Martingales and Profile of Binary Search Trees

B. Chauvin, T. Klein, J-F. Marckert , A. Rouault,

Université de Versailles
45 Avenue des Etats Unis
78035 Versailles Cedex
France

Abstract

We are interested in the asymptotic analysis of the binary search tree (BST) under the random permutation model. Via an embedding in a continuous time model, we get new results, in particular the asymptotic behavior of the profile.

Key words. Binary search tree, fragmentation, branching random walk, convergence of martingales, probability tilting.

A.M.S. Classification. `60J25, 60J80, 68W40, 60C05, 60G42, 60G44.`

1 Introduction

This paper deals mainly with two classical models of binary trees processes: the binary search tree process and the Yule tree process.

• A labeled binary search tree (LBST) is a structure used in computer science to store totally ordered data. At time 0 the LBST is reduced to a leaf without label. Each unit of time, a new item is inserted in a leaf of the tree. This leaf is then replaced by an internal node with two leaves. We are interested in the sequence of underlying unlabeled trees $(\mathcal{T}_n)_n$ induced by this construction. We call this sequence the binary search tree process, or BST process.

• The Yule tree process $(\mathbb{T}_t)_t$ is a continuous time (unlabeled) binary tree process in which each leaf behaves independently from the other ones (at time 0, the tree $\mathbb{T}_0$ is reduced to a leaf). After an (random) exponential time, a leaf has two children. Due to the lack of memory of the exponential distribution, each leaf is equally likely the first one to produce children.

Under a suitably chosen random model of data (the random permutation model), the two models of trees are deeply related. In the Yule tree process, let τ_n be the random time when the $n+1$th leaf appears. Under the random permutation model the link between the two models is the following one: the process $(\mathbb{T}_{\tau_n})_n$ has the same law as $(\mathcal{T}_n)_n$. This allows the construction of the BST process and the Yule tree process on the same probability space on which $(\mathcal{T}_n)_n = (\mathbb{T}_{\tau_n})_n$. This embedding of the BST process into a continuous time model allows to use independence properties between subtrees in the Yule tree process (it is a kind of Poissonization). Many functionals of the BST can then be derived using known results on the Yule tree. An interesting quantity is the profile of $\mathcal{T}_n$ which is the sequence $(U_k(n))_{k\geq 0}$ where $U_k(n)$ is the number of leaves of $\mathcal{T}_n$ at level k. Here, in (21), the martingale family $(\mathcal{M}_n(z))_n$ – the Jabbour's martingale – which encodes the profile of $(\mathcal{T}_n)_n$ is shown to be strongly related to the martingale family $(M_t(z))_t$ that encodes the profile of $(\mathbb{T}_t)_t$.

The aim of the present paper is to revisit the study of $(\mathcal{M}_n(z))_n$ using the embedding. For $z > 0$, we recover very quickly the behavior of the limit $\mathcal{M}_\infty(z)$: positive when $z \in (z_c^-, z_c^+)$, zero when $z \notin [z_c^-, z_c^+]$. In the critical cases $z = z_c^\pm$ the behavior was unknown. We prove that $\mathcal{M}_\infty(z_c^\pm) = 0$ a.s. and get the convergence of the derivative. The limits $\mathcal{M}'_\infty(z)$ and $\mathcal{M}_\infty(z)$ satisfy a splitting formula (40) which, for $z = 1$ gives the Quicksort equation (Corollary 3.6). Thus, the embedding method is the key tool for proving and enlarging convergence results on the BST martingale (Theorem 3.3) and its derivative (Theorem 3.5). See the companion paper [16] for complements.

The paper is organized as follows. After the definition of the models in Section 2, we explore some consequences of the embedding. In particular, in 2.5.1 we exhibit a family of uniform random variables attached to the nodes of the Yule tree. These random variables give, for every node u, the limiting proportion of leaves issued from u among those issued from its parent. A similar property holds for the embedded BST. In subsection 2.5.2, the appearance of uniform variables as limiting proportion of leaves is explained on a LBST model. In Section 3, we study the convergence, as $n \to \infty$, of the BST martingale $\mathcal{M}_n(z)$.

Thanks to this method, we are able in Section 4 to describe the asymptotic behavior of the profile $U_k(n)$ when $k \simeq 2z \log n$ in the whole range $z \in (z_c^-, z_c^+)$. Previously, the result was known only on a sub-domain where the L^2 method works ([15]).

Finally, in Section 5, an other point of view is investigated. In biasing the evolution rules of the BST – a random line of descent is distinguished, and the evolution of the nodes belonging to this branch is different from the other ones – it appears that the behavior of the nodes on the distinguished line gives information on the whole tree.

2 The models

2.1 Binary search trees

For a convenient definition of trees we are going to work with, let us first define

$$\mathbb{U} = \{\emptyset\} \cup \bigcup_{n \geq 1} \{0,1\}^n$$

the set of finite words on the alphabet $\{0,1\}$ (with $\emptyset$ for the empty word). For u and v in $\mathbb{U}$, denote by uv the concatenation of the word u with the word v (by convention we set, for any $u \in \mathbb{U}$, $\emptyset u = u$). If $v \neq \emptyset$, we say that uv is a descendant of u and u is an ancestor of uv. Moreover $u0$ (resp. $u1$) is called left (resp. right) child of u.

A *complete binary tree* T is a finite subset of $\mathbb{U}$ such that

$$\begin{cases} \emptyset \in T \\ \text{ if } uv \in T \text{ then } u \in T\,, \\ u1 \in T \Leftrightarrow u0 \in T\,. \end{cases}$$

The elements of T are called *nodes*, and $\emptyset$ is called the *root* ; $|u|$, the number of letters in u, is the *depth* of u (with $|\emptyset| = 0$). Write **BinTree** for the set of complete binary trees.

A tree $T \in$ **BinTree** can be described by giving the set ∂T of its *leaves*, that is, the nodes that are in T but with no descendants in T. The nodes of $T \backslash \partial T$ are called *internal* nodes.

We now introduce labeled binary search trees (LBST), that are widely used to store totally ordered data (the monograph of Mahmoud [31] gives an overview of the state of the art).

Let A be a totally ordered set of elements named keys and for $n \geq 1$, let $(x_1, ..., x_n)$ be picked up without replacement from A. The LBST built from these data is the complete binary tree in which each internal node is associated with a key belonging to $(x_1, ..., x_n)$ in the following way: the first key x_1 is assigned to the root. The next key x_2 is assigned to the left child of the root if it is smaller than x_1, or it is assigned to the right child of the root if it is larger than x_1. We proceed further inserting key by key recursively. After the n first insertion, one has a labeled binary tree in which n nodes own a label: these nodes are considered as internal nodes. One adds $n+1$ (unlabeled) leaves to this structure in order to get a labeled complete binary tree with n internal nodes.

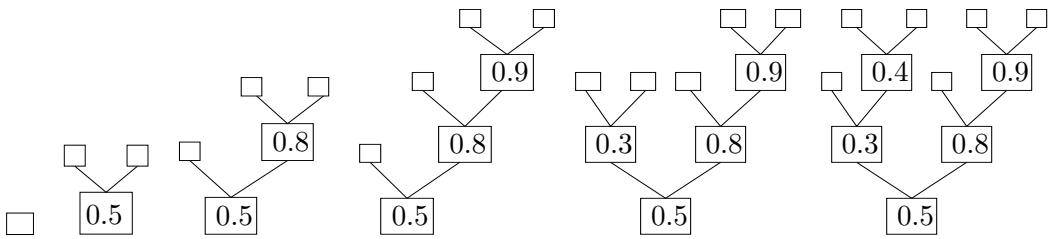

Figure 1: BST built with the sequence of data 0.5, 0.8, 0.9, 0.3, 0.4 (empty squares are leaves).

To study the shape of these trees for large n, it is classical to introduce a random model. One usually assumes that the successively inserted data $(x_i)_{i\geq 1}$ are i.i.d. random variables with a continuous distribution F. Under this model, let us call the LBST $L_n^{(F)}$; it is a random variable taking values in the set of complete binary trees in which each internal node has a label in [0,1]. The sequence $(L_n^{(F)}, n \geq 0)$ is a Markov chain.

We are, in fact, mainly interested in the underlying tree $\mathcal{T}_n^{(F)}$ of $L_n^{(F)}$, i.e. the tree that has the same arborescent structure of $L_n^{(F)}$, but that has no label. We set

$$\big(\mathcal{T}_n^{(F)}, n \geq 0\big) := \big(\,\mathrm{UNDER}(L_n^{(F)}), n \geq 0\big);$$

by construction $\mathcal{T}_n^{(F)}$ is a complete binary tree.

For every $n \geq 1$, the string $x_1, .., x_n$ induces a.s. a permutation σ_n such that $x_{\sigma_n(1)} < x_{\sigma_n(2)} < \cdots < x_{\sigma_n(n)}$. Since the x_i are exchangeable, σ_n is uniformly distributed on the set $\mathcal{S}_n$ of permutations of $\{1, .., n\}$. Since this claim is not sensitive to F we will assume, for the sake of simplicity, that F is the uniform distribution on $[0, 1]$, and we write from now L_n instead of $L_n^{(F)}$ and $\mathcal{T}_n$ instead of $\mathcal{T}_n^{(F)}$. This is the so-called random permutation model. Again by exchangeability, σ_n is independent of the vector $(x_{\sigma_n(1)}, x_{\sigma_n(2)}, \ldots, x_{\sigma_n(n)})$ and we have

$$\begin{aligned} P\big(x_{n+1} \in (x_{\sigma_n(j)}, x_{\sigma_n(j+1)}) \,|\, \sigma_n\big) &= P\big(x_{n+1} \in (x_{\sigma_n(j)}, x_{\sigma_n(j+1)})\big) \\ &= P(\sigma_{n+1}(j+1) = n+1) = (n+1)^{-1} \end{aligned}$$

for every $j \in \{0, 1, .., n\}$, where $x_{\sigma_n(0)} := 0$ and $x_{\sigma_n(n+1)} := 1$. This relation ensures the consistency of the sequence $(\sigma_n)_n$.

One can also express this property with the help of the sequential ranks of the permutation: the random variables $R_k = \sum_{j=1}^{k} \mathbb{1}_{x_j \leq x_k}, k \geq 1$ are independent and R_k is uniform on $\{1, \ldots, k\}$ (see for instance Mahmoud [31], section 2.3), so that $P(R_{n+1} = j+1 \,|\, R_1, .., R_n) = (n+1)^{-1}$.

In terms of binary search tree, this means that the insertion of the $n+1$st key in the tree with n internal nodes is uniform among its $n+1$ leaves. In other words, in the random permutation model, the sequence $(\mathcal{T}_n)_{n\geq 0}$ is a Markov chain on **BinTree** defined by $\mathcal{T}_0 = \{\emptyset\}$ and

$$\begin{aligned} \mathcal{T}_{n+1} &= \mathcal{T}_n \cup \{D_n 0, D_n 1\}\,, \\ P(D_n = u \,|\, \mathcal{T}_n) &= (n+1)^{-1}, \qquad u \in \partial\mathcal{T}_n\,; \end{aligned} \tag{1}$$

the leaf D_n of $\mathcal{T}_n$ is the random node where the $n+1$-st key is inserted, its level is d_n.

The difference of the rule evolutions of L_n (that depends deeply on the values $x_1, \ldots, x_n$ already inserted) and $\mathcal{T}_n$ (that depends of nothing) is similar to Markov chain in random environment (L_n is the quenched Markov chain and $\mathcal{T}_n$ the annealed one).

This Markov chain model is a particular case ($\alpha = 1$) of the diffusion-limited aggregation (DLA) on a binary tree, where a constant α is given and the growing of the tree is random with probability of insertion at a leaf u proportional to $|u|^{-\alpha}$ (Aldous-Shields [1], Barlow-Pemantle-Perkins [6]).

Here are few known results about the evolution of BST. First, the saturation level h_n and the height H_n,

$$h_n = \min\{|u| : u \in \partial\mathcal{T}_n\} \quad , \quad H_n = \max\{|u| : u \in \partial\mathcal{T}_n\} \tag{2}$$

grow logarithmically (see for instance Devroye [17])

$$\text{a.s.} \quad \lim_{n\to\infty} \frac{h_n}{\log n} = c' = 0.3733... \qquad \lim_{n\to\infty} \frac{H_n}{\log n} = c = 4.31107...\,; \tag{3}$$

the constants c' and c are the two solutions of the equation $\eta_2(x) = 1$ where

$$\eta_\lambda(x) := x \log \frac{x}{\lambda} - x + \lambda, \qquad x \geq 0\,, \tag{4}$$

is the Cramer transform of the Poisson distribution of parameter λ. Function η_2 reaches its minimum at $x = 2$. It corresponds to the rate of propagation of the depth of insertion: $\frac{d_n}{2\log n} \xrightarrow{P} 1$. More precise asymptotics for H_n can be found in [19], [36], [37], [28].

Detailed information on $\mathcal{T}_n$ is provided by the whole profile

$$U_k(n) := \#\{u \in \partial\mathcal{T}_n, |u| = k\} \quad , \quad k \geq 1\,, \tag{5}$$

that counts the number of leaves of $\mathcal{T}_n$ at each level. Notice that $U_k(n) = 0$ for $k > H_n$ and for $k < h_n$. To get asymptotic results, it is rather natural to encode the profile by the so-called polynomial level $\sum_k U_k(n) z^k$, whose degree is H_n. Jabbour [15, 25] proved a remarkable martingale property for these random polynomials. More precisely, for $z \notin \frac{1}{2}\mathbb{Z}^- = \{0, -1/2, -1, -3/2, \cdots\}$ and $n \geq 0$, let

$$\mathcal{M}_n(z) := \frac{1}{C_n(z)} \sum_{k\geq 0} U_k(n) z^k = \frac{1}{C_n(z)} \sum_{u\in\partial\mathcal{T}_n} z^{|u|}\,, \tag{6}$$

where $C_0(z) = 1$ and for $n \geq 1$,

$$C_n(z) := \prod_{k=0}^{n-1} \frac{k+2z}{k+1} = (-1)^n \binom{-2z}{n}\,, \tag{7}$$

and let $\mathcal{F}_{(n)}$ be the σ-field generated by all the events $\{u \in \mathcal{T}_j\}_{j \leq n, u \in \mathbb{U}}$. Then $(\mathcal{M}_n(z), \mathcal{F}_{(n)})_n$ is a martingale to which, for the sake of simplicity, we refer from now as the BST martingale. If $z > 0$, this positive martingale is a.s. convergent; the limit $\mathcal{M}_\infty(z)$ is positive a.s. if $z \in (z_c^-, z_c^+)$, with

$$z_c^- = c'/2 = 0.186..., \quad z_c^+ = c/2 = 2.155... \tag{8}$$

and $\mathcal{M}_\infty(z) = 0$ for $z \notin [z_c^-, z_c^+]$ (Jabbour [25]). This martingale is also the main tool to prove that, properly rescaled around $2 \log n$, the profile has a Gaussian limiting shape (see Theorem 1 in [15]).

2.2 Fragmentation, Yule tree process and embedding

The idea of embedding discrete models (such as urn models) in continuous time branching processes goes back at least to Athreya-Karlin [4]. It is described in Athreya and Ney ([5], section 9) and it has been recently revisited by Janson [26]. For the BST, various embeddings are mentioned in Devroye [17], in particular those due to Pittel [35], and Biggins [12, 13]. Here, we work with a variant of the Yule process, taking into account the tree (or "genealogical") structure.

First, let us define a fragmentation process $(F(t))_{t \geq 0}$ of the interval $(0, 1)$ as follows: we set $I_\emptyset = (0, 1)$ and for $u = u_1 u_2 ... u_k \in \mathbb{U}$, set I_u the interval

$$I_u = \Big(\sum_{j=1}^{k} u_j 2^{-j}, 2^{-k} + \sum_{j=1}^{k} u_j 2^{-j} \Big).$$

Hence, each element u of $\mathbb{U}$ encodes a subinterval I_u of $(0, 1)$ with dyadic extremities.

We set $F(0) = I_\emptyset = (0, 1)$. An exponential $\tau_1 \sim \mathbf{Exp}(1)$ random variable is associated with $F(0)$. At time τ_1, the process $F.$ jumps, the interval $(0, 1)$ splits in the middle into two parts and $F(\tau_1) = ((0, 1/2), (1/2, 1)) = (I_0, I_1)$. After each jump time τ, the fragments of $F(\tau)$ behave independently of each other. Each fragment I_u splits after a $\mathbf{Exp}(1)$-distributed random time into two fragments: I_{u0} and I_{u1}. Owing to the lack of memory of the exponential distribution, when n fragments are present, each of them will split first equally likely.

We define now the Yule tree process as an encoding of the fragmentation process. The idea is to interpret the two fragments I_{u0} and I_{u1} issued from I_u as its two children: I_{u0} is considered as the left fragment and I_{u1} the right one; like this, we obtain a binary tree structure (see Fig. 2). An interval with length 2^{-k} corresponds to a leaf at depth k in the corresponding tree structure; the size of fragment I_u is $2^{-|u|}$. More formally, we define the tree $\mathbb{T}_t$ thanks to its set of leaves

$$\partial \mathbb{T}_t = \{u, I_u \in F(t)\}. \tag{9}$$

We call Yule tree process, the **BinTree**-valued random process $(\mathbb{T}_t)_{t \geq 0}$. Both processes $(\mathbb{T}_t)_{t \geq 0}$ and $(F(t))_{t \geq 0}$ are pure jump Markov processes. Each process $(\mathbb{T}_t)_{t \geq 0}$ and $(F(t))_{t \geq 0}$ can be viewed as an encoding of the other one, using (9) and:

$$F(t) = \{I_u, u \in \partial \mathbb{T}_t\}.$$

The counting process $(N_t)_{t \geq 0}$ that gives the number of leaves in $\mathbb{T}_t$,

$$N_t := \# \partial \mathbb{T}_t, \tag{10}$$

is the classical Yule (or binary fission) process (Athreya-Ney [5]).

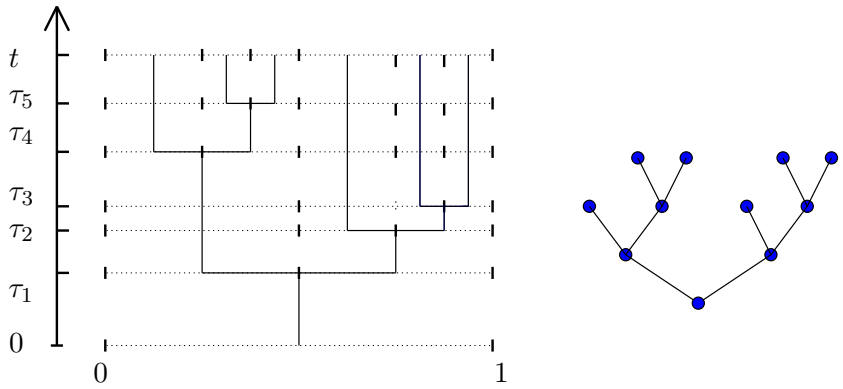

Figure 2: Construction of the Yule tree from the fragmentation.

Let $0 = \tau_0 < \tau_1 < \tau_2 < ...$ be the successive jump times of $\mathbb{T}$. (or of $(F(.))$,

$$\tau_n = \inf\{t : N_t = n + 1\}\,. \tag{11}$$

The following proposition allows us to build the Yule tree process and the BST on the same probability space. This observation was also made in Aldous-Shields [1] section 1, (see also Kingman [27] p.237 and Tavaré [40] p.164 in other contexts).

Lemma 2.1 *a) The jump time intervals $(\tau_n - \tau_{n-1})_n$ are independent and satisfy:*

$$\tau_n - \tau_{n-1} \sim \mathbf{Exp}(n) \text{ for any } n \geq 1, \tag{12}$$

where $\mathbf{Exp}(\lambda)$ is the exponential distribution of parameter λ.

b) The processes $(\tau_n)_{n\geq 1}$ and $\left(\mathbb{T}_{\tau_n}\right)_{n\geq 1}$ are independent.

c) The processes $\left(\mathbb{T}_{\tau_n}\right)_{n\geq 0}$ and $\left(\mathcal{T}_n\right)_{n\geq 0}$ have the same distribution.

Proof: (a) is a consequence of the fact that the minimum of n independent random variables $\mathbf{Exp}(1)$-distributed is $\mathbf{Exp}(n)$-distributed. (b) comes from the independence of jump chain and jump times. Since the initial states and evolution rules of the two Markov chains $\mathbb{T}_{\tau_n}$ and $\mathcal{T}_n$ are the same ones, (c) holds true. ■

Convention: (A unique probability space) From now, we consider that the fragmentation process, the Yule tree process and the BST process are built on the same probability space. Particularly, on this space, we have

$$\left(\mathbb{T}_{\tau_n}\right)_{n\geq 0} = \left(\mathcal{T}_n\right)_{n\geq 0}. \tag{13}$$

We say that the BST process is embedded in the Yule tree process. We define the filtration $(\mathcal{F}_t)_{t\geq 0}$ by $\mathcal{F}_t = \sigma(F(s), s \leq t)$. On the unique probability space, the sigma algebra $\mathcal{F}_{(n)}$ is equal to $\sigma(F(\tau_1), \ldots, F(\tau_n))$.

If we consider the measure valued process $(\rho_t)_{t\geq 0}$ defined by

$$\rho_t = \sum_{u\in\partial\mathbb{T}_t} \delta_{-\log_2|I_u|} = \sum_{u\in\partial\mathbb{T}_t} \delta_{|u|}\,, \tag{14}$$

we obtain a continuous time branching random walk. The set of positions is $\mathbb{N}_0 = \{0, 1, 2, \cdots\}$. Each individual has an $\mathbf{Exp}(1)$ distributed lifetime and does not move. At his death, he disappears and is replaced by two children, whose positions are both their parent's position shifted by 1. The set of individuals alive at time t is $\partial\mathbb{T}_t$ and the position of individual u is simply $|u|$.

This is a particular case of the following general fact: the empirical measure of the logarithm of the size of fragments in homogeneous fragmentations with finite dislocation measures is a branching random walk (this idea goes back to Aldous and Shields [1] Section 7f and 7g).

2.3 Martingales and connection

The classical[1] family of "additive" martingales associated with the Yule process, parameterized by θ in $\mathbb{R}$ (sometimes in $\mathbb{C}$) and indexed by $t \geq 0$, is here given by

$$m(t, \theta) := \sum_{u \in \partial\mathbb{T}_t} \exp(\theta|u| - tL(\theta)),$$

where

$$L(\theta) = 2e^{\theta} - 1 \tag{15}$$

(see [41], [29], and [9] for the fragmentation). For easier use, we set $z = e^{\theta}$ and then consider the family of $(\mathcal{F}_t, t \geq 0)$-martingales

$$M(t, z) := m(t, \log z) = \sum_{u \in \partial\mathbb{T}_t} z^{|u|} e^{t(1-2z)}. \tag{16}$$

In particular $M(t, 1/2) = 1$ and $M(t, 1) = e^{-t} N_t$.

The embedding formula (13) allows to connect the family of BST martingales $(\mathcal{M}_n, \mathcal{F}_{(n)})_n$ to the family of Yule martingales $(M(t, z), \mathcal{F}_t)_t$. If we observe the martingale $(M(., z)$ at the stopping times $(\tau_n)_n$, we can "extract" (Proposition 2.2 below) the space component $\mathcal{M}_n(z)$ and a time component

$$\mathscr{C}_n(z) := e^{\tau_n(1-2z)} C_n(z)\,. \tag{17}$$

Notice that $\big(\mathscr{C}_n(z)\big)_n$ is $\mathcal{F}_{\tau_n}$-adapted.

A classical result (see Athreya-Ney [5] or Devroye [17] 5.4) says that, a.s., $e^{-t} N_t$ converges when $t \to +\infty$, and

$$\xi := \lim_{t \to \infty} e^{-t} N_t \ \sim \mathbf{Exp}(1)\,. \tag{18}$$

Since $\lim_n \tau_n = \infty$ a.s. (see Lemma 2.1 a)) we get from (11) and (18),

$$\text{a.s. } \lim_n n e^{-\tau_n} = \xi\,. \tag{19}$$

Proposition 2.2 (martingale connection) *Let us assume $z \in \mathbb{C} \setminus \frac{1}{2}\mathbb{Z}^-$.*

1) The family $\big(\mathscr{C}_n(z)\big)_{n \geq 0}$ is a martingale with mean 1, and

$$a.s. \lim_n \mathscr{C}_n(z) = \frac{\xi^{2z-1}}{\Gamma(2z)}. \tag{20}$$

Moreover, if $\Re z$, the real part of z, is positive, the convergence is in L^1.

[1] In general $|u|$ is replaced by the position X_u and $L(\theta) = \beta(E \int e^{\theta x} Z(dx) - 1)$ where β is the parameter of lifetime and Z is the offspring point process; here $\beta = 1$ and $Z = 2\delta_1$.

2) The two martingales $(\mathscr{C}_n(z))_{n\geq 0}$ *and* $(\mathcal{M}_n(z))_{n\geq 0}$ *are independent and*

$$M(\tau_n, z) = \mathscr{C}_n(z)\mathcal{M}_n(z)\,. \tag{21}$$

Proof: 1) The martingale property comes from Lemma 2.1 a). The Stirling formula gives the very useful estimate:

$$C_n(z) \sim \frac{n^{2z-1}}{\Gamma(2z)}\,, \tag{22}$$

which yields (20) owing to (19).

2) The second claim comes from (13) and (16), the independence comes from Lemma 2.1 b). ■

Proposition 2.2 allows us to transfer known results about the Yule martingales to BST martingales, thus giving very simple proofs of known results about the BST martingale and also getting much more. In particular, in Theorem 3.3 2), we give the answer to the question asked in [25], about critical values of z, with a straightforward argument.

2.4 Limiting proportions of nodes

Let us study some meaningful random variables arising as a.s limits and playing an important role in the results of Section 3. These variables describe the evolution of relative sizes of subtrees in Yule and BST models.

2.4.1 In the Yule tree

First, we recall a classical identity: let ξ_a and ξ_b be two independent, $\mathbf{Exp}(1)$-distributed random variables. We have

$$\frac{\xi_a}{\xi_a+\xi_b} \sim \mathcal{U}[0,1], \text{ the uniform distribution on } [0,1].$$

For every $u \in \mathbb{U}$, let $\tau^{(u)} = \inf\{t : u \in \mathbb{T}_t\}$ be the time (a.s. finite) at which u appears in the Yule tree, and for $t > 0$, set

$$\mathbb{T}_t^{(u)} = \{v \in \mathbb{U} : uv \in \mathbb{T}_{t+\tau^{(u)}}\}$$

the tree process growing from u. In particular, set

$$N_t^{(u)} = \#\partial\mathbb{T}_t^{(u)}\,.$$

For $t > \tau^{(u)}$, the number of leaves at time t in the subtree issued from node u is $n_t^{(u)} := N_{t-\tau^{(u)}}^{(u)}$. The branching property and (18) give that a.s. for every $u \in \mathbb{U}$

$$\lim_{t\to\infty} e^{-t}N_t^{(u)} = \xi_u \quad , \quad \lim_{t\to\infty} e^{-t}n_t^{(u)} = \xi_u\, e^{-\tau^{(u)}}\,, \tag{23}$$

where ξ_u is distributed as ξ i.e. $\mathbf{Exp}(1)$. Moreover, if u and v are not in the same line of descent, the r.v. ξ_u and ξ_v are independent. Since, for $t > \tau^{(u)}$,

$$n_t^{(u)} = n_t^{(u0)} + n_t^{(u1)} \quad \text{and} \quad \tau^{(u0)} = \tau^{(u1)}, \tag{24}$$

a small computation yields

$$\frac{n_t^{(u0)}}{n_t^{(u)}} \xrightarrow{a.s.} U^{(u0)} := \frac{\xi_{u0}}{\xi_{u0}+\xi_{u1}}, \qquad \frac{n_t^{(u1)}}{n_t^{(u)}} \xrightarrow{a.s.} U^{(u1)} := 1 - U^{(u0)} = \frac{\xi_{u1}}{\xi_{u0}+\xi_{u1}}, \tag{25}$$

which allows to attach a $\mathcal{U}([0,1])$ r.v. to each node of $\mathbb{U}$. In particular we set

$$U := U^{(0)} = \frac{\xi_0}{\xi_0+\xi_1} \tag{26}$$

so that

$$\xi := \xi_\emptyset = e^{-\tau_1}(\xi_0+\xi_1) \quad , \quad \xi_0 = U\xi e^{\tau_1} \quad , \quad \xi_1 = (1-U)\xi e^{\tau_1} \, . \tag{27}$$

If $u0$ and $u1$ are brother nodes, we have $U^{(u1)} + U^{(u0)} = 1$. We claim that if a finite set of nodes $v_1, \ldots, v_k$ does not contain any pair of brothers, the corresponding r.v. $U^{(v_1)}, \ldots, U^{(v_k)}$ are independent. When none of the v_j is an ancestor of another ("stopping line" property) it is a consequence of the branching property. In the general case, it is sufficient to prove that $U^{(u)}$ is independent of $(U^{(v)}, v < u)$. To simplify the reading, let us give the details only for $|u| = 2$, for instance $u = 00$. We have, from (24)

$$U^{(00)} = \frac{\xi_{00}}{\xi_{00}+\xi_{01}} \quad , \quad U^{(0)} = \frac{(\xi_{00}+\xi_{01})e^{-\tau^{(00)}+\tau^{(0)}}}{(\xi_{00}+\xi_{01})e^{-\tau^{(00)}+\tau^{(0)}} + (\xi_{10}+\xi_{11})e^{-\tau^{(10)}+\tau^{(1)}}} \, .$$

Actually, from the branching property, ξ_{00} and ξ_{01} are independent of $\xi_{10}, \xi_{11}, \tau^{(00)}, \tau^{(0)}, \tau^{(10)}, \tau^{(1)}$. Moreover since ξ_{00} and ξ_{01} are independent and $\mathbf{Exp}(1)$ distributed, then $\xi_{00}/(\xi_{00}+\xi_{01})$ and $(\xi_{00}+\xi_{01})$ are independent, which allows to conclude that $U^{(00)}$ and $U^{(0)}$ are independent.

Finally, multiplying along the line of the ancestors of a node u, we get the representation

$$a.s. \lim_{t\to\infty} \frac{n_t^{(u)}}{N_t} = \prod_{v<u} U^{(v)} \, , \tag{28}$$

where the random variables $(U^{(v)})_{v\in\mathbb{U}}$ satisfy the claim.

This is of course related to multiplicative cascade models. See [16].

2.4.2 In the BST

It is straightforward to see that, by embedding, the property of the above subsection holds true for limiting proportions of nodes in the BST, as $n \to \infty$.

This property is also known in the LBST (this gives another proof of this property in the BST). Let us now sketch the argument for LBST.

Assume x_1 fixed. Consider the tree L_n after insertion of the n data $x_2, \ldots, x_{n+1}$. Let $K(n) := \#\{i, i \in [\![2, n+1]\!], x_i \le x_1, \}$ be the number of nodes in the left subtree rooted in $\emptyset$. Since the x_i are i.i.d., $\mathcal{U}[0,1]$, the conditional distribution of $K(n)$ on x_1, is a binomial $B(n, x_1)$. Hence, by the strong law of large numbers,

$$\frac{K(n)}{n} \xrightarrow[n]{a.s.} x_1.$$

Now, the subtree t_0 rooted in $u = 0$ and the subtree t_1 rooted in $u = 1$ are "copies" of L_n. The subtree t_0 is built with the random variables in the list $x_2, x_3, \ldots,$ that are smaller than x_1 (t_1 is

build with the ones that are larger than x_1). In particular, the label x_{t_0} of the root of t_0 is the first value among $x_2, x_3, \ldots$, smaller than x_1. It is easy to check that x_{t_0} is uniform on $[0, x_1]$, therefore it has the following representation: $x_{t_0} = x_1 U$ where U is uniform on $[0,1]$ and does not depend on the value x_1. Hence, the asymptotic proportion of nodes in the subtree t_{00} is $x_1 U$ while it is $x_1(1-U)$ in t_{01} (what happens in the subtree t_1 is totally independent).

This iterative construction of the LBST explains why it enjoys the same property as (28) in the Yule process, and so does the sequence of underlying BST $\mathcal{T}_n$. This is a strong, which means a.s., version of the analogy between BST and branching random walks, first given by Devroye [17].

3 Convergence of martingales

In this section are given the main results about the asymptotic behaviors of the Yule and BST martingales. The martingale connection (Proposition 2.2) allows to express the links between the limits.

3.1 Additive martingales

Theorem 3.1 gives an answer to a natural question asked in [15] about the domain *in the complex plane* where the BST martingale is L^1-convergent and uniformly convergent. Theorem 3.4 gives the optimal L^1 domain on $\mathbb{R}$.

Theorem 3.1 *For $1 < q < 2$, let $\mathcal{V}_q := \{z : \sup_t \mathbb{E}|M(t,z)|^q < \infty\}$. Then $\mathcal{V}_q = \{z : \ f(z,q) > 0\}$ with*

$$f(z,q) := 1 + q(2\Re z - 1) - 2|z|^q \,. \tag{29}$$

If we denote $\mathcal{V} := \cup_{1<q<2} \mathcal{V}_q$, we have :

a) As $t \to \infty$, $\{M(t,z)\}$ converges, a.s. and in L^1, uniformly on every compact C of $\mathcal{V}$.

b) As $n \to \infty$, $\{\mathcal{M}_n(z)\}$ converges, a.s. and in L^1, uniformly on every compact C of $\mathcal{V}$.

Proof: a) is proved in [11] Theorem 6 (see also [9]).

b) We will prove

$$\lim_N \sup_{n \geq N} \mathbb{E} \sup_{z \in C} |\mathcal{M}_n(z) - \mathcal{M}_N(z)| = 0 \,, \tag{30}$$

which implies the uniform L^1 convergence and, since $(\sup_{z \in C} |\mathcal{M}_n(z) - \mathcal{M}_N(z)|)_{n \geq N}$ is a sub-martingale, this will imply also the a.s. uniform convergence[2]. From the martingale connection (Proposition 2.2), we have

$$\mathcal{M}_n(z) - \mathcal{M}_N(z) = \mathbb{E}[M(\tau_n, z) - M(\tau_N, z) | \mathcal{F}_{(n)}]$$

so that taking supremum and expectation we get

$$\mathbb{E} \sup_{z \in C} |\mathcal{M}_n(z) - \mathcal{M}_N(z)| \leq \mathbb{E} \left(\sup_{z \in C} |M(\tau_n, z) - M(\tau_N, z)| \right) \,.$$

[2]For the uniform a.s. convergence, it is possible to give a proof directly from [11]

Taking again the supremum in n we get

$$\sup_{n\geq N} \mathbb{E} \sup_{z\in C} |\mathcal{M}_n(z) - \mathcal{M}_N(z)| \leq \mathbb{E} \sup_{n\geq N} \left(\sup_{z\in C} |M(\tau_n, z) - M(\tau_N, z)| \right) \leq \mathbb{E}\Delta_n \,, \tag{31}$$

where we have set $\Delta_n := \sup_{T\geq\tau_n} (\sup_{z\in C} |M(T,z) - M(\tau_n, z)|)$. Since $M(t,z)$ converges a.s. uniformly, we have a.s. $\lim_n \Delta_n = 0$. Moreover, by the triangle inequality $\Delta_n \leq 2\Delta_0$, and by the proof of Proposition 1 in [9], Δ_0 is integrable. The dominated convergence theorem gives $\lim_n \mathbb{E}\Delta_n = 0$ and (30) holds, which ends the proof of Theorem 3.1. ■

Remark 3.2 *As usual the L^1 convergence in a) of the above theorem comes from a L^q bound (for some $1 < q \leq 2$); more precisely, following the steps in [8] section 2.4, the quantity*

$$\beta_t(\lambda) := (M(t,z) - 1)\, e^{t(2z-1)}$$

satisfies

$$E \mid \beta_t(z) \mid^q \leq e^{tq(2\Re z-1)} \int_0^t \exp\left(-sf(z,q)\right) ds \qquad \text{for } 1 < q \leq 2. \tag{32}$$

Theorem 3.3 *Let us assume $z \in (z_c^-, z_c^+)$.*

1) We have the **limit martingale connection** *:*

$$a.s. \quad M(\infty, z) = \frac{\xi^{2z-1}}{\Gamma(2z)}\, \mathcal{M}_\infty(z)\,, \tag{33}$$

where the exponential variable ξ is defined in (18).

2) We have the following two splitting formulas:

a) for the Yule process,

$$M(\infty, z) = ze^{(1-2z)\tau_1} \left(M_0(\infty, z) + M_1(\infty, z)\right) \quad a.s. \tag{34}$$

where $M_0(\infty, z)$ and $M_1(\infty, z)$ are independent, distributed as $M(\infty, z)$ and independent of τ_1.

b) for the BST,

$$\mathcal{M}_\infty(z) = z\left(U^{2z-1}\mathcal{M}_{\infty,(0)}(z) + (1-U)^{2z-1}\mathcal{M}_{\infty,(1)}(z)\right) \tag{35}$$

where $U \sim \mathcal{U}([0,1])$ is defined in (26), $\mathcal{M}_{\infty,(0)}(z), \mathcal{M}_{\infty,(1)}(z)$ are independent (and independent of U) and distributed as $\mathcal{M}_\infty(z)$.

Proof: 1) is a consequence of (20) and the martingale connection (21).

2) a) For $t > \tau_1$ we have the decomposition

$$M(t,z) = ze^{(1-2z)\tau_1} \left[M^{(0)}(t-\tau_1, z) + M^{(1)}(t-\tau_1, z)\right] \tag{36}$$

where for $i = 0, 1$

$$M^{(i)}(s,z) = \sum_{u\in\partial\mathbb{T}_s^{(i)}} z^{|u|} e^{s(1-2z)}\,,$$

and $\mathbb{T}^{(i)}$ is defined in Section 2.4.

b) Take $t = \tau_n$ in (36), condition on the first splitting time τ_1, apply the branching property, let $n \to \infty$ and apply the limit martingale connection (33) to get

$$\frac{\xi^{2z-1}}{\Gamma(2z)}\mathcal{M}_\infty(z) = ze^{(1-2z)\tau_1}\left(\frac{\xi_0^{2z-1}}{\Gamma(2z)}\mathcal{M}_{\infty,(0)}(z) + \frac{\xi_1^{2z-1}}{\Gamma(2z)}\mathcal{M}_{\infty,(1)}(z)\right) \tag{37}$$

where ξ_0 and ξ_1 come from section 2.4, which yields b) with the help of (27). ■

The following theorem gives the behavior in the remaining cases

Theorem 3.4 *For $z \in (0,\infty) \setminus (z_c^-, z_c^+)$, then a.s. $\lim_t M(t,z) = 0$ and $\lim_n \mathcal{M}_n(z) = 0$.*

Proof: The continuous time result is in [11] (see also [9]); it remains to use again the martingale connection (21). ■

3.2 Derivative martingales

From the above section, we deduce that the derivatives

$$M'(t,z) := \frac{d}{dz}M(t,z), \quad \mathcal{M}_n'(z) := \frac{d}{dz}\mathcal{M}_n(z) \tag{38}$$

are martingales which are no longer positive. They are called the derivative martingales. Their behaviors are ruled by the following theorem.

Theorem 3.5 *1) For $z \in (z_c^-, z_c^+)$, the martingales $(M'(t,z), t \geq 0)$ and $(\mathcal{M}_n'(z), n \geq 0)$ are convergent a.s.. Let us call $M'(\infty,z)$ and $\mathcal{M}_\infty'(z)$ their limits.*

a) We have the (derivative martingale) connection:

$$M'(\infty,z) = \frac{\xi^{2z-1}}{\Gamma(2z)}\left(\mathcal{M}_\infty'(z) + 2\left(\log\xi - \frac{\Gamma'(2z)}{\Gamma(2z)}\right)\mathcal{M}_\infty(z)\right) \qquad a.s. \tag{39}$$

where $\xi \sim \mathbf{Exp}(1)$ is defined in (18) and is independent of $\mathcal{M}_\infty(z)$ and $\mathcal{M}_\infty'(z)$.

b) We have the splitting formula:

$$\begin{aligned}\mathcal{M}_\infty'(z) &= zU^{2z-1}\mathcal{M}_{\infty,(0)}'(z) + z(1-U)^{2z-1}\mathcal{M}_{\infty,(1)}'(z)\\ &+ 2z\left(U^{2z-1}\log U\right)\mathcal{M}_{\infty,(0)}(z) + 2z\left((1-U)^{2z-1}\log(1-U)\right)\mathcal{M}_{\infty,(1)}(z)\\ &+ z^{-1}\mathcal{M}_\infty(z)\end{aligned} \tag{40}$$

where $U \sim \mathcal{U}([0,1])$ is defined in (26), and the r.v. $\mathcal{M}_{\infty,(0)}'(z)$ and $\mathcal{M}_{\infty,(1)}'(z)$ are independent (and independent of U) and distributed as $\mathcal{M}_\infty'(z)$.

2) a) The martingales $(M'(t,z_c^-), t \geq 0)$ and $(\mathcal{M}_n'(z_c^-), n \geq 0)$ (resp. $(M'(t,z_c^+), t \geq 0)$ and $(\mathcal{M}_n'(z_c^+), n \geq 0)$) are convergent a.s.. Their limits denoted by $M'(\infty,z_c^-)$ and $\mathcal{M}_\infty'(z_c^-)$ (resp. $M'(\infty,z_c^+)$ and $\mathcal{M}_\infty'(z_c^+)$) are positive (resp. negative) and satisfy

$$\mathbb{E}\big(M'(\infty,z_c^-)\big) = \mathbb{E}(\mathcal{M}_\infty(z_c^-)) = +\infty, \tag{41}$$

$$\mathbb{E}\big(M'(\infty,z_c^+)\big) = \mathbb{E}(\mathcal{M}_\infty(z_c^+)) = -\infty. \tag{42}$$

b) $M'(\infty, z_c^{\pm})$ and $\mathcal{M}'_\infty(z_c^{\pm})$ satisfy equations similar to (33), (34) and (35):

$$M'(\infty, z_c^{\pm}) = \frac{\xi^{2z_c^{\pm}-1}}{\Gamma(2z_c^{\pm})}\, \mathcal{M}'_\infty(z_c^{\pm}) \tag{43}$$

$$M'(\infty, z_c^{\pm}) = z_c^{\pm} e^{(1-2z_c^{\pm})\tau_1} \left(M'_0(\infty, z_c^{\pm}) + M'_1(\infty, z_c^{\pm})\right) \tag{44}$$

$$\mathcal{M}'_\infty(z_c^{\pm}) = z_c^{\pm}\Big(U^{2z_c^{\pm}-1}\mathcal{M}'_{\infty,(0)}(z_c^{\pm}) + (1-U)^{2z_c^{\pm}-1}\mathcal{M}'_{\infty,(1)}(z_c^{\pm})\Big) \quad a.s.. \tag{45}$$

Proof: 1) For $z \in (z_c^-, z_c^+)$ the a.s. convergence of $M'(t,z)$ is a consequence of the uniform convergence of $M(t,z)$ (by Theorem 3.1) and analyticity. Taking derivatives in the martingale connection (21) gives

$$M'(\tau_n, z) = \left[\frac{C'_n(z)}{C_n(z)} - 2\tau_n\right] \mathscr{C}_n(z)\mathcal{M}_n(z) + \mathscr{C}_n(z)\mathcal{M}'_n(z)\,. \tag{46}$$

Using (19) again and

$$\frac{C'_n(z)}{C_n(z)} = \sum_{j=0}^{n-1} \frac{2}{j+2z} \quad , \quad \frac{\Gamma'(x)}{\Gamma(x)} = \lim_n \Big(\log n - \sum_{j=0}^{n-1} \frac{1}{x+j}\Big)\,,$$

we get

$$\text{a.s. } \lim_n \left[\frac{C'_n(z)}{C_n(z)} - 2\tau_n\right] = 2\left[-\frac{\Gamma'(2z)}{\Gamma(2z)} + \log\xi\right]. \tag{47}$$

We conclude that $\mathcal{M}'_n(z)$ converges and that $\mathcal{M}'_\infty(z)$ satisfies (39) which proves a).

To prove b), we differentiate (36) with respect to z

$$M'(t,z) = (z^{-1} - 2\tau_1)M(t,z) + ze^{(1-2z)\tau_1}\left[M^{(0)'}(t-\tau_1, z) + M^{(1)'}(t-\tau_1, z)\right],$$

and we use the same technique as above: take $t = \tau_n$, let $n \to \infty$, apply (39) and its analogs with $(M'^{(i)}, \mathcal{M}^{(i)}, \mathcal{M}'^{(i)}, \xi_i)_{i=0,1}$ instead of $(M', \mathcal{M}, \mathcal{M}', \xi)$, and use (27).

2) For $z = z_c^{\pm}$, the a.s. convergence of the martingales $M'(t,z)$ and the signs of the limits are proved in [9], and so is the relation

$$\mathbb{E}M'(\infty, z_c^-) = -\mathbb{E}M'(\infty, z_c^+) = \infty\,.$$

Relation (43) is a consequence of (46) and (47), since $\mathcal{M}_\infty(z_c^{\pm}) = 0$.

Formula (45) of 2) is straightforward from (40) since $\mathcal{M}_\infty(z_c^{\pm}) = 0$. Formula (43) is (39) for $z = z_c^{\pm}$. ■

An easy but interesting consequence of (40) is obtained in the following corollary, just taking $z = 1$ in (39) and (40) (remember that $\mathcal{M}_n(1) \equiv 1$). The distributional (weaker) version of (49) below is the subject of a broad literature (see for instance Fill, Janson, Devroye, Neininger, Rösler, Rüschendorf [22, 23, 18, 32, 39, 38]) and some properties of the distribution of $\mathcal{M}'_\infty(1)$ remain unknown.

Corollary 3.6 *We have*

$$M'(\infty,1) = \xi\left(\mathcal{M}'_\infty(1) + 2\left(\log\xi + \gamma - 1\right)\right) \quad a.s.\,, \tag{48}$$

where γ is the Euler constant, and $\mathcal{M}'_\infty(1)$ satisfies the a.s. version of the **Quicksort** *equation:*

$$\mathcal{M}'_\infty(1) = U\mathcal{M}'_{\infty,(0)}(1) + (1-U)\mathcal{M}'_{\infty,(1)}(1) + 2U\log U + 2(1-U)\log(1-U) + 1\,, \tag{49}$$

where as above, $\mathcal{M}'_{\infty,(0)}(1)$ and $\mathcal{M}'_{\infty,(1)}(1)$ are independent (and independent of U), distributed as $\mathcal{M}'_\infty(1)$ and $U \sim \mathcal{U}([0,1])$.

4 Convergence of profiles

4.1 Random measures and profiles

Recall that the profile of the tree $\mathcal{T}_n$ is the sequence

$$U_k(n) = \#\{u \in \partial\mathcal{T}_n : |u| = k\}, \quad k \geq 1\,,$$

and that, according to (3), for every $\epsilon > 0$, there exists a.s. n_0 such that for $n \geq n_0$,

$$U_k(n) = 0 \quad \text{for} \quad k \notin [(c'-\epsilon)\log n, (c+\epsilon)\log n]\,.$$

It means that the convenient scaling for k is $(\log n)^{-1}$. We are interested in the asymptotic behavior of $U_k(n)$ for $k \cong x\log n$ and x fixed in (c', c). It is well known that

$$\mathbb{E}\big(U_k(n)\big) = \frac{2^k}{n!}S_n^{(k)}$$

where $S_n^{(k)}$ is the Stirling number of the first kind. By analysis of singularities, Hwang ([24]) got an asymptotic estimate; for any $\ell > 0$ as $n \to \infty$ and $k \to \infty$ such that $r = k/\log n \leq \ell$:

$$\mathbb{E}\,U_k(n) = \frac{(2\log n)^k}{k!\,n\,\Gamma(r)}(1+o(1))\,. \tag{50}$$

We deduce easily that for any $\ell > 0$:

$$\mathbb{E}\,U_k(n) = \frac{n^{1-\eta_2(\frac{k}{\log n})}}{\Gamma(\frac{k}{\log n})\sqrt{2\pi k}}(1+o(1))\,, \tag{51}$$

where $o(1)$ is uniform for $k/\log n \leq \ell$ and η_2 was defined in (4).

Jabbour in [25] introduced the random measure counting the levels of leaves in $\mathcal{T}_n$

$$r_n := \sum_k U_k(n)\delta_k = \sum_{u\in\partial\mathcal{T}_n} \delta_{|u|}\,.$$

He proved that for $x \in (2, c)$

$$\text{a.s.} \quad \lim_{n\to\infty} \frac{1}{\log n}\,\log r_n(]x\log n, \infty[) \quad = \quad 1 - \eta_2(x) \tag{52}$$

and that the same result holds for $x \in (c', 2)$, replacing $]x \log n, \infty[$ by $]0, x \log n[$.

At the level of random variables, Jabbour & al. proved in [15] that

$$\text{a.s.} \quad \lim_n \sup_{k/\log n \in [1.2,\ 2.8]} \left(\frac{U_k(n)}{\mathbb{E}U_k(n)} - \mathcal{M}_\infty\big(\frac{k}{2\log n}\big) \right) = 0\,. \tag{53}$$

Since their approach relies on L^2 estimations of $\mathcal{M}_n(z)$ with $z = k/(2\log n)$, they guessed that the range $[1.2,\ 2.8]$ may be extended to $I := (2-2^{1/2}, 2+2^{1/2}) = (0.585...\ , 3.414...)$ which corresponds to the maximal interval in z of L^2 convergence. In the following subsection we extend the validity of the above result to (compact subsets of) the entire interval $(c', c) = (0.373...\ , 4.311...)$.

This type of result is very reminiscent of sharp large deviations in branching random walks ([10], [11], [41]). Actually, we use the embedding method and results on the Yule process, as a branching random walk in continuous time. The random measure counting the levels of leaves in the Yule tree is

$$\rho_t = \sum_{u \in \partial \mathbb{T}_t} \delta_{|u|}\,,$$

(recall (14)). With the notations of [41], the exponential rate of growing is ruled by the function

$$x \mapsto L^\star(x) := \sup_\theta \theta x - L(\theta) = \eta_2(x) - 1\,,$$

where the function L is defined in (15). This allows to define three areas:
– for $x \in]c', c[$, $\eta_2(x) < 1$, so there are in mean about $e^{(1-\eta_2(x))t}$ leaves at level $\simeq xt$. Call this interval $]c', c[$ "supercritical area".
– for $x \in [0, c'[\cup]c, \infty[$, $\eta_2(x) > 1$, so there are in mean about $e^{(1-\eta_2(x))t}$ leaves at level $\simeq xt$. Call this set "subcritical area".
– call the set $\{c', c\}$ "critical area" .

More precisely (Theorem 1' p. 909 [41]), for x in the supercritical area,

$$\lim_{t\to\infty} \sqrt{t}\ e^{tL^\star(x)} \rho_t([xt]) = \sqrt{\frac{(L^\star)''(x)}{2\pi}}\ M(\infty, x/2) \quad \text{a.s.}\,. \tag{54}$$

It is now tempting to replace t by τ_n and $\rho_t([xt])$ by $\rho_{\tau_n}([x\log n]) = r_n([x\log n])$. To validate this, we need some uniformity in x in (54). In [11], Biggins obtained such a result. However, it was in the non-lattice case, so we give in the next subsection a complete proof.

4.2 Main result

The aim of this section is to prove the following result

Theorem 4.1 *Almost surely, for any compact subset K of (c', c)*

$$\lim_n \sup_{k:(k/\log n)\in K} \Big(\frac{U_k(n)}{\mathbb{E}\big(U_k(n)\big)} - \mathcal{M}_\infty\big(\frac{k}{2\log n}\big) \Big) = 0\,. \tag{55}$$

Proof: The following lemma, whose proof is postponed yields an asymptotic uniform behavior for $\rho_t(k)$.

Lemma 4.2 *Almost surely, for any compact C of (z_c^-, z_c^+),*

$$\lim_{t\to\infty} \sup_{k\geq 1, z\in C} z^k \sqrt{t} e^{t(1-2z)} \big[\rho_t(k) - M(\infty, z) e^{-t} \frac{(2t)^k}{k!}\big] = 0\,. \tag{56}$$

Let C be a compact subset of (z_c^-, z_c^+). From Lemma 4.2, we know that

$$\rho_t(k) = M(\infty, z)e^{-t}\frac{(2t)^k}{k!} + o(1)z^{-k}t^{-1/2}e^{-t(1-2z)}\,.$$

Recall that $o(1)$ is uniform in k and in $z \in C$. If $\mathcal{P}^{(\lambda)}$ stands for the Poisson law with parameter λ, notice that a $\mathcal{P}^{(2t)}$ appears in the previous expression. Using a change of probability from $\mathcal{P}^{(2t)}$ to $\mathcal{P}^{(2tz)}$, we get

$$\rho_t(k) = z^{-k}t^{-1/2}e^{-t(1-2z)}\Big[t^{1/2}M(\infty, z)\mathcal{P}^{(2tz)}(k) + o(1)\Big]\,.$$

Using the local limit theorem [33], we have

$$\lim_{\lambda\to\infty}\sup_k\Big|\sqrt{2\pi\lambda}\;\mathcal{P}^{(\lambda)}(k) - \exp\Big(-\frac{(k-\lambda)^2}{2\lambda}\Big)\Big| = 0\,.$$

Now, we set $\lambda = 2tz$ with $z \in C$ which yields

$$\lim_{t\to\infty}\sup_{z\in C}\sup_k\Big|\sqrt{4\pi tz}\;\mathcal{P}^{(2tz)}(k) - \exp\Big(-\frac{(k-2tz)^2}{4tz}\Big)\Big| = 0.$$

Hence,

$$\rho_t(k) = A_t(k,z)\Big[\exp\Big(-\frac{(k-2tz)^2}{4tz}\Big)M(\infty, z) + \Big((4\pi z)^{1/2} + M(\infty, z)\Big)o(1)\Big], \tag{57}$$

with

$$A_t(k,z) := \frac{e^{-t(1-2z)}}{z^k(4\pi tz)^{1/2}}\,.$$

Remembering that $U_k(n) = \rho_{\tau_n}(k)$, we take $t = \tau_n$ and $z = \dfrac{k}{2\log n}$ in (57). Using (19) again and the estimate (51), we get

$$\frac{A_{\tau_n}(k,z)}{[\mathbb{E}U_k(n)]\;\xi^{1-2z}\Gamma(2z)} = 1 + o(1) \quad , \quad \exp\Big(-\frac{(k-2\tau_n z)^2}{4\tau_n z}\Big) = 1 + o(1)\,.$$

Now we apply the limit martingale connection (33) and notice that

$$\sup_{z\in C}\Big((4\pi z)^{1/2} + M(\infty, z)\Big) < \infty$$

and we conclude

$$U_k(n) = [\mathbb{E}U_k(n)]\mathcal{M}_\infty(z)(1 + o(1))\,,$$

with $z = k/(2\log n)$ and $o(1)$ uniform in $z \in C$. ■

4.3 Proof of Lemma 4.2

We use the following lemma, which is the continuous time version of Lemma 5 in [11]. Its proof can be managed with the same arguments, replacing Lemma 6 there, by Remark 3.2. We omit the details.

Lemma 4.3 *For any $z_0 \in (z_c^-, z_c^+)$ there exists $r > 0$ for which $z_c^- < z_0 - r < z_0 + r < z_c^+$ and such that a.s.*

$$\lim_{t\to\infty} \sup_{z\in[z_0-r,z_0+r]} \int_{-\pi}^{\pi} \sqrt{t} \mid M(t, ze^{i\eta}) - M(\infty, z) \mid e^{-2tz(1-\cos\eta)} d\eta = 0\,. \tag{58}$$

Write

$$M(t,z) = e^{t(1-2z)} \sum_k \rho_t(k) z^k\,,$$

and the Fourier inversion formula yields

$$\rho_t(k) = \frac{e^{-t(1-2z)} z^{-k}}{2\pi} \int_{-\pi}^{\pi} M(t, ze^{i\eta}) e^{-2tz(1-e^{i\eta})} e^{-ik\eta} d\eta$$

and, owing to Lemma 4.3

$$2\pi\rho_t(k) e^{t(1-2z)} z^k \sqrt{t} = M(\infty, z)\sqrt{t} \int_{-\pi}^{\pi} e^{-2zt(1-e^{i\eta})} e^{-ik\eta} d\eta + o(1)$$

with $o(1)$ uniform in k and in z in any compact subset of (z_c^-, z_c^+). Now, from the Cauchy formula we get that

$$\int_{-\pi}^{\pi} e^{-2zt(1-e^{i\eta})} e^{-ik\eta} d\eta = 2\pi e^{-2zt} \frac{(2zt)^k}{k!}\,,$$

yielding (56), which ends the proof. ■

5 Tagged branches and biased trees

We now introduce the tilting or biasing method in the setting of the Yule and BST processes.

This procedure consists in marking at random a special "ray" or branch of the Yule tree. The special ray of the BST is the spine of the marked Yule tree observed in the splitting times. One then applies to this special ray some evolution rule, different from the other branches. The result of this change of probability is that the whole tree owns a different behavior. This method is usual and fruitful in modern developments on branching processes, and also in the study of fragmentation processes ([9, 14, 30]). The introduction of the tilting method in the setting of BST provides new tools to study some characteristics of the BST.

The martingales studied above are the right change of probability to pass from a tilted model to the non-tilted model: they appear as Radon-Nikodym derivatives. The parameter z, present in the martingales $(M(t,z))_{t\geq 0}$ and $(\mathcal{M}_n(z))_{n\geq 0}$, allows to tune the growing of the special ray, changing in a visible way the shape of the (Yule or BST) tree.

5.1 Tilted fragmentation and biased Yule tree

First at all, let us enlarge the probability space of the fragmentation process introduced in Section 2.2. Let us denote by $\mathcal{F}_t$ the σ-algebra of the interval fragmentation process $F(.)$ up to time t and V be a $\mathcal{U}([0,1])$ r.v. independent of the filtration $(\mathcal{F}_t)_{t\geq 0}$.

Since $\mathbb{P}(V \in \{k2^{-j}, 0 \leq k \leq 2^j, j \in \mathbb{N}, k \in \mathbb{N}\}) = 0$, we may define $\mathbb{P}$-a.s. for every t a unique $S(t) \in \mathbb{U}$ such that $I_{S(t)}$ is an interval of $F(t)$ and $V \in I_{S(t)}$. In other words, $S(t)$ is the element of $\mathbb{U}$ encoding the fragment containing V, its depth is $s(t) := |S(t)|$, the length of $I_{S(t)}$ is $2^{-s(t)}$ and

$$\mathbb{P}(S(t) = u \mid \mathcal{F}_t) = 2^{-|u|} \ , \ u \in \partial \mathbb{T}_t \tag{59}$$

(it is equivalent to choose a fragment at random with probability equal to its length, it is the classical size-biasing setting).

Now we build the process $(\widetilde{\mathbb{T}}_t)_{t\geq 0}$ of marked binary Yule trees associated with the pair $(F(.), S(.))$. The only change with Section 2.2 is the role played by the random variable V (missing in Section 2.2). During the construction of the Yule tree, at any given time t, each leaf in $\mathbb{T}_t$ corresponds to an interval in the fragmentation $F(t)$. For every t we mark the leaf $S(t)$ of $\mathbb{T}_t$ that corresponds to the interval $I_{S(t)}$ that contains V. We obtain a marked tree called $(\widetilde{\mathbb{T}}_t)_{t\geq 0}$. Thus, the set of nodes marked during $[0,t]$ are the prefixes of $S(t)$. We call spine the process $S(.)$.

In fact, given $\widetilde{\mathbb{T}}_t$, one can recover $(F(t), S(t))$. Moreover, with the whole process $(\widetilde{\mathbb{T}}_t)_{t\geq 0}$ one can a.s. recover V:

$$V = \bigcap_{t\geq 0} I_{S(t)}.$$

As a consequence of the general theory of homogeneous fragmentations (see Bertoin [7]) or by a direct computation, we see that $(s(t), t \geq 0)$ is an homogeneous Poisson process with parameter 1. In particular, if

$$\mathcal{E}(t,z) := (2z)^{s(t)} e^{t(1-2z)} \tag{60}$$

then $\mathbb{E}\ \mathcal{E}(t,z) = 1$. Conditionally on $\widehat{\mathcal{F}}_r = \mathcal{F}_r \vee \sigma(S(r), s \leq r)$, the restriction of the fragmentation $F(. + r)$ to the interval $I_{S(r)}$ is distributed as a rescaling of $F(.)$ by a factor $2^{-s(r)}$, which entails that $\left(\mathcal{E}(t,z), \widehat{\mathcal{F}}_t\right)_{t\geq 0}$ is a martingale. By the size biasing scheme (59) and the definition (16) we get

$$M(t,z) = \mathbb{E}\left[\mathcal{E}(t,z) \mid \mathcal{F}_t\right]. \tag{61}$$

Hence, the Yule martingale appears to be a projection of the martingale $\mathcal{E}$ (which is a spine-measurable function) on the σ-algebra containing only the underlying binary tree.

Coming back to the discrete time, set $\mathbf{Spine}_n := S(\tau_n)$ and $s_n := |\mathbf{Spine}_n|$. Notice that the underlying unmarked tree $\mathbb{T}_{\tau_n}$ is $\mathcal{F}_{(n)}$-measurable.

Applying (59) at the $(\mathcal{F}_t, t \geq 0)$ stopping time τ_n, we get for every leaf $u \in \partial\mathcal{T}_n$ (and $k \geq 1$) :

$$\begin{aligned} \mathbb{P}(\mathbf{Spine}_n = u \mid \mathcal{F}_{(n)}) &= 2^{-|u|}, \\ \mathbb{P}(s_n = k \mid \mathcal{F}_{(n)}) &= U_k(n) 2^{-|k|} \ . \end{aligned} \tag{62}$$

Thus, for fixed n, to draw at random the marked tree $\tilde{\mathbb{T}}_{\tau_n}$, one may choose at first a binary tree $\mathcal{T}_n$, and then pick the marked leaf according to the conditional distribution (62). Let $\widehat{\mathcal{F}}_{(0)}$ be the trivial σ-algebra, and for $n \geq 1$ let $\widehat{\mathcal{F}}_{(n)}$ be the σ-algebra obtained from $\mathcal{F}_{(n)}$ by adjunction of

$S(\tau_1), ..., S(\tau_n)$. Let us consider $\mathcal{E}_n(z) := \mathbb{E}\Big[\mathcal{E}(\tau_n, z) \mid \widehat{\mathcal{F}}_{(n)}\Big]$ (with $\mathcal{E}_0(z) := 1$). From Lemma 2.1 a) we have $\mathbb{E}(e^{\tau_n(1-2z)}) = C_n(z)^{-1}$ hence

$$\mathcal{E}_n(z) = (2z)^{S_n}\, C_n(z)^{-1}. \tag{63}$$

From the martingale property of $\mathcal{E}(t, z)$ and the definition of $\mathcal{E}_n(z)$ we see that $\big(\mathcal{E}_n(z), \widehat{\mathcal{F}}_{(n)}\big)$ is a martingale. Like in (61), we get easily

$$\mathcal{M}_n(z) = \mathbb{E}\left[\mathcal{E}_n(z) \mid \mathcal{F}_{(n)}\right], \tag{64}$$

so that the martingales $M(t, z)$ and $\mathcal{M}_n(z)$ are obtained from the "exponential martingales" $\mathcal{E}(z, t)$ and $\mathcal{E}_n(z)$ by projection.

Moreover the martingale connection (21) may be seen as the projection on $\widehat{\mathcal{F}}_{(n)}$ of the relation

$$\mathcal{E}(\tau_n, z) = \mathscr{C}_n(z)\mathcal{E}_n(z)\,.$$

Note that one may also obtain $\mathcal{M}_n(z)$ as

$$\mathcal{M}_n(z) = \mathbb{E}(M(\tau_n, z)|\mathcal{F}_{(n)});$$

this is a kind of integration with respect to the time. All these martingales are precisely the main tool to tilt probabilities. In particular we define $\mathbb{P}^{(2z)}$ on $(\widehat{\mathcal{F}}_t, t \geq 0)$ by

$$\mathbb{P}^{(2z)}_{|_{\widehat{\mathcal{F}}_t}} = \mathcal{E}(t, z)\ \mathbb{P}_{|_{\widehat{\mathcal{F}}_t}}\,, \tag{65}$$

By projection on $(\mathcal{F}_t, t \geq 0)$, (65) yields

$$\mathbb{P}^{(2z)}_{|_{\mathcal{F}_t}} = M(t, z)\ \mathbb{P}_{|_{\mathcal{F}_t}}\,. \tag{66}$$

If ${}^d\mathbb{P}$ (resp. ${}^d\mathbb{P}^{(2z)}$) is the restriction of $\mathbb{P}$ (resp. $\mathbb{P}^{(2z)}$) to $\vee_n \widehat{\mathcal{F}}_{(n)}$, the discrete versions of the above relations are

$${}^d\mathbb{P}^{(2z)}_{|_{\widehat{\mathcal{F}}_{(n)}}} = \mathcal{E}_n(z)\ {}^d\mathbb{P}_{|_{\widehat{\mathcal{F}}_{(n)}}}\quad,\quad {}^d\mathbb{P}^{(2z)}_{|_{\mathcal{F}_{(n)}}} = \mathcal{M}_n(z)\ {}^d\mathbb{P}_{|_{\mathcal{F}_{(n)}}}\,. \tag{67}$$

It turns out that $\mathbb{P}^{(2z)}$ can be seen as a probability on marked Yule trees. This is the object of the following subsection.

5.2 A biased Yule tree

Recall the construction of the Yule tree process $(\mathbb{T}_t)_{t\geq 0}$ given in Section 2.2. Each leaf u of the current Yule tree owns a $\mathbf{Exp}(1)$-distributed clock. At its death, u becomes an internal node, and two leaves $u0$ and $u1$ appear (with new $\mathbf{Exp}(1)$, independent of the other ones).

Let us consider now a model of marked binary tree $(\mathbb{T}^\star_t)_{t\geq 0}$ defined as follows.

In $\mathbb{T}^\star_t$ there are now two kinds of nodes: marked and unmarked. We denote by (v, m) the node v if it is marked, and by $(v, \bar{m})$ the node v if it is unmarked. At time 0, $\mathbb{T}^\star_0 = \{(\emptyset, m)\}$.

Each unmarked leaf owns a $\mathbf{Exp}(1)$-distributed clock. The marked leaf owns a $\mathbf{Exp}(2z)$-distributed clock. Now the evolution of the tilted Yule tree is as follows:

• when an unmarked leaf u dies, u becomes an unmarked internal node, and two unmarked leaves $(u0, \bar{m})$ and $(u1, \bar{m})$ appear.

• when the marked leaf u dies, u becomes a marked internal node. Two leaves $u0$ and $u1$ appear. We mark at random $u0$ or $u1$ (equally likely) and let the other one unmarked.
The marked nodes form a branch in the tree. The behavior of the marked branch depends on the value of $2z$. If $2z > 1$ then, the growing of the marked branch is faster than the other ones, when $2z < 1$, the growing of the marked branch is slower. The depth of the marked leaf follows a Poisson process of rate $2z$. Notice that we have already met this Poisson process in the proof of Theorem 4.1.

It turns out that under $\mathbb{P}^{(2z)}$, the process $(\tilde{\mathbb{T}}_t)_{t\geq 0}$ has the same distribution as the process $(\mathbb{T}_t^\star)_{t\geq 0}$ (consider the spine as the marked leaf). For the underlying branching random walk this construction is classical [3],[14],[20],.... For the fragmentation it can be found in [9].

Remark Denote by $S^\star(t)$ the marked leaf in $\mathbb{T}_t^\star$, and consider

$$V^\star = \bigcap_{t\geq 0} I_{S^\star(t)}.$$

By symmetry of the splittings, the random variable $V^\star$ is $\mathcal{U}[0,1]$-distributed and independent of the process of the underlying unmarked trees derived from $(\mathbb{T}_t^\star)_{t\geq 0}$. Note $F^\star(.)$ the fragmentation process that is associated with the unmarked process derived from $\mathbb{T}_t^\star$.

The a.s. bijection between $(F(.), S(.))$ and $(F(.), V)$ (under $\mathbb{P}^{(1)}$) explained in the beginning of Section 5.1 is also valid between $(F^\star(.), S^\star(.))$ and $(F^\star(.), V^\star)$. Hence, the law on marked fragmentation $(F(.), S(.))$ under $\mathbb{P}^{(2z)}$ (defined by (65)) is the law of $(F^\star(.), S^\star(.))$. It follows that, under $\mathbb{P}^{(2z)}$ one may also build the spine by choosing at first a uniform random variable V and follow the fragment containing V. This is not true in general when using the tilting method. Usually, at each splitting of the marked fragment M, one has to choose the new marked fragment among the children of M, according to a rule depending on the size of these fragments. It cannot be summed up by the drawing of a random variable V, once for all as in our case, where sizes are equal.

According to the representation by $(\mathbb{T}_t^\star)_{t\geq 0}$, the Yule tree owns a natural decomposition according to the marked branch. Let u be a node of the marked branch. One of the nodes $u0$ or $u1$ does not belong to this marked branch. Assume that it is $u0$. Then, (up to a change of the time origin),
• the subtree rooted in $u0$ is a copy of the untilted Yule tree;
• the subtree rooted in $u1$ is a copy of the tilted Yule tree.
We can also see this process as a branching process with immigration, as presented in [40] (see also [34] chap. 10 and [21]).

5.3 A biased BST model

The tilted Yule tree can also been stopped at time τ_n of the creation time of the nth internal node. Let $\widetilde{\mathcal{T}}_n$ be the obtained marked binary search tree. The discrete evolution is as follows:
$\widetilde{\mathcal{T}}_n$ is a complete binary tree with $2n+1$ nodes, in which one leaf is marked and the n other ones are unmarked. Knowing $\widetilde{\mathcal{T}}_n$, the marked tree $\widetilde{\mathcal{T}}_{n+1}$ is as follows:
we choose the marked leaf with probability $2z/(n+2z)$ and each unmarked one with probability $1/(n+2z)$.
• If the chosen leaf v is unmarked, then v becomes an unmarked internal node and two unmarked leaves $v0$ and $v1$ are created.
• If the chosen leaf v is marked, v becomes a marked internal node. Two leaves $v0$ and $v1$ appear. One marks at random $v0$ or $v1$ (equally likely) and let the other ones unmarked.

We note $\mathbb{Q}^{(2z)}$ for the law on the marked binary search tree process $(\widetilde{\mathcal{T}}_n)_n$ under this model of evolution.

Once again, the BST can be decomposed along the marked branch. The speed of growing of the marked branch depends on the value of $2z$. One may also interpret the size of the subtrees rooted on the tilted branch as tables in a Chinese restaurant (see Barbour & al. [2], Pitman [34]), and obtain, like this, new explanations of the behavior of the size of the subtrees rooted on the marked branch.

As in the previous subsection, we denote by (v, m) a marked node and $(v, \bar{m})$ an unmarked node. The dynamics we described above yields the following conditional probabilities:

if $(v\bar{m}) \in \partial\widetilde{\mathcal{T}}_n$, then

$$\mathbb{Q}^{(2z)}(\mathbf{Spine}_{n+1} = \mathbf{Spine}_n, \widetilde{\mathcal{T}}_{n+1} = \widetilde{\mathcal{T}}_n \cup \{(v0, \bar{m}), (v1, \bar{m})\} | \widetilde{\mathcal{T}}_n) = \frac{1}{n+2z}$$

If $(v, m) \in \partial\widetilde{\mathcal{T}}_n$, (i.e. $\mathbf{Spine}_n = v$), then

$$\mathbb{Q}^{(2z)}(\mathbf{Spine}_{n+1} = v0, \widetilde{\mathcal{T}}_{n+1} = \widetilde{\mathcal{T}}_n \cup \{(v0, m), (v1, \bar{m})\} | \widetilde{\mathcal{T}}_n) = \frac{1}{2}\frac{2z}{n+2z}$$

similarly,

$$\mathbb{Q}^{(2z)}(\mathbf{Spine}_{n+1} = v1, \widetilde{\mathcal{T}}_{n+1} = \widetilde{\mathcal{T}}_n \cup \{(v0, \bar{m}), (v1, m)\} | \widetilde{\mathcal{T}}_n) = \frac{1}{2}\frac{2z}{n+2z}\,.$$

Summing up, we have for any marked tree $\widetilde{t}_{n+1}$ with $n+1$ nodes that can be obtained from $\widetilde{\mathcal{T}}_n$ by one insertion

$$\mathbb{Q}^{(2z)}(\widetilde{\mathcal{T}}_{n+1} = \widetilde{t}_{n+1} | \widetilde{\mathcal{T}}_n) = \frac{z^{s_{n+1}-s_n}}{n+2z} \tag{68}$$

and

$$\mathbb{Q}^{(1)}(\widetilde{\mathcal{T}}_{n+1} = \widetilde{t}_{n+1} | \widetilde{\mathcal{T}}_n) = \frac{(1/2)^{s_{n+1}-s_n}}{n+1}.$$

Thus, by iterative construction,

$$\frac{\mathbb{Q}^{(2z)}}{\mathbb{Q}^{(1)}}\Big|_{\hat{\mathcal{F}}_n} = \prod_{j=0}^{n-1} \frac{(2z)^{s_{j+1}-s_j}(j+1)}{j+2z} = (2z)^{s_n}\, C_n(z)^{-1} = \mathcal{E}_n(z)\,.$$

Hence, $\mathbb{Q}^{(2z)}$ is absolutely continuous with respect to $\mathbb{Q}^{(1)}$, with the Radon-Nikodym derivative announced in (67). Since $\mathbb{Q}^{(1)}$ and ${}^d\mathbb{P}^{(1)}$ (the non-biased models) are identical, the law of $(\widetilde{\mathcal{T}}_n)_n$ under $\mathbb{Q}^{(2z)}$ is ${}^d\mathbb{P}^{(2z)}$.

One finds an analogous result (in another context) and its proof in Lemma 1 and 2 of [14].

5.4 Spine evolution

Thanks to the previous subsections, it appears that under ${}^d\mathbb{P}^{(2z)}$

$$s_n = 1 + \sum_{1}^{n-1} \epsilon_k \tag{69}$$

where $(\epsilon_k)_{k\geq 1}$ are independent and for every $k \geq 1$, ϵ_k is a Bernoulli random variable with parameter $\frac{2z}{k+2z}$; (we use the notation $\epsilon_k \sim \mathbf{Be}(\frac{2z}{k+2z})$).

Proposition 5.1 *For any parameter $z > 0$,*

1) (strong law)

$$\lim \frac{s_n}{\log n} = 2z\,, \qquad {}^d\mathbb{P}^{(2z)} - a.s.. \tag{70}$$

2) (central limit theorem) The distribution of $\dfrac{s_n - 2z\log n}{\sqrt{2z\log n}}$ *under* ${}^d\mathbb{P}^{(2z)}$ *converges to a standard normal distribution* $\mathcal{N}(0,1)$.

3) (large deviations) The family of distributions of $(s_n,\ n>0)$ *under* ${}^d\mathbb{P}^{(2z)}$ *satisfies the large deviation principle on* $[0,\infty)$ *with speed* $\log n$ *and rate function* η_{2z} *where the function* η_λ *is defined in (4).*

Proof: 1) and 2) are consequences of known results on sums of independent r.v. (see [33]). Notice also that $s_n - \mathbb{E}^{(2z)}(s_n)$ is a martingale.

3) is a consequence of Gärtner-Ellis theorem. ■

Once again, this proposition shows that under the biased model, the BST evolves rather differently that under the usual model. For example, the marked leaf depth is about $2z\log n$. So, for $z > z_c^+$, the marked leaf is higher that the height of the non-biased BST.

5.5 Depth of insertion

In introducing the BST model, we defined the sequence $(D_n, n\geq 0)$ as the successive inserted nodes and $d_n = |D_n|$ (see (1)). In continuous time, we set $\eta(t) = \inf\{s>t:\ \mathbb{T}_s \neq \mathbb{T}_t\}$ for the first time of growing after t, and $D(t) = \mathbb{T}_{\eta(t)} \setminus \mathbb{T}_t$ for the node of the coming insertion.

Let us stress on the difference between the spine processes $(s_n, n\geq 0)$ and $(s(t), t\geq 0)$ and the insertion processes $(d_n, n\geq 0)$ and $(d(t), t\geq 0)$.

The (marginal) distribution of d_n is given in Jabbour [25] (see also Mahmoud [31])

$$\mathbb{E}z^{d_n} = \frac{C_n(z)}{n+1} = \frac{(2z)(2z+1)\cdots(2z+n-1)}{(n+1)!} \tag{71}$$

so that

$$d_n \overset{law}{=} 1 + \sum_1^{n-1} \varepsilon_k\,,$$

where $(\varepsilon_k)_{k\geq 1}$ are independent and for every $k\geq 1$, $\varepsilon_k \sim \mathbf{Be}(\frac{2}{k+2})$.

Proposition 5.2 *The following convergences hold:*
(i) (d_n) *satisfies a law of large numbers:*

$$\frac{d_n}{2\log n} \xrightarrow{P} 1; \tag{72}$$

(ii) it satisfies a central limit theorem:

$$\frac{d_n - 2\log n}{\sqrt{2\log n}} \overset{law}{\Longrightarrow} \mathcal{N}(0,1)\,. \tag{73}$$

(iii) We have

$$a.s.\ \liminf_n \frac{d_n}{2\log n} = \frac{c'}{2} = z^- \quad , \quad a.s.\ \limsup_n \frac{d_n}{2\log n} = \frac{c}{2} = z^+\,. \tag{74}$$

Note that (iii) of course, implies that (ii) is not an almost sure convergence.
Proof: The arguments to prove (i) and (ii) are classical; (iii) is a consequence of (3). ■

For the Yule tree, we did not find the distribution of $d(t)$ in the literature. Let us give the joint distribution of $(N_t, d(t))$ (for t fixed).

Since $\{N_t = n+1\} = \{\tau_n \le t < \tau_{n+1}\}$, we have $\mathbb{E}(z^{d(t)} s^{N_t}) = \sum_0^\infty (\mathbb{E} z^{d_n}) \mathbb{P}(N_t = n+1) s^{n+1}$. Since the distribution of N_t is geometric of parameter e^{-t}, and owing to (71) we get

$$\mathbb{E}(z^{d(t)} s^{N_t}) = \sum_0^\infty \frac{(2z)(2z+1)\cdots(2z+n-1)}{(n+1)!} e^{-t}(1-e^{-t})^n s^{n+1} \tag{75}$$

$$= \frac{\big(1 - s(1-e^{-t})\big)^{1-2z} - 1}{(e^t - 1)(2z-1)} . \tag{76}$$

Taking $s = 1$, we get the marginal of $d(t)$

$$\mathbb{E} z^{d(t)} = \frac{e^{t(2z-1)} - 1}{(e^t - 1)(2z-1)} .$$

Transforming these generating functions into Fourier transforms, it is now easy to conclude that

Proposition 5.3 *As $t \to \infty$,*

$$\left(N_t e^{-t}, \frac{d(t) - 2t}{\sqrt{2t}} \right) \stackrel{law}{\Longrightarrow} (\xi, G)$$

where ξ is defined in (19) and G is $\mathcal{N}(0,1)$ and independent of ξ.

Remark: For the same reasons as in (74), we have

$$\text{a.s. } \liminf_t \frac{d(t)}{2t} = \frac{c'}{2} = z^- \quad , \quad \text{a.s. } \limsup_t \frac{d(t)}{2t} = \frac{c}{2} = z^+ .$$

Under the change of probability $\mathbb{P}^{(2z)}$ (or using Kolmogorov equations) the distribution of N_t is given by:

$$\mathbb{E}^{(2z)} \phi^{N_t} = \mathbb{E}\big[(2z)^{s(t)} e^{t(1-2z)} \phi^{N_t}\big] \tag{77}$$

$$= e^{t(1-2z)} \big[\phi + \sum_0^\infty (\mathbb{E}(2z)^{s_n}) \mathbb{P}(N_t = n+1) \phi^{n+1}\big] \tag{78}$$

$$= e^{t(1-2z)} \big[\phi e^{-t} + e^{-t} \sum_{n=2}^\infty (1-e^{-t})^{n-1} \prod_0^{n-2} \frac{j+2z}{j+1} \phi^n\big] \tag{79}$$

$$= \phi \left[\frac{e^{-t}}{1 - \phi(1-e^{-t})} \right]^{2z} ; \tag{80}$$

where ϕ is any real in $[0,1]$. Hence, under $\mathbb{P}^{(2z)}$, the r.v. $N_t - 1$ is a negative binomial of order $2z$ and parameter e^{-t}. As $t \to \infty$, the $\mathbb{P}^{(2z)}$ distribution of $e^{-t} N_t$ converges to a $\gamma(2z)$-distributed random variable. Actually we have for every z, t, h

$$\mathbb{E}[N_{t+h} | \mathcal{F}_t] = (N_t - 1)\mathbb{E} N_h + \mathbb{E}^{(2z)} N_h \tag{81}$$

$$= e^h (N_t - 1) + 1 + 2z(e^h - 1) . \tag{82}$$

This implies that $(e^{-t}[N_t - 1 + 2z])_{t\geq 0})$ is a $\mathbb{P}^{(2z)}$ martingale. If $2z > 1$, it is positive hence convergent $\mathbb{P}^{(2z)}$ a.s. If $2z < 1$, then $(e^{-t}N_t)_{t\geq 0})$ is a positive supermartingale, hence $\mathbb{P}^{(2z)}$ a.s. convergent. ■

References

[1] D. Aldous and P. Shields. A diffusion limit for a class of randomly-growing binary trees. *Probab. Theory Related Fields*, 79:509–542, 1988.

[2] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach.* EMS Monographs in Mathematics. European Mathematical Society (EMS), Zurich, 2003.

[3] K. Athreya. Change of measures for Markov chains and the L log L theorem for branching processes. *Bernoulli*, 6:323–338, 1999.

[4] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.

[5] K. B. Athreya and P. E. Ney. *Branching processes.* Springer-Verlag, New York, 1972.

[6] M.T. Barlow, R. Pemantle, and E.A. Perkins. Diffusion-limited aggregation on a tree. *Probab. Theory Relat. Fields*, 107:1–60, 1997.

[7] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Related Fields*, 121(3):301–318, 2001.

[8] J. Bertoin. The asymptotic behavior of fragmentation processes. *J. Europ. Math. Soc.*, 5(4):395–416, 2003.

[9] J. Bertoin and A. Rouault. Discretization methods for homogeneous fragmentations. Preprint available at http://front.math.ucdavis.edu/math.PR/0409545, September 2004.

[10] J. D. Biggins. Growth rates in the branching random walk. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 48:17–34, 1979.

[11] J. D. Biggins. Uniform convergence of martingales in the branching random walk. *Ann. Probab.*, 20(1):137–151, 1992.

[12] J. D. Biggins. How fast does a general branching random walk spread? In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 19–39. Springer, New York, 1997.

[13] J. D. Biggins and D. R. Grey. A note on the growth of random trees. *Statist. Probab. Lett.*, 32(4):339–342, 1997.

[14] J.D. Biggins and A.E. Kyprianou. Measure change in multitype branching. *Adv. Appl. Prob.*, 36(2):544–581, 2004.

[15] B. Chauvin, M. Drmota, and J. Jabbour-Hattab. The profile of binary search trees. *Ann. Appl. Prob.*, 11:1042–1062, 2001.

[16] B. Chauvin and A. Rouault. Connecting Yule process, bisection and binary search trees via martingales. To appear in Journal of the Iranian Statistical Society, available at http://www.math.uvsq.fr/∼rouault/rouaultart/bst3.pdf, 2003.

[17] L. Devroye. Branching processes and their applications in the analysis of tree structures and tree algorithms. In M. Habib et al., editor, *Probabilistic Methods for Algorithmic Discrete Mathematics*. Springer, 1998.

[18] L. Devroye, J.A. Fill, and R. Neininger. Perfect simulation from the quicksort limit distribution. *Electronic Communications in Probability*, 5:95–99, 2000.

[19] M. Drmota. Stochastic analysis of tree-like data structures. *Proc. R. Soc. Lond.*, A460(27):271–307, 2004.

[20] J. Englander and A. E. Kyprianou. Local extinction versus local exponential growth for spatial branching processes. *Annals of Probab.*, 32(1A):78–99, 2004.

[21] S. Feng and F.M. Hoppe. Large deviation principles for some random combinatorial structures in population genetics and brownian motion. *The Annals of Prob.*, 8:975–994, 1998.

[22] J.A. Fill and S. Janson. Approximating the limiting quicksort distribution. In *Special Issue of Analysis on Algorithms*, volume 19, pages 376–406, 2001.

[23] J.A. Fill and S. Janson. Quicksort asymptotics. In *Special Issue of Analysis on Algorithms*, volume 44, pages 4–28, 2002.

[24] H.K. Hwang. Asymptotic expansions for the Stirling numbers of the first kind. *J. Combin. Theory Ser. A*, 71(2):343–351, 1995.

[25] J. Jabbour-Hattab. Martingales and large deviations for binary search trees. *Random Structure and Algorithms*, 19:112–127, 2001.

[26] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004.

[27] J.F.C. Kingman. The coalescent process. *Stochastic Process. Appl.*, 13:235–248, 1982.

[28] P.L. Krapivsky and S.T. Majumdar. Travelling waves, front selection, and exact nontrivial exponents in random fragmentation problem. *Phys. Review Letters*, 85(26):5492–5495, 2000.

[29] A. E. Kyprianou. A note on branching Lévy processes. *Stochastic Process. Appl.*, 82(1):1–14, 1999.

[30] R. Lyons. A simple path to Biggins' martingale convergence for the branching random walk. In P. Jagers K.B. Athreya, editor, *Classical and Modern Branching Processes*, volume 84, pages 217–222. IMA Volumes in Mathematics and its Applications, Springer, 1997.

[31] H. Mahmoud. *Evolution of Random Search Trees*. John Wiley, New York, 1992.

[32] R. Neininger and L. Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Annals of App. Probab.*, 14(1):378–418, 2004.

[33] V.V. Petrov. *Sums of independent random variables*. Springer Verlag, 1975.

[34] J. Pitman. Cours de Saint-Flour 2002: Combinatorial Stochastic Processes. available at http://stat-www.berkeley.edu/users/pitman/bibliog.html.

[35] B. Pittel. On growing random binary trees. *J. Math. Anal. Appl.*, 103(2):461–480, 1984.

[36] B. Reed. The height of a random binary search tree. *Journal of the ACM*, 50(3):306–332, 2003.

[37] J.M. Robson. Constant bounds on the moments of the height of binary search trees. *Theor. Computer Sci.*, 276:435–444, 2002.

[38] U. Rösler. A limit theorem for "quicksort". *RAIRO, Inform. Théor. Appl.*, 25(1):85–100, 1991.

[39] U. Rösler. On the analysis of stochastic divide and conquer algorithms. *Algorithmica*, 29(1-2):238–261, 2001. Average-case analysis of algorithms (Princeton, NJ, 1998).

[40] S. Tavaré. The birth process with immigration, and the genealogical structure of large populations. *J. Math. Biol.*, 25(2):161–168, 1987.

[41] K. Uchiyama. Spatial growth of a branching process of particles living in R^d. *Ann. Probab.*, 10(4):896–918, 1982.