Asymptotic of trees with a prescribed degree sequence and applications

Nicolas Broutin* Jean-François Marckert†

October 24, 2011

The second author is partially supported by the ANR-08-BLAN-0190-04 A3.

Abstract

Let t be a rooted tree and $n_i(t)$ the number of nodes in t having i children. The degree sequence $(n_i(t), i \geq 0)$ of t satisfies $\sum_{i\geq 0} n_i(t) = 1 + \sum_{i\geq 0} i n_i(t) = |t|$, where $|t|$ denotes the number of nodes in t. In this paper, we consider trees sampled uniformly among all trees having the same degree sequence $\mathbf{s}$; we write $\mathbb{P}_{\mathbf{s}}$ for the corresponding distribution. Let $\mathbf{s}(\kappa) = (n_i(\kappa), i \geq 0)$ be a list of degree sequences indexed by κ corresponding to trees with size $\mathsf{n}_\kappa \to +\infty$. We show that under some simple and natural hypotheses on $(\mathbf{s}(\kappa), \kappa > 0)$ the trees sampled under $\mathbb{P}_{\mathbf{s}(\kappa)}$ converge to the Brownian continuum random tree after normalisation by $\mathsf{n}_\kappa^{1/2}$. Some applications concerning Galton–Watson trees and coalescence processes are provided.

1 Introduction

Let t be a rooted tree and $n_i(t)$ the number of nodes in t having i children. The sequence $(n_i(t), i \geq 0)$ is called the degree sequence of t, and satisfies $\sum_{i\geq 0} n_i(t) = \sum_{i\geq 0} i n_i(t) = |t|$, the number of nodes in t.

The aim of this paper is to study trees chosen under $\mathbb{P}_{\mathbf{s}}$, the uniform distribution on the set of trees with specified degree sequence $\mathbf{s} = (n_i, i \geq 0)$, and then size $|\mathbf{s}| := \sum_{i\geq 0} n_i$. More precisely, a sequence of degree sequences $(\mathbf{s}(\kappa), \kappa \geq 0)$ with $\mathbf{s}(\kappa) = (n_i(\kappa), i \geq 0)$, corresponding to trees with size $\mathsf{n}_\kappa := |\mathbf{s}(\kappa)| \to +\infty$ is given, and the investigations concern the limiting behavior of tree under $\mathbb{P}_{\mathbf{s}(\kappa)}$.

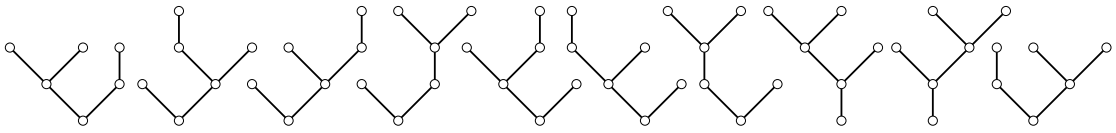

Figure 1: The 10 trees of $\mathbb{T}_{\mathbf{s}}$ for the degree sequence $\mathbf{s} = (3, 1, 2, 0, 0, \dots)$.

We now introduce some notation valid all along the paper. We denote by $\mathbf{p}(\kappa) = (p_i(\kappa), i \geq 0)$ the degree distribution under $\mathbb{P}_{\mathbf{s}(\kappa)}$:

$$p_i(\kappa) = \frac{n_i(\kappa)}{\mathsf{n}_\kappa}, \tag{1}$$

*Projet Algorithms, Inria Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay - France. Email: nicolas.broutin@inria.fr

†LaBRI, Université de Bordeaux - CNRS, 351 cours de la Libération, 33405 Talence - France. Email: marckert@labri.fr

and by

$$\sigma_\kappa^2 := \sum_{i\geq 1} \frac{n_i(\kappa)}{\mathsf{n}_\kappa} i^2 - \left(1 - \frac{1}{\mathsf{n}_\kappa}\right)^2 \tag{2}$$

the associated variance. The maximum degree of any tree with degree sequence $\mathbf{s}(\kappa)$ is

$$\Delta_\kappa = \max\{i : n_i(\kappa) > 0\}.$$

All along the paper $\mathbf{p} = (p_i, i \geq 0)$ is a distribution with mean 1, and variance $\sigma_{\mathbf{p}}^2 \in (0, +\infty) = \sum_{i\geq 0} i^2 p_i - 1 \in (0, \infty)$. In the following theorem, the main result of the present paper, the convergence $\mathbf{p}(\kappa) \Rightarrow \mathbf{p}$ is the standard pointwise convergence, $p_i(\kappa) \to p_i$, as $\kappa \to \infty$, which is equivalent to the convergence in distribution of $\mathbf{p}(\kappa)$ to $\mathbf{p}$.

Theorem 1. *Let $(\mathbf{s}(\kappa), \kappa \geq 0)$ be a sequence of degree sequences such that $\mathsf{n}_\kappa \to +\infty$, $\Delta_\kappa = o(\mathsf{n}_\kappa^{1/2})$, $\mathbf{p}(\kappa) \Rightarrow \mathbf{p}$ with $\sigma_\kappa^2 \to \sigma_{\mathbf{p}}^2$, that is convergence of second moment. Let $\mathbf{t}$ be a plane tree chosen under $\mathbb{P}_{\mathbf{s}(\kappa)}$ and let $d_{\mathbf{t}}$ be the graph distance in $\mathbf{t}$. Under $\mathbb{P}_{\mathbf{s}(\kappa)}$, when $\kappa \to +\infty$, $(\mathbf{t}, \sigma_\kappa \mathsf{n}_\kappa^{-1/2} d_{\mathbf{t}})$ converges in distribution to Aldous's continuum random tree (encoded by twice a Brownian excursion), in the Gromov–Hausdorff sense.*

First observe that the very strong result of Haas and Miermont [26] about the asymptotics of Markov branching trees that has been used to give asymptotics for random trees in a wide variety of settings does not apply in the present case of trees with a prescribed degree sequence. Indeed, the subtrees of a given node are not independent given their sizes when one fixes the degree sequence.Our approach uses the observation by Marckert and Mokkadem [36] that all natural encodings of the trees are asymptotically proportional. In particular, the *height process* or the *contour process* both encoding the metric structure of the tree resemble the *depth-first queue process* encoding the sequence of degrees observed when performing a depth-first traversal. This fact was used by Marckert and Mokkadem [36] to give an alternative proof of Aldous' result in the case of Galton–Watson trees conditioned on the total progeny under some moment condition (Bennies and Kersting [13] also observed this phenomenon).

One of the crucial questions underlying our work is that of the universality of the convergence of random trees to the continuum random tree (CRT). We are motivated by the metric structure of graphs with a prescribed degree sequence. Introduced by Bender and Canfield [12] and by Bollobás [20] in the form of the configuration model, these graphs have received a lot of attention since the first tight analysis of the size of connected components by Molloy and Reed [38, 39]. This is mainly because the model allows for a lot of flexibility in the degree sequence. In particular, the model provides a construction of random graphs with degree sequences that may match the observations in large real-world networks.

Of course, random graphs with a prescribed degree sequence are much more complex than trees with a prescribed degree sequence, but there is no doubt that the analysis of trees is a first step towards the identification of the metric structure of the corresponding graphs. Indeed, recent results of Joseph [31] show that under some moment condition, the sizes of the connected components of random graphs with a prescribed *critical* degree sequence are similar to those of Erdős–Rényi $G(n, p)$ random graphs [21, 24, 30]: they may be asymptotically described in terms of the lengths of the excursions of a Brownian motion with parabolic drift above its current minimum, as demonstrated by Aldous [9]. On the other hand, the metric structure of $G(n, p)$ inside the critical window has recently been identified in terms of modifications of Brownian CRT by Addario-Berry et al. [2, 3]. In other words, the present analysis is one more building block towards an invariance principle for scaling limits of random graphs, i.e., that critical random graphs with a prescribed degree sequence have (under a suitable moment condition on the degree distribution) the

same scaling limit (as sequence of compact metric spaces) as classical random graphs [3]. This is at least what is suggested by the results of Bhamidi et al. [18], van der Hofstad [44] and Joseph [31].

Moreover, in the same way that uniform random trees or forests may be seen as the results of coagulation/fragmentation processes involving particles [41, 42], trees with a prescribed degree sequence appear naturally in similar aggregation processes. The model where particles have constrained valence may appear more "physically" grounded. The relevant underlying coalescing procedure is the additive coalescent [10, 15], a Markov process whose dynamics are such that particles merge at a rate proportional to the sum of their masses/sizes. The additive coalescent is the aggregation process appearing in Knuth's modification of Rényi's parking problem [28, 43] or the hashing with linear probing [17, 22]. The reader may find more information about coagulation/fragmentation processes in the monograph by Bertoin [16] or the recent survey by Berestycki [14].

The model $\mathbb{P}_{\mathbf{s}}$ is related to Galton–Watson trees [11, 27], also called simply generated trees in the combinatorial literature, by a simple conditioning: the distribution $\mathbb{P}_{\mathbf{s}}$ coincides with the distribution of the family tree $\mathbf{t}$ of a Galton–Watson process with offspring distribution $(\nu_i, i \geq 0)$ (which must satisfies $\nu_i > 0$ if $n_i > 0$) conditioned on $\{n_i(\mathbf{t}) = n_i, i \geq 0\}$. In this sense, the distribution ν plays a role of secondary importance, and $\mathbb{P}_{\mathbf{s}}$ appears to be a model of combinatorial nature, fairly stranger to the world of Galton–Watson processes. Nevertheless, we will see that Theorem 1 implies the following result of Aldous (stated in a slightly different form in [6]) (see also [6–8, 33, 36]).

Proposition 2 (Aldous [6]). *Let $\mu = (\mu_i, i \geq 0)$ be a distribution with mean $m_\mu = 1$ and variance $\sigma_\mu^2 \in (0, +\infty)$, and let $\mathbb{P}_\mu$ be the distribution of a Galton–Watson tree with offspring distribution μ. Along the subsequence $\{n \ : \mathbb{P}_\mu(|\mathbf{t}| = n) > 0\}$, under $\mathbb{P}_\mu(\ .\ ||\mathbf{t}| = n)$*

$$\left(\frac{H_{\mathbf{t}}(nx)}{\sqrt{n}}\right)_{x\in[0,1]} \xrightarrow[n\to\infty]{(law)} \frac{2}{\sigma_\mu}\mathsf{e}$$

where e *denotes a standard Brownian excursion, the convergence holding in the space $\mathcal{C}[0,1]$ equipped with the topology of uniform convergence.*

We will see that this theorem may be seen indeed as a consequence of Theorem 1; the argument morally relies on the fact that under $\mathbb{P}_\mu(\ .\ ||\mathbf{t}| = n)$, the empirical degree sequence satisfies with probability going to 1, the hypothesis of Theorem 1 (this is stated in Lemma 11). The proof of this theorem is postponed until Section 6.

Note also a result of Kortchemski [32] that has a flavor similar to our Theorem 1 (although neither implies the other): he proves that Galton–Watson trees conditioned to have all their node degrees in a subset A of the support of the measure μ has a limiting behavior depending on A.

PLAN OF THE PAPER. In Section 2 we introduce precisely the model of trees we consider. Section 3 is devoted a useful backbone decomposition for these trees. We then prove our main result, the convergence of rescaled trees to the continuum random trees, in Section 4. Finally, the application to coagulation processes with particles with constrained valence is developed in Section 7.

2 Trees with prescribed degree sequence

We here define formally the combinatorial object discussed in this paper. For convenience we write $\mathbb{N} = \{1, 2, \dots\}$ for the set of positive natural numbers. First recall some definitions related to standard

rooted plane trees. Let $\mathcal{U} = \bigcup_{n\geq 0} \mathbb{N}^n$ be the set of finite words on the alphabet $\mathbb{N}$, where $\mathbb{N}^0 = \{\varnothing\}$, and $\varnothing$ denotes the empty word. Denote by uv the concatenation of u and v; by convention $\varnothing u = u\varnothing = u$.

A subset T of $\mathcal{U}$ is a *plane tree* (see Figure 2) if it contains $\varnothing$ (called the root), if it stable by prefix (if $uv \in T$ for u and v in $\mathcal{U}$, then $u \in T$) and if ($uk \in T$ for some $k > 1$ and $u \in U$) then $uj \in T$ for j in $\{1, \ldots, k\}$. This last condition appears necessary to get a unique tree with a given genealogical structure. The set of plane trees will be denoted by $\mathbb{T}$.

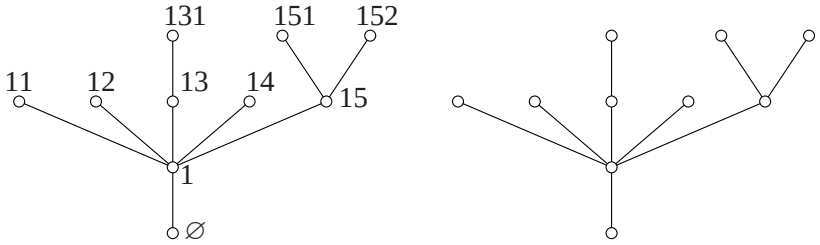

Figure 2: Usual representation of the plane tree $\{\varnothing, 1, 11, 12, 13, 14, 15, 131, 151, 152\}$

Notice that the lexicographical order $<$ on $\mathcal{U}$, also named the depth-first order, induces a total order on any tree t; this is of prime importance for the encodings of t we will present. For $t \in \mathbb{T}$, and $u \in t$, let $c_t(u) = \max\{i \ : ui \in t\}$ be the number of children of u in t. The depth of u in t, its number of letters as a word in $\mathcal{U}$, is denoted $|u|$. The notation $|t|$ refers to the cardinality of t, its number of nodes including the root $\varnothing$.

With a tree $t \in \mathbb{T}$, one can associate its degree sequence $\mathbf{s}(t) = (n_i(t), i \geq 0)$, where $n_i(t) = \#\{u \in t : c_t(u) = i\}$ is the number of nodes with degree i in t. For a fixed degree sequence $\mathbf{s}$, write $\mathbb{T}_\mathbf{s}$ for the set of trees $t \in \mathbb{T}$ such that $\mathbf{s}(t) = \mathbf{s}$, and let $\mathbb{P}_\mathbf{s}$ be the uniform distribution on $\mathbb{T}_\mathbf{s}$. To investigate the shape of random trees under $\mathbb{P}_\mathbf{s}$, we will use the usual encodings: *height process* H and *depth-first walk* S (or Łukasiewicz path). These encodings are defined by first fixing their values at the integral points, and then linear interpolation in between (See Figure 3). For a tree $t \in \mathbb{T}$, let $\tilde{u}_1 = \varnothing < \tilde{u}_2 < \cdots < \tilde{u}_{|t|}$ denote the nodes of t in the lexicographic order on $\mathcal{U}$. Then we define $H = H_t$ by $H(i) = |\tilde{u}_{i+1}|$, $S = S_t$ by $S_t(i) = \sum_{j=0}^{i}(c_t(\tilde{u}_j) - 1)$; the process H_t is defined on $[0, |t| - 1]$ and S_t on $[0, |t|]$. To simplify the normalisations in the sequel, set $H_t(|t|) := 0$.

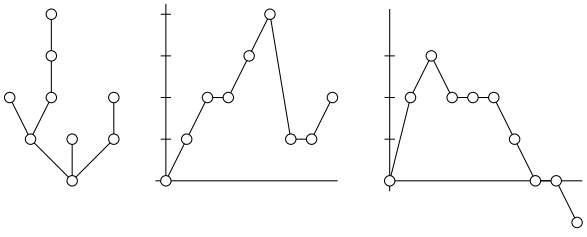

Figure 3: A plane tree $t \in \mathbb{T}$, its height process H_t and Łukasiewicz walk S_t.

Theorem 3. *Under the hypothesis of Theorem 1, under* $\mathbb{P}_{\mathbf{s}(\kappa)}$,

$$\left(\frac{H_{\mathbf{t}}(x\mathsf{n}_\kappa)}{\mathsf{n}_\kappa^{1/2}}, \frac{S_{\mathbf{t}}(x\mathsf{n}_\kappa)}{\mathsf{n}_\kappa^{1/2}}\right)_{x\in[0,1]} \xrightarrow[\kappa\to\infty]{} \left(\frac{2}{\sigma_{\mathbf{p}}}\mathsf{e}, \sigma_{\mathbf{p}}\mathsf{e}\right)$$

in distribution in the space $\mathcal{C}([0,1],\mathbb{R}^2)$ *of continuous functions from* $[0,1]$ *with values in* $\mathbb{R}^2$*, equipped with the supremum distance.*

Note that the condition that $\sigma_{\mathbf{p}}^2 > 0$ is necessary: it ensures that $p_0 = \lim_{\kappa\to\infty} n_0(\kappa)/\mathsf{n}_\kappa > 0$ and that large trees are not close to a linear tree, where most of the nodes have degree one. A tree $t \in \mathbb{T}$ can also be seen as a metric space when equipped with the graph distance d_t. A consequence of Theorem 3 is that, under $\mathbb{P}_{\mathbf{s}(\kappa)}$, the metric space

$$\left(\mathbf{t}, \frac{\sigma_\kappa}{\sqrt{\mathsf{n}_\kappa}} d_{\mathbf{t}}\right)$$

converges to the continuum random tree encoded by 2e in the sense of Gromov–Hausdorff distance between equivalence classes of compact metric spaces. The fact that the convergence of the height process implies that for the Gromov–Hausdorff topology is well known, see for example Lemma 2.3 in Le Gall [34]. So, in particular, to prove Theorem 1 it suffices to prove Theorem 3.

Remark. One can define other models of random trees with a prescribed degree sequence: for example, *rooted labelled trees*. Let $\mathbb{Q}_{\mathbf{s}(k)}$ be the uniform distribution on those with degree sequence $\mathbf{s}(k)$. Since labelled trees have a canonical ordering (using an order on the labels to order the children of each node), forgetting the labels, they can be seen as plane trees with the same degree sequence, inducing a distribution $\mathbb{P}'_{\mathbf{s}(k)}$ on the set of plane trees. By a simple counting argument, it turns out that $\mathbb{P}'_{\mathbf{s}(k)} = \mathbb{P}_{\mathbf{s}(k)}$. This situation is drastically different from the general case, since the projection of uniform labelled trees on plane tree (that is without fixing the degree sequence) does not induce the uniform distribution on plane trees.

In conclusion, Theorem 1 is also valid for the model of labelled trees with a prescribed degree sequence.

3 Combinatorial considerations: a backbone decomposition

In this section we develop a decomposition of trees under $\mathbb{P}_{\mathbf{s}(k)}$ along a branch. It is essentially the usual *backbone decomposition* for Galton–Watson trees Lyons et al. [see, e.g., 35] transposed under $\mathbb{P}_{\mathbf{s}(k)}$. The decomposition amounts to describing the structure of the branch from the root to a distinguished node u, together with the (ordered) forest formed by the trees rooted at the neighbors of that branch.

FOREST WITH A GIVEN DEGREE SEQUENCE. A forest $\mathsf{f} = (t_1, \dots, t_k)$ is a finite sequence of trees; its degree sequence $\mathbf{s}(\mathsf{f}) = \sum_{i=1}^k \mathbf{s}(t_i)$ is the (component-wise) sum of the degree sequences of the trees which compose it. If $\mathbf{s} = (n_i, i \geq 0)$ is the degree sequence of a forest f, then the number of roots of f is given by $r = |\mathbf{s}| - \sum_{i\geq 0} i n_i$. Let $\mathbb{F}_{\mathbf{s}}$ be the set of forests of (r ordered) plane trees having degree sequence $\mathbf{s}$. We have

$$\#\mathbb{F}_{\mathbf{s}} = \frac{r}{|\mathbf{s}|}\binom{|\mathbf{s}|}{(n_i, i\geq 0)} = \frac{r}{|\mathbf{s}|}\cdot\frac{|\mathbf{s}|!}{\prod_{i\geq 0} n_i!}. \tag{3}$$

THE CONTENT OF A BRANCH. Let t be a plane tree, and let $u = i_1 \dots i_{|u|}$ be one of its node, where $i_j \in \mathbb{N}$ for any j. For $j \leq |u|$, write $u_j = i_1 \dots i_j$, the ancestor of u having depth j (with the convention $u_0 = \varnothing$, the root of t). The set $[\![\varnothing, u]\!] = \{u_j \ : j < |u|\}$ is called the branch of u (notice that u is excluded). For any $i \geq 0$, the number of ancestors of u having i children is written

$$M_i(u,t) = \#\{v \ : v \text{ strict ancestors of } u, c_t(v) = i\}.$$

We refer to $\mathbf{M}(u,t) = (M_i(u,t), i \geq 0)$ as the composition of the branch. Note that we necessarily have $M_0(u,t) = 0$. Clearly if $u \in t$, then

$$|u| = \sum_{i \geq 1} M_i(u,t) = |\mathbf{M}(u,t)|. \tag{4}$$

Further let $\mathsf{LR}(u,t)$ (for left or right) be the number of nodes off the path $[\![\varnothing, u]\!]$, that is the number of nodes that are children of a node in $[\![\varnothing, u]\!]$ without being themselves in $[\![\varnothing, u]\!]$ (See Figure 4)

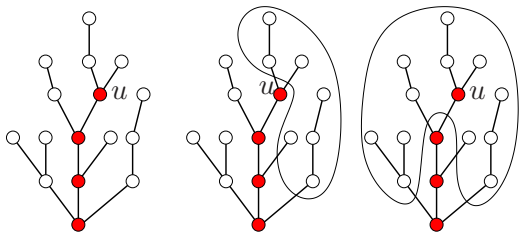

Figure 4: A tree t with a marked node u; the sets in the two right-hand side pictures show the nodes counted by $\mathsf{R}(u,t)$ and in $\mathsf{LR}(u,t)$.

Let also $\mathsf{R}(u,t)$ be the number of such nodes counted by $\mathsf{LR}(u,t)$ that lie to the right of the path. A node v is counted in $\mathsf{R}(u,t)$ if it is a child of some u_i, for $i \in \{0, \dots, |u|-1\}$, and satisfies $v > u_{i+1}$ in the lexicographic order on $\mathcal{U}$. Therefore

$$\mathsf{LR}(u,t) = \sum_{j=0}^{|u|-1} (c_t(u_j) - 1) = \sum_{i \geq 0} M_i(u,t)(i-1)$$
$$\mathsf{R}(u,t) = \sum_{j=0}^{|u|-1} (c_t(u_j) - i_{j+1}).$$

Let $\tilde{u}_1 = \varnothing < \tilde{u}_2 < \dots < \tilde{u}_{|t|}$ be the nodes of t, in increasing lexicographic order. Then $H_t(k) = |\tilde{u}_{k+1}|$ and $S_t(k) = \mathsf{R}(\tilde{u}_k, t) + c_t(\tilde{u}_k)$, so that the discrepancy between H_t and S_t can be accessed using the number of nodes to the right of the paths to $\tilde{u}_i$, $i = 1, \dots, |t|$. This observation lies at the heart of our approach.

The set of plane trees with degree sequence $\mathbf{s}$ and a distinguished node (marked plane trees) is denoted by $\mathbb{T}_{\mathbf{s}}^{\bullet} = \{(t,u) \ : t \in \mathbb{T}_{\mathbf{s}}, u \in t\}$, and the uniform distribution on this set is denoted $\mathbb{P}_{\mathbf{s}}^{\bullet}$. Under $\mathbb{P}_{\mathbf{s}}^{\bullet}$, a marked tree (t,u) is distributed as (t',u') where t' is a tree sampled under $\mathbb{P}_{\mathbf{s}}$ and u' is a uniformly random node in t'. We now decompose a marked tree (t,u) along the branch $[\![\varnothing, u]\!]$. First, consider the structure of this branch, that we call the contents:

$$\mathsf{Cont}(t,u) := \big((c_t(u_0), i_1), \dots, (c_t(u_{|u|-1}), i_{|u|})\big).$$

We write $J^{\mathbf{m}}$ for the set of potential vectors $\mathsf{Cont}(t,u)$ when the composition of the branch $[\![\varnothing,u]\!]$ is $\mathbf{M}(u,t)=\mathbf{m}$. Besides, notice that

$$|J^{\mathbf{m}}| = \binom{|\mathbf{m}|}{(m_i, i\geq 1)} \prod_{i\geq 1} i^{m_i}. \tag{5}$$

Since, if $\mathsf{Cont}(u,t)\in J^{\mathbf{m}}$ then $\mathsf{LR}(u,t)=\sum_{i\geq 0}(i-1)m_i$, we will use the following notation:

$$\mathsf{LR}(\mathbf{m}) := \sum_{i\geq 0}(i-1)m_i.$$

THE FOREST OFF A DISTINGUISHED PATH. For a tree t and any node $v\in t$, let $t_v=\{w\ : vw\in t\}$ be the subtree of t rooted at v. The sequence of trees $\mathsf{F}(t,u)=(t_v, v\in \mathsf{LR}(u,t))$ is the forest constituted by the subtrees of t rooted at the vertices belonging to $\mathsf{LR}(u,t)$, and sorted according to the rank of their root for the lexicographic order.

The decomposition which associates $(\mathsf{Cont}(t,u),\mathsf{F}(t,u))$ to a marked tree (t,u) is clearly one-to-one. The following proposition characterizes the distributions of $\mathbf{M}(\mathbf{u},\mathbf{t})$, $\mathsf{Cont}(\mathbf{u},\mathbf{t})$, and $\mathsf{R}(\mathbf{u},\mathbf{t})$ when $(\mathbf{t},\mathbf{u})$ is sampled under $\mathbb{P}^{\bullet}_{\mathbf{s}}$.

Proposition 4. *Let $\mathbf{s}=(n_0,n_1,\dots)$ be a degree sequence and let $\mathbf{m}=(m_0,m_1,\dots)$ be such that $m_0=0$, and $m_i\leq n_i$ for any $i\geq 1$. Let $(\mathbf{t},\mathbf{u})$ be chosen according to $\mathbb{P}^{\bullet}_{\mathbf{s}}$.*
(a) We have

$$\mathbb{P}^{\bullet}_{\mathbf{s}}\left(\mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m}\right) = \frac{|\mathsf{LR}(\mathbf{m})|\,|\mathbf{m}|!\,|\mathbf{s}-\mathbf{m}|!}{|\mathbf{s}|!\,|\mathbf{s}-\mathbf{m}|}\cdot\prod_{i\geq 1}\binom{n_i}{m_i} i^{m_i}.$$

(b) Moreover, for any vector $C\in J^{\mathbf{m}}$,

$$\mathbb{P}^{\bullet}_{\mathbf{s}}\left(\mathsf{Cont}(\mathbf{u},\mathbf{t})=C \mid \mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m}\right) = 1/\#J^{\mathbf{m}}.$$

(c) For any $x\geq 0$, and $\mathbf{m}$ such that $\mathbb{P}^{\bullet}_{\mathbf{s}}(\mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m})>0$,

$$\mathbb{P}^{\bullet}_{\mathbf{s}}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t})-\frac{\sigma^2_{\mathbf{s}}}{2}|\mathbf{u}|\right|\geq x \;\middle|\; \mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m}\right) = \mathbb{P}\left(\left|\sum_{j\geq 1}\sum_{k=1}^{m_j} U_j^{(k)} - \frac{\sigma^2_{\mathbf{s}}}{2}|\mathbf{m}|\right|\geq x\right) \tag{6}$$

where the $U_j^{(k)}$ are independent random variables, $U_j^{(k)}$ is uniform in $\{0,\dots,j-1\}$ and where $\sigma^2_{\mathbf{s}}$ is the variance associated with $(p_i=n_i/|\mathbf{s}|,\, i\geq 0)$ (as done on (2)).

Proof. Since the backbone decomposition is a bijection, we have for any vector $C\in J^{\mathbf{m}}$, we have

$$\begin{aligned}\mathbb{P}^{\bullet}_{\mathbf{s}}\left(\mathsf{Cont}(\mathbf{u},\mathbf{t})=C\right) &= \frac{\#\mathbb{F}_{\mathbf{s}-\mathbf{m}}}{|\mathbf{s}|\cdot\#\mathbb{F}\mathbf{s}}\\ &= \frac{|\mathsf{LR}(\mathbf{m})|}{|\mathbf{s}|-|\mathbf{m}|}\binom{|\mathbf{s}|-|\mathbf{m}|}{(n_i-m_i, i\geq 0)}\Big/\binom{|\mathbf{s}|}{(n_i, i\geq 0)},\end{aligned}$$

by the expression for the number of forests in (3). As $\mathbb{P}^{\bullet}_{\mathbf{s}}\left(\mathsf{Cont}(\mathbf{u},\mathbf{t})=C\right)$ is independent of $C\in J^{\mathbf{m}}$, it suffices to multiply by $\#J^{\mathbf{m}}$ in order to get $\mathbb{P}^{\bullet}_{\mathbf{s}}\left(\mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m}\right)$. After simplification, this yields the

first statement in (a), and then (b). Now, (b) implies that for any $R \geq 0$, any composition $\mathbf{m}$ for which $\mathbb{P}^\bullet_{\mathbf{s}}(\mathbf{M}(\mathbf{u},\mathbf{t}) = \mathbf{m}) > 0$, we have

$$\mathbb{P}^\bullet_{\mathbf{s}}\left(\mathsf{R}(\mathbf{u},\mathbf{t}) = R \mid \mathbf{M}(\mathbf{u},\mathbf{t}) = \mathbf{m}\right) = \mathbb{P}\left(\sum_{j\geq 1}\sum_{k=1}^{m_j} U_j^{(k)} = R\right),$$

where the $U_j^{(k)}$ are independent random variables, and $U_j^{(k)}$ is uniform in $\{0,\dots,j-1\}$. This implies assertion (c) and completes the proof. □

4 Convergence of uniform trees to the CRT: Proof of Theorem 3

4.1 The general approach

Our approach uses the phenomenon observed in Marckert & Mokkadem [36] in the case of critical Galton–Watson tree (having a variance): under some mild assumptions the Łukasiewicz path $S_{\mathbf{t}}$ and the height process $H_{\mathbf{t}}$ are asymptotically proportional, that is, up to a scalar normalisation, the difference between these processes converge to the zero function. It turns out that a similar phenomenon occurs when the degree sequence is prescribed, and this is the basis of our proof.

In order to prove Theorem 3 we proceed in two steps: the first one consists in showing that the depth-first walk $S_{\mathbf{t}}$ associated to a tree sampled under $\mathbb{P}_{\mathbf{s}(\kappa)}$ converges to a Brownian excursion. The process $S_{\mathbf{t}}$ is much easier to deal with than $H_{\mathbf{t}}$, since $S_{\mathbf{t}}$ is essentially a random walk conditioned to stay non-negative, and forced to end up at the origin. We provide the details in Section 4.2 below. The core of the work lies in the second step, which consists in proving that $S_{\mathbf{t}}$ and $H_{\mathbf{t}}$ are indeed close, uniformly on $[0,1]$. More precisely, by Theorem 3.1 p. 27 of [19], the following proposition is sufficient to show that $\mathsf{n}_\kappa^{-1/2} 2 S_{\mathbf{t}}(\mathsf{n}_\kappa \cdot)$ and $\mathsf{n}_\kappa^{-1/2}\sigma_\kappa^2 H_{\mathbf{t}}(\mathsf{n}_\kappa\cdot)$ have the same limit in $(\mathcal{C}[0,1], \|_\infty)$.

Proposition 5. *Under the hypothesis of Theorem 1, there exists $c_\kappa = o(\mathsf{n}_\kappa^{1/2})$ such that, as $\kappa \to \infty$,*

$$\mathbb{P}_{\mathbf{s}(\kappa)}\left(\sup_{x\in[0,1]}\left|S_{\mathbf{t}}(x\mathsf{n}_\kappa) - \frac{\sigma_\kappa^2}{2} H_{\mathbf{t}}(x\mathsf{n}_\kappa)\right| \geq c_\kappa\right) \xrightarrow[\kappa\to\infty]{} 0.$$

In order to prove Proposition 5, we recall the representation of S_t and H_t in terms of $\mathsf{R}(u,t)$ and $|u|$: for $\tilde{u}_1 = \varnothing < \tilde{u}_2 < \cdots < \tilde{u}_{|t|}$ the nodes of t in the lexicographic order, one has $S_t(k) = \mathsf{R}(\tilde{u}_k, t) + c_t(\tilde{u}_k)$ and $H_t(k) = |\tilde{u}_{k+1}|$. A non-uniform version of the claim in Proposition 5 is the following:

Proposition 6. *Assume the hypothesis of Theorem 1. Let $(\mathbf{t},\mathbf{u})$ chosen under $\mathbb{P}^\bullet_{\mathbf{s}(\kappa)}$. There exists $c_\kappa = o(\mathsf{n}_\kappa^{1/2})$ such that,*

$$\mathbb{P}^\bullet_{\mathbf{s}(\kappa)}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t}) - \frac{\sigma_\kappa^2}{2}|\mathbf{u}|\right| \geq c_\kappa\right) \xrightarrow[\kappa\to\infty]{} 0.$$

From Proposition 6 one can control the discrepancy between the two processes at any finite number of random points in $[0,1]$ using the union bound. So Proposition 6 implies convergence of the finite dimensional distributions. To complete the proof of Proposition 5, it suffices to show that the under the sequence

of distribution $\mathbb{P}_{\mathbf{s}(\kappa)}$, the rescaled difference of the two processes $\mathsf{n}_\kappa^{-1/2}(S_{\mathbf{t}}(x\mathsf{n}_\kappa) - \frac{\sigma_\kappa^2}{2} H_{\mathbf{t}}(x\mathsf{n}_\kappa))$ is tight [see, e.g. 19]. Since the rescaled version of $S_{\mathbf{t}}$ (see Section 4.2) are tight, it suffices to prove that the rescaled versions of $H_{\mathbf{t}}$ are tight as well.

Remark. Under slightly stronger assumptions on the degree sequences, it is possible to control the discrepancy between the height process and the Łukasiewicz path at *every point* in $\{0, 1, \ldots, \mathsf{n}_\kappa - 1\}$. More precisely it would be possible to show that

$$\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)} \left(\left| \mathsf{R}(\mathbf{u}, \mathbf{t}) - \frac{\sigma_\kappa^2}{2} |\mathbf{u}| \right| \geq c_\kappa \right) = o(1/\mathsf{n}_\kappa). \tag{7}$$

Using the union bound, this yields the convergence of the rescaled height process to a Brownian excursion, as a random function in $\mathcal{C}[0,1]$. One is easily convinced that with the optimal assumptions for Theorem 3, the bound in (7) might just not be true.

We now move on to the ingredients of the proof: we first give the details of the convergence of $\mathsf{n}_\kappa^{-1/2} S_{\mathbf{t}}(\ \cdot\ \mathsf{n}_\kappa)$ to a Brownian excursion in Section 4.2, then we prove tightness for $\mathsf{n}_\kappa^{-1/2} H_{\mathbf{t}}\ (\cdot\ \mathsf{n}_\kappa)$ in Section 4.3. The longer proof of Proposition 6 is delayed until Section 5.

4.2 Convergence of the Łukasiewicz walk

In this section, we give the details of the proof of the convergence of the depth-first walk under $\mathbb{P}_{\mathbf{s}(\kappa)}$ towards the Brownian excursion.

Lemma 7. *Assume the hypothesis of Theorem 1. Under* $\mathbb{P}_{\mathbf{s}(\kappa)}$,

$$\left(\frac{S_{\mathbf{t}}(x\mathsf{n}_\kappa)}{\sigma_\kappa \mathsf{n}_\kappa^{1/2}} \right)_{x \in [0,1]} \xrightarrow[\kappa \to +\infty]{(law)} \mathsf{e}$$

as random functions in $\mathcal{C}[0,1]$.

Proof. Let $\mathbf{c} = \{c_1, c_2, \ldots, c_{\mathsf{n}_\kappa}\}$ be a multiset of n_κ integers whose distribution is given by $\mathbf{s}(\kappa)$. When t is sampled under $\mathbb{P}_{\mathbf{s}(\kappa)}$, the Łukasiewicz path S is easily related to the partial sums of the elements samples uniformly at random from the set $\mathbf{c}$: Let $\pi = (\pi_1, \pi_2, \ldots, \pi_{\mathsf{n}_\kappa})$ be a uniformly random permutation of $\{1, 2, \ldots, \mathsf{n}_\kappa\}$, and for $j \in \{1, \ldots, \mathsf{n}_\kappa\}$, define

$$W_\pi(j) = \sum_{i=1}^{j} (c_{\pi_i} - 1).$$

Theorem 20.7 of Aldous [5] (see also Theorem 24.1 in [19]) ensures that, when $\Delta_\kappa = o(\sqrt{\mathsf{n}_\kappa})$,

$$\left(\frac{W_\pi(s\mathsf{n}_\kappa)}{\sigma_\kappa \mathsf{n}_\kappa^{1/2}} \right)_{s \in [0,1]} \xrightarrow[\kappa \to +\infty]{(law)} \mathsf{b},$$

in $\mathcal{C}[0,1]$, where $\mathsf{b} = (\mathsf{b}(s), s \in [0,1])$ is a standard Brownian bridge.

The increments of the walk $(W_\pi(j), 0 \le j \le \mathsf{n}_\kappa)$ satisfy $x_{\pi_i} - 1 \ge -1$ for every i (such walks are sometimes called *left-continuous*). Furthermore, one finally has $W_\pi(\mathsf{n}_\kappa) = -1$. The cycle lemma [23] ensures that there is a unique way to turn the process W_π into an excursion by shifting the increments cyclically. Extend the definition of the permutation in such a way that, for $j \in \{\mathsf{n}_\kappa + 1, \dots, 2\mathsf{n}_\kappa\}$, we have $\pi_j := \pi_{j-\mathsf{n}_\kappa}$. Let j_π be the unique integer (the location of the first minimum of the walk W_π) in $\{1, \dots, \mathsf{n}_\kappa\}$ such that $W_\pi(j + j_\pi) - W_\pi(j_\pi)$ is an excursion in the following sense:

$$\tilde{S}_\pi(j) := W_\pi(j + k_\pi) - W_\pi(j_\pi) \ge 0 \quad \text{for } j < \mathsf{n}_\kappa \qquad \text{and } \tilde{S}_\pi(\mathsf{n}_\kappa) = -1.$$

Then, for $\mathbf{t}$ uniformly chosen in $\mathbb{T}_{\mathbf{s}(\kappa)}$,

$$(S_{\mathbf{t}}(j), 0 \le j \le \mathsf{n}_\kappa) \overset{d}{=} (\tilde{S}_\pi(j), 0 \le j \le \mathsf{n}_\kappa),$$

for π a random permutation of $\{1, \dots, \mathsf{n}_\kappa\}$. Since the Brownian bridge b has almost surely a unique minimum, the claim follows by the mapping theorem [19]. □

4.3 Tightness for the height process

The rescaled height process under $\mathbb{P}_{\mathbf{s}}(\kappa)$ is the process in $\mathcal{C}[0,1]$, $h_\kappa = \mathsf{n}_\kappa^{-1/2} H(\ \cdot\ \mathsf{n}_\kappa)$. In this section, we prove that the family $(h_\kappa, \kappa > 0)$ is tight (we will omit the κ when unnecessary). Since $h_\kappa(0) = 0$, the following lemma is sufficient to prove tightness [see, e.g., 19].

Let ω_h be the modulus of continuity of the rescaled height process h: for $\delta > 0$

$$\omega_h(\delta) = \sup_{|t-s| \le \delta} |h(s) - h(t)|.$$

Lemma 8. *Under the hypothesis of Theorem 1, for any $\epsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ such that, for all κ large enough,*

$$\mathbb{P}_{\mathbf{s}(\kappa)}(\omega_h(\delta) > \epsilon) < \eta.$$

The bound we provide consists in reducing the bounds on the variations of h to bounds on the variations of the Łukasiewicz path S, which is known to be tight since it converges in distribution (Lemma 7). The underlying ideas are due to Addario-Berry et al. [4] and Addario-Berry [1] to prove Gaussian tail bounds for the height and width of Galton–Watson trees and random trees with a prescribed degree sequence, respectively.

For a plane tree $t \in \mathbb{T}$, let t^- be the mirror image of t, or in other words, the tree obtained by flipping the order of the children of every node. Then, we let $S_t^- := S_{t^-}$ be the *reverse depth-first walk*. Observe that the mirror flip is a bijection, so that S_t and S_t^- have the same distribution under $\mathbb{P}_{\mathbf{s}(\kappa)}$.

Proof of Lemma 8. In this proof, we identify the nodes of a tree t and their index in the lexicographic order; so in particular, we write $H_t(u)$ for the height of a node u in t, and we write $|u - v| \le \delta$ to mean that u and v are within δ in the lexicographic order.

Consider a tree t and two nodes u and v. Write $u \wedge v$ for the first common ancestor of u and v in t. In the following we write $u \preceq v$ to mean that u is an ancestor of v in t ($u = v$ is allowed). Then,

$$|H_t(u) - H_t(v)| \leq |H_t(u) - H_t(u \wedge v)| + |H_t(v) - H_t(u \wedge v)|, \tag{8}$$

so that it suffices to bound variations of H_t between two nodes on the same path to the root:

$$\sup_{|u-v| \leq \delta \mathsf{n}_\kappa} |H_t(u) - H_t(v)| \leq 2 \sup_{w \preceq u, |u-w| \leq \delta \mathsf{n}_\kappa} |H_t(u) - H_t(w)|$$

Now, observe that, for $w \preceq u$, every node v on the path between w and u which has degree more than one contributes at least one to the number of nodes off the path between w and u:

$$1 + \sum_{w \preceq v \preceq u} (c_t(v) - 1) \geq H_t(u) - H_t(w) - \sum_{w \preceq v \preceq u} \mathbf{1}_{\{c_t(v)=1\}}$$

However, one may also bound this same number of nodes in terms of the depth-first walk S_t, and the reverse depth-first walk S_t^-:

$$1 + \sum_{w \preceq v \preceq u} (c_t(v) - 1) \leq S_t(u) - S_t(w) + S_t^-(u) - S_t^-(w) + c_t(w). \tag{9}$$

In other words, we have

$$\begin{aligned}
\sup_{|u-v| \leq \delta \mathsf{n}_\kappa} |H_t(v) - H_t(u)| &\leq 2 \sup_{|u-w| \leq \delta \mathsf{n}_\kappa, w \preceq u} |S_t(u) - S_t(w)| + 2 \sup_{|u-w| \leq \delta \mathsf{n}_\kappa, w \preceq u} |S_t^-(u) - S_t^-(w)| \\
&\quad + \max_w c_t(w) + \sup_{|u-w| \leq \delta \mathsf{n}_\kappa} \sum_{w \preceq v \preceq u} \mathbf{1}_{\{c_t(v)=1\}} \\
&\leq 2 \sup_{|u-w| \leq \delta n} |S_t(u) - S_t(w)| + 2 \sup_{|u-w| \leq \delta \mathsf{n}_\kappa} |S_t^-(u) - S_t^-(w)| \\
&\quad + \Delta_\kappa + \sup_{|u-w| \leq \delta \mathsf{n}_\kappa} \sum_{w \preceq v \preceq u} \mathbf{1}_{\{c_t(v)=1\}} \\
&\leq 2\mathsf{n}_\kappa^{1/2} \omega_s(\delta) + 2\mathsf{n}_\kappa^{1/2} \omega_{s-}(\delta) + \Delta_\kappa + \sup_{|u-w| \leq \delta \mathsf{n}_\kappa} \sum_{w \preceq v \preceq u} \mathbf{1}_{\{c_t(v)=1\}},
\end{aligned} \tag{10}$$

where ω_s and ω_{s^-} denote the moduli of continuity of the rescaled Łukasiewicz path $\mathsf{n}_\kappa^{-1/2} S_t$ and $\mathsf{n}_\kappa^{-1/2} S_t^-$, respectively.

The first three terms in (10) are easy to bound since $\Delta_\kappa = o(\sqrt{\mathsf{n}}_\kappa)$ and S_t and S_t^- are tight under $\mathbb{P}_{\mathbf{s}(\kappa)}$. The only term remaining to control is the one concerning the number of nodes of degree one:

$$Y_t(\delta) := \sup_{|u-w| \leq \delta \mathsf{n}_\kappa} \sum_{w \preceq v \preceq u} \mathbf{1}_{\{c_t(v)=1\}}.$$

To bound $Y_t(\delta)$ we relate the distribution of trees under $\mathbb{P}_{\mathbf{s}(\kappa)}$ to those under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$, where $\mathbf{s}(\kappa)^\star = (n_0^\star, n_1^\star, \dots)$ is obtained from $\mathbf{s}(\kappa)$ by removing all nodes of degree one, i.e., $n_1^\star = 0$ and $n_i^\star = n_i$ for every $i \neq 1$. Then, in a tree $t^\star$ sampled under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$, one has $Y_{t^\star}(\delta) = 0$. Recall also that $\Delta_\kappa = o(\sqrt{\mathsf{n}}_\kappa)$.

Now, for a sum of three terms to be at least ϵ, at least one term must exceed $\epsilon/3$. So for every $\epsilon, \delta > 0$, there exists a $\delta > 0$ such that, for all κ large enough,

$$\begin{aligned}\mathbb{P}_{\mathbf{s}(\kappa)^\star}(\omega_h(\delta) \geq \epsilon) &\leq \mathbb{P}_{\mathbf{s}(\kappa)^\star}(2\omega_s(\delta) > \epsilon/3) + \mathbb{P}_{\mathbf{s}(\kappa)^\star}(2\omega_{s^-}(\delta) > \epsilon/3) \\ &= 2\mathbb{P}_{\mathbf{s}(\kappa)^\star}(6\omega_s(\delta) \geq \epsilon) < \eta,\end{aligned}$$

since, under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$, S_t and S_t^- have the same distribution and $\mathsf{n}_\kappa^{-1/2} S_t$ is tight, and since $\mathbb{P}_{\mathbf{s}(\kappa)^\star}(\Delta_\kappa \mathsf{n}_\kappa^{-1/2} \geq \varepsilon/3)$ is zero for κ large enough. This proves that $\mathsf{n}_\kappa^{-1/2} H_t$ is tight under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$.

Now, we can couple the trees sampled under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$ and $\mathbb{P}_{\mathbf{s}}(\kappa)$. Since the nodes of degree one do not modify the tree structure, a tree t under $\mathbb{P}_{\mathbf{s}(\kappa)}$ may be obtained by first sampling $t^\star$ using $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$, and then placing the nodes of degree one uniformly at random : precisely, this insertion of nodes is done inside the edges of $t^\star$ (plus a phantom edge below the root). Given any ordering of the edges of $t^\star$ (plus the one below the root), the vector $(X_1^\star, \dots, X^\star_{\mathsf{n}_\kappa - n_1(\kappa)})$ of numbers of nodes of degree one falling in these edges is such that

$$(X_1^\star, \dots, X^\star_{\mathsf{n}_\kappa - n_1(\kappa)}) \stackrel{d}{=} \text{Multinomial}\left(n_1(\kappa); \frac{1}{\mathsf{n}_\kappa - n_1(\kappa)}, \dots, \frac{1}{\mathsf{n}_\kappa - n_1(\kappa)}\right).$$

Conversely, $t^\star$ is obtained from t by removing the nodes of degree one, so that t and $t^\star$ can be thought as random variables in the same probability space under $\mathbb{P}_{\mathbf{s}(\kappa)}$. To bound $Y_t(\delta)$, observe that it is unlikely that adding the nodes of degree one in this way creates too long paths.

In fact, "the length of paths" is expected to be multiplied by $1 + q_\kappa$ for $q_\kappa = n_1(\kappa)/(\mathsf{n}_\kappa - n_1(\kappa))$. Let $\alpha = 2 + q_\kappa$, and fix $\delta > 0$ such that $\mathbb{P}_{\mathbf{s}(\kappa)^\star}(\omega_h(\delta) \geq \epsilon/\alpha) < \eta/2$; such a $\delta > 0$ exists since the height process is tight under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$. Note that since we add nodes in the construction of t under $\mathbb{P}_{\mathbf{s}(\kappa)}$ from $t^\star$ under $\mathbb{P}_{\mathbf{s}(\kappa)^\star}$, nodes that are within $\delta \mathsf{n}_\kappa$ in t are also within $\delta \mathsf{n}_\kappa$ in $t^\star$. Write $h^\star$ for the rescaled height process obtained from $t^\star$, the tree associated with t by deletion of all nodes of degree one (the rescaling stays $\sqrt{\mathsf{n}_\kappa}$). We have,

$$\begin{aligned}\mathbb{P}_{\mathbf{s}(\kappa)}(\omega_h(\delta) \geq \epsilon) &\leq \mathbb{P}_{\mathbf{s}(\kappa)}(\omega_{h^\star}(\delta) \geq \epsilon/\alpha) + \mathbb{P}_{\mathbf{s}(\kappa)}(\omega_h(\delta) \geq \epsilon\,,\ \omega_{h^\star}(\delta) \leq \epsilon/\alpha) \\ &\leq \mathbb{P}_{\mathbf{s}(\kappa)^\star}(\omega_h(\delta) \geq \epsilon/\alpha) + \mathbb{P}_{\mathbf{s}(\kappa)}(\omega_h(\delta) \geq \epsilon \mid \omega_{h^\star}(\delta) \leq \epsilon/\alpha) \\ &\leq \mathbb{P}_{\mathbf{s}(\kappa)^\star}(\omega_h(\delta) \geq \epsilon/\alpha) + \delta \mathsf{n}_\kappa^2 \mathbb{P}\left(\sum_{i=1}^{\epsilon\sqrt{\mathsf{n}_\kappa}/\alpha} (1 + X_i^\star) \geq \varepsilon\sqrt{\mathsf{n}_\kappa}\right) \\ &\leq \eta/2 + \delta \mathsf{n}_\kappa^2 \mathbb{P}\left(\sum_{i=1}^{\epsilon\sqrt{\mathsf{n}_\kappa}/\alpha} X_i \geq \varepsilon\sqrt{\mathsf{n}_\kappa}(1 - 1/\alpha)\right),\end{aligned}$$

where the X_i are i.i.d. Binomial$(n_1, 1/(\mathsf{n}_\kappa - n_1))$ random variables. The last line follows from the standard fact that the numbers $(X_i^\star)$ obtained from a sampling *without replacement* (of the $n_1(\kappa)$ nodes of degree one) are more concentrated than their counterpart (X_i) coming from a sampling *with replacement* [5].

Now, the sum in the right-hand side is itself a binomial random variable:

$$\sum_{i=1}^{\epsilon\sqrt{\mathsf{n}_\kappa}/\alpha} X_i \stackrel{d}{=} \text{Binomial}\left(\epsilon\sqrt{\mathsf{n}_\kappa} n_1/\alpha, \frac{1}{\mathsf{n}_\kappa - n_1}\right)$$

whose mean is $\varepsilon\sqrt{n_k}q_\kappa/(2+q_\kappa)$ when $\varepsilon\sqrt{\mathsf{n}_\kappa}(1-1/\alpha) = \varepsilon\sqrt{\mathsf{n}_\kappa}(1+q_\kappa)/(2+q_\kappa)$. By Chernoff's bound, using that q_κ converges, it follows that for some constant $c>0$ valid for κ large enough,

$$\mathbb{P}_{\mathbf{s}(\kappa)}(\omega_h(\delta)\geq\epsilon) \leq \eta/2 + \delta\mathsf{n}_\kappa^2 e^{-c\sqrt{\mathsf{n}_\kappa}/\epsilon}.$$

Finally, for all κ large enough, with this value for δ, we have $\mathbb{P}_{\mathbf{s}(\kappa)}(\omega_h(\delta)\geq\epsilon)<\eta$, which completes the proof. □

5 Finite dimensional distributions: Proof of Proposition 6

5.1 A roadmap to Proposition 6: identifying the bad events

Our approach consists in showing that if the event in Proposition 6 occurs, then one of the following three events must occur: (1) either the depth $|\mathbf{u}|$ of node $\mathbf{u}$ is unusually large, (2) or the content of the branch $[\![\varnothing,\mathbf{u}]\!]$ is atypical, (3) or the number of nodes to the right of the path is not what it should be, despite of the length $|\mathbf{u}|$ and content $\mathbf{M}(\mathbf{u},\mathbf{t})$ being typical.

We will then prove that those simpler events are unlikely. For $h\geq 0$, and two sequences $a=(a_\kappa,\kappa\geq 0)$, and $b=(b_\kappa,\kappa\geq 0)$ we define families of sets $A_{h,a,b}$ as follows. Given a sequence of degree distribution $(\mathbf{s}(\kappa),\kappa\geq 0)$,

$$A_{h,a,b}(\kappa) := \left\{ \mathbf{m} \,:\, |\mathbf{m}|=h, \left| \left(\sum_{i\geq 0} m_i \frac{i-1}{2}\right) - \frac{h\sigma_\kappa^2}{2}\right| \leq a_\kappa, \sum_{i\geq 1} m_i i^2 \leq b_\kappa \right\}.$$

If $\mathbf{m}\in A_{h,a,b}(\kappa)$ then $|\mathbf{m}|=h$, and $\mathbf{m}$ corresponds to the content of a branch $[\![\varnothing,u]\!]$ such that $|u|=h$. The set $A_{h,a,b}(\kappa)$ are designed to contain most typical contents of a branch of length h under $\mathbb{P}_{\mathbf{s}(\kappa)}$, provided the choices for the sequences a and b are suitable. The decomposition of the bad event we have outlined above is then expressed formally by

$$\begin{aligned}
\mathbb{P}^\bullet_{\mathbf{s}(\kappa)}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t}) - \frac{\sigma_\kappa^2}{2}|\mathbf{u}|\right| \geq c_\kappa\right) \leq\ & \mathbb{P}^\bullet_{\mathbf{s}(\kappa)}(|\mathbf{u}|\geq x\sqrt{\mathsf{n}_\kappa}) \\
&+ \mathbb{P}^\bullet_{\mathbf{s}(\kappa)}\left(\mathsf{LR}(\mathbf{u},\mathbf{t})\geq x\sqrt{\mathsf{n}_\kappa}\right) \\
&+ \mathbb{P}^\bullet_{\mathbf{s}(\kappa)}\left(|\mathbf{u}|\vee\mathsf{LR}(\mathbf{u},\mathbf{t})\leq x\sqrt{\mathsf{n}_\kappa}, \mathbf{M}(\mathbf{u},\mathbf{t})\notin \bigcup_{h\leq x\sqrt{\mathsf{n}_\kappa}} A_{h,a,b}(\kappa)\right) \\
&+ \sum_{\substack{h\leq x\sqrt{\mathsf{n}_\kappa}\\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}^\bullet_{\mathbf{s}(\kappa)}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t}) - \frac{\sigma_\kappa^2}{2}|\mathbf{u}|\right| \geq c_\kappa, \mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m}\right). \quad (11)
\end{aligned}$$

Proving Proposition 6 reduces to proving that, for some choice of a_κ, b_κ, c_κ and x, every term in the right-hand side above tends to zero as $\kappa\to\infty$. The bound on the first term is a direct consequence of the Gaussian tail bounds for the height of trees recently proved by Addario-Berry [1] in the very setting we use:

$$\mathbb{P}^\bullet_{\mathbf{s}(\kappa)}(|\mathbf{u}|\geq x\sqrt{\mathsf{n}_\kappa}) \leq \mathbb{P}_{\mathbf{s}(\kappa)}\left(\max_{u\in t}|u|\geq x\sqrt{\mathsf{n}_\kappa}\right) \leq \exp(-cx^2/\sigma_\kappa^2), \quad (12)$$

for a universal constant $c > 0$ and all sufficiently large κ. The second term is bounded using the depth-first walk S and the reverse depth-first walk S^-, as in the proof of Lemma 8:

$$\begin{aligned}\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}(\mathsf{LR}(\mathbf{u},\mathbf{t}) \geq x\sqrt{\mathsf{n}_\kappa}) &\leq \mathbb{P}_{\mathbf{s}(\kappa)}\left(\max_{0\leq k\leq \mathsf{n}_\kappa}\{S(k)+S^-(k)\}+\Delta_\kappa \geq x\sqrt{\mathsf{n}_\kappa}\right)\\ &\leq 2\mathbb{P}_{\mathbf{s}(\kappa)}\left(\max_{0\leq k\leq \mathsf{n}_\kappa} S(k) \geq \frac{x}{3}\sqrt{\mathsf{n}_\kappa}\right),\end{aligned}$$

for all κ large enough, since $\Delta_\kappa = o(\mathsf{n}_\kappa)$ and S and S^- have the same distribution under $\mathbb{P}_{\mathbf{s}(\kappa)}$. We finish using the tightness of $\mathsf{n}_\kappa^{-1/2} S(\mathsf{n}_\kappa .)$ under $\mathbb{P}_{\mathbf{s}(\kappa)}$; more precisely, we have

$$\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}(\mathsf{LR}(\mathbf{u},\mathbf{t}) \geq x\sqrt{\mathsf{n}_\kappa}) \leq 16\cdot 9\cdot \frac{\sigma_\kappa^2}{x^2}, \tag{13}$$

by Lemma 20.5 of [5]. The bounds on the two remaining terms are stated in Lemmas 9 and 10, the proof of which appear in Sections 5.2 and 5.3, respectively.

Lemma 9. *Since $\Delta_k = o(\mathsf{n}_\kappa)$ there exists ε_κ such that $\Delta_\kappa \leq \varepsilon_\kappa\sqrt{\mathsf{n}_\kappa}$, with $0 < \varepsilon_\kappa \to 0$. Let $a_\kappa = \varepsilon_\kappa^{1/4}\sqrt{\mathsf{n}_\kappa}$ and $b_\kappa = \varepsilon_\kappa^{1/2}\mathsf{n}_\kappa$. Then, for every $x > 0$, and all κ large enough,*

$$\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(|\mathbf{u}| \vee \mathsf{LR}(\mathbf{u},\mathbf{t}) \leq x\sqrt{\mathsf{n}_\kappa}, \mathbf{M}(\mathbf{u},\mathbf{t}) \notin \bigcup_{h\leq x\sqrt{\mathsf{n}_\kappa}} A_{h,a,b}(\kappa)\right) \leq 6x^2e^{x^2}\exp\left(-\frac{\varepsilon_\kappa^{-1/2}}{2x(\sigma_\kappa^2+1)+2}\right).$$

Lemma 10. *Since $\Delta_k = o(\mathsf{n}_\kappa)$ there exists ε_κ such that $\Delta_\kappa \leq \varepsilon_\kappa\sqrt{\mathsf{n}_\kappa}$, with $0 < \varepsilon_\kappa \to 0$ and $\varepsilon_\kappa^{-3/4} = o(\mathsf{n}_\kappa)$ as $\kappa \to \infty$. Let $a_\kappa = \varepsilon_\kappa^{1/4}\sqrt{\mathsf{n}_\kappa}$, $b_\kappa = \varepsilon_\kappa^{1/2}\mathsf{n}_\kappa$, and $c_\kappa = \varepsilon_\kappa^{1/8}\sqrt{\mathsf{n}_\kappa}$. Then, for all κ large enough,*

$$\sum_{\substack{h\leq x\sqrt{\mathsf{n}_\kappa}\\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t}) - \frac{\sigma_\kappa^2}{2}|\mathbf{u}|\right| \geq c_\kappa, \mathbf{M}(\mathbf{u},\mathbf{t}) = \mathbf{m}\right) \leq 2e^{-\varepsilon_\kappa^{-1/2}}. \tag{14}$$

Before proceeding with the proofs of these two lemmas, we indicate how to use them in order to complete the proof of Proposition 6. Let ε_κ be such that $\Delta_\kappa \leq \varepsilon_\kappa\sqrt{\mathsf{n}_\kappa}$, with $\varepsilon_\kappa \to 0$ as $\kappa \to \infty$. Then, set $a_\kappa = \varepsilon_\kappa^{1/4}\sqrt{\mathsf{n}_\kappa}$, $b_\kappa = \varepsilon_\kappa^{1/2}\mathsf{n}_\kappa$ and $c_\kappa = \varepsilon_\kappa^{1/8}\sqrt{\mathsf{n}_\kappa}$. Let now $\epsilon > 0$ be arbitrary. Pick $x > 0$ large enough such that, for all κ large enough,

$$\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}(|\mathbf{u}| \geq x\sqrt{\mathsf{n}_\kappa}) + \mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}(\mathsf{LR}(\mathbf{u},\mathbf{t}) \geq x\sqrt{\mathsf{n}_\kappa}) < \epsilon/2.$$

The bounds in (12) and (13), and the fact that $\sigma_\kappa^2 \to \sigma_{\mathbf{p}}^2$ ensure that this is possible. The value for x being fixed, Lemmas 9 and 10 now make it possible to choose κ_0 large enough such that, for all $\kappa \geq \kappa_0$, the two remaining terms in the right-hand side of (11) also sum to at most $\epsilon/2$. Thus, for all $\kappa \geq \kappa_0$, we have

$$\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t}) - \frac{\sigma_\kappa^2}{2}|\mathbf{u}|\right| \geq c_\kappa\right) < \epsilon,$$

which completes the proof, since ϵ was arbitrary.

5.2 The content of a branch is very likely typical: Proof of Lemma 9

We now prove that, on the event that $|\mathbf{u}|$ and $\mathsf{LR}(\mathbf{u},\mathbf{t})$ are not too large, the content of the branch $[\![\varnothing,\mathbf{u}]\!]$ is typical with high probability.

We start by rewriting the probability of interest using Proposition 4:

$$\begin{aligned}
&\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(|\mathbf{u}|\vee\mathsf{LR}(\mathbf{u},\mathbf{t})\le x\sqrt{\mathsf{n}_\kappa},\mathbf{M}(\mathbf{u},\mathbf{t})\notin\bigcup_{h\le x\sqrt{\mathsf{n}_\kappa}}A_{h,a,b}(\kappa)\right)\\
&=\sum_{h\le x\sqrt{\mathsf{n}_\kappa}}\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(|\mathbf{u}|=h,\mathsf{LR}(\mathbf{u},\mathbf{t})\le x\sqrt{\mathsf{n}_\kappa},\mathbf{M}(\mathbf{u},\mathbf{t})\notin A_{h,a,b}(\kappa)\right)\\
&=\sum_{h\le x\sqrt{\mathsf{n}_\kappa}}\sum_{\substack{|\mathbf{m}|=h\\ \mathbf{m}\notin A_{h,a,b}(\kappa),\mathsf{LR}(\mathbf{m})\le x\sqrt{\mathsf{n}_\kappa}}}\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(|\mathbf{u}|=h,\mathbf{M}(\mathbf{u},\mathbf{t})=\mathbf{m}\right)\\
&=\sum_{h\le x\sqrt{\mathsf{n}_\kappa}}\sum_{\substack{|\mathbf{m}|=h\\ \mathbf{m}\notin A_{h,a,b}(\kappa),\mathsf{LR}(\mathbf{m})\le x\sqrt{\mathsf{n}_\kappa}}}\frac{|\mathsf{LR}(\mathbf{m})|\,h!\,(\mathsf{n}_\kappa-h)!}{\mathsf{n}_\kappa!(\mathsf{n}_\kappa-h)}\prod_{i\ge1}\binom{n_i}{m_i}i^{m_i}. \qquad (15)
\end{aligned}$$

where, for short, we have written n_i instead of $n_i(\kappa)$. We now reduce the right-hand side to an expected value with respect to multinomial random variables. Let $(P_i,i\ge1)$ be multinomial with parameters h and $(in_i/\mathsf{n}_\kappa,i\ge1)$. Then, for any $\mathbf{m}=(0,m_1,m_2,\dots)$ such that $|\mathbf{m}|=h$, we have

$$\mathbb{P}\left((P_i,i\ge1)=(m_i,i\ge1)\right)=\frac{h!}{\prod_{i\ge1}m_i!}\cdot\prod_{i\ge1}\left(\frac{in_i}{\mathsf{n}_\kappa}\right)^{m_i}.$$

Now, since $(1-x)^{-1}\le\exp(2x)$ for $|x|\le1/2$, we have for all $h\le x\sqrt{\mathsf{n}_\kappa}$, and all κ large enough,

$$\frac{(\mathsf{n}_\kappa-h)!\mathsf{n}_\kappa^h}{\mathsf{n}_\kappa!}\le\prod_{i=0}^{h-1}\frac{1}{1-i/\mathsf{n}_\kappa}\le\prod_{i=0}^{h-1}e^{2i/\mathsf{n}_\kappa}\le e^{x^2}.$$

Note also that, for every $i\ge1$, we have $n_i!\le n_i^{m_i}(n_i-m_i)!$, so that, rewriting (15) in terms of events with respect to $(P_i,i\ge1)$, we obtain

$$\begin{aligned}
&\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(|\mathbf{u}|\vee\mathsf{LR}(\mathbf{u},\mathbf{t})\le x\sqrt{\mathsf{n}_\kappa},\mathbf{M}(\mathbf{u},\mathbf{t})\notin\bigcup_{h\le x\sqrt{\mathsf{n}_\kappa}}A_{h,a,b}(\kappa)\right)\\
&=\sum_{h\le x\sqrt{\mathsf{n}_\kappa}}\sum_{\substack{|\mathbf{m}|=h\\ \mathbf{m}\notin A_{h,a,b}(\kappa),\mathsf{LR}(\mathbf{m})\le x\sqrt{\mathsf{n}_\kappa}}}\frac{\mathsf{LR}(\mathbf{m})}{\mathsf{n}_\kappa-h}\cdot\frac{(\mathsf{n}_\kappa-h)!\mathsf{n}_\kappa^h}{\mathsf{n}_\kappa!}\prod_{i\ge1}\frac{n_i!}{n_i^{m_i}(n_i-m_i)!}\cdot\mathbb{P}\left((P_i,i\ge1)=(m_i,i\ge1)\right)\\
&\le\sum_{h\le x\sqrt{\mathsf{n}_\kappa}}\frac{2x}{\sqrt{\mathsf{n}_\kappa}}e^{x^2}\sum_{\substack{|\mathbf{m}|=h\\ \mathbf{m}\notin A_{h,a,b}(\kappa)}}\mathbb{P}\left((P_i,i\ge1)=(m_i,i\ge1)\right)\\
&\le 2x^2e^{x^2}\sup_{h\le x\sqrt{\mathsf{n}_\kappa}}\mathbb{P}((P_i,i\ge1)\notin A_{h,a,b}(\kappa)).
\end{aligned}$$

Now, we decompose the set of $\mathbf{m}$ in the right-hand side so as to obtain bad events that are individually simpler to deal with

$$\mathbb{P}^{\bullet}_{\mathbf{s}(\kappa)}\left(|\mathbf{u}|\vee\mathsf{LR}(\mathbf{u},\mathbf{t})\le x\sqrt{\mathsf{n}_\kappa},\mathbf{M}(\mathbf{u},\mathbf{t})\notin A_{h,a,b}(\kappa)\right)\le 2x^2e^{x^2}\sup_{h\le x\sqrt{\mathsf{n}_\kappa}}(\zeta_1+\zeta_2)$$

where

$$\zeta_1 = \mathbb{P}\left(\left|\left(\sum_{i\geq 1} P_i \frac{i-1}{2}\right) - \frac{h\sigma_\kappa^2}{2}\right| \geq a_\kappa\right) \qquad \text{and} \qquad \zeta_2 = \mathbb{P}\left(\sum_{i\geq 1} i^2 P_i > b_\kappa\right).$$

We now bound the terms ζ_1 and ζ_2 individually.

THE FIRST TERM ζ_1. Observe first, that

$$\mathbb{E}\left[\sum_{i\geq 1} P_i \frac{i-1}{2}\right] = \frac{h\sigma_\kappa^2}{2},$$

so that bounding ζ_1 consists in bounding the deviations of (a function of) a multinomial vector. However, one can write

$$\sum_{i\geq 1} P_i \cdot \frac{i-1}{2} - h\frac{\sigma_\kappa^2}{2} \stackrel{d}{=} \sum_{j=1}^{h} (B_j - \mathbb{E}B_j),$$

where B_j, $j = 1, \ldots, h$, are i.i.d. random variables taking value $(i-1)/2$ with probability in_i/n_κ, for $i \geq 1$. Now, the sums $\sum_{j=1}^{\ell}(B_i - \mathbb{E}[B_j])$, $\ell = 0, 1, \ldots, h$, form a martingale. We bound their deviations using a concentration inequality from [37] (Theorem 3.15), which says that if S is a sum of independent random variable $X_1 + \cdots + X_n$ such that $\mathbb{E}(S) = \mu$, $\mathrm{var}(S) = V$, and if for all k $X_k - \mathbb{E}(X_k) \leq b$, then $\mathbb{P}(S - \mu \geq t) \leq e^{-t^2/(2V(1+bt/(3V))}$. The variance of B_j may be bounded as follows:

$$\mathrm{var}(B_j) \leq \mathbb{E}[B_j^2] = \sum_{i\geq 1} \frac{(i-1)^2}{4} \frac{in_i}{\mathsf{n}_\kappa} \leq \Delta_\kappa \sum_{i\geq 1} \frac{i-1}{4} \frac{in_i}{\mathsf{n}_\kappa} \leq \Delta_\kappa \sigma_\kappa^2/4,$$

for all κ large enough. Now, since $\max\{|B_j - \mathbb{E}(B_j)| : j = 0, \ldots, h\} \leq \Delta_\kappa$, one has, for $h \leq x\sqrt{\mathsf{n}_\kappa}$,

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{j=1}^{h}(B_j - \mathbb{E}B_j)\right| \geq a_\kappa\right) &\leq 2\exp\left(-\frac{a_\kappa^2}{2h\Delta_\kappa\sigma_\kappa^2/4 + 2\Delta_\kappa a_\kappa/3}\right) \\ &\leq 2\exp\left(-\frac{a_\kappa^2}{x\sqrt{\mathsf{n}_\kappa}\Delta_\kappa\sigma_\kappa^2}\right),\end{aligned}$$

for all κ large enough, since $a_\kappa = \varepsilon_\kappa^{1/4}\sqrt{\mathsf{n}_\kappa} = o(\sqrt{\mathsf{n}_\kappa})$. It follows that, for every $h \leq x\sqrt{\mathsf{n}_\kappa}$, we have

$$\zeta_1 = \sup_{h\leq x\sqrt{\mathsf{n}_\kappa}} \mathbb{P}\left(\left|\left(\sum_{i\geq 1} P_i \frac{i-1}{2}\right) - \frac{h\sigma_\kappa^2}{2}\right| \geq a_\kappa\right) \leq 2\exp\left(-\frac{\varepsilon_\kappa^{-1/2}}{x\sigma_\kappa^2}\right). \tag{16}$$

THE THIRD TERM ζ_2. We bound ζ_2 using the idea we used when bounding ζ_1: one can express the event in terms of independent random variables B_j, $j = 1, \ldots, h$, where B_j takes value i^2 with probability in_i/n_κ. Observe first that

$$\mathbb{E}\left[\sum_{i\geq 1} i^2 P_i\right] = \mathbb{E}\left[\sum_{j=1}^{h} B_j\right] = h\sum_{i\geq 1} i^2 \cdot \frac{in_i}{\mathsf{n}_\kappa} \leq h\Delta_\kappa(\sigma_\kappa^2 + 1).$$

So, we have

$$\mathbb{P}\left(\sum_{i\geq 1} i^2 P_i > b_\kappa\right) = \mathbb{P}\left(\sum_{j=1}^{h} B_j > b_\kappa\right) \leq \mathbb{P}\left(\sum_{j=1}^{h}(B_j - \mathbb{E}[B_j]) > \frac{b_\kappa}{2}\right),$$

for all κ large enough, since $h\Delta_\kappa \leq x\varepsilon_\kappa \mathsf{n}_\kappa = o(\varepsilon_\kappa^{1/2}\mathsf{n}_\kappa) = o(b_\kappa)$. The right-hand side above can be bounded using the martingale inequality in [37] (Theorem 3.15). We note that the variance of B_j satisfies

$$\mathrm{var}(B_j) \leq \mathbb{E}[B_j^2] = \sum_{i\geq 1} i^4 \cdot \frac{in_i}{\mathsf{n}_\kappa} \leq \Delta_\kappa^3(\sigma_\kappa^2 + 1).$$

Since $\max\{|B_i| : i = 1, \ldots, h\} \leq \Delta_\kappa^2$, it follows by McDiarmid's inequality that

$$\begin{aligned}\zeta_2 \leq \mathbb{P}\left(\sum_{j=1}^{h}(B_j - \mathbb{E}[B_j]) > \frac{b_\kappa}{2}\right) &\leq \exp\left(-\frac{b_\kappa^2/4}{2x\sqrt{\mathsf{n}_\kappa}\Delta_\kappa^3(\sigma_\kappa^2+1) + 2\Delta_\kappa^2 b_\kappa/3}\right) \\ &\leq \exp\left(-\frac{b_\kappa}{2(x(\sigma_\kappa^2+1)+1/3)\Delta_\kappa^2}\right) \\ &= \exp\left(-\frac{\varepsilon_\kappa^{-3/2}}{2x(\sigma_\kappa^2+1)+2/3}\right), \qquad (17)\end{aligned}$$

for all κ large enough, since $\Delta_\kappa\sqrt{\mathsf{n}_\kappa} = o(b_\kappa)$.

To complete the proof, it suffices to combine the bounds in (16)–(17), and observe that they imply the claim for κ large enough, since the upper bound in (16) is much smaller than the one in (17).

5.3 The structure of a branch with typical content: Proof of Lemma 10

Finally, we consider the probability that the structure of a branch is not what one expects, in spite of the length and content being close to the typical values. The left hand side in (14) is bounded by

$$\sup_{\substack{h\leq x\sqrt{\mathsf{n}_\kappa}\\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}^\bullet_{\mathbf{s}(\kappa)}\left(\left|\mathsf{R}(\mathbf{u},\mathbf{t}) - \frac{\sigma_\kappa^2}{2}|\mathbf{u}|\right| \geq c_\kappa \;\middle|\; \mathbf{M}(\mathbf{u},\mathbf{t}) = \mathbf{m}\right) = \sup_{\substack{h\leq x\sqrt{\mathsf{n}_\kappa}\\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}\left(\left|\frac{\sigma_\kappa^2}{2}h - \sum_{j\geq 1}\sum_{k=1}^{m_j} U_j^{(k)}\right| \geq c_\kappa\right),$$

by Proposition 4 (3), where $U_j^{(k)}$ are independent random variables with $U_j^{(k)}$ uniform on $\{0, 1, \ldots, j-1\}$. By the triangle inequality, the quantity in the right-hand side above is at most

$$\sup_{\substack{h\leq x\sqrt{\mathsf{n}_\kappa}\\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}\left(\left|\sum_{j\geq 1} m_j\frac{j-1}{2} - \sum_{j\geq 1}\sum_{k=1}^{m_j} U^{(k)}(j)\right| \geq c_\kappa - \left|\frac{\sigma_\kappa^2 h}{2} - \sum_{j\geq 1} m_j\frac{j-1}{2}\right|\right). \qquad (18)$$

By definition of $A_{h,a,b}(\kappa)$, and since $c_\kappa > 2a_\kappa$ for all κ large enough, the quantity in (18) is bounded by

$$\sup_{\substack{h\leq x\sqrt{\mathsf{n}_\kappa}\\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}\left(\left|\sum_{j\geq 1} m_j\frac{j-1}{2} - \sum_{j\geq 1}\sum_{k=1}^{m_j} U_j^{(k)}\right| \geq \frac{c_\kappa}{2}\right).$$

Now, since all the random variables $U_j^{(k)}$, $j \geq 1$, $k = 1, \ldots, m_j$ are symmetric about their respective mean $(j-1)/2$, one obtains using Chernoff's bounding method

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{j\geq 1} m_j \frac{j-1}{2} - \sum_{j\geq 1}\sum_{k=1}^{m_j} U_j^{(k)}\right| \geq \frac{c_\kappa}{2}\right) &\leq 2 \inf_{t\geq 0} e^{-tc_\kappa/2} \mathbb{E}\left[e^{t\sum_{j\geq 1}\sum_{k=1}^{m_j}\left(U_j^{(k)} - \frac{(j-1)}{2}\right)} \right] \\
&= 2 \inf_{t\geq 0} e^{-tc_\kappa/2} \prod_{j\geq 1} \left(\frac{\sinh(tj/2)}{j\sinh(t/2)}\right)^{m_j} \\
&\leq 2 \inf_{t\geq 0} \exp\left(-t\frac{c_\kappa}{2} + \sum_{j\geq 1} m_j \left(\frac{j^2t^2}{24} - \frac{t^2}{24} + \frac{t^4}{2880}\right)\right) \\
&\leq 2 \inf_{t\in(0,1)} \exp\left(-t\frac{c_\kappa}{2} + \sum_{j\geq 1} m_j \left(\frac{j^2t^2}{24} - \frac{t^2}{48}\right)\right) \qquad (19) \\
&\leq 2 \inf_{t\in(0,1)} \exp\left(-t\frac{c_\kappa}{2} + \frac{t^2}{24}\sum_{j\geq 1} m_j j^2\right).
\end{aligned}$$

Here the third line follows from the bounds $\log(\sinh(s)) \leq \log(s) + s^2/6$ and $\log(\sinh(s)) \geq \log(s) + s^2/6 - s^4/180$ valid for $s \geq 0$. Finally, we obtain

$$\begin{aligned}
\sup_{\substack{h\leq x\sqrt{n_\kappa} \\ \mathbf{m}\in A_{h,a,b}(\kappa)}} \mathbb{P}\left(\left|\sum_{j\geq 1} m_j \frac{j-1}{2} - \sum_{j\geq 1}\sum_{k=1}^{m_j} U_j^{(k)}\right| \geq \frac{c_\kappa}{2}\right) &\leq 2 \inf_{t\in(0,1)} \exp\left(-t\frac{c_\kappa}{2} + \frac{t^2 b_\kappa}{24}\right) \\
&\leq 2e^{-3c_\kappa^2/(2b_\kappa)},
\end{aligned}$$

upon choosing $t = 6c_\kappa/b_\kappa$, which is indeed in $(0,1)$ for κ large enough (we restricted the range of t in (19)). This completes the proof since $3c_\kappa^2/(2b_\kappa) = 3\varepsilon_\kappa^{-3/4}/2 \geq \varepsilon_\kappa^{-1/2}$, for all κ large enough.

6 The limit of rescaled Galton–Watson trees: Proof of Proposition 2

Denote by $\widehat{\mathbf{s}}_\mathbf{t} := (\widehat{n}_i(\mathbf{t}), i \geq 0)$ the empirical degree sequence of $\mathbf{t}$, let

$$\begin{aligned}
\widehat{\mu}_i &= \widehat{n}_i(\mathbf{t})/|\mathbf{t}|, \\
\widehat{\sigma}^2 &= \sum_{i\geq 0} i^2 \frac{\widehat{n}_i(\mathbf{t})}{|\mathbf{t}|} - \frac{|\mathbf{t}|-1}{|\mathbf{t}|} \\
\widehat{\Delta} &= \max\{i : \widehat{n}_i > 0\}.
\end{aligned}$$

Write $\mathbb{P}_\mu^n(\,\cdot\,) = \mathbb{P}_\mu^n$. In what follows, all the assertions containing " $\mathbb{P}_\mu^n$" are to be understood "for n such that $\mathbb{P}_\mu(|\mathbf{t}| = n) > 0$"; similarly, the limit with respect to $\mathbb{P}_\mu^n$ are to be understood in the same manner, along subsequences included in $\{n : \mathbb{P}_\mu(|\mathbf{t}| = n) > 0\}$.

Lemma 11. *Assume that μ has mean 1 and variance $\sigma_\mu^2 \in (0, +\infty)$. Then under $\mathbb{P}_\mu^n$,*

$$(\hat{\mu}, \widehat{\sigma}^2, \widehat{\Delta}/\sqrt{n}) \xrightarrow[n]{(d)} (\mu, \sigma_\mu^2, 0), \tag{20}$$

where the convergence holds in the space $\mathcal{M}(\mathbb{N}) \times \mathbb{R} \times \mathbb{R}$ equipped with the product topology.

$\mathcal{M}(\mathbb{N})$ is the set of probability measures on $\mathbb{N}$. The topology on $\mathcal{M}(\mathbb{N})$ is metrizable, for example, by the distance

$$D(\nu, \nu') = \sum_{i \geq 0} d_{\mathrm{TV}}(\nu[i], \nu'[i])$$

where $\nu[i]$ is the distribution of the ith first marginals under ν and d_{TV} is the distance in total variation. Since here the limit is the deterministic measure μ, it suffices to show that, for all i, $\hat{\mu}_i \to \mu_i$ in probability as $n \to \infty$ With D it is easy to construct a metric on $\mathcal{M}(\mathbb{N}) \times \mathbb{R} \times \mathbb{R}$ making of this space a Polish space. Hence, by the Skohorod theorem there exists a probability space where versions of $(\hat{\mu}, \widehat{\sigma}^2, \widehat{\Delta}/\sqrt{n})$ under $\mathbb{P}_\mu^n$ converges almost surely to $(\mu, \sigma_\mu^2, 0)$. On this probability space, the hypothesis of Theorem 1 holds almost surely, and then its conclusion, which is a limit in distribution, holds.

Proof of Lemma 11. The claim is about properties of the degree sequence of Galton–Watson trees conditioned on their total progeny. We first provide a way to construct the degree sequence. Consider the Łukasiewicz walk S_n associated with a tree $\mathbf{t}$ under $\mathbb{P}_\mu^n$; the degree sequence of the tree $\mathbf{t}$ is essentially (just shift by one) the empirical distribution of the increments of S_n. More precisely, consider first a random walk $W = (W_k, k = 0, \ldots, n)$, with i.i.d. increments $X_k = W_k - W_{k-1}, k = 1, \ldots, n$ with distribution

$$\nu_i = \mathbb{P}(X_k = i) = \mu_{i+1} \quad i \geq -1;$$

then $S = (S_0, \ldots, S_n)$ is distributed as W conditioned on $W \in A_{-1}^+(n)$ where

$$A_{-1}^+(n) = \{w = (w_0, \ldots, w_n) \ : \ w_0 = 0, w_k \geq 0, 1 \leq k < n, w_n = -1\}$$

is the set of discrete excursions of length n.

Write $K_i = \#\{k : X_k = i - 1\}$, and $K = (K_i, i \geq 0)$. Then, if $W \in A_{-1}^+(n)$, the sequence $K = (K_i, i \geq 0)$ is distributed as the degree sequence of a tree under $\mathbb{P}_\mu^n$. In other words, we have

$$\mathbb{P}(K \in B \mid W \in A_{-1}^+(n)) = \mathbb{P}_\mu^n((\widehat{n}_i(\mathbf{t}), i \geq 0) \in B).$$

By the rotation principle, we may remove the positivity condition :

$$\mathbb{P}(K \in B \mid W_n = -1) = \mathbb{P}_\mu^n((\widehat{n}_i(\mathbf{t}), i \geq 0) \in B).$$

Our aim is now to show that the condition that W is a bridge imposed by $W_n = -1$ does not completely wreck the properties of W in the following sense: let $\mathcal{F}_k = \sigma(W_0, \ldots, W_k)$ be the σ-field generated by the k first W_i; then there exists a constant $c \in (0, \infty)$ such that for any n large enough, and for any event $B \in \mathcal{F}_{\lfloor n/2 \rfloor}$ one has

$$\mathbb{P}(B \mid W_n = -1) \leq c\, \mathbb{P}(B). \tag{21}$$

That is: any event B in $\mathcal{F}_{\lfloor n/2\rfloor}$ with a very small probability for a standard (unconditioned) random walk also has a small probability in the bridge case (conditional on $W_n = -1$). The argument proving this claim is given in Janson and Marckert [29], page 662 and goes as follows:

$$\begin{aligned}\mathbb{P}(B \,|\, W_n = -1) &= \sum_x \mathbb{P}(B \,|\, W_{\lfloor n/2\rfloor} = x, W_n = -1) \cdot \frac{\mathbb{P}(W_{\lfloor n/2\rfloor} = x, W_n = -1)}{\mathbb{P}(W_n = -1)} \\ &= \sum_x \mathbb{P}(B \,|\, W_{\lfloor n/2\rfloor} = x)\mathbb{P}(W_{\lfloor n/2\rfloor} = x) \cdot \frac{\mathbb{P}(W_{n-\lfloor n/2\rfloor} = -x-1)}{\mathbb{P}(W_n = -1)}.\end{aligned}$$

It then suffices to (a) observe that $\sup_x \mathbb{P}(W_{n-\lfloor n/2\rfloor} = -x-1) \le c/\sqrt{n}$ for some constant $c_1 \in (0,\infty)$ [40, Theorem 2.2 p. 76], and (b) use a local limit theorem to show that $\mathbb{P}(W_n = -1) \ge c_2\sqrt{n}$, for some constant $c_2 \in (0,\infty)$ and all n large enough [25, page 233]. This gives the result in (21) with $c = c_1/c_2$.

Now using that the increments $(X_1, \dots, X_n)$ under $\mathbb{P}(\,\cdot\,|\,W_n = -1)$ are exchangeable, any concentration principle for the first half of them easily extends to the second half (the easy details are omitted). Consider the degree sequence induced by the first half of the walk: let $K_i^{1/2} = \#\{k : X_k = i-1, k \le \lfloor n/2\rfloor\}$, and note that the $K_i^{1/2}$ are $\mathcal{F}_{\lfloor n/2\rfloor}$-measurable. For W (that is, with no conditioning), we have

$$\frac{1}{\lfloor n/2\rfloor}\sum_{i\ge 0} K_i^{1/2}(i-1)^2 = \frac{1}{\lfloor n/2\rfloor}\sum_{j=1}^{\lfloor n/2\rfloor}(X_j-1)^2 \xrightarrow[n\to\infty]{} \mathbb{E}[(X_1-1)^2] = \sigma_\mu^2 \tag{22}$$

by the law of large number, since X_i owns a (finite) moment of order 2. Hence, for any $\varepsilon > 0$, writing

$$Ev(\varepsilon) = \left\{ \left| \frac{1}{\lfloor n/2\rfloor}\sum_{i\ge 0} K_i^{1/2}(i-1)^2 - \sigma_\mu^2 \right| \ge \varepsilon \right\},$$

we have $\mathbb{P}(Ev(\varepsilon)) \to 0$ and thus, according to the bound in (21), $\mathbb{P}(Ev(\varepsilon)|W_n = -1) \to 0$, as $n \to \infty$. Using the argument twice (one for each half of the walk) yields convergence $\widehat{\sigma}^2 \to \sigma_\mu^2$ in probability as $n \to \infty$.

The same argument also proves that

$$\mathbb{P}\left(\left| \frac{K_i^{1/2}}{\lfloor n/2\rfloor} - \mu_i \right| \ge \varepsilon \,\middle|\, W_n = -1 \right) \to 0,$$

which yields $\widehat{\mu}_i \to \mu_i$ in probability.

The fact that $\widehat{\Delta} = o(\sqrt{n})$ (in probability) under $\mathbb{P}_\mu^n$ is also a consequence of the convergence of the sum given in (22). To see this, let $C(\alpha) = \{k \,:\, \mathbb{P}(X_1^2 \ge k) \ge \alpha/k\}$. Since $\mathbb{E}[X_1^2] = \sum_{k\ge 0} k\mathbb{P}(X_1^2 \ge k) < +\infty$, we have $\#C(\alpha) < +\infty$ for any $\alpha > 0$. In particular, for any $\varepsilon > 0$,

$$\#\{n : n\mathbb{P}(X_1^2 \ge \varepsilon n) \ge \alpha/\varepsilon\} < +\infty.$$

Taking $\alpha = \varepsilon\varepsilon'$, one obtains that $\#\{n : n\mathbb{P}(X_1^2 \ge \varepsilon n) \ge \varepsilon'\} < +\infty$, which implies that

$$\mathbb{P}(\max\{X_i : i \le n/2\} \ge \varepsilon\sqrt{n}) \le n\mathbb{P}(X_1^2 \ge \varepsilon n) \xrightarrow[n\to\infty]{} 0.$$

So under the unconditioned law one has $\widehat{\Delta} = o(\sqrt{n})$; we complete the proof using the bound in (21). □

7 Application to constrained coalescing processes

In this final section, we discuss an application of Theorem 1 to a coalescence process with particles having constrained valences.

The famous additive coalescent [10, 15, 16, 41, 42] can be seen as arising from the following natural microscopic description. Consider a set of n distinct particles $\{1, 2, \dots, n\}$. The particles are initially free, and form n clusters; the clusters are organized as rooted trees. The clusters merge according to the following dynamics. At each step, choose a particle u uniformly at random; it belongs to some cluster T rooted at r. Choose uniformly a second cluster $T' \neq T$, with root r'. Add an edge between r' and u to obtain a new cluster rooted at r. At each step, the system consists of a forest of general rooted labelled trees (an acyclic graph on $\{1, 2, \dots, n\}$ with a distinguished node per connected component). The process stops after $n-1$ steps, when the system consists of a single rooted labelled tree. The final tree is then uniform among all rooted labelled trees.

One can similarly define a system of coalescing particles where the degrees would be constrained. Different algorithms might be used, depending on the precise way the uniform choices are made, that yield *a priori* different trees.

LABELLED PARTICLES. Consider the set of particles $\{1, 2, \dots, n\}$, and a set of degrees $c_1 \leq c_2 \leq \cdots \leq c_n$. Write $\mathbf{s} = (n_i, i \geq 0)$ for the associated degree sequence, $n_i = \#\{j : c_j = i\}$. Assign randomly the particles a degree. For instance, this can be done using a random permutation $\sigma = (\sigma(1), \dots, \sigma(n))$ of $\{1, 2, \dots, n\}$ and assigning degree $c_{\sigma(i)}$ to particle i. Think now of the particle i as initially having edges to $c_{\sigma(i)}$ free slots that can each contain a single particle. The particles will now merge to form clusters. Each cluster is represented by a tree with a distinguished vertex (the root). Initially, each particle sits in a tree containing a single node (which is then also the root). Proceed with the following algorithm to merge the particles, as long as there are free slots left:

- Pick a free slot s uniformly at random; say it is bound to particle p lying in the cluster rooted at r.
- Pick *another* cluster, uniformly at random, rooted at some node r'.
- Merge the two clusters by assigning r' to the free slot s; this creates an edge between the particles p and r', and removes the slot s from the set of free slots. The new cluster is rooted at r.

At every iteration, precisely one slot is filled and the process stops after $n-1$ steps. The process yields a random tree *labelled* tree T_n^L.

The labelled tree T_n^L is uniform in the set of labelled trees having the same specified degree sequence. To see this, just consider the encoding of the process by the final labelled tree, together with a labelling of the edge indicating their order of appearance. At iteration $i \in \{1, \dots, n-1\}$, there are $n-i$ free slots left and $n-i+1$ connected components, so that the probability that any couple free slot/other connected component is precisely

$$\frac{1}{(n-i)^2}.$$

Overall, the probability to obtain any particular pairing free slots/particles together with a history is

$$\prod_{i=1}^{n} \frac{1}{(n-i)^2} = \frac{1}{(n-1)!^2}.$$

The same particle adjacency —hence the same labelled tree— is obtained by the $\prod_{j=1}^{n} c_j!$ ways to pair the free slots with particles; and for any labelled tree there are exactly $(n-1)!$ distinct histories. Finally, among the $n!$ ways to assign the labels to particles in the first place, $\prod_{i\geq 0} n_i!$ correspond to the degree/label pattern of the tree, it follows that the probability of seeing any labelled tree after $n-1$ iterations is precisely

$$\frac{\prod_{i\geq 0} n_i!}{n!} \times \frac{1}{(n-1)!^2} \times (n-1)! \times \prod_{i=1}^{n} c_i! = \frac{\prod_{i\geq 0} i!^{n_i}}{(n-1)!} \times \binom{n}{(n_i, i\geq 0)}^{-1}, \tag{23}$$

which depends only on the degree sequence, so that trees with the same degree sequence are chosen uniformly. (This is also, as it should, the inverse of the number of labelled trees with degree sequence given by $\mathbf{s} = (n_i, i \geq 0)$ [42, Example 6.2.2].)

UNLABELLED PARTICLES. Consider a degree sequence in the form of $\mathbf{s} = (n_i, i \geq 0)$ where n_i denotes the number of nodes of degree i. For $c_1 \leq c_2 \leq \cdots \leq c_n$ of size n. So $\sum_{i\geq 0} c_i = n-1$. As before, we think of the particles as having empty slots, but since there are no labels, we impose that the slots of any given particle be ordered. The particles then merge according to the same algorithm, in order to distinguish particles use the canonical labelling giving label i to the particle with degree c_i. After forgetting the canonical labelling, the process yields a plane tree T_n.

Again, the plane tree T_n is uniform among all plane trees with the correct degree sequence. The arguments are similar, only simpler, to those we used in the labelled case. Since, for a given plane tree, there are $\prod_{i\geq 0} n_i!$ ways to assign the canonical labels to the nodes, the probability to obtain any given plane tree is

$$\prod_{i\geq 0} n_i! \times \frac{1}{(n-1)!^2} \times (n-1)! = n \binom{n}{(n_i, i\geq 0)}^{-1}$$

In these coalescing particle systems, one of the parameters of interest is the metric structure of the cluster (structure of the "molecule") eventually obtained after all particles have coalesced into a single component. In the unrestricted case, the metric structure is described by the CRT of Aldous. Our result shows that the quenched version, conditional on the degree sequence, is also valid under reasonable conditions on the degree sequence imposed. Results for Galton–Watson trees conditioned on the size only are recovered by sampling the degree sequence.

For instance, to recover the unrestricted version of the merging process, one can sample n independent Poisson(1) random variables, and keep them if their sum equals $n-1$; the n exchangeable values obtained are then the degrees $C_1, C_2, \ldots, C_n$ of the n particles.

References

[1] L. Addario-Berry. The height and width of random tree with a fixed, "finite variance" degree sequence. 2011.

[2] L. Addario-Berry, N. Broutin, and C. Goldschmidt. Critical random graphs: limiting constructions and distributional properties. *Electronic Journal of Probability*, 15:741–774, 2010. arXiv:0908.3629 [math.PR].

[3] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 2010. to appear.

[4] L. Addario-Berry, L. Devroye, and S. Janson. Sub-gaussian tail bounds for the width and height of conditioned Galton–Watson trees. arXiv:1011.4121 [math.PR], 2010.

[5] D. Aldous. Exchangeability and related topics. *Ecole d'été de probabilités de Saint-Flour, XIII*, 1117:1–198, 1983.

[6] D. Aldous. The continuum random tree II: an overview. In M. Barlow and N. Bingham, editors, *Stochastic Analysis*, pages 23–70. Cambridge University Press, 1991.

[7] D. Aldous. The continuum random tree. I. *The Annals of Probability*, 19:1–28, 1991.

[8] D. Aldous. The continuum random tree III. *The Annals of Probability*, 21:248–289, 1993.

[9] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25:812–854, 1997.

[10] D. Aldous and J. Pitman. The standart additive coalescent. *The Annals of Probability*, 26:1703–1726, 1998.

[11] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer, Berlin, 1972.

[12] E. Bender and E. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24:296–307, 1978.

[13] J. Bennies and G. Kersting. A random walk approach to Galton-Watson trees. *J. Theoret. Probab.*, 13(3): 777–803, 2000. ISSN 0894-9840.

[14] N. Berestycki. Recent progress in coalescent theory. *Ensaios Matematicos*, 16:1–193, 2009.

[15] J. Bertoin. A fragmentation process connected to Brownian motion. *Probability Theory and Related Fields*, 117: 289–301, 2000.

[16] J. Bertoin. *Random fragmentation and coagulation processes*. Cambridge University Press, Cambridge, 2006.

[17] J. Bertoin and G. Miermont. Asymptotics in Knuth's parking problem for caravans. *Random Structures and Algorithms*, 29:38–55, 2006.

[18] S. Bhamidi, R. van der Hofstad, and J. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. arXiv:0909.1472 [math.PR], 2009.

[19] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. ISBN 0-471-19745-9. A Wiley-Interscience Publication.

[20] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1:311–316, 1980.

[21] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2001.

[22] P. Chassaing and G. Louchard. Phase transition for parking blocks, Browian excursion and coalescence. *Random Structures & Algorithms*, 21:76–119, 2002.

[23] A. Dvoretzky and T. Motzkin. A problem of arrangements. *Duke Mathematical Journal*, 14:305–313, 1947.

[24] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[25] B. V. Gnedenko and A. N. Kolmogorov. *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954. Translated and annotated by K. L. Chung. With an

Appendix by J. L. Doob.

[26] B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. 2010.

[27] T. E. Harris. *The Theory of Branching Processes*. Springer, Berlin, 1963.

[28] P. Hemmer. The random parking problem. *Journal of Statistical Physics*, 57:865–869, 1989.

[29] S. Janson and J.-F. Marckert. Convergence of discrete snakes. *J. Theoret. Probab.*, 18(3):615–647, 2005. ISSN 0894-9840.

[30] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley, New York, 2000.

[31] A. Joseph. The component sizes of a critical random graph with pre-described degree sequence. ArXiv:1012.2352 [math.PR], 2010.

[32] I. Kortchemski. Invariance principles for conditioned galton-watson trees. *Arxiv:1110.2163*, 2011.

[33] J.-F. Le Gall. The uniform random tree in a Brownian excursion. *Probability Theory and Related Fields*, 96: 369–383, 1993.

[34] J.-F. Le Gall. Random real trees. *Ann. Fac. Sci. Toulouse Math. (6)*, 15(1):35–62, 2006. ISSN 0240-2963.

[35] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of the $L \log L$ criteria for mean behavior of branching processes. *The Annals of Probability*, (23):1125–1138, 1995.

[36] J. Marckert and A. Mokkadem. The depth first processes of Galton-Watson trees converge to the same Brownian excursion. *Ann. Probab.*, 31(3):1655–1678, 2003. ISSN 0091-1798.

[37] C. McDiarmid. Concentration. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed, editors, *Probabilistic methods for Algorithmic Discrete Mathematics*, volume 16 of *Algorithms and Combinatorics*, pages 195–248, Berlin, 1998. Springer.

[38] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161–179, 1995.

[39] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7:295–305, 1998.

[40] V. V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1995. ISBN 0-19-853499-X. Sequences of independent random variables, Oxford Science Publications.

[41] J. Pitman. Coalescent random forests. *Journal of Combinatorial Theory, Series A*, 85:165–193, 1999.

[42] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer, Berlin, 2006.

[43] A. Rényi. On a one-dimensional problem concering random space-filling. *Publ. Math. Inst. Hungar. Acad. Sci.*, 3:109–127, 1958.

[44] R. van der Hofstad. Critical behavior in inhomogeneous random graphs. arXiv:0902.0216 [math.PR], 2009.