Epidemics on critical random graphs: limits and continuum descriptions

David Clancy, Jr.

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Soumik Pal, Chair

Krzysztof Burdzy

Zhen-Qing Chen

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Epidemics on critical random graphs:
limits and continuum descriptions

David Clancy, Jr.

Chair of the Supervisory Committee:
Professor Soumik Pal
Department of Mathematics

Understanding how diseases spread through populations is vital for mitigation efforts. For any disease at hand, the specifics of how a disease spreads through a community depends on many factors: how the disease is transmitted, how contagious the disease is, how long are individuals in the community contagious, what is the community structure of the population, are there any mitigation efforts taken, among many others. One way to sidestep the complications from studying a particularly disease at hand is to develop mathematical tools that work for understanding the behavior of epidemics on random community structures with relatively simple disease transmission mechanism. This is the topic of the present thesis.

More precisely, we model the community structure by a random graph G_n on n vertices. We make the simplifying assumption that each individual in the population (represented by a vertex in the graph) who becomes infected is infected for just a single day and is then forever cured of the disease. However, during that day the infected individual transmits the disease to each of their susceptible neighbors. We are interested in the large n behavior of

$$Z_n(t) = \#\{\text{infected vertices in } G_n \text{ on day } t\}, \qquad t = 0, 1, \cdots .$$

It turns out, the behavior of the epidemic over time corresponds to understanding certain aspects of the *geometry* of the connected components of the random graph.

For many random graphs models, in particular the Erdős-Rényi and configuration models

which we consider, the geometry of the connected components undergoes a quite dramatic change depending on a parameter R_0. The parameter R_0 is the average number of susceptible individuals infected by a single infected individual. The connected components, and hence the number of individuals infected in the epidemic, exhibit very different properties based on if $R_0 < 1$, $R_0 = 1$ or $R_0 > 1$. When $R_0 < 1$ one typically sees that the number of individuals infected is at most logarithmic in the population size n and when $R_0 > 1$ one typically sees that the largest possible epidemic infects order n many individuals.

The situation is quite different when $R_0 = 1$ where one often sees that each of the largest possible outbreaks when just a single individual is infected will affect $\Theta(n^\gamma)$ many individuals for some $\gamma \in (0,1)$ depending only on the random graph and not on the individuals selected. This is where we focus our attention and such random graphs G_n are called *critical*.

A popular approach for studying critical random graphs has emerged over the past few decades. It turns out that for many models of critical random random graphs (see Chapter 3 for more details and references) the metric space structure of the largest connected components of critical random graphs has some limiting description in terms of a *continuum random graph*. Unfortunately, the standard techniques used to describe the metric space structure do not immediately imply that the process Z_n has some limiting description as $n \to \infty$.

In this thesis we show that the standard approach for proving convergence of the connected components of a critical random graph to a limiting metric space, to say $\mathscr{G}$, is essentially enough to show that the process Z_n has some large n scaling limit, to say Z. This then gives a precise connection between properties of the metric space $\mathscr{G}$ and properties of the limiting process Z. For example, understanding compactness of the metric space $\mathscr{G}$ becomes understanding whether or not Z has compact support.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

ACKNOWLEDGMENTS

I would like to begin by thanking Soumik Pal for being my advisor. In particular, I'm extremely grateful for his helping me grow as a mathematician and as a person. I would not have accomplished what I have professionally if it were not for his constant encouragement and patience while I was (and am) struggling with research and learning new material. Knowing that research success behaves like the local time at zero of Brownian motion has proved quite helpful several times. His comments on my mathematics has also helped me to become a better writer and communicator of math.

I would also like to thank the other members of my committee, Krzysztof Burdzy, Zhen-Qing Chen and Matthew Lorig. I'm thankful for Krzysztof Burdzy helping me learn stochastic calculus during my first year of graduate school and Zhen-Qing Chen for helping me learn about Gaussian free fields during my last. I would also like to thank David Aldous for explaining some of his earlier work on random trees which significantly aided my later research.

I would like to thank Wai Tong (Louis) Fan for sparking my interest in probability theory during our mutual time at the University of North Carolina at Chapel Hill. I doubt I would have studied probability if my undergraduate research project did not happen. In this vein, I would also like to thank the anonymous writer of the 2016 Algebra Prelim. My failing that exam made my choice to study probability theory that much easier.

A separate thanks goes to Chris Hoffman and Bianca Viray for running their "Communicating Math Effectively" course a few years back. That course forced me to think about my research in a different way, taught me how to present my research both during presentations and on paper. It proved invaluable during my search for postdocs.

I would also like to thank several others at the University of Washington who have aided in my understanding of innumerable mathematical concepts - both by teaching me these tools and asking pressing questions while I was explaining these tools to them - or for emotional support when research was not as productive. These include (in alphabetical order) Clayton Barnes, Mark Bennett, Shuntao Chen, Joonyong Choi, Nico Courts, Nikolaos Eptaminitakis, Noah Forman, Carl-Erik Gauthier, Kirill Golubnichiy, Graham Gordon, Yiping Hu, Amzi Jeffs, Dami Lee, Xiangqian Meng, Tim Mesikepp, Reed Meyerson, Andrea Ottolini, Jacob Richey, Sam Roven, Anthony Sanchez, David Simmons, Josh Southerland, Tuomas Tajakka, Yang Yu, and Yizhe Zhu. I don't know if I could have made it through my early years of graduate school if it weren't for my dear friends in (or at least formerly in) Seattle: Joe Dailey, Eddie Elizondo, Maryl Evans, Peter Marcy, Emily Perkins, James Saindon, Clarisse Schneider, Russell Sutter, and Matt Wawiorka.

I would like to acknowledge my partner, Katherine Tully. You have provided me with immense emotional support throughout my time in graduate school. You have reassured me when I've struggled, listened when I've complained, and have encouraged me when I think I have an idea on how to solve a problem - even if you had no idea what I was talking about. I could not have done it without you.

Last, but certainly not least, I would like to dedicate this thesis to my parents Kathleen and David Clancy. They've been a part of my life since the very start! They made sure that I did my homework on time, would pick me up from high school when I had to stay late for academic clubs and so many other things along the way. If it were not for them, I do not believe I would have made it to college, let alone finish my Ph.D. As my mom says, this is their Ph.D. too.

DEDICATION

to Katherine

Chapter 1

INTRODUCTION

This thesis is about mathematically understanding disease spread in critical regimes where each infected person infects, on average, roughly a single susceptible individual. Often in this situation we can not only describe the evolution of the number of individuals infected over time as the size of the population grows large, but we can actually describe some limiting behavior of the history of the disease itself. More precisely the history of the disease spread is captured by a random graph G with some special marked vertices corresponding to the initially infected individuals. Edges in G represent the infection of one individual (the vertex further from those marked vertices) by another (the one close to the marked vertices). Many ways have been proposed in the literature for understanding the behavior of random graphs as their size grows large, and one of these ways is to describe the limiting metric structure of the graph by so-called *continuum random graphs*.

Apart from their connection to modeling disease spread, random graphs have become an area of increasing importance because, in part, they can be used to model abstract interactions of elements in large data sets, which have become increasingly easier to obtain [173]. We therefore spend a descent amount of time describing some of the tools that have become popular in the literature while leaving some of the technicalities of the proofs to cited references.

1.1 Brief History of Mathematical Epidemiology

The use of mathematics to model the behavior of disease spread dates back to the work of William Farr in 1840 involving a smallpox outbreak in the late 1830's in England, see [162] for a historical overview of the mathematical theory. Observing data of the number of smallpox deaths over successive 3-month periods, Farr approximated this data with a bell-curve in order to predict the number of smallpox deaths in the coming months. In the 1910's, Sir Ronald Ross, who discovered the link between malaria and mosquitoes, and Hudson [160, 161] modeled the behavior of epidemics by differential equations. This built off of difference equations used by Ross in his second edition of his book *The Prevention of Malaria* [159] under the labeling of "Theory of Happenings," which does not appear to have caught on.

A more general approach was taken by Kermack and McKendrick in [115] in 1927 and then later developed in [116–119] through the 1930's. Their set-up in [115] is worth quoting at length:

> One (or more) infected person is introduced into a community of individuals, more or less susceptible to the disease in question. The disease spreads from the affected to the unaffected by contact infection. Each infected person runs through the course of his sickness, and finally is removed from the number of those who are sick, by recovery or by death. The chances of recovery or death vary from day to day during the course of his illness. The chances that the affected may convey infection to the unaffected are likewise dependent upon the stage of the sickness. As the epidemic spreads, the number of unaffected members of the community becomes reduced. Since the course of an epidemic is short compared with the life of an individual, the population may be considered as remaining constant, except in as far as it is modified by deaths due to the epidemic disease itself. In the course of time the epidemic may come to an end.

Kermack and McKendrick proceed to obtain a system of non-linear differential equations by discretizing time and informally taking limits. They allow for infected individuals to become more (or less) contagious over time as well as their recovery rate to vary over time. While we will not describe their general differential equations; however, when the rates are constant their equations become

$$\frac{dS}{dt} = -\beta SI, \qquad \frac{dI}{dt} = +\beta SI - \gamma I, \qquad \frac{dR}{dt} = +\gamma I,$$

where S, I and R are respectively the number of susceptible individuals, the number infected individuals and number individuals who have recovered or died (i.e. *r*emoved) and β, γ are constants.

The model quoted above by Kermack and McKendrick is described probabilistically. They use the phrase "[t]he chances" to describe both recovery and infections. However, they do not explicitly use probability to derive their equations. This is because the equations they use to model the discretized behavior over time are the expected values of a corresponding stochastic model.

A stochastic model was used by Lowell J. Reed and Wade Hampton Frost for a course they taught at Johns Hopkins in the second quarter of the 20th century; however, their work went unpublished [1]. Their model is described as follows. Infection is spread from person-to-person through some direct contact. Individuals are infected for a single day, during which they are contagious, but afterwards they are forever cured of the disease. Each day, each non-infected individual has some fixed and constant over time probability p of coming into contact with an infected individual and consequently becoming infected. An infected individual can infect many others (that is they can be *superspreaders*), they can infect no others, or something in between.

In either the set-up for the Kermack-McKendrick model or the Reed-Frost model above, we can create a directed graph G representing the disease transmission. Namely, we label all the individuals in the in the population by $1, \cdots, n$, and then a directed edge $i \to j$ represents that individual i infected individual j. Since the descriptions of the disease spread in both models involve the day-to-day interactions of the individuals in the population, the graph G is discovered over time in a breadth-first manner. We could, alternatively, forget the direction of the edges in G but instead specify the initially infected vertices and consider $(G, \{\rho_1, \cdots, \rho_k\})$ where $\rho_1, \cdots, \rho_k$ are k distinct elements of $\{1, \cdots, n\}$ representing the initially infected individuals. We will work with this second interpretation.

Conversely, given a graph G and vertices $\rho_1, \cdots, \rho_k$ in G there is a corresponding epidemic where the k chosen vertices are the initially infected individuals. Namely, on day zero the vertices $\rho_1, \cdots, \rho_k$ are infected and on day t the infected vertices are

$$\{v \in G : \min_{1 \leq j \leq k} d(v, \rho_j) = t\},$$

where $d(u, v)$ is the graph distance between two vertices u and v.

This gives a direct connection between disease transmission and graphs. In particular, the disease transmission terminates corresponds precisely with the termination of the breadth-first exploration of the graph. However, if we explore a graph starting from a vertex $\rho \in G$ in either a depth-first or breadth-first manner we only discover information of the connected component of G containing the vertex ρ. Therefore, the behavior of disease transmission is directly related to the *structure of the connected components of* G.

1.1.1 Overview of the Rest of the Chapter

In Section 1.2, we begin by describing the simplest random graph models and preliminary information about the structure of their connected components when the size of the graph $n = \#G$ tends towards infinity. As we will see there is often a simple characteristic of the random graph which describes the threshold for a connected component of G to have $\Theta_{\mathbb{P}}(n)$ many vertices - that is with probability tending towards 1 as $n \to \infty$ there is a connected component of G with order $\Theta(n)$ many vertices. This threshold is directly related to branching processes. While for most of the chapter we focus on the simplest case of the Erdős-Rényi random graph, we do point to extensions to other more general and more realistic graph models.

The structure of the connected components - or more precisely the behavior of epidemics on the connected components - at the respective threshold is the main work of my thesis and so we will spend quite a bit of time on this. For many models of critical random graphs there are known limits of the largest connected components *as metric spaces*. These are described in detail in Chapter 2 and 3 starting with the pioneering work of Aldous [9] on the *continuum random tree*.

1.2 The Erdős-Rényi random graph

The stochastic nature of Reed-Frost model means that the corresponding (undirected) graph G is, itself, *random*. The fixed probability p that a infected individual infects a non-infected individual at a particular step in the process implies the each possible pair of edges in the exploration of the graph appears with probability p and each edge is independent of each other.

In fact, the Reed-Frost model naturally leads to the study of the random graph $G(n,p)$ on n vertices labeled by $[n] := \{1, \cdots, n\}$ and each edge (i,j) is included in $G(n,p)$ with probability p. This graph was introduced by Gilbert in [93] in a different context involving routing telephone calls[1]. We should point out that the graph studied by Gilbert in [93] is not exactly the graph constructed from the Reed-Frost model; however, the disease transmission will be the same. The difference between the two models is subtle, but the graph constructed from the Reed-Frost model does not contain edges (i,j) corresponding to two vertices i,j which never get infected at any point or are both infected on the same day. On the other hand, the graph $G(n,p)$ contains this edge with probability p. These possible additional edges do not contribute to the disease transmission however and so we will ignore these differences.

Similar random graph models were introduced and studied around the same time, but neither were initially concerned with disease propagation. In [22], the authors studied a random (multi)graph on n vertices where exactly M of the $\binom{n}{2}$ edges are chosen at random and *with replacement* where the authors were concerned with understanding the number of connected components of this graph. At almost the same time as Gilbert, Erdős and Rényi published their seminal papers [84, 85] describing the structure of random graphs $G(n,M)$ on n vertices containing exactly M *distinct* edges where each of the $\binom{\binom{n}{2}}{M}$ graphs is chosen with equal probability. They were chiefly concerned with identifying conditions on the growth rate of $M = M(n)$ as $n \to \infty$ where various combinatorial properties of the graph hold with probabilities that can be computed. It is now common to call both $G(n,p)$ and $G(n,M)$ an Erdős-Rényi random graph and many properties of $G(n,M)$ remain true for $G(n,p)$ where $\binom{n}{2}p \approx M$ [138].

In terms of the Reed-Frost epidemic, there are two natural questions to consider as $n \to \infty$. If there are n people in the population, what is the probability that everyone eventually contracts the disease and what is the probability that εn many individuals contract the

[1]Gilbert worked for Bell Labs at the time so the concern of network communication should not be that surprising. Later in [94], Gilbert studied a random graph with an application to disease spread. It was constructed from a Poisson point process in the plane where atoms of the process are the vertices of an infinite graph and an edge appears between two vertices if and only if they are within some fixed distance R of each other.

illness at some point? Clearly, the answers to these questions depend on the behavior of $p = p(n)$ as $n \to \infty$. Using the connections between graphs and epidemics these two questions relate to (1) the graph G being connected, and (2) the graph G possessing a connected component which contains a positive proportion of all the vertices. Using the connection between the Reed-Frost model and the Erdős-Rényi random graph, $G(n,p)$, these two questions can be rephrased and refined as follows

1. Can we identify conditions on $p = p(n)$ such that we can compute
$$\lim_{n\to\infty} \mathbb{P}\left(G(n,p) \text{ is connected}\right)?$$

2. Let $A(n,p)$ denote the size of the largest connected component of $G(n,p)$. Also let $f(n)$ be some given function such that $f(n) \to \infty$. Can we identify conditions on $p = p(n)$ such that $f(n)/A(n,p)$ and $A(n,p)/f(n)$ are tight random variables? Equivalently, can we identify conditions on $p = p(n)$ such that given any $\varepsilon > 0$ can we find a $\delta > 0$ such that
$$\mathbb{P}\left(\delta < \frac{A(n,p)}{f(n)} < \delta^{-1}\right) \geq 1 - \varepsilon, \qquad \text{for all } n \text{ sufficiently large?} \tag{1.1}$$

The second condition is stated more succinctly using Big-O notation. Let $((X_n, Y_n); n \geq 1)$ be sequence of pairs of random elements of $\mathbb{R} \times \mathbb{R}_+$. We say that $X_n = O_{\mathbb{P}}(Y_n)$ if $|X_n|/Y_n$ is tight, we say that $X_n = o_{\mathbb{P}}(Y_n)$ if $X_n/Y_n \overset{d}{\Longrightarrow} 0$ and we say that $X_n = \Theta_{\mathbb{P}}(Y_n)$ if both $X_n = O_{\mathbb{P}}(Y_n)$ and $Y_n = O_{\mathbb{P}}(X_n)$. The second condition is now just identifying conditions on $p(n)$ for a given function $f(n)$ such that $A(n,p) = \Theta_{\mathbb{P}}(f(n))$.

We will briefly touch on both of these phenomena before moving on to understanding the so-called "critical behavior" of the connected components of the Erdős-Rényi random graph. For this graph model there is a sharp threshold for connectivity and there is a phase transition for the emergence of a giant componetn of size $\Theta_{\mathbb{P}}(n)$.

1.2.1 Threshold for connectivity

The threshold for connectivity is known to be $p = p(n) = \frac{\log n}{n}$. More precisely

Theorem 1.2.1 (Erdős-Rényi [84]). *Suppose that* $p = p(n) = n^{-1}(\log n + c)$ *for some* $c \in \mathbb{R}$. *Then*

$$\mathbb{P}(G(n,p) \text{ is connected}) \longrightarrow e^{-e^{-c}}, \qquad \text{as } n \to \infty.$$

We will not provide a full proof of this result but instead provide some intuition for why this holds. A full proof can be found in Theorem 7.3 in [40].

Note that $G(n,p)$ is disconnected if and only if there is some connected component in $G(n,p)$ with at most $\lfloor \frac{n}{2} \rfloor$ many vertices. It can be shown that when $p = n^{-1}(\log(n) + c)$ that

$$\mathbb{P}\left(G(n,p) \text{ has a component of size } r \text{ where } 2 \le r \le \frac{n}{2}\right) \to 0 \quad \text{ as } n \to \infty.$$

Every vertex in a component with at least 2 vertices has to have degree ≥ 1, so the above convergence can be rephrased as follows: With probability tending towards 1 as $n \to \infty$ the graph $G(n,p)$ is made up of a single connected component with $\geq \frac{n}{2}$ many vertices and components that are just isolated vertices.

Let $X_0(n,p)$ denote the number of isolated vertices in $G(n,p)$. The proof of Theorem 1.2.1 will be finished if we can show that $X_0(n,p) \stackrel{d}{\Longrightarrow} \text{Poisson}(e^{-c})$. Clearly we have

$$X_0(n,p) = \sum_{i=1}^{n} 1_{[\deg(i)=0]}.$$

Now $\deg(i) \stackrel{d}{=} \text{Bin}(n-1,p)$ and so

$$\mathbb{E}\left[X_0(n,p)\right] = \sum_{i=1}^{n} \mathbb{E}\left[1_{[\deg(i)=0]}\right] = n\mathbb{P}(\text{Bin}(n-1,p) = 0). = n(1-p)^{n-1}.$$

It is also not hard to show that as $n \to \infty$

$$n(1-p)^{n-1} = n \exp\left(-(\log n + c) + O\left(\frac{(\log n)^2}{n}\right)\right) \longrightarrow e^{-c}$$

and hence

$$\mathbb{E}[X_0(n,p)] \longrightarrow e^{-c}.$$

One can show that

$$\mathbb{E}\left[\binom{X_0(n,p)}{k}\right] = \sum_{i_1 < i_2 < \cdots < i_k} \mathbb{P}\left(\text{vertices } i_1, \cdots, i_k \text{ are isolated}\right) = \binom{n}{k}(1-p)^{k(n-k)+\binom{k}{2}}$$

because there are $\binom{n}{k}$ many ways to choose the k vertices, and there are $k(n-k)+\binom{k}{2}$ many possible possible edges which can include any vertex $i_1, \cdots, i_k$. It is not difficult to prove

$$\mathbb{E}\left[\binom{X_0(n,p)}{k}\right] = \binom{n}{k}(1-p)^{k(n-k)+\binom{k}{2}} \longrightarrow \frac{e^{-ck}}{k!} = \mathbb{E}\left[\binom{\text{Poisson}(c)}{k}\right].$$

1.2.2 Emergence of a Giant Component

In this section we discuss conditions on $p = p(n)$ such that (1.1) holds almost surely with $f(n) = n$.

To begin with recall how in the previous section we relied on the fact that for any vertex $i \in G(n,p)$ its degree $\deg(i) \overset{d}{=} \text{Bin}(n-1,p)$. In particular, when $p = p(n) = \frac{c}{n}$ for some $c \geq 0$ we can observe that

$$\text{Bin}\left(n-1, \frac{c}{n}\right) \overset{d}{\Longrightarrow} \text{Poisson}(c),$$

and so each vertex i in the graph $G(n, c/n)$ for large n should have approximately Poisson(c) many neighbors. Also note that the probability that any two of the neighbors of vertex i, say j_1 and j_2 share an edge is c/n which tends to zero as $n \to \infty$. That is we should not expect i is a part of a triangle. This means that the induced subgraph $H(i,1)$ with vertices $\{v \in G(n,p) : d(i,v) \leq 1\}$ is a tree with probability $(1 - o(1))$.

Similarly, each of the neighbors j of the vertex i should have approximately Poisson(c) *excluding* the vertex i. Within the induced subgraph $H(i,2)$ with vertices $\{v \in G(n,p) : d(i,v) \leq 2\}$, the probability that we see a cycle will be order $\frac{1}{n}$. That is the induced subgraph $H(i,2)$ is with high probability a tree, where the root vertex i has approximately Poisson(c) many neighbors and each vertex j has approximately Poisson(c) many neighbors excluding i. This is precisely the first two levels of a Galton-Watson tree with offspring distribution Poisson(c) offspring distribution [2].

One can actually prove something stronger [61, Proposition 2.6]. Namely, locally, as $n \to \infty$ the Erdős-Rényi random graph converges (in some mathematically precise way) to a Galton-Watson branching tree with offspring distribution Poisson(c). It turn out this size of the giant component in the Erdős-Rényi random graph $G(n, cn^{-1})$ is directly related

[2] See Section 2.2.2.6 for a definition of Galton-Watson trees.

to the almost sure extinction of a Galton-Watson branching processes with a Poisson(c) branching mechanism.

Recall that $Z = (Z(t); t = 0, 1, 2, \cdots)$ is a Galton-Watson branching process with offspring distribution μ where μ is a probability measure on $\mathbb{N}_0 = \{0, 1, 2, \cdots\}$ if Z is a Markov chain on $\mathbb{N}_0$ such that

$$(Z(t+1)|Z(t)) \overset{d}{=} \sum_{j=1}^{Z(t)} \xi_j,$$

where $(\xi_j; j = 1, 2, \cdots) \overset{\text{i.i.d.}}{\sim} \mu$. Note that Z is absorbed upon hitting 0, which we call the extinction of Z. A classical result in branching processes [20, Section I.5] says that if $\mu(\{1\}) < 1$ then

$$\mathbb{P}(Z \text{ never becomes extinct}) = \begin{cases} 0 & : \sum_{k\geq 0} k\mu(k) \leq 1 \\ p > 0 & : \sum_{k\geq 0} k\mu(k) > 1 \end{cases}.$$

In particular, a Galton-Watson branching tree with offspring distribution Poisson(c) becomes extinct almost surely if and only if $c \leq 1$ and has a positive probability of never going extinct if and only if $c > 1$.

This is precisely the phase transition for the Erdős-Rényi random graph $G(n, cn^{-1})$ to have a component of order $\Theta_{\mathbb{P}}(n)$. That is there is drastically different behavior of the largest connected component of $G(n, cn^{-1})$ depending on whether $c > 1$ or not. If fact, the following holds:

Theorem 1.2.2 (Erdős-Rényi [85])**.** *Fix $c > 0$. For the Erős-Rényi random graph $G(n, cn^{-1})$, there is the following trichotomy:*

1. *If $c > 1$ then the largest connected component has $\Theta_{\mathbb{P}}(n)$ many vertices. The corresponding disease outbreak starting from 1 individual will infect $O_{\mathbb{P}}(n)$ many individuals.*

2. *If $c = 1$ then the largest connected component has $\Theta_{\mathbb{P}}(n^{2/3})$ many vertices. The corresponding disease outbreak starting from 1 individual will infect $O_{\mathbb{P}}(n^{2/3})$ many individuals.*

3. *If $c < 1$ then the largest connected component has $\Theta_{\mathbb{P}}(\log n)$ many vertices. The corresponding disease outbreak starting from 1 individual will infect $O_{\mathbb{P}}(\log n)$ many individuals.*

The respective regimes are called *supercritical, critical,* and *subcritical.* The change from $\Theta_{\mathbb{P}}$ to $O_{\mathbb{P}}$ is because the first individual infected need not be in the largest connected component. Indeed, let V be a uniform vertex in $G(n, cn^{-1})$, independent of $G(n, cn^{-1})$. Conditionally on the graph $G(n, cn^{-1})$, the vertex V is in the largest connected component with probability $A(n,p)/n$ which is $\Theta_{\mathbb{P}}(1)$ (resp. $o_{\mathbb{P}}(1)$) when $c > 1$ (resp. $c \leq 1$).

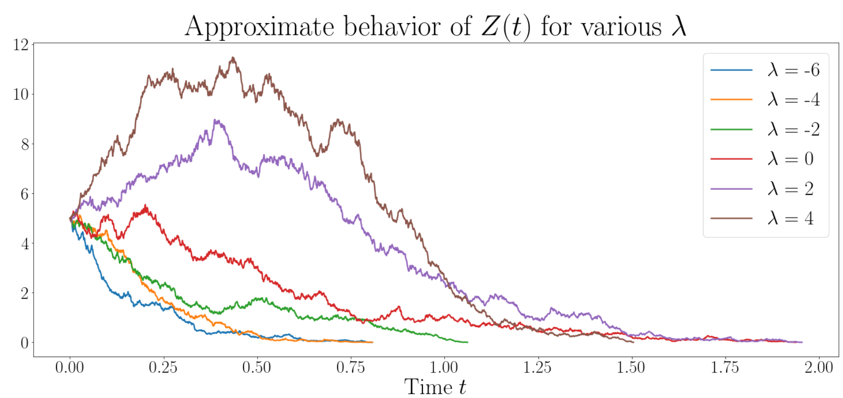


Figure 1.1: Approximate behavior of the critical and near-critical Reed-Frost model, which corresponds to the critical and near-critical Erdős-Renyi random graph.

There is actually much more that we can say about the structure of the largest connected components in the critical and near-critical regime. It turns out that the geometry of the random graphs has some nice limiting features, namely the largest connected components each have a scaling limit called continuum random graph. The limit takes place in the space of metric spaces equipped with measures with some technical restrictions. In order to properly formulate this convergence, we will take a brief aside on how to measure how close two compact pointed metric measure spaces are. If we view the graph as telling us the history of the disease spread once the epidemic is over, the continuum random graph limit gives us a notion of the history of disease spread in the continuum after taking an $n \to \infty$

limit. The limiting descriptions of epidemics on the critical and near-critical Erdős-Rényi random graphs are quite different than the critical epidemics on models of random graphs super-spreaders (compare Figure 1.1 and Figure 1.2 below).

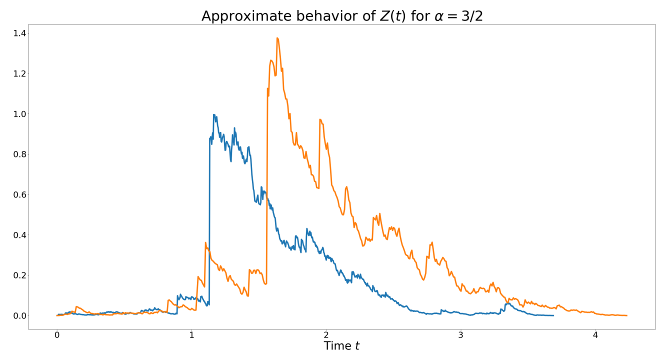


Figure 1.2: Approximate behavior of an epidemic on a random graph model exhibiting super-spreading phenomena. The number of individuals each person infects is in the domain of attraction of an $\alpha = 3/2$ stable random variable.

1.2.3 Overview of Rest of Thesis

The rest of the thesis is organized as follows.

Chapter 2 discusses the technical tools popular in the literature for describing continuum limits of random trees and random graphs. We start this chapter with a brief overview of Aldous's breakthrough continuum random tree trilogy [9–11] which essentially starts the probabilistic study of these continuum objects. Much of technical tools developed over the subsequent decades are then expounded on afterwards. No original work is present in this chapter and it is meant to serve as an introduction to many of the technical tools this field.

Chapter 3 discusses some of the scaling limits for critical and near-critical random graph models. Namely, we discuss the asymptotic component size for the Erdős-Rényi random graph, the configuration model and an inhomogeneous random graph model. Since much of the work on these random graph models involve an exploration introduced by Aldous [12], we spend some time discussing that exploration along with providing heuristics for why that holds. Since Chapter 6 relates random trees to results stemming from random matrix theory, we conclude Chapter 3 with a brief review of some results on eigenvalues of random matrices.

Chapter 4 focuses on the (near-)critical Erdős-Renyi random graph and borrows heavily from my work in [57]. For this model, we can analyze the aforementioned Reed-Frost model on a population of size n with $p = p(n) = (1 + o(1))n^{-1}$ when there are a large number of individuals infected at time $t = 0$.

The resulting description of the limiting disease transmission discussed in Chapter 4, almost breaks down when just a single individual is infected at day 0. We say "almost" because the results still apply; however, the limiting processes are trivial. Chapter 5 describes some general methods to overcome the problem of describing the limiting properties of disease outbreaks on critical random graphs when just a single individual is infected on day 0. This is my work in [58].

The basis of the research project initiated in [57] and described in Chapter 4, was motivated by a question posed by David Aldous during a presentation on my work in [56]. Namely, in [56] I use a breadth-first description of certain random forest models [8, 72] to describe and generalize a result originating in random matrix theory [97, 126]. In [12], Aldous used a breadth-first exploration of the near-critical Erdős-Rényi random graph to obtain distributional limits for the size of the largest connected components. Aldous asked whether or not my results in [56] can be used to describe the near-critical Erdős-Rényi random graph. In Chapter 6, I present my work in [56].

Chapter 2

BEGINNING OF CONTINUUM LIMITS

We provide a brief overview of convergence of Aldous's description for the large uniform random trees via an embedding into ℓ^1. This description describes large random objects in terms of a continuum object, the continuum random tree. Afterwards, we describe some of the technical tools present in the literature that are used to describe large random discrete graphs as metric spaces. It turns out that the simplest descriptions of the limiting continuum random graphs obtained by limits of large connected graphs follows by first looking at a particular spanning trees which essentially captures of the local geometry of the graphs and, with just a little more randomness, can be used to capture the entirety of the components geometry. That is, we will start with describing the behavior of random trees on n vertices and then we will move to describing random graphs.

Since this field starts with the work of Aldous [9], this is where we begin in Section 2.1. Aldous does most of the construction of continuum limits in the space ℓ^1 of absolutely summable sequences. In Section 2.2 we discuss the more abstract way of showing convergence of scaling limits thas become more prevalent in the literature.

2.1 The Continuum Random Tree, Aldous' construction

As we briefly described (and informally justified) above, the structure of the Erdős-Rényi random graph $G(n, cn^{-1})$ is closely related to Galton-Watson branching tree with Poisson(c) offspring law. Consider a tree $\mathcal{T}$ generated as follows. Initially, $\mathcal{T}$ contains a single vertex (0) called the root, this vertex gives birth to Poisson(c) many children $(0,1),\cdots,(0,\ell)$ for some ℓ. The vertices $(0,j)$ are said to be in generation 1. Subsequently, each vertex $(0,u_1,\cdots,u_n)$ at generation n gives birth to Poisson(c) many children (independently of everything else) and the children are labeled $(0,u_1,\cdots,u_n,1),\cdots,(0,u_1,\cdots,u_n,r)$ for some $r\geq 0$.

Let $\mathcal{T}_n$ be the tree $\mathcal{T}$ conditioned on having exactly n vertices, and we will *relabel* the vertices by labeling (0) by 1 and randomly assigning each of the remaining $n-1$ vertices in $\mathcal{T}$ a unique label in $\{2,\cdots,n\}$. It turn out that $\mathcal{T}_n$ is uniformly distributed on all n^{n-2} labeled trees on vertices $\{1,\cdots,n\}$ which are rooted at $\rho := 1$. This gives a *rooted labeled tree* on n vertices.

In [9], Aldous was concerned, in part, on the metric space structure of a uniform tree $\mathcal{T}_n$ on n labelled vertices. The metric geometry of $\mathcal{T}_n$ does not depend on the labels and so we can view $(\mathcal{T}_n,\rho)$ as a (nonuniform) random unlabelled rooted tree on n vertices. This gives us two distinct ways to view the tree $\mathcal{T}_n$.

In a prior work [7], Aldous gives a simple probabilistic construction of $\mathcal{T}_n$ when $n \geq 2$ is fixed. We will temporarily work with just some n fixed and so we omit additional n's from the notation until we begin to take the $n\to\infty$ limit. The construction Aldous gives is as follows:

Algorithm 1. Fix $n \geq 2$.

> Take a root vertex 1 and let $U_2, U_3, \cdots, U_n$ be independent uniform random variables on $\{1,2,\cdots,n\}$
>
> For $2 \leq i \leq n$, connect vertex i to $U_i \wedge (i-1)$.

If we forget the labels, this construction gives something equal in law to $(\mathcal{T}_n,\rho)$ viewed as a random rooted unlabeled tree. We can include labels if in the last step of the algorithm we, instead, relabelled all the vertices uniformly at random and re-rooted at the vertex 1. In that situation, we get something equal in law to $(\mathcal{T}_n,\rho)$ viewed as a random rooted (at 1) labeled tree.

This algorithm naturally gives us a probabilistic way to embed $(\mathcal{T}_n,\rho)$ into the metric (Banach) space $\ell^1 = \{(x_i : i \geq 1) : \sum_i |x_i| < \infty\}$ as follows. Let $z_i = (0,\cdots,0,1,0,\cdots)$ be the element of ℓ^1 with zeros everywhere except the i^{th} location, which contains a 1. This is a basis of the Banach space ℓ^1. Let $(U_i; 2 \leq i \leq n)$ be the uniform random variables on $\{1,\cdots,n\}$ as above and, conditionally gives $(U_i; 2 \leq i \leq n)$ inductively define

$(J_i; i = 1, \cdots, n)$ by $J_1 = 1$ and

$$J_i = \begin{cases} J_{i-1} + 1 & : U_i < i - 1 \\ J_{i-1} & : U_i \geq i - 1 \end{cases}.$$

Given $(U_i; 2 \leq i \leq n)$ and $(J_i; 1 \leq i \leq i)$ we define elements $V_i \in \ell^1$ inductively by $V_1 = 0$ and

$$V_i = V_{U_i \wedge (i-1)} + z_{J_i}, \qquad i \geq 2.$$

Now in both $\{V_1, \cdots, V_n\}$ and the tree constructed via the Algorithm 1 (which is prior to forgetting the labels) we see that element V_i (resp. vertex i) is distance 1 from element $V_{U_i \wedge (i-1)}$ (resp. $U_i \wedge (i-1)$). The choice of working in ℓ^1 with the ℓ^1-norm means the the metric geometry of $\mathcal{T}_n$ is the same as $\mathcal{S}_n := \{V_1, \cdots, V_n\} \subsetneq \ell^1$. That is $\mathcal{T}_n$ and $\mathcal{S}_n$ is isometric and the root $\rho \in \mathcal{T}_n$ is mapped to $V_1 = 0$ under this isometry.

We will now begin to discuss $n \to \infty$ limits, so we will include the additional subscript n to indicate dependence on n. For example, $J_{n,i}$ will replace J_i. Looking at the construction of the sequence $(J_{n,i}; i = 1, \cdots, n)$, a relevant thing to note is that $J_{n,i} \neq J_{n,i-1}$ if and only if $U_{n,i} < i - 1$. Similarly, $V_{n,i} = V_{n,i-1} + z_{J_i}$ if and only if $U_{n,i} \wedge (i-1) = i - 1$ if and only if $J_{n,i} = J_{n,i-1}$.

Now let $(C_{n,j}, B_{n,j})$ for $j \geq 1$ be defined as follows. The integer variables $1 < C_{n,1} < C_{n,2} < \cdots$ are simply

$$\{C_{n,1}, C_{n,2}, \cdots\} = \{i \in \{2, \cdots, n\} : U_i < (i-1)\}.$$

Set $B_{n,j} = U_{n,C_{n,j}}$ It is easy to see by the preceding paragraph that for $\ell \geq 1$

$$\{i : J_{n,i} = \ell\} = \{C_{n,\ell-1}, C_{n,\ell-1} + 1, \cdots, C_{n,\ell} - 1\}, \qquad C_{n,0} := 1.$$

Since vertices i and $i-1$ share an edge if and only if $J_i = J_{i-1}$, we should be able to describe Algorithm 1 for generating $(\mathcal{T}_n, \rho)$ using the random variables $C_{n,j}$ and $B_{n,j}$. Indeed, this can be done as follows.

Algorithm 2. Fix $n \geq 2$.

> Generate $C_{n,j}$ and $B_{n,j}$ as above and consider the path graph P_n on vertices $1, \cdots, n$ where each edge is of the form $\{i, i+1\}$. Root the tree at 1.

For each $j \geq 1$, remove the edge $\{C_{n,j}-1, C_{n,j}\}$ and replace it with $\{C_{n,j}, B_{n,j}\}$.

Both Algorithms 1 and 2 give isometric (unlabelled) trees. The importance of Algorithm 2 is that we can construct the rooted tree $(\mathcal{T}_n, \rho)$ by breaking the stick P_n according to $C_{n,j}$ and gluing the resulting segments back together using $B_{n,j}$.

The key observation in [9] is that $(C_{n,j}, B_{n,j})$ are the atoms of a random measure which has a scaling limit as $n \to \infty$. Namely define the random measure m_n on $\mathbb{R}_+$ by

$$m_n(A) = \sum_{j\geq 1} 1_{[C_{n,j}\in A]} = \sum_{i\in A} 1_{[U_i < i-1]}$$

is a random measure on $\mathbb{R}_+$ with mean

$$\mathbb{E}[m_n(A)] = \sum_{i\in A} \mathbb{P}(U_i < i-1) = \sum_{i\in A} \frac{i-2}{n}.$$

In particular, we see that for the measure $\mu_n((a,b]) := \mathbb{E}[m_n((n^{1/2}a, n^{1/2}b])]$ that as $n\to\infty$

$$\mu_n((a,b]) = \sum_{i\in(n^{1/2}a, n^{1/2}b]\cap\mathbb{N}_0} \frac{i-2}{n} \Longrightarrow \int_a^b t\,dt,$$

by, for example, convergence of Riemann sums. Similarly, it is not difficult to see that

$$\left((n^{-1/2}C_{n,1}, n^{-1/2}B_{n,1}), (n^{-1/2}C_{n,2}, n^{-1/2}B_{n,2}), \cdots\right) \stackrel{d}{\Longrightarrow} ((C_1,B_1),(C_2,B_2),\cdots),$$

where $(C_j; j\geq 1)$ are the jump times of a Poisson point process with rate $t\,dt$ and, conditionally given C_j, by $B_j \sim \mathrm{Unif}(0, C_j)$. Thus, there should be some continuum analog of Algorithm 2. Given the sequence $((C_j, B_j); j\geq 1)$ define $C_0 = B_0 = 0$ and define the function $p: \mathbb{R}_+ \to \ell^1$ with $p(0) = 0$ and

$$p(t) = p(B_i) + (x - C_i)z_{i+1}, \qquad t\in(C_i, C_{i+1}], i\geq 0.$$

Now for each $t \geq 0$, we get a random subset $\mathscr{S}_t := p([0,t]) \subset \ell^1$ and we can equip this space with a probability measure μ_t which is the pushforward under p of the uniform measure on $[0,t]$. Aldous shows that there exists a random compact set $\mathscr{S}$ and a random measure on $\mathscr{S}$, μ, such that almost surely

$$\mathscr{S}_t \to \mathscr{S}, \qquad \text{and} \quad \mu_t \to \mu, \qquad \text{as } t\to\infty \tag{2.1}$$

where the convergence of subsets is with respect to the Hausdorff metric of compact subsets of ℓ^1 and the converges of measures is with respect to the topology of weak convergence of probability measures on ℓ^1.

Let us go back to the interpretation of $\mathcal{T}_n$ as a random subset $\mathcal{S}_n$ of ℓ^1. Let also, ν_n denote the uniform measure on (the discrete space) $\mathcal{S}_n$. Also recall that we scale both $(C_{n,j}, B_{n,j})$ by $n^{-1/2}$ in order to obtain the limit $((C_j, B_j); j \geq 1)$ and let $\text{scale}_c : \ell^1 \to \ell^1$ be defined by $\text{scale}_c((x_i; i \geq 1)) = (cx_i; i \geq 1)$. Aldous proves the following theorem

Theorem 2.1.1 (Aldous [9]). *Suppose that $\mathcal{T}_n$ is a uniformly chosen labeled tree on n vertices rooted at 1. Let $\mathcal{S}_n = \{V_1, \cdots, V_n\}$ be the isometric representation of $\mathcal{T}_n$ as a subset of ℓ^1 and let ν_n be the uniform measure on $\mathcal{S}_n$. Then, for the $(\mathscr{S}, \mu)$ defined by (2.1)*

$$\left(\text{scale}_{n^{-1/2}}(\mathcal{S}_n), (\text{scale}_{n^{-1/2}})_{\#}\, \nu_n\right) \xrightarrow{d} (\mathscr{S}, \mu)\,.$$

The convergence of $\mathcal{S}_n$ is weakly in the space of compact subsets of ℓ^1 with the Hausdorff topology and the converge of ν_n is weakly with respect to the weak topology on probability measures on ℓ^1.

The limiting object $\mathscr{S}$ is now called the *Brownian continuum random tree*. Shortly we will see why this deserves the moniker "Brownian," but for not this seems justified by the $n^{-1/2}$ scaling for an object of size n and we will also see that it deserves the name "tree" because it shares many properties of discrete trees.

Let us also note that $\mathscr{S}$ is a random subset of ℓ^1 and the above convergence is weak convergence in the Hausdorff topology (see Section 2.2.1 below) on the space of closed subsets of ℓ^1. In particular, this relies on the topology of ℓ^1 and a particular embedding into ℓ^1 of the random trees $\mathcal{T}_n$ as a random subset denoted by $\mathcal{S}_n$ above. The modern approach to understanding the large n behavior of random trees and random graphs on n vertices often does not rely on any particular embedding into some ambient space. Instead, the graphs and their continuum limits are described as random metric spaces.

That is, we now view $\mathscr{S}(\omega)$ together with the specified point $0 \in \mathscr{S} \subset \ell^1$ (its root) and its metric d as a representative of isomorphism class of pointed compact metric spaces. This makes $\mathscr{S}$ a random element in some space of (isomorphism classes) pointed compact metric spaces. In fact, as we will describe in Section 2.2.1 below, the metric space $\mathscr{S}$ can be viewed as a random element in some Polish space and therefore the tool of weak convergence becomes available. We can also view $\mathscr{S}$ as a measured metric space by including information about the measure μ as well, and then $(\mathscr{S}, 0, d, \mu)$ becomes an element in a space of (isomorphism classes of) pointed compact measured metric spaces (which can also be topologized as a Polish space).

These tools were not present in the probability literature when Aldous described the Brownian continuum random tree so much of Aldous' pioneering work in [9–11] relied on a particular embedding into ℓ^1.

2.1.1 What is the metric measure space $(\mathscr{S}, \mu)$?

The construction of the limiting subset $\mathscr{S}$ does not give us a good intuition for what it looks like. This problem is separate from embedding $\mathscr{S}$ into an infinite dimensional Banach space ℓ^1. In [11], Aldous give us a way to interpret this space using a (random) continuous function $f(t)$ which we will shortly describe.

To get a preliminary sense of the geometry of $\mathscr{S}$, we will first show that $\mathscr{S}$ is "tree-like" in the sense that between every two points $x, y \in \mathscr{S}$ there is a unique path $[[x, y]] \subset \mathscr{S}$ containing both x and y which is isometric to a closed interval of size $\|x - y\|_{\ell^1}$. Note that in the pre-limit spaces $\mathscr{S}_t$ of (2.1) every point $x \in \mathscr{S}_t$ is of the form

$$x = (x_1, x_2, \cdots, x_m, 0, \cdots) \in \ell^1, \qquad x_i \geq 0.$$

This is by construction using the Poisson point process $(C_j; j \geq 1)$. Thus there is some special path, denoted $[[0, x]] \subset \mathscr{S}_t$, of length $\|x\|_{\ell^1}$ which connects 0 to $(x_1, 0, \cdots)$ to $(x_1, x_2, 0, \cdots)$ etc. Moreover, this path $[[0, x]]$ exists every element $x \in \ell^1$ and is always an isometric embedding of $[0, \|x\|_{\ell^1}] \hookrightarrow \ell^1$. For every $x \in \mathscr{S}$ the path $[[0, x]] \subset \mathscr{S}$ and $x = (x_1, x_2, \cdots)$ is such that $x_j \geq 0$ for all j. This is a feature of discrete trees. Namely, given any vertex $v \in \mathcal{T}$ a rooted connected tree, there exists a unique path $v_0 = \text{root}, v_1, \cdots, v_j = v$ where v_i and v_{i-1} share an edge for each $i = 1, 2, \cdots, j$.

Now we see that for each $x, y \in \mathscr{S}_t$, the paths $[[0, x]] \cap [[0, y]] = [[0, b(x, y)]]$ where $b(x, y)$

is the *branch point* between x and y defined by

$$b(x,y) = (x_1 \wedge y_1, x_2 \wedge y_2, \cdots).$$

Since b is continuous on $(x,y) \in \ell^1 \times \ell^1$ of the form $x = (x_i), y = (y_i)$ with $x_i, y_i \geq 0$ we get

$$[[0,x]] \cap [[0,y]] = [[0, b(x,y)]], \qquad \forall x, y \in \mathscr{S}.$$

Let $f_x : [0, \|x\|_{\ell^1}] \to \mathscr{S}$ be the unique isometry whose image is $[[0,x]]$. This can easily be seen to exist for $x_t \in \mathscr{S}_t$ and for all $x = \lim_t x_t$ as $\lim_t f_{x_t}(s \wedge \|x_t\|_{\ell_s})$. For each $a \in [[0,x]] \setminus \{0,x\}$ there is some time t_a such that $f_x(t_a) = a$ and define $[[a,x]] = f_x([t_1, \|x\|_{\ell^1}])$. This is simply a path of length $\|x - a\|_{\ell^1}$ lying completely in $\mathscr{S}$. Using branch points, we see that $[[x,y]] := [[b(x,y), x]] \cup [[b(x,y), y]]$ is image of a path from x to y lying completely in $\mathscr{S}$ and of length $\|x - y\|_{\ell^1}$. It is also not hard to see that this is the *unique* path that does this by using convergence in the Hausdorff topology in (2.1).

We will call a subset[1] $\mathscr{S} \subset \ell^1$ a *continuum tree* if between each pair of points $x, y \in \mathscr{S}$ there is a subset $[[x,y]] \subset \mathscr{S}$ containing both x and y isometric to to a closed interval of length $\|x - y\|_{\ell^1}$ and this is unique in the sense that given any continuous injection $f : [0,1] \to \mathscr{S}$ such that $f(0) = x$ and $f(1) = y$ then $f([0,1]) = [[x,y]]$.

2.1.2 Sampling the CRT

Aldous gives an alternative way to think about the tree $(\mathscr{S}, \mu)$ by repeatedly sampling elements of $\mathscr{S}$, according to the measure μ, and forming a family of subtrees of $\mathscr{S}$. More formally, conditionally given $(\mathscr{S}, \mu)$, we select i.i.d. random elements $V_1, V_2, \cdots \in \mathscr{S}$ with common law μ. Given just the first k of these elements $V_1, \cdots, V_k$, we can create a tree with k leaves where each edge in this tree also has an edge length. That is, define the tree

$$r(\mathscr{S}; \{V_1, \cdots, V_k\}) := \bigcup_{j=1}^{k} [[0, V_j]] \subset \mathscr{S}.$$

What does this subset look like? When $k = 1$, $r(\mathscr{S}; \{V_1\})$ is simply isometric to a closed line segment of length $\|V_1\|_{\ell^1}$. When $k = 2$, then

$$r(\mathscr{S}; \{V_1, V_2\}) = [[0, b(V_1, V_2)]] \cup [[b(V_1, V_2), V_1]] \cup [[b(V_1, V_2), V_2]].$$

[1]We will shortly remove the condition that it is a subset of ℓ^1.

Moreover, each of those intervals intersect at a single point $b(V_1, V_2)$. That is $r(\mathscr{S}; \{V_1, V_2\})$ is isometric to three line segments of lengths $\|b(V_1, V_2)\|_{\ell^1}$, $\|b(V_1, V_2) - V_1\|_{\ell^1}$ and $\|b(V_1, V_2) - V_2\|_{\ell^1}$ joined at a single point. That is $r(\mathscr{S}; \{V_1, V_2\})$ is a rooted tree (rooted at 0) with two leaves V_1, V_2 where each edge has an edge length. This continues to hold in general, the space $r(\mathscr{S}; \{V_1, \cdots, V_k\}$ is a rooted tree (rooted at 0) with k leaves where each edge has some length.

The tree $\mathscr{S}$ has the property that each internal node (i.e. one of the branch points that is not 0) in the tree $r(\mathscr{S}; \{V_1, \cdots, V_k\})$ is degree 3 (having two children) and the root may have 1 or two children. This can be see by the stick-breaking construction of $\mathscr{S}_t$ which has this property as well. Graph theoretically, Aldous called a rooted tree with k leaves, each internal node has 2 children and each edge has some associated length a *proper k-tree*. That is the tree $r(\mathscr{S}, \{V_1, \cdots, V_k\})$ is a random proper k-tree.

By exchangeability, if we fix any $k \geq 1$ and any $j < k$ and we select any random subset $i_1, i_2, \cdots, i_j$ of distinct elements of $[k]$ then

$$r(\mathscr{S}; \{V_1, \cdots, V_j\}) \stackrel{d}{=} r(\mathscr{S}; \{V_{i_1}, V_{i_2}, \cdots, V_{i_j}\})$$

because $(V_1, \cdots, V_j) \stackrel{d}{=} (V_{i_1}, \cdots, V_{i_j})$. That is a sub-sampling way of generating $r(\mathscr{S}; \{V_1, \cdots, V_j\})$ from $r(\mathscr{S}; \{V_1, \cdots, V_k\})$ for $k > j$. This gives us some consistency of the family of proper k-trees $(r(\mathscr{S}; \{V_1, \cdots, V_k\}); k \geq 1)$. We say that sequence of random proper k trees $(\mathscr{R}(k); k \geq 1)$ is *consistent* if for each $k \geq j \geq 1$ given $\mathscr{R}(k)$ we select j distinct leaves $L_1^k, L_2^k, \cdots, L_j^k$ uniformly among all such choices then the proper j-tree $r(\mathscr{R}(k); \{L_1^k, \cdots, L_j^k\})$, obtained from $\mathscr{R}(k)$ by keeping just the root and each path going from the root to a leaf L_i^k for $i \in [j]$, satisfies

$$r(\mathscr{R}(k); \{L_1^k, \cdots, L_j^k\}) \stackrel{d}{=} \mathscr{R}(j). \tag{2.2}$$

In some sense consistent families of random proper k-trees $(\mathscr{R}(k); k \geq 1)$ are (almost) equivalent to the existence of a continuum random tree (CRT for short) $\mathscr{S} \subset \ell^1$ equipped with a probability measure μ such that $r(\mathscr{S}; \{V_1, \cdots, V_k\}) \stackrel{d}{=} \mathscr{R}(k)$ as proper k-trees. The parenthetical almost is included because there is an additional and more technical condition that we have to put on both $(\mathscr{R}(k); k \geq 1)$ and the pair $(\mathscr{S}, \mu)$. For a more precise statement of this equivalence and its proof, see Theorem 3 in [11].

2.1.3 Encoding the CRT

Now let us add some additional structure to the consistent family of random proper k-trees $(\mathscr{R}(k); k \geq 1) = (r(\mathscr{S}; \{V_1, \cdots, V_k\}); k \geq 1)$. Namely, we wish to embed (non-isometrically) $\mathscr{R}(k)$ into the upper half-plane for each $k \geq 1$ which gives $\mathscr{R}(k)$ and additional ordering structure, i.e. there is some way to distinguish which child was born first. Using an independent and exchangeable ordering of $\{1, 2, \cdots, n\}$ along with the consistency of $\mathscr{R}(k)$ as proper k-trees, there is some way of making (2.2) hold in the sense of rooted planar proper k-trees. The precise details of this construction here is not important, but the important thing is that their is some representation of the random rooted proper k-trees $\mathscr{R}(k)$ as planar trees with edge lengths.

We can encode this metric space structure of any rooted planar tree $\mathcal{T}$ using the contour process $f_{\mathcal{T}} : \mathbb{R}_+ \to \mathbb{R}_+$. Namely, we start with the root, ρ, at time $t = 0$ and set $f_{\mathcal{T}}(0) = 0$. Suppose that at time t, we have defined the values of $f_{\mathcal{T}}(0), \ldots, f_{\mathcal{T}}(t)$ and we are currently at vertex u. At time $t + 1$, we go to the left-most child of u that has not already been explored (if such a vertex exists). Else we go to the parent of the vertex u which must exist if $u \neq \rho$. Call this new vertex v and set $f_{\mathcal{T}}(t + 1) = d(\rho, v)$. If we have finished exploring all the children of u then we set $f_{\mathcal{T}}(s) = 0$ for all $s \geq t$. Lastly, linearly interpolate[2].

Each edge $\{u, v\}$ in $\mathcal{T}$ is "explored" twice in this process. Once when going from parent to child and second when going from child to parent. The respective increments of the function $f_{\mathcal{T}}$ are identical and so this clearly encodes the metric structure of $\mathcal{T}$. This does not naturally encode its mass structure however. If $\mathcal{T}$ has precisely n vertices, then $\mathcal{T}$ has $n - 1$ edges and so $f_{\mathcal{T}}(t) = 0$ for all $t \geq 2n - 2$.

A more important feature of this contour process $f_{\mathcal{T}}$ is the following. Let u, v be vertices in $\mathcal{T}$ and let $b(u, v)$ denote their most recent common ancestor - the discrete analog of the branch point defined for the CRTs $\mathscr{S}$ above. Suppose that t is the first time we explore u and s is the first time we explore v, then

$$d(\rho, b(u, v)) = \min\{f_{\mathcal{T}}(r) : s \wedge t \leq r \leq s \vee t\}. \tag{2.3}$$

[2] We are not following the convention used by Aldous in [11] which is just a minor change. Aldous starts at $t = 1$ instead of $t = 0$, but this affects nothing.

This is simply because $b(u,v)$ is the vertex closest to ρ that is visited while traveling between u and v according to the ordering on the tree and that distance is capture precisely by the right-hand-side above. In particular,

$$d(u,v) = f_{\mathcal{T}}(s) + f_{\mathcal{T}}(t) - 2 \min_{s\wedge t \le r \le s\vee t} f_{\mathcal{T}}(r).$$

How does this order structure and (2.3) relate back to the CRTs $(\mathscr{S}, \mu)$ in (2.1)? With Theorem 13 of [11], Aldous provides a way to take certain continuous function $f : [0,1] \to \mathbb{R}_+$ and construct a function $F : [0,1] \to \ell^1$ such that $\mathscr{T}_f := F([0,1])$ is a continuum tree and

$$d_{\ell^1}(F(s), F(t)) = f(t) + f(s) - 2 \min_{s\le r\le t} f(r), \qquad s \le t. \tag{2.4}$$

The conditions on the continuous function f are that (1) $f(t) = 0$ for at most 1 $t \in (0,1)$; (2) the strict local minimas of f are dense; (3) if $t_1 < t_2$ are strict local minima with $f(t_1) = f(t_2)$ then there is some $s \in (t_1, t_2)$ with $f(s) < f(t_1)$; and (4) the collection of one-sided minmas have Lebesgue measure zero. In this case we say that f *encodes the tree* $\mathscr{T}_f$. Also, the function F induces a measure μ_f on $\mathscr{T}_f$ given by $\mu_f := F_\# \operatorname{Leb}|_{[0,1]}$. With Theorem 15 in [11], Aldous shows that under some additional hypothesis involving a partial order that the continuum random tree $(\mathscr{S}, \mu)$ has a representation as a random tree encoded by a random function f. Lastly, with Corollary 22 in [11], Aldous shows that the random encoding function is equal in law to twice a standard Brownian excursion. At almost the same time Le Gall [130] provides essentially the same encoding result as Corollary 22 in [11] without any reference to embedding the tree encoded by twice a standard Brownian excursion into ℓ^1.

Theorem 2.1.2 (Aldous [11], Le Gall [130])**.** *Let $e = (e(t); t \in [0,1])$ be a standard Brownian excursion and $\mathscr{T}_{2e} \subset \ell^1$ be the CRT encoded by $2e$. The limit space $(\mathscr{S}, \mu)$ in (2.1) is equal in law to $(\mathscr{T}_{2e}, \mu_{2e})$.*

By looking at (2.4) and Theorem 2.1.2 we see that we can recover certain properties of the mass structure of $\mathscr{S}$ using a standard Brownian excursion. For example,

$$\left(\mu\left(v \in \ell^1 : d(0,v) \le x\right); x \ge 0\right) \overset{d}{=} \left(\int_0^1 1_{[2e(t)\le x]}\, dt; x \ge 0\right). \tag{2.5}$$

2.1.4 Universality of The CRT

The motivation for looking at Algorithm 1 was that $\mathcal{T}_n$ was constructed from a Poisson(c) branching process conditioned on having total progeny n. Since the conditioning does not depend on the parameter c, we will only consider the case where $c = 1$ (i.e. a critical condition). More generally, we can consider $\mathcal{T}_n$ to be a Galton-Waston branching tree with offspring distribution ξ conditioned on having exactly n vertices. This conditioning will make sense for all n sufficiently large provided that $\gcd\{j : \mathbb{P}(\xi = j) > 0\} = 1$. We will call ξ *critical* if $\mathbb{E}[\xi] = 1$ and we will further suppose that $\sigma^2 = \mathrm{Var}(\xi) \in (0, \infty)$.

Just as the large n behavior of a simple random walk with step distribution $(\xi - 1)$ depends asymptotically only on σ by Donsker's theorem, the large n behavior of $\mathcal{T}_n$ depends only on σ. Namely, with Theorem 23 in [11] (see also item 1 in the Technical Remarks afterwards) Aldous shows the following.

Theorem 2.1.3 (Aldous [11]). *Let $\mathcal{T}_n$ be a Galton-Watson branching tree with offspring distribution ξ satisfying $\mathbb{E}[\xi] = 1$, $0 < \sigma^2 = \mathrm{Var}(\xi) < \infty$ and $\gcd\{j : \mathbb{P}(\xi = j) > 0\} = 1$. Then there exists a representation of $\mathcal{T}_n$ as a finite subset of ℓ^1, call this $\mathscr{S}_n$, equipped with a uniform measure μ_n, such that jointly as $n \to \infty$*

$$\mathrm{scale}_{\sigma n^{-1/2}} \mathscr{S}_n \overset{d}{\Longrightarrow} \mathscr{S}, \qquad (\mathrm{scale}_{\sigma n^{-1/2}})_{\#}\mu_n \overset{d}{\Longrightarrow} \mu,$$

where $(\mathscr{S}, \mu)$ is the limit as in (2.1) and the topology is as in Theorem 2.1.1

Note the scaling in space is by $\sigma n^{-1/2}$ and note $(\sigma n)^{-1/2}$ as appears in the central limit theorem. This correct scaling was observed by Aldous by appeal to a particular example $(\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = \frac{1}{2})$ and a universality result.

2.1.5 Height profiles, part 1

Now that have the convergence of the conditioned Galton-Watson branching trees $\mathcal{T}_n$ converging after rescaling towards the tree $\mathscr{T}_{2e}$ by Theorems 2.1.2 and 2.1.3 it seems quite natural that the height profile of $\mathcal{T}_n$ will have a scaling limit as well. Namely, let $Z_n = (Z_n(h); h \geq 0)$ be defined by

$$Z_n(h) = \#\{v \in \mathcal{T}_n : d(\rho, v) = h\}, \qquad \rho \in T_n \text{ is the root.}$$

By looking at the scaling present in Theorem 2.1.3 we want to scale the edges of $\mathcal{T}_n$ by $\sigma n^{-1/2}$ so distance t from the root in $\mathscr{S}$ should be roughly look like distance $\frac{\sqrt{n}}{\sigma}t$ in the Galton-Watson tree $\mathcal{T}_n$. Moreover, the rescaled convergence in Theorem 2.1.3 of $\mathcal{T}_n$ towards a *compact* limit suggests that $\mathcal{T}_n$ is actually of radius $O_{\mathbb{P}}(\sqrt{n})$ (and this is indeed true, see [11] and references therein). Thus the sequence of random processes $\left(n^{-1/2}Z_n(\lfloor\sigma^{-1}n^{1/2}t\rfloor); t\geq 0\right)$ should be a tight-sequence in $\mathbb{D}(\mathbb{R}_+)$ and have a scaling limit described in terms of either a Brownian excursion or, equivalently, the Brownian CRT $(\mathscr{S},\mu)$. This was a conjecture by Aldous in [10]; however, the weak convergence arguments described above are too weak to establish this. It was established using techniques from analytic combinatorics in [67].

2.2 Technical Tools

We will now spend some time flushing out some of the technical tools developed since Aldous' work mentioned above.

2.2.1 Gromov-Hausdorff-Prohorov Distance between metric spaces

Informally speaking, if we want to understand what the path of a Brownian motion $B = (B(t); t\in[0,1])$ looks like on $\mathbb{R}^2$ we may think of running a simple random walk on the scaled integer lattice $\delta\mathbb{Z}^2$ run for time $c\delta^{-2}$. This would also work if we replace $\mathbb{Z}^2$ with any lattice $\Lambda\subset\mathbb{R}^2$ with only the constant c changing. This is the case because (1) Donsker's theorem tells us that a simple random walk on $\delta\Lambda$ run at speed δ^{-2} converges to a (multiple) of Brownian motion but (2) in some imprecise (for now) way the scaled lattice $\delta\Lambda$ is close to $\mathbb{R}^2$ for δ small. This last point appears to have originated in the work of Edwards [83] concerning approximating the (possibly infinite) dimensional configuration space of a physical system by a smooth finite-dimensional Riemannian manifold. We will focus on a different formulation due shortly afterwards to Gromov [100].

2.2.1.1 Hausdorff Distance

Let $\mathcal{X}$ be a metric space and let $A, B \subset \mathcal{X}$ be two closed subsets. The *Hausdorff*[3] *distance* between A and B is

$$d_{\mathrm{H}}(A,B) := \inf\{\varepsilon > 0 : A \subset B^{\varepsilon}, B \subset A^{\varepsilon}\}, \quad U^{\varepsilon} := \bigcup_{x \in U} B_{\varepsilon}(x).$$

where $B_{\varepsilon}(x) = \{y \in \mathcal{X} : d(x,y) \leq \varepsilon\}$ is the closed ball centered at x of radius ε. It is well-known that d_H is a metric on the collection of all closed subsets of $\mathcal{X}$. Let $F(\mathcal{X})$ denote the collection of all closed subsets of $F(\mathcal{X})$. The metric space $F(\mathcal{X})$ inherits many properties from $\mathcal{X}$: $F(\mathcal{X})$ is complete if $\mathcal{X}$ is complete; $F(\mathcal{X})$ is separable if $\mathcal{X}$ is separable; $F(\mathcal{X})$ is compact if $\mathcal{X}$ is compact; $F(\mathcal{X})$ is bounded if $\mathcal{X}$ is bounded. Some of these statements are proved in [47, Section 7.3] while the rest are easy to demonstrate. It is also easy to see that $\mathcal{X} \hookrightarrow F(\mathcal{X})$ embeds isometrically into $F(\mathcal{X})$ and so those prior *if*'s are actually *iff*'s.

2.2.1.2 Compact Gromov-Hausdorff Distance

The idea of Gromov in [100] was to use the Hausdorff distance to compare two metric spaces $\mathcal{X}$ and $\mathcal{Y}$ which need not be subsets of the same metric space. Thinking back to the construction of Aldous described in Section 2.1, we considered *rooted* trees $\mathcal{T}_n$ viewed as metric spaces. That is we will instead deal with *pointed metric spaces* which, for now, we assume are compact. That is $\mathcal{X} = (\mathcal{X}, d_{\mathcal{X}}, \mathsf{o}_{\mathcal{X}})$ is a compact metric space together with a distinguished point $\mathsf{o}_{\mathcal{X}} \in \mathcal{X}$, and similarly for $\mathcal{Y}$.

We will say $\mathcal{X}$ and $\mathcal{Y}$ are isometric as compact pointed metric spaces if there exists a bijective isometry $f : \mathcal{X} \to \mathcal{Y}$ such that $f(\mathsf{o}_{\mathcal{X}}) = \mathsf{o}_{\mathcal{Y}}$. We will say that f is an isomorphism of pointed metric spaces. We will let $\mathfrak{N}_c$ denote the collection of all isometry classes of compact metric spaces. We use the symbol $\mathfrak{N}$ to signify that these spaces are equipped with the *n*ull measure.

The idea of Gromov was to embed $\mathcal{X}$ and $\mathcal{Y}$ into a separate metric space $\mathcal{Z}$ isometrically and then measure the Hausdorff distance between their respective images as well as the

[3] This distance has also been called the Pompeiu–Hausdorff distance due to its origins in the work of Pompeiu; however, this terminology is not as common. See the short historical account [36] for more information.

distance between their roots. That is

$$d_{\mathrm{GH},c}(\mathcal{X},\mathcal{Y}) = \inf\left\{d_{\mathrm{H}}(f(\mathcal{X}),g(\mathcal{Y})) + d_{\mathcal{Z}}(f(\mathsf{o}_{\mathcal{X}}),g(\mathsf{o}_{\mathcal{Y}}))\right\},$$

where the infimum is taken over all isometric embeddings $f:\mathcal{X}\to\mathcal{Z}$ and $g:\mathcal{Y}\to\mathcal{Z}$ for a metric space $\mathcal{Z}=(\mathcal{Z},d_{\mathcal{Z}})$. We call the metric $d_{\mathrm{GH},c}$ the compact Gromov-Hausdorff (c-GH or compact GH for short) distance. Some references (including Gromov's work [100]) in the literature take the maximum of the two quantities as opposed to the sum. This makes a difference only in terms of constants (mostly in terms of certain Lipschitz functions into or out of the space $\mathfrak{N}_c$) because the two such metrics are equivalent. Because of this difference, we will state certain lemmas just in terms of some Lipschitz constant L as opposed to specifying precisely what L is.

In order to have weak convergence of elements in $\mathfrak{N}_c$ we need to know that $\mathfrak{N}_c$ is (topologically) a Polish space. The function $d_{\mathrm{GH},c}$ is a metric on $\mathfrak{N}_c$ and turns $\mathfrak{N}_c$ into a Polish space.

Theorem 2.2.1. *The space $(\mathfrak{N}_c, d_{\mathrm{GH},c})$ is a Polish space.*

A proof of the above theorem can be found in [2] and [121]. In particular, see Section 3.4 in [121].

2.2.1.3 Prohorov Distance

If we wish to incorporate information on the measure of, say, $\mathcal{T}_n$ then we will need some way to compare how close two measures are with some metric. When two measures are defined on the same space, there are several ways to do this; however, we will focus on the Prohorov metric.

Let $\mathcal{X}$ be a Polish space and let $\mathscr{M}_f(\mathcal{X})$ denote the collection of finite (positive) Borel measures on $\mathcal{X}$. The Prohorov distance between two measures $\mu,\nu\in\mathscr{M}_f(\mathcal{X})$ is defined by

$$d_{\mathrm{P}}(\mu,\nu) = \inf\left\{\varepsilon>0 : \mu(A)\le\nu(A^{\varepsilon})+\varepsilon, \nu(A)\le\mu(A^{\varepsilon})+\varepsilon,\ \forall A\subset\mathcal{X}\text{ Borel}\right\}.$$

It is well-known that if $\mu_n,\mu\in\mathscr{M}_f(\mathcal{X})$ and $d_{\mathrm{P}}(\mu_n,\mu)\to 0$ then for all bounded continuous functions $f:\mathcal{X}\to\mathbb{R}$

$$\int_{\mathcal{X}} f\,d\mu_n \to \int_{\mathcal{X}} f\,d\mu.$$

That is, the Prohorov metric metrizes the topology of weak convergence of finite measures on $\mathcal{X}$. See Chapter 1, Section 6 in [35] for the case of probability measures and Section 4.1 in [111] in general.

2.2.1.4 Compact Gromov-Hausdorff-Prohorov Distance

We recall that with Aldous's original construction described Section 2.1, Aldous shows that a uniformly chosen rooted random labeled trees $\mathcal{T}_n$ viewed as a finite metric space equipped $n^{-1/2}$ times the graph distance and also equipped with the uniform measure has some embedding as a metric measure space into ℓ^1. Once embedded into ℓ^1, Aldous shows that this embedding converges weakly in the Hausdorff topology of compact subsets of ℓ^1 and its measure converges weakly with respect to the Prohorov metric on $\mathscr{M}_f(\ell^1)$. The compact Gromov-Hausdorff-Prohorov (c-GHP or compact GHP) topology is a succinct way of describing this topology without an explicit embedding into ℓ^1 or any other Polish space.

We will say that a four-tuple $\mathcal{X} = (\mathcal{X}, d_\mathcal{X}, \mathsf{o}_\mathcal{X}, \mu_\mathcal{X})$ is a compact pointed measured metric space (c-PMM space) if $(\mathcal{X}, d_\mathcal{X}, \mathsf{o}_\mathcal{X}) \in \mathfrak{N}_c$ and $\mu_\mathcal{X} \in \mathscr{M}_f(\mathcal{X})$. We will say that two c-PMM spaces $\mathcal{X}$ and $\mathcal{Y}$ are isomorphic if there exists a bijective isometry $f : \mathcal{X} \to \mathcal{Y}$ such that f is an isomorphism of pointed metric spaces and $f_\# \mu_\mathcal{X} = \mu_\mathcal{Y}$. We let $\mathfrak{M}_c$ denote the collection of isomorphism classes of c-PMM spaces.

The compact Gromov-Hausdorff-Prohorov distance between two c-PMM spaces $\mathcal{X}$ and $\mathcal{Y}$ is

$$d_{\mathrm{GHP},c}(\mathcal{X}, \mathcal{Y}) = \inf \left\{ d_\mathrm{H}(f(\mathcal{X}), g(\mathcal{X})) + d_\mathcal{Z}(f(\mathsf{o}_\mathcal{X}), g(\mathsf{o}_\mathcal{Y})) + d_\mathrm{P}(f_\# \mu_\mathcal{X}, g_\# \nu_\mathcal{Y}) \right\},$$

where the infimum is taken over all isometric embeddings $f : \mathcal{X} \to \mathcal{Z}$ and $g : \mathcal{Y} \to \mathcal{Z}$ into a comment metric space $\mathcal{Z}$. Again, some authors take the maximum as opposed to the sum in the definition, c.f. [121] and [2]. This makes only a trivial difference, most importantly for us in terms of Lipschitz constants for certain maps. The following theorem is proved in [2] and [121].

Theorem 2.2.2. *The space* $(\mathfrak{M}_c, d_{\mathrm{GHP},c})$ *is a Polish space.*

2.2.1.5 The Gromov-Hausdorff-Prohorov Topology

Up to now we have been comparing the distances between two compact pointed measured metric spaces. It turns out that sometimes it is useful to work with a way to measure the distance between two non-compact metric spaces. This was first introduced in [2] for the case of *length spaces* where the distance between two points x and y in a set $\mathcal{X}$ is given by

$$d(x, y) = \inf L(\gamma)$$

where $L(\gamma)$ is the length of the continuous curve $\gamma : [0, 1] \to \mathcal{X}$ with $\gamma(0) = x$ and $\gamma(1) = y$. We follow the presentation in [121].

We will say that a quadruple $(\mathcal{X}, d_{\mathcal{X}}, \mathsf{o}_{\mathcal{X}}, \mu_{\mathcal{X}})$ is a *pointed metric measure* (PMM) space if each of the following hold

1. $(\mathcal{X}, d_{\mathcal{X}})$ is a metric space such that every bounded set has compact closure in $\mathcal{X}$ (i.e. $(\mathcal{X}, d_{\mathcal{X}})$ has the Heine-Borel property);

2. $\mu_{\mathcal{X}}$ is a Borel measure on $\mathcal{X}$ such that $\mu_{\mathcal{X}}$ assigns finite mass to bounded sets;

3. $\mathsf{o}_{\mathcal{X}} \in \mathcal{X}$.

Note that each PMM space $\mathcal{X}$ is Polish but not every Polish space is a PMM space according to this definition.

We will say that two PMM spaces $\mathcal{X}$ and $\mathcal{Y}$ are isomoprhic if there exists an isometry $f : \mathcal{X} \to \mathcal{Y}$ such that $f(\mathsf{o}_{\mathcal{X}}) = \mathsf{o}_{\mathcal{Y}}$ and $f_{\#}\mu_{\mathcal{X}} = \mu_{\mathcal{Y}}$. We will say that f is an isomoprhism of PMM spaces or a PMM-isomorphism for short. The collection of all isometry classes PMM spaces will be denoted by $\mathfrak{M}$.

We say that $\mathcal{X}$ is a PMM-subspace of $\mathcal{Y}$ if there exists some isometric embedding[4] $\iota : \mathcal{X} \to \mathcal{Y}$ such that $\iota(\mathsf{o}_{\mathcal{X}}) = \mathsf{o}_{\mathcal{Y}}$ and $\iota_{\#}\mu_{\mathcal{X}} \leq \mu_{\mathcal{Y}}$. We denote this by $\mathcal{X} \preceq \mathcal{Y}$. There are particularly important PMM subspaces of $\mathcal{X}$ denoted by $\mathcal{X}^{(r)}$. The space $\mathcal{X}^{(r)}$ is defined as

[4] This is slighlty different convention between what Khezeli defines in [121] wherein the author requires that $\mathcal{X}$ be a subset of $\mathcal{Y}$ and not isometrically embedding into $\mathcal{Y}$. Since we deal with isomorphism classes of PMM spaces this makes no significant difference.

a set by

$$\mathcal{X}^{(r)} = \bar{B}_r(\mathsf{o}_\mathcal{X}) := \{x \in \mathcal{X} : d_\mathcal{X}(x, \mathsf{o}_\mathcal{X}) \le r\}$$

as the closed ball of radius r centered at $\mathsf{o}_\mathcal{X}$. The space $\mathcal{X}^{(r)}$ is turned into a PMM space by equipping it with the same metric as $\mathcal{X}$ and with the measure $\mu_{\mathcal{X}^{(r)}}(A) = \mu_\mathcal{X}(A \cap \mathcal{X}^{(r)})$.

In [121], the author defines the function $a_\varepsilon(\mathcal{X}, \mathcal{Y})$ on $\mathfrak{M} \times \mathfrak{M}$ for each $\varepsilon \in (0, 1]$ by

$$a_\varepsilon(\mathcal{X}, \mathcal{Y}) = \inf \left\{ d_{\mathrm{GHP},c} \left(\mathcal{X}^{(\varepsilon^{-1})}, \mathcal{Y}' \right) : \mathcal{Y}^{(\varepsilon^{-1}-\varepsilon)} \preceq \mathcal{Y}' \preceq \mathcal{Y} \right\}.$$

In words a_ε informally measures how close the compact PMM space $\mathcal{X}^{(\varepsilon^{-1})}$ is to a small perturbation of $\mathcal{Y}^{(\varepsilon^{-1}-\varepsilon)}$. The Gromov-Hausdorff-Prohorov metric (GHP metric for short) is defined in [121] by

$$d_{\mathrm{GHP}}(\mathcal{X}, \mathcal{Y}) = \inf \{\varepsilon \in (0, 1] : a_\varepsilon(\mathcal{X}, \mathcal{Y}) \vee a_\varepsilon(\mathcal{Y}, \mathcal{X}) \le \varepsilon/2\}$$

where $\inf \emptyset = 1$. The following theorem is proved in [121].

Theorem 2.2.3. *The space* $(\mathfrak{M}, d_{\mathrm{GHP}})$ *is Polish.*

2.2.1.6 Gromov-Weak topology and Gromov-Vague topology

Before describing the Gromov-weak topology, we will work briefly by analogy. Consider what it takes to prove convergence in the Skorohod space $\mathbb{D}([0, 1])$ of, say, $X_n \overset{d}{\Longrightarrow} X$ where one can not easily appeal to a well-known results on convergence of these particular stochastic processes. The first thing one might try is to demonstrate that (1) the finite dimensional distributions X_n converge towards the finite dimensional distributions of X and that (2) the sequence X_n is tight. Many times proving (1) is easier than proving (2). The Gromov-weak topology, introduced in [98], is a way to formalize this notion of convergence of finite-dimensional distributions of random metric spaces where the metric spaces also differ.

Recall that above in Section 2.1.2 we described a way to sample finite-dimensional distributions in the Brownian CRT $\mathscr{S}$ to form a family of proper k-tree $(\mathscr{R}(k); k \ge 1)$. The information that the proper k-trees captures are the distances between different leaves and the root in the trees - that is it captures information about how the metric behaves on $\binom{k+1}{2}$ many pairs of vertices. In the final paragraph therein we briefly mentioned that Aldous shows in Theorem 3 in [11] that, in some sense, consistent families of proper k-trees are equivalent to a CRT $(\mathscr{S}, \mu)$ (not necessarily the Brownian CRT) where $\mathscr{S} \subset \ell^1$. This is a very helpful for identifying certain limits of random trees or describing properties of known CRTs; however, this relies quite heavily on the being able to embed the CRTs into ℓ^1. The additional conditions that Aldous puts on the consistent families (which we have not discussed) allows for a particular embedding into ℓ^1.

One way to overcome the problem of embedding a tree into ℓ^1 was to instead use the compact GHP or compact GH topologies discussed above, or even the (not compact) GHP topology if we do not know the limiting space is compact. While there have been many tools introduced over the decades since Aldous's work in [9–11] that simplify the work of demonstrating weak convergence in these three topologies these results are still sometimes difficult to use in practice.

In [98], the authors introduced a separate topology on the space of isomorphism classes of Polish spaces equipped with probability measures (without a specified point and with a different notion of isomorphism) which they call the Gromov-weak topology. These spaces were extensively studied by Gromov [101]. This topology was slightly modified in [136] to include the specification of a root in the metric space and finite measures as opposed to probability measures. Finally, in [21] the authors extended this topology to the space collection of pointed Polish spaces equipped with a locally finite measure, that is, a measure which assigns finite mass to bounded sets. Focusing on the latter point, let $\mathcal{X} = (\mathcal{X}, d_\mathcal{X}, \mathsf{o}_\mathcal{X}, \mu_\mathcal{X})$ be a quadruple such that $(\mathcal{X}, d_\mathcal{X})$ is a Polish space, $\mathsf{o}_\mathcal{X} \in \mathcal{X}$ and $\mu_\mathcal{X}$ is a Borel measure on $\mathcal{X}$ which assigns finite mass to bounded sets, i.e. $\mu_\mathcal{X}(\bar{B}_r(x)) < \infty$ for all $r > 0$ and $x \in \mathcal{X}$.

To describe the isomorphism classes, we recall that the support of a measure μ defined on a topological space $\mathcal{Z}$ is defined as

$$\operatorname{spt}(\mu) = \operatorname{spt}(\mu; \mathcal{Z}) = \mathcal{Z} \setminus \{x \in \mathcal{Z} : \exists x \in U \subset \mathcal{Z} \text{ open s.t. } \mu(U) = 0\}.$$

That is $x \notin \operatorname{spt}(\mu)$ if and only if there exists some open $U \subset \mathcal{Z}$ containing x such that $\mu(U) = 0$. We will say that two pointed Polish spaces with boundedly finite measures $\mathcal{X}$ and $\mathcal{Y}$ are weak-isomorphic if there exists a $f : \operatorname{spt}(\mu_\mathcal{X}) \cup \{\mathsf{o}_\mathcal{X}\} \to \operatorname{spt}(\mu_\mathcal{Y}) \cup \{\mathsf{o}_\mathcal{Y}\}$ such that $f_\# \mu_\mathcal{X} = \mu_\mathcal{Y}$ and $f(\mathsf{o}_\mathcal{X}) = \mathsf{o}_\mathcal{Y}$. We will call such an f a weak-isomorphism. In particular, $(\mathcal{X}, d_\mathcal{X}, \mathsf{o}_\mathcal{X}, \mu_\mathcal{X})$ is weak-isomorphic to $(\operatorname{spt}(\mu_\mathcal{X}), d_\mathcal{X}, \mathsf{o}_\mathcal{X} \mu_\mathcal{X})$. We let $\mathfrak{X}$ denote the collection of weak-isomorphism classes of pointed Polish spaces with boundedly finite measures. We let $\mathfrak{X}_f$ denote the collection of weak-isomormpshism classes of pointed Polish space with finite measures. Also let $\mathfrak{X}_{\text{loc}}$ denote the collection of weak-isomorphism classes of pointed Polish spaces with boundedly finite measures and such that there is some representative of that isomorphism class, say $\mathcal{X}$, such that bounded sets have compact closure.

Given a measure $\mu = \mu_{\mathcal{X}}$ on $\mathcal{X}$, we let $\mu^{\otimes m}$ denote the m-fold product measure on $\mathcal{X}^m$. For each $m = 0, 1, \cdots$ let

$$\mathbb{R}_+^{\binom{m+1}{2}} = \{(r_{i,j}; 0 \le i < j \le m) : r_{i,j} \in \mathbb{R}_+\}.$$

Given an element $\mathcal{X} \in \mathfrak{X}_f$ and an $m = 0, 1, \cdots$ we define $\nu_m(\mathcal{X}; -)$ as the measure on $\mathbb{R}_+^{\binom{m+1}{2}}$ by

$$\nu_m\left(\mathcal{X}; \prod_{0 \le i < j \le m} A_{i,j}\right) = \int_{\mathcal{X}^n} \prod_{0 \le i < j \le m} 1_{[d_{\mathcal{X}}(x_i, x_j) \in A_{i,j}]} \mu_{\mathcal{X}}^{\otimes m}(dx_1, dx_2, \cdots, dx_m)$$

where $x_0 := \mathsf{o}_{\mathcal{X}}$. The following theorem is proved in [98] (see also Section 2.1 in [136]).

Theorem 2.2.4. *Define the Gromov-weak topology on the space $\mathfrak{X}_f$ by saying $\mathcal{X}_n \to \mathcal{X}$ if and only if for all $m = 0, 1, \cdots$ the finite measures on $\mathbb{R}_+^{\binom{m+1}{2}}$ $\nu_m(\mathcal{X}_n; -) \Longrightarrow \nu_m(\mathcal{X}; -)$ in the sense of weak convergence of finite measures.*

The Gromov-weak topology is metrized by the metric

$$d_{\text{GP},f}(\mathcal{X}, \mathcal{Y}) = \inf\{d_{\text{P}}(f_{\#}\mu_{\mathcal{X}}, g_{\#}\mu_{\mathcal{Y}}) + d_{\mathcal{Z}}(f(\mathsf{o}_{\mathcal{X}}), g(\mathsf{o}_{\mathcal{Y}}))\}$$

where the infimum is taken over all isometric embeddings $f : \text{spt}(\mu_{\mathcal{X}}) \cup \{\mathsf{o}_{\mathcal{X}}\} \to \mathcal{Z}$ and $g : \text{spt}(\mu_{\mathcal{Y}}) \cup \{\mathsf{o}_{\mathcal{Y}}\} \to \mathcal{Z}$.

The space $\mathfrak{X}_f$ equipped with Gromov-weak topology is Polish, i.e. there exists a metric which metrizes the Gromov-weak topology and makes $\mathfrak{X}_f$ complete and separable.

We remark that the metric defined above is called the pointed Gromov-Prohorov distance in [136] wherein the authors take the infimum over metric spaces $\mathcal{Z}$ which are obtained via equipping $\mathcal{X} \sqcup \mathcal{Y}$ with (pseudo)-metrics which agree with $d_{\mathcal{X}}$ on $\text{spt}(\mu_{\mathcal{X}}) \cup \{\mathsf{o}_{\mathcal{X}}\}$ and $d_{\mathcal{Y}}$ on $\text{spt}(\mu_{\mathcal{Y}}) \cup \{\mathsf{o}_{\mathcal{Y}}\}$ and the maps f (resp. g) are the canonical embeddings and then taking a quotient to make the pseudo-metric a bonafide metric. It is not too difficult to see why these two notions are equivalent.

In [21], the authors described some localization procedure. Namely, let $\mathcal{X} \in \mathfrak{X}$ and for each $r \geq 0$ let $\mathcal{X}^{(r)} \in \mathfrak{X}_f$ be defined as the weak-isomrophism class of

$$\left(\mathcal{X}, d_{\mathcal{X}}, \circ_{\mathcal{X}}, \mu_{\mathcal{X}}\left(\cdot \cap \bar{B}_r(\circ_{\mathcal{X}})\right)\right).$$

Note that we need note replace $\mathcal{X}$ with $\bar{B}_r(\circ_{\mathcal{X}})$ as we did for the localization procedure for the GHP topology because the support of $\mu(\cdot \cap F)$ for any closed set F is contained inside F.

The metric d_{GP} on $\mathfrak{X}$ (denoted by $d^{\#}_{\mathrm{GP}}$ in [21]) is defined by

$$d_{\mathrm{GP}}(\mathcal{X}, \mathcal{Y}) = \int_0^\infty e^{-r}\left(1 \wedge d_{\mathrm{GP},c}(\mathcal{X}^{(r)}, \mathcal{Y}^{(r)}\right) dr.$$

It is shown in [21] that d_{GP} is a metric on $\mathfrak{X}$ and the topology it generates is called the *Gromov-vague* topology. The following result, which we state as a theorem here, is proved in Section 4 of [21].

Theorem 2.2.5. *The space $(\mathfrak{X}, d_{\mathrm{GP}})$ is Polish. With this topology, $\mathfrak{X}_{\mathrm{loc}} \subset \mathfrak{X}$ is a Borel subset and $(\mathfrak{X}_{\mathrm{loc}}, d_{\mathrm{GP}})$ is Lusin but not Polish.*

Let us finish with a remark that there is a gap between the Gromov-vague and the Gromov-weak topologies on $\mathfrak{X}_f$. Let $\mathcal{X}_n = (\mathbb{R}_+, d, 0, \mu_n)$ where d is the standard Euclidean metric and $\mu_n = \mathrm{Leb}\,|_{[0,1]} + \frac{1}{n}\,\mathrm{Leb}\,|_{[0,n]}$. Also let $\mathcal{X} = (\mathbb{R}_+, d, 0, \mu)$ where $\mu = \mathrm{Leb}\,|_{[0,1]}$.

Note that for any metric space $\mathcal{Z}$ such that $\mathrm{spt}(\mu_n)$, $\mathrm{spt}(\mu)$ embed into via f_n and f respectively, must have that

$$d_{\mathrm{P}}((f_n)_{\#}\mu_n, f_{\#}\mu) = 1,$$

because the two measures have different total masses. So $\liminf d_{\mathrm{GP},f}(\mathcal{X}_n, \mathcal{X}) \geq 1$ as $n \to \infty$; that is X_n does not converge to $\mathcal{X}$ in the Gromov-weak topology.

However, for each n and each $r \geq 0$ note that

$$d_{\mathrm{P}}(\mu_n(\cdot \cap [0,r]), \mu(\cdot \cap [0,r]) \leq \frac{r}{n}.$$

That means that

$$d_{\mathrm{GP}}(\mathcal{X}_n, \mathcal{X}) \leq \int_0^\infty e^{-r}(1 \wedge r/n)\, dr \longrightarrow 0,$$

by dominated convergence. Hence $\mathcal{X}_n \to \mathcal{X}$ in the Gromov-vague topology.

2.2.2 Discrete Tree, Discrete Graphs and Random Models

In this section we will describe several encodings of discrete trees and discrete graphs as metric measure spaces. Because we care about these discrete spaces at metric measure spaces as opposed to viewed as combinatorial objects we will deal with particular embeddings in the upper half-plane that are natural for describing scaling limits of random trees.

2.2.2.1 Planar Trees

We mostly follow the approach in [131]. Consider the set of labels

$$\mathcal{U} := \bigcup_{n=0}^{\infty} \mathbb{N}^n, \qquad \mathbb{N} = \{1, 2, \cdots\}$$

and $\mathbb{N}^0 := \{\mathtt{o}\}$. An element $u = (u_1, \cdots, u_n) \in \mathbb{N}^n \subset \mathcal{U}$ is said to be a height n and we will often abbreviate and write $u_1 u_2 \cdots u_n$. We will denote the height of $u \in \mathcal{U}$ as $\mathbf{hgt}(u)$. Given an elements $u = u_1 \cdots u_n$ and $j \in \mathbb{N}$ we write $uj = u_1 \cdots u_n j$ as the concatenation. We say that elements uj are *children* of u and u is the *parent* of uj. For each $u \in \mathcal{U} \setminus \{\mathtt{o}\}$ we denote its parent by $\pi(u)$. This is with the convention that $\mathtt{o}j = j$ and $\pi(j) = \mathtt{o}$.

A rooted plane tree $\mathtt{t}$ is a finite subset of $\mathcal{U}$ which satisfies the following three conditions

1. $\mathtt{o} \in \mathtt{t}$;
2. If $u \in \mathtt{t} \setminus \{\mathtt{o}\}$ then $\pi(u) \in \mathtt{t}$;
3. If $u \in \mathtt{t}$ then there exists a number $\chi_{\mathtt{t}}(u) \in \{0, 1, \cdots\}$ such that $uj \in \mathtt{t}$ if and only if $j \le \chi_{\mathtt{t}}(u)$.

We call $\mathtt{o} \in \mathtt{t}$ the root. The set $\mathtt{t}$ is viewed as a graphical tree by saying that u, v share an edge if and only if $u = \pi(v)$ or $v = \pi(u)$. We will implicitly view $\mathtt{t}$ as a metric space with the graph metric and as a measure space with the counting measure (which is not a probability measure!).

We could have also allowed $\mathtt{t}$ to be an infinite set without much difficulty. The space $\mathtt{t}$ would still be locally compact (provided we keep assumption 3) and have a locally finite measure. However, several of the combinatorial identities we describe later require finiteness of the set $\mathtt{t}$ in order to be a bijective relationship.

Partial orderings on $\mathcal{U}$ For our purposes there are two main total orderings on $\mathcal{U}$ which are vitally important because they allow us to encode certain structure results about the random tree (and, in fact, random graph) models.

The first ordering is often called the lexicographical ordering of $\mathcal{U}$, but what we will call the *depth-first ordering*. We will denote this by $\stackrel{\text{DF}}{\prec}$. For example,

$$\mathtt{o} \stackrel{\text{DF}}{\prec} 1 \stackrel{\text{DF}}{\prec} 12 \stackrel{\text{DF}}{\prec} 13 \stackrel{\text{DF}}{\prec} 2314 \stackrel{\text{DF}}{\prec} 3.$$

Given a tree $\mathtt{t}$ with $\#\mathtt{t} = n$ we will write $\mathtt{o} = u_0^{\text{DF}} \stackrel{\text{DF}}{\prec} u_1^{\text{DF}} \stackrel{\text{DF}}{\prec} \cdots \stackrel{\text{DF}}{\prec} u_{n-1}^{\text{DF}}$ the ordering of all elements of $\mathtt{t}$. If we allowed $\#\mathtt{t} = +\infty$ then one can construct examples where the depth-first ordering of all the $\mathtt{t}$ cannot be indexed by the natural numbers. This poses a problem for describing certain limits. As we will see, and has been apparent in the literature since [132, 133], the depth-first ordering is vitally important for describing (in terms of scaling limits) the metric structure of $\mathtt{t}$.

The second ordering is the *breadth-first ordering* which will be denoted by $\stackrel{\text{BF}}{\prec}$. This ordering is defined as follows. If $\mathbf{hgt}(u) < \mathbf{hgt}(v)$ then $u \stackrel{\text{BF}}{\prec} v$ while if $\mathbf{hgt}(u) = \mathbf{hgt}(v)$ then $u \stackrel{\text{BF}}{\prec} v$ if $u \stackrel{\text{DF}}{\prec} v$. The partial ordering of the elements above become

$$\mathtt{o} \stackrel{\text{BF}}{\prec} 1 \stackrel{\text{BF}}{\prec} 3 \stackrel{\text{BF}}{\prec} 12 \stackrel{\text{BF}}{\prec} 13 \stackrel{\text{BF}}{\prec} 2314.$$

Given a tree $\mathtt{t}$ with $\#\mathtt{t} = n$ we will write $\mathtt{o} = u_0^{\text{BF}} \stackrel{\text{BF}}{\prec} u_1^{\text{BF}} \stackrel{\text{BF}}{\prec} \cdots \stackrel{\text{BF}}{\prec} u_{n-1}^{\text{BF}}$ the ordering of all elements of $\mathtt{t}$ in a breadth-first manner. There is no trouble in indexing problem for the breadth-first ordering of trees $\mathtt{t}$ with infinitely many vertices provided that hypothesis 3 is kept. As we will see, and has been apparent in the literature since [49] (although the idea existed before, see e.g. [14]), the breadth-first ordering is vitally important for describing (in terms of scaling limits) the mass structure of $\mathtt{t}$.

2.2.2.2 Encodings of Planar Trees

We now describe a very simple way to encode the structure of random trees. This is the so-called Łukasiewicz path or Łukasiewicz walk.

Let us fix a tree $\mathtt{t}$ and recall the notatoin of $\chi_{\mathtt{t}}(u)$. Define the functions $X_{\mathtt{t}}^{\text{DF}} =$

$(X_{\mathtt{t}}^{\mathrm{DF}}(k); k = 0, 1, \cdots, \#\mathtt{t})$ and $X_{\mathtt{t}}^{\mathrm{BF}} = (X_{\mathtt{t}}^{\mathrm{BF}}(k) : k = 0, 1, \cdots, \#\mathtt{t})$ defined by

$$X_{\mathtt{t}}^{\mathrm{DF}}(k) = \sum_{i=0}^{k-1} \left(\chi_{\mathtt{t}}(u_i^{\mathrm{DF}}) - 1\right) \quad \text{and} \quad X_{\mathtt{t}}^{\mathrm{BF}}(k) = \sum_{i=0}^{k-1} \left(\chi_{\mathtt{t}}(u_i^{\mathrm{BF}}) - 1\right).$$

We call $X_{\mathtt{t}}^{\mathrm{DF}}$ (resp. $X_{\mathtt{t}}^{\mathrm{BF}}$) the depth-first (resp. breadth-first) Łukasiewicz path (or walk) of the tree $\mathtt{t}$.

Clearly, these two Łukasiewicz paths are uniquely defined by the tree $\mathtt{t}$. Also, $\chi_{\mathtt{t}}(u) \geq 0$ so $X_{\mathtt{t}}^{\mathrm{DF}}(k) - X_{\mathtt{t}}^{\mathrm{DF}}(k-1) \geq -1$, and likewise for the breadth-first Łukasiewicz walk. Such functions are often said to be *downward skip-free walks.* A simple counting argument also reveals that

$$X_{\mathtt{t}}^{\mathrm{DF}}(k) \geq 0, \qquad \text{if and only if} \quad k \in \{0, 1, \cdots, \#\mathtt{t} - 1\}$$

and $X_{\mathtt{t}}^{\mathrm{DF}}(\#\mathtt{t}) = -1$. This last fact is perhaps the easiest to see because $\sum_{j=0}^{\#\mathtt{t}-1} \chi_{\mathtt{t}}(u_j^{\mathrm{DF}})$ counts the number of children in the tree and there is only one vertex in $\mathtt{t}$ which is not a child, namely $\mathtt{o}$. We will call a function $f : \{0, 1, \cdots, n\} \to \mathbb{Z}$ such that $f(0) = 0$, $f(k) \geq 0$ if and only if $k < n$ and $f(n) = -1$ a *discrete excursion of duration n.* The following lemma is trivial, but can be found in e.g. [131].

Lemma 2.2.6. *Let $f : \{0, 1, \cdots, n\} \to \mathbb{Z}$ be a discrete excursion of duration n and suppose that f is also a downward skip-free. Then there exists a unique tree $\mathtt{t}^{\mathrm{DF}}$ (resp. $\mathtt{t}^{\mathrm{BF}}$) whose depth-first Łukasiewicz path (resp. breadth-first Łukasiewicz path) is f.*

There are other encodings that uniquely describe the metric space structure of tree $\mathtt{t}$. We will mention the *height process* of the tree, denoted by $H_{\mathtt{t}} = (H_{\mathtt{t}}(k); k \geq 0)$. Namely, we define

$$H_{\mathtt{t}}(k) = \mathbf{hgt}(u_k^{\mathrm{DF}}), \quad k \geq 0, \mathbf{hgt}(u_k^{\mathrm{DF}}) = 0 \forall k = 0, 1, \cdots, \#\mathtt{t} - 1.$$

Clearly, process encodes the metric structure of the tree that is quite useful for proving continuum limits.

A remarkable fact is that $H_{\mathtt{t}}$ can be constructed via the process $X_{\mathtt{t}}^{\mathrm{DF}}$ in some measurable way.

Lemma 2.2.7. *Let* $\mathtt{t}$ *be a rooted planar tree and let* $X_{\mathtt{t}}^{\rm DF}$ *be its depth-first Łukasiewicz path and let* $H_{\mathtt{t}}$ *be its height process. Then, for each* $n \in \{0, 1, \cdots, \#\mathtt{t} - 1\}$,

$$H_{\mathtt{t}}(n) = \# \left\{ k \in \{0, 1, \cdots, n-1\} : X_{\mathtt{t}}^{\rm DF}(k) = \min_{k \le j \le n} X_{\mathtt{t}}^{\rm DF}(j) \right\}. \tag{2.6}$$

While the proof of this fact is not that difficult, it is not terribly insightful nor does it play a vital role in some subsequent proof. Therefore, we do not include a proof. It is, however, vitally important for the construction of the continuum limits that we will eventually describe. For a proof we refer the reader to [133] or the survey [131]. This is called the exploration process in [133] and some subsequent works; however, the exploration process is now used to describe a different object.

Another useful quantity is the so-called contour function of the tree $\mathtt{t}$. This we will denote by $W_{\mathtt{t}} = (W_{\mathtt{t}}(t); t \in [0, 2(\#\mathtt{t} - 1)])$. This is defined as follows. Suppose that there are p many leaves (vertices with $\chi_{\mathtt{t}}(u) = 0$) in $\mathtt{t}$ and let $\ell^1, \cdots, \ell_p$ be leaves in depth-first order. For any leaves ℓ and ℓ' the path from $\mathtt{o}$ to ℓ and the path from $\mathtt{o}$ to ℓ' will be identical until some branching point $b_{\mathtt{t}}(\ell, \ell')$. More formally, suppose the unique path from $\mathtt{o}$ to ℓ is $\mathtt{o} = v_0, v_1, v_2, \cdots, v_i = \ell$ and the unique path from $\mathtt{o}$ to ℓ' is $\mathtt{o} = u_0, u_1, \ldots, u_j = \ell'$. There is some maximal $r < i \wedge j$ such that $u_k = v_k$ if and only if $k \le r$. The branch point $b_{\mathtt{t}}(\ell, \ell') = u_r = v_r$. In terms of genealogy, $b_{\mathtt{t}}(\ell, \ell')$ is the last common ancestor of ℓ and ℓ'. The contour function $W_{\mathtt{t}}$ is the unique piece-wise linear function with $W_{\mathtt{t}}(0) = 0$ and slopes of ± 1 whose consecutive local extremes occur at at the values

$$\mathbf{hgt}(\ell^1),\ \mathbf{hgt}(b_{\mathtt{t}}(\ell^1, \ell_2)),\ \mathbf{hgt}(\ell_2)), \cdots, \mathbf{hgt}(b_{\mathtt{t}}(\ell_{p-1}, \ell_p), \quad \mathbf{hgt}(\ell_p), 0$$

An important result that we will describe informally is the following. If $\gamma_n \to \infty$ but $\gamma_n/n \to 0$ and $\frac{\gamma_n}{n} H_{\mathtt{t}}(\lfloor n \cdot \rfloor) \to h$ for some *continuous* function h as the size of the tree $\mathtt{t} = \mathtt{t}_n$ as size $n \to \infty$, then $\frac{\gamma_n}{n} W_{\mathtt{t}}(\lfloor 2n \cdot \rfloor) \to h$ as well. That is for large trees $\mathtt{t}$ the contour process and the height process roughly differ by just a time-change. See Remark 3.2 in [71] and Section 1.6 in [131].

2.2.2.3 The Height Profile

We recall that the motivation in this work for studying random trees and random graphs was to understand disease propagation. Through the connection with graphs, the number of people infected on day $t = 0, 1, \cdots$ was simply the number of vertices at distance t from the collection of vertices $\{\rho_1, \cdots, \rho_k\}$ representing the first k infected individuals. When just a single individual is infected on a planar tree, the number of people infected is called the *height profile* $Z_{\mathtt{t}} = (Z_{\mathtt{t}}(h); h = 0, 1, \cdots)$ defined by

$$Z_{\mathtt{t}}(h) = \#\{v \in \mathtt{t} : \mathbf{hgt}(v) = h\} = \#\{0 \le k \le \#\mathtt{t} - 1 : H_{\mathtt{t}}(k) = h\}.$$

An identity realized in [14] in a context of so called $\mathbf{p}$-trees, but later used in [49, 50] to classify continuous state branching processes with immigration and their multitype generalizations, is the discrete Lamperti transform.

Lemma 2.2.8. *The height profile $Z_{\mathtt{t}}$ solves the discrete difference equation*

$$Z_{\mathtt{t}}(h+1) = Z_{\mathtt{t}}(0) + X_{\mathtt{t}}^{\mathrm{BF}} \circ C_{\mathtt{t}}(h), \qquad C_{\mathtt{t}}(h) = \sum_{\ell=0}^{h} Z_{\mathtt{t}}(\ell),$$

with initial condition $Z_{\mathtt{t}}(0) = 1$.

Proof. Note that $Z_{\mathtt{t}}(h+1) = C_{\mathtt{t}}(h+1) - C_{\mathtt{t}}(h)$. Moreover, every vertex except for $\mathtt{o}$ of height at most $h+1$ has a parent of height at most h. Similarly, all vertices at height at most h are labelled by $u_0^{\mathrm{BF}}, \cdots, u_{C_{\mathtt{t}}(h)-1}^{\mathrm{BF}}$. Thus

$$\begin{aligned} C_{\mathtt{t}}(h+1) &= 1 + \sum_{v \in \mathtt{t}:\mathbf{hgt}(v) \le h} \chi_{\mathtt{t}}(v) = 1 + \sum_{j=0}^{C_{\mathtt{t}}(h)-1} \chi_{\mathtt{t}}(u_j^{\mathrm{BF}}) \\ &= 1 + \sum_{j=0}^{C_{\mathtt{t}}(h)-1} \left(\chi_{\mathtt{t}}(u_j^{\mathrm{BF}}) \; - 1\right) + C_{\mathtt{t}}(h) \\ &= 1 + X_{\mathtt{t}}^{\mathrm{BF}} \circ C_{\mathtt{t}}(h) + C_{\mathtt{t}}(h). \end{aligned}$$

This proves the desired claim. □

2.2.2.4 Forests

Much of the above caries directly over to studying planar rooted forests made up of k many trees rooted planar trees where the k many trees are also ordered. We can view a forest $\mathtt{f}$

constructed from k trees $\mathtt{t}_j$ with a rooted planar tree by setting

$$\mathtt{f} = \{\mathtt{o}\} \cup \bigcup_{j=1}^{k} \{ju : u \in \mathtt{t}_j\}.$$

We make a slight alteration to the definition of the ordering of vertices in $\mathtt{f}$ and the height. Namely, $\mathtt{o}$ is omitted from both the depth-first and breadth-first labelings of the vertices in $\mathtt{f}$ so $1 = 1\mathtt{o} \in \{1u : u \in \mathtt{t}_1\} = u_0^{\mathrm{DF}} = u_0^{\mathrm{BF}}$. We also define the height of a vertex $ju \in \mathtt{f}$ as the height of $u \in \mathtt{t}_j$.

We observe that these slight changes result in minor alterations to the analysis done above for trees; however we will not spell all of these out here. The things to note are that the height processes of a forest is just the concatenation of the height processes of the trees, the depth-first Łukasiewicz paths of the forest are just the concatenation of the depth-first Łukasiewicz paths of the trees. The height profile of $\mathtt{f}$ is just the sum of the height profiles of the trees. The breadth-first Łukasiewicz path of $\mathtt{f}$ is not a simple functional of the breadth-first Łukasiewicz paths of the trees $\mathtt{t}_j$ - the increments become interlaced. However, often the probabilistic models used to generate random forests have an underlying exchangeable probabilistic structure allowing one to conclude that $X_{\mathtt{f}}^{\mathrm{DF}} \overset{d}{=} X_{\mathtt{f}}^{\mathrm{BF}}$. Importantly, Lemma 2.2.8 (and its proof) carries over with only a change in the initial condition.

With this convention, we can allow an *infinite* collection of trees $(\mathtt{t}_i; i \geq 1)$ where the resulting depth-first constructions carry over with no problem.

2.2.2.5 Moving to graphs

The above descriptions work perfectly fine for the definition of planar trees $\mathtt{t}$ given above. However, not all random tree models will naturally fit into this category and certainly graphs will not, in general, be trees. The reader may therefore wonder how do we bridge this gap in the case of graphs. We state the description used by Spencer [166] although similar ideas were used by Aldous at almost the same time in a different context [12]. Before continuing we recall the definition of surplus of a connected graph G. The surplus $s(G)$ is defined by

$$s(G) = \#E(G) - \#V(G) + 1.$$

Trees are the only connected graphs of surplus 0.

Let us focus on the restricted situation of consider a labelled connected graph G on the vertices $\{1, \cdots, n\}$. We view this graph as rooted at the vertex 1 and we perform a breadth-first search on this graph and generate a breadth-first spanning tree. We do this as follows. Maintain two ordered lists. An active stack, $\mathcal{A}$, of discovered vertices where at time 0 the active stack $\mathcal{A} = (1)$ and a dead stack, $\mathcal{D}$ which is initially empty. At each time $t = 0, 1, \cdots$ label the vertex x at the top of the stack by v_t^{BF}. Order the neighbors of v_t^{BF} in $[n] \setminus (\mathcal{A} \cup \mathcal{D})$ from smallest label to largest label and add these vertices to the bottom of the stack $\mathcal{A}$ in that order, and call these vertices the (breadth-first) children of v_t^{BF}. Lastly, remove v_t^{BF} from $\mathcal{A}$ and place it in $\mathcal{D}$.

Let us consider a particular graph on $n = 7$ vertices. The edges in the graph are 13, 15, 16, 27, 34, 35, and 37. This is a connected graph on 7 vertices and has 7 edges, i.e. its a graph of surplus $s(G) = 1$. The stacks evolve as follows:

$t = 0$: $\mathcal{A} = (1)$, $\mathcal{D} = ()$ and breadth-first children vertices are 3, 5, and 6;

$t = 1$: $\mathcal{A} = (3, 5, 6)$, $\mathcal{D} = (1)$ and breadth-first children vertices are 4 and 7;

$t = 2$: $\mathcal{A} = (5, 6, 4, 7)$, $\mathcal{D} = (1, 3)$ and there are no breadth-first children;

$t = 3$: $\mathcal{A} = (6, 4, 7)$, $\mathcal{D} = (1, 3, 5)$ and there are no breadth-first children;

$t = 4$: $\mathcal{A} = (4, 7)$, $\mathcal{D} = (1, 3, 5, 6)$ and there are no breadth-first children vertices;

$t = 5$: $\mathcal{A} = (7)$, $\mathcal{D} = (1, 3, 5, 6, 4)$ and breadth-first children vertices just 2;

$t = 6$: $\mathcal{A} = (2)$, $\mathcal{D} = (1, 3, 5, 6, 4, 7)$ and there are no breadth-first children vertices;

$t = 7$: $\mathcal{A} = (0)$ and $\mathcal{D}$ is length n.

Given G we can do the above procedure to get a rooted tree $T^{\mathrm{BF}} = T_G^{\mathrm{BF}} \subset G$ which comes with some breadth-first ordering. Let $\chi^{\mathrm{BF}}(x)$ denote the number of breadth-first children of the vertex x obtained in the above procedure. Let $X_G^{\mathrm{BF}} = (X_G^{\mathrm{BF}}(k); k = 0, 1, \cdots, n)$ be defined by

$$X_G^{\mathrm{BF}}(k) = \sum_{j=0}^{k-1} \left(\chi^{\mathrm{BF}}(v_j^{\mathrm{BF}}) - 1\right). \tag{2.7}$$

This is easily seen to be a downward skip-free excursion of duration n and hence uniquely characterizes a planar tree $\mathtt{t}$ by Lemma 2.2.6; however, this will not play an important role for us now. Note however, that $X_G^{\mathrm{BF}}(k)$ is 1 less than the size of the active stack at time $t = k$. Indeed, this is true for $t = 0$ and the increments $\chi^{\mathrm{BF}}(v_j^{\mathrm{BF}}) - 1$ is simply the change in the size of the active stack at each step.

There is, of course, an analogous depth-first search of the graph G which results in a depth-first walk X_G^{BF} defined in an analogous way to (2.7). We do not spell all the details, but the main difference is that the vertices in $[n] \setminus (\mathcal{A} \cup \mathcal{D})$ discovered at each step are added to the *top* of the stack $\mathcal{A}$ as opposed to the bottom.

Spencer in [166] gives a way to count how many graphs G with surplus $s(G) = k$ have the same breadth-first tree T^{BF}. Which we will state as a lemma.

Lemma 2.2.9. *Let T be a labeled tree on $[n]$ rooted at 1 and fix $k = 0, 1, \cdots$. The number of graphs G with surplus k with $T_G^{\mathrm{BF}} = T$ is $\binom{A^{\mathrm{BF}}(T)}{k}$ where $A^{\mathrm{BF}}(T)$ is*

$$A^{\mathrm{BF}}(T) = \sum_{j=0}^{n-1} X_T^{\mathrm{BF}}(j)$$

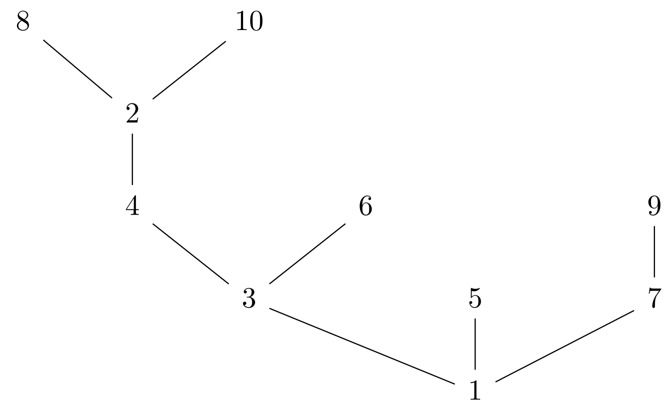
where X_T^{BF} is the walk described by (2.7).

The same formula holds with DF *replacing* BF *in every instance above.*

The idea of the proof is the following. As we look at the neighbors of vertex v_t^{BF} some of these will be in the set $\mathcal{A}$ and $\mathcal{D}$. Those in neighbors in $\mathcal{A}$ correspond to surplus edges that are newly discover at time t (call the *breadth-first back-edges* or *bf back-edges*) while neighbors in $\mathcal{D}$ are surplus edges discovered at some previous time. Since the stack $\mathcal{A}$ at time t has $X_G^{\mathrm{BF}}(t) + 1$ many vertices, one of which is v_t^{BF}, there are $X_T^{\mathrm{BF}}(t)$ many possible bf back-edges. A similar idea works for the depth-first walk which allows us to count the possible depth-first back-edges (df back-edges).

Let us explain how this works with a particular example. The tree T has exactly 10 vertices.

$T =$



$\cdot$ (2.8)

The Łukasiewicz paths are graphed in Figure 2.1. One can easily compute $A^{\text{BF}}(T) = 13$ and $A^{\text{DF}}(T) = 18$. The possible back-edges that can be included in a graph G with $T_G^{\text{BF}} = T$ or $T_G^{\text{DF}} = T$ are

$$\text{bf back-edges} = \{35, 37, 57, 54, 56, 74, 76, 46, 49, 69, 62, 92, 8(10)\}$$

$$\text{df back-edges} = \left\{ \begin{array}{l} 35, 37, 46, 45, 47, 26, 25, 27, 8(10), 86, \\ 85, 87, (10)6, (10)5, (10)7, 65, 67, 57 \end{array} \right\}.$$

Above the back-edge ab appears before cd if a appears before c in the relevant exploration or $a = c$ and b appears before d in the relevant exploration.

Relating metric structure of G and T_G^{DF} We will now describe briefly the idea behind encoding the geometry of G using T_G^{DF} which is described in [5] as well as the companion paper [4].

Let's suppose that we have a labeled tree T rooted at 1 and labeled by $[n]$. Suppose also G is a rooted labeled graph uniformly chosen among all graphs with $T_G^{\text{DF}} = T$ and $s(G) = k$. We know that there are $\binom{A^{\text{DF}}(T)}{k}$. If T is as in (2.8) the there are $\binom{18}{k}$ such graphs. In general, the number of such surplus k graphs is also precisely the number sets of points in $\mathbb{Z}_+^2$ where $\mathbb{Z}_+ := \{0, 1, \cdots\}$ that lie strictly below the graph of the function X_T^{DF} and have size $s(G) = 2$. Let $\mathcal{Q} \subset \mathbb{Z}_+^2$ be such a set, i.e.

$$\mathcal{Q} \subset \left\{(i, j) \in \mathbb{Z}_+^2 : 0 \leq j < X_T^{\text{DF}}(i)\right\}, \qquad \#\mathcal{Q} = k. \tag{2.9}$$

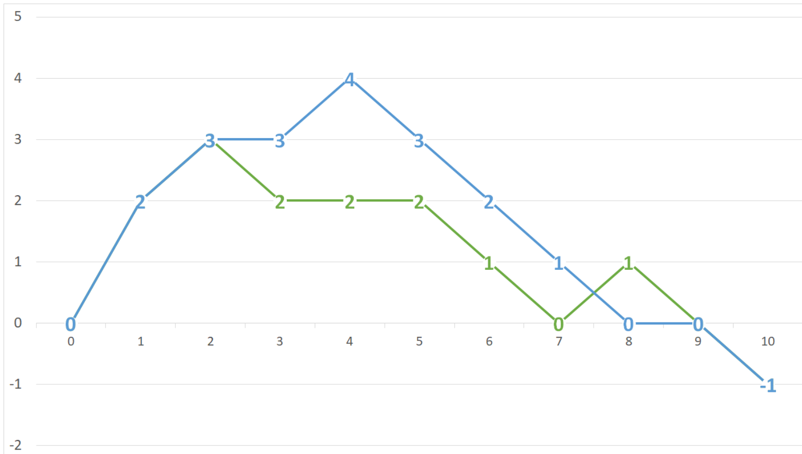


Figure 2.1: The two Łukasiewicz paths (with linear interpolation) for the example tree in (2.8). The depth-first walk is in blue and the breadth-first walk is in green.

Let $\mathcal{Q}$ be a uniformly chosen such set. We generate $G = G^{\mathrm{DF}}(T, \mathcal{Q})$ from T as follows. For each $(i,j) \in \mathcal{Q}$, let $\ell = \min\{t > i : X_T^{\mathrm{DF}}(t) = j\}$. Add an edge between the vertices v_ℓ^{DF} and v_i^{DF}. This gives a bijection between points sets $\mathcal{Q}$ with the above restrictions and labeled rooted connected graphs G with surplus $s(G) = k$ and with $T_G^{\mathrm{DF}} = T$.

We can do the exact same procedure with the breadth-first walks as well; however, this is much less common in the literature for technical reasons. See the discussion in [147] for more details.

Graphs not labeled by $[n]$ We note that the previous discussion constructed graphs which were labeled by $[n]$. However, the construction of G from a Łukasiewicz walk and a point set $\mathcal{Q} \subset \mathbb{Z}_+^2$ via $G^{\mathrm{DF}}(T, \mathcal{Q})$ works in general. For several random graph models there is a "natural" way to explore the connected components of a graph and build a random Łukasiewicz path X^{DF} or X^{BF} and the point set $\mathcal{Q}$ simultaneously.

2.2.2.6 Random Models

Gatlon-Watson Trees Fix a probability measure μ on the non-negative integers $\mathbb{N}_0$. A Galton-Watson branching process $Z = (Z(t); t \geq 0)$ is a time-homogeneous Markov chain such that

$$(Z(t+1)|Z(t)) \stackrel{d}{=} \sum_{j=1}^{Z(t)} \xi_j, \qquad \xi_j \stackrel{i.i.d.}{\sim} \mu.$$

There is a natural genealogical interpretation of this process using a random rooted planar tree $\mathcal{T}$. Namely, we start from a single individuals o who is the only individual alive at time $t = 0$. This individual gives birth ξ_{o} many children, where $\xi_{\mathsf{o}} \sim \mu$. These children will be labeled by $1, 2, \cdots, \xi_{\mathsf{o}}$ together they will constitute generation 1. Subsequently, each individual u alive in generation t gives birth to an independent number of children labeled uj for $1 \leq j \leq \xi_u$ where $\xi_u \sim \mu$. This continues until there are no individuals are present at generation t, i.e. whenever $Z(t)$ is absorbed at zero.

It is easy to see that the resulting collection $\mathcal{T}$ of individuals forms a rooted planar tree if and only if the associated branching process Z is eventually absorbed at zero. It is well known, [20], that this occurs a.s. if and only if $\sum_k k\mu(k) \leq 1$ and $\mu(1) < 1$. We will say that a measure μ with $\sum k\mu(k) < 1$ (resp. $= 1, > 1$) is *subcritical* (resp. *critical*, *supercritical*). Given a (sub)critical distribution μ, the random rooted planar tree T constructed above will be called a *Galton-Watson tree* with *offspring distribution* μ and we will write $\mathcal{T} \sim \mathrm{GW}(\mu)$.

A more direct way to construct the tree $\mathcal{T} \sim \mathrm{GW}(\mu)$ is using (random) Łukasiewicz paths $X_{\mathcal{T}}^{\mathrm{DF}}$ or $X_{\mathcal{T}}^{\mathrm{BF}}$. That is, let $(\xi_j; j \geq 0)$ be i.i.d. random variables with common law μ. Since $(\xi_i - 1) \geq -1$ for all i, the random process $X = (X(t); t = 0, 1, \cdots, T_1)$

$$X(t) = \sum_{i=0}^{t-1} (\xi_i - 1), \qquad T_1 = \min\left\{ t : \sum_{j=0}^{t-1} (\xi_i - 1) = -1 \right\}.$$

is a downward skip-free discrete excursion of (random) duration T_1. Given such a random walk X and first hitting time T, we can generate a $\mathrm{GW}(\mu)$ tree $\mathcal{T}^{\mathrm{DF}}$ whose depth-first Łukasiewicz path is X. Similarly, we could have generated a $\mathcal{T}^{\mathrm{BF}}$ whose breadth-first Łukasiewicz path is the random walk X.

Trees with a Given Degree Sequence Let $\mathtt{t}$ be a random rooted planar tree and for each $i \geq 0$ let $\mathtt{n}_\mathtt{t}^i = \#\{v \in \mathtt{t} : \chi_\mathtt{t}(v) = i\}$. We call $\mathtt{n}_\mathtt{t} := (\mathtt{n}_\mathtt{t}^i; i \geq 0)$ the *degree sequence* of $\mathtt{t}$. It is easy to see that

$$\sum_{i\geq 0} \mathtt{n}_\mathtt{t}^i = \#\mathtt{t} = 1 + \sum_{i\geq 0} i\mathtt{n}_\mathtt{t}^i.$$

Indeed, the first sum counts all the vertices in $\mathtt{t}$ by counting the number of vertices with exactly i children and the second summation counts all the vertices who are the child of a vertex with exactly i children. Since $\mathtt{o}$ is not counted in the second sum, we must include the "1+" term. Conversely, given any sequence $N = (\mathtt{n}^i; i \geq 0)$ with $\sum_i \mathtt{n}^i = 1 + \sum_i i\mathtt{n}^i$ there exists a tree $\mathtt{t}$ with $\mathtt{n}_\mathtt{t} = \mathtt{n}$ as we will shortly see. We will call such sequences $\mathtt{n}$ *degree sequences.*

Given a degree sequence $\mathtt{n}$ define the $(\mathtt{c}_0 \geq \mathtt{c}_1 \geq \cdots \geq \mathtt{c}_{\mathtt{s}-1})$ where $\mathtt{s} = \sum_i \mathtt{n}^i$ uniquely by requiring that $\mathtt{n}^i = \#\{j : \mathtt{c}_j = i\}$. Clearly, $\sum_{i=0}^{t-1}(\mathtt{c}_i - 1) \geq 0$ for all $t < \mathtt{s}$ and $\sum_{i=0}^{\mathtt{s}-1}(\mathtt{c}_i - 1) = -1$ and so there exists a tree $\mathtt{t}$ with degree sequence $\mathtt{n}$.

Now given a measure $\mathcal{T} \sim \mathrm{GW}(\mu)$, we get

$$\mathbb{P}(\mathcal{T} = \mathtt{t}) = \prod_{u\in\mathtt{t}} \mu(\chi_\mathtt{t}(u)) = \prod_{i\geq 0} \mu(k)^{\mathtt{n}_\mathtt{t}^i}, \qquad \forall \mathtt{t}.$$

That is the probability that a Galton-Watson tree is equal to a given tree $\mathtt{t}$ depends *only* on degree sequence of the tree. Said in another way, conditionally on $\mathtt{n}_\mathcal{T} = \mathtt{n}$, a $\mathrm{GW}(\mu)$ tree $\mathcal{T}$ is uniform over all trees $\mathtt{t}$ with $\mathtt{n}_\mathtt{t} = \mathtt{n}$; i.e.

$$\left(\mathcal{T} \middle| \mathtt{n}_\mathcal{T} = \mathtt{n}\right) \overset{d}{=} \mathrm{Unif}\left\{\mathtt{t} : \mathtt{n}_\mathtt{t} = \mathtt{n}\right\}.$$

The natural next question is: What is the distribution of $\mathcal{T}$ distributed uniformly among all rooted planar trees $\mathtt{t}$ with $\mathtt{n}_\mathtt{t} = \mathtt{n}$? Let $\mathbb{T}$ denote the collection of a rooted planar trees and let $\mathbb{T}_\mathtt{n}$ denote the collection of all rooted planar trees with a given degree sequence. Let $\mathtt{c} = (\mathtt{c}_i; i \in [N])$ denote the child sequence associated with $\mathtt{n}$. We will describe how to generate a random tree $\mathcal{T} \sim \mathrm{Unif}(\mathbb{T}_\mathtt{n})$ using a simple combinatorial lemma whose proof can be found in [170] and [154, Section 6.1].

Lemma 2.2.10. *Let $(x_i; i = 0, 1, \cdots, n-1)$ be integers such that $x_i \geq -1$ and $\sum_{i=0}^{n-1} x_i =$*

$-k$ where $k \in \{1, 2, \cdots, n\}$. Then there are precisely k many $j \in \{0, 1, \cdots, n-1\}$ such that

$$\sum_{i=0}^{t-1} x_{i+j} > -k \qquad \forall t < n$$

where we interpret the subscript $i+j$ as $i+j \mod n$.

The cyclic shift lemma above allows us to easily construct a tree $\mathcal{T} \sim \text{Unif}(\mathbb{T}_{\mathtt{n}})$. Indeed, let $\mathtt{c} = (\mathtt{c}_i)$ be the child sequence associated with $\mathtt{n}$ and set $\mathtt{s} = \sum_i \mathtt{n}^i$. Generate $\pi \in \mathfrak{S}_{\mathtt{s}}$ be a uniformly distributed element in the symmetric group on the letters $\{0, 1, \cdots, \mathtt{s}-1\}$ and let $\xi_i = \mathtt{c}_{\pi(i)}$. Let τ be the cyclic shift $\tau(i) = i+1 \mod n$ and note that by Lemma 2.2.10 there exists a unique j such that $(\xi_{\tau^j(i)}; i = 0, 1 \cdots, \mathtt{s}-1)$ such that

$$X_{j,\pi}(t) = \sum_{i=0}^{t-1} (\xi_{\tau^j(i)} - 1),$$

is a downward skip-free discrete excursion of duration $\mathtt{s}$. Given this (random) downward skip-free discrete excursion we can let $\mathcal{T}$ be the tree whose depth-first (or breadth-first) Łukasiewicz path is $X_{j,\pi}$. This tree $\mathcal{T} \sim \text{Unif}(\mathbb{T}_{\mathtt{n}})$. See Lemma 7 in [46].

2.2.3 $\mathbb{R}$-trees and $\mathbb{R}$-graphs

2.2.3.1 $\mathbb{R}$-trees

An important feature of the Brownian CRT of Aldous described in Section 2.1 was that between every two points $x, y \in \mathscr{S}$ there existed a unique path $[[x, y]]$ between x and y. This is a feature shared by discrete trees as well and what is formalized with the notion of $\mathbb{R}$-trees or real trees.

In general, let $\mathcal{X}$ be a metric space. A segment in $\mathcal{X}$ is a set I isometric to an interval $[0, \ell]$ for some ℓ, i.e. $I = \varphi([0, \ell])$ for an isometry φ. The endpoints of I are $\varphi(0)$ and $\varphi(\ell)$.

A boundedly compact[5] metric space $\mathcal{T}$ is called an $\mathbb{R}$-tree (or a real tree) if for each $x, y \in \mathcal{T}$ there exists a segment $[[x, y]]$ with endpoints x and y such that for any continuous injective function $f : [0, 1] \to \mathcal{T}$ with $f(0) = x$ and $f(1) = y$ then $f([0, 1]) = [[x, y]]$. This

[5]The restriction to locally compact is not necessary in the definition. There is a well-developed theory of $\mathbb{R}$-trees which do not require this condition; however, we will only focus on these spaces because, for example, we are interested in the Gromov-Hausdorff(-Prohorov) convergence of these objects.

definition differs slightly from [87, Section 3.3] who does not require local compactness and states the second condition on the metric space in terms of unions of segments. However, these two definitions are equivalent for locally compact $\mathbb{R}$-trees. We will say that $\mathcal{T}$ is rooted if it is comes with additional specified point $\mathsf{o} \in \mathcal{T}$.

If $\varphi : [0, d(x, y)] \to \mathcal{T}$ is the isometric embedding of an interval to $[[x, y]]$ with $\varphi(0) = x$ and $\varphi(d(x, y)) = y$ it is convenient to write $[[x, y[[$, $]]x, y]]$, $]]x, y[[$ as the respective images of $[0, \ell), (0, \ell]$ and $(0, \ell)$ with $\ell := d(x, y)$ under map φ.

We will let $\mathfrak{T}^0 \subset \mathfrak{N}$ denote the collection (isomoporphism classes) of pointed $\mathbb{R}$-trees which are locally compact metric spaces and let $\mathfrak{T} \subset \mathfrak{M}$ denote the collection (isomorphism classes) of $\mathbb{R}$-trees which are PMM spaces. Similarly, define $\mathfrak{T}^0_c \subset \mathfrak{N}_c$ and $\mathfrak{T}_c \subset \mathfrak{M}_c$.

The following theorem is proved in [88] in the case of the compact GH topology, but the proof works in general. See also [2], Section 3.

Theorem 2.2.11. *The following spaces are Polish.*

1. $(\mathfrak{T}^0_c, d_{\mathrm{GH},c})$;
2. $(\mathfrak{T}^0, d_{\mathrm{GH}})$;
3. $(\mathfrak{T}_c, d_{\mathrm{GHP},c})$;
4. $(\mathfrak{T}, d_{\mathrm{GHP}})$.

There are several topological features of rooted $\mathbb{R}$-trees that we should mention before continuing. These are the collection of leafs in $\mathcal{T}$ and branch points in $\mathcal{T}$. The leafs of $\mathcal{T} = (\mathcal{T}, \mathsf{o})$, denoted by $\mathrm{Lf}(\mathcal{T})$, is defined by

$$\mathrm{Lf}(\mathcal{T}) = \{x \in \mathcal{T} \setminus \{\mathsf{o}\} : \mathcal{T} \setminus \{x\} \text{ is connected}\}.$$

The branch points of $\mathcal{T}$, denoted by $\mathrm{Br}(\mathcal{T})$, is defined by

$$\mathrm{Br}(\mathcal{T}) = \{x \in \mathcal{T} \setminus \{\mathsf{o}\} : \mathcal{T} \setminus \{x\} \text{ has at least 3 connected components}\}.$$

We also define the complement of $\mathrm{Lf}(\mathcal{T})$ as the *skeleton* of $\mathcal{T}$, denoted by $\mathrm{Sk}(\mathcal{T}) = \mathcal{T} \setminus \mathrm{Lf}(\mathcal{T})$. By the uniqueness of the segments for $\mathbb{R}$-trees, one can prove that $x \in \mathrm{Sk}(\mathcal{T})$ if $x \in [[\mathsf{o}, z[[$ for some $z \in \mathcal{T}$.

2.2.3.2 Coding ℝ-trees

A question that naturally arises while studying real trees (or discrete to continuum limits) is how do we represent the real trees? Broadly speaking there have been three main approaches to this problem: (1) a stick-breaking construction; (2) a consistent family of proper k-trees construction; and (3) a coding by a continuous function.

The last one is, in the author's opinion, the most natural construction and this is the one that we focus on in this section. The originates in the work of Aldous [11] on the Brownian CRT (see also [130]); however a similar idea appeared in the work of Le Gall [129] to study superprocesses. Since then, there have been a plethora of results that aid in proving (weak) convergence in the GHP or compact GHP topology. Namely, as we will see, the problem is reduced to proving (weak) convergence in the space $C(\mathbb{R}_+ \to \mathbb{R}_+)$.

We follow the treatment in [131]. Let $\mathscr{H}$ denote the collection of functions $h \in C_c(\mathbb{R}_+ \to \mathbb{R}_+)$ such that $h(0) = 0$. For such an $h \in \mathscr{H}$, define $\zeta(h) = \sup\{t : h(t) > 0\}$. The set $\mathscr{H}$ is easily seen to be a Borel subset of $C(\mathbb{R}_+)$.

Let a continuous function $h : [0, \zeta] \to \mathbb{R}_+$ where $\zeta = \zeta(h) > 0$ such that $h(\zeta_h) = 0$ and $h(0) = 0$.

Let us define a function $d_h : [0, \zeta]^2 \to \mathbb{R}_+$ define by

$$d_h(s,t) = h(s) + h(t) - 2 \inf_{r\in[s,t]} h(r), \qquad [s,t] = [t,s] \text{ if } s > t.$$

It is not difficult to see that for $s, t, u \in [0, \zeta]$ that

$$d_h(s,t) \le d_h(s,u) + d_h(u,t),$$

and so d_h is a psuedo-metric on $[0, \zeta]$.

We can define a quotient space $\mathcal{T}_h$ and a quotient map q_h by

$$q_h : [0,\zeta] \to \mathcal{T}_h, \qquad \mathcal{T}_h := [0,\zeta]/\sim$$

where $\sim$ is the smallest equivalence relationship that makes $s \sim t$ if $d_h(s,t) = 0$. The psuedo-metric d_h then because a bonafide metric on $\mathcal{T}_h$. Moreover, the map $q_h \to \mathcal{T}_h$ is a continuous map. Indeed, if $t_n \to t \in [0, \zeta]$ then $d_h(t_n, t) \to 0$ by the continuity of h. Hence $\mathcal{T}_h$ is a compact space. If we root $\mathcal{T}_h$ at $\mathsf{o} := q_h(0)$ then, in fact, $\mathcal{T}_h \in \mathfrak{T}_c^0$ is a compact

$\mathbb{R}$-tree. The proof of this fact is not terribly complicated so we refer the reader to Section 2 of [131] or Lemma 2.1 in [71].

Similarly, the tree $\mathcal{T}_h$ comes equipped with a natural measure $\mu_h := (q_h)_\# \operatorname{Leb}|_{[0,\zeta]}$ which turns $\mathcal{T}_h$ into a rooted real tree with a measure, i.e. $\mathcal{T}_h \in \mathfrak{T}_c$. We call $\mathcal{T}_h$ the tree *coded* by the function h.

A quite useful fact is the following result which we state as a theorem.

Theorem 2.2.12. *Let $h, g \in \mathscr{H}$ and let $\mathcal{T}_h, \mathcal{T}_g$ be the respective trees coded by h and g. Then, there exists a universal constant L such that*

$$\tilde{d}(\mathcal{T}_g, \mathcal{T}_h) \le L \left(\sup_{t\ge 0} |g(t) - h(t)| + |\zeta(g) - \zeta(h)| \right),$$

where $\tilde{d}$ is the compact GH, compact GHP, or GHP metric.

Consequently, the map $h \mapsto \mathcal{T}_h$ is measurable when $\mathscr{H}$ equipped with the trace Borel σ-algebra as a subset of $C(\mathbb{R}_+)$ and $\mathcal{T}_h$ is viewed in as an element in $\mathfrak{N}_c, \mathfrak{N}, \mathfrak{M}_c, \mathfrak{M}, \mathfrak{T}^0_c, \mathfrak{T}_c, \mathfrak{T}^0$, or $\mathfrak{T}$.

For a proof of the above theorem see [131] for the compact GH metric, [3] for the compact GHP metric and the GHP metric follows from comparison results between GHP metric and compact GHP metric for elements in $\mathfrak{M}_c$ in [121]. We mention that these various references allow one to explicitly write down a constant L that works; however, this does depend on the convention of taking maximums or sums in the definition of the compact GH or compact GHP metric. We've omitted the actual constant for brevity.

2.2.3.3 $\mathbb{R}$-graphs

A metric space $\mathcal{G} = (\mathcal{G}, d)$ is an $\mathbb{R}$-graph if locally it is an $\mathbb{R}$-tree. That is, given any $x \in \mathcal{G}$, there exists an $\varepsilon = \varepsilon(x) > 0$ such that $\{y \in \mathcal{G} : d(x, y) \le \varepsilon\}$ is an $\mathbb{R}$-tree. We will not describe in full detail all of the theory of $\mathbb{R}$-graphs, but will instead describe the construction of $\mathbb{R}$-graphs from two functions and point set. This follows the presentation in [30]. Before doing so, we refer the reader back to Section 2.2.2.5 for a discussion on constructing discrete graphs from excursions.

Let $g : [0, \zeta] \to \mathbb{R}_+$, $h : [0, \zeta] \to \mathbb{R}_+$ and let $\mathcal{Q} \subset \mathbb{R}_+^2$ be a discrete point set. Right now, each of these are deterministic. We suppose that h is continuous and that g is càdlàg and possesses no negative jumps. From the discussion in the previous section we can construction $\mathcal{T}_h$ as an $\mathbb{R}$-tree. Define the point set $\mathcal{Q} \cap g$ by

$$\mathcal{Q} \cap g = \{(t, y) \in \mathcal{Q} : 0 \le t \le \zeta, 0 \le y < g(t)\}.$$

Now suppose that $\#(\mathcal{Q} \cap g) = k \ge 0$ and let us enumerate the points by (t_i, y_i) for $i \in [k]$. For each $i \in [k]$, define s_i by

$$s_i = \min\{u > t_i : g(u) = y_i\}.$$

Recalling the discussion in Section 2.2.2.5 on the metric space structure constructed from T_G^{DF}, we added an edge between the vertices corresponding to the times t_i and s_i with the above notation. If we imagine the edge shrinking as the size of the graph grows large, then the distance between t_i and s_i with this newly added edge will shrink towards zero. That is, in the continuum we do not add an edge, but we instead quotient the space $\mathcal{T}_h$ so that the elements t_i and s_i are identified.

Namely, let $\sim$ be the smallest equivalence relationship on $\mathcal{T}_h$ such that for each $1 \le i \le k$ the pair of $q_h(t_i) \sim q_h(s_i)$. Here $q_h : [0, \zeta] \to \mathcal{T} = \mathcal{T}_h$ is the canonical quotient map. The quotient space

$$\mathcal{G} = \mathcal{G}(g, h, \mathcal{Q}) = \mathcal{T}/\sim$$

is an $\mathbb{R}$-graph. Of course, we can make $\mathcal{G}$ into a compact PMM space by rooting it at the image of $\mathsf{o} \in \mathcal{T}$ under the quotient and equipping it with the pushforward of measure $\mu_{\mathcal{T}}$ under the quotient map $\mathcal{T} \to \mathcal{G}$.

Observe that the quotient map $\mathcal{T} \to \mathcal{G}$ changes the metric structure of the space. Importantly, it can change the distance to the root, but it is easy to see that $d_{\mathcal{T}}(t_i, \mathsf{o}) \ge d_{\mathcal{G}}(t_i, \mathsf{o})$, where in both cases we identified t_i with its image in both $\mathcal{T}$ and $\mathcal{G}$. If we view $\mathcal{G}$ as capturing the history of the spread of a disease occurring in the continuum (as the discrete disease spread through discrete graphs), then the construction of $\mathcal{G}(g, h, \mathcal{Q})$ via the above method changes the cumulative number (more accurately mass) of individuals infected by time $t \geq 0$ for some values of t; however, in general we have the following inequality

$$\mu_{\mathcal{G}}(\bar{B}_t(\circ)) \geq \mu_{\mathcal{T}}(\bar{B}_t(\circ)), \qquad \forall t \geq 0.$$

2.2.4 Stochastic Processes

A key ingredient in describing continuum limits of random graphs and random trees was developed by Le Gall and Duquesne in [73], extending prior work of Le Gall and Le Jan [132, 133]. This is with the use of *height process* (sometimes called ψ-height processes) built up from a certain class of Lévy processes which do not jump downward. The processes appear as limits of downward skip free random walks, and since these walks never decrease by more than 1 their natural scaling limits do not jump downward.

2.2.4.1 Lévy processes

We will provide a brief overview of spectrally positive Lévy processes which appear as the scaling limits of downward skip free random walks with i.i.d. increments. More details and proofs of the statements below can be found in Chapter VII of Bertoin's monograph [26].

A (possibly killed) spectrally positive Lévy process $X = (X(t); t \geq 0)$ is a Lévy process which contains no negative jumps. Its Laplace transform exists and uniquely characterizes the process X. The Laplace transform must be of the form

$$\mathbb{E}\left[\exp(-\lambda X(t))\right] = \exp(t\Psi(\lambda)) \qquad \forall \lambda > 0. \tag{2.10}$$

Moreover, the function Ψ must be of the form

$$\Psi(\lambda) = -\kappa + \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} \left(e^{-\lambda r} - 1 + \lambda r 1_{[r<1]}\right) \pi(dr) \tag{2.11}$$

where $\kappa, \beta \geq 0$, $\alpha \in \mathbb{R}$ and π is a Radon measure on $(0,\infty)$ such that $\int_{(0,\infty)} (1 \wedge r^2)\, \pi(dr) < \infty$. Conversely, for each such Ψ, there exists a spectrally positive Lévy process with such a Laplace transform.

In the sequel we will restrict our attention to the case where $\kappa = 0$, where there is no killing of the Lévy process X at an exponential rate.

An important class of spectrally positive Lévy processes are subordinators. These are Lévy processes with non-decreasing sample paths. A subordinator Y has a Laplace transform of the form

$$\mathbb{E}\left[\exp(-\lambda Y(t))\right] = \exp(-t\Phi(\lambda)), \qquad \forall \lambda > 0, \tag{2.12}$$

where Φ is of the form

$$\Phi(\lambda) = \kappa' + \alpha' \lambda - \int_{(0,\infty)} (e^{-\lambda r} - 1)\, \nu(dr) \tag{2.13}$$

with $\kappa', \alpha' \geq 0$ and ν is a Radon measure with $\int_{(0,\infty)} (1 \wedge r)\, \nu(dr) < \infty$. In the special case where $\Phi(\lambda) = \delta\lambda$ for some $\delta > 0$, which makes the subordinator $Y(t) = \delta t$.

2.2.4.2 Continuous state branching processes

Continuous state branching processes with immigration (CBI processes) arise as scaling limits of discrete Galton-Watson processes, see [114]. A CBI process $Z = (Z(t); t \geq 0)$ is a Feller process on $[0, \infty]$ which is absorbed at ∞. As shown by [114], the Laplace transform of Z is of the form

$$\mathbb{E}_x\left[\exp\{-\lambda Z(t)\}\right] = \exp\left[-xu(t,\lambda) - \int_0^t \Phi(u(s,\lambda))\, ds\right], \qquad \forall \lambda > 0$$

where u is the unique non-negative solution to the integral equation

$$u(t,\lambda) + \int_0^t \Psi(u(s,\lambda))\, ds = \lambda,$$

for functions Ψ and Φ. The function Ψ is called the branching mechanism and must be of the form (2.11) and the function Φ is called the immigration rate and must be of the form (2.13). Conversely, given any two functions Ψ and Φ, there exists a CBI process with branching mechanism Ψ and immigration rate Φ. For simplicity, we will use $\mathrm{CBI}_x(\Psi, \Phi)$ to refer to the law of a CBI process starting from $x \geq 0$ and satisfies the above Laplace transform.

This is a classical result of Kawazu and Watanabe [114]: continuous state branching processes with immigration are in one-to-one correspondence with pairs of Lévy processes (X, Y) satisfying (2.10) and (2.12). The bijection described there is in terms of the Laplace transforms of the respective processes. A path-wise identification does exist, thanks to the work of Caballero, Pérez Garmendia and Uribe Bravo [49].

The authors of [49] examine existence and uniqueness to differential equations of the form

$$h' = x + f \circ h + g, \quad h(0) = 0, \quad h' \geq 0$$

for certain càdlàg functions f and g. In particular, they show that if X is a spectrally positive Lévy process with Laplace exponent $(-\Psi)$ and Y is an independent subordinator with Laplace exponent Φ then the unique càdlàg solution to

$$Z(t) = x + X\left(\int_0^t Z(s)\,ds\right) + Y(t) \tag{2.14}$$

is a $\mathrm{CBI}_x(\Psi, \Phi)$ process. When Y is identically 0, the Lévy process X in (2.14) is stopped upon hitting $-x$. The path-wise relationship when $Y = 0$ was observed by Lamperti [128], although proved later by Silverstein [163].

The authors of [49] also provide results that are useful for weak convergence. We will often, when dealing with weak converge, assume that Ψ is *conservative*, meaning [99] Ψ satisfies

$$\int_{0+} \frac{1}{|\Psi(u)|}\,du = \infty. \tag{2.15}$$

In some sense the most general assumption on Ψ that we can make for weak convergence arguments, see [114]. This also appears in weak convergence results involving height processes [73, Chapter 2].

2.2.4.3 The Ψ-height process

In this section we recall properties of the Ψ height process. These processes were introduced by Le Gall and Le Jan in [133], and further examined in [73, 132]. They are the continuous time analog of a contour process for a discrete tree. We recall some of there properties, but do not endeavor to state things in full generality. To this end, we assume that

$$\Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (e^{-\lambda r} - 1 + \lambda r)\,\pi(dr) \tag{2.16}$$

with $\alpha, \beta \geq 0$ and π a Radon measure with the stronger integrability condition $\int_0^\infty (r \wedge r^2)\, \pi(dr) < \infty$. We make the assumption that Ψ is conservative, i.e. it satisfies equation (2.15), and that further Ψ satisfies

$$\int_1^\infty \frac{1}{\Psi(u)}\, du < \infty. \tag{2.17}$$

Both of these assumptions on Ψ are slight restrictions on a general theory, but they imply that a Lévy process X with Laplace exponent $(-\Psi)$ has paths of infinite variation, non-negative jumps, and does not drift towards $+\infty$.

Let X be a Lévy process with Laplace exponent $(-\Psi)$ satisfying the above restrictions. The height process at time t is a way to "measure" the size of the set

$$\left\{ s \leq t : X(s-) = \inf_{s \leq r \leq t} X(r) \right\}. \tag{2.18}$$

This is exactly the continuum analog of (2.6) for the discrete height profile.

This is done through a time-reversal approach. For each $t > 0$, define $\widehat{X}^{(t)} = \left(\widehat{X}^{(t)}(s); s \in [0,t]\right)$ by $\widehat{X}^{(t)}(s) = X(t) - X((t-s)-)$ and the corresponding supremum process by $\widehat{S}^{(t)}(s) = \sup_{r\in[0,s]} \widehat{X}^{(t)}(r)$. The value $H(t)$ is a normalization of the local time at 0 of the process $\widehat{X}^{(t)} - \widehat{S}^{(t)}$. By [73, Theorem 1.4.3] H has a continuous modification if and only if the condition in (2.17) is satisfied. Whenever (2.17) is satisfied, we will implicitly assume that H is this modification. For more information on the height process see [73].

The process H possesses a family of local times $L = (L_t^a; t \geq 0, a \geq 0)$ which almost surely satisfies the occupation density formula:

$$\int_0^t g(H(r))\, dr = \int_0^\infty g(a) L_t^a\, da, \qquad \forall g \in C_c(\mathbb{R}),\, t \geq 0. \tag{2.19}$$

See [73, Proposition 1.3.3]. There is also a Ray-Knight theorem for these processes as shown in [73, Theorem 1.4.1]. Namely, define $T_x = \inf\{t : L_t^0 = x\} = \inf\{t : X(t) = -x\}$ then

$$\left(L_{T_x}^a; a \geq 0\right) \sim \mathrm{CBI}_x(\Psi, 0). \tag{2.20}$$

Similar Ray-Knight theorems were obtained by Warren [178] involving sticky Brownian motion.

The Ray-Knight theorem above was extended by Lambert [127] and Duquesne [72] to allow for some immigration. Lambert's work is slightly more general; however Duquesne's work contains some better approximation results for the local time. For $\delta > 0$ and $x \geq 0$ define the left-height process $\overleftarrow{H} = (\overleftarrow{H}(t); t \geq 0)$ by

$$\overleftarrow{H} = H(t) + \frac{1}{\delta}(L_t^0 - x)_+, \tag{2.21}$$

where $(x)_+ = \max(x, 0)$. By [72], there exists a jointly measurable family of local times $\overleftarrow{L} = \left(\overleftarrow{L}_t^a; t \geq 0, v \geq 0\right)$ which is continuous and increasing in t which satisfies equation (2.19) with $\overleftarrow{H}$ and $\overleftarrow{L}$ replacing H and L. Since $\overleftarrow{H}(t) \to \infty$ as $t \to \infty$, we can define $\overleftarrow{L}_\infty^a = \lim_{t\to\infty} \overleftarrow{L}_t^a$ exists and is finite almost surely due to (2.15). Moreover, the Ray-Knight theorem, see [72, Theorem 1.2] and [127, Theorem 5], states

$$\left(\overleftarrow{L}_\infty^a; a \geq 0\right) \sim \mathrm{CBI}_x(\Psi, \Phi), \qquad \Phi(\lambda) = \delta\lambda. \tag{2.22}$$

In the case where the branching process is a squared Bessel process and the height process is a reflected Brownian, a similar result was shown in [134].

2.2.4.4 Excursions

The discussions above about spectrally positive Lévy processes and height processes implicitly involved working under the *probability* measure $\mathbb{P}$ governing the motion of X. However, if we denote $I(t) = \inf_{r\leq t} X(r)$ as the running infimum of X we can see (informally) by (2.18) that $H(t) = 0$ when $X(t) = I(t)$. That is H has the same excursion intervals as $X - I$ and, in fact, the process H can be defined under the excursion measure of $X - I$. See [73, Chapter 1] for more details.

Working under (2.16) and (2.17) and mostly following the presentation in [74] and [71] (with the latter mostly focused on the stable case) the excursions are constructed as follows. The process $X - I$ is a strong Markov process for which 0 is a regular point for $X - I$, that is $\inf\{t > 0 X(t) - I(t) = 0\} = 0$ with probability 1, and, moreover, $-I$ acts as a local time at zero of $X - I$ at zero. We can decompose $\{s \geq 0 : X(s) - I(s) > 0\}$ by

$$\bigcup_{i\in\mathcal{I}} (g_i, d_i) = \{s \geq 0 : X(s) - I(s) > 0\} = \{s \geq 0 : H(s) > 0\}$$

for some non-empty and disjoint intervals (g_i, d_i) of length $\zeta_i = d_i - g_i$. We can let $e_i = (e_i(t); t \in [0, \zeta_i])$ and $h_i = (h_i(t); t \in [0, \zeta_i])$ be defined by

$$e_i(t) = X(g_i + t) - I(g_i) \quad \text{and} \quad h_i(t) = H(g_i + t).$$

We can define a point measure, $\mathcal{N}(dt, de, dh)$, on $\mathbb{R}_+ \times \mathbb{D}(\mathbb{R}_+) \times C(\mathbb{R}_+)$ by

$$\mathcal{N}(dt, de, dh) = \sum_{i \in \mathcal{I}} \delta_{(-I(g_i), e_i, h_i)}(dt, de, dh).$$

The measure $\mathcal{N}$ is a Poisson random measure with intensity measure $dt \otimes N(de, dh)$, and we will sometime abuse notation and write $N(de)$ (resp. $N(dh)$) and call this the law of X (resp. H) under the its excursion measure. Similarly, we will let ζ denote the lifetime of an excursion e or h under the measure N.

By general theory for measures [112, Chapter 3], we can condition $N(de, dh)$ on the value of ζ generate conditioned excursions (e, h) on the event $\zeta = x$ for deterministic x. The resulting measure $N(de, dh|\zeta = x)$ is a probability measure, but is often difficult to handle in general case. However, when X is a stable Lévy process, i.e. $\Psi(\lambda) = c\lambda^\alpha$ for some $\alpha \in (1, 2]$ and some $c > 0$, conditioning on the value of ζ becomes much simpler thanks to Itô's excursion theory for Brownian motion [156] ($\alpha = 2$) and the α-stable analogs due to Chaumont [54] ($1 < \alpha < 2$). See also [26, Chatper VIII].

Stable Lévy processes of index α have the following scaling relationship:

$$(X(t); t \geq 0) \stackrel{d}{=} \left(\lambda^{1/\alpha} X(t/\lambda); t \geq 0\right), \qquad \forall \lambda > 0.$$

This can be readily read off from the Laplace transform. We now fix an $\alpha \in (1, 2]$ and let $\mathscr{S}_\beta^{(\lambda)} : \mathbb{D}(\mathbb{R}_+) \to \mathbb{D}(\mathbb{R}_+)$, for each $\lambda, \beta > 0$ denote the scaling operation

$$\mathscr{S}_\beta^{(\lambda)}(w)(t) = \lambda^{1/\beta} w(t/\lambda), \qquad t \geq 0.$$

That we scale down time by a factor of λ and we scale up space by a factor of $\lambda^{1/\beta}$.

The law of an α-stable Lévy process X is preserved under the map $\mathscr{S}_\alpha^{(\lambda)}$ for all $\lambda > 0$; however the law of H is *not* preserved under $\mathscr{S}_\alpha^{(\lambda)}$ except in the Brownian case $\alpha = 2$. In the general α-stable case, the corresponding height process H has its law preserved under the scaling $\mathscr{S}_{\frac{\alpha}{\alpha-1}}^{(\lambda)}$:

$$(H(t); t \geq 0) \stackrel{d}{=} \left(\lambda^{\frac{\alpha-1}{\alpha}} H(t/\lambda); t \geq 0\right).$$

Modulo details about a sub-sequential limit, this scaling relationship can be see from Lemma 1.2.1 in [73]. For a more complete description see Section 3.1 in [70].

Using Poisson point processes it is not difficult to see that $N(de, dh|\zeta > t)$ is the image of $N(de, dh|\zeta > \lambda t)$ under the map $(w_1, w_2) \mapsto \left(\mathscr{S}_\alpha^{(\lambda)}(w_1), \mathscr{S}_{\frac{\alpha}{\alpha+1}}^{(\lambda)}(w_2)\right)$. As is shown in [54], see also [26, Chapter VIII], and [70] the conditional law $N(de, dh|\zeta = x)$ is the pushfoward of $N(de, dh|\zeta > t)$ by the (random) map

$$(w_1, w_2) \mapsto \left(\mathscr{S}_\alpha^{(x/\zeta)}(w_1), \mathscr{S}_{\frac{\alpha}{\alpha+1}}^{(x/\zeta)}(w_2)\right).$$

While the above gives a description for the excursion measure, Chaumont [54] provides a nice path-wise construction of $e \sim N(de|\zeta = 1)$ from just the process X. This was extended by Duquesne [70] to include the height process. Namely, let $g_1 = \sup\{t \le 1 : X(t) - I(t) = 0\}$ and $d_1 = \inf\{t \ge 1 : X(t) - I(t) = 0\}$. The interval (g_1, d_1) is the interval straddling time 1, and the corresponding excursions of $X - I$ and H which straddle are

$$\tilde{X}(t) = X(t + g_1) - I(g_1), \quad \tilde{H}(t) = H(t + g_1) \qquad \text{for } t \in [0, d_1 - g_1].$$

We denote by $\zeta_1 = d_1 - g_1$ the (random) length of the excursions. The path-wise construction of $(e, h) \sim N(de, dh|\zeta = x)$ is as follows.

Theorem 2.2.13. *Fix an $x > 0$ and let $\mathbf{e}_x = (\mathbf{e}_x(t); t \in [0, x])$ and $\mathbf{h}_x = (\mathbf{h}_x(t); t \in [0, x])$ be defined by*

$$\mathbf{e}_x(t) = \left(\frac{x}{\zeta_1}\right)^{\frac{1}{\alpha}} \tilde{X}\left(\frac{\zeta_1 t}{x}\right), \quad \textit{and} \quad \mathbf{h}_x(t) = \left(\frac{x}{\zeta_1}\right)^{\frac{\alpha-1}{\alpha}} \tilde{H}\left(\frac{\zeta_1 t}{x}\right).$$

Then $(\mathbf{e}_x, \mathbf{h}_x) \sim N(de, dh|\zeta = x)$.

2.2.4.5 Exchangeable increment processes

In the discrete setting of Section 2.2.2 we saw that Galton-Waston branching processes conditioned on their child sequence is uniformly distributed over all rooted planar trees with said child sequence, provided that the child sequence occurs with positive probability. The continuum analog of the Łukasiewicz path and its associated discrete height process is an excursion e and its corresponding height process h. These processes are “distributed” according to the measure $N(de, dh)$. The continuum analog of the Łukasiewicz path for trees with a given child sequence turns out to be a transformation of a processes with exchangeable increments.

We recall the random elements $(\xi_j; j = 1, 2, \cdots)$ with values in some Polish space $\mathcal{X}$ are said to be *exchangeable* if $(\xi_j; j \ge 1) \overset{d}{=} (\xi_{\pi j}; j \ge 1)$ for any permutation $\pi : \mathbb{N} \to \mathbb{N}$ which fixes all but finitely many coordinates. A classical result, originally due in some version to de Finetti, states that $(\xi_j; j \ge 1)$ are exchangeable in $\mathcal{X}$ if there is some random measure on $\mathcal{X}$, μ, such that conditionally given μ the random variables $(\xi_j; j \ge 1)$ are independent with law μ. A proof of this result can be found in Chapter 27 of [112].

A real-valued process $X = (X(t); t \in [0, 1])$ is said to have exchangeable increments if its increments over disjoint intervals of different lengths are exchangeable random variables. In the case where $X \in D([0, 1])$, we can phrase this as the following. For each $n = 1, 2, \cdots$ the random variables $X(\frac{j}{n}) - X(\frac{j-1}{n})$ for $j \in [n]$ are exchangeable. It is easy to see that by the independence of the increments of Lévy processes imply that Lévy processes have exchangeable increments.

There is a classification theorem of Kallenberg [110] which describes the distribution of exchangeable increment processes. We refer to [110] for a full statement of the result and restrict ourselves to the case where X does not possess negative jumps.

Theorem 2.2.14. *Let X be a process with exchangeable increments, càdlàg paths, and almost surely does not jump downwards. Then there exists*

1. $\alpha \in \mathbb{R}$ *,*$\sigma > 0$*,* $\beta_1 \ge \beta_2 \ge \cdots \ge 0$ *random variables with* $\sum_j \beta_j^2 < \infty$*;*

2. *a sequence of i.i.d.* $\mathrm{Unif}(0, 1)$ *random variables* $(U_j; j \ge 1)$*; and*

3. *a Brownian bridge* $B^{\mathrm{br}} = (B^{\mathrm{br}}(t); t \in [0, 1])$

where each item is independent of the others and

$$X(t) = \alpha t + \sigma B^{\mathrm{br}}(t) + \sum_j \beta_j (1_{[U_j \le t]} - t).$$

Such processes X are the càdlàg scaling limits of discrete processes with exchangeable increments. We will refer to such a càdlàg process X as an EI (exchangeable increment) process. We know that the Łukasiewicz path of a tree terminates at the value -1 but this is not an EI process; however, thanks to Lemma 2.2.10 and the discussion afterwards, the Łukasiewicz path of a tree with a given degree sequence is constructed from a discrete process with exchangeable increments. That is the Łukasiewicz path of a uniformly distributed tree $\mathtt{t}$ with a given degree sequence is obtained by a cyclic shift of a downward skip-free walk with exchangeable increments which starts at 0 and terminates at some time at -1 (which in general will not be the first hitting time of -1). We call such a process a downward skip-free bridge.

Any càdlàg scaling limit of a downward skip-free bridge that has exchangeable increments must be an EI process X^{br} with paths in $D([0,1])$ such that $X^{\mathrm{br}}(0) = X\mathrm{br}(1) = 0$. We will call such a process an *EI bridge* (hence the br in the superscript). In this situation, we have a deterministic sequence of degree sequences and so the resulting bridge satisfies

$$X^{\mathrm{br}}(t) = \theta_0 B^{\mathrm{br}}(t) + \sum_{i=1}^{\infty} \theta_i (1_{[U_i \le t]} - t), \qquad \theta_0 \ge 0, \theta_1 \ge \theta_2 \ge \cdots \ge 0 \text{ and } \sum_i \theta_i^2 < \infty,$$

where $\theta = (\theta_0, \theta_1, \cdots)$ is deterministic. In the sequel we will write $X^{\theta,\mathrm{br}}$ for the bridge described above. Again, we point out that the processes we are considering only jump upwards.

Of course, this is a bridge and not an excursion. That is, in general, $X^{\mathrm{br}}(t)$ attains non-positive values for $t \in (0,1)$ with positive probability. The analog of the cyclic shift lemma (Lemma 2.2.10) in the continuum is the the so-called *Vervaat transform* named after Vervaat [174]. We define the Vervaat transform of a càdlàg function $f : [0,1] \to \mathbb{R}_+$ with no-negative jumps satisfying $f(0) = f(1) = 0$ such does not jump downwards as

$$\mathscr{V}(f)(t) = f(t + \rho \mod 1) - \inf_u f(u), \qquad \rho = \inf\left\{t : f(t) \wedge f(t-) = \inf_u f(u)\right\}.$$

In words, the Vervaat transform $\mathscr{V}(f)$ swaps the pre- and post-infimum sections of the curve f. We note that the Vervaat transform of a bridge need not start from 0. That is for some functions f, $\mathscr{V}(f)(0) > 0$. Moreover, the $\mathscr{V}(f)(t) \ge 0$ for all $t \in [0,1]$ and can take the value 0 for some $t \in (0,1)$ in the case where f obtains its infimum several times.

It turns out [19] that when X is an EI bridge without negative jumps and has *infinite variation* then $\mathscr{V}(X^{\theta,\mathrm{br}})(t) = 0$ if and only if $t \in \{0,1\}$. See [19] and [122]. We will now restrict our attention this case. The values of $\theta = (\theta_0, \theta_1, \cdots)$ that result in processes with infinite varation can be seen to be

$$\theta_0 > 0 \qquad \text{or} \qquad \sum_i \theta_i = +\infty.$$

In this case, we will define

$$X^{\theta,\mathrm{ex}} = \mathscr{V}\left(X^{\theta,\mathrm{br}}\right).$$

The resulting process is an excursion without negative jumps in the case that θ satisfies (2.2.4.5).

In the case of EI processes there is no notion of local time, so it is not immediately clear what the right way to measure the set

$$\left\{s \le t : X^{\theta,\mathrm{ex}}(s-) = \inf_{s \le r \le t} X^{\theta,\mathrm{ex}}(r)\right\}$$

in this case. In [14] the authors came up with a method that is useful for describing inhomogeneous tree. The basic idea is the following. For each θ_j, $j \ge 1$, there is a corresponding jump time t_j such that $X^{\theta,\mathrm{ex}}(t_j) - X^{\theta,\mathrm{ex}}(t_j-) = \theta_j$. This time is unique provided that the θ_i are distinct. Let T_j be defined by

$$T_j = \inf\{t > t_j : X^{\theta,\mathrm{ex}}(t) = X^{\theta,\mathrm{ex}}(t_j-)\}$$

be the first return time. Then define

$$R_j^\theta(u) = \begin{cases} \inf_{t_j \le r \le u} X^{\theta,\mathrm{ex}}(r) - X^{\theta,\mathrm{ex}}(t_j-) & : u \in [t_j, T_j] \\ 0 & : \text{else} \end{cases}.$$

Here $R_j^\theta(u) = 0$ if $\theta_j = 0$. The height process $H^\theta = (H^\theta(t); t \in [0,1])$ is defined by $H^\theta(t) = X^{\theta,\mathrm{ex}}(t) - \sum_{i=1}^\infty R_i^\theta(t)$.

Chapter 3

BRIEF LITERATURE REVIEW

The literature on scaling limits of random trees and random graphs as metric spaces is quite young (spanning roughly three decades) although contains a plethora of significant results. In this chapter we will give a brief review of some of this literature, referencing some major results, as well has providing some heuristics for why the result should hold.

3.1 Scaling limits of Random trees

3.1.1 Galton-Watson forests

Recall from Section 2.2.2.6 that for any (sub)critical probability measure μ on the non-negative integers, we can define a random tree $\mathcal{T} \sim \mathrm{GW}(\mu)$ which is almost surely a finite planar tree. The tree $\mathcal{T}$ can be constructed by either taking its depth-first or breadth-first Łukasiewicz path to be the a random walk with increments equal in law to $(\xi - 1)$ where $\xi \sim \mu$. That is set $X = (X(t); t = 0, 1, \cdots, T_1)$

$$X(t) = \sum_{j=0}^{t-1} (\xi_i - 1), \qquad T_1 = \min\{t : X(t) = -1\}, \qquad \xi_i \overset{i.i.d.}{\sim} \mu \tag{3.1}$$

and $\mathcal{T}$ can be constructed by setting X as either the depth-first or breadth-first Łukasiewicz path. We will focus on the depth-first construction.

We can also consider a forest of planar trees $(\mathcal{T}_1, \mathcal{T}_2, \cdots)$ constructed from the infinite duration random walk $X = (X(t); t = 0, 1, \cdots)$ defined by (3.1) where $\mathcal{T}_j$ is defined by setting first setting $T_j = \min\{t : X(t) = -j\}$ and defining depth-first Łukasiewicz path of $\mathcal{T}_j$, $(X^{\mathrm{DF}}_{\mathcal{T}_j}(t); t = 0, 1, \cdots, T_j - T_{j-1})$ as

$$X^{\mathrm{DF}}_{\mathcal{T}_j}(t) = X(T_{j-1} + t) + (j - 1).$$

Similarly, we can see that the height process, $H_{\mathcal{T}_j} = (H_{\mathcal{T}_j}(t); t = 0, \cdots, T_j - T_{j-1})$ of $\mathcal{T}_j$

defined by (2.6) can be recognized as the excursion of

$$H(t) = \# \left\{ k \in \{0, 1, \cdots, n-1\} : X(k) = \min_{k \le j \le n} X(t) \right\} \tag{3.2}$$

between times T_{j-1} and T_j. Similarly, one can define the contour process $W_{\mathcal{T}_j} = (W_{\mathcal{T}_j}(t); t = 0, 1, \cdots, T_j - T_{j-1})$ for each of these trees $\mathcal{T}_j$ which are obtained from a single process $W = (W(t); t = 0, 1, \cdots)$.

Now consider a sequence of (sub)critical probability measures on the non-negative integers $(\mu_n; n \ge 1)$ and a collection of downward skip-free random walks

$$X_n(t) = \sum_{j=0}^{t-1} (\xi_{n,j} - 1), \qquad (\xi_{n,j}; j = 0, 1, \cdots) \overset{i.i.d.}{\sim} \mu_n. \tag{3.3}$$

Further, from these walks $X_n = (X_n(t); t = 0, 1, \cdots)$ we can define a height process H_n via (3.2). Also define $W_n = (W_n(t); t = 0, 1, \cdots)$ as the concatenation of the contour processes of $(\mathcal{T}_j; j \ge 1)$.

Fix a divergent sequence $\gamma_n \to \infty$. A classical result of Skorohod [165, Theorem 2.7] states that a necessary and sufficient condition for

$$\left(n^{-1} X_n(\lfloor n\gamma_n t \rfloor); t \ge 0\right) \overset{d}{\Longrightarrow} (X(t); t \ge 0)$$

towards a Lévy process $X = (X(t); t \ge 0)$ in the Skorohod space is the much simpler condition that

$$n^{-1} X_n(\lfloor n\gamma_n \rfloor) \overset{d}{\Longrightarrow} X(1). \tag{3.4}$$

Recall from Section 2.2.4.3, that under certain conditions on the Laplace transform Ψ of X we can construct a Ψ-height process $H = (H(t); t \ge 0)$. Namely, we assume that

$$\Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (e^{-\lambda r} - 1 + \lambda r)\pi(dr) \tag{3.5}$$

where $\alpha, \beta \ge 0$ and $(r \wedge r^2)\pi(dr)$ is a finite measure on $(0, \infty)$ and that satisfies both

$$\int_{0+} \frac{1}{|\Psi(u)|} \, du = +\infty, \qquad \text{and} \tag{3.6}$$

$$\int^{\infty} \frac{1}{\Psi(u)} \, du < \infty. \tag{3.7}$$

Condition (3.7) implies that X has paths of infinite variation [26] and condition (3.6) is required for weak convergence results involving branching processes. Under (3.5), (3.6) and (3.7) it is possible to construct a Ψ-height process H with continuous sample paths.

Unfortunately, the above condition on Ψ are not sufficient to imply that H_n (or W_n) possesses a scaling limit as well. There is an additional technical assumption on the generating function of μ_n that is needed for some technical reasons. For completeness we will include it. Let g_n denote the generating function of μ_n and recursively define $g_n^{(k)} = g_n^{(k-1)} \circ g_n$ denote the k^{th}-fold composition of g_n. We can now state a corollary of Theorems 2.3.1 and 2.4.1 in [73]

Theorem 3.1.1 (Duquesne & Le Gall [73]). *Suppose that* (3.4) *holds where X is a spectrally positive Lévy process which satisfies* (3.5), (3.6) *and* (3.7). *If, in addition, for every $\delta > 0$ fixed*

$$\liminf_{n\to\infty} \left(g_n^{(\lfloor \gamma_n \delta \rfloor)}(0) \right)^n > 0,$$

then

$$\left(\left(n^{-1} X_n(\lfloor n\gamma_n t \rfloor), \gamma_n^{-1} H_n(\lfloor n\gamma_n t \rfloor), \gamma_n^{-1} W_n(\lfloor 2n\gamma_n t \rfloor) \right); t \geq 0 \right) \stackrel{d}{\Longrightarrow} ((X(t), H(t), H(t)); t \geq 0)$$

in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^3)$.

Remark 3.1.1. A particular application or modification of Theorem 3.1.1 appears in many subsequent results describing limits of random tree/graph models [4, 5, 44, 45, 59, 72] to name a few.

Remark 3.1.2. The condition on the generating function g_n is automatically satisfied if $\mu_n = \mu$ for all n. In this case the limiting process $X(t)$ is an α-stable Lévy process. See Chapter 2 in [73].

Remark 3.1.3. We included the time-scaling as $\lfloor n\gamma_n t \rfloor$ instead of the perhaps more natural $\lfloor nt \rfloor$ to make the space scaling simpler. In the stable case where $\Psi(\lambda) = c\lambda^\alpha$ for some $\alpha \in (1, 2]$ the scaling can be replaced with

$$\left(\left(n^{-\frac{1}{\alpha}} X_n(\lfloor nt \rfloor), n^{-\frac{\alpha-1}{\alpha}} H_n(\lfloor nt \rfloor), n^{-\frac{\alpha-1}{\alpha}} W_n(\lfloor 2nt \rfloor) \right); t \geq 0 \right) \stackrel{d}{\Longrightarrow} ((X(t), H(t), H(t)); t \geq 0).$$

3.1.2 Height Profiles of Forests

As discussed in Section 2.2.2.3, in particular Lemma 2.2.8, height profiles of planar trees can be obtained from the breadth-first Łukasiewicz walk by solving a difference equation. Since for each $j = 1, 2, \cdots$ the value $T_j < \infty$ a.s. there are infinitely many planar trees $\mathcal{T}_j$ generated by the random walk X defined in (3.1) as in the previous section. In turn, there is no breadth-first walk that encodes all of trees $(\mathcal{T}_j; j \geq 1)$. However, for each fixed $k = 1, 2, \cdots$, we can consider a the forest $\mathcal{F}_k$ constructed from $(\mathcal{T}_j; j = 1, \cdots, k)$ and since the increments $(\xi_i - 1)$ of X^{DF} are exchangeable (in fact i.i.d.) the breadth-first Łukasiewicz path of $\mathcal{F}_k$, denoted by $X^{\mathrm{BF}} = (X^{\mathrm{BF}}(t); t = 0, 1, \cdots, T_k)$ is equal in law to $X^{\mathrm{DF}}(t \wedge T_k)$.

In particular, Lemma 2.2.8 implies the following:

Lemma 3.1.2. *For each $k \geq 1$, let $Z_k = (Z_k(h); h = 0, 1, \cdots)$ denote the height profile of $\mathcal{F}_k$ defined by*

$$Z_k(h) = \#\left\{v \in \mathcal{T}_j : j \leq k, d(\circ_{\mathcal{T}_j}, v) = h\right\}.$$

Then Z_k solves

$$Z_k(h) = k + X^{\mathrm{BF}} \circ C_k(h-1), \qquad C_k(h) = \sum_{\ell=0}^{h} Z_k(\ell),$$

and, in particular, $Z_k \overset{d}{=} \tilde{Z}_k$ where

$$\tilde{Z}_k(h) = k + X^{\mathrm{DF}} \circ \tilde{C}_k(h), \qquad \tilde{C}_k(h) = \sum_{\ell=0}^{h} \tilde{Z}_k(\ell).$$

We can similarly consider height profiles $Z_{k,n}$ constructed from the first k trees in the forest generated by the downward skip-free random walks X_n in (3.3). A corollary of Theorem 3.1.1, Lemma 3.1.2 and the time-change results of [49] imply

Theorem 3.1.3 (Duquesne & Le Gall [73], Caballero et. al. [49])**.** *Fix an $x \geq 0$ and let $k = k(n) = \lfloor \gamma_n x \rfloor$. Jointly with the convergence in Theorem 3.1.1,*

$$\left(n^{-1} Z_{k,n}(\lfloor \gamma_n t \rfloor); t \geq 0\right) \overset{d}{\Longrightarrow} (Z(t); t \geq 0) \tag{3.8}$$

where Z is the unique càdlàg solution to

$$Z(t) = x + \tilde{X}\left(\int_0^t Z(s)\, ds\right), \qquad \tilde{X} \overset{d}{=} X.$$

More generally, if (3.4) *holds then* (3.8) *still holds.*

Let us remark that when $k = k(n) = 1$ for all n, (the "$x = 0$" situation) the limiting solution Z is trivial: $Z(t) = 0$ for all $t \geq 0$. Obtaining non-trivial limits in this situations (under appropriate conditioning) was the subject of my work in [58] along with various other works [14, 19, 120] involving discrete to continuum limits and [108, 146] involving just the continuum.

3.1.3 Conditioned GW trees

Recall from Chapter 2 that Aldous' constructed the Brownian continuum random tree as the metric space scaling limit of $\mathcal{T}_n$ which was a uniformly chosen labeled tree on n vertices rooted at the vertex $1 \in \mathcal{T}_n$. That is $\mathcal{T}_n$ was a graph on exactly n vertices. Theorem 3.1.1 does imply the aforementioned result of Aldous, because the trees constructed from the excursions of X_n are of *random* sizes depending on the distribution μ_n. Moreover, with probability 1 a Brownian motion $B(t)$ will not have any excursion of length exactly 1. Instead, more careful analysis is required to prove scaling limits for conditioned Galton-Watson trees.

Conditional limit theorems exist for conditioned Galton-Watson whose offspring distribution μ lies in the domain of attraction of a stable law. The Brownian case $\alpha = 2$ is due to Marckert and Mokkadem [142] (under an exponential moment assumption) and the general $\alpha \in (1, 2)$ case is due to Duquesne [70]. See also [123, 124].

Theorem 3.1.4 (Marckert & Mokkadem [142], Duquesne [70])**.** *Let μ be a probability measure on the non-negative integers such that $\gcd(k : \mu(k) > 0) = 1$ and let $\mathcal{T}_n \sim (\mathrm{GW}(\mu) | \#\mathcal{T} = n)$.*

1. *[$\alpha = 2$ case] If $\mathbb{E}[\xi] = 1$ and $\mathrm{Var}(\xi) = \sigma^2 \in (0, \infty)$ where $\xi \sim \mu$ then*

$$\left(\left(n^{-1/2} X^{\mathrm{DF}}_{\mathcal{T}_n}(\lfloor nt \rfloor), n^{-1/2} H_{\mathcal{T}_n}(\lfloor nt \rfloor), n^{-1/2} W_{\mathcal{T}_n}(\lfloor 2nt \rfloor)\right); t \in [0,1]\right) \overset{d}{\Longrightarrow} \left(\left(\sigma \mathbf{e}(t), \frac{2}{\sigma}\mathbf{e}(t), \frac{2}{\sigma}\mathbf{e}(t)\right); t \in [0,1]\right),$$

   where $\mathbf{e}$ is a standard Brownian excursion.

2. *[$\alpha \in (1,2)$ case] If $\mu([k,\infty)) = k^{-\alpha}\ell(k)$ for some slowly varying function ℓ, and $\alpha \in (1,2)$ then*

$$\left(\left(n^{-1/\alpha}X^{\mathrm{DF}}_{\mathcal{T}_n}(\lfloor nt\rfloor), n^{-(\alpha-1)/\alpha}H_{\mathcal{T}_n}(\lfloor nt\rfloor), n^{-(\alpha-1)/\alpha}W_{\mathcal{T}_n}(\lfloor 2nt\rfloor)\right); t\in[0,1]\right)$$
$$\stackrel{d}{\Longrightarrow} \left((\mathbf{e}(t), \mathbf{h}(t), \mathbf{h}(t)) ; t \in [0,1]\right),$$

where $(\mathbf{e},\mathbf{h}) \sim N(de, dh|\zeta = 1)$ for N the excursion measure for an α-stable Lévy process.

As a consequence of the above theorem and Theorem 2.2.12 is the following:

Corollary 3.1.5. *Let $\mathcal{T}_n$ be as in Theorem 3.1.4. View $\mathcal{T}_n$ as a rooted tree with graph distance d_n and with the uniform measure on the vertices μ_n. Then*

$$\left(\mathcal{T}_n, \mathtt{o}_n, n^{-\frac{\alpha-1}{\alpha}}d_n, \mu_n\right) \stackrel{d}{\Longrightarrow} c\cdot\mathscr{T}^{(\alpha)} \qquad \text{in GHP},$$

where $c\cdot\mathscr{T}^{(\alpha)}$ is a (up-to a scaling of the metric depending on the offspring law), the α-stable continuum random tree encoding by an excursion $\mathbf{h} \sim N(dh|\zeta = 1)$ of unit duration of the height process H.

3.1.4 Trees with a given degree sequence

Recall from Section 2.2.2.6 that conditioned Galton-Watson trees are mixtures of trees with a given degree sequence so a natural question to ask is if there is an analog of Theorem 3.1.4 for trees with a given degree sequence. Here there are mixed results, because in certain situations there is a not a well-defined theory of height processes for exchangeable increment processes. Trees with a given degree sequence are closely related to so-called **p**-trees which were studied in [14, 51].

We will state the following theorem whose proof can be found in [18].

Theorem 3.1.6 (Antuncio Hernández and Uribe Bravo [18])**.** *Let $\mathtt{s}_n \to \infty$ and let $\mathtt{c}^n = (\mathtt{c}^n_0 \ge \mathtt{c}^n_1 \ge \cdots \ge \mathtt{c}^n_{\mathtt{s}_n-1})$ be a child sequence for a tree $\mathtt{t}_n$. Let X_n be either the depth-first or breadth-first Łukasiewicz path of $\mathtt{t}_n$. Assume that*

1. *There exists $b_n \to \infty$ such that for all $i \ge 1$;*

$$\frac{\mathtt{c}^n_i}{b_n} \longrightarrow \theta_i \ge 0;$$

2. *There exists some* $0 \le \theta_0 < \infty$ *such that*

$$\frac{1}{b_n^2}\sum_i (i-1)^2 \mathtt{n}_i^n \longrightarrow \theta_0 + \sum_{i=1}^{\infty}\theta_i^2, \qquad \mathtt{n}_i^n = \#\{j : \mathtt{c}_j^n = i\};$$

and

3. *Either* $\theta_0 > 0$ *or* $\sum_i \theta_i = +\infty$*.*

Then

$$\left(b_n^{-1}X_n(\lfloor \mathtt{s}_n t\rfloor); t\in[0,1]\right) \overset{d}{\Longrightarrow} \left(\mathscr{V}(X^{\mathrm{br}}(t); t\in[0,1]\right)$$

where $\mathscr{V}$ *is the Vervaat transform of the EI bridge*

$$X^{\mathrm{br}}(t) = \theta_0 B^{\mathrm{br}}(t) + \sum_{i=1}^{\infty}\theta_i(1_{[U_i\le t]} - t), \qquad U_i \overset{i.i.d.}{\sim} \mathrm{Unif}(0,1).$$

Note that the above theorem only deals with the Łukasiewicz path of trees and not with the height process or contour process. Some results involving processes close to height processes have been obtained in this direction, see [46] and [27]. In lieu of the description in terms of height processes, we will state the following theorem in terms of random trees. Before we do this, we will need to define the process H^θ.

H^θ is constructed from $X^{\mathrm{exc}} := \mathscr{V}(X^{\mathrm{br}})$ as follows. For each $i \ge 1$ t_i as the for the location of jump of X^{exc} of size $\theta_i > 0$ (with ties broken arbitrarily). Set $T_i := \inf\{r \in (t_i, 1] : X^{\mathrm{exc}}(r) = X^{\mathrm{exc}}(t_i-)\}$ which exists because the EI bridge X^{br} does not jump downward. Set, for each $i \ge 1$,

$$R_i(u) = \begin{cases} \inf_{t_i\le u\le u} X^{\mathrm{exc}}(s) - X^{\mathrm{exc}}(t_i-) & : u \in [t_i, T_i] \\ 0 & : \text{ else} \end{cases},$$

which is null for all i such that $\theta_i = 0$. Lastly set

$$H^\theta(t) = X^{\mathrm{exc}}(t) - \sum_{i\ge 1} R_i(t).$$

Theorem 3.1.7 (Berzunza Ojeda, Holmgren & Thévenin [27])**.** *Suppose, in addition to the conditions of Theorem 3.1.6,* $\sum_{i=0}^{\infty}\theta_i^2 = 1$*,* $\sum_i \theta_i < \infty$ *and* $\theta_0 > 0$*. Let* μ_n *denote* *the uniform probability measure on* $\mathtt{t}_n$ *and let* $\mathtt{o}$ *denote the root of* $\mathtt{t}_n$*. If* $b_n^2/\mathtt{s}_n \to 1$ *and* $\sup_n \mathtt{n}_1^n/\mathtt{s}_n < 1$ *then*

$$(\mathtt{t}_n, \mathtt{o}, \frac{1}{b_n}d, \mu_n) \xrightarrow{d} \mathcal{T}_H \qquad \text{in cGHP topology}$$

where $H = H^\theta$ *was constructed above.*

3.2 Critical Erdős-Rényi Random Graph

We now move to describing the (near-)critical Erdős-Rényi random graph $G(n, n^{-1}+\lambda n^{-4/3})$ by exploring the graph in a manner that Aldous does in [12]. This type of encoding (and an essentially equivalent depth-first encoding) has become quite popular in describing asymptotic properties of random graph models at and near-criticality [13, 59, 62, 109] to name a few.

Consider the following exploration process of the graph $G_n^\lambda := G(n, n^{-1} + \lambda n^{-4/3})$. We start at time $t = 0$ by exploring vertex $v_0 := 1$ and let $\mathscr{N}_0 = \{v_1, \cdots, v_k\}$ denote the collection of neighbors of v_0 labeled in some arbitrary way and let $\mathcal{A}_0 = (v_1, \cdots, v_k)$ denote a stack of vertices. We write $\chi(v_0) = k = \#\mathscr{N}_0$. Set $X_n(0) = 0$ and $X_n(1) = \chi(v_0) - 1$.

At time $t = 1, 2, \cdots$, we will have some stack $A_{t-1} = (v_t, v_{t+1}, \cdots, v_{t+\ell})$ for some ℓ and we will set

$$\mathscr{N}_t = \{v \in G_n^\lambda : v \text{ is a neighbor of some vertex } v_s, s \le t\}$$

be the collection of all neighbors of some vertex $v_0, v_1, \cdots, v_t$. If $A_{t-1} = ()$ we will just select v_t to be the minimum vertex in $[n] \setminus \{v_0, \cdots, v_{t-1}\}$. By construction, these collections $\mathscr{N}_t$ are nested: $\mathscr{N}_0 \subset \mathscr{N}_1 \subset \cdots$. Define $\chi(v_t) = \#(\mathscr{N}_t \setminus \mathscr{N}_{t-1})$ as the number of newly discovered neighbors by the vertex v_t and set

$$X_n(t) = \sum_{i=0}^{t-1} (\chi(v_i) - 1)$$

and update the stack A_{t-1} to become

$$A_t = (v_{t+1}, \cdots, v_{t+\ell}, v_{t+\ell+1}, \cdots v_{t+\ell+\chi(v_t)})$$

by adding the newly discovered neighbors of v_t to the end of the stack A_{t-1} and deleting the vertex v_t.

This exploration captures essentially all of the structure of the Erdős-Rényi random graph (or really any random graph model). Begin by noting that the newly discovered vertices are always neighbors of the vertex v_t and hence will always lie in the same connected component of v_t. Conversely, every vertex v in a connected component containing v_t will be added to the stack $A.$ after time t and before the next time that the stack $A.$ becomes empty. In particular, the stack $A_s = ()$ is empty precisely at the times s such that we have finished exploring a connected component of G_n^λ. If we let $\mathscr{C}_n^*(1), \mathscr{C}_n^*(2), \cdots,$ denote the connected components of G_n^λ listed in the order in which they appear we can recover the times $\{s < n : A_s = ()\}$ by

$$\{s \le n : A_{s-1} = ()\} = \left\{ \sum_{j=1}^k \#\mathscr{C}_n^*(j) : k \ge 0 \right\}.$$

Indeed, note that we label the vertices in the first connected component by $v_0 = 1, v_1, \cdots, v_{\#\mathscr{C}_n^*(1)-1}$ and so $A_{s-1} = ()$ after we finish exploring the vertex labeled $\mathscr{C}_n^*(1) - 1$. That same counting argument also allows us to recognize

$$\left\{ t = 1, \cdots : X_n(t) < \min_{s<t} X_n(s) \right\} = \left\{ \sum_{j=1}^k \#\mathscr{C}_n^*(j) : k \ge 0 \right\} \tag{3.9}$$

If $\mathscr{C}_n(i)$ is the i^{th} largest connected component in G_n^λ then $(\mathscr{C}_n^*(i); i \ge 1)$ are a size-biased reordering of the connected components $(\mathscr{C}_n(i); i \ge 1)$.

It is also not difficult to see that the excursions of X_n above its running minimum encode a breadth-first spanning tree in each connected component of $\mathscr{C}_n^*(j)$, that is the excursions of X_n encode the breadth-first Łukasiewicz path of a breadth-first spanning tree in the connected components of G_n^λ.

Observe that $\deg(v_t) \ge \chi(v_t)$ and there is equality precisely when the stack $A_{t-1} = ()$ was empty, which is when we begin exploring a new component. When v_t is not the first vertex explored in a connected component, then $\deg(v_t) \ge \chi(v_t) + 1$ because $\chi(v_t)$ does not count the vertex that discovered v_t at some time $s < t$. Also note that there does not have to be equality in this case. Let us re-examine what happens at time t when we explore the vertex v_t. There are three collections of vertices $[n] \setminus \{v_t\}$: vertices that have already been completely explored and removed from stacks, vertices in the stack $A_{t-1} = (v_t, \cdots, v_{t+\ell})$

that are not v_t, and vertices that we have not yet discovered. These collections, with the labeling above, are respectively

$$S_1(t) = \{v_0, \cdots, v_{t-1}\}, \qquad S_2(t) = \{v_{t+1}, \cdots, v_{t+\ell}\}, \quad \text{and} \quad S_3(t) = [n] \setminus \{v_0, \cdots, v_{t+\ell}\}. \tag{3.10}$$

When $A_{t-1} = ()$ and v_t is the first vertex explored in a connected component are precisely those times t such that v_t has no neighbors in $S_1(t)$ since the vertex that discovered v_t in this exploration will always be in that set $S_1(t)$. Similarly, when $A_{t-1} = ()$ the set $S_2(t)$ will be empty but the converse need not hold because $A_{t-1} = (v_t)$ is allowed. We also observe that

$$|\#S_2(t) - (X_n(t) - \min_{s\le t} X_n(s))| \le 1.$$

Note that for every cycle in G_n^λ we can find an edge $\{v_t, v_u\}$ with $t < u$ the vertex v_t will have a neighbor v_u of v_t in and $v_u \in S_2(t)$ and $v_t \in S_1(u)$. Indeed, a cycle is discovered precisely when a vertex v_t is explored and we find an edge connected v_t to some already discovered - but not yet explored - vertex v_u. We will say that vertex v_t *discovers* a surplus edge in this situation. Let $N_n(t)$ be defined as the number of surplus edges discovered by the time we *start* exploring v_t so that

$$N_n(t+1) = N_n(t) + \#\{\text{surplus edges discovered by } v_t\}.$$

We now state the following lemma.

Lemma 3.2.1. *Let $\mathcal{F}_n(t) = \sigma(X_n(s) : s \le t)$. Then*

$$X_n(t+1) - X_n(t) + 1 | \mathcal{F}_n(t) \overset{d}{=} \text{Bin}\left(n - t - \#S_2(t), n^{-1} + \lambda n^{-4/3}\right),$$

and

$$N_n(t+1) - N_n(t) | \mathcal{F}_n(t) \overset{d}{=} \text{Bin}\left(\#S_2(t), n^{-1} + \lambda n^{-4/3}\right).$$

From here it follows that

$$\begin{aligned}\mathbb{E}\left[X_n(t+1) - X_n(t)|\mathcal{F}_n(t)\right] &= (n - t - \#S_2(t))(n^{-1} + \lambda n^{-4/3}) - 1\\ &\approx 1 - tn^{-1} + \lambda n^{-1/3} - 1\\ &= \frac{tn^{-2/3} + \lambda}{n^{1/3}}\\ \mathrm{Var}(X_n(t+1) - X_n(t)|\mathcal{F}_n(t)) &\approx 1,\end{aligned}$$

where in both approximations above, we can ignore the contribution of $\#S_2(t)$. This suggests that limits appear when we rescale time by $n^{2/3}$ and space by $n^{1/3}$ which is just Brownian scaling. Comparing the above martingale functional central limit approximations and (3.10) led Aldous to the following theorem

Theorem 3.2.2. *For each fixed $\lambda \in \mathbb{R}$, the following convergence holds in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$*

$$\left((n^{-1/3}X_n(\lfloor n^{2/3}t\rfloor), N_n(\lfloor n^{2/3}t\rfloor)); t \geq 0\right) \overset{d}{\Longrightarrow} \left((X^\lambda(t), N(t)); t \geq 0\right)$$

where

$$X^\lambda(t) = B(t) + \lambda t - \frac{1}{2}t^2$$

and N is a Poisson process with rate $X^\lambda(t) - \min_{s\leq t} X^\lambda(s)$.

3.2.1 Asymptotic Sizes of Connected Components

Returning to (3.9), we can see that the largest connected components of the near-critical graph G_n^λ appear as the longest excursions (above the minimum) of the breadth-first walk $X_n(t)$. Since the time-scaling of the walk X_n is $\lfloor n^{2/3}t\rfloor$, this implies that the largest connected components of G_n^λ are of order $n^{2/3}$. That is

$$(n^{-2/3}\#\mathscr{C}_n(j); j \in [k]) \overset{d}{\Longrightarrow} (\gamma_j; j \in [k])$$

for some non-degenerate random variables γ_j. Technically, this relies on some additional path properties of Brownian motion with parabolic drift which we do not dwell on here, see Lemma 7 in [12].

Looking as well at the process N_n, we observe that N_n increases at time t whenever we discover a surplus edge in the connected component $\mathscr{C}_n(j) \ni v_t$. Since the process N_n converges after appropriate rescaling of time towards an inhomogeneous Poisson process, the encoding described above not only encodes the asymptotic sizes of the connected components but also describes their surplus.

Formally, we have the following. Consider the following decomposition:

$$\bigcup_{j=1}^{\infty}(g_i, d_i) = \{t : X^{\lambda}(t) > \min_{s\leq t} X^{\lambda}(s)\}$$

ordered in such a way that $\gamma_j := d_j - g_j$ is decreasing. There should be an additional dependence on λ indicated in the notatoin but we omit this for clarity. This can be done because it can be shown [12, 13] that there are only finitely many excursion intervals (g_i, d_i) of length $\gamma_i > \delta$ for each fixed $\delta > 0$. Let $\sigma_j := N(d_j) - N(g_j)$ denote the increment of N over the interval (g_j, d_j). The

Theorem 3.2.3 (Aldous, [12])**.** *For each fixed λ, let $(\mathscr{C}_n(j); j \geq 1)$ be the largest connected components of the near-critical Erdős-Rényi random graph G_n^{λ}. Then, as $n \to \infty$,*

$$\Big((n^{-2/3}\#\mathscr{C}_n(j), s(\mathscr{C}_n(j))); j \geq 1\Big) \xrightarrow{d} ((\gamma_j, \sigma_j); j \geq 1)\,,$$

where the convergence of the first coordinate occurs in ℓ^2.

3.2.2 Scaling limits of the connected components

The work by Aldous in [12] on the sizes of the connected components of the critical Erdős-Rényi random graph were strengthened by Addario-Berry, Broutin and Goldschmidt in [4, 5].

In order to describe their result and outline why it holds, we recall the construction of a connected graph G from a rooted planar tree T. As described in Section 2.2.2.5 in more detail, given a rooted planar tree T and a fixed surplus k there are $\binom{A^{\mathrm{DF}}(T)}{k}$ many connected rooted graphs G such that $T_G^{\mathrm{DF}} = T$, where

$$A^{\mathrm{DF}}(T) = \sum_{t=0}^{\#T-1} X_T^{\mathrm{DF}}(t)$$

with X_T^{DF} is the depth-first Łukasiewicz path of the graph T. This is what happens when we had a deterministic number of surplus edges and when we specify what the depth-first tree T_G^{DF} will be. This is not the case with the connected components $(\mathscr{C}_n(j); j \geq 1)$ because, for example, the surplus $s(\mathscr{C}_n(j))$ is a random variable - in fact, conditionally on $\#\mathscr{C}_n(j)$ the surplus is a binomial random variable.

Consider, instead, the following model. Fix some $p \in (0,1)$ and some integer $k \geq 1$ and generate $\tilde{T}_k^p$ among all trees labeled by $[k]$ and rooted at 1 with probability

$$\mathbb{P}\left(\tilde{T}_k^p = T\right) \propto (1-p)^{A^{\mathrm{DF}}(T)}$$

where T is a labeled tree on $[k]$ and rooted at 1. The tree $\tilde{T}_k^p$ has precisely $A^{\mathrm{DF}}(\tilde{T}_k^p)$ many surplus edges that can be added to form a graph G who's depth-first tree spanning tree is $\tilde{T}_k^p$. Include each possible surplus edge independently with probability p to form a graph $\tilde{G}_k^p$. This generates a random rooted graph on the vertices $[k]$.

To connect $\tilde{G}_k^p$ with the connected components $\mathscr{C}_n(i)$, we will let $K_n = (K_n(i); i \geq 1)$ denote the sizes of the connected components $\mathscr{C}_n$. That is $K_n(i) = \#\mathscr{C}_n(i)$. As shown in [5], the connected component $\mathscr{C}_n(i)$ conditioned on K_n and is distributed as $\tilde{G}_{K_n(i)}^{n^{-1}+\lambda n^{-4/3}}$ modulo a relabeling of the vertices of $\mathscr{C}_n(i)$ by $[K_n(i)]$ in some exchangeable manner, by, for example labeling the vertices of $\mathscr{C}_n(i)$ by $[K_n(i)]$ in the unique way that preserves the ordering of the original labeling. This result does not actually rely on $p = p(n) = n^{-1} + \lambda n^{-4/3}$. Consequently, understanding the metric space structure of finitely macroscopic connected components $\mathscr{C}_n(1), \cdots, \mathscr{C}_n(j)$ relies on understanding: (1) the asymptotics of $K_n(i)$; (2) understanding the metric space scaling limits of $\tilde{T}_{K_n(i)}^p$; and (3) understanding how $\tilde{G}_{K_n(i)}^p$ is constructed from $\tilde{T}_{K_n(i)}^p$ along with how this relates to the GHP topology.

Item (1) is handled by the result of Aldous cited above (Theorem 3.2.3). To understand item (2), we recall that, the metric space scaling limit of the a sequence of random trees can be obtained by proving that the sequence of height processes converges to a continuous limit and by applying Theorem 2.2.12. Moreover, as a metric space, a random tree uniformly chosen among all trees labeled by $[k]$ and rooted at 1 is distributed as a Poisson(1) Galton-Watson tree conditioned on having k many vertices. That is $\mathcal{T}_k \sim (\mathrm{GW}(\mathrm{Poisson}(1)\,|\#\mathcal{T} = k)$. Comparing the change of measure in the definition of $\tilde{T}_k^p$ makes the following lemma clear. For simplicity, we will focus on just the largest connected component.

Lemma 3.2.4. *Let $X_k = (X_k(t); t = 0, 1, \cdots, k)$ denote the depth-first Lukasiewicz path of a Poisson(1) Galton-Watson tree conditioned on having size k and let H_k denote the corresponding height process. Let $\tilde{X}_k^p = (\tilde{X}_k^p(t); t = 0, 1, \cdots, k)$, resp. $\tilde{H}_k^p = (\tilde{H}_k^p(t); t = 0, 1, \cdots, k)$, denote the depth-first Lukasiewicz path, resp. height process, of the tree $\tilde{T}_k^p$. Then*

$$\mathbb{E}\left[F(\tilde{X}_k^p, \tilde{H}_k^p)\right] = \frac{\mathbb{E}\left[(1-p)^{\sum_{t=0}^{k-1} X_k(t)} F(X_k, H_k)\right]}{\mathbb{E}\left[(1-p)^{\sum_{t=0}^{k-1} X_k(t)}\right]}.$$

Recall from Theorem 3.1.4 above that as $k \to \infty$

$$\left(k^{-1/2} X_k(\lfloor kt \rfloor), k^{-1/2} H_k(\lfloor kt \rfloor)); t \in [0,1]\right) \stackrel{d}{\Longrightarrow} ((\mathbf{e}(t), 2\mathbf{e}(t)); t \in [0,1]) \tag{3.11}$$

where $\mathbf{e}$ is a standard Brownian motion and by Theorem 3.2.3

$$n^{-2/3} K_n(1) \stackrel{d}{\Longrightarrow} \gamma_1$$

where γ_1 is the length of the longest excursion of a Brownian motion with parabolic drift. Using Skorohod's representation theorem, we can suppose that $n^{-2/3} K_n(1) \to \gamma_1$ a.s. for all $i \in [j]$ and we will write $k = k_n := K_n(1)$ and $\gamma = \gamma_1$.

We note that $kn^{-2/3} \to \gamma$ and so

$$\left(\left(\frac{k}{\gamma_1}\right)^{-1/2} X_k(\lfloor \gamma^{-1} kt \rfloor), \left(\frac{k}{\gamma}\right)^{-1/2} H_k(\lfloor \gamma^{-1} kt \rfloor); t \in [0, \gamma]\right)$$
$$\stackrel{d}{\Longrightarrow} (\mathbf{e}^{(\gamma)}(t), 2\mathbf{e}^{(\gamma)}(t); t \in [0, \gamma_i]),$$

where $\mathbf{e}^{(\gamma)}|\gamma$ is a Brownian excursion of duration γ.

Returning to the change of measure in Lemma 3.2.4, we see that we can reconstruct $\tilde{G}_{k_n}^{n^{-1}+\lambda n^{-4/3}}$ from a change of measure and the random walk X_k. The change of measure depends on $\sum_{t=0}^{k-1} X_k(t)$ and $p = n^{-1} + \lambda n^{-4/3}$. In particular, note that $k_n p^{2/3} \to \gamma$ and

$$k^{-3/2} \sum_{t=0}^{k-1} X_k(t) \stackrel{d}{\Longrightarrow} \gamma^{-3/2} \int_0^\gamma \mathbf{e}^{(\gamma)}(t)\, dt \stackrel{d}{=} \int_0^1 \mathbf{e}^{(1)}(t)\, dt.$$

Therefore, after an application of Skorohod's representation theorem and assuming that this convergence actually holds almost surely,

$$(1-p)^{\sum_{t=0}^{k-1} X_k(t)} \approx (1-p)^{\frac{k^{3/2}}{\gamma^{3/2}} \int_0^\gamma \mathbf{e}^{(\gamma)}(t)\, dt} \longrightarrow \exp\left(\int_0^\gamma \mathbf{e}^{(\gamma)}(t)\, dt\right),$$

where the error in approximation is negligible, see [5]. Together with a uniform integrability argument, the authors of [5] were able to show

Theorem 3.2.5 (Addario-Berry, Broutin, Goldschimdt [5]). *Let $(\tilde{X}^p_k, \tilde{H}^p_k)$ be as in Lemma 3.2.4. Suppose that $kp^{2/3} \to \gamma$. Then*

$$\left(\left(\frac{k}{\gamma}\right)^{-1/2} \tilde{X}^p_k(\lfloor \gamma^{-1} kt \rfloor), \left(\frac{k}{\gamma}\right)^{-1/2} \tilde{H}^p_n(\lfloor \gamma_n^{-1} kt \rfloor); t \in [0,\gamma]\right) \overset{d}{\Longrightarrow} (\tilde{\mathbf{e}}^{(\gamma)}, 2\tilde{\mathbf{e}}^{(\gamma)})$$

where

$$\mathbb{E}\left[F(\tilde{\mathbf{e}}^{(\gamma)})\right] \propto \mathbb{E}\left[F(\mathbf{e}^{(\gamma)}) \int_0^\gamma \mathbf{e}^{(\gamma)}(t)\, dt\right].$$

In particular, if $\tilde{T}^p_k$ is a tree whose depth-first Łukasiewicz path is $\tilde{X}^p_k$, d is the graph distance on $\tilde{T}^p_k$ and μ_k is the couning measure on $\tilde{T}^p_k$ then

$$\left(\tilde{T}^p_k, 1, (k/\gamma)^{-1/2} d, (k/\gamma)\mu\right) \overset{d}{\Longrightarrow} \mathcal{T}_{2\tilde{\mathbf{e}}^{(\gamma)}}$$

weakly in cGHP topology.

In the construction of $\tilde{G}^p_k$ from $\tilde{T}^p_k$, recall that each possible surplus edge was included independently with probability p. Recalling the discussion in Section 2.2.2.5, in particular the discussion around equation (2.9), each possible surplus edge is uniquely represented by a point $(i,j) \in \mathbb{Z}^2_+$ with $0 \le j < \tilde{X}^p_k(i)$. Therefore, the graph $\tilde{G}^p_k$ can be constructed directly from $\tilde{X}^p_k$ and a binomial point set $\mathcal{Q}\{(i,j) : 0 \le j < \tilde{X}^p_k(i)\}$ where

$$\mathbb{P}\left((i,j) \in \mathcal{Q}\right) = p$$

and this occurs independently over (i,j). It is not hard to see that

$$\left\{\left((k/\gamma)^{-1} i, (k/\gamma)^{-1/2} j\right) : (i,j) \in \mathcal{Q}\right\} \overset{d}{\Longrightarrow} \mathcal{P} \cap \tilde{\mathbf{e}}^{(\gamma)}$$

where $\mathcal{P}$ is a Poisson point process on $\mathbb{R}^2_+$ with Lebesgue intensity, the $\mathcal{P} \cap f := \{(x,y) \in \mathcal{P} : 0 \le y \le f(x)\}$ and the convergence is in the sense of vague convergence of counting measures.

When viewing Theorem 3.2.5 together with the convergence of the point set above, the following theorem should not be too surprising.

Theorem 3.2.6 (Addario-Berry, Broutin, Goldshmidt [5] and Addario-Berry, Broutin, Goldschmidt and Miermont [6][1]). *Let $\mathscr{C}_n(1)$ be the largest connected component of $G(n, n^{-1}+\lambda n^{-4/3})$. Root $\mathscr{C}_n(1)$ at its minimally labeled vertex, and equip it with the graph distance d and counting measure μ_n. Then*

$$(\mathscr{C}_n(1), \mathtt{o}, n^{-1/3}d, n^{-2/3}\mu_n) \overset{d}{\Longrightarrow} \mathcal{G}\left(\tilde{\mathbf{e}}^{(\gamma)}, 2\tilde{\mathbf{e}}^{(\gamma)}, \mathcal{P}\right) \qquad \text{in cGHP}$$

where $\gamma = \gamma_1$ is the length of the longest excursion of the Brownian motion with parabolic drift.

In fact, the above theorem is not as strong as the results proved in [5]. The authors actually prove that the connected components converge jointly in the product space equipped with an ℓ^4-type norm. This allows for some more detailed analysis.

3.2.3 Height profiles of the connected components

The authors of [4, 5] do not discuss the convergence of the height profiles $Z_{n,1} = (Z_{n,1}(t); t = 0, 1, \cdots)$ of

$$Z_{n,1}(t) = \#\{v \in \mathscr{C}_n(1) : d(v, \mathtt{o}) = t\}.$$

It was first established in [147] using a breadth-first construction of the critical components; however, the convergence follows as a simple corollary of a more general result in my work [58] discussed in Chapter 5 below.

3.2.4 The Standard Multiplicative Coalescence

Let us discuss a different way to generate the rescaled component size $n^{-2/3}\#\mathscr{C}_n(j)$ of the near-critical Erdős-Rényi random graph $G(n, p)$ which has proved to be valuable for many random graph models. For each vertex $i \in [n]$ we assign a vertex i mass $x_i > 0$. Now for each $q \geq 0$ construct the graph $\mathcal{G}(\mathtt{x}, q)$ on $[n]$ where

$$\mathbb{P}\left(\{i, j\} \text{ is an edge in } \mathcal{G}(\mathtt{x}, q)\right) = 1 - \exp\{-q x_i x_j\}$$

[1]In [5] the authors only prove convergence of the connected component in the Gromov-Hausdorff topology, not the GHP topology. The extension to the GHP topology follows from Theorem 2.2.12 whose origin for the GHP topology appears to have originated in [6].

independently over all pairs of edges. We interpret the weight x_i as telling us the propensity for the vertex i to find new neighbors in the graph $\mathcal{G}(\mathtt{x}, q)$.

In the scaling of Theorem 3.2.3, each connected component is scaled by a factor of $n^{-2/3}$ which can be alternatively viewed as assigning each $i \in G(n,p)$ a mass $x_i = n^{-2/3}$. Setting $q = q(\lambda)$ as the unique solution to

$$1 - \exp(qn^{-4/3}) = n^{-1} + \lambda n^{-4/3}$$

the ordered weights of the connected components of the graph $G(n, n^{-1} + \lambda n^{-4/3})$ scaled by $n^{-2/3}$ agree with those of $\mathcal{G}(\mathtt{x}, q)$ without any additional scaling. We can also see that

$$q \approx n^{1/3} + \lambda.$$

While this connection between the Erdős-Rényi random graph and the inhomogeneous random graph $\mathcal{G}(\mathtt{x}, q)$ holds for nice choices of weights $\mathtt{x}$ more interesting phenomena can occur, see [13].

If we view the parameter $\lambda \in \mathbb{R}$ as a time-parameter, then we have the following interpretation of the masses of connected components of $\mathcal{G}(\mathtt{x}, n^{-1/3} + \lambda)$:

1. The largest connected component has mass converging to 0 as $\lambda \to -\infty$;

2. The largest connected component has mass converging to $+\infty$ while the second largest component has mass converging to 0 as $\lambda \to +\infty$.

That is as λ goes from $-\infty$ to $+\infty$ we see that the masses of components start at dust (all being zero) and merge as λ increases and as $\lambda \to \infty$ there is a single giant connected component and the rest of the components are dust again.

A remarkable fact proved by Aldous [12] and later extended in [13] is that this interpretation of the near-critical Erdős-Rényi random graph and the inhomogeneous model has a $n \to \infty$ limiting description in terms of a Brownian motion with parabolic drift which is called the *standard multiplicative coalescence*, which we now briefly describe.

Let us fix some vector $\mathtt{x} \in \ell^2_\downarrow = \{(x_1, x_2, \cdots) : x_1 \geq x_2 \geq \cdots \geq 0, \sum_i x_i^2 < \infty\}$ and let $\mathcal{X}(\mathtt{x}, t) = (\mathcal{X}_i(\mathtt{x}, t) : i \geq 1)$ denote the ordered masses of clusters of the graph $\mathcal{G}(\mathtt{x}, t)$

on countably many vertices. Whenever $\mathtt{x}$ has finite length, i.e. $x_i = 0$ for all sufficiently large i, then $\mathcal{X}(\mathtt{x}, t)$ is a continuous time Markov chain on a finite collection of spaces whose dynamics are the following

> each pair of clusters (x, y) merges at rate xy to form a single new cluster of size $x + y$,

where the term cluster refers to the masses $\mathcal{X}_i(\mathtt{x}, t)$. Aldous [12] shows that this process can be defined for starting vectors $\mathtt{x} \in \ell^2_\downarrow$ and not just finite length vectors and that process $(\mathcal{X}(\mathtt{x}, t); t \geq 0)$ is a Feller process in $\ell^2_\downarrow$. The *standard* multiplicative coalescence is the process $(\mathcal{X}(t); -\infty < t < +\infty)$ whose distribution at time t is the ordered lengths of excursions of $X^t(s) = B(s) + ts - \frac{1}{2}s^2$.

3.3 Other Critical Random Graph Models and Results

In this section we will recall some subsequent results on critical and near-critical random graphs and their asymptotics. Much of the work on the critical inhomogeneous random graph models discussed above and below were motivated by conjectures of statistical physicists in the early 2000's on the behavior of large random graphs [41, 42, 104]. These works suggest the following asymptotic behavior: Let $(\mathcal{G}_n; n \geq 1)$ be a sequence of *critical* random graphs on n vertices and let D_n denote the degree of a uniformly chosen vertex in $\mathcal{G}_n$. Denote by $\mathscr{C}_n = \mathscr{C}_n(1)$ denote the largest connected component in $\mathcal{G}_n$. If $\mathbb{P}(D_n \geq k) \approx ck^{1-\tau}$ for some $\tau > 3$ where the precise meaning of $\approx$ is intentionally omitted then, as $n \to \infty$, it is believed that

1. $[\tau > 4]$ $\#\mathscr{C}_n = \Theta(n^{2/3})$ and the typical distance between two vertices in $\mathscr{C}_n$ is $\Theta(n^{1/3})$; and

2. $[\tau \in (3, 4)]$ $\#\mathscr{C}_n = \Theta(n^{-(\tau-2)/(\tau-1)})$ and the typical distance between two vertices in $\mathscr{C}_n$ is order $n^{-(\tau-3)/(\tau-1)}$.

Observe that for the Erdős-Rényi random graph, while note falling into the power-law case with $\tau > 4$, does still exhibit the same scaling as Case 1 above. Mathematically this conjecture has been verified for several notable graph models which we will explain below.

We will often phrase results analogous to Theorem 3.2.3 in terms of the excursions sizes of some process above its running minimum. We will now establish some notation to eliminate some redundancy later on. Consider a càdlàg function f such that $f(0) = 0$, $f(t) \to -\infty$ as $t \to \infty$ and $f(t) - f(t-) \geq 0$ for all t. We call an interval (l, r) an excursion interval of f is $f(x) > f(l)$ for all $x < \ell$ and $r = \inf\{x : f(x) < f(\ell)\}$ and we will call $r - l$ the length of the excursion interval. Given an excursion interval (l, r) we call $f_{(l,r)} : [0, r-l] \to \mathbb{R}_+$ by $f_{(l,r)}(x) := f(l + x) - f(l)$ the excursion of f over (l, r). We note that $f_{(l,r)}$ is an excursion above its past minimum and $f_{(l,r)}(x)$ can be 0 on (l, r). Whenever it makes sense, we let $((l_i, r_i); i \geq 1)$ denote the collection of all excursion intervals of f listed in decreasing order and let $f_i = f_{(l_i,r_i)}$. We will also define

$$\mathcal{E}^{\downarrow}(f) = (f_i; i \geq 1) \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^{\infty}, \qquad \Gamma^{\downarrow}(f) = (r_i - l_i; i \geq 1) \in \ell^{\infty}_{\downarrow}.$$

Also, given an element $\mathtt{x} \in \mathbb{R}_+^{\infty}$ we write

$$\sigma_p(\mathtt{x}) = \sum_i x_i^p \in [0, \infty].$$

Finally given a graph G on, say, n vertices which is equipped with some finite measure $\mathfrak{m}$ we will write

$$\mathscr{W}(G) = (\mathfrak{m}(\mathscr{C}(1)), \mathfrak{m}(\mathscr{C}(2)), \cdots, \mathfrak{m}(\mathscr{C}(k)), 0, 0, \cdots) \in \ell^{\infty}_{\downarrow}$$

as the ordered sequence of masses of the connected components of G, listed in decreasing order with ties broken arbitrarily. Similarly, if each connected component $\mathscr{C}(j)$ has a specified vertex $\mathsf{o}_j \in \mathscr{C}(j)$ we will denote

$$\text{COMP}(G) := \big((\mathscr{C}(i), \mathsf{o}_i, d, \mathfrak{m}|_{\mathscr{C}(i)}); i \geq 1\big)$$

as the ordered collection of PMM spaces where d represents the graph distance and for any PMM space $\mathcal{M}$ we will denote

$$\text{scale}(a, b)(\mathcal{M}, \mathsf{o}, d_{\mathcal{M}}, \mu_{\mathcal{M}}) = (\mathcal{M}, \mathsf{o}, a d_{\mathcal{M}}, b\mu_{\mathcal{M}})$$

as the scaled metric space.

3.3.1 Critical Inhomogeneous Random Graphs

In the last section, we discussed an inhomogeneous random graph $\mathcal{G}(\mathtt{x}, q)$ which under certain tuning of the parameters is identical to the Erdős-Rényi random graph modulo a re-weighting of the vertices by a fact of $n^{-2/3}$. To indicate the parameter n, we will write $\mathtt{x}^n = (x_i^n; i \geq 1)$.

The first extension to a situation were $x_i^n \neq n^{-2/3} 1_{[i\leq n]}$ actually appears in Aldous' work [12]. Recall that the vertex i in $\mathcal{G}(\mathtt{x}, q)$ is assigned a mass x_i. He showed the following theorem

Theorem 3.3.1 (Aldous [12])**.** *If* $(\mathtt{x}^n; n \geq 1)$ *is a sequence of finite positive weight vectors such that*

$$\frac{\sigma_3(x^n)}{(\sigma_2(\mathtt{x}^n))^3} \longrightarrow 1$$
$$\frac{x_j^n}{\sigma_2(\mathtt{x}^n)} \longrightarrow 0, \qquad \forall j \geq 1$$
$$\sigma_2(\mathtt{x}^n) \longrightarrow 0,$$

then for each fixed $t \in \mathbb{R}$

$$\mathscr{W}(\mathcal{G}(\mathtt{x}^n, \frac{1}{\sigma_2(\mathtt{x}^n)} + t)) \longrightarrow \Gamma^{\downarrow}(B(s) + ts - \frac{1}{2}s^2), \qquad in\ \ell^2_{\downarrow}.$$

The proof of the above theorem relies on an exploration of the graph $\mathcal{G}(\vec{x}, q)$ in a manner very similar to the exploration of the Erdős-Rényi random graph described in the previous section.

Shortly afterwards, Aldous and Limic [13] described more collection of graph models whose component masses exhibit scaling limits. Namely, they show

Theorem 3.3.2 (Aldous and Limic [13])**.** *Suppose that* $(\mathtt{x}^n; n \geq 1)$ *are a sequence of finite length vector such that*

$$\frac{\sigma_3(x^n)}{(\sigma_2(\mathtt{x}^n))^3} \longrightarrow \kappa + \sum_j c_j^3$$
$$\frac{x_j^n}{\sigma_2(\mathtt{x}^n)} \longrightarrow c_j, \forall j \geq 1$$
$$\sigma_2(\mathtt{x}^n) \longrightarrow 0,$$

where $\kappa \in \mathbb{R}_+$ *and* $\mathtt{c} \in \ell^3_\downarrow$. *If either* $\kappa > 0$ *or* $\mathtt{c} \in \ell^3_\downarrow \setminus \ell^2_\downarrow$ *then*

$$\mathscr{W}\left(\mathcal{G}\left(\mathtt{x}^n, \frac{1}{\sigma_2(\mathtt{x}^n)} + t\right)\right) \stackrel{d}{\Longrightarrow} \Gamma^\downarrow(X^{\kappa,t,\mathtt{c}}),$$

where

$$X^{\kappa,t,\mathtt{c}}(s) = \sqrt{\kappa}W(s) + ts - \frac{\kappa}{2}s^2 + \sum_{j=1}^\infty \left(c_j 1_{[\xi_j \le s]} - c_j^2 s\right) \tag{3.12}$$

for a collection $(\xi_i; i \ge 1)$ *of independent exponential random variables with respective rates* c_i.

We recall that the measure equipped on $\mathcal{G}(\mathtt{x}, q)$ is the measure assigning weights x_i to vertex i and so the theorems above do not tell us information about the number of vertices in the connected components of the inhomogeneous random graph. This is a subtle difference, but the sizes of the connected components do not follow from the work of [12] or [13], but would be useful for understanding component structure of the random graphs. Such results were obtained by Bhamidi et. al. in [32, 33]. We will not write out their results in detail.

Convergence of the connected components of the inhomogeneous random graph $\mathcal{G}(\mathtt{x}, q)$ under certain assumptions on $\mathtt{x}$ and on q were obtained in [30] where the limiting connected are analogous to those constructed for the critical Erdős-Rényi random graph. See also [28] where the edge probability is replace by $1-\exp(-k(x_i, x_j))$ for certain kernels $k : \mathbb{R}^2_+ \to \mathbb{R}_+$. Both of these deal with the Brownian results which are technically easier to handle because the construction of height processes a la [73] is more straight-forward (although still not trivial). See also the discussion of the configuration model below where descriptions in the case of *random* degree distributions was solved before the deterministic degree sequences.

In [44, 45] the authors describe the continuum limits of the critical inhomogeneous random graph *outside* of the Erdős-Rényi universality class. There are a few technical assumptions similar to those appearing in Theorem 3.1.1 and the construction of height processes which will will not mention; however, the assumptions on the weight sequences $\mathtt{x}^n$ can are stated as follows. There are two sequences a_n, b_n such that $x_1^n = O(a_n)$, a_n, $\frac{b_n}{a_n} \to \infty$ and $\frac{b_n}{a_n^2} \to \beta_0 \in [0, \infty)$ and

$$\frac{b_n}{a_n}\left(1 - \frac{\sigma_2(\mathtt{x}^n)}{\sigma_1(\mathtt{x}^n)}\right) \to \alpha \in \mathbb{R}, \quad \frac{b_n}{a_n^2}\frac{\sigma_3(\mathtt{x}^n)}{\sigma_1(\mathtt{x}^n)} \to \beta + \kappa \sum_j c_j^3, \quad \frac{x_j^n}{a_n} \to c_j,$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$, $\kappa \in (0,\infty)$ and $\mathtt{c} \in \ell^3_\downarrow$. They show

Theorem 3.3.3 (Broutin, Duquesne, Wang [44, 45])**.** *Let $\mathcal{G}_n = \mathcal{G}(\mathtt{x}^n, \frac{1}{\sigma_1(\mathtt{x}^n)})$. Under the assumptions mentioned above along with some technical assumptions not explicitly stated above, there exists limiting metric spaces $(\mathscr{G}_i; i \geq 1)$ such that*

$$\mathrm{scale}(\frac{a_n}{b_n}, \frac{1}{b_n}) \cdot \mathrm{COMP}(\mathcal{G}_n) \overset{d}{\Longrightarrow} (\mathscr{G}_i; i \geq 1),$$

in the product GHP topology. Here $\mathcal{G}_n$ is equipped with either the measure assigning weight x_i^n to vertex i or the counting measure. The graphs $\mathscr{G}_i$ are constructed from (a time-change and scaling) of the process in (3.12).

3.3.2 The configuration model

The configuration model was introduced by Bollobás in [39] to describe some asymptotic formulas appearing asymptotic enumeration results for regular random graphs. Fix a sequence of non-negative integers $\mathbf{d}^n = (d_1, \cdots, d_n)$. The graph $\mathrm{CM}(\mathbf{d}^n)$ is a random graph on the vertex set $[n]$ where each vertex i has degree $\deg(i) = d_i$ counted with multiplicity. Probabilistically the graph can be constructed as follows. View each vertex $i \in [n]$ having d_i many half-edges stemming from the vertex, and one then uniformly chooses a pairing of the $\sum_j d_j$ many half-edges. The edges in $\mathrm{CM}(\mathbf{d}^n)$ are then formed by turning each pair of half-edges into a bona-fide edge in $\mathrm{CM}(\mathbf{d}^n)$.

Apart from the combinatorial aspects of this model, it has become popular in recent decades because the vertices $\mathbf{d}^n$ can be chosen to match the degree distributions appearing in many real-world networks. See [173, Chapter 1] and references therein. Unfortunately, while the description of $\mathrm{CM}(\mathbf{d}^n)$ as a multigraph is straightforward; however, its construction *conditioned* on being a simple graph is not straight forward.

One of the first result on component structure for the configuration model was the threshold for a giant connected component of Molloy and Reed [148, 149]. In terms of the disease interpretation of the model, this is simply stating the probability of a large infection when just a single individual is infected at time $t = 0$.

Theorem 3.3.4 (Molloy and Reed [148, 149])**.** *Let $\mathbf{d}^n$ be a sequence of degree sequences* *and let* $\theta = \theta(\mathbf{d}^n)$ *be defined by*

$$\theta(\mathbf{d}^n) = \frac{\mathbb{E}\left[d_{V_n}(d_{V_n}-1)\right]}{\mathbb{E}[d_{V_n}]}, \qquad P(V_n = j) = \frac{1}{n}, \forall j \in [n].$$

Let $\mathscr{C}_n(1)$ *denote the largest connected component of* $\mathrm{CM}(\mathbf{d}^n)$. *Under certain additional assumptions on* $\mathbf{d}^n$ *not stated,*

$$\#\mathscr{C}_n(1) = \Theta_{\mathbb{P}}(n) \qquad \textit{if and only if} \qquad \theta(\mathbf{d}^n) > 1.$$

The value $\theta(\mathbf{d}^n) = 1$ is the critical value, where the behavior of the largest connected component drastically changes. The first work on the large n properties of connected components of the critical configuration model is due to Riordan [157] under the assumption that the degrees are bounded. Riordan shows that component sizes and surplus for the bounded degree configuration model falls into the same universality class as the critical Erdős-Rényi random graph. The result in [157] is similar to a result due to Joseph [109] for unbounded degree distributions but where additional randomness was included by taking the degree sequence itself to be random. Namely, suppose that $\mathbf{D}^n = (D_1, \cdots, D_n)$ are i.i.d. random variables are such that $\theta(D_1) := \mathbb{E}[D_1(D_1-1)]/\mathbb{E}[D_1] = 1$. Observe that $\theta(D_1)$ is simply the mean of the *size-biased* distribution D^* defined by

$$\mathbb{P}(D^* = k) = \frac{k\mathbb{P}(D_1 = k)}{\mathbb{E}[D_1]}.$$

Joseph [109] proved the following

Theorem 3.3.5 (Joseph [109]). *Let* $(\mathscr{C}_n(j); j \geq 1)$ *denote the connected components of* $\mathrm{CM}(\mathbf{D}^n)$ *where* $\theta(D_1) = 1$. *Then*

1. *Suppose that* $\mathbb{P}(D_1 = 1) < 1$ *and* $\mathbb{E}[D_1^3] < \infty$ *and let* $\mu = \mathbb{E}[D_1]$ *and* $\beta = \mathbb{E}[D_1(D_1 - 1)(D_1-2)]$. *Then*

$$\left(n^{-2/3}\#\mathscr{C}_n(1), n^{-2/3}\#\mathscr{C}_n(2), \cdots\right) \stackrel{d}{\Longrightarrow} \Gamma^{\downarrow}\left(\sqrt{\frac{\beta}{\mu}}W(t) - \frac{\beta}{2\mu^2}t^2\right) \qquad \textit{in } \ell^2$$

   for a Brownian motion W.

   The same result holds when the multigraph $\mathrm{CM}(\mathbf{D}^n)$ *is conditioned on being simple.*

2. Suppose that $\mathbb{P}(D_1 = k) \sim ck^{-\tau}$ *for some* $\tau \in (3,4)$ *and* $c > 0$ *and let* $\mu = \mathbb{E}[D_1]$*. Denote by* $X(t)$ *the process with independent increments and Laplace transform*

$$\mathbb{E}\left[e^{-\lambda X(t)}\right] = \exp\left(\int_0^t ds \int_0^\infty dx(e^{-\lambda x} - 1 + \lambda x)\frac{c}{\mu x^{\tau-1}}e^{-xs/\mu}\right),$$

for $\lambda, t \geq 0$*. Then*

$$n^{-(\tau-2)/(\tau-1)}\left(\#\mathscr{C}_n(1), \#\mathscr{C}_n(2), \cdots\right) \stackrel{d}{\Longrightarrow} \Gamma^{\downarrow}\left(X(t) - \frac{c\Gamma(4-\tau)}{(\tau-3)(\tau-2)\mu^{\tau-2}}t^{\tau-2}\right)$$

in ℓ^2*.*

The above theorem essentially confirms one aspect conjecture by statistical physicists stated at the beginning of this Section under the additional assumption that the degrees themselves are random. The result of Joseph does not establish the conjecture for the behavior of distances between two typical points even in the configuration model; however, his exploration of the graph together with the analogous results on random trees [73] strongly suggests that the distances scale appropriately. Joseph's result in the heavy-tailed also deals with only multigraphs instead of simple graphs. We will return to this after a brief foray into the work on deterministic degree sequences.

This random degree distribution assumption is not made by Riordan in [157] and, in fact, Riordan actual describes a critical window for the bounded degree configuration model. Riordan's work falls much more in line with the work of Dhara et. al. in their two works [62, 63]. All three of these works describe the critical window for the configuration model with deterministic degrees provided that the empirical distribution of the degrees satisfy some asymptotic properties.

Theorem 3.3.6 (Riordan [157], Dhara et. al. [62, 63])**.** *For each* $n \geq 1$*, let* $\mathbf{d}^n = (d_1, d_2, \cdots, d_n)$ *be a degree sequence listed in decreasing order. Let* D_n *denote the degree of a uniformly chosen vertex in* $\mathrm{CM}(\mathbf{d}^n)$*. Suppose that*

$$D_n \stackrel{d}{\Longrightarrow} D, \qquad \text{as } n \to \infty,$$

for some random variable D *such that* $\mathbb{P}(D = 1) > 0$ *and that*

$$\theta(\mathbf{d}^n) = \frac{\mathbb{E}[D_n(D_n-1)]}{\mathbb{E}[D_n]} = 1 + \lambda r_n^{-1} + o(r_n^{-1})$$

for some fixed $\lambda \in \mathbb{R}$ *and sequence* $r_n \to \infty$*.*

1. *Suppose that* $\mathbb{E}[D_n^3] \to \mathbb{E}[D^3] < \infty$ *as* $n \to \infty$ *and* $r_n = n^{1/3}$. *Then as* $n \to \infty$

$$n^{-2/3} \left(\#\mathscr{C}_n(1), \#\mathscr{C}_n(2), \cdots\right) \overset{d}{\Longrightarrow} \Gamma^{\downarrow}\left(\frac{\sqrt{\eta}}{\mu} W(t) + \lambda t - \frac{\eta}{2\mu^3} t^2\right) \qquad \text{in } \ell^2,$$

where $\mu = \mathbb{E}[D]$, *and* $\eta = \mu\mathbb{E}[D^3] - \mathbb{E}[D^2]^2$.

2. *Let* $L(x)$ *be a slowly-varying function and* $\tau \in (3,4)$. *Suppose that (1) for each fixed* $i \geq 1$, *that*

$$\frac{d_i}{n^{1/(\tau-1)} L(n)} \to c_i > 0, \qquad \mathsf{c} = (c_i; i \geq 1) \in \ell^3_\downarrow \setminus \ell^2_\downarrow;$$

(2) that $\mathbb{E}[D_n] \to \mu := \mathbb{E}[D]$, $\mathbb{E}[D_n^2] \to \mathbb{E}[D^2]$ *and*

$$\lim_{K\to\infty} \limsup_{n\to\infty} (n^{1/(\tau-1)} L(n))^{-3} \sum_{j \geq K+1} d_j^3 = 0;$$

and (3) that $r_n = n^{(\tau-3)/(\tau-1)}/L(n)^2$. *Then*

$$\frac{L(n)}{n^{(\tau-2)/(\tau-1)}} \left(\#\mathscr{C}_n(1), \mathscr{C}_n(2), \cdots\right) \overset{d}{\Longrightarrow} \Gamma^{\downarrow}\left(X^{0,t,\mathsf{c}}\right), \qquad \text{in } \ell^2,$$

where $X^{\kappa,t,\mathsf{c}}$ *is as in* (3.12) *where* $t = \lambda \mu^{1/(1-\tau)}$.

The above results continue to hold conditionally given $\mathrm{CM}(\mathbf{d}^n)$ *is simple.*

Let us note that the above theorem establishes the conjecture for component sizes of the critical and near-critical configuration model, with deterministic degree sequences; however, it does not establish the description of the distance between two typical points within the largest connected components.

To understand the typical distances between vertices, we need to know the limiting metric space structure of the critical configuration model. This is due to Conchon-Kerjan and Goldschmidt for the case $\tau \in (3,4)$ and *random* degree distributions. Before we state their result in the following theorem, we recall some properties of excursions and continuum random graphs. Recall that we can generate unit-duration excursions $(\mathbf{e}, \mathbf{h})$ of the (X, H) where X is an $\alpha \in (1,2]$ stable Lévy process and H is its associated height process. In the Brownian case, $\alpha = 2$, the excursion $\mathbf{e}$ is taken to be a standard Brownian excursion and $\mathbf{h} = 2\mathbf{e}$. In the α-stable case ($\alpha \neq 2$) the excursion $\mathbf{e}$ depends on the constant C in $\mathbb{E}[\exp(-\lambda X(t))] = \exp(Ct\lambda^\alpha)$.

In the construction of continuum random graphs $\mathcal{G}(\mathbf{e}, \mathbf{h}, \mathscr{P})$ there is a point-set $\mathscr{P} \cap \mathbf{e}$ which plays an important role in the quotienting procedure. In particular, $\#\mathscr{P} \cap \mathbf{e}$ is the number of vertex identifications made in the tree $\mathcal{T}_{\mathbf{h}}$. This is the *surplus* of the continuum random graph $\mathcal{G} = \mathcal{G}(\mathbf{e}, \mathbf{h}, \mathscr{P})$ and we will denote this by $s(\mathcal{G})$.

For $\alpha \in (1, 2]$, we will call $\mathscr{G}^{(\alpha,k)}$ the α-stable continuum random graph[2] with fixed surplus k by the graph

$$\mathscr{G}^{(\alpha,k)} := \mathcal{G}(\tilde{\mathbf{e}}^{(k)}, \tilde{\mathbf{h}}^{(k)}, \mathscr{P}^{(k)})$$

where $\mathscr{P}^{(k)}$ are k uniformly chosen points in $\{(t, x) \in [0, 1] \times \mathbb{R}_+ : 0 \leq x \leq \tilde{\mathbf{e}}^{(k)}(t)\}$ and

$$\mathbb{E}\left[F(\tilde{\mathbf{e}}^{(k)}, \tilde{\mathbf{h}}^{(k)})\right] = \frac{\mathbb{E}\left[F(\mathbf{e}, \mathbf{h})\left(\int_0^1 \mathbf{e}(t)\, dt\right)^k\right]}{\mathbb{E}\left[\left(\int_0^1 \mathbf{e}(t)\, dt\right)^k\right]}$$

where $(\mathbf{e}, \mathbf{h})$ as above.

Theorem 3.3.7 (Conchon-Kerjan & Goldschmidt [59])**.** *Suppose that $\mathbf{D}^n$ is as in Theorem 3.3.5 and for each connected component $\mathscr{C}_n(j)$ in $\mathrm{CM}(\mathbf{D}^n)$, let $\mathsf{o}_j \in \mathscr{C}_n(j)$ be a randomly chosen vertex with probability proportional to its degree. Let $\mu_{i,n}$ denote the counting measure on $\mathscr{C}_n(i)$. In the respective critical regimes:*

1. *As $n \to \infty$, in the product GHP topology*

$$\left(\left(\mathscr{C}_n(i), \mathsf{o}_i, n^{-1/3} d, n^{-2/3} \mu_{i,n}\right); i \geq 1\right) \xrightarrow{d} (\mathscr{G}_i; i \geq 1),$$

where $\mathscr{G}_i$ continuum random graphs.

2. *As $n \to \infty$, in the product GHP topology*

$$\left(\left(\mathscr{C}_n(i), \mathsf{o}_i, n^{-(\alpha-1)/(\alpha+1)} d, n^{-\alpha/(\alpha+1)} \mu_{i,n}\right); i \geq 1\right) \xrightarrow{d} (\mathscr{G}_i; i \geq 1), \qquad \alpha = \tau - 2,$$

for continuum metric spaces $\mathscr{G}_i$.

---

[2]This graph depends on the constant C in the Laplace transform of the stable process X; however, changing the constant C only rescales the metric in the graph.

Moreover, the graphs $(\mathscr{G}_i; i \geq 1)$ *are i.i.d. conditionally given* $(\zeta_i = \mu_i(\mathscr{G}_i), k_i = s(\mathscr{G}_i))_{i\geq 1}$ *with law*

$$(\mathscr{G}_j | (\zeta_i, k_i)_{i\geq 1}) \overset{d}{=} \operatorname{scale}(\zeta_j^{(\alpha-1)/\alpha}, \zeta_j) \cdot \left(\mathscr{G}^{(\alpha,k_i)}\right).$$

Remark 3.3.1. Each of the works of [59, 62, 63, 109, 157] explore the connected components of the configuration model in a manner very similar to how Aldous explores the near-critical Erdős-Rényi random graph in [12]. There are, of course, subtle differences related to particular method of constructing the configuration model which is highlighted in Chapter 5 below.

3.4 Random Matrix Theory

The results in Chapter 6, recalled from [58], elaborate on some results about random trees that originally appeared in the study of random matrices [97, 126]. To give some of a background on this result, we will briefly mention some results in this line of work.

3.4.1 Why Eigenvalues and Gaussian Ensembles?

One of the motivations for the early study of random matrices comes from physics, namely understanding energy levels of atomic nuclei. The energy levels of a quantum system are supposed to be described by the eigenvalues of the Hamiltonian operator H, which is some Hermitian operator on some infinite dimensional Hilbert space. Since one is concerned with the eigenvalues and not the spectrum of the operator H, physicists were concerned with understanding the behavior of eigenvalues of $N \times N$ matrices where N is large. For many systems the exact Hamiltonian H is not known, and so instead one considers the entries of the matrix H to be *random* subject to certain physical symmetry constraints. See Chapter 1, Section 1 of [144] and references therein for more physical motivations.

The symmetry of physical systems, see [80], led early authors to consider $N\times N$ matrices $X_N = (X_N(\ell, r) : \ell, r \in [N])$ with entries in some (skew) field $\mathbb{F}$ taking particular forms. The three (skew) fields considered are the field of real numbers $\mathbb{R}$, the field of complex numbers $\mathbb{C}$, and the skew field (non-commutative) quaternions $\mathbb{H}$. These matrices took the form [76]

$$X_N(\ell, r) = \sum_{\alpha=0}^{\beta-1} X_{N,\alpha}(\ell, r) e_\alpha \tag{3.13}$$

where $e_0 = 1$, $e_1 = \mathbf{i} \in \mathbb{C}$, $e_2 = \mathbf{j}, e_3 = \mathbf{k} \in \mathbb{H}$ are particular elements in the (skew) field $\mathbb{F}$, $\beta = \dim_{\mathbb{R}}(\mathbb{F})$ is the dimension of $\mathbb{F}$ as an $\mathbb{R}$-vector space and $X_{N,\alpha}(\ell, r)$ are real numbers There is the additional constraint that the matrices are Hermitian over the field $\mathbb{F}$ which means

$$X_{N,0}(\ell, r) = X_{N,0}(r, \ell) \qquad \text{and} \qquad X_{N,\alpha}(\ell, r) = -X_{N,\alpha}(r, \ell) \text{ for } \alpha \neq 0. \tag{3.14}$$

The Gaussian orthogonal/unitary/symplectic ensemble (GOE/GUE/GSE) is the family of $N \times N$ Hermitian matrices $X_N = (X_N(\ell, r); \ell, r \in [N])$ with independent (subject to (3.14)) entries in $\mathbb{R}/\mathbb{C}/\mathbb{H}$ with entries defined by (3.13) where $X_{N,\alpha}(\ell, r)$ are independent and distributed as

$$X_{N,0}(\ell, \ell) \sim \mathscr{N}\left(0, \frac{2}{\beta}\right), \qquad \text{and} \qquad X_{N,\alpha}(\ell, r) \sim \mathscr{N}\left(0, \frac{1}{\beta}\right)$$

where $\beta = 1/\beta = 2/\beta = 4$.

It can be shown, [76] and Theorem 2.5.2 in [17], that if $\lambda_1^N \leq \lambda_2^N \leq \cdots \leq \lambda_N^N$ are the eigenvalues of GO/U/SE they're joint law has a density on the set $\{x_1 \geq x_2 \geq \cdots \geq x_N\}$

$$\mathbb{P}\left(\lambda_1^N \in dx_1, \cdots, \lambda_N^N \in dx_N\right) \propto \left(\prod_{i<j} |x_i - x_j|^\beta\right) \exp\left(-\frac{\beta}{4} \sum_{j=1}^N x_j^2\right) dx_1\, dx_2 \cdots dx_N \tag{3.15}$$

where $\beta = 1,\ 2,\ 4$ respectively.

For each $\beta = 1, 2, 4$ there is a family of real $N \times N$ tridiagonal matrices whose eigenvalues match the eigenvalues of GO/U/SE matrices. In fact, there is actually a generalization of the Gaussian ensembles to all $\beta > 0$ and not just $\beta = 1, 2, 4$ due to Dumitriu and Edelman [69] which we now state.

Theorem 3.4.1 (Dumitriu & Edelman [69]). *Fix a $\beta > 0$. Let*

$$M_N^\beta = \frac{1}{\sqrt{\beta}} \begin{pmatrix} \sqrt{2}g_1 & \chi_{(N-1)\beta} & 0 & \dots & \\ \chi_{(N-1)\beta} & \sqrt{2}g_2 & \chi_{(N-2)\beta} & 0 & \dots \\ 0 & \chi_{(N-2)\beta} & \sqrt{2}g_3 & \ddots & \\ \vdots & & \ddots & \ddots & \\ & & & & \chi_\beta \\ & & & \chi_\beta & \sqrt{2}g_N \end{pmatrix}, \qquad \text{for } \beta > 0, N \geq 1, \tag{3.16}$$

where $\chi_\beta, \chi_{2\beta}, \cdots \chi_{(N-1)\beta}$ are independent χ-random variables which are indexed by their parameters and $g_1, \dots, g_N$ are independent standard normal random variables. Then the eigenvalues of M_N^β have density (3.15).

We call such matrices the Gaussian β ensemble (GβE).

3.4.2 The Semi-circle Law

A major object of study in random matrix theory is the study of eigenvalues of $N \times N$ matrices. We will focus mostly on real symmetric $N \times N$ matrices, which we will denote for a fixed N by $X_N = (X_N(i,j); i,j \in [N])$. We will also denote its N real eigenvalues by $\lambda_1^N \geq \lambda_2^N \geq \dots \geq \lambda_N^N$, or, equivalently, the (scaled) empirical measure

$$\rho_N(dx) = \frac{1}{N}\sum_{j=1}^N \delta_{\lambda_j^N/\sqrt{N}}(dx), \qquad \int f(x)\,\rho_N(dx) = \frac{1}{N}\sum_{j=1}^N f\left(\frac{\lambda_j^N}{\sqrt{N}}\right).$$

As can be visually seen from Figure 3.1, the eigenvalues for a particular sample appear to mimic a particular function. This is not a coincidence, but is a general fact due originally to Wigner [181]. This is so-called semi-circle law stated in the next theorem. The proof of this result, under the following formulation, can be found in Chapter 2 of [17] and [69].

Theorem 3.4.2 (Wigner [181]). *Let Y and Z be mean zero random variables with* $\mathrm{Var}(Z) = 1$. *Suppose that $(X_N(i,j) = X_N(j,i); i > j, N \geq 1)$ are i.i.d. copies of Z and $(X_N(i,i); i \geq 1, N \geq 1)$ are i.i.d. copies of Y. If*

$$\mathbb{E}[|Z|^k] + \mathbb{E}[|Y|^k] < \infty, \qquad \forall k \geq 1,$$

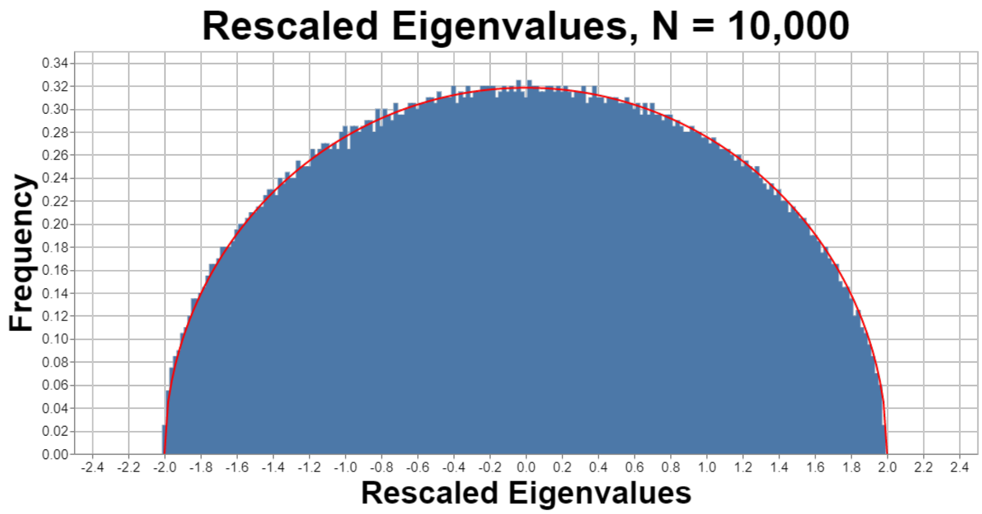


Figure 3.1: In blue: Histogram of the eigenvalues of a 10,000×10,000 matrix. In red: The semi-circle $\frac{1}{2\pi}\sqrt{4-x^2}$.

then

$$\rho_N \xrightarrow[N\to\infty]{\mathbb{P}} \rho(dx) := \frac{\sqrt{4-x^2}}{2\pi}\, 1_{[-2\leq x\leq 2]}\, dx,$$

where the measures are topologized with the topology of weak convergence of finite measures.

The same convergence holds if X_N is the $N\times N$ $G\beta E$ matrix.

Recall that one of the early motivations for studying random matrices was to understand the spacing of energy levels of some complicated quantum systems, but that the above theorem does not quite give us this. What Theorem 3.4.2 does give us is that the proportion of eigenvalues that fall outside of $[-(2+\varepsilon)\sqrt{N},(2+\varepsilon)\sqrt{N}]$ tends to zero as $N\to\infty$ for each fixed $\varepsilon>0$. In particular, the N eigenvalues are spaced in an interval of size roughly $4\sqrt{N}$ and hence should the spacing between adjacent eigenvalues should be of order $N^{-1/2}$. These are concerned with the local properties of eigenvalues, and there are essentially two regimes that are studied: the *bulk* and the *edge*. The bulk eigenvalues are those eigenvalues which fluctuate around $u\sqrt{N}$ for $u\in(-2,2)$ and the edge are those concerned with what happens for the largest (or smallest) eigenvalues. Since Chapter 6 is related to the edge eigenvalues and so we will mostly focus on the behavior edge statistics of the Gaussian β

ensemble.

3.4.3 The Edge Statistics of GβE

Let M_N^β be as in (3.16) be in the $N \times N$ GβE and denote its eigenvalues by $\lambda_{\beta,1}^N \geq \cdots \geq \lambda_{\beta,N}^N$. Note that for each fixed $k \geq 1$, the largest k eigenvalues should all roughly of size $2\sqrt{N}$ by the semi-circle law; however, it does not tell us the order of $(2\sqrt{N} - \lambda_{\beta,1}^N)$ as $N \to \infty$. That is, the semi-circle law does not tell us the fluctuations of the largest eigenvalues.

The first result dealing with such fluctuations were due to Tracy and Widom in [171] for the GUE case $\beta = 2$ and then shortly after in [172] for the GO/SE cases $\beta = 1, 4$. While they were concerned with a different scaling of the eigenvalues, these two theorems are formulated as follows

Theorem 3.4.3 (Tracy & Widom [171, 172]). *For $\beta \in \{1, 2, 4\}$ there exists a distribution function $F_\beta(x)$, called the Tracy-Widom(β) distribution such that*

$$F_\beta(s) = \lim_{n\to\infty} \mathbb{P}\left(N^{1/6}(\lambda_{\beta,1}^N - 2\sqrt{N}) \leq s\right).$$

In particular, for each $\beta \in \{1, 2, 4\}$ both $(\lambda_{\beta,1}^N - 2\sqrt{N}) = \Theta(n^{-1/6})$ and there exists a limit random variable $\Lambda_{\beta,1}$ such that

$$N^{1/6}(2\sqrt{N} - \lambda_{\beta,1}^N) \overset{d}{\Longrightarrow} \Lambda_{\beta,1}, \qquad \text{as } N \to \infty. \tag{3.17}$$

About a decade after the results by Tracy and Widom, Sutton [169] and then jointly with Edelman [82] gave a heuristic argument that the $N \times N$ re-centered and re-scaled matrix

$$\tilde{M}_N^\beta := N^{1/6}\left(2\sqrt{N} I_{N\times N} - M_N^\beta\right)$$

can be approximated in the $N \to \infty$ limit by stochastic operator

$$\mathscr{H}^\beta f := \left(-\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} W_x'\right) f, \qquad f(0) = 0, \quad f \in L^2(\mathbb{R}_+), \tag{3.18}$$

where W_x' is a white noise. If we let The operator $\mathscr{H}^\beta$ is called the *stochastic Airy operator*. This connection was made precise in the work of Ramírez, Rider and Virág [155].

Let $\Lambda_{\beta,1}^N \leq \Lambda_{\beta,2}^N \leq \cdots \Lambda_{\beta,N}^N$ denote the eigenvalues of $\tilde{M}_N^\beta$ listed in *increasing* order. Note that $\Lambda_{\beta,j} = N^{1/6}(2\sqrt{N} - \lambda_{\beta,j}^N)$ for each $j \in [N]$. Consequently, $\Lambda_{\beta,1}^N \overset{d}{\Longrightarrow} \Lambda_{\beta,1}$ for each

$\beta \in \{1, 2, 4\}$ as in (3.17). It can be shown [155] that the operator $\mathscr{H}^\beta$ has well-defined smallest eigenvalues which can be defined by $\Lambda_{\beta,1} \leq \Lambda_{\beta,2} \leq \cdots$. The connection between the GβE and the stochastic Airy operator is made precise in the following theorem:

Theorem 3.4.4 (Ramírez, Rider & Virág [155]). *For each fixed k, and $\beta > 0$ then jointly as $N \to \infty$*

$$(\Lambda_{\beta,1}^N, \cdots, \Lambda_{\beta,k}^N) \overset{d}{\Longrightarrow} (\Lambda_{\beta,1}, \cdots, \Lambda_{\beta,k})\,.$$

The above theorem holds for more general joint eigenvalue distributions as well, see [125] and references therein for a more precise formulation.

Let us remark that the above theorem holds for all $\beta > 0$ and not just $\beta \in \{1, 2, 4\}$ as was the case for the Tracy-Widom distribution. This situation was the consideration of Gorin and Shkolnikov in their work [97].

Therein, they showed that the operator $\frac{1}{2}\mathscr{H}^\beta$ generates a semigroup $(\mathcal{U}^\beta(t); t \geq 0)$ of trace class integral operators on $L^2(\mathbb{R}_+)$ defined via a kernel $K^\beta(x, y; t)$ for $t > 0$. Let us make this connection more precise. For each t, let $B^{x,y,t} = (B^{x,y,t}(s); s \in [0, t])$ denote a Brownian bridge from x to y of duration t. Let $(L_v^{x,y,t}; v \in \mathbb{R})$ denote a continuous version of the local time of $B^{x,y,t}$ at time t and level v. Finally, let $W = (W(v); v \geq 0)$ denote an independent standard Brownian motion. The kernel $K^\beta(x, y; t)$ is defined by

$$K^\beta(x, y; t) = \frac{e^{-(x-y)^2/2t}}{\sqrt{2\pi t}} \mathbb{E}_B \left[1_{[B^{x,y,t} \geq 0]} \exp\left\{ -\frac{1}{2} \int_0^t B^{x,y,t}(s)\, ds + \frac{1}{\sqrt{\beta}} \int_0^\infty L_v^{x,y,t}\, dW(v) \right\} \right],$$

where $\mathbb{E}_B$ means that the expectation is taken only with respect to $B^{x,y,t}$ and $f \geq 0$ means $f(s) \geq 0$ for all s. Some care has to be taken to define the stochastic integral above in order to consider multiple operators $\mathcal{U}^\beta(t_1), \cdots, \mathcal{U}^\beta(t_m)$, [113].

Theorem 3.4.5 (Gorin & Shkolnikov [97]). *Let $(e^{-\frac{t}{2}\mathscr{H}^\beta}; t \geq 0)$ denote the semi-group generated by $\frac{1}{2}\mathscr{H}^\beta$. One can couple the Brownian motions W in their respective definitions to make*

$$\mathbb{P}\left(e^{-\frac{t}{2}\mathscr{H}^\beta} = \mathcal{U}^\beta(t) \right) = 1, \qquad \forall t \geq 0,$$

where

$$\left(\mathcal{U}^\beta(t) f \right)(x) = \int_0^\infty K^\beta(x, y; t) f(y)\, dy.$$

For an extension of this result see the joint work by Lamarre and Shkolnikov [126]. The connection with a Brownian excursion, and not a Brownian bridge, will be discussed in Chapter 6.

Chapter 4

EPIDEMICS ON CRITICAL ERDŐS-RÉNYI RANDOM GRAPH

4.1 Introduction

Recall that the Erdős-Rényi random graph $G(n,p)$ is a random graph on n vertices labeled $\{1,\cdots,n\}$ where each edge $\{i<j\}$ is included independently with probability p. We imagine the graph $G(n,p)$ as modeling a population on n vertices through which a disease spreads in discrete time. On day $t=0$, there are k many individuals infected and the rest of the population is susceptible. On each day $t=0,1,\cdots$, each infected individual infects their susceptible neighbors who become infected on day $t+1$ and each infected individual on day t recovers on day $t+1$.

Let us write $\{\rho_1,\cdots,\rho_k\}$ for the k many infected individuals on day zero. Since relabeling vertices does not change the edge probability, we can take these k vertices to be the vertices labeled $1,\cdots,k$; however, we prefer to stick with former choice. If we equip $G(n,p)$ with the graph distance, then the vertices who are infected on day t are simply those vertices

$$\left\{v\in G(n,p):\min_{j\in[k]} d(v,\rho_j)=t\right\}.$$

If we let

$$Z_n^k(t)=\#\left\{v\in G(n,p):\min_{j\in[k]} d(v,\rho_j)=t\right\}, \tag{4.1}$$

then it is not hard to convince yourself that Z_n^k is not Markov because, for example, the more people who are initially infected at time $t-1$ limiting the number of individuals who can become infected on day $t+1$.

However, we can consider the $\mathbb{Z}^2$-valued (Z_n^k,C_n^k), where $C_n^k(t)=\sum_{\ell=0}^t Z_n^k(\ell)$, and see that this is a Markov chain. More explicitly, we can see that transitions completely determined by

$$\left(Z_n^k(t+1)|Z_n^k(t)=z,C_n^k(t)=c\right)\overset{d}{=}\operatorname{Bin}\left(n-c,1-(1-p)^z\right).$$

Indeed, conditioning on $Z_n^k(t) = z$ and $C_n^k(t) = c$ implies that there are $n - c$ susceptible individuals remaining in the population. Each of those $n-c$ individuals shares an edge with each of the z infected individuals independently with probability p. Thus, the probability the a particular susceptible vertex does *not* share an edge with any of the z infected individuals (and hence does not become infected on day $t+1$) is $(1-p)^z$.

We are interested in understanding the behavior of the SIR epidemic in within the critical window [12] of the Erdős-Rényi random graph. That is when $p = p(n) = n^{-1} + \lambda n^{-4/3}$. The main theorem is the following

Theorem 4.1.1. *Fix a $\lambda \in \mathbb{R}$ and a sequence $k = k(n)$ such that $kn^{-1/3} \to x \geq 0$ as $n \to \infty$. Suppose that $p = p(n) = n^{-1} + \lambda n^{-4/3}$. Then, as $n \to \infty$, the following weak convergence holds in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$:*

$$\left(n^{-1/3} Z_n^k(\lfloor n^{1/3} t \rfloor); t \geq 0\right) \stackrel{d}{\Longrightarrow} (\mathbf{Z}(t); t \geq 0),$$

where $\mathbf{Z}$ is the unique strong solution to the stochastic integral equation

$$\mathbf{Z}(t) = x + \int_0^t \sqrt{\mathbf{Z}(s)}\, dW(s) + \left(\lambda - \frac{1}{2}\int_0^t \mathbf{Z}(s)\, ds\right) \int_0^t \mathbf{Z}(s)\, ds \tag{4.2}$$

for a standard Brownian motion W.

We mention that similar results using simpler, continuous-time, epidemic models were studied in [66, 164]. We prove the above theorem by establishing the hypotheses of the martingale functional central limit theorem [86, Chapter 7] in Lemmas 4.2.2 and 4.2.3 below.

The formulation of the limiting process $\mathbf{Z}$ in Theorem 4.1.1 is not a standard stochastic differential equation because of the factor of $\left(\int_0^t \mathbf{Z}(s)\, ds\right)^2$. However, this is not as surprising when we consider the following lemma.

Lemma 4.1.2. *The process $\mathbf{Z}$ solving (4.2) is equal in law to the pathwise unique solution to*

$$\mathbf{Z}(t) = x + \mathbf{X}^\lambda\left(\int_0^t \mathbf{Z}(s)\, ds\right), \qquad \mathbf{X}^\lambda = B(t) + \lambda t - \frac{1}{2}t^2, \tag{4.3}$$

where B is a linear Brownian motion.

The process $\mathbf{X}^\lambda$ is a Brownian motion with parabolic drift and appears in the seminal work of Aldous [12] on the Erdős-Rényi random graph.

Before moving onto the proof of Theorem 4.1.1 and Lemma 4.1.2, let us mention what happens when just a single individual is initially infected, i.e. $k = 1$. This is a much more natural setting when there is initially just a single source of the disease. Obviously in this case, the limiting process $\mathbf{Z}$ solves (4.3); however, the unique solution to the resulting stochastic equation in (4.3) (with $x = 0$) is the process $\mathbf{Z}(t) \equiv 0$. This may lead the reader to believe that the $(n^{1/3}, n^{1/3})$ time-space scaling is not the correct scaling window to examine for this situation; however, based on the results we prove in Chapter 5 we believe this is incorrect. We will show in a much more general context in the next chapter that when just a single individual is infected in one of the *largest* connected components then there is a non-trivial limit of the resulting process $\mathbf{Z}$.

4.2 Technical Lemmas

In this section we provide lemmas necessary to go from the convergence in [66, 164] of a continuous-time epidemic model to the statement presented in Theorem 4.1.1. This is hinted at in [177, Appendix 2] as well although not proved. We fix a $\lambda \in \mathbb{R}$ and let $\mathcal{G}_n = G(n, n^{-1} + \lambda n^{-4/3})$ denote an Erdős-Rényi random graph, and let Z_n^k be defined by (4.1). For convenience, we let $C_n^k = \big(C_n^k(h); h = 0, 1, \cdots\big)$ be defined by

$$C_n^k(h) = \sum_{j=0}^{h} Z_n^k(j). \tag{4.4}$$

In terms of the graph $\mathcal{G}_n$, $C_n^k(h)$ represents the number of vertices within distance h of the k randomly selected vertices $\{\rho_n(1), \cdots, \rho_n(k)\}$. In terms of the SIR model, $C_n^k(h)$ represents the number of individuals who have contracted the disease at or before "time" h.

From the correspondence of the Reed-Frost model and the Erdős-Rényi random graph, we know that $(Z_n^k(h), C_n^k(h))$ is a Markov chain with state space

$$\mathcal{S} = \{(z, c) \in \mathbb{Z}^2 : z, c \geq 0\}$$

which is absorbed upon hitting the line $\{(0, c) : c \geq 0\}$. Moreover, the conditional distribution of $Z_n^k(h+1)$ given $\big(Z_n^k(h), C_n^k(h)\big)$ is

$$\left(Z_n^k(h+1)\middle| Z_n^k(h)=z, C_n^k(h)=c\right) \stackrel{d}{=} \begin{cases} \mathrm{Bin}(n-c, q(n,z)) & : z>0, c<n \\ 0 & : \text{else} \end{cases}, \tag{4.5}$$

where $q(n,z)$ is defined by

$$q(n,z) = 1 - \left(1 - n^{-1} - \lambda n^{-4/3}\right)^z. \tag{4.6}$$

The joint conditional distribution of $(Z_n^k(h+1), C_n^k(h+1))$ is easily deduced from equations (4.5) and (4.4).

4.2.1 Asymptotics for binomial statistics

We begin by examining the binomial random variables

$$\beta(n,z,c) \stackrel{d}{\sim} \mathrm{Bin}(n-c, q(n,z)),$$

where $q(n,z)$ is defined by (4.6). Examining the convergence in Theorem 4.1.1, we'll study various statistics of $\beta(n,z,c)$ as $n\to\infty$ with $z = O(n^{1/3})$ and $c = O(n^{2/3})$.

We define the following statistics

$$\mu(n,z,c) = \mathbb{E}\left[\beta(n,z,c)\right], \quad \sigma^2(n,z,c) = \mathrm{Var}\left[\beta(n,z,c)\right], \quad \kappa(n,z,c) = \mathbb{E}\left[(\beta(n,z,c)-z)^4\right]. \tag{4.7}$$

The main purpose of this subsection is to establish the following lemma:

Lemma 4.2.1. *Fix an $r>0$ and $T>0$ and let*

$$\Omega_n = \Omega(n,r,T) = \left\{(z,c)\in\mathbb{Z}^2 : 0\le z\le n^{1/3}r, 0\le c\le n^{2/3}Tr\right\}.$$

Then, as $n\to\infty$, the following bounds hold:

$$\begin{aligned} &\sup_{\Omega_n}\left|\mu(n,z,c) - z - n^{-1/3}z(\lambda - n^{-2/3}c)\right| = O(n^{-1/3}) \\ &\sup_{\Omega_n}\left|\sigma^2(n,z,c) - z - n^{-1/3}z(\lambda - n^{-2/3}c)\right| = O(n^{-1/3}) \\ &\sup_{\Omega_n}\left|\kappa(n,z,c)\right| = O(n^{2/3}) \end{aligned} \tag{4.8}$$

In particular,

$$\begin{aligned} &\sup_{\Omega_n} |\mu(n,z,c) - z| = O(1) \\ &\sup_{\Omega_n} \left|\sigma^2(n,z,c) - z\right| = O(1) \end{aligned}, \qquad \text{as } n \to \infty.$$

Proof. We prove the statements in equation (4.8), since the latter bounds easily follow from the more detailed asymptotics.

We start with the expansion of $\mu(n,z,c)$. The binomial theorem gives

$$\begin{aligned} \mu(n,z,c) &= (n-c)\left(1-(1-n^{-1}-\lambda n^{-4/3})^z\right) \\ &= (n-c)\left(z(n^{-1}+\lambda n^{-4/3}) - \sum_{j=2}^{z}\binom{z}{j}(-1)^j(n^{-1}+\lambda n^{-4/3})^j\right) \\ &= z + n^{-1/3}z(\lambda - n^{-2/3}c) - \lambda n^{-4/3}zc - (n-c)\sum_{j=2}^{z}\binom{z}{j}(-1)^j(n^{-1}+\lambda n^{-4/3})^j. \end{aligned}$$

For n sufficiently large, we can obtain the bounds

$$\begin{aligned} \left|\mu(n,z,c) - z - n^{-1/3}z(\lambda - n^{-2/3}c)\right| &\le |\lambda| n^{-4/3}zc + (n-c)\left|\sum_{j=2}^{z}\binom{z}{j}(-1)^j(n^{-1}+\lambda n^{-4/3})^j\right| \\ &\le |\lambda| n^{-4/3}zc + n\sum_{j=2}^{z}\binom{z}{j}(n^{-1}+\lambda n^{-4/3})^j \\ &\le |\lambda| n^{-4/3}zc + n\sum_{j=2}^{z}\left(\frac{2ez}{n}\right)^j. \end{aligned}$$

In the second and third inequality above, we used the bound $0 < n^{-1}+\lambda n^{-4/3} \le 2n^{-1}$ and the bound $\binom{m}{k} \le (em)^k$.

Taking the supremum over Ω_n, gives

$$\begin{aligned} \sup_{\Omega_n}\left|\mu(n,z,c) - z - n^{-1/3}z(\lambda - n^{-2/3})\right| &\le n^{-1/3}|\lambda| Tr^2 + n\sum_{j=2}^{n^{1/3}r}\left(\frac{2er}{n^{2/3}}\right)^j \\ &\le n^{-1/3}|\lambda| Tr^2 + n\left(\frac{4e^2r^2n^{-4/3}}{1-2ern^{-2/3}}\right) \\ &\le \left(|\lambda| Tr^2 + 8e^2r^2\right)n^{-1/3} = O(n^{-1/3}). \end{aligned}$$

This proves the desired expansion and bound for $\mu(n,z,c)$.

We now examine the bounds for $\sigma^2(n, z, c)$. Again, we use the binomial theorem

$$\begin{aligned}
\sigma^2(n, z, c) &= \mu(n, z, c)(1 - n^{-1} - \lambda n^{-4/3})^z \\
&= \mu(n, z, c) \left(1 - z \left(n^{-1} + \lambda n^{-4/3} \right) + \sum_{j=2}^{z} \binom{z}{j} (-1)^j (n^{-1} + \lambda n^{-4/3})^j \right) \\
&= \mu(n, z, c) - \mu(n, z, c) z (n^{-1} + \lambda n^{-4/3}) + \mu(n, z, c) \sum_{j=2}^{z} \binom{z}{j} (-1)^j (n^{-1} + \lambda n^{-4/3})^j.
\end{aligned} \tag{4.9}$$

We can then use the previous asymptotic bounds for $\mu(n, z, c)$ to get

$$\sup_{\Omega_n} \left| \sigma^2(n, z, c) - z - n^{-1/3} z (\lambda - n^{-2/3}) \right| \le \sup_{\Omega_n} |\sigma^2(n, z, c) - \mu(n, z, c)| + O(n^{-1/3}).$$

We can bound the first term on the right-hand side as we did for $\mu(n, z, c)$ above. We get

$$\begin{aligned}
\sup_{\Omega_n} \left| \sigma^2(n, z, c) - \mu(n, z, c) \right| &\le \sup_{\Omega_n} \left| \mu(n, z, c) z (n^{-1} + \lambda n^{-4/3}) + n \sum_{j=2}^{z} \binom{z}{j} (n^{-1} + \lambda n^{-4/3})^j \right| \\
&\le \sup_{\Omega_n} \left(2rn^{-2/3} \mu(n, z, c) + n \sum_{j=2}^{z} \binom{z}{j} (n^{-1} + \lambda n^{-4/3})^j \right) \\
&\le 2rn^{-2/3} \sup_{\Omega_n} \left(|\mu(n, z, c) - z - n^{-1/3} z (\lambda - n^{-2/3} c)| + |z + n^{-1/3} z (\lambda - n^{-2/3} c)| \right) \\
&\quad + O(n^{-1/3}) \\
&= 2rn^{-2/3} \left(O(n^{-1/3}) + O(n^{1/3}) \right) + O(n^{-1/3}) \\
&= O(n^{-1/3}).
\end{aligned}$$

In the first line we used the bound $0 < n^{-1} + \lambda n^{-4/3} \le 2n^{-1}$ for large enough n, and $\mu(n, z, c) \le n$, the second inequality used the bounds $(n^{-1} + \lambda n^{-4/3}) z \le 2rn^{-2/3}$ on Ω_n. The third inequality used the previously derived bound of $\sum_{j=2}^{z} \binom{z}{j} (n^{-1} + \lambda n^{-4/3}) \le 8e^2 r^2 n^{-4/3}$ which holds for n sufficiently large.

To show the bound for $\kappa(n,z,c)$, we expand it as follows

$$\begin{aligned}\kappa(n,z,c) &= \mathbb{E}\left[(\beta(n,z,c)-\mu(n,z,c))^4\right] + 4\mathbb{E}\left[(\beta(n,z,c)-\mu(n,z,c))^3\right](\mu(n,z,c)-z) \\ &\quad + 6\mathbb{E}\left[(\beta(n,z,c)-\mu(n,z,c))^2\right](\mu(n,z,c)-z)^2 \\ &\quad + 4\mathbb{E}\left[\beta(n,z,c)-\mu(n,z,c)\right](\mu(n,z,c)-z)^3 \\ &\quad + (\mu(n,z,c)-z)^4 \\ &=: \kappa_4(n,z,c) + 4\kappa_3(n,z,c) + 6\kappa_2(n,z,c) + 0 + \kappa_0(n,z,c).\end{aligned}$$

We now show that $\kappa_j(n,z,c)$ for $j = 0,2,3,4$ have the desired bound.

By the approximations for $\mu(n,z,c)$ it is easy to see that

$$\sup_{\Omega_n} |\kappa_0(n,z,c)| = O(1), \qquad \text{as } n \to \infty.$$

Similarly, we can use the approximations for both $\mu(n,z,c)$ and $\sigma^2(n,z,c)$ to arrive at

$$\begin{aligned}\sup_{\Omega_n} |\kappa_2(n,z,c)| &= \sup_{\Omega_n} \left|\sigma^2(n,z,c)(\mu(n,z,c)-z)^2\right| \\ &\le O(1)\cdot \sup_{\Omega_n}\left(\left|\sigma^2(n,z,c) - z - n^{-1/3}z(\lambda - n^{-2/3}c)\right| + |z + n^{-1/3}z(\lambda - n^{-2/3}c)|\right) \\ &= O(1)\cdot\left(O(n^{-1/3}) + O(n^{1/3})\right) = O(n^{1/3}).\end{aligned}$$

Using the third central moment of a binomial random variable gives

$$\kappa_3(n,z,c) = \sigma^2(n,z,c)(1-2q(n,z))(\mu(n,z,c)-z).$$

A similar expansion as that for $\kappa_2(n,z,c)$ shows that $\sup_{\Omega_n}|\sigma^2(n,z,c)| = O(n^{1/3})$, and the other two terms are $O(1)$ over Ω_n and hence

$$\sup_{\Omega_n} |\kappa_3(n,z,c)| = O(n^{1/3}).$$

By the fourth central moment for a binomial random variable, we have

$$\begin{aligned}|\kappa_4(n,z,c)| &= \sigma^2(n,z,c)\left|1 + 3(n-2-c)(q(n,z)-q(n,z)^2)\right| \\ &\le \sigma^2(n,z,c)\left(\left|1+3(n-c)(q(n,z)-q(n,z)^2)\right| + 2|(q(n,z)-q(n,z)^2)|\right) \\ &\le \sigma^2(n,z,c)\left(3 + 3\sigma^2(n,z,c)\right).\end{aligned}$$

Hence, by the bound for $\sigma^2(n,z,c)$

$$\sup_{\Omega_n} |\kappa_4(n,z,c)| = O(n^{2/3}).$$

This proves the desired claim.

□

4.2.2 Martingale estimates

In this section we verify the conditions of the martingale functional central limit theorem, as found in [86, Theorem 7.4.1]. Before moving onto the lemma, we establish some notation.

We let

$$\mathcal{F}_n^k(h) = \sigma\left(Z_n^k(j) : j \le h\right)$$

denote the filtration generated by Z_n^k. We let

$$Z_n^k(h) = k + M_n^k(h) + B_n^k(h),$$

be the Doob decomposition of Z_n^k into an $\{\mathcal{F}_n^k(h)\}_{h\ge 0}$-martingale M_n^k and a predictable process B_n^k. Similarly, we let Q_n^k be the unique increasing process which makes $(M_n^k(h))^2 - Q_n^k(h)$ an $\{\mathcal{F}_n^k(h)\}_{h\ge 0}$-martingale. That is

$$\begin{aligned} B_n^k(h) &= \sum_{\ell=0}^{h-1} \mathbb{E}\left[Z_n^k(\ell+1) - Z_n^k(\ell) | \mathcal{F}_n^k(\ell)\right] \\ Q_n^k(h) &= \sum_{\ell=0}^{h-1} \mathbb{E}\left[\left(Z_n^k(\ell+1) - Z_n^k(\ell)\right)^2 | \mathcal{F}_n^k(\ell)\right] - \mathbb{E}\left[Z_n^k(\ell+1) - Z_n^k(\ell) | \mathcal{F}_n^k(\ell)\right]^2. \end{aligned}$$

We define the following rescaled processes

$$\begin{aligned} &\tilde{Z}_n^k(t) = n^{-1/3} Z_n^k(\lfloor n^{1/3} t\rfloor) \qquad \tilde{C}_n^k(t) = n^{-2/3} C_n^k(\lfloor n^{1/3} t\rfloor) \\ &\tilde{M}_n^k(t) = n^{-1/3} M_n^k(\lfloor n^{1/3} t\rfloor) \qquad \tilde{B}_n^k(t) = n^{-1/3} B_n^k(\lfloor n^{1/3} t\rfloor) \qquad \tilde{Q}_n^k(t) = n^{-2/3} Q_n^k(\lfloor n^{1/3} t\rfloor) \end{aligned}. \tag{4.10}$$

Also define $\tau_n^k(r) = \inf\{t : \tilde{Z}_n^k(t) \vee \tilde{Z}_n^k(t-) \ge r\}$ and $\hat{\tau}_n^k(r) = n^{-1/3}\inf\{h : Z_n^k(h) \ge n^{1/3} r\}$.

Lemma 4.2.2. *Fix any $r > 0$, $T > 0$ and $x > 0$. Let $k = k(n) = \lfloor n^{1/3} x\rfloor$. The following limits hold*

1. $\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{Z}_n^k(t)-\tilde{Z}_n^k(t-)|^2\right]=0.$

2. $\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{B}_n^k(t)-\tilde{B}_n^k(t-)|^2\right]=0.$

3. $\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{Q}_n^k(t)-\tilde{Q}_n^k(t-)|\right]=0.$

4. $\sup_{t\le T\wedge\tau_n^k(r)}\left|\tilde{Q}_n^k(t)-\int_0^t \tilde{Z}_n^k(s)\,ds\right| \longrightarrow 0,$ *in probability as* $n\to\infty$.

5. $\sup_{t\le T\wedge\tau_n^k(r)}\left|\tilde{B}_n^k(t)-\int_0^t (\lambda-\tilde{C}_n^k(s))\tilde{Z}_n^k(s)\,ds\right| \longrightarrow 0,$ *in probability as* $n\to\infty$.

Proof. In order to show (1), we prove the stronger claim

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{Z}_n^k(t)-\tilde{Z}_n^k(t-)|^4\right]=0.$$

To show this, note the following string of inequalities

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{Z}_n^k(t)-\tilde{Z}_n^k(t-)|^4\right] &\le n^{-4/3}\mathbb{E}\left[\sup_{h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))} |Z_n^k(h+1)-Z_n^k(h)|^4\right] \\
&\le n^{-4/3}\sum_{h=0}^{\lfloor n^{1/3}(T\wedge\hat{\tau}_n^k(r))\rfloor} \mathbb{E}\left[|Z_n^k(h+1)-Z_n^k(h)|^4\right] \\
&\le n^{-4/3}\sum_{h=0}^{\lfloor n^{1/3}(T\wedge\hat{\tau}_n^k(r))\rfloor} \sup_{\Omega_n}\mathbb{E}\left[\mathbb{E}\left[(Z_n^k(h+1)-Z_n^k(h))^4\middle| Z_n^k(h)=z, C_n^k(h)=c\right]\right] \\
&\le Tn^{-1}\sup_{\Omega_n}\mathbb{E}\left[(\beta(n,z,c)-z)^4\right]=O(n^{-1/3}).
\end{aligned}$$

In the third inequality above, we used the tower property and on fact that for $h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))$ both $Z_n^k(h)\le n^{1/3}r$ and $C_n^k(h)\le n^{2/3}Tr$. The fourth inequality used the Markov property of (Z_n^k, C_n^k). The convergence then holds by the asymptotic result for $\kappa(n,z,c)$ shown in Lemma 4.2.1

To verify (2), we begin by noting that

$$B_n^k(h)=\sum_{j=0}^{h-1}\mathbb{E}\left[(Z_n^k(j+1)-Z_n^k(j))|\mathcal{F}_n^k(j)\right],$$

and hence

$$\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{B}_n^k(t)-\tilde{B}_n^k(t-)|^2 \le n^{-2/3} \sup_{h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))} \left|\mathbb{E}\left[Z_n^k(h+1)-Z_n^k(h)\middle|\mathcal{F}_n^k(h)\right]\right|^2.$$

We also note that almost surely on $h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))$

$$\mathbb{E}\left[Z_n^k(h+1)-Z_n^k(h)\middle|\mathcal{F}_n^k(h)\right] \le \sup_{\Omega_n} |\mathbb{E}[(\beta(n,z,c)-z)]| = O(1) \text{ as } n\to\infty,$$

by Lemma 4.2.1. Hence,

$$\begin{aligned}\mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{B}_n^k(t)-\tilde{B}_n^k(t-)|^2\right] &\le \mathbb{E}\left[n^{-2/3} \sup_{h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))} \left|\mathbb{E}\left[Z_n^k(h+1)-Z_n^k(h)\middle|\mathcal{F}_n^k(h)\right]\right|^2\right] \\ &\le n^{-2/3}\cdot O(1) = O(n^{-2/3}),\end{aligned}$$

which argues (2).

We next show (3). We begin by noting that

$$Q_n^k(h) = \sum_{j=0}^{h-1} \mathbb{E}\left[\left(Z_n^k(j+1)-Z_n^k(j)\right)^2\middle|\mathcal{F}_n^k(j)\right] - \mathbb{E}\left[Z_n^k(j+1)-Z_n^k(j)\middle|\mathcal{F}_n^k(j)\right]^2.$$

Hence

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\le T\wedge\tau_n^k(r)} |\tilde{Q}_n^k(t)-\tilde{Q}_n^k(t-)|\right] \\ &\quad\le n^{-2/3}\mathbb{E}\left[\sup_{h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))} \mathbb{E}[(Z_n^k(h+1)-Z_n^k(h))^2|\mathcal{F}_n^k(h)] + \mathbb{E}[Z_n^k(h+1)-Z_n^k(h)|\mathcal{F}_n^k(h)]^2\right] \\ &\quad\le n^{-2/3}\mathbb{E}\left[\sup_{h\le n^{1/3}(T\wedge\hat{\tau}_n^k(r))} \sup_{\Omega_n} \left|\mathbb{E}[(\beta(n,z,c)-z)^2] + (\mu(n,z,c)-z)^2\right|\right] \\ &\quad= n^{-2/3}\sup_{\Omega_n} \left|\sigma^2(n,z,c) + 2(\mu(n,z,c)-z)^2\right| = O(n^{-1/3}).\end{aligned}$$

In the last equalities, we Lemma 4.2.1 and the observation that $\sup_{\Omega_n}\sigma^2(n,z,c) = O(n^{1/3})$.

To argue claim (4), we observe

$$\begin{aligned}\left(Q_n^k(h+1)-Q_n^k(h)\right) &= \mathbb{E}\left[(\beta(n,Z_n^k(h),C_n^k(h)) - Z_n^k(h))^2|\mathcal{F}_n^k(h)\right] \\ &\qquad - (\mu(n,Z_n^k(h),C_n^k(h)) - Z_n^k(h))^2 \\ &= \sigma^2(n,Z_n^k(h),C_n^k(h)).\end{aligned}$$

Therefore,

$$\begin{aligned}
\sup_{h\le n^{1/3}(T\wedge\hat\tau_n^k(r))}\left|Q_n^k(h)-\sum_{j=0}^{h-1}Z_n^k(j)\right| &= \sup_{h\le n^{1/3}(T\wedge\hat\tau_n^k(r))}\left|\sum_{j=0}^{k-1}\sigma^2(n,Z_n^k(j),C_n^k(j))-Z_n^k(j)\right|\\
&\le \sup_{h\le n^{1/3}(T\wedge\hat\tau_n^k(r))}\sum_{j=0}^{h-1}|\sigma^2(n,Z_n^k(j),C_n^k(j))-Z_n^k(j)|\\
&\le \sum_{j=0}^{n^{1/3}T}\sup_{\Omega_n}|\sigma^2(n,z,c)-z|\\
&= O(n^{1/3}).
\end{aligned}$$

In the third inequality above, we used the previously discussed bounds on $Z_n^k(h)$ and $C_n^k(h)$ for all h such that $h\le n^{1/3}(T\wedge\hat\tau_n^k(r))$ and in the last term, we used the bound for $\sigma^2(n,z,c)-z$ on Ω_n given by Lemma 4.2.1. Hence

$$\begin{aligned}
\sup_{t\le T\wedge\tau_n^k(r)}\left|\tilde Q_n^k(t)-\int_0^t\tilde Z_n^k(s)\,ds\right| \le{}& \sup_{t\le T\wedge\tau_n^k(r)}\left|\tilde Q_n^k(t)-n^{-2/3}\sum_{j=0}^{\lfloor n^{1/3}t\rfloor}Z_n^k(j)\right|\\
&+\sup_{t\le T\wedge\tau_n^k(r)}\left|n^{-2/3}\sum_{j=0}^{\lfloor n^{1/3}t\rfloor}Z_n^k(j)-\int_0^t\tilde Z_n^k(s)\,ds\right|.
\end{aligned}\tag{4.11}$$

We can bound the first term, using the bound for $Q_n^k(h)-\sum_{j=0}^{h-1}Z_n^k(j)$ from above, to get

$$\begin{aligned}
\sup_{t\le T\wedge\tau_n^k(r)}\left|\tilde Q_n^k(t)-n^{-2/3}\sum_{j=0}^{\lfloor n^{1/3}t\rfloor}Z_n^k(j)\right| &\le n^{-2/3}\sup_{h\le n^{1/3}(T\wedge\hat\tau_n(r))}\left|Q_n^k(h)-\sum_{j=0}^{h-1}Z_n^k(j)\right|\\
&= O(n^{-1/3}).
\end{aligned}$$

We can bound the second term as follows

$$\begin{aligned}
&\sup_{t\le T\wedge\tau_n^k(r)}\left|n^{-2/3}\sum_{j=0}^{\lfloor n^{1/3}t\rfloor}Z_n^k(j)-\int_0^t\tilde Z_n^k(s)\,ds\right|\\
&\le \sup_{t\le T\wedge\tau_n^k(r)}\left|n^{-2/3}\int_0^{\lfloor n^{1/3}t\rfloor+1}Z_n^k(\lfloor u\rfloor)\,du-\int_0^t\tilde Z_n^k(s)\,ds\right|\\
&\le \sup_{t\le T\wedge\tau_n^k(r)}\left|\int_0^{n^{-1/3}(\lfloor n^{1/3}t\rfloor+1)}\tilde Z_n^k(s)\,du-\int_0^t\tilde Z_n^k(s)\,ds\right|\\
&\le r\sup_{t\le T}\left|t-n^{-1/3}(\lfloor n^{1/3}t\rfloor+1)\right|\longrightarrow 0.
\end{aligned}$$

The above bounds hold almost surely. This proves the convergence in (4).

We lastly establish (5). We begin by noting that

$$B_n^k(h+1) - B_n^k(h) = \mathbb{E}\left[Z_n^k(h+1) - Z_n^k(h)|\mathcal{F}_n^k(h)\right] = \mu(n, Z_n^k(h), C_n^k(h)) - Z_n^k(h).$$

Hence,

$$B_n^k(h) = \sum_{j=0}^{h-1} \mu(n, Z_n^k(j), C_n^k(j)) - Z_n^k(j).$$

Therefore, almost surely we have

$$\begin{aligned}
&\sup_{h \leq n^{1/3}(T \wedge \hat{\tau}_n^k(r))} \left| B_n^k(h) - \sum_{j=0}^{h-1} n^{-1/3} Z_n^k(j)(\lambda - n^{-2/3} C_n^k(j)) \right| \\
&\qquad = \sup_{h \leq n^{1/3}(T \wedge \hat{\tau}_n^k(r))} \left| \sum_{j=0}^{h-1} (\mu(n, Z_n^k(j), C_n^k(j)) - Z_n^k(j)) - n^{-1/3} Z_n^k(j)(\lambda - n^{-1/3} C_n^k(j)) \right| \\
&\qquad \leq \sup_{h \leq n^{1/3}(T \wedge \hat{\tau}_n^k(r))} \sum_{j=0}^{h-1} \left| \mu(n, Z_n^k(j), C_n^k(j)) - Z_n^k(j) - n^{-1/3} Z_n^k(j)(\lambda - n^{-1/3} C_n^k(j)) \right| \\
&\qquad \leq T n^{1/3} \sup_{\Omega_n} \left| \mu(n, z, c) - z - n^{-1/3} z(\lambda - n^{-2/3} c) \right| = O(1).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sup_{t \leq T \wedge \tau_n^k(r)} \left| \tilde{B}_n^k(t) - \int_0^t (\lambda - \tilde{C}_n^k(s)) \tilde{Z}_n^k(s)\, ds \right| \\
&\quad \leq \sup_{h \leq n^{1/3}(T \wedge \hat{\tau}_n^k(r))} \left| n^{-1/3} B_n^k(h) - n^{-1/3} \sum_{j=0}^{h-1} n^{-1/3} Z_n^k(j)(\lambda - n^{-2/3} C_n^k(j)) \right| \\
&\qquad + \sup_{t \leq T \wedge \tau_n^k(r)} \left| n^{-1/3} \sum_{j=0}^{\lfloor n^{1/3} t \rfloor} n^{-1/3} Z_n^k(j)(\lambda - n^{-2/3} C_n^k(j)) - \int_0^t (\lambda - \tilde{C}_n^k(s)) \tilde{Z}_n^k(s)\, ds \right|.
\end{aligned}$$

By factoring out an $n^{-1/3}$ from the first term on the right-hand side, it is easy to see that term is $O(n^{-1/3})$ almost surely. Examining the second term on the right-hand side, we get almost surely

$$\begin{aligned}
&\sup_{t\le T\wedge\tau_n^k(r)}\left|n^{-1/3}\sum_{j=0}^{\lfloor n^{1/3}t\rfloor} n^{-1/3}Z_n^k(j)(\lambda-n^{-2/3}C_n^k(j))-\int_0^t(\lambda-\tilde{C}_n^k(s))\tilde{Z}_n^k(s)\,ds\right|\\
&=\sup_{t\le T\wedge\tau_n^k(r)}\left|n^{-1/3}\int_0^{\lfloor n^{1/3}t\rfloor+1} n^{-1/3}Z_n^k(\lfloor u\rfloor)(\lambda-n^{-2/3}C_n^k(\lfloor u\rfloor))\,du-\int_0^t(\lambda-\tilde{C}_n^k(s))\tilde{Z}_n^k(s)\,ds\right|\\
&=\sup_{t\le T\wedge\tau_n^k(r)}\left|\int_0^{n^{-1/3}(\lfloor n^{1/3}t\rfloor+1)}\tilde{Z}_n^k(s)(\lambda-\tilde{C}_n^k(s))\,ds-\int_0^t\tilde{Z}_n^k(s)(\lambda-\tilde{C}_n^k(s))\,ds\right|\\
&=(|\lambda|+Tr)r\sup_{t\le T}\left|t-n^{-1/3}(\lfloor n^{1/3}t\rfloor+1)\right|\to 0.
\end{aligned}$$

This proves the lemma.

□

4.2.3 Existence and uniqueness lemma

Using the functional central limit machinery found in [86, Chapter 7], it is not difficult to argue Theorem 4.1.1 from Lemma 4.2.2 along with the following existence and uniqueness lemma:

Lemma 4.2.3. *Fix a $\lambda\in\mathbb{R}$ and an $x\ge 0$.*

1. *There exists a unique strong solution to the following stochastic differential equation*

$$\begin{aligned}
d\mathbf{Z}(t)&=\sqrt{\mathbf{Z}(t)}dW(t)+(\lambda-\mathbf{C}(t))\,\mathbf{Z}(t)\,dt, \qquad \mathbf{Z}(0)=x\\
d\mathbf{C}(t)&=\mathbf{Z}(t)\,dt, \qquad \mathbf{C}(0)=0,
\end{aligned} \tag{4.12}$$

   which is absorbed upon $\mathbf{Z}$ hitting zero.

2. *Given a weak solution $(\mathbf{Z},\mathbf{C})$ to the equation (4.12), then on an enlarged probability space there exists a Brownian motion B such that $(\mathbf{Z},\mathbf{C})$ solves*

$$\mathbf{Z}(t)=x+\mathbf{X}^\lambda\left(\mathbf{C}(t)\wedge T_{-x}\right), \tag{4.13}$$

   where $\mathbf{X}^\lambda(t)=B(t)+\lambda t-\frac{1}{2}t^2$ and $T_{-x}=\inf\{t:\mathbf{X}^\lambda=-x\}$.

3. *Given* $\mathbf{X}^\lambda(t) = B(t) + \lambda t - \frac{1}{2}t^2$ *for a Brownian motion* B *there exists a path-wise unique solution* $(\mathbf{Z}, \mathbf{C})$ *where* $\mathbf{C}(t) = \int_0^t \mathbf{Z}(s)\, ds$ *and such a solution is a weak solution to* (4.12).

Remark 4.2.1. We observe that the SDE in equation (4.12) does not have a $\frac{1}{2}$, while in the integrated form found in equation (4.2) there is such a term. This is because

$$\int_0^t \mathbf{C}(s)\, \mathbf{Z}(s)\, ds = \int_0^t \mathbf{C}(s)\, d\mathbf{C}(s) = \frac{1}{2}\mathbf{C}(t)^2.$$

Proof. The strong existence and uniqueness in first item follows from the Yamada-Watanabe theorem [184, Theorem 1]. The absorption upon $\mathbf{Z}$ is obvious by stopping $(\mathbf{Z}, \mathbf{C})$ upon $\mathbf{Z}$ hitting zero and observing that this stopped process still solves (4.12).

The path-wise existence found in the third item follows from known theorems on random time-changes. See, for example [86, Chapter VI, Section 1], [48] or [49, Section 2].

Now suppose that $(\mathbf{Z}, \mathbf{C})$ solves (4.12) for a Brownian motion W. We observe that the quadratic variation of $\mathbf{Z}$ is given by

$$\langle \mathbf{Z} \rangle(t) = \int_0^t \mathbf{Z}(s)\, ds.$$

Define the process $M(t) = \int_0^t \sqrt{\mathbf{Z}(s)}\, dW(s)$. Define $V(t) = \inf\{s : \mathbf{C}(s) > t\}$ with the convention that $\inf \emptyset = \infty$.

Hence, by the Dambis, Dubins-Schwarz theorem [156, Chapter V, Theorem 1.7], on an enlarged probability space, there exists a Brownian motion $\tilde{B}$ such that process

$$B(t) = \begin{cases} M(V(t)) & : t < \int_0^\infty \mathbf{Z}(s)\, ds \\ M(\infty) + \tilde{B}_{t - \langle M \rangle(\infty)} & : t \geq \int_0^\infty \mathbf{Z}(s)\, ds. \end{cases}$$

is a Brownian motion.

We then have for $t < \int_0^\infty \mathbf{Z}(s)\, ds$:

$$\begin{aligned} \mathbf{Z}(V(t)) &= x + M(V(t)) + \int_0^{V(t)} (\lambda - \mathbf{C}(s))\mathbf{Z}(s)\, ds \\ &= x + B(t) + \int_0^t (\lambda - s)\ ds \\ &= x + \mathbf{X}^\lambda(t). \end{aligned}$$

Observe $t < \int_0^\infty \mathbf{Z}(s)\,ds$ occurs if and only if $\mathbf{Z}(V(t)) > 0$. Indeed, since $\mathbf{Z}$ is continuous, non-negative and absorbed upon reaching zero then by [156, Lemma 0.4.8] $t \le \mathbf{C}(V(t)) = \int_0^{V(t)} \mathbf{Z}(s)\,ds < \int_0^\infty \mathbf{Z}(s)\,ds$ if and only if $\int_{V(t)}^\infty \mathbf{Z}(s)\,ds > 0$ which occurs if and only if $\mathbf{Z}(V(t)) > 0$.

Hence, we can rewrite the above string of equalities as

$$\mathbf{Z}(V(t)) = x + \mathbf{X}^\lambda(t \wedge T_{-x}),$$

which now holds for all t. Indeed, if $t \ge \int_0^\infty \mathbf{Z}(s)\,ds = \int_0^\zeta \mathbf{Z}(s)\,ds$ then $\mathscr{V}(t) = \mathbf{C}(\zeta) = T_{-x}$. Since V and $\mathbf{C}$ are two-sided inverses of each other prior to $V(t) = \infty$, the above equation implies equation (4.13) holds.

Reversing the above steps, gives the implication in the third item. $\square$

Lemma 4.2.4. *Let* $(\mathbf{Z}, \mathbf{C})$ *be a solution of* (4.12). *Then almost surely*

$$\zeta := \inf\{t : \mathbf{Z}(t) = 0\} < \infty.$$

Proof. We let $\mathbf{X}^\lambda(t)$ be the process define in Lemma 4.2.3(2) and such that $(\mathbf{Z}, \mathbf{C})$ solves (4.13). We can write $\mathbf{C}$ as a function of just the process $\mathbf{X}^\lambda$. Indeed let T_{-x} be the first hitting time of $-x$ of $\mathbf{X}^\lambda(t)$, then

$$\mathbf{C}(t) = \inf\left\{s : \int_0^s \frac{1}{x + \mathbf{X}^\lambda(u \wedge T_{-x})}\,du = t\right\}.$$

This is a simple calculus exercise and the proof can be found in [49, Section 2] and in [86, Chapter VI, Section 1].

We note that almost surely

$$I := \int_0^{T_{-x}} \frac{1}{x + \mathbf{X}^\lambda(u \wedge T_{-x})}\,du < \infty.$$

Indeed, this is true if we replace $\mathbf{X}^\lambda$ with a Brownian motion B and Girsanov's theorem [156, Chapter VIII] implies that it is true almost surely for the $\mathbf{X}^\lambda$ as well. Hence

$$\zeta = \inf\{t : x + \mathbf{X}^\lambda(\mathbf{C}(t)) = 0\} = \inf\{s : \mathbf{C}(s) = T_{-x}\} = I < \infty.$$

$\square$

4.3 A Self-Similarity Result

We first observe the following relationship in λ for the process $\mathbf{X}^\lambda$. Namely,

$$\left(\left(\mathbf{X}^\lambda(t_0+t)-\mathbf{X}^\lambda(t_0); t\geq 0\right)\Big|\mathbf{X}^\lambda(t_0)=\inf_{s\leq t_0}\mathbf{X}^\lambda(s)\right)\stackrel{d}{=}\left(\mathbf{X}^{\lambda-t_0}(t); t\geq 0\right).$$

This observation was used by Aldous [12] to simplify the description of the (time-inhomogeneous) excursion measure of $\mathbf{X}^\lambda$ at time t to the excursion measure of $\mathbf{X}^{\lambda-t}$ at time 0. See also [5].

A similar result will hold in our situation as well. We state it in the following theorem

Theorem 4.3.1. *Let $\mathbf{Z}_x^\lambda(t), \mathbf{C}_x^\lambda(t)$ denote the solution to*

$$\begin{aligned} d\mathbf{Z}_x^\lambda(t) &= \sqrt{\mathbf{Z}_x^\lambda(t)}\, dW(t) + \left(\lambda - \mathbf{C}_x^\lambda(t)\right)\mathbf{Z}_x^\lambda(t)\, dt, \qquad \mathbf{Z}_x^\lambda(0)=x \\ d\mathbf{C}_x^\lambda(t) &= \mathbf{Z}_x^\lambda(t)\, dt \qquad \mathbf{C}_x^\lambda(0)=0 \end{aligned}.$$

Then the following self-similarity result holds for any $t_0>0$, $z>0$ and $\mu>0$:

$$\left(\left(\left(\mathbf{Z}_x^\lambda(t_0+t), \mathbf{C}_x^\lambda(t_0+t)\right); t\geq 0\right)\Big|\mathbf{Z}_x^\lambda(t_0)=z, \mathbf{C}_x^\lambda(t_0)=\mu\right)\stackrel{d}{=}\left(\left(\mathbf{Z}_z^{\lambda-\mu}(t), \mathbf{C}_z^{\lambda-\mu}(t)\right); t\geq 0\right) \tag{4.14}$$

Proof. The proof follows from the decomposition in Lemma 4.2.3, particularly in equation (4.13). Namely, there exists a Brownian motion with parabolic drift $\mathbf{X}^\lambda(t)$ such that

$$\mathbf{Z}_x^\lambda(t) = x + \mathbf{X}^\lambda(\mathbf{C}_x^\lambda(t)\wedge T_{-x}).$$

We also observe that

$$\begin{aligned} \mathbf{X}^\lambda(s_0+s) &= B(s_0+s)+\lambda(s_0+s)-\frac{1}{2}(s_0+s)^2 \\ &= B(s_0)+B(s_0+s)-B(s_0)+\lambda s_0+\lambda s-\frac{1}{2}s_0^2-s_0s-\frac{1}{2}s^2 \\ &= \mathbf{X}^\lambda(s_0)+\left(B(s_0+s)-B(s_0)+(\lambda-s_0)s-\frac{1}{2}s^2\right) \\ &= \mathbf{X}^\lambda(s_0)+\tilde{\mathbf{X}}^{\lambda-s_0}(s), \end{aligned}$$

for a process $\tilde{\mathbf{X}}^{\lambda-s_0} \stackrel{d}{=} \mathbf{X}^{\lambda-s_0}$ which is independent of $\sigma\left\{\mathbf{X}^\lambda(u); u \le s_0\right\}$. Hence, we have

$$\begin{aligned}
\mathbf{Z}_x^\lambda(t_0+t) &= x + \mathbf{X}^\lambda\left(\mathbf{C}_x^\lambda(t_0+t)\right)\\
&= x + \mathbf{X}^\lambda\left(\mathbf{C}_x^\lambda(t_0) + \int_0^t \mathbf{Z}_x^\lambda(t_0+s)\,ds\right)\\
&= x + \mathbf{X}^\lambda(\mathbf{C}_x^\lambda(t_0)) + \tilde{B}\left(\int_0^t \mathbf{Z}_x^\lambda(t_0+s)\,ds\right)\\
&\qquad + \left(\lambda - \mathbf{C}_x^\lambda(t_0)\right)\int_0^t \mathbf{Z}_x^\lambda(t_0+s)\,ds - \frac{1}{2}\left(\int_0^t \mathbf{Z}_x^\lambda(t_0+s)\,ds\right)^2,
\end{aligned}$$

where $\tilde{B}$ is a Brownian motion independent of $\sigma\{\mathbf{X}^\lambda(u) : u \le \mathbf{C}(t_0)\}$. Hence, conditionally on $\mathbf{Z}_x^\lambda(t_0) = z$ and $\mathbf{C}_x^\lambda(t_0) = \mu$ gives

$$\mathbf{Z}_x^\lambda(t_0+t) = z + \tilde{\mathbf{X}}^{\lambda-\mu}\left(\int_0^t \mathbf{Z}_x^\lambda(t_0+s)\,ds\right)$$

By Lemma 4.2.3, this is equivalent to the statement in (4.14) . □

4.4 A More General Asymptotic Regime

As observed by Bollobás in [40], the asymptotic order of largest component of the Erdős-Rényi random graph $G(n, n^{-1} + \lambda\log(n)^{1/2}n^{-4/3})$ is $n^{2/3}(\log n)^{1/2}$ as $n \to \infty$. Actually, he proves a much more general result, but we will not state that fully here. We instead examine a more general asymptotic regime.

We consider any sequence of real numbers θ_n such that

$$\theta_n = o(n^{1/3}), \qquad \text{and} \qquad \lim_{n\to\infty}\theta_n = \infty. \tag{4.15}$$

We also fix a $\lambda \in \mathbb{R}$ and let

$$\mathcal{G}_n^\theta = G(n, n^{-1} + \lambda\theta_n n^{-4/3}).$$

To distinguish the notation, we let $Z_n^{\theta,k}(h)$ denote the height profile of $\mathcal{G}_n^\theta$ starting from k uniformly chosen vertices. With this notation, we can state the following theorem:

Theorem 4.4.1. *Fix $x > 0$. Suppose that θ_n satisfies (4.15) and $k = k(n) = \lfloor\theta_n^2 n^{1/3}x\rfloor$. Then the following convergence holds in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$*

$$\left(\frac{1}{\theta_n^2 n^{1/3}} Z_n^{\theta,k}\left(\lfloor\theta_n^{-1}n^{1/3}t\rfloor\right); t \ge 0\right) \Longrightarrow (z(t); t \ge 0)\,, \tag{4.16}$$

where z solves the deterministic equation

$$z(t) = f\left(\int_0^t z(s)\,ds\right), \qquad f(t) = x + \lambda t - \frac{1}{2}t^2. \tag{4.17}$$

The proof follows from lemmas similar to the lemmas found in Section 4.2.3. Before stating those lemmas, we make some comments on the solution $z(t)$ found in (4.17). We have already mentioned that

$$c(t) = \int_0^t z(s)\,ds = \inf\{s : \int_0^s \frac{1}{f(u)}\,du = t\}.$$

See also, [49, Section 2] and [86, Section 6.1] for more details on time changes. We have "$= t$" instead of "$> t$" because the inverse is actually a two-sided inverse. Indeed, since $\int_0^{t_0} \frac{1}{f(u)}\,du = \infty$ where $t_0 = \lambda + \sqrt{2x+\lambda^2}$ is the largest root of $f(t)$, the function c is strictly increasing continuous function $c : [0,\infty) \to [0, \lambda + \sqrt{2x+\lambda^2})$. The function c can actually be explicitly computed:

$$c(t) = \lambda + \sqrt{2x+\lambda^2}\tanh\left(\frac{\sqrt{2x+\lambda^2}}{2}t + \operatorname{arctanh}\left(\frac{-\lambda}{\sqrt{2x+\lambda^2}}\right)\right).$$

We also make comments on the scaling found in Theorem 4.4.1. In order to describe this scaling, we introduce the diameter of the graph $\mathcal{G}_n^\theta$, as

$$\mathscr{D}_n^\theta = \mathscr{D}_n^{\theta_n} = \max_{u,v\in\mathcal{G}_n^\theta}\left\{d(u,v) : d(u,v) < \infty\right\}.$$

The trivial observation is that $Z_n^{\theta,k}(h) > 0$ implies that $\mathscr{D}_n^\theta \ge h$. A result of Łuczak [139, Theorem 11(iii)] implies when $\lambda < 0$ that

$$\mathscr{D}_n^\theta = \frac{\log(2\theta_n^3) + O(1)}{-\log(1-\theta_n n^{-1/3})}$$

with high probability as $n \to \infty$. There is a typo in the statement of Theorem [139, Theorem 11(iii)], he writes an $\log(2\varepsilon^2 n)$ term when there should be an $\log(2\varepsilon^3 n)$ term. In the supercritical ($\lambda > 0$) regime, it appears that the work of Ding, Kim, Lubetzky and Peres [64, 65] provide more precise results. Namely, they show [65, Theorem 1.1] that if $\mathscr{C}_n^\theta$ is the largest component of $\mathcal{G}_n^\theta$, for $\lambda > 0$, then with high probability

$$\operatorname{diam}(\mathscr{C}_n^\theta) = (3+o(1))n^{1/3}\theta_n^{-1}\log(\theta_n^3) \qquad \text{as } n \to \infty.$$

Even more precise asymptotic result in this regime can be found in [158], again in the supercritical regime when $\lambda > 0$.

Both of these results on the asymptotic diameter $\mathscr{D}_n^\theta$ suggest the proper "time" scaling in Theorem 4.4.1 should be $\theta_n^{-1} n^{1/3} \log(\theta_n) t$ as compared with $\theta_n^{-1} n^{1/3} t$; however, this is not the correct scaling to obtain a non-trivial limit.

4.4.1 Lemmas

In the connection to the Reed-Frost model of epidemics, it is easy to see that the analog of (4.5) becomes the following

$$\left(Z_n^{\theta,k}(h+1) \middle| Z_n^{\theta,k}(h) = z, C_n^{\theta,k}(h) = c\right) \stackrel{d}{=} \begin{cases} \text{Bin}\left(n-c, q_\theta(n,z)\right) & : z > 0, c < n \\ 0 & : \text{else} \end{cases},$$

where $q_\theta(n,z)$ is defined as

$$q_\theta(n,z) = 1 - \left(1 - n^{-1} - \lambda\theta_n n^{-4/3}\right)^z$$

The analog of Lemma 4.2.1 becomes the following

Lemma 4.4.2. *Let $\beta_\theta(n,z,c)$ denote a Bin$(n-c, q_\theta(n,z))$ random variable. Let $\mu_\theta, \sigma_\theta^2, \kappa_\theta$ denote the statistics in (4.7) with β_θ replacing β. Fix $r > 0$ and $T > 0$ and define*

$$\Omega_n^\theta = \Omega_n^\theta(n,r,T) := \left\{(z,c) \in \mathbb{Z}^2 : 0 \le z \le n^{1/3}\theta_n^2 r, 0 \le c \le n^{2/3}\theta_n r T\right\}$$

then the following bounds hold

$$\sup_{\Omega_n^\theta} \left|\mu_\theta(n,z,c) - z - n^{-1/3} z(\lambda\theta_n - n^{-2/3}c)\right| = O\left(\theta_n^4 n^{-1/3} + 1\right)$$
$$\sup_{\Omega_n^\theta} \left|\sigma_\theta^2(n,z,c) - z - n^{-1/3} z(\lambda\theta_n - n^{-2/3}c)\right| = O\left(\theta_n^4 n^{-1/3} + 1\right)$$
$$\sup_{\Omega_n^\theta} \left|\kappa_\theta(n,z,c)\right| = O(\theta_n^{12} + \theta_n^8 n^{1/3} + \theta_n^4 n^{2/3})$$

Proof. The proofs of the convergence of μ_θ and σ_θ^2 follow from the same argument as in the proof of Lemma 4.2.1, and we omit it here.

We do argue the result for κ_θ since it is much more involved computationally. We again use the expansion:

$$\begin{aligned}
\kappa_\theta(n,z,c) &= \mathbb{E}\left[(\beta_\theta(n,z,c)-\mu_\theta(n,z,c))^4\right] + 4\mathbb{E}\left[(\beta_\theta(n,z,c)-\mu_\theta(n,z,c))^3\right](\mu_\theta(n,z,c)-z) \\
&\quad + 6\mathbb{E}\left[(\beta_\theta(n,z,c)-\mu_\theta(n,z,c))^2\right](\mu_\theta(n,z,c)-z)^2 \\
&\quad + 4\mathbb{E}\left[\beta_\theta(n,z,c)-\mu_\theta(n,z,c)\right](\mu_\theta(n,z,c)-z)^3 \\
&\quad + (\mu_\theta(n,z,c)-z)^4 \\
&=: \kappa_{4,\theta}(n,z,c) + 4\kappa_{3,\theta}(n,z,c) + 6\kappa_{2,\theta}(n,z,c) + 0 + \kappa_{0,\theta}(n,z,c).
\end{aligned}$$

We can use the bound for μ_θ and Minkowski's inequality to get

$$\begin{aligned}
&\sup_{\Omega_n^\theta} |\kappa_{0,\theta}(n,z,c)|\,(\mu_\theta(n,z,c)-z)^4 \\
&\qquad \le \left[\sup_{\Omega_n^\theta}\left|n^{-1/3}z(\lambda\theta_n - n^{-2/3}c)\right| + O(\theta_n^4 n^{-1/3} + \theta_n^2 n^{-2/3})\right]^4 \\
&\qquad \le C\left(\sup_{\Omega_n^\theta}|n^{-1/3}z(\lambda\theta_n - n^{-2/3}c)|^4 + O(\theta_n^{16}n^{-4/3} + \theta_n^8 n^{-8/3})\right) \\
&\qquad = O\left(\theta_n^{12} + \theta_n^{16}n^{-4/3} + \theta_n^8 n^{-8/3}\right) \le O(\theta_n^{12})
\end{aligned}$$

where in the last inequality we used the bounds in (4.15).

The next three follow from the bounds below. They are easy to verify using the original bounds of σ_θ^2 and μ_θ, and computations similar to the one above:

$$\begin{aligned}
\sup_{\Omega_n^\theta}|\mu_\theta(n,z,c)-z| &= O\left(\theta_n^3\right) \\
\sup_{\Omega_n^\theta}\left|\sigma_\theta^2(n,z,c)\right| &= O\left(\theta_n^2 n^{1/3} + \theta_n^3 + \theta_n^4 n^{-1/3}\right) \\
&= O\left(\theta_n^2 n^{1/3}\right)
\end{aligned}$$

Using the same expansions as in Lemma 4.2.1, we have

$$\begin{aligned}
\sup_{\Omega_n^\theta}|\kappa_{2,\theta}(n,z,c)| &= O(\theta_n^6)\times O(\theta_n^2 n^{1/3}) = O(\theta_n^8 n^{1/3}) \\
\sup_{\Omega_n^\theta}|\kappa_{3,\theta}(n,z,c)| &= O(\theta_n^2 n^{1/3})\times O(\theta_n^3) = O(\theta_n^8 n^{1/3}) \\
\sup_{\Omega_n^\theta}|\kappa_{4,\theta}(n,z,c)| &= O(\theta_n^2 n^{1/3})^2 = O(\theta_n^4 n^{2/3})
\end{aligned}$$

This proves the desired bounds. □

One can use the bounds in the lemma above to prove an analog of Lemma 4.2.2. We first establish some notation. We now let $\mathcal{F}_n^{\theta,k}(h) = \sigma(Z_n^{\theta,k}(j), j \leq h)$ be the filtration generated by $Z_n^{\theta,k}$ and let $Z_n^{\theta,k}(h) = M_n^{\theta,k}(h) + B_n^{\theta,k}(h)$ be the decomposition of $Z_n^{\theta,k}$ into an $\mathcal{F}_n^{\theta,k}(h)$-martingale $M_n^{\theta,k}$ and a process $B_n^{\theta,k}$. We also let $Q_n^{\theta,k}$ be the process which makes $(M_n^{\theta,k}(h))^2 - Q_n^{\theta,k}(h)$ an $\mathcal{F}_n^{\theta,k}(h)$-martingale. Define the rescaled processes, in comparison to (4.10),

$$\begin{aligned}
&\tilde{Z}_n^{\theta,k}(t) = \theta_n^{-2} n^{-1/3} Z_n^{\theta,k}(\lfloor \theta_n^{-1} n^{1/3} t \rfloor) \quad \tilde{C}_n^{\theta,k}(t) = \theta_n^{-1} n^{-2/3} C_n^{\theta,k}(\lfloor \theta_n^{-1} n^{1/3} t \rfloor) \\
&\tilde{M}_n^{\theta,k}(t) = \theta_n^{-2} n^{-1/3} M_n^{\theta,k}(\lfloor \theta_n^{-1} n^{1/3} t \rfloor) \quad \tilde{B}_n^{\theta,k}(t) = \theta_n^{-2} n^{-1/3} B_n^{\theta,k}(\lfloor \theta_n^{-1} n^{1/3} t \rfloor) \quad . \\
&\tilde{Q}_n^{\theta,k}(t) = \theta_n^{-4} n^{-2/3} Q_n^{\theta,k}(\lfloor \theta_n^{-1} n^{1/3} t \rfloor).
\end{aligned} \tag{4.18}$$

Also define $\tau_n^{\theta,k}(r) = \inf\{t : \tilde{Z}_n^{\theta,k}(t) \vee \tilde{Z}_n^{\theta,k}(t-) > r\}$ and $\hat{\tau}_n^{\theta,k}(r) = \theta_n n^{-1/3} \inf\{k : Z_n^{\theta,k}(h) > \theta_n^2 n^{1/3} r\}$.

The analog of Lemma 4.2.2 is the following lemma. The proof is omitted since it is similar to the proof of Lemma 4.2.2.

Lemma 4.4.3. *Fix any $r > 0$, $T > 0$ and $x > 0$. Let $k = k(n) = \lfloor \theta_n^2 n^{1/3} x \rfloor$. The following limits hold*

1. $\lim_{n\to\infty} \mathbb{E}\left[\sup_{t \leq T \wedge \tau_n^{\theta,k}(r)} |\tilde{Z}_n^{\theta,k}(t) - \tilde{Z}_n^{\theta,k}(t-)|^2\right] = 0.$

2. $\lim_{n\to\infty} \mathbb{E}\left[\sup_{t \leq T \wedge \tau_n^{\theta,k}(r)} |\tilde{B}_n^{\theta,k}(t) - \tilde{B}_n^{\theta,k}(t-)|^2\right] = 0.$

3. $\lim_{n\to\infty} \mathbb{E}\left[\sup_{t \leq T \wedge \tau_n^{\theta,k}(r)} |\tilde{Q}_n^{\theta,k}(t) - \tilde{Q}_n^{\theta,k}(t-)|\right] = 0.$

4. $\sup_{t \leq T \wedge \tau_n^{\theta,k}(r)} \left|\tilde{Q}_n^{\theta,k}(t)\right| \longrightarrow 0,$ *as $n \to \infty$ almost surely*

5. $\sup_{t \leq T \wedge \tau_n^{\theta,k}(r)} \left|\tilde{B}_n^{\theta,k}(t) - \int_0^t (\lambda - \tilde{C}_n^{\theta,k}(s)) \tilde{Z}_n^{\theta,k}(s)\, ds\right| \xrightarrow{P} 0,$ *as $n \to \infty$.*

Finally, using the machinery of [86, Chapter 7], in particular Theorem 7.4.1, Lemma 4.4.1 follows from Lemma 4.4.3.

4.5 Breakdown with single source

Let us describe briefly the breakdown when $k = 1$ many individuals are infected on day $t = 0$ and informally describe why we should be able to overcome this issue. Theorem 4.1.1 still holds, however, the limiting stochastic process $\mathbf{Z} = (\mathbf{Z}(t); t \geq 0)$ is the unique solution to

$$\mathbf{Z}(t) = \mathbf{X}^\lambda \left(\int_0^t \mathbf{Z}(s)\, ds \right).$$

It is easy to see that since $\mathbf{Z}(t) = 0$ for all t is a solution it must be the only solution. Informally, this means that the first infected vertex is not in a macroscopic connected component which is of order $n^{2/3}$ [12].

However, as is shown in [5] and extended in [6], if we let $\mathscr{C}_n(\lambda)$ denote the large connected component of $G(n, n^{-1} + \lambda n^{-4/3})$ rooted at some vertex $\rho \in \mathscr{C}_n(\lambda)$ then we can view $\mathscr{C}_n(\lambda)$ as a PMM space with root ρ, with the rescaled graph metric $d_n = n^{-1/3} d$ and with the rescaled counting measure $\mu_n(dx) = n^{-2/3} \sum_{v \in \mathscr{C}_n(\lambda)} \delta_v(dx)$ and then

$$(\mathscr{C}_n(\lambda), \rho, d_n, \mu_n) \overset{d}{\Longrightarrow} \mathscr{M}_1(\lambda) = (\mathscr{M}, \mathsf{o}, d, \mu)$$

weakly in the compact GHP topology where $\mathscr{M}_1(\lambda)$ is some fractal metric spaces. Now informally, if the first infected vertex is the root $\rho \in \mathscr{C}_n(\lambda)$ then the total number of individuals infected by day $n^{1/3} r$ is

$$n^{-2/3} \sum_{\ell=0}^{n^{1/3} r} \#\{v \in \mathscr{C}_n(\lambda) : d(v, \rho) = \ell\} = \mu_n \left(B_r(\rho) \right)$$

where B_r denote the ball in $\mathscr{C}_n(\lambda)$ equipped with the rescaled graph distance d_n. By the results in [121], for Lebesgue almost every $r \geq 0$ we have

$$\mu_n(B_r(\rho)) \overset{d}{\Longrightarrow} \mu(B_r(\mathsf{o})).$$

Hence we should should expect that if $\hat{Z}_n(\ell) = \#\{v \in \mathscr{C}_n(\lambda) : d(v, \rho) = \ell\}$ then

$$\left(\int_0^r n^{-1/3} \hat{Z}_n(\lfloor n^{1/3} t \rfloor)\, dt \right) \approx \left(n^{-2/3} \sum_{\ell=0}^{\lfloor n^{1/3} r \rfloor} \hat{Z}_n(\ell) \right) \overset{d}{\Longrightarrow} (\mu(B_r(\mathsf{o})); r \geq 0) \qquad \text{in } \mathbb{D}(\mathbb{R}_+)$$

where $\approx$ in this case means that the distance between the two processes is asymptotically negligible.

This suggests that if $n^{-1/3}\hat{Z}_n(\lfloor n^{1/3}\cdot\rfloor)$ converges to a limit $\hat{Z}$ then $\frac{d}{dr}\mu(B_r(\mathsf{o}))$ exists and is equal to $\hat{Z}(r)$. While the above reasoning does not imply that this is true, we will see in the next chapter that the informal arguments in this subsection can be formalized and turned into a proof.

Chapter 5

EPIDEMICS ON CONNECTED COMPONENTS OF CRITICAL RANDOM GRAPHS

5.1 Introduction

Consider the following simple susceptible-infected-recovered (SIR) model of disease spread in discrete time. On day 0, a single individual becomes infected with a disease. On day 1, that single infected individual comes into contact with some random number (possibly zero) of non-infected individuals and transmits the disease. After transmitting the disease to others, this initial infected individual is cured and can never catch the disease again. On subsequent days each infected individual does the same thing: they come into contact with some non-infected individuals, transmit the disease but then are cured. The study of how the disease spreads over time naturally gives rise to a graph [24] constructed in a breadth-first order, see Figure 5.1 for an example of a small outbreak and Figure 5.2 for an example of a larger outbreak. The individuals are represented by vertices, and an edge between two vertices represents that a vertex closer to the source transmitted the disease to the other. Knowing the graph and the source tells us more information than the number of individuals infected on a particular day, it tells us the history of how the disease spread from individual from individual.

The size of the outbreak then corresponds to the size of a connected component in the graph and, more importantly for our work, the number of people infected on day $h = 0, 1, \cdots$ is just the number of vertices at distance h from a root vertex corresponding to the initially infected individual. Let $Z_n(h)$ represent the number of people infected on day $h \geq 0$ when the total population is of size n. The process $Z_n(h)$ is just the *height profile* of the component containing the initially infected individual. We are interested in the describing $n \to \infty$ scaling limits of $Z_n(h)$ for the macroscopic outbreaks for certain critical random graphs which exhibit a "super-spreader" phenomena - that is they possess vertices with

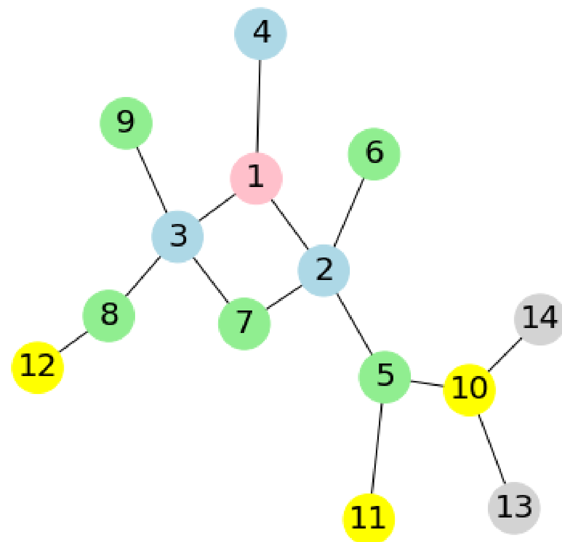


Figure 5.1: A small outbreak. Here, on day 0 the vertex labeled 1 is infected. The vertex 1 transmits the disease to vertices 2, 3 and 4 (in blue) who become the infected population on day 1. The vertices infected on day 1 will infect the green vertices (5 through 9) who are infected on day 2. This continues with the yellow vertices becoming infected on day 3, and the grey vertices on day 4.

large degree.

A classical probabilistic model in this area is the so-called Reed-Frost model, where each individual comes into contact with every non-infected individual independently with probability p. It is not hard to see that the corresponding graph is the Erdős-Rényi random graph $G(n,p)$ where each edge is independently added with probability p. This object is well-studied, and we know that in the critical window $p = p(n) = n^{-1} + \lambda n^{-4/3}$ the size of the macroscopic outbreaks are of order $n^{2/3}$ [12]. Within this critical window each vertex has approximately Poisson(1) many neighbors, so in particular it has light tails. In turn, the process $Z_n(h)$ corresponding to the largest component has a scaling limit and that limit is a continuous process [147]. We stress that this is not because we are looking only at an epidemic started from a single individual. The same can be said if we infect $O(n^{1/3})$

individuals on day 0 [57].

To capture some super-spreading phenomena we focus mostly on the configuration model with a heavy-tailed degree distribution: $\mathbb{P}\deg(i) = k) \sim ck^{-(2+\alpha)}$ for some $\alpha \in (1, 2)$, along with some other technical assumptions dealing with criticality. The configuration model is a graph on n vertices chosen randomly over all graphs with a prescribed degree sequence. See Chapter 7 of [173] for an introduction to this model. We omit the case $\alpha = 2$ because this model falls within the same universality class as the critical Erdős-Rényi random graph $G(n, n^{-1} + \lambda n^{-4/3})$ [28, 59] and so, up to some scaling factors, the structure of the processes $Z_n(h)$ on largest components (which correspond to the largest possible outbreaks) will be asymptotically the same as those in the Erdős-Rényi random graph. In the asymptotic regime we study, the largest outbreaks are of order $O(n^{\frac{\alpha}{\alpha+1}})$ and scaling limits of $Z_n(h)$ will possess positive jumps. These positive jumps come from presence of the super-spreading individuals.

We also restrict our focus to critical regimes. One reason is general principle that what happens at a phase transition is often interesting. Another is that while there are some important results on the structure of the largest components of the critical heavy-tailed configuration model [59, 109], there is not much information on the structure of the disease outbreaks. In this vein, there are results in the literature on the behavior of the largest outbreak when initially only a single individual is infected. While studying a model similar to ours where edges are kept with probability $p \in [0, 1]$ but are otherwise deleted, the authors of [43] show that there is a parameter R_0 such that if $R_0 \leq 1$ then only outbreaks of size $o(n)$ as $n \to \infty$ can occur whereas if $R_0 > 1$ there is a positive probability that an outbreak of size $O(n)$ occurs as $n \to \infty$. See also [107, 148, 149]. A continuous time analog of that model was studied in [38] and there the authors show that there is a similar phase transition between outbreaks of size $o(n)$ and outbreaks which are of size $O(n)$ with positive probability. Those authors also describe some of the large n behavior of $Z_n(t)$ (the number of individuals infected at a continuous time $t \geq 0$) conditionally on having an outbreak of size $O(n)$, but they do not provide information for what happens at the phase transition. We hope to fill in this gap in the literature.

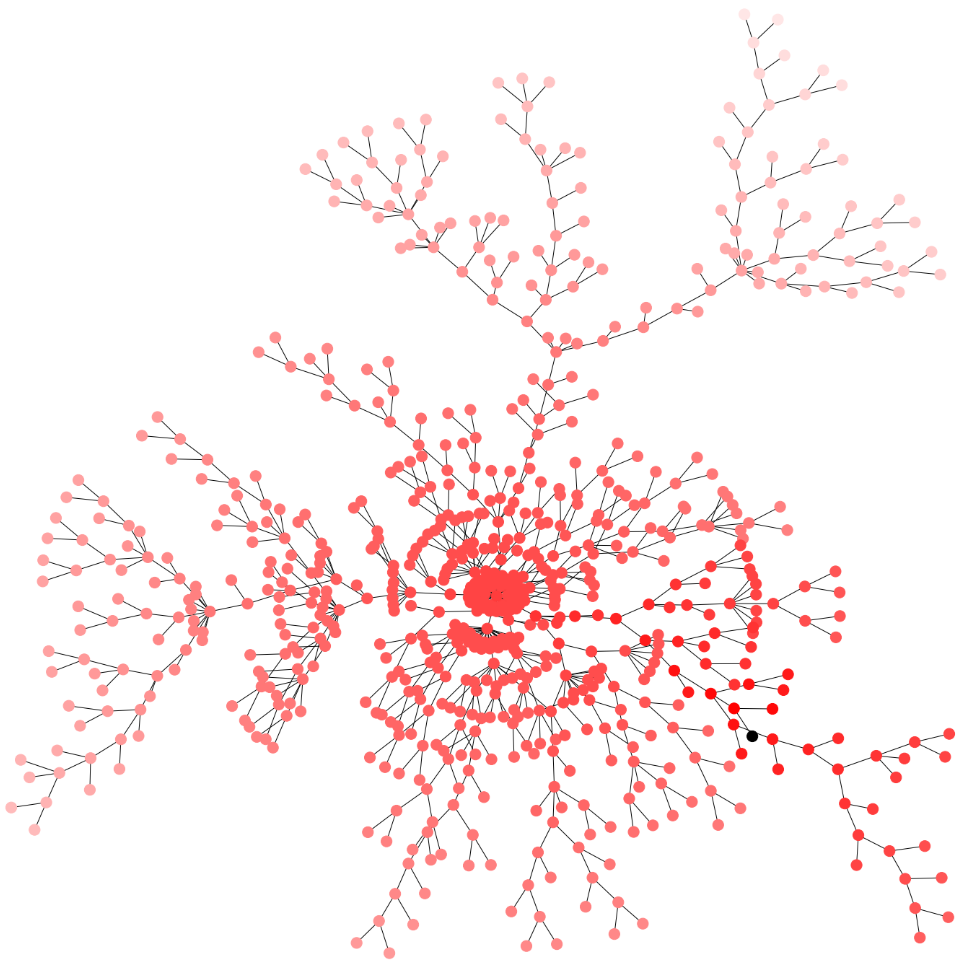


Figure 5.2: A simulation of the largest outbreak on a configuration model with heavy-tailed degree distribution with $\alpha = 3/2$. This component has 735 vertices, while the entire graph has 70,000. The black node is the first vertex to be infected, and then darker shades indicate that the corresponding vertex infected earlier in the outbreak. Most of the vertices have small degree (≤ 3); however, there are some vertices with large degree. The large red blob in the middle of the image comes from a vertex of relatively large degree, i.e. a super-spreader. We can also see that there is another super-spreader depicted just below that red blob.

5.1.1 Weak convergence results

Let us discuss a little more formally the configuration model. Before doing so, we recall that a multi-graph can have multiple edges and self-loops while a simple graph does not contain multiple edges nor self-loops. In terms of our approach to studying epidemics, self-loops and multiple edges do not make any physical sense because, for example, an infected individual cannot reinfect themself.

Given $\mathbf{d}^n = (d_1, \cdots, d_n)$ a finite sequence of strictly positive integers $d_j \geq 1$, the configuration model $M(\mathbf{d}^n)$ is the random multi-graph chosen randomly over all multi-graphs G on the vertex set $[n] := \{1, \cdots, n\}$ where the degree (counted with multiplicity) of vertex j is $\deg(j) = d_j$. In order to construct such a multi-graph we need $\sum_{j=1}^n d_j$ to be even, and two algorithms for its construction will be discussed in Section 5.5.1. We say that any such graph G has degree sequence $\mathbf{d}^n$.

A priori it may not be possible to construct a simple graph on with degree sequence $\mathbf{d}^n$ because, for example, a single vertex may have degree $d_i > \sum_{j\neq i} d_j$. However, if there is a simple graph with degree sequence $\mathbf{d}^n$, then conditionally on the event $\{M(\mathbf{d}^n) \text{ is simple}\}$ the graph is uniformly distribution over all simple graphs with degree sequence $\mathbf{d}^n$ [173, Proposition 7.15]. Moreover for the asymptotic regime we study it makes no difference [59] whether or not we examine simple graphs or multi-graphs so we will just say "graph."

One aspect of randomness for the configuration model comes from taking the graph to be randomly constructed over all graphs with a fixed deterministic degree sequence. Another comes from taking the degree sequence itself to be random, say, with a common distribution ν on $\{1, 2, \cdots\}$. We then generate the graph conditionally given this degree distribution. That is we generate $M(\mathbf{d}^n)$ where d_j are i.i.d. with common law ν. We may have to replace d_n with $d_n + 1$ to obtain the proper parity; however, this does not affect the analysis [59]. To distinguish between these two situations we will write $M_n(\nu)$ instead of $M(\mathbf{d}^n)$.

We focus on the degree distributions studied by Joseph [109] and Conchon-Kerjan and Goldschmidt [59]:

$$\lim_{k\to\infty} k^{(\alpha+2)}\nu(k) = c \in (0, \infty), \qquad \mathbb{E}[d_1] = \delta \in (1, 2), \qquad \mathbb{E}[d_1^2] = 2\mathbb{E}[d_1], \tag{5.1}$$

for some $\alpha \in (1, 2)$. The third statement about the second moment and the mean imply that we are examining the random graph at criticality [107, 148, 149]. This means that there is no giant component, i.e. there is no single component which contains a positive proportion of the total number of vertices. Instead, there are macroscopic components which are of order $O(n^{\frac{\alpha}{\alpha+1}})$.

In order to obtain scaling limits for a height profile represent the number of people infected on day h, we would either need to look at the case where a significant number of individuals are infected on day zero, or focus on the largest possible outbreaks. We focus on the latter situation and hence decompose the graph $M_n(\nu)$ into its connected components $G_n^1, G_n^2, \cdots$ where they are indexed so

$$\#G_n^1 \geq \#G_n^2 \geq \cdots .$$

In order to know how a disease spreads through G_n^i, we need to know its source. We will start the spread from a single vertex ρ_n^i chosen with probability proportional to its degree, and we will say that the component G_n^i is *rooted* at the vertex ρ_n^i.

The selection of ρ_n^i is a size-biased sample and not a uniform sample, but this is for good reason. In terms of how a disease spreads through a community, vertices with higher degree have more neighbors from whom they can catch the disease and so we should expect these vertices to be infected earlier in the outbreak. This has been observed in a survey of how influenza (seasonal or the H1N1 variant) spread through Harvard in 2009 [55]. Researchers surveyed two sets of students twice-weekly to see when they developed flu-like symptoms. One set was a random sample of all students and the other was a sample of friends nominated by this original set. The set of friends was size-biased sample of the students at Harvard and not a uniform sample. Sometimes called “the friendship paradox,” this is just the observation that the average number of friends of friends is always greater than the average number of friends [91]. In the study of influenza, the set of friends showed flu-like symptoms earlier than the uniform random sample. See also [92].

Of course, there are only a finite number K_n, say, of connected components which correspond to each of the outbreaks. To simplify the presentation we set G_n^i for $i > K_n$ as the graph on a single vertex with no edges and rooted at its only vertex.

The i^{th} largest possible outbreak is then described by the process $Z_{n,i} = (Z_{n,i}(h) : h =$

$0, 1, \cdots)$ defined by

$$Z_{n,i}(h) = \#\{v \in G_n^i : d(\rho_n^i, v) = h\}, \tag{5.2}$$

where $d(-,-)$ is the graph distance on G_n^i. In terms of the graph, $Z_{n,i}$ is the height profile of the component G_n^i. Our first result is the joint convergence of the processes $Z_{n,i}$ to a time-change of some excursion processes $\tilde{e}_i = (\tilde{e}_i(t); t \geq 0)$. The processes $\tilde{e}_i$, for $i \geq 1$, are the excursions above past minima of a certain stochastic process $\tilde{X}$ obtained by an exponential tilting of a spectrally positive α-stable processes. See Section 5.3.1 for more information on these processes.

Theorem 5.1.1. *Fix some $\alpha \in (1, 2)$, and some distribution ν satisfying (5.1). In the product Skorohod topology on $\mathbb{D}(\mathbb{R}_+)^\infty$, the following convergence holds*

$$\left(\left(n^{-\frac{1}{\alpha+1}} Z_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t \rfloor); t \geq 0\right); i \geq 1\right) \overset{d}{\Longrightarrow} ((Z_i(t); t \geq 0); i \geq 1),$$

where Z_i is the unique càdlàg solution to

$$Z_i(t) = \tilde{e}_i \circ C_i(t), \qquad C_i(t) = \int_0^t Z_i(s)\, ds, \qquad \inf\{t > 0 : C_i(t) > 0\} = 0,$$

where $\tilde{e}_i = (\tilde{e}_i(t) : t \geq 0)$ are defined in equation (5.12) and depend on the value α, c and δ in (5.1).

5.1.2 A single macroscopic outbreak and the α-stable graph

More has been said about the graph components G_n^i in the literature. Joseph [109] has argued that the size of the component G_n^i scaled down by $n^{\frac{\alpha}{\alpha+1}}$ converges to a random variable ζ_i for each $i \geq 1$, which in fact can be seen to be $\zeta_i = \inf\{t > 0 : \tilde{e}_i(t) = 0\}$. Conchon-Kerjan and Goldschmidt [59] generalize Joseph's results and show that the graph G_n^i itself has a scaling limit which is a random rooted compact measured metric space $\mathcal{M}_i = (\mathcal{M}_i, \mathrm{d}_i, \rho_i, \mu_i)$. Here d_i is a metric on $\mathcal{M}_i$, $\rho_i \in \mathcal{M}_i$ is a specified element and μ_i is a finite Borel measure on $\mathcal{M}_i$.

This means not only does height profile (the number of infected people on day h) converge Theorem 5.1.1, but there is some limiting continuum structure of the components G_n^i which is represented by these continuum spaces $(\mathcal{M}_i; i \geq 1)$. The standard construction of these continuum limits, first obtained in the critical Erdős-Rényi case in [4, 5], are constructions in a depth-first manner: the spaces are obtained by gluing together k pairs of points on a continuum random tree with a depth-first selection of the pairs. This gluing procedure changes the distance from the origin and therefore it is non-trivial to argue convergence of the height profiles similar to Theorem 5.1.1 from the results of [59].

The processes $\tilde{e}_i = (\tilde{e}_i(t); t \geq 0)$ in Theorem 5.1.1 and the graph $\mathcal{M}_i$ are of a random length and mass, respectively. That is

$$\tilde{e}_i(t) > 0 \qquad \text{if and only if} \quad t \in (0, \zeta_i)$$

for some random ζ_i and, moreover, $\zeta_i = \mu_i(\mathcal{M}_i)$. This does complicate the analysis somewhat; however, conditionally given the values $(\zeta_i; i \geq 1)$ the excursions $\tilde{e}_i$ (resp. the spaces $\mathcal{M}_i$) are independent and are described by a scaling of an excursion (resp. metric measure space) of unit length (resp. unit mass) [59]. Therefore in order to understand the scaling limit $Z_1 = (Z_1(t); t \geq 0)$ of a single macroscopic outbreak $Z_{n,1}$, we can study the structure of a the process $\tilde{e}_1$ conditioned on $\zeta_1 = 1$.

To do this we let $\mathbf{e} = (\mathbf{e}(t); t \in [0,1])$ denote a standard excursion [54] of a spectrally positive α-stable Lévy process $X = (X(t); t \geq 0)$. To simplify our proofs, we will work with situation where the Laplace transform of X satisfies

$$\mathbb{E}\left[\exp(-\lambda X(t))\right] = \exp\left(A\lambda^\alpha t\right), \qquad \forall \lambda, t \geq 0, \quad \text{where} \quad A = \frac{c\Gamma(2-\alpha)}{\delta\alpha(\alpha-1)}, \tag{5.3}$$

for c, δ defined in (5.1). We remark that this excursion depends on the value A; however, the results also hold for any value of A by using scaling properties of Lévy processes and their associated height processes.

We also recall from above that $\mathcal{M}_i$ are obtained by gluing together a finite collection of pairs of points in a continuum tree. This is the surplus of the continuum random graph $\mathcal{M}_i$. We will let $\mathscr{G}^{(\alpha,k)} = (\mathscr{G}^{(\alpha,k)}, d, \rho, \mu)$ denote the graph $\mathcal{M}_1$ conditioned on $\mu_1(\mathcal{M}_1) = 1$ and having surplus k. A precise construction of this object will be delayed until Section 5.3.4, but it suffices to say that it will be constructed from an excursion $\mathbf{e}^{(k)} = (\mathbf{e}^{(k)} : t \in [0,1])$ defined by the polynomial tilting

$$\mathbb{E}\left[f(\mathbf{e}^{(k)}; t \in [0,1])\right] \propto \mathbb{E}\left[\left(\int_0^1 \mathbf{e}(t)\, dt\right)^k f(\mathbf{e}; t \in [0,1])\right]. \tag{5.4}$$

The continuum object $\mathscr{G}^{(\alpha,k)}$, in our case, represents the limiting structure of the history of the disease spread. With this, we can ask several questions in the hope that this will shed light on the structure of $M_n(\nu)$. What is the structure of the disease outbreak, i.e. what is the height profile of the graph $\mathscr{G}^{(\alpha,k)}$? When does a uniformly chosen person get infected, or when do a finite number of uniformly chosen individuals get infected? When does the outbreak die out, that is, when is the last person infected? In terms of the graph $\mathscr{G}^{(\alpha,k)}$ this is asking what is the radius of the graph $\mathscr{G}^{(\alpha,k)}$. What's the most number of people infected at any one time or, in terms of the continuum graph, what is the distribution of the width of $\mathscr{G}^{(\alpha,k)}$? There are many more different questions we can ask, but we have answers to these questions.

We start by answer the first question: what is the height profile of the graph $\mathscr{G}^{(\alpha,k)}$, from which the others will follow by analysis of an integral equation.

Theorem 5.1.2. *Fix $k \geq 0$, and $\alpha \in (1,2)$. Let $\mathscr{G}^{(\alpha,k)}$ be the α-stable continuum random graph, constructed from a spectrally positive Lévy process with Laplace exponent (5.3) and rooted at a point $\rho \in \mathscr{G}^{(\alpha,k)}$.*

1. *Let $B(x,t)$ is the closed ball of radius t centered at x. The process $\mathbf{c} = (\mathbf{c}(t); t \geq 0)$ defined by $\mathbf{c}(t) = \mu(B(\rho,t))$ is absolutely continuous and*

$$\mathbf{c}(t) = \int_0^t \mathbf{z}(s)\,ds$$

   for a càdlàg process $\mathbf{z} = (\mathbf{z}(t); t \geq 0)$.

2. *The process $(\mathbf{z}(t); t \geq 0) \stackrel{d}{=} (z(t); t \geq 0)$ where z is the unique càdlàg solution to*

$$z(t) = \mathbf{e}^{(k)}\left(\int_0^t z(s)\,ds\right), \qquad \inf\left\{t > 0 : \int_0^t z(s)\,ds > 0\right\} = 0.$$

We can now answer all of the other questions once we know the height profile.

Corollary 5.1.3. *1. The radius of the graph $\mathscr{G}^{(\alpha,k)}$ is given by*

$$\sup_{v \in \mathscr{G}^{(\alpha,k)}} d(\rho, v) \stackrel{d}{=} \int_0^1 \frac{1}{\mathbf{e}^{(k)}(s)}\,ds.$$

2. The width of the graph $\mathscr{G}^{(\alpha,k)}$ is given by

$$\sup_{t\geq 0} \mathbf{z}(t) \stackrel{d}{=} \sup_{t\in[0,1]} \mathbf{e}^{(k)}(t).$$

3. Let $V \in \mathscr{G}^{(\alpha,k)}$ be distributed according to the mass measure μ, and let U denote a uniform random variable on $(0,1)$. Then

$$d(\rho, V) \stackrel{d}{=} \int_0^U \frac{1}{\mathbf{e}^{(k)}(s)}\, ds.$$

4. More generally, for any $n \geq 1$, let $V_1, \cdots, V_n$ denote random points distributed according to μ on $\mathscr{G}^{(\alpha,k)}$. Let $R_{(1)} \leq R_{(2)} \leq \cdots R_{(n)}$ denote the order statistics of $d(\rho, V_1), \cdots, d(\rho, V_n)$. Let $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$ denote the order statistics for an i.i.d. sample of n uniform random variables. Then

$$\left(R_{(1)}, \cdots, R_{(n)}\right) \stackrel{d}{=} \left(\int_0^{U_{(1)}} \frac{1}{\mathbf{e}^{(k)}(s)}\, ds, \cdots, \int_0^{U_{(n)}} \frac{1}{\mathbf{e}^{(k)}(s)}\, ds\right).$$

After discussing the integral equation involved in the statement of part 2 of Theorem 5.1.2, which is called the Lamperti transform in the literature, we will show how Corollary 5.1.3 follows from Theorem 5.1.2 in Section 5.3.2.

5.1.3 Relation to other works and proof structure

Epidemics on random graphs are important for many areas in the applied sciences, see [25, 152, 175, 176] and references therein for a non-exhaustive collection of such works. One difficulty in describing the limiting behavior comes from analyzing the influence that the specific degree distribution has on the local structure of the graph.

One approach to overcoming this issue is by using a *mean-field* approach [25]. A typical approximation is in continuous time where each infected vertex v is infected for an exponential time, and infects its neighbors at independent exponential rates. On homogeneous networks the behavior of $Z(t)$, the number of people infected at time $t \in \mathbb{R}_+$, is modeled by the ordinary differential equation

$$\frac{dz(t)}{dt} = \lambda\mu z(t)(1 - z(t)) \qquad \text{where} \quad z(t) = \frac{1}{n}Z(t).$$

A more careful analysis can be done on heterogeneous networks, where one can track the proportion of vertices with degree k infected at a certain time. A remarkable thing is that this approach, while losing a lot of information about the specific local structure, it can be used to find heuristics on the proper scaling of the graphs or epidemic, see [152].

A more detailed approach to studying heterogeneous networks was taken by Volz in [176], and rigorously proved in [60] under a fifth moment condition. In these works the population of size n is broken into 3 compartments - the susceptible, the infected and the recovered - and individuals in one compartment are moved to another compartment (i.e. an infected individual recovers or a susceptible individual is infected) at certain exponential rates. The global changes in the proportional size of the outbreak is described, to first order, by just the size of the respective compartments and the degree distribution. Here the limiting structure is described by a system of deterministic ordinary differential equations, which depend on the degree distribution. Deterministic limiting equations, perhaps after some random time T_0, were also obtained in [106] under a second moment condition.

Our approach is different and takes it idea from studies of height profiles of random trees and branching processes. Particularly, we focus on the approach implicit in [14], and later studied in [18, 49, 50] and we use the so-called Lamperti transform. This transform was originally used for a path-by-path bijection between continuous state branching processes and a certain class of Lévy processes. This transform was originally stated by Lamperti in [128], but was proved later by Silverstein [163]. See also [48].

We can describe the discretized version in our situation as follows. Instead of looking at the total number of people infected on day h, we look at the number of individuals that person v_j infects, when v_j is the j^{th} individual who contracts the disease. This corresponds to a breadth-first ordering of the underlying connected component of the graph. Call this number of newly infected individuals χ_j, and let X be the breadth-first walk

$$X(k) = \sum_{j=1}^{k} (\chi_j - 1).$$

It was this walk on the Erdős-Rényi random graph that Aldous used in [12] to describe the scaling limits of the component sizes of $G(n, p)$ in the critical window, and an analogous walk was used by Joseph [109] for the configuration model.

An interesting property of this walk is that the number of people infected on day h solves the difference equation

$$Z(h) = Z(0) + X\left(C(h-1)\right), \qquad C(h) = \sum_{j=0}^{h} Z(j).$$

As far as the author is aware, the first instance of this identity can be found in [14] with a slightly more complicated formulation. See the Introduction of [49] for a proof of this equality. The authors of [49, 50] studied the scaled convergence of solutions of the above equation (with the addition of an immigration term) to its continuum analog

$$Z(t) = X\left(\int_0^t Z(s)\,ds\right).$$

Unfortunately, there is not a unique solution to this integral equation when $X(0) = 0$ and so proving weak convergence is quite difficult. For certain models of random trees one can prove a weak convergence result [14, 18, 120], and also works for the Erdős-Rényi random graph when $X(0) > 0$ as was shown in Chapter 4 which borrows from [57].

We overcome the uniqueness problem by arguing that the rescaled processes $Z_{n,i}(h)$ are tight, and we further show that each subsequential weak limit must be of a particular form. This approach to overcoming the uniqueness problem was used in [18] to study trees with a certain degree distribution, as opposed to graphs with a given degree distribution in the present situation. While, at first, these two discrete models may seem related, the proofs are quite different. In [18], the authors use a combinatorial transformation of the tree to show that that subsequential limits must be of a particular form. We, instead, show that this follows automatically once we know the the underlying graph converges to a measured metric space. In turn, in Section 5.6 we discuss how our abstract convergence results described in Section 5.2 can be applied to the rank-1 inhomogeneous model [13, 31, 44, 45].

5.2 General Weak Convergence Results

5.2.1 General Weak Convergence Approach

Let us now discuss the general set up for our weak convergence arguments. In the introduction we discussed the epidemic, which can be realized as the height profile of a connected component of a random graph. Explicitly those graphs were viewed as a metric space, but we implicitly equipped them with the counting measure. We phrase our results in terms of more general measures on random graphs, which will likely be useful in inhomogeneous models in [13, 44, 45]. The epidemiological interpretation of considering non-uniform measures is not immediately clear; however, we could think of the unequal mass of vertices as measuring the size of a clique in a community which was reduced to single vertex.

A major assumption of these results is the convergence of graphs as measured metric spaces. We delay a more detailed discussion of this topic until Section 5.3.3. For now it suffices to say that we can equip the space $\mathfrak{M}$ of (equivalent classes) of pointed measured metric spaces with additional boundedness assumptions with a metric which turns $\mathfrak{M}$ into a Polish space. This metric is called the Gromov-Hausdorff-Prohorov metric, and we will denote it by d_{GHP}.

We will denote a generic element of $\mathfrak{M}$ as $\mathcal{M} = (\mathcal{M}, \rho, \mathrm{d}, \mu)$ where $(\mathcal{M}, \mathrm{d})$ is a metric space such that bounded sets have compact closure, $\rho \in \mathcal{M}$ is a specified point and μ is a Borel measure on $\mathcal{M}$ such that bounded sets have finite mass. For each $\alpha > 0, \beta \geq 0$ define the scaling operation by

$$\text{scale}(\alpha, \beta)\mathcal{M} := (\mathcal{M}, \rho, \alpha \cdot \mathrm{d}, \beta \cdot \mu).$$

Now let G denote a connected graph on, say, n vertices, with $\rho \in G$ a specified vertex. We view G as a measured metric space with graph distance and $\mathfrak{m}$ a finite measure such that each vertex has strictly positive mass. As we did implicitly before, we explore the graph in a breadth-first manner. The precise way in which this is done can vary depending on the graph model, but we assume that the vertices are labeled by $v_1, \cdots, v_n$ such that if $i < j$ then $d(\rho, v_i) \leq d(\rho, v_j)$. This trivially implies that $\rho = v_1$. This labeling can be viewed as an indexing of each individual who gets infected, so that if person B got infected after person A, then person A has a smaller index than person B.

We now discuss an underlying tree structure and breadth-first walk for the graph, which draws inspiration from the breadth-first tree and walk in [13]. The tree is constructed by looking at which vertices v_i infects in the graph G. More formally, we will say that vertex v_j is the child of v_i if $\{v_i, v_j\}$ is an edge in G, but $\{v_l, v_j\}$ is *not* an edge for all $l < i$. This implies $i < j$. In most models with a breadth-first exploration, v_j will be a child of v_i if vertex v_j is discovered while exploring the vertex attached to v_i.

We also suppose that there is some breadth-first walk X_G^{BF} as well:

$$X_G^{\mathrm{BF}}(\tau(k)) = X_G^{\mathrm{BF}}(\tau(k-1)) - \mathfrak{m}(v_k) + \sum_{u \text{ child of } v_k} \mathfrak{m}(u), \qquad \tau(k) = \sum_{j\le k} \mathfrak{m}(v_j), \tag{5.5}$$

and with $X_G^{\mathrm{BF}}(0) = 0$. How the process X_G^{BF} behaves on the intervals $(\tau(k-1), \tau(k))$ will play no important role in this paper. The breadth-first walk used by Aldous and Limic in their classification of the multiplicative coalescence [13] satisfies equation (5.5). Later an analogous walk satisfying (5.5) was used in [14] to describe the inhomogeneous continuum random tree and extend Jeulin's identity [108]. When $\mathfrak{m}$ is a uniform measure on G, this walk will be the breadth-first Łukasiewicz path [131].

Importantly for us, the walk X_G^{BF} encodes the masses and tree structure of $v_1, \cdots, v_n$. However there is no clean functional amenable to scaling limits which allows us to reconstruct the genealogical structure from this breadth-first walk.

We now define the height profile of G by

$$Z_G(h) = \mathfrak{m}\left\{v \in G : d(\rho, v) = h\right\}.$$

It will be useful to define its cumulative sum as well:

$$C_G(h) = \sum_{j=0}^{h} Z_G(j) = \mathfrak{m}\left\{v \in G : d(\rho, v) \le h\right\}.$$

As observed in [14, equations (13-14)], $Z_G(h)$ solves the following difference equation:

$$Z_G(h+1) = Z_G(0) + X_G^{\mathrm{BF}} \circ C_G(h). \tag{5.6}$$

To describe what happens in the $n \to \infty$ limit, let $(G_n; n \ge 1)$ be a sequence of connected random graphs on a finite number of vertices, viewed as a measured metric spaces where G_n is equipped with the measure $\mathfrak{m}_n$. We write X_n^{BF} for the breadth-first walk $X_{G_n}^{\mathrm{BF}}$. We prove the following in Section 5.4.

Theorem 5.2.1. *Suppose that there exists a sequence $\alpha_n \to \infty$, and that $\gamma_n := \mathfrak{m}_n(G_n) \to \infty$ a.s. In addition assume:*

1. *In the Skorohod space* $\mathbb{D}([0,1],\mathbb{R})$*, the following weak convergence holds*

$$\left(\frac{\alpha_n}{\gamma_n} X_n^{\mathrm{BF}}(\gamma_n t); t \in [0,1]\right) \stackrel{d}{\Longrightarrow} (X(t); t \in [0,1])\,,$$

   where X *is a process such that almost surely* $X(0) = X(1) = 0$*,* $X(t) > 0$ *for all* $t \in (0,1)$ *and* $X(t) - X(t-) \geq 0$ *for all* t*;*

2. *There exists a random pointed measured metric space* $\mathcal{M} = (\mathcal{M}, \rho, \mathrm{d}, \mu)$ *which is locally compact and has a boundedly finite measure such that*

$$\mathrm{scale}(\alpha_n^{-1}, \gamma_n^{-1}) G_n \stackrel{d}{\Longrightarrow} \mathcal{M},$$

   weakly in the Gromov-Hausdorff-Prohorov topology.

3. *For each* $\varepsilon > 0$*,* $\mu(B(\rho,\varepsilon) \setminus \{\rho\}) > 0$ *for all* $\varepsilon > 0$*.*

4. $\frac{\alpha_n}{\gamma_n} \sup_{v \in G_n} \mathfrak{m}_n(v) \to 0$ *as* $n \to \infty$ *in probability.*

Then

1. *There is joint convergence in* $\mathbb{D}(\mathbb{R}_+,\mathbb{R})^2$*:*

$$\left(\left(\frac{\alpha_n}{\gamma_n} Z_n(\lfloor \alpha_n t \rfloor); t \geq 0\right), \left(\frac{1}{\gamma_n} C_n(\lfloor \alpha_n t \rfloor); t \geq 0\right)\right) \stackrel{d}{\Longrightarrow} ((Z(t); t \geq 0), (C(t); t \geq 0))$$

   where Z *and* C *are the unique càdlàg solution to*

$$Z(t) = X \circ C(t), \qquad C(t) = \int_0^t Z(s)\,ds, \qquad \inf\{t : C(t) > 0\} = 0; \tag{5.7}$$

2. *The measure* μ *on* $\mathcal{M}$ *satisfies*

$$(\mu(B(\rho,t)); t \geq 0) \stackrel{d}{=} (C(t); t \geq 0)$$

Let us make some important remarks on the assumptions in Theorem 5.2.1. Assumption (1) is the convergence of the breadth-first walk, which is required in order to have a description of the limiting process Z as described above, barring some stochastic analysis tools that can be used in particular cases [153]. Assumptions (2) and (3) are how we overcome any possible uniqueness problems that were identified in [18] (see Proposition 5.3.1 below). Particularly, assumption (3) allows for the classification of the limit C satisfying $\inf\{t : C(t) > 0\} = 0$. Lastly, assumption (4) is so that the term $Z_n(0)\overset{d}{\Longrightarrow}0$ as $n \to \infty$. Without this assumption we are left to deal with a simpler situation to which we can use the known weak convergence results in [49].

As the reader may guess, this formulation will not be helpful for the proof of Theorem 5.1.1 nor in the study of any of the macroscopic outbreaks for random graphs. Instead, the above theorem works only with a single macroscopic component. In order to prove Theorem 5.1.1 we must develop a joint convergence result where each of the macroscopic components of a graph converge to some limiting graphs structure. This is something that appears quite often in the literature on continuum random graphs, dating back to the celebrated result of Addario-Berry, Broutin and Goldschmidt [5]. We now suppose that we have a sequence of graphs $(G_n; n \geq 1)$ on a finite number of vertices with a measure $\mathfrak{m}_n$. For each n we denote the connected components of G_n as $(G_n^i; i \geq 1)$, ordered so that

$$\mathfrak{m}_n\left(G_n^1\right) \geq \mathfrak{m}_n\left(G_n^2\right) \geq \cdots .$$

Again, for convenience we will say that G_n^i is a graph on a single vertex where the vertex has mass 0 for all $i > K_n$. We view each of the components as a measured metric space with graph distance, and we select a vertex ρ_n^i from each component to start the breadth-first walks. Here we write $X_{n,i}^{\mathrm{BF}}$ for the breadth-first walk on G_n^i which, by assumption, satisfies equation (5.5) with the obvious notation changes. Additionally we extend it by constancy to be a function on all of $\mathbb{R}_+$:

$$X_{n,i}^{\mathrm{BF}}(t) = X_{n,i}^{\mathrm{BF}}(\mathfrak{m}_n(G_n^i)) \qquad \forall t \geq \mathfrak{m}_n(G_n^i).$$

Let $Z_{n,i} = (Z_{n,i}(h))$ be the height profile of the i^{th} component G_n^i. They solve an equation analogous to (5.6) with the obvious notation change.

We prove the following

Theorem 5.2.2. *Suppose there exists two sequence $\alpha_n \to \infty$ and $\gamma_n \to \infty$ such that*

1. *In the product Skorohod space* $\mathbb{D}^\infty$ *the following weak convergence holds:*

$$\left(\left(\frac{\alpha_n}{\gamma_n} X^{\mathrm{BF}}_{n,i}(\gamma_n t); t \ge 0\right); i \ge 1\right) \stackrel{d}{\Longrightarrow} \left((X_i(t); t \ge 0)\,; i \ge 1\right)$$

where almost surely, X_i *does not possess negative jumps and there exists a* $\zeta_i > 0$ *such that* $X(t) > 0$ *if and only if* $t \in (0, \zeta_i)$*.*

2. *There exists a sequence of pointed measured metric spaces* $\mathcal{M}_i = (\mathcal{M}_i, \rho_i, d_i, \mu_i)$ *which is locally compact and has a boudedly finite measure such that*

$$\left(\mathrm{scale}(\alpha_n^{-1}, \gamma_n^{-1}) G_n^i; i \ge 1\right) \stackrel{d}{\Longrightarrow} (\mathcal{M}_i; i \ge 1)$$

weakly in the product Gromov-Hausdorff-Prohorov topology.

3. *Suppose that* $\mu_i(B(\rho_i, \varepsilon) \setminus \{\rho_i\}) > 0$ *for all* $\varepsilon > 0$*.*

4. $\frac{\alpha_n}{\gamma_n} \sup_{v \in G_n} \mathfrak{m}_n(v) \to 0$ *as* $n \to \infty$ *in probability.*

Then

1. *In the product Skorohod topology*

$$\left(\left(\frac{\alpha_n}{\gamma_n} Z_{n,i}(\lfloor \alpha_n t \rfloor), \frac{1}{\gamma_n} C_{n,i}(\lfloor \alpha_n t \rfloor); t \ge 0\right); i \ge 1\right) \stackrel{d}{\Longrightarrow} \left((Z_i(t), C_i(t); t \ge 0)\,; i \ge 1\right),$$

where (Z_i, C_i) *is the unique càdlàg solution to*

$$Z_i(t) = X_i \circ C_i(t), \qquad C_i(t) = \int_0^t Z_i(s)\, ds, \qquad \inf\{t : C_i(t) > 0\} = 0.$$

2. *For each* $i \ge 1$*,*

$$(\mu_i(B(\rho_i, t)); t \ge 0) \stackrel{d}{=} (C_i(t); t \ge 0)\,.$$

5.2.2 Compactness Corollaries

Let us begin with the first corollary, which follows from Theorem 5.2.1 and a result in [18] recalled in Proposition 5.3.1 below.

Corollary 5.2.3. *If the hypotheses of Theorem 5.2.1 are met, then*

$$\int_{0+} \frac{1}{X(s)}\, ds < \infty \qquad a.s.$$

The above corollary avoids a hypothesis in Theorem 1 in [18], but this comes at the expense of assuming convergence in the Gromov-Hausdorff-Prohorov topology of an underlying metric space, which is a difficult hypothesis to verify. The inverse of Corollary 5.2.3 is interesting, because it gives a necessary condition for convergence in the Gromov-Hausdorff-Prohorov topology.

For certain models of random trees and random graphs, determining compactness of the candidates for limiting metric space is difficult. This has been a particular problem for the inhomogeneous continuum random trees introduced by Aldous, Camarri and Pitman [16, 51]. These trees are characterized by a parameter $\theta = (\theta_0, \theta_1, \cdots)$ and in [14], the authors showed that boundedness of the continuum random tree is equivalent to the almost sure finiteness of an integral $\int_0^1 \frac{1}{X(s)}\, ds$. A question was posed in [14] to develop useful criteria for compactness of the ICRT and determine if boundedness implied compactness. This problem was open for 16 years, but appears to be solved very recently in [37].

It is in this vein that we state the next corollary. It is a more abstract version of part (1) of Corollary 5.1.3 above, and follows from part (2) in Theorem 5.2.1 and Proposition 5.3.1.

Corollary 5.2.4. *Let $\mathcal{M}$ and X be as in Theorem 5.2.1, assume the hypotheses of Theorem 5.2.1 are met, and let $\operatorname{spt}(\mu) \subset \mathcal{M}$ denote the topological support of the measure μ. Then*

$$\sup_{v \in \operatorname{spt}(\mu)} d(\rho, v) \overset{d}{=} \int_0^1 \frac{1}{X(s)}\, ds.$$

5.3 Preliminaries

5.3.1 Lévy processes, height processes, excursions

In this section we briefly recall the construction of Ψ-height processes and their excursions which were discussed in Section 2.2.4.3. For more in depth discussion on the height processes and their excursions see the works of Le Gall, Le Jan and Duquense in [73, 132, 133] and for information about spectrally positive Lévy processes, see Bertoin's monograph [26].

Let $X = (X(t); t \geq 0)$ denote a spectrally postive, i.e. no negative jumps, Lévy process, and let $-\Psi$ denote its Laplace transform:

$$\mathbb{E}\left[\exp\left\{-\lambda X(t)\right\}\right] = \exp\left(t\Psi(\lambda)\right).$$

In order to discuss Ψ-height processes, we restrict our attention in this situation to have Ψ is of the form

$$\Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (e^{-\lambda r} - 1 + \lambda r 1_{[r<1]})\,\pi(dr),$$

where $\alpha \geq 0$, $\beta \geq 0$, $(r \wedge r^2)\,\pi(dr)$ is a finite measure along with

$$\beta > 0 \qquad \text{or} \qquad \int_{(0,1)} r\pi(dr) = \infty.$$

The last assumption occurs if and only if the paths of X have infinite variation almost surely.

The Ψ-height process $H = (H(t); t \geq 0)$ is a way to give a measure (in a local time sense) to the set

$$\{s \in [0,t] : X(s) = \inf_{s \leq r \leq t} X(r)\}. \tag{5.8}$$

Slightly more formally, under the the additional assumption that

$$\int_1^\infty \frac{1}{\Psi(\lambda)}\,d\lambda = \infty,$$

there exists a *continuous* process $H = (H(t); t \geq 0)$ such that for all $t \geq 0$ then

$$H(t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t 1_{[X(s) - \inf_{r \in [s,t]} X(r) \leq \varepsilon]}\,ds, \tag{5.9}$$

where the limit is in probability. See [132] and [73, Section 1.2] for more details. In the case where $\beta > 0$, the process H can be seen [73, equation (1.7)] to satisfy

$$H(t) = \frac{1}{\beta} \,\mathrm{Leb}\left\{\inf_{s \leq r \leq t} X(r) : s \in [0,t]\right\}.$$

In particular, when X is a standard Brownian motion then

$$H(t) = 2\left(X(t) - \inf_{s \leq t} X(s)\right)$$

is twice a reflected Brownian motion.

One can also do this same procedure to the excursions of X. That is, if $I(t) = \inf_{s\le t} X(s)$ is the running infimum of X then the process $-I$ acts as a (Markovian) local time at level 0 for the reflected process $X - I$ [26, Chapter IV]. Moreover, by looking at $T(y) = \inf\{t \ge 0 : I(t) < -y\}$, we can talk about the excursions of $X - I$ between times $T(y-)$ and $T(y)$. As well [73, Section 1.1.2], it is possible to define the height process H for the excursions of X above its running infimum. The associated excursion measure will be denoted by N. To avoid confusion, we will write (e, h) for (X, H) under the excursion measure N.

5.3.1.1 *Stable Processes and Tilting*

We now restrict our attention to the stable case where

$$\Psi(\lambda) = A\lambda^\alpha, \qquad \alpha \in (1, 2]. \tag{5.10}$$

The process X satisfies the scaling [26]

$$(X(t); t \ge 0) \overset{d}{=} \left(k^{-1/\alpha} X(kt); t \ge 0\right), \qquad \forall k > 0.$$

Similarly, the height process H satisfies the scaling

$$(H(t); t \ge 0) \overset{d}{=} \left(k^{(1-\alpha)/\alpha} H(kt); t \ge 0\right), \qquad \forall k > 0,$$

which can be derived from (5.9).

Remark 5.3.1. By scaling the Lévy process X, the constant A in (5.10) can be taken to equal 1 and this is typically done in the literature. We will not do this when proving Theorems 5.1.1 or 5.1.2 in order to simplify the presentation. By using scaling properties for both X and H, it is possible to prove the results in Theorem 5.1.2 and Corollary 5.1.3 continue to hold when $A = 1$.

As originally observed by Aldous [12], one can encode the size of components of a random graph by a certain walk which possesses a scaling limit of the form $X(t) + f(t)$ where X is a Lévy process and $f(t)$ is a deterministic drift term. Aldous first proved this [12] within the critical window of Erdős-Rényi random graph where X is a Brownian motion and f is a quadratic function. This later extended to the α-stable case by Joseph [109] on the configuration model where X is a stable Lévy process.

For the α-stable case ($\alpha \in (1,2)$), Conchon–Kerjan and Goldschmidt [59] described the process in [109] via an exponential tilting of a Lévy process. That is they examine an α-stable process X and its associated height process H of the form and define $\tilde{X}$ and $\tilde{H}$ by

$$\mathbb{E}\left[F(\tilde{X},\tilde{H};[0,t])\right] = \mathbb{E}\left[\exp\left(-\frac{1}{\delta}\int_0^t s\,dX(s) - A\frac{t^{\alpha+1}}{(\alpha+1)\delta^\alpha}\right)F(X,H;[0,t])\right] \qquad (5.11)$$

where A is as in (5.10) and F is a function on the paths of upto time t. The A in our notation is $\frac{C_\alpha}{\delta}$ in the notation of [59].

The excursions of the process $\tilde{X}$ before time $t > 0$ can be described via the absolute continuity relationship in (5.11) and the excursions of X prior to time t. What is very useful for us is that all the excursions of

$$\tilde{R}(t) = \tilde{X}(t) - \tilde{I}(t)$$

above zero can be ordered by decreasing length [59, Lemma 3.5]. That is the lengths of the excursion intervals, $(\zeta_i; i \geq 1)$, can be indexed such that $\zeta_1 \geq \zeta_2 \geq \cdots \geq 0$. Corresponding the values ζ_i, there is an excursion interval (g_i, d_i) of length $d_i - g_i = \zeta_i$ such that $\tilde{R}(g_i) = \tilde{R}(d_i) = 0$ and $\tilde{R}(t) > 0$ for all $t \in (g_i, d_i)$. We define the excursion $\tilde{e}_i = (e_i(t); t \geq 0)$ by

$$\tilde{e}_i(t) = \tilde{R}((g_i + t) \wedge d_i), \qquad t \geq 0. \qquad (5.12)$$

These are the excursion which appear in Theorem 5.1.1.

We also let $\tilde{h}_i = (\tilde{h}_i(t); t \geq 0)$ be the excursion of $\tilde{H}$ which straddles (g_i, d_i) defined by

$$\tilde{h}_i(t) = \tilde{H}\left((g_i + t) \wedge d_i\right).$$

5.3.1.2 Normalized excursions and tilting

We now recall Chaumont's path construction of a normalized excursion of a spectrally positive α-stable Lévy process X. See [54] or [26, Chapter VIII] for more details on this. This allows for an simple description of the conditioning the excursion measure $N(\cdot|\zeta = x)$, for a fixed constant (deterministic) $x > 0$ and ζ is the duration of the excursion. These results also hold in the Brownian case $\alpha = 2.$, and we refer to Chapter XII of [156] for that treatment.

Define $\widehat{g}_1$ and $\widehat{d}_1$ by

$$\widehat{g}_1 = \sup\{s \le 1 : X(s) = I(s)\}, \qquad \widehat{d}_1 = \inf\{s > 1 : X(s) = I(s)\},$$

and define

$$\mathbf{e}(t) = \frac{1}{(\widehat{d}_1 - \widehat{g}_1)^{1/\alpha}} \left(X(\widehat{g}_1 + (\widehat{d}_1 - \widehat{g}_1)t) - X(\widehat{g}_1) \right), \qquad t \in [0,1]. \tag{5.13}$$

The normalized excursion $\mathbf{e} = (\mathbf{e}(t); t \in [0,1])$ has duration $\zeta = 1$, and its law is $N(\cdot|\zeta = 1)$. We obtain, the law $N(\cdot|\zeta = x)$ by scaling. Namely, set

$$\mathbf{e}_x(t) = x^{1/\alpha}\mathbf{e}(x^{-1}t), \qquad t \in [0,x],$$

and then $\mathbf{e}_x = (\mathbf{e}_x(t); t \in [0,x])$ has law $N(\cdot|\zeta = x)$.

This can also be done under the conditioning on the lifetime of the excursion of the height process H. See [70] or [146] for more information. We denote $\mathbf{h} = (\mathbf{h}(t); t \in [0,1])$ as the height process under the measure $N(\cdot|\zeta = 1)$ and (by the scaling for the height process) we write

$$\mathbf{h}_x(t) = x^{(\alpha-1)/\alpha}\mathbf{h}(x^{-1}t), \qquad t \in [0,x].$$

The normalized excursions of $\tilde{X}$ and $\tilde{H}$ are trickier to handle because the process $\tilde{X}$ does not have stationary increments. However, there is a relatively simple way of describing these in terms of an exponential tilting of the excursions $\mathbf{e}$ and $\mathbf{h}$ similar to Aldous' description in [12] in the Brownian case. We define the tilted processes denoted by, $\tilde{\mathbf{e}}_x^{(\delta)}$ and $\tilde{\mathbf{h}}_x^{(\delta)}$, by

$$\mathbb{E}\left[F(\tilde{\mathbf{e}}_x^{(\delta)}, \tilde{\mathbf{h}}_x^{(\delta)})\right] = \frac{\mathbb{E}[\exp(\frac{1}{\delta}\int_0^x \mathbf{e}_x(t)\,dt)F(\mathbf{e}_x, \mathbf{h}_x)]}{\mathbb{E}[\exp(\frac{1}{\delta}\int_0^x \mathbf{e}_x(t)\,dt)]} \tag{5.14}$$

When $x = 1$ or $\delta = 1$ we omit it from notation. The excursions $\tilde{\mathbf{e}}_x^{(\delta)}$ and $\tilde{\mathbf{h}}_x^{(\delta)}$ are shown in [59] to be the excursions $(\tilde{e}_i, \tilde{h}_i)$ conditioned on their duration being exactly x.

Remark 5.3.2. To clear up any confusion between $\tilde{\mathbf{e}}^{(\delta)}$ defined in (5.14) and $\mathbf{e}^{(k)}$ defined in (5.4), we note that we use the tilde $\tilde{}$ to denote tilting associated with an exponential tilting of an excursion. We do not include a tilde when discussing the polynomial tilting in (5.4).

5.3.2 Lamperti Transform

The Lamperti transform relates continuous state branching processes and Lévy processes via a time-change. This relationship dates back to a path-by-path relationship observed by Lamperti [128], although only proved later by Silverstein [163]. More recently the authors of [49] gave a path-by-path transformation between certain pairs of Lévy processes and continuous state branching processes with immigration. The bijective relationship was known before the path-by-path connection as well, see [114]. For more information on this transformation see [48] for a description in the continuum, see [49, 50] for scaling limits related to continuous state branching processes and their generalizations affine processes, and see [18] for a scaling limits involving a similar situation of non-uniqueness of the limiting equation.

We will focus on the transform applied to excursions. Given a non-decreasing function $c : \mathbb{R}_+ \to \mathbb{R}$ denote its right-hand derivative by D_+c, i.e.

$$D_+c(t) = \lim_{\varepsilon \downarrow 0} \frac{c(t+\varepsilon) - c(t)}{\varepsilon}.$$

We now define the Lamperti transform and the Lamperti pair.

Definition 1. Given a càdlàg function $f \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$ let

$$\iota(t) = \int_0^t \frac{1}{f(s)}\, ds.$$

Define the right-continuous inverse of ι, denoted by c^0, by

$$c^0(t) = \inf\{s \geq 0 : \iota(s) > t\},$$

with the convention $\inf \emptyset = \inf\{t > 0 : f(t) = 0\}$. The *Lamperti transform* of f is the function $h^0 = f \circ c^0$ and we call the pair (h^0, c^0) the *Lamperti pair* associated to f.

Hopefully the choice of notating the Lamperti pair by (h^0, c^0) will be clear after the statement of the next proposition, which we recall from [18] while introducing a trivial scaling argument and fixing a typo:

Proposition 5.3.1. *([18, Proposition 2]) Let $f \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$ with non-negative jumps. Assume that $f(t) = 0$ if and only if $t \in \{0\} \cup [\zeta, \infty)$ for some $\zeta \in (0, \infty)$. Let (h^0, c^0) denote* *the Lamperti pair associated to f. Then, solutions to*

$$c(0) = 0, \qquad D_+ c = f \circ c \tag{5.15}$$

can be characterized as follows:

1. *If* $\displaystyle\int_{0+} \frac{1}{f(s)}\, ds = \infty$ *then* $h = c = 0$ *is the unique solution to* (5.15).

2. *If* $\displaystyle\int_{0+} \frac{1}{f(s)}\, ds < \infty$ *then* c^0 *is not identically zero,* $D_+ c^0 = h^0$*, and* c^0 *solves* (5.15). *Furthermore, solutions to* (5.15) *are a one-parameter family* $(c^\lambda; \lambda \in [0, \infty])$ *given by*

   $$c^\lambda(t) = c^0\left((t-\lambda)_+\right), \qquad (x)_+ := x \vee 0.$$

   In addition,

   (a) *If* $\displaystyle\int^{\zeta-} \frac{1}{f(s)}\, ds = \infty$ *then* c^0 *is strictly increasing with* $\lim_{t\to\infty} c^0(t) = \zeta$.

   (b) *If* $\displaystyle\int^{\zeta-} \frac{1}{f(s)}\, ds < \infty$ *then* c^0 *is strictly increasing until reaching* $\iota(1) = \displaystyle\int_0^{\zeta} \frac{1}{f(s)}\, ds$.

The above proposition states that all the solutions to (5.15) are determined by time-shifts of the Lamperti pair associated with f, or is identically zero. As we will see in the sequel, a major part of the proof of Theorem 5.2.1 is showing that every subsequential (weak) limit of the $\tilde{C}_n$ is of the form C^0 and not a time-shift, C^Λ, of C^0 for some random Λ.

With Proposition [18] recalled, we can prove Corollary 5.1.3 from Theorem 5.1.2.

Proof of Corollary 5.1.3. We begin by observing that

$$\sup_{v\in \operatorname{spt}(\mu)} \mathrm{d}(\rho, v) \overset{d}{=} \int_0^1 \frac{1}{\mathbf{e}^{(k)}(s)}\, ds$$

follows from Theorem 5.1.2 by an application of conclusion (2)(b) in Proposition 5.3.1. To replace the support of the measure μ with the graph $\mathscr{G}^{(\alpha,k)}$ we observe that

$$\sup_{v\in \operatorname{spt}(\mu)} \mathrm{d}(\rho, v) = \sup_{v\in \mathscr{G}^{(\alpha,k)}} \mathrm{d}(\rho, v).$$

Indeed to prove this equality observe that the leafs of the graph $\mathscr{G}^{(\alpha,k)}$ are dense in both the support of the measure μ and the graph $\mathscr{G}^{(\alpha,k)}$ which follows from analogous results for the continuum random trees [70, 74] and the observation that the exponential tilting in the construction of the graphs does not change this almost sure statement.

Part (2) trivially follows from Theorem 5.1.2 and the observation that $\mathbf{c}$ increases from 0 to 1 as t ranges from 0 to ∞. We restrict the rest our proof to part (3), the argument of which will imply part (4) with minor modifications.

We recall the well-known fact that if X is a real random variable taking values in (a, b) with cumulative distribution function F which is strictly increasing on (a, b), then $X \sim F^{-1}(U)$ where U is a standard uniform random variable and $F^{-1}(y) = \inf\{t : F(t) > y\}$ is the right-continuous inverse. Typically this is stated with the left-continuos inverse of F; however, when F is strictly increasing these two inverses agree on (a, b).

Now, conditionally given $\mathscr{G}^{(\alpha,k)}$, Theorem 5.1.2 implies that

$$\mathbb{P}(\mathrm{d}(\rho, V) \le t|\mathscr{G}^{(\alpha,k)}) = \mu(B(\rho, t)) = \mathbf{c}(t).$$

Thus,

$$\mathrm{d}(\rho, V) \stackrel{d}{=} \inf\{t : \mathbf{c}(t) > U\}, \qquad U \sim \mathrm{Unif}(0, 1).$$

However, the process $\mathbf{c}$ is equal in distribution to $c(t) = \int_0^t z(s)\, ds$ where z is as in part (2) of Theorem 5.1.2. It is easy to see by examining the discussion of the Lamperti transform above, that

$$\mathbf{c}(t) = \inf\left\{u : \int_0^u \frac{1}{\mathbf{e}^{(k)}(s)}\, ds > t\right\} \wedge 1.$$

See also the discussion preceding Proposition 2 in [18] and Chapter 6 of [86] as well. The result now follows by taking another inverse.

The proof of part (4) is a trivial generalization involving order statistics. □

5.3.3 Lemmas involving Convergence of Metric Spaces

Recall the definitions of PMM spaces and PMM isometries from Section 2.2.1 discussed above.

We now prove the following simple lemma, which we cannot find in the existing literature. This will be used in the proof of Theorem 5.2.1. We denote by $B(y, r)$, the closed ball of radius r centered at y in the appropriate metric space.

Lemma 5.3.2. *Let* $\mathcal{M}_n = (\mathcal{M}_n, \rho_n, \mathrm{d}_n, \mu_n)$*,* $\mathcal{M} = (\mathcal{M}, \rho, \mathrm{d}, \mu)$ *be random elements of* $\mathfrak{M}$ *such that*

$$\mathcal{M}_n \overset{d}{\Longrightarrow} \mathcal{M}, \qquad \text{as random elements of } (\mathfrak{M}, d_{\mathrm{GHP}}),$$

and μ *on* $\mathcal{M}$ *is almost surely not the zero measure. Then for all but countably many* $r \in (0, \infty)$*:*

$$\mu_n(B(\rho_n, r)) \overset{d}{\Longrightarrow} \mu(B(\rho, r)), \qquad \text{as real numbers.}$$

Both convergences above can be replaced with almost sure convergence as well.

We prove this by first appealing to a deterministic lemma.

Lemma 5.3.3. *Let* $\mathcal{M}_n \to \mathcal{M}$ *in* d_{GHP} *and suppose that the measure* μ *on* $\mathcal{M}$ *is not the zero measure. Let* r *be a radius such that*

$$\mu\{x \in \mathcal{M} : \mathrm{d}(\rho, x) = r\} = 0.$$

Then

$$\mu_n(B(\rho_n, r)) \to \mu(B(\rho, r)).$$

Proof. By Theorem 3.16 in [121], it suffices to consider the compact case where $\mathcal{M}_n, \mathcal{M} \in \mathfrak{M}_c$ and $\mathcal{M}_n \to \mathcal{M}$ with respect to the d^c_{GHP} metric and that $r = \sup_{x\in\mathcal{M}} \mathrm{d}(\rho, x)$.

We recall that, for metric spaces X and Y, a function $f : X \to Y$ is an ε-isometry if f is measurable and

$$\sup\{|d(x_1, x_2) - d(f(x_1), f(x_2))| : x_1, x_2 \in X\} \le \varepsilon$$

and for all $y \in Y$ there exists some $x \in X$ such that $d(y, f(x)) < \varepsilon$.

By Theorem 3.18 in [121], there exists a sequence $\varepsilon_n \to 0$ and a sequence of functions $f^n : \mathcal{M}_n \to \mathcal{M}$ such that f^n is an ε_n-isometry and such that

$$f^n_{\#}\mu_n \to \mu,$$

with respect to the weak-* topology of measures on $\mathcal{M}$, that is convergence of the integrals against compactly supported continuous functions. However, $1_{\mathcal{M}}$ is continuous and compactly supported since $\mathcal{M}$ is compact. So the following convergence holds in because of convergence in the weak-* topology:

$$\begin{aligned}\mu_n(\mathcal{M}_n) &= \int_{\mathcal{M}_n} 1_{\mathcal{M}_n}\, d\mu_n = \int_{\mathcal{M}_n} 1_{\mathcal{M}} \circ f^n(x)\, \mu_n(dx) \\ &= \int_{\mathcal{M}} 1_{\mathcal{M}}(x)(f^n_{\#}\mu_n)(dx) \to \int_{\mathcal{M}} 1_{\mathcal{M}}\, d\mu = \mu(\mathcal{M}).\end{aligned}$$

Therefore, there is no loss in generality in assuming that the measures μ_n and μ are probability measures, since we can just rescale the measures by their (non-zero) total mass. Since weak-* convergence of probability measures on a compact space is simply weak convergence of probability measures, the desired convergence holds by Portmanteau. □

Proof of Lemma 5.3.2. By Lemma 5.3.3, we have shown that the map $\Phi_r : \mathfrak{M} \to \mathbb{R}$ by

$$\Phi_r(\mathcal{M}) = \mu(B(\rho, r))$$

is continuous at each $\mathcal{M}$ such that $\mu(\{x \in \mathcal{M} : \mathrm{d}(\rho, x) = r\}) = 0$.

Now given a random element $\mathcal{M}$ with law $\mathbb{P}$, we just need to show

$$\{r : \mathbb{P}\left[\mu(\{x \in \mathcal{M} : \mathrm{d}(\rho, x) = r\}) > 0\right] > 0\} \quad \text{is countable.}$$

This follows from the same argument that random processes in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ cannot have a uncountably many jump-times which occur with strictly positive probability. The proof of that latter statement can be found in Section 13 of [35], but is omitted here.

□

5.3.4 Continuum random trees and continuum random graphs

We refer the reader back to Section 2.2.3 for the construction of continuum random graphs from a continuous function $h : [0, \zeta] \to \mathbb{R}_+$, a càdlàg function $g : [0, \zeta] \to \mathbb{R}_+$ without negative jumps and a point set $\mathcal{Q} \subset \mathbb{R}^2_+$ having finitely many points within any compact set.

Let us now describe the graphs $\mathscr{G}^{(\alpha,k)}$ that we mentioned in the introduction. See [59, 96] for more information on these graphs. We first define the α-stable graph $\mathscr{G}^{(\alpha)}$

where we let the surplus be a random non-negative integer. The graphs $\mathscr{G}^{(\alpha)}$ are the graphs $\mathcal{G}(\tilde{\mathbf{h}}_1^{(1)}, \tilde{\mathbf{e}}_1^{(1)}, \mathscr{P})$ for a Poisson random measure $\mathscr{P}$ on $\mathbb{R}_+^2$ with Lebesgue intensity. The Poisson point process $\mathscr{P}$ has only a finite number of points (t, y) such that $0 \le y \le \tilde{\mathbf{e}}_1^{(1)}(t)$, and this is the surplus of the random graph $\mathscr{G}^{(\alpha)}$. The graph $\mathscr{G}^{(\alpha,k)}$ is just the graph $\mathscr{G}^{(\alpha)}$ conditioned on having fixed surplus k. This conditioning on the number of points of $\mathscr{P}$ which lie under the curve $\tilde{\mathbf{e}}_1^{(1)}$ changes the exponential tilting in (5.14) to the polynomial tilting in (5.4).

For more information on random trees and graphs see [8, 9, 11, 131] for the Brownian CRT, see [70, 73, 146] for the stable-trees, see [5, 28] for the Brownian random graph and [59, 96] for the stable graph.

5.4 *Proofs of Weak Convergence Results*

We now turn our attention to proving the abstract weak convergence results: Theorems 5.2.1 and 5.2.2. To simplify the notation in the proof of Theorem 5.2.1, we write $X_n(\cdot) = Z_n(0) + X_n^{\mathrm{BF}}(\cdot)$. By assumption (4) in Theorem 5.2.1, $Z_n(0) \to 0$ in probability and so assumption (1) in Theorem 5.2.1 holds with X_n replacing X_n^{BF} by Slutsky's theorem. Moreover, changing (5.6) to match this notation, the process Z_n solves

$$Z_n(h+1) = X_n \circ C_n(h), \qquad C_n(h) = \sum_{j=0}^{h} Z_n(j), \qquad C_n(-1) = 0.$$

We define the rescalings:

$$\begin{aligned} \tilde{Z}_n(t) &= \frac{\alpha_n}{\gamma_n} Z_n(\lfloor \alpha_n t \rfloor) \\ \tilde{C}_n(t) &= \frac{1}{\gamma_n} C_n(\lfloor \alpha_n t \rfloor) \\ \tilde{X}_n(t) &= \frac{\alpha_n}{\gamma_n} X_n(\gamma_n t). \end{aligned}$$

We begin by proving the tightness of $\tilde{X}_n$ and $\tilde{C}_n$.

Proposition 5.4.1. *Under the assumptions of Theorem 5.2.1, and the above notation, the sequence $((\tilde{C}_n, \tilde{X}_n); n \ge 1)$ is tight in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})^2$. Moreover, any subsequential limit of $(\tilde{C}_n, \tilde{X}_n; n \ge 1)$, say (C, X), must satisfy*

$$C(t) = \int_0^t X \circ C(s)\, ds.$$

Proof. We alter the proof of Proposition 7 in [18]. That proof involves a linear interpolation of C_n instead, which makes their proof slightly simpler. The differences are easily overcome using compactness results in Billingsley's monograph [35].

Because tightness of marginals implies tightness of the pair of random elements, in order to show the tightness claimed, it suffices to show that $(\tilde{C}_n; n \geq 1)$ is tight, since we assume that $\tilde{X}_n$ converges weakly and is therefore tight. Towards this end, observe that $\tilde{C}_n$ is uniformly bounded:

$$0 \leq \tilde{C}_n(t) \leq \frac{1}{\gamma_n}\mathfrak{m}_n(G_n) = 1. \tag{5.16}$$

We now set $t > s$. We have

$$\begin{aligned}
\tilde{C}_n(t) - \tilde{C}_n(s) &= \frac{1}{\gamma_n}(C_n(\lfloor \alpha_n t \rfloor) - C_n(\lfloor \alpha_n s \rfloor)) \\
&= \frac{1}{\gamma_n} \sum_{h=\lfloor \alpha_n s \rfloor + 1}^{\lfloor \alpha_n t \rfloor} Z_n(h) \\
&\leq \frac{1}{\gamma_n} \int_{\alpha_n s - 1}^{\alpha_n t + 1} Z_n(\lfloor u \rfloor)\, du \\
&= \frac{1}{\gamma_n} \int_{\alpha_n s - 1}^{\alpha_n t + 1} X_n \circ C_n(\lfloor u \rfloor)\, du \\
&= \frac{\alpha_n}{\gamma_n} \int_{s - \alpha_n^{-1}}^{t + \alpha_n^{-1}} X_n \circ C_n(\lfloor \alpha_n u \rfloor)\, du \\
&\leq (t - s + 2\alpha_n^{-1}) \|\tilde{X}_n\|,
\end{aligned}$$

where

$$\|f(t)\| = \sup_{t \in [0,1]} |f(t)|.$$

Define the functions

$$w(f; I) := \sup_{s,t \in I} |f(t) - f(s)|, \qquad f \in \mathbb{D}(\mathbb{R}_+), I \subset \mathbb{R},$$

and, for $\delta > 0$,

$$w'_N(f; \delta) := \inf_{\{t_i\}} \max_{1 \leq i \leq v} w(f; [t_{i-1}, t_i)), \qquad N = 1, 2, \cdots$$

where the infimum is taken over all partitions $0 = t_0 < t_1 < \cdots < t_v = N$ such that $t_i - t_{i-1} > \delta$ for $1 \leq i < v$.

From the above string of inequalities, for any integer $N > 0$ and any $\delta > 0$, we have

$$w'_N(\tilde{C}_n;\delta) \le 2\left(\delta + \alpha_n^{-1}\right)\|\tilde{X}_n\|, \qquad \forall n \ge 1,$$

Moreover, for any fixed $\delta > 0$, there exists an n_0 sufficiently large such that

$$2(\delta + \alpha_n^{-1})\|\tilde{X}_n\| \le 4\delta\|\tilde{X}_n\|, \qquad \forall n \ge n_0,$$

since $\alpha_n \to \infty$. Fix $\varepsilon > 0$ and an integer N. Applying Theorem 13.2 in [35] gives

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\left(w'_N(\tilde{C}_n;\delta) \ge \varepsilon\right) \le \lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\left(\|\tilde{X}_n\| \ge \frac{\varepsilon}{4\delta}\right) = 0.$$

Hence, by Theorem 16.8 in [35], the process $\tilde{C}_n$ is tight.

The statement about the form of the subsequential weak limits follows as in the proof of Proposition 7 in [18] with little alteration. □

5.4.1 Proofs of Theorems 5.2.1 and 5.2.2

We now move to describe more accurately the possible subsequential limits in Proposition 5.4.1. By Proposition 5.3.1 and Proposition 5.4.1, the subsequential limits $(\tilde{C}_{n_\ell}, \tilde{X}_{n_\ell}) \overset{d}{\Longrightarrow} (C, X)$ must be of the form $C(t) = C^0((t-\Lambda)_+)$ for some (random) $\Lambda := \inf\{t : C(t) > 0\} \in [0,\infty]$ where C^0 is the Lamperti transform of the X. We desire to show that $\Lambda = 0$ almost surely.

By the Skorokhod representation theorem and by possible taking a further subsequence, we can assume that we are working on a probability space such that both

$$\text{scale}(\alpha_{n_\ell}^{-1}, \gamma_{n_\ell}^{-1})G_{n_\ell} \to \mathcal{M}, \qquad \text{and} \qquad (\tilde{C}_{n_\ell}, \tilde{X}_{n_\ell}) \to (C, X)$$

occur almost surely in their respective topologies: the first convergence is with respect to the pointed Gromov-Hausdorff-Prohorov topology and the second convergence is with respect to the product topology on the $\mathbb{D}\times\mathbb{D}$. We write $\tilde{G}_{n_\ell} = (\tilde{G}_{n_\ell}, \rho_{n_\ell}, \tilde{d}, \tilde{\mathfrak{m}}_{n_\ell})$ for $\text{scale}(\alpha_{n_\ell}^{-1}, \gamma_{n_\ell}^{-1})G_n^i$. By Lemma 5.3.2, we have for all but countably many $t > 0$,

$$\begin{aligned}\tilde{C}_{n_\ell}(t) &= \frac{1}{\gamma_{n_\ell}}\mathfrak{m}_{n_\ell}\{v \in G_{n_\ell} : d(v,\rho_{n_\ell}) \le \alpha_{n_\ell} t\} \\ &= \tilde{\mathfrak{m}}_{n_\ell}\left(\{v \in \tilde{G}_{n_\ell} : \tilde{d}(v,\rho_{n_\ell}) \le t\}\right) \longrightarrow \mu(\{x \in \mathcal{M} : d(\rho, x) \le t\}).\end{aligned}$$

Similarly, by the convergence of $\tilde{C}_{n_\ell}$ in $\mathbb{D}$ we have for all but countably many $t > 0$

$$\tilde{C}_{n_\ell}(t) \to C(t).$$

By a standard diagnolization argument there exists a sequence $t_m \downarrow 0$ such that

$$C(t_m) = \mu\left(\bar{B}(\rho, t_m)\right) > 0.$$

The inequality follows from Assumption (3) in Theorem 5.2.1.

Hence

$$\Lambda = \inf\{t : C(t) > 0\} = 0.$$

Therefore every subsequential weak limit for $(\tilde{C}_n, \tilde{X}_n)$ must be of the form (C^0, X) where C^0 is part of the Lamperti pair associated with X. Along with looking at Proposition 5.3.1, we have proved the following:

Proposition 5.4.2. *Let (Z^0, C^0) be the Lamperti pair of X. Then, under the assumptions of Theorem 5.2.1, the following weak convergence holds*

$$\left(\tilde{C}_n, \tilde{X}_n\right) \Longrightarrow \left(C^0, X\right),$$

in the product Skorokhod space $\mathbb{D} \times \mathbb{D}$.

Moreover, since C^0 is not identically zero X must satisfy

$$\int_{0+} \frac{1}{X(s)}\, ds < \infty.$$

We now finish the proof of Theorem 5.2.1.

Proof of Theorem 5.2.1. The proof of Proposition 5.4.2 gives the proof of conclusion (2) of Theorem 5.2.1, and so we finish the proof of part (1).

By Proposition 5.4.1 and Proposition 5.4.2, and the Skorohod representation theorem, we can assume that we are working on a probability space such that

$$\left(\tilde{C}_n, \tilde{X}_n\right) \longrightarrow \left(C^0, X\right) \qquad \text{a.s..}$$

Then, a result of Wu [183, Theorem 1.2] which extends a result of Whitt [180], the following convergence holds in the $\mathbb{D}$:

$$\tilde{Z}_n = \tilde{X}_n \circ \tilde{C}_n \longrightarrow X \circ C^0 \qquad \text{a.s..}$$

Using Proposition 5.3.1 part (2), we observe that

$$Z^0 := X \circ C^0 = D_+ C^0.$$

That is (Z^0, C^0) is the Lamperti pair associated with X and in $\mathbb{D}^2$

$$(\tilde{Z}_n, \tilde{C}_n) \stackrel{d}{\Longrightarrow} (Z^0, C^0).$$

□

The proof of Theorem 5.2.1 can be easily extended to joint convergence of the of finitely many graphs G_n^i of random masses $\mathfrak{m}_n(G_n^i)$ as $n \to \infty$. Since the graph G_n^i are ordered by decreasing mass, the excursion lengths also decrease: $\zeta_1 \geq \zeta_2 \geq \cdots$. The only part that changes is (5.16) is replaced with an analogous tightness bound on

$$\max_{j \leq N} \frac{1}{\gamma_n} \mathfrak{m}_n(G_n^j) = \frac{1}{\gamma_n} \mathfrak{m}_n(G_n^1)$$

This will yield a proof of Theorem 5.2.2. The details are omitted.

5.5 The Configuration Model

In this section we focus on the applications to the configuration model when one specifies a critical degree distribution ν in the domain of attraction of a stable law. We will focus on the case $\alpha \in (1, 2)$, although the Brownian case $\alpha = 2$ can be obtained by these methods. The results can easily be altered to cover the $\alpha = 2$ as well, by instead considering the case where ν has finite third moment at the critical point (see the definition of θ in (5.19) below) and omitting the cases $\nu(2) < 1$ and $\nu(0) > 0$.

We will be using the results of Joseph [109] and Conchon-Kerjan and Goldschmidt [59] on scaling limits related to the configuration model. The latter reference provides a metric space scaling limit for the components of the graph at the point of criticality $\theta = 1$, where θ is defined in (5.19). This allows us to utilize Theorem 5.2.2. Similar results in the $\alpha = 2$ case were obtained prior to Joseph, see Riordan's work [157] and also [29].

5.5.1 *Preliminaries: The configuration model and convergence*

Let us describe briefly the configuration model, some of the associated walks on the graphs, and their scaling limits. For a more detailed account of the configuration model, see Chapter 7 of [173].

The multigraph $M(\mathbf{d}^n)$ is a random graph on vertex set $(i; i \in [n])$ where the vertex i has degree (counted with multiplicity) d_i. We can construct this graph by viewing the vertices i as hubs with d_i *half-edges* jutting out from the vertex i. We then pair half-edges uniformly at random to create a multigraph. Given a multigraph G, we have [173, Prop 7.7]

$$\mathbb{P}(M(\mathbf{d}^n) = G) = \frac{1}{\left(-1 + \sum_{j=1}^n d_j\right)!!} \times \frac{\prod_{j=1}^n d_j!}{\prod_{j=1}^n \operatorname{loop}(i) \times \prod_{1\le i<j\le n} \operatorname{mult}(i,j)!},$$

where $\operatorname{loop}(i)$ is the number of self-loops at vertex i and $\operatorname{mult}(i,j)$ is the number of edges between i and j. Below we describe two different algorithms for how to construct the multigraph and describe associated walks. It is also convenient to assume that the d_i half-edges connected to a vertex i are ordered, so that we can talk about the "least" half-edge. We remark that this random construction described above is taken from a deterministic sequence of half-edges $\mathbf{d}^n$, later on we will take the vertex degrees to be random.

We describe two algorithms for the construction in a manner quite similar to Joseph [109]. We partition the $\sum_{j=1}^n d_j$ half-edges into three disjoint subsets: the set $\mathcal{S}$ of sleeping half-edges, the set $\mathcal{A}$ of active half-edges, and the set $\mathcal{D}$ of dead half-edges. We call the set $\mathcal{S} \cup \mathcal{A}$ the collection of alive half-edges. Initially all half-edges are sleeping.

5.5.1.1 *Breadth-first construction*

We construct a graph $M^{\mathrm{BF}}(\mathbf{d}^n)$ (we initially include a BF to specify the construction) as follows:

To initialize at step 1, we pick a sleeping half-edge uniformly at random. Label the corresponding vertex as v_1 and declare all of the half-edges attached to v_1 as active.

Suppose that we have just finished step j. There are three possibilities: (1) $\mathcal{A} \neq \emptyset$, (2) $\mathcal{A} = \emptyset$ and $\mathcal{S} \neq \emptyset$, (3) all half-edges are dead.

In case 1, we proceed as follows:

1. Let i be the **smallest** integer k such that there exists an active half-edge attached to v_k.

2. Pick the least half-edge l from all active half-edges attached to v_i.

3. Kill l, that is, remove it from $\mathcal{A}$ and add it to $\mathcal{D}$.

4. Choose uniformly at random from all living half-edges r and pair it with l, that is, add an edge between the vertex v_i (which is attached to l) and the corresponding vertex connected to r.

5. If r is sleeping, then we have discovered a new vertex. Label this new vertex v_{m+1} where we have discovered the vertices $v_1, \cdots, v_m$ up to this point. Declare all the half-edges of v_{m+1} are active.

6. Kill r.

In case 2, we have finished exploring a connected component of $M^{\mathrm{BF}}(\mathbf{d}^n)$. We proceed by picking a sleeping half-edge uniformly at random. We then label the corresponding vertex v_{m+1} if we've discovered vertices $v_1, \cdots, v_m$ up to this point, and we declare all the half-edges connected to v_{m+1} as active.

In case 3, we have explored the entire graph and we are done.

The above is the breadth-first construction of the multigraph $M^{\mathrm{BF}}(\mathbf{d}^n)$. In the sequel, we denote the ordering of the vertices in the exploration/construction above by $v_1^{\mathrm{BF}}, \cdots, v_n^{\mathrm{BF}}$.

Remark 5.5.1. While the above algorithm gives a breadth-first construction of the graph $M(\mathbf{d}^n)$, observe that this can also be used to explore the graph $M(\mathbf{d}^n)$. Indeed, if in step (4), we selected the half-edge r which is connected to l instead of sampling it uniformly, then we would have explored the graph and obtain the an equal in distribution ordering of the vertices.

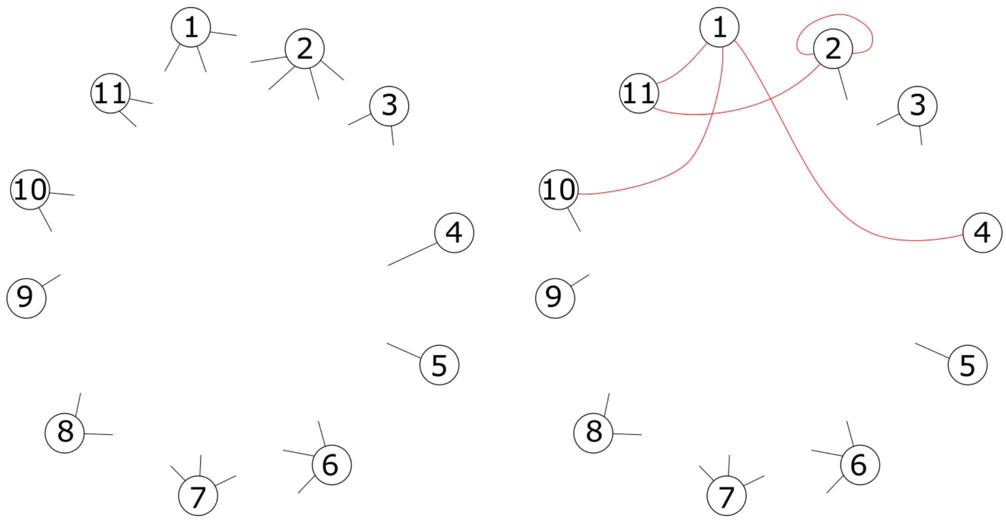


Figure 5.3: Left: The initial collection of 11 vertices with half-edges appearing from the center. Right: The structure of the breadth-first constructed graph after initially selected a half-edge connected to vertex 11. The edges were added in this order $\{11, 1\}, \{11, 2\}, \{1, 4\}, \{1, 10\}, \{2, 2\}$. The next half-edge to be explored is the remaining half-edge jutting out form vertex 2.

In case 1 above, it is possible that we match two half-edges l with r where r is already active. We call the corresponding edge in the multigraph $M^{\mathrm{BF}}(\mathbf{d}^n)$ a *breadth-first (bf) backedge*.

We let $F^{\mathrm{BF}}(\mathbf{d}^n)$ denote the forest constructed from the multigraph $M^{\mathrm{BF}}(\mathbf{d}^n)$ obtained by splitting all bf backedges into two half-edges and adding two leaves to each of these half-edges. More formally, if the multigraph $M^{\mathrm{BF}}(\mathbf{d}^n)$ has a bf backedge between vertices v_l, v_r. Remove that edge from the multigraph and add two vertices v_l' and v_r' and add an edge between both pairs (v_l, v_l') and (v_r, v_r'). Continue this until all bf backedges are removed and replaced.

Remark 5.5.2. This algorithm can also be used to mark where the new leafs occur within a breadth-first exploration of the forest $F^{\mathrm{BF}}(\mathbf{d}^n)$. When we first find backedge between half-edges l and r in $M^{\mathrm{BF}}(\mathbf{d}^n)$ we replace it with two new leafs. Then as we are exploring the half-edge l, we find a new leaf and do not "see" the half-edge r. This means we do not kill that half-edge in step (6). This means we will eventually choose half-edge r in step (2) and find second new leaf for this bf backedge. We can then label these vertices $u_1^{\mathrm{BF}}, \cdots, u_p^{\mathrm{BF}}$ for some $p \geq n$. See Figure 5.4 for an example of how this is done for the breadth-first construction.

5.5.1.2 Depth-first construction

We construct a graph $M^{\mathrm{DF}}(\mathbf{d}^n)$ (we initially include a DF to specify the construction) as follows:

We initialize at step 1 as before, we pick a sleeping half-edge uniformly at random. Label the corresponding vertex as v_1 and declare all of the half-edges attached to v_1 as active. The only thing that changes on subsequent steps is that in Case 1 we replace part (1) with

(1') Let i be the **largest** integer k such that there exists an active half-edge attached to v_k.

This above is the depth-first construction of the multigraph $M^{\mathrm{DF}}(\mathbf{d}^n)$. This changes the order in which we find vertices and label them, so we will denote the new ordering and labeling by $v_1^{\mathrm{DF}}, \cdots, v_n^{\mathrm{DF}}$. We analogously construct the depth-first forest, by removing depth-first (df) backedges and replacing them with leaves.

5.5.1.3 Symmetry between constructions

Recall that the multigraph $M(\mathbf{d}^n)$ is taken to be uniform over all possible pairings of half-edges. The following proposition is trivial.

Proposition 5.5.1. *For any degree sequence $\mathbf{d}^n = (d_1, \cdots, d_n)$ the graphs $M^{\mathrm{BF}}(\mathbf{d}^n)$ and $M^{\mathrm{DF}}(\mathbf{d}^n)$ are equal in distribution. We write both as $M(\mathbf{d}^n)$.*

The symmetry between the constructions allow us to look at two random walks which turn out being equal in distribution. We write $\deg(v)$ for the degree (counted with multiplicity) of the vertex v in a graph, $M(\mathbf{d}^n), F^{\mathrm{DF}}(\mathbf{d}^n)$, etc., which is clear from context. Recall that $(v_j^{\mathrm{BF}}; j \in [n])$ and $(v_j^{\mathrm{DF}}; j \in [n])$ are the vertices in the multigraph $M(\mathbf{d}^n)$ labeled in two distinct ways acording the breadth-first exploration or depth-first explorations respectively. Define the following two walks:

$$S^{\mathrm{BF}}_{\mathbf{d}^n}(k) = \sum_{j=1}^{k} (\deg(v_j^{\mathrm{BF}}) - 2), \qquad S^{\mathrm{DF}}_{\mathbf{d}^n}(k) = \sum_{j=1}^{k} (\deg(v_j^{\mathrm{DF}}) - 2). \tag{5.17}$$

We wish to define analogous walks on the forests $F^{\mathrm{BF}}(\mathbf{d}^n)$ and $F^{\mathrm{DF}}(\mathbf{d}^n)$, to which we remind the reader of Remark 5.5.2. Hence we have a labeling all the vertices of the forests as (u_j^{BF}) for the forest $F^{\mathrm{BF}}(\mathbf{d}^n)$ with the breadth-first exploration and (u_j^{DF}) for the forest $F^{\mathrm{DF}}(\mathbf{d}^n)$ with the depth-first exploration.

For each connected component of the multigraph and hence forest, there is a vertex discovered when the collection of active vertices $\mathcal{A}$ was empty, we call those vertices *roots*. If u is a root in a forest, write $\chi(u) = \deg(u)$, otherwise write $\chi(u) = \deg(u) - 1$. The value of $\chi(u)$ is precisely the number of children that vertex u has in the forest in which it lives. For j sufficiently large, there will be no vertex of u_j^{BF} or u_j^{DF}. This will not matter for our scaling limits, but for completeness we define u_j^{BF} and u_j^{DF} as root vertices of components with a single vertex, and therefore $\chi(u_j^{\mathrm{BF}}) = \chi(u_j^{\mathrm{DF}}) = 0$ for sufficiently large j. Define the walks

$$X^{\mathrm{BF}}_{\mathbf{d}^n}(k) = \sum_{j=1}^{k} (\chi(u_j^{\mathrm{BF}}) - 1), \qquad X^{\mathrm{DF}}_{\mathbf{d}^n}(k) = \sum_{j=1}^{k} (\chi(u_j^{\mathrm{DF}}) - 1). \tag{5.18}$$

As is shown in Section 5.1 of [59], the distribution of the depth first walk $X^{\mathrm{DF}}_{\mathbf{d}^n}$ can be reconstructed from $S^{\mathrm{DF}}_{\mathbf{d}^n}$. This was done for when the degree sequence is taken to be random; however it works for a deterministic degree sequence as well. A trivial alteration of that algorithm can be used to construct $X^{\mathrm{BF}}_{\mathbf{d}^n}$ from the walk $S^{\mathrm{BF}}_{\mathbf{d}^n}$ and moreover, we can couple these two constructions to see that backedges have a particular correspondence. We summarize this construction later in the appendix. We write this as the following lemma.

Lemma 5.5.2. *1. For any degree sequence $\mathbf{d}^n$. The breadth-first walks and the depth-first walks are equal in distribution. That is*

$$(S^{\mathrm{BF}}_{\mathbf{d}^n}(k); k \ge 0) \stackrel{d}{=} (S^{\mathrm{DF}}_{\mathbf{d}^n}(k); k \ge 0), \qquad \textit{and} \qquad (X^{\mathrm{BF}}_{\mathbf{d}^n}(k); k \ge 0) \stackrel{d}{=} (X^{\mathrm{DF}}_{\mathbf{d}^n}(k); k \ge 0).$$

2. There exists a coupling of $F^{\mathrm{DF}}(\mathbf{d}^n)$ and $F^{\mathrm{BF}}(\mathbf{d}^n)$ such that for all $j < i$ a df-backedge *appears between* u_j^{DF} *and* u_i^{DF} *if and only if a bf-backedge appears between* u_j^{BF} *and* u_i^{BF}*.*

In particular, there exists a coupling of X_n^{BF} *and* X_n^{DF} *such that*

$$\inf_{i\le k} X_n^{\mathrm{BF}}(i) = \inf_{i\le k} X_n^{\mathrm{DF}}(i).$$

The last part of the above lemma tells us something that will be used several times in the sequel, under this coupling the distribution excursions of $X_{\mathbf{d}^n}^{\mathrm{BF}}$ and $X_{\mathbf{d}^n}^{\mathrm{DF}}$ above the running infimum have the same length.

5.5.1.4 Random degree distribution

In this subsection we describe some of what happens when we take the sequence $\mathbf{d}^n = (d_1, \cdots, d_n)$ to be from i.i.d. samples from a distribution ν on $\mathbb{N} = \{1, 2, \cdots\}$. We take $(d_j; j \ge 1)$ to be an i.i.d. sequence with common distribution ν. In order to guarantee that a multigraph with degree sequence $\mathbf{d}^n = (d_1, \cdots, d_n)$ exists, we replace d_n with $d_n + 1$ if the sum has the wrong parity.

Recall from the introduction that we write $M_n(\nu)$ for the random graphs with a random degree distribution and forest. We do the same when referencing the forests, i.e. we will write $F_n^{\mathrm{BF}}(\nu)$ instead of $F^{\mathrm{BF}}(\mathbf{d}^n)$ when the degree sequence is random. We do not emphasize this dependence on ν when describing the random walks, instead we replace subscript $\mathbf{d}^n$ with just n, i.e. we write X_n^{DF} instead of $X_{\mathbf{d}^n}^{\mathrm{DF}}$. If ν has finite variance, then it can be shown (see [173, Section 7.6]) there is positive probability that the multigraph is simple, i.e. contains no self-loops nor multiple edges, and it has an explicit asymptotic formula

$$\lim_{n\to\infty} \mathbb{P}(M_n(\nu) \text{ is simple}) = \exp\left(-\frac{\theta}{2} - \frac{\theta^2}{4}\right), \qquad \text{where} \quad \theta = \theta(\nu) = \frac{\mathbb{E}[d_1(d_1-1)]}{\mathbb{E}[d_1]}. \tag{5.19}$$

Moreover, when $M_n(\nu)$ is simple, it is uniformly distributed over all simple graphs with degree distribution $\mathbf{d}^n$. We write $G_n(\nu)$ for the graph $M_n(\nu)$ conditioned on it being simple.

Under certain conditions on ν, the value of θ also tells us something about the critical behavior of the graph which dates back to Molloy and Reed's work [148, 149]. See also [107]. That is if V_n denotes the size of the largest component of the multigraph $M_n(\nu)$ there is a phase-transition which occurs:

1. If $\theta(\nu) > 1$ then $V_n/n \xrightarrow{p} a(\nu)$ for a deterministic constant $a(\nu) > 0$.

2. If $\theta(\nu) \le 1$ then $V_n/n \xrightarrow{p} 0$.

We will now restrict our attention to the case where ν is as in (5.1) and observe that the right-most assumption in (5.1) is $\mathbb{E}[d_1^2] = 2\mathbb{E}[d_1]$, which is equivalent to $\theta = 1$.

In this setting Joseph [109] gives a scaling limit of a depth-first walk for the multigraph $M_n(\nu)$, which is very slightly different than what we wrote as S_n^{DF}. That work was extended by Conchon-Kerjan and Goldschmidt in [59]. We now recall the scaling limit in the latter reference. Let $\tilde{X} = (\tilde{X}(t); t \ge 0)$ and $\tilde{H} = (\tilde{H}(t); t \ge 0)$ be defined by the change of measure in (5.11). We write X for a Lévy process with Laplace exponent (5.10) and H is its associated height process. We define the process $J_n = (J_n(k); k \ge 0)$

$$J_n(k) = \#\{j \in \{0, 1, \cdots, k-1\} : S_n^{\mathrm{DF}}(j) = \inf_{j \le \ell \le k} S_n(\ell)\}, \tag{5.20}$$

which is a discretization of (5.8).

Theorem 5.5.3 (Joseph [109], Conchon-Kerjan - Goldschmidt [59])**.** *Fix some $\alpha \in (1, 2)$. Let ν be a distribution satisfying (5.1) and write $A = \frac{c\Gamma(2-\alpha)}{\delta\alpha(\alpha-1)}$. Using the notation above, the following joint convergence holds in $\mathbb{D}^2$:*

$$\left(n^{-\frac{1}{\alpha+1}} S_n(\lfloor n^{\frac{\alpha}{\alpha+1}} t\rfloor), n^{-\frac{\alpha-1}{\alpha+1}} J_n(\lfloor n^{\frac{\alpha}{\alpha+1}} t\rfloor); t \ge 0\right) \overset{d}{\Longrightarrow} \left(\tilde{X}(t), \tilde{H}(t); t \ge 0\right).$$

A similar result is obtained in [32] under a finite third moment condition on the measure ν where the limiting process is rescaling of Brownian motion with a different parabolic drift.

Crucially for their descriptions of the limiting graphs, the authors of [59] also develop the excursion theory for the process $\tilde{X}$ (which is notated as $\tilde{L}$ in that work). Proposition 3.9 in [59] shows that the excursions of $\tilde{X}, \tilde{H}$, conditioned on their length being exactly x are distributed as $(\tilde{\mathbf{e}}_x^{(\delta)}, \tilde{\mathbf{h}}_x^{(\delta)})$ defined in (5.14).

5.5.1.5 Continuum Graph Limits

We now heuristically describe how the authors of [59] obtain their metric space scaling limit. Let H_n be the height process on the forest $F_n^{\mathrm{DF}}(\nu)$. That is $H_n(k)$ is the distance in $F_n^{\mathrm{DF}}(\nu)$

from vertex u_k^{DF} to the root in its connected component. This process H_n satisfies [131]:

$$H_n(k) = \#\{j \in \{0, \cdots, k-1\} : X_n^{\mathrm{DF}}(j) = \inf_{j \le \ell \le k} X_n^{\mathrm{DF}}(\ell)\}.$$

To examine the components of the graph $M_n(\nu)$, the authors of [59] look at the collection of excursions of the processes X_n^{DF} and H_n. These are defined as follows:

$$\begin{aligned} \widehat{X}_{n,i}^{\mathrm{DF}}(k) &= X_n^{\mathrm{DF}}(\sigma_n(i-1)+k) - X_n^{\mathrm{DF}}(\sigma_n(i-1)), \\ \widehat{H}_{n,i}(k) &= H_n(\sigma(i-1)+k), \end{aligned} \qquad k = 0, \cdots, \sigma_n(i) - \sigma_n(i-1)$$

where $\sigma_n(i) = \inf\{j : X_n^{\mathrm{DF}}(j) = -i\}$ is the first hitting time of level $-i$. The process $\widehat{X}_{n,i}^{\mathrm{DF}}$ starts at zero and is non-negative until it hits level -1 at time $k = \sigma_n(i) - \sigma_n(i-1)$. The process $\widehat{H}_{n,i}$ is strictly positive for $k = 1, \ldots, \sigma_n(i) - \sigma_n(i-1) - 1$. We extend both of these by constancy for $k > \sigma_n(i) - \sigma_n(i-1)$. These processes encode the tree structure [131] of the i^{th} connected component of the forest $F_n^{\mathrm{DF}}(\nu)$. By the construction of $M_n^{\mathrm{DF}}(\nu)$, this orders the components of the forest $F_n^{\mathrm{DF}}(\nu)$ in a manner size-biased by the number of edges in the component. There are further only a finite number of indexes i such that $\sigma_n(i) - \sigma_n(i-1) \neq 1$ since for sufficiently large i the i^{th} component of $F_n^{\mathrm{DF}}(\nu)$ is simply an isolated vertex.

To study the large components of the graph, we instead reorder the excursions by decreasing lengths with ties broken arbitrarily. Denote this new ordering by omitting the "widehat" notation: $(X_{n,i}^{\mathrm{DF}}; i \ge 1) = \Big((X_{n,i}^{\mathrm{DF}}(k); k \ge 0); i \ge 1\Big)$ and $(H_{n,i}; i \ge 1) = ((H_{n,i}(k); k \ge 0); i \ge 1)$.

The i^{th} excursion $X_{n,i}^{\mathrm{DF}}$ may not tell us information about the i^{th} largest component of $M_n(\nu)$, G_n^i. This is because the forest $F_n^{\mathrm{DF}}(\nu)$ contains additional vertices which could change the ordering of the components. I.e. if G_n^i had 10 vertices and 0 df backedges and component G_n^{i+1} had 9 vertices and 2 df backedges then the corresponding components in $F_n^{\mathrm{DF}}(\nu)$ will have 10 and 13 vertices respectively. In turn, their indices will appear in the opposite order. We also note that the excursions of the process X_n^{DF} do not identically correspond to the excursions of the process S_n^{DF} discussed previously. While this may cause some problems in the discrete, in the large n limit neither of these problems are relevant as we will shortly explain.

Before turning to the scaling limits, let T_n^1 be the largest connected component, i.e. tree, contained in $F_n^{\mathrm{DF}}(\nu)$. This is encoded by $X_{n,1}^{\mathrm{DF}}$ and $H_{n,1}$. This tree may contain df backedges and these backedges appear in pairs. For concreteness, suppose that there are $m \geq 1$ of these pairs. These can be indexed by $(l_1, r_1), \cdots, (l_m, r_m)$. This means that the l_i^{th} vertex explored in the depth-first exploration of the largest component of T_n^1 will be paired with the r_i^{th} vertex explored in the corresponding component of $M_n(\nu)$. See Figure 5.5 for the analogous pairs for the breadth-first labeling of the component in Figure 5.4. We now define $\mathscr{P}_{n,1}^{\mathrm{DF}}$ as the collection of points

$$\mathscr{P}_{n,1}^{\mathrm{DF}} = \left\{ \left(n^{-\frac{\alpha}{\alpha+1}} l_i, n^{-\frac{\alpha}{\alpha+1}} r_i \right) : i = 1, \cdots, m \right\}.$$

When there are no df backedges present, just define the set as the empty set. We call this set the set of *marks*, and we can do the same thing to each of the other connected components as well to get sets $\mathscr{P}_{n,i}^{\mathrm{DF}}$.

An important step in how the authors of [59] proved the components $G_n^1, G_n^2, \cdots$ have scaling limits was showing scaling limits of the processes $X_{n,i}^{\mathrm{DF}}$ and $H_{n,i}$ and of the set of marks $\mathscr{P}_{n,i}^{\mathrm{DF}}$. The convergence of the sets $\mathscr{P}_{n,i}^{\mathrm{DF}}$ is with respect to the vague topology of its associated counting measure. Namely they prove [59, Proposition 5.16]

$$\begin{aligned} \left(\left(n^{-\frac{1}{\alpha+1}} X_{n,i}^{\mathrm{DF}}(\lfloor n^{\frac{\alpha}{\alpha+1}} t \rfloor); t \geq 0 \right), \left(n^{-\frac{\alpha-1}{\alpha+1}} H_{n,i}(\lfloor n^{\frac{\alpha}{\alpha+1}} t \rfloor); t \geq 0 \right), \mathscr{P}_{n,i}^{\mathrm{DF}}; i \geq 1 \right) \\ \overset{d}{\Longrightarrow} \left(\tilde{e}_i, \tilde{h}_i, \mathscr{P}_i; i \geq 1 \right), \end{aligned} \tag{5.21}$$

for some discrete sets $(\mathscr{P}_i; i \geq 1)$. Here the convergence in the first two coordinates is with respect to the Skorohod topology and the convergence in the third coordinate is with respect to the vague topology of its associated counting measure and then the product topology is taken over the index $i \geq 1$.

The limiting set $\mathscr{P}_i$ can be described as follows. Let $(\mathcal{Q}_i; i \geq 1)$ denote an i.i.d. collection of Poisson point processes on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $\frac{1}{\delta}$ Leb. Only finitely many of these points $(s, y) \in \mathcal{Q}_i$ will satisfy $0 \leq y \leq \tilde{e}_i(s)$, and index these as $(s^1, y^1), \cdots, (s^m, y^m)$ for some m. The set $\mathscr{P}_i$ is then the collection

$$\mathscr{P}_i = \{(s^p, t^p) : p = 1, \cdots, m\} \qquad t^p = \inf\{u \geq s^p : \tilde{e}_i(u) \leq y^p\}.$$

This, in turn, allowed them to show that the ordered sequence of components of $M_n(\nu)$ and $G_n(\nu)$ by conditioning converge after proper rescaling in a product Gromov-Hausdorff-Prohorov topology to the sequence of continuum random graphs

$$(\mathcal{M}_i; i \geq 1) = \left(\mathcal{G}(\tilde{h}_i, \tilde{e}_i, \mathcal{Q}_i); i \geq 1\right), \tag{5.22}$$

where $\mathcal{Q}_i$ were defined above.

Let us summarize these results in a theorem for easy reference.

Theorem 5.5.4. *(Conchon-Kergan - Goldschmidt [59]) Let $(G_n^i; i \geq 1)$ denote the components of the critical random graph $M_n(\nu)$ ordered by decreasing number of vertices and viewed as pointed measured metric spaces.*

Under the assumptions of Theorem 5.5.3, the convergence in (5.21) holds in the product topology (product over the index k). Jointly with (5.21), the weak convergence

$$\left(\mathrm{scale}(n^{-\frac{\alpha-1}{\alpha+1}}, n^{-\frac{\alpha}{\alpha+1}})G_n^i; i \geq 1\right) \stackrel{d}{\Longrightarrow} (\mathcal{M}_i; i \geq 1) \tag{5.23}$$

holds with respect to the product Gromov-Hausdorff-Prohorov topology, where the sequence $(\mathcal{M}_i; i \geq 1)$ is distributed as (5.22).

Remark 5.5.3. Theorem 1.1 in [59] does not state the joint convergence between equations (5.21) and (5.23); however, the proof of said theorem shows there is joint convergence.

These recalled results from [59] are on the convergence for *depth-first* objects. However, symmetry between depth-first and breadth-first constructions described above in Lemma 5.5.2 allow us to have similar results for the analogous *breadth-first* object. Consequently, if we let $X_{n,i}^{\mathrm{BF}}$ be the breadth-first walk of the i^{th} largest component of $F_n^{\mathrm{BF}}(\nu)$, or, equivalently stated, it is the i^{th} longest excursion of X_n^{BF} above its running minimum, then

$$\left(\left(n^{-\frac{1}{\alpha+1}} X_{n,i}^{\mathrm{BF}}(\lfloor n^{\frac{\alpha}{\alpha+1}} t \rfloor); t \geq 0\right); i \geq 1\right) \stackrel{d}{\Longrightarrow} (\tilde{e}_i^*; i \geq 1), \tag{5.24}$$

where $(\tilde{e}_i^*; i \geq 1) \stackrel{d}{=} (\tilde{e}_i; i \geq 1)$. More importantly for our work, there are auxiliary processes described in [59, pg. 30, 32] and recalled in the appendix that can easily be defined in the same way for a breadth-first construction. In particular, these auxiliary processes include the collection of marks $\mathscr{P}_{n,k}^{\mathrm{DF}}$ recalled above. Therefore, we can extend the convergence

(5.24) by using Lemma 5.5.2 to include collection of marks in an analogous way to the depth-first marks:

Lemma 5.5.5. *There exists a finite set $\mathscr{P}_{n,i}^{\mathrm{BF}} \subset \mathbb{R}_+^2$ of marks corresponding to the i^{th} largest component of $F_n^{\mathrm{BF}}(\nu)$ which keep track of the bf backedges such that*

$$\left(X_{n,i}^{\mathrm{BF}}, \mathscr{P}_{n,i}^{\mathrm{BF}}; i \geq 1\right) \stackrel{d}{=} \left(X_{n,i}^{\mathrm{DF}}, \mathscr{P}_{n,i}^{\mathrm{DF}}; i \geq 1\right).$$

We now state the following lemma:

Lemma 5.5.6. *Couple X_n^{BF} and X_n^{DF} as in Lemma 5.5.2 so they have the same excursion intervals. Then, under the assumptions of Theorem 5.5.4, any joint subsequential limit of (5.21), (5.23) and (5.24) satisfies for each fixed $i \geq 1$:*

$$\zeta(\tilde{e}_i^*) = \zeta(\tilde{e}_i) = \mu_i(\mathcal{M}_i), \tag{5.25}$$

almost surely.

Proof. Under this conditioning, the excursion intervals of X_n^{DF} and X_n^{BF} have the same length. The equality $\zeta(\tilde{e}_i) = \mu_i(\mathcal{M}_i)$ follows from [59], and that implies $\zeta(\tilde{e}_i^*) = \mu_i(\mathcal{M}_i)$ as well. See also [109]. □

We have now gathered most of the required ingredients and background to prove Theorems 5.1.1 and 5.1.2 using the approach in Theorem 5.2.2. The last thing we'll verify is that Assumption 3 holds in Theorems 5.2.1 and 5.2.2. By scaling of the α stable graph [59], we focus on the case that the total mass equals 1.

Proposition 5.5.7. *Fix $\alpha \in (1,2)$. Let $(\tilde{\mathbf{e}}_x^{(\delta)}, \tilde{\mathbf{h}}_x^{(\delta)})$ be defined as in (5.14). Let $\mathscr{G} = \mathcal{G}(\tilde{\mathbf{h}}_x^{(\delta)}, \tilde{\mathbf{e}}_x^{(\delta)}, \mathcal{Q})$ denote the continuum random graph where $\mathcal{Q}$ is a Poisson point process with intensity $\frac{1}{\delta}\mathrm{Leb}$ for some $\delta > 0$. Then, almost surely,*

$$\mu(B(\rho,t) \setminus \{\rho\}) > 0, \qquad \forall t > 0.$$

The same holds for the graphs $\mathcal{M}_i$ appearing in (5.22).

Proof. The same statement holds for the graphs $\mathcal{M}_i$ hold by conditioning on their mass $\mu_i(\mathcal{M}_i)$ [59, Theorem 1.2].

We use the tree $\mathcal{T}_{\tilde{\mathbf{h}}_x^{(\delta)}}$ as a measured metric space, and when confusion might arise we will use subscripts to specify whether we are dealing with the tree or the graph.

We observe that from the quotient map

$$q : \mathcal{T}_{\tilde{\mathbf{h}}_x^{(\delta)}} \longrightarrow \mathcal{G}(\tilde{\mathbf{h}}_x^{(\delta)}, \tilde{\mathbf{e}}_x^{(\delta)}, \mathcal{Q})$$

in the construction of the random graph satisfies the following:

$$\mathrm{d}_{\mathscr{G}}(\rho, q(x)) \leq \mathrm{d}_{\mathcal{T}_{\tilde{\mathbf{h}}_x^{(\delta)}}}(\rho, x), \qquad \forall x \in \mathcal{T}_{\tilde{\mathbf{h}}_x^{(\delta)}}.$$

Consequently,

$$\mu_{\mathscr{G}}\left(B_{\mathscr{G}}(\rho, t)\right) \geq \mu_{\mathcal{T}_{\tilde{\mathbf{h}}_x^{(\delta)}}}\left(B_{\mathcal{T}_{\tilde{\mathbf{h}}_x^{(\delta)}}}(\rho, t)\right) = \int_0^1 1_{[\tilde{\mathbf{h}}_x^{(\delta)}(s) \in (0,t)]}\, ds.$$

Since the process $\tilde{\mathbf{h}}_x^{(\delta)}$ is non-negative and almost surely not identically zero. The result follows easily. □

5.5.2 Proof of Theorem 5.1.1

We now prove Theorem 5.1.1.

Proof of Theorem 5.1.1. Throughout the proof all limits will be as n or a subsequence of n goes towards infinity.

Make the components of the graph G_n^i, as measured metric spaces with graph distance and the measure of each vertex is one.

The processes $Z_{n,i}$ measure the number of vertices infected on day h, which is simply the number of vertices at distance h from ρ_i in G_n^i:

$$Z_{n,i}(h) = \#\{v \in G_n^i : d(v, \rho_i) = h\}$$

and the process $C_{n,i} = (C_{n,i}(h); h \geq 0)$ denote its running sum:

$$C_{n,i}(h) = \sum_{j=0}^{h} Z_{n,i}(j) = \mathfrak{m}_i(\{v \in G_n^i : d(v, \rho_i) \leq h\}).$$

These processes measure something close to the height profile on the components of the forests $F_n^{\mathrm{BF}}(\nu)$; however it is not exactly the same because of the addition of new leaves. This complicates a direct application of Theorem 5.2.2.

Let us write $Z^*_{n,i}(h)$ is the discrete Lamperti transform of $X^{\mathrm{BF}}_{n,i}$:

$$Z^*_{n,i}(h+1) = 1 + X^{\mathrm{BF}}_{n,i} \circ C^*_{n,i}(h), \qquad C^*_{n,i}(h) = \sum_{j=0}^{h} Z^*_{n,i}(j). \tag{5.26}$$

This process $Z^*_{n,i}$ measures the number of vertices at height h in the i^{th} largest component of the forest $F_n^{\mathrm{BF}}(\nu)$; see the discussion around (5.6) and more generally [49]. Said another way, the values of $Z^*_{n,i}(h)$ and $Z_{n,i}(h)$ for a fixed h only differ by the number of new leaves at height h in the component of the forest $F_n^{\mathrm{BF}}(\nu)$.

The total number of such additional vertices is twice the number of bf backedges (which is the number of df backedges as well). Therefore, for a fixed index i, the number of bf backedges in G_n^i is a tight sequence in the index n of random variables. Indeed, a stronger statement is true. By Proposition 5.12 in [59] the weak convergence

$$\#\{\text{bf backedges in } G_n^i\} \overset{d}{\Longrightarrow} \text{Poisson}\left(\frac{1}{\delta}\int_0^{\zeta(\tilde{e}_i)} \tilde{e}_i(t)\,dt\right),$$

where $\tilde{e}_i$ is as in (5.21). Now for each i we can bound the difference between $Z_{n,i}(h)$ and $Z^*_{n,i}(h)$ for each i uniformly in h. Indeed

$$\begin{aligned}\sum_{h\ge 0} |Z_{n,i}(h) - Z^*_{n,i}(h)| &= \sum_{h\ge 0} \left|\#\{v \in G_n^i : d(v,\rho_i) = h\} - Z^*_{n,i}(h)\right| \\ &= 2\cdot \#\{\text{bf backedges in } G_n^i\} \le \kappa_{n,i}\end{aligned}$$

where for each $i \ge 1$ the sequence $(\kappa_{n,i}; n\ge 1)$ is a tight sequence of random variables.

By Slutsky's theorem, in order to prove the rescaled convergence of $(Z_{n,i}; i\ge 1)$ to the desired limit, we just need to prove the convergence of $(Z^*_{n,i}; i \ge 1)$ under the same scaling regime to the same limiting processes. Indeed, their difference, when rescaled by $n^{-\frac{1}{\alpha+1}}$ converges in probability to the zero path in the Skorohod space:

$$\left(n^{-\frac{1}{\alpha+1}}\kappa_{n,i}; t\ge 0\right) \overset{d}{\Longrightarrow} (0; t\ge 0) \qquad \text{as } n\to\infty.$$

By Proposition 5.4.1, for any fixed integer N

$$\left(\left(\left(\left(n^{-\frac{\alpha}{\alpha+1}} C^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t\rfloor); t\ge 0\right), \left(n^{-\frac{1}{\alpha+1}} X^{\mathrm{BF}}_{n,i}(\lfloor n^{\frac{\alpha}{\alpha+1}} t\rfloor); t\ge 0\right)\right); i \in [N]\right); n\ge 1\right)$$

viewed as a sequence in $\mathbb{D}(\mathbb{R}_+,\mathbb{R})^{2N}$, is tight. Consequently, in the product topology over the index $i\ge 1$ the sequence

$$\left(\left(\left(\left(n^{-\frac{\alpha}{\alpha+1}}C^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}}t\rfloor);t\ge 0\right),\left(n^{-\frac{1}{\alpha+1}}X^{\mathrm{BF}}_{n,i}(\lfloor n^{\frac{\alpha}{\alpha+1}}t\rfloor);t\ge 0\right)\right);i\ge 1\right);n\ge 1\right) \quad (5.27)$$

is tight in $(\mathbb{D}^2)^\infty$. Additionally, by Proposition 5.4.1 any subsequential limit, say

$$(((C_i(t);t\ge 0),(X_i(t);t\ge 0));i\ge 1), \quad (5.28)$$

must satisfy $C_i(t)=\int_0^t X_i\circ C_i(s)\,ds$. Moreover, the sequence $(X_i;i\ge 1)\overset{d}{=}(\tilde{e}^*_i;i\ge 1)\overset{d}{=}(\tilde{e}_i;i\ge 1)$. In particular subsequential limits of (5.27) are classified by a time-shift as in Proposition 5.3.1.

By Theorem 5.5.4 and Lemma 5.5.2, we know that the convergences in (5.21), (5.23) and (5.24) hold. By a tightness argument, we can assume that sequence converge jointly along a subsequence, which we will denote by the index n.

Observe the sequence

$$\left(\left(\left(n^{-\frac{\alpha}{\alpha+1}}C_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}}t\rfloor);t\ge 0\right);i\ge 1\right);n\ge 1\right)$$

is tight in $\mathbb{D}^\infty$. Indeed, this easily follows the tightness of

$$\left(\left(\left(n^{-\frac{\alpha}{\alpha+1}}C^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}}t\rfloor);t\ge 0\right);i\ge 1\right);n\ge 1\right)$$

in $\mathbb{D}^\infty$ discussed above and the bounds

$$|C_{n,i}(h)-C^*_{n,i}(h)|\le\sum_{j\ge 0}|Z_{n,i}(j)-Z^*_{n,i}(j)|\le\kappa_{n,i}.$$

Let us work on a subsequence of (5.27) which converges to (5.28). Call this index n_j. Then, by the previous paragraph,

$$\left(\left(n_j^{-\frac{\alpha}{\alpha+1}}C_{n_j,i}(\lfloor n_j^{\frac{\alpha-1}{\alpha+1}}t\rfloor);t\ge 0\right);i\ge 1\right)\overset{d}{\Longrightarrow}((C_i(t);t\ge 0);i\ge 1)$$

for the same processes C_i in (5.28). However, $C_{n,i}(h)=\#\{v\in G^i_n:d(v,\rho_i)\le h\}$ is just the measure of the ball of radius h in G^i_n and so an application of Lemma 5.3.3 implies

$$n_j{}^{-\frac{\alpha}{\alpha+1}}C_{n_j,i}(\lfloor n_j^{\frac{\alpha-1}{\alpha+1}}t\rfloor)\overset{d}{\Longrightarrow}\mu_i(B(\rho,t)),\qquad \text{for all but countably many } t>0, \quad (5.29)$$

where μ_i is the mass measure on the scaling limit of the graph component G_n^i. Hence $C_i(t)$ must satisfy:

$$C_i(t) \stackrel{d}{=} \mu_i(B(\rho, t)), \qquad \text{for Lebsgue a.e. } t > 0.$$

It follows easily from Proposition 5.5.7 that

$$\mathbb{P}(C_i(t) > 0, \forall t > 0) = 1, \qquad \forall i \geq 1.$$

Indeed $C_i(t)$ is non-decreasing, and we can find a countable dense set of $t > 0$ such that $P(C_i(t) > 0) = 1$. Hence, by Propositions 5.3.1 and $(X_i; i \geq 1) \stackrel{d}{=} (\tilde{e}_i; i \geq 1)$, we get $((C_i, X_i); i \geq 1) \stackrel{d}{=} ((C_i, \tilde{e}_i); i \geq 1)$ where (Z_i, C_i) is the Lamperti pair associated with $\tilde{e}_i$. Since this works for any subsequential limit, we conclude that the original sequence converges:

$$\left(\left(n^{-\frac{\alpha}{\alpha+1}} C_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t \rfloor; t \geq 0\right); i \geq 1\right) \stackrel{d}{\Longrightarrow} (C_i; i \geq 1).$$

A similar proof of Proposition 5.4.2 yields the joint convergence

$$\left(\left(n^{-\frac{1}{\alpha+1}} Z^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t \rfloor); t \geq 0\right), \left(n^{-\frac{\alpha}{\alpha+1}} C^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t \rfloor); t \geq 0\right); i \geq 1\right) \stackrel{d}{\Longrightarrow} ((Z_i, C_i); i \geq 1) \tag{5.30}$$

where (Z_i, C_i) is the Lamperti pair associated with the excursion $\tilde{e}_i$.

This proves the desired claim. □

Before turning to proof of Theorem 5.1.2, we state and prove the following lemma:

Lemma 5.5.8. *Couple the depth-first and breadth-first walks as in Lemma 5.5.2. Under the assumptions of Theorem 5.5.3, and using the notation in (5.26). There is joint convergence in distribution along a subsequence of the index $n \geq 1$ of the collection*

$$\Bigg(\Bigg(\left(n^{-\frac{1}{\alpha+1}} Z^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t \rfloor); t \geq 0\right), \left(n^{-\frac{\alpha}{\alpha+1}} C^*_{n,i}(\lfloor n^{\frac{\alpha-1}{\alpha+1}} t \rfloor); t \geq 0\right),$$
$$\left(n^{-\frac{1}{\alpha+1}} X^{\mathrm{BF}}_{n,i}(\lfloor n^{\frac{\alpha}{\alpha+1}} t \rfloor); t \geq 0\right), \mathrm{scale}(n^{-\frac{\alpha-1}{\alpha+1}}, n^{-\frac{\alpha}{\alpha+1}}) G^i_n, \# \mathscr{P}^{\mathrm{BF}}_{n,i}\Bigg); i \geq 1\Bigg)_{n \geq 1}$$

towards

$$((Z_i, C_i, \tilde{e}^*_i, \mathcal{M}_i, \mathrm{sur}(\mathcal{M}_i)); i \geq 1),$$

where

1. $\mathcal{M}_i$ *are as in Theorem* 5.5.4 *and* (5.22)*;*

2. (Z_i, C_i) *is the Lamperti pair associated with the excursion* $\tilde{e}_i^*$*;*

3. *The process* $C_i(t) = \mu_i(B(\rho, t))$ *for almost all (and hence all)* $t \geq 0$*;*

4. *The excursions* $(\tilde{e}_i^*; i \geq 1) \stackrel{d}{=} (\tilde{e}_i; i \geq 1)$ *in the construction of* $\mathcal{M}_i$*, and, in particular, the length of the excursion* $\tilde{e}_i^*$ *is the mass of the space* $\mathcal{M}_i$*, i.e.*

$$\zeta(\tilde{e}_i^*) = \mathcal{M}_i;$$

5. *The random variable* $\operatorname{sur}(\mathcal{M}_i)$*, the surplus of the space* $\mathcal{M}_i$ *is*

$$\operatorname{sur}(\mathcal{M}_i) \stackrel{d}{=} \operatorname{Poisson}\left(\frac{1}{\delta}\int_0^{\zeta(\tilde{e}_i^*)} \tilde{e}_i^*(t)\,dt\right);$$

6. *Lastly, conditionally on the length of excursion* $\zeta_i := \zeta(\tilde{e}_i^*)$ *and the surplus values* $R_i := \operatorname{sur}(\mathcal{M}_i)$*, the graph* $\mathcal{M}_i$ *satisfies*

$$\mathcal{M}_i \stackrel{d}{=} \operatorname{scale}\left(\zeta_i^{(\alpha-1)/\alpha}, \zeta_i\right)\mathscr{G}^{(\alpha, R_i)}$$

Proof. These are tight random variables in each of the marginals, so joint convergence along a subsequence is standard.

Item 1 follows from the referenced theorem.

Item 2 follows from the proof of Theorem 5.1.1 and the identity in distribution $(\tilde{e}_i; i \geq 1) \stackrel{d}{=} (\tilde{e}_i^*; i \geq 1)$ previously seen.

Item 3 follows from the proof of Theorem 5.1.1, particularly around (5.29).

Item 4 is from Lemma 5.5.6.

Item 5 follows from Theorem 5.5 and Proposition 5.12 in [59] along with the equality $\#\mathscr{P}_{n,1}^{\mathrm{DF}} = \#\mathscr{P}_{n,i}^{\mathrm{BF}}$.

Item 6 follows from the proof of Theorem 1.2 in [59]. □

5.5.3 *Proof of Theorem 5.1.2*

Proof of Theorem 5.1.2. The big content of this proof is to show that we can condition on the length $\zeta(\tilde{e}_i^*)$ of the excursion $\tilde{e}_i^*$ and the surplus of the graphs $\mathcal{M}_i$ by using Lemma 5.5.8 and scaling results for the excursions proved in [59].

By Lemma 5.5.8, we know that we can write the height profile of the graph $\mathcal{M}_i$ (which is of random mass) as the process Z_i where (Z_i, C_i) is the Lamperti pair associated with the excursion $\tilde{e}_i^*$. In fact, we know

$$((\mu_i(B(\rho_i, v); v \geq 0), \mu_i(\mathcal{M}_i), \operatorname{sur}(\mathcal{M}_i))_{i\geq 1} \stackrel{d}{=} ((C_i(v); v \geq 0), \zeta(\tilde{e}_i^*), R_i)_{i\geq 1}$$

where (Z_i, C_i) is the Lamperti pair associated with the excursion $\tilde{e}_i^*$ and $R_i \sim \text{Poisson}\left(\frac{1}{\delta}\int_0^{\zeta(\tilde{e}_i^*)} \tilde{e}_i^*(t)\, dt\right)$.

Conditioning on the values of $\zeta(\tilde{e}_1^*)$ and R_1 gives

$$\left((\mu_1(B(\rho_1, v)); v \geq 0)\middle|\mu_1(\mathcal{M}_1) = 1, \operatorname{sur}(\mathcal{M}_1) = k\right) \stackrel{d}{=} \left((C_1(t); t \geq 0)\middle|\zeta(\tilde{e}_1^*) = 1, R_1 = k\right). \tag{5.31}$$

We can use the *proof* of Theorem 1.2 in [59] to handle this conditioning on the right-hand side and the statement of Theorem 1.2 in [59] to handle the left-hand side.

The conditioning in the proof of Theorem 1.2 in [59] gives

$$\mathbb{E}\left[g(\tilde{e}_i^*)\middle|R_i = k, \zeta(\tilde{e}_i) = 1\right] = \mathbb{E}[g(\mathbf{e}^{(k)})]$$

for all positive functionals g and where $\mathbf{e}^{(k)} = (e^{(k)}(t); t \in [0,1])$ is defined in (5.4). In particular this holds for $i = 1$. Recall from Section 5.3.2, that c_1 is simply a functional of $\tilde{e}_1^*$. Therefore, conditionally on the values of $\zeta(\tilde{e}_1^*)$ and R_1 we have

$$\left((C_1(t); t \geq 0)\middle|\zeta(\tilde{e}_1^*) = 1, R_1 = k\right) \stackrel{d}{=} \left(\mathbf{c}^{(k)}(t); t \geq 0\right)$$

where $(\mathbf{z}^{(k)}, \mathbf{c}^{(k)})$ is the Lamperti pair associated with the excursion $\mathbf{e}^{(k)}$.

The left-hand side of (5.31) is easy to condition with part (6) of Lemma 5.5.8. Conditionally on the values of $\mu_1(\mathcal{M}_1)$ and $\operatorname{sur}(\mathcal{M}_1)$ (which is precisely the conditioning described above for c_1) the metric spaces $\mathcal{M}_1$ satisfies

$$\left(\mathcal{M}_1\middle|\mu_1(\mathcal{M}_1) = 1, \operatorname{sur}(\mathcal{M}_1) = k\right) \stackrel{d}{=} \mathscr{G}^{(\alpha,k)}.$$

Hence

$$(\mu_{\mathscr{G}(\alpha,k)}(B(\rho,v)); v \geq 0) \overset{d}{=} \left(\mathbf{c}^{(k)}(v); v \geq 0\right).$$

An application of Proposition 5.3.1 completes the proof. □

5.6 Discussion

In this work we showed convergence of the height profiles for the macroscopic components of a certain class of critical random graphs. We did this by looking at the height profile of these graphs and we relied on the weak convergence results that exist in the literature on some encoding stochastic processes. We observe that these techniques can likely be extended to other graph models appearing in the literature.

For example, the work of Broutin, Duquesne and Wang [44, 45] provides the rescaled convergence under certain conditions of the *rank-1 inhomogeneous model* associated to a weight sequence $\mathtt{w} = (w_1, \cdots, w_n)$. That graph, whose asymptotics were studied by Aldous and Limic in [13], is a graph on n vertices where edges are added independently with probability

$$\mathbb{P}\left(\text{edge}\{i,j\} \text{ is included}\right) = 1 - \exp\left(-w_i w_j / q\right)$$

for some parameter $q > 0$. This graph goes by other names as well: the Poisson random graph [32, 150] and the Norros-Reittu model [32]. See also [31, 33] and Section 6.8.2 of [173] for more information. The resulting limiting processes and graphs are related to Lévy-type processes (sometimes called Lévy processes without replacement) constructed from spectrally positive Lévy processes which are not stable. As in [5, 59], Broutin, Duquesne and Wang show convergence of the graphs as metric spaces by using a depth-first descriptions. However, there is also convergence of the breadth-first walks [13], and so proving convergence of the height profiles should similar to the proof of Theorems 5.1.1 and Theorems 5.2.2.

Using the results in the literature on Galton-Watson trees conditioned on having a fixed size [11, 70, 131, 142] one can recover the Jeulin identity [108] and its α-stable extension due to Miermont [146] from our Theorem 5.2.1 as well. The proofs in [108, 146] do not rely on weak convergence arguments. For proofs using weak-convergence arguments more in-line with the results of this papers see Kersting's work [120], or joint work of Angtuncio and

Uribe Bravo [18]. See also [14] for a weak convergence result in a slightly weaker topology.

More generally, under certain conditions (see Theorem 2.3.1 in [73]) on the offspring distribution, there is convergence of Galton-Watson forests to continuum forests encoded by spectrally positive Lévy processes. Under these assumptions, one can use a modification of Lemma 4.8 in [143] or Lemma 5.8 in [59], one should be able to prove a Jeulin-type identity for excursions for non-stable Lévy processes and their associated height processes by a simple application of Theorem 5.2.2. As far as the author is aware, such results are not present in the literature.

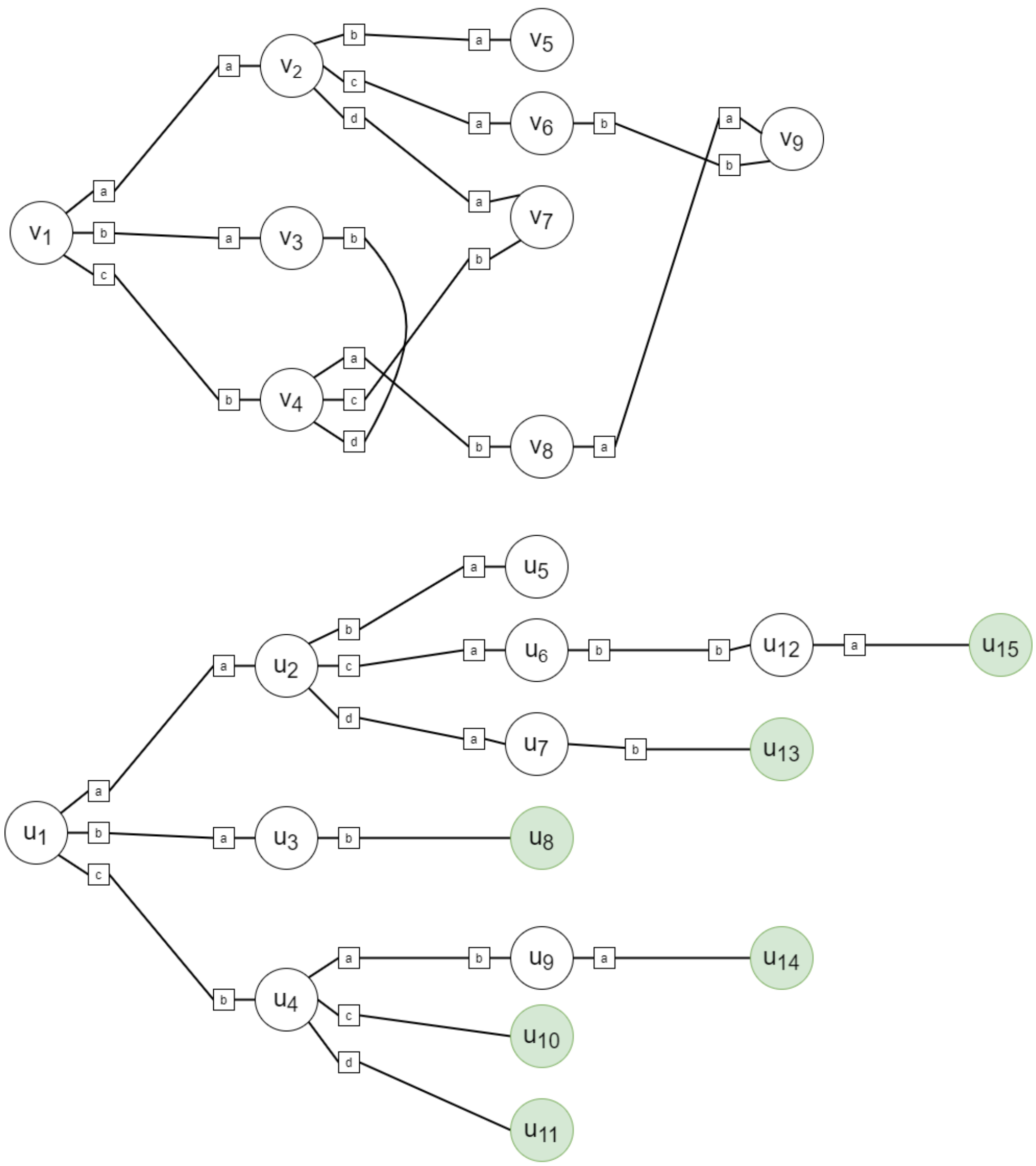


Figure 5.4: The first component of $M^{\mathrm{BF}}(\mathbf{d}^n)$ (top) and the corresponding first component of $F^{\mathrm{BF}}(\mathbf{d}^n)$ (bottom). The circles are the vertices, the labeled squares are the ordering of the half-edges connected to each hub with $a < b < c < d$. The three bf backedges in this graph connect half-edge (v_3, b) to (v_4, d), half-edge (v_4, c) to (v_7, b), and half-edge (v_8, a) to (v_9, a). The new-leafs are vertices $u_8, u_{10}, u_{11}, u_{13}, u_{14}, u_{15}$ in green and they are ordered according to the ordering of the half-edge to which it is connected.

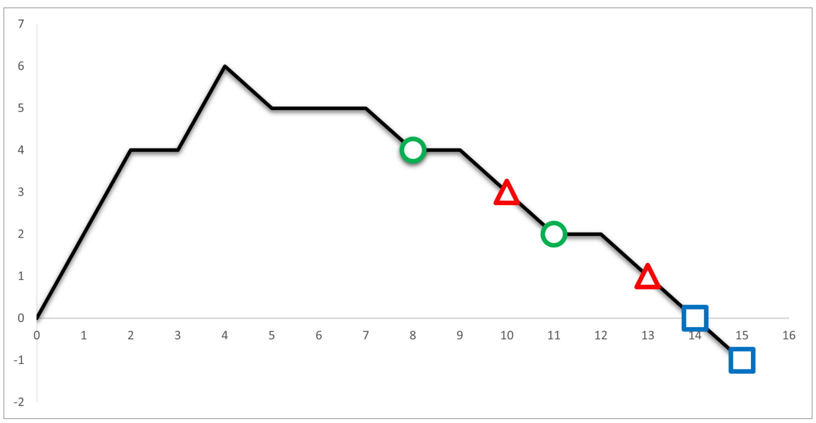


Figure 5.5: The excursion $X_{n,i}^{\mathrm{BF}}$ associated with the component shown in Figure 5.4 with the marks $\mathscr{P}_{n,i}^{\mathrm{BF}}$ included. The vertices u_8^{BF} and u_{11}^{BF} are paired, u_{10}^{BF} and u_{13}^{BF} are paired, and u_{14}^{BF} and u_{15}^{BF} are paired. These are represented by the green circles, red triangles and blue squares above.

Chapter 6

THE GORIN-SHKOLNIKOV IDENTITY AND ITS RANDOM TREE GENERALIZATION

6.1 Introduction

Let $e = (e_t; t \in [0,1])$ denote a standard Brownian excursion, or, equivalently, a 3-dimensional Bessel bridge from 0 to 0 of unit duration, see [156, Chapter XII]. This process is a semi-martingale with quadratic variation t, and so possesses a family of local times $L = (L_t^v; v \geq 0, t \in [0,1])$ which satisfies almost surely the occupation time formula:

$$\int_0^t g(e_r)\,dr = \int_0^\infty g(v) L_t^v\,dv, \qquad \forall g \text{ bounded Borel measurable.}$$

In this particular case, there exists a jointly continuous modification of L for which the above formula holds almost surely.

Using topics in random matrix theory Gorin and Shkolnikov [97] obtained, as a corollary of one of their main results, the following interesting distributional identity:

$$\int_0^1 e_t\,dt - \frac{1}{2}\int_0^\infty \left(L_1^v\right)^2\,dv \overset{d}{=} \mathscr{N}\left(0, \frac{1}{12}\right). \tag{6.1}$$

Gorin and Shkolnikov were studying Gaussian β ensembles and their eigenvalues, which are briefly discussed in Section 3.4. Namely, they were concerned with the eigenvalues of the tri-diagonal matrices of the form

$$M_N^\beta = \frac{1}{\sqrt{\beta}} \begin{pmatrix} \sqrt{2}g_1 & \chi_{(N-1)\beta} & 0 & \dots & \\ \chi_{(N-1)\beta} & \sqrt{2}g_2 & \chi_{(N-2)\beta} & 0 & \dots \\ 0 & \chi_{(N-2)\beta} & \sqrt{2}g_3 & \ddots & \\ \vdots & & \ddots & \ddots & \\ & & & & \chi_\beta \\ & & & \chi_\beta & \sqrt{2}g_N \end{pmatrix}, \qquad \text{for } \beta > 0, N \geq 1, \tag{6.2}$$

where $\chi_\beta, \chi_{2\beta}, \cdots, \chi_{(N-1)\beta}$ are independent χ-distributed random variables which are indexed by their parameters and $g_1, \ldots, g_N$ are independent standard normal random variables. The ordered eigenvalues, $\lambda_1^{(N)} \geq \lambda_2^{(N)} \geq \cdots \geq \lambda_N^{(N)}$, of M_N^β in (6.2) are of particular interest for their connection to β-ensembles. The celebrated work of Dumitriu and Edelman [69] shows that joint distribution of the eigenvalues has a density proportional to

$$\prod_{1 \leq i < j \leq N} (x_i - x_j)^\beta \prod_{k=1}^{N} e^{-\beta x_k / 4}.$$

Ramírez, Rider and Virág [155] relate a rescaling and recentering of largest eigenvalues, i.e. the *edge*, of M_N^β to the eigenvalues of the so-called stochastic Airy operator. In particular they show for each fixed $k = 1, 2, \cdots$ that

$$\Lambda_{j,N} := N^{1/6} \left(2\sqrt{N} - \lambda_j^{(N)} \right), \qquad j \in \{1, 2, \cdots, k\}$$

converge jointly in distribution as $N \to \infty$ to eigenvalues $\Lambda_1 \leq \Lambda_2 \leq \cdots$ of the operator

$$\mathscr{H}^\beta f := \left(-\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} W'_x \right) f, \qquad f(0) = 0, \qquad f \in L^2(\mathbb{R}_+)$$

where W' is a white noise on $\mathbb{R}_+ := [0, \infty)$. In turn, Gorin and Shkolnikov showed the powers of the matrix $\frac{1}{2\sqrt{N}} M_N^\beta$ converge in a certain operator-theoretic sense to the semigroup generated by $\frac{1}{2}\mathscr{H}^\beta$. See Theorem 2.1 in [97] for a proper formulation of this convergence.

Proposition 2.7 in [97] relates the eigenvalues of $\mathscr{H}^\beta$ to the functional in (6.1) by

$$\mathbb{E}\left[\sum_{j \geq 1} e^{T \Lambda_j / 2} \right] = \sqrt{\frac{2}{\pi T^3}} \mathbb{E}\left[\exp\left(-\frac{T^{3/2}}{2} \left(\int_0^1 e_t \, dt - \frac{1}{\beta} \int_0^\infty (L_1^v)^2 \, dv \right) \right) \right].$$

By comparing the Laplace transform in the right-hand side with the prior work of Okounkov [151] when $\beta = 2$, Gorin and Shkolnikov were able to prove the equality in distribution (6.1).

Lamarre and Shkolnikov [126] connected a *spiked* version of the matrix model in [97] to a reflected Brownian bridge as well. If we let $B^{|\mathrm{br}|} = (B_t^{|\mathrm{br}|}; t \in [0, 1])$ denote a reflected Brownian bridge and let $L = (L_t^v; v \geq 0, t \in [0, 1])$ denote its local time, then the work of [126] relates eigenvalues of a spiked operator to the random variable

$$A_\beta = \sqrt{12} \left(\int_0^1 B_t^{|\mathrm{br}|} \, dt - \frac{1}{\beta} \int_0^\infty (L_1^v)^2 \, dv \right). \tag{6.3}$$

Lamarre and Shkolnikov are able to give an explicit formulation of the moment generating function of A_β only in the case $\beta = 2$, which evaluates to

$$\mathbb{E}\left[A_2^{2n-1}\right] = -\frac{2^n(2n-1)!}{4(n-1)!}\sqrt{6\pi}, \qquad n = 1, 2, \cdots, 7.$$

This leads Lamarre and Shkolnikov "to believe that A_2 admits an interesting combinatorial interpretation." This work is devoted to giving one such combinatorial interpretation using random trees or random forests which can be generalized to the infinite forest models of [8, 72]. We should mention that there are many relationships between the area of a Brownian excursion and limits appearing in enumerative combinatorics. For a good survey on such connections, Janson's survey [105] is an excellent resource.

Shortly after Gorin and Shkolnikov posted their work on the arXiv, Hariya [102] gave a path-wise interpretation and proof of the normality result (6.1). Hariya's proof relies heavily on the Jeulin identity [108]. Define

$$H(x) = \int_0^x L_1^y \, dy = \int_0^1 1_{[e_t \le x]} \, dt,$$

and let $H^{-1} = (H^{-1}(t); t \in [0,1])$ denote its right-continuous inverse. Jeulin's identity is the following identity in distribution

$$\left(\frac{1}{2} L_1^{H^{-1}(t)}; t \in [0,1]\right) \stackrel{d}{=} (e_t; t \in [0,1]) \,.$$

Lamarre and Shkolnikov [126] use Hariya's idea and some results of Pitman [153] to show

$$\left(\frac{1}{\sqrt{12}} A_2 \middle| L_1^0 = x\right) \stackrel{d}{=} \mathcal{N}\left(-\frac{x}{4}, \frac{1}{12}\right). \tag{6.4}$$

For more information on this time-change including its proof, see Jeulin's original work [108, pg. 264] or a random tree interpretation and generalization in [14, 146, 153]. The paper [14] conveys this transformation succinctly: The Jeulin identity "can be roughly interpreted as the width of the layer of the tree containing [a] vertex [...] where the vertices are labeled [...] in breadth-first order."

Because of their connection to random matrix theory and the study of the stochastic Airy semigroup, the distributional properties of the random variables A_β in (6.3) and the random variables

$$\int_0^1 e_t \, dt - \frac{1}{\beta} \int_0^\infty \left(L_1^v\right)^2 \, dv$$

for each $\beta > 0$ are of interest. Apart from the case $\beta = 2$, where tools from stochastic calculus are useful [102, 126], it is not obvious what approach to studying these random variables will be fruitful. In this article, we present a discrete forest model in order to understand these random variables which can yield a wide class of results involving the difference of an integral of a stochastic process and a constant multiple of an integral of its squared local time. See Corollaries 6.1.4 and 6.3.3 below. Some of these techniques hold even outside of a Brownian regime and *do not* rely on the stochastic calculus techniques of [102, 126]. However, these stochastic calculus techniques can be used when we are inside the Brownian regime. We develop the techniques in [102, 126] further with Theorems 6.1.5 and 6.1.6.

6.1.1 Random tree and branching process interpretation

We begin by giving the discrete interpretation of the Gorin-Shkolnikov identity in (6.1) and its generalization in (6.4) in terms of two statistics on a random forests. We then show that these same statistics have a functional limit for a wide class of random forest models.

Consider the vertex set $[n] := \{1, 2, \cdots, n\}$. By a forest on $[n]$, we mean a cycle-free graph on the vertices $[n]$. We say that a forest $\mathfrak{f}$ is *rooted* if each connected component has a distinguished vertex called its root. Roots will be denoted by the letter ρ. We equip the forests with the graph distance, denoted by $\operatorname{dist}(-,-)$. Given a forest $\mathfrak{f}$ on $[n]$ we can define two statistics on the graph, measuring height and width. The first statistic, denoted by $\mathbf{ht}$, is the height of the vertex, i.e. the distance from the root. The second statistic counts the number of "cousin" vertices, i.e. vertices at the same height; we denote this by $\mathbf{csn}$. If v is in the same component as the root ρ, then the two statistics are

$$\mathbf{ht}(v) = \operatorname{dist}(v, \rho), \qquad \mathbf{csn}(v) = \#\{w \in \mathfrak{f} : \mathbf{ht}(v) = \mathbf{ht}(w), v \neq w\}. \tag{6.5}$$

The statistics are portrayed graphically in Figure 6.1 below.

The following theorem describes the scaling relationship.

Theorem 6.1.1. *Fix a sequence k_n such that $2k_n/\sqrt{n} \to x \geq 0$. For each n, let $\mathfrak{f}^{(n)}$ denote a rooted forest on the vertex $[n]$ with k_n roots uniformly chosen among all such forests. Then*

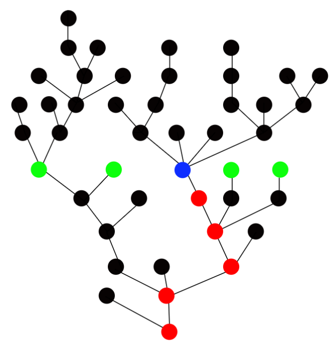


Figure 6.1: A pictorial representation of **ht** and **csn**. The green vertices are the cousins of the blue vertex, and the number of red vertices represents the height of the blue vertex. In this example, $\mathbf{ht}(v) = 5$ and $\mathbf{csn}(v) = 4$ where v is the blue vertex.

the following weak convergence holds as $n \to \infty$

$$\frac{1}{2n^{3/2}} \sum_{v \in \mathfrak{f}^{(n)}} \mathbf{ht}(v) - \frac{1}{n^{3/2}} \sum_{v \in \mathfrak{f}^{(n)}} \mathbf{csn}(v) \overset{(d)}{\Longrightarrow} \mathscr{N}\left(\frac{-x}{4}, \frac{1}{12}\right).$$

The above theorem relies on the connection between random trees and forests and excursions and reflected Brownian bridges in the literature. See, for example, [9, 153] and references therein for more details on this connection. It also relies heavily on the extension of Jeulin's identity implicit in the work of Pitman [153], which relates the local time of a reflected Brownian bridge with a time-change of a 3-dimensional Bessel bridge. We discuss some extensions of the stochastic calculus approach used in [102, 126] below. Before moving to that section, we discuss a generalization using branching processes.

In the study of branching processes, there is a breadth-first model similar to the Jeulin identity. It is called the Lamperti transform originating in the work [128]. Consider a genealogical structure with immigration depicted in Figure 6.2 below.

In this picture, the vertices (or individuals) come in two types. Using the language in [72], the *mutant* individuals are the unlabeled white vertices and the *non-mutant* vertices are the labeled black vertices. The mutant vertices are simply a convenient way to introduce immigration and play little to no role in much of our analysis. The non-mutant vertices are

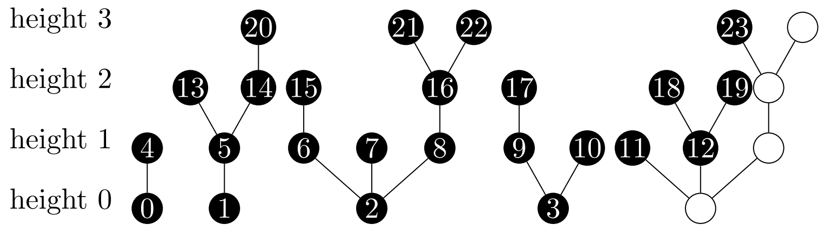


Figure 6.2: The indices of the breadth-first labeling on an immigration forest of 24 non-mutant vertices.

labeled by $w_0, w_1, \cdots$ in a breadth-first order with the convention that the immigrant vertices are labeled last in each generation. The mutant vertices are unlabeled. Let χ_j denote the number of children of the vertex w_j, and let η_h denote the number of immigrants which arrive at height (or generation) h. In Figure 6.2, the forest has 4 non-mutant individuals at height 0, the sequence of χ_j's begins $\chi_0 = 1$, $\chi_1 = 1$, $\chi_2 = 3$, etc. and the sequence of η_h's begins $\eta_1 = 2$, $\eta_2 = 0$, and $\eta_3 = 1$. Conversely, given a number k of non-mutant vertices at height 0 (i.e. the roots) and two sequence of non-negative integers $(\chi_j; j = 0, 1, \cdots)$ and $(\eta_h; h = 1, 2, \cdots)$ one can inductively construct a forest like the one in Figure 6.2. See Section 6.2.1 for more information.

If we let c_h denote the number of non-mutant vertices appearing at or before height h, then with the conventions $c_{-1} = 0$ the vertices at generation exactly h are indexed by $c_{h-1}, c_{h-1} + 1, \cdots, c_h - 1$. From here it follows

$$c_{h+1} = c_h + \sum_{j=c_{h-1}}^{c_h - 1} \chi_j + \eta_{h+1}.$$

Let z_h denote the number of non-mutant vertices at height exactly h, let $x_m = \sum_{j=0}^{m-1}(\chi_j - 1)$ and let $y_h = \sum_{j=1}^{h} \eta_j$. As observed in [49], we can recover the successive generation sizes by solving the discrete equation

$$z_{h+1} = k + x_{c_h} + y_{h+1}, \qquad c_h = \sum_{j=0}^{h} z_j.$$

This is the *discrete Lamperti transform*, which was rigorously studied in [49] and generalized by the same authors in [50]. The authors described continuous time analogs and developed robust limit theorems. In particular, the authors of [49] show that if $X = (X_t; t \geq 0)$ is a Lévy process with no negative jumps and $Y = (Y_t; t \geq 0)$ is an independent subordinator with Laplace transforms satisfying

$$\mathbb{E}[\exp(-\lambda X_t)] = \exp(t\Psi(\lambda)), \qquad \mathbb{E}[\exp(-\lambda Y_t)] = \exp(-t\Phi(\lambda)) \qquad \lambda > 0$$

then there exists a unique càdlàg solution to the equation

$$Z_t = x + X_{C_t} + Y_t, \qquad C_t = \int_0^t Z_s\, ds, \tag{6.6}$$

which exists until some possible explosion time. Moreover the process $Z = (Z_t; t \geq 0)$ is a continuous state branching process started from x with branching mechanism Ψ and immigration rate Φ. We denote the law of Z solving (6.6) for these Lévy processes X and Y by $\mathrm{CBI}_x(\Psi, \Phi)$. See Section 6.2.4 for more details on these processes. However, throughout this work we will work under the following assumptions on Ψ and Φ:

Assumption 1. The function Ψ is of the form

$$\Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (e^{-\lambda r} - 1 + \lambda r 1_{[r<1]})\, \pi(dr),$$

for some $\alpha \in \mathbb{R}$, $\beta \geq 0$ and a Radon measure π such that $(1 \wedge r^2)\pi(dr)$ is a finite measure. We assume that Ψ in addition is *conservative*, meaning [99]

$$\int_0^\varepsilon \frac{1}{|\Psi(u)|}\, du = \infty, \qquad \forall \varepsilon > 0. \tag{6.7}$$

The subordinator $Y = (Y_t; t \geq 0)$ is $Y_t = \delta t$ for all $t \geq 0$, equivalently $\Phi(\lambda) = \delta\lambda$. □

The first part of Assumption 1 is equivalent to X being a real-valued Lévy process which has no negative jumps and is not killed, see [26]. The second part of Assumption 1 is in some sense the most general assumption on Ψ that we can make for weak convergence arguments, see [114].

As briefly mentioned above, we can construct a forest, $\mathfrak{f}$, from two sequences $(\chi_j; j = 0, 1, \cdots)$ and $(\eta_h; h = 1, 2, \cdots)$ along with a non-negative integer k. We say that $\mathfrak{f}$ is a *Galton-Watson immigration forest* with offspring distribution μ and immigration distribution ν starting from k roots if the sequences $(\chi_j) \overset{i.i.d.}{\sim} \mu$ and $(\eta_h) \overset{i.i.d.}{\sim} \nu$ and there are k non-mutant individuals at height 0. We use the notation $\mathfrak{f} \sim \mathrm{GWI}_k(\mu,\nu)$. We also label the non-mutant vertices by $w_0, w_1, \cdots$ in a breadth-first manner. See Section 6.2.1 and Definition 2 therein for a more in-depth description of this model.

We can again define two statistics, **ht** and **csn** by (6.5). We will examine the following two processes,

$$K_p^{\mathfrak{f}} = \sum_{j=0}^{p-1} \mathbf{csn}(w_j), \qquad \text{and} \qquad J_p^{\mathfrak{f}} = \sum_{j=0}^{p-1} \mathbf{ht}(w_j). \tag{6.8}$$

We refer to $K^{\mathfrak{f}} = (K_p^{\mathfrak{f}}; p = 0,1,\cdots)$ as the cumulative breadth-first cousin process and $J^{\mathfrak{f}} = (J_p^{\mathfrak{f}}; p = 0,1,\cdots)$ as the cumulative breadth-first height process.

We will be concerned about distributional limits of the processes $K^{\mathfrak{f}}$ and $J^{\mathfrak{f}}$, for sequences of forests $(\mathfrak{f}^{(n)}; n \ge 1)$. For ease of notation we write $K^{*,n}$ instead $K^{\mathfrak{f}^{(n)}}$, along with similar notation shifts for other processes.

Let $(\mu_n; n \ge 1)$ and $(\nu_n; n \ge 1)$ be any sequence of probability measures on $\mathbb{N}_0 = \{0,1,\cdots\}$ which satisfy Assumption 2 below.

Assumption 2. There exists a sequence $(\gamma_n; n \ge 1)$ of non-negative integers and a real-valued Lévy process $X = (X_t; t \ge 0)$ with non-negative jumps such that its negative Laplace exponent Ψ is conservative (i.e. satisfies (6.7)). We assume

$$\frac{1}{n}\sum_{j=0}^{n\gamma_n - 1} (\chi_j^{*,n} - 1) \overset{(d)}{\Longrightarrow} X_1, \qquad \text{as } n \to \infty,$$

where $(\chi_j^{*,n}; j \ge 1)$ are i.i.d. with common distribution μ_n. Simultaneously, for the same sequence γ_n, we suppose

$$\frac{1}{n}\sum_{j=1}^{\gamma_n} \eta_j^{*,n} \overset{p}{\longrightarrow} \delta,$$

for some non-random constant $\delta > 0$ where $(\eta_j^{*,n}; j \ge 1)$ are i.i.d. with common distribution ν_n.

There are many examples of sequences of distributions $(\mu_n; n \ge 1)$ and $(\nu_n; n \ge 1)$ which satisfy Assumption 2. Some examples are included below. In the following examples and throughout the work, we write $[x]$ for the integer part of a real number x.

1. For any $\mu_n = \mu$ with mean 1 and variance $\sigma^2 < \infty$ and $\nu_n = \nu$ with mean $\delta > 0$,

$$\frac{1}{n}\sum_{j=0}^{n^2-1}\left(\chi_j^{*,n} - 1\right) \overset{(d)}{\Longrightarrow} \mathscr{N}(0,\sigma^2), \qquad \text{and} \qquad \frac{1}{n}\sum_{j=1}^{n}\eta_j^{*,n} \overset{p}{\longrightarrow} \delta,$$

by the central limit theorem and the weak law of large numbers. Hence, Assumption 2 holds for $\gamma_n = n$ and $X_t = \sigma B_t$ for a standard Brownian motion B (i.e. $\Psi(\lambda) = \frac{\sigma^2}{2}\lambda^2$).

2. For a non-centered normal distributions, take a fixed $a \in \mathbb{R}$ with $\mu_n = \text{Poisson}(1 + an^{-1})$ for n sufficiently large, $\nu_n = \nu$ having mean $\delta > 0$ and $\gamma_n = n$, then

$$\frac{1}{n}\sum_{j=0}^{n^2-1}\left(\chi_j^{*,n} - 1\right) \overset{(d)}{\Longrightarrow} \mathscr{N}(a,1), \qquad \text{and} \qquad \frac{1}{n}\sum_{j=1}^{n}\eta_j^{*,n} \overset{p}{\longrightarrow} \delta.$$

This corresponds to $X_t = B_t + at$ and $\Psi(\lambda) = -a\lambda + \frac{1}{2}\lambda^2$.

3. Outside of a Brownian regime, we can take $\mu_n = \mu$ where $\lim_{k\to\infty} k^{1+\alpha}\mu(k) = c > 0$ for some $\alpha \in (1,2)$ and $\sum_{k=0}^{\infty} k\mu(k) = 1$. Then, for $\gamma_n = [n^{\alpha-1}]$ we have

$$\frac{1}{n}\sum_{j=0}^{n\gamma_n-1}\left(\chi_j^{*,n} - 1\right) \overset{(d)}{\Longrightarrow} X_1,$$

where X is a spectrally positive α stable Lévy process and $\Psi(\lambda) = \tilde{c}\lambda^\alpha$ when $\lambda \geq 0$ for some constant $\tilde{c}$. See, for example, [75, Section 3.7]. Taking, for example, $\nu_n = \text{Poisson}(\delta n^{2-\alpha})$ for some $\delta > 0$ implies

$$\frac{1}{n}\sum_{j=1}^{\gamma_n}\eta_j^{*,n} \overset{p}{\longrightarrow} \delta.$$

We now state the following generalization of the Gorin-Shkolnikov identity, at least as described in the Theorem 6.1.1 interpretation of the result.

Theorem 6.1.2. *Suppose that μ_n and ν_n satisfy Assumption 2 and fix some $x \geq 0$. For each n, let $\mathfrak{f}^{(n)}$ denote a $\mathrm{GWI}_{[nx]}(\mu_n,\nu_n)$ forest. Let $K^{*,n}$ (resp. $J^{*,n}$) denote the cumulative breadth-first cousin (resp. height) process. Then, in the Skorohod space $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$ the following weak convergence holds*

$$\left(\frac{\delta}{n\gamma_n^2}J^{*,n}_{[n\gamma_n t]} - \frac{1}{n^2\gamma_n}K^{*,n}_{[n\gamma_n t]}; t \geq 0\right) \overset{(d)}{\Longrightarrow} \left(-xt - \int_0^t X_u\,du; t \geq 0\right),$$

where X is the Lévy process in Assumption 2.

The above theorem relies on the following lemma.

Lemma 6.1.3. *Suppose that Z is a $\mathrm{CBI}_x(\Psi,\Phi)$ process, where Ψ and Φ satisfy Assumption 1. Define $V=(V_t;t\ge 0)$ by*

$$V_t = \inf\left\{r\ge 0 : \int_0^r Z_u\,du > t\right\}. \tag{6.9}$$

Then

$$\left(\int_0^{V_t} (Z_u-\delta u)\,Z_u\,du; t\ge 0\right) \stackrel{d}{=} \left(xt+\int_0^t X_u\,du; t\ge 0\right).$$

It may not be immediately clear that Theorem 6.1.2 is related to the integral identity in (6.1). We hope to illuminate the connection with this example, based on the work by Duquesne [72]. Consider the sequences of measures $(\mu_n;n\ge 1)$ and $(\nu_n;n\ge 1)$ with $\mu_n=\mu=\mathrm{Poisson}(1)$ and $\nu_n=\nu=\mathrm{Poisson}(\delta)$ for each $n\ge 1$. For these measures Assumption 2 is clearly satisfied with the $\gamma_n=n$, and the Lévy process X a Brownian motion B.

We let $\mathfrak{f}^{(n)}\sim \mathrm{GWI}_{[nx]}(\mu,\nu)$. As shown by Duquesne [72], there exists an encoding (in a depth-first manner) of the non-mutant vertices in $\mathfrak{f}^{(n)}$ by a height process $H^n=(H^n_j;j=0,1,\cdots)$ such that

$$\#\{w\in\mathfrak{f}^{(n)}:\mathbf{ht}(w)=h,\ w\text{ is non-mutant}\} = \#\{j\ge 0: H^n_j=h\}.$$

Moreover, Duquesne shows that

$$\left(\frac{1}{n}H^n_{[n^2t]};t\ge 0\right)\xrightarrow{(d)}\left(\overleftarrow{H}_t;t\ge 0\right),$$

where

$$\overleftarrow{H}_t = 2|W_t| + \frac{1}{\delta}\left(L^0_t(W)-x\right)_+,$$

for a Brownian motion W with its local time at level zero and time t being $L_t^0(W)$ and $(x)_+:=\max(0,x)$. Duquesne [72] also shows that the process $\overleftarrow{H}$ possesses a jointly measurable (jointly continuous in this situation [156, Theorem VI.1.7]) family of local times $L(\overleftarrow{H})=(L^y_t(\overleftarrow{H});t\ge 0,y\ge 0)$ satisfying the occupation time formula

$$\int_0^t g(\overleftarrow{H}_r)\,dr = \int_0^\infty g(y)L^y_t(\overleftarrow{H})\,dy,\qquad \forall g\in C_c(\mathbb{R}_+)\quad \text{a.s.}$$

We remark that we **always** take this definition of local time even if the process is a semi-martingale with quadratic variation not identically t, which is the case in this present example.

Then Theorem 6.1.2, Lemma 6.1.3 and the Ray-Knight theorem [72] imply the following corollary:

Corollary 6.1.4. *Define* $V = (V_t; t \geq 0)$ *by*

$$V_t = \inf\left\{ r \geq 0 : \int_0^r L_\infty^y(\overleftarrow{H})\, dy > t \right\}.$$

Then

$$\left(\delta \int_0^\infty \overleftarrow{H}_r 1_{[\overleftarrow{H}_r \leq V_t]}\, dr - \int_0^{V_t} \left(L_t^y(\overleftarrow{H}) \right)^2 dy; t \geq 0 \right) \stackrel{d}{=} \left(-xt - \int_0^t B_s\, ds; t \geq 0 \right). \quad (6.10)$$

In particular, for each t,

$$\delta \int_0^\infty \overleftarrow{H}_r 1_{[\overleftarrow{H}_r \leq V_t]}\, dr - \int_0^{V_t} \left(L_t^y(\overleftarrow{H}) \right)^2 dy \stackrel{d}{=} \mathscr{N}\left(-xt, \frac{t^3}{3} \right).$$

This corollary can be generalized further with Corollary 6.3.3, which includes examples of processes $\overleftarrow{H}$ where the right-hand side is a spectrally positive Lévy process with certain additional constraints. The integral relationship in equation (6.10) is analogous to the Gorin-Skholnikov identity in the sense that both relate a linear combination of the area under a curve and the integral of the squared local time to a normal distribution. Of course, the formulation in Corollary 6.1.4 gives a Gaussian process and not just a single normal distribution. This process-level generalization is further studied using stochastic calculus with Theorem 6.1.5.

6.1.2 Stochastic Calculus Approach

While Corollary 6.1.4 follows from Lemma 6.1.3 it can also be obtained by stochastic calculus without the appeal to Lamperti transform in [49]. Similar methods were used by Hariya [102] and Lamarre and Shkolnikov [126] to obtain the normality results in (6.1) and (6.4). We extend their methods in this paper.

Suppose the Lévy process in Lemma 6.1.3 is of the form $X_t = \sigma B_t + at$ for $\sigma > 0$ and $a \in \mathbb{R}$. Then an elementary calculation yields $xt + \int_0^t (\sigma B_s + as)\, ds$ is a Gaussian process with mean $\mu_t = xt + \frac{a}{2}t^2$ and covariance function $\Gamma(t_1, t_2) = \sigma^2 \left(\frac{t_1^2 t_2}{2} - \frac{t_1^3}{6} \right)$ for $t_1 \leq t_2$. We can extend this to a more general Gaussian structure. To do this examine the stochastic differential equation:

$$\begin{aligned} dZ_v &= g(C_v)\sqrt{Z_v}\, dW_v + (c + f(C_v)Z_v)\, dv, \quad && Z_0 = x > 0 \\ dC_v &= Z_v\, dv && C_0 = 0 \end{aligned} \tag{6.11}$$

where g is a function such that g^2 is Lipschitz and $0 < \varepsilon < g(t) < M < \infty$, c is a constant such that $c \geq \frac{1}{2}\sup_t g(t)$ and f is a continuous function and W is a standard Brownian motion. In Section 6.5.2 we prove the weak existence of a solution to (6.11) by a sequence of time-changes.

Theorem 6.1.5. *Suppose that (Z, C) is a weak solution to the stochastic differential equation in (6.11). Let $V_t = \inf\{r \geq 0 : C_r > t\}$. Define the process $X = (X_t; t \geq 0)$ by*

$$X_t = \int_0^{V_t} (Z_v - cv) Z_v\, dv.$$

Then X is a Gaussian process with continuous sample paths. Its mean is given by

$$\mu_t = xt + \int_0^t (t - s) f(s)\, ds$$

and its covariance function is given by

$$\Gamma(t_1, t_2) = \int_0^{t_1} (t_2 - s)(t_1 - s) g^2(s)\, ds, \qquad t_1 \leq t_2.$$

We can also study stochastic differential equations of the form

$$\begin{aligned} dZ_v &= a\sqrt{Z_v}\, dW_v + \left(c + f(C_v)Z_v - \frac{Z_v^2}{1 - C_v} \right) dv, \qquad && Z_0 = x \geq 0 \\ dC_v &= Z_v\, dv, && C_0 = 0, \end{aligned} \tag{6.12}$$

for a standard Brownian motion W, constants $c > \frac{1}{2}a^2$ and f a continuous function. As shown by Pitman [153], the terminal local time of a reflected Brownian bridge $(L_1^v; v \geq 0)$ conditioned on the event $L_1^0 = x$ is a weak solution to the stochastic differential equation

$$dZ_v = 2\sqrt{Z_v}\, dW_v + \left(4 - \frac{Z_v^2}{1 - \int_0^v Z_s\, ds} \right) dv \qquad Z_0 = x.$$

In Section 6.5.1 we prove the weak existence of a solution to (6.12). See also Leuridans' work [135] on the conditioned Ray-Knight theorem. In Section 6.6 we extend Lemma 6.1.3 to the SDEs in (6.12). We prove the following theorem:

Theorem 6.1.6. *Suppose that (Z, C) is a weak solution to the stochastic differential equation in (6.12). Then*

$$\int_0^\infty \left\{2Z_v^2 - cvZ_v\right\} \, dv \stackrel{d}{=} \mathscr{N}\left(\mu, \sigma^2\right),$$

where $\mu = x + \int_0^1 (1-s)f(s)\, ds$ and $\sigma^2 = \dfrac{a^2}{3}$.

The above theorem has the following corollary, due to [97] when $x = 0$ and [126] when $x > 0$. See also [102].

Corollary 6.1.7. *Let $B^{|\mathrm{br}|} = (B_t^{|\mathrm{br}|}; t \in [0,1])$ denote a reflected Brownian bridge and let $L = (L_t^v; t \in [0,1], v \geq 0)$ denote its local time profile. Then, conditionally on $L_1^0 = x \geq 0$,*

$$\int_0^1 B_t^{|\mathrm{br}|}\, dt - \frac{1}{2}\int_0^\infty \left(L_1^v\right)^2 \, dv \stackrel{d}{=} \mathscr{N}\left(-\frac{x}{4}, \frac{1}{12}\right).$$

6.1.3 Another Example

We do not claim that this is the best discrete interpretation of the Gorin-Shkolnikov identity found in [97] and its generalization in [126]. However, the model presented in this article presents a generalization to a wide class of random forests. There are other discrete models which are asymptotically related to Brownian excursions and Brownian bridges. We briefly make note of one such example, which may be of use in order to understand A_β for other values $\beta \in (0, \infty) \setminus \{2\}$.

It is well known that a Brownian excursion can be constructed from a Brownian bridge by placing the origin at the bridges absolute minimum, see [174]. One extension of the result can be found in the work of Chassaing and Janson [52]. In their analysis, they use another combinatorial object of study called parking functions. Certain classes of parking functions are in bijection with labeled trees on the vertex set $\{0, 1, \cdots, n\}$ rooted at 0. See [53] and references therein for more details on the specifics of the bijection. In the discrete setting, the analog of the average height of a vertex in a rooted tree is the so-called displacement for parking functions, and the average number of cousin vertices is the average number of "cars" that, at some point, want to park in a certain parking space. In this work we do not endeavor to prove weak limits for these discrete analogs; however, we do state what the limiting objects should be. Going back to the belief that there may be an "interesting combinatorial interpretation" of the quantity A_2, parking functions may be another way to find a combinatorial interpretation of the results.

In this section we deviate from the notation in the last section. We let $B^{\mathrm{br}} = (B^{\mathrm{br}}_t; t \in [0,1])$ denote a Brownian bridge, and let $e = (e_t; t \in [0,1])$ denote a standard Brownian excursion. We take the (deterministic) periodic extensions of these functions so that they are defined on all of $\mathbb{R}$. Fix an $x \geq 0$, define the following processes

$$X^{(1)}_t = B^{\mathrm{br}}_t - xt + \sup_{t-1 \leq s \leq t} \left(xs - B^{\mathrm{br}}_s \right)$$

$$X^{(2)}_t = e_t - xt + \sup_{0 \leq s \leq t} \left(xs - e_s \right).$$

The process $X^{(2)}$ is just a reflected Brownian excursion with negative drift $-x$.

We now let X denote either of the processes above. The occupation measure of X is equal in law to the occupation measure of a reflected Brownian bridge conditioned on its local time at zero being $x/2$. See [52, Cor. 2.3]. The presence of the factor of $\frac{1}{2}$ comes from the different conventions for the local time at 0 used in [153] and [52]. Since, almost surely the occupation measure for a reflected Brownian bridge has a continuous density with respect to the Lebesgue measure, the same holds for the occupation measure of X. Thus we get the following corollary of Corollary 6.1.7.

Corollary 6.1.8. *Let X denote either $X^{(1)}$ or $X^{(2)}$ above. Let $(L^v_1; v \geq 0)$ denote a continuous version of the density of the occupation measure for X. Then,*

$$\int_0^1 X_t \, dt - \frac{1}{2} \int_0^\infty (L^v_1)^2 \, dv \overset{d}{=} \mathscr{N}\left(-\frac{x}{8}, \frac{1}{12} \right).$$

Parking functions and random forests are not the only combinatorial models that have asymptotics related to the Brownian excursion or Brownian and their local times. As mentioned previously, numerous quantities in graph enumerations are related to either a Brownian excursion or its integrals, see [105] and references therein. See also [15] for the connection between these processes and the asymptotics of uniform random mappings.

6.1.4 Overview of The Paper

In Section 6.2.1, we discuss the discrete forest model underlying our weak convergence results in Theorem 6.1.2. This model is based on a random forest model used by Aldous [8] and

Duquesne [72]. Afterward, in Section 6.2.2 we discuss in more detail the various processes defined on the random forest. In Sections 6.2.3 and 6.2.4 we discuss Lévy processes and continuous state branching processes respectively. Importantly, we describe the path-by-path relationship between these two classes from [49]. Section 6.2.5 contains information on Ψ-height processes and their Ray-Knight theorems.

After discussing these preliminaries we move to prove integral relationships for Lévy processes and their relationships with continuous state branching processes with immigration in Section 6.3. This brief discussion allows us to prove Corollary 6.3.3, which is the continuum random tree generalization of the Gorin-Shkolnikov identity.

In Section 6.4 we prove our weak convergence results which rely on the integral results in Section 6.3. The proofs of Theorems 6.1.1 and 6.1.2 are found in Section 6.4.1. In Section 6.6 we prove the normality results in Theorems 6.1.5 and 6.1.6 after studying the properties of solutions to (6.11) or (6.12) in Section 6.5.

Acknowledgments

The author thanks David Aldous for a helpful discussion on the random tree interpretation of Jeulin's identity. The author thanks Soumik Pal for continued guidance during this project and providing helpful comments on clarifying several arguments in the paper. I would also like to thank the two anonymous referees for their careful reading and detailed comments on the presentation and proofs, including an idea for simplifying the proof of Lemma 6.6.1.

6.2 Preliminaries

6.2.1 Forest Constructions

In this section we describe the forest "picture" underlying some of our weak convergence arguments later. This model is based on a model used by Aldous in [8] and then by Duquesne in [72]. However, since we will not be proving convergence of the contour functions, we do use a labeling which is convenient for our analysis and is not useful for any type of genealogical analysis which can be found in [72].

In this work, a forest will mean a locally finite planar graph on the vertex set

$$\{w_j, m_j : j = 0, 1, \cdots\}$$

which has a finite number of connected components which themselves are rooted planar trees. The vertices of the form w_j will be called the *non-mutant vertices* (depicted in Figure 6.2 as black vertices) and the vertices of the form m_j will be called *mutant vertices* (depicted as the white unlabeled vertices in Figure 6.2). When referring to a forest $\mathfrak{f}$, we will let $\mathscr{V}(\mathfrak{f})$ denote the vertex set of $\mathfrak{f}$ and $\mathscr{E}(\mathfrak{f})$ denote the edge set.

We will construct our random forests in a breadth-first approach from two given sequences of non-negative integers $(\chi_j; j = 0, 1, \cdots)$ and $(\eta_j; j = 1, 2, \cdots)$ along with a number $k = 0, 1, 2, \cdots$. At each height there will be two types of vertices, mutant and non-mutant vertices. There will only be one mutant vertex in each generation, and will be omitted from most of the objects we count on our forest. Children of mutant vertices will appear last in the breadth-first ordering when restricting to each height.

Each χ_j will represent the number of offspring the j^{th} non-mutant vertex will have. Each η_j will represent the number of immigrants that will appear in generation j. The immigrants will be non-mutant offspring of the mutant vertex from the previous generation.

We start with defining $\mathfrak{f}_0$ to be the graph with vertex set

$$\mathscr{V}(\mathfrak{f}_0) = \{w_0, \ldots, w_{k-1}, m_0\}$$

and edge set $\mathscr{E}(\mathfrak{f}_0) = \emptyset$. The vertex m_0 will be the mutant vertex. These vertices will be the roots of the random forest, and will consequently be the vertices of height 0. For each $h \geq 1$ we construct $\mathfrak{f}_h$ from $\mathfrak{f}_{h-1}$ by adding vertices of height h. There are two cases which we must consider: (1) there are no non-mutant vertices at height $h-1$ in $\mathfrak{f}_{h-1}$ or (2) the non-mutant vertices at height $h-1$ in $\mathfrak{f}_{h-1}$ are $\{w_j : j = a, a+1, \cdots, b-1\}$ for some natural numbers $a < b$.

In Case (1), we add η_h non-mutant vertices and a single mutant vertex as offspring of the mutant vertex m_{h-1}. The collection of non-mutant vertices of $\mathfrak{f}_{h-1}$ is of the form

$$\{w_j : j = 0, 1, \cdots, b-1\}$$

for some $b \geq 0$. When $b = 0$ the above set is empty. We let $\mathfrak{f}_h$ be the graph with vertex set $\mathscr{V}(\mathfrak{f}_h) := \mathscr{V}(\mathfrak{f}_{h-1}) \cup \{w_j : j = b, b+1, \cdots, b+\eta_h - 1\} \cup \{m_h\}$ and the edge set

$$\mathscr{E}(\mathfrak{f}_h) := \mathscr{E}(\mathfrak{f}_{h-1}) \cup \{(m_{h-1}, w_j) : j = b, b+1, \cdots, b+\eta_h - 1\} \cup \{(m_{h-1}, m_h)\}.$$

The height of the vertices of $\mathfrak{f}_h$ will be the distance to the root, which inductively implies for all $v \in \mathfrak{f}_h \setminus \mathfrak{f}_{h-1}$ that $\mathbf{ht}(v) = h$. We remark that if $\eta_h = 0$ then the only vertex in $\mathfrak{f}_h \setminus \mathfrak{f}_{h-1}$ is the vertex m_h.

In Case (2), we add χ_j non-mutant vertices as offspring for each non-mutant vertex w_j in generation $h-1$. We also add η_h non-mutant vertices and a single mutant vertex as the offspring of m_{h-1}. We go into more detail in order to describe the breadth-first labeling. There are non-mutant vertices at height $h-1$ in $\mathfrak{f}_{h-1}$. These vertices will be of the form $\{w_j : j = a, \cdots, b-1\}$ for some $a < b$. Each vertex w_j gives birth to χ_j children at height h and the vertex m_{h-1} will give birth to η_h non-mutant children and 1 mutant child. Thus the vertex set for $\mathfrak{f}_h$ will be

$$\mathscr{V}(\mathfrak{f}_h) = \mathscr{V}(\mathfrak{f}_{h-1}) \cup \left\{ w_j : j = b, b+1, \cdots, b + \sum_{i=a}^{b-1} \chi_i + \eta_h - 1 \right\} \cup \{m_h\}.$$

Moreover, for each $\ell = 0, 1, \ldots, b-a-1$ we add edges between a vertex $w_{a+\ell}$ and w_j for $j = b + \sum_{i=a}^{a+\ell-1} \chi_i, b + \sum_{i=a}^{a+\ell-1} \chi_i + 1, \cdots, b + \sum_{i=a}^{a+\ell-1} \chi_i + \chi_{a+\ell} - 1$. We also add edges between m_{h-1} and m_h along with w_j for $j = b + \sum_{i=a}^{b-1} \chi_i, b + \sum_{i=a}^{b-1} \chi_i + 1, \cdots, \sum_{i=a}^{b-1} \chi_i - 1 + \eta_h$. Again, the vertices in $\mathfrak{f}_h \setminus \mathfrak{f}_{h-1}$ have distance h from a root.

We continue this process ad infinitum, and define $\mathfrak{f} := \bigcup_{h \geq 0} \mathfrak{f}_h$. By construction, for $i < j$ we must have $\mathbf{ht}(w_i) \leq \mathbf{ht}(w_j)$.

Definition 2. A forest $\mathfrak{f}$ constructed as above starting from k non-mutant vertices is called **Galton-Watson immigration forest** (GWI forest for short) when the integer sequence $(\chi_j; j = 0, 1, \cdots)$ are i.i.d. and the sequence $(\eta_h; h = 1, 2, \cdots)$ are i.i.d. and are independent of (χ_j). We call the common distribution, say μ, of (χ_j) the offspring distribution, and we call the common distribution, say ν, of (η_h) the immigration rate. We denote the law of $\mathfrak{f}$ by $\mathrm{GWI}_k(\mu, \nu)$.

6.2.2 Processes defined on the forest

In this section we define several processes on the forest, and describe some of the relationships between the various processes and the breadth-first labeling of the vertices. Let $\mathfrak{f}$ be a forest constructed as above. We recall that the forest $\mathfrak{f}$ is rooted, and that the height of a vertex is defined as the distance from the vertex to a root in the graph distance.

Define the *height profile* of the forest $\mathfrak{f}$ as the process $Z^{\mathfrak{f}} = \left(Z^{\mathfrak{f}}_h : h \in \mathbb{N}_0\right)$ by

$$Z^{\mathfrak{f}}_h = \#\{\text{non-mutant vertices } v \in \mathfrak{f} \text{ with } \mathbf{ht}(v) = h\}.$$

This counts the number of non-mutant vertices at height h in the forest $\mathfrak{f}$. We also define the *cumulative height profile* of $\mathfrak{f}$ as the process $C^{\mathfrak{f}} = (C^{\mathfrak{f}}_h; h \in \mathbb{N}_0)$ by

$$C^{\mathfrak{f}}_h = \sum_{j=0}^{h} Z^{\mathfrak{f}}_j,$$

and its right-continuous inverse $V^{\mathfrak{f}} = (V^{\mathfrak{f}}_r; r = 0, 1, \cdots)$ by

$$V^{\mathfrak{f}}_r = \inf\{h \geq 0 : C^{\mathfrak{f}}_h > r\}.$$

We observe that $C^{\mathfrak{f}}_h$ is the index of the first vertex at height $h+1$. This follows easily from the convention that we start indexing at 0. Hence, $V^{\mathfrak{f}}_r = \mathbf{ht}(w_r)$ and so $Z^{\mathfrak{f}}_{V^{\mathfrak{f}}_r}$ is the width of the layer of the forest containing the vertex w_r.

Lastly, we refer the reader back to (6.8) for a definition of the cumulative breadth-first cousin process $K^{\mathfrak{f}}$ and the cumulative breadth-first height process $J^{\mathfrak{f}}$.

6.2.3 Lévy processes

We will provide a brief overview of spectrally positive Lévy processes. More details and proofs of the statements below can be found in Chapter VII of Bertoin's monograph [26].

A (possibly killed) spectrally positive Lévy process $X = (X_t; t \geq 0)$ is a Lévy process which contains no negative jumps. Its Laplace transform exists and uniquely characterizes the process X. The Laplace transform must be of the form

$$\mathbb{E}\left[\exp(-\lambda X_t)\right] = \exp(t\Psi(\lambda)) \qquad \forall \lambda > 0. \tag{6.13}$$

Moreover, the function Ψ must be of the form

$$\Psi(\lambda) = -\kappa + \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} \left(e^{-\lambda r} - 1 + \lambda r 1_{[r<1]}\right) \pi(dr) \tag{6.14}$$

where $\kappa, \beta \geq 0$, $\alpha \in \mathbb{R}$ and π is a Radon measure on $(0, \infty)$ such that $\int_{(0,\infty)} (1 \wedge r^2)\, \pi(dr) < \infty$. Conversely, for each such Ψ, there exists a spectrally positive Lévy process with such a Laplace transform.

A particular class of spectrally positive Lévy processes are subordinators. These are Lévy processes with increasing sample paths. A subordinator Y has a Laplace transform of the form

$$\mathbb{E}\left[\exp(-\lambda Y_t)\right] = \exp(-t\Phi(\lambda)), \qquad \forall \lambda > 0, \tag{6.15}$$

where Φ is of the form

$$\Phi(\lambda) = \kappa' + \alpha'\lambda - \int_{(0,\infty)} (e^{-\lambda r} - 1)\, \nu(dr) \tag{6.16}$$

with $\kappa', \alpha' \geq 0$ and ν is a Radon measure with $\int_{(0,\infty)} (1 \wedge r)\, \nu(dr) < \infty$. In our work, we concern ourselves with the case where $\Phi(\lambda) = \delta\lambda$ for some $\delta > 0$, which makes the subordinator $Y_t = \delta t$.

6.2.4 Continuous state branching processes

Continuous state branching processes with immigration (CBI processes) arise as scaling limits of discrete Galton-Watson processes, see [114]. A CBI process $Z = (Z_t; t \geq 0)$ is a Feller process on $[0, \infty]$ which is absorbed at ∞. We denote by $\mathbb{E}_x[-]$ the expectation conditionally given $Z_0 = x \geq 0$. As shown by [114], the Laplace transform of Z is of the form

$$\mathbb{E}_x\left[\exp\{-\lambda Z_t\}\right] = \exp\left[-xu(t,\lambda) - \int_0^t \Phi(u(s,\lambda))\, ds\right], \qquad \forall \lambda > 0$$

where u is the unique non-negative solution to the integral equation

$$u(t,\lambda) + \int_0^t \Psi(u(s,\lambda))\, ds = \lambda,$$

for functions Ψ and Φ. The function Ψ is called the branching mechanism and must be of the form (6.14) and the function Φ is called the immigration rate and must be of the form

(6.16). Conversely, given any two such functions Ψ and Φ, there exists a CBI process with branching mechanism Ψ and immigration rate Φ. For simplicity, we will use $\mathrm{CBI}_x(\Psi, \Phi)$ to refer to the law of a CBI process starting from $x \geq 0$ and satisfies the above Laplace transform.

By [114], we know that continuous state branching processes with immigration are in one-to-one correspondence with pairs of Lévy processes (X, Y) satisfying (6.13) and (6.15). The bijection described there is in terms of the Laplace transforms of the respective processes. A path-wise identification does exist, thanks to the work of Caballero, Pérez Garmendia and Uribe Bravo [49]. As mentioned previously, the authors of [49] show that if X is a spectrally positive Lévy process with Laplace exponent $(-\Psi)$ and Y is an independent subordinator with Laplace exponent Φ then a càdlàg solution to (6.6) exists, is unique, and is a $\mathrm{CBI}_x(\Psi, \Phi)$ process. When Y is identically 0, the Lévy process X in (6.6) is stopped upon hitting $-x$. The path-wise relationship when $Y = 0$ was observed by Lamperti [128], although proved later by Silverstein [163].

6.2.5 The Ψ-height process

Height processes were introduced by Le Gall and Le Jan in [133], and further examined in [73, 132] and we refer back to Section 2.2.4.3 for a more detailed summary of their properties. Recall from therein that we can define so-called Ψ-height processes under the more restrictive assumptions:

$$\Psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (e^{-\lambda r} - 1 + \lambda r)\, \pi(dr) \tag{6.17}$$

with $\alpha, \beta \geq 0$ and π a Radon measure with the stronger integrability condition $\int_0^\infty (r \wedge r^2)\, \pi(dr) < \infty$; the assumption that Ψ is conservative, i.e. it satisfies equation (6.7); and that further Ψ satisfies

$$\int_1^\infty \frac{1}{\Psi(u)}\, du < \infty. \tag{6.18}$$

All of these assumptions on Ψ are slight restrictions on a general theory, but they imply that a Lévy process X with Laplace exponent $(-\Psi)$ has paths of infinite variation, non-negative jumps, and does not drift towards $+\infty$.

Under the above conditions (6.17), and (6.7) there exists a process H which encodes a forest of Lévy trees which has a continuous modification if and only if (6.18) holds. See [73, Theorem 1.4.3] for the continuity statement. Whenever (6.18) is satisfied, we will assume that H is this modification. For more information on height processes see [73].

The process H possesses a family of local times $L = (L_t^a; t \geq 0, a \geq 0)$ which almost surely satisfies the occupation density formula:

$$\int_0^t g(H_r)\, dr = \int_0^\infty g(a) L_t^a \, da, \qquad \forall g \in C_c(\mathbb{R}),\, t \geq 0. \tag{6.19}$$

See [73, Proposition 1.3.3]. There is also a Ray-Knight theorem for these processes as shown in [73, Theorem 1.4.1]. Namely, define $T_x = \inf\{t : L_t^0 = x\} = \inf\{t : X_t = -x\}$ then

$$\left(L_{T_x}^a; a \geq 0\right) \sim \mathrm{CBI}_x(\Psi, 0).$$

Similar Ray-Knight theorems were obtained by Warren [178] involving sticky Brownian motion.

The Ray-Knight theorem above was extended by Lambert [127] and Duquesne [72] to allow for some immigration. Lambert's work is slightly more general; however Duquesne's work contains some better approximation results for the local time. For $\delta > 0$ and $x \geq 0$ define the left-height process $\overleftarrow{H} = (\overleftarrow{H}_t; t \geq 0)$ by

$$\overleftarrow{H}_t = H_t + \frac{1}{\delta}(L_t^0 - x)_+. \tag{6.20}$$

By [72], there exists a jointly measurable family of local times $\overleftarrow{L} = \left(\overleftarrow{L}_t^a; t \geq 0, a \geq 0\right)$ which is continuous and increasing in t and satisfies equation (6.19) with $\overleftarrow{H}$ and $\overleftarrow{L}$ replacing H and L. Since $\overleftarrow{H}_t \to \infty$ as $t \to \infty$, the limit $\overleftarrow{L}_\infty^a := \lim_{t\uparrow\infty} \overleftarrow{L}_t^a$ is finite almost surely due to (6.7). Moreover, the Ray-Knight theorem, see [72, Theorem 1.2, Remark 3.2] and [127, Theorem 5], states

$$\left(\overleftarrow{L}_\infty^a; a \geq 0\right) \sim \mathrm{CBI}_x(\Psi, \Phi), \qquad \Phi(\lambda) = \delta\lambda. \tag{6.21}$$

In the case where the branching process is a squared Bessel process and the height process is a reflected Brownian, a similar result was shown in [134].

6.3 Integral Relationships for CBIs

In this section we describe various properties of continuous state branching processes and their implications for random trees. We will mostly work under the assumption that Ψ is a conservative branching mechanism; however, when discussing the implications for left-height processes we must work under slightly stricter assumptions. We recall that Ψ being conservative is equivalent, see [99], to the almost sure finiteness of a $\mathrm{CBI}_x(\Psi, 0)$ process.

Let us start by gathering certain properties of processes related to a $\mathrm{CBI}_x(\Psi, \Phi)$ process Z. If $\Phi(\lambda) = \delta\lambda$ for some $\delta > 0$ then, by Theorem 2 in [49], we can and do assume Z to be the unique càdlàg solution to

$$Z_t = x + X_{C_t} + \delta t, \qquad C_t = \int_0^t Z_s \, ds, \tag{6.22}$$

where X is a Lévy process with Laplace exponent $(-\Psi)$. Moreover, using this representation of Z along with Lemma 3 and Corollary 5 in [49] and the strong Markov property for Z the following lemma is easy to see.

Lemma 6.3.1. *Let Z be a $\mathrm{CBI}_x(\Psi, \Phi)$ process where Ψ is a conservative branching mechanism (i.e. satisfies (6.7) and (6.14)) and $\Phi(\lambda) = \delta\lambda$ for some $\delta > 0$. Then $C_t := \int_0^t Z_s \, ds$ is strictly increasing and almost surely $C_t \to \infty$ as $t \to \infty$.*

The above lemma tells us that the process V defined in equation (6.9) is actually the two-sided inverse of C. Moreover, since C diverges towards infinity almost surely, the value of V_r is finite for each $r \geq 0$.

We now move to the proof of Lemma 6.1.3.

Proof of Lemma 6.1.3. We assume that we are working on a probability space where Z has the path-by-path representation in (6.22), for a Lévy process X. The change of variable formula implies

$$\int_0^{V_r} F(u, Z_u) \, dC_u = \int_0^r F(V_u, Z_{V_u}) \, du$$

for locally bounded functions F. Moreover, since $dC_u = Z_u \, du$, we can claim

$$\int_0^{V_r} (Z_u - \delta u) Z_u \, du = \int_0^r (Z_{V_u} - \delta V_u) \, du.$$

However, by (6.22) and the fact that V is the two-sided inverse of C,

$$Z_{V_u} = x + X_{C_{V_u}} + \delta V_u = x + X_u + \delta V_u.$$

The result desired claim now easily follows. □ □

In fact the above proof easily implies the following corollary as well.

Corollary 6.3.2. *Under the conditions for Lemma 6.1.3 and for any* $n = 1, 2, \cdots$

$$\int_0^{V_r} (Z_u - \delta u)^n Z_u \, du \stackrel{d}{=} \int_0^r (x + X_u)^n \, du,$$

where X is a Lévy process with Laplace exponent $(-\Psi)$.

6.3.1 The continuum random tree interpretation

Here we must strengthen the assumptions put onto Ψ in order to guarantee that there is a continuous Ψ-height process. As discussed in Section 6.2.5, we require that Ψ satisfies (6.7), (6.17) and (6.18).

Let $\overleftarrow{H} = (\overleftarrow{H}_t; t \geq 0)$ be defined as in (6.20) and let $\overleftarrow{L} = (\overleftarrow{L}_t^a; a \geq 0, t \geq 0)$ denote its local time. We know from (6.21) that the process $\left(\overleftarrow{L}_\infty^a; a \geq 0\right)$ is a $\text{CBI}_x(\Psi, \Phi)$ process where $\Phi(\lambda) = \delta\lambda$. Define $V_r = \inf\left\{x \geq 0 : \int_0^x \overleftarrow{L}_\infty^y \, dy > r\right\}$ and then Lemma 6.1.3 implies

$$\left(\delta \int_0^{V_r} a \overleftarrow{L}_\infty^a \, da - \int_0^{V_r} \left(\overleftarrow{L}_\infty^a\right)^2 da; r \geq 0\right) \stackrel{d}{=} \left(-xr - \int_0^r X_u \, du; r \geq 0\right).$$

However, the occupation time formula implies that almost surely the left-most integral above can be recognized as

$$\int_0^{V_r} a \overleftarrow{L}_\infty^a \, da = \int_0^\infty \overleftarrow{H}_t 1_{[\overleftarrow{H}_t \leq V_r]} \, dt.$$

Hence, we can conclude the following:

Corollary 6.3.3. *Let $\overleftarrow{H}$ be defined as in* (6.20) *with Ψ satisfying* (6.7), (6.18) *and* (6.17) *and let $\overleftarrow{L}$ denote its local time. Then*

$$\left(\delta \int_0^\infty \overleftarrow{H}_t 1_{[\overleftarrow{H}_t \leq V_r]} - \int_0^{V_r} \left(\overleftarrow{L}_\infty^a\right)^2 da; r \geq 0\right) \stackrel{d}{=} \left(-xr - \int_0^r X_u \, du; r \geq 0\right)$$

where X is a Lévy process with Laplace exponent $(-\Psi)$.

6.4 Weak Convergence

Throughout this section we will let $(\mathfrak{f}^{(n)}; n \geq 1)$ denote a sequence of forests where each $\mathfrak{f}^{(n)}$ is a $\mathrm{GWI}_{[nx]}(\mu_n, \nu_n)$-forest. We also assume that μ_n and ν_n satisfy Assumption 2, and $(\gamma_n; n \geq 1)$ is the sequence of integers specified therein.

With this we can now prove the following joint convergence lemma:

Lemma 6.4.1. *If μ_n and ν_n satisfy Assumption 2, then the following joint convergence holds in the product of the Skorokhod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})^3$*

$$\left(\left(\frac{1}{n}Z^{*,n}_{[\gamma_n s]}\right)_{s\geq 0}, \left(\frac{1}{n\gamma_n}C^{*,n}_{[\gamma_n t]}\right)_{t\geq 0}, \left(\frac{1}{\gamma_n}V^{*,n}_{[n\gamma_n r]}\right)_{r\geq 0}\right)$$
$$\overset{(d)}{\Longrightarrow} \left((Z_s)_{s\geq 0}, \left(\int_0^t Z_s\,ds\right)_{t\geq 0}, \left(\inf\left\{t : \int_0^t Z_s\,ds > r\right\}\right)_{r\geq 0}\right), \tag{6.23}$$

where Z is a $\mathrm{CBI}_x(\Psi, \Phi)$ process.

Proof. We remark that by [72, Theorem 1.4], the convergence of rescaling of $Z^{*,n}$ converges to a $\mathrm{CBI}_x(\Psi, \Phi)$ process, i.e.

$$\left(\frac{1}{n}Z^{*,n}_{[\gamma_n s]}; s \geq 0\right) \overset{(d)}{\Longrightarrow} (Z_s; s \geq 0).$$

We also observe

$$\begin{aligned}
\frac{1}{n\gamma_n}C^{*,n}_{[\gamma_n t]} &= \frac{1}{n\gamma_n}\sum_{h=0}^{[\gamma_n t]} Z^{*,n}_h \\
&= \frac{1}{\gamma_n}\int_0^{[\gamma_n t]+1} \frac{1}{n}Z^{*,n}_{[u]}\,du \\
&= \int_0^{([\gamma_n t]+1)/\gamma_n} \frac{1}{n}Z^{*,n}_{[\gamma_n s]}\,ds.
\end{aligned}$$

Since $([\gamma_n t] + 1)/\gamma_n \to t$ locally uniformly, the joint convergence

$$\left(\left(\frac{1}{n}Z^{*,n}_{[\gamma_n s]}\right)_{s\geq 0}, \left(\frac{1}{n\gamma_n}C^{*,n}_{[\gamma_n t]}\right)_{t\geq 0}\right)$$
$$\overset{(d)}{\Longrightarrow} \left((Z_s)_{s\geq 0}, \left(\int_0^t Z_s\,ds\right)_{t\geq 0}\right),$$

will follow once we argue the continuity of the map $f \mapsto (f, \int_0^\cdot f(s)\,ds)$.

To see that $f \mapsto \int_0^\cdot f(s)\,ds$ is continuous we prove that $\mathbb{D}$ continuously embeds into L^1_{loc}. Indeed, suppose that $f_n \to f$ in $\mathbb{D}$ and let t be a continuity point of f. For $g \in \mathbb{D}(\mathbb{R}_+, \mathbb{R})$, set $\|g\|_{[0,t]} = \sup_{r\le t} |g(r)|$. Then there exists a sequence $\tau_n : [0,t] \to [0,t]$ such that $\|\tau_n - id\|_{[0,t]} \vee \|f\circ\tau_n - f_n\|_{[0,t]} \to 0$ as $n\to\infty$. Then

$$\begin{aligned}\int_0^t |f_n(s) - f(s)|\,ds &\le \int_0^t |f_n(s) - f(\tau_n(s))|\,ds + \int_0^t |f(\tau_n(s)) - f(s)|\,ds\\ &\le t\|f_n - f\circ\tau_n\|_{[0,t]} + \int_0^t |f(\tau_n(s)) - f(s)|\,ds.\end{aligned}$$

The first term can easily be handled since $f_n \to f$ in $\mathbb{D}$ and the second term converges to 0 as $n\to\infty$ by dominated convergence. This proves the continuity of the embedding $\mathbb{D} \hookrightarrow L^1_{\text{loc}}$. Since $f \mapsto \int_0^\cdot f(s)\,ds$ is a continuous map from $L^1_{\text{loc}}(\mathbb{R}_+) \to C(\mathbb{R}_+)$ and $C(\mathbb{R}_+)$ embeds continuously into $\mathbb{D}$, we have shown the desired continuity.

The convergence of the third coordinate in equation (6.23) follows from [180, Theorem 7.2] and Lemma 6.3.1. Indeed, define the set $E \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ to be the collection of functions f which are unbounded from above with $f(0) \ge 0$ and equip this set with the Skorokhod J_1 topology. Then the map $E \to E$ defined by $(f(t); t\ge 0) \mapsto (\inf\{t \ge 0 : f(t) > r\}; r \ge 0)$ is measurable and it is continuous on the set of strictly increasing functions. □ □

6.4.1 Proofs of Theorem 6.1.2 and Theorem 6.1.1

Proof of Theorem 6.1.2. Throughout this proof we refer to Z as a $\text{CBI}_x(\Psi,\Phi)$ process, the quantity $C_t = \int_0^t Z_s\,ds$ and V defined in (6.9). In what follows we sometimes write the index in the process as $Z^{*,n}(h)$ instead of as a subscript $Z^{*,n}_h$, and similar remarks hold for the other processes.

We recall that the first index of a vertex in $\mathfrak{f}^{(n)}$ at height $h+1$ is $C^{*,n}_h$. Therefore

$$\begin{aligned}K^{*,n}(C^{*,n}_h) &= \sum_{j=0}^{C^{*,n}_h - 1} \mathbf{csn}(w_j) = \sum_{\ell=0}^{h} Z^{*,n}_\ell (Z^{*,n}_\ell - 1)\\ &= \sum_{\ell=0}^{h} \left(Z^{*,n}_\ell\right)^2 - C^{*,n}_h.\end{aligned}$$

Similarly, we can see that

$$J^{*,n}(C^{*,n}_h) = \sum_{\ell=0}^{h} \ell Z^{*,n}_\ell.$$

Consequently,

$$\frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}_{[\gamma_n t]}) = \frac{1}{n^2\gamma_n}\sum_{\ell=0}^{[\gamma_n t]}\left(Z^{*,n}_\ell\right)^2 - \frac{1}{n^2\gamma_n}C^{*,n}_{[\gamma_n t]}$$
$$= \int_0^{([\gamma_n t]+1)/\gamma_n}\left(\frac{1}{n}Z^{*,n}_{[\gamma_n s]}\right)^2 ds - \frac{1}{n^2\gamma_n}C^{*,n}_{[\gamma_n t]}$$

and

$$\frac{1}{n\gamma_n^2}J^{*,n}(C^{*,n}_{[\gamma_n t]}) = \int_0^{([\gamma_n t]+1)/\gamma_n}\frac{[\gamma_n s]}{\gamma_n}\cdot\frac{1}{n}Z^{*,n}_{[\gamma_n s]}\,ds.$$

This easily implies the weak convergence in $\mathbb{D}(\mathbb{R}_+,\mathbb{R})^2$

$$\left(\left(\frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}_{[\gamma_n t]}); t\ge 0\right), \left(\frac{1}{n\gamma_n^2}J^{*,n}(C^{*,n}_{[\gamma_n t]}); t\ge 0\right)\right)$$
$$\overset{(d)}{\Longrightarrow}\left(\left(\int_0^t Z_s^2\,ds; t\ge 0\right), \left(\int_0^t sZ_s\,ds; t\ge 0\right)\right).$$

Moreover this convergence is joint with the convergence in (6.23), and hence by a lemma on pg. 151 in [35]

$$\left(\frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})) - \frac{\delta}{n\gamma_n^2}J^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})); r\ge 0\right)$$
$$\overset{(d)}{\Longrightarrow}\left(\int_0^{V_r}(Z_s-\delta s)Z_s\,ds; r\ge 0\right).$$

By Lemma 6.1.3 and Slutsky's theorem, the result follows if we can show

$$\left(\left|\frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})) - \frac{1}{n^2\gamma_n}K^{*,n}_{[n\gamma_n r]}\right|; r\ge 0\right)\overset{(d)}{\Longrightarrow}\mathbf{0} := (0; r\ge 0) \tag{6.24}$$

and

$$\left(\left|\frac{1}{n\gamma_n^2}J^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})) - \frac{1}{n\gamma_n^2}J^{*,n}_{[n\gamma_n r]}\right|; r\ge 0\right)\overset{(d)}{\Longrightarrow}\mathbf{0}.$$

To argue the above convergences, we observe that $C^{*,n}(V^{*,n}_r - 1)\le r < C^{*,n}(V^{*,n}_r)$. Since $K^{*,n}$ is increasing, we have

$$\left|\frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})) - \frac{1}{n^2\gamma_n}K^{*,n}_{[n\gamma_n r]}\right|$$
$$\le\left|\frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})) - \frac{1}{n^2\gamma_n}K^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]}-1))\right|$$
$$\le\frac{1}{n^2\gamma_n}\left(Z^{*,n}(V^{*,n}_{[n\gamma_n r]})\right)^2. \tag{6.25}$$

The presence of the square in the second inequality follows from the fact that

$$K^{*,n}(C^{*,n}(h)) - K^{*,n}(C^{*,n}(h-1)) = Z_h^{*,n} \cdot (Z_h^{*,n} - 1) \leq \left(Z_h^{*,n}\right)^2.$$

The desired convergence in (6.24) follows from

$$\left(\frac{1}{n^2\gamma_n}\left(Z^{*,n}(V^{*,n}_{[n\gamma_n r]})\right)^2; r \geq 0\right) \overset{(d)}{\Longrightarrow} \mathbf{0}.$$

The above convergence then holds by Lemma 6.4.1 and standard weak convergence arguments for time-changes (see e.g. a lemma on pg. 151 in [35]). Indeed, we have

$$\left(n^{-1}Z^{*,n}\left(V^{*,n}_{[n\gamma_n r]}\right); r \geq 0\right) \overset{(d)}{\Longrightarrow} \left(Z\left(\inf\{t : \int_0^t Z_s\,ds > r\}\right); r \geq 0\right). \tag{6.26}$$

Since $\gamma_n \to \infty$ by Assumption 2, the stated convergence to zero holds.

Reasoning as in (6.25) we get the upper bound

$$\left|\frac{1}{n\gamma_n^2}J^{*,n}(C^{*,n}(V^{*,n}_{[n\gamma_n r]})) - \frac{1}{n\gamma_n^2}J^{*,n}_{[n\gamma_n r]}\right| \leq \frac{1}{n\gamma_n^2}V^{*,n}_{[n\gamma_n r]} \cdot Z^{*,n}(V^{*,n}_{[n\gamma_n r]}).$$

By Lemma 6.4.1 and the convergence in (6.26), the following weak convergence holds

$$\left(\frac{1}{n\gamma_n^2}V^{*,n}_{[n\gamma_n r]} \cdot Z^{*,n}(V^{*,n}_{[n\gamma_n r]}); r \geq 0\right) \overset{(d)}{\Longrightarrow} \mathbf{0}.$$

The result now follows. □ □

Similar arguments can yield Theorem 6.1.1. However, the authors of [97] and [126] are interested in distributional properties of the quantity

$$\int_0^1 B_t^{|\mathrm{br}|,x} - \frac{1}{\beta}\int_0^\infty \left(L_{1,x}^v\right)^2\,dv, \quad \forall \beta > 0$$

where $B^{|\mathrm{br}|,x}$ is a reflected Brownian bridge conditioned on its local time at level zero and time 1 being exactly x and $L_{1,x}^v$ is the local time of $B^{|\mathrm{br}|,x}$ at time 1 and level v. Hence, we prove the following proposition, which by taking $\beta = 2$ and using Theorem 1.13 in [126] yields the formulation in Theorem 6.1.1.

Proposition 6.4.2. *Suppose that $\mathfrak{f}^{(n)}$ is a uniformly chosen rooted labeled forest on n vertices with k_n roots where $2k_n/\sqrt{n} \to x \geq 0$ as $n \to \infty$. Then the following convergence in distribution holds*

$$\frac{1}{2n^{3/2}}\sum_{v\in\mathfrak{f}^{(n)}}\mathbf{ht}(v) - \frac{2}{\beta n^{3/2}}\sum_{v\in\mathfrak{f}^{(n)}}\mathbf{csn}(v) \overset{(d)}{\Longrightarrow} \int_0^1 B_t^{|\mathrm{br}|,x}\,dt - \frac{1}{\beta}\int_0^\infty (L_{1,x}^v)^2\,dv,$$

where $B^{|\mathrm{br}|,x}$ and $L_{1,x}^v$ are as above.

Proof. This proof follows from arguments similar to Theorem 6.1.2 and the results of Pitman [153] describing a conditional version of a result by Drmota and Gittenberger [68]. Alternatively, when $x = 0$ we can use Jeulin's identity along with a prior work by Drmota and Gittenberger [67].

Let $Z^{*,n} = (Z_h^{*,n}; h = 0, 1, \cdots)$ denote the height profile of $\mathfrak{f}^{(n)}$. Then, by Theorems 4 and 7 in [153], we have

$$\left(\frac{2}{\sqrt{n}} Z^{*,n}_{[2\sqrt{n}v]}; v \geq 0\right) \stackrel{(d)}{\Longrightarrow} \left(L^v_{1,x}; v \geq 0\right),$$

where the convergence above is weak convergence in the Skorokhod space.

Hence,

$$\begin{aligned}
\frac{1}{n^{3/2}} \sum_{v \in \mathfrak{f}^{(n)}} \mathbf{ht}(v) &= \frac{1}{n^{3/2}} \sum_{h=0}^{\infty} h Z_h^{*,n} = \frac{1}{n^{3/2}} \int_0^\infty [u] Z^{*,n}_{[u]}\, du \\
&= \int_0^\infty \frac{[2\sqrt{n}v]}{\sqrt{n}} \cdot \frac{2}{\sqrt{n}} Z^{*,n}_{[2\sqrt{n}v]}\, dv \stackrel{(d)}{\Longrightarrow} 2 \int_0^\infty v L^v_{1,x}\, dv
\end{aligned}$$

and, similarly

$$\begin{aligned}
\frac{1}{n^{3/2}} \sum_{v \in \mathfrak{f}^{(n)}} \mathbf{csn}(v) &= \frac{1}{n^{3/2}} \sum_{h=0}^{\infty} (Z_h^{*,n})(Z_h^{*,n} - 1) \\
&= \frac{1}{n^{3/2}} \sum_{h=0}^{\infty} \left\{ (Z_h^{*,n})^2 - Z_h^{*,n} \right\} \\
&= \frac{1}{2} \int_0^\infty \left(\frac{2}{\sqrt{n}} Z^{*,n}_{[2\sqrt{n}v]} \right)^2 dv - \frac{1}{\sqrt{n}} \\
&\stackrel{(d)}{\Longrightarrow} \frac{1}{2} \int_0^\infty (L^v_{1,x})^2\, dv.
\end{aligned}$$

The result now easily follows. □ □

6.5 SDE Results

In this section we discuss the existence and uniqueness of solutions to SDEs of the form (6.12) and (6.11). We first study the situation with equation (6.12) and then move onto studying equation (6.11).

6.5.1 *Analysis of* (6.12)

We begin by studying a fairly different stochastic differential equation. Let $\tilde{B}$ be a Brownian motion on the filtered probability space $(\Omega, \mathscr{F}_t, \tilde{P})$ and let Y_t be the unique strong solution to the stochastic differential equation

$$dY_t = a d\tilde{B}_t + \left(\frac{c}{Y_t} - \frac{Y_t}{1-t} \right) dt, \qquad Y_0 = x \geq 0, \quad t \in [0,1]$$

where $a > 0$ and $c \geq \frac{a^2}{2}$ are constants. The process Y is a times a Bessel bridge of dimension $\delta = \frac{2c}{a^2} + 1 \geq 2$ and so a unique strong solution exists by properties of Bessel bridges. See Section XI.3 in [156] and [95] for more properties about Bessel bridges.

Now let $f : \mathbb{R}_+ \to \mathbb{R}$ be a continuous function and define M to be the continuous martingale $M_t := \frac{1}{a}\int_0^t f(s)\, d\tilde{B}_s$. Observe that $\mathcal{E}(M_t)$, where $\mathcal{E}$ is the Doléans-Dade exponential, is a positive continuous local martingale. Hence, by Girsanov's theorem [156, Theorem VIII.1.7], the measures $P := \mathcal{E}(M_t) \cdot \tilde{P}$ and $\tilde{P}$ are equivalent on $\mathscr{F}_t$ for each $t \geq 0$. Moreover, $B_t := \tilde{B}_t - \frac{1}{a}\int_0^t f(s)\, ds$ is a P-Brownian motion and so

$$dY_t = a dB_t + \left(\frac{c}{Y_t} + f(t) - \frac{Y_t}{1-t} \right) dt, \qquad Y_0 = x \geq 0, \quad t \in [0,1]. \tag{6.27}$$

Using [156, Theorem IX.1.11], we have the following lemma:

Lemma 6.5.1. *There is weak existence and uniqueness in law to equation* (6.27) *for constants $a > 0$, $c \geq \frac{a^2}{2}$ and continuous function f.*

Moreover, the law of a solution Y to (6.27) *is equivalent to the law of $aR = (aR_t; t \in [0,1])$ where R is a Bessel bridge of dimension $\frac{2c}{a^2} + 1$.*

Using facts about Bessel bridges and the equivalence of measures described in the above lemma, we obtain the following corollary. See, for example, [156, Chapter XI], [95]. In particular, the derivation of Equation (2.6) in [102] generalizes for δ-dimensional Bessel processes when $\delta > 2$.

Corollary 6.5.2. *Suppose Y is a solution to* (6.27) *with $c \geq a^2/2$ and f a continuous function. Then the following hold almost surely.*

1. $\displaystyle\int_0^1 \frac{1}{Y_t}\, dt < \infty$ *if* $c > \frac{a^2}{2}$ *and* $\displaystyle\int_0^t \frac{1}{Y_s}\, ds < \infty$ *for all* $t \in [0,1)$ *when* $x > 0$ *and* $c = \frac{a^2}{2}$.

2. $\displaystyle\int_0^1 \frac{Y_t}{1-t}\,dt < \infty$ *if* $c > \frac{a^2}{2}$.

3. $Y_t > 0$ *for all* $t \in (0,1)$.

4. *The process* $V_t := \int_0^t \frac{1}{Y_s}\,ds$ *for* $t \in [0,1]$ *is strictly increasing and bounded for* $c > \frac{a^2}{2}$. *It is strictly increasing and locally bounded on* $[0,1)$ *when* $c = \frac{a^2}{2}$ *and* $x > 0$.

When $c > \frac{a^2}{2}$, we can define the right-continuous inverse $\tau_v := \inf\{t : V_t > v\}$ for $v \in [0, V_1)$. Since V is strictly increasing τ is actually a two-sided inverse of V. Define the process $Z_v = Y_{\tau_v}$ and observe on the event $\{V_1 > t\}$ and for $v \in [0,t)$

$$\begin{aligned} Z_v &= x + aB_{\tau_v} + \int_0^{\tau_v} \left(\frac{c}{Y_s} + f(s) - \frac{Y_s}{1-s} \right) ds \\ &= x + aB_{\tau_v} + \int_0^{v} \left(\frac{c}{Y_{\tau_u}} + f(\tau_u) - \frac{Y_{\tau_u}}{1-\tau_u} \right) d\tau_u. \end{aligned}$$

We observe that

$$d\tau_u = Y_{\tau_u}\,du = Z_u\,du$$

and, by Proposition V.1.5 and Theorem V.1.6 in [156], we can write $B_{\tau_v} = \int_0^v \sqrt{Z_v}\,dW_v$ for a Brownian motion W. Hence, for $v \in [0, V_1)$

$$Z_v = x + \int_0^v a\sqrt{Z_v}\,dW_v + \int_0^v \left(c + f\left(\int_0^u Z_s\,ds \right) Z_u - \frac{Z_u^2}{1 - \int_0^u Z_s\,ds} \right) du. \tag{6.28}$$

Finally, we observe that $\lim_{v\uparrow V_1} Z_v = \lim_{t\uparrow 1} Y_t = 0$ almost surely. Similarly, given a process Z satisfying (6.28) the process $Y_t = Z_{V_t}$ with $V_t = \inf\{u : \int_0^u Z_s\,ds > t\}$ for $t \in [0,1)$ solves the stochastic differential equation (6.27).

We have thus argued the following proposition:

Proposition 6.5.3. *For* f *a continuous function there is weak existence for the stochastic differential equation* (6.12) *when* $c > \frac{a^2}{2}$ *and* $x \geq 0$.

Moreover, if Z *is any such solution then* $Y_t = Z_{V_t}$ *where* $V_t = \inf\{u : \int_0^u Z_s\,ds > t\}$ *for* $t \in [0,1)$ *solves* (6.27).

6.5.2 Analysis of (6.11)

Throughout this section we assume that $f : \mathbb{R}_+ \to \mathbb{R}$ is a continuous function and $g : \mathbb{R}_+ \to \mathbb{R}$ is a function such that g^2 is Lipschitz and there is some $\varepsilon > 0$ and $M < \infty$ such that $\varepsilon \le g \le M$.

We begin by observing that since g^2 is Lipschitz and bounded below, the function $\frac{1}{g^2}$ is Lipschitz. Now define the function $h(t) = \inf\{v : \int_0^v g^2(s)\, ds > t\}$, and observe that this function is continuous since g^2 is strictly positive and moreover, it is the unique solution to

$$h'(t) = \frac{1}{g^2(h(t))}, \qquad h(0) = 0.$$

Observe that for each $c \ge 0$ the function $\tilde{b}(t) = \frac{2c}{g^2(t)} + 1$ is bounded and continuous. If we define

$$b(t) = \tilde{b}(h(t)),$$

then $b(t)$ is also continuous.

Since $\tilde{b}$ is continuous, by [89, Proposition 1] and [156, Theorem IX.3.5], for each $c \ge 0$ there exists a unique strong solution to the stochastic differential equation

$$dX_t = 2\sqrt{X_t}\, d\tilde{B}_t + b(t)\, dt, \qquad X_0 = x \ge 0,$$

where $\tilde{B}$ is a standard Brownian motion on some filtered probability space $(\Omega, \mathscr{F}_t, \tilde{P})$. But this means that there is a weak solution to the stochastic differential equation

$$\begin{aligned} dX_t^{(1)} &= 2\sqrt{X_t^{(1)}}\, d\tilde{B}_t + \tilde{b}(X_t^{(2)})\, dt, \qquad && X_0^{(1)} = x \ge 0 \\ dX_t^{(2)} &= \frac{1}{g^2(X_t^{(2)})} dt, && X_0^{(2)} = 0, \end{aligned}$$

since $b(t) = \tilde{b}(h(t))$ and $X_t^{(2)} = h(t)$. Since g satisfies $\varepsilon \le g \le M$, by [156, Proposition IX.1.13], there is a weak solution to the stochastic differential equation:

$$\begin{aligned} dX_t^{(1)} &= 2g(X_t^{(2)})\sqrt{X_t^{(1)}}\, d\tilde{B}_t + g^2(X_t^{(2)})\tilde{b}(X_t^{(2)})\, dt, \qquad && X_0^{(1)} = x \ge 0 \\ dX_t^{(2)} &= dt, && X_0^{(2)} = 0. \end{aligned}$$

But this means that $X^{(1)}$ solves

$$dX_t = 2g(t)\sqrt{X_t}\, d\tilde{B}_t + \left(2c + g^2(t)\right)\, dt, \qquad X_0 = x \ge 0. \tag{6.29}$$

The next lemma states that the process X can be bounded below by a deterministic time-change of a squared Bessel process $S = (S(t); t \geq 0)$. We recall [156] that a squared Bessel process of dimension $\delta \geq 0$ and starting from $x \geq 0$ is the unique strong solution of the stochastic differential equation

$$dS(t) = 2\sqrt{S(t)}\, dB_t + \delta\, dt, \qquad S(0) = x,$$

for a standard Brownian motion B.

Lemma 6.5.4. *Suppose X is a solution to (6.29) with respect to a Brownian motion $\tilde{B}$ and started from $x \geq 0$. Fix any $0 \leq \delta \leq \inf_t \left(\frac{2c}{g^2(t)} + 1\right)$. Then on the same probability space $(\Omega, \mathscr{F}, P)$ there exists a δ-dimensional squared Bessel process $S = (S(t); t \geq 0)$ starting from x such that*

$$P\left(S\left(\int_0^t g^2(r)\, dr\right) \leq X_t, \forall t\right) = 1.$$

Proof. We prove this lemma by a time-change. Let $\tau_t = \inf\{s : \int_0^s g^2(r)\, dr > t\}$ and define the process R by $R_t = X_{\tau_t}$. Observe

$$\begin{aligned} R_t &= x + \int_0^{\tau_t} 2g(r)\sqrt{X_r}\, d\tilde{B}_r + \int_0^{\tau_t} \left(2c + g^2(r)\right)\, dr \\ &= x + M_{\tau_t} + \int_0^t \left(\frac{2c}{g^2(\tau_s)} + 1\right)\, ds \end{aligned}$$

where $M_t = \int_0^t 2g(r)\sqrt{X_r}\, d\tilde{B}_r$ and we used the change of variable $r = \tau_s$ in the drift integral. We now observe

$$\langle M_{\tau_\cdot}\rangle_t = \langle M\rangle_{\tau_t} = \int_0^{\tau_t} 4g^2(r)X_r\, dr = \int_0^t 4X_{\tau_s}\, ds = \int_0^t 4R_s\, ds,$$

and so by the Dambis-Dubins-Schwarz theorem [156, Theorem V.1.6] there is a Brownian motion B such that

$$R_t = x + \int_0^t 2\sqrt{R_s}\, dB_s + \int_0^t \left(\frac{2c}{g^2(\tau_s)} + 1\right)\, ds.$$

For each $\delta \geq 0$ and with respect to this Brownian motion B, there exists a unique strong solution to the stochastic differential equation

$$dS_t = 2\sqrt{S_t}\, dB_t + \delta\, dt, \qquad S_t = x.$$

The comparison theorems [156, Theorem IX.3.7] imply that

$$P(S_t \leq R_t, \forall t) = 1$$

for any $0 \leq \delta \leq \inf_t \left(\frac{2c}{g^2(t)} + 1\right)$. The desired claim now follows. □ □

Arguments similar to Lemma 6.5.1 and Corollary 6.5.2 give the following lemma, the details of which are omitted:

Lemma 6.5.5. *Suppose f is a continuous function and g^2 is a Lipschitz function with $0 < \varepsilon \leq g \leq M < \infty$. There exists a weak solution to the stochastic differential equation:*

$$dX_t = 2g(t)\sqrt{X_t}\,dB_t + \left(2c + g^2(t) + 2f(t)\sqrt{X_t}\right)\,dt, \qquad X_0 = x \geq 0.$$

Moreover, for any such solution the following hold almost surely:

1. *If $c > \frac{1}{2}\sup_t g^2(t)$ and $x \geq 0$ or $c = \frac{1}{2}\sup_t g^2(t)$ and $x > 0$, then $\int_0^t X_s^{-1/2}\,ds < \infty$ for each $t < \infty$.*

2. *If $c \geq \frac{1}{2}\sup_t g^2(t)$ and $x > 0$ then $\inf\{t : X_t = 0\} = \infty$.*

Observe that if $X_0 = x > 0$ then X_t almost surely never reaches 0. Therefore, we can apply Itô's rule to $Y_t = \sqrt{X_t}$ and see

$$\begin{aligned} dY_t &= \frac{1}{2\sqrt{X_t}}dX_t - \frac{1}{8X_t^{3/2}}d\langle X\rangle_t \\ &= g(t)dB_t + \left(\frac{c + \frac{1}{2}g^2(t)}{\sqrt{X_t}} + f(t)\right)\,dt - \frac{1}{8X_t^{3/2}}(4g^2(t)X_t)\,dt \\ &= g(t)dB_t + \left(\frac{c}{Y_t} + f(t)\right)\,dt. \end{aligned}$$

Hence we have argued the following lemma:

Lemma 6.5.6. *Suppose f is a continuous function and g^2 is a Lipschitz function with $0 < \varepsilon \leq g \leq M < \infty$. Assume that $c \geq \sup_t \frac{1}{2}g^2(t)$ and $x > 0$. There exists a weak solution to the stochastic differential equation*

$$dY_t = g(t)\,dB_t + \left(\frac{c}{Y_t} + f(t)\right)\,dt, \qquad Y_0 = x > 0. \tag{6.30}$$

Moreover, any such solution is strictly positive and so the process $V_t = \int_0^t \frac{1}{Y_s}\,ds$ *is continuous and strictly increasing.*

Finally, we use a time change and obtain the following proposition.

Proposition 6.5.7. *Suppose* f *is a continuous function and* g^2 *is a Lipschitz function with* $0 < \varepsilon \le g \le M < \infty$*. Assume that* $c \ge \sup_t \frac{1}{2} g^2(t)$ *and* $x > 0$*. Then there exists a weak solution to (6.11). Moreover, for any such solution* Z*, the process* $C_t = \int_0^t Z_s\,ds$ *is strictly increasing.*

Proof. The proof of existence is omitted, since it follows from the arguments similar to those in Proposition 6.5.3.

We now argue that C_t is strictly increasing. To this end let (Z, C) be a solution to (6.11) with respect to some Brownian motion W. To argue that C is strictly increasing, we argue its derivative Z is strictly positive. To this end, we argue by contradiction and suppose that $\tau = \inf\{t : Z_t = 0\} < \infty$. First $\tau > 0$ by continuity and the fact $Z_0 = x > 0$. Hence, $Z_t > 0$ for $t \in [0, \tau)$ and hence C_t is strictly increasing for $t \in [0, \tau)$ and we set $h = C_\tau > 0$.

We now define the right-continuous inverse of C as $V_t = \inf\{r : C_r > t\}$ and define $Y_t = Z_{V_t}$. Observe that for $t \in [0, h)$ we have

$$\begin{aligned} Y_t &= x + \int_0^{V_t} g(C_u)\sqrt{Z_u}\,dW_u + \int_0^{V_t} (c + f(C_u)Z_u)\,du \\ &= x + \tilde{M}_{V_t} + \int_0^t \left(\frac{c}{Z_{V_u}} + f(u) \right) dt \\ &= x + \tilde{M}_{V_t} + \int_0^t \left(\frac{c}{Y_u} + f(u) \right) dt \end{aligned}$$

where $\tilde{M}_t = \int_0^t g(C_u)\sqrt{Z_u}\,dW_u$.

Moreover, we have that

$$\langle \tilde{M}_{V.} \rangle_t = \langle \tilde{M} \rangle_{V_t} = \int_0^{V_t} g^2(C_u) Z_u\,du = \int_0^t g^2(t)\,dt.$$

Hence, by Dambis-Dubins-Schwarz [156, Theorem V.1.6] there is a Brownian motion B on the interval $[0, h)$ such that

$$Y_t = x + \int_0^t g(u)\,dB_u + \int_0^t \left(\frac{c}{Y_u} + f(u) \right) ds, \qquad t \in [0, h).$$

Moreover, $V_t = \int_0^t \frac{1}{Y_s}\,ds$ is continuous and strictly increasing by Lemma 6.5.6 and the stochastic process Y is strictly positive and uniformly bounded away from zero on bounded time intervals. Hence

$$Z_\tau = \lim_{r\uparrow h} Z_{V_r} = \lim_{r\uparrow h} Y_r > 0.$$

This gives the desired contradiction and we conclude that Z_t is always strictly positive and hence $C_t = \int_0^t Z_u\,du$ is strictly increasing. □ □

6.6 Proofs of Normality Results

6.6.1 Proof of Theorem 6.1.6

Throughout this subsection we fix a $c > \frac{a^2}{2}$ and $x \geq 0$ and let Y and Z be related as in Section 6.5.1. In particular, we know that Y and Z solve (6.27) and (6.12), respectively, and are related by $Y_t = Z_{V_t}$ where $V_t = \int_0^t \frac{1}{Y_s}\,ds$. As we have seen the right-continuous inverse of V_t is equal to $\int_0^t Z_u\,du$.

Proof of Theorem 6.1.6. The proof is quite similar to the proofs in [102, Theorem 1.1] and [97, Theorem 1.11].

Since Y solves (6.27), we observe

$$\begin{aligned}\int_0^1 Y_u\,du &= x + \int_0^1 aB_u\,du + \int_0^1 \left(\int_0^u \left\{\frac{c}{Y_s} + f(s) - \frac{Y_s}{1-s}\right\} ds\right) du \\ &= x + \int_0^1 aB_t\,dt + \int_0^1 cV_u\,du + \int_0^1 \left(\int_0^u \left\{f(s) - \frac{Y_s}{1-s}\right\} ds\right) du.\end{aligned}$$

Corollary 6.5.2 allows us to apply Fubini's theorem to the $\frac{Y_s}{1-s}$ integrand to obtain:

$$\int_0^1 \int_0^u \frac{Y_s}{1-s}\,ds\,du = \int_0^1 ds \left\{\int_s^1 \frac{Y_s}{1-s}\,du\right\} = \int_0^1 Y_s\,ds.$$

Hence we have

$$\int_0^1 (2Y_u - cV_u)\,du = x + a\int_0^1 B_s\,ds + \int_0^1 \int_0^u f(s)\,ds\,du$$

It follows that

$$\int_0^1 \left(2Z_{V_t} - cV_t\right)\,dt \stackrel{d}{=} \mathcal{N}\left(\mu, \sigma^2\right)$$

where $\mu = x + \int_0^1 (1-s) f(s)\, ds$ and $\sigma^2 = \frac{a^2}{3}$. By the observation that $Z_v = 0$ for all $v \geq V_1$, and a change of variable we get

$$\int_0^\infty (2Z_v - cv) Z_v \, dv \stackrel{d}{=} \mathcal{N}\left(\mu, \sigma^2\right),$$

which gives the desired claim. □ □

6.6.2 *Proof of Theorem 6.1.5*

Proof of Theorem 6.1.5. We observe that given a solution (Z, C) to (6.11), we can define $Y_t = Z_{V_t}$ where V_t satisfies (6.9) and then Y is a weak solution to (6.30). This follows from the proof of Proposition 6.5.7.

Thus

$$dY_t = g(t)\, dB_t + \left(\frac{c}{Y_t} + f(t)\right) dt, \qquad Y_0 = x.$$

We recall that $V_t = \displaystyle\int_0^t \frac{1}{Y_s}\, ds$ since

$$dV_t = \frac{1}{Z_{V_t}}\, dt = \frac{1}{Y_t}\, dt.$$

Also, $V_t = \int_0^t \frac{1}{Y_s}\, ds$ is finite for each $t \in [0, \infty)$ by Lemma 6.5.6. Then for each $t \geq 0$ we have

$$\begin{aligned} \int_0^t Y_s\, ds &= xt + \int_0^t \int_0^s g(r)\, dB_r ds + \int_0^t \int_0^s \left(\frac{c}{Y_r} + f(r)\right) dr\, ds \\ &= xt + \int_0^t \int_0^s g(r)\, dB_r\, ds + \int_0^t cV_r\, dr + \int_0^t \int_0^s f(r)\, dr\, ds. \end{aligned}$$

Hence

$$\int_0^t (Y_s - cV_s)\, ds = xt + \int_0^t \int_0^s g(r)\, dB_r\, ds + \int_0^t \int_0^s f(r)\, dr\, ds.$$

We observe that the left-hand side is equal to

$$\int_0^t \left(Z_{V_s} - cV_s\right)\, ds = \int_0^{V_t} (Z_v - cv) Z_v\, dv =: X_t.$$

We just need to show that

$$X_t = xt + \int_0^t \int_0^s f(r)\, dr\, ds + \int_0^t \int_0^s g(r)\, dB_r\, ds$$

has the desired mean and covariance structure. It is easy to see

$$\mathbb{E}[X_t] = xt + \int_0^t \int_0^s f(r)\,dr\,ds = xt + \int_0^t (t-s)f(s)\,ds,$$

while the covariance structure follows from the following lemma. □ □

Lemma 6.6.1. *Let $g \in L^2_{\text{loc}}(\mathbb{R}_+)$, let B be a standard Brownian motion and define $G(h) = \int_0^h \int_0^u g(s)\,dB_s\,du$. Then $G(h) = \int_0^h (h-u)\,g(u)\,dB_u$, and consequently it is a centered Gaussian process such that for $h_1 \le h_2$,*

$$\mathbb{E}[G(h_1)G(h_2)] = \int_0^{h_1} (h_2 - s)(h_1 - s) g^2(s)\,ds.$$

Proof. The claim that $G(h) = \int_0^h (h-u)g(u)\,dB_u$ is simply the stochastic Fubini theorem [**?** , Theorem 65].

The covariance structure now follows by Itô's isometry: for any $f_1, f_2 \in L^2_{\text{loc}}(\mathbb{R}_+)$

$$\mathbb{E}\left[\int_0^{h_1} f_1(s)\,dB_s \int_0^{h_2} f_2(t)\,dB_t\right] = \int_0^{h_1 \wedge h_2} f_1(s)\,f_2(s)\,ds.$$

The result follows letting $f_1(u) = (h_1 - u)g(u)$ and $f_2(u) = (h_2 - u)g(u)$. □ □

BIBLIOGRAPHY

[1] Helen Abbey. An examination of the reed-frost theory of epidemics. *Human biology*, 24(3):201, 1952.

[2] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18:no. 14, 21, 2013.

[3] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit. Exit times for an increasing Lévy tree-valued process. *Probab. Theory Related Fields*, 159(1-2):357–403, 2014.

[4] L. Addario-Berry, N. Broutin, and C. Goldschmidt. Critical random graphs: limiting constructions and distributional properties. *Electron. J. Probab.*, 15:no. 25, 741–775, 2010.

[5] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Related Fields*, 152(3-4):367–406, 2012.

[6] Louigi Addario-Berry, Nicolas Broutin, Christina Goldschmidt, and Grégory Miermont. The scaling limit of the minimum spanning tree of the complete graph. *Ann. Probab.*, 45(5):3075–3144, 2017.

[7] David Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.*, 3(4):450–465, 1990.

[8] David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991.

[9] David Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

[10] David Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.

[11] David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[12] David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.

[13] David Aldous and Vlada Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.*, 3:No. 3, 59 pp., 1998.

[14] David Aldous, Grégory Miermont, and Jim Pitman. The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity. *Probab. Theory Related Fields*, 129(2):182–218, 2004.

[15] David Aldous and Jim Pitman. Brownian bridge asymptotics for random mappings. *Random Structures Algorithms*, 5(4):487–512, 1994.

[16] David Aldous and Jim Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields*, 118(4):455–482, 2000.

[17] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[18] Osvaldo Angtuncio and Gerónimo Uribe Bravo. On the profile of trees with a given degree sequence. *arXiv e-prints*, page arXiv:2008.12242, August 2020.

[19] Osvaldo Angtuncio Hernández and Gerónimo Uribe Bravo. Dini derivatives and regularity for exchangeable increment processes. *Trans. Amer. Math. Soc. Ser. B*, 7:24–45, 2020.

[20] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[21] Siva Athreya, Wolfgang Löhr, and Anita Winter. The gap between Gromov-vague and Gromov-Hausdorff-vague topology. *Stochastic Process. Appl.*, 126(9):2527–2553, 2016.

[22] T. L. Austin, R. E. Fagen, W. F. Penney, and John Riordan. The number of components in random linear graphs. *Ann. Math. Statist.*, 30:747–754, 1959.

[23] Nicolas Bacaër. *A short history of mathematical population dynamics*. Springer-Verlag London, Ltd., London, 2011.

[24] Andrew Barbour and Denis Mollison. Epidemics and random graphs. *Lect. Notes Biomath.*, 86:86–89, 01 1990.

[25] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *J. Theoret. Biol.*, 235(2):275–288, 2005.

[26] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[27] Gabriel Berzunza Ojeda, Cecilia Holmgren, and Paul Thévenin. Convergence of trees with a given degree sequence and of their associated laminations. *arXiv e-prints*, page arXiv:2111.07748, November 2021.

[28] Shankar Bhamidi, Nicolas Broutin, Sanchayan Sen, and Xuan Wang. Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph. *arXiv e-prints*, page arXiv:1411.3417, November 2014.

[29] Shankar Bhamidi and Sanchayan Sen. Geometry of the vacant set left by random walk on random graphs, Wright's constants, and critical random graphs with prescribed degrees. *Random Structures Algorithms*, 56(3):676–721, 2020.

[30] Shankar Bhamidi, Sanchayan Sen, and Xuan Wang. Continuum limit of critical inhomogeneous random graphs. *Probab. Theory Related Fields*, 169(1-2):565–641, 2017.

[31] Shankar Bhamidi, Remco van der Hofstad, and Sanchayan Sen. The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. *Probab. Theory Related Fields*, 170(1-2):387–474, 2018.

[32] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.*, 15:no. 54, 1682–1703, 2010.

[33] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.*, 40(6):2299–2361, 2012.

[34] M. J. Bienaymé. *De la loi de multiplication et de la durée des familles probabilités.* Cosson, Paris, 1845.

[35] Patrick Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[36] T. Birsan and D. Tiba. One hundred years since the introduction of the set distance by Dimitrie Pompeiu. In *System modeling and optimization*, volume 199 of *IFIP Int. Fed. Inf. Process.*, pages 35–39. Springer, New York, 2006.

[37] Arthur Blanc-Renaudie. Compactness and fractal dimensions of inhomogeneous continuum random trees. *arXiv e-prints*, page arXiv:2012.13058, December 2020.

[38] Tom Bohman and Michael Picollelli. SIR epidemics on random graphs with a fixed degree sequence. *Random Structures Algorithms*, 41(2):179–214, 2012.

[39] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[40] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[41] Lidia A Braunstein, Sergey V Buldyrev, Reuven Cohen, Shlomo Havlin, and H Eugene Stanley. Optimal paths in disordered complex networks. *Physical review letters*, 91(16):168701, 2003.

[42] Lidia A Braunstein, Zhenhua Wu, Yiping Chen, Sergey V Buldyrev, Tomer Kalisky, Sameet Sreenivasan, Reuven Cohen, Eduardo López, Shlomo Havlin, and H Eugene Stanley. Optimal path and minimal spanning trees in random weighted networks. *International Journal of Bifurcation and Chaos*, 17(07):2215–2255, 2007.

[43] Tom Britton, Svante Janson, and Anders Martin-Löf. Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Adv. in Appl. Probab.*, 39(4):922–948, 2007.

[44] Nicolas Broutin, Thomas Duquesne, and Minmin Wang. Limits of multiplicative inhomogeneous random graphs and Lévy trees: Limit theorems. *arXiv e-prints*, page arXiv:2002.02769, February 2020.

[45] Nicolas Broutin, Thomas Duquesne, and Minmin Wang. Limits of multiplicative inhomogeneous random graphs and Lévy trees: limit theorems. *Probab. Theory Related Fields*, 181(4):865–973, 2021.

[46] Nicolas Broutin and Jean-François Marckert. Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures Algorithms*, 44(3):290–316, 2014.

[47] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[48] M. Emilia Caballero, Amaury Lambert, and Gerónimo Uribe Bravo. Proof(s) of the Lamperti representation of continuous-state branching processes. *Probab. Surv.*, 6:62–89, 2009.

[49] M. Emilia Caballero, José Luis Pérez Garmendia, and Gerónimo Uribe Bravo. A Lamperti-type representation of continuous-state branching processes with immigration. *Ann. Probab.*, 41(3A):1585–1627, 2013.

[50] M. Emilia Caballero, José Luis Pérez Garmendia, and Gerónimo Uribe Bravo. Affine processes on $\mathbb{R}_+^m \times \mathbb{R}^n$ and multiparameter time changes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1280–1304, 2017.

[51] Michael Camarri and Jim Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab.*, 5:no. 2, 18, 2000.

[52] Philippe Chassaing and Svante Janson. A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at 0. *Ann. Probab.*, 29(4):1755–1779, 2001.

[53] Philippe Chassaing and Jean-François Marckert. Parking functions, empirical processes, and the width of rooted labeled trees. *Electron. J. Combin.*, 8(1):Research Paper 14, 19, 2001.

[54] L. Chaumont. Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bull. Sci. Math.*, 121(5):377–403, 1997.

[55] Nicholas A. Christakis and James H. Fowler. Social network sensors for early detection of contagious outbreaks. *PLOS ONE*, 5(9):1–8, 09 2010.

[56] David Clancy, Jr. The Gorin-Shkolnikov identity and its random tree generalization. *J. Theoret. Probab.*, 34(4):2386–2420, 2021.

[57] David Clancy, Jr. A new relationship between Erdős-Rényi graphs, epidemic models and Brownian motion with parabolic drift. *arXiv e-prints*, page arXiv:2006.06838, June 2020.

[58] David Clancy, Jr. Epidemics on critical random graphs with heavy-tailed degree distribution. *arXiv e-prints*, page arXiv:2104.05826, April 2021.

[59] Guillaume Conchon–Kerjan and Christina Goldschmidt. The stable graph: the metric space scaling limit of a critical random graph with i.i.d. power-law degrees. *arXiv e-prints*, page arXiv:2002.04954, February 2020.

[60] Laurent Decreusefond, Jean-Stéphane Dhersin, Pascal Moyal, and Viet Chi Tran. Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Probab.*, 22(2):541–575, 2012.

[61] Amir Dembo and Andrea Montanari. Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.*, 24(2):137–211, 2010.

[62] Souvik Dhara, Remco van der Hofstad, Johan S. H. van Leeuwaarden, and Sanchayan Sen. Heavy-tailed configuration models at criticality. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):1515–1558, 2020.

[63] Souvik Dhara, Remco van der Hofstad, Johan SH Van Leeuwaarden, and Sanchayan Sen. Critical window for the configuration model: finite third moment degrees. *Electronic Journal of Probability*, 22:1–33, 2017.

[64] Jian Ding, Jeong Han Kim, Eyal Lubetzky, and Yuval Peres. Diameters in supercritical random graphs via first passage percolation. *Combin. Probab. Comput.*, 19(5-6):729–751, 2010.

[65] Jian Ding, Jeong Han Kim, Eyal Lubetzky, and Yuval Peres. Anatomy of a young giant component in the random graph. *Random Structures Algorithms*, 39(2):139–178, 2011.

[66] R. G. Dolgoarshinnykh and Steven P. Lalley. Critical scaling for the SIS stochastic epidemic. *J. Appl. Probab.*, 43(3):892–898, 2006.

[67] Michael Drmota and Bernhard Gittenberger. On the profile of random trees. *Random Structures Algorithms*, 10(4):421–451, 1997.

[68] Michael Drmota and Bernhard Gittenberger. Strata of random mappings—a combinatorial approach. *Stochastic Process. Appl.*, 82(2):157–171, 1999.

[69] Ioana Dumitriu and Alan Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[70] Thomas Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.*, 31(2):996–1027, 2003.

[71] Thomas Duquesne. The coding of compact real trees by real valued functions. *arXiv Mathematics e-prints*, page math/0604106, April 2006.

[72] Thomas Duquesne. Continuum random trees and branching processes with immigration. *Stochastic Process. Appl.*, 119(1):99–129, 2009.

[73] Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.

[74] Thomas Duquesne and Jean-François Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.

[75] Rick Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.

[76] Freeman J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Mathematical Phys.*, 3:1191–1198, 1962.

[77] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Mathematical Phys.*, 3:140–156, 1962.

[78] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. II. *J. Mathematical Phys.*, 3:157–165, 1962.

[79] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. III. *J. Mathematical Phys.*, 3:166–175, 1962.

[80] Freeman J. Dyson. The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics. *J. Mathematical Phys.*, 3:1199–1215, 1962.

[81] Freeman J. Dyson and Madan Lal Mehta. Statistical theory of the energy levels of complex systems. IV. *J. Mathematical Phys.*, 4:701–712, 1963.

[82] Alan Edelman and Brian D Sutton. From random matrices to stochastic operators. *Journal of Statistical Physics*, 127(6):1121–1165, 2007.

[83] David A. Edwards. The structure of superspace. In *Studies in topology (Proc. Conf., Univ. North Carolina, Charlotte, N. C., 1974; dedicated to Math. Sect. Polish Acad. Sci.)*, pages 121–133, 1975.

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

[85] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[87] Steven N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005.

[88] Steven N. Evans, Jim Pitman, and Anita Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.

[89] Gabriel Faraud and Stéphane Goutte. Bessel bridges decomposition with varying dimension: applications to finance. *J. Theoret. Probab.*, 27(4):1375–1403, 2014.

[90] William Farr. Progress of epidemics. *Second report of the Registrar General of England and Wales*, pages 16–20, 1840.

[91] Scott L. Feld. Why your friends have more friends than you do. *American Journal of Sociology*, 96(6):1464–1477, 1991.

[92] Manuel Garcia-Herranz, Esteban Moro, Manuel Cebrian, Nicholas A. Christakis, and James H. Fowler. Using friends as sensors to detect global-scale contagious outbreaks. *PLOS ONE*, 9(4):1–7, 04 2014.

[93] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959.

[94] E. N. Gilbert. Random plane networks. *J. Soc. Indust. Appl. Math.*, 9:533–543, 1961.

[95] Anja Göing-Jaeschke and Marc Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[96] Christina Goldschmidt, Bénédicte Haas, and Delphin Sénizergues. Stable graphs: distributions and line-breaking construction. *arXiv e-prints*, page arXiv:1811.06940, November 2018.

[97] Vadim Gorin and Mykhaylo Shkolnikov. Stochastic airy semigroup through tridiagonal matrices. *Ann. Probab.*, 46(4):2287–2344, 2018.

[98] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields*, 145(1-2):285–322, 2009.

[99] D. R. Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probability*, 11:669–677, 1974.

[100] Mikhael Gromov. Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.*, (53):53–73, 1981.

[101] Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[102] Yuu Hariya. A pathwise interpretation of the gorin-shkolnikov identity. *Electron. Commun. Probab.*, 21:Paper No. 52,6, 2016.

[103] T. E. Harris. First passage and recurrence distributions. *Trans. Amer. Math. Soc.*, 73:471–486, 1952.

[104] Shlomo Havlin, Lidia A Braunstein, Sergey V Buldyrev, Reuven Cohen, Tomer Kalisky, Sameet Sreenivasan, and H Eugene Stanley. Optimal path in random networks with disorder: A mini review. *Physica A: Statistical Mechanics and its Applications*, 346(1-2):82–92, 2005.

[105] Svante Janson. Brownian excursion area, wright's constants in graph enumeration, and other brownian areas. *Probab. Surv.*, 4:80–145, 2007.

[106] Svante Janson, Malwina Luczak, and Peter Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures Algorithms*, 45(4):726–763, 2014. [Paging previously given as 724–761].

[107] Svante Janson and Malwina J. Luczak. A new approach to the giant component problem. *Random Structures Algorithms*, 34(2):197–216, 2009.

[108] Th. Jeulin and M. Yor, editors. *Grossissements de filtrations: exemples et applications*, volume 1118 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985. Papers from the seminar on stochastic calculus held at the Université de Paris VI, Paris, 1982/1983.

[109] Adrien Joseph. The component sizes of a critical random graph with given degree sequence. *Ann. Appl. Probab.*, 24(6):2560–2594, 2014.

[110] Olav Kallenberg. Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 27:23–36, 1973.

[111] Olav Kallenberg. *Random measures, theory and applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017.

[112] Olav Kallenberg. *Foundations of modern probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham, [2021] ©2021. Third edition [of 1464694].

[113] Rajeeva L. Karandikar. On pathwise stochastic integration. *Stochastic Process. Appl.*, 57(1):11–18, 1995.

[114] Kiyoshi Kawazu and Shinzo Watanabe. Branching processes with immigration and related limit theorems. *Teor. Verojatnost. i Primenen.*, 16:34–51, 1971.

[115] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[116] William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 138(834):55–83, 1932.

[117] William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics. iii.—further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 141(843):94–122, 1933.

[118] William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics iv. analysis of experimental epidemics of the virus disease mouse ectromelia. *Epidemiology & Infection*, 37(2):172–187, 1937.

[119] William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics: V. analysis of experimental epidemics of mouse-typhoid; a bacterial disease conferring incomplete immunity. *Epidemiology & Infection*, 39(3):271–288, 1939.

[120] Götz Kersting. On the Height Profile of a Conditioned Galton-Watson Tree. *arXiv e-prints*, page arXiv:1101.3656, January 2011.

[121] Ali Khezeli. Metrization of the Gromov-Hausdorff (-Prokhorov) topology for boundedly-compact metric spaces. *Stochastic Process. Appl.*, 130(6):3842–3864, 2020.

[122] F. B. Knight. The uniform law for exchangeable and Lévy process bridges. Number 236, pages 171–188. 1996. Hommage à P. A. Meyer et J. Neveu.

[123] Igor Kortchemski. A simple proof of Duquesne's theorem on contour processes of conditioned Galton-Watson trees. In *Séminaire de Probabilités XLV*, volume 2078 of *Lecture Notes in Math.*, pages 537–558. Springer, Cham, 2013.

[124] Igor Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé-Galton-Watson trees. *Probab. Theory Related Fields*, 168(1-2):1–40, 2017.

[125] Manjunath Krishnapur, Brian Rider, and Bálint Virág. Universality of the stochastic airy operator. *Communications on Pure and Applied Mathematics*, 69(1):145–199, 2016.

[126] Pierre Yves Gaudreau Lamarre and Mykhaylo Shkolnikov. Edge of spiked beta ensembles, stochastic Airy semigroups and reflected Brownian motions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(3):1402–1438, 2019.

[127] Amaury Lambert. The genealogy of continuous-state branching processes with immigration. *Probab. Theory Related Fields*, 122(1):42–70, 2002.

[128] John Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73:382–386, 1967.

[129] Jean-François Le Gall. Brownian excursions, trees and measure-valued branching processes. *Ann. Probab.*, 19(4):1399–1439, 1991.

[130] Jean-François Le Gall. The uniform random tree in a Brownian excursion. *Probab. Theory Related Fields*, 96(3):369–383, 1993.

[131] Jean-François Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.

[132] Jean-François Le Gall and Yves Le Jan. Branching processes in Lévy processes: Laplace functionals of snakes and superprocesses. *Ann. Probab.*, 26(4):1407–1432, 1998.

[133] Jean-François Le Gall and Yves Le Jan. Branching processes in Lévy processes: the exploration process. *Ann. Probab.*, 26(1):213–252, 1998.

[134] Jean-François Le Gall and Marc Yor. Excursions browniennes et carrés de processus de bessel. *C. R. Acad. Sci. Paris Sér. I Math.*, 303(3):73–76, 1986.

[135] Christophe Leuridan. Le théorème de Ray-Knight à temps fixe. In *Séminaire de Probabilités, XXXII*, volume 1686 of *Lecture Notes in Math.*, pages 376–396. Springer, Berlin, 1998.

[136] Wolfgang Löhr, Guillaume Voisin, and Anita Winter. Convergence of bi-measure $\mathbb{R}$-trees and the pruning process. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(4):1342–1368, 2015.

[137] M. Lothaire. *Combinatorics on words.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997. With a foreword by Roger Lyndon and a preface by Dominique Perrin, Corrected reprint of the 1983 original, with a new preface by Perrin.

[138] Tomasz Łuczak. On the equivalence of two basic models of random graphs. In *Random graphs '87 (Poznań, 1987)*, pages 151–157. Wiley, Chichester, 1990.

[139] Tomasz Łuczak. Random trees and random graphs. In *Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997)*, volume 13, pages 485–500, 1998.

[140] Jan Łukasiewicz. *Selected works.* North-Holland Publishing Co., Amsterdam-London; PWN-Polish Scientific Publishers, Warsaw, 1970. Edited by L. Borkowski, Studies in Logic and the Foundations of Mathematics.

[141] Thomas Malthus. *An Essay on the Principle of Population.* Bensley, 1803.

[142] Jean-François Marckert and Abdelkader Mokkadem. The depth first processes of Galton-Watson trees converge to the same Brownian excursion. *Ann. Probab.*, 31(3):1655–1678, 2003.

[143] James B. Martin and Balázs Ráth. Rigid representations of the multiplicative coalescent with linear deletion. *Electron. J. Probab.*, 22:Paper No. 83, 47, 2017.

[144] Madan Lal Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[145] Madan Lal Mehta and Freeman J. Dyson. Statistical theory of the energy levels of complex systems. V. *J. Mathematical Phys.*, 4:713–719, 1963.

[146] Grégory Miermont. Self-similar fragmentations derived from the stable tree. I. Splitting at heights. *Probab. Theory Related Fields*, 127(3):423–454, 2003.

[147] Grégory Miermont and Sanchayan Sen. On breadth-first constructions of scaling limits of random graphs and random unicellular maps. *Random Structures & Algorithms*, n/a(n/a).

[148] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pages 161–179, 1995.

[149] Michael Molloy and Bruce Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, 7(3):295–305, 1998.

[150] Ilkka Norros and Hannu Reittu. On a conditionally Poissonian graph process. *Adv. in Appl. Probab.*, 38(1):59–75, 2006.

[151] Andrei Okounkov. Generating functions for intersection numbers on moduli spaces of curves. *Int. Math. Res. Not.*, (18):933–957, 2002.

[152] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Rev. Modern Phys.*, 87(3):925–979, 2015.

[153] Jim Pitman. The SDE solved by local times of a brownian excursion or bridge derived from the height profile of a random tree or forest. *Ann. Probab.*, 27(1):261–283, 1999.

[154] Jim Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[155] José A. Ramírez, Brian Rider, and Bálint Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.*, 24(4):919–944, 2011.

[156] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[157] Oliver Riordan. The phase transition in the configuration model. *Combin. Probab. Comput.*, 21(1-2):265–299, 2012.

[158] Oliver Riordan and Nicholas Wormald. The diameter of sparse random graphs. *Combin. Probab. Comput.*, 19(5-6):835–926, 2010.

[159] Ronald Ross. *The prevention of malaria*. England, 1911.

[160] Ronald Ross. An application of the theory of probabilities to the study of a priori pathometry.—part i. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 92(638):204–230, 1916.

[161] Ronald Ross and Hilda P Hudson. An application of the theory of probabilities to the study of a priori pathometry.—part ii. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 93(650):212–225, 1917.

[162] Robert E Serfling. Historical review of epidemic theory. *Human biology*, 24(3):145–166, 1952.

[163] M. L. Silverstein. A new approach to local times. *J. Math. Mech.*, 17:1023–1054, 1967/1968.

[164] Florian Simatos. State space collapse for critical multistage epidemics. *Adv. in Appl. Probab.*, 47(3):715–740, 2015.

[165] A. V. Skorohod. Limit theorems for stochastic processes with independent increments. *Teor. Veroyatnost. i Primenen.*, 2:145–177, 1957.

[166] Joel Spencer. Enumerating graphs and Brownian motion. *Comm. Pure Appl. Math.*, 50(3):291–294, 1997.

[167] J. F. Steffensen. Om sandsynligheden for at afkommet uddør. *Matematisk tidsskrift. B*, pages 19–23, 1930.

[168] J. F. Steffensen. Deux problèmes du Calcul des Probabilités. *Ann. Inst. H. Poincaré*, 3(3):319–344, 1933.

[169] Brian David Sutton. *The stochastic operator approach to random matrix theory.* PhD thesis, Massachusetts Institute of Technology, 2005.

[170] Lajos Takács. A generalization of the ballot problem and its application in the theory of queues. *J. Amer. Statist. Assoc.*, 57:327–337, 1962.

[171] Craig A. Tracy and Harold Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1):151–174, 1994.

[172] Craig A. Tracy and Harold Widom. On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.*, 177(3):727–754, 1996.

[173] Remco van der Hofstad. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017.

[174] Wim Vervaat. A relation between Brownian bridge and Brownian excursion. *Ann. Probab.*, 7(1):143–149, 1979.

[175] E. Volz and L. A. Meyers. Susceptible-infected-recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society B: Biological Sciences*, 274:2925–2934, 2007.

[176] Erik Volz. SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.*, 56(3):293–310, 2008.

[177] Bengt von Bahr and Anders Martin-Löf. Threshold limit theorems for some epidemic processes. *Adv. in Appl. Probab.*, 12(2):319–349, 1980.

[178] Jonathan Warren. Branching processes, the Ray-Knight theorem, and sticky Brownian motion. In *Séminaire de Probabilités, XXXI*, volume 1655 of *Lecture Notes in Math.*, pages 1–15. Springer, Berlin, 1997.

[179] H. W. Watson and Francis Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.

[180] Ward Whitt. Some useful functions for functional limit theorems. *Math. Oper. Res.*, 5(1):67–85, 1980.

[181] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)*, 62:548–564, 1955.

[182] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

[183] Biao Wu. On the weak convergence of subordinated systems. *Statist. Probab. Lett.*, 78(18):3203–3211, 2008.

[184] Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971.