so the limit kernel $K(x, dy)$ exists. Moreover, for $\theta \in \mathbb{R}^d$, we have

$$
\begin{aligned}
m(x,\theta) &= \int_{\mathbb{R}^{\mathbb{N}}} e^{\langle\theta, a(y)\rangle} K(x, dy) \\
&= \exp\left(\theta_0 - \theta_1 + \frac{x_1}{1-x_0}(e^{-\theta_1} - 1) + \sum_{k=2}^{d} \frac{kx_k}{1-x_0}(e^{\theta_{k-1}-\theta_k} - 1) + \frac{(d+1)x_{d+1}}{1-x_0}(e^{\theta_d} - 1)\right).
\end{aligned}
$$

Fix $0 < r_0 < 1$ and $0 < r < \infty$ and let

$$S = \{x \in \mathbb{R}^{\mathbb{N}} : |x_0| < r_0, 0 < x_1 < r, |x_k| < r \text{ for } k \geq 2\}.$$

Then there exists $\eta_0 < \infty$ such that

$$\sup_{x \in S} \sup_{\|\theta\|_d \leq \eta_0} m(x,\theta) < \infty,$$

which fulfils Condition 1. Now,

$$
\begin{aligned}
m^N(x,\theta) &= N\mathbb{E}\left[\exp\left(\frac{\theta_0}{N} - \frac{\theta_1}{N}Y_1 - \frac{\theta_1}{N} + \sum_{k=2}^{d}\left(\frac{\theta_{k-1}}{N} - \frac{\theta_k}{N}\right)Y_k + \frac{\theta_d}{N}Y_{d+1}\right)\right] \\
&= N\exp\left\{\frac{1}{N}(\theta_0 - \theta_1) + (Nx_1 - 1)\log\left(1 + (e^{-\theta_1/N} - 1)\frac{1}{N - Nx_0}\right)\right. \\
&\qquad + \sum_{k=2}^{d} Nx_k \log\left(1 + (e^{(\theta_{k-1}-\theta_k)/N} - 1)\frac{k}{N - Nx_0}\right) \\
&\qquad \left. + Nx_{d+1}\log\left(1 + (e^{\theta_d/N} - 1)\frac{(d+1)}{N - Nx_0}\right)\right\}
\end{aligned}
$$

and so

$$\sup_{x \in S^N} \sup_{\|\theta\|_d \leq \eta_0} \left|\frac{m^N(x, N\theta)}{N} - m(x,\theta)\right| \to 0$$

as $N \to \infty$, which is Condition 2.

Now turn to Condition 3. For $x \in S^N$,

$$
\begin{aligned}
b_0^N(x) &= 1 \\
b_1^N(x) &= \frac{2x_2}{1-x_0} - \frac{(x_1 - \frac{1}{N})}{1-x_0} - 1 \\
b_k^N(x) &= \frac{(k+1)x_k}{1-x_0} - \frac{kx_k}{1-x_0} \quad \text{for } k \geq 2,
\end{aligned}
$$

which clearly converges to $b$ defined for $x \in S$ by

$$\begin{aligned} b_0(x) &= 1 \\ b_1(x) &= \frac{2x_2}{1-x_0} - \frac{x_1}{1-x_0} - 1 \\ b_k(x) &= \frac{(k+1)x_k}{1-x_0} - \frac{kx_k}{1-x_0} \quad \text{for } k \geq 2. \end{aligned}$$

We have already demonstrated that there is a unique solution to this system in $\ell^\infty$ and by taking $r_0$ and $r$ large enough in the definition of $S$, we can ensure that the solution lies in $S$ until $t^*$, where $t^* = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$.

We have $X_0^{k,N} = \frac{\lfloor N\beta_k \rfloor}{N} \to \beta_k$ deterministically, uniformly in $k \geq 1$, and that $X_0^{0,N} = 0$. Hence, Condition 4 is trivially satisfied. For any $d$, $\sup_{0\leq t\leq t^*} \sup_{k \geq d+1} |x_k(t)| \leq \sum_{k=d+1}^\infty \beta_k$, which is decreasing in $d$ because $\sum_{k=1}^\infty \beta_k < \infty$. Thus, given $\epsilon > 0$, we can find $d$ such that $\sum_{k=d+1}^\infty \beta_k < \epsilon$, which is Condition 5. Likewise, for any $d$,

$$\sup_{0 \leq t \leq T^N} \sup_{k \geq d+1} \left| X_t^{k,N} \right| \leq \frac{1}{N} \sum_{k=d+1}^\infty \lfloor N\beta_k \rfloor \leq \sum_{k=d+1}^\infty \beta_k$$

and so the same bounds apply, giving Condition 6.

Truncating at level $d$, we obtain

$$\begin{aligned} \hat{b}_k(\hat{x}, \bar{x}) - \hat{b}_k(\hat{x}, 0) &= 0 \quad \text{for } 0 \leq k \leq d-1 \\ \hat{b}_d(\hat{x}, \bar{x}) - \hat{b}_d(\hat{x}, 0) &= \frac{(d+1)}{1-x_0} x_{d+1} \end{aligned}$$

and

$$\begin{aligned} \hat{b}_k^N(\hat{x}, \bar{x}) - \hat{b}_k^N(\hat{x}, 0) &= 0 \quad \text{for } 0 \leq k \leq d-1 \\ \hat{b}_d^N(\hat{x}, \bar{x}) - \hat{b}_d^N(\hat{x}, 0) &= \frac{(d+1)}{1-x_0} x_{d+1}. \end{aligned}$$

Thus, for $x, y \in \mathbb{R}^d$,

$$\|\hat{b}(x,0) - \hat{b}(y,0)\|_d \leq \frac{(d+1)}{(1-r_0)} \|x-y\|_d$$

and so Condition 7 is fulfilled. Furthermore,

$$\begin{aligned} \sup_{0\leq t\leq t^*} \|e(t)\|_d &\leq \frac{(d+1)}{(1-r_0)} \sum_{k=d+1}^\infty \beta_k \\ \sup_{0\leq t\leq T^N} \|E_t^N\|_d &\leq \frac{(d+1)}{(1-r_0)} \sum_{k=d+1}^\infty \beta_k \end{aligned}$$

and so Conditions 8 and 9 also hold. Thus, the result follows from Theorem 2.2. $\square$

**Corollary 3.2.** *The proportion of vertices identifiable in the hypergraph converges in probability to $t^*$.*

*Proof.* As we have assumed that $\beta'(t)+\log(1-t)$ has no zeros in $[0,t^*)$, it follows directly from Theorem 3.1 that $T^N \xrightarrow{p} t^*$. □

## 3.2 General starting conditions

The same analysis will go through as long as for any $\delta > 0$ and all $k \geq 1$,

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\left|X_0^{k,N} - \beta_k\right| > \delta\right) < 0. \tag{3.2}$$

A sufficient condition is that for all $k \geq 1$,

$$\Lambda^k(\theta) = \log \mathbb{E}\left[e^{\theta X_0^{k,N}}\right]$$

should have a limit

$$\Lambda^k(\theta) = \lim_{N\to\infty} N^{-1}\Lambda_N^k(N\theta)$$

as an extended real number, with 0 in the interior of $\{\theta : \Lambda^k(\theta) < \infty\}$. This entails the Gärtner–Ellis Theorem (see Dembo and Zeitouni [13], p.44), which implies (3.2). Only minor alterations then need to be made to the proof of Theorem 3.1 to make it work for these starting conditions also.

We have demonstrated that a large class of distributions for the initial number of hyperedges result in the same limiting proportion of identifiable vertices as in the Poisson random hypergraph. Moreover, we have learnt the limiting number of $k$-edges per vertex at all stages of the process of identification.

# Chapter 4

# Distribution of the Patches

## 4.1 System

In Section 1.6 of the Introduction, we discussed Darling and Norris' derivation of a fluid limit path for the number of patches on the Poisson random hypergraph at each stage of the process of identification. Here we seek to decompose this limit. How are the patches distributed over the vertices present?

We shall approach this problem by looking at the evolution of the infinite-dimensional process $(X_n^N)_{n\geq 0}$ where

$$X_n^N = (X_n^{0,N}, X_n^{1,N}, X_n^{2,N}, \dots)$$

and $X_n^{i,N}$ is the number of vertices with $i$ patches on them just after the $n$th vertex deletion (the superscript $N$ is to emphasise the fact that we are working with the Poisson random hypergraph on $N$ vertices). Let $B_1^N, B_2^N, \dots, B_N^N$ be the numbers of patches on vertices $1, 2, \dots, N$ respectively; these random variables are independent and identically distributed with Poisson($\beta_1$) distribution. We have

$$X_0^{i,N} = \sum_{k=1}^{N} \mathbb{1}_{\{B_k^N = i\}}$$

for all $i \geq 0$. At each stage of the process, we pick a patch at random and delete it and all others on the same vertex. If that was the $n$th vertex deletion, imagine we re-label the remaining vertices $1, \dots, N-n$ in a random order. Now call the number of patches which are added to vertex $k$ (caused by the collapse of 2-edges) $A_n^{k,N}$, for $k = 1, \dots, N-n$. As discussed in the Introduction, we have $A_n^{k,N} \sim \text{Poisson}(\lambda_2(N,n))$, where $A_n^{1,N}, \dots, A_n^{N-n,N}$ are independent and identically distributed. They are independent (but not identically distributed) for different values of $n$ and are independent of $B_1^N, B_2^N, \dots, B_N^N$.

We take the better understanding of identifiability in the Poisson random hypergraph as our primary motivation, but we observe that the above model can be translated into a rather nice urn model in the classical style of Feller [17]. Imagine that we have an urn and balls of $N$ different colours (so that colours correspond to vertices and balls to patches). Put $B_i^N$ balls of colour $i$ into the urn, for $i = 1, 2, \dots, N$. Now pick a ball uniformly at random and remove it along with all other balls of the same colour from the urn. This colour is now 'dead' and will take no further part. Now randomly re-label the colours and add $A_1^{k,N}$ balls of colour $k$ for $k = 1, 2, \dots, N-1$. Repeat this process, with $A_n^{k,N}$ balls of colour $k$ added at stage $n$, until the urn becomes empty. Now $X_n^{i,N}$ is the number of colours in the urn with $i$ balls after $n$ deletions, for $i \geq 1$, and $X_n^{0,N}$ is the number of colours which have not yet appeared in the urn by just after the $n$th deletion. (Note that the number of deleted colours is $N - \sum_{i=1}^{\infty} X_n^{i,N}$.) Considered thus, we have a generic process which could model different situations. We give here three examples, with questions that we might like to answer.

*Example 1: Stochastic epidemic.* A population of $N$ individuals is attacked by a fatal disease, carried by certain 'particles'. At each epoch, one infected individual dies, with individuals being chosen with probability proportional to the number of infective particles they have been exposed to. When an individual dies, a random number of infective particles are released and attack the other individuals. When there are no longer any infected individuals, the disease dies out. What proportion of the population is killed? At all stages of the epidemic, how many individuals are afflicted and how much exposure have they had?

*Example 2: Queueing system.* $N$ tour guides are waiting to show people around an attraction. Visitors are equally likely to choose each one of them, with a random number joining a queue for each guide in each time interval. Once every time interval, a tour leaves, the guide to lead it being chosen with probability proportional to the number of individuals in his queue. Once a guide has left, he does not return and, if there are no customers left, the remaining guides all go home for the day. What proportion of the guides give tours? At each stage, how many people are waiting for tours?

*Example 3: Nuclear reaction with by-products.* There are $N$ types of particle. Each particle present decays at rate 1 and, when it decays, takes all particles of the same type with it and causes random numbers of the other types of particle to enter the system. Once a type of particle has been destroyed, that type can no longer enter the system. How many particles of each type are present at each stage?

Example 3 is physically unrealistic, but it will be useful to us as a way of thinking about this model in Section 4.5. With these examples in mind, we will impose slightly less restrictive conditions on the distributions of the $B_i^N$'s and $A_n^{k,N}$'s (see Section 4.2), but will retain the independence we have in the Poisson random hypergraph model. We will also allow $\beta''$ to be a more general function than the power series with non-negative coefficients that we have in the hypergraph.

Let $\mathcal{F}_n^N = \sigma(X_m^N : m \leq n)$ be the natural filtration of the process $(X_n^N)_{n\geq 0}$. Then, given $\mathcal{F}_n^N$, $X_{n+1}^N$ depends only on $X_n^N$ and the random variables $A_n^{1,N}, \ldots, A_n^{N-n,N}$, which are independent of $\mathcal{F}_n^N$. Thus $(X_n^N)_{n\geq 0}$ is a Markov chain. Note that, almost surely, only finitely many of the co-ordinates of $X_n^N$ are non-zero.

## 4.2 Conditions

Suppose that $B_1^N, B_2^N, \ldots, B_N^N$ are independent and identically distributed non-negative random variables with

$$B_1^N \xrightarrow{d} \text{Poisson}(\beta_1) \tag{4.1}$$

as $N \to \infty$. Assume that the function $\beta''$ is non-negative and Lipschitz continuous on $[0,1]$ and is the differential of a function $\beta'$ with $\beta'(0) = \beta_1$. Define $\lambda = \sup_{0\leq t\leq 1} \beta''(t)$. To avoid unnecessary complication, we will assume that the function $\beta'(t) + \log(1-t)$ has no zeros in $[0, t^*)$, where $t^* = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$, and that $t^* < 1$. Suppose that $A_n^{1,N}, \ldots, A_n^{N-n,N}$ are independent and identically distributed non-negative random variables, independent but not necessarily identically distributed for different values of $n$. Let $p_k(N,n) = \mathbb{P}\left(A_n^{1,N} = k\right)$ for $k \geq 0$ and $p_{>0}(N,n) = \mathbb{P}\left(A_n^{1,N} > 0\right)$. Let $r < 1$. Then we assume that there exists $\eta > 0$ such that

$$N \sum_{k=2}^{\infty} e^{\theta k} \left( \sup_{0\leq n\leq \lfloor Nr \rfloor} p_k(N,n) \right) \to 0 \tag{4.2}$$

as $N \to \infty$, uniformly in $|\theta| \leq \eta$, and

$$N p_1(N, Nt) \to \beta''(t) \tag{4.3}$$

as $N \to \infty$, uniformly for $t \in \left\{\frac{1}{N}, \frac{2}{N}, \dots, \frac{\lfloor Nr \rfloor}{N}\right\}$.

**Proposition 4.1.** *Conditions (4.2) and (4.3) imply that*

$$N \log \mathbb{E}\left[e^{\theta A^{1,N}_{Nt}}\right] \to \beta''(t)(e^\theta - 1) \tag{4.4}$$

*uniformly in $|\theta| \leq \eta$ and $t \in \left\{\frac{1}{N}, \frac{2}{N}, \dots, \frac{\lfloor Nr \rfloor}{N}\right\}$, as $N \to \infty$.*

*Proof.* We have

$$\begin{aligned} N \log \mathbb{E}\left[e^{\theta A^{1,N}_{Nt}}\right] &= N \log \left(\sum_{k=0}^{\infty} e^{\theta k} p_k(N, Nt)\right) \\ &= N \log \left(1 + p_1(N, Nt)(e^\theta - 1) + \sum_{k=2}^{\infty} (e^{\theta k} - 1) p_k(N, Nt)\right) \\ &\to \beta''(t)(e^\theta - 1) \end{aligned}$$

uniformly in $|\theta| \leq \eta$ and $t \in \left\{\frac{1}{N}, \frac{2}{N}, \dots, \frac{\lfloor Nr \rfloor}{N}\right\}$, as $N \to \infty$, by (4.2) and (4.3). $\square$

Condition (4.1) enables us to prove a simple result about $X_0^{k,N}$ for $k \geq 0$.

**Proposition 4.2.** *For each $k \geq 0$ and all $\epsilon > 0$,*

$$\limsup_{N \to \infty} N^{-1} \log \mathbb{P}\left(\left|N^{-1} X_0^{k,N} - e^{-\beta_1} \beta_1^k / k!\right| > \epsilon\right) < 0.$$

*Proof.* We have

$$X_0^{k,N} = \sum_{i=1}^{B} \mathbb{1}_{\{B_k^N = i\}} \sim \mathrm{Bin}(N, \mathbb{P}\left(B_1^N = k\right))$$

and so, because $\mathbb{P}\left(B_1^N = k\right) \to e^{-\beta_1} \beta_1^k / k!$, the result follows from the Gärtner–Ellis Theorem (see p.44 of Dembo and Zeitouni [13]). $\square$

## 4.3 Heuristic analysis of the process

We wish to prove a fluid limit result for $(X_n^N)_{n \geq 0}$ so we will begin, as usual, by examining the one-step conditional increments. Recall that we must also deal with running out of patches. Temporarily, abbreviate $p_k(N, n+1)$ to $p_k$. Then, conditional on at least one patch remaining at time $n$,

$$
\begin{aligned}
\mathbb{E}\left[X_{n+1}^{0,N} - X_n^{0,N} | \mathcal{F}_n^N\right] &= -p_{>0} X_n^{0,N} \\
\mathbb{E}\left[X_{n+1}^{1,N} - X_n^{1,N} | \mathcal{F}_n^N\right] &= p_1 X_n^{0,N} - p_{>0} X_n^{1,N} - \frac{p_0 X_n^{1,N}}{\sum_{j=1}^{\infty} j X_n^{j,N}} \\
\mathbb{E}\left[X_{n+1}^{2,N} - X_n^{2,N} | \mathcal{F}_n^N\right] &= p_2 X_n^{0,N} + p_1 X_n^{1,N} - p_{>0} X_n^{2,N} - \frac{2p_0 X_n^{2,N} + p_1 X_n^{1,N}}{\sum_{j=1}^{\infty} j X_n^{j,N}} \\
&\vdots \\
\mathbb{E}\left[X_{n+1}^{k,N} - X_n^{k,N} | \mathcal{F}_n^N\right] &= \sum_{i=0}^{\infty} p_{k-i} X_n^{i,N} - p_{>0} X_n^{k,N} - \frac{\sum_{i=1}^{k} i p_{k-i} X_n^{i,N}}{\sum_{j=1}^{\infty} j X_n^{j,N}} \\
&\vdots
\end{aligned}
$$

The second line, for example, arises as follows: at the $(n+1)$th epoch, we pick a patch which is the only patch on its vertex with probability $X_n^{1,N} / \sum_{j=1}^{\infty} j X_n^{j,N}$. Suppose we did pick such a patch. Then we have $X_n^{1,N} - 1$ vertices with one patch left, each of which is promoted to being a vertex with more than one patch with probability $p_{>0}$. There are also vertices which had no patches at time $n$. Each of these vertices is promoted to having one patch at time $n+1$ with probability $p_1$. Alternatively, suppose we initially picked a patch which came from a vertex with more than one patch. Then we have $X_n^{1,N}$ vertices with one patch, each of which is promoted to having more than one patch with probability $p_{>0}$. The vertices with no patches behave as before. So, putting this together,

$$
\mathbb{E}\left[X_{n+1}^{1,N} | \mathcal{F}_n^N\right] = p_1 X_n^{0,N} + p_0 (X_n^{1,N} - 1) \frac{X_n^{1,N}}{\sum_{j=1}^{\infty} j X_n^{j,N}} + p_0 X_n^{1,N} \left(1 - \frac{X_n^{1,N}}{\sum_{j=1}^{\infty} j X_n^{j,N}}\right)
$$

which gives the desired expression on re-arrangement.

Now, by (4.2) and (4.3), as $N \to \infty$, $Np_{>0}(N, Nt) \to \beta''(t)$, $Np_1(N, Nt) \to \beta''(t)$, $p_0(N, Nt) \to 1$ and for $k \geq 2$, $Np_k(N, Nt) \to 0$. So, using the fluid scaling, any deterministic limit must, heuristically, satisfy the following system of differential equations:

$$\begin{aligned}
\dot{x}_0(t) &= -\beta''(t)x_0(t)\\
\dot{x}_1(t) &= \beta''(t)x_0(t) - \beta''(t)x_1(t) - \frac{x_1(t)}{\sum_{i=1}^{\infty} ix_i(t)}\\
\dot{x}_2(t) &= \beta''(t)x_1(t) - \beta''(t)x_2(t) - \frac{2x_2(t)}{\sum_{i=1}^{\infty} ix_i(t)}\\
&\vdots\\
\dot{x}_k(t) &= \beta''(t)x_{k-1}(t) - \beta''(t)x_k(t) - \frac{kx_k(t)}{\sum_{i=1}^{\infty} ix_i(t)} \qquad (4.5)\\
&\vdots
\end{aligned}$$

Expression (4.1) gives initial conditions

$$x_k(0) = \frac{e^{-\beta_1}\beta_1^k}{k!} = \frac{e^{-\beta'(0)}(\beta'(0))^k}{k!},\ k \geq 0,$$

and so it is clear from (4.5) that we must have $x_0(t) = \exp(-\beta'(t))$. In order to find the rest of the solution, consider the Poisson random hypergraph. There, we have $A_1^{n,N} \sim \text{Poisson}(\lambda_2(N,n))$ with $N\lambda_2(N,Nt) \to \beta''(t)$. Moreover, because of the splitting and addition properties of Poisson random variables, each remaining vertex has a Poisson number of patches and the numbers on different vertices are independent. But what is the parameter? We know by Theorem 1.6 that the total number, $P_n^N$, of patches on the hypergraph after $n$ deletions satisfies

$$N^{-1}P_{\lfloor Nt \rfloor}^N \xrightarrow{p} (1-t)(\beta'(t) + \log(1-t)),$$

where the $(1-t)$ factor arises because we are normalising by $N$ rather than $(N - \lfloor Nt \rfloor)$, which is the number of remaining vertices. Thus, we must have $\beta'(t) + \log(1-t)$ as the asymptotic parameter and by a law of large numbers heuristic, we expect that

$$\begin{aligned}
N^{-1}X_{\lfloor Nt \rfloor}^{k,N} &\approx (1-t)\frac{1}{N - \lfloor Nt \rfloor}X_{\lfloor Nt \rfloor}^{k,N}\\
&\xrightarrow{p} (1-t)\frac{\exp(-\beta'(t) - \log(1-t))(\beta'(t) + \log(1-t))^k}{k!}\\
&= \frac{\exp(-\beta'(t))(\beta'(t) + \log(1-t))^k}{k!}.
\end{aligned}$$

It can, indeed, be checked that

$$x_k(t) = \frac{\exp(-\beta'(t))(\beta'(t) + \log(1-t))^k}{k!}, \quad k \geq 0. \qquad (4.6)$$

is a solution to the system (4.5).

The purpose of the rest of this chapter is to make the above argument rigorous, culminating in Corollary 4.9 which states that (4.6) is indeed the desired fluid limit for $(X_n^N)_{n\geq 0}$.

## 4.4 System of differential equations

We have demonstrated that a solution to the system (4.5) exists; we must show that it is unique (at least in some appropriate space). Our basic strategy is to use the Contraction Mapping Theorem. In order to do this, we must perform some transformations on the equations to make $\dot{x}(t)$ into a locally Lipschitz function of $x(t)$.

We have $\dot{x}(t) = g(x(t))/h(x(t))$ where $g_j(x(t)) = \beta''(t)(x_{j-1}(t) - x_j(t)) \sum_{i=1}^{\infty} ix_i(t) - jx_j(t)$ for $j \geq 1$ and $h(x(t)) = \sum_{i=1}^{\infty} ix_i(t)$. Define $\tau : \mathbb{R}^+ \to [0,1]$ by the differential equation

$$\dot{\tau}(t) = h(x(\tau(t))), \quad \tau(0) = 0$$

and let $y(t) = x(\tau(t))$, which implies that $\dot{y}(t) = \dot{x}(\tau(t))\dot{\tau}(t) = g(y(t))$. Hence, the transformed system of equations is

$$\begin{aligned} \dot{y}_0(t) &= -\beta''(\tau(t))y_0(t)\sum_{i=1}^{\infty} iy_i(t), \\ \dot{y}_j(t) &= \beta''(\tau(t))(y_{j-1}(t) - y_j(t))\sum_{i=1}^{\infty} iy_i(t) - jy_j(t), \quad j \geq 1. \end{aligned} \tag{4.7}$$

We are now running time at a rate proportional to the number of patches on the hypergraph.

Note that if $y$ solves this new system then $y_0(t) = \exp(-\beta'(\tau(t)))$. Because $\dot{\tau}(t) = \sum_{i=1}^{\infty} iy_i(t) \geq 0$, $\tau(t)$ is strictly increasing in $t$ until the first time that $\sum_{i=1}^{\infty} iy_i(t)$ hits 0. Let $s^* = \inf\{t \geq 0 : \dot{\tau}(t) = 0\}$. Then $\tau$ is invertible until $s^*$. Thus, any solution to (4.7) corresponds to one in the original system (4.5) via $x(t) = y(\tau^{-1}(t))$, until $t = s^*$.

We now perform a simple scaling which makes these equations easier to deal with. Define

$z_j(t) = e^{jt}y_j(t)$ for $j \geq 0$ and $z_{-1}(t) = \tau(t)$. Then,

$$\dot{z}_{-1}(t) = \sum_{k=1}^{\infty} ke^{-kt}z_k(t)$$

$$\dot{z}_0(t) = -\beta''(z_{-1}(t))z_0(t)\sum_{k=1}^{\infty} ke^{-kt}z_k(t)$$

and, for $j \geq 1$,

$$\dot{z}(t) = \beta''(z_{-1}(t))\left(e^t z_{j-1}(t) - z_j(t)\right)\sum_{k=1}^{\infty} ke^{-kt}z_k(t). \tag{4.8}$$

The system is now in a form amenable to the Contraction Mapping Theorem method.

**Theorem 4.3.** *Take $z(t) = (z_{-1}(t), z_0(t), z_1(t), \dots)$ and $t_0 < \infty$ to be a fixed time. Define the function $f : \mathbb{R}^+ \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ co-ordinatewise by*

$$f_{-1}(t, z(t)) = \sum_{k=1}^{\infty} ke^{-kt}z_k(t)$$

$$f_0(t, z(t)) = -\beta''(z_{-1}(t))z_0(t)\sum_{k=1}^{\infty} ke^{-kt}z_k(t)$$

*and, for $i \geq 1$,*

$$f_i(t, z(t)) = \beta''(z_{-1}(t))\left(e^t z_{i-1}(t) - z_i(t)\right)\sum_{k=1}^{\infty} ke^{-kt}z_k(t).$$

*Write $F(t, z(t)) = z(0) + \int_0^t f(s, z(s))ds$. Then the function $F$ has a unique fixed point $z(t)$ in*

$$S = \Bigg\{x : [0, t_0 \wedge s^*] \to \mathbb{R}^{\mathbb{N}} \text{ such that}$$
$$\sup_{0\leq t\leq t_0\wedge s^*}\left(|x_{-1}(t)| + |x_0(t)| + \sum_{j=1}^{\infty} je^{-jt}|x_j(t)|\right) < \infty\Bigg\},$$

*which satisfies $\dot{z}(t) = f(t, z(t))$.*

First we need two lemmas whose proofs consist of routine analysis and are omitted.

**Lemma 4.4.** *Take $x = (x_{-1}, x_0, x_1, \dots)$. Define a norm on $\mathbb{R}^{\mathbb{N}}$ by $\|x\|_t = |x_{-1}| + |x_0| + \sum_{j=1}^{\infty} je^{-jt}|x_j|$. This norm induces a metric $d_t$ such that $d_t(x, y) = \|x - y\|_t$. Let $S_t = \{x \in \mathbb{R}^{\mathbb{N}} : \|x\|_t < \infty\}$. Then $(S_t, d_t)$ is a complete metric space.*

**Lemma 4.5.** *Define a norm on functions $x : [0, t_0 \wedge s^*] \to \mathbb{R}^{\mathbb{N}}$ by $\|x\| = \sup_{0 \le t \le t_0} \|x(t)\|_t$ and a corresponding metric by $d_S(x, y) = \sup_{0 \le t \le t_0} d_t(x(t), y(t))$. Let $S = \{x : [0, t_0 \wedge s^*] \to \mathbb{R}^{\mathbb{N}}$ such that $\|x\| < \infty\}$. Then $(S, d_S)$ is a complete metric space.*

We can now prove Theorem 4.3.

*Proof of Theorem 4.3.* We wish to show that $F$ is locally Lipschitz. Take

$$\begin{aligned} x(t) &= (x_{-1}(t), x_0(t), x_1(t), \dots) \\ y(t) &= (y_{-1}(t), y_0(t), y_1(t), \dots). \end{aligned}$$

Then

$$\begin{aligned} &\|f(t, x(t)) - f(t, y(t))\|_t \\ &\quad = |f_{-1}(t, x(t)) - f_{-1}(t, y(t))| + |f_0(t, x(t)) - f_0(t, y(t))| \\ &\qquad + \sum_{j=1}^{\infty} j e^{-jt} |f_j(t, x(t)) - f_j(t, y(t))| \\ &\quad = \left| \sum_{j=1}^{\infty} j e^{-jt} \left(x_j(t) - y_j(t)\right) \right| + \beta''(x_{-1}(t)) \left| x_0(t) \sum_{j=1}^{\infty} j e^{-jt} x_j(t) - y_0(t) \sum_{j=1}^{\infty} j e^{-jt} y_j(t) \right| \\ &\qquad + \sum_{k=1}^{\infty} k e^{-kt} \left| \beta''(x_{-1}(t)) \left(e^t x_{k-1}(t) - x_k(t)\right) \sum_{j=1}^{\infty} j e^{-jt} x_j(t) \right. \\ &\qquad\qquad \left. - \beta''(y_{-1}(t)) \left(e^t y_{k-1}(t) - y_k(t)\right) \sum_{j=1}^{\infty} j e^{-jt} y_j(t) \right| \\ &\quad \le \left(1 + 4\lambda \|x(t)\|_t^2 + 4\lambda \|x(t)\|_t \|y(t)\|_t + 4\lambda \|y(t)\|_t^2\right) \|x(t) - y(t)\|_t \end{aligned}$$

by repeated application of the triangle inequality, noting that the Lipschitz constant of $\beta''$ must be at most $\lambda$. Therefore,

$$\begin{aligned} &\|F(t, x(t)) - F(t, y(t))\|_t \\ &\qquad \le \|x(0) - y(0)\|_t + \int_0^t \|f(s, x(s)) - f(s, y(s))\|_t \, ds \\ &\qquad \le \|x(0) - y(0)\|_t + \int_0^t \|f(s, x(s)) - f(s, y(s))\|_s \, ds \\ &\qquad \le \|x(0) - y(0)\|_t + \left(1 + 4\lambda(\|x(t)\|_t + \|y(t)\|_t)^2\right) \int_0^t \|x(s) - y(s)\|_s \, ds. \end{aligned} \tag{4.9}$$

Assuming that $x_k(t) \geq 0$ for all $k \geq -1$, we know that any solution to the original system (4.5) will satisfy

$$x_k(t) \leq \frac{\exp(-\beta'(t))(\beta'(t))^k}{k!}, \quad k \geq 0$$

because the right-hand side is the maximal solution to the differential inequality obtained by omitting the final term $-x_k(t)/\sum_{j=1}^{\infty} jx_j(t)$ from the expression for $\dot{x}_k(t)$ in (4.5). This inequality corresponds to

$$y_k(t) \leq \frac{\exp(-\beta'(\tau(t)))(\beta'(\tau(t)))^k}{k!}, \quad k \geq 0$$

via $y(t) = x(\tau(t))$ and to

$$z_k(t) \leq \frac{\exp(kt - \beta'(\tau(t)))(\beta'(\tau(t)))^k}{k!}, \quad k \geq 0$$

in the version we are interested in, (4.8). For $t = 0$, these expressions hold with equality and also $z_{-1}(0) = 0$. Thus, we may restrict attention to the set

$$\begin{aligned} R = \Big\{ & x \in S : x_{-1}(0) = 0,\ x_i(0) = \exp(-\beta'(0))(\beta'(0))^i/i!,\ i \geq 0,\ x_{-1}(t) \leq 1, \\ & x_0(t) = \exp(-\beta'(x_{-1}(t))),\ 0 \leq x_k(t) \leq \frac{\exp(kt - \beta'(x_{-1}(t)))(\beta'(x_{-1}(t)))^k}{k!},\ k \geq 0 \Big\} \end{aligned}$$

as we will not find a fixed point outside it. For $x, y \in R$, we have $\|x(0) - y(0)\| = 0$ and

$$\begin{aligned} \|x\| &\leq \sup_{0 \leq t \leq t_0 \wedge s^*} \left( |x_{-1}(t)| + |x_0(t)| + \sum_{i=1}^{\infty} i \frac{\exp(-\beta'(x_{-1}(t)))(\beta'(x_{-1}(t)))^i}{i!} \right) \\ &= \sup_{0 \leq t \leq t_0 \wedge s^*} \big( |x_{-1}(t)| + \exp(-\beta'(x_{-1}(t))) + \beta'(x_{-1}(t)) \big) \\ &\leq 2 + \beta_1 + \lambda \end{aligned}$$

and likewise for $\|y\|$. So equation (4.9) becomes

$$\|F(t, x(t)) - F(t, y(t))\|_t \leq (1 + 16\lambda(2 + \beta_1 + \lambda)^2) \int_0^t \|x(s) - y(s)\|_s \, ds.$$

Hence, by induction, taking $K = 1 + 16\lambda(2 + \beta_1 + \lambda)^2$ and applying $F$ iteratively,

$$\|F^n(t, x(t)) - F^n(t, y(t))\|_t \leq \frac{K^n t^n}{n!} \|x - y\|$$

and so

$$\|F^n(\cdot, x) - F^n(\cdot, y)\| \leq \frac{K^n (t_0 \wedge s^*)^n}{n!} \|x - y\|$$

which is a contraction for sufficiently large $n$. So, by the Contraction Mapping Theorem, $F$ has a unique fixed point which is a local solution to the differential equation. $\square$

Note that the condition imposed by $S$ is a natural one: it merely states that the average number of patches per vertex should be finite. We have shown that the system (4.8) has a unique solution which corresponds to a unique solution to (4.7) via $y_j(t) = e^{-jt} z_j(t)$ for $j \geq 0$. Moreover from (4.6) this means that we must have

$$y_j(t) = \exp(-\beta'(\tau(t)))(\beta'(\tau(t)) + \log(1-\tau(t)))^j / j!$$

where $\tau(t)$ is the unique solution to

$$\dot{\tau}(t) = (1-\tau(t))(\beta'(\tau(t)) + \log(1-\tau(t))).$$

Note that this implies that $\tau(s^*) = t^* = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$.

## 4.5 Putting the process into continuous time

The time-change in the previous section was essential in making the system of equations manageable. For the process $(X_n^N)_{n \geq 0}$ itself, we will need a Lipschitz property, at least for a truncated version. Thus, we now embed the process in continuous time (which we need to do anyway in order to use the fluid limit methods of Chapter 2) in a way analogous to the transformation of time we made in the system of differential equations to obtain (4.7). Therefore, we seek to run the process at a rate proportional to the number of patches on the hypergraph.

Using the idea of nuclear decay in Section 4.1, imagine each patch on the hypergraph exploding at rate 1. When a patch explodes, it destroys the vertex and all other patches on it. When such an explosion has just occurred, a random number of new patches is added to the remaining vertices (according to the same distributions as previously discussed). Let $T_t^N$ be the proportion of the $N$ vertices destroyed by time $t$ and, for $k \geq 0$, let $Y_t^{k,N}$ be the proportion of the $N$ vertices with $k$ patches on at time $t$. Let $Y_t^{-1,N} = T_t^N$ for ease of notation later. Note that we no longer need to stop the process when there are no patches left because new events then occur at rate 0. Take $t_0$ to be a fixed finite time and consider the pure Markov jump process $(Y_t^N)_{0 \leq t \leq t_0}$ where

$$Y_t^N = (Y_t^{-1}, Y_t^{0,N}, Y_t^{1,N}, \ldots).$$

We shall investigate the possible transitions of this process, conditional on

$$Y_t^N = (y_{-1}, y_0, y_1, \dots).$$

For the moment, notation will be eased by letting $p_i = p_i(N, Ny_{-1}+1)$. Let

$$M_i \sim \text{Multinomial}(Ny_i; \underbrace{0, \dots, 0}_{i+1}, p_0, p_1, \dots),$$

and suppose the $M_i$ are independent for different values of $i$. Let

$$M_i' \sim \text{Multinomial}(Ny_i - 1; \underbrace{0, \dots, 0}_{i+1}, p_0, p_1, \dots),$$

independently of $M_j$ for $j \neq i$. (Note that $M_i, M_i' \in \mathbb{R}^{\mathbb{N}}$ for all $i \geq 0$.) Then $M_i$ tells us what happens to the $Ny_i$ vertices with $i$ patches after a jump event if the patch which explodes does not belong to a vertex with $i$ patches; $M_i'$ tells us what happens to the $Ny_i - 1$ vertices with $i$ patches if a patch on a vertex with $i$ patches did explode. Thus, $Y_t^N = (y_{-1}, y_0, y_1, \dots)$ becomes

$$\begin{aligned}
&\frac{1}{N}e_{-1} + \frac{1}{N}\sum_{\substack{i=0\\ i\neq 1}}^{\infty} M_i + \frac{1}{N}M_1' \quad \text{at rate } Ny_1\\
&\frac{1}{N}e_{-1} + \frac{1}{N}\sum_{\substack{i=0\\ i\neq 2}}^{\infty} M_i + \frac{1}{N}M_2' \quad \text{at rate } 2Ny_2\\
&\qquad\vdots\\
&\frac{1}{N}e_{-1} + \frac{1}{N}\sum_{\substack{i=0\\ i\neq k}}^{\infty} M_i + \frac{1}{N}M_k' \quad \text{at rate } kNy_k
\end{aligned} \tag{4.10}$$
$$\vdots$$

where $e_{-1} = (1, 0, 0, \dots)$ (the 1 is in the position with index $-1$). It is clear that $(Y_t^N)_{t\geq 0}$ is, indeed, a Markov process. Letting $p_i = p_i(N, NY_t^{-1,N}+1)$, the conditional increments obey

$$N\mathbb{E}\left[Y_{t+dt}^{-1,N} - Y_t^{-1,N} \middle| \mathcal{F}_t^N\right] = \sum_{i=1}^{\infty} i\, Y_t^{i,N}\, dt$$

$$N\mathbb{E}\left[Y_{t+dt}^{0,N} - Y_t^{0,N} \middle| \mathcal{F}_t^N\right] = -Np_{>0} Y_t^{0,N} \sum_{i=1}^{\infty} i\, Y_t^{i,N}\, dt$$

and, for $k \geq 1$,

$$
\begin{aligned}
N\mathbb{E}&\left[Y_{t+dt}^{k,N} - Y_t^{k,N} \middle| \mathcal{F}_t^N\right] \\
&= \left(N\sum_{i=1}^{\infty} iY_t^{i,N}\right)\left[p_k Y_t^{0,N} + p_{k-1}Y_t^{1,N} + \cdots + p_1 Y_t^{k-1,N} - p_{>0}Y_t^{k,N}\right] dt \\
&\quad - \left[p_{k-1}Y_t^{i,N} + 2p_{k-2}Y_t^{2,N} + \cdots + (k-1)p_1 Y_t^{k-1,N} + k\,p_0 Y_t^{k,N}\right] dt.
\end{aligned}
$$

If we consider the limit as $N \to \infty$ in these equations, with the correct scaling of time, we obtain

$$
\begin{aligned}
\dot{y}_{-1}(t) &= \sum_{i=1}^{\infty} i\, y_i(t) \\
\dot{y}_0(t) &= -\beta''(y_{-1}(t))y_0(t)\sum_{i=1}^{\infty} i\, y_i(t)
\end{aligned}
$$

and, for $k \geq 1$,

$$
\dot{y}_k(t) = \beta''(y_{-1}(t))(y_{k-1}(t) - y_k(t))\sum_{i=1}^{\infty} i\, y_i(t) - k\, y_k(t),
$$

which is the form of the differential equations in (4.7), with which we shall work from now on.

## 4.6 The no-deletions process

An essential part of the convergence proof will be bounding part of the process $(Y_t^N)_{t\geq 0}$ by another process which is simpler and, in some sense, dominates it. It will be convenient, for this section, to think of $(Y_t^N)_{t\geq 0}$ as representing the evolution of the model of balls in an urn outlined in Section 4.1 (thus, $Y_t^{k,N}$ is the proportion of the $N$ colours with $k$ balls in the urn at time $t$, for $k \geq 0$). Now consider the following modification: instead of removing balls and then adding in new ones of different colours, just add the new balls. In this way, there will always be more balls in the modified system. Refer to this as the *no-deletions process*. We will construct the no-deletions process on the same probability space as the process with deletions and couple them so that we may make comparisons.

We will take $(\tilde{Y}_t^N)_{t\geq 0}$ to be the no-deletions process corresponding to $(Y_t^N)_{t\geq 0}$, so that $\tilde{Y}_t^{k,N}$ will be the proportion of colours in the urn with $k$ balls, for $k \geq 0$. We construct

$(\tilde{Y}_t^N)_{t\geq 0}$ as follows. Take two urns and put an identical number $B_j^N$ of balls of colour $j$ into each urn, $j = 1, \dots, N$. The first urn is the 'no-deletions' urn and the second is the 'deletions' urn.

Define a probability distribution on $\mathbb{Z}^+$ by

$$\tilde{p}_0^N = \min_{0\leq n\leq \lfloor Nr\rfloor} p_0(N,n)$$

and, for $k \geq 1$,

$$\tilde{p}_0^N + \tilde{p}_1^N + \cdots + \tilde{p}_k^N = \min_{0\leq n\leq \lfloor Nr\rfloor} \{p_0(N,n) + p_1(N,n) + \cdots + p_k(N,n)\}.$$

Note that this really is a distribution as

$$\sum_{k=0}^{\infty} \tilde{p}_k^N = \lim_{m\to\infty} \min_{0\leq n\leq \lfloor Nr\rfloor} \left\{\sum_{k=0}^{m} p_k(N,n)\right\} = 1.$$

**Proposition 4.6.** *We have*

$$N\tilde{p}_1^N \to \lambda \tag{4.11}$$

*as $N \to \infty$. Suppose a random variable $\tilde{A}^N$ has distribution $(\tilde{p}_k^N)_{k\geq 0}$ (i.e. $\mathbb{P}\left(\tilde{A}^N = k\right) = \tilde{p}_k^N$, for $k \geq 0$). Then*

$$N \log \mathbb{E}\left[e^{\theta \tilde{A}^N}\right] \to \lambda(e^\theta - 1) \tag{4.12}$$

*uniformly in $|\theta| \leq \eta$ as $N \to \infty$.*

*Proof.* We have

$$\begin{aligned}
N\tilde{p}_1^N &= N \min_{0\leq n\leq \lfloor Nr\rfloor} \{p_0(N,n) + p_1(N,n)\} - N \min_{0\leq n\leq \lfloor Nr\rfloor} \{p_0(N,n)\} \\
&= N \min_{0\leq n\leq \lfloor Nr\rfloor} \left\{1 - \sum_{k\geq 2} p_k(N,n)\right\} - N \min_{0\leq n\leq \lfloor Nr\rfloor} \left\{1 - p_1(N,n) - \sum_{k\geq 2} p_k(N,n)\right\} \\
&= N \max_{0\leq n\leq \lfloor Nr\rfloor} \left\{p_1(N,n) + \sum_{k\geq 2} p_k(N,n)\right\} - N \max_{0\leq n\leq \lfloor Nr\rfloor} \left\{\sum_{k\geq 2} p_k(N,n)\right\} \\
&\to \lambda
\end{aligned}$$

as $N \to \infty$, by (4.3). Furthermore,

$$N \log \mathbb{E}\left[e^{\theta \tilde{A}^N}\right] = N \log \left(1 + \tilde{p}_1^N (e^\theta - 1) + \sum_{k\geq 2} (e^{\theta k} - 1)\tilde{p}_k^N\right)$$

and

$$N\sum_{k\geq 2} e^{\theta k}\tilde{p}_k^N \leq N\sum_{k\geq 2} e^{\theta k}\sup_n(p_k(N,n)).$$

The result now follows from (4.2) and (4.3). □

Define random variables $\tilde{A}_n^{i,N}$ on the same probability space as the $A_n^{i,N}$'s with $\tilde{A}_n^{i,N}$ independent and identically distributed, with the same distribution as $\tilde{A}^N$ and with $A_n^{i,N} \leq \tilde{A}_n^{i,N}$ a.s. for all $n$ and $i$ (Proposition 4.6 shows that this is possible).

- Pick each ball in the no-deletions urn at rate 1. Suppose that the colour picked has $b$ balls in the deletions urn and $\tilde{b}$ balls in the no-deletions urn, where $b \leq \tilde{b}$.
- If $b = 0$, just add a number of balls distributed as $\tilde{A}^N$ to each colour in the no-deletions urn, leave the deletions urn alone and return to the start.
- If $b > 0$, flip a coin with probability $b/\tilde{b}$ of heads.
- If the coin shows heads, delete the colour from the deletions urn. If this is the $n$th deletion that has been performed, add $A_n^{i,N}$ balls to colour $i$ in the deletions urn for $1 \leq i \leq N-n$. Also add $\tilde{A}_n^{i,N}$ balls to colour $i$ in the non-deletions urn, for $1 \leq i \leq N$. Now return to the start.
- If the coin shows tails, do not delete any colours and add a number of balls distributed as $\tilde{A}^N$ to each colour in the no-deletions urn. Don't touch the deletions urn. Return to the start.

The above construction results in a no-deletions process which behaves as follows: a number of balls (distributed as $\tilde{A}^N$) is added to each colour at a rate equal to the number of balls in the urn. In addition, we have coupled the deletions and no-deletions processes such that the number of colours with more than $d$ balls in the urn is always greater in the non-deletions process. This fact will be essential in Section 4.8.

## 4.7 Fluid limit for the no-deletions process

We will prove a fluid limit for a finite-dimensional summary of $(\tilde{Y}_t^N)_{t\geq 0}$. Let $W_t^{k,N} = \tilde{Y}_t^{k,N}$ for $0 \leq k \leq d$ and let $W_t^{d+1,n} = \sum_{j=d+1}^{\infty} j\tilde{Y}_t^{j,N}$; then $W_t^N = (W_t^{0,N}, W_t^{1,N}, \ldots, W_t^{d+1,N})$ is a $(d+2)$-dimensional pure jump Markov process.

Condition on $W_t^N = (w_0, w_1, \dots, w_{d+1})$. For $i = 0, 1, \dots, d$ let

$$\tilde{M}^i \sim \text{Multinomial}(Nw_i; \underbrace{0, \dots, 0}_{i}, \tilde{p}_0^N, \tilde{p}_1^N, \tilde{p}_2^N, \dots) \in \mathbb{R}^{\mathbb{N}}.$$

Let $\tilde{M}_k^i$ be the $k$th co-ordinate of $\tilde{M}^i$, where we number from 0. Then $\tilde{M}_k^i$ is the number of colours which start with $i$ balls in the urn and end up with $k$ balls after a transition event. Let $L^N$ be the sum of $N\left(1 - \sum_{j=0}^d w_j\right)$ independent copies of $\tilde{A}^N$. This is the number of new balls added to colours which already have strictly more than $d$ balls. Thus, at rate $N\left(\sum_{j=1}^d jw_j + w_{d+1}\right)$, $W_t^N = (w_0, w_1, \dots, w_{d+1})$ becomes

$$\frac{1}{N}\left(\sum_{i=0}^d \tilde{M}_0^i,\ \sum_{i=0}^d \tilde{M}_1^i,\ \dots,\ \sum_{i=0}^d \tilde{M}_d^i,\ Nw_{d+1} + \sum_{i=0}^d \sum_{k=d+1}^\infty k\tilde{M}_k^i + L^N\right). \tag{4.13}$$

Let $\tilde{\mathcal{F}}_t^N = \sigma(W_s^N : s \le t)$. Then the conditional increments of $(W_t^N)_{t\ge 0}$ are as follows:

$$N\mathbb{E}\left[W_{t+dt}^{0,N} - W_t^{0,N} \middle| \tilde{\mathcal{F}}_t^N\right] = -N\tilde{p}_{>0}^N W_t^{0,N}\left(\sum_{j=1}^d jW_t^{j,N} + W_t^{d+1,N}\right) dt,$$

for $1 \le k \le d$,

$$\begin{aligned} &N\mathbb{E}\left[W_{t+dt}^{k,N} - W_t^{k,N} \middle| \tilde{\mathcal{F}}_t^N\right] \\ &= N\left(\tilde{p}_k^N W_t^{0,N} + \tilde{p}_{k-1}^N W_t^{1,N} + \cdots + \tilde{p}_1^N W_t^{k-1,N} - \tilde{p}_{>0}^N W_t^{k,N}\right)\left(\sum_{j=1}^d jW_t^{j,N} + W_t^{d+1,N}\right) dt \end{aligned}$$

and

$$\begin{aligned} &N\mathbb{E}\left[W_{t+dt}^{d+1,N} - W_t^{d+1,N} \middle| \tilde{\mathcal{F}}_t^N\right] \\ &= N\left(\left(1 - \sum_{j=0}^d W_t^{j,N}\right)\mathbb{E}\left[\tilde{A}^N\right] + W_t^{0,N}\sum_{j=d+1}^\infty j\tilde{p}_j^N + \cdots + W_t^{d,N}\sum_{j=1}^\infty (d+j)\tilde{p}_j^N\right) \\ &\quad \left(\sum_{i=1}^d iW_t^{i,N} + W_t^{d+1,N}\right) dt. \end{aligned}$$

Heuristically, we expect the fluid limit, $w(t)$, to satisfy

$$\begin{aligned}
\dot{w}_0(t) &= -\lambda w_0(t)\left(\sum_{j=1}^{d} j w_j(t) + w_{d+1}(t)\right) \\
\dot{w}_k(t) &= \lambda(w_{k-1}(t) - w_k(t))\left(\sum_{j=1}^{d} j w_j(t) + w_{d+1}(t)\right) \quad \text{for } 1 \le k \le d \\
\dot{w}_{d+1}(t) &= \lambda\left(1 - \sum_{j=0}^{d} w_j(t) + (d+1)w_d(t)\right)\left(\sum_{i=1}^{d} i w_i(t) + w_{d+1}(t)\right) \qquad (4.14)
\end{aligned}$$

with initial conditions $w_k(0) = \frac{e^{-\beta_1}\beta_1^k}{k!}$ for $0 \le k \le d$ and $w_{d+1}(t) = \sum_{j=d+1}^{\infty} j e^{-\beta_1}\beta_1^j/j!$. Then this system has unique solution

$$\begin{aligned}
w_k(t) &= e^{-\beta_1 e^{\lambda t}}(\beta_1 e^{\lambda t})^k/k!, \quad 0 \le k \le d \\
w_{d+1}(t) &= \beta_1 e^{\lambda t} - \sum_{k=1}^{d} k\, e^{-\beta_1 e^{\lambda t}}(\beta_1 e^{\lambda t})^k/k!. \qquad (4.15)
\end{aligned}$$

Let $\|\cdot\|_{d+1}$ denote the supremum norm such that

$$\|x\|_{d+1} = \sup_{0 \le k \le d+1} |x_k|.$$

Fix $t_0 < \infty$.

**Theorem 4.7.** *For all $\delta > 0$,*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left(\sup_{0 \le t \le t_0} \|W_t^N - w(t)\|_{d+1} > \delta\right) < 0.$$

*Proof.* We will use Theorem 1.9 from the Introduction and the notation from that section. (Note that there we had a $d$-dimensional process whereas here we have a $(d+1)$-dimensional process.) The Lévy kernel of $(W_t^N)_{t\ge 0}$ is essentially determined by (4.13). The corresponding Laplace transform is

$$\begin{aligned}
m^N(w,\theta) = N\left(\sum_{j=1}^{d} j w_j + w_{d+1}\right) \exp\Bigg\{ &\sum_{i=0}^{d} N w_i \log\left(\sum_{k=i}^{d} \tilde{p}_{k-i}^N e^{\theta_k/N} + \sum_{k=d+1}^{\infty} \tilde{p}_{k-i}^N e^{k\theta_{d+1}/N}\right) \\
&- \sum_{i=0}^{d} \theta_i w_i + N\left(1 - \sum_{k=0}^{d} w_k\right) \log \mathbb{E}\left[\exp(\theta_{d+1}\tilde{A}^N/N)\right] \Bigg\}.
\end{aligned}$$

Let $S = \{w \in \mathbb{R}^{d+2} : \sup_{0\le k\le d+1} |w_k| < r\}$, where $0 < r < \infty$ is a fixed constant.

We must check Conditions (C1) to (C5). Now, if $e_i$ is the unit vector with a 1 in position $i$ and zeros in positions $j = 0, 1, \ldots, i-1, i+1, \ldots, d+1$, we have

$$\tilde{M}^i - N w_i e_i \xrightarrow{d} (\underbrace{0, 0, \ldots, 0}_{i}, -P_i, P_i, 0, \ldots)$$

where $P_i \sim \text{Poisson}(\lambda w_i)$. Moreover, by (4.12), $L^N$ converges to a $\text{Poisson}(\lambda(1-\sum_{k=0}^{d} w_k))$ distribution. Thus, there exists a limit kernel $K(w, dy)$ such that

$$m(w,\theta) = \left(\sum_{j=1}^{d} j w_j + w_{d+1}\right) \exp\left(\lambda \sum_{i=0}^{d-1} w_i (e^{\theta_{i+1}-\theta_i} - 1) + \lambda w_d (e^{(d+1)\theta_{d+1}-\theta_d} - 1)\right.$$
$$\left. + \lambda \left(1 - \sum_{k=0}^{d} w_k\right)(e^{\theta_{d+1}} - 1)\right).$$

Clearly we can find a constant $\eta_0 > 0$ such that

$$m(w,\theta) < \infty \quad \text{for all } w \in S,\ \|\theta\|_{d+1} \leq \eta_0,$$

which is Condition (C1).

Furthermore, letting $S^N = S \cap \frac{1}{N}\mathbb{Z}^{d+1}$,

$$\sup_{w \in S^N} \sup_{\|\theta\|_{d+1} \leq \eta_0} \left| N^{-1} m^N(w, N\theta) - m(w,\theta) \right| \to 0$$

as $N \to \infty$, which is Condition (C2).

Setting $b(w) = m'(w, 0)$, it is easily checked that we obtain Equation (4.14). We must check that $b$ is Lipschitz in $S$. Let $x, y \in S$. Then

$$\|b(x) - b(y)\|_{d+1} \leq 4r(d+1)^2 d^2 \|x - y\|_{d+1}$$

by repeated use of the triangle inequality. So $b$ is, indeed, Lipschitz on $S$ and $\dot{w}(t) = b(w(t))$ has the unique solution given in (4.15). Hence, Condition (C3) is satisfied. Moreover, if we take $r$ large enough in the definition of the set $S$ then we can ensure that $w(t)$ stays in $S$ at least until $t_0$. Condition (C4) is trivially true.

Finally, it follows from Proposition 4.2 that Condition (C5) is true, namely

$$\limsup_{N \to \infty} N^{-1} \log \mathbb{P}\left(\|W_0^N - w(0)\|_{d+1} > \epsilon\right) < 0.$$

Thus, we may apply Theorem 1.9 to obtain the desired result. □

## 4.8 Fluid limit for the deletions process

Recall that $(Y_t^N)_{0\leq t\leq t_0}$ is the deletions process and that $(\tilde{Y}_t^N)_{0\leq t\leq t_0}$ is the no-deletions process with $(d+2)$-dimensional summary $(W_t^N)_{0\leq t\leq t_0}$, as discussed in the previous section. We have identified balls with patches and colours with vertices. Define the *remainder* $R_t^N = \sum_{j=d+1}^{\infty} j\, Y_t^{j,N}$, the number of patches on vertices with more than $d+1$ patches. By the coupling of the deletions and no-deletions processes, we have

$$R_t^N \leq W_t^{d+1,N} \quad \text{a.s..}$$

We have already derived a system of differential equations which any fluid limit must satisfy and have shown that it has a unique solution in the set

$$\left\{ z : [0, t_0 \wedge s^*] \to \mathbb{R}^{\mathbb{N}} \text{ such that } \sup_{0\leq t\leq t_0\wedge s^*} \left( |z_{-1}(t)| + |z_0(t)| + \sum_{j=1}^{\infty} j|z_j(t)| \right) < \infty \right\}.$$

As previously discussed, we need not worry about a stopping time for when the hypergraph no longer has any patches, so we will simply work on the compact time interval $[0, t_0 \wedge s^*]$.

**Theorem 4.8.** *Let $(y(t))_{0\leq t\leq t_0\wedge s^*}$ be defined by*

$$\begin{aligned} y_{-1}(t) &= \tau(t) \\ y_k(t) &= \exp(-\beta'(\tau(t)))(\beta'(\tau(t)) + \log(1-\tau(t)))^k/k!, \;\; k \geq 0 \end{aligned}$$

*where $\tau(t)$ is the unique solution to*

$$\dot{\tau}(t) = (1-\tau(t))(\beta'(\tau(t)) + \log(1-\tau(t))).$$

*Then $(y(t))_{0\leq t\leq t_0\wedge s^*}$ is the unique solution to equations (4.7) and for every $\delta > 0$, we have*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left( \sup_{0\leq t\leq t_0\wedge s^*} \|Y_t^N - y(t)\| > \delta \right) < 0,$$

*where $\|y\| = \sup_{k\geq -1} |y_k|$.*

*Proof.* We use the notation of Chapter 2 but replace $x$ by $y$ and vectors are now numbered from $-1$ rather than 1. So here $\|x\|_d = \sup_{-1\leq k\leq d} |x_k|$. We need to check Conditions 1 to 9 of Theorem 2.2. The Lévy kernel $K^N(y, dz)$ of $(Y_t^N)_{t\geq 0}$ is determined by (4.10). The multinomial vectors there scale correctly to give independent Poisson random variables in the limit:

$$M_i - Ny_ie_i \xrightarrow{d} (\underbrace{0,0,\dots,0}_{i}, -P_i, P_i, 0, \dots)$$

as $N \to \infty$, where $P_i \sim \text{Poisson}(\beta''(y_{-1})y_i)$. So the limit kernel $K(y,dz)$ exists. Recall that the function $a : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{d+2}$ is such that $a_k(z) = z_k$ for $k = -1, 0, 1, \dots, d$. Then

$$\begin{aligned}
m(y,\theta) &= \int_{\mathbb{R}^{\mathbb{N}}} e^{\langle \theta, a(z)\rangle} K(y,dz) \\
&= \sum_{k=1}^{d} ky_k \exp\left\{ \sum_{j=0}^{d-1} y_j\beta''(y_{-1})(e^{\theta_{j+1}-\theta_j} - 1) + y_d\beta''(y_{-1})(e^{-\theta_d} - 1) - \theta_k + \theta_{-1} \right\} \\
&\quad + \sum_{k=d+1}^{\infty} ky_k \exp\left\{ \sum_{j=0}^{d-1} y_j\beta''(y_{-1})(e^{\theta_{j+1}-\theta_j} - 1) + y_d\beta''(y_{-1})(e^{-\theta_d} - 1) + \theta_{-1} \right\}.
\end{aligned}$$

Fix $0 < r < \infty$. Let

$$S = \left\{ y \in \mathbb{R}^{\mathbb{N}} : |y_{-1}| < r, \quad 0 < y_0 < r, \quad 0 < \sum_{k=1}^{\infty} ky_k < r \right\}.$$

Then there exists $\eta_0 < \infty$ such that

$$\sup_{y\in S} \sup_{\|\theta\|_d \le \eta_0} m(y,\theta) < \infty,$$

and Condition 1 holds. Now, retaining the abbreviation $p_k = p_k(N, Ny_{-1}+1)$,

$$\begin{aligned}
m^N(y,\theta) &= \int_{\mathbb{R}^{\mathbb{N}}} e^{\langle \theta, a(z)\rangle} K^N(y,dz) \\
&= N\sum_{k=1}^{d} ky_k \exp\left\{ \sum_{\substack{j=0\\ j\ne k}}^{d} \left( Ny_j \log\left( \sum_{i=j}^{d} p_{i-j}e^{\theta_i/N} + \sum_{i=d+1}^{\infty} p_{i-j} \right)\right) \right. \\
&\qquad \left. + (Ny_k - 1)\log\left( \sum_{i=k}^{d} p_{i-k}e^{\theta_i/N} + \sum_{i=d+1}^{\infty} p_{i-k} \right) - \sum_{j=0}^{d} y_j\theta_j + \frac{\theta_{-1}}{N} \right\} \\
&\quad + N\sum_{k=d+1}^{\infty} ky_k \exp\left\{ \sum_{j=0}^{d} \left( Ny_j \log\left( \sum_{i=j}^{d} p_{i-j}e^{\theta_i/N} + \sum_{i=d+1}^{\infty} p_{i-j} \right)\right) \right. \\
&\qquad \left. - \sum_{j=0}^{d} y_j\theta_j + \frac{\theta_{-1}}{N} \right\}
\end{aligned}$$

and so by (4.2) and (4.3),

$$\sup_{y\in S^N}\sup_{\|\theta\|_d\le\eta_0}\left|N^{-1}m^N(y,N\theta)-m(y,\theta)\right|\to 0$$

as $N\to\infty$, satisfying Condition 2. We have

$$\begin{aligned} b^N_{-1}(y) &= \sum_{i=1}^{\infty} iy_i \\ b^N_0(y) &= -Np_{>0}y_0\sum_{i=1}^{\infty} iy_i \end{aligned}$$

and, for $k\ge 1$,

$$\begin{aligned} b^N_k(y) = {} & (Np_ky_0+Np_{k-1}y_1+\dots+Np_1y_{k-1}-Np_{>0}y_k)\sum_{i=1}^{\infty} iy_i \\ & -(p_{k-1}y_1+2p_{k-2}y_2+\dots+(k-1)p_1y_{k-1}+kp_0y_k). \end{aligned}$$

This clearly converges to $b$ defined by

$$\begin{aligned} b_{-1}(y) &= \sum_{i=1}^{\infty} iy_i \\ b_0(y) &= -\beta''(y_{-1})y_0\sum_{i=1}^{\infty} iy_i \end{aligned}$$

and, for $k\ge 1$,

$$b_k(y)=\beta''(y_{-1})(y_{k-1}-y_k)\sum_{i=1}^{\infty} iy_i-ky_k.$$

Moreover, we have already demonstrated that there is a unique solution to $\dot{y}(t)=b(y(t))$ up until the first time $y(t)$ leaves $S$. By taking $r$ large enough, we can ensure that the only way this can happen is if $\sum_{k=1}^{\infty} ky_k(t)$ hits 0. But this cannot happen before $s^*$. Hence, Condition 3 is satisfied.

Now, Proposition 4.2 implies that

$$\limsup_{N\to\infty} N^{-1}\log\mathbb{P}\left(\|\hat{Y}^N_0-\hat{y}(0)\|_d>\delta\right)<0.$$

This gives Condition 4. Condition 5 is trivially true as

$$\sup_{0\le t\le t_0\wedge s^*}\sup_{k\ge d+1}|y_k(t)|\le\sup_{0\le t<t^*}\sup_{k\ge d+1}\frac{\exp(-\beta'(t))(\beta'(t)+\log(1-t))^k}{k!}$$

where $t^* = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$. For the process, we have

$$\sup_{0 \leq t \leq t_0} \sup_{k \geq d+1} \left| Y_t^{k,N}(t) \right| \leq \sup_{0 \leq t \leq t_0} R_t^N \leq \sup_{0 \leq t \leq t_0} W_t^{d+1,N} \text{ a.s.,}$$

by the no-deletions process coupling. But $(W_t^{d+1,N})_{t \geq 0}$ has a fluid limit $(w_{d+1}(t))_{t \geq 0}$ and $\sup_{0 \leq t \leq t_0} w_{d+1}(t) \to 0$ as $d \to \infty$. Hence, for each $\delta > 0$, there exists $d$ sufficiently large that

$$\limsup_{N \to \infty} N^{-1} \log \mathbb{P}\left( \sup_{0 \leq t \leq t_0} \sup_{k \geq d+1} \left| Y_t^{k,N} \right| > \delta \right) < 0,$$

which is Condition 6.

Now truncate at level $d$. Then

$$\begin{aligned}
\hat{b}_{-1}(\hat{y}, 0) &= \sum_{k=1}^{d} k y_k \\
\hat{b}_0(\hat{y}, 0) &= -\beta''(y_{-1}) y_0 \sum_{k=1}^{d} k y_k \\
\hat{b}_k(\hat{y}, 0) &= \beta''(y_{-1})(y_{k-1} - y_k) \sum_{i=1}^{d} i y_i - k y_k, \quad \text{for } 1 \leq k \leq d.
\end{aligned}$$

So, for $\hat{x}, \hat{y}$ the first $d+2$ co-ordinates of $x, y \in S$, by repeated use of the triangle inequality, we obtain

$$\|\hat{b}_k(\hat{x}, 0) - \hat{b}_k(\hat{y}, 0)\|_d \leq K(r, \lambda) \|\hat{x} - \hat{y}\|_d$$

where $K(r, \lambda)$ is a constant depending only on $r$ and $\lambda$. Thus, Condition 7 is verified.

Now,

$$\begin{aligned}
\|e(t)\|_d \leq \max &\left\{ \sum_{i=d+1}^{\infty} i y_i(t),\ \beta''(y_{-1}(t)) y_0(t) \sum_{i=d+1}^{\infty} i y_i(t), \right. \\
&\left. \sup_{1 \leq k \leq d} \left| \beta''(y_{-1}(t))(y_{k-1}(t) - y_k(t)) \sum_{i=d+1}^{\infty} i y_i(t) \right| \right\} \\
&\leq C(r, \lambda) \sum_{i=d+1}^{\infty} i y_i(t),
\end{aligned}$$

where $C(r, \lambda)$ is a constant depending only on $r$ and $\lambda$. Thus, $\sup_{0 \leq t \leq t_0 \wedge s^*} \|e(t)\|_d$ can be made arbitrarily small by taking $d$ sufficiently large, which gives Condition 8.

Finally, taking $p_k = p_k(N, NY_t^{-1,N} + 1)$, for $N$ sufficiently large, we have

$$\begin{aligned}\|E_t^N\|_d &\leq \max\Bigg\{\sum_{i=d+1}^{\infty} iY_t^{i,N},\ Np_{>0}Y_t^{0,N}\sum_{i=d+1}^{\infty} iY_t^{i,N},\\ &\qquad \sup_{1\leq k\leq d}\left|\left(Np_kY_t^{0,N}+\cdots+Np_1Y_t^{k-1,N}-Np_{>0}Y_t^{k,N}\right)\sum_{i=d+1}^{\infty} iY_t^{i,N}\right|\Bigg\}\\ &\leq C'(\lambda)\sum_{i=d+1}^{\infty} iY_t^{i,N},\end{aligned}$$

where $C'(\lambda)$ is a constant depending only on $\lambda$, using the bound $Y_t^{t,N} \leq 1$ for all $k \geq 0$. Once again, we can use the fact that

$$\sum_{i=d+1}^{\infty} iY_t^{i,N} \leq W_t^{d+1,N} \quad \text{a.s.}$$

and so, given $\delta > 0$, using the fluid limit result for $(W_t^{d+1,N})_{t\geq 0}$, we can find a $d$ sufficiently large that

$$\limsup_{N\to\infty} N^{-1}\log\mathbb{P}\left(\sup_{0\leq t\leq t_0}\|E_t^N\|_d > \delta\right) < 0.$$

This is Condition 9 and so the result follows. □

## 4.9 Interpretation

Firstly, we will give an illustration of the effect of the time-change. Recall from the Introduction that $p(t) = (1-t)(\beta'(t) + \log(1-t))$ is the fluid limit for the total number of patches on the hypergraph. Let $\tilde{p}(t) = p(\tau(t))$. Figures 1.4 and 1.5 of the Introduction gave a comparison of two possible types of behaviour of $p(t)$. More precisely, we took $\beta'(t) = 6.5\gamma(0.1+0.9t)^5$ for $\gamma = 140$ and $\gamma = 143$. If we solve the differential equation for $\tau(t)$ numerically we obtain Figures 4.1 and 4.2 as the time-changed versions of Figures 1.4 and 1.5.

Let us now discuss the implications of Theorem 4.8 for the original model (i.e. in discrete time, without the time-change). Our original discrete-time process is the jump-chain of the continuous-time process $(Y_t^N)_{t\geq 0}$ of Section 4.5, multiplied by $N$. Thus we have the following corollary of Theorem 4.8:

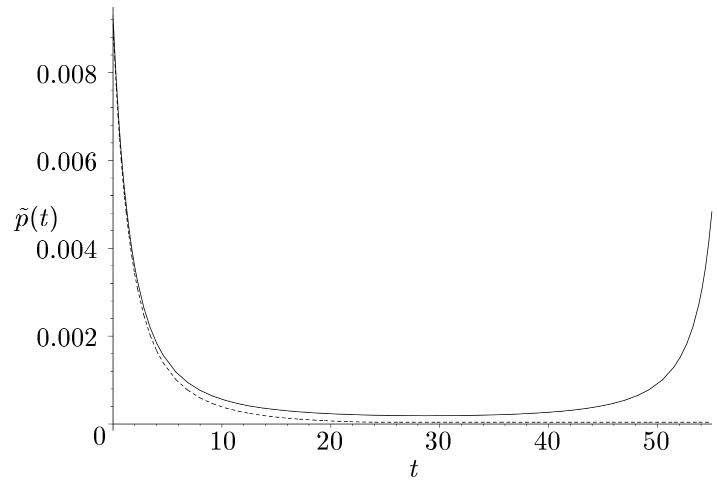


Figure 4.1: The leftmost section of the graph of $\tilde{p}(t)$ against $t$. The dashed line is the case $\gamma = 140$ and the solid line is the case $\gamma = 143$.

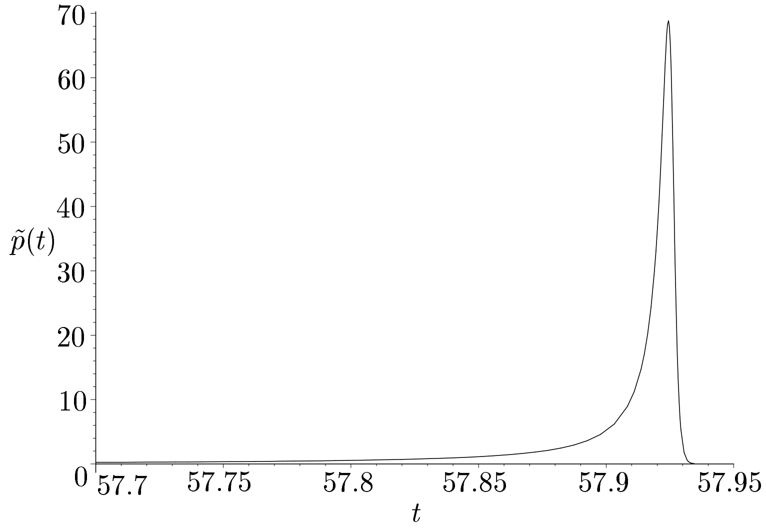


Figure 4.2: The main feature of the graph of $\tilde{p}(t)$ against $t$ for $\gamma = 143$.

**Corollary 4.9.** *Let $(x(t))_{t\geq 0}$ be defined by*

$$x_k(t) = \exp(-\beta'(t))(\beta'(t) + \log(1-t))^k/k!,\ \ k \geq 0,$$

*as at (4.6). Take $(X_n^N)_{n\geq 0}$ to be the discrete-time Markov process of Section 4.1. Then for $t_0 < \infty$ and all $\delta > 0$,*

$$\limsup_{N\to\infty} N^{-1} \log \mathbb{P}\left( \sup_{0\leq t\leq t^*\wedge\tau^{-1}(t_0)} \left\| N^{-1}X^N_{\lfloor Nt\rfloor} - x(t)\right\| > \delta \right) < 0,$$

*where $\|x\| = \sup_{k\geq 0} |x_k|$.*

*Proof.* For ease of notation, let $t_1 = t^* \wedge \tau^{-1}(t_0)$ and note that, as $t^* < 1$, we have $t_1 < 1$. Then

$$\begin{aligned} &\sup_{0\leq t\leq t_1} \left\| N^{-1}X^N_{\lfloor Nt\rfloor} - x(t)\right\| \\ &\qquad \leq \sup_{t\in\{0,1/N,\dots,\lfloor Nt_1\rfloor/N\}} \left\| N^{-1}X^N_{Nt} - x(t)\right\| + \sup_{0\leq t\leq t_1} \|x(t) - x(\lfloor Nt\rfloor/N)\|. \end{aligned} \tag{4.16}$$

Now recall that $Y_t^{-1,N} = T_t^N$, the proportion of vertices deleted by time $t$. Then,

$$\begin{aligned} &\sup_{t\in\{0,1/N,\dots,\lfloor Nt_1\rfloor/N\}} \left\| N^{-1}X^N_{Nt} - x(t)\right\| \\ &\leq \sup_{t\in\{0,1/N,\dots,\lfloor Nt_1\rfloor/N\}} \sup_{s:T_s^N=t} \|Y_s^N - x(t)\| \\ &\leq \sup_{t\in\{0,1/N,\dots,\lfloor Nt_1\rfloor/N\}} \sup_{s:T_s^N=t} \|Y_s^N - y(s)\| + \sup_{t\in\{0,1/N,\dots,\lfloor Nt_1\rfloor/N\}} \sup_{s:T_s^N=t} \|x(\tau(s)) - x(T_s^N)\|, \end{aligned}$$

recalling that $y(s) = x(\tau(s))$. But this last expression is equal to

$$\sup_{0\leq s\leq s^*\wedge t_0} \|Y_s^N - y(s)\| + \sup_{0\leq s\leq s^*\wedge t_0} \|x(\tau(s)) - x(T_s^N)\|. \tag{4.17}$$

Now we know that the first of these terms converges exponentially fast in probability to 0, as does $\sup_{0\leq s\leq s^*\wedge t_0} \|T_s^N - \tau(s)\|$, by Theorem 4.8. So the result will follow from (4.16) and (4.17) if we can show that $x$ is Lipschitz in $\|\cdot\|$ i.e. that there exists a constant $L$ such that for $0 \leq t, u \leq t_1$,

$$\|x(t) - x(u)\| \leq L|t-u|.$$

Let $C = \sup_{0\le s\le t_1} |\beta'(s) + \log(1-s)|$. Then

$$\begin{aligned}
&\|x(t) - x(u)\| \\
&\quad = \sup_{k\ge 0} \left| \exp(-\beta'(t))(\beta'(t) + \log(1-t))^k/k! - \exp(-\beta'(u))(\beta'(u) + \log(1-t))^k/k! \right| \\
&\quad \le \left|\exp(-\beta'(t)) - \exp(-\beta'(u))\right| \sup_{k\ge 0} \frac{C^k}{k!} \\
&\qquad\quad + \sup_{k\ge 0} \left|\beta'(t) + \log(1-t) - \beta'(u) - \log(1-u)\right| \sup_{k\ge 1} \frac{C^{k-1}}{k!} \\
&\quad \le L|t-u|,
\end{aligned}$$

using the Mean Value Theorem, where

$$L = \left( \sup_{0\le s\le t_1} \left|\beta''(s)e^{-\beta'(s)}\right| + \sup_{0\le s\le t_1} \left|\beta''(s)\right| + \frac{1}{1-t_1} \right) \max\left\{ \sup_{k\ge 0} \frac{C^k}{k!}, \sup_{k\ge 1} \frac{C^{k-1}}{k!} \right\}.$$

Hence result. □

Now, if $s^* < \infty$, we can conclude that Corollary 4.9, in fact, holds all the way up to $t^*$. If $s^* = \infty$ then for any $t_0 < \infty$ we will have $\tau^{-1}(t_0) < t^*$ and so the result does not hold all the way up to $t^*$. Note that the latter is the generic case: if $\beta(t)$ is the power series obtained from the parameters of a Poisson random hypergraph then $s^* = \infty$.

We conclude with the results of some simulations. We took $N = 10\,000$ and

$$\begin{aligned}
B_1^N &\sim \text{Bernoulli}(\beta'(0)) \\
A_1^{n,N} &\sim \text{Bernoulli}(N^{-1}\beta''(n/N)),
\end{aligned}$$

with $\beta'(t) = 975(0.1 + 0.9t)^5$. Figure 4.3 gives $x_k(t)$ for various values of $k$ and Figure 4.4 gives an overview of $(x_k(t) : k \ge 0,\ t \ge 0)$.

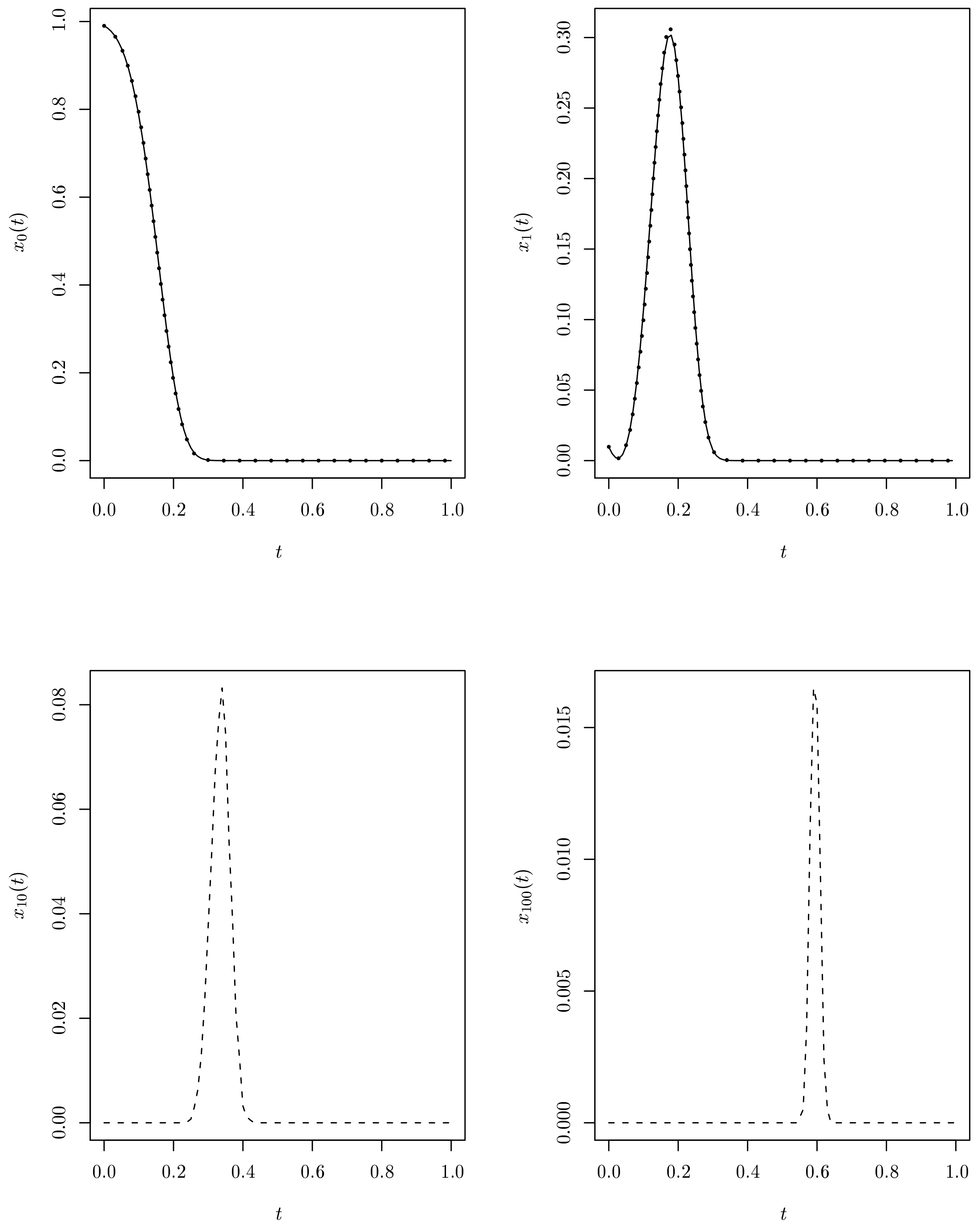


Figure 4.3: $x_k(t)$ for $k = 0, 1, 10, 100$. In the top two graphs, the dotted line is the result of simulation and the solid line is the theoretical curve. In the bottom two graphs, the dashed line is the result of simulation.

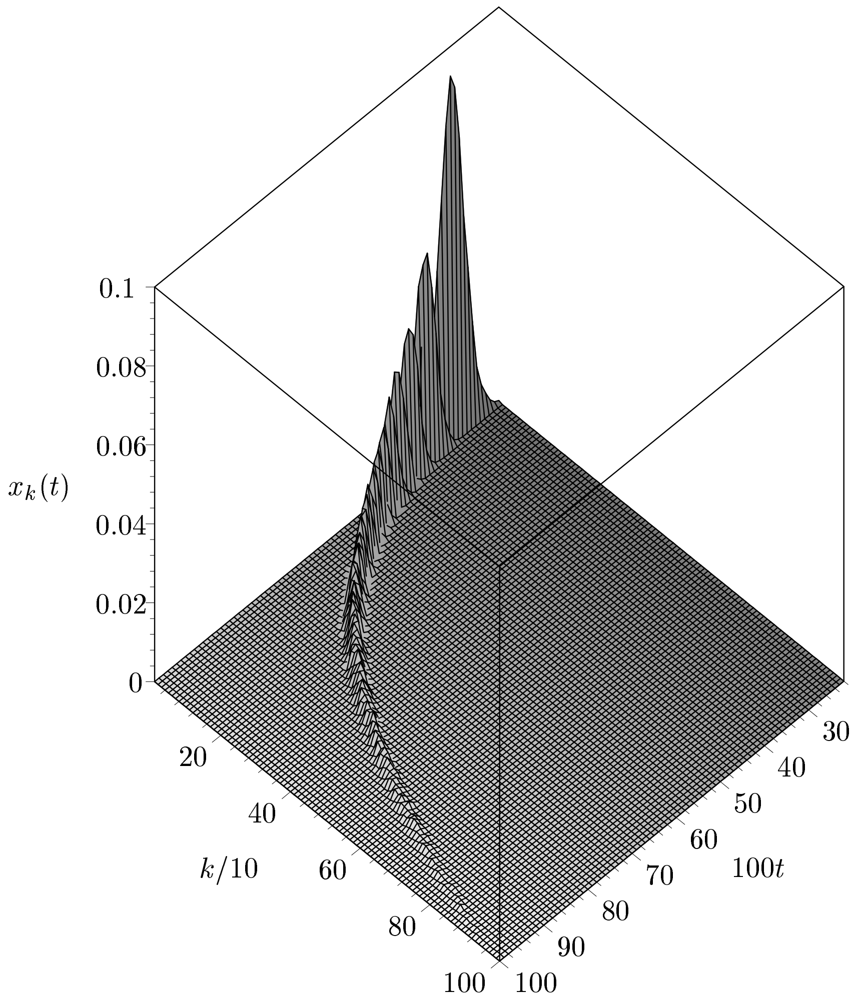


Figure 4.4: A summary of $(x_k(t) : k \geq 0,\ t \geq 0)$. (Note the scalings on the axes.) The surface is less smooth than it should be because we have been forced to sample from our simulated results.

# Chapter 5

# Essential Edges

## 5.1 Introduction

In this chapter, we look in more detail at the structure of the identifiable hyperedges. Throughout, we will assume that there is a largest size of hyperedge, say $M$, and so $\beta(t)$, the generating function of the parameters of the hypergraph, is a polynomial.

We say that a hyperedge is *essential* if removing it reduces the number of identifiable vertices. Let $\mathcal{E}(N,k)$ be the number of essential $k$-edges in the Poisson random hypergraph on $N$ vertices. The purpose of this chapter is to prove the following law of large numbers:

**Theorem 5.1.** *Under the assumptions of Theorem 1.6, with* $t^* = \inf\{t \geq 0 : \beta'(t) + \log(1-t) < 0\} \wedge 1$,

$$\frac{1}{N}\sum_{k=1}^{N} \mathcal{E}(N,k) \to -(1-t^*)\log(1-t^*)$$

*in probability as* $N \to \infty$.

Thus, the limit of $N^{-1}H_N$ given in Theorem 1.6 splits into two parts: $\beta(t^*)$ corresponding to non-essential identifiable hyperedges and $-(1-t^*)\log(1-t^*) = (1-t^*)\beta'(t^*)$ corresponding to essential hyperedges.

## 5.2 Essential edges in random graphs

In order to provide intuition about essential edges, we consider first the graph case, in which $\beta(t) = \beta_1 t + \beta_2 t^2$. Asymptotically, the edge-structure of the Poisson random graph behaves in the same way as that of the more-commonly studied binomial model $\mathbb{G}(N, p)$, with $p = 2\beta_2/N$. We give a simple calculation for essential patches and then a heuristic derivation for essential 2-edges based on known results for the core.

In the random graph, a patch is essential if and only if it is the only patch on a component. This enables us to prove part of our limiting result concerning the expected number of essential patches in an elementary way. Fix a vertex $v$ and let $D_N$ be the size of the domain of $v$. By Theorem 1.7, $D_N \xrightarrow{d} D$. We note that the limiting Borel$(2\beta_2)$ distribution for $D$ has (possibly degenerate) probability generating function $F(s)$ which is the solution to

$$F(s) = s\exp(2\beta_2(F(s) - 1))$$

in the range $0 \le s \le 1$ (see Harris [19] p.32). Let $A$ denote the event that $v$ has a patch and that this patch is essential. Then, as there is a Poisson$(n\beta_1)$ number of patches on a set of vertices of size $n$,

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{P}(A|D_N)] = \mathbb{E}\left[\beta_1 e^{-\beta_1 D_N}\right] \to \beta_1 F(e^{-\beta_1}).$$

Now,

$$F(e^{-\beta_1}) = \exp(-\beta_1 + 2\beta_2(F(e^{-\beta_1}) - 1)).$$

But $F(e^{-\beta_1}) = 1 - t^*$ is a solution to this equation and so

$$\mathbb{P}(A) \to \beta_1(1 - t^*).$$

Hence, $N^{-1}\mathbb{E}[\mathcal{E}(N, 1)]$ converges to $\beta_1(1 - t^*)$.

An edge in a random graph with patches is essential if and only if removing it splits its component into two disconnected components, exactly one of which has a patch on it. In particular, edges in cycles are not essential; an edge in a path between two patches is not essential; edges in components with no patches cannot be essential (they are not even identifiable). All edges in a tree-component with a single patch are essential.

We now describe a connection with the core of a random graph, that is the maximal subgraph with minimum degree 2. The core of a random graph is only of size $\Theta(N)$ if the random graph is supercritical (i.e. $2\beta_2 > 1$), so we work with that case.

It is known (see Theorem 1.4) that the giant component consists of its own core and a *mantle* of trees, each sprouting from a different vertex of the core. The proportion of vertices in the giant core is a.a.s. $\theta - 2\beta_2\theta(1-\theta)$ and the proportion of vertices in the mantle is $2\beta_2\theta(1-\theta)$, where $\theta = \inf\{t \geq 0 : 2\beta_2 t + \log(1-t) < 0\}$. Note that $\theta = t^*$ when we have $\beta_1 = 0$ i.e. no patches. We can imagine instead that we have an $o(N)$ number of patches, enough to ensure that a.a.s. one lands on the giant component (see Section 6.1 for a more detailed discussion). Then the process of identification basically picks out the giant component (and a few other smaller tree-like components which we may neglect). Thus, the proportion of identifiable vertices is $t^*$ which corresponds to the size of the giant component. Since the edges over vertices in the tree sections of the mantle are certainly essential, we see that these tree sections give us effectively all of the essential edges. Thus, because the mantle is like a forest, the number of edges and vertices in it are approximately equal and so the size of the mantle scaled by $1/N$ is the same as the number of essential edges scaled by $1/N$. Thus, $N^{-1}\mathbb{E}[\mathcal{E}(N,2)]$ converges to $2\beta_2 t^*(1-t^*)$.

## 5.3 Convergence of expectations

Let $V^*(\Lambda)$ be the set of identifiable vertices in the hypergraph $\Lambda$. Let $\Lambda\mathbb{1}_{\{A\}^C}$ be the Poisson random hypergraph $\Lambda$ with any hyperedges over the set $A$ removed. We say that a set $A \subseteq V$ is *essential for* $\Lambda$ if $\Lambda(A) = 1$ and $V^*(\Lambda\mathbb{1}_{\{A\}^C}) \neq V^*(\Lambda)$. Thus a set is essential if and only if it has an essential hyperedge over it.

**Proposition 5.2.** *A set $A$ is essential if and only if $\Lambda(A) = 1$ and $|A \setminus V^*(\Lambda\mathbb{1}_{\{A\}^C})| = 1$.*

*Proof.* It is clear that a set cannot be essential if it has no hyperedges over it. Also, it cannot be essential if it has more than one hyperedge over it. Recall that the order of deletion does not affect the set of identifiable edges. Suppose $|A \setminus V^*(\Lambda\mathbb{1}_{\{A\}^C})| = 0$. Then everything in $A$ is identifiable without the hyperedge over $A$ and so if we were to re-introduce the hyperedge, it would not be essential. If $|A \setminus V^*(\Lambda\mathbb{1}_{\{A\}^C})| \geq 2$ then replacing the hyperedge over $A$ will not make the two or more vertices identifiable and so $A$ cannot have been essential. $\square$

**Lemma 5.3.** *Under the assumptions of Theorem 1.6,*

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{N}\mathcal{E}(N,k)\right] \to (1-t^*)\beta'(t^*).$$

*Proof.* The numbers of hyperedges on distinct subsets of $V$ are independent and so $\Lambda\mathbb{1}_{\{A\}^C}$ has the same law as $\Lambda$ conditioned on $\Lambda(A)=0$. Suppose $|A|=k$ and use the shorthand $V^*=V^*(\Lambda)$. Then, by Proposition 5.2,

$$\begin{aligned}
\mathbb{P}\,(A \text{ is essential}) &= \mathbb{P}\left(\Lambda(A)=1, |A\setminus V^*(\Lambda\mathbb{1}_{\{A\}^c})|=1\right)\\
&= \mathbb{P}\,(\Lambda(A)=1)\,\mathbb{P}\left(|A\setminus V^*(\Lambda\mathbb{1}_{\{A\}^c})|=1\right)\\
&= \mathbb{P}\,(\Lambda(A)=1)\,\mathbb{P}\,(|A\setminus V^*|=1|\Lambda(A)=0)\\
&= \mathbb{P}\,(\Lambda(A)=1)\,\frac{\mathbb{P}\,(|A\setminus V^*|=1)}{\mathbb{P}\,(\Lambda(A)=0)} \quad \text{as } |A\setminus V^*|=1 \text{ implies } \Lambda(A)=0\\
&= \frac{N\beta_k}{\binom{N}{k}}\mathbb{P}\,(|A\setminus V^*|=1)\,.
\end{aligned}$$

Now, $A$ must contain $k-1$ identifiable vertices and one non-identifiable vertex and symmetry implies that, given $|V^*|$, all vertices are equally likely to be identifiable. So,

$$\begin{aligned}
\mathbb{P}\,(|A\setminus V^*|=1||V^*|) &= \frac{\binom{|V^*|}{k-1}(N-|V^*|)}{\binom{N}{k}}\\
&= k\left(1-\frac{|V^*|}{N}\right)\frac{|V^*|}{(N-1)}\frac{(|V^*|-1)}{(N-2)}\cdots\frac{(|V^*|-k+2)}{(N-k+1)}\\
&\xrightarrow{p} k(1-t^*)(t^*)^{k-1}
\end{aligned}$$

as $|V^*|/N \xrightarrow{p} t^*$. Thus, by bounded convergence,

$$\mathbb{P}\,(|A\setminus V^*|=1) = \mathbb{E}[\mathbb{P}\,(|A\setminus V^*|=1||V^*|)] \to k(1-t^*)(t^*)^{k-1}.$$

Now, taking $A_k$ to be a set of size $k$,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{N}\mathcal{E}(N,k)\right] &= \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{N}\sum_{\substack{A\subseteq V,\\ |A|=k}}\mathbb{1}_{\{\text{A is essential}\}}\right]\\
&= \frac{1}{N}\sum_{k=1}^{N}\binom{N}{k}\mathbb{P}\,(A_k \text{ is essential})\\
&= \mathbb{E}\left[\sum_{k=1}^{N}\beta_k\frac{\binom{|V^*|}{k-1}(N-|V^*|)}{\binom{N}{k}}\right].
\end{aligned}$$

The total number of edges on the hypergraph has a Poisson$(N\beta(1))$ distribution, and is always larger than the number of essential edges, $\sum_{k=1}^{N}\mathcal{E}(N,k)$. But $N^{-1}$Poisson$(N\beta(1))$ is bounded in $L^2$ and, therefore, uniformly integrable. Thus, as $N\to\infty$,

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{N}\mathcal{E}(N,k)\right] \to (1-t^*)\beta'(t^*). \qquad \square$$

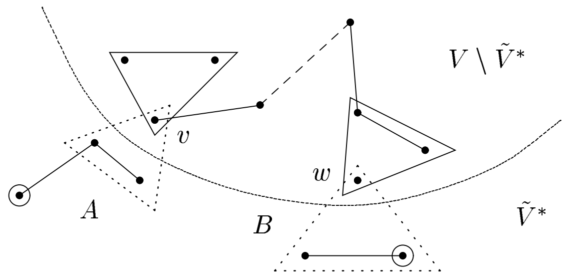


Figure 5.1: Example of two essential edges. The dashed curve indicates the boundary between $\tilde{V}^*$ and $V \setminus \tilde{V}^*$.

## 5.4 Asymptotic independence

The key point is to show that the events $\{A \text{ is essential}\}$ and $\{B \text{ is essential}\}$ are asymptotically independent, for distinct sets $A$ and $B$. Once we have done this, Theorem 5.1 can be proved in much the same way as the weak law of large numbers for independent random variables.

We now proceed as in the proof of Lemma 5.3. Suppose that $|A| = |B| = k$ where $A \neq B$. Then we have

$$\begin{aligned}
&\mathbb{P}\left(A \text{ is essential}, B \text{ is essential}\right) \\
&\qquad = \mathbb{P}\left(\Lambda(A) = 1, \Lambda(B) = 1, |A \setminus V^*(\Lambda \mathbb{1}_{\{A\}^c})| = 1, |B \setminus V^*(\Lambda \mathbb{1}_{\{B\}^c})| = 1\right).
\end{aligned}$$

Now, if we could deal with $\Lambda \mathbb{1}_{\{A,B\}^c}$ instead of $\Lambda \mathbb{1}_{\{A\}^c}$ and $\Lambda \mathbb{1}_{\{B\}^c}$ then we would be able to proceed easily by saying that $\Lambda \mathbb{1}_{\{A,B\}^c}$ has the same law as $\Lambda$ conditioned on $\Lambda(A) = 0$, $\Lambda(B) = 0$. However, we must then deal explicitly with the cases where $A$ contributes towards the identifiability of $B$, or vice versa. For ease of notation, write $\tilde{V}^*$ for $V^*(\Lambda \mathbb{1}_{\{A,B\}^c})$. We commence with some examples. In Figure 5.1, $|A \setminus \tilde{V}^*| = 1$ and $|B \setminus \tilde{V}^*| = 1$. Both $A$ and $B$ are essential as long as $v$ is not in the domain of $w$ in $V \setminus \tilde{V}^*$ and $w$ is not in the domain of $v$ in $V \setminus \tilde{V}^*$ (if the dashed edge is present then $w$ is identifiable from $v$ and so $B$ is not essential). In Figure 5.2, both $A$ and $B$ are essential but $|B \setminus \tilde{V}^*| = 2$. The important point here is that precisely one element $u$ of $B \setminus \tilde{V}^*$ is identifiable from $v$ and the other is not.

In general, in order to have $|A \setminus V^*(\Lambda \mathbb{1}_{\{A\}^c})| = 1$ and $|B \setminus V^*(\Lambda \mathbb{1}_{\{B\}^c})| = 1$ we must always have either $|A \setminus \tilde{V}^*| = 1$ or $|B \setminus \tilde{V}^*| = 1$ because otherwise none of $V \setminus \tilde{V}^*$ would

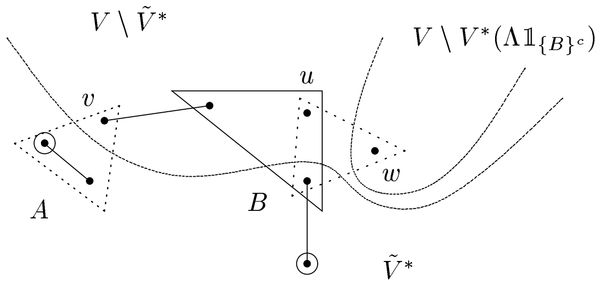


Figure 5.2: Example of two essential edges. The dashed curves indicate the boundaries between $\tilde{V}^*$ and $V \setminus \tilde{V}^*$ and between $V \setminus \tilde{V}^*$ and $V \setminus V(\Lambda \mathbb{1}_{\{B\}^c})$.

be identifiable when we reintroduce the edges over $A$ and $B$. Suppose, without loss of generality, that $A \setminus \tilde{V}^* = \{v\}$. Then we must also have that all but one of the vertices in $B \setminus \tilde{V}^*$ are in the domain of $v$. Take $\tilde{\mathcal{D}}(v)$ to be the domain of $v$ in the collapsed hypergraph on $V \setminus \tilde{V}^*$, $\mathcal{D}(v)$ to be the domain of $v$ in the collapsed hypergraph on $V \setminus V^*$ and $\{w\} = B \setminus \tilde{V}^*$ when $|B \setminus \tilde{V}^*| = 1$. Assume that $v \neq w$. Then,

$$
\begin{aligned}
&\mathbb{P}\left(\Lambda(A) = 1, \Lambda(B) = 1, |A \setminus V^*(\Lambda \mathbb{1}_{\{A\}^c})| = 1, |B \setminus V^*(\Lambda \mathbb{1}_{\{B\}^c})| = 1\right) \\
&= \mathbb{P}\Big(\Lambda(A) = 1, \Lambda(B) = 1, |A \setminus \tilde{V}^*| = 1, |B \setminus \tilde{V}^*| = 1, v \notin \tilde{\mathcal{D}}(w), w \notin \tilde{\mathcal{D}}(v)\Big) \\
&\qquad + \sum_{i=2}^{k} \mathbb{P}\Big(\Lambda(A) = 1, \Lambda(B) = 1, |A \setminus \tilde{V}^*| = 1, |B \setminus \tilde{V}^*| = i, |\tilde{\mathcal{D}}(v) \cap B \setminus \tilde{V}^*| = i - 1\Big) \\
&\qquad + \sum_{i=2}^{k} \mathbb{P}\Big(\Lambda(A) = 1, \Lambda(B) = 1, |B \setminus \tilde{V}^*| = 1, |A \setminus \tilde{V}^*| = i, |\tilde{\mathcal{D}}(w) \cap A \setminus \tilde{V}^*| = i - 1\Big) \\
&= \mathbb{P}\left(\Lambda(A) = 1\right) \mathbb{P}\left(\Lambda(B) = 1\right) \Bigg[\mathbb{P}\Big(|A \setminus \tilde{V}^*| = 1, |B \setminus \tilde{V}^*| = 1, v \notin \tilde{\mathcal{D}}(w), w \notin \tilde{\mathcal{D}}(v)\Big) \\
&\qquad\qquad + 2\sum_{i=2}^{k} \mathbb{P}\Big(|A \setminus \tilde{V}^*| = 1, |B \setminus \tilde{V}^*| = i, |\tilde{\mathcal{D}}(v) \cap B \setminus \tilde{V}^*| = i - 1\Big)\Bigg].
\end{aligned}
$$

We now use the fact that $\Lambda \mathbb{1}_{\{A,B\}^c}$ has the same law as $\Lambda$ conditioned on $\Lambda(A) = 0$ and

$\Lambda(B) = 0$ to show that this is equal to

$$\left(\frac{N\beta_k}{\binom{N}{k}} \exp\left(-N\beta_k / \binom{N}{k}\right)\right)^2$$
$$\times \Bigg[ \mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = 1, v \notin \mathcal{D}(w), w \notin \mathcal{D}(v) | \Lambda(A) = 0, \Lambda(B) = 0\right)$$
$$+ 2 \sum_{i=2}^{k} \mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = i, |\mathcal{D}(v) \cap B \setminus V^*| = i - 1 | \Lambda(A) = 0, \Lambda(B) = 0\right) \Bigg].$$

We have that $|A \setminus V^*| = 1, |B \setminus V^*| = 1$ implies $\Lambda(A) = 0,\ \Lambda(B) = 0$ and so this equals

$$\left(\frac{N\beta_k}{\binom{N}{k}} \exp\left(-N\beta_k / \binom{N}{k}\right)\right)^2 \Bigg[ \frac{\mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = 1, v \notin \mathcal{D}(w), w \notin \mathcal{D}(v)\right)}{\mathbb{P}\left(\Lambda(A) = 0\right) \mathbb{P}\left(\Lambda(B) = 0\right)}$$
$$+ 2 \sum_{i=2}^{k} \frac{\mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = i, |\mathcal{D}(v) \cap B \setminus V^*| = i - 1, \Lambda(A) = 0, \Lambda(B) = 0\right)}{\mathbb{P}\left(\Lambda(A) = 0\right) \mathbb{P}\left(\Lambda(B) = 0\right)} \Bigg].$$

Finally, we can cancel $\mathbb{P}\left(\Lambda(A) = 0\right) \mathbb{P}\left(\Lambda(B) = 0\right)$ and $\exp(-N\beta_k / \binom{N}{k})^2$ to obtain

$$\begin{aligned} &\mathbb{P}\left(\Lambda(A) = 1, \Lambda(B) = 1, |A \setminus V^*(\Lambda \mathbb{1}_{\{A\}^c})| = 1, |B \setminus V^*(\Lambda \mathbb{1}_{\{B\}^c})| = 1\right) \\ &= \left(\frac{N\beta_k}{\binom{N}{k}}\right)^2 \Bigg[ \mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = 1, v \notin \mathcal{D}(w), w \notin \mathcal{D}(v)\right) \\ &\quad + 2 \sum_{i=2}^{k} \mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = i, |\mathcal{D}(v) \cap B \setminus V^*| = i - 1, \Lambda(A) = 0, \Lambda(B) = 0\right) \Bigg]. \end{aligned} \tag{5.1}$$

For the moment, we will assume that

$$\mathbb{P}\left(v \notin \mathcal{D}(w), w \notin \mathcal{D}(v)\right) \to 1 \tag{5.2}$$
$$\mathbb{P}\left(|\mathcal{D}(v) \cap B \setminus V^*| \geq 1\right) \to 0 \tag{5.3}$$

as $N \to \infty$ which, in particular, means that we may discard the second term in (5.1). These results will be proved in the next section and reflect the fact that the collapse algorithm will not die out while the vertex domains remain supercritical. Finally, as we will also show later,

$$\mathbb{P}\left(|A \setminus V^*| = 1, |B \setminus V^*| = 1\right) \to (k(t^*)^{k-1}(1 - t^*))^2$$

and so

$$\begin{aligned}\mathbb{P}\left(A \text{ is essential}, B \text{ is essential}\right) &\sim \left(\frac{N\beta_k}{\binom{N}{k}} k(t^*)^{k-1}(1-t^*)\right)^2 \\ &= \mathbb{P}\left(A \text{ is essential}\right)\mathbb{P}\left(B \text{ is essential}\right).\end{aligned}$$

## 5.5 Proof of Theorem 5.1

We wish to prove the statements (5.2) and (5.3). As a first step, we prove

**Lemma 5.4.** *Let $\gamma(t) = (1-t)\beta''(t)$. Then*

$$\gamma(t^*) \leq 1.$$

*Proof.* For $t \in [0,1]$, define $f(t) = 1 - t - \exp(-\beta'(t))$. Then

$$f'(t) = \beta''(t)\exp(-\beta'(t)) - 1.$$

Now $t^* = \inf\{t \geq 0 : f(t) < 0\}$ and so $f'(t^*) = (1-t^*)\beta''(t^*) - 1 \leq 0$. Hence result. $\square$

As we shall soon see, the 2-edge parameter for the collapsed hypergraph is approximately $\frac{1}{2}\gamma(t^*)$ and so this lemma says that the collapsed hypergraph is nearly subcritical.

**Lemma 5.5.** *Suppose that $v, w \notin V^*$ are chosen uniformly at random. Recall that $\mathcal{D}(v)$ is the domain of $v$ in the collapsed hypergraph on $V \setminus V^*$ and that $|\mathcal{D}(v)| = D(v)$. Then*

$$\mathbb{P}\left(w \in \mathcal{D}(v)\right) \to 0$$

*as $N \to \infty$.*

*Proof.* Suppose $|V^*| = n$ and let $\Lambda^{V^*}$ be the hypergraph obtained from $\Lambda$ by removing all of the vertices in $V^*$. Then $\Lambda^{V^*}$ is the collapsed hypergraph. For $v \in V \setminus V^*$,

$$\Lambda^{V^*}(\{v\}) = 0$$

as $v$ cannot be identifiable. For $A \subseteq V \setminus V^*$ with $|A| = j \geq 2$, conditional on $V^*$,

$$\Lambda^{V^*}(A) = \sum_{\substack{B \supseteq A, \\ B \setminus V^* = A}} \Lambda(B) \sim \text{Poisson}\left(N \sum_{i=0}^{n} \beta_{i+j}\binom{n}{i} \Big/ \binom{N}{i+j}\right)$$

and these random variables are independent. So $\Lambda^{V^*}$ is a new Poisson random hypergraph on $N-n$ vertices and with parameters $\beta_k(n,N)$ where

$$\beta_k(n,N) = \begin{cases} 0 & \text{if } k=1 \\ \frac{N}{N-n}\binom{N-n}{k}\sum_{i=0}^{n}\beta_{i+k}\binom{n}{i}/\binom{N}{i+k} & \text{if } k \geq 2. \end{cases}$$

Since $\beta$ is a polynomial, it is straightforward to obtain the estimate

$$|\beta_2(n,N) - \tfrac{1}{2}\gamma(\tfrac{n}{N})| \leq C/N$$

for some constant $C < \infty$ and all $n \leq N$. Furthermore,

$$\beta_k(n,N) \leq \frac{1}{k!}\beta^{(k)}(1)$$

for all $k \geq 2$ and all $n \leq N$.

Now let $C_t$ have the distribution of the size of the domain of a vertex in a Poisson random hypergraph with parameters $\beta_1 = 0$, $\beta_2 = \frac{1}{2}\gamma(t) + C/N$ and $\beta_k = \frac{1}{k!}\beta^{(k)}(1)$ for $k \geq 3$. Then by Theorem 1.7 and an obvious comparison argument,

$$C_t \xrightarrow{d} \text{Borel}(\gamma(t))$$

as $N \to \infty$. Also, if $|V^*| = Nt$ then $D(v)$ is stochastically dominated by $C_t$.

For any $\delta > 0$, we can choose $\epsilon > 0$ small enough and $R < \infty$ large enough that

$$\mathbb{P}\left(\text{Borel}(\gamma(t^*+\epsilon)) \leq R\right) \geq 1-\delta$$

because $\gamma(t^*) \leq 1$ and $\gamma$ is continuous. Hence,

$$\begin{aligned}
\mathbb{P}\left(w \in \mathcal{D}(v)\right) &= \mathbb{P}\left(w \in \mathcal{D}(v), \left|\frac{|V^*|}{N} - t^*\right| \leq \epsilon\right) + \mathbb{P}\left(w \in \mathcal{D}(v), \left|\frac{|V^*|}{N} - t^*\right| > \epsilon\right) \\
&\leq \mathbb{P}\left(w \in \mathcal{D}(v), D(v) \leq R, \left|\frac{|V^*|}{N} - t^*\right| \leq \epsilon\right) \\
&\qquad + \mathbb{P}\left(D(v) > R, \left|\frac{|V^*|}{N} - t^*\right| \leq \epsilon\right) + \mathbb{P}\left(\left|\frac{|V^*|}{N} - t^*\right| > \epsilon\right) \\
&\leq \frac{R}{N(1-t^*-\epsilon)} + \mathbb{P}\left(C_{t^*+\epsilon} > R\right) + \mathbb{P}\left(\left|\frac{|V^*|}{N} - t^*\right| > \epsilon\right) \\
&\to \mathbb{P}\left(\text{Borel}(\gamma(t^*+\epsilon)) > R\right) \quad \text{as } N \to \infty
\end{aligned}$$

But this last quantity is less than $\delta$ and so we are done. □

Now we are ready to prove Theorem 5.1.

*Proof of Theorem 5.1.* Write $\beta(t) = \sum_{k=1}^{M} \beta_k t^k$. Observe that

$$\frac{1}{N}\sum_{k=1}^{M}\mathcal{E}(N,k) = \sum_{k=1}^{M}\frac{1}{N}\sum_{\substack{A\subseteq V,\\ |A|=k}} \mathbb{1}_{\{A \text{ is essential}\}}$$

and that if each of the $M$ terms in the sum converges in probability then so does the sum. Take $\epsilon > 0$ and fix $1 \leq k \leq M$. Then, by Chebyshev's inequality,

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{N}\sum_{\substack{A\subseteq V,\\ |A|=k}}(\mathbb{1}_{\{A \text{ is essential}\}} - (1-t^*)k\beta_k(t^*)^{k-1})\right| > \epsilon\right)\\
&\leq \frac{1}{\epsilon^2}\mathbb{E}\left[\left(\frac{1}{N}\sum_{\substack{A\subseteq V,\\ |A|=k}}(\mathbb{1}_{\{A \text{ is essential}\}} - (1-t^*)k\beta_k(t^*)^{k-1})\right)^2\right]\\
&= \frac{1}{\epsilon^2}\frac{1}{N^2}\sum_{\substack{A,B\subseteq V, A\neq B\\ |A|=|B|=k}}\left\{\mathbb{P}\left(A \text{ is essential, } B \text{ is essential}\right) - ((1-t^*)k\beta_k(t^*)^{k-1})^2\right\}\\
&\quad + \frac{1}{\epsilon^2}\left\{\frac{1}{N^2}\mathbb{E}[\mathcal{E}(N,k)] + 2(1-t^*)k\beta_k(t^*)^{k-1}\left[(1-t^*)k\beta_k(t^*)^{k-1} - \frac{1}{N}\mathbb{E}[\mathcal{E}(N,k)]\right]\right\}
\end{aligned}$$

where

$$\frac{1}{\epsilon^2}\left\{\frac{1}{N^2}\mathbb{E}[\mathcal{E}(N,k)] + 2(1-t^*)k\beta_k(t^*)^{k-1}\left[(1-t^*)k\beta_k(t^*)^{k-1} - \frac{1}{N}\mathbb{E}[\mathcal{E}(N,k)]\right]\right\} \to 0$$

as $N \to \infty$, by Lemma 5.3. From (5.1),

$$\begin{aligned}
&\frac{1}{N^2}\sum_{\substack{A,B\subseteq V, A\neq B\\ |A|=|B|=k}}\mathbb{P}\left(A \text{ is essential, } B \text{ is essential}\right)\\
&= \frac{1}{N^2}\sum_{\substack{A,B\subseteq V, A\neq B\\ |A|=|B|=k}}\frac{N^2\beta_k^2}{\binom{N}{k}^2}\Bigg\{\mathbb{P}\left(|A\setminus V^*| = 1, |B\setminus V^*| = 1, v\notin\mathcal{D}(w), w\notin\mathcal{D}(v)\right)\\
&\quad + 2\sum_{i=2}^{k}\mathbb{P}\left(|A\setminus V^*| = 1, |B\setminus V^*| = i, |\mathcal{D}(v)\cap B\setminus V^*| = i-1, \Lambda(A) = 0, \Lambda(B) = 0\right)\Bigg\}
\end{aligned}$$

Now, by Lemma 5.5, $\mathbb{P}\left(v\notin\mathcal{D}(w), w\notin\mathcal{D}(v)\right) \to 1$ and $\mathbb{P}\left(|\mathcal{D}(v)\cap B\setminus V^*| \geq 1\right) \to 0$ as

$N \to \infty$. Thus, we are really interested in

$$\begin{aligned}
&\frac{\beta_k^2}{\binom{N}{k}^2} \sum_{\substack{A,B\subseteq V, A\neq B\\ |A|=|B|=k}} \mathbb{P}\left(|A\setminus V^*|=1, |B\setminus V^*|=1\right)\\
&\quad= \frac{\beta_k^2}{\binom{N}{k}^2} \sum_{\substack{A,B\subseteq V, A\cap B=\emptyset\\ |A|=|B|=k}} \mathbb{P}\left(|A\setminus V^*|=1, |B\setminus V^*|=1\right)\\
&\qquad+ \frac{\beta_k^2}{\binom{N}{k}^2} \sum_{i=1}^{k-1} \sum_{\substack{A,B\subseteq V, |A\cap B|=i\\ |A|=|B|=k}} \mathbb{P}\left(|A\setminus V^*|=1, |B\setminus V^*|=1\right) \qquad (5.4)
\end{aligned}$$

In the case where $A$ and $B$ are disjoint, by a similar argument to that used in the proof of Lemma 5.3, we have

$$\mathbb{P}\left(|A\setminus V^*|=1, |B\setminus V^*|=1\right) = \mathbb{E}\left[\frac{(N-|V^*|)(N-|V^*|-1)\binom{|V^*|}{k-1}\binom{|V^*|-k+1}{k-1}}{\binom{N}{k}\binom{N-k}{k}}\right].$$

If $|A\cap B| = i$ for $1 \le i \le k-1$, then the two non-identifiable vertices must lie outside the intersection (otherwise they are the same vertex and so $v = w$). Thus,

$$\begin{aligned}
&\mathbb{P}\left(|A\setminus V^*|=1, |B\setminus V^*|=1\right)\\
&\quad= \mathbb{E}\left[\frac{(N-|V^*|)\binom{|V^*|}{k-i-1}\binom{|V^*|-k+i+1}{i}(N-|V^*|-1)\binom{|V^*|-k+1}{k-i-1}}{\binom{N}{k}\binom{k}{i}\binom{N-k}{k-i}}\right].
\end{aligned}$$

So (5.4) is equal to

$$\begin{aligned}
&\frac{\beta_k^2}{\binom{N}{k}^2}\mathbb{E}\left[(N-|V^*|)(N-|V^*|-1)\binom{|V^*|}{k-1}\binom{|V^*|-k+1}{k-1}\right]\\
&\quad+ \frac{\beta_k^2}{\binom{N}{k}^2}\sum_{i=1}^{k-1}\mathbb{E}\left[(N-|V^*|)\binom{|V^*|}{k-i-1}\binom{|V^*|-k+i+1}{i}(N-|V^*|-1)\binom{|V^*|-k+1}{k-i-1}\right]
\end{aligned}$$

As $\frac{|V^*|}{N} \to t^*$ in probability, the first of these two terms converges to

$$(k\beta_k(1-t^*)(t^*)^{k-1})^2,$$

by bounded convergence. It remains to show that the second term converges to 0. We have

$$\begin{aligned}
&\sum_{i=1}^{k-1} \mathbb{E}\left[\frac{(N-|V^*|)\binom{|V^*|}{k-i-1}\binom{|V^*|-k+i+1}{i}(N-|V^*|-1)\binom{|V^*|-k+1}{k-i-1}}{\binom{N}{k}^2}\right] \\
&\quad \leq \sum_{i=1}^{k-1} \frac{(k!)^2}{i!((k-i-1)!)^2(N-k+i+1)!} \\
&\quad \to 0
\end{aligned}$$

as $N \to \infty$. The result follows. $\square$

# Chapter 6

# Critical Random Hypergraphs

## 6.1 Introduction

The emergence of the giant component in a random graph is now a well-understood phenomenon (see Janson, Łuczak and Ruciński [20]). The purpose of this chapter is to demonstrate that an analogous, but richer, phenomenon occurs in random hypergraphs. Recall from the Introduction that we have assumed that $\beta(1) < \infty$ and that we define $x(t) = \inf\{y \geq 0 : t\beta'(y) + \log(1-y) < 0\}$. Our starting point is Theorem 1.8, which concerns the behaviour of the proportion of identifiable vertices in a Poisson random hypergraph process when the limit path $x(t)$ has a discontinuity. An initial discussion of the hypergraph process will help our understanding of the static hypergraph.

Here we are interested in a particular example of a discontinuity in the function $x(t)$, which was first exhibited in Darling, Levin and Norris [11]. We will restrict our attention to the case $\beta_1 = 0$, $\beta_2 = 1/2$ (i.e. there are no patches and the underlying random graph becomes critical at $t = 1$. Note that, for the moment, $x(t)$ does not represent a proportion of identifiable vertices because there are no patches). Then $x(t)$ is zero until $t = 1$ at which point it becomes positive. If $\beta_3 > 1/6$ then there is a discontinuity at $t = 1$ but not if $\beta_3 < 1/6$. If $\beta_3 = 1/6$ then the discontinuity occurs if $\beta_4 > 1/12$ but not if $\beta_4 < 1/12$. In general, the discontinuity occurs if there exists a $k < \infty$ such that $\beta_k > 1/k(k-1)$ and for all $2 \leq j \leq k-1$, $\beta_j = 1/j(j-1)$. If $\beta_k < 1/k(k-1)$ and for all $2 \leq j \leq k-1$, $\beta_j = 1/j(j-1)$ then $x(t)$ is continuous. Figure 6.1 shows the two possibilities.

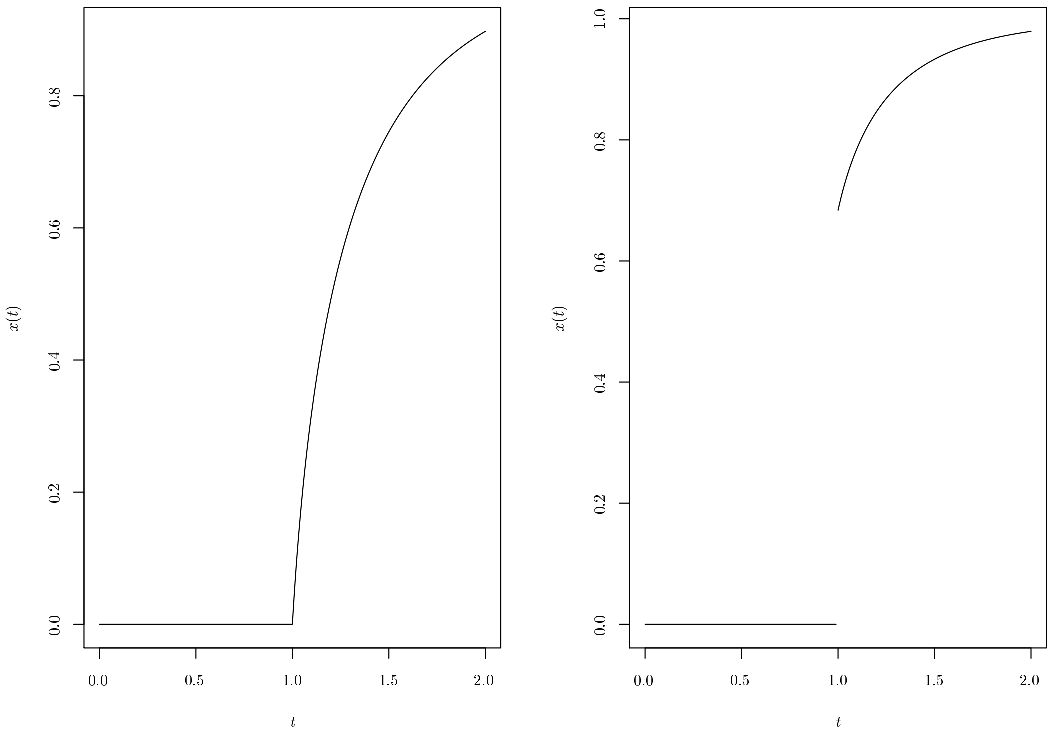


Figure 6.1: On the left, $\beta_2 = 1/2$, $\beta_3 = 1/10$ and there is no discontinuity; on the right, $\beta_2 = 1/2$, $\beta_3 = 1/3$ and there is a discontinuity in $x(t)$ at $t = 1$.

If we add a single patch to a Poisson random hypergraph process at time 0 then the limiting proportion of identifiable vertices is 0 until some random time after $t = 1$ when the proportion jumps up and follows the path of $x(t)$. The delay is most easily explained by considering the case of a Poisson random graph: we must wait until the patch is joined into the giant component (which a.a.s. exists and is unique for any $t > 1$), at which point it identifies that component, i.e. a positive proportion of the vertices.

In a hypergraph, the set of vertices made identifiable when we add a single patch is the domain of the vertex under the patch. While the largest domain is not an easy object to study, it is reasonable to discuss a giant domain, meaning one whose size is a strictly positive proportion of the total number of vertices. If there exists a giant 2-edge component then the domains of all of its vertices are certainly giant domains. Thus, there must exist a time when the domain of our single patched vertex becomes a giant domain, which gives the time when the limiting proportion of identifiable vertices jumps to follow the path of $x(t)$.

In a random graph, even if our patch happened to be on the largest component at time 1, we would identify only a zero proportion of the vertices because the largest component is of size $\Theta(N^{2/3})$ (as is reflected by $x(1) = 0$). If, instead of a single patch, we put a number $\omega(N)$ of patches on the graph, where $\omega(N) \to \infty$ as $N \to \infty$ but $\omega(N)/N \to 0$, we still would not attain a positive proportion of the vertices. This is also true for any critical random hypergraph such that $x(t)$ is continuous at 1. (To see this, note that if we add a number $\epsilon N$ of patches, we identify an asymptotic proportion of the vertices given by Theorem 1.6. It is easily shown that this proportion converges to 0 as $\epsilon \to 0$. But $\omega(N) < \epsilon N$ for any $\epsilon > 0$ as long as $N$ is sufficiently large and so, by monotonicity, adding $\omega(N)$ patches must yield a limiting proportion 0.) However, if $x(t)$ is discontinuous, does there exist a number $\omega(N)$ of patches such that we identify $\Theta(N)$ vertices at $t = 1$ (i.e. attain the upper point of this discontinuity)? We will show that such a $\omega(N)$ exists, revealing that there can be a significant difference between the behaviour at criticality in a random graph and in a random hypergraph.

## 6.2 Results

Consider a static Poisson random hypergraph with $\beta_1 = 0$ and $\beta_2 = 1/2$. Define $\alpha(k) = (2k-4)/(2k-3)$.

(i) If there exists $k$ such that for all $2 \leq j \leq k-1$, $\beta_j = 1/j(j-1)$ and $\beta_k < 1/k(k-1)$ then adding $o(N)$ patches will never result in the identification of a positive proportion of the vertices. The domains of vertices are of size $\mathcal{O}(N^{\alpha(k)})$.

(ii) If there exists $k$ such that for all $2 \leq j \leq k-1$, $\beta_j = 1/j(j-1)$ and $\beta_k > 1/k(k-1)$ then adding $CN^{\alpha(k)/2}$ patches for any $0 < C < \infty$ gives positive probability of identifying a positive proportion of the vertices. Moreover, $N^{1/2}$ patches will always suffice to identify a positive proportion, for any value of $k$.

(iii) In the preceding case, the proportion identified converges in probability to $t^* = \inf\{y \geq 0 : \beta'(y) + \log(1-y) < 0\} \wedge 1$.

We do not know what happens if $\beta_j = 1/j(j-1)$ for all $j \geq 2$.

## 6.3 Breadth-first walk

We will perform our analysis by adding a patch to the hypergraph, collapsing as far as possible and then adding another patch whenever needed to keep the process of collapse going. In order to track what is going on, we construct an extension to hypergraphs of the usual breadth-first walk on graphs. (Note that this is a different extension to that used in Schmidt-Pruzan and Shamir [29].)

Consider any hypergraph on $N$ vertices, with no patches and an arbitrary numbering of the vertices. Then we may define the *breadth-first ordering* as follows:

- Take the lowest-numbered vertex, call it $v(1)$ and put a patch on it. Define the *children* of vertex $v(1)$ to be those vertices connected to it by a 2-edge. Suppose that $v(1)$ has $c(1)$ children. Number them $v(2), \ldots, v(c(1)+1)$, retaining the ordering of the original labels. Now collapse the patch on $v(1)$, leaving patches on all of its children and any higher-order edges collapsed onto their remaining vertices.

- Now look at $v(2)$. Label its children as $v(c(1)+2), \ldots, v(c(1)+c(2)+1)$, where, in general, we define the children of a vertex to be those vertices it is has a 2-edge to which have not already been numbered. Note that some of these children may only just have appeared as a result of the collapse of vertex $v(1)$. For example, in Figure 6.2, $v(3)$ is the child of $v(2)$ but only becomes visible as such after the deletion of $v(1)$.

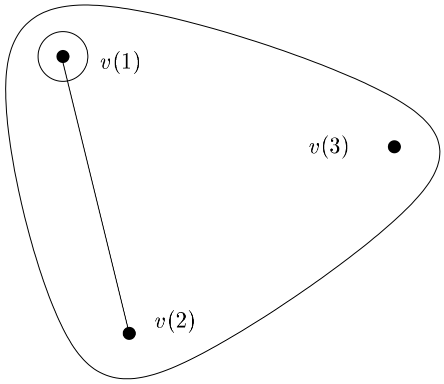


Figure 6.2: Children of vertices can appear during the process of collapse.

- Continue in this way, collapsing the vertices in numerical order (so the next one to consider is $v(3)$). When we run out of patches, pick the next lowest-numbered vertex in the old ordering, put a patch on it and proceed as before. The process terminates when there are no more vertices to consider.

So, loosely speaking, we number within levels of an underlying tree before moving further from the 'root', the only complication being that the children of $v(i)$ may only all be visible after the deletion of vertex $v(i-1)$.

Now we can define a walk $(z(i))_{0 \leq i \leq N}$ on the integers associated with this hypergraph by

$$\begin{aligned} z(0) &= 0 \\ z(i) &= z(i-1) + c(i) - 1, \quad i \geq 1, \end{aligned}$$

where, as before, $c(i)$ is the number of children of vertex $v(i)$ in the breadth-first ordering. Then $i$ is the number of vertex deletions (we will also refer to this as a *time*) and $z(i)$ is the number of patched vertices on the hypergraph after the $i$th vertex deletion, minus the number of patches added before the $(i+1)$th deletion to keep the process going. The process $(z(i))_{0 \leq i \leq N}$ is called the breadth-first walk.

Thus, for a *random* hypergraph on $N$ vertices, we obtain a random walk, $(Z_N(i))_{i \geq 0}$, on the integers which summarises information about the hypergraph. (In the sequel, we will refer to the random number of children of vertex $v(i)$ as $C_N(i)$.) Most importantly, the number of vertices which are identifiable from the patches we add are coded as excursions above past minima in the breadth-first walk. This is because the breadth-first walk picks out an underlying forest structure, with each tree naturally having one more vertex than it has edges. For more details, see Aldous [1].

It will be useful later to have some notation for the number of patches added to the hypergraph so far. Let $P_N(0) = 1$ and, for $i \geq 1$,

$$P_N(i) = 1 - \min_{j \leq i} Z_N(j).$$

Then $P_N(i)$ is the number of patches added before the $(i+1)$th vertex deletion. $Z_N(i) + P_N(i)$ is the actual number of patches on the hypergraph just after the deletion of $v(i)$ and is always greater than or equal to 1. Thus, we have the alternative representation

$$P_N(i) = 1 + \sum_{i=1}^{N-1} \mathbb{1}_{\{Z_N(i-1)+P_N(i-1)=1,\ C_N(i)=0\}}. \tag{6.1}$$

Let $\alpha(k) = (2k-4)/(2k-3)$ and

$$W^k(t) = B(t) + \frac{1}{k-1}(k(k-1)\beta_k - 1)t^{k-1},$$

where $(B(t))_{t \geq 0}$ is a standard Brownian motion. Then our main result is the following:

**Theorem 6.1.** *Suppose that $(Z_N(i))_{i \geq 0}$ is the breadth-first walk on the Poisson random hypergraph with no patches (i.e. $\beta_1 = 0$) and that there exists a $k \geq 3$ such that for all $2 \leq j \leq k-1$, $\beta_j = 1/j(j-1)$. Rescale by defining*

$$\bar{Z}_N(t) = N^{-\alpha(k)/2} Z_N(\lfloor N^{\alpha(k)} t \rfloor).$$

*Then*

$$\bar{Z}_N \xrightarrow{d} W^k$$

*as $N \to \infty$ in $D[0, \infty)$.*

Note that here the convergence is uniform on compact time-intervals. The proof of this result is deferred to Section 6.6 to enable us to first interpret its implications for the hypergraph.

## 6.4 Consequences

If there exists $k$ such that $\beta_j = 1/j(j-1)$ for all $2 \leq j \leq k-1$ and $\beta_k < 1/k(k-1)$ , then the limit process has a negative drift which increases in magnitude with time. Thus, the process keeps hitting its previous minima, on average resulting in smaller and smaller numbers of identifiable vertices per patch added. This is very like what we see in Theorem 3 of Aldous [1] where the components of a random graph appear in size-biased order. In the critical random graph case, $\beta_3 = 0 < 1/6$ and the components are of size $\mathcal{O}(N^{2/3})$, as is well-known. However, in the random hypergraph, there is a whole series of critical scalings ($N^{2/3}$, $N^{4/5}$, $N^{6/7}$, $\ldots$) which can be attained by suitable adjustments of the parameters $\beta_3, \beta_4, \ldots$. Thus, the random hypergraph demonstrates much richer behaviour than the random graph.

If there exists $k$ such that $\beta_j = 1/j(j-1)$ for all $2 \leq j \leq k-1$ and $\beta_k > 1/k(k-1)$, then the process $W^k$ has positive drift and so there is a (random) last time that it hits its own minimum. This signals the start of a giant excursion which is too big to be measured on the scale of $N^{\alpha(k)}$. We wish to prove that the domain which this excursion represents is, in fact, of size comparable to $N$. In order to do this, we will show that the giant excursion has length at least $N\delta$ for all sufficiently small $\delta > 0$. This will then allow us to prove a fluid limit theorem for the breadth-first walk; the length of the excursion of the fluid limit above 0 gives us the asymptotic size of the giant set of identifiable vertices. We will also discuss the fluctuations of the breadth-first walk around this fluid limit.

Before proceeding, it remains to explain how to convert our results for the breadth-first walk into the situation where we add $\omega(N)$ patches to the hypergraph right at the start, rather than one by one. Suppose we are in the case where $\beta_k > 1/k(k-1)$ and $\beta_j = 1/j(j-1)$ for all $2 \leq j \leq k-1$. Thus, adding $\Theta(N^{\alpha(k)/2})$ patches one by one will be sufficient to cause the breadth-first walk to have a giant excursion. Then, taking $\omega(N) = CN^{\alpha(k)/2}$ patches right at the start (for some constant $C > 0$) gives positive probability of identifying a giant set of vertices. Furthermore, as $\alpha(k) \to 1$ as $k \to \infty$, $\Theta(N^{1/2})$ patches will always suffice to identify a giant set of vertices.

Figures 6.3 to 6.5 are the results of some simulations to illustrate various cases of Theorem 6.1. In each of them, $N = 100\,000$, $\beta_2 = 1/2$ and $\beta_k = 0$ for $k \geq 4$.

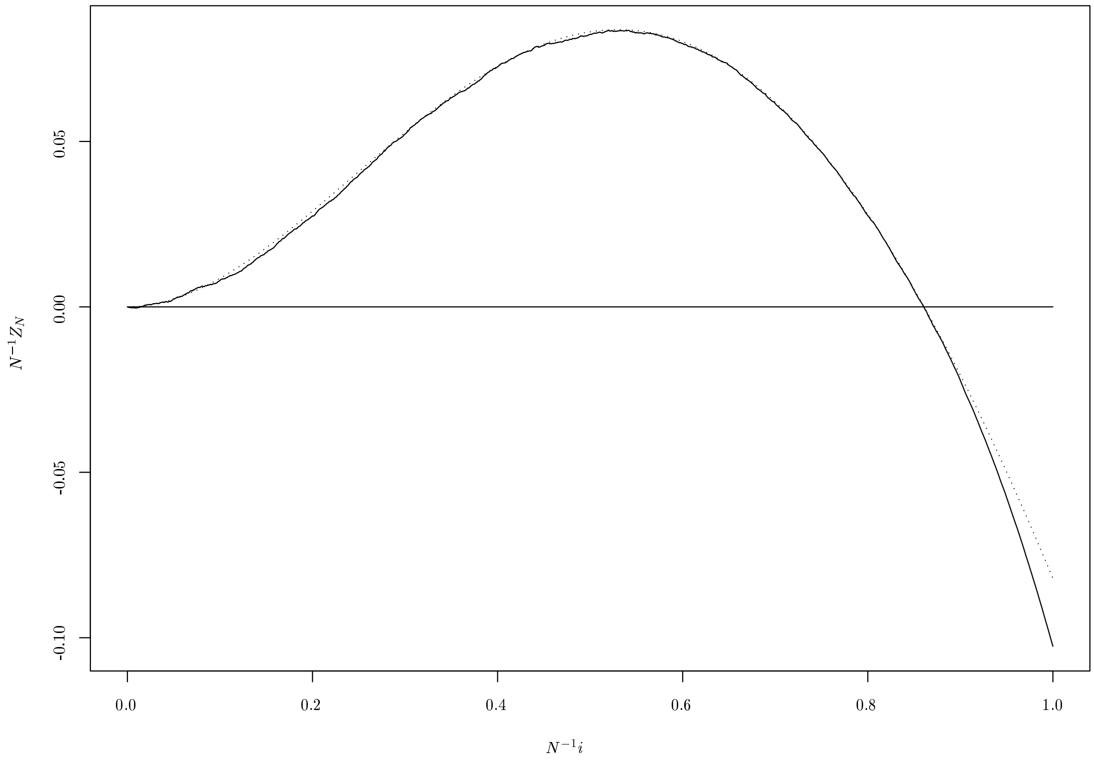


Figure 6.3: Here $\beta_3 = 1/2 > 1/6$ and, scaled by $1/N$ on both axes, the breadth-first walk has a large excursion of size $\Theta(N)$. The (barely visible) dotted line is the fluid limit for the breadth-first walk (see Theorem 6.4).

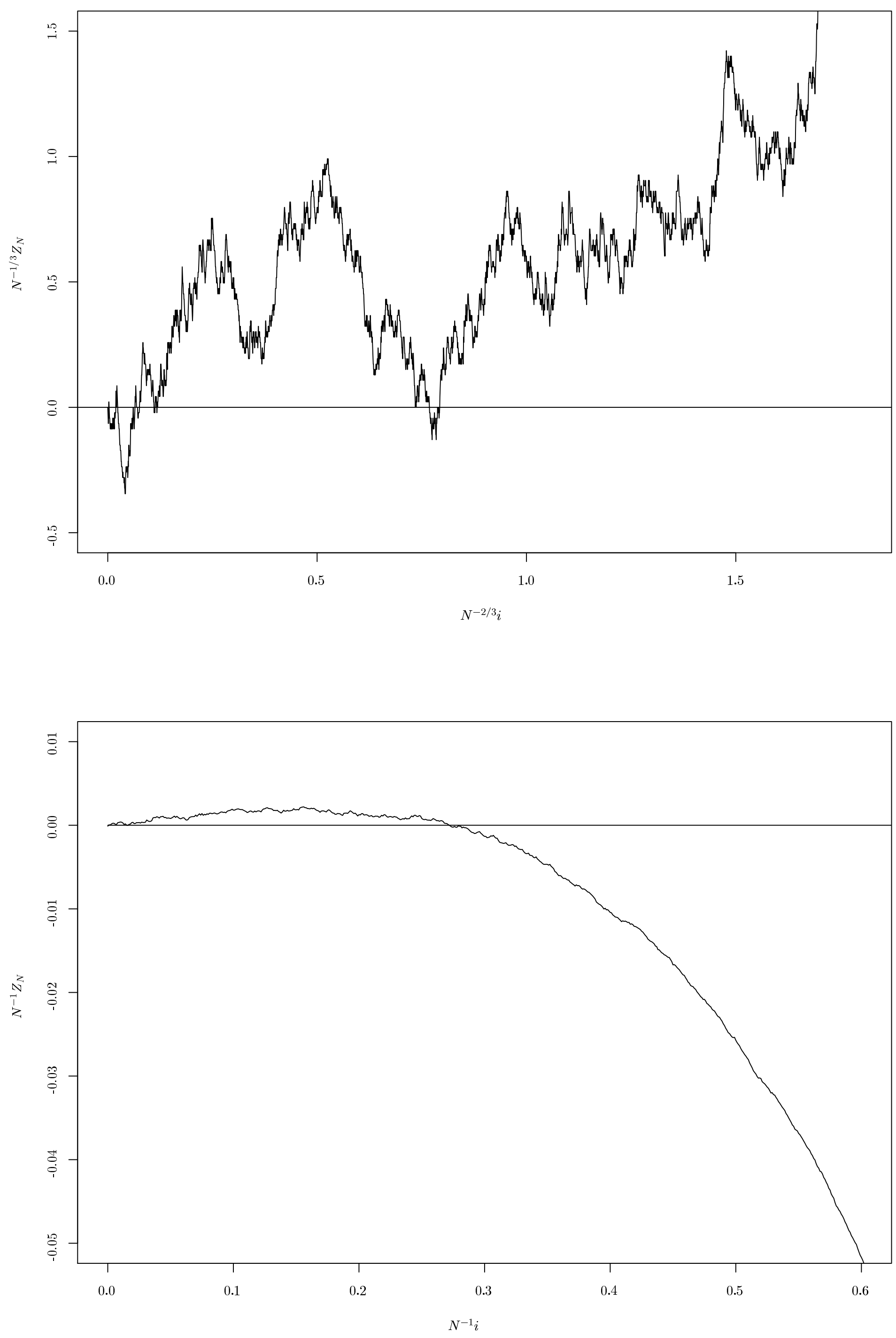


Figure 6.4: Case $\beta_3 = 1/5 > 1/6$. Top: the Brownian limit has just positive drift. Bottom: the macroscopic scale.

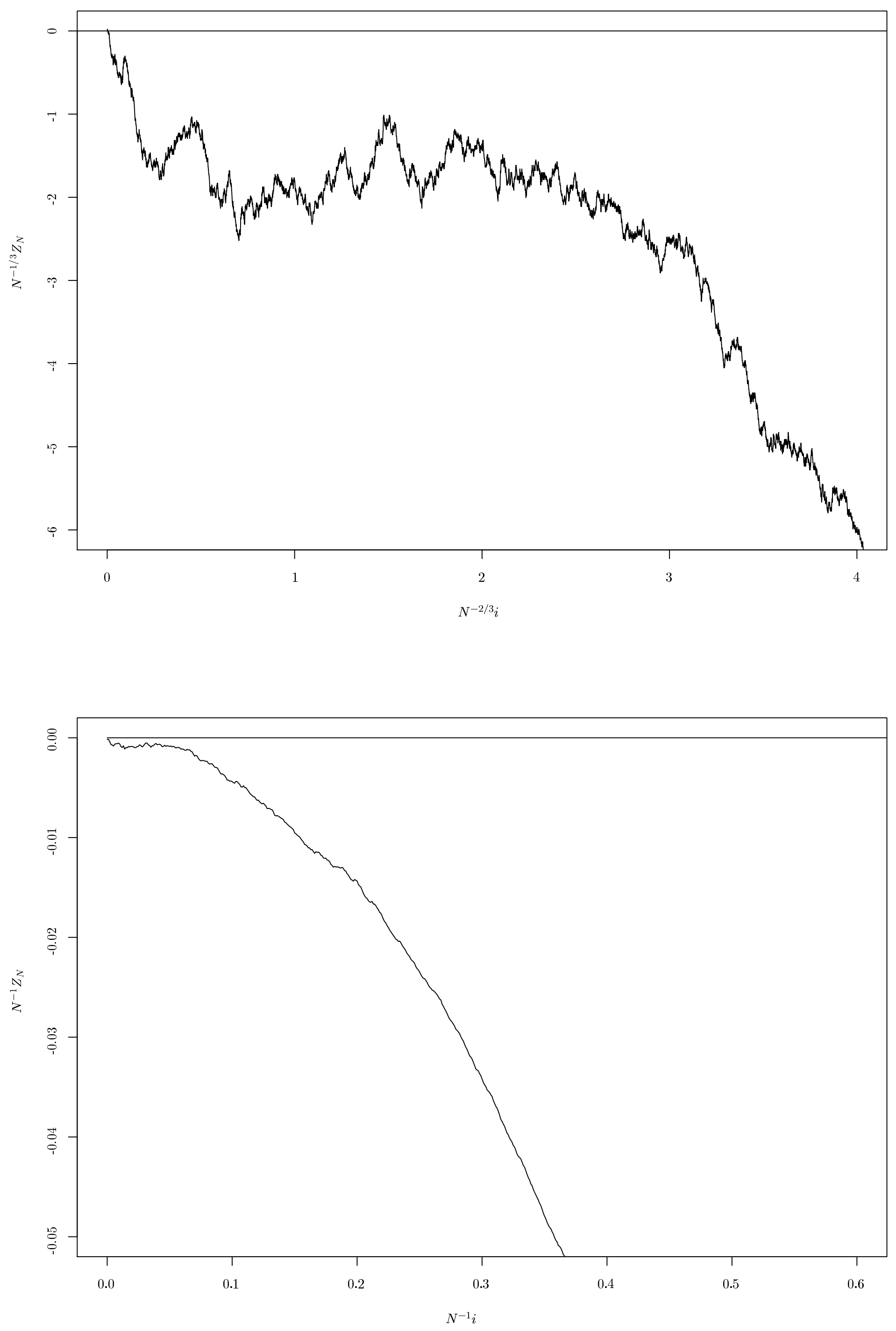


Figure 6.5: Case $\beta_3 = 1/20 < 1/6$. Top: the Brownian limit, with negative drift. Bottom: the macroscopic scale.

## 6.5 The giant set of identifiable vertices

For ease of notation, define $\mu_k = k(k-1)\beta_k - 1$. We will now fix $k \geq 3$ and look at the case $\mu_k > 0, \mu_j = 0$ for $2 \leq j \leq k-1$ in more detail. First, we state a proposition which will be of use to us later:

**Proposition 6.2.** *For $W^k$ defined as in the previous section,*

$$\mathbb{P}\left(W^k(R^2) > R\right) \to 1$$

*as $R \to \infty$.*

*Proof.* We have

$$\begin{aligned}\mathbb{P}\left(W^k(R^2) > R\right) &= \mathbb{P}\left(B(R^2) + \frac{\mu_k}{k-1}R^{2(k-1)} > R\right) \\ &= 1 - \Phi\left(1 - \frac{\mu_k}{k-1}R^{2k-3}\right) \\ &\to 1\end{aligned}$$

as $R \to \infty$, where $\Phi$ is the standard Normal distribution function. $\square$

Thus, it is a corollary of Theorem 6.1 that the event $\{Z_N(\lceil N^{\alpha(k)}R^2\rceil) > RN^{\alpha(k)/2}\}$ has asymptotically high probability in the sense that

$$\lim_{R\to\infty}\lim_{N\to\infty}\mathbb{P}\left(Z_N(\lceil N^{\alpha(k)}R^2\rceil) > RN^{\alpha(k)/2}\right) = 1. \tag{6.2}$$

Recall that $P_N(i) = 1 - \min_{j\leq i} Z_N(i)$ is the number of patches added before the $(i+1)$th deletion to keep the process going. Then, by Theorem 6.1 and the Continuous Mapping Theorem (Corollary 3.1.9 of Ethier and Kurtz [16]), $N^{-\alpha(k)/2}P_N(\lfloor N^{\alpha(k)}t\rfloor) \xrightarrow{d} -\inf_{0\leq s\leq t} W^k(s)$. Because of the positive drift of $W^k$, we have

$$\mathbb{P}\left(\inf_{s\geq 0} W^k(s) < -R\right) \to 0$$

as $R \to \infty$ and so it is clear that

$$\lim_{R\to\infty}\lim_{N\to\infty}\mathbb{P}\left(P_N(\lceil N^{\alpha(k)}t\rceil) < RN^{\alpha(k)/2}\right) = 1 \tag{6.3}$$

for any value of $t$.

Define

$$S_N^R = \inf\{i \geq N^{\alpha(k)}R^2 : Z_N(i) \leq 0\}.$$

**Theorem 6.3.** *For some $\delta > 0$,*

$$\lim_{R\to\infty} \lim_{N\to\infty} \mathbb{P}\left(S_N^R \leq N\delta\right) = 0.$$

Essentially, by time $\lceil N^{\alpha(k)}R^2 \rceil$, the process $Z_N$ is, with high probability, in an excursion above its last minimum of length $\Theta(N)$.

Define

$$z(t) = 1 - t - \exp(-\beta'(t)) \tag{6.4}$$

and recall that $t^* = \inf\{t \geq 0 : z(t) < 0\} \wedge 1$. Assume that there are no zeros of the function $z(t)$ in $(0, t^*)$ and that $t^* < 1$. Let $0 < \tilde{t} < t^*$ and let $\tilde{Z}_N$ be a modified version of $Z_N$ such that no more patches are added after time $\lfloor N\tilde{t} \rfloor$. Thus, $\tilde{P}_N(\lfloor Nt \rfloor) = P_N(\lfloor N\tilde{t} \rfloor)$ for all $t \geq \tilde{t}$ and the first time that $\tilde{Z}_N$ goes below its past-minimum after time $\lfloor N\tilde{t} \rfloor$, it stops evolving and stays constant. Let $\tilde{z}(t) = z(t \wedge t^*)$. Theorem 6.3 allows us to prove a fluid limit for this modified version of $Z_N$.

**Theorem 6.4.** *For all $\epsilon > 0$,*

$$\lim_{N\to\infty} \mathbb{P}\left(\sup_{0\leq t\leq 1} \left|\frac{1}{N}\tilde{Z}_N(\lfloor Nt \rfloor) - \tilde{z}(t)\right| > \epsilon\right) = 0.$$

This implies that for any $0 < \sigma < t^*$,

$$\sup_{0\leq t\leq \sigma} \left|\frac{1}{N}Z_N(\lfloor Nt \rfloor) - z(t)\right| \xrightarrow{p} 0.$$

See Figure 6.3 for an illustration of the fluid limit path. Note that the fluid limit visibly ceases to hold after $t^*$.

In addition to Theorem 6.4, we have a Functional Central Limit Theorem, which describes the fluctuations of the breadth-first walk around its fluid limit.

**Theorem 6.5.** *For any $0 < \sigma < t^*$,*

$$\frac{1}{\sqrt{N}}\left(Z_N(Nt) - Nz(t)\right)_{0\leq t\leq\sigma} \xrightarrow{d} (X_t)_{0\leq t\leq\sigma},$$

*where*

$$X_t = \exp(-\beta'(t)) \int_0^t \exp(\beta'(s))dG_s$$

*and $(G_t)_{t\geq 0}$ is a Gaussian process such that if $(B_t)_{t\geq 0}$ is a standard Brownian motion then $(G_t)_{t\geq 0} \sim (B_{z(t)+t})_{t\geq 0}$.*

*Remark:* This result is consistent with Theorem 6.1; the scaling there gives a zoomed-in version.

## 6.6 Proofs

### Proof of Theorem 6.1

Our method of proof follows that of Theorem 3 of Aldous [1].

**Lemma 6.6.** *Let*

$$\lambda_2(N,i) = N\sum_{j=0}^{i}\beta_{j+2}\binom{i}{j}\Big/\binom{N}{j+2}$$

*and* $\rho(N,i) = 1-\exp(-\lambda_2(N,i))$. *Then, given* $Z_N(i-1)$ *and* $P_N(i-1)$, *the distribution of* $C_N(i)$ *is*

$$\mathrm{Bin}(N-(i-1)-Z_N(i-1)-P_N(i-1), \rho(N,i-1)).$$

*Proof.* Recall from Section 1.6 that for a deterministic set $S$ of the vertices with $|S| = i$, we have that the random hypergraph, $\Lambda'$, obtained by deleting $S$ and collapsing any hyperedges with vertices in $S$ down onto their remaining vertices, is a new Poisson random hypergraph. Furthermore, for $A \subseteq V\backslash S$ with $|A| = k$ we have that the Poisson distribution of $\Lambda'(A)$ has parameter

$$\lambda_k(N,i) = N\sum_{j=0}^{i}\beta_{k+j}\binom{i}{j}\Big/\binom{N}{k+j}.$$

When we perform the breadth-first walk on the hypergraph we do not delete a deterministic set of vertices. We aim to show, nonetheless, that when we have deleted $i$ vertices, the probability that the number of 2-edges over the pair $\{v(i+1), w\}$, for any vertex $w$ we have not yet looked at, is Poisson$(\lambda_2(N,i))$, independently for all such $w$. Start with a Poisson random hypergraph and let $\mathcal{L}$ be the set consisting of a list of the vertices. Let $\mathcal{R}$ be the set of vertices we have reached (empty to start with). Perform the following version of the breadth-first numbering algorithm:

1. Remove a vertex at random from $\mathcal{L}$, call it $v(1)$ and add it to $\mathcal{R}$.
2. Examine the number of hyperedges over the sets $\{v(1), w\}$, $w \in \mathcal{L}$. The children of $v(1)$ are those $w$ such that this number of hyperedges is 1 or greater.

3. Retaining the original ordering of the vertex-labels, call the children $v(2), \dots, v(C_N(1)+1)$ where $C_N(1)$ is the number of children of $v(1)$. Remove $v(2), \dots, v(C_N(1)+1)$ from the list $\mathcal{L}$ and add them to $\mathcal{R}$.

4. Suppose we have already found the children of vertices $v(1), \dots, v(i-1)$. If there is a vertex $v(i)$ in the set $\mathcal{R}$, go to step 5. Otherwise, take a randomly-chosen vertex from $\mathcal{L}$, call that $v(i)$, add it to $\mathcal{R}$ and go to step 5.

5. The children of $v(i)$ are those $w \in \mathcal{L}$ such that the number of original hyperedges over the set $\{v(i), w\} \cup A$ is 1 or greater for at least one set $A \subseteq \{v(1), \dots, v(i-1)\}$. (In our original version of the breadth-first ordering, $A$ would have been collapsed by now and so $w$ really is a child.)

6. Re-name the children as before, remove them from $\mathcal{L}$ and add them to $\mathcal{R}$. Increment $i$ and repeat from step 4.

Observe that, before we find the children of $v(i)$, we do not look at the sets $\{v(i), w\} \cap A$ where $w \in \mathcal{L}$ and $A \subseteq \{v(1), \dots, v(i-1)\}$. Thus, in order to find the children, we need only consider random variables which are independent of what we have seen so far. So, if we imagine deleting $v(1), \dots, v(i-1)$, the 2-edge parameter in the remaining hypergraph is, indeed, $\lambda_2(N, i-1)$. Thus, any particular 2-edge is present with probability $\rho(N, i-1) = 1 - \exp(-\lambda_2(N, i-1))$.

Finally, we need to find the number of vertices eligible to be children of $v(i)$:

$$\begin{aligned}
\#\{\text{vertices which cannot be a child of } v(i)\} &= |\mathcal{R}| \\
&= C_N(1) + \dots + C_N(i-1) + P_N(i-1) \\
&= Z_N(i-1) + (i-1) + P_N(i-1).
\end{aligned}$$

and so $C_N(i) \sim \mathrm{Bin}(N - (i-1) - Z_N(i-1) - P_N(i-1), \rho(N, i-1))$. □

In the statement of Theorem 6.1, we have used the floor function to interpolate between integer-valued time-points. Here, we will prove that the process $(Z_N(i))_{0 \le i \le N}$ converges with a different interpolation but this will be equivalent to the theorem as stated. Let $(E_{i,j} : 1 \le i \le N,\ 1 \le j \le N - (i-1) - Z_N(i-1) - P_N(i-1))$ be a family of independent $\mathrm{Exp}(\lambda_2(N, i-1))$ random variables and set

$$Z_N(i-1+u) = Z_N(i-1) - u + \sum_{j=1}^{\substack{N-(i-1)- \\ Z_N(i-1)-P_N(i-1)}} \mathbb{1}_{\{E_{i,j} \le u\}}, \quad 0 \le u \le 1.$$

Take the filtration $\mathcal{F}_t^N = \sigma(Z_N(u) : u \leq t)$. This filtration is spatial in that it tells us what we can 'see' on the hypergraph at time $t$. Imagine that the vertex $v(i)$ is deleted at time $i-1$ (recall that $Z_N(i)$ is the number of patches on the hypergraph *after* the deletion of $v(i)$, adjusted for patches we have introduced). Imagine that the patches on the children of $i$ appear one by one in the interval $(i-1, 1]$. There are $N-(i-1)-Z_N(i-1)-P_N(i-1)$ possible children of $v(i)$ (i.e. vertices we have not yet reached) and each of them is actually a child with probability $\rho(N, i-1) = 1-\exp(-\lambda_2(N, i-1))$. Imagine each of these potential children having an $\mathrm{Exp}(\lambda_2(N, i-1))$ random variable associated with it; then the ones which actually are children are those whose exponential random variable is less than 1.

Define new processes by the standard decompositions

$$Z_N = M_N + A_N \tag{6.5}$$

$$M_N^2 = R_N + Q_N \tag{6.6}$$

where $M_N(0) = A_N(0) = R_N(0) = Q_N(0) = 0$, $M_N$ and $R_N$ are martingales, $A_N$ is a continuous adapted bounded variation process and $Q_N$ is a continuous adapted increasing process.

We will show that, for fixed $t_0$,

$$\frac{1}{N^{\alpha(k)/2}} \sup_{t \leq N^{\alpha(k)} t_0} \left| A_N(t) - \frac{\mu_k}{(k-1)} \frac{t^{k-1}}{N^{k-2}} \right| \xrightarrow{p} 0 \tag{6.7}$$

$$\frac{1}{N^{\alpha(k)}} Q_N(N^{\alpha(k)} t_0) \xrightarrow{p} t_0 \tag{6.8}$$

$$\frac{1}{N^{\alpha(k)}} \mathbb{E}\left[ \sup_{t \leq N^{\alpha(k)} t_0} |M_N(t) - M_N(t-)|^2 \right] \to 0. \tag{6.9}$$

Equivalently, on rescaling, we will show that

$$\sup_{t \leq t_0} \left| \bar{A}_N(t) - \frac{\mu_k}{k-1} t^{k-1} \right| \xrightarrow{p} 0$$

$$\bar{Q}_N(t_0) \xrightarrow{p} t_0$$

$$\mathbb{E}\left[ \sup_{t \leq t_0} |\bar{M}_N(t) - \bar{M}_N(t-)|^2 \right] \to 0$$

where $\bar{A}_N(t) = N^{-\alpha(k)/2} A_N(t)$, $\bar{M}_N(t) = N^{-\alpha(k)/2} M_N(t)$ and $\bar{Q}_N(t) = N^{-\alpha(k)} Q_N(t)$. By the Martingale Central Limit Theorem (Theorem 7.1.4(b) of Ethier and Kurtz [16]), Conditions (6.8) and (6.9) are sufficient to prove that

$$\bar{M}_N \xrightarrow{d} B,$$

where $B$ is a standard Brownian motion. In conjunction with (6.7), this implies that

$$(\bar{Z}_N(t))_{t\geq 0} \xrightarrow{d} \left(B(t) + \frac{\mu_k}{k-1}t^{k-1}\right)_{t\geq 0}.$$

Now, since $A_N$ is continuous, the jumps of $M_N$ are those of $Z_N$, which are of size 1 by construction and so (6.9) is obvious. So we just need to find the explicit forms of the martingale decompositions.

Define $P_N(t) = P_N(\lfloor t \rfloor) = 1 - \inf_{s\leq \lfloor t \rfloor} Z_N(s)$, a continuous-time version of number of patches added to keep the process going. Let

$$a_N(t)dt = \mathbb{P}\left(\text{some new child appears during } [t, t+dt] \mid \mathcal{F}_t^N\right).$$

Then, as $Z_N$ has drift of rate $-1$ and jumps of $+1$,

$$A_N(t) = \int_0^t (a_N(s) - 1)ds \tag{6.10}$$

$$Q_N(t) = \int_0^t a_N(s)ds. \tag{6.11}$$

(Heuristically, this is because

$$\begin{aligned} A_N(t+dt) - A_N(t) &= \mathbb{E}\left[Z_N(t+dt) - Z_N(t)|\mathcal{F}_t^N\right] \\ &= -dt + \mathbb{P}\left(\text{some new child appears during } [t, t+dt] \mid \mathcal{F}_t^N\right) \\ Q_N(t+dt) - Q_N(t) &= \mathbb{E}\left[(Z_N(t+dt) - Z_N(t))^2|\mathcal{F}_t^N\right] \\ &= \mathbb{P}\left(\text{some new child appears during } [t, t+dt] \mid \mathcal{F}_t^N\right).) \end{aligned}$$

Thus,

$$Q_N(t) = A_N(t) + t. \tag{6.12}$$

Define $\lambda_2(N,t) = \lambda_2(N, \lfloor t \rfloor)$. So, for $0 \leq u \leq 1$,

$$\begin{aligned} &\#\{\text{vertices at time } i-1+u \text{ which cannot be a child of } v(i)\} \\ &= Z_N(i-1) + (i-1) + P_N(i-1) + \sum_{j=1}^{\substack{N-(i-1)- \\ Z_N(i-1)-P_N(i-1)}} \mathbb{1}_{\{E_{i,j}\leq u\}} \\ &= Z_N(i-1+u) + (i-1+u) + P_N(i-1+u) \end{aligned}$$

and so, at time $t$, the time until the next child appears is the minimum of $N-t-Z_N(t)-P_N(t)$ independent $\text{Exp}(\lambda_2(N,t))$ random variables, and so is exponential with parameter $(N-t-Z_N(t)-P_N(t))\lambda_2(N,t)$. Hence,

$$a_N(t) = (N-t-Z_N(t)-P_N(t))\lambda_2(N,t).$$

Then we have the following lemma:

**Lemma 6.7.** *Fix $0 \le r < 1$. Then for $N$ sufficiently large and all $0 \le t \le rN$,*

$$\left| N\lambda_2(N,t) - \left( \sum_{i=0}^{k-3} (t/N)^i + k(k-1)\beta_k (t/N)^{k-2} \right) \right| = \mathcal{O}\left( (\log N)^2/N + (t/N)^{k-1} \right) \tag{6.13}$$

$$\left| a_N(t) - 1 - \mu_k (t/N)^{k-2} \right| \le \beta(r) \frac{|Z_N(t)| + P_N(t)}{N} + \mathcal{O}\left( (\log N)^2/N + (t/N)^{k-1} \right). \tag{6.14}$$

*Proof.* Lemma 1.5 says that there exists a constant $C < \infty$ such that

$$|N\lambda_2(N,t) - \beta''(t/N)| \le C(\log N)^2/N \tag{6.15}$$

for all $N \in \mathbb{N}$ and all $t \in \mathbb{Z} \cap [0, rN]$. Also, for $0 \le t \le rN$,

$$\begin{aligned} \left| \beta''(t/N) - \left( \sum_{i=0}^{k-3} (t/N)^i + k(k-1)\beta_k (t/N)^{k-2} \right) \right| &\le (t/N)^{k-1} \sup_{0 \le s < r} \beta^{(k+1)}(s) \\ &= (t/N)^{k-1} \beta^{(k+1)}(r) \end{aligned}$$

by Taylor's theorem and (6.13) follows. Furthermore, using (6.13),

$$\begin{aligned} a_N(t) &= \left( 1 - \frac{t}{N} - \frac{Z_N(t) + P_N(t)}{N} \right) N\lambda_2(N,t) \\ &= \left( 1 - \frac{t}{N} - \frac{Z_N(t) + P_N(t)}{N} \right) \left( \sum_{i=0}^{k-3} (t/N)^i + k(k-1)\beta_k (t/N)^{k-2} \right) \\ &\quad + \mathcal{O}((\log N)^2/N + (t/N)^{k-1}) \end{aligned}$$

and so

$$\left| a_N(t) - 1 - \mu_k (t/N)^{k-2} \right| \le \beta(r) \frac{|Z_N(t)| + P_N(t)}{N} + \mathcal{O}((\log N)^2/N + (t/N)^{k-1})$$

for sufficiently large $N$ and (6.14) holds. □

Thus, on integrating $a_N(t) - 1 - \mu_k(t/N)^{k-2}$ and using the fact that $P_N(t) = 1 - \inf_{s \le \lfloor t \rfloor} Z_N(s)$, we obtain that for some constant $C < \infty$,

$$\left| A_N(t) - \frac{\mu_k}{(k-1)} \frac{t^{k-1}}{N^{k-2}} \right| \le \frac{2Ct \max_{s \le t} |Z_N(s)|}{N} + \mathcal{O}\left( \frac{t(\log N)^2}{N} + \frac{t^k}{N^{k-1}} \right).$$

We wish to prove that

$$\frac{1}{N^{\alpha(k)/2}} \sup_{t \le N^{\alpha(k)} t_0} \left| A_N(t) - \frac{\mu_k}{(k-1)} \frac{t^{k-1}}{N^{k-2}} \right| \xrightarrow{p} 0$$

and so it will be sufficient to prove that

$$\frac{1}{N^{1-\alpha(k)/2}} \sup_{t \le N^{\alpha(k)} t_0} |Z_N(t)| \xrightarrow{p} 0$$

or the stronger statement that $\frac{1}{N^{\alpha(k)/2}} \sup_{t \le N^{\alpha(k)} t_0} |Z_N(t)|$ is stochastically bounded as $N \to \infty$.

Fix a large constant $K$ and let $T_N = \inf\{t \ge 0 : |Z_N(t)| > K N^{\alpha(k)/2}\} \wedge (N^{\alpha(k)} t_0)$. Then

$$\begin{aligned}
\mathbb{E}|Z_N(T_N)| &\le \mathbb{E}|M_N(T_N)| + \mathbb{E}|A_N(T_N)| \\
&\le \sqrt{\mathbb{E}\left[M_N^2(T_N)\right]} + \mathbb{E}|A_N(T_N)| \\
&= \sqrt{\mathbb{E}[Q_N(T_N)]} + \mathbb{E}|A_N(T_N)| \\
&= \sqrt{\mathbb{E}[A_N(T_N) + T_N]} + \mathbb{E}|A_N(T_N)| \\
&\le \sqrt{\mathbb{E}|A_N(T_N)| + N^{\alpha(k)} t_0} + \mathbb{E}|A_N(T_N)|.
\end{aligned}$$

where the equality on the third line is by the Optional Stopping Theorem applied to (6.6) and the fourth line is from (6.12). Now, there exists a constant $C < \infty$ such that

$$\begin{aligned}
\mathbb{E}|A_N(T_N)| &\le \mathbb{E}\left[ \int_0^{T_N} |a_N(t) - 1| dt \right] \\
&\le \mathbb{E}\left[ \int_0^{N^{\alpha(k)} t_0} \mu_k \left( \frac{t}{N} \right)^{k-2} dt \right] + \mathbb{E}\left[ \int_0^{T_N} C \frac{P_N(t) + |Z_N(t)|}{N} dt \right] + \mathcal{O}\left( N^{\frac{k-3}{2k-3}} \right) \\
&\le \mathbb{E}\left[ \frac{2CT_N}{N} \sup_{t < T_N} |Z_N(t)| \right] + \mathcal{O}\left( N^{\alpha(k)/2} \right) \\
&\le \frac{2CN^{\alpha(k)} t_0 K N^{\alpha(k)/2}}{N} + \mathcal{O}\left( N^{\alpha(k)/2} \right) \\
&= 2Ct_0 K N^{\frac{k-3}{2k-3}} + \mathcal{O}(N^{\alpha(k)/2}).
\end{aligned}$$

Hence, $\mathbb{E}|Z_N(T_N)| \leq C'KN^{\frac{k-3}{2k-3}} + C'N^{\alpha(k)/2}$ for some constant $C'$ and so, by Markov's inequality,

$$\begin{aligned}\mathbb{P}\left(N^{-\alpha(k)/2} \sup_{t\leq N^{\alpha(k)}t_0} |Z_N(t)| > K\right) &= \mathbb{P}\left(|Z_N(T_N)| > KN^{\alpha(k)/2}\right)\\ &\leq \frac{\mathbb{E}|Z_N(T_N)|}{KN^{\alpha(k)/2}}\\ &\leq \frac{C'}{K} + C'N^{-\frac{1}{2k-3}}\end{aligned}$$

which gives the required stochastic boundedness as $N \to \infty$.

Finally, we need to show that

$$\frac{1}{N^{\alpha(k)}} Q_N(N^{\alpha(k)}t_0) \xrightarrow{p} t_0.$$

But since $Q_N(t) = A_N(t) + t$, this follows immediately from

$$\frac{1}{N^{\alpha(k)/2}} \sup_{t\leq N^{\alpha(k)}t_0} \left| A_N(t) - \frac{\mu_k}{(k-1)} \frac{t^{k-1}}{N^{k-2}} \right| \xrightarrow{p} 0.$$

## Proof of Theorem 6.3

Return now to the discrete-time setting of Theorem 6.3. The proof is quite involved and so we will begin by outlining our method. The random walk $Z_N$, started from above $RN^{\alpha(k)/2}$ at time $R^2N^{\alpha(k)}$, is close to dominating the breadth-first walk associated with the tree produced by a Galton–Watson branching process with Poisson offspring distribution of mean $1 + \frac{1}{2}\mu_k N^{-\alpha(k)/2} R^{2(k-2)}$. This branching process is (just) supercritical and so the associated breadth-first walk has both positive drift and the virtue of identically distributed step sizes. We will essentially work with that process instead, using it as a lower bound for our original breadth-first walk. If, with high probability, the lower bound goes for $N\delta$ steps without hitting 0 then so does our original breadth-first walk. In order to show that the smaller random walk is unlikely to come back to 0, we use an exponential supermartingale argument and the Optional Stopping Theorem.

Fix $R > 0$ and work on the set

$$\Omega_{N,R} = \{Z_N(\lceil N^{\alpha(k)}R^2\rceil) > RN^{\alpha(k)/2}, P_N(\lceil N^{\alpha(k)}R^2\rceil) < RN^{\alpha(k)/2}\}.$$

As we have already seen at (6.2) and (6.3), this set has asymptotically high probability. It will be useful to keep track of the times when $(Z_N(i))_{\lceil N^{\alpha(k)}R^2\rceil \leq i \leq N}$ goes above and below

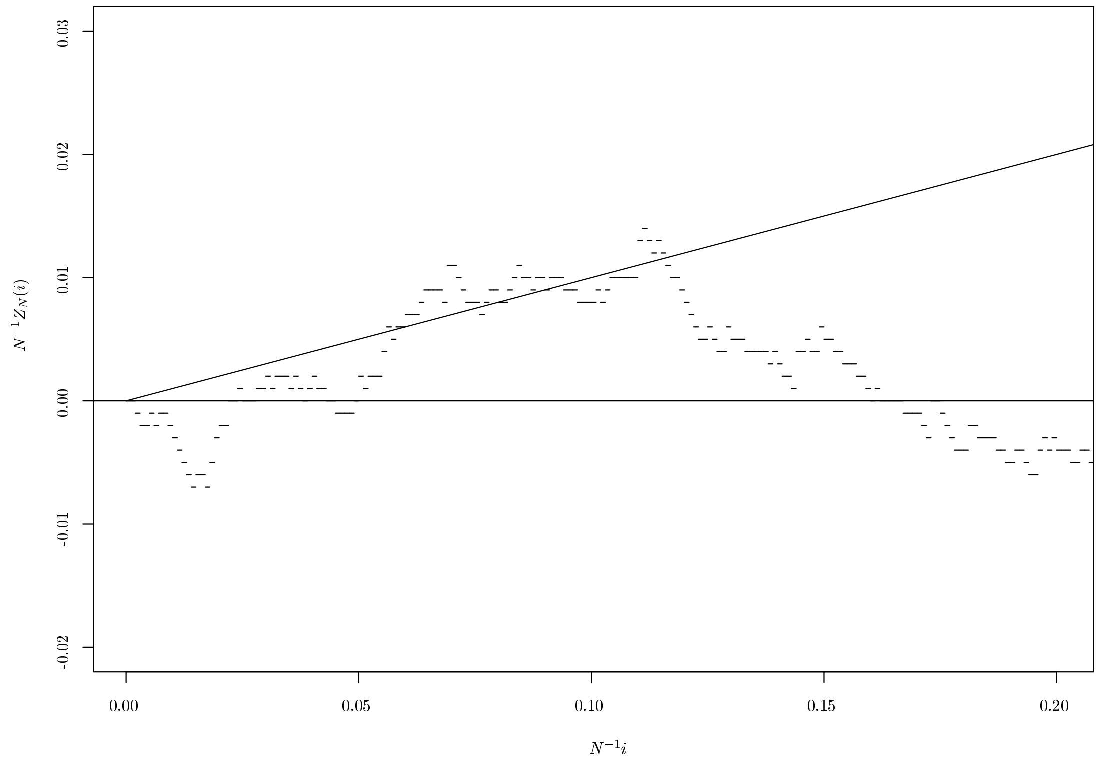


Figure 6.6: Simulation of the breadth-first walk with $N = 1000$, $\beta_2 = 0.5$, $\beta_3 = 0.2$ and $\beta_j = 0$ for $j \geq 4$ (i.e. $k = 3$). Also shown is the line $f(i) = \frac{1}{2}\mu_3 i$.

the line $f(i) = \frac{1}{2}\mu_k i^{k-2}N^{-k+3}$, as in Figure 6.6. To this end, introduce two sequences of stopping times, $(\tau_n)_{n\geq 1}$ and $(\tau'_n)_{n\geq 1}$, such that

$$\tau_1 = \inf\left\{i \geq \lceil N^{\alpha(k)}R^2\rceil : Z_N(i) \geq f(i)\right\}$$

and, for $n \geq 1$,

$$\begin{aligned}\tau'_n &= \inf\left\{i > \tau_n : Z_N(i) < f(i)\right\}\\ \tau_{n+1} &= \inf\left\{i > \tau'_n : Z_N(i) \geq f(i)\right\}.\end{aligned}$$

Let $\Omega_N(i) = \{\lceil N^{\alpha(k)}R^2\rceil \leq i < \tau_1\} \cup \left(\bigcup_{n\geq 1}\{\tau'_n \leq i < \tau_{n+1}\}\right)$, the event that $Z_N(i)$ is below the line. Define $\mathcal{F}_i^N = \sigma(Z_N(j) : j \leq i)$.

**Lemma 6.8.** *For sufficiently large $N$, there exists a $\theta < 0$ such that $L_N(i) = \exp(\theta Z_N(i))$ is a supermartingale on the disjoint sets of times*

$$\{\lceil N^{\alpha(k)}R^2\rceil, \ldots, \tau_1\}, \{\tau'_1, \ldots, \tau_2\}, \{\tau'_2, \ldots, \tau_3\}, \ldots$$

*That is, $(L_N(i))_{0\leq i\leq N}$ is an integrable adapted process and*

$$\mathbb{E}\left[L_N(i)|\mathcal{F}_{i-1}^N\right] \leq L_N(i-1)$$

*whenever $i - 1 \in \{\lceil N^{\alpha(k)}R^2\rceil, \ldots, \tau_1 - 1\} \cup \left(\bigcup_{n\geq 1}\{\tau'_n, \ldots, \tau_{n+1} - 1\}\right)$.*

*Proof.* Consider the conditional moment generating function of an increment of $Z_N$: by Lemma 6.6,

$$\begin{aligned}\phi_i(N,\theta) &= \mathbb{E}\left[\exp(\theta(Z_N(i) - Z_N(i-1)))|\mathcal{F}_{i-1}^N\right]\\ &= \exp\{(N - (i-1) - Z_N(i-1) - P_N(i-1))\log(1 + \rho(N, i-1)(e^\theta - 1)) - \theta\}\\ &\leq \exp\{(N - (i-1) - Z_N(i-1) - P_N(i-1))\rho(N, i-1)(e^\theta - 1) - \theta\}.\end{aligned}$$

By Lemma 1.5, there exists a constant $C < \infty$ such that

$$N\rho(N, i-1) \geq \beta''(\tfrac{i}{N}) - C(\log N)^2/N.$$

On the set $\Omega_{N,R} \cap \Omega_N(i-1)$, we have $P_N(i-1) < RN^{\alpha(k)/2}$ and so, for sufficiently large $N$,

$$\begin{aligned}&(N - (i-1) - Z_N(i-1) - P_N(i-1))\rho(N, i-1)\\ &\quad\geq \left(1 - \tfrac{i}{N} - \tfrac{1}{2}\mu_k(\tfrac{i}{N})^{k-2}\right)\beta''(\tfrac{i}{N}) - KN^{-\frac{k-1}{2k-3}}\end{aligned}$$

for some constant $K < \infty$ depending on $R$. The first non-zero derivative of $(1 - s - \frac{1}{2}\mu_k s^{k-2})\beta''(s)$ at 0 is the $(k-2)$th which is $\frac{1}{2}(k-2)!\mu_k$. Thus, there exists a $\delta > 0$ such that $(1-s-\frac{1}{2}\mu_k s^{k-2})\beta''(s)$ is increasing on $[0, \delta]$. Therefore, for $i \leq \delta N$ on $\Omega_{N,R} \cap \Omega_N(i-1)$, we have

$$\begin{aligned}&\left(1 - \tfrac{i}{N} - \tfrac{1}{2}\mu_k (\tfrac{i}{N})^{k-2}\right) \beta''(\tfrac{i}{N}) \\ &\geq \left(1 - N^{\alpha(k)-1}R^2 - \tfrac{1}{2}\mu_k N^{(\alpha(k)-1)(k-2)}R^{2(k-2)}\right) \beta''(N^{\alpha(k)-1}R^2)\end{aligned}$$

by putting in $i = N^{\alpha(k)}R^2$. Now, expanding $\beta''(s)$ as $1 + s + \cdots + s^{k-3} + k(k-1)\beta_k s^{k-2}$ and incurring an error of size $\mathcal{O}(s^{k-1})$ we see that the right-hand side is bounded below by

$$1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)} - K'N^{-\frac{k-1}{2k-3}}.$$

for some constant $K' < \infty$. Thus,

$$\phi_i(N, \theta) \leq \exp\left\{\left(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}\right)(e^\theta - 1) - \theta + K'N^{-\frac{k-1}{2k-3}}(1 - e^\theta)\right\}.$$

If we had the breadth-first walk on the tree produced by a branching process with Poisson offspring distribution of mean $1 + \frac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}$, we would have

$$\exp\left\{\left(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}\right)(e^\theta - 1) - \theta\right\}$$

as the conditional moment generating function of an increment. Thus, we will effectively use this simpler process as a 'lower bound' for our original process.

Now, let $\bar{\theta}$ be the value of $\theta$ which minimises

$$\exp\left\{\left(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}\right)(e^\theta - 1) - \theta\right\},$$

so that it is easily seen that

$$\bar{\theta} = -\log(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)})$$

(and so, trivially, $\bar{\theta} < 0$). For $i \leq N\delta$ on $\Omega_{N,R} \cap \Omega_N(i-1)$,

$$\begin{aligned}\phi_i(N, \bar{\theta}) &\leq \exp\left\{\left(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}\right)(e^{\bar{\theta}} - 1) - \bar{\theta} + K'N^{-\frac{k-1}{2k-3}}(1 - e^{\bar{\theta}})\right\} \\ &= \exp\left\{\log(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}) - \tfrac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}\right. \\ &\qquad \left. + \frac{\frac{1}{2}\mu_k K'N^{-1}R^{2(k-2)}}{1 + \frac{1}{2}\mu_k N^{-\alpha(k)/2}R^{2(k-2)}}\right\} \\ &\leq \exp\{-C_2 N^{-\alpha(k)} + C_1 N^{-1}\} \\ &\leq 1,\end{aligned}$$

for some constants $C_1, C_2$ and sufficiently large $N$. Hence, for $\theta = \bar{\theta}$ and sufficiently large $N$, $\mathbb{E}\left[L_N(i)|\mathcal{F}^N_{i-1}\right] \leq L_N(i-1)$ on the set $\Omega_{N,R} \cap \Omega_N(i-1)$. $\square$

**Lemma 6.9.** *We have*

$$\lim_{N\to\infty} \mathbb{P}\left(S^R_N \leq N\delta | \Omega_{N,R}\right) \leq \exp\left(-\tfrac{1}{2}\mu_k R^{2k-1}\right).$$

*Proof.* $Z_N$ cannot hit 0 when it is above the line $f(i) = \frac{1}{2}\mu_k i^{k-2} N^{-k+3}$. Thus, if $Z_N$ does hit 0 before time $N\delta$, it must occur before time $\tau_1$ or between times $\tau'_i$ and $\tau_{i+1}$ for some $i \geq 1$ and so

$$\begin{aligned}
&\mathbb{P}\left(S^R_N \leq N\delta \;\middle|\; \Omega_{N,R}\right) \\
&\quad \leq \mathbb{P}\left(S^R_N \leq \tau_1 \wedge (N\delta) \;\middle|\; \Omega_{N,R}\right) \\
&\qquad + \sum_{i=1}^{\infty} \mathbb{P}\left(S^R_N \leq \tau_{i+1} \wedge (N\delta) \;\middle|\; \Omega_{N,R}, \tau_i < S^R_N \wedge (N\delta), \tau'_i < N\delta\right) \\
&\quad \leq \mathbb{P}\left(S^R_N \leq \tau_1 \wedge (N\delta) \;\middle|\; \Omega_{N,R}\right) \\
&\qquad + \sum_{i=\lceil N^{\alpha(k)} R^2 \rceil}^{\infty} \mathbb{P}\left(S^R_N \leq T_i \wedge (N\delta) \middle| \; \Omega_{N,R}, Z_N(i-1) \geq f(i-1), Z_N(i) < f(i)\right)
\end{aligned}$$

where $T_i = \inf\{j \geq i : Z_N(j) > f(j)\}$ and each term in the above summation expresses the probability of a downcrossing from the line $f(i)$ to 0. It will turn out that only the term $\mathbb{P}\left(S^R_N \leq \tau_1 \wedge (N\delta) \mid \Omega_{N,R}\right)$ makes a significant contribution (intuitively because the process is much closer to 0 at time $\lceil N^{\alpha(k)} R^2 \rceil$ than it is at $\tau'_i$ for any $i \geq 1$).

As $S^R_N \wedge \tau_1 \wedge (N\delta)$ is a bounded stopping time, by the Optional Stopping Theorem we obtain that

$$\begin{aligned}
L_N(\lceil N^{\alpha(k)} R^2 \rceil) \mathbb{1}_{\Omega_{N,R}} &\geq \mathbb{E}\left[L_N(S^R_N \wedge \tau_1 \wedge (N\delta)) \mathbb{1}_{\Omega_{N,R}} \middle| \; \mathcal{F}^N_{\lceil N^{\alpha(k)} R^2 \rceil}\right] \\
&\geq \mathbb{P}\left(S^R_N \leq \tau_1 \wedge (N\delta) \middle| \; \mathcal{F}^N_{\lceil N^{\alpha(k)} R^2 \rceil}\right) \mathbb{1}_{\Omega_{N,R}}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{P}\left(S^R_N \leq \tau_1 \wedge (N\delta) \middle| \; \Omega_{N,R}\right) &\leq \mathbb{E}\left[L_N(\lceil N^{\alpha(k)} R^2 \rceil) \;\middle|\; \Omega_{N,R}\right] \\
&= \mathbb{E}\left[\exp(\bar{\theta} Z_N(\lceil N^{\alpha(k)} R^2 \rceil)) \;\middle|\; \Omega_{N,R}\right] \\
&\leq \exp(\bar{\theta} R N^{\alpha(k)/2}) \\
&= \exp(-\log(1 + \tfrac{1}{2}\mu_k N^{-\alpha(k)/2} R^{2(k-2)}) R N^{\alpha(k)/2}) \\
&= \exp(-\tfrac{1}{2}\mu_k R^{2k-1} + \mathcal{O}(N^{-\alpha(k)/2})) \\
&\to \exp(-\tfrac{1}{2}\mu_k R^{2k-1})
\end{aligned}$$

as $N \to \infty$. By a similar argument,

$$\begin{aligned}
&\mathbb{P}\left(S_N^R \le T_i \wedge (N\delta) \middle|\, \Omega_{N,R}, Z_N(i-1) \ge f(i-1), Z_N(i) < f(i)\right)\\
&\qquad \le \mathbb{E}\left[L_N(i) \;\middle|\; \Omega_{N,R}, Z_N(i-1) \ge f(i-1), Z_N(i) < f(i)\right]\\
&\qquad \le \mathbb{E}\left[L_N(i) \;\middle|\; \Omega_{N,R}, Z_N(i) = \lceil f(i-1)\rceil - 1\right]\\
&\qquad \le C \exp\left(\tfrac{1}{2}\mu_k \bar{\theta} i^{k-2} N^{-k+3}\right),
\end{aligned}$$

where $C$ is a constant and the second inequality holds because $Z_N$ can step down by at most 1 and so the smallest that $Z_N(i)$ can be and still have had $Z_N(i-1)$ above the line is $\lceil f(i-1)\rceil - 1$. For $i \ge \lceil N^{\alpha(k)} R^2 \rceil$,

$$\begin{aligned}
&\exp\left(\frac{1}{2}\mu_k \bar{\theta} i^{k-2} N^{-k+3}\right)\\
&\quad \le \exp\left(\frac{1}{2}\mu_k \bar{\theta} N^{-\frac{k-3}{2k-3}} R^{2(k-3)} i\right)\\
&\quad = \exp\left(-\frac{1}{2}\mu_k N^{-\frac{k-3}{2k-3}} R^{2(k-3)} \log\left(1 + \frac{1}{2}\mu_k N^{-\alpha(k)/2} R^{2(k-3)}\right) i\right)\\
&\quad = \left[\left(1 + \frac{1}{2}\mu_k N^{-\alpha(k)/2} R^{2(k-3)}\right)^{-\frac{1}{2}\mu_k N^{-\frac{k-3}{2k-3}} R^{2(k-3)}}\right]^i.
\end{aligned}$$

Let

$$g(N,R) = \left(1 + \frac{1}{2}\mu_k N^{-\alpha(k)/2} R^{2(k-3)}\right)^{-\frac{1}{2}\mu_k N^{-\frac{k-3}{2k-3}} R^{2(k-3)}}.$$

Then

$$\begin{aligned}
&\sum_{i=\lceil N^{\alpha(k)} R^2\rceil}^{\infty} \mathbb{P}\left(S_N^R \le T_i \wedge (N\delta) \middle|\, \Omega_{N,R}, Z_N(i-1) \ge f(i-1), Z_N(i) < f(i)\right)\\
&\quad \le C \sum_{i=\lceil N^{\alpha(k)} R^2\rceil}^{\infty} g(N,R)^i\\
&\quad = \frac{C g(N,R)^{\lceil N^{\alpha(k)} R^2\rceil}}{1 - g(N,R)}.
\end{aligned}$$

This behaves essentially like $N^{\frac{2k-5}{2k-3}} \exp(-N^{\frac{1}{2k-3}})$ and so converges to 0 as $N \to \infty$. Hence,

$$\lim_{N\to\infty} \mathbb{P}\left(S_N^R \le N\delta | \Omega_{N,R}\right) \le \exp(-\tfrac{1}{2}\mu_k R^{2k-1}).$$

$\square$

Now note that

$$\mathbb{P}\left(S_N^R \le N\delta\right) \le \mathbb{P}\left(S_N^R \le N\delta \mid \Omega_{N,R}\right) + \mathbb{P}\left(\Omega_{N,R}^c\right)$$

and so

$$\lim_{N\to\infty} \mathbb{P}\left(S_N^R \le N\delta\right) \le \exp(-\tfrac{1}{2}\mu_k R^{2k-1}) + \mathbb{P}\left(W^k(R^2) \le R\right) + \mathbb{P}\left(-\inf_{s\ge 0} W^k(s) \ge R\right)$$

which converges to 0 as $R \to \infty$. Theorem 6.3 follows.

## Proof of Theorem 6.4

We will use the fluid limit methods from Section 1.7 of the Introduction. We need to be working with a pure jump Markov process. Thus, we will take $(\nu_t^N)_{t\ge 0}$ to be a a Poisson process of rate $N$ and then prove a fluid limit for the three-dimensional process

$$N^{-1}(\nu_t^N, Z_N(\nu_t^N), P_N(\nu_t^N))_{0\le t\le 1}.$$

This will imply the theorem as stated.

Let $X_t^N = (X_t^{1,N}, X_t^{2,N}, X_t^{3,N})$ where $X_t^{1,N} = N^{-1}\nu_t^N$, $X_t^{2,N} = N^{-1}Z_N(\nu_t^N)$ and $X_t^{3,N} = N^{-1}P_N(\nu_t^N)$. We need to check that Conditions (C1) to (C5) of Theorem 1.9 hold. The process $X^N$ is naturally defined on $I^N = \{x \in \mathbb{R}^3 : Nx_1 \in \mathbb{Z}^+, Nx_2 \in \mathbb{Z}, Nx_3 \in \mathbb{Z}^+\}$. Let $S = \{x \in \mathbb{R}^3 : |x_1| < r_1, 0 < x_2 < r_2, |x_3| < r_3\}$ for constants $r_1, r_2, r_3 < \infty$ and let $S^N = I^N \cap S$. Let $K^N(x, dy)$ be the Lévy kernel of $X^N$. Then, using the representation (6.1) of the evolution of $P_N$ and Lemma 6.6, $N^{-1}K^N(x,\cdot)$ is the law of

$$N^{-1}\left(1, B_N - 1, \mathbb{1}_{\{x_2+x_3=1/N,\ B_N=0\}}\right)$$

where $B_N \sim \mathrm{Bin}(N - Nx_1 - Nx_2 - Nx_3, \rho(N, Nx_1))$. Thus, $K^N(x, dy)$ has Laplace transform

$$m_1^N(x,\theta) = Ne^{\theta/N}$$

$$m_2^N(x,\theta) = N\exp((N - Nx_1 - Nx_2 - Nx_3)\log(1 + \rho(N, Nx_1)(e^{\theta/N} - 1)) - \theta/N)$$

$$m_3^N(x,\theta) = \begin{cases} Ne^{\theta/N}(1-\rho(N,Nx_1))^{N-Nx_1-1} + N(1-(1-\rho(N,Nx_1))^{N-Nx_1-1}) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } x_2 + x_3 = 1/N \\ N \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{if } x_2 + x_3 \ne 1/N. \end{cases}$$

Using Lemma 1.5, $B_N \xrightarrow{d} \text{Poisson}((1-x_1-x_2-x_3)\beta''(x_1))$ as $N \to \infty$ and so there exists a limit kernel $K(x,dy)$ with Laplace transform

$$\begin{aligned}
m_1(x,\theta) &= e^{\theta} \\
m_2(x,\theta) &= \exp((1-x_1-x_2-x_3)\beta''(x_1)(e^{\theta}-1)-\theta) \\
m_3(x,\theta) &= \begin{cases} (e^{\theta}-1)\exp(-(1-x_1)\beta''(x_1))+1 & \text{if } x_2 = -x_3 \\ 1 & \text{if } x_2 \neq -x_3. \end{cases}
\end{aligned}$$

Furthermore, there exists $\eta_0 > 0$ such that $m(x,\theta) < \infty$ for all $x \in S$ and $\|\theta\| \leq \eta_0$ and also such that

$$\sup_{x \in S^N} \sup_{\|\theta\| < \eta_0} \left| N^{-1} m^N(x, N\theta) - m(x,\theta) \right| \to 0,$$

where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^3$. Thus, Conditions (C1) and (C2) are satisfied.

Let $T_N = \inf\{t \geq 0 : X_t^N \notin S^N\}$. Fix a large $R < \infty$. We will prove the fluid limit result in three time intervals: $[0, R^2 N^{\alpha(k)-1}]$, $[R^2 N^{\alpha(k)-1}, \delta]$ and $[\delta, 1]$.

- *Time interval* $[0, R^2 N^{\alpha(k)-1}]$:

Until time $R^2 N^{\alpha(k)-1}$, $Z_N$ is in the Brownian regime of Theorem 6.1 and so

$$\sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} \left| N^{-1} Z_N(\lfloor Nt \rfloor) - z(t) \right| \leq \sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} \left| N^{-1} Z_N(Nt) \right| + \sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} |z(t)|$$

which converges to 0 in probability as $N \to \infty$, regardless of the value of $R$. Similarly,

$$\sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} \left| N^{-1} P_N(Nt) \right| \xrightarrow{p} 0.$$

It is elementary that $\sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} |N^{-1}\nu_t^N - t| \xrightarrow{p} 0$. Thus,

$$\sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} \left\| X_t^N - x(t) \right\| \xrightarrow{p} 0 \tag{6.16}$$

where $x(t) = (t, z(t), 0)$.

- *Time interval* $[R^2 N^{\alpha(k)-1}, \delta]$:

Expression (6.16) provides us with Condition (C5) of the fluid limit theorem. (Note that the convergence in (6.16) is not exponentially fast and so we will not use the full power of the methods in Section 1.7.) Suppose now that we fix $N$ and $R$ and work on the set

$$\Omega_{N,R} = \left\{ X^{2,N}_{\lceil R^2 N^{\alpha(k)-1} \rceil} > R N^{\alpha(k)/2-1}, X^{3,N}_{\lceil R^2 N^{\alpha(k)-1} \rceil} < R N^{\alpha(k)/2-1} \right\}.$$

Define $T_N^R = \inf\{t \geq R^2 N^{\alpha(k)-1} : X_t^N \notin S\}$ (which, for $r_1$, $r_2$ and $r_3$ large enough in the definition of $S$ is the time that $X^{2,N}$ first hits 0; this is a Poissonized equivalent of $S_N^R$ of Section 6.5). On $\Omega_{N,R}$, $X_{\lfloor Nt \rfloor}^{3,N}$ is constant on the interval $[R^2 N^{\alpha(k)-1}, T_N^R]$ and so

$$\sup_{0 \leq t \leq T_N^R} \left| N^{-1} X_{\lfloor Nt \rfloor}^{3,N} \right| \xrightarrow{p} 0$$

also. Now, setting $b(x) = m'(x, 0)$, we have

$$\begin{aligned}
b_1(x) &= 1 \\
b_2(x) &= (1 - x_1 - x_2 - x_3)\beta''(x_1) - 1 \\
b_3(x) &= \exp(-(1 - x_1)\beta''(x_1))\mathbb{1}_{\{x_2 = -x_3\}}.
\end{aligned}$$

By taking $r_1, r_2$ and $r_3$ large enough, the differential equation $\dot{x}_t = b(x(t))$ has unique solution $x(t) = (t, 1 - t - \exp(-\beta'(t)), 0) = (t, z(t), 0)$ in $S$. Thus Condition (C3) is satisfied and so, by Theorem 1.9, for all $\epsilon > 0$,

$$\lim_{N \to \infty} \mathbb{P}\left( \sup_{R^2 N^{\alpha(k)-1} \leq t \leq T_N^R \wedge \delta} \|X_t^N - x(t)\| > \epsilon \;\middle|\; \Omega_{N,R} \right) = 0.$$

Now, for small enough $\delta > 0$ we have that $\delta < t^*$ and $\lim_{R \to \infty} \lim_{N \to \infty} \mathbb{P}\left(S_N^R > N\delta\right) = 1$ and so we will also have $\lim_{R \to \infty} \lim_{N \to \infty} \mathbb{P}\left(T_N^R > \delta\right) = 1$. Then

$$\begin{aligned}
&\lim_{N \to \infty} \mathbb{P}\left( \sup_{R^2 N^{\alpha(k)-1} \leq t \leq \delta} \left\|X_t^N - x(t)\right\| > \epsilon \right) \\
&\quad \leq \lim_{N \to \infty} \mathbb{P}\left( \sup_{R^2 N^{\alpha(k)-1} \leq t \leq T_N^R \wedge \delta} \left\|X_t^N - x(t)\right\| > \epsilon \,, T_N^R > \delta \;\middle|\; \Omega_{N,R} \right) \\
&\qquad + \lim_{N \to \infty} \mathbb{P}\left(T_N^R \leq \delta \mid \Omega_{N,R}\right) + \lim_{N \to \infty} \mathbb{P}\left(\Omega_{N,R}^c\right)
\end{aligned}$$

where the first of the three terms on the right-hand side is 0 as we have just proved.

Thus, for all $R < \infty$ and all $\epsilon > 0$, we have that

$$\begin{aligned}
\lim_{N \to \infty} \mathbb{P}\left( \sup_{0 \leq t \leq \delta} \left\|X_t^N - x(t)\right\| > \epsilon \right) &\leq \lim_{N \to \infty} \mathbb{P}\left( \sup_{0 \leq t \leq R^2 N^{\alpha(k)-1}} \left\|X_t^N - x(t)\right\| > \epsilon \right) \\
&\quad + \lim_{N \to \infty} \mathbb{P}\left( \sup_{R^2 N^{\alpha(k)-1} \leq t \leq \delta} \left\|X_t^N - x(t)\right\| > \epsilon \right)
\end{aligned}$$

and so we can take the limit as $R \to \infty$ on the right-hand side to obtain

$$\begin{aligned}
\lim_{N \to \infty} \mathbb{P}\left( \sup_{0 \leq t \leq \delta} \left\|X_t^N - x(t)\right\| > \epsilon \right) &\leq \lim_{R \to \infty} \lim_{N \to \infty} \mathbb{P}\left(T_N^R \leq \delta \mid \Omega_{N,R}\right) + \lim_{R \to \infty} \lim_{N \to \infty} \mathbb{P}\left(\Omega_{N,R}^c\right) \\
&= 0.
\end{aligned}$$

• *Time interval* $[\delta, 1]$*:*
Suppose we now start from time $\delta$. We have just shown that the initial value converges in probability. As $z(\delta) > 0$, we may as well take $\tilde{t} = \delta$. Re-define $T_N = \inf\{t \geq \delta : X^N \notin S^N\}$ and let $\tilde{X}^N$ be such that $\tilde{X}_t^{1,N} = X_t^{1,N}$, $\tilde{X}_t^{2,N} = X_{t\wedge T_N}^{2,N}$ and $\tilde{X}_t^{3,N} = X_\delta^{3,N}$. Minor modifications to our argument in the previous section lead us to the limiting function $\tilde{x}(t) = (t, z(t \wedge t^*), 0)$ and the conclusion that for all $\epsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left( \sup_{\delta \leq t \leq 1} \left\| \tilde{X}_t^N - \tilde{x}(t) \right\| > \epsilon \right) = 0,$$

where we have that $T_N \xrightarrow{p} t^*$ by our assumption that there are no zeros of $z(t)$ in $(0, t^*)$. Thus, finally,

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}\left( \sup_{0\leq t\leq 1} \left| \tilde{X}_t^{2,N} - \tilde{z}(t) \right| > \epsilon \right) &\leq \lim_{N\to\infty} \mathbb{P}\left( \sup_{0\leq t\leq \delta} \left| X_t^{2,N} - z(t) \right| > \epsilon \right) \\ &\quad + \lim_{N\to\infty} \mathbb{P}\left( \sup_{\delta\leq t\leq 1} \left| \tilde{X}_t^{2,N} - z(t\wedge t^*) \right| > \epsilon \right) \\ &= 0.\end{aligned}$$

### Proof of Theorem 6.5

Use the interpolation from the proof of Theorem 6.1 and the Doob decompositions contained therein.

**Lemma 6.10.** *Choose* $0 < \sigma < t^*$*. Then*

$$\frac{1}{\sqrt{N}} \left( Z_N(Nt) - A_N(Nt) \right)_{0\leq t\leq \sigma} \xrightarrow{d} (G(t))_{0\leq t\leq\sigma}$$

*as* $N \to \infty$*, where* $G$ *is the time-changed Brownian motion such that if* $B$ *is a standard Brownian motion,* $(G(t))_{t\geq 0} \sim (B(t + z(t)))_{t\geq 0}$*.*

*Proof.* By the Martingale Central Limit Theorem (Theorem 7.1.4(b) of Ethier and Kurtz [16]), it is sufficient to show that

$$\frac{1}{N} Q_N(Nt) \xrightarrow{p} z(t) + t$$

and that

$$\frac{1}{N} \mathbb{E}\left[ \sup_{0\leq t\leq N\sigma} \left( M_N(t) - M_N(t-) \right)^2 \right] \to 0,$$

the latter being obvious from the fact that $M_N$ cannot jump by more than 1. Now, by (6.12), $Q_N(t) = A_N(t) + t$ and so it will be sufficient to show that $\frac{1}{N}A_N(Nt) \xrightarrow{p} z(t)$. By (6.10) and the definition of $z$ at (6.4),

$$\begin{aligned}
&\frac{1}{N}A_N(Nt) - z(t) \\
&\quad = \frac{1}{N}\int_0^{Nt} ((N - s - Z_N(s) - P_N(s))\lambda_2(N, s) - 1)ds - \int_0^t \dot{z}(s)ds \\
&\quad = \int_0^t \left(\left(1 - s - \frac{1}{N}Z_N(Ns) - \frac{1}{N}P_N(Ns)\right) N\lambda_2(N, Ns) - \beta''(s)(1 - s - z(s))\right) ds,
\end{aligned}$$

for $t \leq \sigma$. There exists a constant $C < \infty$ such that

$$\begin{aligned}
\left|\frac{1}{N}A_N(Nt) - z(t)\right| \leq \left|\int_0^t (1 - s)(N\lambda_2(N, Ns) - \beta''(s))ds\right| + C\int_0^t \left|\frac{1}{N}Z_N(Ns) - z(s)\right| ds \\
+ C\int_0^t \left|\frac{P_N(Ns)}{N}\right| ds.
\end{aligned}$$

By Lemma 1.5, there exists a $K < \infty$ such that the right-hand side is bounded above by

$$K\frac{(\log N)^2}{N} + C\int_0^t \left|\frac{1}{N}Z_N(Ns) - z(s)\right| ds + C\int_0^t \left|\frac{P_N(Ns)}{N}\right| ds.$$

As $\mu_k > 0$, by Theorems 6.1 and 6.3, only $\mathcal{O}(N^{\alpha(k)/2})$ patches need to be added before the breadth-first walk begins its giant excursion and so, as $\alpha(k) < 1$, we must have

$$\sup_{0 \leq t \leq \sigma} \left|\frac{P_N(Nt)}{\sqrt{N}}\right| \xrightarrow{p} 0$$

as $N \to \infty$. Thus, $\left|N^{-1}A_N(Nt) - z(t)\right|$ converges to 0 in probability, uniformly in $0 \leq t \leq \sigma$, by Theorem 6.4. $\square$

Now define

$$\begin{aligned}
X_t^N &= \frac{1}{\sqrt{N}}(Z_N(Nt) - Nz(t)) \\
G_t^N &= \frac{1}{\sqrt{N}}(Z_N(Nt) - A_N(Nt)).
\end{aligned}$$

Then we know already from Lemma 6.10 that $(G_t^N)_{0 \leq t \leq \sigma} \xrightarrow{d} (G_t)_{0 \leq t \leq \sigma}$ and we wish to prove that $(X_t^N)_{0 \leq t \leq \sigma} \xrightarrow{d} (X_t)_{0 \leq t \leq \sigma}$. Now,

$$\begin{aligned}
X_t^N &= G_t^N + \sqrt{N}\int_0^t (a_N(Ns) - 1)ds - \sqrt{N}\int_0^t \dot{z}(s)ds \\
&= G_t^N + E_t^N - \int_0^t X_s^N \beta''(s)ds,
\end{aligned}$$

where

$$E_t^N = \int_0^t \left(\sqrt{N}(1-s) - \frac{1}{\sqrt{N}}(Z_N(Ns) + P_N(Ns))\right)(N\lambda_2(N, Ns) - \beta''(s))ds$$
$$- \int_0^t \frac{1}{\sqrt{N}} P_N(Ns)\beta''(s)ds.$$

Observe that, because $1 \leq Z_N + P_N \leq N$, we have

$$\begin{aligned}|E_t^N| &\leq \int_0^t 2\sqrt{N}|N\lambda_2(N, Ns) - \beta''(s)|ds + \beta''(\sigma)t \sup_{0\leq s\leq t}\left|\frac{P_N(Ns)}{\sqrt{N}}\right| \\ &\leq 2t\frac{(\log N)^2}{\sqrt{N}} + \beta''(\sigma)t \sup_{0\leq s\leq t}\left|\frac{P_N(Ns)}{\sqrt{N}}\right|.\end{aligned}$$

At the end of the proof of Lemma 6.10, we showed that $\sup_{0\leq t\leq\sigma}\left|\frac{P_N(Nt)}{\sqrt{N}}\right| \xrightarrow{p} 0$ as $N \to \infty$ and so $\sup_{0\leq t\leq\sigma}|E_t^N| \xrightarrow{p} 0$.

Now, there exists a continuous function (of *paths*) $F : C_0[0,\sigma] \to C_0[0,\sigma]$ such that $X^N = F(G^N + E^N)$. Let $X = F(G)$ i.e. let $(X_t)$ satisfy

$$dX_t = dG_t - X_t\beta''(t)dt.$$

Then $(G_t^N + E_t^N)_{0\leq t\leq\sigma} \xrightarrow{d} (G_t)_{0\leq t\leq\sigma}$ and so, by continuity of $F$ and the Continuous Mapping Theorem (Corollary 3.1.9 of Ethier and Kurtz [16]), $(X_t^N)_{0\leq t\leq\sigma} \xrightarrow{d} (X_t)_{0\leq t\leq\sigma}$. By using an integrating factor to solve the above stochastic differential equation, we see that

$$X_t = \exp(-\beta'(t))\int_0^t \exp(\beta'(s))dG_s.$$

# Bibliography


[1] David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25(2):812–854, 1997.

[2] David Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[3] Krishna B. Athreya and Peter E. Ney. *Branching Processes*, volume 196 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1972.

[4] A. D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1992. Oxford Science Publications.

[5] Claude Berge. *Hypergraphs*, volume 45 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1989. *Combinatorics of Finite Sets*, Translation of *Hypergraphes*.

[6] Béla Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.

[7] Béla Bollobás. *Combinatorics*. Cambridge University Press, Cambridge, 1986. *Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability*.

[8] Béla Bollobás. *Random Graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[9] Hong Chen and David D. Yao. *Fundamentals of Queueing Networks*, volume 46 of *Applications of Mathematics*. Springer-Verlag, New York, 2001. *Performance, Asymptotics and Optimization*, Stochastic Modelling and Applied Probability.

[10] R. W. R. Darling. Fluid limits of pure jump Markov processes: a practical guide. Preprint math.PR/0210109, 2002.

[11] R. W. R. Darling, D. A. Levin, and J. R. Norris. Continuous and discontinuous phase transitions in hypergraph processes. Preprint, 2002.

[12] R. W. R. Darling and J. R. Norris. Vertex identifiability in large random hypergraphs. Preprint math.PR/0109020, 2002.

[13] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 1998.

[14] Pierre Duchet. Hypergraphs. In Graham et al. [18], pages 381–432.

[15] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei (Publ. Math. Inst. Hungar. Acad. Sci.)*, 5:17–61, 1960.

[16] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[17] William Feller. *An Introduction to Probability Theory and its Applications. Vol. 1.* John Wiley & Sons Inc., New York, third edition, 1968.

[18] R. L. Graham, M. Grötschel, and L. Lovász, editors. *Handbook of Combinatorics. Vol. 1, 2.* Elsevier Science B.V., Amsterdam, 1995.

[19] Theodore E. Harris. *The Theory of Branching Processes*, volume 119 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1963.

[20] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[21] Michał Karoński and Tomasz Łuczak. The phase transition in a random hypergraph. *Journal of Computational and Applied Mathematics*, 142(1):125–135, 2002. Special issue: Probabilistic methods in combinatorics and combinatorial optimization.

[22] Malwina J. Luczak. A quantitative law of large numbers via exponential martingales. Preprint, 2003.

[23] Tomasz Łuczak. Component behavior near the critical point of the random graph process. *Random Structures and Algorithms*, 1(3):287–310, 1990.

[24] Tomasz Łuczak. Phase transition phenomena in random discrete structures. *Discrete Mathematics*, 136(1-3):225–242, 1994.

[25] Jim Pitman. Combinatorial stochastic processes. To appear in *Lectures on Probability Theory and Statistics (Saint-Flour 2002)*. Springer-Verlag.

[26] Boris Pittel. On tree census and the giant component in sparse random graphs. *Random Structures & Algorithms*, 1(3):311–342, 1990.

[27] Boris Pittel, Joel Spencer, and Nicholas Wormald. Sudden emergence of a giant $k$-core in a random graph. *Journal of Combinatorial Theory. Series B*, 67(1):111–151, 1996.

[28] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition, 1999.

[29] Jeanette Schmidt-Pruzan and Eli Shamir. Component structure in the evolution of random hypergraphs. *Combinatorica*, 5(1):81–94, 1985.

[30] Adam Shwartz and Alan Weiss. *Large Deviations for Performance Analysis*. Stochastic Modeling Series. Chapman & Hall, London, 1995. *Queues, Communications and Computing*, With an appendix by Robert J. Vanderbei.

[31] George F. Simmons. *Introduction to Topology and Modern Analysis*. McGraw-Hill Book Co. Inc., New York, 1963.

[32] Ward Whitt. *Stochastic-Process Limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002. *An Introduction to Stochastic-Process Limits and Their Application to Queues*.

[33] Nicholas C. Wormald. Differential equations for random processes and random graphs. *The Annals of Applied Probability*, 5(4):1217–1235, 1995.