Almost sure asymptotics for the random binary search tree

Matthew I. Roberts[†]

Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VI
Case courrier 188, 4, Place Jussieu, 75252 Paris, France.
E-mail: *matthew.roberts@upmc.fr* .

We consider a (random permutation model) binary search tree with n nodes and give asymptotics on the $\log \log$ scale for the height H_n and saturation level h_n of the tree as $n \to \infty$, both almost surely and in probability. We then consider the number F_n of particles at level H_n at time n, and show that F_n is unbounded almost surely.

Keywords: Binary search tree, Random permutation model, Yule tree, Quicksort, Branching random walk

1 Introduction and main results

Consider the complete rooted binary tree $\mathbb{T}$. We construct a sequence $\mathbb{T}_n$, $n = 1, 2, \ldots$ of subtrees of $\mathbb{T}$ recursively as follows. $\mathbb{T}_1$ consists only of the root. Given $\mathbb{T}_n$, we choose a leaf u uniformly at random from the set of all leaves of $\mathbb{T}_n$ and add its two children to the tree to create $\mathbb{T}_{n+1}$. Thus $\mathbb{T}_{n+1}$ consists of $\mathbb{T}_n$ and the children $u1$, $u2$ of u, and contains in total $2n + 1$ nodes, including $n + 1$ leaves. We call this sequence of trees $(\mathbb{T}_n)_{n \geq 1}$ the binary search tree.
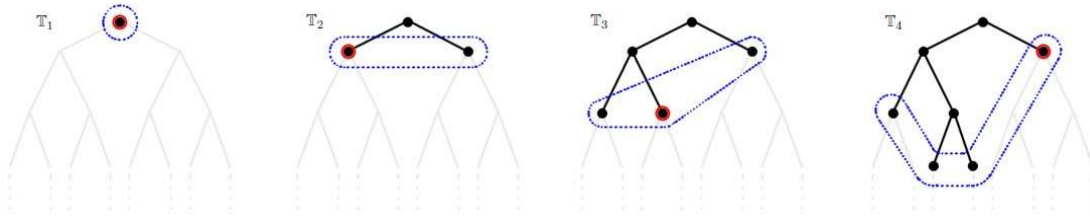

Fig. 1: An example of the beginning of a binary search tree: at each stage, we choose uniformly at random from amongst the available leaves and add the children of the chosen leaf to the tree.

This model has various equivalent descriptions: for example one may construct $\mathbb{T}_n$ by successive insertions into $\mathbb{T}$ of a uniform random permutation of $\{1, \ldots, n\}$. For a more detailed explanation of this and other constructions see Reed [9].

One interesting quantity in this model is the height H_n of the tree $\mathbb{T}_n$ — that is, the greatest generation amongst all nodes of $\mathbb{T}_n$ (where the root is defined to have generation 0); so $H_1 = 0$, $H_2 = 1$, $H_3 = 2$, and H_4 is either 2 (with probability 1/3) or 3 (with probability 2/3). Another is the saturation level h_n, defined to be the greatest complete generation of $\mathbb{T}_n$ — that is, the greatest generation k such that all nodes of generation k are present in $\mathbb{T}_n$ (so $h_1 = 0$, $h_2 = 1$, $h_3 = 1$ and h_4 is 1 with probability 2/3 and 2 with probability 1/3). These two quantities, H_n and h_n, have been studied extensively. Pittel [8] showed that there exist constants c and γ such that $H_n/\log n \to c$ and $h_n/\log n \to \gamma$ almost surely, and gave bounds on the values of c and γ. Devroye [4] calculated c exactly by showing that $H_n/\log n \to c$ in probability as $n \to \infty$; and Reed [9] showed that for the same c and another known constant d, $\mathbb{E}[H_n] = c\log n - d\log\log n + O(1)$. Drmota [5] and Reed [9] also showed that $\mathrm{Var}H_n = O(1)$.

Our first aim in this article is to prove the following theorem.

Theorem 1 *Let a be the solution to*

$$2(a-1)e^a + 1 = 0, \quad a > 0$$

and let

$$b := 2ae^a$$

(we get $a \approx 0.76804$ and $b \approx 3.31107$). Then

$$\frac{1}{2} = \liminf_{n\to\infty} \frac{b\log n - aH_n}{\log\log n} < \limsup_{n\to\infty} \frac{b\log n - aH_n}{\log\log n} = \frac{3}{2}$$

almost surely and

$$\frac{b\log n - aH_n}{\log\log n} \xrightarrow{\mathbb{P}} \frac{3}{2} \quad as \quad n \to \infty.$$

Of course, a and b agree with the constants c and d mentioned above in the sense that $c = b/a$ and $d = 3/2a$. By the same methods, we obtain a similar theorem concerning h_n.

Theorem 2 *Let α be the solution to*

$$2(\alpha+1)e^{-\alpha} - 1 = 0, \quad \alpha > 0$$

and let

$$\beta := 2\alpha e^{-\alpha}$$

(we get $\alpha \approx 1.6783$ and $\beta \approx 0.6266$). Then

$$\frac{1}{2} = \liminf_{n\to\infty} \frac{\alpha h_n - \beta\log n}{\log\log n} < \limsup_{n\to\infty} \frac{\alpha h_n - \beta\log n}{\log\log n} = \frac{3}{2}$$

almost surely and

$$\frac{\alpha h_n - \beta\log n}{\log\log n} \xrightarrow{\mathbb{P}} \frac{3}{2} \quad as \quad n \to \infty.$$

This shows in particular that the lower bound given by Pittel [8] is the correct growth rate for the saturation level h_n on the $\log$ scale.

Other aspects of the binary search tree model also give interesting results. The article by Chauvin *et al.* [3], for example, tracks the number of leaves at certain levels of the tree, called the profile of the tree, via convergence theorems for polynomial martingales associated with the system.

We are also interested in how many leaves are present at level H_n of the tree at time n. We call the set of particles at this level the *fringe* of the tree, and call the size of the fringe F_n, so that $F_1 = 1$, $F_2 = 2$, $F_3 = 2$, and F_4 is 2 with probability $2/3$ or 4 with probability $1/3$. Note that the word "fringe" has been used also in a different context by, for example, Drmota *et al.* [6]. Trivially $F_n \in \{2, 4, 6, \ldots\}$ for all $n \geq 2$, and (given that $H_n \to \infty$ almost surely, which is a simple consequence of Theorem 1) $\liminf_{n\to\infty} F_n = 2$ almost surely. We are able to prove the following result.

Theorem 3 *We have*
$$\limsup_{n\to\infty} F_n = \infty$$
almost surely.

Further work on the behaviour of F_n in the limit as $n \to \infty$ is underway.

Our main tool throughout is the relationship between binary search trees and an extremely simple continuous time branching random walk, called the Yule tree. This relationship is well-known — see Aldous & Shields [1] and Chauvin *et al.* [3]. The hard work required for Theorem 1 is then done for us by a remarkable result of Hu & Shi [7]. We introduce the Yule tree model in Section 2 before proving Theorems 1 and 2 in Section 3. Finally we study F_n, and in particular prove Theorem 3, in Section 4.

2   The Yule tree

Consider a branching random walk in continuous time with branching rate 1, starting with one particle at the origin, in which if a particle with position x branches it is replaced by two children with position $x - 1$. That is:

- We begin with one particle at 0;

- All particles act independently;

- Each particle lives for a random amount of time, exponentially distributed with parameter 1;

- Each particle has a position x which does not change throughout its lifetime;

- At its time of death, a particle with position x is replaced by two offspring with position $x - 1$.

We call this process a Yule tree. If a particle v is a descendant of u (that is, there exists $k \in \{1, 2, \ldots\}$ and particles $u = u_1, u_2, \ldots, u_k = v$ such that u_j is an offspring of u_{j-1} for each j), then we write $u \leq v$. Let $N(t)$ be the set of particles alive at time t, and for a particle $u \in N(t)$ define $X_u(t)$ to be the position of u at time t. Let $M(t)$ denote the smallest of these positions at time t, and $S(t)$ the largest — that is,

$$M(t) := \inf\{X_u(t) : u \in N(t)\}$$

and

$$S(t) := \sup\{X_u(t) : u \in N(t)\}.$$

We note that if we look at the Yule tree model only at integer times, then we have a discrete-time branching random walk. On the other hand, we have the following simple relationship between the Yule tree process and the binary search tree process.

Lemma 4 *Let $T_1 = 0$ and for $n \geq 2$ define*

$$T_n := \inf\{t > T_{n-1} : N(t) \neq N(T_{n-1})\}$$

so that the times T_n are the birth times of the branching random walk. Then we may construct the Yule tree process and the binary search tree process on the same probability space, such that

$$-M(T_n) = H_n \quad \forall n \geq 1$$

and

$$-S(T_n) = h_n + 1 \quad \forall n \geq 1$$

almost surely.

Proof: By the memoryless property of the exponential distribution, at any time t the probability that a particular particle $u \in N(t)$ will be the next to branch is exactly $1/\#N(t)$. Thus, if we consider the sequence of genealogical trees produced by the Yule tree process at the times $T_j, j \geq 1$, we have exactly the binary search tree process: particles in $N(t)$ correspond to leaves in the binary search tree. Clearly the position of a particle in the Yule tree process is -1 times its height in the genealogical tree, so we may build the Yule tree process and binary search tree process on the same probability space and then $-M(T_n) = H_n$ and $-S(T_n) = h_n + 1$ for all $n \geq 1$ (almost surely). □

We would like to study $(H_n, n \geq 1)$ via knowledge of $(M(t), t \geq 0)$, and similarly for h_n and $S(t)$, and hence it will be important to have control over the times T_n. It is well-known that T_n is close to $\log n$. We give a simple martingale proof, as seen in Athreya & Ney [2].

Lemma 5 *There exists an almost surely finite random variable ζ such that*

$$T_n - \log n \to \zeta \quad \text{almost surely as } n \to \infty;$$

and hence for any $\delta > 0$ we may choose $K \in \mathbb{N}$ such that

$$\mathbb{P}\left(\limsup_{n\to\infty} |T_n - \lfloor \log n \rfloor| > K\right) < \delta$$

and

$$\limsup_{n\to\infty} \mathbb{P}(|T_n - \lfloor \log n \rfloor| > K) < \delta.$$

Proof: For each $n \geq 1$, let $V_n := n(T_n - T_{n-1})$. Then the random variables V_n, $n \geq 1$ are independent and exponentially distributed with parameter 1. Define

$$X_n := \sum_{j=1}^{n} \frac{V_j - 1}{j} = T_n - \sum_{j=1}^{n} j^{-1}.$$

Then X_n is clearly a zero-mean martingale; and

$$\mathbb{E}[X_n^2] = \sum_{j=1}^{n} \frac{\text{Var}(V_j)}{j^2} \leq \sum_{j=1}^{\infty} j^{-2} < \infty$$

so by the martingale convergence theorem X_n converges almost surely (and in L^2) to some almost surely finite limit X. But it is well-known that

$$\sum_{j=1}^{n} j^{-1} - \log n$$

converges to some finite, deterministic constant (Euler's constant). This is enough to complete the proof of the first statement in the Lemma, and the consequences are standard. □

We mentioned above that, if we look at the Yule tree only at integer times, we see a discrete-time branching random walk. Since discrete-time branching random walks are more widely studied than their continuous-time counterparts (in particular the theorem that we would like to apply is stated only in discrete-time), it will be helpful to have some information about the branching distribution of the discrete model. This is a standard, well-known calculation, but we include it for completeness.

Lemma 6 *We have*

$$\mathbb{E}\left[\sum_{u \in N(1)} e^{-\theta X_u(1)} \right] = \exp(2e^{\theta} - 1).$$

Proof: Let

$$E_\theta(t) = \mathbb{E}\left[\sum_{u \in N(t)} e^{-\theta X_u(t)} \right],$$

and for $s, t \geq 0$ and a particle $u \in N(t)$ define $N_u(t; s)$ to be the set of descendants of particle u alive at time $t + s$: that is, $N_u(t; s) := \{v \in N(t + s) : u \leq v\}$. Then by the Markov property,

$$
\begin{aligned}
E_\theta(t + s) &= \mathbb{E}\left[\sum_{u \in N(t+s)} e^{-\theta X_u(t+s)}\right] \\
&= \mathbb{E}\left[\sum_{u \in N(t)} e^{-\theta X_u(t)} \sum_{v \in N_u(t;s)} e^{-\theta(X_v(t+s) - X_v(t))}\right] \\
&= \mathbb{E}\left[\sum_{u \in N(t)} e^{-\theta X_u(t)} \mathbb{E}\left[\sum_{v \in N_u(t;s)} e^{-\theta(X_v(t+s) - X_v(t))} \,\middle|\, \mathcal{F}_t\right]\right] \\
&= \mathbb{E}\left[\sum_{u \in N(t)} e^{-\theta X_u(t)} E_\theta(s)\right] \\
&= E_\theta(t) E_\theta(s).
\end{aligned}
$$

We deduce that for $s, t > 0$,

$$
\frac{E_\theta(t + s) - E_\theta(t)}{s} = E_\theta(t)\left(\frac{E_\theta(s) - 1}{s}\right)
$$

and

$$
\frac{E_\theta(t - s) - E_\theta(t)}{-s} = E_\theta(t - s)\left(\frac{E_\theta(s) - 1}{s}\right).
$$

It is easily checked that $E_\theta(t)$ is continuous in t, and hence if $E'_\theta(0+)$ exists then by the above we have that $E_\theta(t)$ is continuously differentiable and for all $t > 0$

$$
E'_\theta(t) = E_\theta(t) E'_\theta(0+).
$$

Since $E_\theta(0) = 1$ this entails that

$$
E_\theta(t) = \exp(E'_\theta(0+)t).
$$

Now, for small t,

$$
\begin{aligned}
E_\theta(t) &= \mathbb{P}(\text{first split after } t) + 2e^\theta \mathbb{P}(\text{first split before } t) + o(t) \\
&= 1 - t + 2te^\theta + o(t)
\end{aligned}
$$

so that $E'(0+) = 2e^\theta - 1$, and hence $E_\theta(t) = \exp((2e^\theta - 1)t)$. Taking $t = 1$ completes the proof. $\qquad \square$

These simple properties of the Yule tree will allow us to prove our main theorem.

3   Proof of Theorems 1 and 2

We would like to apply the following theorem of Hu and Shi [7]. Their Theorem 1.2 was proved for a large class of branching random walks, and we translate it to our particular simple case. Their conditions (1.1) and (1.2) are trivially satisfied; and their Remark (ii) (page 745) shows that, when recentred, our case also satisfies (1.3). Translating into our notation, we obtain the following result.

Theorem 7 (Hu & Shi [7]) *Define*

$$\psi(\theta) := \mathbb{E}\left[\sum_{u \in N(1)} e^{-\theta X_u(1)}\right].$$

If θ^* satisfies

$$\frac{\theta^* \psi'(\theta^*)}{\psi(\theta^*)} = \log \psi(\theta^*), \quad \theta^* > 0,$$

then

$$\frac{1}{2} = \liminf_{n \to \infty} \frac{\theta^* M(n) + n \log \psi(\theta^*)}{\log n} < \limsup_{n \to \infty} \frac{\theta^* M(n) + n \log \psi(\theta^*)}{\log n} = \frac{3}{2}$$

and

$$\frac{\theta^* M(n) + n \log \psi(\theta^*)}{\log n} \xrightarrow{\mathbb{P}} \frac{3}{2} \quad as \quad n \to \infty.$$

In view of this result, our method of proof for Theorems 1 and 2 is unsurprising: we know that the times T_n are near $\log n$ for large n, and we may use the monotonicity of H_n and h_n — together with the flexibility offered by the $\log \log$ scale — to ensure that nothing else can go wrong. It may be possible to extend this method of proof to cover more general trees, where the same monotonicity property does not necessarily hold, via a Borel-Cantelli argument. This would only introduce unneccessary complications in our case.

Proof of Theorem 1: We show first the statement involving the limsup; the proofs of the other statements are almost identical.

It is immediate from Lemma 6 that a in Theorem 1 corresponds to θ^* in Theorem 7, and that b corresponds to $\log \psi(\theta^*)$. Fix $\delta > 0$. Choose $K \in \mathbb{N}$ such that

$$\mathbb{P}(\limsup_{n \to \infty} |T_n - \lfloor \log n \rfloor| > K) < \delta$$

— this is possible by Lemma 5. For each $n \geq 1$, let $j_n = \lfloor \log n \rfloor - K$. We use the abbreviation "i.o." to mean "infinitely often" — that is, for a sequence of measurable sets U_n, $\{U_n \text{ i.o.}\}$ represents the event

$\limsup_{n \to \infty} U_n$. For any $\varepsilon > 0$, using the fact that $M(t)$ is non-increasing,

$$\mathbb{P}(aM(T_n) + b\log n > (3/2 + \varepsilon)\log\log n \ \text{i.o.})$$
$$\leq \mathbb{P}(\{aM(T_n) + b\log n > (3/2 + \varepsilon)\log\log n, |T_n - \lfloor\log n\rfloor| \leq K\} \ \text{i.o.})$$
$$+ \mathbb{P}(|T_n - \lfloor\log n\rfloor| > K \ \text{i.o.})$$
$$< \mathbb{P}\bigg(aM(j_n) > -bj_n + (3/2 + \varepsilon)\log j_n + (bj_n - b\log n)$$
$$+ (3/2 + \varepsilon)(\log\log n - \log j_n) \ \text{i.o.}\bigg) + \delta$$
$$\leq \mathbb{P}(aM(j_n) > -bj_n + (3/2 + \varepsilon/2)\log j_n \ \text{i.o.}) + \delta$$
$$\leq \delta$$

by Theorem 7. Taking a union over $\varepsilon > 0$ tells us that

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{aM(T_n) + b\log n}{\log\log n} > \frac{3}{2}\right) \leq \delta;$$

but since $\delta > 0$ was arbitrary we deduce that

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{aM(T_n) + b\log n}{\log\log n} > \frac{3}{2}\right) = 0.$$

This completes the proof of the upper bound, since $H_n = -M(T_n)$. The proof of the lower bound is similar. We let $i_n = \lfloor\log n\rfloor + K$ and use the abbreviation "ev." to mean "eventually" (that is, for all large n; so $\{U_n \ \text{ev.}\}$ represents the event $\liminf_{n\to\infty} U_n$). For any $\varepsilon \in (0, 3/2)$,

$$\mathbb{P}(aM(T_n) + b\log n < (3/2 - \varepsilon)\log\log n \ \text{ev.})$$
$$\leq \mathbb{P}(\{aM(T_n) + b\log n < (3/2 - \varepsilon)\log\log n, |T_n - \lfloor\log n\rfloor| \leq K\} \ \text{ev.})$$
$$+ \mathbb{P}(|T_n - \lfloor\log n\rfloor| > K \ \text{i.o.})$$
$$< \mathbb{P}\bigg(aM(i_n) < -bi_n + (3/2 - \varepsilon)\log i_n + (bi_n - b\log n)$$
$$+ (3/2 - \varepsilon)(\log\log n - \log i_n) \ \text{ev.}\bigg) + \delta$$
$$\leq \mathbb{P}(aM(i_n) < -bi_n + (3/2 - \varepsilon/2)\log i_n \ \text{ev.}) + \delta$$
$$\leq \delta$$

by Theorem 7. As with the upper bound, taking a union over $\varepsilon > 0$, and then letting $\delta \to 0$, tells us that

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{aM(T_n) + b\log n}{\log\log n} < \frac{3}{2}\right) = 0$$

and hence combining with the upper bound we obtain

$$\limsup_{n\to\infty} \frac{b\log n - aH_n}{\log\log n} = \frac{3}{2}.$$

almost surely. The proof of the statement involving the liminf is almost identical, and we omit it for the sake of brevity. The convergence in probability is also similar: one considers for example that

$$\limsup_{n\to\infty} \mathbb{P}\left(aM(T_n) + b\log n > (3/2 + \varepsilon)\log\log n\right)$$

$$\leq \limsup_{n\to\infty} \mathbb{P}\left(aM(T_n) + b\log n > (3/2 + \varepsilon)\log\log n, \ |T_n - \lfloor\log n\rfloor| \leq K\right)$$

$$+ \limsup_{n\to\infty} \mathbb{P}\left(|T_n - \lfloor\log n\rfloor| > K\right)$$

$$< \limsup_{n\to\infty} \mathbb{P}\left(aM(T_n) + b\log n > (3/2 + \varepsilon)\log\log n, \ |T_n - \lfloor\log n\rfloor| \leq K\right) + \delta$$

and uses the statement about convergence in probability in Theorem 7 to show that the probability in the last line above converges to zero for any $\varepsilon > 0$. Then since $\delta > 0$ was arbitrary we must have

$$\limsup_{n\to\infty} \mathbb{P}\left(aM(T_n) + b\log n > (3/2 + \varepsilon)\log\log n\right) = 0.$$

The lower bound is, again, similar. □

Proof of Theorem 2: Consider a slightly altered Yule tree model, where each particle gives birth to two children whose position is that of their parent *plus* 1, instead of minus 1. If we couple this model with the usual Yule tree model in the obvious way, then clearly the minimal position of a particle in the altered model is equal to -1 times the maximal position in the usual model. Thus if we let $\hat{M}(t)$ be the minimal position in the altered model, it suffices to show that

$$\frac{1}{2} = \liminf_{n\to\infty} \frac{\alpha\hat{M}(T_n) - \beta\log n}{\log\log n} < \limsup_{n\to\infty} \frac{\alpha\hat{M}(T_n) - \beta\log n}{\log\log n} = \frac{3}{2}$$

and

$$\frac{\alpha\hat{M}(T_n) - \beta\log n}{\log\log n} \xrightarrow{\mathbb{P}} \frac{3}{2} \quad \text{as} \quad n \to \infty.$$

Lemma 6 (substituting $\hat{\theta} := -\theta$, say) tells us that for the altered model, α in Theorem 2 corresponds to θ^* in Theorem 7, and that $-\beta$ corresponds to $\log\psi(\theta^*)$. The rest of the proof proceeds exactly as in the proof of Theorem 1. □

4 The size of the fringe, F_n

We are now interested in the size of the fringe of the tree: how many leaves lie at level H_n at time n. Recall that we called this quantity F_n.

We will show that F_n is unbounded almost surely, but first we need a short lemma. For this lemma we consider again the Yule tree model, and call the set of particles with position $M(t)$ the *frontier* of the Yule tree at time t — recall that this is the set of particles with minimal position at time t, so as we saw earlier the frontier of the Yule tree corresponds to the fringe of the binary search tree. Define $\tilde{F}_t$ to be the number of particles at the frontier at time t,
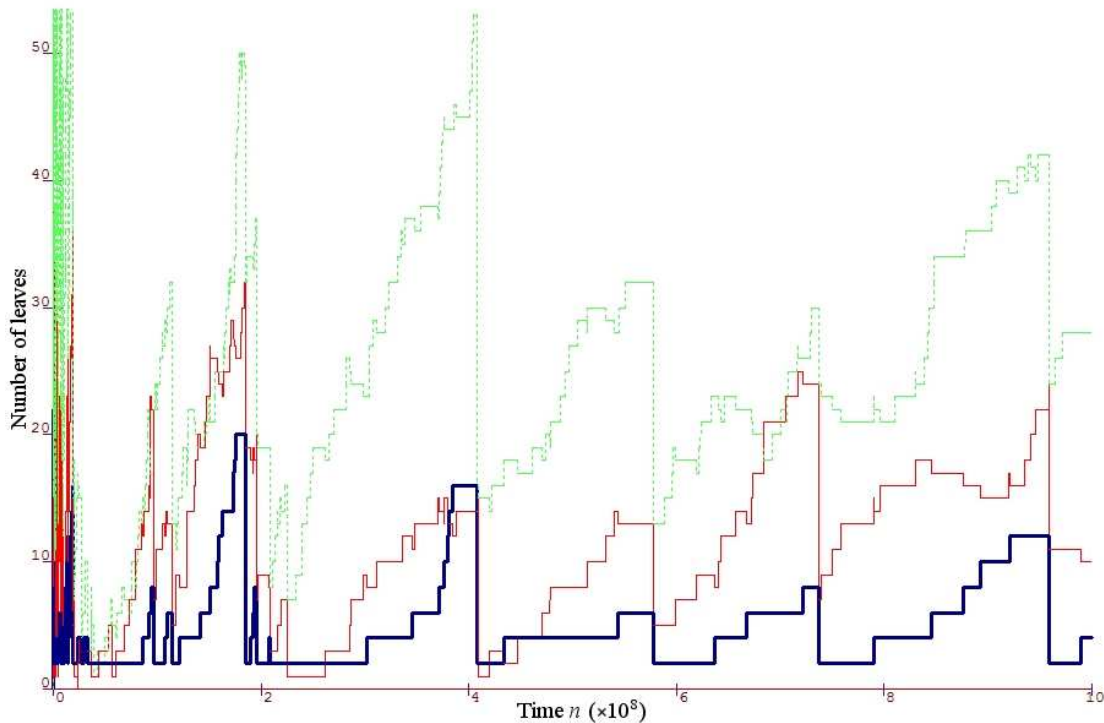
$$\tilde{F}_t := \#\{u \in N(t) : X_u(t) = M(t)\}.$$

Fig. 2: The top three levels of a binary search tree run for 10^9 steps. The thick blue line shows the size of the fringe, F_n, which is the number of leaves at level H_n; the thin red line shows the number of leaves at level $H_n - 1$; and the dashed green line shows the number of leaves at level $H_n - 2$.

Lemma 8 *If $M(t) < -\lfloor \log_2(2k) \rfloor$ and $\tilde{F}_t = 2k$, then there is at least one particle that is not at the frontier at time t, but which is within distance $\lfloor \log_2(2k) \rfloor$ of the frontier — that is, its position is in the interval $[M(t) + 1, M(t) + \lfloor \log_2(2k) \rfloor]$.*

Proof: Clearly at some time before t there was a particle which had position $M(t) + \lfloor \log_2(2k) \rfloor$; and hence at some time there were at least 2 particles with this position, since particles (except the root) arrive in pairs. At time t, either these particles have at least one descendant not at the frontier, in which case we are done (as particles cannot move in the positive direction); or all their descendants are at the frontier. So, for a contradiction, suppose that all their descendants *are* at the frontier at time t. Then there must be $2 \times 2^{\lfloor \log_2(2k) \rfloor}$ particles at the frontier (since a movement of distance 1 yields 2 new particles, and hence a movement of distance $\lfloor \log_2(2k) \rfloor$ yields $2^{\lfloor \log_2(2k) \rfloor}$ new particles; and this holds for each of the two initial particles). But

$$2 \times 2^{\lfloor \log_2(2k) \rfloor} = 2^{\lfloor \log_2(2k) \rfloor + 1} > 2^{\log_2(2k)} = 2k$$

so there are strictly more than $2k$ particles at the frontier. This is a contradiction — there are exactly $2k$ particles at the frontier, by assumption — and hence our claim holds.   □

We now prove Theorem 3, which we recall says that $\limsup_{n\to\infty} F_n = \infty$ almost surely.

Proof of Theorem 3: Again consider the continuous time Yule tree. By the relationship between the Yule tree and the binary search tree seen in Section 2, $\tilde{F}_t$ and F_n have the same paths up to a time change, and hence it suffices to show that $\limsup \tilde{F}_t = \infty$ almost surely.

The idea is as follows: suppose we have $2k$ particles at the frontier. By Lemma 8, there is a particle close to the frontier; and this particle has probability greater than some strictly positive constant of having 2 of its descendants make it to the frontier before the $2k$ already there branch. So if we have 2k particles infinitely often, then we have $2k + 2$ particles infinitely often. We make this argument rigorous below.

For any $t > 0$ and $k \in \mathbb{N}$, define

$$\tau_1^{(2k)} := \inf\{s > 0 : M(s) < -\lfloor \log_2(2k) \rfloor \text{ and } \tilde{F}_s = 2k\}$$

and for each $j \geq 1$

$$\sigma_j^{(2k)} := \inf\{s > \tau_j^{(2k)} : \tilde{F}_s \neq 2k\}$$

and

$$\tau_{j+1}^{(2k)} := \inf\{s > \sigma_j^{(2k)} : \tilde{F}_s = 2k\}.$$

Then $\tau_j^{(2k)}$ is the jth time that we have $2k$ particles at the frontier and at least distance $\log_2(2k)$ from the origin. We show, by induction on k, that for any $k \in \mathbb{N}$

$$\tau_j^{(2k)} < \infty \quad \text{almost surely, for all } j \in \mathbb{N}. \tag{1}$$

Trivially, since $M(t) \to -\infty$ almost surely (which is true since $H_n \to \infty$ almost surely), we have $\tau_j^{(2)} < \infty$ almost surely for all $j \in \mathbb{N}$ and so (1) holds for $k = 1$. Suppose now (1) holds for some $k \geq 1$.

By Lemma 8, for any j, at time $\tau_j^{(2k)}$ there is at least 1 particle that is not at the frontier but is within distance $\lfloor \log_2(2k) \rfloor$ of the frontier. Let $A_j^{(k)}$ be the event that the descendants of this particle reach level $M(\tau_j^{(2k)})$ before any of the $2k$ particles already at that level branch. Then the events $A_1^{(k)}, A_2^{(k)}, A_3^{(k)}, \ldots$ are independent by the strong Markov property. Also, since all particles branch at rate 1, for each j the probability of $A_j^{(k)}$ is certainly at least the probability that the sum of $\lfloor \log_2(2k) \rfloor$ independent, rate 1 exponential random variables is less that the minimum of $2k$ independent, rate 1 exponential random variables. This is some strictly positive number, γ_k say.

Now, at time $\tau_j^{(2k)}$ — which is finite for each j, by our induction hypothesis — there are $2k$ particles at the frontier. One of two things can happen: either two more particles join them and we reach $2k + 2$ particles at the frontier, or one of the $2k$ branches before this happens and we have a new frontier with 2 particles. Call the first event, that two more particles reach the frontier before any of the $2k$ already there branch, $B_j^{(k)}$. Then $A_j^{(k)} \subseteq B_j^{(k)}$ since the event that some pair makes it to the frontier before the $2k$ branch contains the event that descendants of our particular particle make it to the frontier before the $2k$

branch. Thus

$$\mathbb{P}\left(\limsup_{m\to\infty} B_m^{(k)}\right) \geq \mathbb{P}\left(\limsup_{m\to\infty} A_m^{(k)}\right) = \mathbb{P}\left(\bigcap_{n\geq 1}\bigcup_{m\geq n} A_m^{(k)}\right)$$

$$= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{m\geq n} A_m^{(k)}\right) = \lim_{n\to\infty}\lim_{N\to\infty} \mathbb{P}\left(\bigcup_{m=n}^{N} A_m^{(k)}\right)$$

$$\geq \lim_{n\to\infty}\lim_{N\to\infty}\left(1 - (1-\gamma_k)^{N-n+1}\right) = 1.$$

But the event $\limsup_{m\to\infty} B_m^{(k)}$ is exactly the event that we have $2k+2$ particles at the frontier infinitely often — and thus (using again that $M(t) \to -\infty$ almost surely) we have that $\tau_j^{(2k+2)}$ is finite almost surely for all j. Hence by induction we have proved that (1) holds for each k. Our result follows.   □

Acknowledgements

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

[3] B. Chauvin, T. Klein, J-F. Marckert, and A. Rouault. Martingales and profile of binary search trees. *Electron. J. Probab.*, 10(12):420–435, 2005.

[4] L. Devroye. A note on the height of binary search trees. *J. Assoc. Comput. Mach.*, 33(3):489–498, 1986.

[5] M. Drmota. An analytic approach to the height of binary search trees II. *J. ACM*, 50(3):333–374, 2003.

[6] M. Drmota, B. Gittenberger, A. Panholzer, H. Prodinger, and M.D. Ward. On the shape of the fringe of various types of random trees. *Math. Methods Appl. Sci.*, 32(10):1207–1245, 2009.

[7] Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, 37(2):742–789, 2009.

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

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