

# Thermodynamic Limit for Large Random Trees.

Yuri Bakhtin\*

September 17, 2008

## Abstract

We consider Gibbs distributions on finite random plane trees with bounded branching. We show that as the order of the tree grows to infinity, the distribution of any finite neighborhood of the root of the tree converges to a limit. We compute the limiting distribution explicitly and study its properties. We introduce an infinite random tree consistent with these limiting distributions and show that it satisfies a certain form of the Markov property. We also study the growth of this tree and prove several limit theorems including a diffusion approximation.

## 1 Introduction

Various kinds of random trees have been studied in the literature. In this note we consider simply generated random (plane rooted) trees also known as branching processes conditioned on the total population (CBP), see [Ald91b]. Our initial motivation was a study of the secondary structure statistics for large RNA molecules, see [BH08a] and [BH08b]. Hence we prefer the language of statistical mechanics, so that we treat these distributions as Gibbs ensembles on trees, see the description of the setting in Section 2.

The first goal of this paper is to prove that as the order of the tree grows to infinity, the distribution induced by the Gibbs measure converges to that of an infinite discrete tree that we explicitly describe in detail (Sections 2 to 5). The computation of this “thermodynamic limit” belongs to the category of discrete limits of CBP according to the terminology introduced in [Ald91b], and our result (as well as the limiting object) appears to be new. In particular, it does not involve any rerooting procedures like the one introduced in [Ald91a]. We prove the result above for the bounded branching (or out-degree) case, although it should hold true under less restrictive assumptions.

The limiting infinite discrete tree is a more sophisticated object than a classical Galton–Watson tree. In particular, it dies out with zero probability and the progenies of distinct vertices are not independent. However, it turns out that the limit tree is Markov in a natural sense, and the the

---

\*School of Mathematics, Georgia Tech, Atlanta GA, 30332-0160; email:bakhtin@math.gatech.edu, 404-894-9235 (office phone), 404-894-4409(fax)

Markov transition probability is explicitly computed in Section 5. In Section 6 we notice that the number of vertices at a given distance  $n$  from the root also form a Markov chain if  $n$  is understood as a time parameter. We prove that under linear scaling this Markov chain satisfies a limit theorem with the limit given by a gamma distribution. In Section 7 we strengthen this result and show that a functional limit theorem holds with weak convergence to a diffusion process on the positive semi-line with constant drift and diffusion proportional to the square root of the space coordinate. Since this process (under the name of local time for Bessel(3) process) also serves as a scaling limit of the “height profile” for CBP itself, see [Ald91b, Conjecture 7] and [Git98], we can say that the infinite Markov random tree that we construct belongs to the same universality class as the original CBP.

There are several natural and interesting problems arising in connection with our results. One is, obviously, strengthening them to give an alternative to [Git98] proof of the scaling limit in Aldous’s Conjecture 7. Another one is to use our approach to study finer details of the random tree rather than the height profile. Our heuristic computation (see Section 8) shows that the limit can be described as a solution of an SPDE w.r.t. to a Brownian sheet.

## 2 The setting and first results on thermodynamic limit

Let us recall that plane trees (or, ordered trees) are rooted trees such that subtrees at any vertex are linearly ordered. In other words, two plane trees are considered equal if there is a bijection between the vertices of the two trees such that it preserves the parent — child relation on the vertices and preserves the order of the child subtrees of any vertex. Figure 1 shows all plane trees on 4 vertices.

We fix  $D \in \mathbb{N}$  and introduce  $\mathbb{T}_N = \mathbb{T}_N(D)$ , the set of all plane trees on  $N$  vertices such that the branching number (i.e. the number of children) of each vertex does not exceed  $D$ . Since we are going to consider a Gibbs distribution on  $\mathbb{T}_N$ , we have to assign an energy value to each tree. We assume that an energy value  $E_i \in \mathbb{R}$  is assigned to every  $i \in \{0, \dots, D\}$ , and the energy of the tree  $T$  is defined via

$$E(T) = \sum_{v \in V(T)} E_{\deg(v)} = \sum_{i=0}^D \chi_i(T) E_i,$$

where  $V(T)$  denotes the set of vertices of the tree  $T$ ,  $\deg(v)$  denotes the branching number of vertex  $v$ , and  $\chi_i(T)$  is the number of vertices of degree  $i$  in  $T$ .

Now we fix an inverse temperature parameter  $\beta > 0$  and define a probability measure  $\mu_N$  on  $\mathbb{T}_N$  by

$$\mu_N\{T\} = \frac{e^{-\beta E(T)}}{Z_N},$$

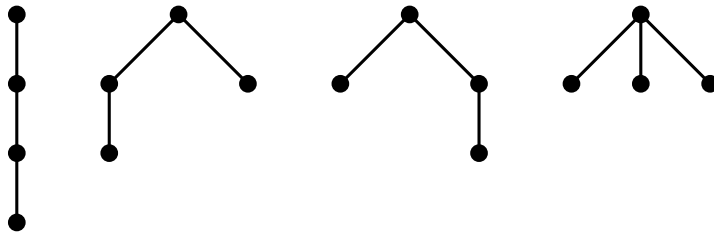


Figure 1: Four different plane trees of order 4.

where the normalizing factor (partition function) is defined by

$$Z_N = \sum_{T \in \mathbb{T}_N} e^{-\beta H(T)}.$$

First, we are going to demonstrate that the above model admits a thermodynamic limit, i.e. the sequence of measures  $(\mu_N)_{N \in \mathbb{N}}$  has a limit in a certain sense as  $N \rightarrow \infty$ . Secondly, we study several curious properties of the limiting infinite random trees.

For each vertex  $v$  of a tree  $T \in \mathbb{T}_N$  its height  $h(v)$  is defined as the distance to the root of  $T$ , i.e. the length of the shortest path connecting  $v$  to the root along the edges of  $T$ . The height of a finite tree is the maximum height of its vertices.

Let  $n \in \mathbb{N}$ . For any plane tree,  $\pi_n T$  denotes the neighborhood of the root of radius  $n$ , i.e. the subtree of  $T$  spanned by all vertices with height not exceeding  $n$ .

For any  $n$  and sufficiently large  $N$ , the map  $\pi_n$  pushes the measure  $\mu_N$  on  $\mathbb{T}_N$  forward to the measure  $\mu_N \pi_n^{-1}$  on  $S_n$ , the set of all trees with height  $n$ .

**Theorem 1** *For each  $n \in \mathbb{N}$ , the measures  $\mu_N \pi_n^{-1}$  on  $S_n$  converge in total variation, as  $N \rightarrow \infty$ , to a measure  $P_n$ .*

A proof of this theorem will be given in Section 3. At this point we prefer to introduce more definitions that will allow us to describe the limiting measures  $P^n$ .

We define

$$\Delta = \left\{ p = (p_0, \dots, p_D) \in [0, 1]^{D+1} : \sum_{i=0}^D p_i = 1, \sum_{i=0}^D i p_i = 1 \right\},$$

and let

$$J(p) = -H(p) + \beta E(p), \quad p \in \Delta.$$

where

$$H(p) = - \sum_{i=0}^D p_i \ln p_i$$

is the entropy of the probability vector  $p \in \Delta$ , and

$$E(p) = \sum_{i=0}^D p_i E_i$$

is the associated energy.

The function  $J$  is used to construct the rate function in the Large Deviation Principle for large plane trees, see [BH08a],[BH08b].

It is strictly convex and its minimum value on  $\Delta$  is attained at a unique point  $p^*$ . Using Lagrange's method, we find that

$$\ln p_i^* + 1 + \beta E_i + \lambda + i\mu = 0, \quad i = 0, 1, \dots, D,$$

where  $\lambda$  and  $\mu$  are the Lagrange multipliers. So we see that

$$p_i^* = C e^{-\beta E_i} \rho^i, \quad i = 0, 1, \dots, D, \quad (1)$$

where  $C = e^{-1-\lambda}$ , and  $\rho = e^{-\mu}$ . In particular,

$$p_i^* > 0, \quad i = 0, 1, \dots, D. \quad (2)$$

Notice that  $\rho$  can be characterized as a unique solution of

$$\sum_{i=0}^D e^{-\beta E_i} \rho^i = \sum_{i=0}^D i e^{-\beta E_i} \rho^i,$$

and  $C$  may be defined via

$$\frac{1}{C} = \sum_{i=0}^D e^{-\beta E_i} \rho^i = \sum_{i=0}^D i e^{-\beta E_i} \rho^i. \quad (3)$$

We denote  $J^* = J(p^*)$  and  $\sigma = e^{J^*}$ . For a tree  $\tau \in S_n$ , we introduce

$$\bar{E}(\tau) = \sum_{\substack{v \in V(\tau) \\ h(v) < n}} E_{\deg(v)}. \quad (4)$$

Notice that the summation above excludes the highest level of the tree.

**Theorem 2** *For any  $n \in \mathbb{N}$ , the limiting probability measure  $P_n$  is given by*

$$P_n\{\tau\} = Q_n k \rho^k \sigma^m e^{-\beta \bar{E}(\tau)} \quad (5)$$

where the tree  $\tau \in S_n$  is assumed to have  $k$  vertices of height  $n$  and  $m$  vertices of height less than  $n$ . The constant  $Q_n$  is a normalizing factor.

We give a proof of Theorems 1 and 2 in the next Section 3. In Section 4 we compute the value of  $Q_n$  explicitly. In Section 5 we shall see that our convergence results may be interpreted as convergence to an infinite random tree.

### 3 Proof of Theorems 1 and 2

For both theorems it is sufficient to check that for any  $n$  and any two trees  $\tau_1, \tau_2 \in S_n$ ,

$$\lim_{N \rightarrow \infty} \frac{\mu_N \pi_n^{-1}\{\tau_1\}}{\mu_N \pi_n^{-1}\{\tau_2\}} = \frac{k_1 e^{-\beta \bar{E}(\tau_1)} \rho^{k_1} \sigma^{m_1}}{k_2 e^{-\beta \bar{E}(\tau_2)} \rho^{k_2} \sigma^{m_2}}, \quad (6)$$

where we assume that  $\tau_1$  has  $k_1$  vertices of height  $n$ , and  $m_1$  vertices of height less than  $n$ ;  $\tau_2$  has  $k_2$  vertices of height  $n$ , and  $m_2$  vertices of height less than  $n$ .

The energy of each tree  $T$  with  $\pi_n T = \tau_1$  is composed of contributions from the vertices of the tree  $\pi_{n-1} \tau_1$  (we call this contribution  $\bar{E}(\tau_1)$ , see (4)) and the contribution from the plane forest on  $N - m_1$  vertices with  $k_1$  connected components. The same applies to  $\tau_2$ .

Let us recall (see e.g. Theorem 5.3.10 in [Sta99]) that the number of plane forests on  $N$  vertices with  $k$  components and  $r_0, r_1, \dots, r_D$  vertices with branching numbers, respectively,  $0, 1, \dots, D$  is

$$\frac{k}{N} \binom{N}{r_0, r_1, \dots, r_D}$$

if  $r_0 + \dots + r_D = N$ ,  $r_1 + 2r_2 + \dots + Dr_D = N - k$ , and 0 otherwise.

Therefore,

$$\begin{aligned} \frac{\mu_N \pi_n^{-1} \{\tau_1\}}{\mu_N \pi_n^{-1} \{\tau_2\}} &= \frac{e^{-\beta \bar{E}(\tau_1)} \sum_{r \in \Delta(N, m_1, k_1)} \frac{k_1}{N - m_1} \binom{N - m_1}{r_0, r_1, \dots, r_D} e^{-\beta E(r)}}{e^{-\beta \bar{E}(\tau_2)} \sum_{r \in \Delta(N, m_2, k_2)} \frac{k_2}{N - m_2} \binom{N - m_2}{r_0, r_1, \dots, r_D} e^{-\beta E(r)}} \\ &= \frac{e^{-\beta \bar{E}(\tau_1)} I_1(N)}{e^{-\beta \bar{E}(\tau_2)} I_2(N)}. \end{aligned} \quad (7)$$

Here

$$\Delta(N, m, k) = \{r \in \mathbb{Z}_+^{D+1} : r_0 + \dots + r_D = N - m, \\ r_1 + 2r_2 + \dots + Dr_D = N - m - k\},$$

and  $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ .

Fix any  $\varepsilon > 0$  and define

$$\Delta(N, m, k, \varepsilon) = \left\{ r \in \Delta(N, m, k) : \left| \frac{r}{N - m} - p^* \right| < \varepsilon \right\}.$$

We claim that

$$I_1(N) = I_1(N, \varepsilon)(1 + o(1)), \quad N \rightarrow \infty, \quad (8)$$

where

$$I_1(N, \varepsilon) = \sum_{r \in \Delta(N, m_1, k_1, \varepsilon)} \frac{k_1}{N - m_1} \binom{N - m_1}{r_0, r_1, \dots, r_D} e^{-\beta E(r)}.$$

In fact, using Stirling's formula we see that if  $r_i \neq 0$  for all  $i = 0, \dots, D$ ,

$$\begin{aligned} & \frac{k_1}{N - m_1} \binom{N - m_1}{r_0, r_1, \dots, r_D} e^{-\beta E(r)} \\ &= \frac{k_1 (N - m_1)^{N - m_1 - \frac{1}{2}} e^{-\beta E(r)} e^{\frac{\theta_{N - m_1}}{12(N - m_1)} - \frac{\theta_{r_0}}{12r_0} - \dots - \frac{\theta_{r_D}}{12r_D}}}{(2\pi)^{\frac{D}{2}} r_0^{r_0 + \frac{1}{2}} \dots r_D^{r_D + \frac{1}{2}}} \\ &= \frac{k_1 e^{-(N - m_1)J(\frac{r}{N - m_1})}}{((N - m_1)r_0 \dots r_D)^{\frac{1}{2}}} \cdot \frac{e^{\frac{\theta_{N - m_1}}{12(N - m_1)} - \frac{\theta_{r_0}}{12r_0} - \dots - \frac{\theta_{r_D}}{12r_D}}}{(2\pi)^{\frac{D}{2}}}, \end{aligned}$$

with  $0 < \theta_j < 1$  for all  $j \in \mathbb{N}$ , and it is clear that due to the strong convexity of  $J$ , the terms corresponding to the values of  $\frac{r}{N-m_1}$  that are close to its minimum  $p^*$  grow exponentially faster (as  $N \rightarrow \infty$ ) than the terms corresponding to elements of  $\Delta(N, m_1, k_1) \setminus \Delta(N, m_1, k_1, \varepsilon)$ , and there are polynomially many elements in the latter set. This argument can be easily extended to the case where  $r_i = 0$  for some  $i$ , which completes the proof of our claim (8).

Let us now introduce

$$I_2(N, \varepsilon) = \sum_{r \in \Delta'(N, \varepsilon)} \frac{k_2}{N - m_2} \binom{N - m_2}{r_0, r_1, \dots, r_D} e^{-\beta E(r)}.$$

Here  $\Delta'(N, \varepsilon)$  is the image of  $\Delta(N, m_1, k_1, \varepsilon)$  under the map

$$b : \Delta(N, m_1, k_1, \varepsilon) \rightarrow \Delta(N, m_2, k_2)$$

defined by

$$b(r) = (r_0 + (k_2 - k_1), r_1 - (k_2 - k_1) - (m_2 - m_1), r_2, r_3, \dots, r_D).$$

Notice that  $b$  establishes a bijection between  $\Delta(N, m_1, k_1, \varepsilon)$  and  $\Delta'(N, \varepsilon)$  for sufficiently small  $\varepsilon$  and sufficiently large  $N$ .

Using exactly the same reasoning as for  $I_1$ , we see that

$$I_2(N) = I_2(N, \varepsilon)(1 + o(1)), \quad N \rightarrow \infty. \quad (9)$$

Equations (7),(8),(9) imply now that

$$\begin{aligned} \frac{\mu_N \pi_n^{-1} \{\tau_1\}}{\mu_N \pi_n^{-1} \{\tau_2\}} &= \frac{e^{-\beta \bar{E}(\tau_1)} I_1(N, \varepsilon)}{e^{-\beta \bar{E}(\tau_2)} I_2(N, \varepsilon)} (1 + o(1)) \\ &= \frac{k_1 e^{-\beta \bar{E}(\tau_1)}}{k_2 e^{-\beta \bar{E}(\tau_2)}} \cdot \frac{\sum_{r \in \Delta(N, m_1, k_1, \varepsilon)} a_{1,r}}{\sum_{r \in \Delta(N, m_1, k_1, \varepsilon)} a_{2,r}} (1 + o(1)), \quad N \rightarrow \infty, \end{aligned} \quad (10)$$

where

$$a_{1,r} = \binom{N - m_1}{r_0, r_1, \dots, r_D} e^{-\beta E(r)},$$

and

$$a_{2,r} = \binom{N - m_2}{r_0 + (k_2 - k_1), r_1 - (k_2 - k_1) + m_2 - m_1, r_2, \dots, r_D} e^{-\beta E(b(r))}.$$

Assuming that  $k_1 \geq k_2$  and  $m_1 \geq m_2$  (all the other cases can be treated in the same way), we get

$$\frac{a_{1,r}}{a_{2,r}} = \frac{((r_1 - (k_2 - k_1) - (m_2 - m_1)) \dots (r_1 + 1))}{(N - m_2) \dots (N - m_1 + 1) \cdot (r_0 \dots (r_0 + (k_2 - k_1) + 1))} R,$$

where

$$R = R(k_1, m_1, k_2, m_2) = e^{\beta(E_0 - E_1)(k_2 - k_1) - \beta E_1(m_2 - m_1)}.$$

Due to the definition of  $\Delta(N, m, k, \varepsilon)$ ,

$$\frac{a_{1,r}}{a_{2,r}} \leq \frac{((p_1^* + \varepsilon)(N - m_1) - (k_2 - k_1) - (m_2 - m_1))^{-(k_2 - k_1) - (m_2 - m_1)}}{(N - m_1)^{-(m_2 - m_1)}((p_0^* - \varepsilon)(N - m_1) + (k_2 - k_1))^{-(k_2 - k_1)}} R,$$

so that

$$\begin{aligned} \limsup_{N \rightarrow \infty} \sup_{r \in \Delta(N, m, k, \varepsilon)} \frac{a_{1,r}}{a_{2,r}} &\leq (p_1^* + \varepsilon)^{-(m_2 - m_1)} \left( \frac{p_1^* + \varepsilon}{p_0^* - \varepsilon} \right)^{-(k_2 - k_1)} R \\ &\leq \left( \frac{e^{-\beta E_1}}{p_1^* + \varepsilon} \right)^{m_2 - m_1} \left( \frac{(p_0^* - \varepsilon)e^{\beta(E_0 - E_1)}}{p_1^* + \varepsilon} \right)^{k_2 - k_1} \end{aligned} \quad (11)$$

In the same way,

$$\liminf_{N \rightarrow \infty} \inf_{r \in \Delta(N, m, k, \varepsilon)} \frac{a_{1,r}}{a_{2,r}} \geq \left( \frac{e^{-\beta E_1}}{p_1^* - \varepsilon} \right)^{m_2 - m_1} \left( \frac{(p_0^* + \varepsilon)e^{\beta(E_0 - E_1)}}{p_1^* - \varepsilon} \right)^{k_2 - k_1} \quad (12)$$

Since the choice of  $\varepsilon$  is arbitrary, relations (10),(11), and (12) imply that

$$\lim_{N \rightarrow \infty} \frac{\mu_N \pi_n^{-1}\{\tau_1\}}{\mu_N \pi_n^{-1}\{\tau_2\}} = \frac{k_1 e^{-\beta \bar{E}(\tau_1)}}{k_2 e^{-\beta \bar{E}(\tau_2)}} \left( \frac{e^{-\beta E_1}}{p_1^*} \right)^{m_2 - m_1} \left( \frac{p_0^* e^{\beta(E_0 - E_1)}}{p_1^*} \right)^{k_2 - k_1}. \quad (13)$$

Using (1), we see that

$$\frac{p_0^* e^{\beta(E_0 - E_1)}}{p_1^*} = \frac{1}{\rho}. \quad (14)$$

A direct computation based on (1) implies

$$H(p^*) = -\ln(C\rho) + \beta E(p^*).$$

Therefore,

$$\frac{e^{-\beta E_1}}{p_1^*} = \frac{1}{C\rho} = e^{-J(p^*)} = \frac{1}{\sigma}. \quad (15)$$

Now, (6) is an immediate consequence of (13),(14), and (15).  $\square$

## 4 Consistency and the precise value of $Q_n$

We begin with the following consistency property:

**Theorem 3** *The family of measures  $(P_n)_{n \in \mathbb{N}}$  is consistent, i.e. for any  $n$  and any  $\tau \in S_n$*

$$P_n\{\tau\} = \sum_{\substack{\tau' \in S_{n+1} \\ \pi_n \tau' = \tau}} P_{n+1}\{\tau'\}.$$

PROOF: This theorem is a direct consequence of the limiting procedure in Theorem 1. However, it is interesting to derive it from the specific form of  $P_n$  provided by Theorem 2.

Let us assume that  $\tau \in S_n$ , and  $\tau$  has  $n$  vertices of height  $k$  and  $m$  of height less than  $n$ .

$$\begin{aligned}
& \sum_{\substack{\tau' \in S_{n+1} \\ \pi_n \tau = \tau}} P_{n+1}\{\tau'\} \\
&= Q_{n+1} \sum_{i_1, \dots, i_k=0}^D e^{-\beta(\bar{E}(\tau) + E_{i_1} + \dots + E_{i_k})} (i_1 + \dots + i_k) \rho^{i_1 + \dots + i_k} \sigma^{m+k} \\
&= Q_{n+1} e^{-\beta \bar{E}(\tau)} \sigma^{m+k} \sum_{i_1, \dots, i_k=0}^D e^{-\beta(E_{i_1} + \dots + E_{i_k})} (i_1 + \dots + i_k) \rho^{i_1 + \dots + i_k} \\
&= Q_{n+1} e^{-\beta \bar{E}(\tau)} \sigma^{m+k} k \sum_{i_1=0}^D (i_1 \rho^{i_1} e^{-\beta E_{i_1}}) \sum_{i_2=0}^D (\rho^{i_2} e^{-\beta E_{i_2}}) \dots \sum_{i_k=0}^D (\rho^{i_k} e^{-\beta E_{i_k}}) \\
&= Q_{n+1} e^{-\beta \bar{E}(\tau)} \sigma^{m+k} k \frac{1}{C} \left( \frac{1}{C} \right)^{k-1}.
\end{aligned}$$

In this calculation we denoted by  $i_1, \dots, i_k$  the branching numbers of the vertices of height  $n$ . We used the definition of  $P_n$  in the first identity. The second identity is just a convenient rearrangement. The third one follows from the symmetry in the factor  $(i_1 + \dots + i_k)$ . In the last identity we used (3) and the fact that  $p^* \in \Delta$ . Identity (15) implies

$$\frac{1}{C} = \frac{\rho}{\sigma}, \quad (16)$$

so that

$$\sum_{\substack{\tau' \in S_{n+1} \\ \pi_n \tau = \tau}} P_{n+1}\{\tau'\} = Q_{n+1} e^{-\beta \bar{E}(\tau)} \sigma^{m+k} k \sigma^{-k} \rho^k = \frac{Q_{n+1}}{Q_n} P_n\{\tau\}. \quad (17)$$

Since this holds true for all  $\tau \in S_n$ , we can conclude that  $Q_n = Q_{n+1}$  which completes the proof.  $\square$

Identity (17) means that the constant  $Q = Q_n$  in Theorem 2 is the same for all  $n$ . Choosing  $n = 1$  we can compute it using (5):

$$1 = Q \sum_{k=1}^D k e^{-\beta E_k} \rho^k \sigma^1 = \frac{Q\sigma}{C}.$$

A more precise version of Theorem 2 easily follows:

**Theorem 4** *Let  $C$  be defined by (3). For each  $n$ , the limiting probability measure  $P_n$  is given by*

$$P^n\{\tau\} = C k e^{-\beta \bar{E}(\tau)} \rho^k \sigma^{m-1},$$

where the tree  $\tau \in S_n$  is assumed to have  $k$  vertices of height  $n$  and  $m$  vertices of height less than  $n$ .



## 5 The limiting random tree

Let  $S_\infty$  be the set of infinite plane trees with branching number bounded by  $D$ . Theorem 3 along with the classical Daniell—Kolmogorov Consistency theorem (see [Bil99]) allows us to introduce a measure  $P_\infty$  on  $S_\infty$  consistent with measures  $P_n$  for all  $n$ . Intuitively this is clear, but to make it precise we need to introduce a coding of plane trees. We have chosen one of several possible coding schemes.

Let  $T$  be a plane tree (finite or infinite) with branching bounded by  $D$ . Then  $T$  has a finite number  $r_n \leq D^n$  of vertices of any given height  $n$ . Let us say that all vertices of the same height  $n$  form the  $n$ -th level of the tree. The vertices of  $n$ -th level are naturally ordered and can be enumerated by numbers from 1 to  $r_n$  (except for the case when there are no vertices at  $n$ -th level at all). Each of  $r_n$  vertices of the  $n$ -th level has a parent at the level  $n - 1$ . Denote the number received by the parent of  $l$ -th vertex of the  $n$ -th level under the described enumeration by  $g_{n,l}$ . If  $r_n < l \leq D^n$  we set  $g_{n,l} = 0$ . Then the  $n$ -th level  $n$  can be encoded by a vector

$$g_n = (g_{n,1}, \dots, g_{n,D^n}) \in \{0, 1, \dots, D^{n-1}\}^{D^n},$$

and the whole tree can be identified with the sequence of levels

$$(g_1, g_2, \dots) \in \mathbb{X} = \prod_{n=1}^{\infty} \{0, 1, \dots, D^{n-1}\}^{D^n},$$

so that the space  $\mathbb{T}$  of all plane trees (finite or infinite) with branching bounded by  $D$  can be identified with a subset of  $\mathbb{X}$ .

**Theorem 5** *There is a unique measure  $P_\infty$  on  $\mathbb{T}$  such that it is consistent with measures  $P^n$ :*

$$P_\infty \pi_n^{-1} = P^n.$$

*This measure is concentrated on  $S_\infty$ .*

PROOF: The first statement follows from Theorem 3 and the Consistency theorem. The second statement is a consequence of the fact that for each  $n \in \mathbb{N}$ ,  $P_n$  is concentrated on trees with positive number of vertices at  $n$ -th level.  $\square$

The space  $\mathbb{X}$  is compact in the product topology. Therefore, the convergence of finite-dimensional distributions established in Theorem 1 and the classical Prokhorov theorem (see e.g. [Bil99]) imply the following result:

**Theorem 6** *As  $N \rightarrow \infty$ , measures  $P_N$  viewed as measures on  $\mathbb{X}$  converge weakly to  $P_\infty$  in the product topology.*

This statement shows that there is a limiting object for the random trees that we consider. This object is an infinite random tree. For any  $n \in \mathbb{N}$ , the first  $n$  levels of this random tree are distributed according to  $P_n$ .

Let us now embed the space  $\mathbb{X}$  into  $\bar{\mathbb{X}} = (\mathbb{Z}_+^{\mathbb{N}})^{\mathbb{Z}_+}$  filling up all the unused coordinates with zeros. The measure  $P_\infty$  can be treated as a measure on  $\bar{\mathbb{X}}$  thus generating a  $\mathbb{Z}_+^{\mathbb{N}}$ -valued process  $(X_n)_{n=0}^\infty$  with discrete time. This process along with the associated random tree is visualized on Figure 2.

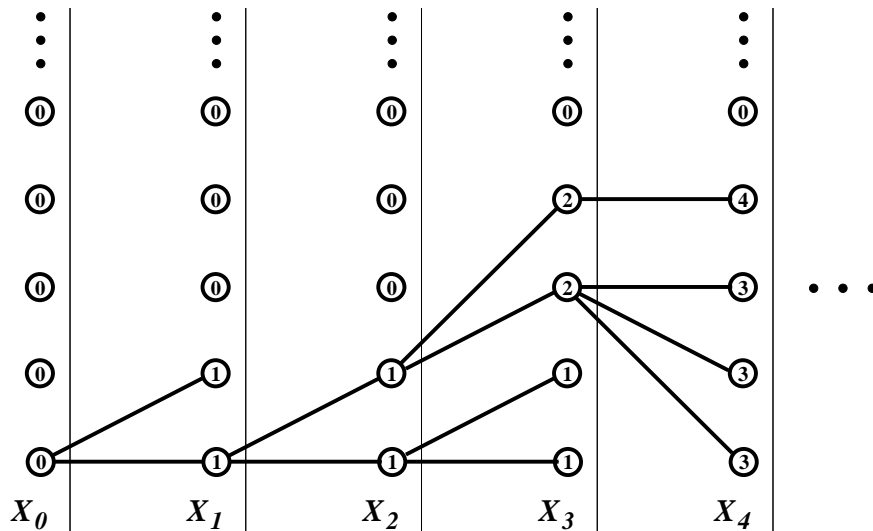


Figure 2: A realization of the process  $(X_n)$  and the associated random tree

For any  $n$ ,  $X_n$  describes how the  $n$ -th level of the tree is built upon the  $n - 1$ -th one.

For a level  $g \in \mathbb{Z}_+^{\mathbb{N}}$ , we denote by  $|g|$  the number of non-zero entries in  $g$  (i.e. the number of vertices at the level). We also introduce

$$E(g) = \sum_{i=1}^{\infty} E_{\#\{j: g_j=i\}},$$

the energy induced by level  $g$  at its parent level. Notice that the sum in the r.h.s. is in fact finite. For two levels  $g$  and  $g'$  we write  $g \triangleleft g'$  if  $\max_i g'_i \leq |g|$ .

Theorem 4 immediately implies the following result:

**Theorem 7** *The process  $(X_n)$  defined above is Markov with transition probability*

$$\mathbb{P}\{X_{n+1} = g' \mid X_n = g\} = \begin{cases} \frac{|g'|}{|g|} e^{-\beta E(g')} \rho^{|g'| - |g|} \sigma^{|g|}, & g \triangleleft g', \\ 0, & \text{otherwise.} \end{cases}$$

## 6 A limit theorem for the size of $n$ -th level

Let us introduce  $Y_n = |X_n|$ , the random number of vertices at  $n$ -th level. The following statement is a direct consequence of Theorem 4 or Theorem 7:

**Theorem 8** *The process  $(Y_n)_{n=0}^\infty$  is Markov with transition probability*

$$\mathbb{P}\{Y_{n+1} = k' \mid Y_n = k\} = \frac{k'}{k} \rho^{k'-k} \sigma^k \sum_{\substack{0 \leq i_1, \dots, i_k \leq D \\ i_1 + \dots + i_k = k'}} e^{-\beta(E_{i_1} + \dots + E_{i_k})}.$$

The next theorem shows that in fact  $Y_n$  grows linearly in time. Let

$$\mu = B_2 - 1, \quad (18)$$

where

$$B_n = \sum_{i=0}^D i^n p_i^*, \quad n \in \mathbb{N}.$$

Then  $\mu > 0$  being the variance of  $p^*$ , a nondegenerate distribution.

**Theorem 9**

$$\frac{Y_n}{n} \xrightarrow[\mu]{L} \Gamma,$$

where  $\Gamma$  is a random variable with density

$$p(t) = \begin{cases} te^{-t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

PROOF: Let us find the Laplace transform of the distribution of  $Y_n$ :

$$L_n(s) = \mathbb{E}e^{sY_n}, \quad s \leq 0,$$

(this definition differs from the traditional one by a sign change of the argument) and prove that for any  $x \leq 0$ ,

$$\lim_{n \rightarrow \infty} L_n\left(\frac{x}{n}\right) = \frac{1}{\left(1 - \frac{\mu x}{2}\right)^2} =: L_\infty(x), \quad (19)$$

the r.h.s. being the Laplace transform of

$$p(t) = \frac{4t}{\mu^2} e^{-\frac{2t}{\mu}},$$

the density of the r.v.  $\frac{\mu}{2}\Gamma$ . This will imply the desired result, see e.g. [Kal86, Appendix 5] for various statements on Laplace transforms.

Theorem 8 and (16) imply

$$\begin{aligned} \mathbb{E}\left[e^{sY_{n+1}} \mid Y_n = k\right] &= \sum_{k'} \frac{e^{sk'} k'}{k} \rho^{k'-k} \sigma^k \sum_{\substack{0 \leq i_1, \dots, i_k \leq D \\ i_1 + \dots + i_k = k'}} e^{-\beta(E_{i_1} + \dots + E_{i_k})} \\ &= \frac{\sigma^k}{k \rho^k} \sum_{0 \leq i_1, \dots, i_k \leq D} (i_1 + \dots + i_k) \rho^{i_1 + \dots + i_k} e^{s(i_1 + \dots + i_k)} e^{-\beta(E_{i_1} + \dots + E_{i_k})} \\ &= w(s)v(s)^{k-1}, \end{aligned}$$

where

$$v(s) = \sum_{i=0}^D p_i^* e^{si} = \sum_{i=0}^D C \rho^i e^{-\beta E_i} e^{si},$$

and

$$w(s) = \sum_{i=0}^D ip_i^* e^{si} = \sum_{i=0}^D Ci\rho^i e^{-\beta E_i} e^{si} = v'(s).$$

Therefore,

$$L_{n+1}(s) = \mathbb{E}w(s)v(s)^{Y_n-1} = \frac{w(s)}{v(s)} \mathbb{E}e^{\ln v(s)Y_n} = z(s)L_n(f(s)), \quad (20)$$

where

$$f(s) = \ln v(s), \quad (21)$$

and

$$z(s) = \frac{w(s)}{v(s)} = f'(s).$$

Both  $z$  and  $f$  are analytic functions. An elementary calculation shows that

$$f(s) = s + \frac{\mu}{2}s^2 + r(s),$$

and

$$\ln z(s) = \mu s + q(s),$$

where  $\mu = w'(0) - 1$  was introduced in (18) and

$$|r(s)| \leq c|s|^3, \quad |q(s)| \leq c|s|^2 \quad (22)$$

for some  $c > 0$  and all  $s \leq 0$ .

From now on,  $x \leq 0$  is fixed. Using (20) and the obvious identity

$$L_0(s) = e^s,$$

we can write

$$L_n\left(\frac{x}{n}\right) = z\left(\frac{x}{n}\right) \cdot z\left(f\left(\frac{x}{n}\right)\right) \cdot z\left(f^2\left(\frac{x}{n}\right)\right) \cdot \dots \cdot z\left(f^{n-1}\left(\frac{x}{n}\right)\right) e^{f^n\left(\frac{x}{n}\right)},$$

where

$$f^k(x) = \underbrace{f \circ f \dots \circ f}_k(x), \quad k \geq 0,$$

so that we have to study the numbers  $(x_{n,k})_{n \in \mathbb{N}, k=0, \dots, n}$  defined by

$$x_{n,k} = f^k\left(\frac{x}{n}\right).$$

We shall compare  $(x_{n,k})_{n \in \mathbb{N}, k=0, \dots, n}$  to  $(y_{n,k})_{n \in \mathbb{N}, k=0, \dots, n}$  defined by

$$y_{n,k} = \frac{1}{\frac{n}{x} - \frac{\mu}{2}k}.$$

For fixed  $n$ , both sequences  $(x_{n,k})$  and  $(y_{n,k})$  are negative and increasing in  $k$ . Therefore

$$|x_{n,k}| \leq |x_{n,0}| = \frac{|x|}{n},$$

and

$$|y_{n,k}| \leq |y_{n,0}| = \frac{|x|}{n}.$$

Let us prove that for sufficiently large  $n$  and any  $k$  between 0 and  $n$ ,

$$|y_{n,k} - x_{n,k}| \leq k \left( \frac{\mu^2}{4} + c \right) \left( \frac{|x|}{n} \right)^3. \quad (23)$$

This is certainly true for  $k = 0$ . For the induction step, we write

$$|y_{n,k} - x_{n,k}| \leq |y_{n,k} - f(y_{n,k-1})| + |f(y_{n,k-1}) - f(x_{n,k-1})| = I_1 + I_2.$$

A straightforward computation based on (21) shows that

$$|I_1| = \left| \frac{\mu^2}{4} y_{n,k-1}^2 y_{n,k} + r(y_{n,k-1}) \right| \leq \left( \frac{\mu^2}{4} + c \right) \left( \frac{|x|}{n} \right)^3$$

Since  $|f'(s)| \leq 1$  for all sufficiently small  $s$ , we see that

$$|I_2| \leq |y_{n,k-1} - x_{n,k-1}|.$$

Combining these estimates we see that

$$|y_{n,k} - x_{n,k}| \leq \left( \frac{\mu^2}{4} + c \right) \left( \frac{|x|}{n} \right)^3 + |y_{n,k-1} - x_{n,k-1}|,$$

and our claim (23) follows. It immediately implies that

$$|y_{n,k} - x_{n,k}| \leq \frac{C}{n^2} \quad (24)$$

for some  $K = K(x)$ , sufficiently large  $n$  and all  $k$ . We can now write

$$\begin{aligned} \ln L_n \left( \frac{x}{n} \right) &= \sum_{k=0}^{n-1} \ln z(x_{n,k}) + x_{n,n} \\ &= \sum_{k=0}^{n-1} \mu x_{n,k} + \sum_{k=0}^{n-1} q(x_{n,k}) + x_{n,n} \\ &= \sum_{k=0}^{n-1} \mu y_{n,k} + \sum_{k=0}^{n-1} \mu (x_{n,k} - y_{n,k}) + \sum_{k=0}^{n-1} q(x_{n,k}) + x_{n,n} \\ &= I_1 + I_2 + I_3 + I_4. \end{aligned}$$

It is straightforward to see that  $\lim_{n \rightarrow \infty} I_2 + I_3 + I_4 = 0$ . The first term

$$I_1 = \mu \sum_{k=0}^{n-1} \frac{1}{\frac{x}{n} - \frac{\mu}{2}k} = \mu x \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{1 - \frac{\mu x}{2} \frac{k}{n}}$$

can be viewed as a Riemann integral sum, so that

$$\lim_{n \rightarrow \infty} \ln L_n \left( \frac{x}{n} \right) = \mu x \int_0^1 \frac{du}{1 - \frac{\mu x}{2} u} = -2 \ln \left( 1 - \frac{\mu x}{2} \right),$$

which immediately implies (19).  $\square$

## 7 A functional limit theorem

In this section we prove the following theorem on diffusion approximation for the process  $Y$ :

**Theorem 10** *Let*

$$Z_n(t) = \frac{Y_{[nt]}}{n}, \quad n \in \mathbb{N}, t \in \mathbb{R}_+.$$

*Then, as  $n \rightarrow \infty$ , the distribution of  $Z_n$  converges weakly in the Skorokhod topology in  $D[0, \infty)$  to the unique nonnegative weak solution  $Z$  of the stochastic Itô equation*

$$\begin{aligned} dZ(t) &= \mu dt + \sqrt{\mu Z(t)} dW(t), \\ Z(0) &= 0. \end{aligned}$$

PROOF: Since the initial point  $Z(0) = 0$  is an “entrance and non-exit” singular point for the positive semi-axis (see the classification of singular points in [IM74]), the existence and uniqueness of a solution nonnegative for all positive times is guaranteed. Let us define

$$b(x) \equiv \mu, \quad \text{and } a(x) = \mu(x \vee 0), \quad x \in \mathbb{R},$$

and extend the equation above to the negative semi-axis by

$$dZ(t) = b(Z(t))dt + \sqrt{a(Z(t))}dW(t).$$

An obvious argument shows that there is no solution starting at 0 and being negative for some  $t > 0$ . Therefore the weak existence and uniqueness in law hold for (7). According to Section 5.4B of [KS98], this existence and uniqueness is equivalent to the well-posedness of the martingale problem associated with  $b$  and  $a$ .

We will use Theorem 4.1 from [EK86, Chapter 7] on diffusion approximation. The coefficients  $a, b$  were defined on the whole real line so as the theorem applies directly, with no modification. We proceed to check its conditions.

We must find processes  $A_n$  and  $B_n$  with the following properties:

1. Trajectories of  $A_n$  and  $B_n$  are in  $D[0, \infty)$ .
2.  $A_n$  is nondecreasing.
3.  $M_n = Z_n - B_n$  and  $M_n^2 - A_n$  are martingales with respect to the natural filtration generated by  $Z_n, A_n, B_n$ .
4. For every  $T > 0$  the following holds true:

$$\lim_{n \rightarrow \infty} \mathbf{E} \sup_{t \leq T} |Z_n(t) - Z_n(t-)|^2 = 0, \quad (25)$$

$$\lim_{n \rightarrow \infty} \mathbf{E} \sup_{t \leq T} |A_n(t) - A_n(t-)| = 0, \quad (26)$$

$$\lim_{n \rightarrow \infty} \mathbf{E} \sup_{t \leq T} |B_n(t) - B_n(t-)|^2 = 0, \quad (27)$$

$$\sup_{t \leq T} \left| B_n(t) - \int_0^t b(Z_n(s)) ds \right| \xrightarrow{\mathbf{P}} 0, \quad n \rightarrow \infty, \quad (28)$$

$$\sup_{t \leq T} \left| A_n(t) - \int_0^t a(Z_n(s)) ds \right| \xrightarrow{\mathbf{P}} 0, \quad n \rightarrow \infty. \quad (29)$$

We shall need the following lemma:

**Lemma 1**

$$\begin{aligned}
\mathbb{E}[Y_{j+1}|Y_j = k] &= \mu + k, \\
\mathbb{E}[Y_{j+1}^2|Y_j = k] &= B_3 + 3(k-1)B_2 + (k-1)(k-2), \\
\mathbb{E}[Y_{j+1}^3|Y_j = k] &= B_4 + 4(k-1)B_3 + 6(k-1)(k-2)B_2 + 3(k-1)B_2^2 \\
&\quad + (k-1)(k-2)(k-3), \\
\mathbb{E}[Y_{j+1}^4|Y_j = k] &= B_5 + 5(k-1)B_4 + 10(k-1)(k-2)B_3 + 10(k-1)B_3B_2 \\
&\quad + 15(k-1)(k-2)B_2^2 + 10(k-1)(k-2)(k-3)(k-4).
\end{aligned}$$

PROOF: For the first of these identities, we write

$$\begin{aligned}
\mathbb{E}[Y_{j+1}|Y_j = k] &= \frac{\sigma^k}{k\rho^k} \sum_{0 \leq i_1, \dots, i_k \leq D} (i_1 + \dots + i_k)^2 \rho^{i_1 + \dots + i_k} e^{-\beta(E_{i_1} + \dots + E_{i_k})} \\
&= \frac{1}{k} \left[ k \left( \sum_{i_1=0}^D i_1^2 C \rho^{i_1} e^{-\beta E_{i_1}} \right) \left( \sum_{i_2=0}^D C \rho^{i_2} e^{-\beta E_{i_2}} \right)^{k-1} \right. \\
&\quad \left. + k(k-1) \left( \sum_{i_1=0}^D i_1 C \rho^{i_1} e^{-\beta E_{i_1}} \right)^2 \left( \sum_{i_2=0}^D C \rho^{i_2} e^{-\beta E_{i_2}} \right)^{k-2} \right] \\
&= \frac{1}{k} (kB_2 + k(k-1)) = B_2 + k - 1 \\
&= \mu + k,
\end{aligned}$$

where we used the symmetry of the terms  $(i_1^2 + \dots + i_k^2)$ ,  $i_1 i_2 + i_1 i_3 + \dots + i_{k-1} i_k$  and (16). Next,

$$\begin{aligned}
\mathbb{E}[Y_{j+1}^2|Y_j = k] &= \frac{\sigma^k}{k\rho^k} \sum_{0 \leq i_1, \dots, i_k \leq D} (i_1 + \dots + i_k)^3 \rho^{i_1 + \dots + i_k} e^{-\beta(E_{i_1} + \dots + E_{i_k})} \\
&= \frac{1}{k} (kB_3 + 3k(k-1)B_2 + k(k-1)(k-2)) \\
&= B_3 + 3(k-1)B_2 + (k-1)(k-2),
\end{aligned}$$

and the other two identities in the statement of the lemma can be obtained in a similar way.  $\square$

Returning to the proof of the functional limit theorem, let us find the coefficient  $B_n(t)$  first. The process  $Z_n$  is constant on any interval of the form  $[j/n, (j+1)/n)$ . Due to Lemma 1,

$$\mathbb{E} \left[ Z_n \left( t + \frac{1}{n} \right) \middle| Z_n(t) \right] = Z_n(t) + \mu \frac{1}{n}, \quad (30)$$

so that we can set  $B_n(t) = \mu[nt]/n$  to satisfy the martingale requirement on  $M_n = Z_n - B_n$ . Notice that with this choice of  $B_n$ , relations (27) and (28) are easily seen to be satisfied. Lemma 1 also implies

$$\begin{aligned}
\mathbb{E} \left[ Z_n^2 \left( t + \frac{1}{n} \right) \middle| Z_n(t) \right] \\
= \frac{B_3 + 3(nZ_n(t) - 1)B_2 + (nZ_n(t) - 1)(nZ_n(t) - 2)}{n^2}, \quad (31)
\end{aligned}$$

so that for  $t \in \frac{1}{n}\mathbb{Z}$ ,

$$\mathbf{E} \left[ M_n^2 \left( t + \frac{1}{n} \right) - M_n^2(t) \middle| Z_n(t) \right] = \frac{1}{n} \mu Z_n(t) + \frac{1}{n^2} (B_3 - B_2^2 - B_2 + 1).$$

Therefore we can set

$$A_n(t) = \sum_{j: \frac{j}{n} \leq t} \left( \frac{\mu}{n} Z_n \left( \frac{j}{n} \right) + \frac{[nt]}{n^2} (B_3 - B_2^2 - B_2 + 1) \right),$$

to satisfy the martingale requirement on  $M_n^2 - A$ . Notice that  $A_n$  is nondecreasing since

$$\begin{aligned} \frac{1}{n} \mu Z_n(t) + \frac{1}{n^2} (B_3 - B_2^2 - B_2 + 1) &\geq \frac{1}{n^2} (B_2 - 1) + \frac{1}{n^2} (B_3 - B_2^2 - B_2 + 1) \\ &\geq \frac{1}{n^2} (B_3 - B_2^2) \geq 0, \end{aligned}$$

where the last inequality follows from the Cauchy–Schwartz inequality and  $B_1 = 1$ . So properties 1–3 are satisfied, and (29) follows from the definitions of  $a$  and  $A$ , and the convergence

$$\lim_{n \rightarrow \infty} \sup_{t \leq T} \sum_{j: \frac{j}{n} \leq t} \frac{1}{n^2} (B_2 - 1) + \frac{1}{n^2} (B_3 - B_2^2 - B_2 + 1) = 0.$$

To prove (26) we use the definition of  $A$  to write

$$\mathbf{E} \sup_{t \leq T} |A_n(t) - A_n(t-)| \leq \frac{\mu}{n} \mathbf{E} \sup_{\frac{j}{n} \leq T} Z_n \left( \frac{j}{n} \right) + \frac{1}{n^2} (B_3 - B_2^2 - B_2 + 1)$$

so it suffices to prove that  $\mathbf{E} \sup_{\frac{j}{n} \leq T} Z_n \left( \frac{j}{n} \right)$  is bounded. The definition of  $M_n$ , Lyapunov's inequality and Doob's maximal inequality for submartingales imply that for some  $c > 0$ :

$$\begin{aligned} \mathbf{E} \sup_{\frac{j}{n} \leq T} Z_n \left( \frac{j}{n} \right) &\leq \mu T + \mathbf{E} \sup_{\frac{j}{n} \leq T} \left| M_n \left( \frac{j}{n} \right) \right| \\ &\leq \mu T + c \sqrt{\mathbf{E} M_n^2(T)} \\ &\leq \mu T + c \sqrt{2(\mathbf{E} Z_n^2(T) + \mu^2 T^2)}. \end{aligned}$$

Lemma 1 implies that  $\mathbf{E} Z_n^2(T)$  has a limit, as  $n \rightarrow \infty$ , so that (26) is verified.

A lengthy but elementary calculation based on Lemma 1 shows that

$$\mathbf{E}[(Y_{j+1} - Y_j)^4 | Y_n] \leq c(Y_j^2 + 1)$$



for some constant  $c > 0$ , so that we can write

$$\begin{aligned} \mathbf{E} \sup_{t \leq T} (Z_n(t) - Z_n(t-))^2 &\leq \left[ \mathbf{E} \sup_{t \leq T} (Z_n(t) - Z_n(t-))^4 \right]^{1/2} \\ &\leq \left[ \frac{1}{n^4} \sum_{j: \frac{j}{n} \leq T} \mathbf{E} (Y_{j+1} - Y_j)^4 \right]^{1/2} \\ &\leq \left[ \frac{c}{n^4} \sum_{j \leq nT} \mathbf{E} (Y_j^2 + 1) \right]^{1/2} \end{aligned}$$

Since Lemma 1 implies that for some constant  $c_1 > 0$ ,

$$\mathbf{E} (Y_j^2 + 1) \leq c_1 j^2, \quad j \in \mathbb{N},$$

we conclude that

$$\mathbf{E} \sup_{t \leq T} (Z_n(t) - Z_n(t-))^2 \leq \sqrt{\frac{c}{n^4} \cdot n \cdot c_1 n^2 T^2} \rightarrow 0, \quad n \rightarrow \infty,$$

and the proof of the theorem is complete.  $\square$

## 8 Diffusion limit for finer structure of the random tree

In this section we present a non-rigorous and sketchy description for the diffusion limit of the infinite Markov random tree itself rather than its width given by  $Y_n$  at time  $n$ . Let us fix any time  $n_0$  and divide all  $Y_{n_0}$  vertices into  $r$  nonempty disjoint groups. For any  $n \geq n_0$  denote the progeny of  $i$ -th group at time  $n$  by  $V_{i,n}$ .

We want to study the coevolution of  $(V_{1,n}, \dots, V_{r,n})$ . Though each  $V_{i,n}$  is not a Markov process, it is elementary to see that the whole vector is a homogeneous Markov process. We would like to compute the diffusion limit for this vector under an appropriate rescaling:

$$\frac{1}{n} (V_{1,[nt]}, \dots, V_{r,[nt]})$$

We need to find the local drift and diffusion coefficients for the limiting process. Let  $j_1 + \dots + j_r = k$ . Then computations similar to Lemma 1 produce

$$\mathbf{E} \left[ \frac{V_{1,m+1}}{n} - \frac{j_1}{n} \mid \frac{1}{n} (V_{1,[nt]}, \dots, V_{r,[nt]}) = \frac{1}{n} (j_1, \dots, j_r) \right] = \mu \frac{j_1/n}{k/n} \frac{1}{n},$$

so, by symmetry, the local limit drift is

$$b_i(v) = \mu \frac{v_i}{v_1 + \dots + v_r}.$$

Similarly, the diagonal terms for local diffusion:

$$\begin{aligned} & \mathbb{E} \left[ \left( \frac{V_{1,m+1}}{n} - \frac{j_1}{n} \right)^2 \middle| \frac{1}{n}(V_{1,[nt]}, \dots, V_{r,[nt]}) = \frac{1}{n}(j_1, \dots, j_r) \right] \\ &= \mu \frac{j_1}{n} \frac{1 + \frac{B_3 - 3B_2 + 2}{k}}{n}, \end{aligned}$$

and

$$a_{ii}(v) = \mu v_i.$$

For the off-diagonal terms a computation produces

$$\mathbb{E}[(V_{1,m+1} - j_1)(V_{2,m+1} - j_2) | V_{1,m} = j_1, \dots, V_{r,m} = j_m] = 0,$$

so that

$$a_{ij} \equiv 0, \quad i \neq j.$$

So, the limiting equations are

$$dV_i(t) = \mu \frac{V_i(t)}{\sum_j V_j(t)} dt + \sqrt{\mu V_i(t)} \mathbf{1}_{\{V_i > 0\}} dW_i(t).$$

Let us introduce cumulative counts

$$U_j = V_1 + \dots + V_j.$$

Then  $Z(t) = U_r(t)$ , and

$$\begin{aligned} dU_j &= \mu \frac{U_j}{U_r} dt + \sqrt{\mu U_1} \mathbf{1}_{\{U_1 > 0\}} dW_1 + \dots \\ &\quad \dots + \sqrt{\mu(U_j - U_{j-1})} \mathbf{1}_{\{U_j - U_{j-1} > 0\}} dW_j \quad (32) \end{aligned}$$

Then, for each  $0 \leq u_1 \leq \dots \leq u_r$ , we can solve this equation with initial data

$$(U_1(t_0), \dots, U_r(t_0)) = (u_1, \dots, u_r),$$

which gives a random nondecreasing map

$$\Phi = \Phi_{t_0} : u \mapsto (U(t))_{t \geq t_0}. \quad (33)$$

Here  $u$  runs through the set  $\{u_1, \dots, u_r\}$ . It is clear though that if we insert another point  $u'$  between 0 and  $u^* = u_r$ , then solving the stochastic equation above for the modified set of initial points, we see that the new solution map is a monotone extension of the old one. Adding points of a countable dense set one after another, we can extend the solution map onto  $u \in [0, u^*]$ . So, for each  $u^* \geq 0$  we are able to define a random monotone continuous maps  $\Phi : [0, u^*] \rightarrow \mathbb{R}_+^{[t_0, \infty)}$ .

Our last point is to represent these solution maps via stochastic integrals w.r.t. a Brownian sheet  $(W(x, t))_{t, x \geq 0}$ , i.e. a continuous Gaussian random field with zero mean and

$$\text{cov}(W(x_1, t_1), W(x_2, t_2)) = (x_1 \wedge x_2)(t_1 \wedge t_2), \quad x_1, x_2, t_1, t_2 \geq 0.$$

Equations (32) imply that  $\Phi(u, t)$ ,  $t \geq t_0$ ,  $u \in [0, u^*]$  is equal in law to the solution of the following SPDE:

$$d\Phi(u, t) = \mu \frac{\Phi(u, t)}{\Phi(u^*, t)} dt + \int_{x \in \mathbb{R}} \mathbf{1}_{[0, \sqrt{\mu\Phi(u, t)}} W(dx \times dt),$$

$$\Phi(u, t_0) = u, \quad u \in [0, u^*].$$

A rigorous treatment of the limiting solution  $\Phi$ , and a precise convergence statement will appear elsewhere.

## References

- [Ald91a] David Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.
- [Ald91b] David Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.
- [BH08a] Yuri Bakhtin and Christine Heitsch. Large deviations for random trees. *J. Stat. Phys.*, 132(3):551–560, 2008.
- [BH08b] Yuri Bakhtin and Christine Heitsch. Large deviations for random trees and the branching of RNA secondary structures. 2008. Submitted to *Bulletin of Mathematical Biology*.
- [Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.
- [Git98] Bernhard Gittenberger. Convergence of branching processes to the local time of a Bessel process. In *Proceedings of the Eighth International Conference “Random Structures and Algorithms” (Poznan, 1997)*, volume 13, pages 423–438, 1998.
- [IM74] Kiyosi Itô and Henry P. McKean, Jr. *Diffusion processes and their sample paths*. Springer-Verlag, Berlin, 1974. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.
- [Kal86] Olav Kallenberg. *Random measures*. Akademie-Verlag, Berlin, fourth edition, 1986.
- [KS98] Ioannis Karatzas and Steven E. Shreve. *Methods of mathematical finance*, volume 39 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1998.
- [Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.