Scaling limits of random trees and random graphs

Christina Goldschmidt

Department of Statistics and Lady Margaret Hall, University of Oxford,
goldschm@stats.ox.ac.uk,
http://www.stats.ox.ac.uk/~goldschm

Abstract. In the last 30 years, random combinatorial structures and their scaling limits have formed a flourishing area of research at the interface between probability and combinatorics. In this mini-course, I aim to show some of the beautiful theory that arises when considering scaling limits of random trees and graphs.

Trees are fundamental objects in combinatorics and the enumeration of different classes of trees is a classical subject. In the first section, we will take as our basic object the genealogical tree of a critical Galton–Watson branching process. (As well as having nice probabilistic properties, this class turns out to include various natural types of random combinatorial tree in disguise.) In the same way as Brownian motion is the universal scaling limit for centred random walks of finite step-size variance, it turns out that all critical Galton–Watson trees with finite offspring variance have a universal scaling limit, Aldous' Brownian continuum random tree.

The simplest model of a random network is the Erdős–Rényi random graph: we take n vertices, and include each possible edge independently with probability p. One of the most well-known features of this model is that it undergoes a phase transition. Take $p = c/n$. Then for $c < 1$, the components have size $O(\log n)$, whereas for $c > 1$, there is a giant component, comprising a positive fraction of the vertices, and a collection of components of size $O(\log n)$. (This statements hold with probability tending to 1 as $n \to \infty$.) In the second section, we will focus on the critical setting, $c = 1$, where the largest components have size of order $n^{2/3}$, and are "close" to being trees, in the sense that they have only finitely many more edges than a tree with the same number of vertices would have. We will see how to use a careful comparison with a branching process in order to derive the scaling limit of the critical Erdős–Rényi random graph.

In the final section, we consider the setting of a critical random graph generated according to the configuration model with independent and identically distributed degrees. Here, under natural conditions we obtain the same scaling limit as in the Erdős–Rényi case (up to constants).

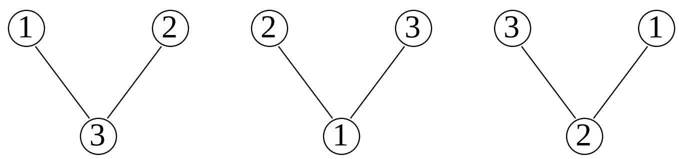
Keywords: scaling limits, random graphs, random trees, $\mathbb{R}$-trees, Brownian continuum random tree

These are (somewhat expanded) lecture notes for a 3-hour long mini-course. The principal aim is to give an idea of the intuition behind the main results, rather than fully rigorous proofs. Some of the ideas are further explored in exercises.

1 Galton–Watson trees and the Brownian continuum random tree

1.1 Uniform random trees

In order to build up some intuition, we start with perhaps the simplest model of a random tree. Let $\mathbb{T}_n$ be the set of (unordered) trees labelled by $[n] := \{1, 2, \ldots, n\}$, and let $\mathbb{T} = \cup_{n \geq 1} \mathbb{T}_n$. For example, $\mathbb{T}_3$ consists of the following trees.



Cayley's formula says that $|\mathbb{T}_n| = n^{n-2}$. We let T_n be a tree chosen uniformly at random from the n^{n-2} elements of $\mathbb{T}_n$. Our first aim is to understand what T_n "looks like" for large n. In order to do this, it will be useful to have an algorithm for building T_n.

It will be technically easier to deal with $\mathbb{T}_n^\bullet$, the set of elements of $\mathbb{T}_n$ with a single distinguished vertex, called the *root*. Given a uniform element of $\mathbb{T}_n^\bullet$, we obtain a uniform element of $\mathbb{T}_n$ by simply forgetting the root.

The Aldous–Broder algorithm
Start from the complete graph on $[n]$.

- Pick a uniform vertex from which to start; this acts as a root.
- Run a simple random walk $(S_k)_{k \geq 0}$ on the graph (i.e. at each step, move to a vertex distinct from the current one, chosen uniformly at random).
- Whenever the walk visits a new vertex, keep the edge along which it was reached.
- Stop when all vertices have been visited.

Claim: the resulting rooted tree is uniform on $\mathbb{T}_n^\bullet$.

The random walk $(S_k)_{k \geq 0}$ has a uniform stationary distribution, and is reversible, so that we may talk about a stationary random walk $(S_k)_{k \in \mathbb{Z}}$. The dynamics of this random walk give rise to Markovian dynamics on $\mathbb{T}_n^\bullet$. In order to see this, let τ_k be the tree constructed from the random walk started at time k (which is rooted at S_k). For $1 \leq i \leq n$, let $\sigma_k(i) = \inf\{j \geq k : S_j = i\}$. Then the tree τ_k has edges

$$\{S_{\sigma_k(i)-1}, i\} \text{ for } 1 \leq i \leq n \text{ such that } \sigma_k(i) > k.$$

Now notice that $\sigma_{k+1}(i) \geq \sigma_k(i)$ for all $1 \leq i \leq n$. So conditionally on the tree τ_k, the tree τ_{k+1} must be independent of $\tau_{k-1}, \tau_{k-2}, \ldots$.

Since the random walk is stationary, the tree must be also. It remains to prove that its distribution is uniform on $\mathbb{T}_n^\bullet$.

Exercise 1 *Consider the* time-reversed chain *(which must have the same stationary distribution). For $\tau, \tau' \in \mathbb{T}_n^\bullet$, write $q(\tau, \tau')$ for the transition probability from τ to τ' for the time-reversed chain.*

1. *Argue that the chain is irreducible on $\mathbb{T}_n^\bullet$.*
2. *Show that for fixed τ, $q(\tau, \tau') = 0$ or $1/(n-1)$.*
3. *Show that for fixed τ', $q(\tau, \tau') = 0$ or $1/(n-1)$.*
4. *It follows that $Q = (q(\tau, \tau'))_{\tau, \tau' \in \mathbb{T}_n^\bullet}$ is a doubly stochastic matrix. Deduce that the stationary distribution must be uniform.*

A variant algorithm due to Aldous

Note that nothing changes if we permute all of the vertex-labels uniformly. So we may as well just do the labelling at the very end. Also, there are steps on which we do not add a new edge at all because the vertex to which the walk moves has already been visited. (Indeed, steps on which we add new edges are separated by geometrically-distributed numbers of steps on which we add no edges.) We may as well suppress this "wandering around" inside the structure we have already built.

- Start from a single vertex labelled 1.
- For $2 \leq i \leq n$, connect vertex i to vertex V_i by an edge, where

$$V_i = \begin{cases} i-1 & \text{with probability } 1 - \frac{i-2}{n-1} \\ k & \text{with probability } \frac{1}{n-1} \text{ for } 1 \leq k \leq i-2. \end{cases}$$

- Uniformly permute the vertex labels.

We may think of this algorithm as growing a sequence of paths with consecutive vertex-labels (of random lengths), with such a path ending whenever we reach a vertex labelled i which connects to $V_i \neq i-1$. The first such path has length

$$C_1^n := \inf\{i \geq 2 : V_i \neq i-1\}.$$

Our first glimpse into the scaling behaviour of T_n is given by the following simple proposition.

Proposition 1 *We have*

$$\frac{C_1^n}{\sqrt{n}} \xrightarrow{d} C_1$$

as $n \to \infty$, where $\mathbb{P}(C_1 > x) = \exp(-x^2/2)$, $x \geq 0$.

Proof. We have

$$\mathbb{P}\left(n^{-1/2}C_1^n > x\right) = \mathbb{P}\left(C_1^n \geq \lfloor x\sqrt{n}\rfloor + 1\right) = \prod_{i=1}^{\lfloor x\sqrt{n}\rfloor - 2}\left(1 - \frac{i}{n-1}\right).$$

Taking logarithms and then Taylor expanding, we have

$$\begin{aligned} -\log \mathbb{P}\left(n^{-1/2}C_1^n > x\right) &= -\sum_{i=1}^{\lfloor x\sqrt{n}\rfloor - 2} \log\left(1 - \frac{i}{n-1}\right) \\ &= \sum_{i=1}^{\lfloor x\sqrt{n}\rfloor - 2} \frac{i}{n-1} + o(1) \\ &= \frac{(\lfloor x\sqrt{n}\rfloor - 2)(\lfloor x\sqrt{n}\rfloor - 1)}{2(n-1)} + o(1) \to \frac{x^2}{2}, \end{aligned}$$

as $n \to \infty$. □

Once we have built the first path of consecutive labels, we pick a uniform random point along it and start growing a second path of uniform labels, etc.

Imagine now that edges in the tree have length 1. Formally, we do this by thinking of T_n as a metric space, where the points of the space are the vertices and the metric is given by the graph distance, for which we write d_n. The proposition suggests that, in order to get some sort of nice limit as $n \to \infty$, we should rescale the graph distance by $n^{-1/2}$.

Here is what turns out to be the limiting version of the algorithm.

Line-breaking construction
Let $C_1, C_2, \ldots$ be the points of an inhomogeneous Poisson process on $[0,\infty)$ of intensity measure $t\mathrm{d}t$. (In particular, we have $\mathbb{P}(C_1 > x) = \exp\left(-\int_0^x t\mathrm{d}t\right) = \exp(-x^2/2)$.) For each $i \geq 1$, conditionally on C_i, let $J_i \sim \mathrm{U}[0, C_i)$. Cut $[0,\infty)$ into intervals at the points given by the C_i's and, for $i \geq 1$, glue the line-segment $[C_i, C_{i+1})$ to the point J_i. (In particular, if we think of this as gradually building up a tree branch by branch, we glue $[C_i, C_{i+1})$ to a point chosen uniformly from the tree built so far.) Think of the union of all of these line-segments as a path metric space, and take its completion. This is (one somewhat informal definition of) the *Brownian continuum random tree (CRT)* $\mathcal{T}$. Write d for its metric.

Theorem 2 (Aldous [5], Le Gall [29]) *As* $n \to \infty$*,*

$$\left(T_n, \frac{1}{\sqrt{n}}d_n\right) \xrightarrow{d} (\mathcal{T}, d).$$

In order to make sense of this convergence, we need a topology on metric spaces, which we will discuss in Section 1.4. For the purposes of the present discussion, let us observe that one way to prove this theorem has at its heart the following joint convergence: if C_k^n is the kth element of the set $\{i \geq 2 : V_i \neq i-1\}$ and $J_k^n = V_{C_k^n}$ then

$$\left(\frac{1}{\sqrt{n}}(C_1^n, J_1^n), \frac{1}{\sqrt{n}}(C_2^n, J_2^n), \ldots\right) \xrightarrow{d} ((C_1, J_1), (C_2, J_2), \ldots).$$

(See Theorem 8 of Aldous [5].) We will later take a different approach in order to prove a more general version of Theorem 2.

1.2 Ordered trees and their encodings

We will henceforth find it easier to work with rooted, ordered trees i.e. those with a distinguished vertex (the root) and such that the *children* of a vertex (its neighbours which are further away from the root) have a given planar order. We will use the standard *Ulam–Harris labelling* by elements of $\mathbb{U} := \cup_{n=0}^{\infty} \mathbb{N}^n$ where, by convention, $\mathbb{N}^0 := \{\emptyset\}$.

Write $\mathbf{T}$ for the set of finite rooted ordered trees. (To an element of $t \in \mathbb{T}$ we may associate a canonical element of $\mathbf{t} \in \mathbf{T}$ by rooting at the vertex labelled 1 in t and then embedding the children of a vertex in t in left-to-right order by increasing label.)

We will find it convenient to encode elements of $\mathbf{T}$ by discrete functions in two different ways. For $\mathbf{t} \in \mathbf{T}$ with n vertices, let $v_0, v_1, \ldots, v_{n-1}$ be the vertices listed in lexicographical order (so that, necessarily, $v_0 = \emptyset$). Let d denote the graph distance on $\mathbf{t}$. We define the *height function* of $\mathbf{t}$ to be $(H(i), 0 \leq i \leq n-1)$, where

$$H(i) := d(v_0, v_i), \quad 0 \leq i \leq n-1.$$

We imagine visiting the vertices of the tree in lexicographical order and simply recording the distance from the root at each step. It is straightforward to recover $\mathbf{t}$ from its height function.

Now let $K(i)$ be the number of children of v_i, for $i \geq 0$, and define the *depth-first walk* (or *Łukasiewicz path*) of $\mathbf{t}$ to be $(X(i), 0 \leq i \leq n)$, where $X(0) := 0$ and

$$X(i) := \sum_{j=0}^{i-1} (K(j) - 1).$$

Again, we imagine visiting the vertices in lexicographical order, but this time we keep track of a *stack* of vertices which we "know about", but have not yet visited. At time 0, we are at the vertex v_0. Whenever we leave a vertex, we become aware of its children (if any) and add them to the stack. We add the children to the stack in reverse lexicographical order, so that the lexicographically smallest child of the vertex we have just left sits at the top at the stack. We also choose a new vertex to visit by taking the one from the top of the stack (it is then removed from the stack). Then for $0 \leq i \leq n-1$, the value $X(i)$ records the size of the stack when we visit vertex v_i.

See Figure 1 for an example.

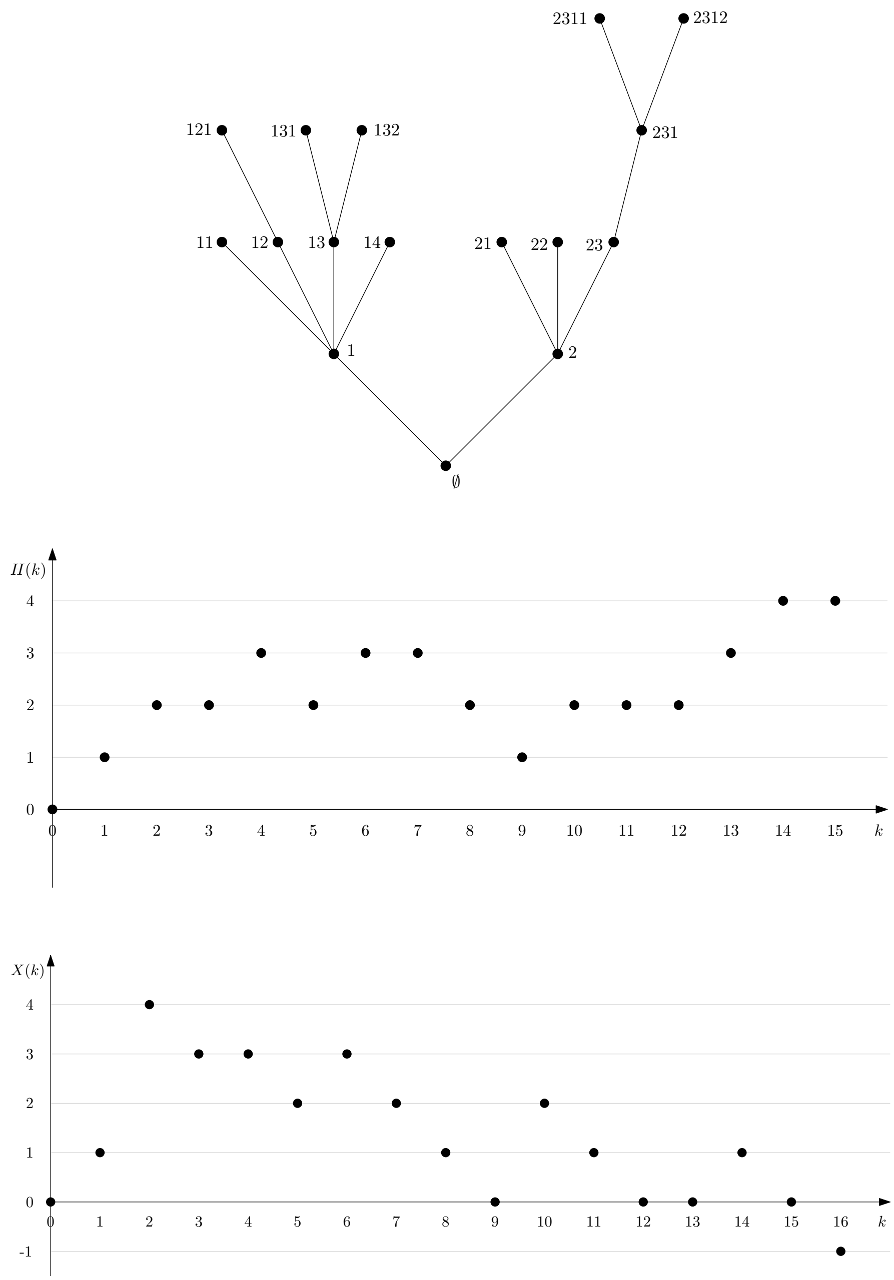


Fig. 1. A rooted ordered tree, its height function and its depth-first walk.

We observe straight away that $X(n) = \sum_{i=0}^{n-1} K(i) - n = -1$, since every vertex is the child of some other vertex, except v_0. On the other hand, for $i < n$, there is some non-negative number of vertices on the stack, so that $X(i) \geq 0$.

We shall now show that X also encodes $\mathbf{t}$ (see Proposition 1.2 of Le Gall [29] for a formal proof).

Proposition 3 *For* $0 \leq k \leq n-1$,

$$H(k) = \# \left\{ 0 \leq i \leq k-1 : X(i) = \min_{i \leq j \leq k} X(j) \right\}.$$

Proof (sketch). For any subtree of the original tree, the value of X once we have just finished exploring it is one less than its value when we visited the root of the subtree, whereas within the subtree, X takes at least its value at the root. Now, the height of v_k is equal to the number of subtrees we have begun but not completed exploring at the step before we reach v_k. The roots of these subtrees are times i before $k-1$ such that X has not yet gone lower by step k i.e.

$$H(k) = \# \left\{ 0 \leq i \leq k-1 : X(i) = \min_{i \leq j \leq k} X(j) \right\},$$

as desired. □

1.3 Galton–Watson trees

Let us now take $T \in \mathbf{T}$ to be random by letting it be the family tree of a Galton–Watson branching process, with offspring distribution $(p_i)_{i \geq 0}$ i.e. each vertex gets a random number of children with distribution $(p_i)_{i \geq 0}$, independently of all other vertices. Let N be the total progeny (i.e. the number of vertices in the tree). We will impose the conditions $p_1 < 1$ and $\sum_{i \geq 0} i p_i \leq 1$, under which $N < \infty$ almost surely. To avoid complicating the statements of our results, except where otherwise stated, we shall also assume that for all n sufficiently large, $\mathbb{P}(N = n) > 0$.

Proposition 4 *Let* $(R(k), k \geq 0)$ *be a random walk with* $R(0) = 0$ *and step distribution* $\nu(i) = p_{i+1}$, $i \geq -1$. *Set*

$$M = \inf\{k \geq 0 : R(k) = -1\}.$$

Then

$$(X(k), 0 \leq k \leq N) \stackrel{d}{=} (R(k), 0 \leq k \leq M).$$

See Proposition 1.5 and Corollary 1.6 of Le Gall [29] for a careful proof. So the depth-first walk of a (sub-critical or critical) Galton–Watson tree is a stopped random walk, which is a rather natural object from a probabilistic perspective. It turns out that many of the most natural combinatorial models of random trees are actually conditioned critical Galton–Watson trees.

Exercise 2 *Let T be a Galton–Watson tree with Poisson(1) offspring distribution and total progeny N.*

1. *Fix a particular rooted ordered tree $\mathbf{t}$ with n vertices having numbers of children $c_v, v \in \mathbf{t}$. What is $\mathbb{P}(T = \mathbf{t})$?*
2. *Condition on the event $\{N = n\}$. Assign the vertices of T a uniformly random labelling by $[n]$, and let $\tilde{T}$ be the labelled tree obtained by forgetting the ordering and the root. Show that $\tilde{T}$ has the same distribution as T_n, a uniform random tree on n vertices.*
   Hint: it suffices to show that the probability of obtaining a particular tree $\mathbf{t}$ is a function of n only.

Exercise 3 *Let T be a Galton–Watson tree with offspring distribution $(p_k)_{k\geq 0}$ and total progeny N.*

(a) Show that if $p_k = 2^{-k-1}, k \geq 0$ then, conditional on $N = n$, T is uniform on the set of ordered rooted trees with n vertices.
(b) Show that if $p_0 = 1/2$ and $p_2 = 1/2$ then, conditional on $N = 2n + 1$, T is uniform on the set of binary trees with n vertices of degree 3 and $n + 1$ leaves.

Suppose now that $\sum_{i=1}^{\infty} ip_i = 1$ and $\sigma^2 := \sum_{i=1}^{\infty}(i-1)^2 p_i \in (0, \infty)$. Write $(X^n(k), 0 \leq k \leq n)$ for the depth-first walk of our Galton–Watson tree conditioned on $N = n$.

Theorem 5 *As $n \to \infty$,*

$$\frac{1}{\sigma\sqrt{n}}(X^n(\lfloor nt \rfloor), 0 \leq t \leq 1) \xrightarrow{d} (e(t), 0 \leq t \leq 1),$$

where $(e(t), 0 \leq t \leq 1)$ is a standard Brownian excursion.

Using the fact that the depth-first walk of a Galton–Watson tree is a stopped random walk, this follows from a conditioned version of Donsker's invariance principle (Theorem 2.6 of Kaigh [27]). A highly non-trivial consequence of Theorem 5 is that, up to a scaling constant, the same is true for H^n, the conditioned height process.

Theorem 6 *As $n \to \infty$,*

$$\frac{\sigma}{\sqrt{n}}(H^n(\lfloor nt \rfloor), 0 \leq t \leq 1) \xrightarrow{d} 2(e(t), 0 \leq t \leq 1),$$

The Brownian CRT, which we encountered in Section 1.1 via the line-breaking construction, is the tree encoded (in a sense to be made precise in the next section) by $2(e(t), 0 \leq t \leq 1)$.

See Sections 1.3 and 1.5 of Le Gall [29] for a complete proof of this theorem. We will give a sketch proof, not for the case of a single tree conditioned to have size n, but rather for a sequence of i.i.d. unconditioned critical Galton–Watson trees. (It is technically easier not to have to deal with the conditioning.) We encode this "forest" via the (shifted) concatenation of the depth-first walks of its trees: a new tree starts every time X reaches a new minimum. (We must be a little careful now with our interpretation of the quantity $X(k)$: it is the number of vertices on the stack, minus the number of components we have completely explored.) We take H to be defined, as before, via

$$H(k) = \#\left\{0 \leq i \leq k-1 : X(i) = \min_{i \leq j \leq k} X(j)\right\}. \tag{1}$$

Donsker's theorem easily gives

$$\frac{1}{\sigma\sqrt{n}}(X(\lfloor nt \rfloor), t \geq 0) \xrightarrow{d} (W(t), t \geq 0),$$

where W is a standard Brownian motion. The analogue of Theorem 6 is then as follows.

Theorem 7 *As $n \to \infty$,*

$$\frac{\sigma}{\sqrt{n}}(H^n(\lfloor nt \rfloor), t \geq 0) \xrightarrow{d} 2\left(W(t) - \inf_{0 \leq s \leq t} W(s), t \geq 0\right).$$

(The right-hand side has the same distribution as twice a reflecting Brownian motion $(|W(t)|, t \geq 0)$. We interpret this as encoding a forest of continuous trees, each corresponding to an excursion away from 0.)

The random walks which occur as depth-first walks of Galton–Watson trees have the special property that they are *skip-free to the left*, which means that they have step distribution concentrated on $\{-1, 0, 1, 2, \ldots\}$. It turns out that these random walks have particularly nice properties, some of which we explore in the next exercise.

Exercise 4 *Let $(X(k), k \geq 0)$ be a random walk with step distribution $\nu(k)$, $k \geq -1$. Assume that $\sum_{k \geq -1} k\nu(k) = 0$ and that $\sum_{k \geq -1} k^2\nu(k) = \sigma^2 < \infty$. Suppose $X(0) = 0$ and let $T = \inf\{k \geq 1 : X(k) \geq 0\}$. This is called the* first weak (ascending) ladder time*. The random walk is recurrent, and so $T < \infty$ a.s. and it follows that the* first weak ladder height *$X(T)$ is finite a.s.*

If the first step of the random walk is to 0 or above, then $T = 1$ and $X(T)$ is simply the new location of the random walk.

If the first step is to -1, on the other hand, things are more involved. In general, the random walk may now make several excursions which go below -1 and stay below it before returning to -1. Finally the walk leaves -1, perhaps initially going downwards, but eventually reaching $\{0, 1, 2, \ldots\}$ without hitting -1 again*. Indeed, using the strong Markov property, we can see that the random walk makes a geometrically distributed number of excursions which return to -1 before it hits $\{0, 1, \ldots\}$, where the parameter of the geometric distribution is (by translation-invariance) $\mathbb{P}(X(T) > 0)$.*

1. *By conditioning on the first step of the random walk, and using the above considerations, show that for* $k \geq 0$,
$$\mathbb{P}(X(T) = k) = \nu(k) + \nu(-1)\mathbb{P}(X(T) = k+1 | X(T) > 0).$$
2. *Show that for* $k \geq 0$,
$$\mathbb{P}(X(T) = k) = \sum_{j=0}^{\infty} \left(\frac{\nu(-1)}{\mathbb{P}(X(T) > 0)} \right)^j \nu(k+j).$$
3. *Show directly that*
$$\sum_{k=0}^{\infty} \bar{\nu}(k) = 1$$
where $\bar{\nu(k)} = \sum_{j=k}^{\infty} \nu(j)$.
4. *Using the fact that* $\sum_{k=0}^{\infty} \mathbb{P}(X(T) = k) = 1$, *deduce carefully that we must have*
$$\mathbb{P}(X(T) > 0) = \nu(-1),$$
and hence that $\mathbb{P}(X(T) = k) = \bar{\nu}(k)$ *for* $k \geq 0$.
Hint: you may want to use a probability generating function.
5. *Finally, show that* $\mathbb{E}[X(T)] = \sigma^2/2$.

This calculation is inspired by one by Jean-François Marckert & Abdelkader Mokkadem in [32]. They credit the argument to Feller.

We will use the result of the exercise to prove that the height process converges in the sense of finite-dimensional distributions.

Proposition 8 *For any* $m \geq 1$ *and any* $0 \leq t_1 \leq t_2 \leq \ldots \leq t_m < \infty$,
$$\frac{1}{\sqrt{n}} (H(\lfloor nt_1 \rfloor), \ldots, H(\lfloor nt_m \rfloor)) \xrightarrow{d} \frac{2}{\sigma} \left(W(t_1) - \inf_{0 \leq s \leq t_1} W(s), \ldots, W(t_m) - \inf_{0 \leq s \leq t_m} W(s) \right).$$

Proof. Let $S(n) = \sup_{0 \leq k \leq n} X(k)$ and $I(n) = \inf_{0 \leq k \leq n} X(k)$. Let us introduce the time-reversed random walk, which takes the same jumps but in the opposite order:
$$\hat{X}^n(k) = X(n) - X(n-k).$$
Then
$$(\hat{X}^n(k), 0 \leq k \leq n) \stackrel{d}{=} (X(k), 0 \leq k \leq n).$$
Hence,
$$\begin{aligned} H(n) &= \# \left\{ 0 \leq k \leq n-1 : X(k) = \inf_{k \leq j \leq n} X(j) \right\} \\ &= \# \left\{ 1 \leq i \leq n : X(n-i) = \inf_{0 \leq \ell \leq i} X(n-\ell) \right\} \\ &= \# \left\{ 1 \leq i \leq n : \hat{X}^n(i) = \sup_{0 \leq \ell \leq i} \hat{X}^n(\ell) \right\}. \end{aligned}$$

By analogy, define

$$J(n) = \# \left\{ 1 \leq i \leq n : X(i) = \sup_{0 \leq \ell \leq i} X(\ell) \right\} = \# \left\{ 1 \leq i \leq n : X(i) = S(i) \right\}.$$

Note that

$$\sup_{0 \leq k \leq n} \hat{X}^n(k) = X(n) - \inf_{0 \leq k \leq n} X(k) = X(n) - I(n).$$

It follows that *for each fixed* n,

$$(S(n), J(n)) \stackrel{d}{=} (X(n) - I(n), H(n)).$$

Now define $T_0 = 0$ and $T_k = \inf\{i > T_{k-1} : X(i) = S(i)\}$, $k \geq 1$. Then the random variables $\{X(T_{k+1}) - X(T_k), k \geq 0\}$ are i.i.d. by the strong Markov property. By Exercise 4, they have mean $\sigma^2/2$.

We now claim that

$$\frac{H(n)}{X(n) - I(n)} \xrightarrow{p} \frac{2}{\sigma^2}, \tag{2}$$

as $n \to \infty$. To see this, write

$$\begin{aligned} S(n) = \sum_{k \geq 1 : T_k \leq n} (S(T_k) - S(T_{k-1})) &= \sum_{k=1}^{J(n)} (S(T_k) - S(T_{k-1})) \\ &= \sum_{k=1}^{J(n)} (X(T_k) - X(T_{k-1}). \end{aligned}$$

Since $J(n) \to \infty$ as $n \to \infty$, by the Strong Law of Large Numbers we have

$$\frac{S(n)}{J(n)} \to \mathbb{E}\left[X(T_1)\right] = \frac{\sigma^2}{2} \quad \text{a.s.}$$

as $n \to \infty$. Since $(S(n), J(n)) \stackrel{d}{=} (X(n) - I(n), H(n))$ for each n, we deduce that

$$\frac{X(n) - I(n)}{H(n)} \xrightarrow{p} \frac{\sigma^2}{2}$$

as $n \to \infty$. Now, we know that

$$\frac{1}{\sqrt{n}} (X(\lfloor nt \rfloor), t \geq 0) \xrightarrow{d} \sigma (W(t), t \geq 0)$$

and so, by the continuous mapping theorem,

$$\begin{aligned} &\frac{1}{\sqrt{n}} \left(X(\lfloor nt_1 \rfloor) - I(\lfloor nt_1 \rfloor), \ldots, X(\lfloor nt_m \rfloor) - I(\lfloor nt_m \rfloor) \right) \\ &\quad \xrightarrow{d} \sigma \left(W(t_1) - \inf_{0 \leq s \leq t_1} W(s), \ldots, W(t_m) - \inf_{0 \leq s \leq t_m} W(s) \right). \end{aligned}$$

The result then follows by using (2). □

1.4 $\mathbb{R}$-trees encoded by continuous excursions

We now turn to our notion of a continuous tree.

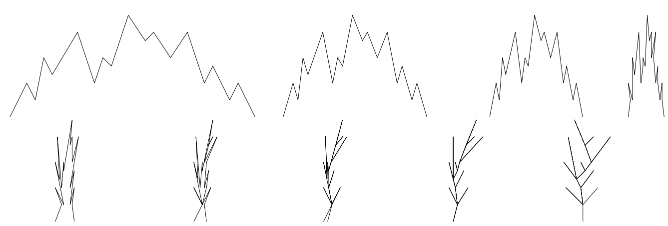
Definition 9 *A compact metric space (T, d) is an* $\mathbb{R}$-tree *if the following conditions are fulfilled for every pair $x, y \in T$:*

- *There exists a unique isometric map $f_{x,y} : [0, d(x, y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x, y)) = y$. We write $[[x, y]] := f_{x,y}([0, d(x, y)])$.*
- *If g is a continuous injective map $[0, 1] \to T$ such that $g(0) = x$ and $g(1) = y$ then $g([0, 1]) = [[x, y]]$.*

A *continuous excursion* is a continuous function $h : [0, \zeta] \to \mathbb{R}_+$ such that $h(0) = h(\zeta) = 0$ and $h(x) > 0$ for $x \in (0, \zeta)$, for some $0 < \zeta < \infty$. We will build all of our $\mathbb{R}$-trees from such functions. For a continuous excursion h, define first a pseudo-metric on $[0, \zeta]$ via

$$d_h(x, y) = h(x) + h(y) - 2 \inf_{x \wedge y \leq x \leq x \vee y} h(z).$$

Then define an equivalence relation by $x \sim y$ iff $d_h(x, y) = 0$, and let $\mathcal{T}_h$ be given by the quotient $[0, \zeta]/\sim$. Intuitively, we put glue on the underside of the function h and then imagine squashing the function from the right: whenever two parts of the function at the same height and with glue on them meet, they stick together. See below (distances in the tree should be interpreted vertically).



In particular, local minima of the function become branch-points of the tree.

Theorem 10 *For any continuous excursion h, $(\mathcal{T}_h, d_h)$ is an $\mathbb{R}$-tree.*

For a proof, see Theorem 2.2 of Le Gall [29]. For $t \in [0, \zeta]$, we write p_h for the canonical projection $[0, \zeta] \to \mathcal{T}_h$. It is usual to think of the tree as rooted at $\rho = p_h(0) = p_h(\zeta)$, the equivalence class of 0. It will be useful later to have a measure μ_h on $\mathcal{T}_h$ which is given by the push-forward of the Lebesgue measure on $[0, \zeta]$.

Now let $\mathcal{M}$ be the space of isometry classes of compact metric spaces. We endow $\mathcal{M}$ with the Gromov–Hausdorff distance. To define this, let (X, d) and (X', d') be elements of $\mathcal{M}$. A *correspondence* R is a subset of $X \times X'$ such that for all $x \in X$, there exists $x' \in X'$ such that $(x, x') \in R$ and vice versa. The *distortion* of R is

$$\operatorname{dis}(R) = \sup\{|d(x, y) - d'(x', y')| : (x, x'), (y, y') \in R\}.$$

The *Gromov–Hausdorff distance* between (X, d) and (X', d') is then given by

$$d_{\mathrm{GH}}((X,d),(X',d')) = \frac{1}{2} \inf_R \mathrm{dis}(R),$$

where the infimum is taken over all correspondences R between X and X'. Importantly, $(\mathcal{M}, d_{\mathrm{GH}})$ is a Polish space. (See Burago, Burago and Ivanov [17] for much more about the Gromov–Hausdorff distance.)

We define the Brownian CRT to be $(\mathcal{T}_{2e}, d_{2e})$, where $e = (e(t), 0 \le t \le 1)$ is a standard Brownian excursion.

1.5 Convergence to the Brownian CRT

Let T_n be a critical Galton–Watson tree with finite offspring variance $\sigma^2 > 0$, conditioned to have total size n, and let d_n be the graph distance on T_n.

Theorem 11 (Aldous [7], Le Gall [29]) *As $n \to \infty$,*

$$\left(T_n, \frac{\sigma}{\sqrt{n}} d_n\right) \xrightarrow{d} (\mathcal{T}_{2e}, d_{2e}).$$

Proof. (I learnt this proof from Grégory Miermont.) By Skorokhod's representation theorem, we can find a probability space on which the convergence

$$\frac{\sigma}{\sqrt{n}}(H^n(\lfloor nt \rfloor), 0 \le t \le 1) \to 2(e(t), 0 \le t \le 1)$$

occurs almost surely (in the uniform norm). As usual, write $v_0, v_1, \ldots, v_{n-1}$ for the vertices of T_n in lexicographical order. Then $(T_n, \frac{\sigma}{\sqrt{n}} d_n)$ is isometric to $\{0, 1, \ldots, n-1\}$ endowed with the distance

$$d^n(i,j) = \frac{\sigma}{\sqrt{n}} d_n(v_i, v_j).$$

Define a correspondence R_n between $\{0, 1, \ldots, n-1\}$ and $[0,1]$ by setting $(i, s) \in R_n$ if $i = \lfloor ns \rfloor$; we also declare that $(n-1, 1) \in R_n$. Now endow $[0,1]$ with the pseudo-metric d_{2e}. We will bound $\mathrm{dis}(R_n)$. Note first that if we write $u \wedge v$ for the most recent common ancestor of u and v, then

$$d_n(v_i, v_j) = d_n(v_0, v_i) + d_n(v_0, v_j) - 2 d_n(v_0, v_i \wedge v_j).$$

By definition,

$$d_n(v_0, v_i) = H^n(i).$$

Moreover, it is not hard to see that

$$\left| d_n(v_0, v_i \wedge v_j) - \min_{i \le k \le j} H^n(k) \right| \le 1.$$

Now suppose that $(i,s),(j,t) \in R_n$ with $s \le t$. Then

$$\begin{aligned} |d^n(i,j) - d_{2e}(s,t)| &= \Bigg| \frac{\sigma}{\sqrt{n}} \left(H^n(\lfloor ns \rfloor) + H^n(\lfloor nt \rfloor) - 2 d_n(v_0, v_i \wedge v_j) \right) \\ &\qquad - \left(2e(s) + 2e(t) - 4 \min_{s \le u \le t} e(u) \right) \Bigg| \\ &\le \Bigg| \frac{\sigma}{\sqrt{n}} \left(H^n(\lfloor ns \rfloor) + H^n(\lfloor nt \rfloor) - 2 \min_{s \le u \le t} H^n(\lfloor nu \rfloor) \right) \\ &\qquad - \left(2e(s) + 2e(t) - 4 \min_{s \le u \le t} e(u) \right) \Bigg| + \frac{2\sigma}{\sqrt{n}}. \end{aligned}$$

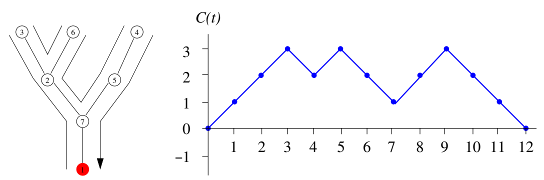
The right-hand side converges to 0 uniformly in $s, t \in [0,1]$. Since

$$d_{\mathrm{GH}} \left(\left(T_n, \frac{\sigma}{\sqrt{n}} d_n \right), (\mathcal{T}_{2e}, d_{2e}) \right) \le \frac{1}{2} \mathrm{dis}(R_n),$$

the result follows. □

There are several steps along the way to the proof of Theorem 11 which, due to a lack of time, I have omitted. The following exercise is intended to lead you through a complete proof in one special case, assuming only Kaigh's theorem on the convergence of a random walk excursion.

Exercise 5 *We have discussed the depth-first walk and the height function of a tree. A third encoding which is often used is the so-called* contour function $(C(i), 0 \le i \le 2(n-1))$*. For a tree* $\mathbf{t} \in \mathbf{T}$*, we imagine a particle tracing the outline of the tree from left to right at speed 1. (The picture below is for a labelled tree, with a planar embedding given by the labels.) Notice that we visit every vertex apart from the root* $\emptyset$ *a number of times given by its degree.*



Let T_n *be a Galton–Watson tree with offspring distribution* $p(k) = 2^{-k-1}, k \ge 0$*, conditioned to have total progeny* $N = n$*, as in Exercise 3(a). Let* $(C^n(i), 0 \le i \le 2(n-1))$ *be its contour function. It will be convenient to define a somewhat shifted version: let* $\tilde{C}^n(0) = 0$*,* $\tilde{C}^n(2n) = 0$ *and, for* $1 \le i \le 2n-1$*,* $\tilde{C}^n(i) = 1 + C(i-1)$*.*

1. *Show that* $(\tilde{C}^n(i), 0 \le i \le 2n)$ *has the same distribution as a simple symmetric random walk (i.e. a random walk which makes steps of* $+1$ *with probability* $1/2$ *and steps of* -1 *with probability* $1/2$*) conditioned to return to the origin for the first time at time* $2n$*.*
   Hint: first consider the unconditioned Galton–Galton–WatsonWatson tree with this offspring distribution.

2. *It's straightforward to interpolate linearly to get a continuous function* $\tilde{C}^n : [0, 2n] \to \mathbb{R}_+$. *Let* $\tilde{T}^n$ *be the* $\mathbb{R}$*-tree encoded by this linear interpolation. Show that*
$$d_{\mathrm{GH}}(T_n, \tilde{T}_n) \leq \frac{1}{2}.$$
Hint: notice that T_n considered as a metric space has only n points, whereas $\tilde{T}_n$ is an $\mathbb{R}$-tree and consists of uncountably many points. Draw a picture and find a correspondence.
3. *Suppose that we have continuous excursions* $f : [0,1] \to \mathbb{R}_+$ *and* $g : [0,1] \to \mathbb{R}_+$ *which encode* $\mathbb{R}$*-trees* $\mathcal{T}_f$ *and* $\mathcal{T}_g$. *For* $t \in [0,1]$, *let* $p_f(t)$ *be the image of* t *in the tree* $\mathcal{T}_f$ *and similarly for* $p_g(t)$. *Define a correspondence*
$$R = \{(x,y) \in \mathcal{T}_f \times \mathcal{T}_g : x = p_f(t),\ y = p_g(t) \text{ for some } t \in [0,1]\}.$$
Show that $\mathrm{dis}(R) \leq 4\|f - g\|_\infty$.
Hint: recall how the metric in an $\mathbb{R}$-tree is related to the function encoding it.
4. *Observe that the variance of the step-size in a simple symmetric random walk is* 1. *Hence, by Theorem 5, we have*
$$\frac{1}{\sqrt{2(n-1)}}(C^n(2(n-1)t), 0 \leq t \leq 1) \xrightarrow{d} (e(t), 0 \leq t \leq 1) \qquad (*)$$
as $n \to \infty$. *Use this, (b) and (c) to prove directly that* $(T_n, \frac{1}{\sqrt{n}} d_n)$ *converges to a constant multiple of the Brownian CRT in the Gromov-Hausdorff sense.*
Hint: you may want to use Skorokhod's representation theorem in order to work on a probability space where the convergence (*) occurs almost surely.

This approach is taken from Jean-François Le Gall & Grégory Miermont's lecture notes [31].

1.6 Properties of the Brownian CRT

A relatively straightforward extension of Theorem 11 shows that, in the appropriate topology (that generated by the Gromov–Hausdorff–Prokhorov distance; see Abraham, Hoscheit and Delmas [1] for a definition), the metric space $(T_n, \frac{\sigma}{\sqrt{n}} d_n)$ endowed additionally with the uniform measure on the vertices of T_n converges to $(\mathcal{T}_{2e}, d_{2e})$ endowed with the measure μ_{2e} (the push-forward of the Lebesgue measure on $[0,1]$). In consequence, we refer to μ_{2e} as the *uniform measure* on $\mathcal{T}_{2e}$.

Consider picking points according to μ_{2e}. We may generate a sample from μ_{2e} simply by taking $p_{2e}(U)$ where $U \sim \mathrm{U}[0,1]$ (recall that p_{2e} is the projection $[0,1] \to \mathcal{T}_{2e}$). It turns out that $p_{2e}(U)$ is almost surely a leaf. (This may seem surprising at first sight. But we can think of it as saying, for example, that every vertex of a uniform random tree is at distance $o(\sqrt{n})$ from a leaf.) It is also the case that the rooted Brownian CRT $(\mathcal{T}_{2e}, d_{2e}, \rho)$ is invariant in distribution under random re-rooting at a point sampled from μ_{2e} (this follows because the same property is true for the uniform random tree T_n).

For fixed $k \geq 1$, let $X_1, X_2, \ldots, X_k$ be leaves of $\mathcal{T}_{2e}$ sampled according to μ_{2e}. Then the subtree of $\mathcal{T}_{2e}$ spanned by the set of points $\{\rho, X_1, \ldots, X_k\}$ has exactly the same distribution as the tree produced at step k in the line-breaking construction discussed in Section 1.1 [5].

The Brownian CRT has many other fascinating properties: for example, it is a random fractal, with Hausdorff and Minkowski dimension both equal to 2, almost surely [24, 30].

2 The critical Erdős–Rényi random graph

We now turn to perhaps the simplest and best-known model of a random graph. Take n vertices labelled by $[n]$ and put an edge between any pair of them independently with probability p, for some fixed $p \in [0,1]$. We write $G(n,p)$ for the resulting random graph. We will be interested in the connected components of $G(n,p)$ and, in particular, in their size and structure.

2.1 The phase transition and component sizes in the critical window

Let $p = c/n$ for some constant $c > 0$. The following statements hold with probability tending to 1 as $n \to \infty$:

- if $c < 1$, the largest connected component of $G(n,p)$ has size $\Theta(\log n)$;
- if $c > 1$, the largest connected component has size $\Theta(n)$ and the others are all of size $O(\log n)$.

In the latter case, we refer to the largest component as the *giant*.

Let us give an heuristic explanation for this phenomenon. We think about exploring the graph in a depth-first manner, which we will make more precise later. Firstly consider the vertex labelled 1. It has a $\text{Bin}(n-1, c/n) \approx \text{Po}(c)$ number of neighbours, say K. Consider the lowest-labelled of these neighbours. Conditionally on K, it itself has a $\text{Bin}(n-K-1, c/n)$ number of new neighbours. This distribution is still well-approximated by $\text{Po}(c)$, as long as $K = o(n)$. So we may think of exploring vertex by vertex and approximating the size of the component that we discover by the total progeny of a Galton–Watson branching process with $\text{Po}(c)$ offspring distribution, as long as the total number of vertices we have visited remains small relative to the size of the graph. If $c \leq 1$, such a branching process dies out with probability 1, which corresponds to obtaining a small component containing vertex 1. A similar argument will then work in subsequent components. If $c > 1$, there is positive probability that the branching process will survive. The branching process approximation holds good until we first explore a component which does not "die out"; this ends up being the giant component.

We will focus here on the critical case $c = 1$ or, more precisely, on the *critical window*: $p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$, $\lambda \in \mathbb{R}$. We will show in a moment that here the largest components have sizes on the order of $n^{2/3}$. With a view to later understanding the structure of these components, we will also track the *surplus* of each one, that is the number of edges it has more than a tree with the same number of vertices would: a component with m vertices and k edges has surplus $k - m + 1$.

Let us fix $\lambda \in \mathbb{R}$ and let $C_1^n, C_2^n, \ldots$ be the component sizes of $G\left(n, \frac{1}{n} + \frac{\lambda}{n^{4/3}}\right)$, listed in decreasing order, and let $S_1^n, S_2^n, \ldots$ be the corresponding surpluses.

Theorem 12 (Aldous [8]) *As* $n \to \infty$,

$$\left(\frac{1}{n^{2/3}}(C_1^n, C_2^n, \ldots), (S_1^n, S_2^n, \ldots)\right) \xrightarrow{d} ((C_1, C_2, \ldots), (S_1, S_2, \ldots))$$

where the limit has an explicit description to be given below. Convergence for the first sequence takes place in

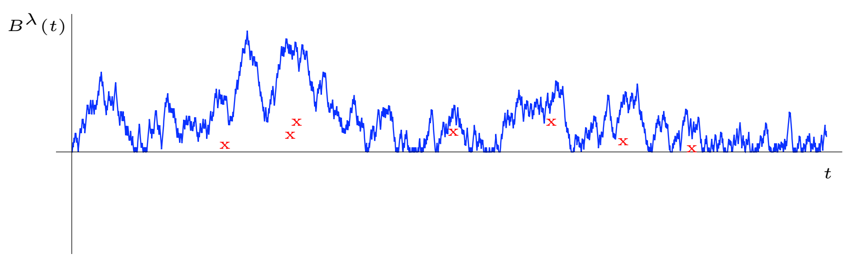
$$\ell^2_{\downarrow} := \left\{\mathbf{x} = (x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots \geq 0, \sum_{i=1}^{\infty} x_i^2 < \infty\right\}$$

with the usual ℓ^2 *distance; for the second sequence, it is in the product topology.*

To describe the limit, let

$$W^{\lambda}(t) = W(t) + \lambda t - \frac{t^2}{2}, \quad t \geq 0,$$

where $(W(t), t \geq 0)$ is a standard Brownian motion. Let $B^{\lambda}(t) = W^{\lambda}(t) - \inf_{0 \leq s \leq t} W^{\lambda}(s)$ be the the process reflected at its minimum. Now draw the graph of the process B^{λ} and decorate it with the points of a rate 1 Poisson process in the plane, keeping only those points which fall between the x-axis and the function, as illustrated below. Conditionally on B^{λ}, we let $(P^{\lambda}(t), t \geq 0)$ be the inhomogeneous Poisson process of intensity $B^{\lambda}(t)$ at time t which describes the arrival of the points.



Then $(C_1, C_2, \ldots)$ is the ordered sequence of excursion-lengths above 0 of the process B^{λ} and $(S_1, S_2, \ldots)$ is the sequence of numbers of Poisson points falling in the corresponding excursions.

The key to the proof of this result is the depth-first walk (or a variant thereof: Aldous' original proof rather uses a *breadth-first* walk, but the details are essentially identical). Consider the component containing vertex 1. We explore as we would in a tree, simply ignoring any edges to vertices we have already seen (which make cycles), and using the labels to tell us which is the root of a component (the lowest-labelled vertex) and to provide a left-right ordering for the children of a vertex. When we reach the end of a component, we start exploring the component of the lowest-labelled vertex we have not yet seen, concatenating the depth-first walks. Note that, as in the case of a forest of Galton–Watson trees, we begin a new component every time the depth-first walk attains a new minimum. The following is essentially Theorem 3 of Aldous [8].

Theorem 13 (Aldous [8]) *Let X^n be the depth-first walk associated with $G\left(n, \frac{1}{n} + \frac{\lambda}{n^{4/3}}\right)$ and let N^n be the counting process of surplus edges (so that $N^n(k)$ is the number of surplus edges encountered by step k). Then the following convergences occur jointly:*

$$\frac{1}{n^{1/3}}\left(X^n(\lfloor n^{2/3}t\rfloor), t \geq 0\right) \xrightarrow{d} (W^\lambda(t), t \geq 0)$$

and

$$\left(N^n(\lfloor n^{2/3}t\rfloor), t \geq 0\right) \xrightarrow{d} (P^\lambda(t), t \geq 0)$$

as $n \to \infty$ (here both convergences are in the Skorokhod sense).

Proof (sketch). Write $v_0, v_1, \ldots$ for the vertices in the order we visit them, where $v_0 = 1$. Let $K^n(i)$ be the number of children of v_i, where children are *new* vertices we discover when we explore the neighbours of v_i. Then

$$X^n(k) = \sum_{i=0}^{k-1}(K^n(i) - 1).$$

Write $L^n(k) = -\inf_{0 \leq i \leq k} X^n(i)$. Then $L^n(k)$ is the number of components that we have fully explored before step k. When we visit v_i, there are $X^n(i) + L^n(i)$ vertices on the stack, and i vertices we have already visited. Hence, there are possible edges from v_i to any of the remaining $n - i - 1 - X^n(i) - L^n(i)$ vertices, each of which is present independently with probability $\frac{1}{n} + \frac{\lambda}{n^{4/3}}$. So, given $X^n(i)$ and $L^n(i)$,

$$K^n(i) \sim \mathrm{Bin}\left(n - i - 1 - X^n(i) - L^n(i), \frac{1}{n} + \frac{\lambda}{n^{4/3}}\right),$$

with

$$\begin{aligned}&\mathbb{E}\left[K^n(i) - 1 | X^n(0), \ldots, X^n(i)\right] \\ &\quad = \frac{\lambda}{n^{1/3}} - \frac{(i+1)}{n} - \frac{X^n(i) + L^n(i)}{n} - \frac{\lambda(i+1) + \lambda X^n(i) + \lambda L^n(i)}{n^{4/3}}.\end{aligned}$$

In particular, (X^n, L^n) is a time-inhomogeneous Markov process. Suppose now that $\sup_{0 \leq i \leq k} |X^n(i)| = o(n^{2/3})$ for $k = \Theta(n^{2/3})$. Then we may neglect the last two terms on the right-hand side to obtain that, for $i = \Theta(n^{2/3})$,

$$\mathbb{E}\left[K^n(i) - 1 | X^n(0), \ldots, X^n(i)\right] = \frac{\lambda}{n^{1/3}} - \frac{i}{n} + o(n^{-1/3}).$$

So, for $i = \Theta(n^{2/3})$, it is approximately the case that

$$X^n(i+1) - X^n(i) \sim \text{Po}\left(1 + \frac{\lambda}{n^{1/3}} - \frac{i}{n}\right) - 1.$$

Now let

$$M^n(k) = X^n(k) - \sum_{i=0}^{k-1} \left(\frac{\lambda}{n^{1/3}} - \frac{i}{n}\right) \approx X^n(k) - \frac{\lambda k}{n^{1/3}} + \frac{k^2}{2n}.$$

Then $(M^n(k), k \geq 0)$ is approximately a martingale. Since the step-sizes of X^n have variance approximately equal to 1, we may apply the martingale central limit theorem (Theorem 7.1.4 of [22]) to obtain

$$\left(\frac{1}{n^{1/3}} X^n(\lfloor tn^{2/3} \rfloor) - \lambda t + \frac{t^2}{2}, t \geq 0\right) \xrightarrow{d} (W(t), t \geq 0).$$

It follows by the continuous mapping theorem that

$$\frac{1}{n^{1/3}} \left(X^n(\lfloor tn^{2/3} \rfloor) + L^n(\lfloor tn^{2/3} \rfloor), t \geq 0\right) \xrightarrow{d} \left(W^\lambda(t) - \inf_{0 \leq s \leq t} W^\lambda(s), t \geq 0\right).$$

Turning now to the surplus edges, there is potentially an edge between the vertex v_k we are currently visiting and any one of the $X^n(k) + L^n(k)$ vertices on the stack, which cannot be children of v_k, since they are already children of elements of $\{v_0, \ldots, v_{k-1}\}$. Any such edges which are present thus contribute to the surplus of a component. Since such edges occur independently and with probability $\frac{1}{n} + \frac{\lambda}{n^{4/3}}$, we get a Binomial$(X^n(k) + L^n(k), \frac{1}{n} + \frac{\lambda}{n^{4/3}})$ number of surplus edges from v_k. Now the conditional mean of this random variable, which is approximately $(X^n(\lfloor tn^{2/3} \rfloor) + L^n(\lfloor tn^{2/3} \rfloor))/n$, is on the order of $n^{-2/3}$, and so we see points only on time-intervals of order $n^{2/3}$ steps. But this is precisely the timescale in which we are interested and so taking limits as $n \to \infty$, we obtain a Poisson point process of intensity $W^\lambda(t) - \inf_{0 \leq s \leq t} W^\lambda(s) = B^\lambda(t)$ at time $t \geq 0$. □

In the depth-first walk, we explore the components of the graph in size-biased random order. (Given the component sizes, vertex 1 lies in a component chosen with probability proportional to its size, and similarly for subsequent components.) In order to get from Theorem 13 to Theorem 12, one must show that in going from size-biased order to decreasing order, one doesn't lose track of any of the large components; see Section 2.3 of Aldous [8] for the details.

2.2 Component structures

We saw in Proposition 3 that the depth-first walk of a tree can be used to recover its structure. The excursions of the depth-first walk in the graph setting thus encode a collection of *spanning trees* for the components. In the scaling limit, this depth-first walk becomes a Brownian motion with parabolic drift. Let us first look a little deeper into the properties of this limit process.

By the Cameron–Martin–Girsanov theorem (e.g. Theorem 38.5 in Section IV of [36]), for any bounded measurable test-function $f : \mathcal{C}([0,t],\mathbb{R}) \to \mathbb{R}$, we have

$$\mathbb{E}\left[f(W^\lambda(s), 0 \le s \le t)\right] \\ = \mathbb{E}\left[\exp\left(\int_0^t (\lambda - s)\mathrm{d}W(s) - \frac{1}{2}\int_0^t (\lambda - s)^2 \mathrm{d}s\right) f(W(s), 0 \le s \le t)\right].$$

So W^λ is absolutely continuous with respect to a process which would encode a forest of Brownian CRT's. Consider the process of excursions of W^λ above its running minimum or, equivalently, the excursions of B^λ above 0. One can make sense of a sort of time-inhomogeneous excursion measure for B^λ, where the time-inhomogeneity manifests itself only in the *lengths* of the excursions. The shapes of the excursions themselves are absolutely continuous with respect to those of W. Indeed, let us write $\tilde{e}$ for an excursion of B^λ normalised to have length 1. Recall that e is a normalised Brownian excursion. Then for a bounded measurable test-function $g : \mathcal{C}([0,1],\mathbb{R}_+) \to \mathbb{R}$, we have

$$\mathbb{E}\left[g(\tilde{e})\right] = \frac{\mathbb{E}\left[g(e)\exp\left(\int_0^1 e(u)\mathrm{d}u\right)\right]}{\mathbb{E}\left[\exp\left(\int_0^1 e(u)\mathrm{d}u\right)\right]}.$$

More generally, for an excursion of length x (and an appropriately adjusted test-function), we have the usual Brownian scaling relation:

$$\mathbb{E}\left[g\left(\tilde{e}^{(x)}\right)\right] = \mathbb{E}\left[f\left(\sqrt{x}\tilde{e}(t/x), 0 \le t \le x\right)\right].$$

In summary, it is helpful to think of the reflected process B^λ in two parts: (a) it has excursions above 0 whose ordered lengths are the random variables $(C_1, C_2, \ldots)$ from Theorem 12 and (b), given these lengths, the excursions themselves are conditionally independent, with the ith longest having the same distribution as $\tilde{e}^{(C_i)}$. Part (a) of this decomposition is rather complicated, but taking that as an input, part (b) is comparatively simple.

The limiting picture for a single component corresponding to an excursion of length x is then as follows. We have an excursion $\tilde{e}^{(x)}$ (with law as above). Conditionally on $\tilde{e}^{(x)}$, the number of points falling under the excursion has $\mathrm{Po}\left(\int_0^x \tilde{e}^{(x)}(u)\mathrm{d}u\right)$ distribution. Let $\tilde{\mathcal{T}}^{(x)}$ be the $\mathbb{R}$-tree encoded by $2\tilde{e}^{(x)}$ and write $p : [0,x] \to \tilde{\mathcal{T}}^{(x)}$ for the canonical projection. (We will also write $\tilde{\mathcal{T}}$ in place of $\tilde{\mathcal{T}}^{(1)}$.) A point $(t, y/2)$ under $\tilde{e}^{(x)}$ identifies the vertex $p(t)$ which is at height $2\tilde{e}^{(x)}$ with the vertex at distance y along the path from the root to v in $\tilde{\mathcal{T}}^{(x)}$; we then endow the resulting space with the quotient metric. See Figure 2 for an illustration.

Let $\mathcal{C}_1^n, \mathcal{C}_2^n, \ldots$ be the sequence of components of $G(n, \frac{1}{n} + \frac{\lambda}{n^{4/3}})$, listed in decreasing order of size. In order to keep the notation compact, for $a > 0$, write $a\mathcal{C}_i^n$ for the metric space formed by $\mathcal{C}_i^n$ endowed with the graph distance rescaled by a.

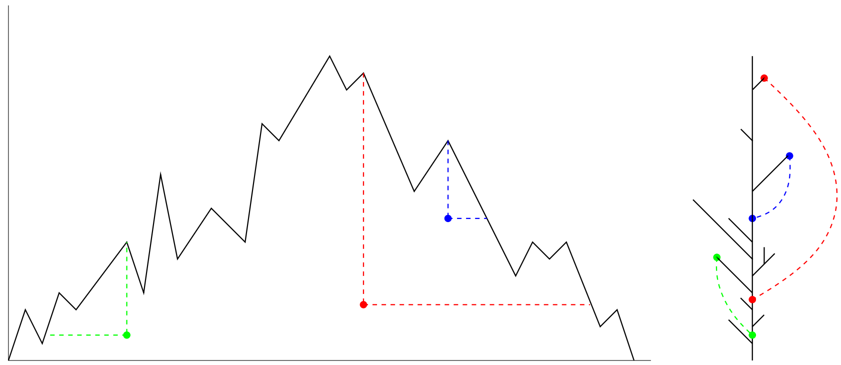


Fig. 2. Left: an excursion with three points. Right: the corresponding $\mathbb{R}$-tree with vertex-identifications (identify any pair of points joined by a dashed line).

Theorem 14 (Addario-Berry, Broutin, G. [3]) *As $n \to \infty$,*

$$n^{-1/3}(\mathcal{C}_1^n, \mathcal{C}_2^n, \ldots) \xrightarrow{d} (\mathcal{C}_1, \mathcal{C}_2, \ldots),$$

where $\mathcal{C}_1, \mathcal{C}_2, \ldots$ is the sequence of random compact metric spaces corresponding to the excursions of Aldous' marked limit process B^λ in decreasing order of length.

The convergence here is with respect to the distance

$$\mathrm{dist}(\mathcal{A}, \mathcal{B}) = \left(\sum_{i=1}^{\infty} d_{\mathrm{GH}}(\mathcal{A}_i, \mathcal{B}_i)^4 \right)^{1/4},$$

where $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \ldots)$ and $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2, \ldots)$ are sequences of compact metric spaces.

Proof (sketch). Consider a component G of $G(n,p)$, conditioned to have a vertex set of size m (we take $[m]$ for simplicity). To any such component, we may associate a canonical spanning tree $T(G)$, called the *depth-first tree* of G: this is the tree we pick out when we do our depth-first walk, for which we write $(X(k), 0 \leq k \leq m)$.

Given a fixed tree $T \in \mathbb{T}_m$, which connected graphs G have $T(G) = T$? In other words, where might we put surplus edges into T such that we don't change the depth-first tree? Call any such edges *permitted*. It is straightforward to see that there are precisely $X(k)$ permitted edges at step k: one between v_k and each of the vertices which have been seen but not yet fully explored. So there are

$$a(T) := \sum_{k=0}^{m-1} X(k)$$

permitted edges in total. We call this the *area* of T. Now let

$$\mathbb{G}_T = \{\text{graphs } G \text{ such that } T(G) = T\}.$$

Then

$$\{\mathbb{G}_T : T \in \mathbb{T}_m\}$$

is a partition of the set of connected graphs on $[m]$. Moreover, $|\mathbb{G}_T| = 2^{a(T)}$, since each permitted edge may either be included or not.

Exercise 6 *Let $\tilde{G}_m^p$ be a connected graph on vertices labelled by $[m]$ generated as follows:*

- *Pick a random tree $\tilde{T}_m^p$ such that*

$$\mathbb{P}\left(\tilde{T}_m^p = T\right) \propto (1-p)^{-a(T)}, \quad T \in \mathbb{T}_m.$$

- *Add each of the $a(\tilde{T}_m^p)$ permitted edges independently with probability p.*

Show that $\tilde{G}_m^p$ has the same distribution as a component of $G(n,p)$ conditioned to have vertex set $[m]$.

It remains now to show that, for $m \sim xn^{2/3}$ and $p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$, we have

- $n^{-1/3}\tilde{T}_m^p \xrightarrow{d} \tilde{\mathcal{T}}^{(x)}$
- the locations of the surplus edges converge to the locations in the limiting picture.

For simplicity, let us take $x = 1$ and $\lambda = 0$, so that $p = m^{-3/2}$. (The general case is similar.) Write $\tilde{X}^m$ for the depth-first walk of $\tilde{T}_m^p$, and let $\tilde{H}^m$ be the corresponding height process defined, as usual, from $\tilde{X}^m$ via the relation (1). Then

$$a(\tilde{T}_m^p) = \int_0^m \tilde{X}^m(\lfloor s \rfloor)\mathrm{d}s = m\int_0^1 \tilde{X}^m(\lfloor mt \rfloor)\mathrm{d}t,$$

by a simple change of variables in the integral.

If T_m is a uniform random tree on $[m]$ and X^m is its depth-first walk, we know from Theorem 5 that

$$(m^{-1/2}X^m(\lfloor mt \rfloor), 0 \le t \le 1) \xrightarrow{d} (e(t), 0 \le t \le 1). \tag{3}$$

Moreover, by Exercise 6, for a bounded continuous test-function f,

$$\begin{aligned}
&\mathbb{E}\left[f(m^{-1/2}\tilde{X}^m(\lfloor mt \rfloor), 0 \le t \le 1)\right] \\
&\quad = \frac{\mathbb{E}\left[f(m^{-1/2}X^m(\lfloor mt \rfloor), 0 \le t \le 1)(1-p)^{-m\int_0^1 X^m(\lfloor mu \rfloor)\mathrm{d}u}\right]}{\mathbb{E}\left[(1-p)^{-m\int_0^1 X^m(\lfloor mu \rfloor)\mathrm{d}u}\right]} \\
&\quad = \frac{\mathbb{E}\left[f(m^{-1/2}X^m(\lfloor mt \rfloor), 0 \le t \le 1)(1-m^{-3/2})^{-m^{3/2}\int_0^1 m^{-1/2}X^m(\lfloor mu \rfloor)\mathrm{d}u}\right]}{\mathbb{E}\left[(1-m^{-3/2})^{-m^{3/2}\int_0^1 m^{-1/2}X^m(\lfloor mu \rfloor)\mathrm{d}u}\right]}.
\end{aligned}$$

We have

$$(1 - m^{-3/2})^{-m^{3/2} \int_0^1 m^{-1/2} X^m(\lfloor mu \rfloor) \mathrm{d}u} \xrightarrow{d} \exp\left(\int_0^1 e(u) \mathrm{d}u \right)$$

as $m \to \infty$, by (3) and the continuous mapping theorem. The sequence of random variables on the left-hand side may be shown to be uniformly integrable (see Lemma 14 of [3]) and so

$$\mathbb{E}\left[f(m^{-1/2} \tilde{X}^m(\lfloor mt \rfloor), 0 \le t \le 1) \right] \to \mathbb{E}\left[f(\tilde{e}) \right]$$

as $m \to \infty$. Similar reasoning then gives that

$$\mathbb{E}\left[f(m^{-1/2} \tilde{H}^m(\lfloor mt \rfloor), 0 \le t \le 1) \right] \to \mathbb{E}\left[f(2\tilde{e}) \right],$$

which implies (by the same argument as in the proof of Theorem 11) that

$$\frac{1}{\sqrt{m}} \tilde{T}_m^p \xrightarrow{d} \tilde{\mathcal{T}}$$

as $m \to \infty$, in the Gromov–Hausdorff sense.

Now consider the surplus edges. It is straightforward to see that there is a bijection between permitted edges and integer points under the graph of the depth-first walk. A point at (k, ℓ) means "put an edge between v_k and the vertex at distance ℓ from the bottom of the stack". Since each permitted edge is present independently with probability p, the surplus edges form a point process, which converges on rescaling to our Poisson point process. Finally, surplus edges always join v_k and a younger child of some ancestor of v_k. In the limit, the distance between a vertex and its children vanishes, so that surplus edges are effectively to ancestors. □

3 Critical random graphs with i.i.d. random degrees

We have seen that degrees in the Erdős–Rényi model are approximately Poisson distributed. In recent years, there has been much interest in modelling settings where this is certainly not the case. We will discuss one popular model which has arisen, with the principal aim of demonstrating that the results in the previous section are universal. We will restrict our attention to component sizes.

3.1 The configuration model

Suppose we wish to generate a graph G_n uniformly at random from those with vertex set $[n]$ and such that vertex i has degree vertex d_i, where $d_i \ge 1$ for $1 \le i \le n$ and $\ell_n = \sum_{i=1}^n d_i$ is even.

Assign d_i half-edges to vertex i. Label the half-edges in some arbitrary way by $1, 2, \ldots, \ell_n$. Then generate a uniformly random pairing of the half-edges to create full edges.

Clearly this may produce self-loops (edges whose endpoints are the same vertex) or multiple edges between the same pair of vertices, so in general the configuration model produces a multigraph, M_n. Assuming that there exists at least one simple graph with the given degrees then, conditionally on the event $\{M_n \text{ is a simple graph}\}$, M_n has the same law as G_n. This is a consequence of the following exercise.

Exercise 7 *Fix a degree sequence $d_1, \ldots, d_n$. Show that the probability of generating a particular multigraph G with these degrees is*

$$\frac{1}{(\ell_n - 1)!!} \frac{\prod_{i=1}^n d_i!}{2^{\mathrm{sl}(G)} \prod_{e \in E(G)} \mathrm{mult}(e)!},$$

where $\ell_n = \sum_{i=1}^n d_i$, $\mathrm{sl}(G)$ is the number of self-loops in G and $\mathrm{mult}(e)$ is the multiplicity of the edge $e \in E(G)$. (We recall the double factorial notation $n!! := \prod_{k=0}^{\lceil n/2 \rceil - 1} (n - 2k)$.)

We will take the degrees themselves to be random: let $D_1, D_2, \ldots, D_n$ be i.i.d. with the same law as some random variable D which has finite variance. We resolve the issue of $\sum_{i=1}^n D_i$ potentially being odd by simply throwing the last half-edge away when we generate the pairing in that case.

Let $\gamma = \mathbb{E}[D(D-1)] / \mathbb{E}[D]$. Then

$$\mathbb{P}(M_n \text{ is simple}) \to \exp\left(-\frac{\gamma}{2} - \frac{\gamma^2}{4}\right) > 0$$

(see Theorem 7.12 of van der Hofstad [25]), so that conditioning on simplicity will make sense for large n.

Theorem 15 (Molloy and Reed [33]) *If $\gamma < 1$ then, with probability tending to 1 as $n \to \infty$, there is no giant component; if $\gamma > 1$ then, with probability tending to 1 as $n \to \infty$, there is a unique giant component.*

Let us give an heuristic argument for why $\gamma = 1$ should be the critical point. An important point is that we may generate the pairing of the half-edges one by one, in any order that is convenient. So we will generate and explore the graph at the same time. Perform a depth-first exploration from an arbitrary vertex. Consider the first half-edge attached to the vertex (the one with the smallest half-edge label). It picks its pair uniformly from all those available, and so in particular it picks a half-edge belonging to a vertex chosen with probability proportional to its degree. The same will be true of subsequent half-edges. As long as we have not explored much of the graph, these degrees should have law close to the *size-biased distribution*, given by

$$\mathbb{P}(D^* = k) = \frac{k\mathbb{P}(D = k)}{\mathbb{E}[D]}, \quad k \geq 1.$$

Hence, we can compare to a branching process with offspring distribution D^*-1, which has expectation

$$\mathbb{E}\left[D^*-1\right]=\frac{\mathbb{E}\left[D^2\right]}{\mathbb{E}\left[D\right]}-1=\gamma.$$

The following exercise gives an idea of why Poisson degrees are particularly nice.

Exercise 8 *Suppose that D is a non-negative integer-valued random variable with finite mean, and let D^* have the size-biased distribution*

$$\mathbb{P}\left(D^*=k\right)=\frac{k\mathbb{P}\left(D=k\right)}{\mathbb{E}\left[D\right]},\quad k\geq 1.$$

Show that $D^*-1\stackrel{d}{=}D$ if and only if D has a Poisson distribution.

3.2 Scaling limit for the critical component sizes

We will henceforth consider the configuration model with the following set-up: the degrees $D_1, D_2, \ldots, D_n$ are i.i.d. with the same distribution as D such that

- $\mathbb{P}\left(D\geq 1\right)=1$, $\mathbb{P}\left(D=2\right)<1$;
- $\gamma=\frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}=1$;
- $\mathbb{E}\left[D^3\right]<\infty$.

We write $\mu=\mathbb{E}\left[D\right]$ and $\beta=\mathbb{E}\left[D(D-1)(D-2)\right]$. We observe immediately that $\mathbb{E}\left[D^*\right]=2$ and that $\mathrm{var}\left(D^*\right)=\beta/\mu$.

The analogue of Theorem 13 in this setting is as follows.

Theorem 16 (Riordan [35], Joseph [26]) *Let $C_1^{;}C_2^n,\ldots$ be the ordered component sizes of M_n or G_n. Then*

$$n^{-2/3}(C_1^n, C_2^n,\ldots)\stackrel{d}{\to}(C_1, C_2,\ldots)$$

in $\ell^2_{\downarrow}$, where the limit is given by the ordered sequence of excursion-lengths above past-minima of the process $(W^{\beta,\mu}(t), t\geq 0)$ defined by

$$W^{\beta,\mu}(t):=\sqrt{\frac{\beta}{\mu}}W(t)-\frac{\beta t^2}{2\mu^2},\quad t\geq 0,$$

where W is a standard Brownian motion.

We give a sketch of a proof of this result, and refer the reader to Theorem 2.1 of [26] for a complete proof. To start with, we make precise the connection between a collection of i.i.d. random variables in size-biased random order and a collection of i.i.d. random variables with the size-biased distribution.

Exercise 9 *Suppose that $D_1, D_2, \ldots, D_n$ are i.i.d. random variables with finite mean μ, and let $(\hat{D}_1^n, \hat{D}_2^n, \ldots, \hat{D}_n^n)$ be the same random variables in size-biased random order. That is, given $D_1, D_2, \ldots, D_n$, let Σ be a permutation of $[n]$ with conditional distribution*

$$\mathbb{P}(\Sigma = \sigma | D_1, D_2, \ldots, D_n) = \frac{D_{\sigma(1)}}{\sum_{j=1}^n D_{\sigma(j)}} \frac{D_{\sigma(2)}}{\sum_{j=2}^n D_{\sigma(j)}} \cdots \frac{D_{\sigma(n)}}{D_{\sigma(n)}}, \quad \sigma \in \mathfrak{S}_n.$$

Then define

$$(\hat{D}_1^n, \hat{D}_2^n, \ldots, \hat{D}_n^n) = (D_{\Sigma(1)}, D_{\Sigma(2)}, \ldots, D_{\Sigma(n)}).$$

Now let $D_1^*, D_2^*, \ldots$ be i.i.d. with the (true) size-biased distribution. Show that for $m < n$ and $d_1, d_2, \ldots, d_m \geq 1$,

$$\mathbb{P}\left(\hat{D}_1^n = d_1, \hat{D}_2^n = d_2, \ldots, \hat{D}_m^n = d_m\right) \\ = \phi_m^n(d_1, d_2, \ldots, d_m)\mathbb{P}(D_1^* = d_1, D_2^* = d_2, \ldots, D_m^* = d_m),$$

where

$$\phi_m^n(d_1, d_2, \ldots, d_m) := \frac{n!\mu^m}{(n-m)!}\mathbb{E}\left[\prod_{i=1}^m \frac{1}{\sum_{j=i}^m d_j + \Delta_{n-m}}\right],$$

and $\Delta_{n-m} \stackrel{d}{=} D_1 + \ldots + D_{n-m}$.

Proof (sketch). We again use a depth-first walk but this time with a stack of unpaired half-edges. Start by picking a vertex with probability proportional to its degree. Declare one of its half-edges to be active and put the rest on the stack. Sample the active half-edge's pair (either on the stack or not) and remove both from further consideration. If we discovered a new vertex, add its remaining half-edges to the top of the stack. Then declare whichever half-edge is now on top of the stack to be active. If ever the stack becomes empty, pick a new vertex with probability proportional to its degree and continue.

In this procedure, we observe the vertex-degrees precisely in size-biased random order. Let $\tilde{X}(0) = 0$ and

$$\tilde{X}^n(k) := \sum_{i=1}^k (\hat{D}_i^n - 2), \quad k \geq 1.$$

Then $\tilde{X}^n$ behaves exactly like the depth-first walk except

- at the start of a component, where we should add $\hat{D}_i^n - 1$ rather than $\hat{D}_i^n - 2$ and
- whenever we pair the active half-edge with one on the stack.

Neither problem shows up in the limit (although showing this properly is somewhat technical). For the purposes of this sketch, we shall ignore the difference. We write $X^*(0) = 0$ and let

$$X^*(i) = \sum_{j=1}^i (D_j^* - 2)$$

be a similar process built instead from i.i.d. size-biased random variables. Note that X^* is a centred random walk with step-variance equal to β/μ. In particular, by Donsker's theorem,

$$n^{-1/3}(X^*(\lfloor n^{2/3}s \rfloor), s \geq 0) \xrightarrow{d} \sqrt{\frac{\beta}{\mu}}(W(s), s \geq 0) \tag{4}$$

as $n \to \infty$. We aim to show that

$$n^{-1/3}(\tilde{X}(\lfloor n^{2/3}s \rfloor), s \geq 0) \xrightarrow{d} (W^{\beta,\mu}(s), s \geq 0).$$

By the Cameron–Martin–Girsanov theorem and integration by parts, for suitable test-functions f,

$$\begin{aligned}
&\mathbb{E}\left[f(W^{\beta,\mu}(s), 0 \leq s \leq t)\right] \\
&\mathbb{E}\left[\exp\left(-\sqrt{\frac{\beta}{\mu^3}}\int_0^t s \mathrm{d}W(s) - \frac{1}{2}\frac{\beta}{\mu^3}\int_0^t s^2 \mathrm{d}s\right) f\left(\sqrt{\frac{\beta}{\mu}}(W(s), 0 \leq s \leq t)\right)\right] \\
&= \mathbb{E}\left[\exp\left(\sqrt{\frac{\beta}{\mu^3}}\int_0^t (W(s) - W(t))\mathrm{d}s - \frac{\beta t^3}{6\mu^3}\right) f\left(\sqrt{\frac{\beta}{\mu}}(W(s), 0 \leq s \leq t)\right)\right].
\end{aligned} \tag{5}$$

Exercise 9 gives us a way to obtain a discrete analogue of this change of measure. Write $x(i) = \sum_{j=1}^{i}(d_j - 2)$. Then we may rewrite

$$\begin{aligned}
&\phi_m^n(d_1, d_2, \ldots, d_m) \\
&= \frac{n!}{(n-m)!n^m}\mathbb{E}\left[\prod_{i=1}^{m}\frac{n\mu}{x(m) - x(i-1) + 2(m-i+1) + \Delta_{n-m}}\right] \\
&= \prod_{i=1}^{m-1}\left(1 - \frac{i}{n}\right)\mathbb{E}\left[\prod_{i=1}^{m}\frac{n\mu}{\Delta_{n-m} + 2(m-i+1) + x(m) - x(i-1)}\right] \\
&= \exp\left(\sum_{i=1}^{m-1}\log\left(1 - \frac{i}{n}\right)\right) \\
&\qquad \times \mathbb{E}\left[\exp\left(-\sum_{i=1}^{m}\log\left(\frac{\Delta_{n-m} + 2(m-i+1) + x(m) - x(i-1)}{n\mu}\right)\right)\right].
\end{aligned}$$

Taylor expanding the logarithms, we get that this is approximately equal to

$$\begin{aligned}
&\exp\left(-\sum_{i=1}^{m-1}\left[\frac{i}{n}+\frac{i^2}{2n^2}\right]-\sum_{i=1}^{m}\left[\frac{2(m-i+1)}{n\mu}-\frac{2(m-i+1)^2}{n^2\mu^2}\right]+m\right)\\
&\qquad\times\exp\left(-\frac{1}{n\mu}\sum_{i=1}^{m}(x(m)-x(i-1))\right)\mathbb{E}\left[\exp\left(-\frac{m}{n\mu}\Delta_{n-m}\right)\right]\\
&\approx\exp\left(\frac{1}{n\mu}\sum_{i=1}^{m-1}(x(i)-x(m))+m-\frac{(2+\mu)m^2}{2\mu n}+\frac{(2+\mu)(2-\mu)m^3}{6\mu^2n^2}\right)\\
&\qquad\times\mathbb{E}\left[\exp\left(-\frac{m}{n\mu}D_1\right)\right]^{n-m}.
\end{aligned}$$

Using the moments of D_1, it is straightforward to show that its Laplace transform has the following asymptotic behaviour:

$$\mathbb{E}\left[\exp(-\theta D_1)\right]=\exp\left(-\theta\mu+\frac{\theta^2\mu(2-\mu)}{2}-\frac{\theta^3}{6}(\beta+4\mu-6\mu^2+2\mu^3)+o(\theta^3)\right),$$

as $\theta\downarrow 0$. Putting all of this together, almost everything cancels and we get that for $m=\lfloor tn^{2/3}\rfloor$,

$$\phi_m^n(D_1^*,D_2^*,\ldots,D_m^*)\approx\exp\left(\frac{1}{n\mu}\sum_{i=1}^{m-1}(X^*(i)-X^*(m))-\frac{\beta t^3}{6\mu^3}\right).$$

More work gets uniform integrability, and then we may conclude using (4) and the continuous mapping theorem that

$$\begin{aligned}
&\mathbb{E}\left[f\left(n^{-1/3}\tilde{X}(\lfloor n^{2/3}s\rfloor),0\le s\le t\right)\right]\\
&=\mathbb{E}\left[\phi_m^n(D_1^*,D_2^*,\ldots,D_m^*)f\left(n^{-1/3}X^*(\lfloor n^{2/3}s\rfloor),0\le s\le t\right)\right]\\
&\to\mathbb{E}\left[\exp\left(\frac{1}{\mu}\int_0^t\sqrt{\frac{\beta}{\mu}}(W(s)-W(t))\mathrm{d}s-\frac{\beta t^3}{6\mu^3}\right)f\left(\sqrt{\frac{\beta}{\mu}}(W(u),0\le s\le t)\right)\right]\\
&\qquad=\mathbb{E}\left[f(W^{\beta,\mu}(s),0\le s\le t)\right],
\end{aligned}$$

where the last equality holds by (5).

Finally, it is possible to show that when exploring M_n, the first loop or multiple edge occurs at a time which is $\gg n^{2/3}$ and so the same distributional convergence holds if we condition on simplicity. □

4 Sources for these notes and suggested further reading

David Aldous' series of papers [5–7] and Jean-François Le Gall's paper [28] on the Brownian CRT remain an excellent source of inspiration. I have used Aldous' approach from [5, 6] in Section 1.1. The survey paper [6] gives what Aldous refers to as The Big Picture and is a great place to start reading about the Brownian CRT. I first learnt much of the material in Section 1 from a wonderful DEA lecture course at Paris VI in 2003 given by Jean-François Le Gall. These notes borrow heavily from his excellent survey of random trees [29].

To learn about generalisations of the Brownian CRT, in particular the so-called Lévy trees, see the monograph of Duquesne and Le Gall [21]. The *stable trees* are the scaling limits of critical Galton–Watson trees with offspring distributions which do not have finite variance, but instead lie in the domain of attraction of an α-stable law, for $\alpha \in (1, 2)$. They have particularly nice properties, including a line-breaking construction [23].

For those looking to learn about random graphs, I warmly recommend Remco van der Hofstad's recent book [25]. David Aldous' paper [8] on the critical Erdős–Rényi random graph and the multiplicative coalescent is essential reading for Section 2.1. Jim Pitman's St-Flour course [34] contains much complementary material. Section 2.2 is based on joint work with Louigi Addario-Berry and Nicolas Broutin [3]. To learn more about the properties of the metric space scaling limit of the Erdős–Rényi random graph, see the companion paper [2]. The results in these two papers played a key role in our proof, jointly with Grégory Miermont, of a scaling limit for the minimum-spanning tree of the complete graph endowed with i.i.d. random edge-weights from a continuous distribution [4].

The component sizes in the configuration model with critical i.i.d. degrees, as treated in Section 3, were studied by Joseph [26]; the component sizes and surpluses in a more general set-up were studied independently by Riordan [35]. See also Dhara, van der Hofstad, van Leeuwaarden and Sen [20]. The sketch proof of Theorem 16 presented here is based on joint work in progress with Guillaume Conchon-Kerjan [18]. The corresponding metric space scaling limit has been proved by Bhamidi and Sen [14]. In recent years, several other critical random graph models have been shown to possess the same scaling limit as the critical Erdős–Rényi random graph (either for the component sizes, or for the full component structures). See [9, 10, 12, 14, 15]. The case where the degree distribution does not have a third moment is more complicated and gives different scaling limits; see [11, 13, 16, 19, 26].

5 Acknowledgements

My work is supported by EPSRC Fellowship EP/N004833/1. I am very grateful to two referees, whose careful reading of these notes considerably improved them. Thanks also to Louigi Addario-Berry for the picture of B^λ in Section 2.1. I would like to thank the organisers for doing a wonderful job of putting on the summer school and for making everyone feel so welcome, and to the participants for being such an interested and involved audience. Special thanks to Omer Angel for his hospitality and to Lior Silberman for the bike!

References

1. Abraham, R., Delmas, J.F., Hoscheit, P.: A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. Electron. J. Probab. **18**, no. 14, 21 (2013)
2. Addario-Berry, L., Broutin, N., Goldschmidt, C.: Critical random graphs: limiting constructions and distributional properties. Electron. J. Probab. **15**, 741–775, paper no. 25 (electronic) (2010)
3. Addario-Berry, L., Broutin, N., Goldschmidt, C.: The continuum limit of critical random graphs. Probab. Theory Related Fields **152**(3–4), 367–406 (2012)
4. Addario-Berry, L., Broutin, N., Goldschmidt, C., Miermont, G.: The scaling limit of the minimum spanning tree of the complete graph. Ann. Probab. **45**(5), 3075–3144 (2017)
5. Aldous, D.: The continuum random tree. I. Ann. Probab. **19**(1), 1–28 (1991)
6. Aldous, D.: The continuum random tree. II. An overview. In: Stochastic analysis (Durham, 1990), *London Math. Soc. Lecture Note Ser.*, vol. 167, pp. 23–70. Cambridge Univ. Press, Cambridge (1991)
7. Aldous, D.: The continuum random tree. III. Ann. Probab. **21**(1), 248–289 (1993)
8. Aldous, D.: Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. **25**(2), 812–854 (1997)
9. Bhamidi, S., Broutin, N., Sen, S., Wang, X.: Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph (2014+). `arXiv:1411.3417 [math.PR]`
10. Bhamidi, S., Budhiraja, A., Wang, X.: The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. Probab. Theory Related Fields **160**(3-4), 733–796 (2014)
11. Bhamidi, S., Dhara, S., van der Hofstad, R., Sen, S.: Universality for critical heavy-tailed network models: metric structure of maximal components (2017). `arXiv:1703.07145 [math.PR]`
12. Bhamidi, S., van der Hofstad, R., van Leeuwaarden, J.S.H.: Scaling limits for critical inhomogeneous random graphs with finite third moments. Electron. J. Probab. **15**, no. 54, 1682–1703 (2010)
13. Bhamidi, S., van der Hofstad, R., van Leeuwaarden, J.S.H.: Novel scaling limits for critical inhomogeneous random graphs. Ann. Probab. **40**(6), 2299–2361 (2012)
14. Bhamidi, S., Sen, S.: Geometry of the vacant set left by random walk on random graphs, Wright's constants, and critical random graphs with prescribed degrees (2016+). `arXiv:1608.07153 [math.PR]`
15. Bhamidi, S., Sen, S., Wang, X.: Continuum limit of critical inhomogeneous random graphs. Probab. Theory Related Fields **169**(1-2), 565–641 (2017)
16. Broutin, N., Duquesne, T., Wang, M.: Limits of multiplicative inhomogeneous random graphs and Lévy trees (2018+). `arXiv:1804.05871 [math.PR]`
17. Burago, D., Burago, Y., Ivanov, S.: A course in metric geometry, *Graduate Studies in Mathematics*, vol. 33. American Mathematical Society, Providence, RI (2001)
18. Conchon-Kerjan, G., Goldschmidt, C.: Stable graphs: the metric space scaling limits of critical random graphs with i.i.d. power-law degrees (2018+). In preparation.
19. Dhara, S., van der Hofstad, R., van Leeuwaarden, J., Sen, S.: Heavy-tailed configuration models at criticality (2016+). `arXiv:1612.00650 [math.PR]`
20. Dhara, S., van der Hofstad, R., van Leeuwaarden, J., Sen, S.: Critical window for the configuration model: finite third moment degrees. Electron. J. Probab. **22**, Paper No. 16, 33 (2017)

21. Duquesne, T., Le Gall, J.F.: Random trees, Lévy processes and spatial branching processes. Astérisque **281**, vi+147 (2002)
22. Ethier, S.N., Kurtz, T.G.: Markov processes: Characterization and convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York (1986)
23. Goldschmidt, C., Haas, B.: A line-breaking construction of the stable trees. Electron. J. Probab. **20**, no. 16, 24 (2015)
24. Haas, B., Miermont, G.: The genealogy of self-similar fragmentations with negative index as a continuum random tree. Electron. J. Probab. **9**, no. 4, 57–97 (2004)
25. van der Hofstad, R.: Random graphs and complex networks: volume I. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (2016)
26. Joseph, A.: The component sizes of a critical random graph with given degree sequence. Ann. Appl. Probab. **24**(6), 2560–2594 (2014)
27. Kaigh, W.: An invariance principle for random walk conditioned by a late return to zero. Ann. Probab. **4**(1), 115–121 (1976)
28. Le Gall, J.F.: The uniform random tree in a Brownian excursion. Probability Theory and Related Fields **96**(3), 369–383 (1993)
29. Le Gall, J.F.: Random trees and applications. Probab. Surv. **2**, 245–311 (2005)
30. Le Gall, J.F.: Random real trees. Ann. Fac. Sci. Toulouse Math. (6) **15**(1), 35–62 (2006)
31. Le Gall, J.F., Miermont, G.: Scaling limits of random trees and planar maps. In: Probability and statistical physics in two and more dimensions, *Clay Math. Proc.*, vol. 15, pp. 155–211. Amer. Math. Soc., Providence, RI (2012)
32. Marckert, J.F., Mokkadem, A.: The depth first processes of Galton-Watson trees converge to the same Brownian excursion. Ann. Probab. **31**(3), 1655–1678 (2003)
33. Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Structures & Algorithms **6**(2-3), 161–180 (1995)
34. Pitman, J.: Combinatorial stochastic processes, *Lecture Notes in Mathematics*, vol. 1875. 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
35. Riordan, O.: The phase transition in the configuration model. Combin. Probab. Comput. **21**(1-2), 265–299 (2012)
36. Rogers, L.C.G., Williams, D.: Diffusions, Markov processes, and martingales. Vol. 2. Cambridge Mathematical Library. Cambridge University Press, Cambridge (2000). Itô calculus, Reprint of the second (1994) edition